

Big Data: Medienkultur im Umbruch

Ramón Reichert

Zusammenfassung

Das Schlagwort „Big Data“ ist in aller Munde – und beschreibt nicht nur wissenschaftliche Datenpraktiken, sondern steht auch für einen gesellschaftlichen Wandel und eine Medienkultur im Umbruch. Welche Einflüsse hat Big Data auf die Gegenwartskultur und ihre Machtverschiebungen? Mit der Annahme, dass die digitalen Medien und Technologien nicht einfach nur neutrale Botschaften übertragen, sondern ein kulturelles Gedächtnis etablieren und eine soziale Wirkmächtigkeit entfalten, können sie als Selbstverständigungsdiskurse der Gesellschaft verstanden werden. Welche Medien und Wissenstechniken ermöglichen folglich die Modellierungen der Big Data und welche Narrative, Bilder und Fiktionen sind hierbei beteiligt?

Wenn wir vor diesem Hintergrund akzeptieren, dass Soziologen in den Konstruktionsprozess der Daten involviert sind und es von daher keine objektivierende Distanz zum Datenmaterial geben kann, dann kann auch im sich wechselseitig bedingenden Verhältnis gefragt werden, auf welche Weise Mediendispositive, technische Infrastrukturen und kulturelle Kontexte die digitale Wissensproduktion und ihre Wissensansprüche formieren. Ich gehe in meinem Beitrag aber auch von einer kulturalistischen These aus und versuche zu argumentieren, dass der Big-Data-Ansatz nicht ausschließlich für eine digitale Wende bei der Objektivierung kollektiver Praktiken steht, sondern selbst in historische Wissens-, Medien-, Bild- und Erzählkulturen eingelagert ist.

“Every day, people are breaking up and entering into relationships on Facebook. When they do, they play songs that personify their mood. With Valentine’s Day just around the corner, we looked at the songs most played by people in the U.S. on Spotify as they make their relationships and breakups ‘Facebook official’.”
(Facebook Data Team 2012)¹

Einführung

In öffentlichen Debatten ist bereits viel spekuliert worden, auf welche Weise Soziale Netzwerke die Zukunft ihrer Mitglieder vorhersehen und planen können. Diese Frage kann jedoch ohne Rekurs auf die Dominanz der angewandten Mathematik und der Medieninformatik nicht ausreichend beantwortet werden. Denn beide Praxis- und Wissensfelder haben mit ihren stochastischen Analysetechniken von Nutzeraktivitäten die digitale Vorhersagekultur der Sozialen Medien im Web 2.0 erst ermöglicht, die es früher in diesem Ausmaß und Machtanspruch noch nicht gegeben hat.

¹ <https://www.facebook.com/data/posts/10152217010993415#!/data?fref=nf>

Vor diesem Hintergrund versucht mein Aufsatz, die neuen Datenstrukturen methodologisch zu betrachten, um einerseits den Positivismus der sehr großen Datenanalysen infrage zu stellen und zweitens den Stellenwert von Theorie in der Onlineforschung kritisch zu sondieren. Infolgedessen plädiere ich für eine *datenkritische* Perspektivierung der Erforschung großer Datensätze und frage nach den sozialen und kulturellen Auswirkungen, die sich durch die Transformationen des wissenschaftlichen Wissens ergeben könnten.

In allen Bereichen der digitalen Internetkommunikation werden heute große Datenmengen (Big Data) generiert:

“More business and government agencies are discovering the strategic uses of large databases. And as all these systems begin to interconnect with each other and as powerful new software tools and techniques are invented to analyze the data for valuable inferences, a radically new kind of ‘knowledge infrastructure’ is materializing.” (Bollier 2010, S. 3)

In der Ära der Big Data hat sich der Stellenwert von Sozialen Netzwerken radikal geändert, denn sie figurieren zunehmend als gigantische Datensammler für die Beobachtungsanordnungen sozialstatistischen Wissens und als Leitbild normalisierender Praktiken. Als Schlagwort steht Big Data für die Überlagerung eines statistisch fundierten Kontrollwissens mit einer medientechnologisch fundierten *Makroorientierung* an der ökonomischen Verwertbarkeit von Daten und Informationen. Die großen Datenmengen werden in verschiedenartigen Wissensfeldern gesammelt: Biotechnologie, Genomforschung, Arbeits- und Finanzwissenschaften, Risiko- und Trendforschung berufen sich in ihren Arbeiten und Studien auf die Ergebnisse der Informationsverarbeitung der Big Data und formulieren auf dieser Grundlage aussagekräftige Modelle über den gegenwärtigen Status und die künftige Entwicklung von sozialen Gruppen und Gesellschaften. In den meisten Fällen geht es bei der Erforschung sehr großer Datenmengen um die Aggregation von Stimmungen und Trends. Diese Datenanalysen und -visualisierungen versammeln aber in der Regel nur faktische Gegebenheiten und lassen die Frage nach den sozialen Kontexten und Motiven außer Acht. Dessen ungeachtet hat sich der Big-Data-Ansatz in den Human-, Sozial- und Kulturwissenschaften mittlerweile etablieren können. Ich möchte im Folgenden die soziale Relevanz der Big-Data-Analyse exemplarisch problematisieren.

Social Media Monitoring

Im Forschungsfeld der *Social Media Data* hat sich mit der Gesundheitsprognostik eine evidenzbasierte Praxis der Prävention herausgebildet, die auf die institutionelle Entwicklung der staatlich-administrativen Gesundheitsvorsorge und auf die Kulturtechniken der Lebensführung Einfluss nimmt. Die Gesundheitsvorsorge beobachtet mit großem Interesse, dass weltweit Millionen von Nutzerinnen und Nutzern täglich mit der Internet-Suchmaschine Google Informationen zum Thema Gesundheit suchen. In Grippezeiten häufen sich die Suchanfragen zur Grippe und die Häufigkeit bestimmter Suchbegriffe kann Anhaltspunkte für die Häufigkeit von Grippeerkrankungen liefern. Studien zum Suchvolumenmuster haben herausgefunden, dass ein

signifikanter Zusammenhang zwischen der Anzahl von grippebezogenen Suchanfragen und der Anzahl von Personen mit tatsächlichen Grippe-symptomen besteht (Freyer-Dugas et. al. 2012 S. 463–469). Dieses epidemiologische Beziehungsgefüge kann zur Frühwarnung vor Epidemien auf Städte, Regionen, Länder und Kontinente ausgedehnt und differenziert dargestellt werden. Mit der epidemiologischen Auswertung von textuellen Clustern und semantischen Feldern erhält das Social Web den Status einer großen Datenbank, die das soziale Leben in seiner Gesamtheit widerspiegelt und damit eine repräsentative Datenquelle für die präventive Gesundheitspolitik darstellt. Die Kommunikationsprozesse in Online-Netzwerken stehen im Fokus staatlicher Biopolitik, die um die Gesundheit der Bevölkerung besorgt ist und spezifische Wissenstechniken und -modelle zur Erforschung der Big Data entwickelt hat, um die Wahrscheinlichkeit der Verbreitung von Krankheiten in absehbarer Zukunft statistisch zu schätzen.

