



*Ausgehend von der Abkürzung o.J. (ohne Jahr) analysiert der Beitrag aus film- und medienwissenschaftlicher Perspektive die Zusammenhänge sowie Bedeutung von Recherchepraktiken, Quelldaten und Klassifikationen (Zuschreibungen) im Forschungsprozess. Daran anschließend werden gegenwärtige Fragen aus dem Feld der Digital Humanities diskutiert, die für eine historisch orientierte Medienforschung sowie transdisziplinäre, situierte Wissenschaftspraxis produktiv werden könnten.<sup>1</sup>*

Gebe ich o.J. ohne Leer- und Anführungszeichen in die Suchmaske von Google ein, verweisen die ersten Treffer auf den ehemaligen US-amerikanischen Football-Star und Schauspieler, der 1995 wegen Mordes an seiner Ex-Frau vor Gericht stand: Orenthal James Simpson – O. J. Simpson. Trotz der erdrückenden Beweislast gelang es der Verteidigung, die Geschworenen davon zu überzeugen, ihn freizusprechen. Der Prozess wurde live im Fernsehen übertragen und sorgte nicht nur wegen der Bekanntheit des Angeklagten, sondern auch aufgrund von Rassismuskwürfen international für Aufmerksamkeit. 2016 erschien die vielfach ausgezeichnete Fernsehserie *The People v. O. J. Simpson: American Crime Story*, die den mehrmonatigen Prozess durch neue Perspektiven auf das Geschehen zu rekonstruieren sucht.

Füge ich ein Leerzeichen in die Suchmaske ein, wie es der Duden vorsieht, o. J., so zeigt mir Google zumindest an dritter Stelle einen Wiktionary-Eintrag zu jener Thematik an, der ich mich eigentlich widmen wollte: der bibliographischen Angabe *ohne Jahr*. Selbst wenn ich die Suchanfrage in Anführungszeichen setze, erhalte ich eine Ergebnisliste, die dem, was mir zu Beginn vorschwebte, nur ansatzweise nahekommt. Nur drei von zehn Einträgen beziehen sich auf die Bedeutung der bibliographischen Abkürzung. Die übrigen verweisen vor allem auf O. J. Simpson. Hinzu kommt ein Marktbericht des US-amerikanischen Landwirtschaftsministeriums aus dem Jahr 1926 sowie das offizielle Register der US-amerikanischen Air Force, Stand 1. Januar 1952. Kurios.

Ausgehend von dieser Suchanfrage und der Abkürzung o.J. möchte ich in dem vorliegenden Beitrag drei Facetten digitaler Filmgeschichtsschreibung aufzeigen, die mir als Film- und Medienwissenschaftlerin für die Auseinandersetzung mit gegenwärtigen Forschungspraktiken relevant zu sein scheinen – und denen wir angesichts der Digitalisierung verstärkt Aufmerksamkeit widmen sollten: 1) Recherchepraktiken, 2) Datenquellen, 3) Datenmodellierungen. Zunächst werde ich skizzieren, welchen Bedingungen die alltägliche Recherchepraxis im Internet und in anderen digitalen Datenbanken unterliegt, um anschließend das Verhältnis von Daten und Quellen

sowie den Komplex von Zuschreibungen im Rahmen von Datenmodellierungen zu beleuchten. Der Beitrag zielt darauf ab, ausschnittshaft gegenwärtige Fragen im Feld der Digital Humanities zur Bedeutung von Daten für die historisch orientierte Medienforschung produktiv zu machen und somit den transdisziplinären Austausch über die Situiertheit von wissenschaftlichem Wissen zu vertiefen.

