

# Erzählungen durch Daten: distanzierte, engagierte, responsable, kollektive und planetare<sup>1</sup>

In Zusammenarbeit mit dem *American Museum of Natural History* wird aktuell die Software *Collectionscope* entwickelt, die die Objekte der Museumssammlung auf individuellen Bildschirmen sichtbar macht. Eine Beta-Version des Projekts verspricht, dass 250.000 ethnografische und archäologische Objekte des Museums mit *Collection-scope* zugänglich gemacht werden.<sup>2</sup> Ich klicke auf die Beta-Version des Programms und sehe auf schwarzem Hintergrund eine Sphäre von weißen Punkten, die mich an einen Sternenhimmel erinnert. Die Punkte kommen näher und verwandeln sich in kleine Vierecke, die schnell Farbe annehmen. Indem sie größer werden, zeigt sich, dass es Bilder sind, die aussehen, als würden sie durch den Weltraum jagen. Wenn die Entwicklung des Programms finalisiert ist, soll man in dieser „random“ genannten Visualisierung Informationen über die abgebildeten Objekte erhalten sowie über die Expeditionen, die die Objekte zum Museum gebracht haben. Über die „random“-Visualisierung hinaus gibt es eine Visualisierung, die die Objekte auf einer Mercatorschen Weltkarte an dem Ort platziert, an dem sie gefunden

- 1 Ich bedanke mich ganz herzlich bei den Teilnehmer\*innen an der ÖGEKW-Tagung in Klagenfurt am Wörthersee 2023, die mit ihren Diskussionsbeiträgen bei meiner Keynote maßgeblich zur Weiterbearbeitung meiner Datenforschung beigetragen haben. Teile der Forschung, die in diesem Aufsatz diskutiert werden, sind bei einer Fellowship am Center for Advanced Internet Studies (CAIS) entstanden.
- 2 <https://amnh-sciviz.github.io/collectionscope/apps/amnh/> (Zugriff: 1.7.2024). Siehe auch John S. Hansen: Cutting Edge and Cutting Corners: Evolving Technology, Expanding Usership, and Responsive Solutions in a Museum Database. In: *Advances in Archaeological Practice*, 7(3), 2019.

wurden. In einer dritten Visualisierung befinden sich die Bilder in einer Zeitröhre mit einer Länge von 170 Jahren. Die letzte Visualisierungsart besteht aus einer Darstellung der abgebildeten Objekte in einer vom Programm bezeichneten Rennstrecke, die die Bilder der Museumsobjekte auf Bahnen zeigt, deren Breite Kontinente und deren Länge die Zeit repräsentieren – eine besondere Art, Ort und Zeit gleichzeitig zu visualisieren.

Wie Paul Dourish<sup>3</sup> darauf hinweist, dass ein Feld in einer Exceltabelle eine Handlungseinheit ist (unit of action), die dafür da ist, bearbeitet zu werden, so sind auch die klickbaren Objekte im *Collectionscope* Handlungseinheiten. Ich kann auf die 250.000 Objekte klicken, sie ansehen und mehr über sie erfahren. Die Visualisierung bietet mir das Gefühl einer transzendenten Macht, 250.000 Objekte zur Hand zu haben und mit ihnen mühelos und konsequenzlos zu tun, was mir angeboten wird. Zugleich begegne ich den Objekten aus einer unendlichen Distanz, die mir niemals eine körperliche und materielle Nähe zu ihnen ermöglichen wird.

Bei meinem letzten Besuch des *American Museum of Natural History* in New York im Jahr 2022 verweilte ich lange bei Objekten chinesischer Wissenschaft und afrikanischer Geschlechterbeziehungen, die bei mir das Nachdenken evozierten. Auch wie die Objekte kuratiert sind, beeindruckte mich: Viele Objekte sind in Vitrinen, andere freistehend oder hängend gestaltet. Es wird mit Farben und Licht in den Räumen gearbeitet und überhaupt bezaubert mich die Anordnung der Räume, die auf mich wie ein Labyrinth wirkt, in dem ich mich verliere. Sie gibt mir ein Gefühl von Demut vor den vielen Gegenständen und macht mich zugleich neugierig, mehr zu entdecken. Zudem bemerkte ich bei meinem jüngsten Besuch, wie die Wandfarbe in großen Blättern von manchen Wänden und Decken hing und abzufallen drohte. Sie erinnerte an den knappen Haushalt öffentlicher kultureller Einrichtungen in den Vereinigten Staaten, der in starkem Kontrast zu dem privatwirtschaftlichen Luxus und Überverbrauch stand, den ich draußen auf den Straßen von New York erlebte. In vielerlei Hinsicht machte das Museum auf sich als Museum aufmerksam und die Objekte präsentierten sich unverkennbar als kuratierte Ausstellungsobjekte.

3 Vgl. Paul Dourish: *The Stuff of Bits: An Essay on the Materialities of Information*. Cambridge 2017.

Die Objekte des Museums oder ihre Abbildungen sowie deren Umgebung sind Daten für meine zwei kurzen Erzählungen aus dem Museumsbesuch sowie aus meiner Nutzung von *Collectionscope*. Sie sind Daten im Sinn von aufbereitetem und rekontextualisiertem Material, das eine lesbare Form angenommen hat und als solche selbst als Material für ihre Weiterbearbeitung gilt, wobei sie bei der Weiterbearbeitung ihren indexikalen Charakter beibehält.<sup>4</sup> Durch ihre unterschiedliche Situierung<sup>5</sup> und ihre unterschiedlichen Arten, mir zu begegnen, ließen die Daten aus *Collectionscope* und die Daten aus dem Museum mich verschiedene Erzählungen zum Ausdruck bringen.

