

Visuelle Datenanalysen, digitale Tools und interdisziplinäre Forschung

Praxeologische Reflexionen

Sarah-Mai Dang

Einleitung

Datenbasierte Forschung eröffnet Film- und Medienwissenschaftler:innen eine Vielzahl an Möglichkeiten, filmhistorische Zusammenhänge zu analysieren, zu reflektieren und zu präsentieren. Sie erfordert aber auch eine große Bandbreite unterschiedlicher Expertisen, die weder eine einzelne Person, noch eine einzelne Disziplin vorweisen kann. Dementsprechend ist eine kollaborative und interdisziplinäre Arbeitsweise unabdinglich. Diese gestaltet sich je nach Konzeption, Methode und Beteiligten des Projekts unterschiedlich und geht mit je spezifischen Herausforderungen einher.

Um das gesamte Spektrum datenbasierter Forschung – „von der Fragestellung über die Informationssuche, Analyse, Interpretation bis zur Repräsentation und Darstellung neuen historischen Wissens“ (Fickers 2020: 158) – und die damit verbundenen Anforderungen zu begreifen, hat der Historiker Andreas Fickers das Konzept der „digitalen Hermeneutik“ entwickelt (ebd.). Dieses zielt darauf ab, die Arbeitsprozesse in ihren multiplen Dimensionen zu reflektieren – samt der infrastrukturellen, materiellen und technologischen Kontexte, in die sie eingebettet sind. Neben der Kritik digitaler bzw. digitalisierter Quellen „oder besser: geschichtswissenschaftliche Datenkritik“ gehört dazu die Analyse der eingesetzten Tools, der zugrundeliegenden Algorithmen und der Repräsentationspolitiken, die in Interfaces zum Ausdruck kommen (ebd: 158f.).

In seinem Plädoyer für ein „Update“ der klassischen Hermeneutik in der Geschichtswissenschaft“ unterstreicht Fickers die neuen Herausforderungen, vor die uns die Digitalisierung bzw. Datafizierung stellt. Angesichts der damit verbundenen Vervielfältigung der Methoden – der Mischung aus analogen und digitalen, qualitativen und quantitativen, Close und Distant Reading- bzw. Viewing-Ansätzen – sind wir, wie er betont, mit einer „methodologische[n] und epistemologische[n] Doppelbelastung“ konfrontiert. Neben etablierten Verfahren, Konzepten und Be-

griffen gilt es nun nämlich zusätzlich Techniken zu erlernen, die bislang fachfremd waren. Wir befinden uns derzeit in einer Übergangsphase, die durch ein transformatives Spannungsverhältnis zwischen tradierten und neuen wissenschaftlichen Begriffen, Konzepten und Methoden gekennzeichnet ist (ebd.).

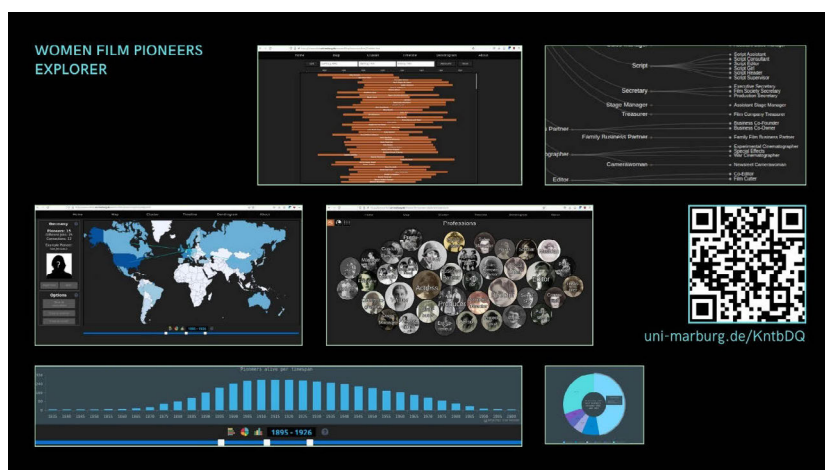
Fickers Feststellungen treffen nicht nur für die Geschichtswissenschaft, sondern auch für andere Disziplinen zu. Die sich durch den zunehmenden Einsatz computationaler Verfahren verändernden Forschungspraktiken, -gegenstände und -ansätze und die damit verbundenen neuen, teils disruptiven Formen der Wissensproduktion werden seit geraumer Zeit fächerübergreifend diskutiert, insbesondere im Kontext der Digital Humanities (z. B. Berry 2012; Schreibman et al. 2004, 2016). Dabei geht es nicht nur um die kritische Analyse technologischer, methodologischer und epistemologischer Umbrüche, sondern auch um die infrastrukturellen, institutionellen, politischen und ökonomischen Aspekte aktueller Entwicklungen – nicht zuletzt auch innerhalb der Universität (z. B. Berry 2022; Oiva und Pawlicka-Deger 2020). Auch für die Film- und Medienwissenschaft gilt es, die derzeitigen Transformationen intensiv zu beobachten, sich den neuen Herausforderungen gewahr zu werden und auch die eigenen Vorgehensweisen offen und selbstkritisch zu reflektieren.

In diesem Sinne werde ich im Folgenden die spezifischen Anforderungen, Arbeitsweisen und Erkenntnisprozesse reflektieren, mit denen ich im Rahmen meiner Forschung zu filmhistorischen Metadatensammlungen und Datenvisualisierungen konfrontiert worden bin. Anhand exemplarischer Fallstudien aus meinem Teilprojekt der von mir geleiteten Forschungsgruppe „Ästhetiken des Zugangs. Datenvisualisierungen in der digitalen Filmgeschichtsschreibung am Beispiel der Forschung zu Frauen*¹ im frühen Kino“ (DAVIF) (2021–2026)² werde ich die Potenziale und Grenzen digitaler Tools und datenbasierter Ansätze im Kontext der Film- und Medienwissenschaft erörtern. Mein Interesse gilt hier – wie auch in der Forschungsgruppe – methodologischen, toolkritischen und praktischen Aspekten datenbasierter Forschung in der Film- und Medienwissenschaft und weniger historischen Erkenntnissen.

-
- 1 Wenn ich von „Frauen“ oder „Männern“ schreibe, beziehe ich mich auf die soziale, kulturelle und politische Konstruktion von Geschlecht als binäre Kategorie, die den untersuchten Datensammlungen zugrunde liegt. Ich hingegen begreife Geschlecht nicht als essentialistische, objektive Kategorie, sondern wie Avtar Brah als eine von vielen Differenzkategorisierungen, deren Bedeutung und Funktion abhängig vom jeweiligen Kontext variieren (Brah 1996).
 - 2 Die Forschungsgruppe wird durch das Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR) im Rahmen der Richtlinie zur Förderung von Forschungs- und Entwicklungsvorhaben zur theoretischen, methodischen und technischen Weiterentwicklung der digitalen Geisteswissenschaften gefördert. Für weitere Informationen zur Forschungsgruppe s. <https://uni-marburg.de/Q8500>.

Der Women Film Pioneers Explorer – Zwischenfazit eines Forschungspraxisseminars

Abb. 1: Die Visualisierungen des Women Film Pioneers Explorer (<https://uni-marburg.de/KntbDQ>) – eine Timeline, ein Dendrogramm, eine Karte, ein Cluster, ein Histogramm und ein Kreisdiagramm (im Uhrzeigersinn) – erlauben es, die Daten des WFPP in einer chronologischen, hierarchischen, geographischen und relationalen Weise zu betrachten.



Der Women Film Pioneers Explorer (WFPE o. J.) war die erste Fallstudie meines Projekts (Dickel et al. 2021). Im Rahmen eines gemeinsam mit dem Informatiker Thorsten Thormählen angeleiteten Forschungspraxisseminars, das im Zuge der Kooperation der DAVIF-Forschungsgruppe mit der Arbeitsgruppe Grafik und Multimedia Programmierung der Philipps-Universität Marburg durchgeführt wurde, erprobten fünf Masterstudierende der Informatik klassische Visualisierungsformen auf der Basis von Daten des Women Film Pioneers Project (WFPP o. J.). Das WFPP wurde in den 1990er Jahren zunächst als Buchprojekt initiiert und 2013 an der Columbia University als kollaborative Online-Plattform gelauncht (Gaines et al. 2013). Das Projekt umfasst über 300 Profile zu Pionierinnen der Stummfilmzeit, die filmographische und biographische Angaben enthalten. Mit seiner umfangreichen Datensammlung zielt das WFPP darauf ab, eine Geschichtsschreibung jenseits eines (männlichen) Masternarrativs zu ermöglichen. Für die feministische Forschung zum frühen Kino stellt die Website eine zentrale Anlaufstelle dar. Auch darüber hinaus ist sie von großem historiographischem Wert, da sie neben den zahlreichen Profilen auch Informationen zu Archiven und weiteren Ressourcen vorhält (Dang 2020).

