

Digitale Daten und Methoden als Erweiterung qualitativer Forschungsprozesse: Herausforderungen und Potenziale aus den Digital Humanities und Computational Social Sciences

Franken, Lina

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:

Franken, L. (2022). Digitale Daten und Methoden als Erweiterung qualitativer Forschungsprozesse: Herausforderungen und Potenziale aus den Digital Humanities und Computational Social Sciences. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 23(2). <https://doi.org/10.17169/fqs-22.2.3818>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see: <https://creativecommons.org/licenses/by/4.0>

Digitale Daten und Methoden als Erweiterung qualitativer Forschungsprozesse. Herausforderungen und Potenziale aus den *Digital Humanities* und *Computational Social Sciences*

Lina Franken

Keywords: Digital Humanities; Computational Social Sciences; qualitative Sozialforschung; Forschungsprozess; digitale Methoden; digitale Daten; wissenssoziologische Diskursanalyse

Zusammenfassung: Obwohl die Arbeit mit digitalen Daten und Methoden die qualitative Sozialforschung erweitern kann, hat sich diese bislang wenig hierfür geöffnet. Zur Erleichterung des Einstiegs können bestehende Werkzeuge etwa aus der Informatik und den Datenwissenschaften genutzt werden, müssen aber für die besonderen Bedürfnisse der qualitativen Forschung angepasst werden. Dabei ist es zentral, nicht nur die digitalen Daten, sondern auch die entsprechenden Methoden kritisch zu analysieren und hinsichtlich der eigenen Forscher:innenperspektive zu reflektieren, um fortlaufend Grenzen und Möglichkeiten für die qualitative Sozialforschung auszuloten. Im Beitrag zeige ich Bezüge zu den *Digital Humanities* und den *Computational Social Sciences* auf. Am Beispiel einer wissenssoziologischen Diskursanalyse, die mit verschiedenen digitalen Verfahren realisiert wurde, modelliere ich den qualitativen Forschungsprozess und veranschauliche, welche Schritte digital erweitert werden können. Im Anschluss diskutiere ich zentrale Herausforderungen und leite daraus Potenziale und mögliche Weiterentwicklungen für digitale Daten und Methoden ab. Der Beitrag schafft eine Grundlage, um eine kreative Integration neuer digitaler Elemente in den qualitativen Forschungsprozess zu etablieren. Ich beleuchte damit den Wandel von Forschungsfragen, -feldern, -methoden und Epistemologien.

Inhaltsverzeichnis

- [1. Digitale Daten und Methoden und ihre Anschlüsse an die qualitative Sozialforschung](#)
- [2. Digitale Quellen und Methoden für kultur- und sozialwissenschaftliche qualitative Forschung – ein Widerspruch?
 - \[2.1 *Digital Humanities*, *Computational Social Sciences* und digitale Methoden\]\(#\)
 - \[2.2 Kultur- und sozialwissenschaftliche Ausgangspunkte: Quellengrundlagen qualitativer Forschung\]\(#\)](#)
- [3. Modellierung qualitativer Forschungsprozesse anhand einer Diskursanalyse](#)
- [4. Erweiterung des Forschungsprozesses in Auswahl, Erhebung, Aufbereitung und Auswertung](#)
- [5. Herausforderungen und Potenziale
 - \[5.1 Herausforderungen\]\(#\)
 - \[5.2 Potenziale und Weiterentwicklungen\]\(#\)](#)
- [6. Fazit: Erweiterungen durch digitale Daten und Methoden in qualitativen Forschungsprozessen](#)

[Danksagung](#)

[Literatur](#)

[Zur Autorin](#)

[Zitation](#)

1. Digitale Daten und Methoden und ihre Anschlüsse an die qualitative Sozialforschung

Menschen produzieren zunehmend Daten in ihrem Alltag (KITCHIN 2014a; LUPTON 2015). So wurde geschätzt, dass im Jahr 2020 global rund 190 Millionen E-Mails pro Minute versendet, 1,3 Millionen Log-Ins bei Facebook getätigt und eine gute Millionen US-Dollar im Online-Shopping ausgegeben wurden (LEWIS 2020). Aber nicht nur die Produktion von Daten steigt durch die Digitalisierung rasant an, auch das menschliche Handeln ist durch die Sammlung, Auswertung und Nutzbarmachung von Daten beeinflusst, Analoges und Digitales ist in Relationen verwoben (BECK 2019 [2015]; BOELLSTORFF 2016). Gleichzeitig wird das "metrische Wir" (MAU 2017) postuliert, das stärker als je zuvor durch Zahlen abzubilden sei. Im wissenschaftlichen Diskurs ist dementsprechend das Narrativ wirkmächtig, dass im Zeitalter des Internets eine schier unendliche Menge von Text-, Bild-, Audio- und Videodaten zur Verfügung stehe, die ihrerseits eine Auswertung nahelegten. Gleichzeitig sei der Zugang zu großen Datenmengen im Digitalen stark vereinfacht und eine schnelle und präzise Analyse mit digitalen Verfahren möglich (DOURISH & GÓMEZ CRUZ 2018; MAYER-SCHÖNBERGER & CUKIER 2013). Daneben werden bestehende qualitative Forschungspraxen infrage gestellt, da sie vielfach nicht mehr ausreichen, um digital durchdrungene Gesellschaften zu analysieren. Zahlreiche Herausforderungen in der Nutzung und Weiterentwicklung digitaler Methoden und Daten müssten angegangen werden, um die für Kultur- und Sozialwissenschaften zweifelsfrei vorhandenen Potenziale ausschöpfen zu können (MOSCONI et al. 2019; PAULUS, JACKSON & DAVIDSON 2017; RUPPERT, LAW & SAVAGE 2013). Es ist, so meine Grundthese im vorliegenden Beitrag, von einer Veränderung und Erweiterung der qualitativen Forschung auszugehen, wenn digitale Verfahren angewendet und digitale Daten verarbeitet werden. Im Folgenden gehe ich deshalb einerseits der Frage nach, welche konkreten Daten und Verfahren qualitativ Forschende für welche Settings nutzen könnten. Andererseits hinterfrage ich, wie sich qualitative Forschungsdesigns verändern, wenn einzelne Schritte digital realisiert werden. [1]

In der Verwendung und Weiterentwicklung digitaler Methoden besteht ein Potenzial für qualitative Forschungspraxen. Es geht dabei um mehr als um digitale oder virtuelle Ethnografie, durch welche eine Entwicklung ethnografischer Methoden für die Erhebung in digitalen Forschungsfeldern vorgenommen wird (HINE 2015; KNOX & NAFUS 2018; PINK et al. 2016). Im Mittelpunkt steht auch nicht die Nutzung eines einzelnen Tools zur digitalen Quellenauswertung, die insbesondere für *Computer-Assisted Qualitative Data Analysis Software* (CAQDAS) reflektiert wird (EVERS, CAPRIOLI, NÖST & WIEDEMANN 2020; GIBBS, FRIESE & MANGABEIRA 2002; MacMILLAN 2005). Vielmehr spreche ich mich für ein Verständnis von digitalen Methoden aus, bei dem tendenziell die starren Grenzen zwischen qualitativen und quantitativen Analysen von Gesellschaft aufgehoben werden. Dabei sollen die qualitativen Perspektiven weder aufgegeben noch untergeordnet werden. Vielmehr frage ich: Wie sind methodische Verfahren und digitale Möglichkeiten ineinander verschränkt und durcheinander bedingt? Welche digitalen Daten und Methoden können in

qualitativen Forschungsperspektiven zielführend genutzt werden, und wie müssen diese weiterentwickelt werden? [2]

Im Folgenden werde ich zunächst erläutern, was unter *Digital Humanities* (DH) und *Computational Social Sciences* (CSS) verstanden wird, wie sich beide unterscheiden bzw. überlagern (Abschnitt 2.1). Danach umreiße ich qualitative Sozialforschung aus kulturwissenschaftlicher Perspektive und gehe darauf ein, wie sich deren Quellengrundlage durch digitale Daten erweitert (Abschnitt 2.2). Am Beispiel einer eigenen Untersuchung zur Digitalisierung im Gesundheitsbereich in Anlehnung an KELLERs wissenssoziologische Diskursanalyse (2011 [2005]) modelliere ich qualitative Forschungsprozesse (Abschnitt 3) und zeige deren Erweiterungen durch digitale Daten und Verfahren auf (Abschnitt 4). Im Anschluss systematisiere ich zentrale Herausforderungen und identifiziere Potenziale digitaler Methoden (Abschnitt 5). Abschließend ziehe ich Schlussfolgerungen zur Verwendung digitaler Daten und Methoden und deren Weiterentwicklung und Anpassung an kultur- und sozialwissenschaftliche Bedürfnisse (Abschnitt 6). [3]

2. Digitale Quellen und Methoden für kultur- und sozialwissenschaftliche qualitative Forschung – ein Widerspruch?

2.1 *Digital Humanities*, *Computational Social Sciences* und digitale Methoden

In den Geistes-, Kultur- und Sozialwissenschaften haben sich in jüngerer Zeit digitale Verfahren mit jeweils unterschiedlichen Schwerpunkten etabliert, wobei dies für die Geisteswissenschaften als DH, für die Sozialwissenschaften als CSS bezeichnet wird. Beide werden verstanden als sogenannte "Brückenfächer" (SAHLE 2015, Kap. 2) oder "Transformationswissenschaften" (JANNIDIS, KOHLE & REHBEIN 2017a, S.XI), die es insbesondere ermöglichen sollen, große Mengen von Daten auszuwerten. Hierfür wird neben digitalen Daten auf neue, digitale Untersuchungsmethoden zurückgegriffen, wobei aufgrund der Anschlussfähigkeit an informatische Verfahren vor allem mit quantitativen Ansätzen gearbeitet wird. Ein Unterschied zwischen den DH und CSS liegt in dem verwendeten Quellenmaterial. Während in den DH vorwiegend mit historischen, retro-digitalisierten Materialien (wie beispielsweise historischen Dokumenten oder Gemälden) gearbeitet wird, stellen Forschende in den CSS gegenwärtig produzierte, digitale Daten in den Mittelpunkt (wie etwa Beiträge auf Twitter oder aus der Medienberichterstattung) (FRANKEN 2022a). Methodologische Entwicklungen und Überlegungen sind jedoch vergleichbar. In beiden Fällen geht das Potenzial weit über die Zusammenführung von *Digital* (im Sinne von Informatik und Informationswissenschaften) und *Humanities* oder *Social Sciences* (im umfassenden Sinne von Geistes- und Kultur- bzw. Sozialwissenschaften) hinaus: [4]

Mit digitalen Methoden wird an den Grenzen etablierter qualitativer Methoden – wie Interview, Beobachtung oder die Analyse audiovisueller Quellen – angesetzt, wie in Abschnitt 4 aufgezeigt wird. Dadurch verändern sich die Art und Weise,

Quellen zu generieren, zu analysieren und Ergebnisse darzustellen (BERKELAAR & FRANCISCO-REVILLA 2018, S.17). Dabei entstehen neue Forschungsmöglichkeiten, wobei SAHLE (2015, Kap. 2) die Verfahren in "low end" und "high end" differenziert hat. Diese Unterteilung impliziert jedoch eine Abwertung der Nutzung von generischen Werkzeugen, denn nicht nur die "Entwicklung spezieller neuer Werkzeuge und Methoden"(a.a.O.) kann als "high end" DH verstanden werden, sondern auch die Erweiterung der eigenen Methodik um DH-Perspektiven geht weit über das hinaus, was als "low" bezeichnet würde. SVENSSON (2016, S.16) unterschied minimale und maximale Lesarten von DH, und damit in "application of computer technology to traditional scholarly work (a minimalist reading) and changing the substance of humanistic matter (a maximalist reading)". Beide machten aber zu wenig deutlich, dass grundlegend unterschiedliche Tätigkeiten mit digitalen Methoden gefasst werden können. Im Folgenden soll deshalb getrennt werden in einerseits eine Erweiterung des methodischen Werkzeugkastens und andererseits einen Experimentierkasten, durch den Neues geschaffen werden kann. Als Werkzeugkasten sind sowohl die DH als auch die CSS zu verstehen, weil durch die Verwendung grundlegender Werkzeuge – also solcher, die in der Informatik bereits etabliert sind und mit denen einfache Operationen durchgeführt werden – Erleichterungen geschaffen werden, indem digitale Methoden, Verfahren, Tools und Standards etabliert und weiterentwickelt werden. So sind etwa häufig gemeinsam verwendete Begriffe in Texten leicht zu identifizieren ("n-Gramme", BUBENHOFER 2017) oder große Datenmengen aus dem Netz zu speichern ("Crawling", BARBARESI 2019, S.30). Mit dem Verständnis als Experimentierkasten entsteht in DH und CSS durch Kollaborationen und Offenheit Neues: Experimente, die erst durch Ausbau der Werkzeuge und deren Kombination für neue Fragestellungen ermöglicht werden. Durch maschinelles Lernen können Forschende beispielsweise implizite Aussagen historischer Quellen herausarbeiten (NELSON 2020) oder aus einer großen Menge Bilder Bildsprachen für spezifische Tageszeiträume durch Anordnung nach Farben oder Helligkeiten deutlich machen (MANOVICH 2020). Diese Experimente werden jedoch erst möglich, nachdem die Werkzeuge auf infrage kommende Forschungsinteressen angepasst, verändert und kombiniert wurden. Durch diese Differenzierung in Werkzeug- und Experimentierkasten wird deutlich, dass ein breites Spektrum an unterschiedlichen Tätigkeiten mit digitalen Methoden gefasst wird. [5]