Die Mehrzahl der Monitoring-Projekte, die große Datenmengen im Social Web untersuchen, wird von Computerlinguistinnen bzw. -linguisten und Informatikerinnen bzw. Informatikern durchgeführt. Generell interpretieren sie die Kommunikation als kollektiv geteilte und kulturspezifische Wissensstrukturen, mit denen Individuen versuchen, ihre Erfahrungen zu interpretieren. Die Erhebung dieser Wissensstrukturen verfolgt den Anspruch, einen sozial differenzierten Einblick in öffentliche Debatten und sozial geteilte Diskursnetze zu erhalten. Die Wissensstrukturen werden hierbei mit Hilfe eines korpuslinguistischen Ansatzes erschlossen. Am Beginn der Forschung steht die Erstellung eines digitalen Korpus, der sich aus begrifflichen Entitäten zusammensetzt, die in der Regel als ‚kanonisch‘ eingestuft werden. So ergeben sich einige Hypothesen erst aus der empirischen Widerständigkeit der Big Data und entwickeln sich erst im Fortgang ihrer Beschreibung. Die Kategorienkataloge suggerieren damit zwar auf den ersten Blick wissenschaftliche Objektivität, andererseits bleibt angesichts der riesigen Datenmengen eine genaue Validierung der Begriffsauswahl, d.h. der interpretativen Selektion der Big Data, oft unklar und vage. Diese Unsicherheit bei der Hypothesenbildung liegt darin begründet, dass das umfangreiche Datenmaterial in keiner Gesamtschau mehr überblickt werden kann und daher auch nicht mehr linguistisch kodiert werden kann. Oft ist die erhobene Datenmenge so umfangreich, dass nach einer ersten Sondierung des Materials weitere Gewichtungen und Einschränkungen zur Komplexitätsreduktion vorgenommen werden müssen. An dieser methodischen Einschränkung des Big-Data-Monitoring wurde kritisiert, dass die erarbeiteten Erkenntnisse nur ein atomistisches Bild der Daten liefern können und daher auf eine Kontextualisierung des Textmaterials und damit auf eine kontextsensitive Interpretation des Zeichengebrauchs weitgehend verzichten müssen (vgl. Boyd und Crawford 2011). Der Vorteil der Dekontextualisierung bei der nach Worthäufigkeiten fahndenden Big-Data-Analyse besteht nach Manovich (vgl. 2012, S. 465) darin, dass die einzelnen Worteinheiten auf eine enthierarchisierte und dezentrale Repräsentation des Wissens hinauslaufen und damit die Möglichkeit alternativer kollektiver Äußerungsgefüge anbieten.

Die Auswertung der Daten der Google-Suche kann auf andere Trendentwicklungen erweitert werden. Mittlerweile gibt es zahlreiche Studien, welche die textuellen Daten der Sozialen Medien untersuchen, um *politische Einstellungen* (vgl. Conover et. al. 2011), *Finanztrends* und *Wirtschaftskrisen* (vgl. Gilbert und Karahalios 2010), *Psychopathologien* (vgl. Wald, Khosh-

goftaar und Sumner 2012, S. 109–115) und *Aufstände* und *Protestbewegungen* (vgl. Yogatama 2012) frühzeitig vorherzusagen. Von einer systematischen Auswertung der Big Data erwarten sich die Prognostikerinnen und Prognostiker eine effizientere Unternehmensführung bei der statistischen Vermessung der Nachfrage- und Absatzmärkte, individualisierte Serviceangebote und eine bessere gesellschaftliche Steuerung. Einen großen politischen Stellenwert hat vor allem die algorithmische Prognostik kollektiver Prozesse. In diesem Konnex ist das Social Web zur wichtigsten Datenquelle zur Herstellung von Regierungs- und Kontrollwissen geworden. Die politische Kontrolle sozialer Bewegungen verschiebt sich hiermit in das Netz, wenn Soziologinnen bzw. Soziologen und Informatikerinnen bzw. Informatiker gemeinsam etwa an der Erstellung eines *Riot Forecasting* mitwirken und dabei auf die gesammelten Textdaten von Twitter-Streams zugreifen:

“Due to the availability of the dataset, we focused on riots in Brazil. Our datasets consist of two news streams, five blog streams, two Twitter streams (one for politicians in Brazil and one for general public in Brazil), and one stream of 34 macroeconomic variables related to Brazil and Latin America.” (Ebd., S. 3)

Big Data bietet eine spezifische Methode und Technologie zur statistischen Datenauswertung, die aus der epistemischen Schnittstelle von Wirtschaftsinformatik und kommerzieller Datenbewirtschaftung hervorgeht und die Bereiche der *Business Intelligence*, des *Data Warehouse*² und des *Data Mining*³ in sich vereint. Die Diskussion um den technologisch-infrastrukturellen und machtstrategischen Stellenwert der Big Data zeigt auf, dass die numerische Repräsentation von Kollektivitäten zu den grundlegenden Operationen digitaler Medien gehört und eine rechnerbasierte Wissenstechnik bezeichnet, mit welcher kollektive Praktiken mathematisch beschreibbar und auf diese Weise quantifizierbar werden. Die Bestimmung der Vielheiten mit Hilfe von numerisch gegliederten Mengenangaben dient in erster Linie der Orientierung und kann als eine Strategie verstanden werden, die kollektive Datenströme in lesbare Datenkollektive übersetzt. In diesem Sinne firmieren in der medialen Öffentlichkeit soziale Netzmedien wie Facebook, Twitter und Google+ als Spiegel der allgemeinen Wirtschaftslage (vgl. Bollen, Mao und Zeng 2011, S. 1–8) oder als prognostischer Indikator von nationalen Gefühlsschwankungen (vgl. Bollen 2011, S. 237–251). In diesem Sinn bilden sie selbst Schauplätze einer populären Aufmerksamkeit und popularisierender Diskurse, die ihnen bestimmte Außenwir-

² Das Data Warehousing ist eine infrastrukturelle Technologie, die zur Auswertung großer Datenbestände dient.

³ Im kommerziellen Bereich etablierte sich der Begriff *Data Mining* für den gesamten Prozess des *Knowledge Discovery in Databases*. *Data Mining* meint die Anwendung von explorativen Methoden auf einen Datenbestand mit dem Ziel der Mustererkennung. Ziel der explorativen Datenanalyse ist über die Darstellung der Daten hinaus die Suche nach Strukturen und Besonderheiten. Sie wird daher typischerweise eingesetzt, wenn die Fragestellung nicht genau definiert ist oder auch die Wahl eines geeigneten statistischen Modells unklar ist. Ihre Suche umfasst, ausgehend von der Datenselektion, alle Aktivitäten, die zur Kommunikation von in Datenbeständen entdeckten Mustern notwendig sind: Aufgabendefinition, Selektion und Extraktion, Vorbereitung und Transformation, Mustererkennung, Evaluation und Präsentation.

kungen – etwa als ein Gradmesser der konjunkturellen Entwicklung der Wirtschaft und der sozialen Wohlfahrt – zuschreiben.

Welche Musik werden eine Milliarde Menschen in Zukunft hören, wenn sie frisch verliebt sind und welche Musik werden sie hören, wenn sie gerade ihre Beziehung beendet haben? Diese Fragestellungen hat das „Facebook Data Team“ im Jahr 2012 zum Anlass genommen, um die Daten von über einer Milliarde Nutzerinnen- und Nutzerprofilen (mehr als 10 Prozent der Weltbevölkerung) und 6 Milliarden Songs des Online-Musikdienstes Spotify mittels einer korrelativen Datenanalyse auszuwerten, die den Grad des gleichgerichteten Zusammenhangs zwischen der Variable „Beziehungsstatus“ und der Variable „Musikgeschmack“ ermittelt.⁴ Diese Prognose über das kollektive Konsumverhalten basiert auf Merkmalsvorhersagen, die mittels Data Mining in einer simplen Kausalbeziehung ausgedrückt werden. Unter Leitung des Soziologen Cameron Marlow erforschte die aus Informatikerinnen bzw. Informatikern, Statistikerinnen bzw. Statistikern und Soziologinnen bzw. Soziologen bestehende Gruppe das statistische Beziehungsverhalten der Facebook-Nutzerinnen und -Nutzer und veröffentlichte am 10. Februar des gleichen Jahres zwei Hitlisten von Songs, die Nutzerinnen und Nutzer hörten, als sie ihren Beziehungsstatus änderten, und nannte sie lapidar „Facebook Love Mix“ und „Facebook Breakup Mix“.⁵ Die Forschergruppe im Back-End destillierte aus der statistischen Ermittlungsarbeit der „Big Data“ (vgl. Wolf et.al. 2011, S. 217–219) nicht nur eine globale Verhaltensdiagnose, sondern transformierte diese auch in eine suggestive Zukunftsansage. Sie lautete: Wir Forschende im Back-End bei Facebook wissen, welche Musik eine Milliarde Facebook-Nutzerinnen und -Nutzer am liebsten hören werden, wenn sie sich verlieben oder trennen.⁶ Unter dem Deckmantel des bloßen Sammelns und Weitergebens von Informationen etabliert die Forschergruppe des „Facebook Data Teams“ eine Deutungsmacht gegenüber den Userinnen und Usern, indem sie die Nutzerinnen und Nutzer im automatisch generierten Update-Modus „What’s going on?“ auffordert, regelmäßig Daten und Informationen zu posten.