Ich möchte hier auf den Unterschied zwischen meinen Erinnerungen aus dem Besuch vor Ort des *American Museum of Natural History* an der Upper West Side in New York und der Erzählung derselben Sammlung im *Collectionscope* hinweisen. Mein Besuch im Museum hat mich dazu bewegt, von Spannungen in der Stadt zu berichten, von der wirtschaftlichen und politischen Situation in den Vereinigten Staaten, von Objekten, bei denen ich bei meinem Besuch verweilt und nachgedacht habe. Die Erzählung meiner Erinnerung ist affektiv geprägt. Im Gegensatz dazu regte *Collectionscope* eine Erzählung aus der Perspektive der gesamten Sammlung und über die räumlichen sowie zeitlichen Herkünfte der Objekte an. Die Darstellung der individuellen Objekte war in meiner Erzählung eher technowissenschaftlich geprägt.

Pierre Bourdieu<sup>6</sup> beklagt, dass Architekturmodelle Gebäude und Städte aus einem Blick von Oben darstellen, den kein Mensch jemals haben wird, denn er erlebt Gebäude und Städte durch seine körperliche Bewegung in, durch und neben diesen Einrichtungen, nicht über ihnen. Dieser Vergleich spiegelt den Unterschied zwischen

- 4 Vgl. Bruno Latour: Zirkulierende Referenz: Bodenstichproben aus dem Urwald am Amazonas. In: Ders.: Die Hoffnung der Pandora: Untersuchungen zur Wirklichkeit der Wissenschaft. Berlin 2002, S. 36–95. Die Objekte – die Daten – verweisen in beiden Fällen auf dasselbe Material, auf die Objekte in der Sammlung: Im *Collectionscope* verweisen die Bilder auf Objekte in der Sammlung im *American Museum of Natural History*, die die gleichen Objekte sind (das gleiche Material), die ich bei meinem Museumsbesuch gesehen habe.
- 5 Vgl. Yannis A. Loukissas: All Data Are Local: Thinking Critically in a Data-Driven Society. Cambridge, MA 2022.
- 6 Vgl. Pierre Bourdieu: Homo Academicus. Berlin 1992.

meinen Erinnerungen an den Besuch im Museum und meiner Erzählung von *Collectionscope* wider. Letztere zeigt die Objekte aus einer distanzierten Perspektive vor dem Raster der Weltkarte und einer linearen Zeitlinie, wobei die Erinnerung an meinen Besuch einen engagierten Charakter hat.

Der Vergleich mag ungerecht angelegt erscheinen. Ich habe mich mit *Collectionscope* nur als Beta-Version beschäftigt und deutlich weniger Zeit mit ihr verbracht als im *American Museum of Natural History*. Die finale Version von *Collectionscope*, die Zugang zu mehr Objekten geben wird, wird mich sicherlich dazu verleiten, mehr Zeit im Programm zu verbringen. Trotzdem ist es auch sicher, dass ich, wenn ich nach New York reise, mir einen Nachmittag für den Museumsbesuch frei nehme und einen Ticketpreis von 28 US-Dollar bezahle, meine Aufmerksamkeit und Emotionalität voll und ganz und über längere Zeit den Objekten widme. Michael Hagner<sup>7</sup> unterstreicht einen ähnlichen Effekt beim Vergleich von physikalischen Bibliotheken und online zugänglicher Open-Access-Literatur. Die Schnelligkeit, mit der ich ein Open-Access-Buch im Internet finden und durchblättern kann, lässt mich weniger bei den einzelnen Quellen verweilen, sie erleben und über sie nachdenken, als wenn ich in eine Bibliothek gehe – und damit Zeit für den Bibliotheksbesuch reserviert habe, um mich mit Büchern zu beschäftigen – und die Bücher aus dem Regal hole, sie in der Hand halte, sie durchblättere und lese.<sup>8</sup>

Bei Bourdieu sowie bei Hagner geht es um verschiedene Erfahrungen, die distanzierter und mentaler oder engagierter und praktischer sind. Bei meinem Vergleich zwischen meinem Besuch im *American Museum of Natural History* und meiner Nutzung von *Collectionscope* habe ich die Erzählungen fokussiert, die die Nutzung der

7 Vgl. Michael Hagner: Vortrag gehalten an der Ruhr-Universität Bochum 2016.

8 Für Diskussionen über den Vergleich digitaler und nichtdigitaler Museumsobjekte siehe Werner Schweibenz: Das Museumsobjekt im Zeitalter seiner digitalen Repräsentierbarkeit. In: Elke Murlasits, Gunther Reisinger (Hg.): *Museum multimedial: audiovisionäre Traditionen in aktuellen Kontexten*. Münster 2012, S. 47–70; Christoph Bareither, Katharina Geis, Sarah Ullrich u. a.: *Digitales Bildkuratieren*. Hildesheim 2023, haben zudem ethnografisch untersucht, wie digitale Anwendungen nicht getrennt von, sondern auch als Teil von verorteten Ausstellungen kombiniert werden.

