

Claudia Fraas, Christian Pentzold

## Big Data vs. Slow Understanding?

### Voraussetzungen und Vorgehen computerunterstützter Analyse transmedialer multimodaler Diskurse

**Abstract:** The paper looks into the practice of a computer-assisted discourse analysis. Centring on the decisions and procedures that go into reconstructing multimodal frames from transmedia discourse, the paper has two aims. For one, it discusses the chances and challenges automatic text analysis has to address in facing the vast amounts of multimodal discourse that emerge in convergent media. Building on that, the paper explains the methodological premises and methodical procedures of a discourse analysis employing the computer-assisted qualitative data analysis software Atlas.ti. As such, it presents an investigation that does not answer the current prevalence of available ‘big data’ with computer power but explores the interplay between an interpretative analysis and technological support. It does so by using material from the discourse on the so-called Handygate affair. There, state authorities collected mobile phone data during the commemorative events of the Dresden bombings in February 2011. They thus created a big data collection which became the topic of public attention and was either framed as illegal and extensive instrument of state surveillance or as an efficient and accurate tool for law enforcement and targeted prosecution.

DOI 10.1515/zgl-2015-0005

- 1 Einleitung: Chancen und Herausforderungen von Big Data
- 2 Forschungspraxis: Die softwareunterstützte Analyse transmedialer multimodaler Diskurse
- 2.1 Konzeptueller und methodologischer Rahmen: Multimodale Frames und der Forschungsstil der Grounded Theory
- 2.2 Methoden: Frameanalyse und Multimodalitätsanalyse
- 2.3 Technologische Unterstützung: Qualitative Datenanalysesoftware
- 3 Beispieldiskurs: Framing Big Data
- 3.1 Korpuserstellung: Sampling und Samplingkriterien
- 3.2 Rekonstruktion: Reflexives und reversibles Kodieren mit Atlas.ti

---

**Prof. Dr. Claudia Fraas:** Institut für Medienforschung, Technische Universität Chemnitz, Thüringer Weg 11, D-09126 Chemnitz, E-Mail: claudia.fraas@phil.tu-chemnitz.de

**Dr. des. Christian Pentzold:** Institut für Medienforschung, Technische Universität Chemnitz und Alexander von Humboldt Institut für Internet und Gesellschaft, Berlin, Thüringer Weg 11, D-09126 Chemnitz, E-Mail: christian.pentzold@phil.tu-chemnitz.de

- 4      Fazit: Nutzen und Grenzen computergestützter Verfahren im Rahmen einer Methodologie der Analyse transmedialer multimodaler Diskurse  
Literatur

## 1 Einleitung: Chancen und Herausforderungen von Big Data

Ein Aspekt der medialen Durchdringung privaten und öffentlichen Lebens sind die Unmengen an Daten, ‚Big Data‘ also, die entstehen, weil ein wachsender Teil alltäglicher und professioneller Aktivitäten mittels digitaler und vernetzter Technologien abläuft, und, damit verbunden, die steigende technologische Kapazität, diese Daten auszuwerten. Für die Analyse von Diskursen erwachsen in dreierlei Hinsicht *Chancen* aus diesen massiven Datenmengen und den technologischen Werkzeugen, sie zu bearbeiten:

*Erstens* können Diskursanalysen auf eine breite Palette an sozusagen ‚natürlich‘ generiertem diskursiven Material zurückgreifen (vgl. Lewis/Zamith/Hermida 2013, 35). Während sich traditionellerweise Diskursanalysen häufig auf massenmediale Kommunikation bzw. professionell hergestellte Texte beschränken und mündliche Anschlusskommunikation erst dokumentieren müssen, können sie nunmehr auf Diskurse zugreifen, deren semantische Beziehungen und intertextuelle Verweise sich medien- und kommunikationsformübergreifend, also transmedial, konstituieren. Damit einher geht die Erwartung, auf diese Weise eine breitere kulturelle Vielfalt und multimodale Varianz an diskursiv zum Ausdruck gebrachten Bedeutungen zu berücksichtigen, die sich abseits von bzw. mit vielfältigen Bezügen zu massenmedialen Diskursen äußern.

*Zweitens* können Diskursanalysen zusätzlich zu den diskursiv verfertigten Texten und Bedeutungsinhalten auf systemgenerierte Daten zurückgreifen, die permanent anfallen und beispielsweise Informationen über die zeitliche und räumliche Dimension diskursiver Aktionen, ihre Häufigkeit und die damit entstehenden Netzwerke zwischen Nutzern und zwischen Dokumenten liefern.

*Drittens* liegen beide Datenarten – explizit diskursiv generierte Aussagen und sozusagen beiläufig erzeugte Informationen – in digitaler Form vor und können maschinell gesucht und gesammelt werden. Entsprechend bieten sie sich insbesondere einer elektronischen Auswertung an, mittels derer sehr große Korpora mit vergleichsweise geringem technologischen und manuellen Aufwand untersucht werden können (vgl. Mautner 2005). In dieser Hinsicht konzentriert sich das Forschungsfeld der Digital Humanities unter anderem darauf, durch automatische Textauswertungen die Menge des analysierbaren Materials zu skalieren und Verfahrenswege für empirisch breit angelegte Untersuchungen zu schaffen (vgl. Manovich 2012). Ähnliche Bemühungen um *text mining* finden sich in den

Computational Social Science. Bereits in den 1990er Jahren sind im Rahmen der Korpus- und Computerlinguistik zudem Ansätze entstanden, die für computer-gestützte diskurslinguistische Untersuchungen angewandt werden können (vgl. Pêcheux 1995, Stubbs 2001, Scharloth/Eugster/Bubenhofner 2013).

Big Data stellt die Diskursanalyse zugleich vor eine Reihe von *Herausforderungen*. Zum einen muss das Zusammenstellen von adäquaten Korpora sechs Faktoren berücksichtigen, die insbesondere im Vergleich zu massenmedialer Kommunikation das Suchen, Sichern und Verfügbarmachen zeichenhaft belegter Diskurse erschweren (vgl. Meier et al. 2010): Weil das elektronisch zugängliche Material *erstens* in seiner Periodizität und *zweitens* in seiner Zentralität variiert, lässt sich keine verlässliche Grundgesamtheit bestimmen. Stets bleiben potentielle diskursive Fragmente der zeitlich und räumlich entgrenzten „mass self-communication“ (Castells 2009, 70) unberücksichtigt und es existiert kein verlässlicher Index (vgl. Lewis/Zamith/Hermida 2013, 40). Dazu tragen *drittens* und *viertens* die variable Sichtbarkeit und Persistenz der Daten bei, weshalb synchron und diachron stets nur empirische Ausschnitte aus dem möglichen Gesamtdiskurs präsent sind und unklar bleibt, inwiefern konkrete Korpora die virtuellen Korpora repräsentieren. Sich zur Lösung dieser Probleme auf Suchmaschinen und zentrale Kommunikationsplattformen zu verlassen, hieße wiederum, deren rechtliche und technologische Entscheidungen zu akzeptieren, welcher Teil eines Diskurses sichtbar und welcher gesperrt sein soll (vgl. Mahrt/Scharkow 2013, 21f.). Damit verbunden sind zudem ethische Fragen, inwiefern Äußerungen ohne Wissen und Zustimmung ihrer Sprecher genutzt werden dürfen, nur weil sie unter Umständen öffentlich zugänglich sind (vgl. boyd/Crawford 2012, 673) – eine Frage die auch in der Einleitung zum Themenheft als Desiderat einer Forschungsethik formuliert wurde. *Fünftens* muss die Konservierung mehrfachkodierten Diskursmaterials mit der Animation diskursrelevanter Texte umgehen, die in statischer Form z. T. ihre modale Kohärenz verlieren, und sie muss *sechstens* die Reaktivität der Zeichenkomplexe beachten, die nicht fertig vorliegen, sondern erst datenbankbasiert in Relation zu nutzer- und gerätebezogenen Einstellungen für Interfaces erzeugt werden (vgl. Pentzold/Fraas/Meier 2013). Weitere Aspekte wie die wachsende Bedeutung von Visualisierungen werden in den anderen Beiträgen des Sammelbandes thematisiert (Kath/Schaal/Dumm, Bubenhofner/Scharloth).

