

# DAS ENDE DES VIRTUELLEN

---

## Digitale Methoden

Aus guten Gründen lässt sich ein ontologischer Unterschied zwischen dem genuin Digitalen und dem Digitalisierten machen: zwischen den Objekten, dem Content, den Verfahren und Umgebungen, die in dem neuen Medium <geboren> wurden, und solchen, die <eingewandert> sind. Sollten die etablierten Forschungsmethoden daher, mit Blick auf Objekte und Inhalte des *Mediums*, verändert werden?

Das hier vorgestellte Forschungsprogramm setzt sich im Kontext dieser Frage zunächst mit <virtuellen Methoden> auseinander, die Standardpraktiken aus den Sozial- und Geisteswissenschaften für die Internetforschung importieren. Die Unterscheidung zwischen dem genuin Digitalen und dem Digitalisierten lässt sich auch auf gegenwärtige Untersuchungsmethoden anwenden. Welche Typen der Internetforschung lassen sich mit digitalisierten Methoden (wie zum Beispiel Online-Befragungen) durchführen und welche mit genuin digitalen Methoden (wie Empfehlungsdiensten oder *folksonomy*)?

Schließlich werde ich vorschlagen, dass die Internetforschung neue Anwendungsgebiete finden kann, wenn man die genuin digitalen Methoden gegenüber den digitalisierten ins Zentrum stellt. Ich werde versuchen, die Aufmerksamkeit von den Problemstellungen, die sich aus der Transformation von analogen zu digitalen Verfahrensweisen ergeben, wegzulenken und stattdessen herauszufinden, wie Forschung *mit* dem Internet über das ausschließliche Studium der Onlinekultur hinausgehen kann. Wie erfasst und analysiert man Hyperlinks, Tags, Suchmaschinenresultate, archivierte Webseiten und andere digitale Objekte? Was kann man vom Umgang der bestehenden Online-Werkzeuge (wie Suchmaschinen und Empfehlungsdiensten) mit ihren Gegenständen lernen? Und wie können solche Verfahren für sozial- und kulturwissenschaftliche Forschungen umgewidmet werden? Letztendlich schlage ich eine Forschungspraxis vor, die auf der Basis von Online-Dynamiken Aussagen über kulturellen Wandel und gesellschaftliche Sachverhalte auch jenseits der Online-Kultur trifft – dazu

führe ich den Begriff *online groundedness* ein: Es geht darum, die <Erdung> von Online-Phänomenen und somit ihre Verankerung in Offline-Prozessen zu berücksichtigen. Das umfassendere Ziel besteht darin, die Methoden der Internetforschung zu überarbeiten und damit einen neuen Studienzweig zu entwickeln: digitale Methoden.

### Digitale Methoden im Kontext der Internetforschung

Bisher dienten die Methoden, die in der Internetforschung zur Anwendung kamen, dazu, die hartnäckige Idee, das Internet sei ein abgetrennter, virtueller Raum, kritisch zu hinterfragen. Diese Idee entstammte dem Diskurs der virtuellen Realität aus den späten 1980er und frühen 1990er Jahren. Das Internet bekam dabei die Rolle eines virtuellen Reichs zugewiesen, das Möglichkeiten eröffne, Bewusstsein, Identität, Körperlichkeit, Gemeinschaft, (Staats-)Bürgerschaft und Politik / soziale Bewegungen neu zu definieren.<sup>1</sup> Als dann aber der Kommunikationswissenschaftler Steve Jones 1999 einen der ersten Versuche unternahm, die verschiedenen Stränge der zeitgenössischen Internetforschung zusammenzuführen, forderte er die ForscherInnen auf, über die Perspektive des Internets als eines eigenen Bereichs hinauszugehen. Damit war die Methodendiskussion eröffnet.<sup>2</sup> Wie konnten SozialwissenschaftlerInnen das Internet studieren, ohne sich auf die Ansätze zu verlassen, die bisher damit assoziiert wurden: Mensch-Computer-Interaktion, Sozialpsychologie und Cyberkulturwissenschaften?<sup>3</sup> In ihrem grundlegenden Werk über Internetnutzung in Trinidad und Tobago widersprachen die Ethnographen Daniel Miller und Don Slater der Idee des Cyberspace als eines eigenen Bereichs, in dem alle, die ihn <bewohnen>, seine identitätsverändernden Konsequenzen erfahren können, unabhängig von dem Ort, an dem sie sich physisch befinden.<sup>4</sup> Slater und Miller verankerten das Internet, indem sie darlegten, wie die Bevölkerung von Trinidad und Tobago sich des Mediums in einer Weise bediente, die ihren eigenen kulturellen Praktiken entspricht.

Das Aufsehen erregende *Virtual Society*-Programm (1997–2002) brachte einen weiteren Wendepunkt in der Internetforschung mit sich. Es entlarvte den Mythos der transformativen Möglichkeiten des Cyberspace durch zahlreiche empirische Studien über Internetnutzer. Das Programm formulierte schließlich fünf «Regeln der Virtualität».<sup>5</sup> Die ForscherInnen argumentierten, dass die Verwendung neuer Medien durch die spezifischen Situationen geprägt ist (Zugangsaspekte) und dass Ängste und Risiken ungleich verteilt sind (Kompetenzaspekte). Diese Themen sind gegenwärtig in der Kritik an der digitalen Kluft (*digital divide*) fest etabliert. Das <Reale> wird durch die virtuellen Interaktionen weniger ersetzt als vielmehr ergänzt; diese stimulieren eher reale Interaktionen, als dass sie Isolation und Verzweiflung mit sich bringen würden. Schließlich zeigte die Untersuchung, dass Identitäten sich aus dem Zusammenspiel von Online- und Offline-Kontexten ergeben. Bezeichnender Weise

<sup>1</sup> John Perry Barlow, *A Declaration of the Independence of Cyberspace*, <http://homes.eff.org/ffbarlow/Declaration-Final.html>, dort datiert Februar 2006, gesehen am 10.07.2011; Michael Benedict, *Cyberspace: Some Proposals*, in: ders. (Hg.), *Cyberspace – First Steps*, Cambridge, MA (Cambridge MIT Press) 1991, 119–224; Julian Dibbell, *My Tiny Life: Crime and Passion in a Virtual World*, New York (Henry Holt) 1998; Howard Rheingold, *Virtual Reality: Exploring the Brave New Technologies*, New York (Summit) 1991; Steven Shaviro, *Money for Nothing: Virtual Worlds and Virtual Economies*, <http://shaviro.com/OtherTexts/MMOs.pdf>, dort datiert 2007, gesehen am 10.7.2011; Allucquere Rosa Stone, *The War of Desire and Technology at the Close of the Mechanical Age*, Cambridge, MA (MIT Press) 1995; Sherry Turkle, *Life on the Screen: Identity in the Age of the Internet*, New York (Simon & Schuster) 1995.

<sup>2</sup> Steve Jones, *Studying the Net: Intricacies and Issues*, in: ders. (Hg.), *Doing Internet Research: Critical Issues and Methods for Examining the Net*, London (Sage) 1999, 1–28.

<sup>3</sup> Christine M. Hine, *Virtual Ethnography*, London (Sage) 2000.

<sup>4</sup> Daniel Miller, Don Slater, *The Internet: An Ethnographic Approach*, Oxford (Berg) 2000.

<sup>5</sup> Steve Woolgar, *Five Rules of Virtuality*, in: ders. (Hg.), *Virtual Society? Technology, Cyberbole, Reality*, Oxford (Oxford University Press) 2002, 1–22.

verwendete das Projekt zu Untersuchung der Gebrauchsformen Verfahren, die man in der Folge als <virtuelle Methoden> charakterisierte. Befragungen, Interviews, Beobachtung und teilnehmende Beobachtung wurden die bevorzugten Untersuchungsmethoden.

Auch die bald darauf folgende kulturwissenschaftliche Nutzungsforschung, die sich auf den Amateur, den Fan und die Figur des <producers> konzentrierte, hatte mit der vermeintlichen Kluft zwischen Realität und Virtualität zu kämpfen und setzte sich kritisch mit den dominanten Vorstellungen von Onlinekultur auseinander.<sup>6</sup> Alles in allem haben die virtuellen Methoden und die Nutzungsforschung in den Sozial- und Kulturwissenschaften somit dazu beigetragen, die Aufmerksamkeit von den Charakteristika des Mediums wegzulenken. Damit wurde zugleich die Möglichkeit eröffnet, sehr viel mehr als nur die Onlinekultur zu erforschen.