Für das Forschungspraxisseminar wurde ein eigens für die DAVIF-Forschungsgruppe angefragter Datensatz verwendet. Er umfasst Information zu allen Pionierinnen, die zum Zeitpunkt der Abfrage im Jahr 2020 im WFPP erfasst waren: ID, Name, Alias, Permalink des einzelnen Profils, Link des Porträtfotos, Berufsbezeichnungen und Hierarchisierung der Berufe, Arbeits-orte bzw. Länder sowie Geburts- und Sterbedatum (Gaines und Columbia University Libraries 2020). Ziel des Forschungspraxisseminars war es, das Erkenntnispotenzial der Datensammlung des WFPP auszuloten und sowohl die weltweiten Einflüsse von Frauen in der Stummfilmzeit und darüber hinaus sichtbarer zu machen als auch den abgefragten Datensatz zu analysieren. Wie ich bereits an anderer Stelle ausführte, wurden durch die Visualisierungen etwa die globalen Verbindungen der Pionierinnen deutlich und die Vielzahl an Berufen, die sie innehatten. So zeigt sich etwa, dass das WFPP Daten zu 15 Frauen vorhält, die zwischen 1895 und 1926 u. a. in Deutschland gearbeitet haben. Gleichzeitig offenbarten sich Lücken und blinde Flecken der Metadatenammlung (Dang 2023a). Nicht zuletzt führten die Visualisierungen dazu, sich mit den Datenquellen und lokalen Datenpraktiken des WFPP auseinanderzusetzen (Dang 2024).

Das im Rahmen der Informatik angesiedelte Forschungspraxisseminar ist interdisziplinär angelegt. Eine Reihe fachfremder Auftraggeber:innen stellen jedes Jahr ihre Anliegen in einem Pitch vor. Die Studierenden entscheiden sich auf dieser Grundlage für ein Projekt, das sie im Laufe von zwei Semestern umsetzen möchten. Im Jahr 2020 nahm ich mit dem WFPP-Datensatz an diesem Pitch teil, fünf Studierende entschieden sich für eine Projektarbeit mit diesen Daten. In monatlichen Treffen mit den Studierenden wurden der Auftrag näher erörtert, die Visualisierungen entwickelt und angepasst. Die Studierenden stellten ihre Fortschritte vor und erhielten Feedback. In die detaillierte technische Umsetzung war ich als Auftraggeberin nicht eingebunden. In den Treffen wurden vornehmlich konzeptionelle und ästhetische Fragen besprochen. Die Studierenden dokumentierten den Arbeitsprozess in groben Zügen auf GitHub (GitHub o. J.), einzelne Schritte wie etwa die Modellierung der Daten wurden dabei nicht festgehalten. Erst im Abschlussgespräch traten die technischen Herausforderungen und implizite Prämissen zu Tage. Zwar wurden kontinuierliche Updates durchgeführt, doch es entwickelte sich kein gemeinsamer Erkenntnisprozess.

Ein kollektiver Raum der Zusammenarbeit im Sinne der von Andreas Fickers, Juliane Tatarinov und Tim van der Heijden beschriebenen „trading zone“ (Fickers et al. 2022: 3) wurde nicht geschaffen. Ein intellektueller und sozialer Austausch zwischen verschiedenen Disziplinen und „communities of practice“ (Lave und Wenger 1991), wie ihn Fickers, Tatarinov und van der Heijden propagieren (Fickers et al. 2022: 2), fand insofern nicht statt. Dies ist zum einen auf die universitären Rahmenbedingungen zurückzuführen, die Studierende dazu anhalten, Seminararbeiten möglichst zügig und effizient abzuschließen. Zum anderen ist dies sicherlich auch

auf die grundsätzlichen Differenzen zwischen Geisteswissenschaftler:innen und Informatiker:innen zurückzuführen, auf ihre unterschiedlichen methodischen und epistemischen Traditionen (ebd.: 4). Während ich mir von der Zusammenarbeit versprach, die vielschichtige Bedeutung der Daten mit Hilfe von Visualisierungen zu verstehen, waren die Studierenden an einer pragmatischen Lösung interessiert, die eine sinnvolle Nutzung der Daten ermöglicht. Verschiedene „communities of practice“ mögen zwar dieselben Tools, Daten und Infrastrukturen nutzen, doch unter völlig verschiedenen Voraussetzungen und mit anderen Bedürfnissen, Fragen und Zielen (ebd.: 5). Daher gilt es prinzipiell bei interdisziplinären Kollaborationen – innerhalb oder außerhalb der Geisteswissenschaften – die Bedingungen der Zusammenarbeit auszuhandeln, wie auch die Medienwissenschaftlerin Eef Masson vor dem Hintergrund ihrer eigenen Erfahrung konstatiert (Masson 2017: 30f.). Um die unterschiedlichen Positionen in einen produktiven Dialog zu bringen, ist viel an Kommunikation und Offenheit notwendig (ebd.). Dafür wird im akademischen Arbeitsalltag allerdings in der Regel zu wenig Zeit eingeräumt.

Für zukünftige Praxisseminare erscheint es sinnvoll, den Fokus stärker auf gemeinsames Forschen und Lernen zu legen, anstatt ein unternehmerisch geprägtes Auftraggeber:in/Auftragnehmer:in-Verhältnis zu reproduzieren. Letzteres, der Idee einer kollegialen Universität eigentlich zuwider laufende Arbeitsweise ist in vielen DH-Projekten, -Labs oder -Zentren stark verbreitet und basiert häufig auf ausgeprägten hierarchischen Arbeitsverhältnissen zwischen Management/Leitung, wissenschaftlichem, technischem und administrativem Personal (Berry 2022: 126f.). Um die kritische Reflexion in Forschungspraxisseminaren zu fördern, sollte die Aufgabenstellung idealerweise im Vorfeld gemeinsam definiert werden, sodass die Interessen, Kenntnisse und Erfahrungen aller Beteiligten stärker berücksichtigt werden. Zudem wäre es produktiv, Studierende aus allen beteiligten Disziplinen einzubeziehen, um den interdisziplinären Austausch zu erleichtern. Nicht zuletzt sollte nicht nur der Stand der Zwischenergebnisse, sondern auch der fortlaufende Arbeitsprozess dokumentiert und reflektiert werden – nicht nur, um Missverständnissen vorzubeugen, sondern auch, um Erkenntnisse im Entstehungsprozess als solche identifizieren zu können. Was für die eine Person wie ein banales technisches Problem erscheinen mag, kann für eine andere von hohem epistemischen Interesse sein – und umgekehrt.

Zwar hat sich die Möglichkeit eines tiefergehenden Austauschs im Rahmen des skizzierten Forschungspraxisseminars nicht ergeben bzw. wurde nicht als solche genutzt, doch das Ergebnis, der Women Film Pioneers Explorer, dient nach wie vor als höchst fruchtbare Ausgangslage, um die Daten des WFPP zu erforschen und filmhistorische Analysen durchzuführen (Dang 2023a). Um die praxeologische Dimension von Datenvisualisierungen intensiver zu untersuchen bzw. um das erkenntnistheoretische Potential, das sich durch die konkrete Umsetzung während des Arbeitsprozesses ergibt, stärker auszuschöpfen und systematischer festzuhal-

ten, habe ich im Anschluss an den Women Film Pioneers Explorer selbst eine Reihe digitaler Tools erprobt.

Explorationen mit OpenRefine, Datawrapper und RawGraphs

Inzwischen gibt es zahlreiche Übersichten zu Visualisierungstools, deren Funktionsweisen und Einsatzbereiche, sowie eine Unmenge an frei zugänglichen Tutorials hierzu im Internet. Empfehlenswert ist beispielsweise The Data Visualisation Catalogue, der einen sehr guten Überblick zu Datenvisualisierungen, ihre Funktionen und die entsprechenden Softwareprogramme gibt (Rebecca, o. J. c). Diese Fülle an Informationen erleichtert einerseits den Zugang zu digitalen Anwendungen. Andererseits erschwert die Bandbreite an Möglichkeiten die Auswahl und Einordnung der Tools – insbesondere, da es bislang nur wenige dezidiert geisteswissenschaftliche, geschweige denn film- oder medienwissenschaftliche Übersichten gibt. Dieser Mangel an fachspezifischen Angeboten liegt auch daran, dass die Tools in der Regel in fachfremden Kontexten mit unterschiedlichen epistemischen Traditionen entwickelt wurden, etwa der Medizin, Geowissenschaften oder Wirtschaftswissenschaften, und deren Funktionen daher nur bedingt übertragbar sind (Masson 2017: 29). Eine nennenswerte Ausnahme stellt das kollaborative Projekt „The Programming Historian“ dar, dessen speziell für Geisteswissenschaftler:innen konzipierte Tutorials ein breites Spektrum an digitalen Tools, Techniken und Arbeitsläufen erläutern, die sowohl in der Forschung als auch in der Lehre Anwendung finden können (z. B. Düring 2015). Neben den allgemeinen Überblicksseiten hat sich dieses Projekt als besonders hilfreich erwiesen.