Software und der Museumsbesuch inspiriert haben: eine distanziertere und eine engagiertere Erzählung. Von Haraway<sup>9</sup> inspiriert, interessiert mich, welche Geschichten unsere Erfahrungen uns verleiten zu erzählen. Denn es geschieht weitgehend durch Erzählungen, dass wir Beziehungen zu anderen menschlichen und nichtmenschlichen Kollektiven eingehen. Erzählungen stellen Beziehungen her und flechten Kollektive. Verschiedene Erzählungen erzeugen verschiedene Beziehungen, oder wie Haraway unterstreicht: „[i]t matters what stories tell stories“<sup>10</sup>. Erzählungen sind durch situierte Erfahrungen und damit durch die Kontexte geprägt, in denen sie entstanden sind. Ein Besuch im *American Museum of Natural History* und die Nutzung von *Collectionscope* situieren Erfahrungen, unterscheiden sich und haben verschiedene Kontexte, weswegen verschiedene Erzählungen daraus entstehen. Allein der Unterschied zwischen „Besuch“ und „Nutzung“ leitet verschiedene Erzählungen ein.

Erzählungen generieren verschiedene Beziehungen: Die distanzierte Erzählung, die durch meine Nutzung von *Collectionscope* entstanden ist, ist für einen fokussierten Austausch über die Sammlung sowie über einzelne Objekte geeignet. Solche Austausche, die auch ihre Dokumentation voraussetzen, sind für spezialisierte, technowissenschaftliche und professionelle Kontexte geeignet, für Expert\*innenrunden oder auch für bürokratische Kontexte, in denen Phänomene der Welt (so gut wie möglich) isoliert und in Übereinstimmung mit gesetzten Methoden, Vorschriften oder Richtlinien behandelt, Probleme fokussiert und Lösungen identifiziert werden. Sie unterstützen und stärken solche Kollektive, solche Arten von Austauschen und solche Lebensformen, die sich durch fokussierte Informationen und Austausche gut entwickeln.

Die engagierte Erzählung aus dem *American Museum of Natural History* ist in ihrer mäandernden Art für Gespräche des alltäglichen Austauschs geeignet, die weniger ergebnisorientiert sind, sondern teilweise für den Austausch selbst da sind, der bestehende Verständnisse durch die erzählten Phänomene ergänzt, neu beleuchtet oder infrage stellt, komplex verwoben mit anderen Phänomenen

9 Donna J. Haraway: Unruhig bleiben: Die Verwandtschaft der Arten im Chthuluzän. Frankfurt a. M. 2018.

10 Ebd., S. 35.

und Ereignissen. Es gibt Teile der Kultur- und Sozialwissenschaften, die diese Art von Erzählungen zu ihrer Expertise gemacht haben, so wie auch oft die Kunst engagierte über distanzierte Erzählungen bevorzugt.<sup>11</sup> Engagierte Erzählungen etablieren einen Austausch, der das Erleben der Phänomene und Ereignisse schätzt und auch wie sie spezifisch von der erzählenden Person erlebt wurden. Isabelle Stengers<sup>12</sup> assoziiert eine solche Erzählart mit *Connoisseurs* oder intelligenten Öffentlichkeiten, die ich als wild bezeichnen würde, weil sie sich ohne Rücksicht auf disziplinäre Grenzen und ohne Orthodoxien wissenschaftlicher oder bürokratischer Begrifflichkeiten und Vorgänge für die Vielfalt der Existenz von und des Umgangs mit den Phänomenen und Ereignissen interessieren. Engagierte Erzählungen unterstützen und stärken solche Kollektive, solche Arten von Austausch und solche Lebensformen, die durch diese wilde Heterogenität aufblühen.

Wie jede distanzierte Unterscheidung ist auch der Vergleich zwischen distanzierter und engagierter Erzählungen nur als Hinweis auf eine differenzierte Beobachtung zu nehmen und nicht als Definition von zwei voneinander getrennt tatsächlich existierenden Erzählformen. Beide sind komplexer, als hier beschrieben und in- und miteinander verwoben. Die Differenzierung ist bei der Entwicklung verschiedener Erzählungen durch Daten und bei der Berücksichtigung nützlich, bestimmte Kollektive und Arten des Austausches zu unterstützen.

In diesem Essay werde ich weitere Untersuchungen darüber vornehmen, welche Erzählungen Daten ermöglichen und welche Beziehungen sie unterstützen. Als Erstes diskutiere ich, wie Datenvisualisierungen multiple Erzählungen erzeugen können. Darauf folgt eine Analyse der möglichen Responsabilität von Daten sowie

11 Philipp Fischer, Gabriele Gramelsberger, Christoph Hoffmann u. a.: *Natures of Data: A Discussion between Biology, History and Philosophy of Science and Art*. Diaphanes. Zürich 2020, S. 12: Sie notieren, dass die (biologische) Wissenschaft sucht, Daten von ihrer verwirrenden Umwelt zu bereinigen, um sie zu befähigen, als isolierte Zahlen zu dienen, während die Kunst an der materiellen Bedingung und an konkreten Gegebenheiten interessiert ist, in denen Daten situiert sind.