Die Zukunftsansagen des „Facebook Data Teams“ sind jedoch nur vordergründig mathematisch motiviert und verweisen auf den performativen Ursprung des Zukunftswissens. Trotz fortgeschrittener Mathematisierung, Kalkülisierung und Operationalisierung des Zukünftigen bezieht das Zukunftswissen seine performative Macht immer auch aus Sprechakten und Ansageordnungen, die sich in literarischen, narrativen und fiktionalen Inszenierungsformen ausdifferenzieren können. In diesem Sinne sind die Bedeutungen im Möglichkeitsraum der Zukunft nicht eindeutig determiniert, sondern erweisen sich vielmehr als ein *aggregatähnliches Wissen*, dessen konsenserzwingende Plausibilität sich nicht in *Wahrheitsdiskursen* und *epistemischen Diskursen* erschöpft, sondern auch von *kulturellen* und *ästhetischen* Kommunikati-

⁴ Facebook Data Science, <https://www.facebook.com/data> (letzter Zugriff: 28.12.2013).

⁵ Unter dem Titel „Facebook Reveals Most Popular Songs for New Loves and Breakups“ äußerte sich „Wired“ begeistert über die neuen Möglichkeiten des Data Minings: www.wired.com/underwire/2012/02/facebook-love-songs/ (letzter Zugriff: 28.12.2013).

⁶ Die kollektive Figur „Wir“ meint in diesem Fall die Forscher im Backend-Bereich und hat futurologische Verschwörungstheorien angeheizt, die das Weltwissen in den Händen weniger Forscherinnen und Forscher vermuten.

onsprozessen und Erwartungshaltungen (*patterns of expectation*) gestützt wird, die Imaginäres, Fiktives und Empirisches in Beziehung setzen.

Das Format der Hitliste und ihrer beliebtesten zehn Songs versucht, durch Vereinfachung komplexe Sachverhalte auf einen Blick darstellbar zu machen. Es handelt sich um ein popularisierendes Zukunftsnarrativ, das eine verhaltensmoderierende, repräsentative und rhetorische Funktion übernehmen und die Zukunftsforschung als unterhaltsame und harmlose Tätigkeit herausstreichen soll. Um in diesem Sinn glaubwürdig zu sein, muss die futurische Epistemologie immer auch auf eine gewisse Weise überzeugend in Szene gesetzt werden, sie muss theatralisch überhöht und werbewirksam inszeniert und erzählt werden, damit sie Aufmerksamkeit generieren kann. Insofern ist den futurischen Aussageweisen immer auch ein Moment der prophetischen Selbst- und Wissensinszenierung inhärent, mit dem die wissenschaftlichen Repräsentantinnen und Repräsentanten den gesellschaftsdiagnostischen Mehrwert der Sozialen Netzwerke unter Beweis stellen wollen (vgl. Doorn 2010, S. 583–602). Soziale Netzmedien agieren heute als Global Player der Meinungsforschung und der Trendanalyse und spielen eine entscheidende Rolle bei der Modellierung von Zukunftsaussagen und futurologischer Wissensinszenierung.

Front End – Back End

Eine der Grundthesen meines Beitrags basiert auf der Annahme, dass das im vorigen Kapitel erörterte Metawissen der Sozialen Netzwerke auf einer asymmetrisch verlaufenden Machtbeziehung beruht, die sich in die technisch-mediale Infrastruktur verlagert hat und das bipolare Schema von Front End und Back End hervorgebracht hat, das in diesem Kapitel untersucht werden soll. In diesem Sinne gehe ich davon aus, dass Facebook nicht nur die zeitgemäßen Anforderungen für Subjektivierungsprozesse in der Ära der neuen Vernetzungskultur markiert, sondern eine im Back-End-Bereich angesiedelte Wissenstechnik, die auf der Grundlage der Daten und Informationen der Facebook-Mitglieder bestimmte Registrierungs-, Klassifizierungs-, Taxierungs- und Ratingverfahren entwickelt.

In unternehmerischer Hinsicht kann Facebook als eine konzernkontrollierte soziale Medienplattform verstanden werden (vgl. Leistert/Röhle 2011, S. 9f.). Darunter versteht man ein digitales Anwendungssystem, das seinen Nutzerinnen und Nutzern Funktionalitäten zum Identitätsmanagement – zur Darstellung der eigenen Person in Form eines Profils zur Verfügung stellt und darüber hinaus die Vernetzung mit anderen Nutzerinnen und Nutzern und damit die Verwaltung eigener Kontakte – ermöglicht.

Um personenzentriertes Wissen über die Userinnen und User herzustellen, offeriert die Social Software von *Facebook* standardisierte E-Formulare für Subjekte, die sich in Selbstbeschreibungs-, Selbstverwaltungs- und Selbstausswertungsprozeduren eigenständig organisieren sollen. Diese elektronischen Formulare sind tabellarisch angeordnete Rastergrafiken, die aus logisch vorstrukturierten Texten mit *slot*- und *filler*-Funktionen bestehen; sie sollen eine einheitliche, standardisierte Inventarisierung, Verwaltung und Repräsentation des Datenmaterials ermöglichen. Die grafische Rasterform und die determinierende Vereinheitlichung der Tabellenfelder

etablieren Standards der informationellen Datenverarbeitung: Sie machen aus den personen-zentrierten Darstellungs- und Erzählformen des Front-End-Bereichs einheitliche Informationsbausteine, die der formallogischen Verarbeitung der Datenbanksysteme im Back-End-Bereich zugeführt werden.

Bereits beim Anlegen eines Accounts haben sich künftige Facebook-Mitglieder einer vielschichtigen Prozedur der Wissenserfassung zu unterwerfen. Bei der Erstanmeldung auf der Internetplattform Facebook werden die Systemnutzerinnen und -nutzer aufgefordert, ihre persönlichen Daten in standardmäßig vorgegebene Erfassungsmasken und dokumentspezifische Datenfelder, die gleichzeitig als Orientierungsinstanzen fungieren, einzutragen. Die Darstellung von Profilen in Tabellenform signalisiert Überblick und liefert ein einfaches Raster der Personenbeschreibung, das zum gemeinsamen Referenzpunkt der Betrachtung, Beurteilung und Entscheidung werden kann (vgl. zur Veralltäglichen von Testverfahren Lemke 2004, S. 123). Der kategoriale Ordnungsanspruch der Selbstthematisierung wird im Enduser-Interface nicht sprachlich, sondern grafisch vollzogen. Das elektronische Datenblatt wird somit zu einer Instanz, welche die Transparenz des Überblickes herstellt und Entscheidungen standardisiert.