Ziel der folgenden Fallstudien war es, wie bereits zuvor beim Women Film Pioneers Explorer, filmhistorische Metadaten zu Frauen und ihren Tätigkeiten des WFPP zu visualisieren, um die feministische Forschung sichtbarer zu machen. Nach umfangreicher Recherche, die eher dem aus der Literaturrecherche bekannten Schneeballsystem entsprach und mit entsprechenden Vor- und Nachteilen verbunden ist, und ersten Tests habe ich mich gegen Tools wie unter anderem Excel, Gephi, Tableau, QGIS, ArcGIS oder Carto und die Anwendung von Programmiersprachen wie R oder Python entschieden. Die Gründe, die gegen die Nutzung sprachen, waren vielfältig: Beispielsweise ist das Erlernen von R oder Python mit einem zeitlichen Aufwand verbunden, der sich für meine Zwecke als unverhältnismäßig hoch herausstellte. Dasselbe gilt für die Software QGIS und ArcGIS, die der Analyse und Visualisierung geographischer Informationen dienen. Bei der letzteren Option handelt es sich um ein kommerzielles Programm, für das Lizenzgebühren erhoben werden, während die erste eine kostenlose Open-Source-Alternative darstellt, die zudem intuitiver zu handhaben ist. Wie auch die cloudbasierte Software Carto schienen mir QGIS und ArcGIS jedoch zu komplex für eine

erste Annäherung an digitale Tools zur Datenvisualisierung. Obwohl alle drei Tools keine Programmierkenntnisse erfordern, sind fundierte Kenntnisse von Geodaten und geographischen Informationssystemen (GIS) notwendig, um aussagekräftige Visualisierungen zu erstellen. Auch die Nutzer:innen der Software Gephi, die im akademischen Kontext ihren Ursprung hat und als eines der bekanntesten Tools für Netzwerkanalysen intensiv in den Digital Humanities genutzt wird, erwartet eine steile Lernkurve. Tableau von Salesforce wiederum, das ebenfalls relativ weit verbreitet ist und eine große Palette an Diagrammen, Grafiken und Karten ermöglicht, ist nicht mit dem von mir verwendeten Linux-Betriebssystem kompatibel. Excel bzw. LibreOffice Calc, die aufgrund ihrer vielseitigen Tabellenfunktionen häufig für grundlegende Datenanalysen und -visualisierungen verwendet werden, gelten hingegen als vergleichsweise wenig ästhetisch. Dies könnte ein Grund dafür sein, weshalb diese traditionelle Software im Kontext der Digital Humanities bislang nur geringe Beachtung findet.

Dies sind nur einige der zahlreichen Tools, mit denen ich in den vergangenen Jahren experimentiert habe. Wegen des kostenfreien Zugangs und der niederschweligen Anwendung habe ich mich schließlich für RawGraphs und Datawrapperr entschieden, um Visualisierungen zu erstellen. Zur Aufbereitung der Daten habe ich OpenRefine benutzt.

OpenRefine – Datenmodellierung und Anreicherung

OpenRefine (OpenRefine o. J.) ist ein kostenfreies Open-Source-Programm, um große Mengen an Daten zu bearbeiten. Mit seinem browserbasierten Interface und seinen zahlreichen Funktionen arbeitet es effizient und transparent und ist relativ leicht zu erlernen. Zudem ist es mit allen Betriebssystemen, auch mit Linux, kompatibel. Nicht zuletzt erlaubt OpenRefine es, die ausgewählten Datensätze mit bereits existierenden Datenbanken wie Wikidata zu verknüpfen, die eigenen Daten zu prüfen, anzureichern und damit die Datenqualität erheblich zu verbessern.

Um die Potentiale und Grenzen eines Tools zu verstehen, ist es grundsätzlich sinnvoll, zunächst mit einem überschaubaren und bekannten Datensatz zu arbeiten, dessen Kontext einem gut vertraut ist. Mit Hilfe von OpenRefine habe ich daher in einem ersten Schritt die durch das WFPP erfassten Daten zu 16 Frauen,³ die zwischen 1895 und 1925 u.a. in Deutschland gearbeitet haben, mit den Geburts- und Sterbeorten aus Wikidata angereichert, um die globalen Migrationsbewegungen

3 Aus irgendeinem Grund wurde eine Pionierin von den Studierenden im Forschungspraxisseminar übersehen und nicht in den Explorer aufgenommen. Es handelt sich um insgesamt 16 Pionierinnen, die u.a. in Deutschland gearbeitet haben und im WFPP zum Zeitpunkt der Visualisierungen erfasst wurden.

der Pionierinnen zu analysieren und sichtbar zu machen. Diese Informationen fehlten bislang in dem *Women Film Pioneers Explorer* bzw. in dessen Datengrundlage. Neben den verschiedenen Arbeitsorten wurde durch die Anreicherung ersichtlich, woher die Pionierinnen kamen und wohin das Leben sie zuletzt geführt hat. Für das Erlernen des Programms war besonders hilfreich ein Webinar von Kristina Rose, Datenkoordinatorin im Digitalen Sammlungsmanagement des DFF – Deutsches Filminstitut & Filmmuseum (Rose 2023), sowie ein Tutorial für ein Einführungsseminar in die Digital Humanities der Medienwissenschaftlerin Miriam Posner (Posner 2017a). Darüber hinaus bietet OpenRefine selbst ein sehr gutes Manual zu allen Funktionen des Programms.

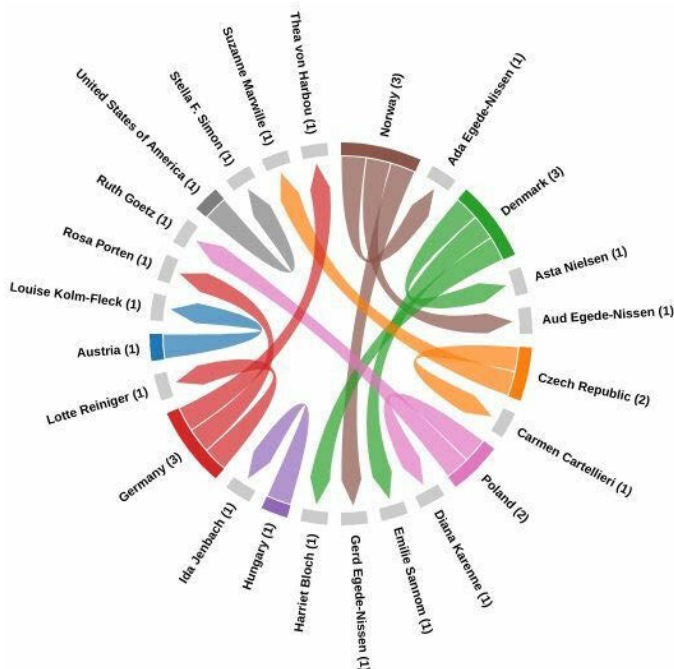
Neben den praktischen Aspekten ist der Arbeitsprozess, in den man sich mit OpenRefine begibt, bemerkenswert. Denn anders als manche Kolleg:innen mit digitalen Tools assoziieren mögen – vor allem diejenigen, die noch nicht mit ihnen gearbeitet haben –, erfolgt die Datenbearbeitung mit OpenRefine (ebenso wie mit anderen Tools) nicht rein automatisch. Zwar ermöglicht das Programm es, große Datenmengen binnen Sekunden zu bearbeiten und damit viele Arbeitsstunden einzusparen. Doch um sich zu vergewissern, dass das Ergebnis den qualitativen Anforderungen entspricht, ist eine manuelle Überprüfung unabdinglich. Dieser Schritt erfordert viel Zeit und Aufmerksamkeit. Er ist meines Erachtens allerdings auch besonders produktiv. Denn diese Art der Auseinandersetzung lenkt den Blick auf einzelne Datensätze, in diesem Fall einzelne Personen der Filmgeschichte. Dopplungen, Ähnlichkeiten und andere Auffälligkeiten werfen wiederum neue Fragen auf und machen neugierig auf die individuellen Lebensgeschichten. Sind die Personen mit demselben Nachnamen verwandt? Warum wurden unterschiedliche Geburtsorte auf Wikidata erfasst? Gleichzeitig öffnet diese Datenarbeit den Blick auf übergeordnete Zusammenhänge. Wenn beispielsweise sowohl „Russia“ als auch „Soviet Union“ als Arbeitsorte genannt werden, wird man sich bewusst, dass Pionierinnen wie Aleksandra Khokhlova oder Esfir Shub sowohl vor 1917 als auch nach der Revolution dort gearbeitet haben und man fragt sich, mit welchen weltgeschichtlichen Umwälzungen sie konfrontiert gewesen sein mussten.