## Recherchepraktiken: Suchanfragen und Programme

Das Eingangsbeispiel verdeutlicht, dass ein einziges Leerzeichen große Auswirkungen auf unsere alltägliche Recherche haben kann. Wie genau und auf welche Weise die unterschiedlichen Suchanfragen das jeweilige Ergebnis beeinflussen, lässt sich allerdings nur bedingt erfassen. Denn die den Ergebnissen zugrundeliegenden Algorithmen, die programmierte Abfolge definierter Schritte, sind bei proprietären Anwendungen normalerweise nicht öffentlich zugänglich. Dass alle drei Anfragevarianten (*o.J.*, *o. J.*, *„o. J.“*) überwiegend Einträge zu O. J. Simpson hervorbringen, ist höchst wahrscheinlich auf die Popularität des Superstars zurückzuführen. Inwiefern es außerdem eine Rolle spielt, dass die Serie seit einigen Jahren auf meiner Watchlist steht, lässt sich nur vermuten. Die Logik, nach welcher kommerzielle Anbieter:innen personenbezogene Daten speichern und entsprechende Nutzungsprofile generieren, die in die Funktionsweise der Applikationen einfließen, ist nach wie vor undurchsichtig. Gleichwohl ist zu bedenken, dass auch ein veröffentlichter, also Open-Source-Code nicht selbsterklärend ist, sondern informatische Kenntnisse und eine nachvollziehbare Dokumentation voraussetzt. Der wachsende Einfluss globaler Digitalkonzerne auf die Informations- und Wissensproduktion gehört meines Erachtens zu den zentralen Aspekten, die es in der Reflexion von Forschungspraktiken und digitalen Infrastrukturen zu berücksichtigen gilt.<sup>2</sup> Jedoch ist auch die Rolle gemeinnütziger Plattformen wie Wikipedia im Kontext wissenschaftlicher Recherchepraktiken und darüber hinaus kritisch zu untersuchen. Denn bei der *freien Enzyklopädie* können wir ebenfalls nur ansatzweise nachvollziehen, welche und wessen Interessen die unzähligen Einträge beeinflussen und nach welcher diskursiven Logik Themen überhaupt Eingang in die nutzer:innengenerierte Online-Plattform finden (Dang 2020: 130f.).

Wie Ted Underwood in seiner Analyse digitaler Forschungspraktiken argumentiert, verschleiert der unscheinbar wirkende Begriff *Suchanfrage* (*search*) die komplexen technologischen und epistemologischen Prämissen, die mit seinen unterschiedlichen

Manifestationen verbunden sind (Underwood 2014: 64). Automatisierte Suchprogramme – sei es zwecks Informationsabfrage via Google oder Volltextanalyse im Zuge des sogenannten *Distant Reading* – sind, anders als die geläufige Bezeichnung *Tools* suggeriert, nämlich nicht als bloße Werkzeuge zu betrachten, die uns allein dazu dienen, das Forschen zu erleichtern. Vielmehr sind sie als ein wichtiger Bestandteil unserer Methodiken und damit unseres gesamten Forschungsdesigns zu verstehen. Denn digitale Tools beeinflussen grundsätzlich, wie wir unsere Gegenstände konturieren und begreifen und welche Fragen wir an sie und mit ihnen stellen (auch Abs., Schäffer 2023). Aus diesem Grund sollten wir die eingesetzten Anwendungen stets auch als Gegenstand unserer Untersuchung berücksichtigen (Dang 2018). Beispielsweise können wir eine Suchanfrage so oft anpassen, bis wir ein Ergebnis erhalten, das wir in etwa erwartet haben oder das zumindest in eine Richtung weist, die wir als produktiv erachten (Underwood 2014: 65ff.). Underwood führt seine Überlegungen mit Blick auf die Volltextsuche als Form des Data Mining aus, doch ist seine Argumentation – wenngleich unter anderen Vorzeichen angesichts der Übermacht von Google – ebenso auf das Eingangsbeispiel dieses Beitrags übertragbar. Einerseits bergen Internetsuchmaschinen die Gefahr, dass wir aufgrund der personalisierten Algorithmen in unserer Filterblase verweilen und in der Regel vorhersehbare Ergebnisse erhalten, oder dass wir gezielt mittels Google manipuliert werden. Andererseits können Suchmaschinen ebenso unverhoffte Treffer produzieren oder assoziative Denkprozesse in Gang setzen, die sich außerhalb des zuvor abgesteckten Rahmens vollziehen, wie im Beispiel der Suchanfrage o.J.. In jedem Fall sollten wir uns bewusst machen, dass wir es trotz des fortwährenden Big-Data-Diskurses weder mit kompletten noch unendlichen Datensätzen zu tun haben. Obwohl wir mit Datenbanken generell eine gewisse Vollständigkeit assoziieren mögen und uns definitives Wissen von Daten versprechen, halten sie weder lückenlose Sammlungen noch neutrale Informationen vor. Film- und Medienhistoriker:innen arbeiten – im Gegenteil – meist mit vereinzelt, fragmentarischen und „höchst interpretationsbedürftigen Datensätzen“, wie Patrick Vonderau (2017: 10) erklärt.

Ebenso wie Archive spiegeln Datenbanken intellektuelle Konventionen und institutionelle Rahmenbedingungen wider, denen verschiedene Vorstellungen und Funktionen eingeschrieben sind. Im Bereich der film- und medienhistorischen Forschung wird beispielsweise hervorgehoben, dass digitale Datenbanken uns den Zugang zu Artefakten erleichtern, Überblicke über große Kollektionen verschaffen,

historische Trends aufzeigen und/oder neue Perspektiven eröffnen können. Gleichzeitig dürfen wir nicht vergessen, dass es sich bei Datenbanken um eine (mehr oder weniger bewusste) Selektion von Informationen handelt, die in ein bestimmtes Ordnungssystem eingeschrieben sind. So definieren Metadaten immer eine partikulare Sicht auf ausgewählte Entitäten, womit andere Attribute oder Objekte zwangsläufig außer Acht geraten (Dang 2018).