Mit seinen vorstrukturierten Applikationen erstellt Facebook spezifische Gestaltungsimperative der Wissenserfassung und -repräsentation persönlicher Daten. Elektronische Formulare stellen Formansprüche, die zunächst die Autorinnen und Autoren betreffen. So können bestimmte Einträge nur auf eine bestimmte Art und Weise vorgenommen werden. Die grafische Autorität setzt sich aus sowohl qualitativen als auch quantitativen Kriterien zusammen und diktiert nicht nur die inhaltlichen Kategorien der Selbstbeschreibung, sondern fordert auch das *vollständige* Ausfüllen des Formulars, mit welchem erst der Vorgang abgeschlossen werden kann. Um im Raster der E-Formulare verortet werden zu können, muss lineares und narratives Wissen in Informationsbausteine zerlegt werden. Diese formimmanente Regeln begründen die Autorität des E-Formulars. Es handelt sich jedoch um eine brüchige und instabile Autorität der grafischen und logischen Struktur, die im *Formularegebrauch* permanent unterwandert werden kann (etwa durch das Erstellen von Fake-Profilen). Die kulturelle Einbettung der Kompilatorinnen und Kompilatoren in historisch bedingte und sozial differenzierte Lektüre-, Schreib-, Erzähl- und Wahrnehmungspraktiken relativiert die expliziten Anweisungen, Belehrungen und Direktiven der formimmanente Regelfassung des E-Formulars. Es gibt also unter allen Umständen eine Vielzahl taktischer Möglichkeiten, ihr formimmanentes Diktat zu unterlaufen. Die Praktiken der Kompilation stellen das Gesamtschema der elektronischen Formulare infrage und können somit widersprüchliche Dateneinträge produzieren.

Es gibt jedoch eine weitere Ausprägung des *Formularegebrauchs*. Sie bietet die Basis stillschweigender Akzeptanz des Formulars als einer adäquaten Form der Aufzeichnung und Darstellung von Daten und Informationen. Die Autorität der elektronischen Wissenserfassung und -repräsentation hängt folglich auch von der *Bereitschaft* der Enduserinnen und -user ab, die Benutzung der elektronischen Formulare als neutral, evident und selbsterklärend anzuerkennen. Einer solchen Akzeptanz liegen historische Lese- und Schreibgewohnheiten (Buchhaltung, Prüfungs- und Testverfahren) zugrunde, die dazu führen, dass die elektronischen Formu-

lare als *gebräuchliche* und *geläufige* Wissensmanuale der empirischen Datenermittlung wiedererkannt werden. Insofern stützt sich die Formautorität elektronischer Formulare weniger auf die individuelle Autorisierung des Dokumentes durch Institutionen, sondern ist von der kulturellen Akzeptanz der Form abhängig. Die fraglose Überzeugungskraft des Formulars basiert wesentlich auf historisch gemachten Erfahrungen mit dieser Form, d.h. dem *Wiedererkennen* der Form.

Mit der *slot-* und *filler-*Funktion des Status-Update und seiner Aufforderung „What’s on your mind?“ wird zusätzlich eine restriktive Wirkung auf narrativ-literarische Selbstbeschreibungen ausgebildet. Anhäufungen von Namen, Halbsätzen und Daten dominieren die Profilauster und bilden Abkürzungen, die letztlich stets auf andere Texte und Kontexte verweisen. Die persönlichen Angaben, Statusmeldungen und Kommentare sind denn auch in weiten Teilen enumerativ. Sie referieren Daten, kumulieren Fundstücke und bieten nur selten stringente Narrative. Die Tabellenform der elektronischen Profile trennt die einzelnen Textbausteine durch ein Strichnetz grafisch ab, isoliert die Texte und weist ihnen mittels Zeilen und Spalten spezifische Leitkontexte zu. Die E-Formulare machen mehr als einen bloßen Kontext der Information verfügbar, denn sie setzen ein *Raster der Erfassung* persönlicher Merkmale ins Werk. Mit diesem Rasterwerk können die Profile miteinander verglichen und einzelne Parameter bestimmten Suchanfragen zugeordnet werden. Die von den Sozialen Netzwerkseiten programmierte Rasterfahndung nach biografischer Information meint in diesem Zusammenhang, dass das Frageprotokoll und die Auskunftgeberinnen und -geber in einem hierarchischen Bezugsrahmen zueinander positioniert sind: Das Dokument soll als solches unveränderlich bleiben und von jedem Einzelnen beliebig oft benutzt werden können. *Durch das schriftlich vorformulierte und grafisch strukturierte Frageschema hat sich die Autorität in den Bereich der Form verschoben*, d.h. dass die formimmanente Festlegung der Reihenfolge und die Art und Weise der Fragen einen kommunikativen Rahmen absteckt.

Das elektronische Formular kann Potenziale formaler Autorität entwickeln, wenn der Umgang mit ihm streng geregelt wird und Verstöße sanktioniert werden. Seine formale Strenge gewährleistet, dass das E-Formular geplante Informationen vorhersehbar übertragen kann. Die immanente Formautorität markiert jedoch keinen feststehenden Status, sondern bleibt anfällig für das Entstehen einer informellen Regelausweitung, die mittels der *kulturellen Praktiken* im Bereich der Anwendung entsteht. Die Formulare der Wissenserfassung zielen zwar darauf ab, den Zufluss an Daten schon im Vorfeld zu standardisieren und zu kategorisieren, doch je weiter sich das Raster in die Grafik der Formalität zurückzieht, d.h. *nicht explizit* wird, desto erfolgreicher scheint es bei der Hervorbringung *intimer Bekenntnisse* zu sein. Es sind die scheinbar „zwanglosen“ Datenfelder, die informellen Spielraum zur persönlichen Gestaltung aufweisen und zur Führung individueller Listen und Tabellen einladen sowie Gelegenheit zu persönlichen Bekenntnissen und intimen Enthüllungen bieten. Vor diesem Hintergrund vermitteln die „freien“ und „ungezwungenen“ Kommunikationsräume auf Facebook eine äußerst tragfähige Motivation zur eigenständigen Wissensproduktion. Der Rückzug des Rasters in den informellen Bereich bewirkt also nicht per se die Desorganisation der Wissensregistriaturen. Im Gegenteil: Die Individualisierung der Datenkumulation suggeriert eine „dezentrale“, „herrschaftslo-

se“ und letztlich „wahrhaftige“ Praxis der Wissensermittlung und „belebt“ in vielen Fällen die Bereitschaft des Subjekts zur Selbstauskunft (vgl. Adelman 2011, S. 127–144).

Im Rezeptionsmodus seiner Mitglieder – aus der Sicht des Front End – markiert Facebook den Aufstieg einer neuen Subjektkultur, die sich vor allem im Erwerb von kultureller und sozialer Distinktion herstellt:

„Entsprechende Definitionen von Zugehörigkeit und Ausschlussstrategien und das Verhandeln der Werte und Normen sowie der davon abgeleiteten Machtrelationen definieren nicht nur die jeweiligen ‚Außengrenzen‘ derselben, sondern auch die Auseinandersetzungen um deren Erhalt.“ (Lummerding 2011, S. 209)

Selbstredend zählen Distinktionsmechanismen zum Grundbestand sozialer Organisation. Mit dem Aufstieg von konzernkontrollierten sozialen Medien wie etwa Facebook werden Distinktionen aber rechnerbasiert mit Hilfe algorithmischer Verfahren generiert. Die rechnerbasierte Datenverarbeitung der von den Userinnen und Usern generierten Daten und Informationen vollzieht Facebook im Back End. Dabei handelt es sich um einen, hinter dem für die Userinnen und User zugänglichen Interface der graphischen Benutzeroberfläche, angesiedelten Rechenraum, der ausschließlich der unternehmensinternen Datenerhebung dient.