Die Arbeit mit datenbasierten Tools wie OpenRefine ist, anders als von einigen angenommen, in der Regel also weder abstrakt oder ausschließlich quantitativ noch rein technisch oder „distanziert“. Stattdessen kann sie ein Eintauchen in die mit den Daten verbundenen Schicksale und historischen Ereignisse ermöglichen. Insofern ist es nicht zwangsläufig der Fall, wie von der Historikerin Kate Hunter befürchtet, dass die Arbeit mit Big Data persönliche Geschichten zugunsten eines Masternarrativs ausblendet (Hunter 2017; Dang 2018: 150f.).

RawGraphs – Chord-Diagramme und Alluvial-Diagramme

Für die Visualisierung der mittels OpenRefine und Wikidata erfassten Geburts- und Sterbeorte der Pionierinnen habe ich RawGraphs (RAWGraphs o. J.) verwendet. Hierbei handelt es sich ebenfalls um ein kostenfreies Open-Source-Tool. RawGraphs ist eine Web-Applikation, die nahezu selbsterklärend funktioniert. Auch für diese Anwendung habe ich mich auf ein Tutorial von Miriam Posner gestützt (Posner 2017b). Wichtig waren zudem das Manual sowie Videotutorials auf YouTube (RAWgraphs, o. J.). Wenngleich das Tool relativ intuitiv funktioniert, stellt es eine Herausforderung dar, sofern man nicht mit Datenvisualisierungen und ihren spezifischen Funktionen und Einsatzbereichen vertraut ist. Zwar erklärt eine Übersicht zu RawGraphs die verschiedenen Visualisierungsarten und -formen, doch welche Daten auf welche Weise und zu welchem Zweck visualisiert werden sollten, erschließt sich einer Autodidaktin erst durch die ausgiebige Erprobung.

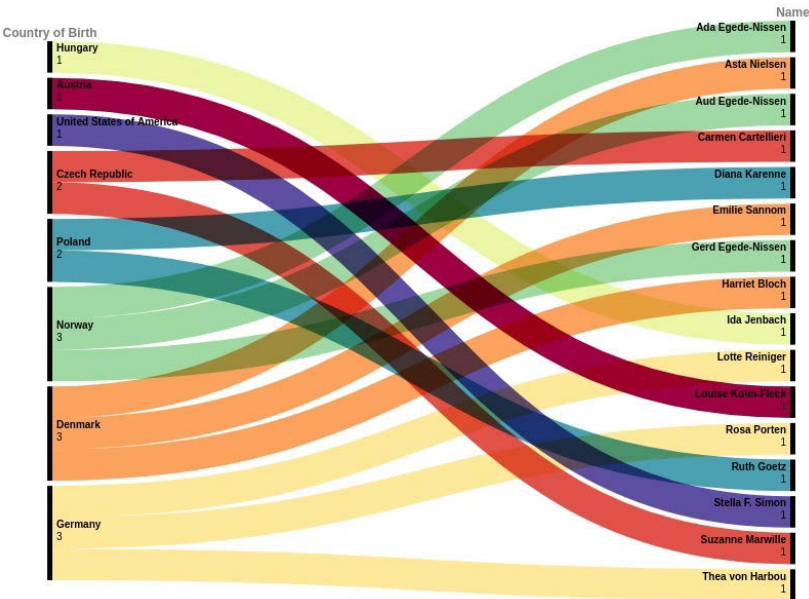
Abb. 2: Chord-Diagramm, erstellt von der Autorin mit Hilfe von RawGraphs auf den durch Wikidata mit Geburts- und Sterbeorten angereicherten Daten des WFPP. Zu sehen sind die Geburtsländer der 16 Frauen, die u.a. in Deutschland gearbeitet haben.



Als nützlich hat sich das Chord-Diagramm herausgestellt (Abb. 2). Mittels dieser Visualisierungsart lassen sich die Beziehungen zwischen Entitäten darstellen und diese bezüglich ihrer Gemeinsamkeiten miteinander vergleichen. Die Breite der Kreissegmente entspricht der Gewichtung der Entität bzw. dem zugewiesenen Wert, die Aufteilung den Anteilen der jeweiligen Entität. Die Farben erleichtern es, die verschiedenen Entitäten zu unterscheiden und die einzelnen Beziehungen zuzuordnen. Wenn zu viele Entitäten dargestellt werden, kann es zu einer Überfrachtung der Visualisierung kommen (Ribeca, o. J. a).

Visualisierungen können nicht nur zur Darstellung von Informationen, sondern auch zur Analyse von Daten dienen. Mithilfe des Chord-Diagramms lässt sich bei den angereicherten WFPP-Daten zum einen zeigen, um welche Frauen es sich konkret handelt. Zudem geht aus der Visualisierung hervor, dass von diesen Frauen, die u.a. in Deutschland gearbeitet haben, drei auch dort geboren wurden: Lotte Reiniger, Rosa Porten und Thea von Harbou. Des Weiteren sehen wir, dass die in Deutschland berühmte Schauspielerinnen und Produzentin Asta Nielsen ursprünglich aus Dänemark kommt. Diese Informationen sind für Filmhistoriker:innen nicht neu, doch stehen sie für die Möglichkeiten, die solche Visualisierungen bieten.

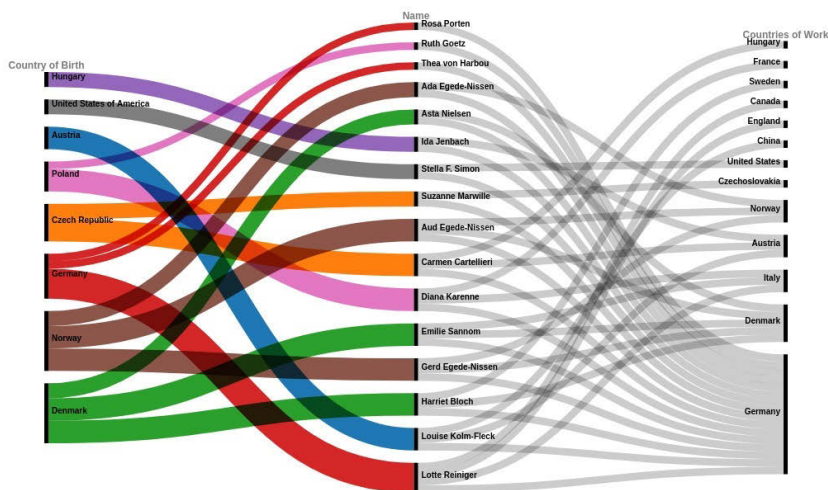
Abb. 3: Alluvial-Diagramm, erstellt von der Autorin mit Hilfe von RawGraphs auf den durch Wikidata mit Geburts- und Sterbeorten angereicherten Daten des WFPP. Zu sehen sind die Geburtsländer der 16 Frauen, die u.a. in Deutschland gearbeitet haben.



Um zu verdeutlichen, in welchen Ländern die Pionierinnen geboren wurden, habe ich mit dem Alluvial-Diagramm experimentiert, gewissermaßen die parallele Variante des Chord-Diagramms (Abb. 3). Hauptsächlich dienen Alluvial-Diagramme dazu aufzuzeigen, wie Daten über mehrere Kategorien verteilt sind. Oftmals werden sie auch benutzt, um zeitliche Veränderungen abzubilden. Die Entitäten werden dabei vertikalen Achsen zugewiesen, die parallel zueinander verlaufen. Die Werte werden durch Blöcke auf jeder Achse dargestellt. Die Höhe eines Blocks steht für die Größe des Clusters und die Höhe einer Verbindung für die Bedeutung bzw. dem zugewiesenen Wert der Entitäten.

Im Vergleich mit dem Chord-Diagramm ist das Alluvial-Diagramm aufgrund der vertikalen Anordnung meines Erachtens für diesen Zweck übersichtlicher. Zudem ist die Reihung der Länder, nämlich gemäß der Anzahl der Referenzen, sofort nachvollziehbar. Neben den oben bereits erwähnten Informationen, offenbart sich im Alluvial-Diagramm die Bandbreite an Herkünften. Dass die sechzehn Frauen, die u. a. in Deutschland gearbeitet haben, aus acht verschiedenen Ländern kommen, ist bemerkenswert.

Abb. 4: Alluvial-Diagramm, erstellt von der Autorin mit Hilfe von RawGraphs auf den durch Wikidata mit Geburts- und Sterbeorten angereicherten Daten des WFPP. Zu sehen sind die Geburtsländer der 16 Frauen und die verschiedenen Länder, in denen sie gearbeitet haben.