Zum anderen hat das Auswerten transmedialer Diskurse mit ihrer kommunikativen Organisation und semiotischen Konstitution umzugehen. Dies bedeutet *erstens*, dass sich das diskursrelevante Material hinsichtlich des Grads an Interaktivität unterscheidet. Die Vielzahl an Kommunikationsformen eröffnen in ihrem technologischen Aufbau und ihrer Nutzung verschiedene Optionen uni-, bi- oder multidirektionaler Kommunikation. Die diskursanalytische

Herausforderung liegt hierbei einmal darin, die Einheiten für Auswahl und Analyse zu bestimmen, und dann darin, auf die damit einhergehenden Formen der Adressierung und Kommunikationsrichtung, der zeitlichen Struktur des Austauschs, der diskursiven Rollen, der interdiskursiven impliziten und expliziten Bezüge sowie der möglichen modalen Ausdrucksweisen und stilistischen Register zu achten (vgl. Holly 2011). *Zweitens* steht die Diskursanalyse vor der Aufgabe, in multimodal konstituierten Diskursen die unterschiedlich kodiert vorliegenden Fragmente mit geeigneten Verfahren zu analysieren. Soll die Diskursanalyse also nicht nur verbalisierte Bedeutungsmuster rekonstruieren, sondern die empirische Vielfalt an multimodalen Zeichenensembles im Blick haben, dann muss auch die indexikalisch, ikonisch und symbolisch realisierte Bedeutungsstiftung von Typografie, Layout und Farbe, von Bewegtbild, Bild-Bild- bzw. Bild-Sprache-Kombinationen zeichenadäquat analysiert werden (vgl. Meier 2011, Fraas/Meier/Pentzold 2013, 15f., Bucher 2013). Schließlich stellen drittens die kommunikative Kreativität und das Tempo, mit dem neue Kommunikationsformen sich formieren, die Diskursanalyse vor die Herausforderung, flexibel ihre Erhebungs- und Auswertungsmethoden anzupassen und deren Adäquanz fortlaufend zu evaluieren (vgl. Karpf 2012, 647).

Im Folgenden soll in der Form eines Werkstattberichts die laufende Forschungspraxis des computerunterstützten Analysierens eines transmedial konstituierten Diskurses exemplarisch und ausschnittsweise vorgestellt werden. Folglich will der Beitrag nicht die Ergebnisse der materialen Diskursanalyse darlegen. Vielmehr soll erläutert werden, wie auf der einen Seite die kommunikationsstrukturelle und modale Variabilität des Materials methodisch reflektiert werden kann und wie auf der anderen Seite das Sichten der Datenmengen und das Sammeln und Auswerten der digitalen Daten durch geeignete Software zu unterstützen ist.

Konkret werden im nächsten Abschnitt das in der Analyse gebrauchte konzeptuelle Verständnis multimodaler Frames und ihr methodologischer Rahmen gemäß der Verfahrenslogik der Grounded Theory eingeführt. Darauf aufbauend werden das computerunterstützte Vorgehen als Kombination linguistischer Frameanalyse und sozialsemiotischer Multimodalitätsanalyse beschrieben sowie der diskursanalytische Einsatz der qualitativen Datenanalyse-Software Atlas.ti erklärt. Der Schwerpunkt der Diskursanalyse lag auf den multimodal generierten Deutungsmustern des Diskurses zur Handygate-Affäre 2011. Thema der Handygate-Affäre war eine Funkzellenabfrage der sächsischen Polizei während der Demonstrationen zum Gedenken des 13. Februar 2011 in Dresden. Die Analyse zeigt also, mithilfe welcher Deutungsmuster diese Form von Big Data als Phänomen und als Problem diskursiv konstituiert wurde.

## 2 Forschungspraxis: Die softwareunterstützte Analyse transmedialer multimodaler Diskurse

Da gesellschaftlich relevante Diskurse zunehmend von Referenzen zwischen professioneller Medienkommunikation und amateurhafter medienvermittelter Kommunikation geprägt sind, ist es diskursanalytisch problematisch, diese transmedialen interdiskursiven Verflechtungen in eine Online- und eine Offline-Sphäre zu trennen. Vielmehr werden Diskurse in thematisch, technologisch und institutionell konvergierenden Medienarrangements realisiert, wobei digitale vernetzte Kommunikation zentraler Bezugspunkt dieser Konvergenzbewegungen ist (vgl. Jenkins 2006).

Entsprechend sind diskursive Bedeutungsstrukturen kommunikationsformadäquat, also in ihren kommunikationsstrukturellen und semiotischen Besonderheiten zu rekonstruieren. Das dazu notwendige gegenstandsadaquate Methodeninstrumentarium adressiert daher Chancen und Herausforderungen des Erhebens und Analysierens transmedialer, multimodaler Diskursmuster. Dazu wurden linguistische Frameanalysen und sozialsemiotische Multimodalitätsanalysen miteinander im methodologischen Rahmen der sozialwissenschaftlich prominenten Grounded Theory verbunden und das Verfahren durch die Qualitative Datenanalyse-Software Atlas.ti unterstützt (vgl. Fraas et al. 2013, Sommer et al. 2013).<sup>1</sup>

### 2.1 Konzeptueller und methodologischer Rahmen: Multimodale Frames und der Forschungsstil der Grounded Theory

Der diskursanalytische Forschungsansatz greift auf linguistische und sozialsemiotische Konzepte zurück, um diskursive Bedeutungsstrukturen zu rekonstruieren, die mittels verschiedener semiotischer Ressourcen hervorgebracht wurden.

Wissensordnungen und ihre Rolle in Verstehens- und Interpretationsprozessen werden in unterschiedlichen Disziplinen häufig mithilfe des Frame-Konzepts beschrieben, so etwa in der Linguistik, der Kognitions- und der Kommunikationswissenschaft, der Politologie, Pädagogik und Soziologie. Der framesemantischen Definition folgend, werden Frames als Repräsentationsformate kognitiver Sinnstrukturen verstanden, die das Kontextualisierungspotential für Konzepte bereit-

---

<sup>1</sup> Der Beitrag fußt auf dem DFG-geförderten Projekt „Methodeninstrumentarium der Datenerhebung und -analyse zur Bestimmung von Online-Diskursen als gesellschaftliche Praktiken“, Professur Medienkommunikation, Technische Universität Chemnitz (FR 1328/5-1; Laufzeit 2009–2012).

stellen, sich in Zeichenverbänden materialisieren und damit Diskursanalysen zugänglich sind (vgl. Fillmore 1985, Konerding 1993, Fraas 1996, Ziem 2008). Wesentlich für das diskursive Framing von Sinnstrukturen sind das Zusammenspiel sprachlicher, visueller bzw. akustischer Zeichenressourcen in kontextgebundenen Bedeutungskonstitutions- und Interpretationsprozessen sowie deren Perspektivierungsleistung hinsichtlich Diskursthemen. Demgemäß finden sich neben Studien zur sprachlichen Repräsentation von Frames vermehrt Arbeiten zum strategischen Einsatz von Sprache-Bild-Korrespondenzen als Visual Framing (vgl. Geise/Lobinger 2013, Meier 2014).

Um eine Diskursanalyse angesichts der forschungspraktischen Chancen und Herausforderungen von Big Data viabel und analytisch plausibel durchführen zu können, sollten das Zusammenstellen des Korpus sowie die Auswahl und Anwendung der Analysemethoden systematisch hergeleiteten Kriterien folgen. Der Rahmen der Grounded Theory bietet sich hierfür an, denn er stellt mit seinen Sampling-, Kodier- und Kategorisierungsstrategien ein reflexiv-zirkuläres, interpretatives Vorgehen zur Verfügung, das Datenerhebung und -auswertung systematisch anleitet (vgl. Glaser/Strauss 1967/2005). Forschungspraktisch heißt dies, dass Erhebung und Analyse sich abwechseln und gegenseitig informieren: Das Auswählen, Kodieren und konzeptuelle Verdichten der Daten bilden, geleitet von den heuristischen Kodierelementen, die Grundlage für die Auswahl weiterer Fragmente, die wiederum kodiert sowie zu Konzepten verdichtet werden, welche neuerlich die Kriterien für das weitere Sampling liefern (vgl. Meier/Pentzold 2010). Dieser rekursive Prozess ermöglicht es zum einen, im Erheben der Daten aus multimodalen transmedialen Diskursfragmenten auf die variable Periodizität, Zentralität, Sichtbarkeit, Persistenz und Konservierbarkeit zu reagieren und die angewandten Verfahren zu schulen. Zum anderen können die Auswertungsmethoden gemäß der variablen kommunikativen Organisation und semiotischen Konstitution des Materials eingesetzt werden. Mit der so orientierten Analyse ist es schließlich möglich, konzeptuell begründete Frames als diskursive Deutungsmuster zu rekonstruieren. Zugleich sind mit der auf diese Weise ermöglichten Flexibilität und methodischen Adaptierbarkeit die beiden Standards für den Einsatz softwarebasierter Instrumente definiert.