Forschende in den DH und CSS fragen danach, wie mit digitaler Technologie und Konzepten der Informatik geistes- und sozialwissenschaftliche Fragestellungen verwendet und angepasst werden können. Solche Methoden, Verfahren und Modelle werden als computationell bezeichnet, denn es geht um die Sammlung, Aufbereitung und Analyse von digitalen Daten mithilfe von Computerprogrammen oder Skripten, die der Untersuchung (nicht nur) der digitalisierten Gesellschaft dienen. Für die Geistes-, Kultur- und Sozialwissenschaften ergibt sich damit die Chance, neu darüber nachzudenken, wie digitale Verfahren kombiniert und verbessert werden können und sollen, um aufkommende Fragestellungen zu beantworten. [6]

Beide Transformationswissenschaften institutionalisieren sich gerade erst. Die DH wurden mit Forschungszentren, Servicestellen in Bibliotheken, Professuren und Studiengängen in den letzten Jahren zunehmend zu einem feststehenden Forschungszusammenhang entwickelt, im Blog des deutschsprachigen Verbandes wurden bereits mehr als 100 Professuren in diesem Kontext gezählt (SAHLE 2019). Einführungswerke in deutscher (JANNIDIS, KOHLE & REHBEIN 2017b) und englischer (SCHREIBMAN, UNSWORTH & SIEMENS 2004) Sprache erscheinen seit mehr als 15 Jahren, als Vorläufer wurden insbesondere die "Humanities Computing" (McCARTY 2005) seit den 1980er Jahre angesehen. Demgegenüber existiert für die CSS (noch) keine deutschsprachige Fachgesellschaft, Konferenzen und auch Professuren und Studiengänge sind erst vereinzelt vorhanden. Die internationale Forschungsgemeinschaft ist diverser und wird stärker aus den Naturwissenschaften geprägt, zentrale Forschungspersönlichkeiten kommen aus der Physik und Mathematik (CIOFFI-REVILLA 2018; CONTE et al. 2012). In den letzten Jahren wurden verstärkt Beiträge aus der Politikwissenschaft (BLÄTTE, BEHNKE, SCHNAPP & WAGEMANN 2018) und Soziologie (LUPTON 2015; MARRS 2017) veröffentlicht. Darüber hinaus bestehen konkurrierende Bezeichnungen wie "Digital Methods" (ROGERS 2019), "Social Research in the Digital Age" (SALGANIK 2018), "Digital Methods for Social Science" (SNEE, HINE, MOREY, ROBERTS & WATSON 2016) oder "Humanwissenschaft des Digitalen" (UNIVERSITÄT KLAGENFURT 2021), bei denen nicht immer trennscharf Methode, Datengrundlage und Untersuchungsgegenstand voneinander abgegrenzt wurden. Werden diese Einzelentwicklungen im Gesamten betrachtet, sind sie weit davon entfernt, marginal oder nur als kleines Fach (ARBEITSSTELLE KLEINE FÄCHER 2018) zu existieren. Sie bewegen sich in die Mitte der Geistes-, Kultur- und Sozialwissenschaften. Forschende einzelner Disziplinen differenzieren sie weiter aus. Das facettenreiche Forschungsfeld ist heterogen und unabgeschlossen, es wird experimentiert und parallel erarbeitet. [7]

Die Erarbeitung von DH und CSS befindet sich in vielen Bereichen noch am Anfang. Entwicklungen seien dabei oft projektgetrieben, hielt SAHLE (2015) fest. Es existiert eine unüberschaubare Zahl von Tools und Daten, die wenig systematisch miteinander vernetzt sind und oft nur in Form von Prototypen erstellt wurden. Aktuell bewegen sich die Forschungen zwischen digitaler Handarbeit und *Big-Data*-Analysen: Unter digitaler Handarbeit verstehe ich in diesem Zusammenhang die Arbeit mit generischen Tools, also in den alltäglichen Gebrauch übergegangene Software wie etwa Textverarbeitung oder Internetbrowser. Dazu kommt die manuelle Arbeit an Datensätzen, die wenig bis gar nicht computationell realisiert wird, etwa in Literaturdatenbanken, Listen von Quellen oder auch der computergestützten Annotation in CAQDAS-Umgebungen (FRANKEN 2020a). Mit *Big-Data*-Analysen werden hingegen einfache Zugänge versprochen, doch viel erfordert enormen Aufwand, und manchmal sind die Ergebnisse frustrierend. In den Geistes-, Kultur- und Sozialwissenschaften besteht in der Breite zudem ein großes Unwissen über digitale Methoden. Wenn bereits digitale Verfahren genutzt werden, wird meist auf einzelne Tools zurückgegriffen. So stellte etwa NELSON (2020) ihre Analyse aufgrund eines *Topic Modeling* (zum Verfahren vgl. Abschnitt 4) an, es erfolgte aber keine

manuelle Lektüre und Annotation der entsprechenden Texte mehr. Das geht mit Komplexitätsreduktion einher, die nicht in Pragmatismus umschlagen darf. Aktuell wird ausgehandelt und ausgelotet, wo digitale Methoden und wo digitale Handarbeit notwendig und sinnvoll sind. Um den Charakter dieses Prozesses zu greifen, hilft ein Verständnis als "liminal space or contact zone" (SVENSSON 2016, S.5). Liminalität wurde, ausgehend von Forschungen zu Ritualen (TURNER 2005 [1969]; VAN GENNEP 1999 [1909]), verstanden als eine Phase des Übergangs, in der etwas Neues dezidiert gestaltet wird. Klassisch angewendet auf die großen Umbrüche im Lebenslauf wie Schulanfang oder Hochzeit gestalten Menschen Liminalität bewusst, um die liminale Phase zwischen alt und neu aktiv zu begehen und sichtbar zu machen – bezogen auf die digitalen Daten und Methoden ist dies der Übergang von etablierten Methoden der Geistes-, Sozial- und Kulturwissenschaften hin zu computationell erweiterten Ansätzen. Wenn qualitative Forschungsperspektiven Teil dieser Entwicklung sein sollen, müssen Forschende sie aktiv in der liminalen Phase einbringen und den Umbruch sichtbar gestalten, um einen Platz im neu Entstehenden zu finden. [8]

2.2 Kultur- und sozialwissenschaftliche Ausgangspunkte: Quellengrundlagen qualitativer Forschung

Qualitative Forschung ist in einem kulturwissenschaftlichen Verständnis geprägt durch die Frage nach Bedeutungen in ihren Kontexten und dem Konstruktionscharakter von Kultur (LINDNER 2003; RECKWITZ 2004). Ich argumentiere aus Perspektive der empirischen Kulturwissenschaft innerhalb des Diskurses der qualitativen Sozialforschung. Mit einem weiten, relationalen Kulturbegriff ist dabei das Sinnverstehen zentral; mit Clifford GEERTZ (1983 [1973]) ist Kultur das "selbstgesponnene Bedeutungsgewebe" (S.9), in welches Menschen verstrickt seien. Gegenwart ist historisch geworden und im Wandel, sodass kulturellem Erbe und Erinnerungspraktiken große Bedeutung zukommen. Allerdings stehen im Mittelpunkt des Forschungsinteresses qualitative Forschung die Gegenwart und jüngste Vergangenheit und konkrete Praktiken – soziales Handeln in den damit verbundenen Narrativen und Wissensrepräsentationen – sowie Machtstrukturen und kulturelle Codes als normative Setzungen, welche durch Kulturanalyse explizit gemacht werden (BECK 2009; LEIMGRUBER 2013). Diese Zusammenhänge hat RECKWITZ (2008, S.201) als "Praxis/Diskurs-Formationen" gefasst, in welchen sowohl symbolische Strukturen als auch subjektive Interpretationen berücksichtigt werden. [9]

In diesem Verständnis setzen Forschende Beziehungen und Dynamiken zwischen Akteur:innen, ihre Interaktion mit Welt als Ereignishaftigkeit, zentral. Auch nicht-menschliche Akteur:innen haben eine Handlungsmacht (BARAD 2003). Dies gilt besonders im Digitalen, denn hier konstruieren technische Geräte und Algorithmen Sinn aktiv mit (REICHERT & RICHTERICH 2015). Die Lesart von Kulturwissenschaften als qualitative Sozialforschung umfasst gleichzeitig eine klare Positionierung in der kritischen Kulturanalyse, durch welche die Repräsentation von Eigenem und Fremdem infrage gestellt wird. Postkoloniale

und feministische Theorieansätze gerade im Kontext der *Science and Technology Studies* (BECK, NIEWÖHNER & SØRENSEN 2012; SISMONDO 2010) sind für diese Forschungen zentral. Kulturwissenschaften zielen damit "towards new cultural insights, as well as towards new ways of thinking about, practicing and presenting ethnographic analyses" (FORTUN et al. 2014, S.635). Kultur zeigt sich in Sprache, Wissen, Symbolen, Materialitäten und Handlungen. Dementsprechend verwenden Kulturwissenschaftler:innen vielfältige Quellen. Mit digitalen Daten wird diese Quellengrundlage erweitert. [10]

In Anlehnung an BAUR und GRAEFF (2021) unterscheide ich für die Geistes-, Kultur- und Sozialwissenschaften in forschungsinduzierte und prozessproduzierte Quellen. Forschungsinduzierte Quellen entstehen durch wissenschaftliche Erhebungen: Surveys oder Interviews, Fotos, Videos oder Beobachtungsprotokolle. Prozessproduzierte Quellen (BAUERNSCHMIDT 2014) sind schon vorliegende Korpora wie etwa historische Archivalien, administrative Daten, Objekte oder Gebäude und nicht zuletzt Produkte der Popularkultur wie Romane, Filme, Comics oder Lieder. Teilweise sind diese Daten bereits als *digital heritage* (SMITH 2013) oder in *digitalen Editionen* (KLUG 2021) aufbereitet worden. Dazu kommen neuartige prozessproduzierte Daten, insbesondere *Trace*-Daten und Daten aus den sozialen Medien. *Trace*-Daten sind Spurendaten wie etwa der Aufenthaltsort, der von einem Smartphone übermittelt wird, oder Suchbegriffe, die im Internet eingegeben werden. Sie sind prozessproduziert, sie entstehen eher beiläufig und manchmal ohne das Wissen oder die Reflexion derjenigen, die sie erzeugen. *Social-Media*-Daten werden zwar von Nutzenden bewusst hergestellt, da die Beiträge wie Text oder Fotos mit anderen Nutzer:innen einer Plattform aktiv geteilt werden, jedoch ist den Produzierenden der Daten dabei nicht zwangsläufig die langfristige Speicherung und das Teilen mit Dritten wie den Plattformbetreibenden oder Forschenden transparent bzw. bewusst. Auch wenn das Bewusstsein über die bei der Nutzung von digitalen Plattformen produzierten Daten gestiegen ist, bleibt die Auswertbarkeit vielen Menschen unbekannt (KROPF 2019; PEETZ 2021). Nicht von Forschenden erzeugt, sind diese ebenfalls Teil der prozessproduzierten Daten, wie Katharina KINDER-KURLANDA (2020) festhielt. In die Kategorie der neuartigen prozessproduzierten Daten fallen auch diverse Internetdaten, die etwa im Medienwandel aus anderen prozessproduzierten Quellen entstehen können (wie beispielsweise Medienberichte), aber auch Blogs oder Foren. Nicht alle diese Daten sind mithilfe digitaler Verfahren direkt zu verarbeiten, vielmehr ist eine Unterscheidung in Datentypen notwendig. Textdaten sind etwa in Form von Medienberichten oder Texten aus dem Internet zu finden, aber auch als Interviewtranskripte oder Beiträge in den sozialen Medien. Bilddaten können aus den sozialen Medien extrahiert, aber auch in Erhebungen selbst erzeugt oder zur Dokumentation von Objekten und Gebäuden in Archiven aufbewahrt werden. Schließlich können audiovisuelle Daten wiederum als Teil von sozialen Medien oder Internetdaten erhoben werden, etwa eingebunden in eine Webseite, oder im Rahmen einer Erhebung entstehen. Zudem werden zahlenförmige Daten wichtig, wie sie etwa in *Trace*-Daten vorliegen. [11]