Um nicht nur zu zeigen, in welchen Ländern die Pionierinnen geboren wurden, sondern auch, wo sie damals arbeiteten, habe ich in einem weiteren Schritt diese Kategorie im Alluvial-Diagramm aufgenommen (Abb. 4). Dieses hat sich für meine Zwecke, die Migrationsbewegungen nachzuzeichnen, als besonders hilfreich erwiesen. Zu sehen ist beispielsweise, dass Lotte Reiniger neben Deutschland in Italien, England und Kanada tätig war. Die Informationen lassen sich einfacher ablesen, wenn man mit dem zugehörigen WFPP-Profil der Autor:innen Frances Guerin und Anke Mebold vertraut ist (Guerin und Mebold 2016). Wie sich anhand der grauen Verbindungen, die aufgrund ihrer Eintönigkeit etwas mühsam nachzuvollziehen sind, beispielhaft zeigen lässt, erfordern aussagekräftige Visualisierungen neben der fachlichen und technologischen Expertise auch Design-Fertigkeiten. Mittels Excel bzw. LibreOffice Calc erstellte Datenvisualisierungen mögen zwar hinsichtlich ihrer Ästhetik grundsätzlich etwas bescheiden wirken, doch können sie in einem abschließenden Vergleich vielleicht doch besser als gedacht ihren Zweck auch im Kontext film- und medienhistoriographischer Forschung erfüllen.

Datawrapper – interaktive Choroplethenkarten

Bei den mittels RawGraphs erstellten Visualisierungen handelt es sich um statische Darstellungen. Sobald sie produziert und heruntergeladen werden, sind sie weder bearbeitbar noch länger online zugänglich. Das Gegenteil ist der Fall bei den mithilfe von Datawrapper kreierten Visualisierungen (Datawrapper o. J.). Um herauszufinden, welche interaktiven Visualisierungen ohne Programmierkenntnisse möglich sind, habe ich exemplarisch mit sogenannten No-Code-Anwendungen wie der Freemium-Version von Datawrapper, ebenfalls eine browserbasierte Web-Applikation, experimentiert. Hierbei handelt es sich um ein kommerzielles Programm für Datenvisualisierungen, das auch in journalistischen Kontexten verwendet wird. Visualisierungen mittels Datawrapper erlauben nicht nur die Einbindung interaktiver Elemente, sondern können darüber hinaus online veröffentlicht werden. Zudem lässt sich der jeweilige Datensatz als Quelle direkt verlinken. Zwar ist er anders als in Repositorien abgelegte Dokumente nicht mit einem Digital Object Identifier (DOI) versehen, doch zumindest ist die Datengrundlage direkt einsehbar. Eine nachvollziehbare und nachnutzbare Bereitstellung der verwendeten Datensätze, etwa mittels Data Papers, wird in film- und medienwissenschaftlichen Kontexten erst allmählich Teil der Forschungspraxis.

Die hier vorgestellte Herangehensweise bewegt sich im eingangs skizzierten Spannungsfeld zwischen fachfremden Ansätzen und der praktischen Erprobung neuer methodischer Tools in film- und medienwissenschaftlichen Kontexten. Da es sich bei Datenvisualisierungen um eine noch nicht etablierte Praxis handelt, lassen sich bestimmte Aspekte wie etwa die Vergabe von persistenten Identifikatoren wie DOIs, die für eine Best bzw. Good Practice ideal wären, nicht immer

umsetzen. Dies liegt an den technischen Rahmenbedingungen des verwendeten No-Code-Tools, das zwar einen niedrigschwelligen Zugang und flexible Gestaltungsmöglichkeiten bietet, jedoch nicht wissenschaftlichen Standards priorisiert. Die Entscheidung, dennoch mit Datawrapper zu arbeiten, ist als produktiver Kompromiss zu verstehen, der sich an der technischen Umsetzbarkeit ausrichtet, um methodisches Neuland zu erkunden und dabei das Vorgehen wissenschaftlich möglichst nachvollziehbar macht.

Auch für die Visualisierungen mittels Datawrapper habe ich den bereits für den Women Film Pioneers Explorer eingesetzten Datensatz des WFPP verwendet. Die Geburtsländer und Arbeitsländer wurden nun mittels einer Choroplethenkarte, auch Flächenkartogramme genannt, visualisiert. Hierzu habe ich den Datensatz mit den über Datawrapper verfügbaren Ländercodes angereichert. Eine weitere geographische Ansicht sollte einen Vergleich mit der Karte des studentischen Explorer ermöglichen – sowohl bezüglich des Ergebnisses als auch des Workflows.

Das Backend von Datawrapper bietet eine Vielzahl an Einstellungsmöglichkeiten, um die Visualisierung gemäß den spezifischen Bedürfnisse anzupassen (Abb. 5). Anders als bei RawGraphs, wo die Ergebnisse jedes Mal aufs Neue ausgespielt werden müssen, erfolgt die Bearbeitung mittels Datawrapper live, sodass die unmittelbaren Auswirkungen sofort ersichtlich sind. Um zu signalisieren, dass sich aus den blinden Flecken, die durch die erste Fallstudie im Women Film Pioneers Explorer zutage getreten sind, Handlungsbedarf ergibt, habe ich mich für eine Farbpalette entschieden, nach der eine gute Quellenlage durch Grün und eine schlechte Quellenlage durch Pink ausgewiesen wird. Anders als beim Explorer, bei dem die weißen Länder sowohl als blinde Flecken als auch als inexistente Filmkultur gedeutet werden können (s. Abb. 1; Dang 2023a: 8f.), soll die pinke Farbgebung den Blick auf die im WFPP unterrepräsentierten Länder Afrikas, Zentralasiens und des Mittleren Ostens lenken. Wie auch im Explorer wird durch die mittels Datawrapper erstellte Choroplethenkarte ersichtlich, dass nur eine einzige Frau in Tunesien durch das WFPP erfasst wurde (Abb. 5).

In einem nächsten Schritt habe ich die Geburtsländer der Pionierinnen visualisiert. Auf den ersten Blick unterscheidet sich diese Karte kaum von jener der Arbeitsländer. Dies ließe vermuten, dass die Frauen in ihren Geburtsländern auch gearbeitet haben. Wie oben erläutert, verweisen die mittels RawGraphs erstellten Chord-Diagramme und Alluvial-Diagramme jedoch auf das Gegenteil. Ein einfacher Kartenvergleich eignet sich augenscheinlich weniger, um Migrationsbewegungen nachzuzeichnen. Hierzu sind weitere konzeptionelle bzw. visuelle Schritte notwendig.

Abb. 5: Screenshot des Backends von Datawrapper mit Kartenansicht der Arbeitsländer. Die interaktive Karte ist online verfügbar unter https://www.Datawrapper.de/_/zxWgo/.

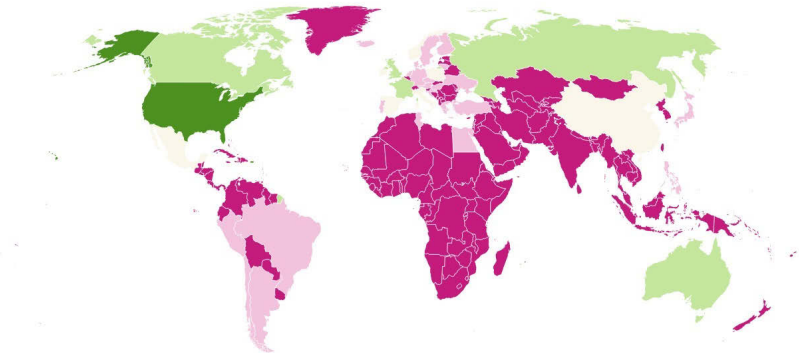


Abb. 6: Kartenansicht der Geburtsländer. Die interaktive Karte ist online verfügbar unter https://www.datawrapper.de/_/1GwsQ/?v=2.

Countries of birth (WFPP) I

This map shows in which countries the women included in the WFPP database were born (steps)

■ < 1 ■ 1–5 ■ 5–10 ■ 10–50 ■ ≥ 50



Map: smd • Source: WFPP, Wikidata • Created with Datawrapper

Was in der Choroplethenkarte auffällt, ist die leichte rosa Einfärbung Ägyptens. In der interaktiven Online-Version wird erkennbar, dass eine Pionierin in Ägypten geboren wurde, die allerdings, so die Ansicht der Arbeitsländer (Abb. 5), nicht dort gearbeitet hat. Ein Blick in den angereicherten Datenansatz verrät, dass es sich hier-

bei um Eval Nil handelt, die in Ägypten geboren worden ist und wohl vor allem in Brasilien arbeitete. Dies bestätigt auch das von Luciana Corrêa de Araújo verfasste Profil im WFPP (Corrêa de Araújo 2020), das wie bereits erwähnt, die Datengrundlage aller hier erprobten Visualisierungen darstellt. Durch das Profil erfahren wir, dass Eva Nil zu den berühmtesten brasilianischen Stars der 1920er Jahre zählt, allerdings keine ihrer Filme erhalten geblieben sind. Darüber hinaus gründete sie eine eigene Produktionsfirma (ebd.).