## 2.2 Methoden: Frameanalyse und Multimodalitätsanalyse

Der hier vorgestellte Ansatz für eine Analyse von transmedialen multimodalen Diskursen versteht Frames als modal differenziert repräsentierte Bedeutungsstrukturen, die konzept- und datengeleitet in deduktiven und induktiven Verfahrensschritten rekonstruiert werden können.

Methodisch wird hierbei einmal auf die Tradition linguistischer Frameanalyse zurückgegriffen, die mit Minsky (1975) Frames als Slot-Filler-Strukturen versteht. Die Slots repräsentieren zum einen Standardwerte (*default values*), die erwartbare Situationen, Prozesse oder Objekte beschreiben. Zum anderen fungieren sie als Platzhalter für konkrete Werte (*fillers*), die in aktuellen Kontexten situationsadäquat aktualisiert werden können. Die für die Analyse benötigten Frameelemente werden in dem hier vorgestellten Ansatz auf der Grundlage des FrameNet-Projekts hergeleitet.<sup>2</sup> Entsprechend der framesemantischen Tradition des Projekts werden Frames definiert als „scriptlike conceptual structure that describes a particular type of situation, object, or event along with its participants“ (Ruppenhofer et al. 2010, 5). Die Frameelemente, welche die Struktur der Frames bilden, stehen für Einheiten „which may or must be present in any instance of a given frame“ (Fillmore/Baker 2010, 324). Dabei wird davon ausgegangen, dass in geschriebenen wie gesprochenen Sprachtexten Schlüsselausdrücke bestimmte Schlüsselkonzepte und damit Frames aktivieren, die die Interpretation dieser Ausdrücke leiten. FrameNet stellt ein hierarchisch aufgebautes semantisches Netz dar, in dem Frameelemente von abstrakteren Frames auf weniger abstrakte vererbt und zum Teil zu eigenen Frames entfaltet werden können. Auf Grund dieser in der Struktur angelegten semantischen Beziehungen zwischen den Frames und Frameelementen können Schlüsselkonzepte, die für die Konstituierung diskursiver Sinnstrukturen eine zentrale Rolle spielen und dementsprechend mit Hilfe sprachlicher Zeichen materialen Ausdruck finden, auf übergeordnete Frames zurückgeführt werden. Diese übergeordneten Frames stellen die Slot-Struktur für die Frames konkreter Ausdrücke zur Verfügung, die in aktuellen Verwendungssituationen kontextuell gefüllt werden können. In der Analyse werden also zunächst Schlüsselkonzepte ermittelt und mit Hilfe von FrameNet die Slot-Struktur der Frames rekonstruiert, die diese Schlüsselkonzepte evozieren. In der ersten Phase der Datenanalyse werden dann sprachliche und bildliche Ausdrücke unterschiedlicher Komplexität (Wörter, Wortgruppen, Textpassagen, Bild- und Video-Ausschnitte) als Filler für die Frameelemente den diskursiv zentralen Schlüsselkonzepten zugeordnet. Forschungspraktisch gestaltet sich dieser Vorgang so, dass auf der Grundlage der Frameelemente, die als Kodes erster Ordnung fungieren (vgl. Abb. 1), Fragen an das Material gestellt und

---

<sup>2</sup> FrameNet entwickelt ein Korpus von hierarchisch organisierten Frame-Netzwerken des gesamten lexikalischen Inventars der englischen Sprache. Vgl. <https://framenet.icsi.berkeley.edu> (letzter Zugriff 02.04.2014). Die Korpuserstellung wurde von Audrey Canaud, die Datenauswertung von Charlotte Fischer unterstützt.

ständig Vergleiche hinsichtlich der zugeordneten Filler gezogen werden. Zugleich werden die Codes auf der Grundlage der entsprechenden Filler ausdifferenziert.

Um die bedeutungsstiftende Korrespondenz visueller und sprachlicher Bedeutungsstiftung analytisch erfassen und somit der Multimodalität in den transmedialen Diskursfragmenten entsprechen zu können, werden im zweiten Schritt Analysefragen zur Kodierung der visuellen Daten generiert. Sie basieren auf den von Halliday (1990) entworfenen und von Kress und van Leeuwen (2001, 2006) für die Analyse von visuellen Zeichenverbänden weiterentwickelten sozialsemiotischen Metafunktionen. Indem die Metafunktionen als sensibilisierende Konzepte für Bilder und Design aufgefasst werden, lassen sich die bedeutungsstiftenden Potenziale visueller Arrangements in Korrespondenz mit anderen verwendeten semiotischen Ressourcen analysieren (vgl. Meier 2011, 2014). Bezogen auf diskursanalytische Fragestellungen heißt dies, mit der *ideational metafunction* diskursiv zentrale visuelle Schlüsselkonzepte zu erheben. Diese Metafunktion operiert ebenso wie Frames auf der konzeptuellen Ebene. Die *interpersonal metafunction* betrifft insbesondere stilistische Praktiken, die die Diskursthemen perspektivieren. Die *textual metafunction* schließlich fokussiert die Komposition der einzelnen modalen Aspekte in den Diskursfragmenten.

Die aus den Framestrukturen und sozialsemiotischen Metafunktionen abgeleiteten Elemente dienen im Sinne der Grounded Theory als heuristische Listen an Codes für einen ersten konzeptuellen Zugriff auf die Daten. Die framebasierten und sozialsemiotischen Elemente werden über das Material gelegt (*top-down* bzw. *corpus-based*). Im weiteren Verlauf der Analyse werden zum einen die Kodierheuristiken datennah ausdifferenziert. Zum anderen werden entsprechend der jeweiligen modalen Kontextualisierungen neue Codes generiert (*bottom-up* bzw. *corpus-driven*). In weiteren Schritten werden konzeptuelle Querverbindungen zwischen den modal verfassten Codes gebildet und diese zu Kategorien gebündelt, was in späteren Phasen der Analyse sowohl das weitere Sampling als auch das Herausarbeiten diskursiver Deutungsmuster anleitet (Abb. 1).

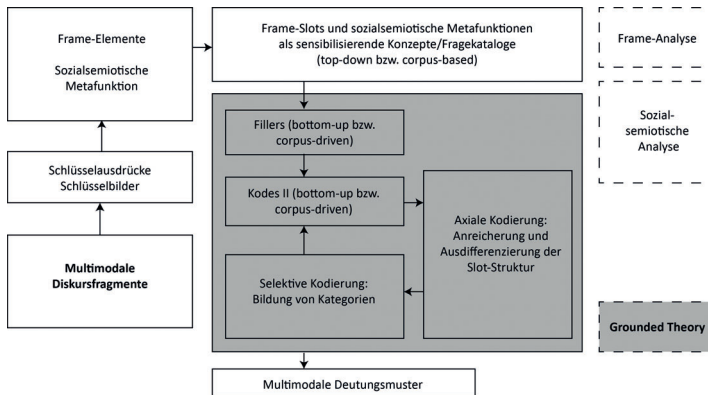


Abbildung 1: Forschungsprozess. Quelle: eigene Darstellung

## 2.3 Technologische Unterstützung: Qualitative Datenanalysesoftware

Frei verfügbare oder proprietäre Qualitative Datenanalysesoftware (QDA-Software, engl. CAQDA, d. h. *computer assisted qualitative data analysis*) bieten, allgemein gesehen, Werkzeuge zum Strukturieren, Organisieren und Codieren von sprachlich und, je nach Kapazität des Programms, auch bildlich bzw. bewegtbildlich vorliegenden Daten. Damit unterstützen die seit den 1980er Jahren entwickelten QDA-Pakete wie Aquad, Atlas.ti, FreeQDA, NUD\*IST, NVivo oder MAXQDA disziplinenübergreifend qualitativ-verstehend angelegte Forschungsvorgaben im Management und in der Interpretation ihrer Daten, vom Sampling bis zur Ergebnisdarstellung.

Der Einsatz von QDA-Software in Diskursanalysen wird einerseits mit gesteigerter Ressourceneffizienz durch einfachen Textzugriff über Suchfunktionen, der Organisation des Erhebungs- und Auswertungsprozesses durch die Dokumentation von analytischen Entscheidungen, Systematisierung von Abläufen und prozessbegleitende Reflexion sowie der Möglichkeit, qualitative und quantifizierende Strategien zu kombinieren, assoziiert. Diese Optionen können forschungspraktisch insbesondere im Umgang mit großen Korpora vorteilhaft genutzt werden. Jedoch wird andererseits auf die begrenzte Flexibilität und projektbezogene Adaptierbarkeit der Programme durch methodisch orientierte Voreinstellungen und damit einhergehende technologisch bedingte Vorgaben verwiesen (vgl. Diaz-Bone/Schneider 2004, Pentzold 2014).