## Datenquellen: Datenbanken und Quellenkritik

Wie seit geraumer Zeit in Anbetracht einer zunehmenden Digitalisierung der Geisteswissenschaften vielerorts herausgestellt wird, sind nicht nur Datenbanken aus medienwissenschaftlicher Sicht näher zu untersuchen (Burkhardt 2015), sondern auch Daten selbst (Gitelman 2013; Schäfer/van Es 2017). Daten sind nicht naturgegeben, wie der lateinische Begriff suggeriert. Sie sind Ergebnis oder vielmehr Interpretation einer historisch-kulturellen Konstellation und daher immer im Kontext zu betrachten. Daten werden unter bestimmten Bedingungen zu einem konkreten Zweck generiert und analysiert. Insofern repräsentieren sie stets einen spezifischen Zugang zur Welt (Wilkinson 2005: 41), der sich nur unter Berücksichtigung des jeweiligen Entstehungs- und Rezeptionskontexts erschließt. Um die daraus hervorgehende Situiertheit von Daten im Auge zu behalten, schlägt Johanna Drucker (2016: 249) vor, von *capta* (im Sinne von *erarbeitet* und *erhoben*) anstatt von Daten (im Sinne von *gegeben*) zu sprechen.

Möchten wir die Implikationen einer Suchanfrage besser verstehen, müssen wir vor diesem Hintergrund neben den Algorithmen von Programmen auch das Ausgangsmaterial, welches einer Datenbank zugrunde liegt, untersuchen. Zuallererst wäre herauszuarbeiten, welches Material überhaupt zu einer Quelle gemacht wird, aus der dann wiederum Daten extrahiert werden. Dabei gälte es zu beachten, welche Informationen unter welchen Prämissen und zu welchem Zweck einer Quelle über Annotationen (also Definitionen, Anmerkungen und Hinweise) oder als Metadaten (etwa Entstehungsdatum, Objektbezeichnung, Produzent:in) hinzugefügt wurden. Und nicht zuletzt stellt sich aus epistemologischer Sicht die Frage, mit welchem Forschungsgegenstand wir es eigentlich zu tun haben, wenn wir Quellen in Form von Daten nutzen, beispielsweise im Rahmen einer computergestützten Präsentation oder Analyse digitaler Artefakte, oder wenn Daten für uns Quellen darstellen, wie bei der Recherche in Datenbanken. Im Fall von filmhistoriographischen Onlineplattformen wie dem *Women Film Pioneers Project (WFPP)*<sup>3</sup>,

das Informationen und weiterführende Ressourcen zu Frauen im Frühen Kino (1890 bis 1920er Jahre) versammelt, wäre systematisch zu untersuchen, auf welcher Basis, also nach welchen Kriterien, die mehr als 300 Profile umfassende Datenbank erstellt wurde: Wie wurde aus historischen und historiographischen Quellen Daten gewonnen, die wiederum als Quellen für weitere Forschungsprojekte dienen? An welchem Punkt wurde entschieden, wer eigentlich als *Woman Film Pioneer* zählt – und wer nicht? Es wäre herauszuarbeiten, was *Woman, Film, Pioneer* mit Blick auf die Datenbank bedeutet und welche Annahmen diesen Begriffen vorausgehen. In anderen Worten: Welche Datenmodellierung liegt der umfänglichen Übersicht des *WFPP* zugrunde und wie ist diese diskursiv zu verorten? Wie wird durch die Plattform Wissen re/präsentiert und Bedeutung generiert?