12 Vgl. Isabelle Stengers: *Another Science is Possible: A Manifesto for a Slow Science*. Cambridge 2018.

von Daten als Akteure kollektiver Datenerzählungen. Am Ende des Essays gehe ich auf die planetare Interdependenz wissenschaftlicher Dateninfrastrukturen und ihre weitgehende Abwesenheit in Erzählungen über wissenschaftliche Dateninfrastrukturen ein.

## Datenvisualisierungen und multiple Erzählungen

Die Datenvisualisierung in Abbildung 1 wurde von dem Künstler Kamel Makhloufi erstellt<sup>13</sup>. Die Daten stammen aus der SIGACT-Datenbank, veröffentlicht von Wikileaks und bereinigt von der britischen Tageszeitung *The Guardian*. Jedes Pixel in den Quadraten steht für eine Person, die zwischen dem Jahr 2004 und dem Jahr 2009 während der Besetzung des Irak durch eine von den Vereinigten Staaten angeführte sogenannte „Koalition der Willigen“ getötet wurde. Die blaue Farbe repräsentiert amerikanische Truppen, die hellgrüne Farbe verweist auf irakische Militärs, die orangefarbenen Pixel stehen für Zivilisten und die schwarzen für Rebellen. Auf der linken Seite sind die Verletzten der jeweiligen Gruppen wie in einem Balkendiagramm dargestellt.

Das Quadrat auf der rechten Seite zeigt die Tötungen im Verlauf der Zeit an. Der Beginn des Jahres 2004 befindet sich am oberen Rand des Quadrats, und die Zeit bewegt sich abwärts bis zum Jahr 2009 am unteren Rand. Beide Visualisierungen reduzieren die tragischen Erfahrungen militärischer Aktionen, die sowohl journalistische wie auch ethnografische Berichte eingefangen haben, was uns noch mal daran erinnert, dass in der empirischen Kulturwissenschaft digitale Daten und Datenvisualisierungen immer nur Ergänzungen zu ethnografischen Berichten sein können und sollten.<sup>14</sup> Mein Interesse an den Visualisierungen von Makhloufi ist jedoch ein anderes: Welche Erzählungen legen sie nahe?

13 Kamel Makhloufi: PIXELLATING THE WAR CASUALTIES IN IRAQ, <https://github.com/melka/iraqwarlogs> (Zugriff: 1.7.2024).

14 Vgl. Martina Klausner: A More-than-digital Anthropology. Ethnographies of Participation and Administration. In: Zeitschrift für Empirische Kulturwissenschaft, 118, 2022, S. 5–24; Martina Klausner, Dennis Eckhardt: Digitalität und Ethnografie: Eine Einführung in Forschungsmethoden für mehr-als-digitale Felder. In: Kulturanthropologie Notizen 85, 2023, S. 2–19.

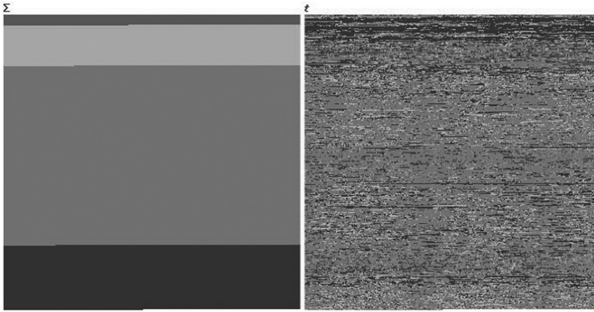


Abb. 1: Kamel Makhloufi (2010): Pixellating the War Casualties in Iraq

Die linke Visualisierung ähnelt einem Balkendiagramm, das Menschen in wenige und grobe Kategorien einordnet. In diesem Fall sind es militärische Kategorien, was nicht verwunderlich ist, da die Visualisierung auf Daten beruht, die das US-amerikanische Militär gesammelt hat. Sie ist hilfreich, um über militärische Erfolge oder Misserfolge zu berichten sowie auch für kritische Auseinandersetzungen und für die Beurteilung der Ungerechtigkeit der Besatzung. Sie erleichtert den Wissensmodus einer distanzierten urteilenden Beobachtung.<sup>15</sup> Sie ist in ihrer Klarheit sehr einfach zu lesen und erleichtert Entscheidungen. Sie schließt eher ein Thema, als dass sie neue Fragen aufwirft.

Die Visualisierung auf der rechten Seite erzählt eine komplexere Geschichte als die linke. Zeitlich geordnet vermittelt die Visualisierung den Eindruck einer chaotischen, vielleicht sogar schmutzigen Auseinandersetzung. In der ersten Periode werden hauptsächlich bewaffnete Kräfte getötet, insbesondere Aufständische. Bald jedoch – wenn wir uns ein Stückchen weiter nach unten bewegen – sind zunehmend Zivilisten unter den Getöteten. Dies – und das ist die bittere Pointe der Erzählung – ändert sich nicht wesentlich, bis das Quadrat, aber nicht die Besatzung endet. Die vier Farben in der Visualisierung auf der rechten Seite sind alle miteinander vermischt und in räumlicher Nähe zueinander, was die Visualisierung weniger dazu geeignet macht, über die quantitativen Differenzen des Tötens von Rebell\*innen, Zivilist\*innen, irakischen und amerikanischen Truppen

15 Vgl. Estrid Sørensen: *The Materiality of Learning: Technology and Knowledge in Educational Practice*. Cambridge 2009.