Wie lässt sich nun auf der Basis von Daten, die (automatisch) durch Software erhoben werden, Nutzungsforschung neu denken? Bislang beruhte Nutzungsforschung weitgehend auf Beobachtung, Interviews und Befragungen. Geprägt war sie zum einen durch das unterschiedliche Arsenal an sozial- und geisteswissenschaftlicher Datenverarbeitung und zum anderen durch die Forschung großer kommerzieller Unternehmen mit ihren enormen Datenbeständen. In gewisser Weise betreiben allerdings auch Google, Amazon und viele andere dominante Web-Anwendungen Nutzungsforschung, auch wenn dieser Begriff selten dafür verwendet wird.

Unter Bezug auf derartige Unternehmungen hat Lev Manovich zu einer methodologischen Wende in der Internetforschung aufgerufen, zumindest was die Datengewinnung betrifft. Sein Vorschlag einer Kulturanalytik (der Name *Cultural Analytics* bezieht sich ausdrücklich auf *Google Analytics*) zielt auf die Errichtung massiver Computerkapazitäten für die geisteswissenschaftliche Sammlung, Aufbewahrung und Analyse von Daten.<sup>7</sup> Ein wesentliches Kennzeichen dieser methodologischen Wende ist, dass sie sich – ungeachtet dessen, wie wertvoll diese Erkenntnisse sein mögen – nicht mehr auf den (ggf. auszuhandelnden) Zugriff auf kommerzielle Datenbestände verlässt (wie beispielsweise die Suchanfragen, die AOL von seinen Nutzern gespeichert hat, oder Linden Labs Protokolle der Aktivitäten von Millionen Nutzern in Second Life oder Sonys Daten aus Everquest) – stattdessen werden eigene Datenbestände geschaffen (etwa durch Digitalisierung von Büchern, Bildern etc.).<sup>8</sup> Kulturanalytik strebt dabei, nicht anders als die computerbasierten Sozialwissenschaften, *big science* an:<sup>9</sup> «Visualizations should be designed to take full advantage of the largest gigapixel wall-size displays available today.»<sup>10</sup>

Diese Forschungsprogramme sind gewissermaßen die Antworten der wissenschaftlichen Community auf die Frage: Was würde Google tun? Bisher hat die Kritik an der <Googlisierung>, die vor allem durch das Eindringen des Suchmaschinenunternehmens in die Bibliothek (das Projekt *Google Books*) ausgelöst wurde, die heimliche Ausdehnung von Google untersucht, sein Geschäftsmodell

<sup>6</sup> Henry Jenkins, *Convergence Culture: Where Old and New Media Collide*, New York (NYU Press) 2006; Andrew Keen, *The Cult of the Amateur: How Today's Internet is Killing Our Culture*, London (Nicholas Brealey) 2007; Axel Bruns, *Blogs, Wikipedia, Second Life, and Beyond: From Production to Produsage*, New York (Peter Lang) 2008.

<sup>7</sup> Lev Manovich, *Cultural Analytics. Analysis and Visualization of Large Cultural Data Sets*, [http://www.manovich.net/cultural\\_analytics.pdf](http://www.manovich.net/cultural_analytics.pdf), dort datiert 30.9.2007, gesehen am 28.1.2009.

<sup>8</sup> Noshir Contractor, *Digital Traces: An Exploratorium for Understanding and Enabling Social Networks*, Präsentation im Rahmen von The Annual Meeting of the American Association for the Advancement of Science (AAAS), 2009.

<sup>9</sup> David Lazer u. a., *Computational Social Science*, in: *Science*, 323, 2009, 721–723.

<sup>10</sup> Lev Manovich, *Software Takes Command*, <http://www.softwarestudies.com/softbook>, dort datiert 20.11.2008, gesehen am 10.7.2011.

und seine Ästhetik, über Informations- und Wissensindustrien hinweg.<sup>11</sup> Der <Google-Effekt> wird auch daran deutlich, dass Surfen und Browsen zunehmend durch Suchen ersetzt wurden.

### «Online Groundedness»: Jenseits von real vs. virtuell

Ich möchte vorschlagen, eine neue Ära der Internetforschung zu begründen, die sich nicht länger mit der Kluft zwischen dem Realen und dem Virtuellen aufhält. Dies betrifft eine Veränderung der Fragen, die in Untersuchungen des Internets gestellt werden. Das Internet wird dabei zum Ort der Erforschung von viel mehr als *bloß* der Onlinekultur. Das Thema ist nicht mehr, wie groß der gesellschaftliche und kulturelle Teil ist, der online ist, sondern eher, wie sich anhand der Nutzung des Internets kulturelle Veränderungen und gesellschaftliche Sachverhalte diagnostizieren lassen. Der konzeptuelle Ausgangspunkt für das Forschungsprogramm ist die Anerkennung der Tatsache, dass das Internet nicht nur ein Untersuchungsgegenstand ist, sondern auch eine Quelle.

Eines der bemerkenswerteren Beispiele ist Google Flu Trends, ein nicht-kommerzielles Projekt auf Google.org, das 2008 gestartet wurde und das lokale Grippeausbrüche antizipiert, indem es Suchmaschinenanfragen nach Grippe, Grippe-symptomen und verwandten Begriffen zählt und daraufhin die Orte <geolokalisiert>, an denen die Abfragen gemacht wurden. Bislang etablierte Methoden der Datengewinnung (Meldungen aus Notaufnahmen in Krankenhäusern) werden dadurch infrage gestellt und das Web wird als ein Medium der Prognostik und Vorsorge entdeckt, das sehr viel enger mit der realen Geografie und Lebenswelt verbunden ist, als man erwarten würde.<sup>12</sup>

Das Digital Methods-Programm schlägt den Begriff *online groundedness* vor, um Forschung zu beschreiben, die dem Medium folgt, die dessen Dynamiken erfasst und fundierte Aussagen über den kulturellen und gesellschaftlichen Wandel trifft. Das umfassendere theoretische Ziel von digitalen Methoden besteht darin, die Beziehungen zwischen dem Web und der Offline-Realität neu zu denken.

### Folge dem Medium: Das Digital Methods Research Program

Warum soll man dem Medium folgen? Ein Ausgangspunkt ist die Einsicht, dass Internetforschung häufig mit instabilen Untersuchungsgegenständen zu tun hat. Diese Instabilität wird oft der Flüchtigkeit von Websites und anderen digitalen Medien zugeschrieben und den Schwierigkeiten, sie zu *fixieren* – um einen Begriff aus der Fotografie zu entlehnen.

Webarchivierung steht ständig vor dem Dilemma, Websites zu erfassen, ohne deren Lebendigkeit zu beeinträchtigen. Ein Ansatz geht dahin, alte Hardware und Software aufzubewahren, um so die Medien als <Untote> zu erhalten. Eine andere Technik, die in Game-Umgebungen praktiziert wird, antwortet

<sup>11</sup> Jean-Noel Jeanneney, *Google and the Myth of Universal Knowledge*, Chicago (University of Chicago Press) 2007; Siva Vaidhyanathan, *Where is this book going? in: The Googlization of Everything Blog*, [www.googlizationofeverything.com/2007/09/where\\_is\\_this\\_book\\_going.php](http://www.googlizationofeverything.com/2007/09/where_is_this_book_going.php), dort datiert 25.9.2007, gesehen am 22.12.2008; Richard Rogers, *The Googlization Question, and the Inculpable Engine*, in: Felix Stalder, Konrad Becker (Hg.), *Deep Search: The Politics of Search Engines*, Edison, NJ (Transaction Publishers) 2009, s. auch: [http://www.goucom.org/publications/full\\_list/rogers\\_inculpable\\_engine\\_2009.pdf](http://www.goucom.org/publications/full_list/rogers_inculpable_engine_2009.pdf), gesehen am 10.7.2011.

<sup>12</sup> Richard Rogers, *The Viagra Files: The Web as Anticipatory Medium*, in: *Prometheus*, 21. Jg., 2/2003, 195–212.

mit Simulation/Emulation auf die Flüchtigkeit und lässt die nostalgische Software, zum Beispiel Atari-Spiele, auf heutiger Hardware laufen. Das Problem der Vergänglichkeit betrifft viel mehr als nur die Frage der Erhaltung. Die Internetforschung wird häufig von Ereignissen des Mediums überrumpelt, zum Beispiel Software-Updates, die Forschungsergebnisse gegenstandslos machen.