Für geistes-, kultur- und sozialwissenschaftliche Forschung insbesondere der empirischen Kulturwissenschaft wird in der Regel eine multimodale Quellengrundlage herangezogen. Auch ohne digitale Methoden müssen qualitativ Forschende aus dieser Quellenvielfalt je nach ihrer Fragestellung auswählen (KOCH & FRANKEN 2020), um eine aussagekräftige und zugleich handhabbare Quellenmenge herauszuarbeiten. Oft experimentieren sie methodologisch, um die untersuchten Phänomene in ihrer Komplexität überhaupt fassen zu können, wie sich in der digitalen Ethnografie als Überführung und Anpassung ethnografischer Methoden in digitale Umgebungen zeigt oder bei der "Walkthrough Method" (LIGHT, BURGESS & DUGUAY 2017), eine Methode zur Erforschung von Apps, bei der systematisch die Schritte der Registrierung und Nutzung von Apps dokumentiert werden. Für die Realisierung digitaler Verfahren bedeutet dieses Auswählen aktuell, dass in der Regel nach Datentypen unterschieden wird. So konzentrierte ich mich in der im vorliegenden Beitrag referenzierten eigenen Diskursanalyse auf Textdaten, da diese bereits gut anhand unterschiedlicher digitaler Verfahren erhoben und ausgewertet werden können. Diese forschungspragmatische Operationalisierung soll nicht generalisierend verstanden werden, denn grundsätzlich sind multimodale Quellengrundlagen selbstverständlich auch für digitale Daten und Verfahren zu berücksichtigen. Wie in klassischen Feldzuschnitten auch, ist die Konzeption einer Studie teilweise begründet durch Zugänge und Unzugänglichkeiten (MARCUS 2009; RIEKER, HARTMANN SCHAELLI & JAKOB 2020). Viel wichtiger als in anderen Erhebungen sind für die Umsetzung digitaler Verfahren allerdings technische Möglichkeiten, mit denen der Feldzugang neu strukturiert wird und durch die pragmatische Lösungen erforderlich werden, um das Quellenmaterial interpretierbar zu machen. Damit komme ich zurück auf die Eingangsbeobachtung der nur scheinbar verfügbaren Quellen, die in einer Systematik für die Nutzung in qualitativen Perspektiven gerade im Digitalen neu strukturiert werden müssen. Ein spezifisches Forschungsprogramm ist notwendig, um Herausforderungen und Potenziale angehen zu können. [12]

3. Modellierung qualitativer Forschungsprozesse anhand einer Diskursanalyse

Um digitale Verfahren verwenden zu können, müssen die Schritte im Forschungsprozess zunächst operationalisiert werden. Dabei hilft eine Modellierung des Forschungsprozesses (FRANKEN 2020b), die ich mit McCARTY als Abstrahierung verstehe: "A 'model' I take to be either a representation of something for purposes of study, or a design for realizing something new" (2004, S.255). McCARTY folgte mit seiner Unterscheidung in ein "model of" und ein "model for" Clifford GEERTZ (1983 [1973], S.93). Auch wenn GEERTZ gerade die Verwobenheit der beiden Ebenen für Modelle von Kultur betonte, ist eine analytische Trennung hilfreich. So unterschied der DH-Literaturwissenschaftler JANNIDIS beispielsweise nicht in "of" und "for", sondern nannte lediglich eine "of"-Funktion: "Ein Modell ist ein Modell von etwas. Ein Modell ist somit immer eine Abbildung; es repräsentiert etwas, ein Ding, einen Begriff, in einem anderen Medium, z.B. Sprache, Bild, oder Ton" (2017, S.100). FLANDERS und JANIDIS (2016, S.230) differenzierten Prozess- und

Datenmodellierung, wobei Letztere auf der Ebene der genauen Festlegung von Datenformaten liege, wie sie oben allgemein für das Datenspektrum der Geistes-, Kultur- und Sozialwissenschaften vorgenommen wurde. Eine Prozessmodellierung als "model of" ist in einigen Disziplinen üblicher als in anderen. In den empirisch und historisch arbeitenden Kultur- und Sozialwissenschaften finden sich in zahlreichen Einführungswerken Grafiken, in denen die notwendigen Forschungsschritte in einer linearen Abfolge oder einem Kreisdiagramm dargestellt werden (etwa in FLICK 2007, S.128; MURI 2014, S.463; REHBEIN 2017, S.342; WITT 2001, §15). Auch im computationellen Denken ist Modellierung allgegenwärtig (DENNING & TEDRE 2019). Denn um einen Algorithmus zu erstellen, müssen die einzelnen Schritte zunächst in Teile zerlegt werden, um das Vorgehen und die Reihenfolge genau festzulegen. Modellierung hilft dabei im doppelten Sinn: Durch sie wird das eigene Tun bewusst und zugleich kann damit gezielt herausgearbeitet werden, wo digitale Daten in qualitative Forschung integriert und wo sie mit digitalen Verfahren analysiert werden bzw. wie dies geschehen kann. [13]

Modellieren hat zum Ziel, Prozesse zu abstrahieren und damit Abweichungen vom idealtypischen Verlauf zugunsten wesentlicher Aspekte auszublenden (FLANDERS & JANNIDIS 2016, S.229; THALLER 2017, S.16). Geistes-, kultur- und sozialwissenschaftliche Forschungsprozesse sind im Grundsatz vergleichbar (siehe Abbildung 1): Es besteht ein Problem, das so beschrieben wird, dass sich daraus eine handhabbare Fragestellung ableiten lässt. Auf dieser Grundlage wählen die Forschenden ein methodologisches Setting und konkrete Methoden, darauf folgen dann die Datenerhebung und Auswertung der Quellen. Schließlich verschriftlichen die Forschenden ihre Ergebnisse und (sie oder andere Forschende) leiten daraus neue Problemstellungen ab. Der dargestellte Prozess ist dann nicht linear, sondern iterativ zyklisch, etwa wenn Forschende im Stadium des Pretests anpassen oder wenn auffällt, dass relevante Quellen noch fehlen. In den DH wurden diese Arbeitsschritte früh als Teil der "Scholarly Primitives" (UNSWORTH 2000) bezeichnet, als "basic functions common to scholarly activity across disciplines, over time, and independent of theoretical orientation" (§2). [14]

Im Folgenden soll eine exemplarische Studie genauer betrachtet werden, um diese idealtypischen Prozesse zu verdeutlichen. Dabei stand die Frage im Mittelpunkt, wie die Akzeptanz von digitaler Gesundheitsversorgung allgemein und von Telemedizin im Speziellen verhandelt wird und welche Problematiken relevante Diskursakteur:innen identifizierten. Die Diskursanalyse entstand im Rahmen des Forschungsverbundes "Automatisierte Modellierung hermeneutischer Prozesse (hermA)" (GAIDYS et al. 2017) und war durch die Kollaboration von Geistes-, Kultur- und Sozialwissenschaftler:innen mit Informatiker:innen und Computerlinguist:innen geprägt. Im Rahmen der Studie wurde, auf die dargestellte Modellierung übertragen, zunächst das Problem beschrieben, dass Telemedizin gesellschaftlich verhandelt wird. Darauf aufbauend wurde die Fragestellung danach entwickelt, welche Diskursakteur:innen in den Aushandlungsprozessen beteiligt sind und wie Akzeptanz diskursiviert wird. Mit der wissenssoziologischen Diskursanalyse (KELLER 2011 [2005]) habe ich als konkrete Methode ein Sampling und Coding

nach der Grounded-Theory-Methodologie (GLASER & STRAUSS 2010 [1967]; CHARMAZ 2014) gewählt. In der Datenerhebung und Auswertung habe ich iterativ die potenziell relevanten multimodalen Quellen auf digital verfügbaren Text reduziert, um mit den Textquellen verschiedene digitale Methoden nutzen zu können. Theoretisches Sampling (STRÜBING 2019) und induktives Vorgehen sowie abduktive Schlüsse (REICHERTZ 2013) waren für die Korpuszusammenstellung unerlässlich. Manuelle Annotation als digitale Handarbeit (FRANKEN 2020a; FRANKEN, KOCH & ZINSMEISTER 2020) war in den mehrfach durchlaufenen Schleifen der Materialerhebung und Analyse zentral für das Verstehen der Kontexte und das Herausarbeiten von Diskurspositionen und -akteur:innen. Es folgte die Verschriftlichung zentraler Ergebnisse, etwa das Verschwinden von Patient:innen aus dem Diskurs (FRANKEN 2022b) und das Formulieren weiterer Problemstellungen etwa nach den diskursiven Setzungen rund um Datenschutz. Der entsprechende Forschungsprozess ist mit exemplarischen digitalen Erweiterungen idealtypisch in Abbildung 1 dargestellt. Mit diesem Ansatz wurden rund 13.000 Texte erhoben, auf knapp 9.000 relevante reduziert und mit den unterschiedlichen digitalen Methoden, die im Folgenden erläutert werden, vorläufig 87 interessante Datensätze identifiziert, die qualitativen Analysen unterzogen wurden.

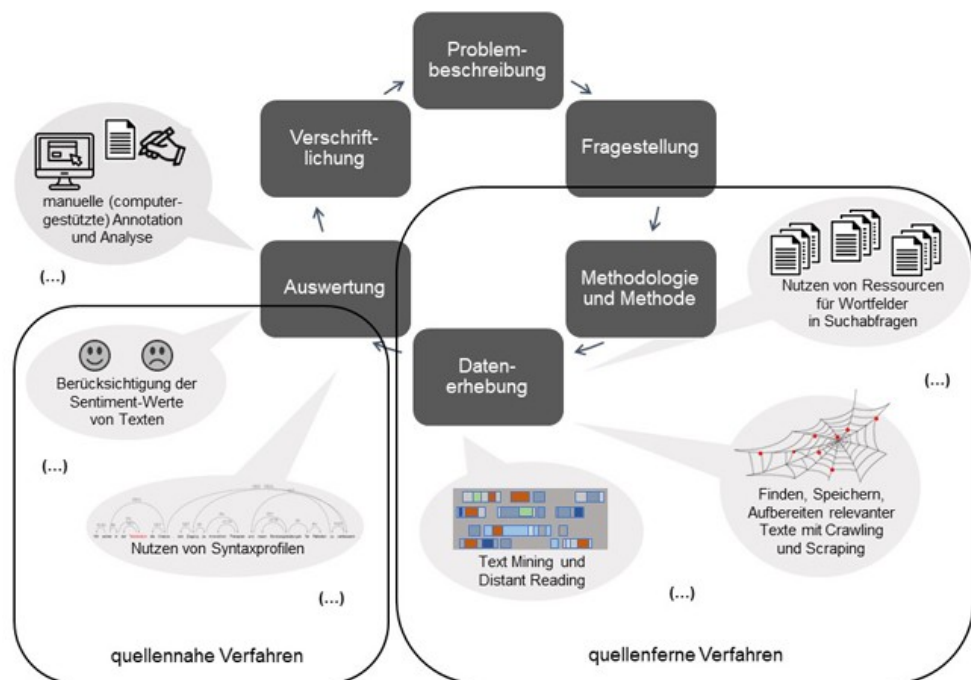


Abbildung 1: qualitativer Forschungsprozess, erweitert um digitale Verfahren [15]

4. Erweiterung des Forschungsprozesses in Auswahl, Erhebung, Aufbereitung und Auswertung

In der erfolgten ersten Modellierung wird am Beispiel der getätigten Diskursanalyse (Abschnitt 3) deutlich, dass Datenauswahl, -erhebung und -auswertung für digitale Verfahren geeignet sind. Mit diesen Verfahren verändern und erweitern Forschende den bisher modellierten Forschungsprozess. Betrachtet man diese Schritte genauer, so fällt zunächst die Datenauswahl ins Auge, die ein Vorschrift der Datenerhebung ist. Hierfür ist Vorwissen über das Untersuchungsfeld notwendig, das Grundlage für die informierte Suche ist. Die Nutzung von Ressourcen wie etwa kontrollierten Vokabularen und Normdaten stellt sich als besonders hilfreich dar. Auf Grundlage der Normdaten können Wortfelder gebildet werden, die mit der manuellen Durchsicht erster Quellen noch angereichert werden können (siehe zum Vorgehen ADELMANN, FRANKEN, GIUS, KRÜGER & VAUTH 2019). Mit Wortfeldern werden sprachliche Zusammenhänge als semantische Einheiten aufgezeigt, die jedoch nicht trennscharf zu ziehen sind (TRIER 1973 [1931]). Durch die Bildung von Wortfeldern wird nicht mehr nach einzelnen Begriffen gesucht, sondern mit einem Set von Begriffen. Die Erstellung dieses Sets ist in großen Teilen digitale Handarbeit: die Durchsicht von Vokabularen und anderen Normdaten nach thematisch passenden Kategorien ebenso wie das Lesen von ersten Quellenausschnitten zur Identifizierung von relevanten Suchbegriffen. [16]