Choroplethenkarten visualisieren flächenbezogene Daten über Färbungen, Schattierungen und Muster. Ein offensichtlicher Nachteil bei der Verwendung von Farben besteht darin, dass man die Werte auf der Karte nicht genau ablesen oder vergleichen kann. Hinzu kommt etwa, dass größere Länder stärker hervorgehoben erscheinen als kleinere, so dass die Wahrnehmung der Werte beeinträchtigt wird. Ein häufiger Fehler bei der Erstellung von Choroplethenkarten, der sowohl den Studierenden als auch mir unterlaufen ist, besteht in der Kodierung von Rohdatenwerten (z. B. absolute Anzahl der Einträge) anstelle der Verwendung normalisierter Werte (z. B. die Berechnung der Anzahl pro Quadratkilometer). Diesen Fehler habe ich nicht behoben, da es mir in der konkreten Anwendung vorrangig um die Erprobung visueller Darstellungsformen sowie um eine praxeologische Reflexion über den Einsatz digitaler Tools ging. In diesem Fall dient er als Ausgangspunkt für eine kritische Auseinandersetzung mit den Herausforderungen datenbasierter Forschung. Perspektivisch ist eine Normalisierung jedoch zu berücksichtigen, da sie die Vergleichbarkeit zwischen Regionen maßgeblich erhöht (Ribeca, o. J. b).

Wie auch im Kontext der Kritischen Kartographie betont wird, sind Karten keine neutralen Repräsentationen, sondern bilden Machtverhältnisse ab und (re)produzieren diese. Beispielsweise basieren die Weltkarten des Women Film Pioneers Explorer und des Datawrapper-Tools auf der weit verbreiteten, aber durchaus umstrittenen Mercator-Projektion. Auch wenn diese aus dem 16. Jahrhundert stammende Karte sich zur Navigation eignen mag, stellt sie die geographischen Formen und Proportionen verzerrt dar. Dies ist vor allem anhand von Russland, Grönland und Alaska zu sehen, die allesamt unverhältnismäßig groß dargestellt werden. Zudem unterstützt die Mercator-Projektion ein eurozentristisches Weltbild, der Globale Norden ist im Vergleich zum Globalen Süden überdimensional proportioniert (Monmonier 2010). Aus diesem Grund müssen auch bei geographischen Visualisierungen stets die politischen, sozialen, ökonomischen und historischen Bedingungen ihrer Entstehung und Nutzung berücksichtigt werden. Wenn es um historische Daten geht, drängt sich insbesondere die Frage auf, welche Zeitlichkeit für die Visualisierung relevant ist. Gilt es, die historischen Verhältnisse adäquat abzubilden, also die zum Zeitpunkt der Datenerhebung geltende geographische Ordnung, oder jenen Auffassungen gerecht zu werden, die gegenwärtig relevant sind?

Um diese Frage zu beantworten, ist es zunächst notwendig zu verstehen, mit welchen Daten man es überhaupt zu tun hat, wer sie wann und zu welchem Zwecke geschaffen hat. Wenn es sich wie im Falle der vorgestellten Fallstudien um recht heterogene Datensammlungen handelt, ist die Datenprovenienz nur ansatzweise nachzuvollziehen. Die Arbeit mit und an Daten wird nach wie vor eher sporadisch dokumentiert. Um ihre Genese dennoch zu verstehen, können beispielsweise Interviews mit den Beteiligten dabei helfen, die Entstehungsbedingungen festzuhalten (Dang et al. 2025).

Toolkritik und zirkuläre Iterationen

Ziel meiner Forschungsgruppe DAVIF ist es, einen Beitrag zur Sichtbarmachung von Forschung in Form von Daten zu Frauen im frühen Kino zu leisten und zugleich die epistemologischen Potenziale und Grenzen digitaler Visualisierungen und Tools im Kontext film- und medienwissenschaftlicher Forschung kritisch zu reflektieren. Während sich in der Durchführung des Forschungspraxisseminars insbesondere zeigte, welche disziplinären Unterschiede zu überbrücken sind, um eine Zusammenarbeit auf Augenhöhe durchzuführen, wurde in der eigenständigen Erprobung digitaler Tools vor allem deutlich, dass aus film- und medienwissenschaftlicher Sicht Visualisierungen nicht das finale Ergebnis sein können. Vielmehr fungieren sie als eine Art Kompass, wie die Filmhistorikerin Clara Auclair erklärt, der uns dabei hilft, die weitere Richtung zu bestimmen und neue Problemstellungen zu identifizieren (Auclair 2024). Je nach Fragestellung, Interesse und Expertise werden im Laufe solch einer methodischen Orientierungssuche sehr unterschiedliche Aspekte relevant: von Quellenkritik und Datenmodellierungen über die technische und konzeptionelle Umsetzung mittels digitaler Tools oder Softwareprogramme bis hin zur fachspezifischen Interpretation und Vermittlung der Ergebnisse.

Es gibt kein allumfassendes Forschungsprojekt, das alle Aspekte dieser digitalen Hermeneutik gleichermaßen berücksichtigen kann oder berücksichtigen sollte. Stattdessen sind dezidierte Richtungsentscheidungen bewusst in den Forschungsprozess einzubinden. Ohne Fokussierung und gezielte Komplexitätsreduktion sind datenbasierte Projekte kaum gewinnbringend durchzuführen. Es geht darum, Herausforderungen zu identifizieren und Arbeitsprozesse iterativ anzugehen. Dabei sind neben den fachlichen immer auch die zeitlichen, technischen und finanziellen Ressourcen zu berücksichtigen. Zudem ist es wichtig, selbstkritisch zu hinterfragen, welche Forschungsfragen von (fachlichem) Interesse sind. Welche und wessen Geschichte(n) möchten wir erzählen? Welches Tool und welche Daten können hierzu dienlich sein?

Es kann durchaus fruchtbar sein, Forschungsfragen an die technologischen Möglichkeiten anzupassen. Digitale Tools können bestehende Perspektiven er-

weitern und neue Aspekte in den Blick rücken, wie auch der Historiker Thomas Durlacher am Beispiel der Netzwerkanalyse konstatiert: „Traditional history often depends on narratives as a main tool to represent the past. Network analysis extends the toolbox of possible representations of the past by use of a formally rigorous theory of networks. The interpretation of what exactly is represented with the help of networks is dependent on the historian“ (Durlacher 2022: 121). Beispielsweise ist mein Interesse an der geographischen Verteilung der Pionierinnen und den Migrationsbewegungen erst durch die explorativen Datenvisualisierungen erwachsen. Insofern ist die Interpretation nicht nur durch meine subjektive Prädisposition geprägt, sondern ebenso durch die verfügbaren Daten und die spezifischen Tools, die ich im Rahmen meiner Recherche sowie unter den gegebenen finanziellen und technischen Bedingungen gefunden und genutzt habe.

Zugleich besteht jedoch die Gefahr, dass Probleme gemäß der Hammer-Nagel-Metapher vorrangig im Hinblick auf die verfügbaren Werkzeuge konzipiert werden, anstatt von einem Desiderat auszugehen und einen methodischen Ansatz zu entwickeln, um dieses zielführend zu bearbeiten. So stellt etwa die Visualisierung filmhistorischer Kontingenz und epistemologischer Unsicherheiten ein bislang ungelöstes Problem für meine Forschungsgruppe dar (Dang 2023b). Mangels passender – oder vielmehr durch die Dominanz bestimmter – Tools sowie begrenzter designkonzeptioneller Expertise habe ich mich zunächst auf vorhandene, niederschwellige Anwendungen konzentriert. Angesichts des erweiterten Methodenspektrums ist es wichtig, wie bereits erwähnt, die erkenntnistheoretischen Annahmen und Zielsetzungen geisteswissenschaftlicher Forschung von naturwissenschaftlich ausgerichteten Ansätzen – und Tools(!) – zu unterscheiden. „The purpose of humanities data research is not empirical validation and hypothesis testing“, wie Eef Masson in Anlehnung an Stephen Ramsay insistiert, „but the development of questions and the discovery of insights (Ramsay 2003: 173). Rather than import questions and methods from the hard sciences, we must develop our own approaches and sensitivities in working with data that will reflect the humanities' traditions.“ (Masson 2017: 15) In der Erprobung digitaler Tools geht es vor allem darum, die eigenen Fragen nicht aus den Augen zu verlieren und sich der jeweiligen disziplinären Prämissen und Herausforderungen bewusst zu werden bzw. bleiben.