Die Idee, <dem Medium (zu) folgen> (statt es zu fixieren), lässt sich als Forschungspraxis auch in Zusammenhang mit einem Begriff beschreiben, der aus dem Journalismus und der Wissenschaftssoziologie entlehnt ist – dem <Scoop>. Wer etwas zuerst veröffentlicht, der landet einen Scoop. Wenn jemand einen Scoop landet, dann hat auch jemand das Nachsehen. Der Wissenschaftssoziologe Michael Lynch hat diesen Begriff auf Situationen angewendet, in denen die Untersuchungsgegenstände zu denselben oder ähnlichen Schlussfolgerungen kommen wie die Forscher und zuerst damit an die Öffentlichkeit gehen. Das Ergebnis ist, dass die «Erforschten das Feld neu konfigurieren, in dem wir unsere Studie ursprünglich situiert hatten».<sup>13</sup>

In der Internetforschung kommt es häufig vor, dass man durch einen Scoop das Nachsehen hat. Industrieanalysten, Regulierungsbehörden und Blogger schaffen ständig neue Begriffe (etwa *googlization*) und kommen zu Ergebnissen, die die fortlaufende akademische Arbeit mitdefinieren. Ich würde aber argumentieren, dass Scoops auch von den Gegenständen selbst kommen können, die permanent neu konfiguriert werden. Zum Beispiel hat man das soziale Netzwerk Facebook als ein relativ abgeschlossenes Community-System bezeichnet, in dem nur «Freunde» die Informationen und Aktivitäten anderer Freunde sehen können. Im März 2009 veränderte Facebook eine Einstellung. Nun können Nutzer ihr Profil für alle anderen Nutzer mit einem Account öffnen; davor war dies nur für Freunde und Freunde von Freunden möglich gewesen. Welche Forschungstypen wurden durch dieses Umlegen eines Schalters bei Facebook <gescoopt>? Welche Forschung profitierte davon?

In theoretischer Perspektive ist <dem Medium zu folgen> eine besondere Form medienspezifischer Forschung. Medienspezifität betrifft nicht nur die Untereinteilung disziplinärer Zuständigkeiten in den Medienwissenschaften in Entsprechung zu den primären Untersuchungsgegenständen: Film, Radio, Fernsehen etc. Zugleich ist es ein Plädoyer dafür, deren ontologische Unterscheidbarkeit ernst zu nehmen, auch wenn die Mittel zur Konstruktion dieser Ontologien differieren.

Die Medienspezifität, die hier gemeint ist, liegt nicht so sehr in McLuhans Beanspruchung der Sinne, den sozialen Formen von Raymond Williams oder in der Materialität von Katherine Hayles oder in den Eigenschaften und Befunden anderer Theoretiker.<sup>14</sup> Vielmehr liegt sie in der Methode. Ich habe das an anderer Stelle als «Web-Epistemologie» beschrieben.<sup>15</sup> Im Web werden Informationen, Wissen und Gemeinschaftlichkeit durch Empfehlungsdienste organisiert – Algorithmen und Scripts, die Ordnungen in URLs, Datenfiles, Freunden etc. schaffen und anbieten.

<sup>13</sup> Michael Lynch, A sociology of knowledge machine, in: *Ethnographic Studies*, 2/1997, 16–38.

<sup>14</sup> Marshall McLuhan, *Understanding Media: The Extensions of Man*, New York (McGraw Hill) 1964; Raymond Williams, *Television: Technology and Cultural Form*, London (Fontana) 1974; Katherine Hayles, Print is Flat, Code is Deep: The Importance of Media-Specific Analysis, in: *Poetics Today*, 25. Jg. 1/2004, 67–90; Alexander Galloway, *Protocol: How Control Exists After Decentralization*, Cambridge, MA (MIT Press) 2004.

<sup>15</sup> Richard Rogers, *Information Politics on the Web*, Cambridge, MA (MIT Press) 2004.

Untersuchungen zur Web-Epistemologie entstanden anfänglich im Zusammenhang mit der Politik von Suchmaschinen.<sup>16</sup> Sie versuchten die Kriterien zu erfassen, anhand derer Quellen von Suchmaschinen zugeordnet wurden. Warum ergab eine Suche mit dem Begriff <Terrorismus> im März 2003 das Weiße Haus, die Central Intelligence Agency (CIA), das Federal Bureau of Investigation (FBI), die konservative Heritage Foundation und führende Nachrichtenorganisationen wie CNN als erste Resultate? Die Antwort hat damit zu tun, wie Hyperlinks behandelt werden. Hyperlinks sind allerdings nur ein digitales Objekt unter anderen: Thread, Tag, PageRank, Wikipedia-Edit, robots.txt, Posts, Kommentare, trackback, pingback, IP Adressen, URLs, whois, timestamp, permalink, social bookmark, Nutzerprofile – diese ungeordnete Liste ließe sich fortsetzen. Mein Vorschlag zielt darauf ab, den spezifischen Umgang des Mediums mit diesen Objekten zu untersuchen und somit die Methode von dem Medium zu lernen.

Im Folgenden möchte ich eine Reihe von Objekten, Anwendungen, Räumen und Plattformen des Mediums vorstellen. Ich möchte zuerst kurz darauf eingehen, wie sie üblicherweise mit digitalisierten Methoden und konzeptuellen Ausgangspunkten von außerhalb des Mediums untersucht werden. Dann möchte ich den Unterschied diskutieren, der sich ergibt, wenn man bei der Untersuchung dem Medium folgt und dabei lernt und dann auch zur Anwendung bringt, wie digitale Gegenstände von Anwendungen behandelt werden, wie Webseiten archiviert werden, wie Suchmaschinen Informationen ordnen und wie Geo-IP-Lokalisierungstechniken Inhalte nach nationalen oder linguistischen Kriterien anbieten. Insgesamt liegt das Ziel dieses <dem Medium folgen> darin, die Internetforschung neu auszurichten und das Internet als Quelle von Daten, Methoden und Technologien zu betrachten.

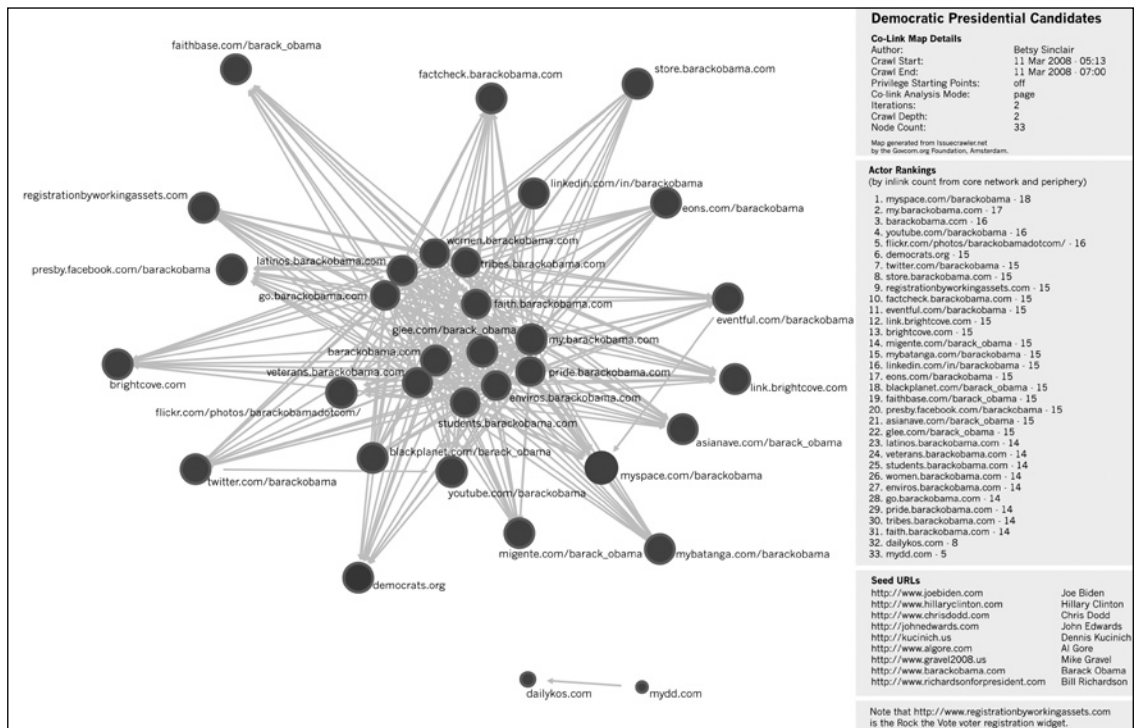
### Der Link

Wie wird der Hyperlink gewöhnlich betrachtet? Es gibt zumindest zwei dominante Ansätze beim Studium von Hyperlinks: die Literaturtheorie des Hertexts und die Theorie sozialer Netzwerke, darunter die *small world*-Theorie und die Pfadtheorie.<sup>17</sup>