In meiner Studie habe ich aus der [Gemeinsamen Normdatei](#) der Deutschen Nationalbibliothek (DNB) (WIECHMANN 2012) alle Unterbegriffe von als relevant identifizierten Begriffen berücksichtigt. Je nach Themensetzung werden dabei mehrere hundert Unterbegriffe zusammengestellt, die in eigenem Nachdenken nicht so umfassend berücksichtigt werden könnten. Erstellt man etwa ein Wortfeld zu Essen und Trinken (STIEMER & VAUTH 2020), so sind mit Datengrundlage schnell umfangreiche Listen von Lebensmitteln vorhanden. Für spezifischere Fälle wie die Telemedizin fallen eher Begriffe auf, die anders nicht bekannt geworden wären, etwa die Telekardiologie. Ausgelesen werden diese Begriffe aus den Vokabularen automatisiert über eine Schnittstelle, welche Standard-Formate ausgibt und seitens der DNB zur Verfügung gestellt wird. Durch die Kombination unterschiedlicher Normdaten und manueller Recherche in ersten Quellen ergibt sich ein umfassendes Wortfeld zum Untersuchungsthema. Im Laufe der Beschäftigung mit dem Thema können Wortfelder wachsen und präzisiert werden. Besonders in großen, unbekannteren Quellenmengen können Forschende so Material identifizieren, das durch einige wenige Suchbegriffe nicht gefunden würde (KOCH & FRANKEN 2020). [17]

Für die Datenerhebung sind digitale Verfahren ebenfalls zielführend einsetzbar. Um Datensätze zu extrahieren, benötigen Forschende ein Verfahren des *Crawling* und *Scraping*, also das Finden, Speichern und Aufbereiten von relevanten Internetquellen (BARBARESI 2019). Auch in diesem Schritt überwiegt zunächst die digitale Handarbeit, und erst in der Folge steht die computationale Generierung von *Big Data* im Mittelpunkt. Für das reflektierte Erzeugen von Startpunkten des *Crawlens* im Netz, sogenannten *Seed-URLs*, ist wiederum das

Wissen einzelner Forscher:innen, manuelles Suchen und Zusammenstellen notwendig. Die weitere Arbeit, die Skalierung, erledigt in der Hauptarbeit ein Algorithmus, der auf Grundlage der Ausgangspunkte Text speichert, aus diesem Text Links extrahiert, den Links folgt und wiederum den Text speichert. Für das thematische *Crawlen* ist die Verwendung der vorab erzeugten Wortfelder und manuell erzeugten Listen von Webadressen als Startpunkte essenziell, um fokussiert vorzugehen (ADELMANN & FRANKEN 2020). Mit dem Verfahren werden schnell große Mengen von Quellen hervorgebracht. Natürlich können auch auf andere Art Daten generiert werden, etwa indem auf die Schnittstellen von sozialen Medien, sogenannte *Application Programming Interfaces* (APIs), zugegriffen und etwa *Tweets* zu einzelnen *Hashtags* gespeichert werden. Im *Scraping* werden die so ausgewählten Quellen dann für das weitere Verfahren gespeichert. Auswahl und Erhebung sind durch dieses Verfahren eng miteinander verbunden. In der Beispielstudie wurden Webseiten von Krankenkassen, Ärzt:innen- und Patient:innenverbänden als *Seed-URLs* verwendet. Davon ausgehend wurden etwa Kund:innenzeitschriften, wissenschaftliche Studien und politische Strategiepapiere, aber auch technische Anleitungen für Praxen von Ärzt:innen und Informationstexte gefunden. Da in einer wissenssoziologischen Diskursanalyse die Identifizierung relevanter Diskurakteur:innen und von deren Positionen zentral ist, eignet sich ein thematisches *Crawling* – ausgehend von durch Vorwissen ermittelten Wortfeldern und Akteur:innengruppen – auch für das Herausarbeiten eines breiten Spektrums an Positionen. Allerdings sind die Daten im Ergebnis nicht mehr oder nur in kleinen Teilen nach den vorab identifizierten Diskursakteur:innen getrennt, da die Links auf den Webseiten sich nicht an Grenzziehungen von Forschenden halten. Dies macht neue Perspektiven etwa der Zuordnung in der weiteren Auswertung notwendig und zeigt den Bedarf an digitaler Handarbeit während und nach computationellen Schritten. Die gewonnenen Daten können durch Quellen wie *Tweets* oder auch qualitative Interviews weiter ergänzt werden. [18]

In der Auswertung mittels digitaler Verfahren ist insbesondere bei Textdaten eine vorhergehende Aufbereitung notwendig. Neben der strukturierten Benennung und Speicherung der großen Datensätze sind je nach den weiteren Schritten auch Fragen danach relevant, welche Texteinheiten getrennt gespeichert werden sollen. Besonders zentral stehen linguistische Verarbeitungen, die unabdingbar für präzise digitale Analyseergebnisse sind. Mit ihnen werden einzelne Wörter identifiziert (*Tokenisierung*), diese auf ihre Grundform zurückgebildet (*Lemmatisierung*) und Wortarten zugeordnet (*Part-of-Speech-Tagging*) oder nach syntaktischen Relationen in den Sätzen aufgeschlüsselt (*Parsing*). Solche Verfahren des *Natural Language Processing* stammen aus der Computerlinguistik und sind in den DH und CSS weit verbreitet (vgl. einführend ANDRESEN & ZINSMEISTER 2019; PUSTEJOVSKY & STUBBS 2012). Sucht man etwa nach Worthäufigkeiten, so ist dies ohne eine *Lemmatisierung* nur sehr begrenzt aussagekräftig. In anderen Zusammenhängen interessiert eher das Aufschlüsseln nach Wortarten oder nach Entitäten, wenn Forschende etwa nachvollziehbar halten möchten, dass "Jens Spahn" und "Bundesgesundheitsminister" synonym zueinander waren – zumindest für einen spezifischen Zeitraum. In der Datenaufbereitung wird aus (in computationeller

Perspektive) unstrukturiertem Text strukturierter Text, der maschinenlesbar und damit mit digitalen Verfahren bearbeitbar ist. Während für einige Verarbeitungsschritte gut nachnutzbare Tools bestehen, sind je nach Fragestellung und Datenstruktur durchaus Aufwände für diese Arbeitsschritte einzuplanen. Nicht zu vernachlässigen ist auch hier die digitale Handarbeit, denn in der Aufbereitung treffen Forschende eine Vielzahl von Entscheidungen zu Verfahren und Umsetzungen im Detail, sie arbeiten mit unterschiedlichen Versionen der Daten und mit entsprechender Dateibenennung, Datenablage und Überblickslisten. [19]

Wenn potenziell forschungsrelevante Daten entsprechend aufbereitet sind, beginnt die erste Sichtung und Auswertung, die mit *Text Mining* und *Distant Reading* unterstützt werden kann, also dem Filtern innerhalb großer Textmengen nach relevanten Passagen (IGNATOW & MIHALCEA 2017; KOCH & FRANKEN 2020; LEMKE & WIEDEMANN 2016). Als konkrete Methoden kommen etwa statistische und computerlinguistische Analysen in Betracht: von einfachen Häufigkeitsverteilungen bis hin zu *Kollokationen*, also der Frage danach, welche Wörter häufig zusammenstehen. Dabei können wiederum Begriffe aus den gebildeten Wortfeldern genutzt werden, um spezifische Zusammenhänge herauszuarbeiten. Auch mit *Topic Modeling* (BLEI 2012; JACOBS & TSCHÖTSCHHEL 2019) können gemeinsam auftauchende Begriffe einbezogen werden. Dabei wird in *Machine-Learning*-Verfahren angenommen, dass Begriffe zu einem Thema gehören, wenn sie häufig zusammenstehen. Da hier keine Begriffe als Setzungen vorab ins Verfahren gegeben werden, eignet es sich besonders zur Unterstützung induktiver und abduktiver Forschungsansätze (NELSON 2020), denn Ergebnis eines *Topic Modeling* ist eine gruppierte Liste von häufig zusammenstehenden Begriffen ohne jegliche Annotation. Dies ist ein grundlegender Unterschied zu den statistischen Verfahren, welche stets nach vorab festgelegten Suchwörtern strukturiert sind. Nach der Nutzung eines das *Topic Modeling* umsetzenden Tools wie etwa "Mallet" ("Machine Learning for Language Toolkit", McCALLUM 2002) als kommandozeilenbasiertem Script, das sich in den DH und CSS zu einem Standard entwickelt hat, folgt wiederum eine Menge digitaler Handarbeit: in diesem Fall vor allem in Form des Durchsehens und Markierens von Listen. Das eigentliche Erstellen der *Topic-Term*-Matrix als Liste von Begriffen, die einem *Topic* zugeordnet werden, dauert mit dem Tool und entsprechend aufbereiteten Texten nur wenige Minuten – die Arbeit liegt (nach der vorhergehenden Datenaufbereitung) in der Herausarbeitung und Nutzung der darin steckenden Aussagen. In einer großen Menge unbekannter Texte werden durch *Topic Modeling* Texte sichtbar, die einzelne *Topics* stark abbilden. Auf Grundlage der entsprechenden *Topic*-Listen wissen Forschende, dass etwa Begriffe wie "Patient", "Daten", "Einwilligung", "Verarbeitung" und "Gesundheitsdaten" oft gemeinsam auftauchen, allerdings nicht im Detail, was in den Texten enthalten ist. Auch analytisch relevante Feinheiten wie die Gender-Form von Begriffen verschwinden. Wenn zunächst auf diejenigen *Topics* reduziert wird, deren Begriffe tatsächlich mit der eigenen Fragestellung zu tun haben (könnten), lassen sich mit diesem Verfahren aussagekräftige Text(teile) im Quellenmaterial identifizieren. Entsprechende Texte können in der Folge manuell gesichtet werden, um mehr über den Themenbereich zu erfahren. *Topic*

Modeling kann also in verschiedenen Dimensionen zum Filtern innerhalb von Textmaterial genutzt werden. Es fehlen dabei allerdings jegliche Kontexte in Texten, gearbeitet wird mit Tabellen und nicht mit dem eigentlichen Quellenmaterial. [20]

Um die Kontexte nachzuvollziehen, ist die Nutzung von syntaktischen Annotationen hilfreich (ANDRESEN et al. 2020). Dabei werden sprachliche Strukturen ausgewertet, um den konkreten Kontext der Wortverwendung zu ermitteln. Dies ermöglicht einen anderen Blick auf das (noch unbekannte oder doch noch nicht intensiv gesichtete) Quellenmaterial: Es können etwa Zuschreibungen zu den bereits bekannten Konzepten sichtbar gemacht werden. So wurde im Beispielmateriale Telemedizin als "Möglichkeit" attribuiert, "bereichern" und "ermöglichen" standen ebenfalls oft im Zusammenhang mit diesem Begriff. Viele der Zuordnungen waren allerdings ohne weiteren Einbezug des Kontextes nicht zu interpretieren. Es fanden sich im *Webcrawling*-Korpus sowohl Zuschreibungen von Telemedizin zu "sichergestellt" als auch zu "gefährdet", was ohne einen Blick in die jeweiligen Textstellen zunächst einmal die Vielschichtigkeit des gesammelten Datenmaterials aufzeigte, das offensichtlich unterschiedliche Diskurspositionen enthielt. Forschende werden hier wiederum auf möglicherweise relevante Texte und Textstellen hingewiesen. Das Erstellen von Syntaxprofilen erlaubt einen anderen Blick auf das Material, und die Ergebnisse können als Filter eingesetzt werden. Diese Zugriffe sind um ein Vielfaches aussagekräftiger als die reinen Häufigkeiten, die CAQDAS-Pakete als *Mixed-Methods*-Ansätze für qualitative Verfahren mit anbieten. [21]

Ein weiteres Verfahren zur Datenauswertung ist die Berücksichtigung von Sentiment-Werten der Texte, also die Klassifizierung von Wortbedeutungen als positiv oder negativ (D'ANDREA, FERRI, GRIFONI & GUZZO 2015; LIU 2015). Dabei wird in Textteilen nach positiv und negativ konnotierten Begriffen gesucht, entweder durch vorab (oft von Dritten) festgelegte Wörterbücher oder auf Grundlage von maschinellem Lernen. In explorierenden Versuchen im Kontext der dargestellten Diskursanalyse hat sich schnell gezeigt, dass in beiden Alternativen (zu) viele Mehrdeutigkeiten in der Bewertung von Begriffen als positiv oder negativ bestanden, sodass die computationally zugeordneten Bedeutungen einen eingeschränkten Erkenntnisgewinn brachten. Erst mit digitaler Handarbeit waren verlässliche Aussagen möglich: Nur mit manueller Überprüfung der einzelnen computationallyen Bewertungen kann entschieden werden, ob die Wortverwendung im konkreten Zusammenhang tatsächlich positiv oder negativ ist. Daraus ergeben sich ähnlich wie im *Topic Modeling* neue methodologische Fragen der Dokumentation und auch der Gütekriterien (STRÜBING, HIRSCHAUER, AYAB, KRÄHNKE & SCHEFFER 2018) für entsprechende Bewertungspraxen. [22]

In der Auswertung ist die manuelle Annotation, also Markierung von zentralen Textteilen und induktive Kategorienentwicklung (FRANKEN et al. 2020) unerlässlich und kann nur durch Computerunterstützung begleitet werden. Annotationen sind zwar auf Wortebene computationally zu lösen – das in Abschnitt 4 genannte Beispiel kann umgesetzt und alle Nennungen von "Jens Spahn" und