zu erzählen. Vielmehr befähigt sie uns dazu, über die militärischen Auseinandersetzungen als ein fortlaufendes Geschehen zu reden, in dem Menschen verschiedener Zugehörigkeiten nicht sauber getrennt agieren, sondern alle gemeinsam in das Elend des Konfliktes verwickelt sind. Ich stelle mir die repräsentierten Personen nicht mehr als Menschen vor – wie bei der linken Visualisierung –, die streng voneinander getrennt sind und nur mit Menschen ihrer eigenen Kategorie zusammenleben, sondern ich sehe vor meinem inneren Auge, wie sie alle auf irakischen Straßen entlanglaufen, in denselben Läden ihr Brot kaufen und auf dem Bürgersteig aneinander vorbeilaufen. Bis sie auch dort durch Bomben oder Schießereien getötet werden. Die Pixel sind durchmischt, die Kategorien weniger prägend. Diese Visualisierung evoziert dichtere Erzählungen und fordert auch mehr Zeit, um sie zu verstehen. Wir müssen uns auf sie einlassen und die Geschichte nacherzählen, um die Visualisierung zu verstehen. Sie lädt weniger zu einer distanziert urteilenden Beobachtung ein und eher zum Nachdenken. Zum Teil auch einfach deswegen, weil sie ungewöhnlich ist.

Es ist faszinierend, wie sehr das Hinzufügen einer einzigen Variable zur Visualisierung – Zeit – die Erzählung verändert. Eher als zwei einzelne Visualisierungen für sich ist Makhloufis Doppeltvisualisierung spannend, weil sie uns zwei verschiedene Geschichten erzählen lässt. Es ist die Differenz zwischen den Erzählungen, die das Nachdenken sowie das Neudenken fördert<sup>16</sup>, getrieben durch die Friktion, die zwischen ihnen entsteht.<sup>17</sup> Datenforschung bietet die Möglichkeit, sehr viele verschiedene Visualisierungen zu erstellen und damit verschiedene Geschichten zu erzählen. Das macht sie m. E. produktiv für die empirische Kulturwissenschaft. Mit unter anderem Annemarie Mol<sup>18</sup> können wir unterstreichen, dass die zwei

16 Vgl. John Dewey: *The Public and Its Problems: An Essay in Political Inquiry*. Athens 2012; Tanja Bogusz: *Experimentalismus und Soziologie. Von der Krisen- zur Erfahrungswissenschaft*. Frankfurt a. M., New York 2018.

17 Vgl. Anna L. Tsing: *Friction. An Ethnography of Global Connections*. Princeton 2005.

18 Vgl. Annemarie Mol: *The Body Multiple: Ontology in Medical Practice*. Durham 2003; Estrid Sørensen, Jan Schank: *Praxeographie [Praxeography]*. In: Susanne Bauer, Torsten Heinemann, Thomas Lemke (Hg.): *Science & Technology Studies: Klassische Positionen und aktuelle Perspektiven*. Frankfurt a. M. 2017, S. 407–428.

verschiedenen datengenerierten Erzählungen nicht auf zwei unterschiedliche Konflikte hinweisen, oder auf verschiedene Aspekte eines Konflikts, zwischen denen entschieden werden kann oder sollte. Vielmehr weisen Doppelt- oder Mehrfachvisualisierungen auf die Multiplizität eines Phänomens hin, die eine eindeutige und vollendete Erzählung unmöglich macht und die Komplexität und Heterogenität von situierten Praktiken sowie von Erinnerungen fokussiert.

### **Von Datenerzählungen zu Responsabilität durch Daten**

Welche Erzählungen legen öffentliche Datensätze nahe? Die Frage ist wichtig, da amtliche Daten unser aller Daten sind und auch eine gewisse Autorität besitzen. Die Frage ergab sich zu Anfang der Forschung einer Studierendengruppe im Jahr 2021. Sie interessierte sich für die Möglichkeit von Geflüchteten, an Informationen auf Webseiten der Stadt Bochum zu gelangen. Der Hintergrund der Frage war die Erfahrung einer Studierenden durch ihre Arbeit mit Geflüchteten, dass öffentliche Webseiten für diese Menschen eine Herausforderung darstellen. Es sollte sich später zeigen, dass es viel mehr als sprachliche Herausforderungen sind. Die Sprache war jedoch der erste Punkt, der die Aufmerksamkeit der Gruppe auf sich zog. Die Studierenden suchten nach Daten über die Sprachkompetenzen von Geflüchteten in Bochum, welche es jedoch nicht zu geben schien. Stattdessen fanden sie Daten im Informationsportal der Stadt zur Staatsangehörigkeit ihrer Bürger\*innen.<sup>19</sup> Obwohl viele – aber nicht alle – Nationalstaaten sich um Einsprachigkeit bemühen, gibt es keine zwingende Übereinstimmung zwischen einer Staatsangehörigkeit und einer Sprachkompetenz. Auch weil viele Menschen Fremdsprachen wie Englisch, Französisch oder Deutsch beherrschen, sagt die Staatsangehörigkeit nur bedingt etwas über die Sprachkenntnisse von Menschen aus.