Datenbasierte Forschung zeichnet sich durch „heads-on and hands-on experiences“ (Fickers et al. 2022: 11) aus. Die Arbeitsprozesse gehen dabei alles andere als linear von statten. Vielmehr zeichnen sie sich durch zirkuläre, sich wiederholende, teils verwirrende und oftmals unvorhersehbare Schritte aus. Die Arbeit mit und an Daten bedeutet, wie Fickers, Tatarinov und van der Heijden unter Bezugnahme auf die Literaturwissenschaftler und Digital Humanists Stephen Ramsay (2014) und Joris van Zundert (2016) ausführen, dass, „[d]epending on how the research question is approached and modified over time, new searches for data have to be made, new

tools to be tested, datasets to be adapted and modified, and visualizations or interpretations to be revised and refined“ (2022: 9).

In diesem Sinne ist Toolkritik nicht als Analyse eines einzelnen Tools zu begreifen, sondern, wie die Medienwissenschaftler:innen Karin van Es, Maranke Wieringa und Mirko Tobias Schäfer herausarbeiten, als Reflexion der praktischen und epistemologischen Implikationen in einem konkreten Anwendungszusammenhang (van Es et al. 2018: 26). Wie beeinflusst ein Tool Fragestellung, Forschungsprozess und nicht zuletzt das Ergebnis und seine Präsentation? Bei der Analyse von Tools wie RawGraphs oder Datawrapper ist außerdem zu berücksichtigen, wie die Visualisierungsexpert:innen Catherine D'Ignazio und Rahul Bhargava hervorheben, dass diese für Nutzer:innen entwickelt worden sind – nicht für Lernende bzw. Forschende. Bei diesen Tools geht es darum, schnell Output generieren zu können und nicht um ein tieferliegendes Verständnis der Funktionslogiken, geschweige denn um eine kritische Analyse (D'Ignazio und Bhargava 2016: 84).

Oftmals wird im Zusammenhang mit digitalen Tools und im Zuge der Algorithmenkritik von einer „Blackbox“ gesprochen – insbesondere dann, wenn der Quellcode nicht offen zugänglich ist. Doch selbst Open-Source-Software garantiert kein vollständiges Verständnis der Funktionsweise, vor allem dann nicht, wenn es an informatischer Expertise mangelt. Zudem wird häufig übersehen, dass auch manuelle bzw. semi-manuelle Forschungsprozesse in vielerlei Hinsicht intransparent bleiben. So wird beispielsweise in wissenschaftlichen Texten normalerweise nur jene Literatur zitiert, die Autor:innen als relevant erachten. Aus welchen Gründen bestimmte Quellen nicht aufgeführt werden, bleibt in der Regel unbegründet. Hinzu kommt, dass die Literaturrecherche häufig über Google oder andere digitale Plattformen erfolgt, deren algorithmische Sortierung und Gewichtung nicht offengelegt sind (Underwood 2014). Auch die Ordnungssysteme von Bibliothekskatalogen folgen bestimmten Klassifikationen, deren Logiken den meisten Film- und Medienwissenschaftler:innen nicht bekannt sind. Doch trotz dieser strukturellen Intransparenzen sind eine sorgfältige Zitationspraxis sowie eine umfassende Literaturrecherche über Suchmaschinen und Bibliotheksportale zentrale Praktiken wissenschaftlichen Arbeitens. Allein aus praktischen Gründen lässt sich nicht jeder Aspekt gleichermaßen analysieren oder dokumentieren. Geht es zunächst um eine Annäherung an ein digitales Tool, so kann bereits das Arbeitsergebnis Aufschluss darüber geben, ob das Tool für das konkrete Forschungsvorhaben geeignet ist – etwa zur Visualisierung filmhistorischer Metadaten – oder ob eine alternative Herangehensweise zielführender wäre. In einem nächsten Schritt ist dann im Sinne einer gestuften Toolkritik zu entscheiden, ob und inwiefern die zugrundeliegenden Algorithmen hinterfragt werden sollten – abhängig von der Fragestellung, dem Erkenntnisinteresse und den eigenen forschungspraktischen Ressourcen. Entscheidend ist – und dies gilt, wie eingangs betont, für den gesamten Forschungsprozess – die politischen, sozialen und ökonomischen Rahmenbedin-

gungen ebenso zu reflektieren wie die technologischen, methodologischen und epistemologischen Implikationen.

Im Grunde gehen die mit datenbasierter Forschung verknüpften Anforderungen im wissenschaftlichen Arbeitsalltag über eine „methodologische und epistemologische Doppelbelastung“ (Fickers 2020: 159) hinaus. Vielmehr handelt es sich angesichts der wachsenden Anforderungen um eine multidimensionale Mehrfachherausforderung. Neben sich wandelnden alten und neuen, teilweise fachfremden Expertisen kommen zunehmend praktische Kompetenzen wie Projektmanagement (planen, organisieren, steuern) und Teamführung (kommunizieren, delegieren, unterstützen und sorgen) hinzu. Einerseits kann eine Forschungsgruppe komplexe Arbeitsvolumen bewerkstelligen, die nicht durch eine einzelne Person bewältigbar sind. Andererseits ist sie als Team mit anderen Herausforderungen konfrontiert als ein individuelles Projekt. Diese werden durch befristete Projektlaufzeiten und kontinuierliche Evaluationsanforderungen weiter verstärkt.

Die Vervielfältigung an Methoden und Arbeitsweisen bietet viel Potenzial. Gleichzeitig erfordert dies eine erweiterte Betrachtung der disziplinären Anforderungen. Nur gemeinsam lässt sich dieser Mehrfachherausforderung begegnen, um Forschungsfragen und Erkenntnisinteressen gemäß einer guten wissenschaftlichen Praxis verfolgen zu können. Es ist daher wichtig, sich innerhalb der Fachgemeinschaften darüber zu verständigen, wie den neuen Entwicklungen bestmöglich Rechnung getragen werden kann. Dies umfasst grundsätzliche Fragen: Wie ist beispielsweise ein ertragreiches und anschlussfähiges Datenprojekt zu konzipieren (und zu beantragen)? Welche Art von Förderungen sind nötig, um nachhaltige Forschung zu ermöglichen? Welche Infrastrukturen benötigen wir, um digitale Arbeiten kollaborativ durchzuführen? Wie können neue Erfordernisse in Curricula eingebunden werden? Welche Expertisen sind überhaupt (noch) erforderlich? Auf welchen gemeinsamen Nenner können wir uns angesichts der zunehmenden Ausdifferenzierung an Forschungsgegenständen und Methoden noch verständigen? Nach welchen Kriterien lässt sich datenbasierte Forschung, die unweigerlich explorativ, iterativ und kollaborativ angelegt ist, bewerten und einordnen – insbesondere, wenn wir unter Arbeitsleistung nicht nur den Forschungssoutput, sondern ebenso den Prozess verstehen.

Acknowledgments

Für die äußerst sorgfältige Lektüre sowie die wertvollen Anregungen zur Literatur und zur Schärfung der Argumentation danke ich herzlich den Herausgeberinnen.