"Bundesgesundheitsminister" sowie weitere Synonyme für ihn können automatisiert markiert werden –, aber für die Analyse von Bedeutungsstrukturen sind nach und während der Filterfunktionen digitaler Verfahren Kopfarbeit und digitale Handarbeit gleichermaßen notwendig. Mit Annotationstools wird der sonst mit Stift und Papier realisierte Arbeitsschritt digital unterstützt. Für die manuelle Annotation gibt es noch immer kein gut für geistes-, kultur- und sozialwissenschaftliche Bedürfnisse ausgebautes *Open-Source*-Tool, sodass hier meist auf proprietäre Lösungen der CAQDAS zurückgegriffen wird, auch wenn dies den Grundsätzen von *Open Science* widerspricht. [23]

Digitale Methoden müssen auf ihren Anwendungsbereich hin differenziert werden. *Topic Modeling* etwa ist gut geeignet, um grobe Unterthemen zu identifizieren. Hingegen ist *Sentiment*-Analyse auf den konkreten Text bezogen, also nutzbar zum Aufspüren von Besonderheiten. Quellenferne Verfahren eignen sich für frühe Phasen der Auswertung, quellennahe hingegen eher für späte Phasen. Die exemplarisch vorgestellten Verfahren zeigen, wie Datenauswahl, -erhebung und -auswertung durch digitale Methoden unterstützt und erweitert werden können. Qualitativ Forschende müssen zwischen Zugängen und Quellen immer wieder wechseln, um fundierte Aussagen zu treffen. Die Auswahl der Herangehensweisen und deren Anordnung bedingen mögliche Erkenntnisse zentral. Vorannahmen sind in allen Schritten vorhanden und oft schon in Tools, Verfahren und Vorgehensweisen eingeschrieben – sie müssen immer wieder transparent gemacht und hinterfragt werden. [24]

Natürlich gibt es noch andere Verfahren etwa aus dem Kontext des *Natural Language Processing* (beispielsweise in der Anwendung von BAKER [2012] oder NELSON [2020]). Auch Visualisierungen von Datensätzen können Interpretationsgrundlage herangezogen werden (DRUCKER 2020; SVENSSON 2016). Der Kontext der jeweiligen Fragestellung ist maßgeblich, um die Anwendbarkeit auf geistes-, kultur- und sozialwissenschaftliche Forschungsinteressen zu bewerten. Für andere Quellengattungen könnte man ein vergleichbares Modell aufsetzen wie das hier für Text entworfene. Die Analyse von Bildern beispielsweise ist bereits weit entwickelt, für audiovisuelle Daten ist das noch schwieriger (Versuche finden sich etwa bei LEGEWIE & NASSAUER [2018]). Eine multimodal kombinierte Quellengrundlage bedeutet aktuell in der Regel ein paralleles Auswerten, denn es gibt kaum Verfahren, die etwa Text-Bild-Kombinationen computationell verarbeiten könnten. [25]

5. Herausforderungen und Potenziale

Was sind Herausforderungen, aber auch Potenziale der Erweiterung geistes-, kultur- und sozialwissenschaftlicher Forschungsprozesse mit digitalen Daten und Methoden? Wie gezeigt wurde, können nicht alle Schritte des Forschungsprozesses mit digitalen Verfahren erweitert werden. Die epistemologisch bestimmten Anteile sind weiterhin von den Forschenden manuell zu bearbeiten. [26]

5.1 Herausforderungen

Ich sehe fünf zentrale Herausforderungen der Anwendung von digitalen Daten und digitalen Methoden als Werkzeugkasten sowie deren Weiterentwicklung im Sinne eines Experimentierkastens. Erstens ist die Korpusgenerierung noch nicht für alle Forschungsbereiche zufriedenstellend. In den DH und den CSS wird oft mit prozessproduzierten Quellen (Abschnitt 2.2) gearbeitet. Studien sind in der Regel begrenzt auf einen Datentypus. Die Qualität von qualitativ empirisch arbeitenden Fächern, aber auch der Geschichtswissenschaft oder der Geografie liegt jedoch gerade in der Kombination verschiedener Quellentypen. Der Arbeitsschritt der Korpuserstellung steht damit vor spezifischen Herausforderungen der digitalen Verfahren zum Finden und Speichern sowie des Einbezugs von forschungsproduzierten Daten in ihrer Vielfalt und Vernetztheit. [27]

Damit hängt zweitens zusammen: Gerade die Forschungscommunity der DH ist sehr zentriert auf Schriftsprache. In den CSS werden zwar Daten aus den sozialen Medien zentral gestellt, doch Zusammenhänge etwa zwischen Text und Bild selten berücksichtigt. Forschung wird stark beschnitten, wenn nicht auch Bild- oder Audiodaten neben Text analysiert werden. Heterogene und multimodale Quellen müssen stärker als bisher einbezogen werden. [28]

Drittens ist die Frage nach den angewendeten DH-Verfahren zentral. Die Auswahl der Herangehensweisen und deren Reihenfolge bedingen mögliche Erkenntnisse und sind eine Herausforderung in der Modellierung und Realisierung des Forschungsprozesses. Eine Begrenzung auf ein Verfahren reicht in der Regel nicht aus, sondern gerade die Kombination von unterschiedlichen quellennahen und -fernen Zugängen bereichert und erweitert Forschungsperspektiven. Fragestellungen werden im Laufe der Projekte verändert, und die methodischen Herausforderungen und Möglichkeiten werden oft erst im Tun deutlich. [29]

Damit zusammenhängend ist viertens relevant: Der oben modellierte Forschungsprozess wird in Auswahl, Erhebung, Aufbereitung und Auswertung verändert, wenn digitale Verfahren genutzt werden. An die Stelle einer individuellen Anordnung der Quellen, die iterativ angepasst und zunehmend in ihrer Komplexität überschaut wird, tritt eine nicht-lineare, granulare Perspektive auf das Material. Strukturiert werden die Einsichten in manuell unüberschaubare Quellensettings vor allem durch die Ausgaben der algorithmischen Verfahren. Es ist auch deshalb immer wieder notwendig, manuell Quellen zu sichten, etwa

linearen Text zu lesen, um die Gesamtschau und Tiefe der Quellen nicht aus dem Blick zu verlieren. Ergebnisse digitaler Verfahren würden, so hielten auch BAKER und LEVON (2015) fest, noch zu oft als neutral und als Gegenmaßnahme zum vermeintlichen *Cherry-Picking* der qualitativen Sozialforschung angesehen. [30]

Die fünfte Herausforderung liegt in den Bedeutungspluralitäten, die geistes-, kultur- und sozialwissenschaftlichen Quellen und ihrer Interpretation immanent sind, aber bisher kaum digital erfasst werden können. Ansätze qualitativer Forschung zeichnen sich besonders dadurch aus, dass keine eindeutigen Antworten bestehen und diese in vielen Fällen auch gar nicht gesucht werden. Formale Kategorien, die in digitalen Verfahren oft zentral sind, sind für diese Fragestellungen nicht immer (einfach) zu finden. Dementsprechend ist hier noch viel Bedarf für Weiterentwicklungen im Experimentierkasten der digitalen Methoden. [31]

Quer zu diesen Herausforderungen liegt die reflexive Ebene, die in den DH zwar betont wird (GEIGER & PFEIFFER 2020), aber noch ausbaufähig ist. Qualitativ Forschende sind für diese Fragen mit ihren Reflexionen zu Methoden und Theoriebildung bestens vorbereitet. Die Hinwendung zu digitalen Methoden macht spezifische kritische Perspektiven auf Daten und Codestrukturen notwendig, auf deren Bedingtheit und Operationalisierbarkeit, auf Begrenztheit der entsprechenden Forschung und Möglichkeiten ihrer Erweiterung, die mangelnde Reflexion vielleicht sogar verhindern können. Insofern ist nach der durch digitale Verfahren veränderten, digitalisierten hermeneutischen Wissensproduktion und damit adressierten Wirklichkeitsausschnitten zu fragen. [32]

5.2 Potenziale und Weiterentwicklungen

Potenziale wurden implizit bereits genannt (Abschnitte 4 und 5.1), denn die Herausforderungen können als mögliche Weiterentwicklungen gespiegelt werden. Durch die Nutzung von Standards und Vokabularen können erstens Korpora informierter zusammengestellt werden. Quellen werden so leichter auffindbar, besser vergleichbar und qualitativ hochwertiger erschlossen. Eine vielversprechende Einführung ist dabei die Praxis des Annotierens: Sie kann als Bindeglied zwischen Korpusaufbereitung und -analyse dienen. Viele Teilschritte sind hier bereits automatisierbar, etwa durch *Named Entity Recognition*, bei der Entitäten wie Personen oder Orte automatisch annotiert werden (IGNATOW & MIHALCEA 2017, S.61). Mit computationellen Annotationen etwa von *Kollokationen* können Strukturen sichtbar gemacht werden, die manuelle Annotation und damit Interpretation evozieren. [33]

Ein Potenzial für die Analyse komplexer relationaler Phänomene und die Entschlüsselung von Bedeutungen ist zweitens die Weiterentwicklung von semantischen Suchtechnologien und Bilderkennungen für kombinierte Analysen. Für multimodale Quellengrundlagen bietet sich die Verwendung von teilautomatischen Verfahren an, die *Humans in the Loop* aktiv einbeziehen, sodass im Sinne einer "Human-Computer Interaction" Mensch und Computer

immer wieder in Schleifen miteinander interagieren (WOLETZ 2016). Solche auch durch multimodale Quellengrundlagen getriebenen Perspektiven tragen zum Experimentierkasten bei: Mit ihnen werden neue Herausforderungen und daraus folgende Weiterentwicklungen in die digital arbeitenden Querschnittsdisziplinen eingebracht. [34]

Der Einbezug von digitalen Verfahren ist drittens gewinnbringend, wenn er reflektiert und in den passenden Teilschritten im Forschungsprozess geschieht. Dies erfolgt durch einen Rückgriff auf den Werkzeugkasten. Die Nutzung von integrativen Arbeitsumgebungen könnte den Prozess des iterativen Hin- und Herspringens zwischen Quellen und Verfahren weiter unterstützen, um nicht zwischen den verschiedenen Tools und Datensätzen den Überblick zu verlieren. Trotz der Notwendigkeit eines flexiblen und benutzendenfreundlichen Tools, das eine freie Kombination und nachvollziehbare Nutzung verschiedener Verfahren ermöglicht, steht die Entwicklung entsprechender Umgebungen noch aus. Bei einer angemessenen Reduktion von Komplexität unter Beibehaltung der Nachvollziehbarkeit ist davon auszugehen, dass die Hemmschwelle zur Anwendung der Verfahren durch ein solches Tool sinken würde. Hierdurch könnten bestehende Ansätze digitaler Methoden weiter in die Breite qualitativer Forschung getragen werden. Die Entwicklung einer entsprechenden Softwarelösung könnte auch die Lücke schließen zwischen digitalen Daten und deren Aufbereitung im Forschungsdatenmanagement, denn die langfristige Speicherung als (ggf.) nachnutzbare Forschungsdaten steht aktuell noch zu unverbunden nach oder neben der eigentlichen Datenauswahl, -erhebung und -analyse. [35]

Die Veränderung des Analyseprozesses kann viertens neue Perspektiven auf das Quellenmaterial hervorrufen. Einblicke mit digitalen Verfahren ermöglichen ein gezielteres Sampling und potenziell andere Erkenntnisse. Denn mit granularen Perspektiven auf Grundlage quellenferner und quellennaher Verfahren können fundierte Entscheidungen darüber getroffen werden, welche Wirklichkeitsausschnitte einer tiefgehenden Analyse unterzogen werden sollen. Gleichzeitig erweitern die Analysen mit digitalen Verfahren die Quellenbasis und ermöglichen Einblicke in Datensätze, die rein qualitativ arbeitend nicht überschaubar wären. Auch ohne eine integrierte Arbeitsumgebung gibt es bereits viele Tools, die dabei einzelne Arbeitsschritte übernehmen oder Unterstützung für diese anbieten. [36]

Die umfassendste Dimension, in welcher der Experimentierkasten für qualitative Forschung genutzt und weiterentwickelt werden kann, ist fünftens die der Bedeutungspluralitäten. Erste Ansätze bestehen bereits. So können etwa *Embeddings*, also die Darstellung der Umgebung eines Begriffes im Kontext des spezifischen Satzes durch Vektoren, genutzt werden, um Polysemie (also Bedeutungsvielfalt) und sprachliche Synonyme anhand ihrer Verwendungsweise im konkreten Satz zu erfassen (WIEDEMANN, REMUS, CHAWLA & BIEMANN 2019). Dies kann ein Zugang sein, mit dem Bedeutungen digital erkannt werden und der das manuelle Auffinden ähnlicher Äußerungen durch ihre spezifischen Verwendungskontexte unterstützt. [37]