Die Nutzerinnen und Nutzer haben folglich nur eine ganz vage Vorstellung davon, welche Arten von Informationen gesammelt und wie sie verwendet werden. So gehen die Formen der Personalisierung zwar auf Informationen zurück, die von den Mitgliedern selbst ausgehen, aber der Prozess der Vorhersage und des Anbietens von Inhalten, die auf vergangenem Verhalten basieren, ist eine Push-Technologie, welche die Mitglieder von Beteiligung und Einflussnahme konsequent ausschließt.

In diesem Zusammenhang möchte ich folgende Fragestellungen aufwerfen: Wie berechnet die soziale Netzwerkseite Facebook die Zukunft seiner Mitglieder? Welche Verfahren der Registrierung, der Berechnung, der Auswertung, der Adressierung verwendet sie zur Herstellung prognostischen Wissens? Wie werden diese Verfahren eingesetzt, um prognostische Modellierungen über das Verhalten von Userinnen und Usern herzustellen?

Happiness Index

Die Glücksforschung nutzt heute vermehrt die Sozialen Netzwerke zur Auswertung ihrer Masendaten. Innerhalb der Big-Data-Prognostik stellt die sogenannte „Happiness Research“ eine zentrale Forschungsrichtung dar. Doch die sozio-ökonomische Beschäftigung mit dem Glück wird überwiegend unter Ausschluss der akademischen Öffentlichkeit durchgeführt. In diesem Zusammenhang warnen einflussreiche Theoretiker wie Lev Manovich (vgl. 2012) und Danah Boyd (vgl. 2011) daher vor einem „Digital Divide“, der das Zukunftswissen einseitig verteilt und zu Machtasymmetrien zwischen Forschenden *innerhalb* und *außerhalb* der Netzwerke führen könnte. Manovich kritisiert den limitierten Zugang zu sozialstatistischen Daten, der von vornherein eine monopolartige Regierung und Verwaltung von Zukunft schafft:

“[...] only social media companies have access to really large social data – especially transactional data. An anthropologist working for Facebook or a sociologist working for Google will have access to data that the rest of the scholarly community will not.” (Manovich 2012: S. 467)

Dieses ungleiche Verhältnis festigt die Stellung der Sozialen Netzwerke als computerbasierte Kontrollmedien, die sich Zukunftswissen entlang einer vertikalen und eindimensionalen Netz-Kommunikation aneignen: (1) Sie ermöglichen einen kontinuierlichen Fluss von Daten (digitale Fußabdrücke), (2) sie sammeln und ordnen diese Daten und (3) sie etablieren geschlossene Wissens- und Kommunikationsräume für Expertinnen bzw. Experten und ihre Expertisen, welche die kollektiven Daten zu Informationen verdichten und interpretieren. Das Zukunftswissen durchläuft folglich unterschiedliche mediale, technologische und infrastrukturelle Schichten, die hierarchisch und pyramidal angeordnet sind:

“The current ecosystem around Big Data creates a new kind of digital divide: the Big Data rich and the Big Data poor. Some company researchers have even gone so far as to suggest that academics shouldn’t bother studying social media data sets – Jimmy Lin, a professor on industrial sabbatical at Twitter argued that academics should not engage in research that industry ‘can do better’.” (Boyd/Crawford 2011)

Diese Aussagen verdeutlichen – neben der faktisch gegebenen technologisch-infrastrukturellen Abschottung des Zukunftswissens –, dass das strategische Entscheidungshandeln im Back-End-Bereich und nicht in der Peer-to-Peer-Kommunikation angelegt ist. Die Peers können zwar in ihrer eingeschränkten Agency die Ergebnisse verfälschen, Fake-Profile anlegen und Nonsense kommunizieren, besitzen aber keine Möglichkeiten der aktiven Zukunftsgestaltung, die über taktische Aktivitäten hinausgehen.

Warum ist eigentlich die Erforschung des Glücks für die Gestaltung des Zukunftswissens so relevant geworden? Die Dominanz der Glücksforschung hat zwei historische Gründe (vgl. Frey/Stutzer 2002, S. 402). Seit der griechischen Antike wird dem Glück eine zentrale Stelle im menschlichen Leben eingeräumt und nach Aristoteles besteht das Ziel alles menschlichen Tuns darin, den Zustand der Glückseligkeit zu erlangen.⁷ Ein weiterer maßgeblicher Diskursstrang ist der seit Jeremy Bentham einflussreich gewordene Utilitarismus der Glücksdiskurse. Mit dem *Greatest Happiness Principle* entwickelte Bentham die Vorstellung, dass das größte zu erreichende Gut das Streben nach dem größtmöglichen Glück für die größtmögliche Anzahl von Menschen bedinge (vgl. Bentham 1977, 393f.). An diese sozio-ökonomische Konzeption des Glücks knüpft die „Happiness Research“ an, die Glück nach rationalem Kalkül als individuellen Nutzen interpretiert und in der Hochrechnung von aggregierten Glücksbekundungen das soziale Wohlbefinden berechnet.

⁷ Dieses unveräußerliche Recht des Menschen auf Glück (*the pursuit of happiness*) nahmen die Vereinigten Staaten von Amerika in die Eröffnungspassage ihrer Unabhängigkeitserklärung auf.

Eine maßgebliche Spielart der futurologischen Prophetie stellt der seit 2007 eingeführte „Facebook Happiness Index“ dar, der anhand einer Wortindexanalyse in den Statusmeldungen die Stimmung der Nutzerinnen und Nutzer sozialempirisch auswertet. Auf der Datengrundlage der Status-Updates errechnen die Netzwerkforscherinnen und -forscher in ihrem „Gross National Happiness Index“ (GNH) das sogenannte „Bruttonationalglück“ von Gesellschaften. Der Soziologe Adam Kramer arbeitete von 2008 bis 2009 bei Facebook und errechnete gemeinsam mit den Mitarbeiterinnen und Mitarbeitern des Facebook Data Team, der Sozialpsychologin Moira Burke, dem Informatiker Danny Ferrante und dem Leiter der Data Science Research Cameron Marlow den Happiness Index. Kramer konnte dabei das intern verfügbare Datenvolumen des Netzwerks nutzen. Er evaluierte die Häufigkeit von positiven und negativen Wörtern im selbstdokumentarischen Format der Statusmeldungen und kontextualisierte diese Selbstaufzeichnungen mit der individuellen Lebenszufriedenheit der Nutzer (*convergent validity*) und mit signifikanten Datenkurven an Tagen, an denen unterschiedliche Ereignisse die Medienöffentlichkeit bewegten (*face validity*):

“‘Gross national happiness’ is operationalized as a standardized difference between the use of positive and negative words, aggregated across days, and present a graph of this metric.” (Kramer 2010, S. 287)

Die von den Soziologinnen und Soziologen analysierten individuellen Praktiken der Selbstsorge werden mit Hilfe von semantischen Wortnetzen letztlich auf die Oppositionspaare „Glück“/„Unglück“ und „Zufriedenheit“/„Unzufriedenheit“ reduziert. Eine binär strukturierte Stimmungslage wird schließlich als Indikator einer kollektiven Mentalität veranschlagt, die auf bestimmte kollektiv geteilte Erfahrungen rekurriert und spezifische Stimmungen ausprägt. Die soziologische Massenerhebung der Selbstdokumentationen (*self reports*) in Sozialen Netzwerken hat bisher die Stimmungslage von 22 Nationalstaaten ermittelt. Mit der wissenschaftlichen Korrelation von subjektiven Befindlichkeiten und bevölkerungsstatistischem Wissen kann der „Happy Index“ nicht nur als Indikator eines „guten“ oder „schlechten“ Regierens gewertet werden, sondern als Kriterium einer möglichen Anpassungsleistung des Politischen an die Wahrnehmungsverarbeitung der Sozialen Netzwerke. In diesem Sinne stellt der „Happy Index“ ein erweitertes Instrumentarium wirtschaftlicher Expansion und staatlicher-administrativer Entscheidungsvorbereitung dar.