Wie gehen, im Vergleich dazu, Suchmaschinen mit Links vor? Man kann sagen, dass sie einen szientometrischen Ansatz verfolgen (und teils den einer Soziologie der Assoziationen). Wie bei der Analyse sozialer Netzwerke steht die Positionierung der Akteure im Mittelpunkt des Interesses; dabei spielt die Distanz zwischen den Akteuren oder die Möglichkeit, einen Akteur durch Netzwerke zu erreichen, eine untergeordnete Rolle. Verbindungen sind vielmehr reputationelle Indikatoren, die Aufschluss über das Standing von Akteuren ermöglichen. Zudem spielt es bei dem Ansatz keine Rolle, ob die Verbindungen zwischen den Akteuren freundschaftlich sind oder eine andere Zweckmäßigkeit besitzen. Akteure können durch die Anzahl der auf sie verweisenden Links und, einen Schritt

<sup>16</sup> Lucas D. Introna, Helen Nissenbaum, *Shaping the Web: Why the Politics of Search Engines Matters*, *The Information Society*, 16. Jg., 3/2000, 1–17.

<sup>17</sup> George P. Landow, *Hyper/Text/Theory*, Baltimore, MD (Johns Hopkins University Press) 1994; Duncan J. Watts, *Small World*, Princeton (Princeton University Press) 1999; Han Woo Park, Mike Thelwall, *Hyperlink Analyses of the World Wide Web: A Review*, in: *Journal of Computer-Mediated Communication*, 8. Jg., 4/2003, s. auch <http://jcmc.indiana.edu/vol8/issue4/park.html>, gesehen am 10.7.2011; Greg Elmer *Profiling Machines*, Cambridge, MA (MIT Press) 2004; Valdis E. Krebs, *Mapping Networks of Terrorist Cells*, in: *Connections*, 24. Jg., 3/2002, 43–52.



weiter, durch die Anzahl der auf sie verweisenden Links auf Seiten, die selbst besonders viele Links erhalten, im grundlegenden Suchmaschinenalgorithmus profiliert werden. Akteure können darüber hinaus aber auch durch eine Untersuchung der spezifischen Links definiert werden, die auf sie verweisen und mit denen sie auf andere verweisen.<sup>18</sup> Die Analyse von Links dient dabei nicht nur der Erfassung der Mikropolitik von Hyperlinks, sie kann auch zur Erhebung komplexer Muster eingesetzt werden. Hier zeigt sich die Unterscheidung zwischen digitalisierten und genuin digitalen Methoden besonders deutlich. Die Open Net-Initiative (ONI) an der Universität von Toronto betreibt Untersuchungen über Internetzensur, indem sie Listen von Websites erstellt (aus Online-Verzeichnissen wie dem *Open Directory Project* oder Yahoo). Danach überprüfen die Forscher, ob diese Seiten in bestimmten Ländern blockiert sind. Das ist eine wichtige Arbeit, die Aufschlüsse über das weltweite Ausmaß staatlicher Internetzensurpraktiken wie auch über die dabei verwendete technische Infrastruktur gibt.<sup>19</sup> In der analytischen Praxis werden Seiten nach Kategorien gruppiert: berühmte Blogger, Regierungsseiten, Menschenrechtsseiten, Humor, Frauenrechte etc. Es gibt ungefähr vierzig Kategorien. Zensurmuster lassen sich dementsprechend nach Seitentyp je nach Land erforschen. Die vollständige Liste der Websites, die pro Land überprüft werden (rund 3.000), stellt eine Auswahl dar und erfasst natürlich nur einen sehr kleinen Teil aller Webseiten und auch nur eine Auswahl der Webseiten zu einem bestimmten Thema.

**Abb. 1** Hyperlink-Analyse von Barack Obamas Online-Kampagne: Verknüpfung von *Obama.com* und den Subsites seiner Wählerschaft mit Obamas Seiten in den Sozialen Medien

<sup>18</sup> Anne Beaulieu, *Sociable Hyperlinks: An Ethnographic Approach to Connectivity*, in: Christine M. Hine (Hg.), *Virtual Methods: Issues in Social Research on the Internet*, Oxford (Berg) 2005, 183–197; Noor Marres, Richard Rogers, *Depluralising the Web, Repluralising Public Debate. The GM Food Debate on the Web*, in: Richard Rogers (Hg.), *Preferred Placement*, Maastricht (Jan van Eyck Editions) 2000, 113–135; Richard Rogers, *Operating Issue Networks on the Web*, in: *Science as Culture*, 11. Jg., 2/2002, 191–214.

<sup>19</sup> Robert Deibert, John Palfrey, Rafal Rohozinski, Jonathan Zittrain (Hg.), *Access Denied: The practice and policy of global Internet filtering*, Cambridge, MA (MIT Press) 2008.



**Abb. 2** Der Niedergang des *directory*: Zunächst auf der Startseite als eigenständiges Feld geführt, wird der *directory* allmählich in «more» und «even more» verbannt. Screenshots der Google.com-Startseite 2000 bis 2007, erstellt mit der Wayback Machine des Internet Archive

Nach welchen Kriterien würde man Webseiten erfassen, wenn man eine Methode wählt, die dem Medium folgt, wenn man davon lernt, wie Suchmaschinen funktionieren (Linkanalyse), und wenn man dies für sozialwissenschaftliche Forschung umwidmet? Meine KollegInnen und ich haben zu der Arbeit der Open Net-Initiative beigetragen, indem wir ein Verfahren eingesetzt haben, das alle Webseiten in einer bestimmten Kategorie durchgeht, die Hyperlinks dieser Seiten erfasst und (durch eine Co-Link-Analyse) zusätzliche Schlüssel-seiten bestimmt, die nicht auf den Listen sind. Ich habe diese Methode als «dynamische URL-Erfassung» bezeichnet (*dynamic URL sampling*), um so den Unterschied zwischen manueller URL-Listenerstellung und stärker automatisierten Techniken zur Auffindung signifikanter URLs hervorzuheben. Sobald die neuen Seiten gefunden sind, werden sie auf Verbindungsstatistiken überprüft (zunächst durch Proxies, später vielleicht durch Maschinen, die sich in den entsprechenden Ländern befinden), um herauszufinden, ob sie blockiert werden. In dem Forschungsprojekt zu «sozialen, politischen und religiösen» Webseiten im Iran gingen ich und meine KollegInnen alle Seiten in dieser ONI-Kategorie durch und durch Hyperlinks-Analyse fanden wir gut dreißig, zuvor unbekannte blockierte Seiten. Wichtig ist dabei, dass es sich auch um eine Analyse auf Seitenebene handelte (und nicht nur auf der Ebene der Hosts). So konnte herausgefunden werden, dass der Iran die Starterseite von BBC News nicht blockiert (wie ONI es festgestellt hatte), sondern nur deren Seite in persischer Sprache. Der Unterschied zwischen den beiden Methoden zur Gewinnung von Website-Listen zur Analyse – manuelle Arbeit an Verzeichnissen und dynamische URL-Erfassung – zeigt den Beitrag medienspezifischer Methoden.



## Die Website

Bisher waren Untersuchungen von Webseiten dominiert von Nutzungs- und Wahrnehmungsforschungen, in denen Versuche zu einer Poetik der Navigation auf so ernüchternde Ideen wie <don't make me think> trafen.<sup>20</sup> Viele Methoden zum Studium von Webseiten gleichen dem Blick über die Schulter von NutzerInnen, bei dem man die Navigation oder die Verwendung einer Suchmaschine beobachtet, woraufhin man später Interviews mit der betroffenen Person führt. Eine populäre Technik, die man den klassischen datenbasierten Ansätzen zurechnen kann, ist *eye tracking*. Daraus ergeben sich Verteilungsmuster, die sogenannten *heat maps*, aus denen man Schlüsse für das Neudesign der Seiten ziehen kann. Google.com hat zum Beispiel seine Dienste von oberhalb des Suchfeldes (*tabs*) in die linke obere Ecke der Seite – in ein Menü – gezogen.