Querliegend zu diesen fünf Potenzialen liegt auch hier die reflexive Ebene. Zentral ist, was mit digitalen Verfahren nicht gefunden wird und was vielleicht weiterhin nur mit digitaler Handarbeit auffindbar ist, mit Kontextwissen der Forschenden und ihrem spezifischen Blick auf die in Relationen und Kontexten repräsentierten Quellen. Daten sind nicht neutral und dürfen auch keine *Black Box* sein. Sie sind unvollständig und bleiben unsauber, sie sind lebendig und in vielen Fällen auch defekt oder nicht interpretierbar (PINK, RUCKENSTEIN, WILLIM & DUQUE 2018). Daten sind in ein Netz von Bedeutungen und unterschiedlicher Handlungsmacht eingebunden und daher nur in ihren diskursiven Konstruktionen und Materialitäten zu verstehen. Welche Fragen qualitative Forscher:innen an sie stellen können, ist neben der Quellenauswahl zentral davon abhängig, welche Analyseverfahren und Tools verwendet werden. Wenn etwa wie in der dargestellten Diskursanalyse die Daten nicht aufwendig für eine Netzwerkanalyse aufbereitet werden, bleibt diese Blickrichtung auf die Quellen verwehrt. Daraus folgt: "To some extent, tools and infrastructure shape the questions we can ask" (SVENSSON 2016, S.16). Für die Analyse solcher Grenzen bieten die *Science and Technology Studies* Anknüpfungspunkte insbesondere mit den *Critical Data and Code Studies*, in welchen Daten und Algorithmen kritisch auf ihre Wirkmächtigkeit hin untersucht werden (BOYD & CRAWFORD 2012; KITCHIN & LAURIAULT 2018). Eine stärkere Einbindung dieser Perspektiven in die Praxis der digitalen Methodenentwicklung und -nutzung kann gewinnbringend für beide Seiten sein. [38]

6. Fazit: Erweiterungen durch digitale Daten und Methoden in qualitativen Forschungsprozessen

Wie sieht also die Erweiterung qualitativer Forschungsprozesse mit den digitalen Daten und Methoden aus, die aktuell unter den Begriffen DH oder CSS subsumiert werden? Es ist notwendig, konkrete Verfahren aus dem Werkzeugkasten zu nutzen *und* sie für Bedürfnisse qualitativ Forschender mit dem Experimentierkasten weiterzuentwickeln. Dabei müssen die Grenzen der digitalen Verfahren im Blick behalten und es muss reflektiert werden, was möglich und zielführend ist. [39]

Der Einsatz digitaler Verfahren, die Nutzung verfügbarer Tools und deren Weiterentwicklung bereichern qualitativ forschendes Vorgehen. Die kreative Integration neuer Elemente in den Forschungsprozess ist für Problemlösungen aus und in den DH und den CSS besonders wichtig. Diese nur als Service-Dienstleistende zu verstehen, wird dem Anspruch nicht gerecht. Es muss dabei aber klar sein, dass die meisten Fragestellungen durch digitale Ansätze nicht vollständig bearbeitbar werden. Es besteht jedoch das Potenzial, Analysen nicht nur zu unterstützen, sondern weiterzuentwickeln. Auch wenn die intensive Durchdringung des Quellenmaterials im hermeneutischen Verstehen als dem Kernschritt qualitativer Analysen künftig häufiger auf digitale Verfahren aufbauen kann, wird der Schwerpunkt dabei weiterhin in der digitalen Handarbeit liegen. Digitale Verfahren können diese an vielen Stellen erleichtern. Gleichzeitig kann die Integration digitaler Verfahren für die Erweiterung qualitativer Forschung

sorgen, insbesondere mit den Prinzipien des Aussortierens uninteressanter und des Auffindens besonders relevanter Materialien. [40]

Potenzial besteht außerdem in der noch besseren Verbindung quantitativer Auswertungen und qualitativer Analyseschritte. Vieles ist in der quantitativen Forschung bereits üblich, weil dort mit anderen Datentypen und auch Datenmengen gearbeitet wird. Deshalb lohnt sich ein neuer Blick auf alte Diskussionen rund um *Mixed Methods* (MORSE & NIEHAUS 2016). Quantitative und qualitative Forschung befinden sich in unterschiedlichen Phasen eines Umbruchs der Forschungsprozesse hin zur Datafizierung in einer Gesellschaft, die zunehmend durch Daten vermittelt ist oder durch Daten erst ermöglicht wird (DOURISH & GÓMEZ 2018). In den quantitativen Forschungen ändert sich stärker die Datengrundlage hin zu *Trace Data* und der Nachnutzung ohnehin entstehender Datensätze, in der qualitativen Forschung ist es mehr der Zugriff auf die Daten selbst, der sich wandelt. Mit digitalen Verfahren können große Mengen von Daten verarbeitet werden, was insbesondere die Datenerhebung verändert. Die in digitalen Auswertungen als besonderes relevant identifizierten Daten können dann qualitativen Verfahren und damit verbundener digitaler Handarbeit zugänglich gemacht werden. Diese Erweiterung kann als Weiterentwicklung von *Mixed-Methods*-Ansätzen verstanden werden. Ich möchte mich dem Geografen und *Big-Data*-Forscher Rob KITCHIN (2014b) anschließen, der eine *Data Driven Science* vorschlägt, bei der deduktive, induktive und abduktive Elemente verbunden werden. Die unbekannt, großen Datenmengen können mit Verfahren wie den oben genannten in einer solchen Kombination genutzt werden, um aus dem Material heraus – also *data driven* – neue Hypothesen zu generieren. Nachfolgend sind diese dann weiterhin zu validieren oder falsifizieren. [41]

Nur wenn konkret Möglichkeiten und Grenzen ausgelotet werden, die digitale Daten und Verfahren qualitativer Forschung aktuell bieten, kann ein Wandel im methodischen Vorgehen vorgebracht werden. Forschenden muss dabei bewusst sein, dass viel digitale Handarbeit in diesen Schritten steckt – und wohl weiterhin stecken wird. Die Verheißungen der *Big-Data*-Analysen, auf Knopfdruck eine fundierte Quellenauswertung vornehmen zu können, widersprechen geistes-, kultur- und sozialwissenschaftlichen Gütekriterien. Jedoch darf die digitale Handarbeit nicht als Ballast wahrgenommen werden, den es zu minimieren gilt, sondern als notwendige und zielführende Beschäftigung mit dem eigenen Quellenmaterial. Digitale Handarbeit gehört ebenso wie die konzeptionell-intellektuellen Auseinandersetzungen zum wissenschaftlichen Arbeiten und kann oder sollte von keinem Algorithmus erledigt werden. Hier zeigt sich die untrennbare Verbundenheit der reflexiven Ebene mit den Daten und Verfahren. [42]

Aus dem Selbstverständnis qualitativer Forschung, den spezifischen Fragestellungen, Quellen und Methoden sowie den Bedürfnissen der Forschenden ergeben sich besondere Ansprüche an die DH und CSS. In die interdisziplinären Zusammenhänge tragen qualitativ Forschende nicht nur neue Perspektiven auf multimodale Quellen und andere Einsätze für etablierte

Verfahren, sondern auch eine erweiterte Reflexion der digital veränderten Wissensproduktion (FRANKEN 2022a). Aufgrund der wachsenden Relevanz digitaler Daten und Methoden in Forschungsfeldern ebenso wie in der Wissenschaftspolitik sollten die Potenziale digitaler Methoden in die Breite der qualitativen Diskussionen eingebracht werden. Nur wenn qualitativ Forschende die Prozesse aktiv gestalten, werden diese im Sinne der entsprechenden Forschungslogiken passend weiterentwickelt werden. Qualitativ Forschende müssen sich somit übergreifend fragen, was die Nutzung digitaler Daten und Verfahren für ihre Forschungsprozesse und Methoden, für Forschungsfelder und Gegenstände sowie für Forschungsfragen und Epistemologien bedeutet. [43]

Danksagung

Für die konstruktive Diskussion von Ideen sowie die Unterstützungen in der technischen Umsetzung der Beispielstudie danke ich insbesondere Benedikt ADELMANN, Melanie ANDRESEN, Evelyn GIUS, Wolfgang MENZEL, Michael VAUTH und Heike ZINSMEISTER. Darina HASHEM und Susanne HOCHMANN haben die Arbeit an der Datengrundlage der Studie tatkräftig unterstützt. Für bereichernde Hinweise und Anregungen zum Weiterdenken an Entwürfen des vorliegenden Textes danke ich sehr herzlich Anne DIPPEL, Moritz FÜRNEISEN, Gunther HIRSCHFELDER, Isabella KÖLZ, Hannah ROTTHAUS, Sarah THANNER und Libuše Hannah VEPREK sowie den Diskutierenden im Kolloquium des Instituts für Soziologie der Ludwig-Maximilians-Universität München. Den anonymen Gutachter:innen sowie der FQS-Redaktion danke ich für zahlreiche gewinnbringende Kommentare und Fragen. Katharina LILLICH und Florian SCHMID haben bei der Finalisierung der Aufsatzversion wertvolle Unterstützung geleistet.

Literatur

Adelmann, Benedikt & Franken, Lina (2020). Thematic web crawling and scraping as a way to form focussed web archives. In Sharon Healy, Michael Kurzmeier, Helena La Pina & Patricia Duffe (Hrsg.), *Book of abstracts. Engaging With Web Archives: "Opportunities, Challenges and Potentialities" (#EWAVirtual)*, Maynooth University Arts and Humanities Institute, Kildare, Irland, 21.-22. September 2020 (S.35-37), https://zenodo.org/record/4058013#X_YtdRYxmM8 [Datum des Zugriffs: 25. Oktober 2021].

Adelmann, Benedikt; Franken, Lina; Gius, Evelyn; Krüger, Katharina & Vauth, Michael (2019). Die Generierung von Wortfeldern und ihre Nutzung als Findeheuristik. Ein Erfahrungsbericht zum Wortfeld "medizinisches Personal". In Patrick Sahle (Hrsg.), *Book of abstracts. 6. Tagung des Verbands Digital Humanities im deutschsprachigen Raum e.V. (DHd 2019). Digital Humanities: multimedial & multimodal*, Universitäten zu Mainz und Frankfurt, 25.-29. März 2019 (S.114-16), <https://zenodo.org/record/4622122> [Datum des Zugriffs: 25. Oktober 2021].

Andresen, Melanie & Zinsmeister, Heike (2019). *Korpuslinguistik*. Tübingen: Narr Francke Attempto.

Andresen, Melanie; Begerow, Anke; Franken, Lina; Gaidys, Uta; Koch, Gertraud & Zinsmeister, Heike (2020). Syntaktische Profile für Interpretationen jenseits der Textoberfläche. In Christoph Schöch (Hrsg.), *Konferenzabstracts. 7. Tagung des Verbands Digital Humanities im deutschsprachigen Raum e.V. (DHd 2020). Spielräume. Digital Humanities zwischen Modellierung und Interpretation*, Universität Paderborn, 2.-6. März 2020 (S.219-223), <https://zenodo.org/record/4621932> [Datum des Zugriffs: 25. Oktober 2021].

Arbeitsstelle Kleine Fächer (2018). *Was ist ein kleines Fach?*, <https://www.kleinefaecher.de/kartierung/was-ist-ein-kleines-fach.html> [Datum des Zugriffs: 25. Oktober 2021].