„Profiling the Future“: Subjektivierungsmodelle

Die operative Erforschung der Big Data weist zwei unterschiedliche Ausrichtungen auf. Einerseits versucht sie, kollektive Figurationen und Tendenzen kollektiver Dynamiken zu modellieren, andererseits geht es ihr darum, aus großen Datenmengen *personenzentrierte* und *zielgruppenspezifische* Merkmalsausprägungen herauszulesen. Im Folgenden möchte ich die historischen Diskursfelder dieser datenbasierten Techniken zur Herstellung subjektzentrierten Wissens herausarbeiten, um das strategische Bezugsverhältnis zwischen Wissen und Macht – oder genauer: zwischen Regierungswissen und der Herausbildung von Subjektivierungsmodellen – akzentuieren zu können.

In seinen Anfängen wurde das Profiling als Bewertungsmethode im Personalausleseverfahren der Testpsychologie in den USA entwickelt (vgl. Giordano 2005). Die standardisierten Verfahren der Testpsychologie zur Ermittlung von Leistungsfähigkeit bilden direkte Vorläufer des Profiling (aber auch der Rasterfahndung). Begriffe wie das „Persönlichkeitsprofil“ oder das „Profiling“ entstammen dem psychologisch-therapeutischen Diskurs und markieren heute Leitdiskurse in den Praxisformen der Selbstthematization. Unter den Vorzeichen des Postfordismus hat sich das Profiling als ein Ökonomisierungs- und Standardisierungsinstrument gesellschaftlich verallgemeinert und ist als eine vielschichtige Such- und Analyse-methode der Informations- bzw. Wissensgesellschaften in Verwendung. Das hohe Ansehen der Selbstevaluation verweist auf zwei soziale Prozesse. Einerseits hat sich die Anzahl der Testparameter und -verfahren und der daran beteiligten Testobjekte mit dem Auftritt der Web-2.0-Interfacetechnologien vervielfältigt, andererseits hat sich – in Abgrenzung zur beruflichen Eignungsdiagnostik – die Evaluationspraxis auch in qualitativer Hinsicht verändert und umfasst heute die gesamte Persönlichkeit und kreativen Potenziale des Subjekts.

Das Profiling wird von Facebook zur prognostischen Verhaltensanalyse eingesetzt (vgl. Klaasen 2007, S. 35). Beim Profiling handelt es sich um eine Wissenstechnik, bei der die Mitglieder in statistische Kollektiva eingeteilt werden. Grundlage dieser Datenaggregation sind demografische, sozio-ökonomische und geografische Faktoren. Die Mitglieder werden nach Alter, Geschlecht, Lebensabschnitt, sozialer Klasse, Bildungsgrad, Einkommen und Wohnort erfasst. Das Profiling erschließt auch Keywords durch Fanangaben und Gruppenmitgliedschaften. Die Aufzeichnungs- und Speicherpraktiken der Keywords werden für die Durchführung von Matchingverfahren herangezogen. Dabei werden zwischen den Geschmacksprofilen Ähnlichkeiten und Übereinstimmungen hergestellt. Eine Gruppe ergibt sich nach dem Prinzip der Übereinstimmung, z.B. formt sich eine Gruppe, wenn sie bestimmte Freundinnen und Freunde, Musikvorlieben, Kinofilme und Reisedestinationen miteinander teilt. Die Keywords basieren auf den Informationen, welche die Mitglieder selbst angegeben haben, also auf Interessen, Aktivitäten, Präferenzen.

Die Auswertungsverfahren des Nutzungsverhaltens zielen auf die Herstellung von Verhaltensvorhersagen, die als Bezugspunkt für die zielgruppenspezifische und interessengebundene Werbung dienen. Seit Januar 2011 transformiert Facebook mit Hilfe der sogenannten „gesponserten Meldungen“ (*sponsored stories*) die Facebook-Aktivitäten normaler Nutzerinnen und Nutzer in eine Werbeanzeige. Das Prinzip dieser User-basierten Marketingmethode ist simpel: Wenn Userinnen und User den „Gefällt-mir“-Button auf der Webseite eines Unternehmens klicken oder über „Facebook Places“, den Ortungsservice des Netzwerks, bei einem Partnerunternehmen von Facebook einchecken, kann sich das Unternehmen die Information zunutze machen und sich automatisiert in der rechten Spalte im neuen Echtzeit-Ticker als Anzeige darstellen lassen. Die Anzeige wird automatisch vermittels einer Social Tracking Software geschaltet und besteht aus dem Logo und dem Link des Unternehmens mitsamt Profilfoto, Name und Aktion des/der Facebook-Nutzers/Nutzerin. Facebook-Nutzerinnen und -Nutzer werden dadurch zu unbezahlten Werbeträgerinnen bzw. -trägern.

Das Web 2.0 mit seinen Social Networks und Communities verspricht daher ein großes prognostisches Potenzial, weil Marketingaktivitäten auf bestimmte Zielgruppen mittels modularer Technologien für User Tracking, Webmining, Profiling, Testing, Optimierung, Ad-Serving und Targeted Advertising abgestimmt werden können. Das Profiling im Web 2.0 verläuft nach dem Prinzip des Closed Circuit. Die Anordnung des Closed Circuit beschreibt ein Aufzeichnungsverfahren, bei der das Eingabemedium direkt mit dem Abbildungsmedium verbunden ist. Bei der Beobachtungsanordnung im Closed Circuit machen die Userinnen und User die Erfahrung der Synchronität ihrer Handlungen. Die sofortige Verfügbarkeit der Datenstrukturen und ihre gleichzeitige Manipulationsmöglichkeit durch das Targeted Advertising ist eine besondere Eigenschaft des Echtzeit-Profilings, das vergangene Nutzungsgewohnheiten von Online-Rezipientinnen und -Rezipienten analysiert (Click Advertising, Graphenanalyse), um zielgerichtete Werbung (Quality Market) für ein künftiges Konsumverhalten zu modellieren. Vor diesem Hintergrund entwickelte Microsoft ein Profiling-System, das soziometrische Daten wie etwa Alter, Geschlecht, Einkommen und Bildung mit möglichst großer Wahrscheinlichkeit ableiten sollte. Der Wirkungsbereich dieser sozialen Software umfasst zwei Bereiche: Als Medium vermittelt sie Prozesse und bewirkt eine Virtualisierung und Entgrenzung von Kommunikation; als Werkzeug greift sie strukturbildend in Zusammenhänge ein, bleibt aber selbst interpretationsbedürftig:

“The information architectures and classification tools that underlie many of the new technologies impacting on front-line practice are designed by a small elite, with decisions on what is represented and what is not.” (Webb 2006, S. 165)

Die Prognosefähigkeit der Sozialen Netzwerke ist davon abhängig, ob es gelingt, die biografisch und demografisch relevanten Daten und Informationen in distinkte und segregierte Bausteine der weiteren Datenverarbeitungen aufzugliedern. Als ein gemischtes Medium muss sich das Profiling zwangsläufig aus heterogenen Repräsentationen zusammensetzen. Es übernimmt das Modell der Prüfung von Persönlichkeitsmerkmalen der älteren Eignungsdiagnostik und macht es zur Sache kollektiver Approbationsleistungen, um seine Wirkungsweisen zu vervielfältigen und zu verstärken.