Daten zu Staatsangehörigkeit sprechen eine rechtliche Sprache. Sie befähigen uns, Erzählungen über Bleibeperspektiven zu erstellen, über Abschiebung und über Flüchtlingsstatus. Solche Erzählungen sind nicht nur isolierte Endprodukte einer Erfahrung

<sup>19</sup> <https://bostatis.bochum.de/#app/mainpage/Bev%C3%B6lkerung/> (Zugriff: 1.7.2024).

oder eines Datensatzes, sondern befähigen zudem, bestimmte Beziehungen zu knüpfen. Haraway<sup>20</sup> führt den Begriff der *Responsabilität* ein, um die Aufmerksamkeit für Beziehungen, Infrastrukturen und Konsequenzen zu schärfen, die Akteur\*innen befähigen, auf eine Anfrage, Aufforderung, Einladung, Angebot usw. zu reagieren oder zu antworten. Staatsangehörigkeit, Bleibeperspektiven, Abschiebung und Flüchtlingsstatus gehören zum Diskurs des Rechts und machen ihre Leser\*innen *responsabel* auf rechtliche, staatliche oder behördliche Themen. Sie ermöglichen es ihren Leser\*innen weniger, darüber zu reden, wie Menschen mit verschiedenen Sprachen befähigt werden, miteinander zu kommunizieren. Damit war die Frage der Studierenden geweckt, ob die fehlenden Daten zu Sprachkenntnissen der Bewohner\*innen Bochums ihr Gegenstück in einer fehlenden Wissensinfrastruktur<sup>21</sup> finden würden, die es den Behörden erleichtert hätte, über die Sprachkompetenzen nachzudenken und damit darüber, wie sie besser mit Menschen mit verschiedenen Sprachkompetenzen kommunizieren können. Die Studierenden zogen in die Stadt, um der (fehlenden) Wissensinfrastruktur zu den Sprachkenntnissen der Bewohner\*innen nachzugehen und entdeckten, dass alle öffentlichen Hinweisschilder der Stadt – auch diejenigen, die den Weg zur Ausländerbehörde zeigten – ausschließlich auf Deutsch waren. Auch wenn diese Hinweise nicht ausreichten, um die Verfasstheit der Wissensinfrastruktur der Stadt Bochum im Detail zu bestimmen, motivierten sie die Studierenden, die *Responsabilität* der Behörden bezüglich der Sprachkompetenzen von Geflüchteten zu fördern. Sie erstellten eine Anleitung dafür, wie Behörden ihre Webseiten für Personen anderer Sprachen als Deutsch zugänglich machen können. Sie trugen damit zu Komponenten einer Wissensinfrastruktur bei, die die Behörde gegenüber Bewohner\*innen als Menschen mit verschiedenen Sprachkompetenzen *responsabel* machen können.<sup>22</sup>

- 20 Vgl. Donna J. Haraway: *Unruhig bleiben: Die Verwandtschaft der Arten im Chthuluzän*. Frankfurt a. M. 2018.
- 21 Vgl. Emma Cardwell, Solveig Joks, John Law: *Colonialism and the Infrastructures of Knowing*. In: Alan Irwin, Ulrike Feld (Hg.): *The Encyclopedia of Science & Technology Studies*. Cheltenham 2022.
- 22 Die Ergebnisse des Projekts können unter [http://datastories.rub.de/sprache\\_datastory/](http://datastories.rub.de/sprache_datastory/) eingesehen werden (Zugriff: 1.7.2024).

Dieser Blick auf Daten betrachtet sie nicht nur als Repräsentationen einer mehr oder weniger distanzierten Wirklichkeit, sondern als Komponenten von Wissensinfrastrukturen, die spezifische Erzählungen und spezifische Verantwortlichkeiten erleichtern und andere erschweren.

### Kollektive Erzählungen durch Daten

Zusammen mit Laura Kocksch führten wir 2019 einen Datasprint in unserem interdisziplinären Graduiertenkolleg SecHuman durch. Die Beteiligten waren Informatiker\*innen sowie geistes- und sozialwissenschaftliche Forscher\*innen des Graduiertenkollegs. Abbildung 2 zeigt eine Netzwerkgrafik mit den Publikationen der Beteiligten zu Cybersicherheit.<sup>23</sup> In der Grafik sind Publikationen verknüpft, in denen dieselben semantischen Einheiten auftreten. Sie zeigt nicht überraschend, dass einige Publikationen in dem sehr heterogenen Graduiertenkollegium weit auseinanderliegen, wobei andere enger semantisch verknüpft sind.

Die farbigen Cluster des Netzwerks sind jedoch andere, als die Beteiligten – inklusive wir selbst – vermutet hatten. Die Publikationen eines Kollegen aus der Kryptografie, der generell als eine zentrale Person mit vielfältigen Beziehungen im Forschungsverbund angesehen wurde, landeten in einer abgesonderten (orangenen) Provinz auf der Karte. Dies bedeutet nicht, dass die Grafik als Beleg für ein Beziehungsgeflecht gelesen werden kann, das unabhängig von den Erfahrungen der Alltagspraktiken im Graduiertenkolleg existiert. Vielmehr nutzten wir die überraschende Visualisierung unserer Publikationsverwandtschaften als Anlass dazu, in einem Datasprint mit den Mitgliedern des Graduiertenkollegs eine neue Erzählung der interdisziplinären Beziehungen im Kolleg kollektiv zu erstellen. Wesentlich war nicht nur, was die Datenvisualisierungen zeigten, sondern auch, dass sie eine andere – visuelle – Sprache verwendeten als das Vokabular unserer wissenschaftlichen Disziplinen, das wir

23 Vgl. Laura A. Kocksch, Estrid Sørensen: Towards a Typology of Interdisciplinarity in Cybersecurity: Trade, Choice, and Agnostic-Antagonist. In: Proceedings of the 2023 New Security Paradigms Workshop (NSPW '23). New York 2023, S. 116–129.