Diese methodologisch abzuwägende Erleichterung und Disponierung zeigt sich exemplarisch an dem Programm Atlas.ti, das in der hier erläuterten Analyse

gebraucht wurde (in Versionsnummer 6). Dazu wurden die im Sampling von online- und offline verfügbarem Material erhobenen schriftlichen Dokumente, Bilddateien, Videodateien und Screenshots von Webseiten in ihrer modalen Orchestrierung in Atlas.ti geladen und ausgewertet. Die ergonomische Gestaltung von Atlas.ti orientiert sich am Verfahrensrahmen der Grounded Theory. Zum einen assistierte sie daher dem Auswerten, indem sie das Sammeln, Speichern und Verwalten der in digitaler Form vorliegenden Daten erleichterte und einen schnellen Zugriff auf Belegstellen und interpretative Deutungen (Kodes) gewährte. Demgemäß unterstützte Atlas.ti zum anderen das kategorienbasierte Erschließen und (Re-)Organisieren des Materials sowie das iterative Revidieren und Entwickeln des Kategoriensystems. Hinzu kamen, wie bei anderen QDA-Paketen mehr oder minder ähnlich, Möglichkeiten, die Datenarten wie Texte, Bilder und Videos in verschiedenen Formaten zu bearbeiten, Daten zu selektieren und zu gruppieren, Kodierungen von Segmenten, Konzept- und Kategorienbezeichnungen zu verändern, Querverweise zwischen Dokumenten, Kodes und reflektierenden Notizen zum Vorgehen und zu Interpretationen (Memos) einzurichten und die Beziehungen zwischen diesen einzelnen Materialien in Netzwerken zu visualisieren (vgl. Friese 2012).

### **3 Beispieldiskurs: Framing Big Data**

Thema der Handygate-Affäre war die Funkzellenabfrage der sächsischen Polizei während der Demonstrationen zum Gedenken des 13. Februar 2011 in Dresden. In Referenz auf dieses Ereignis, seine Umstände, Akteure und Konsequenzen entwickelte sich ein Diskurs zwischen publizistischer Berichterstattung, offiziellen Stellungnahmen von Staat und Polizei, den Positionen politischer und zivilgesellschaftlicher Akteure und Kommentaren in sozialen Medien. Die bezüglich dieses transmedialen und multimodalen Diskurses verfolgte Frage war, mittels welcher Deutungsmuster die massive Datensammlung als sozusagen Big Data in dem transmedialen Diskurs als Phänomen und als Problem konstituiert wurde.

#### **3.1 Korpuserstellung: Sampling und Samplingkriterien**

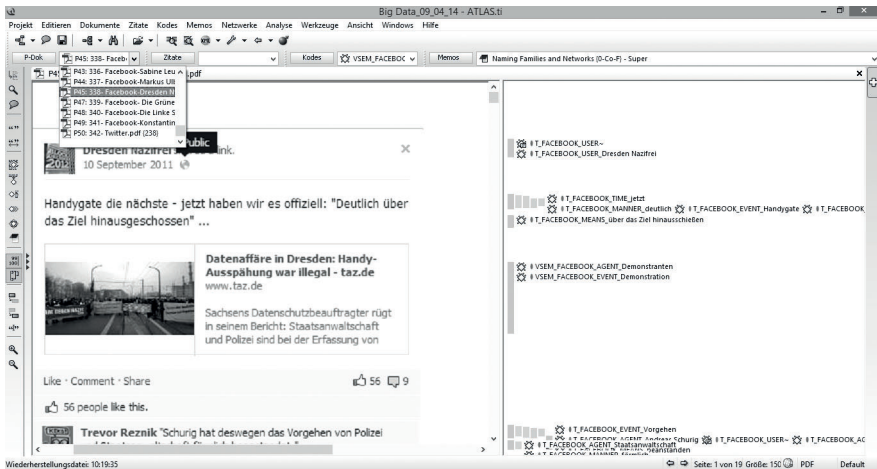
Für das Korpus wurden alle thematisch relevanten Texte aus der publizistischen Berichterstattung in Print- und Online-Medien, von persönlich auftretenden einzelnen und kollektiven Diskursakteuren sowie der Kommunikation auf sozialen Netzwerkseiten und Multimedia-Plattformen erhoben. Auswahleinheiten waren also Artikel in den Printausgaben lokaler sächsischer Tageszeitungen

(*Sächsische Zeitung, Dresdner Neueste Nachrichten, Freie Presse, Leipziger Volkszeitung*) und in überregionalen Tageszeitungen (*SZ, FAZ*) sowie in den Online-Ausgaben dieser Publikationen und zusätzlich den Online-Ablegern der Wochenzeitungen *Die Zeit* und *Der Spiegel*. Hinzu kamen Stellungnahmen, Pressemitteilungen und Webseiten der involvierten staatlichen Stellen, politischen Vertreter und zivilgesellschaftlichen Sprecher sowie implizit-thematisch und explizit-hypertextverlinkt verbundene Blögeinträge, Facebook-Posts, Facebook-Kommentare, Tweets, Foreneinträge, Videos und Kommentare auf YouTube. Die bis zum 12.12.2012 insgesamt 3.031 erhobenen Fragmente wurden im Zeitraum vom 23.02.2011 bis zum 21.10.2011 veröffentlicht.

Um ein konzeptuell repräsentatives Korpus zu erheben, wurden für die Analyse relevante Diskursfragmente kriteriengeleitet berücksichtigt. Zum einen wurden auf inhaltlich-diskursthematischer Ebene mittels Schlüsselbegriffe und Schlüsselbilder systematische Suchmaschinenrecherchen nach diskursrelevanten Themen und Akteuren in ihrer diskurspezifischen Darstellung durchgeführt. Diskursfragmente wurden demnach hinsichtlich folgender Kriterien gesucht und ausgewählt: inhaltliche Schwerpunkte, vertretene Meinungen bzw. Standpunkte, Autoren und erwähnte Akteure. Ziel der inhaltlich-thematischen Orientierung war es, im Sinne einer theoretischen Sättigung des Materials keine weiteren vom Diskursthema abweichenden Inhalte, Positionierungen und Perspektivierungen mehr ermitteln zu können.

Zum anderen wurde das Spektrum diskursiver Repräsentationsformen entsprechend des Kriteriums der Kommunikationsform abgedeckt. Kommunikationsformen wurden als mediale Dispositive verstanden, „die sich auf der Basis verfügbarer technischer Möglichkeiten und sozialer Bedürfnisse allmählich herausbilden und weiterentwickeln“ (Holly 2011, 155). Sie wurden erstens gemäß ihrer Modes und Codes, also der angesprochenen Sinneskanäle und verfügbaren Zeichenarten, zweitens entsprechend ihrer kommunikationsstrukturellen Gegebenheiten der Grade an Kopräsenz, Wechselseitigkeit und Adressierung sowie drittens im Blick auf die zeitliche Struktur hinsichtlich Flüchtigkeit bzw. Verdauerung differenziert. Ziel der auf Kommunikationsformen gerichteten Orientierung war es, keine vom Spektrum diskursiver Repräsentation abweichenden Darstellungen mehr ermitteln zu können.

Alle Diskursfragmente wurden in digitaler Form gespeichert und in Atlas.ti als nummerierte Dokumente übertragen, sodass sie dort einzeln in ihrer jeweiligen modalen Komposition aufgerufen, markiert und mit Codes versehen werden konnten (Abb. 2).



**Abbildung 2:** Screenshot des Bearbeitungsfensters von Atlas.ti und codiertem Dokument.  
Quelle: eigene Darstellung.

### 3.2 Rekonstruktion: Reflexives und reversibles Kodieren mit Atlas.ti

Zum analytischen Einstieg wurde ein Diskursfragment gewählt, das einen inhaltlichen Überblick über das Diskursthema gibt und dazu, so die Annahme, eine Reihe zentraler sprachlicher Ausdrücke, bildlicher Elemente und prominenter Akteure umfasst. Da die *Tagesschau* für eine stabil hohe journalistische Qualität und eine um ausgewogene Darstellung bemühte Berichterstattung steht, wurde deshalb zum Einstieg in die Analysen ein hier gesendeter audiovisueller Nachrichtenbeitrag gewählt (vom 25.07.2011).<sup>3</sup> Er wurde in seine filmischen Einstellungen zerlegt, der Sprechtext wurde transkribiert und jeder Einstellung zugeordnet.