- Baker, Paul (2012). Acceptable bias? Using corpus linguistics methods with critical discourse analysis. *Critical Discourse Studies*, 9(3), 247-256.
- Baker, Paul & Levon, Erez (2015). Picking the right cherries? A comparison of corpus-based and qualitative analyses of news articles about masculinity. *Discourse & Communication*, 9(2), 221-236, <https://doi.org/10.1177/1750481314568542> [Datum des Zugriffs: 25. Oktober 2021].
- Barad, Karen (2003). Posthumanist performativity. Toward an understanding of how matter comes to matter. *Signs: Journal of Women in Culture and Society*, 28(3), 801-831.
- Barbaresi, Adrien (2019). The vast and the focused. On the need for thematic web and blog corpora. *Proceedings, Workshop "Challenges in the Management of Large Corpora" (CMLC 7)*, Cardiff, UK, 22. Juli 2019, <https://doi.org/10.14618/ids-pub-9025> [Datum des Zugriffs: 25. Oktober 2021].
- Bauernschmidt, Stefan (2014). Kulturwissenschaftliche Inhaltsanalyse prozessgenerierter Daten. In Christine Bischoff, Karoline Oehme-Jüngling & Walter Leimgruber (Hrsg.), *Methoden der Kulturanthropologie* (S.415-430). Bern: UTB.
- Baur, Nina & Graeff, Peter (2021). Datenqualität und Selektivitäten digitaler Daten. Alte und neue digitale und analoge Datensorten im Vergleich. In Birgit Blättel-Mink (Hrsg.), *Gesellschaft unter Spannung. Verhandlungen des 40. Kongresses der Deutschen Gesellschaft für Soziologie 2020*, Berlin/digital, 14.-24. September 2020, https://publikationen.sozilogie.de/index.php/kongressband_2020/article/view/1362 [Datum des Zugriffs: 13. März 2022].
- Beck, Stefan (2009). Vergesst Kultur – wenigstens für einen Augenblick! Oder: Zur Vermeidbarkeit der kulturtheoretischen Engführung ethnologischen Forschens. In Sonja Windmüller, Beate Binder & Thomas Hengartner (Hrsg.), *Kultur – Forschung. Zum Profil einer volkskundlichen Kulturwissenschaft* (S.48-68). Berlin: LIT.
- Beck, Stefan (2019 [2015]). Von Praxistheorie 1.0 zu 3.0. Oder: wie analoge und digitale Praxen relationiert werden sollten. *Berliner Blätter. Ethnographische und ethnologische Beiträge*, 81, 9-27.
- Beck, Stefan; Niewöhner, Jörg & Sørensen, Estrid (Hrsg.) (2012). *Science and Technology Studies. Eine sozialanthropologische Einführung*. Bielefeld: transcript.
- Berkelaar, Brenda L. & Francisco-Revilla, Luis (2018). Motivation, evidence and computation. A research framework for expanding computational social science participation and design. In Cathleen Stützer, Martin Welker & Marc Egger (Hrsg.), *Computational social science in the age of big data. Concepts, methodologies, tools, and applications* (S.16-62). Köln: Herbert von Halem.
- Blätte, Andreas; Behnke, Joachim; Schnapp, Kai-Uwe & Wagemann, Claudius (Hrsg.) (2018). *Computational Social Science. Die Analyse von Big Data*. Baden-Baden: Nomos.
- Blei, David M. (2012). Probabilistic topic models. Surveying a suite of algorithms that offer a solution to managing large document archives. *Communications of the ACM*, 55, 77-84, <http://www.cs.columbia.edu/~blei/papers/Blei2012.pdf> [Datum des Zugriffs: 25. Oktober 2021].
- Boellstorff, Tom (2016). For whom the ontology turns. Theorizing the digital real. *Current Anthropology*, 57, 387-406.
- boyd, danah & Crawford, Kate (2012). Critical questions for big data. Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662-679.
- Bubenhof, Noah (2017). Kollokationen, n-Gramme, Mehrworteinheiten. In Kersten Sven Roth, Martin Wengeler & Alexander Ziem(Hrsg.), *Handbuch Sprache in Politik und Gesellschaft* (S.69-93). Berlin: De Gruyter.
- Charmaz, Kathy (2014). *Constructing grounded theory. Introducing qualitative methods*. Los Angeles, CA: Sage.
- Cioffi-Revilla, Claudio (2018). *Introduction to computational social science. Principles and applications* (2. Aufl.). Cham: Springer.
- Conte, Rosaria; Gilbert, Nigel; Bonelli, Giulia; Cioffi-Revilla, Claudio; Deffuant, Guillaume; Kertesz, János; Loreto, Vittorio; Moat, Suzy; Nadal, Jean-Pierre; Sánchez, Angel; Nowak, Andrzej; Flache, Andreas; San Miguel, Maxi & Helbing, Dirk (2012). Manifesto of computational social science. *The European Physical Journal Special Topics*, 214(1), 325-346, <https://doi.org/10.1140/epjst/e2012-01697-8> [Datum des Zugriffs: 25. Oktober 2021].
- D'Andrea, Alessia; Ferri, Fernando; Grifoni, Patrizia & Guzzo, Tiziana (2015). Approaches, tools and applications for sentiment analysis implementation. *International Journal of Computer Applications*, 125(3), 26-33, <http://dx.doi.org/10.5120/ijca2015905866> [Datum des Zugriffs: 25. Oktober 2021].

- Denning, Peter J. & Tedre, Matti (2019). *Computational thinking*. Cambridge, MA: The MIT Press.
- Dourish, Paul & Gómez Cruz, Edgar (2018). Datafication and data fiction. Narrating data and narrating with data. *Big Data & Society*, 5(2), 1-10, <https://doi.org/10.1177/2053951718784083> [Datum des Zugriffs: 25. Oktober 2021].
- Drucker, Johanna (2020). *Visualization and interpretation. Humanistic approaches to display*. Cambridge, MA: The MIT Press.
- Evers, Jeanine; Caprioli, Mauro Ugo; Nöst, Stefan & Wiedemann, Gregor (2020). What is the REFI-QDA standard. Experimenting with the transfer of analyzed research projects between QDA software. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 21(2), Art. 22, <http://dx.doi.org/10.17169/fqs-21.2.3439> [Datum des Zugriffs: 25. Oktober 2021].
- Flanders, Julia & Jannidis, Fotis (2016). Data modeling. In Susan Schreibman, Raymond George Siemens & John Unsworth (Hrsg.), *A new companion to digital humanities* (S.229-237). Chichester: Wiley Blackwell.
- Flick, Uwe (2007). *Qualitative Sozialforschung. Eine Einführung*. Reinbek: Rowohlt.
- Fortun, Kim; Fortun, Mike; Bigras, Erik; Saheb, Tahereh; Costelloe-Kuehn, Brandon; Crowder, Jerome; Price, Daniel & Kenner, Alison (2014). Experimental ethnography online. *Cultural Studies*, 28(4), 632-642.
- Franken, Lina (2020a). Kulturwissenschaftliches digitales Arbeiten. Qualitative Forschung als "digitale Handarbeit"? *Berliner Blätter. Ethnographische und ethnologische Beiträge*, 82, 107-118, <https://www2.hu-berlin.de/ifeeoj/index.php/blaetter/article/view/1069/16> [Datum des Zugriffs: 25. Oktober 2021].
- Franken, Lina (2020b). Methodologie der Zukunft? Automatisierungspotentiale in kulturwissenschaftlicher Forschung. In Dagmar Hänel, Ove Sutter, Ruth Dorothea Eggel, Fabio Freiberg, Andrea Graf, Victoria Huszka & Kerstin Wolff (Hrsg.), *Planen. Hoffen. Fürchten. Zur Gegenwart der Zukunft im Alltag* (S.217-233). Münster: Waxmann.
- Franken, Lina (2022a). Erweiterungen der Digital Humanities durch kulturwissenschaftliche Perspektiven. In Michaela Geierhos (Hrsg.), *Konferenzabstracts. 8. Tagung des Verbands Digital Humanities im deutschsprachigen Raum e.V. (DHd 2022). Kulturen des digitalen Gedächtnisses*, Potsdam/digital. 7.-11. März 2022 (S.101-104), <https://zenodo.org/record/6327985> [Datum des Zugriffs 2. Mai 2022].
- Franken, Lina (2022b/im Druck). Invisible patients? Patients' agency within the discourse on telemedicine. *Curare. Zeitschrift für Medizinethnologie*, 45(2).
- Franken, Lina; Koch, Gertraud & Zinsmeister, Heike (2020). Annotationen als Instrument der Strukturierung. In Julia Nantke & Frederik Schlupkothén (Hrsg.), *Annotation in scholarly editions and research* (S.89-108). Berlin: De Gruyter, <https://doi.org/10.1515/9783110689112-005> [Datum des Zugriffs: 25. Oktober 2021].
- Gaidys, Uta; Gius, Evelyn; Jarchow, Margarete; Koch, Gertraud; Menzel, Wolfgang; Orth, Dominik & Zinsmeister, Heike (2017). Project description – hermA: Automated modelling of hermeneutic processes. *Hamburger Journal für Kulturanthropologie*, 7, 119-123, <https://journals.sub.uni-hamburg.de/hjk/article/view/1213> [Datum des Zugriffs: 25. Oktober 2021].
- Geertz, Clifford (1983 [1973]). *Dichte Beschreibung. Beiträge zum Verstehen kultureller Systeme*. Frankfurt/M.: Suhrkamp.
- Geiger, Jonathan & Pfeiffer, Jasmin (2020). Spielplätze der Theoriebildung in den Digital Humanities. In Christoph Schöch (Hrsg.), *Konferenzabstracts. 7. Tagung des Verbands Digital Humanities im deutschsprachigen Raum e.V. (DHd 2020). Spielräume. Digital Humanities zwischen Modellierung und Interpretation*, Universität Paderborn, 2.-6. März 2020 (S.48-51), <https://zenodo.org/record/4621980> [Datum des Zugriffs: 25. Oktober 2021].
- Gibbs, Graham R.; Friese, Susanne & Mangabeira, Wilma C. (2002). Technikeinsatz im qualitativen Forschungsprozess. Einführung zu FQS Band 3(2). *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 3(2), Art. 8, <http://dx.doi.org/10.17169/fqs-3.2.847> [Datum des Zugriffs: 25. Oktober 2021].
- Glaser, Barney G. & Strauss, Anselm L. (2010 [1967]). *Grounded Theory. Strategien qualitativer Forschung*. Bern: Huber.
- Hine, Christine (2015). *Ethnography for the internet. Embedded, embodied and everyday*. London: Bloomsbury Academic.
- Ignatow, Gabe & Mihalcea, Rada F. (2017). *Text mining. A guidebook for the social sciences*. Los Angeles, CA: Sage.

- Jacobs, Thomas & Tschötschel, Robin (2019). Topic models meet discourse analysis. A quantitative tool for a qualitative approach. *International Journal of Social Research Methodology*, 22(5), 469-485.
- Jannidis, Fotis (2017). Grundlagen der Datenmodellierung. In Fotis Jannidis, Hubertus Kohle & Malte Rehbein (Hrsg.), *Digital Humanities. Eine Einführung* (S.99-108). Stuttgart: J.B. Metzler.
- Jannidis, Fotis; Kohle, Hubertus & Rehbein, Malte (2017a). Warum ein Lehrbuch für Digital Humanities?. In Fotis Jannidis, Hubertus Kohle & Malte Rehbein (Hrsg.), *Digital Humanities. Eine Einführung* (S.XI-XIII). Stuttgart: J.B. Metzler.
- Jannidis, Fotis; Kohle, Hubertus & Rehbein, Malte (Hrsg.) (2017b), *Digital Humanities. Eine Einführung*. Stuttgart: J.B. Metzler.
- Keller, Reiner (2011 [2005]). *Wissenssoziologische Diskursanalyse. Grundlegung eines Forschungsprogramms* (3. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kinder-Kurlanda, Katharina E. (2020). Big Social Media Data als epistemologische Herausforderung für die Soziologie. In Sabine Maasen & Jan-Hendrik Passoth (Hrsg.), *Soziologie des Digitalen – Digitale Soziologie?* (S.109-133). Baden-Baden: Nomos.
- Kitchin, Rob (2014a). *The data revolution. Big data, open data, data infrastructures & their consequences*. Los Angeles, CA: Sage.
- Kitchin, Rob (2014b). Big data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1), 1-12, <https://doi.org/10.1177%2F2053951714528481> [Datum des Zugriffs: 25. Oktober 2021].
- Kitchin, Rob & Lauriault, Tracey P. (2018). Toward critical data studies: Charting and unpacking data assemblages and their work. In Jim Thatcher, Andrew Shears & Josef Eckert (Hrsg.), *Thinking big data in geography. New regimes, new research* (S.3-20). Lincoln, NE: University of Nebraska Press.
- Koch, Gertraud & Franken, Lina (2020). Filtern als digitales Verfahren in der wissenssoziologischen Diskursanalyse. Potentiale und Herausforderungen der Automatisierung im Kontext der Grounded Theory. In Samuel Breidenbach, Peter Klimczak & Christer Petersen (Hrsg.), *Soziale Medien. Interdisziplinäre Zugänge zur Onlinekommunikation* (S.121-138). Wiesbaden: Springer.
- Klug, Helmut W. (Hrsg.) (2021). *KONDE Weißbuch. Kompetenznetzwerk Digitale Edition*, <https://www.digitale-edition.at/> [Datum des Zugriffs: 13. März 2022].
- Knox, Hannah & Nafus, Dawn (Hrsg.) (2018). *Ethnography for a data-saturated world*. Manchester: Manchester University Press.
- Kropf, Jonathan (2019). Recommender Systems in der populären Musik. Kritik und Gestaltungsoptionen. In Jonathan Kropf & Stefan Laser (Hrsg.), *Digitale Bewertungspraktiken. Für eine Bewertungssoziologie des Digitalen* (S.127-163). Wiesbaden: Springer VS.
- Legewie, Nicolas & Nassauer, Anne (2018). YouTube, Google, Facebook: 21st century online video research and research ethics. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 19(3), Art. 32, <https://doi.org/10.17169/fqs-19.3.3130> [Datum des Zugriffs: 25. Oktober 2021].
- Leimgruber, Walter (2013). Entgrenzungen. Kultur – empirisch. In Reinhard Jöhler, Christian Marchetti, Bernhard Tschofen & Carmen Weith (Hrsg.), *Kultur_Kultur. Denken. Forschen. Darstellen. 38. Kongress der Deutschen Gesellschaft für Volkskunde, Tübingen, 21.-24. September 2011* (S.71-85). Münster: Waxmann.
- Lemke, Matthias & Wiedemann, Gregor (Hrsg.) (2016). *Text Mining in den Sozialwissenschaften. Grundlagen und Anwendungen zwischen qualitativer und quantitativer Diskursanalyse*. Wiesbaden: Springer VS.
- Lewis, Lori (2020). Infographic: What happens in an internet minute 2020. *AllAccess.com*, 10. März, <https://www.allaccess.com/merge/archive/31294/infographic-what-happens-in-an-internet-minute> [Datum des Zugriffs: 25. Oktober 2021].
- Light, Ben; Burgess, Jean & Duguay, Stefanie (2017). The walkthrough method. An approach to the study of apps. *New Media & Society*, 20(3), 881-900.
- Lindner, Rolf (2003). Vom Wesen der Kulturanalyse. *Zeitschrift für Volkskunde*, 99, 177-188.
- Liu, Bing (2015). *Sentiment analysis. Mining opinions, sentiments, and emotions*. Cambridge: Cambridge University Press.
- Lupton, Deborah (2015). *Digital sociology*. London: Routledge Taylor & Francis Group.