Abb. 2: Netzwerkgrafik von Publikationen des Graduiertenkollegs

üblicherweise für den Austausch nutzen. Diese verfremdende visuelle Sprache stellte alle Teilnehmenden, die in ihrem wissenschaftlichen Alltag hochspezialisierte Expert\*innen sind, vor ein und dieselbe Herausforderung, Worte zu finden, um über die Verwandtschaft, die die Netzwerkgrafik darstellt, zu erzählen. Die Verfremdung der Datenvisualisierung hat damit einen kollektivierenden Effekt auf die Erzählung und auf das Knüpfen von Beziehungen.

In einem weiteren Datasprint im Rahmen eines Studienprojekts luden unsere Studierenden Bochumer Klima-Akteur\*innen aus der Stadtverwaltung, von Scientists for Future sowie von Ende Gelände ein. Auch hier dienten Daten als Anlass für kollektive Erzählungen und für das Knüpfen neuer Beziehungen. Die Herausforderung in Bochum war, dass es sehr viele Akteur\*innen gab, die sich für Klimaschutz einsetzten, wobei es zwischen ihnen wenig Koordination und Kooperation gab. Unsere Überlegung war, ob man durch

Daten eine kollektive Erzählung anstoßen könnte. Dabei gingen wir einen Schritt weiter als im Sprint zu Cybersicherheit und erstellten einen spekulativen Datensatz. Auf Grundlage von Informationen der Webseiten von 33 Klima-Gruppierungen der Stadt erstellten wir einen Datensatz, der ihre jeweiligen Themen listete und quantitativ schätzte, wie sehr die Gruppierungen sich für die verschiedenen Themen einsetzen. Die Selbstdarstellungen der Webseiten reichten eindeutig nicht dafür aus, die Gruppierungen realistisch zu repräsentieren. Der Datensatz und seine Visualisierungen waren entsprechend spekulativ in dem Sinn, dass sie nicht darauf zielten, die Realität der Gruppierungen darzustellen, sondern als Material fungieren sollten, das die Teilnehmenden verändern und überarbeiten konnten, um dabei eine gemeinsame Geschichte über die Klimaaktivitäten in der Stadt zu erzählen.

Üblicherweise nehmen Datendesigner\*innen an Datasprints teil und ordnen die Daten für die Teilnehmenden an, sodass schöne und oft beeindruckende Visualisierungen erstellt werden. Wir hatten einerseits keine professionellen Datendesigner\*innen im Team, verzichteten jedoch andererseits auch gern darauf, da es genau die gemeinsame Arbeit an den Daten und an den Datenvisualisierungen war, die den Anlass für kollektives Erzählen generieren sollte – darüber wie die Daten am besten visualisiert werden konnten, welche Werte von Relevanz waren, wer sie vertrat usw. Im Anschluss an Isabelle Stengers<sup>24</sup> können wir solche Daten nicht nur spekulativ nennen, sondern sogar auch „idiotische Daten“. Stengers hat den „Idiot“ als eine konzeptuelle Figur eingeführt, die „sich der Einmütigkeit widersetzt, in der sich eine Situation präsentiert, und dadurch das Denken oder Handeln mobilisiert“<sup>25</sup>. Der Idiot kann jedoch nicht erklären, was vor sich geht, denn er „kann weder antworten noch das Thema diskutieren [...] [er] weiß nichts [...] er verlangt die Entschleunigung“<sup>26</sup>. Statt als wissenschaftliche Expert\*innen oder als professionelle Datendesigner\*innen aufzutreten, die eindeutige Daten und

24 Vgl. Isabelle Stengers: *The Cosmopolitical Proposal*. In: Bruno Latour, Peter Weibel (Hg.): *Making Things Public*. Cambridge 2005, S. 994–1003.

25 Ebd., S. 994, eigene Übersetzung aus dem Englischen.

26 Ebd.

Evidenzen vorlegen und damit die Deutungshoheit für sich behalten, legten wir – Studierende und Dozent\*innen – idiotische Daten vor, um das Denken und das Erzählen der Teilnehmenden anzuregen.

Als Akteure in Kooperationsprozessen haben spekulative und idiotische Daten das Potenzial, neue und kollektive Erzählungen anzuregen, die potenziell neue Beziehungen knüpfen können.