Anhand dieser zwei modalen Ebenen des Einstiegstextes wurden Kandidaten für mögliche Schlüsselkonzepte bestimmt. Weil der Beitrag zur Handygate-Affäre zentral die darin von der Polizei betriebenen Aktivitäten des Sammelns von Mobilfunkdaten der Dresdner Demonstranten in ihrer juristischen Begründung und ihren demokratischen Implikationen problematisierte, begann die Analyse des textuellen Materials mit einer Kodierheuristik, die in Bezug zum Schlüsselkonzept SAMMELN und dem dazugehörigen FrameNet-Frame TRANSITIVE ACTION hergeleitet wurde.

<sup>3</sup> <http://www.tagesschau.de/multimedia/sendung/ts28004.html> (Zugriff: 01.04.2014).

Diesem Frame sind in FrameNet die folgenden Frameelemente zugeordnet, aus denen Analysefragen abgeleitet wurden, die in der ersten Phase des Kodierungsprozesses im Sinne der Kodierheuristik an das Datenmaterial zu stellen sind:

AGENT (Person/Institution, die in Bezug auf eine Entität (etwas) handelt): Wer handelt in Bezug worauf? Auf welche Weise handelt er? In welcher Phase der Aktion? Unter welchen Umständen?

PATIENT (Entität, in Bezug auf die gehandelt wird): In Bezug worauf wird gehandelt? Auf welche Weise wird in Bezug darauf gehandelt? In welcher Phase der Aktion? Unter welchen Umständen?

CAUSE (Ereignis, das verursacht, dass in Bezug auf etwas gehandelt wird): Welches Ereignis verursacht, dass in Bezug auf etwas gehandelt wird? Wie wird das Ereignis beschrieben? Auf welche Weise? In welcher Phase der Aktion? Unter welchen Umständen?

DEPICTIVE (Beschreibung Beteiligter an der transitiven Aktion): Wie werden die Akteure beschrieben?

EVENT (Ereignis, in dessen Rahmen die transitive Aktion stattfindet): Im Rahmen welchen Ereignisses handeln die Akteure?

MANNER (alle Beschreibungen der transitiven Aktion, die nicht durch spezifische andere Frameelemente gefasst sind, z. B. sekundäre Effekte oder generelle Beschreibungen): Wie wird die transitive Aktion bezüglich sekundärer Effekte beschrieben?

MEANS (eine vom Akteur ausgeführte Handlung, die ihn ermächtigt, auf das Betreffende einzuwirken): Welche vom Akteur ausgeführte Handlung ermächtigt ihn, auf das Betreffende einzuwirken? Auf welche Weise? In welcher Phase der Aktion? Unter welchen Umständen?

PLACE (Ort, an dem die transitive Aktion stattfindet): Wo findet das Ereignis statt? Wie wird der Ort beschrieben?

RESULT (Ergebnis der transitiven Aktion): Zu welchem Ergebnis führt die beeinflussende Handlung? Wie wird das Resultat beschrieben?

TIME (Zeitpunkt der transitiven Aktion): Zu welcher Zeit fand die transitive Aktion statt?

Für visuelle Inhalte wurden aus den sozialsemiotischen Metafunktionen folgende Codes abgeleitet:

Akteur: Wer ist dargestellt?

Soziale Rolle: Welche sozialen Rollen lassen sich für die Akteure anhand ihres Aussehens rekonstruieren?

Thema, Ereignis, Gegenstand, Situation: Was ist dargestellt? Kameraeinstellung/Perspektive: Welche Kameraeinstellung bzw. Perspektive ist gewählt und welche (Nähe-)Beziehung ergibt sich daraus zur dargestellten Szenerie (Frage nach Totale/Vogelperspektive, Halb-totale, Detail oder Nahaufnahme und der damit verbundenen Beobachterrolle als unbetieilt Überblickender, stark angesprochener Augenzeuge, Beteiligter etc.)?

Bildausschnitt: Welche Ansicht auf Bildobjekte und Kontexte ermöglicht der Bildausschnitt und welche Beziehung lässt sich so zu beidem aufbauen?

Bewegungen: Welche Bewegungen bzw. Dynamiken lassen sich durch bestimmte und imaginäre Linienführungen (Vektoren) im Bild erkennen?

Beziehung der Akteure: Wie ist die vermeintliche Beziehung zwischen dargestellten Akteuren durch Körperhaltung, Größenverhältnisse und Positionierungen auf der Fläche begründbar?

Nähe/Distanz: Wie organisiert die realisierte Verteilung der Bildobjekte im Bildraum Nähe und Distanz, Zugehörigkeiten und Abgrenzung der Elemente zueinander?

Die beiden Kodierheuristiken wurden in Atlas.ti in Listen an Codes übertragen und diese in einem ersten Schritt auf das Einstiegsdokument bezogen. Atlas.ti generiert also weder automatisch Codes noch überträgt es selbstläufig bereitgestellte analytische Paradigmen auf die geladenen Dokumente. Stattdessen bietet es eine nützliche „Arbeitsoberfläche“ (Diaz-Bone/Schneider 2004, 469), mit deren Hilfe Textstellen und Bildausschnitte manuell und nacheinander einzeln in ihrer modalen Form ausgewählt und mit Codes versehen werden können. Der Vorteil, alle analytischen Operationen in dieser einen Oberfläche gebündelt durchzuführen und zu dokumentieren, wurde besonders in dem mit der Analyse befassten Team zum wichtigen Instrument, um kooperativ die Arbeit aufzuteilen und anschlussfähig fortzuführen. Um diese Erleichterung jedoch einzulösen, mussten die beteiligten Kodierer den vom Programm gestellten Anforderungen entsprechen und sowohl die technologischen Einstellungen als auch die Kodierregeln beachten. Die analytischen Optionen wurden also instrumentell effektiv, indem das diskursanalytische Vorgehen an die von Atlas.ti angebotenen Verfahren angepasst wurde.<sup>4</sup> Umgekehrt wurde Atlas.ti in der Diskursanalyse gezielt anders als vorprogrammiert gebraucht. Dieser ‚kontrollierte Missbrauch‘ (vgl. ebd., 474) ist erforderlich, weil die Analyse nicht, wie entsprechend der Grounded Theory eingeplant, in eine gegenstandsnahe erklärende Theorie sozialen Lebens mündete, sondern diskursiv generierte Frames als Deutung sozialer Wirklichkeit rekonstruieren wollte. Entsprechend hatten die eingeführten Codes nicht den Status von Theorieelementen. Vielmehr dienen sie in einem ersten Schritt dazu, heuristisch materiale Ausschnitte in den Diskursfragmenten zu identifizieren, in denen diskursiv relevante Bedeutungen textualisiert bzw. visualisiert wurden. In einem zweiten Schritt wurden die so am Material gewonnenen Codes miteinander verknüpft, um Aussagen über die Struktur der diskursiven Deutungsmuster treffen zu können. Mithilfe von Atlas.ti konnten also mit dem Gliedern, Auszeichnen und Kategorisieren diskursiver Ausschnitte einerhand und dem konzeptuellen Reflektieren und Schließen andererhand schon während des Bearbeitens des Einstiegsdoku-

---

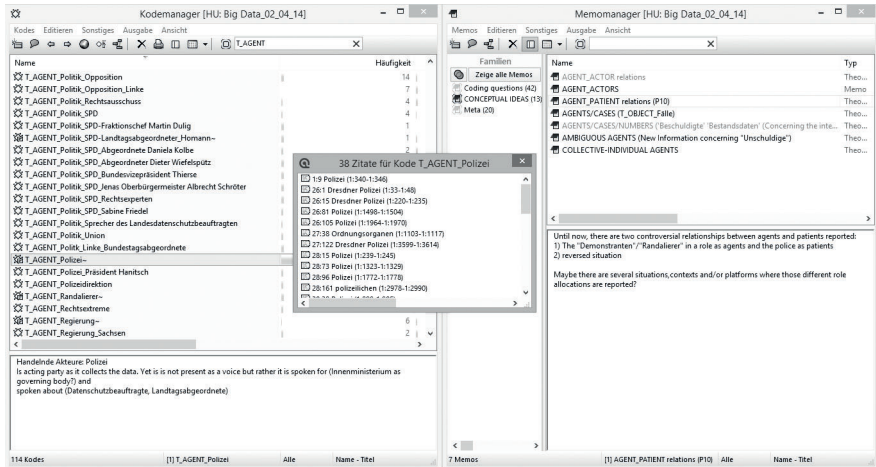
<sup>4</sup> So erfordert Atlas.ti zum Beispiel, dass die Analyse in der Einheit einer so genannten Hermeneutic Unit (HU) stattfindet, in der alle Dokumente und alle während der Analyse erzeugten Daten und Dokumente im Zusammenhang bleiben.

ments zwei einander bedingende und formierende analytische Prozeduren durchgeführt werden. Der Einsatz der QDA-Software Atlas.ti unterstützte die Diskursanalyse dabei in der folgenden Weise:

*Erstens* wurde das Kodieren von Atlas.ti dadurch erleichtert, dass einzelne Passagen über die dazu vergebenen Codes und vice versa aufgefunden wurden. Der analytische Wechsel zwischen deduktiven Hypothesen und induktiv gezogenen interpretativen Schlüssen wurde *zweitens* dadurch erleichtert, dass Codes mehrfach vergeben, variabel umfangreiche Ausschnitte mit beliebig vielen Codes versehen und Codes miteinander in verschiedener Hinsicht in Beziehung gesetzt werden konnten. So war es möglich, analytisch gleichrangige Codes nebenzuordnen, in Hierarchien über- bzw. unterzuordnen, Codes zur Bedingung für andere Codes zu erklären und zu definieren, ob zwei Codes als konzeptuell gleich oder konträr zu behandeln sind. Ebenso konnten Codes in so genannten Familien gruppiert und mit allgemeineren Codes zusammenfassend benannt werden. *Drittens* erlaubte Atlas.ti mit fortschreitendem Durcharbeiten des Materials, die mit einem Code belegten Stellen aufzulisten, um zu prüfen, ob der Code konsistent an konzeptuell vergleichbare diskursive Aussagen vergeben wurde. Andernfalls konnten in Atlas.ti *viertens* die hergestellten Beziehungen von empirischem Material und analytischen Konzepten revidiert werden. Entsprechen war es *fünftens* möglich, die Codes selbst auszudifferenzieren. *Sechstens* konnten in Atlas.ti Memos verfasst werden, um diese schrittweise und gerichtete analytische Rekonstruktion zwischen Diskursfragmenten und konzeptuellen Vorschlägen zu reflektieren. *Siebtens* erfasste Atlas.ti quantitative Informationen zum interpretativen Vorgehen und machte so deskriptive Daten zur Worthäufigkeit, zur Zahl der Codes, der kodierten Stellen, der Memos und ihrer Verknüpfungen miteinander verfügbar (Abb. 3).<sup>5</sup>

---

<sup>5</sup> Hinzu kommt u. a. die Option des Datenexports in XML, sodass die Informationen in anderen Programmen genutzt werden können.



**Abbildung 3:** Screenshot des Bearbeitungsfensters von Atlas.ti: Quellen für einen Code und Memo. Quelle: eigene Darstellung.

Nach dem Kodieren des Einstiegstextes waren, so zeigte es die von Atlas.ti generierte Übersicht, Codes für die textuellen und visuellen Kontextualisierungen von Akteuren, Ereignissen, getätigten Aktionen, die dabei benutzten Entitäten, die sich einstellenden Konsequenzen in ihrer zeitlichen Abfolge und die Orte der Ereignisse besonders häufig vergeben und ausdifferenziert. Auf dieser basalen Ebene wurde zudem ein Code für Zahlenwerte in Bezug zu den behandelten Entitäten eingeführt, da in dem Bericht an mehreren Stellen zahlenmäßige Verhältnisse thematisiert wurden. Entsprechend wurde ebenso ersichtlich, welche Aspekte, die gemäß der Kodierheuristiken potentiell kontextualisierbar gewesen wären, wenig oder gar nicht kontextualisiert worden sind und entsprechend im bisherigen Material wenig oder nicht kodiert werden konnten.

Gemäß der Verfahrenslogik der Grounded Theory, an der sich die diskursanalytische Rekonstruktion und der Einsatz von Atlas.ti orientierte, wurden auf Grundlage der Belege und bisherigen Befunde konzeptuelle Fragen aufgestellt und als Memo hinterlegt:

**Problematisierung:** Inwiefern wird das Sammeln der Mobilfunkdaten diskursiv zum Problem gemacht, welche Ursachen/Gründe, welche Verantwortlichkeiten und welche Konsequenzen werden diskursiv behandelt?

**Diskurspositionen:** Welche Akteure werden eingeführt und wie treten sie in dem Diskurs in Sprecherpositionen, Adressatenpositionen oder in diskursiv konstituierten Subjektpositionen auf (z. B. als Randalierer, kriminelle Vereinigung, Demonstranten, Amtsträger, Opposition, politisch Verantwortliche, Justiz)?

Relationen: Wie kann die Beziehung zwischen den verschiedenen erwähnten bzw. auftretenden Akteuren (z. B. Innenminister Ulbig, Polizeipräsident Hanitsch, Aktionsbündnis Nazifrei!), den fraglichen Daten (z. B. Verbindungsdaten, Bestandsdaten, Gesprächsinhalten) und den Forderungen (z. B. Datenlöschung, Aufklärung, Regierungserklärung, Rücktritt) konzeptualisiert werden?

Legitimität: Wie werden die Rechtmäßigkeit der Aktionen und die Verantwortlichkeiten der Akteure diskursiv erfasst (z. B. als gerichtlich gestattet, verfahrenstechnisch zu verbessern, demokratiefeindlich, grundgesetzwidrig)?

Ereignisfolge: Welche Ereignisse werden in welchen zeitlichen, verfahrenstechnischen und juristischen Bezügen Gegenstände des Diskurses (z. B. des Neonaziaufmarsches, der Gegen demonstration, der Funkzellenabfrage)?

Um bezüglich dieser Fragen zum diskursiv konstituierten Hergang der Ereignisse, der als Sprecher und in diskursiven Rollen eingebundenen Akteure, der Bewertung der Vorgänge und damit schließlich der diskursiv generierten Deutung des Datensammelns weiteres Material zu kodieren, wurden im nächste Analyseschritt die online und offline publizierten Artikel der *Freien Presse* einbezogen. Im Blick auf die offenen Fragen wurde angenommen, dass die so hinzukommenden Diskursfragmente eine Berichterstattung liefern würden, die zum einen als Tageszeitung chronologisch die Affäre begleitet und zum anderen aufgrund des lokalen Bezugs als Nachrichtenwert die Vorgänge und Sachverhalte relativ detailliert darstellt.

Der konsistente Einbezug der Artikel wird durch Atlas.ti dadurch erleichtert, dass die am ersten Dokument entworfenen Codes auch auf alle folgenden Dokumente bezogen werden können. In der Übersicht nach Abschluss dieses Analyseschrittes zeigte das Programm 922 mindestens einmal vergebene Codes. Besonders häufig vergebene und damit, so die Annahme, diskursiv ausgeprägte Codes bezogen sich auf die Akteure und ihre Attribute, die Gründe und Konsequenzen der Datensammlung, Ereignisse und ihren diskursiven Status, Orte und zahlenmäßige Verhältnisse, während die andere Codes seltener an Passagen gesetzt wurden.

Die rekursiven Beziehungen von Datenmaterial und rekonstruierten Konzepten wurden durch Atlas.ti explizit gemacht und sie wurden in Form von gestaltbaren Netzwerken visualisiert. Für die Analyse wurde diese Möglichkeit genutzt, um die strukturellen Beziehungen zwischen den einzelnen Bedeutungsaspekten aus den verschiedenen medialen und modalen Fragmenten sozusagen als Diskursordnung zu erarbeiten. Ausgehend von ihrer Kookkurrenz in Aussagezusammenhängen von Textpassagen, Bildern bzw. Bildausschnitten oder Videoeinstellungen wurden die Codes in Netzwerken miteinander in Beziehung gesetzt, in Familien kategorisiert und diese wiederum konzeptuell benannt.

Im Blick auf diese Forschungspraxis wird deutlich, dass Atlas.ti nicht dazu eingesetzt werden konnte, die interpretative Kompetenz der Kodierer zu übernehmen. Vielmehr aktualisierte sich das forschungspraktische Potential des Programms erst im deutend-verstehenden Wechsel zwischen den aus sehr verschie-

denen medialen Quellen erhobenen multimodalen Äußerungen und der darauf bezogenen Zuordnung von Aussagen und davon abgeleiteten strukturierten Bedeutungszusammenhängen. Atlas.ti organisierte und förderte den fortlaufenden Wechsel deduktiver und induktiver analytischer Praxis, indem es erstens die konzeptuelle Verknüpfung von Passagen und Kodes fixierte, visualisierte und quantifizierte und so den Rückgriff auf einzelne Passagen und den zahlenmäßigen Überblick erleichtert. Damit konnte jede Interpretation auf spezifische Belege zurückgeführt und so kontrolliert werden. Zweitens gewährleistete Atlas.ti die Offenheit dieses Prozesses, weil es nicht nur das Validieren der diskursanalytischen Befunde leichter machte, sondern auch ihr Revidieren. Die Analyse war so in jeder Phase im Prinzip reversibel, sollten sich im wiederholten Rekurs auf das empirische Material plausible Schlüsse und Folgerungen ziehen lassen.