Die Profilbildung enthält Wissenstechniken, die auf binären Unterscheidungen beruhen (z.B. die Geschlechtszugehörigkeit), mit quantitativen Skalierungen operieren (z.B. hierarchische Ranking-Techniken) oder die auf die Erstellung qualitativer Profile abzielen (z.B. das Aufzeigen kreativer Fähigkeiten und Begabungen in „freien“ Datenfeldern). Profile reproduzieren einerseits soziale Normen und bringen andererseits auch neue Formen von Individualität hervor. Sie verkörpern den Imperativ zur permanenten Selbstentzifferung auf der Grundlage bestimmter Auswahlmenüs, vorgegebener Datenfelder und eines Vokabulars, das es den Individuen erlauben soll, sich selbst in einer boomenden Bekenntniskultur zu verorten. Das „bedienerfreundliche“ Profiling besteht in der Regel aus sogenannten Tools, das sind Checklisten, Fragebögen für Selbst-Evaluierung, analytische Rahmen, Übungsabschnitte, Bilanzen, Statistiken mit Kommentar, Datenbanken, Listen von Adressen und pädagogische Module zur Ermittlung individueller Fähigkeiten, Neigungen und Lieblingsbeschäftigungen.

Kommerzielle Suchmaschinen analysieren mittels Behavioural Targeting die Profile ihrer Nutzerinnen und Nutzer. Diese Suchtechnologie erlaubt es, auf verhaltensorientierte Kriterien wie Produkteinstellung, Markenwahl, Preisverhalten, Lebenszyklus zu reagieren und relevante Werbung zu schalten. Das Behavioural Targeting evaluiert kontinuierliche Nutzungsgewohnheiten, private Interessen und demografische Merkmale und erstellt damit ein statistisches Relief pluraler und flexibler Subjektivität (vgl. Castelluccia 2012, S. 21–33). Das wesentliche Merkmal des digitalen Targeting ist der Sachverhalt, dass das Individuum nur noch als dechiffrierbare und transformierbare Figur seiner Brauchbarkeiten in den Blick kommt. Es erzeugt ein multiples und „dividuelles Selbst“ (Deleuze 1993, S. 260), das zwischen Orten, Situationen, Teilsystemen und Gruppen oszilliert – ein Rekurs auf eine personale Identität oder ein Kernselbst ist unter dividuellen Modulationsbedingungen nicht mehr vorgesehen.

Digitales Targeting ist Bestandteil umfassender Such- und Überwachungstechnologien im Netz: Das Data-Mining ist eine Anwendung von statistisch-mathematischen Methoden auf einen spezifischen Datenbestand mit dem Ziel der Mustererkennung und beschränkt sich nicht auf die in der Vergangenheit erhobenen Daten, sondern erfasst und aktualisiert die Daten bei jedem Besuch im Netzwerk erneut in Echtzeit. Die im Internet geläufigen Surveillance-Tools ermöglichen es dem E-Commerce-Business, die jeweiligen Zielgruppen im Internet spezifischer zu identifizieren und gezielter zu adressieren. Das Marketing wächst im Internet zu einer entscheidenden Größe sozialer Regulation und die neuen Kontrollformen bedienen sich des Consumer Profiling. Mit dem digitalen Regime hat sich die computergestützte Rasterfahndung auf die Allgemeinheit ausgeweitet. Professionelle und kommerziell orientierte Consumer Profiler, die sowohl für das Marketing als auch für das E-Recruiting arbeiten, vollziehen eine Transformation des polizeilichen Wissens und sammeln ihr Wissen über die privaten Gewohnheiten der Bürgerinnen und Bürger mit der Akribie geheimdienstlicher Methoden. Bemerkenswert an dieser neuartigen Konstellation ist die emphatische Verankerung der Ökonomisierung des „menschlichen“ Faktors in weiten Bereichen des sozialen Lebens: „Die numerische Sprache der Kontrolle besteht aus Chiffren, die den Zugang zur Information kennzeichnen bzw. die Abweisung. Die Individuen sind ‚dividuell‘ geworden, und die Massen Stichproben, Daten, Märkte oder ‚Banken‘.“ (ebd.)

In der Argumentation von Gilles Deleuze wird nochmals deutlich, dass das numerische Prinzip als Metapher für das Funktionieren neuer gesellschaftlicher und ökonomischer Ordnungsstrukturen verwendet wird. Die neue Sprache der prognostischen Kontrolle besteht – nach Deleuze – aus Nummernkombinationen, Passwörtern oder Chiffren und organisiert den Zugang zu oder den Ausschluss von Informationen und Transaktionen. Soziale Organisationen werden wie Unternehmen geführt und werden nach der numerischen Sprache der Kontrolle kodiert: vom Bildungscontrolling bis zur Rankingliste.

Im Unterschied zur klassisch analogen Rasterfahndung geht es beim digitalen Data Mining nicht mehr um die möglichst vollständige Ausbreitung der Daten, sondern um eine Operationalisierung der Datenmassen, die für prognostische Abfragen und Auswertungen effektiv in Beziehung zueinander gesetzt werden können. Es verändert nicht nur die Wissensgenerierung persönlicher Daten und Informationen, sondern auch die Prozesse sozialer Regle-

mentierung. Insofern erzeugt das computergestützte Behavioural Targeting mehr als eine technische Virtualisierung von Wissensformen, denn es transformiert nachhaltig das Konzept des Raums, was zur Folge hat, dass sich das Raster vom topografischen Raum verflüchtigt und an seine Stelle der topologische Datenraum tritt. Dieser topologische Datenraum steht in Opposition zur Anwendungsschicht, die dem Kommunikationsraum der Nutzerinnen und Nutzer entspricht. Das futurische Wissen (bestehend aus der statistischen Erhebungsmethode des Data Mining, der Visualisierungstechnik des Data Mapping und des systematischen Protokollierungsverfahrens des Data Monitoring) ist konstitutiv aus der Anwendungsschicht ausgeschlossen und den Nutzerinnen und Nutzern nicht zugänglich. Damit basiert das Zukunftswissen der Sozialen Netzwerke auf einer asymmetrisch verlaufenden Machtbeziehung, welche sich in die technische Infrastruktur und in den Aufbau des medialen Dispositivs verlagert hat.

Das Zukunftswissen der Sozialen Netzwerke

Ich habe in den vorangegangenen Kapiteln versucht aufzuzeigen, dass Soziale Netzwerke zu gewichtigen Quellensammlungen für die statistische Massenerhebung aufgestiegen sind und neue Formen wissenschaftlichen Wissens hervorgebracht haben. Ihre gigantischen Datenbanken dienen der systematischen Informationsgewinnung und werden für das Sammeln, Auswerten und Interpretieren von sozialstatistischen Daten und Informationen eingesetzt. In ihrer Funktion als Speicher-, Verarbeitungs- und Verbreitungsmedium von Massendaten haben Soziale Netzwerke umfangreiche Datenaggregate hervorgebracht, die zur Prognose von gesellschaftlichen Entwicklungen herangezogen werden.

Das Zukunftswissen der Sozialen Medien steht aber nicht allen Beteiligten gleichermaßen zur Verfügung. Dieses asymmetrische Verhältnis zwischen gewöhnlichen Nutzerinnen bzw. Nutzern und exklusiven Expertinnen bzw. Experten wurde in der einschlägigen Literatur als „Participatory Gap“ (vgl. Taewoo/Stromer-Galley 2012, S. 133–149) diskutiert. Obwohl es eine neue Form des Regierens und Verwaltens nahe legt, wird das von den Sozialen Netzwerken ermittelte Zukunftswissen von der öffentlichen Diskussion ausgeschlossen.