- MacMillan, Katie (2005). More than just coding? Evaluating CAQDAS in a discourse analysis of news texts. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 6(3), Art. 25, <https://doi.org/10.17169/fqs-6.3.28> [Datum des Zugriffs: 25. Oktober 2021].
- Manovich, Lev (2020). *Cultural analytics*. Cambridge, MA: The MIT Press.
- Marcus, George E. (2009). Introduction. Notes toward an ethnographic memoir of supervising graduate research through anthropology's decades of transformation. In James D. Faubion & George E. Marcus (Hrsg.), *Fieldwork is not what it used to be. Learning anthropology's method in a time of transition* (S.1-34). Ithaca, NY: Cornell University Press.
- Marres, Noortje (2017). *Digital sociology. The reinvention of social research*. Cambridge: Polity.
- Mau, Steffen (2017). *Das metrische Wir. Über die Quantifizierung des Sozialen*. Frankfurt/M.: Suhrkamp.
- Mayer-Schönberger, Viktor & Cukier, Kenneth (2013). *Big data. A revolution that will transform how we live, work and think*. London: Muray.
- McCallum, Andrew Kachites (2002). *MALLET. A machine learning for language toolkit*, <http://mallet.cs.umass.edu/> [Datum des Zugriffs: 25. Oktober 2021].
- McCarty, Willard (2004). Modeling. A study in words and meanings. In Susan Schreibman, John Unsworth & Raymond George Siemens (Hrsg.), *A companion to digital humanities* (S.254-270). Malden, MA: Blackwell.
- McCarty, Willard (2005). *Humanities computing*. Basingstoke: Palgrave Macmillan.
- Morse, Janice M. & Niehaus, Linda (2016). *Mixed method design. Principles and procedures*. London: Taylor and Francis.
- Mosconi, Gaia; Li, Qinyu; Randall, Dave; Karasti, Helena; Tolmie, Peter; Barutzky, Jana; Korn, Matthias & Pipek, Volkmar (2019). Three gaps in opening science. *Computer Supported Cooperative Work (CSCW)*, 28(3-4), 749-789.
- Muri, Gabriela (2014). Triangulationsverfahren im Forschungsprozess. In Christine Bischoff, Karoline Oehme-Jüngling & Walter Leimgruber (Hrsg.), *Methoden der Kulturanthropologie* (S.459-473). Bern: UTB.
- Nelson, Laura K. (2020). Computational grounded theory. A methodological framework. *Sociological Methods & Research*, 49, 3-42, <https://doi.org/10.1177/0049124117729703> [Datum des Zugriffs: 25. Oktober 2021].
- Paulus, Trena M.; Jackson, Kristi & Davidson, Judith (2017). Digital tools for qualitative research: Disruptions and entanglements. *Qualitative Inquiry*, 23(10), 751-756.
- Peetz, Thorsten (2021). Digitalisierte intime Bewertung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 73, 425-450.
- Pink, Sarah; Ruckenstein, Minna; Willim, Robert & Duque, Melisa (2018). Broken data. Conceptualising data in an emerging world. *Big Data & Society*, 5(1), <https://doi.org/10.1177/2053951717753228> [Datum des Zugriffs: 25. Oktober 2021].
- Pink, Sarah; Horst, Heather A.; Postill, John; Hjorth, Larissa; Lewis, Tania & Tacchi, Jo (2016). *Digital ethnography. Principles and practice*. Los Angeles, CA: Sage.
- Pustejovsky, James & Stubbs, Amber (2012). *Natural language annotation for machine learning*. Beijing: O'Reilly.
- Reckwitz, Andreas (2004). Die Kontingenzperspektive der "Kultur". Kulturbegriffe, Kulturtheorien und das kulturwissenschaftliche Forschungsprogramm. In Friedrich Jaeger & Jörn Rüsen (Hrsg.), *Handbuch der Kulturwissenschaften. Band 3: Themen und Tendenzen* (S.1-20). Stuttgart: J.B. Metzler.
- Reckwitz, Andreas (2008). Praktiken und Diskurse. Eine sozialtheoretische und methodologische Relation. In Herbert Kalthoff, Stefan Hirschauer & Gesa Lindemann (Hrsg.), *Theoretische Empirie. Zur Relevanz qualitativer Forschung* (2. Aufl., S.188-209). Frankfurt/M.: Suhrkamp.
- Rehbein, Malte (2017). Informationsvisualisierung. In Fotis Jannidis, Hubertus Kohle & Malte Rehbein (Hrsg.), *Digital Humanities. Eine Einführung* (S.328-342). Stuttgart: J.B. Metzler.
- Reichert, Ramón & Richterich, Annika (2015). Introduction. Digital materialism. *Digital Culture & Society*, 1, 5-17, <https://doi.org/10.25969/mediarep/634> [Datum des Zugriffs: 25. Oktober 2021].
- [Reichert, Jo](#) (2013). *Die Abduktion in der qualitativen Sozialforschung. Über die Entdeckung des Neuen* (2. Aufl.). Wiesbaden: Springer VS.

- Rieker, Peter; Hartmann Schaelli, Giovanna & Jakob, Silke (2020). Zugang ist nicht gleich Zugang. Verläufe, Bedingungen und Ebenen des Feldzugangs in ethnografischen Forschungen. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 21(2), Art. 19, <https://doi.org/10.17169/fqs-21.2.3353> [Datum des Zugriffs: 25. Oktober 2021].
- Rogers, Richard (2019). *Doing digital methods*. London: Sage.
- Ruppert, Evelyn; Law, John & Savage, Mike (2013). Reassembling social science methods: The challenge of digital devices. *Theory, Culture & Society*, 30(4), 22-46.
- Sahle, Patrick (2015). Digital Humanities? Gibt's doch gar nicht!. *Zeitschrift für digitale Geisteswissenschaften*, http://dx.doi.org/10.17175/sb001_004 [Datum des Zugriffs: 25. Oktober 2021].
- Sahle, Patrick (2019). *Professuren für Digital Humanities*, <https://dhd-blog.org/?p=11018> [Datum des Zugriffs: 25. Oktober 2021].
- Salganik, Matthew J. (2018). *Bit by bit. Social research in the digital age*. Princeton, NJ: Princeton University Press.
- Schreibman, Susan; Unsworth, John & Siemens, Raymond George (Hrsg.) (2004). *A companion to digital humanities*. Malden; MA: Blackwell.
- Sismondo, Sergio (2010). *An introduction to science and technology studies* (2. Aufl.). Chichester: Wiley-Blackwell.
- Smith, Rachel Charlotte (2013). Designing heritage for a digital culture. In Wendy Gunn, Ton Otto & Rachel Charlotte Smith (Hrsg.), *Design anthropology. Theory and practice* (S.177-138). London: Bloomsbury.
- Snee, Helene; Hine, Christine; Morey, Yvette; Roberts, Steven & Watson, Hayley (Hrsg.) (2016). *Digital methods for social science. An interdisciplinary guide to research innovation*. Basingstoke: Palgrave Macmillan.
- Stiemer, Haimo & Vauth, Michael (2020). "Mit den Himbeeren stimmte etwas nicht". Die Ernährung in der dystopischen Gegenwartsliteratur. In Jennifer Grünewald & Anja Trittelvitz (Hrsg.), *Ernährung und Identität* (S.203-222). Stuttgart: ibidem.
- Strübing, Jörg (2019). Grounded Theory und Theoretical Sampling. In Nina Baur & Jörg Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S.525-544). Wiesbaden: Springer VS.
- Strübing, Jörg; Hirschauer, Stefan; Ayaß, Ruth; Krähnke, Uwe & Scheffer, Thomas (2018). Gütekriterien qualitativer Sozialforschung. Ein Diskussionsanstoß. *Zeitschrift für Soziologie*, 47(2), 83-100.
- Svensson, Patrik (2016). *Big digital humanities*. Ann Arbor, MI: University of Michigan Press.
- Thaller, Manfred (2017). Digital Humanities als Wissenschaft. In Fotis Jannidis, Hubertus Kohle & Malte Rehbein (Hrsg.), *Digital Humanities. Eine Einführung* (S.13-18). Stuttgart: J.B. Metzler.
- Trier, Jost (1973 [1931]). *Über Wort- und Begriffsfelder*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Turner, Victor (2005 [1969]). *Das Ritual. Struktur und Anti-Struktur*. Frankfurt/M.: Campus.
- Universität Klagenfurt (2021). *Humanwissenschaft des Digitalen*, <https://www.aau.at/digital-age-research-center/humanwissenschaft-des-digitalen/> [Datum des Zugriffs: 25. Oktober 2021].
- Unsworth, John (2000). Scholarly primitives. What methods do humanities researchers have in common, and how might our tools reflect this. *Beitrag, Symposium "Humanities Computing: Formal Methods, Experimental Practice"*, King's College, London, 13. Mai 2000, <https://people.brandeis.edu/~unsworth/Kings.5-00/primitives.html> [Datum des Zugriffs: 25. Oktober 2021].
- van Gennep, Arnold (1999 [1909]). *Übergangsriten*. Frankfurt/M.: Campus.
- Wiechmann, Brigitte (2012). Die Gemeinsame Normdatei (GND). Rückblick und Ausblick. *Dialog mit Bibliotheken*, 24(2), 20-22, <https://d-nb.info/1038598885/34> [Datum des Zugriffs: 25. Oktober 2021].
- Wiedemann, Gregor; Remus, Steffen; Chawla, Avi & Biemann, Chris (2019). Does BERT make any sense? Interpretable word sense disambiguation with contextualized embeddings. In German Society for Computational Linguistics & Language Technology (Hrsg.), *Proceedings of the 15th Conference on Natural Language Processing* (S.161-170), <https://corpora.linguistik.uni->

erlangen.de/data/konvens/proceedings/papers/KONVENS2019_paper_43.pdf [Datum des Zugriffs: 28. April 2022].

Witt, Harald (2001). Forschungsstrategien bei quantitativer und qualitativer Sozialforschung. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 2(1), Art. 8, <https://doi.org/10.17169/fqs-2.1.969> [Datum des Zugriffs: 25. Oktober 2021].

Woletz, Julie (2016). *Human-Computer Interaction. Kulturanthropologische Perspektiven auf Interfaces*. Darmstadt: BÜCHNER.

Zur Autorin

Lina FRANKEN vertritt die Professur für Computational Social Sciences am Institut für Soziologie der Ludwig-Maximilians-Universität München. Zuvor war sie Projektkoordinatorin für den Forschungsverbund "Automatisierte Modellierung hermeneutischer Prozesse – Der Einsatz von Annotationen für sozial- und geisteswissenschaftliche Analysen im Gesundheitsbereich (hermA)" am Institut für Empirische Kulturwissenschaft der Universität Hamburg (2018-2021) und Projektkoordinatorin für das DFG-Projekt "Digitales Portal Alltagskulturen im Rheinland" beim Landschaftsverband Rheinland (2013-2017). Promotion im Fach vergleichende Kulturwissenschaft an der Universität Regensburg 2017, Studium der Volkskunde, neueren Geschichte und Medienwissenschaft an der Universität Bonn. Forschungsschwerpunkte: *Computational Social Sciences* und *Digital Humanities*, Methodologie und digitale Methodenentwicklung, Technisierung und Digitalisierung in Alltag und Wissenschaft, Bildungskulturen und -politik, immaterielles Kulturerbe, Arbeits- und Nahrungskulturen.

Kontakt:

Prof. Dr. Lina Franken

Ludwig-Maximilians-Universität München
Institut für Soziologie
Konradstraße 6 80801, München

E-Mail: lina.franken@soziologie.uni-muenchen.de

URL: https://www.css.soziologie.uni-muenchen.de/personen/leitung/lina_franken/index.html

Zitation

Franken, Lina (2022). Digitale Daten und Methoden als Erweiterung qualitativer Forschungsprozesse. Herausforderungen und Potenziale aus den *Digital Humanities* und *Computational Social Sciences* [43 Absätze]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 23(2), Art. 12, <https://doi.org/10.17169/fqs-22.2.3818>.