Literatur

- Auclair, Clara. 2024. „From Joinville to New Jersey: Early Women Workers Mobilities“. *Vortrag gehalten an der Philipps-Universität Marburg*, Marburg, 15.11.2024.
- Berry, David M., Hg. 2012. „*Understanding Digital Humanities*.“ Palgrave Macmillan.
- Berry, David M. 2022. „Critical Digital Humanities“. In *The Bloomsbury Handbook to the Digital Humanities*, herausgegeben von James O'Sullivan: 125–136. Bloomsbury. <http://dx.doi.org/10.5040/9781350232143.ch-12>.
- Brah, Avtar. 1996. „*Cartographies of Diaspora*.“ Routledge.
- Corrêa de Araújo, Luciana. 2020. „Eva Nil“. In *Women Film Pioneers Project*, herausgegeben von Jane Gaines, Radha Vatsal und Monica Dall'Asta. Columbia University Libraries. Zuletzt abgerufen am 14.01.2025. <https://doi.org/10.7916/d8-qtcy-pd98>.
- Dang, Sarah-Mai. 2018. „Digital Tools & Big Data: Zu gegenwärtigen Herausforderungen für die Film- und Medienwissenschaft am Beispiel der feministischen Filmgeschichte sschreibung“. *MEDIENwissenschaft: Rezensionen | Reviews* 35 (2–3): 142–156. <http://dx.doi.org/10.25969/mediarep/4393>.
- Dang, Sarah-Mai. 2020. „Unknowable Facts and Digital Databases: Reflections on the Women Film Pioneers Project and Women in Film History“. *Digital Humanities Quarterly* 14 (4). <http://digitalhumanities.org/dhq/vol/14/4/000528/000528.html>.
- Dang, Sarah-Mai. 2023a. „Forschung explorieren: Zu den Möglichkeiten digitaler Datenvisualisierungen für die feministische Filmgeschichte sschreibung“. In *Handbuch Digitale Medien und Methoden*, herausgegeben von Laura Niebling, Felix Raczkowski und Sven Stollfuß. Springer VS.
- Dang, Sarah-Mai. 2023b. „Representing the Unknown. A Critical Approach to Digital Data Visualization in the Context of Feminist Film Historiography“. In *How Film Histories Were Made. Material, Methods, Discourses*, herausgegeben von Malte Hagener und Yvonne Zimmermann: 467–493. Amsterdam University Press.
- Dang, Sarah-Mai. 2024. „Managing the Past. Research Data and Film History“. In *Digital Film Historiography. Concepts, Tools, Practices*, herausgegeben von Sarah-Mai Dang, Tim Van der Heijden und Christian Gosvig Olesen. De Gruyter: 135–155. <https://doi.org/10.1515/9783111082486-007>.
- Dang, Sarah-Mai, Pauline Junginger, und Anne Hart, Hg. 2025. „Exploring Film Historical Metadata. Interviews on Data Practices, Data Provenance, and Categorizations.“ In *Film Heritage Collections*. [in Vorbereitung]
- Datawrapper. o. J. <https://www.Datawrapper.de/>.
- Dickel, Henri, Matija Miskovic, Kharazm Noori, Christian Schmidt, Atefeh Soltanifard, Sarah-Mai Dang, und Thorsten Thormählen. 2021. „*Women Film Pioneers Explorer*“. <https://www.online.uni-marburg.de/women-film-pioneers-explore/>.

- D'Ignazio, Catherine, und Rahul Bhargava. 2016. „DataBasic: Design Principles, Tools and Activities for Data Literacy Learners“. *The Journal of Community Informatics* 12 (3). <https://doi.org/10.15353/joci.v12i3.3280>.
- Düring, Marten. 2015. „From Hermeneutics to Data to Networks: Data Extraction and Network Visualization of Historical Sources“. *Programming Historian* 4. <https://doi.org/10.46430/phenoo44>.
- Durlacher, Thomas. 2022. „Philosophical Perspectives on Computational Research Methods in Digital History. The Cases of Topic Modeling and Network Analysis“. In *Digital History and Hermeneutics: Between Theory and Practice*, herausgegeben von Andreas Fickers und Juliane Tatarinov. De Gruyter Oldenbourg. <https://doi.org/10.1515/9783110723991-006>.
- Es, Karin van, Maranke Wieringa, und Mirko Tobias Schäfer. 2018. „Tool Criticism: From Digital Methods to Digital Methodology“. In *WS.2 2018: Proceedings of the 2nd International Conference on Web Studies*. Association for Computing Machinery: 24–27. <https://doi.org/10.1145/3240431.3240436>.
- Fickers, Andreas. 2020. „Update für die Hermeneutik. Geschichtswissenschaft auf dem Weg zur digitalen Forensik?“. *Zeithistorische Forschungen – Studies in Contemporary History* 17 (1): 157–168. <https://doi.org/10.14765/zzf.dok-1765>.
- Fickers, Andreas, Juliane Tatarinov, und Tim van der Heijden. 2022. „Digital History and Hermeneutics – Between Theory and Practice: An introduction“. In *Digital History and Hermeneutics: Between Theory and Practice*, herausgegeben von Andreas Fickers und Juliane Tatarinov: 1–20. De Gruyter Oldenbourg. <https://doi.org/10.1515/9783110723991-001>.
- Gaines, Jane, und Columbia University Libraries. 2020. „Women Film Pioneers Project Biographical Data“. <https://doi.org/10.7916/m4dc-n768>.
- Gaines, Jane, Radha Vatsal, und Monica Dall'Asta. 2013. „Women Film Pioneers Project.“ Columbia University Libraries. Zuletzt abgerufen am 2. Mai 2022. <https://wfpp.columbia.edu/>.
- GitHub. o. J. <https://github.com/MatijaMi/data-vis-women-in-film>.
- Guerin, Frances, und Anke Mebold. 2016. „Lotte Reiniger“. In *Women Film Pioneers Project*, herausgegeben von Jane Gaines, Radha Vatsal und Monica Dall'Asta. Columbia University Libraries. Zuletzt abgerufen am 14.01.2025. <https://doi.org/10.7916/d8-6m3p-mh27>.
- Hunter, Kathryn M. 2017. „Silence in Noisy Archives: Reflections on Judith Allen's ‚Evidence and Silence – Feminism and the Limits of History‘ (1986) in the Era of Mass Digitisation“. *Australian Feminist Studies* 32 (91–92): 202–212. <https://doi.org/10.1080/08164649.2017.1357009>.
- Lave, Jean, und Etienne Wenger. 1991. „*Situated Learning: Legitimate Peripheral Participation*.“ Cambridge University Press.
- Masson, Eef. 2017. „Humanistic Data Research: An Encounter between Epistemic Traditions“. In *The Datafied Society: Studying Culture through Data*, herausgegeben

- von Mirko Tobias Schäfer und Karin van Es: 25–38. Amsterdam University Press.
<https://doi.org/10.1515/9789048531011-004>.
- Monmonier, Mark. 2010. „*Rhumb Lines and Map Wars: A Social History of the Mercator Projection*.“ University of Chicago Press.
- Oiva, Mila und Urszula Pawlicka-Deger, Hg. 2020. „Lab and Slack. Situated Research Practices in Digital Humanities“. Special Issue. *Digital Humanities Quarterly* 14 (3).
- OpenRefine. o. J. <https://openrefine.org/>.
- Posner, Miriam. 2017a. „*Get Started with OpenRefine – Introduction to Digital Humanities*“. Zuletzt abgerufen am 15. Juli 2024. <http://miriamposner.com/classes/dh101f17/tutorials-guides/data-manipulation/get-started-with-openrefine/>.
- Posner, Miriam. 2017b. „*Make a Diagram with RAW – Introduction to Digital Humanities*“. Zuletzt abgerufen am 15. Juli 2024. <http://miriamposner.com/classes/dh101f17/tutorials-guides/data-visualization/make-a-diagram-with-raw/>.
- Ramsay, Stephen. 2003. „Special Section: Reconceiving Text Analysis: Toward an Algorithmic Criticism“. *Literary and Linguistic Computing* 18 (2): 167–174. <https://doi.org/10.1093/lc/18.2.167>.
- Ramsay, Stephen. 2014. „The Hermeneutics of Screwing Around; or What You Do with a Million Books.“ In *Pastplay. Teaching and Learning History with Technology*, herausgegeben von Kevin Kee. University of Michigan Press, 111–120. <https://doi.org/10.3998/dh.12544152.0001.001>.
- RAWgraphs. o. J. YouTube. Zuletzt abgerufen am 15. Juli 2024. https://www.youtube.com/channel/UCuo7DBpCe6PDwaO_3aQsx2A.
- RAWgraphs. o. J. <https://www.rawgraphs.io/>.
- Ribeca, Severino. o. D. a. „*Chord Diagram – Learn About this Chart and Tools to Create it*.“ Zuletzt abgerufen am 7. Januar 2025. https://datavizcatalogue.com/methods/chord_diagram.html.
- Ribeca, Severino. o. J. b. „*Choropleth Map – Learn about this chart and tools to create it*.“ Zuletzt abgerufen am 2. August 2024. <https://datavizcatalogue.com/methods/choropleth.html>.
- Ribeca, Severino. o. J. c. „*The Data Visualisation Catalogue*“. Zuletzt abgerufen am 7. Januar 2025. <https://datavizcatalogue.com/index.html>.
- Rose, Kristina. 2023. „*Webinar: Enriching Film Heritage Data with OpenRefine*.“ European Film Gateways, Vimeo. Zuletzt abgerufen am 14.01.2025. <https://vimeo.com/842110935>.
- Schreibman, Susan, Ray Siemens, und John Unsworth, Hg. 2004. „*Companion to Digital Humanities*.“ Wiley-Blackwell.
- Schreibman, Susan, Ray Siemens, und John Unsworth, Hg. 2016. „*A New Companion to Digital Humanities*.“ Wiley-Blackwell.
- Underwood, Ted. 2014. „Theorizing research practices we forgot to theorize twenty years ago.“ In *Representations* 127 (1): 64–72.

Zundert, Joris J. Van. 2016. „Screwmenetics and Hermenumericals: The Computationality of Hermeneutics“. In *A New Companion to Digital Humanities*, herausgegeben von Susan Schreibman, Ray Siemens und John Unsworth. Wiley-Blackwell: 331–347.