Ein weiterer dominanter Trend in Webseiten-Studien ist die Feature-Analyse. Dabei werden Seiten auf Basis der Gewichtung von Interaktivität, Möglichkeiten für Nutzer-Feedback etc. verglichen und gegenübergestellt.<sup>21</sup> Die Fragen zielen darauf ab, ob eine bestimmte Kombination von Features zu mehr Nutzern oder mehr Aufmerksamkeit führt. In dieser Tradition, vor allem in der Spezialsammlung 9/11, werden Websites häufig für künftige Studienzwecke archiviert. Eine Menge Arbeit geht also in die Archivierung der Seiten, bevor man überhaupt erst mit der Analyse beginnen kann. Eine der wichtigsten kommenden Aufgaben betrifft die Mittel, mit denen Webseiten erfasst und gesichert werden können, um so die Daten sicherzustellen, auf deren Basis Erkenntnisse gewonnen werden können. Das Forschungsprogramm zu digitalen Methoden beschäftigt sich deswegen speziell mit der Webseite als archiviertes Objekt, das vor allem durch die Wayback Machine des Internet Archive besonders leicht zugänglich gemacht wird. Das Forschungsprogramm fragt insbesondere, welche Formen von Webseiten-Untersuchungen durch die Wayback Machine ermöglicht und welche behindert werden können.

Um diese Frage zu beantworten, dekonstruiert die Arbeit zuerst das Internet Archive und seine Wayback Machine. Sie öffnet die *black box*. Die Vorstellung, dass der Untersuchungsgegenstand durch eben die Mittel konstruiert wird, durch die er in den Griff bekommen und methodisch erfasst werden kann, ist ein klassisches Motiv in Wissenschaftssoziologie und -philosophie und anderswo.<sup>22</sup> Die erste Frage ist deshalb: Welche Forschungsmethoden werden durch die spezifische Form, die das Webarchiv angenommen hat, privilegiert und welche werden beschränkt? Wenn man zum Beispiel das Internetarchiv benutzt (archive.org), dann fällt Nutzern, die an Suchmaschinen gewöhnt sind, nicht so sehr die Errungenschaft der Existenz eines archivierten Internets auf; vielmehr sind die Nutzer überrascht darüber, wie das Internet archiviert wird, und noch mehr darüber, wie es durchsucht wird. Man sucht nach einer URL – und nicht nach Schlagwörtern – und man bekommt eine Liste von abgelegten Seiten, die

<sup>20</sup> Steve Krug, *Don't Make Me Think! A Common Sense Approach to Web Usability*, Indianapolis, IN (New Riders) 2000; Anthony Dunne, *Hertzian Tales: Electronic Products, Aesthetic Experience, and Critical Design*, Cambridge, MA (MIT Press) 2005.

<sup>21</sup> Kirsten A. Foot, Steven M. Schneider, *Online Action in Campaign 2000: An Exploratory Analysis of the U.S. Political Web Sphere*, in: *Journal of Broadcast and Electronic Media*, 46. Jg. 2/2002, 222–244.

<sup>22</sup> Karin Knorr-Cetina, *Epistemic Cultures*, Cambridge, MA (Harvard University Press) 1999; Bruno Latour, Steve Woolgar, *Laboratory Life*, Princeton, NJ (Princeton University Press) 1986; Jil Walker, *Feral Hypertext: When Hypertext Literature Escapes Control*, *Proceedings of the Sixteenth ACM conference on Hypertext and Hypermedia*, 6.–9. September 2005, Salzburg, 46–53.

mit der URL in der Vergangenheit verbunden waren. Im Endeffekt organisiert das Internet Archive durch das Interface der Wayback Machine, die Geschichte des Webs als eine Geschichte individueller Seiten.

Mit der Wayback Machine kann man die Evolution einer einzelnen Seite (oder mehrerer Seiten) im Lauf der Zeit studieren, zum Beispiel, indem man Snapshots liest oder sammelt, die gemacht wurden, als die Seite indexiert wurde. Das Digital Methods-Programm hat Verfahren entwickelt, die Geschichte von Seiten durch die Herstellung von Snapshots zu erfassen, die dann zu einem Film nach dem Vorbild der Zeitrafferfotografie zusammengefügt werden.<sup>23</sup> Um zu demonstrieren, wie man das Internet für die Erfassung solcher evolutionärer Geschichten nutzt, machten meine KollegInnen und ich Snapshots der Startseiten von Google von 1998 bis Ende 2007. Die Analyse betraf die kaum merklichen Veränderungen an diesem Interface, insbesondere die Tabs. Wir fanden heraus, dass das Google-Verzeichnis (*directory*), die Organisation des Webs nach Themen durch menschliche Redakteure, immer mehr in den Hintergrund gerückt ist. 2001 wurde es auf der Google-Startseite platziert, 2004 wurde es unter dem *More*-Button rubriziert und 2006 unter *Even More*. Ende 2007 wurde die Option *Even More* entfernt und nun musste man Google durchsuchen, um das Verzeichnis überhaupt finden zu können.<sup>24</sup> Der Niedergang der menschlichen Redakteure, den man in diesem Fall aus der Evolution der Google-Seite herauslesen kann, ist eine Angelegenheit von größerer Bedeutung und hat weitreichende Implikationen für die Sammlung und Ordnung von Informationen. Nach dieser Untersuchung Googles wandten wir uns Yahoo zu, dem ursprüngliche Web-Verzeichnis, und fanden heraus, dass auch dort das Verzeichnis durch einen Server-Algorithmus ersetzt worden war. Indem wir die Ergebnisse einer Suche im Verzeichnis untersuchten, fanden wir auch heraus, dass auf Yahoo die Ergebnisse nicht mehr alphabetisch geordnet werden, also in der egalitären Form der Informations- und Quellenangabe, die von den Enzyklopädien geerbt worden war. Yahoo reiht sein Quellenverzeichnis nun nach Popularität, das heißt allgemeiner: in der wohlbekannten Weise von Empfehlungsdiensten.

### Die Suchmaschinen & die Sphären

Das Studium von Suchmaschinen wurde durch die inzwischen stark kritisierte Veröffentlichung von AOL-Suchmaschinendaten im Jahr 2006 angestoßen. Damals wurden die Suchanfragen von 500.000 Nutzern aus einem Zeitraum von drei Monaten online gestellt. Die Presseberichte über das Ausmaß der Preisgabe intimer Details klangen furchteinflößend und aufgeregt, obwohl die Suchgeschichten anonymisiert worden und geografisch nicht mehr zuordenbar waren (keine IP-Adressen). Man kann die Erkenntnisse aus dem AOL-Fall als eine Verschiebung der Interpretation von Onlinepräsenz verstehen – falls dies weiterhin ein angemessener Begriff ist. Eine Person kann <gegoogelt> werden und seine oder ihre selbstautorisierte Präsenz erscheint dann häufig ganz oder ziemlich weit oben in

<sup>23</sup> Seitenerfassungssoftware wurde davor schon für die Analyse von Wikipedia-Seiten genutzt. So konnte man die Entwicklung von Einträgen studieren und damit auch, wie Wikipedia Wissen schafft.