In diesem Sinn wurden in die Analyse im nächsten Schritt die erhobenen Facebook-Nachrichten, Facebook-Kommentare, YouTube-Kommentare und Twitter-Meldungen einbezogen und dazu in Atlas.ti geladen. Im Kontrast zu journalistischen Darstellungen, so die Annahme, umfassen diese Fragmente eine breitere Palette an Sprechern, die sich in Referenz auf massenmediale Berichte oder offizielle Stellungnahmen positionieren. Dadurch sollte es möglich sein, die bisherigen Kodes im Fortgang der Analyse zu überprüfen und zusätzliche Kontextualisierungen, Akteure und Perspektiven auf die Vorgänge aufzunehmen. Von diesem Analysestand ausgehend werden über weitere Samplingphasen und Kodierstufen zwei hypothetische und miteinander konkurrierende Deutungsmuster von Big Data als entweder außerrechtliches und extensives Instrument staatlicher Überwachung oder als effizientes und präzises Hilfsmittel staatlicher Rechtsausübung analytisch weiter erarbeitet.

## **4 Fazit: Nutzen und Grenzen computergestützter Verfahren im Rahmen einer Methodologie der Analyse transmedialer multimodaler Diskurse**

Ohne Zweifel haben die vergangenen Jahre eine neue Dimension der Verfügbarkeit und Auswertbarkeit großer Datenmengen mit sich gebracht. Entsprechend gestiegen sind das akademische Engagement – nicht zuletzt auch der Forschungsförderung – und die unternehmerischen Bestrebungen, die Auswertungswerkzeuge immer weiter zu verbessern, um von rekurrenten Mustern plausibel auf soziokulturelle Phänomene schließen zu können.

Im Blick auf diese Ambitionen ist es umso wichtiger, anzuerkennen, dass der computergestützten Auswertung großer Datenmengen auch Grenzen gesetzt sind,

vor allem in Bezug auf die Interpretation und Kontextualisierung der Befunde. Nach wie vor sind dazu menschliche Kompetenzen des hermeneutischen Sinnverstehens erforderlich (vgl. Hermanns/Holly 2007). Auch muss anerkannt werden, dass ‚rohe‘ Daten nicht objektiv vorliegen und keine irgendwie natürliche Ressource darstellen, sondern durch Prozesse ihrer Generierung, Speicherung und Interpretation erst zu auswertbaren Daten gemacht werden (vgl. Gitelman/Jackson 2013, 3). Zudem ist zu berücksichtigen, dass in die Auswertungswerkzeuge entsprechend ihrer Rahmenbedingungen Restriktionen und Optionen notwendigerweise einprogrammiert sind und ihre Vorzugsnutzungsmuster so jede Analyse präfigurieren (vgl. Pentzold/Fraas/Meier 2013).

Für die computerbasierte Analyse von sprachlich wie semantisch komplexen Sinnfiguren wie Frames finden sich demgemäß verschiedene Ansätze automatisierter Auswertung, wie etwa das *frame mapping*, eine Methode der Inhaltsanalyse, die computerbasierte Analysen mit multidimensionaler Skalierung kombiniert (vgl. Miller/Riechert 2001). Nachteile dieses Zugangs und ähnlicher Zugänge sind indessen, dass die Worthäufigkeit nicht unbedingt als Indiz für diskursive Präsenz gelten kann, die Interpretation der Befunde lediglich an andere Stellen der Analyse verlagert wird, nur verbales Material Berücksichtigung findet und somit der prinzipiellen Multimodalität der diskursiven Bedeutungskonstitution nicht Rechnung getragen wird. Das Automatisieren diskursanalytischer Verfahren kommt dann insbesondere dort an seine Grenzen, wo die Komplexität ihres Gegenstands sich nicht von vornherein auf ein mit den verfügbaren Computerwerkzeugen bearbeitbares Maß reduziert lässt, sondern im Gegenteil Diskursmuster kommunikationsformadäquat rekonstruiert werden sollen. Konsequenterweise richtet sich die Kritik am Bevorzugen automatischer Textanalyse einmal auf die forschungspraktischen Hürden der unzulänglichen Reliabilität im Erkennen von Humor, Ironie und anderer komplexer zeichenhafter Phänomene sowie kultureller Referenzen. Darüber noch hinaus weist das Problem, inwiefern plausibel von maschinell erfassbaren Einheiten und strukturellen Merkmalen eines Diskurses auf Bedeutungsmuster und soziokulturelle Kontexte insgesamt geschlossen werden kann. Besonders die quantitativ extensiv auswertbaren Elemente wie zum Beispiel Schlüsselbegriffe, Kollokationen, Hyperlinks oder Hashtags bieten nur ansatzweise Anhaltspunkte für ihre kontextuelle Bedeutung. Während zum diskursanalytischen Verstehen zusätzliche Kontextinformationen einbezogen werden müssten und Detailanalysen nötig sind, um die automatisch erstellten Auswertungen zu deuten, befördert das Aggregieren sehr großer Datenmengen sogar die Reduktion und Entkontextualisierung des Materials (vgl. boyd/Crawford 2012, 670f.).

Die naheliegende Lösung für diese methodischen Probleme ist die Kombination von diskursanalytischen Methoden, um so standardisierte und automatisierbare Verfahrensschritte mit einem manuell-verstehenden Vorgehen zu verbinden

(vgl. Spitzmüller/Warnke 2011, 39, Bubenhofer 2013, 109), wie es auch im Modell des Wechselverhältnisses von *close* und *distant reading* in den Beiträgen von Kath/Schaal/Dumm und Lemke/Stulpe präferiert wird. Die Präzision computerbasierter Algorithmen soll hierbei zum systematischen Filtern der Daten gebraucht werden, indem etwa Schlüsselwortsuchen automatisiert und deren Ergebnisse für weitere manuelle Entscheidungen aufbereitet werden. Die wesentliche Frage lautet also, wie es gelingen kann, computergestützte Ansätze zur zuverlässigen Bewältigung von Datenmassen mit qualitativen Methoden zusammenzubringen, um materialbasiert und kontextsensibel analytische Schlüsse ziehen zu können, um also ‚Slow Understanding‘ und Big Data viabel zu verknüpfen. Dazu bleibt die Ausrichtung des Forschungsziels der Hintergrund für die Beurteilung von Nutzen und Grenzen computergestützter Verfahren und damit für Entscheidungen hinsichtlich der zu wählenden methodischen Vorgehensweise.

Für den hier vorgestellten Ansatz heißt dies, dass die Rekonstruktion und Analyse multimodaler Deutungsmuster davon ausgeht, dass sich Diskurse in thematisch, technologisch und institutionell konvergierenden Medienarrangements konstituieren und dementsprechend deren Transmedialität und Multimodalität sowie deren kommunikationsstrukturelle und semiotische Besonderheiten zu berücksichtigen sind. Deshalb ist es analytisch geboten, standardisierte Verfahrensschritte mit einem manuell-verstehenden Vorgehen zu kombinieren, um die Chancen computerbasierter Instrumente für das Erheben und Auswerten der Datenmengen zu nutzen und zugleich die sich ergebenden Herausforderungen des interpretativen Einordnens, Zuordnens und Abstrahierens durch hermeneutische Verstehensprozesse zu adressieren. Ein solcher „blended approach“ (Lewis/Zamith/Hermida 2013, 35) nutzt also die Computerunterstützung, um arbeitsaufwändige Phasen des diskursanalytischen Prozesses, z. B. hinsichtlich Korpuserstellung und Kodierung, zu unterstützen und so humane Ressourcen für die detaillierte Analyse von Sinnzusammenhängen und die Bewertung des Materials zu entlasten. Hierbei sind die folgenden analytischen Bedingungen zu beachten.