Soziale Netzwerke haben der empirischen Sozialforschung neue Möglichkeiten der Quellenerschließung eröffnet. Das Zukunftswissen der Sozialen Netzwerke überlagert zwei Wissensfelder. Die empirische Sozialwissenschaft und die Medieninformatik sind für die Auswertung der medienvermittelten Kommunikation in interaktiven Netzmedien zuständig. Die Sozialforschung sieht in den Kommunikationsmedien der Sozialen Netzwerke eine maßgebliche Kraft für die gesellschaftliche Entwicklung. Ihre Forschungsperspektive auf die informationstechnische Vergesellschaftung in multimedial vernetzten Medien hat ein Koordinatennetz unterschiedlicher Wissensquellen und Wissenstechniken entwickelt, um prognostisches Wissen herzustellen. So wird etwa die Wissensbeschaffung an Suchroboter delegiert, die auf die öffentlichen Informationen zugreifen können. Das Zukunftswissen kann aber auch zur Inszenierung von künftig zu erwartenden Konstellationen der statistischen Datenaggregate verwendet werden, wenn etwa das Facebook Data Team bestimmte Ausschnitte seiner Tätigkeiten auf seiner Webseite popularisiert. In diesem Sinne werden statistische Daten und Informationen in

die Außenrepräsentation der Sozialen Netzwerke eingebaut und erhalten eine zusätzliche performative Komponente.

In seiner Modellierung durchläuft das Zukunftswissen unterschiedliche Felder der Herstellung, Aneignung und Vermittlung und kann als Verfahren, Argumentation und Integration eingesetzt werden. Vor diesem Hintergrund kann das Zukunftswissen als ein heterogenes Wissensfeld angesehen werden, das empirisches, formal-mathematisches, semantisches, psychologisches und visuelles Wissen in sich aufnimmt. Dementsprechend hat sich eine futurische Episteme an die Sozialen Medien angelagert und eine Vielzahl von Planungs- und Beratungspraktiken hervorgebracht, die als Multiplikatoren eines rechnerbasierten Machtgefälles und einer zeitbasierten Herrschaftsordnung auftreten. Vor diesem Hintergrund habe ich in meinem Beitrag darauf aufmerksam machen wollen, dass Prognosetechniken immer auch als Machttechniken angesehen werden können, die sich in medialen Anordnungen und infrastrukturellen Vorgaben im Front-End manifestieren. Das gestiegene Interesse der Markt- und Meinungsforschung an den Trendanalysen und Prognosen der Sozialen Netzwerke verdeutlicht, dass soziale, politische und ökonomische Entscheidungsprozesse hochgradig von der Verfügbarkeit prognostischen Wissens abhängig gemacht werden. Insofern berührt die Plan- und Machbarkeit des Zukunftswissens in unterschiedlichen Gesellschafts-, Lebens- und Selbstentwürfen immer auch die Frage: „Wie ist es möglich, nicht regiert zu werden?“

Literatur

- Adelmann, Ralf (2011): Von der Freundschaft in Facebook: Mediale Politiken sozialer Beziehungen in Social Network Sites. In: Leistert, Oliver & Röhle, Theo (Hrsg.): *Generation Facebook: Über das Leben im Social Net*. Bielefeld: transcript Verlag, S. 127–144.
- Bentham, Jeremy (1977): A Comment on the Commentaries and A Fragment on Government. In: Burns, James H. & Hart, Herbert L. A. (Hrsg.): *The Collected Works of Jeremy Bentham*. London.
- Bollen, Johan (2011): Happiness Is Assortative in Online Social Networks. *Artificial Life* 17 (3). S. 237–251.
- Bollen, Johan; Mao, Huina & Zeng, Xiaojun (2011): Twitter mood predicts the stock market. *Journal of Computational Science* 2 (1), S. 1–8.
- Bollier, David (2010): The promise and peril of big data, Washington, DC: The Aspen Institute, http://www.aspeninstitute.org/sites/default/files/content/docs/pubs/The_Promise_and_Peril_of_Big_Data.pdf [Stand vom 27-12-2013].
- Boyd, Danah & Crawford, Kate (2011): Six Provocations for Big Data. In: *Conference Paper, A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*, September 2011, Oxford: University of Oxford und Online: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1926431 [Stand vom 27-12-2013].

- Castelluccia, Claude (2012): Behavioural Tracking on the Internet: A Technical Perspective. In: Gutwirth, Serge et. al. (Hrsg.): *European Data Protection. Good Health?* New York/London: Springer Verlag, S. 21–33.
- Conover, Michael D.; Goncalves, Bruno; Ratkiewicz, Jacob; Flammini, Alessandro & Menczer, Filippo (2011): *Predicting the Political Alignment of Twitter Users. Proceedings of the 3rd IEEE Conference on Social Computing*, forthcoming. Online: http://cnets.indiana.edu/wpcontent/uploads/conover_prediction_socialcom_pdfexpress_ok_version.pdf, [Stand vom 27-12-2013].
- Deleuze, Gilles (1993): *Unterhandlungen 1972–1990*. Frankfurt am Main: Suhrkamp Verlag.
- Doorn, Niels Van (2010): The ties that bind: the networked performance of gender, sexuality and friendship on MySpace. *New Media & Society* 12 (4), S. 583–602.
- Frey, Bruno S. & Stutzer, Alois (2002): What can economists learn from happiness research? *Journal of Economic Literature* 40 (2), S. 402–435.
- Freyer-Dugas, Andrea et. al. (2012): Google Flu Trends: Correlation With Emergency Department Influenza Rates and Crowding Metrics. *Clinical Infectious Diseases*, 54 (15), S. 463–469.
- Gilbert, Eric & Karahalios, Karrie (2010): Widespread Worry and the Stock Market. In: *4th International AAAI Conference on Weblogs and Social Media (ICWSM)*. Washington, DC: George Washington University.
- Giordano, Gerard (2005): *How testing came to dominate American schools: The history of educational assessment*. New York/Wien: Lang.
- Klaasen, Abbey (2007): The Right Ads at the Right Time – via Yahoo; Web Giant Looks to Offer Behavioral-Targeting Tools outside its Own Properties. *Advertising Age* 3 Februar, S. 34–39.
- Kramer, Adam D. I. (2010): An Unobtrusive Behavioral Model of ‘Gross National Happiness’. In: Association for Computing Machinery (Hrsg.): *Conference on Human Factors in Computing Systems* 28 (3), New York, S. 287–290.
- Leistert, Oliver & Röhle, Theo (Hrsg.) (2011): *Generation Facebook: Über das Leben im Social Net*. Bielefeld: transcript Verlag.
- Lemke, Thomas (2004): Test. *Leviathan. Zeitschrift für Sozialwissenschaft* 32 (1), S. 119–124.
- Lummerding, Susanne (2011): Facebooking. What You Book is What You Get – What Else? In: Leistert, Oliver & Röhle, Theo (Hrsg.) (2011): *Generation Facebook: Über das Leben im Social Net*. Bielefeld: transcript Verlag, S. 199–216.
- Manovich, Lev (2012): Trending: The promises and the challenges of Big Social Data. In: Gold, Matthew K. (Hrsg.): *Debates in the digital humanities*. Minneapolis: University of Minnesota Press, S. 460–475.

- Nam, Taewoo & Stromer-Galley, Jennifer (2012): The Democratic Divide in the 2008 US Presidential Election. *Journal of Information Technology & Politics* 9 (2), S. 133–149.
- Wald, Randall; Khoshgoftaar, Taghi M. & Summer, Chris (2012): Machine Prediction of Personality from Facebook Profiles. In: *13th IEEE International Conference on Information Reuse and Integration*. Washington, S. 109–115.
- Webb, Stephen (2006): *Social Work in a Risk Society. Social and Political Perspectives*. Houndmills: Palgrave Macmillan.
- Wolf, Fredric et. al. (2011): Education and data-intensive science in the beginning of the 21st century. *OMICS: A Journal of Integrative Biology* 15 (4), S. 217–219.
- Yogatama, Dani (2012): *Predicting the Future: Text as Societal Measurement*, Online: http://www.cs.cmu.edu/~dyogatam/Home_files/statement.pdf [Stand vom 27-12-2013].