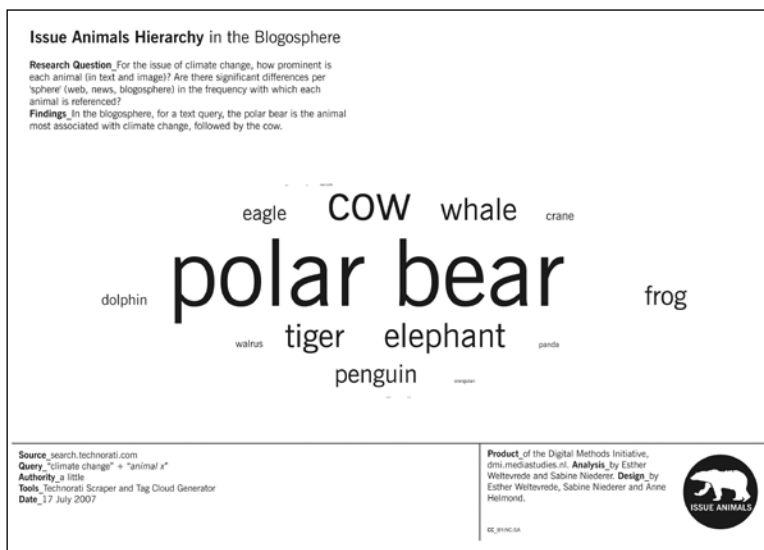
<sup>24</sup> Der *Even More*-Button kehrte 2008 auf die Seite Google.com zurück.

den Ergebnissen. Allgemein gesprochen würde das, was andere über eine Person geschrieben haben, in den Ergebnissen weiter unten erscheinen. Nun aber, da Suchmaschinenanfragen aufbewahrt werden, könnte bei der Bestimmung eines Individuums noch ein dritter Satz an Spuren bedeutsam werden.

Ich habe zuvor einen anderen Zweig der Suchmaschinenforschung erwähnt, der sich mit dem Begriff Googlisierung (*googlization*) zusammenfassen lässt. Es handelt sich dabei um eine Kritik im politisch-ökonomischen Stil, die sich damit beschäftigt, wie Googles Geschäftsmodell, das kostenlosen Service im Austausch gegen Profilinformationen bietet, sich auf andere Industrien und (Software-)Kulturen verbreiten könnte. Ich habe mich mit dieser Kritik an anderer Stelle beschäftigt und dabei versucht, eine Forschungsagenda für Googlisierungs-Forscher zu entwerfen, die sich sowohl auf *front end*- wie *back end*-Googlisierung bezieht. Zur *front end*-Googlisierung würde die Untersuchung der Informationspolitik des Benutzerinterface zählen (inklusive des Niedergangs des von Menschen gestalteten Verzeichnisses). *Back end*-Googlisierung betrifft den Siegeszug der Algorithmen, die Quellen in hierarchischer Ordnung empfehlen und nicht mehr alphabetisch, wie oben erwähnt.

Die Bedeutung des Studiums der neuen, durch Suchmaschinen produzierten, Informationshierarchien sollte auch aus der Perspektive der Nutzungsforschung gesehen werden. Nur ein kleiner Prozentsatz von Nutzern verändert die Voreinstellungen so, dass mehr als zehn Resultate pro Seite erscheinen. Üblicherweise sehen Nutzer sich nur die Resultate auf der ersten Seite an und tendieren dazu, die obersten Resultate auf der ersten Seite anzusehen.<sup>25</sup> Die Macht der Suchmaschinen liegt deswegen in der Kombination ihrer Rankingverfahren (*source inclusion* bei den Topresultaten) mit dem offensichtlichen <Respekt> der Nutzer vor dieser Anordnung (der dazu führt, dass sie nicht weiter nach unten schauen). Googles Modell beruht darüber hinaus auf registrierungsbasierter Interaktivität – das Unternehmen registriert die Präferenzen und die Geschichte der Nutzer, bewahrt sie auf und benutzt sie zunehmend dazu, maßgeschneiderte Resultate zu liefern.

Die unterschiedlichen Anordnungen von Quellen und Dingen in Suchmaschinen sind nicht ausreichend untersucht. Das liegt zu einem großen Teil



**Abb. 3** Sphären-Analyse: Häufigkeit, mit der bestimmte Tiere in Blog-Einträgen zum Klimawandel auftauchen

<sup>25</sup> Armanda Spink, Bernard J. Jansen, *Web Search: Public Searching on the Web*, Dordrecht (Kluwer) 2004.

darán, dass sie nicht gespeichert und für die Forschung zugänglich gemacht werden – abgesehen von den Daten, die AOL veröffentlicht hat, oder anderen ausverhandelten Vereinbarungen mit Suchmaschinenunternehmen. Google stellte zu einem bestimmten Zeitpunkt ein API (*application programming interface*) zur Verfügung, das Datensammlung möglich machte. Man konnte eine beschränkte Zahl von Anfragen pro Tag durchführen und die Resultate anschließend weiterverwenden. Google landete einen Scoop gegenüber Forschern, die sich auf die API verließen, als das Unternehmen den Service Ende 2006 beendete. 2009 wurde er in einer veränderten Form wieder eingeführt, doch nun betonte Google, dass automatisierte Suchabfragen und die permanente Speicherung von Ergebnissen durch die Geschäftsbedingungen untersagt sind. Wie kann man unter solchen Bedingungen Suchmaschinen studieren? Nun durchforsten wir Google, und veröffentlichen einen Kommentar, der unsere Wertschätzung für Googles Duldsamkeit zum Ausdruck bringt.<sup>26</sup>

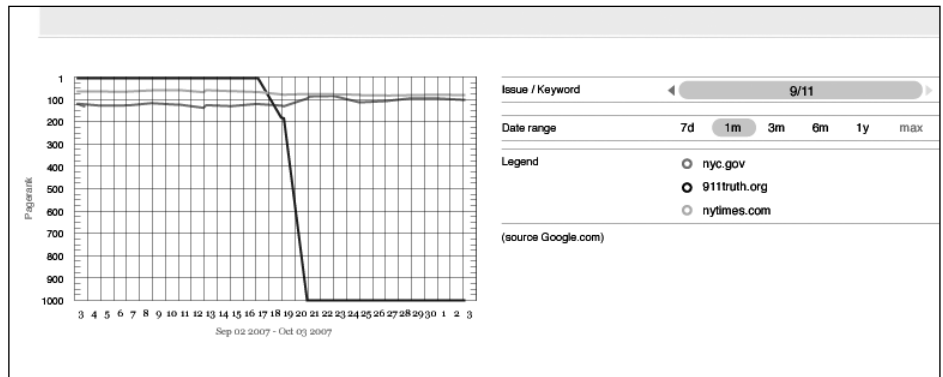
Was lässt sich in Googles Suchmaschinenresultaten finden? Wie ich schon bemerkt habe, sind Suchmaschinen ein entscheidender Punkt des Eintritts in das Netz. Sie sind epistemologische Maschinen in dem Sinn, dass sie Inhalte durchforsten, indexieren, speichern und schließlich ordnen. In einem früheren Text habe ich das Netz, und insbesondere ein auf Suchmaschinen basiertes Netz, als potenziellen Kollisionsraum für alternative Realitätsentwürfe beschrieben.<sup>27</sup> Diese Formulierung ging auf die Arbeit des Soziologen C. Wright Mills zurück, der den Anspruch an die Sozialforschung formulierte: «no less than to present conflicting definitions of reality itself».<sup>28</sup> Positionieren Suchmaschinen alternative Realitätsentwürfe nebeneinander oder entsprechen die Resultate dem Offiziellen und dem Mainstream? Die Speicherung und Analyse von Ergebnissen von Suchabfragen würde es erlauben, solche Fragen zu beantworten. Dies war das Projekt des Softwareprojekts namens *Issue Dramaturg*, das so benannt wurde, weil darin das potenzielle Drama zwischen den Topergebnissen zum Ausdruck kommt. Seiten können aufsteigen oder plötzlich abstürzen. Es ist wichtig, darauf hinzuweisen, dass Spitzenpositionen in Suchmaschinenergebnissen sehr gefragt sind; Organisationen greifen auf Optimierungstechniken zurück, um ihre Sichtbarkeit zu erhöhen.

Bei dem *Issue Dramaturg*-Projekt speicherte mein Team Suchmaschinenresultate von Google für die Suche <9/11> und weiterer Stichwörter. Dabei hatten wir zwei Dinge im Auge. Das erste betrifft die Hierarchisierung von Quellen, wie oben beschrieben. Welche <gewinnen> im Wettbewerb um die Position unter den Topresultaten für bestimmte Suchen? Das andere Vorhaben bestand darin, die Positionierung bestimmter Quellen zu erfassen. Diesen Versuch in Sachen Suchmaschinenforschung habe ich als <Quellendistanz> (*source distance*) bezeichnet. Wie weit sind zum Beispiel so signifikante Akteure wie die Stadtregierung von New York oder die *New York Times* bei einer Suche nach 9/11 von den obersten Ergebnissen entfernt? Sind solche Quellen prominent oder erscheinen sie in unmittelbarer Nachbarschaft zu Quellen, die offizielle

<sup>26</sup> Die Notiz findet sich auf der Creditseite des *Issue Dramaturg*, <http://issuedramaturg.issuercrawler.net/>, gesehen am 10.7.2011.