*Erstens* wurde für das vorgestellte Methodeninstrumentarium aufgrund der Inhomogenität und Dynamik des empirischen Materials die Entscheidung getroffen, die Analysen im Verfahrensrahmen der Grounded Theory durchzuführen, da diese Sampling-, Kodier- und Kategorisierungsstrategien für ein reflexiv-zirkuläres, interpretatives Vorgehen zur Verfügung stellt, das Datenerhebung und -auswertung systematisch anleitet. Damit wurde zugleich entschieden, mit dem Verfahren des Kodierens zu arbeiten, um das empirische Material aufzuteilen, zuzuordnen, zu verknüpfen und daraus analytische Schlüsse abzuleiten. *Zweitens* musste berücksichtigt werden, dass hermeneutische Prozesse nicht über Gebühr systematisiert werden können. Die Kulturtechnik, mit Software umgehen zu können, erfordert eher reaktives Bedienen als prüfend-verstehendes Durchdringen, wie es hermeneu-

tische Verfahren verlangen. Dies bringt für humane Bediener die Gefahr, im Sinne der Programmlogiken zu reagieren und die Vorzüge qualitativ-verstehender Verfahren gegen analytische Raster und Kodierschemata einzutauschen. *Drittens* war zu beachten, dass der Einsatz von Softwaretools technische Voraussetzungen erfüllen muss und erst durch die Einbindung in komplexere Werkzeuge und Arrangements funktioniert. Je nach Analyseanforderungen musste der Computer durch andere Geräte wie Server und Festplatten und durch Werkzeuge für Bildbearbeitung, Schnittprogramme usw. ergänzt werden. All dies führt vor Augen, dass allein vom Einsatz technischer Hilfsmittel und Verfahren noch keine Erleichterung der Diskursanalyse erwartet werden kann – es bleiben die knappen humanen Ressourcen an Aufmerksamkeit und die Kompetenz verstehend-interpretativer Auseinandersetzung, die sich durch Computerunterstützung nicht übernehmen oder exponentiell steigern ließen.

## Literatur

- Boyd, Danah/Crawford, Kate (2012): Critical questions for Big Data. In: *Information, Communication & Society* 15, 662–679.
- Bubenhof, Noah (2013): Quantitativ informierte qualitative Diskursanalyse. In: Kersten Roth, Carmen Spiegel (Hgg.): *Angewandte Diskurslinguistik*. Berlin: Akademie Verlag, 109–134.
- Bucher, Hans-Jürgen (2013): Online-Diskurse als multimodale Netzwerk-Kommunikation. In: Claudia Fraas, Stefan Meier, Christian Pentzold (Hgg.): *Online-Diskurse. Theorien und Methoden transmedialer Online-Diskursforschung*. Köln: Halem, 57–101.
- Diaz-Bone, Rainer, Schneider, Werner (2004): Qualitative Datenanalyse-Software in der sozialwissenschaftlichen Diskursanalyse. In: Keller, Reiner et al. (Hgg.): *Handbuch Sozialwissenschaftliche Diskursanalyse*. Bd. 2. Wiesbaden: VS, 457–494.
- Fillmore, Charles (1985): Frames and the semantics of understanding. In: *Quaderni di Semantica* 6, 222–254.
- Fillmore, Charles/Baker, Christopher (2010): A frame approach to semantic analysis. In: Bernd Heine/Heiko Narrog (Hgg.): *The Oxford Handbook of Linguistic Analysis*. Oxford: Oxford University Press, 313–339.
- Fraas, Claudia (1996): *Gebrauchswandel und Bedeutungsvarianz in Textnetzen*. Tübingen: Narr.
- Fraas, Claudia/Meier, Stefan/Pentzold, Christian (2013): Perspektiven einer interdisziplinären Diskursforschung. In: Diess. (Hgg.): *Online-Diskurse*. a. a. O., 7–34.
- Fraas, Claudia/Meier, Stefan/Pentzold, Christian/Sommer, Vivien (2013): Diskursmuster – Diskurspraktiken. In: Claudia, Fraas, Stefan Meier, Christian Pentzold (Hgg.): *Online-Diskurse*. a. a. O., 102–135.
- Friese, Susanne (2012): *Qualitative data analysis with ATLAS.ti*. London: Sage.
- Geise, Stephanie, Lobinger, Katharina (Hgg.) (2013): *Visual Framing*. Köln: Halem.
- Gitelman, Lisa/Jackson, Virginia (2013): Introduction. In: Lisa Gitelman (Hg.): *'Raw Data' is an oxymoron*. Cambridge: MIT Press, 1–14.
- Glaser, Barney/Strauss, Anselm (1967/2005): *Grounded Theory*. Bern: Huber.
- Halliday, Michael (1990): *Spoken and written language*. Oxford: Oxford University Press.

- Hermanns, Fritz/Holly, Werner (Hgg.)(2007): *Linguistische Hermeneutik*. Tübingen: Niemeyer.
- Holly, Werner (2011): *Medien, Kommunikationsformen, Textsortenfamilien*. In: Stephan Habscheid (Hg.): *Textsorten, Handlungsmuster, Oberflächen*. Berlin, New York: de Gruyter, 144–163.
- Jenkins, Henry (2006): *Convergence culture*. New York: New York University Press.
- Karpf, David (2012): *Social science research methods in internet time*. In: *Information, Communication & Society* 15, 639–661.
- Konerdig, Klaus-Peter (1993): *Frames und lexikalisches Bedeutungswissen*. Tübingen: Niemeyer.
- Kress, Gunter/van Leeuwen, Theo (2001): *Multimodal discourse*. London: Routledge.
- Kress, Gunter/van Leeuwen, Theo (2006): *Reading images*. London: Routledge.
- Lewis, Seth/Zamith, Rodrigo/Hermida, Alfred (2013): *Content analysis in an era of big data*. In: *Journal of Broadcasting & Electronic Media* 57, 34–52.
- Mahrt, Merja/Scharkow, Michael (2013): *The value of big data in digital media research*. In: *Journal of Broadcasting & Electronic Media* 57, 20–33.
- Manovich, Lev (2012): *The promises and the challenges of big social data*. In: Matthew Gold (Hg.): *Debates in the digital humanities*. Minneapolis: University of Minnesota Press, 460–475.
- Mautner, Gerlinde (2005): *Time to get wired. Using web-based corpora in critical discourse analysis*. In: *Discourse & Society* 16, 809–828.
- Meier, Stefan (2011): *Multimodalität im Diskurs: Konzept und Methode einer multimodalen Diskursanalyse*. In: Reiner Keller et al. (Hgg.): *Handbuch Sozialwissenschaftliche Diskursanalyse*. Bd. 1. Wiesbaden: VS, 499–532.
- Meier, Stefan (2014): *Zur Sozialesemiotik visueller Medienkultur und konvergenter Designpraxis*. Bielefeld: transcript, i. E.
- Meier, Stefan, Pentzold, Christian (2010): *Theoretical Sampling als Auswahlstrategie für Online-Inhaltsanalysen*. In: Martin Welker/Carsten Wunsch (Hgg.): *Die Online-Inhaltsanalyse*. Köln: Halem, 124–143.
- Meier, Stefan et al. (2010): *Auswahlverfahren für Online-Inhalte*. In: Martin Welker, Martin/Carsten Wunsch (Hg.): *Die Online-Inhaltsanalyse*. a. a. O., 103–123.
- Miller, Mark/Riechert, Bonnie (2001): *The spiral of opportunity and frame resonance*. In: Reese, Stephen/Gandy, Oscar/Grant, August (Hgg.): *Framing public life*. Mahwah: Erlbaum.
- Minsky, Marvin (1975): *A framework for representing knowledge*. In: Winston, Patrick (Hg.): *The psychology of computer vision*. New York: McGraw-Hill, 211–278.
- Pentzold, Christian/Fraas, Claudia/Meier, Stefan (2013): *Online-mediale Texte: Kommunikationsformen, Affordanzen, Interfaces*. In: *ZGL* 41, 81–101.
- Pentzold, Christian (2014): *Datenanalysesoftware*. In: Wrana, Daniel et al. (Hgg.): *DiskursNetz*. Berlin: Suhrkamp, 62–63.
- Pêcheux, Michel (1995): *Automatic discourse analysis*. Amsterdam: Rodopi.
- Ruppenhofer, Josef et al. (2010). *FrameNet II: Extended Theory and Practice*. Online: <http://framenet2.icsi.berkeley.edu/docs/r1.5/book.pdf> (03.04.2014).
- Scharloth, Joachim/Eugster, David/Bubenhofer, Noah (2013): *Das Wuchern der Rhizome. Linguistische Diskursanalyse und Data-driven Turn*. In: Dietrich Busse, Wolfgang Teubert (Hgg.): *Linguistische Diskursanalyse*. Wiesbaden: VS, 345–380.
- Sommer, Vivien/Fraas, Claudia/Meier, Stefan/Pentzold, Christian (2013): *Qualitative Online-Diskursanalyse*. In: Claudia, Fraas/Stefan Meier/Christian Pentzold (Hgg.): *Online-Diskurse*. a. a. O., 258–284.
- Spitzmüller, Jürgen, Warnke, Ingo (2011): *Diskurslinguistik*. Berlin, New York: de Gruyter.
- Stubbs, Michael (2001): *Words and phrases. Corpus studies of lexical semantics*. Oxford: Blackwell.
- Ziem, Alexander (2008). *Frames und sprachliches Wissen*. Berlin, New York: de Gruyter.