## **Die planetare Interdependenz von Daten in Erzählungen miteinbeziehen**

Bewegen wir uns etwas von den Daten selbst weg und betrachten die Infrastruktur der Daten. Die Digitalisierung von Daten – sowohl Daten, die wir in Anlehnung an Rogers<sup>27</sup> digital geborene Daten nennen können, wie auch analoge Daten, die später digitalisiert werden – erfordert eine umfassende Infrastruktur der Speicherung sowie der Prozessierung von Daten. Auch wenn Begriffe wie „Cloud“ sowie das „Virtuelle“ Erzählungen über Daten als immateriell voranbringen, ist es mittlerweile allgemein bekannt, dass Daten umfassende Mengen an Strom verbrauchen sowie auch Rohstoffe, die in großem Umfang von der Oberfläche des Planeten abgebaut werden.<sup>28</sup> Weil digitale Daten das Teilen und die Wiederverwendung von Forschungsdaten ermöglichen und die EU mit Nachdruck eine Erzählung davon vorantreibt<sup>29</sup>, werden Dateninfrastrukturen weiter und weiter ausgebaut. Bei Berufungsverhandlungen werden zunehmend Dateninfrastrukturen gefordert, Institute richten Serverräume und HPC-Cluster (Hochleistungsrechner) ein und Universitäten bauen eigene größere, leistungsfähigere Data Centers – auch in Universitätsverbänden –, welche als Bausteine für Erzählungen über die Wettbewerbsfähigkeit der Universitäten dienen. Diese Einsichten stammen aus meiner aktuellen Forschung zu universitären Data Centers, die im Rahmen des SFB 1567 „Virtuelle Lebenswelten“<sup>30</sup> durchgeführt wird und

27 Vgl. Richard Rogers: *Digital Methods*. Cambridge 2015.

28 Vgl. Kate Crawford: *An Atlas of AI*. New Haven 2021; Tung-Hui Hu: *A Pre-History of the Cloud*. Cambridge 2016.

29 Klaus Hoeyer, Sara Green, Andrea Martani u. a.: Health in data space. Formative and Experiential Dimensions of Cross-Border Health Data Sharing. In: *Big Data & Society* 11 (1), 2024.

30 <https://www.virtuelle-lebenswelten.de> (Zugriff: 1.7.2024).

zudem zeigt, dass sowohl die technischen Infrastrukturen deutscher Universitäten, ihre Beschaffungsinfrastrukturen, die Arbeitsteilung zwischen wissenschaftlichem und technischem Personal, die Wissenschaftsfreiheit und nicht zuletzt der Wettbewerb in der Wissenschaft dafür sorgen, dass die planetare Interdependenz von Forschungsdateninfrastrukturen weitgehend unbeachtet in Erzählungen über universitäre und wissenschaftliche Dateninfrastrukturen bleibt.<sup>31</sup> Dateninfrastruktur wird an Universitäten vor allem im Bereich der Lebens-, Technik- und Naturwissenschaften ausgebaut und Erzählungen über Dateninfrastrukturen sind hier besonders konsolidiert. Dateninfrastrukturen der empirischen Kulturwissenschaft sind weit weniger ausgebaut, zeigen entsprechend eine deutlich bessere Klimabilanz und sind (noch) weniger pfadabhängig, was die Entwicklung von Dateninfrastruktur angeht. Sie besitzt damit noch eine Flexibilität, die es ihr eher als anderen Wissenschaften ermöglicht, die planetare Interdependenz von Dateninfrastruktur in ihre Erzählungen von ihren wissenschaftlichen Dateninfrastrukturen einzubauen und diese auch beim Ausbau dieser zu beachten.

### **Erzählende Daten in der empirischen Kulturwissenschaft**

In den Technik-, Natur- und Lebenswissenschaften sowie in den Computational Social Sciences und Digital Humanities werden digitale Daten überwiegend als Grundlagen für wissenschaftliche Evidenz verstanden.<sup>32</sup> Ich habe anhand mehrerer Beispiele argumentiert, dass Daten nicht nur repräsentieren, sondern auch Erzählungen erzeugen. Als Grundlage für Erzählungen bieten digitale Daten neues Material

31 Vgl. Leman Çelik, Estrid Sørensen: Scientific engagements with data infrastructures and their planetary matter. In: Joachim Baur, Claus-Marco Dieterich, Michaela Haibl, Gudrun König, Jens Stöcker, Markus Tauschek, Jan Watzlawik (Hg.): *Analysen des Alltags. Komplexität, Konjunktur, Krise. Tagungsband des 44. DGEKW-Kongress in Dortmund* (im Erscheinen); Estrid Sørensen, Stefan Laser: *Towards Artful Sustainable Integration of IT Infrastructures*. In: Patricia Jankowski, Anja Höfner, Marja L. Hoffmann u. a. (Hg.): *Shaping Digital Transformation for a Sustainable Society. Contributions from Bits & Bäume*. Berlin 2023, S. 87–90.

32 Vgl. Lina Franken: *Digitale Methoden für qualitative Forschung. Computationelle Daten und Verfahren*. Münster 2023.

für die empirischen Kulturwissenschaften, das sowohl dazu dienen kann, neue Geschichten zu erzählen, neue Beziehungen zu knüpfen und Wissensinfrastrukturen zu formen sowie kollektive Erzählungen zu generieren. Zudem kann – und sollte – die empirische Kulturwissenschaft die planetare Interdependenz von Dateninfrastrukturen in ihre Erzählungen über Daten integrieren. Weil die Verwendung von digitalen Daten noch relativ jung in der empirischen Kulturwissenschaft ist und im Fachbereich kaum Pfadabhängigkeit erfährt, bestehen weite Gestaltungsmöglichkeiten darüber, wie Erzählungen mit ihnen geformt werden.

---

*Whether in quantitative, semantic, photographic or other forms, digital data are usually treated as representations of worlds that are separate from the data. These are distant data. The chapter argues that data also tell stories. Story-telling data are engaged data that lend themselves to ethnographic research. Data can not only represent and tell stories in different ways with distant and engaged data; they also enable different social and material relations and impede others. The choice between distant and engaged data is thus also a question of worlding.*