<sup>27</sup> Rogers, *Information Politics*.

<sup>28</sup> C. Wright Mills, *The Sociological Imagination*, Harmondsworth (Penguin) 1971, 212; Richard Rogers, Noor Marres, French scandals on the Web, and on the streets: A small experiment in stretching the limits of reported reality, in: *Asian Journal of Social Science*, 30. Jg., 2/2002, 339–353.



oder vertraute Ansichten hinterfragen? Neben der Stadtverwaltung von New York und der *New York Times* haben wir als weiteren Akteur das 9/11 Truth Movement genau beobachtet (911truth.org). Über Monate im Zeitraum zwischen März und September 2007 tauchte die Seite des 9/11 Truth Movements in den Top 5 der Resultate der Suchabfrage 9/11 auf, während die anderen beiden erst nach den ersten fünfzig Treffern kamen. Mitte September 2007, um den Jahrestag des Ereignisses herum, kam es zu einem Drama. 911truth.org fiel zunächst steil auf eine Position um 200 ab und schließlich aus den ersten 1.000 Ergebnissen raus – die maximale Ergebniszahl, die Google gibt. Wir gehen davon aus, dass das einer der ersten umfassend dokumentierten Fälle der offensichtlichen Entfernung einer Webseite durch Google ist – von einer Top 5-Platzierung über sechs Monate hinweg zu einem Ranking unterhalb der Top 1.000.<sup>29</sup> Der Fall führt zu Fragen nach der Stabilität und Volatilität von Suchmaschinen-ergebnissen und eröffnet ein neues Forschungsfeld.

**Abb. 4** Eine Website verschwindet: *911truth.org* ist offenkundig aus der Ergebnisliste zur Suchabfrage «9/11» bei *Google.com* entfernt worden.

### Soziale Netzwerke und Postdemografie

«We define social networking websites here as sites where users can create a profile and connect that profile to other profiles for the purposes of making an explicit personal network.»<sup>30</sup> So beginnt eine vom Pew Internet & American Life Project in Auftrag gegebene Untersuchung der Nutzung von Seiten wie MySpace und Facebook durch amerikanische Teenager. Dazu wurden Umfragen durchgeführt: 91 Prozent der Antwortenden nutzen die Seiten, um «Freundschaften zu managen». Weniger als ein Viertel nutzen sie zum «Flirten». Weitere führende Untersuchungen von sozialen Netzwerken im Web betreffen Themen wie Selbstpräsentation, Statusmanagement online, die unterschiedlichen «sozialen Klassen» von Usern auf MySpace und Facebook und die Beziehungen zwischen Freunden im richtigen Leben und «Friend»-Freunden.<sup>31</sup>

Ein anderer Arbeitsbereich, der häufig aus der Softwareentwicklung kommt, betrifft die Verwendung der massenhaft anfallenden Daten in Online-Profilen, vor allem zu Interessen und Vorlieben von Nutzern. Für diesen Arbeitsbereich

<sup>29</sup> Rogers, *Googlication*.

<sup>30</sup> Amanda Lenhart, Mary Madden, *Social Networking Websites and Teens*, Pew Internet, Washington, DC, 2007, <http://www.pewinternet.com>.

<sup>31</sup> Danah Boyd, Nicole Ellison, *Social network sites: Definition, history, and scholarship*, in: *Journal of Computer-Mediated Communication*, 13. Jg. 1/2007, <http://jcmc.indiana.edu/vol13/issue1/boyd.ellison.html>, gesehen am 10.7.2011.

el**f**riendo beta [FAQ](#) | [ABOUT](#)

YOUR NEW PROFILE BASED ON PSYCHOLOGY

General	psychology, music, philosophy, photography, writing
Music	mb, alternative, radiohead, r&b, kom, classical
Movies	action, seven, fight club, interview with the vampire
Books	interview with the vampire, the joy luck club, alfred adler
Status	single
Here for	friends, networking
Body type	5' 4" / slim / slender
Religion	other
Zodiac Sign	capricorn
Smoke / Drink	no / no
Children	someday
Education	in collage
Orientation	straight
Hometown	nashville
Income	000
Television	six feet under, mtv

**Abb. 5** Postdemografische Analyse: Analyse der Geschmacksbeziehungen von MySpace-NutzerInnen, die angeben, ein allgemeines Interesse an Psychologie zu haben

satz zu der herkömmlichen Weise, in der die Demografie Gruppen, Märkte und Wähler in einem soziologischen Sinn organisiert. Sie bezeichnet auch eine theoretische Abkehr von der <biopolitischen> Verwendung demografischer Forschung (zur Regierung von Körpern) und davon, wie Postdemografie <in-fopolitisch> eingesetzt wird, also dazu, bestimmte Informationen bestimmten Leuten zuzuspielen oder zu empfehlen.<sup>32</sup> Der Begriff Postdemografie ist auch offen für neue Methoden der Untersuchung sozialer Netzwerke, in denen die traditionellen demografischen Kategorien Ethnie, Alter, Einkommen und Bildungsniveau (oder Ableitungen davon wie Klasse) sukzessive durch Geschmack, Interessen, Vorlieben, Gruppen, angenommene Einladungen, installierte Apps und andere Informationen ersetzt werden, aus denen ein Online-Profil, und das damit einhergehende Drumherum, besteht.

Das Vorhaben der Postdemografie besteht in einem Beitrag zur Internetforschung, die von den Profilanalitikern und Rechercheuren lernt, die das Datenmaterial von sozialen Netzwerkseiten für weitere Analysen oder Softwareentwicklung, wie zum Beispiel Mash-ups, sammeln und abschöpfen.<sup>33</sup> Wie machen soziale Netzwerkseiten ihre Daten für Profiling zugänglich? Man kann sich zum Beispiel unter dem Entwicklermenü-Punkt bei Facebook einloggen und die Felder, die in dem API verfügbar sind, sehen. Beispielskripts werden zur Verfügung gestellt, wie zum Beispiel <Freunde von Nutzer Nummer x>, wobei man x in diesem Fall selber ist. Die zugänglichen Skripts folgen also der Kultur der Privatsphäre, das heißt, dass die Nutzer entscheiden, was beim Profiling zu sehen ist. Wenn viele Nutzer Zugang gewähren, indem sie <Ich stimme zu> bei den Applikationen Dritter anklicken, dann wird die Sache für das Profiling interessanter.

Eine andere Variante von Profilbildungspraktiken ist nicht an persönlichen Daten per se interessiert, sondern eher an Geschmack und vor allem Geschmacksbeziehungen. Man kann eine Menge Profilbildungsaktivitäten in die Kategorie der entpersonalisierten Datenanalyse stecken, dazu zählt Ama-

<sup>32</sup> Michel Foucault, *The History of Sexuality Vol.1: The Will to Knowledge*, London (Penguin) 1998; Rogers, *Information Politics*.

<sup>33</sup> Nicht-Nutzer verweisen auf Profiler. Profiler können natürlich auch Nutzer der Plattformen sein und in den meisten Fällen sind sie es wahrscheinlich auch, denn das Verständnis dafür, was sich auswerten lässt, wie es analysiert werden sollte und wie es sich für Mash-ups kombinieren ließe, erwächst aus der Nutzung, also aus einem zumindest minimalen Aktivitätsniveau.

zons wegweisendes Empfehlungssystem, bei dem es nicht sonderlich relevant ist, welche Person außerdem ein bestimmtes Buch gekauft hat, sondern nur, dass Leute das getan haben. Kundenkarten von Supermärkten und die Datenbanken, in denen Einkaufsgeschichten gespeichert werden, beruhen auch auf entpersonalisierter Informationsanalyse. Wie bei Amazon sind dabei die Anzahl der gekauften Gegenstände und die Beziehungen zwischen diesen Gegenständen wichtig (welche Chips mit welchem Softdrink). Populäre Produkte werden daraufhin gepusht. Bestimmte Produktkombinationen können im Regal vortsortiert werden.

Die bedeutendsten postdemografischen Maschinen sind natürlich, auch wenn sie sich selbst nicht als solche beschreiben, die Plattformen der sozialen Netzwerke selber. Sie sammeln die Vorlieben der Nutzer und präsentieren sie anderen – seien es andere Freunde oder Leute, die einfach nach anderen Leuten Ausschau halten, oder Profiler. Ich möchte hier kurz eine Software beschreiben, die mein Forschungsteam auf Grundlage der großen Sammelstelle MySpace geschrieben hat, und die postdemografischen analytischen Praktiken darlegen, die daraus resultierten.

Elfriendo.com ist das Ergebnis einer Reflexion über die Verwendung von Profilen auf der sozialen Netzwerkplattform MySpace. Bei Elfriendo.com gibt man ein einzelnes Interesse ein und das Programm erstellt ein neues Profil auf Basis von Profilen anderer Leute, die ebenfalls dieses Interesse eingetragen haben. Man kann die Kompatibilität von Interessen vergleichen, also ob ein oder mehrere Interessen, Musikstücke, Filme, Fernsehsendungen, Bücher und Helden mit anderen kompatibel sind. Elfriendo.com beantwortet solche Fragen durch die Analyse einer Reihe von befreundeten Profilen und des Vergleichs von Interessen, die auf diesen Profilen genannt werden. Ein Film, eine Fernsehsendung etc. haben damit ein zusammengesetztes Profil, das aus weiteren Interessensnennungen besteht. (Der Rapper Eminem taucht im Februar 2009 sowohl in den zusammengesetzten Profilen zu Christentum und Islam auf.) Man kann mit dem Programm sogar eine postdemografische Analyse durchführen und bekommt dabei einen Eindruck von den relationalen Vorlieben, ähnlich wie sie in sozialen Netzwerken analysiert werden.<sup>34</sup>

Hier ist es wichtig festzuhalten, dass MySpace leichter zugänglich ist als Facebook. Es bildet keinen so exklusiven <ummauerten Garten>, denn es erlaubt dem Profiler, die Profile der Freunde eines Users (und die Profile von deren Freunden) zu sehen, ohne selbst mit ihnen befreundet zu sein. Man kann zum Beispiel alle Freunde von Barack Obama und John McCain sehen und die Profile ihrer jeweiligen Freunde analysieren. Die Software zählt die Gegenstände, Themen und Namen, die die Freunde unter Interessen, Musik, Filme, Fernsehsendungen, Bücher und Helden aufgelistet haben. Was ergibt sich aus dieser Zählpraxis von Geschmacksbeziehungen? Die Resultate ergeben klare Bilder der Anhänger der beiden Präsidentschaftskandidaten 2008. So zum Beispiel, dass das Kompatibilitätslevel zwischen den Interessen der Freunde der

<sup>34</sup> Man bekommt nur einen ersten Eindruck davon, wie diese Analyse durchgeführt wird, und von den Erkenntnissen, die gewonnen werden können, denn Elfriendo.com erfasst nur die Top-100-Profile. Es erlaubt damit nur erste Hinweise, während ein tatsächlicher Auswertungsprozess stärker fundierte Erkenntnisse ergeben würde.

beiden Kandidaten insgesamt niedrig ist. Die beiden Gruppen teilen wenige Interessen. Die Vorlieben der Freunde der Kandidaten sind bei Filmen, Musik, Büchern und Helden nicht kompatibel. Bei den Fernsehsendungen liegt die Kompatibilität allerdings bei 16 Prozent. Es scheint spezifische Medienprofile für bestimmte Gruppen der Freunde der Kandidaten zu geben. Freunde von Obama schauen die *Daily Show*, während Freunde von McCain *Family Guy*, *Top Chef* und *America's Next Top Model* schauen. Beide Gruppen schauen *Lost*. Die Ergebnisse lassen sich in Begriffen der Wählerpostdemografie diskutieren. Die Beschreibung von Wählerprofilen beruht auf Medienvorlieben und nicht auf Bildungsniveaus, Einkommen und anderen Standardindikatoren.

### **Schlussfolgerung. Das Ende des Virtuellen – Grounding-Ansprüche online**

Mein Ziel ist es, eine Transformation in Gang zu setzen, die das Was und Warum von Forschung unter Verwendung des Internets betrifft. Der erste Schritt besteht darin, die Diskussion von den Beschränkungen des Virtuellen (Wie viel von Gesellschaft und Kultur ist online?) wegzulenken und auf die Beschränkungen der gegenwärtigen Methode (Wie können mit dem Internet Kultur und Gesellschaft untersucht werden?) hinzulenken.

Ich würde gerne mit einer kurzen Reflexion dieser Beschränkungen in der Internetforschung enden und einen Vorschlag für eine Erneuerung vorbringen. Erstens ist das Ende des Cyberspace und seine Ortlosigkeit (wie auch das Ende des Virtuellen als ein eigener Bereich) beklagenswert für bestimmte Forschungsansätze. Die Kluft zwischen dem Realen und dem Virtuellen war auf eine gewisse Weise für spezifische Forschungspraktiken hilfreich.<sup>35</sup> Das Konzept Cyberspace verlieh beispielsweise dem Projekt einer umfassenden Archivierung von Webseiten Plausibilität, während die zunehmende Verankerung des Web im Offline-Leben und die national definierten Webseiten das Archiv schrumpfen lassen.

Ich habe darauf Wert gelegt, dass man von den Methoden lernen kann, die in dem Medium selbst Anwendung finden. Dadurch verschob sich die Diskussion einer Theorie der Medienspezifität von der Ontologie (Eigenschaften und Features) auf die Epistemologie (Methode). Das Internet und – spezifischer – das Web haben ihre ontologischen Objekte, wie den Link oder den Tag. Webepistemologie besteht unter anderem in der Untersuchung, wie diese genuin digitalen Gegenstände von Netzanwendungen behandelt werden. Die Einsichten, die aus solchen Untersuchungen gewonnen werden können, führen zu wichtigen methodologischen Unterscheidungen und auch zu Erkenntnissen über die Zwecke von Internetforschung. Was die methodologische Unterscheidung anlangt, kann man heutige Internetmethoden in solche einteilen, die dem Medium folgen (und den dominanten Techniken, die in Erstellung und Ordnung von Information, Wissen und Gesellschaftlichkeit Anwendung finden), und in solche, die bestehende Methoden remediatisieren oder digitalisieren.

<sup>35</sup> Das Ende des Cyberspace war auch nicht hilfreich für Projekte, die auf dem klassischen Internetfeature des anonymen Users beruhen. Zum Beispiel verbieten es Organisationen und Regierungen ihren Mitarbeitern, während der Arbeitszeit an Wikipedia-Einträgen zu arbeiten, denn diese redaktionellen Eingriffe lassen sich örtlich zurückverfolgen und könnten zu Skandalen führen.



Die Methodenunterschiede können zu gewichtigen Ergebnissen führen. Ein Grund für die ungedeihliche Entwicklung des Webarchivierens könnte in der Wahl einer digitalisierten Methode (redaktionelle Auswahl) gegenüber einer digitalen (automatische Datenerfassung) liegen, wie sie bei dem ursprünglichen Internet Archive-Projekt zur Anwendung kam, bei dem Seiten gespeichert wurden, die von Usern beim Surfen erreicht wurden. Ich habe den Begriff <digitale Methoden> so verwendet, dass Forscher den Wert und die Ergebnisse des einen Ansatzes mit dem anderen vergleichen können. Der Vorzug der dynamischen URL-Erfassung gegenüber dem redaktionellen Modell erweist sich dabei als der Erforschung von Internetzensur zuträglich, wie ich erörtert habe.

Drittens und abschließend habe ich dargelegt, dass das Internet ein Ort der Forschung ist, der über Onlinekultur und ihre User hinausgeht. Mit dem Ende der Kluft zwischen dem Realen und dem Virtuellen lässt sich das Internet als Quelle von Daten über Gesellschaft und Kultur neu denken, wozu diese im Detail auch nützlich sein mögen. Dies erfordert nicht nur eine Neuperspektivierung des Internets, sondern auch der Methoden, die eine Verankerung und <Erdung> der Forschungsergebnisse möglich machen. Die Zielsetzungen der Internetforschung erleben eine markante Verschiebung, wenn online Erkenntnisse gewonnen werden, die auch jenseits des Internets Relevanz beanspruchen: Erforscht werden nicht mehr ausschließlich das Internet und seine Nutzer; vielmehr können Kultur und Gesellschaft *mit dem Internet* erforscht werden.

---

Aus dem Englischen von Berthold Rebhandl