## Datenmodellierungen: Standardisierungen und Klassifikationen

Die Frage nach der Datenmodellierung rückt gegenwärtig immer stärker ins Zentrum film- und medienhistorischer Forschung im Feld der Digital Humanities. Unter Datenmodellierung ist, allgemein gesprochen, ein formales Verfahren zu verstehen, das darauf abzielt, Daten für Menschen und/oder Computer verständlich, also les- und nutzbar aufzubereiten.<sup>4</sup> So gesehen stellt ein Datenmodell nach Julia Flanders und Fotis Jannidis (2019a: 28) eine Repräsentation von etwas durch jemanden mit einem definierten Ziel zu einem konkreten Zeitpunkt dar. Hinsichtlich der Ausrichtung eines jeweiligen Modells wird gemeinhin zwischen *curation-driven* und *research-driven* unterschieden. Während eine *curation-driven* Modellierung vor allem die potenzielle Nachnutzung in unterschiedlichen Zusammenhängen fokussiert und infolgedessen Interoperabilität durch Standards zu gewährleisten sucht, zielt eine *research-driven* Modellierung in erster Linie auf ein konkretes Forschungsinteresse ab, das aus einem spezifischen disziplinären Kontext herrührt (Flanders/Jannidis 2019b: 86). Allerdings lassen sich diese Definitionen nicht immer scharf voneinander trennen. Denn auch für theoretisch-konzeptionell ausgerichtete Modellierungen spielt die Anschlussfähigkeit eine Rolle – und Datenkurator:innen verfolgen wiederum Ziele, denen diskursiv verankerte Annahmen vorausgehen. Beide Ansätze sind mit der Herausforderung konfrontiert, zwischen den Vor- und Nachteilen standardisierter und situierter Verfahren abzuwägen. So sind etwa Fragen nach der Auswahl oder Vereinheitlichung von Daten zulasten der Heterogenität und Mehrdeutigkeit des Materials zu berücksichtigen (Dang/Hirsbrunner 2019).

Wie Katrin Moeller in diesem Zusammenhang jedoch bemerkt, wird Standardisierungsinitiativen zu Unrecht eher wenig Innovationspotenzial zugesprochen. Denn eigentlich handelt es sich dabei um eine höchst anspruchsvolle und produktive Angelegenheit (Moeller 2019: 19). Am Beispiel von modernen und historischen Berufsklassifikationen erläutert sie, weshalb eine möglichst gute Umsetzung der FAIR-Data-Prinzipien (*findable, accessible, interoperable, reusable*) eine zwar sehr aufwendige, aber durchaus erstrebenswerte, wenn nicht notwendige wissenschaftliche Aufgabe im Zuge der Digitalisierung der Geisteswissenschaften beziehungsweise der Geschichtswissenschaft darstellt. Entsprechend sei diese zu kreditieren (ebd.: 17f.). Die Aneignung und Erarbeitung methodischen und theoretischen Wissens um Forschungsdatenmanagement ist ein wesentlicher Schritt, um fachübergreifende, forschungsrelevante Standards zu etablieren und somit eine Qualitätssicherung im Sinne guter wissenschaftlicher Praxis (Stichwort: Transparenz) zu gewährleisten (ebd.). Formalisierte Verfahren führen nach Moeller interessanterweise nicht unbedingt zu Einschränkungen der Forschung, wie von manchen Wissenschaftler:innen befürchtet. Idealerweise erlauben sie uns, Informationsverlust vorzubeugen und eine differenzierte Verortung von Datenquellen vorzunehmen. Voraussetzung dafür sei, dass bei Datenmodellierungen sowohl die a) originären Varianten, b) nachträglichen Normierungen und c) persistenten, übergeordneten Standards als auch die d) flexiblen, forschungsorientierten Klassifikationen vorgehalten werden (ebd.: 34ff.).

Für die Berufsbezeichnung *Böttcher* führt Moeller exemplarisch die originären Varianten (a) *Fassbinde*, *Fassbinderin*, *Fassbendergesell* und *Fassbendermeister* an. Auf der Ebene der Normschreibung (b) heiße es entsprechend *Binder*, *Binderin*, *Bindergeselle*, *Bindermeister*; und nach der Standardisierung beziehungsweise Klassifikation der Berufe des Statistischen Bundesamts *Böttcher* (c) und *Böttchermeister* (d). Während die Varianten in Verbindung mit den Normschreibungen eine historische Einordnung nachvollziehbar machen, könnten die Klassifikationen je nach Zweck und Interesse dynamisch angepasst werden. Berufsbezeichnungen seien hier insofern höchst relevant, als dass sie nicht nur Auskunft über eine Tätigkeit geben, sondern auch über den sozialen Status einer Person und ihren Lebenslauf sowie die hierarchischen Strukturen einer Gesellschaft (ebd.: 23). Berufsbezeichnungen ebenso wie Tätigkeitsprofile divergieren je nach kulturellem und sozialem Kontext und von Sprache zu Sprache. Entsprechende Standardisierungen setzen daher komplexe Übersetzungsleistungen von Raum

und Zeit im engen Austausch der Fachcommunity mit Gedächtnisinstitutionen voraus. Wie historische Transformationen zwecks projektübergreifenden Austauschs erfasst werden können, bis zu welchem Grad etwa automatisiert vorgegangen werden kann oder an welchen Stellen eine manuelle Erkennung notwendig ist, gehört zu den aktuellen Fragen von Standardisierungsbestrebungen.

Mit Blick auf das *Women Film Pioneers Project* wäre angesichts des vorhandenen Datenmaterials beispielsweise sowohl aus filmhistorischer als auch aus archivarischer Sicht zu überlegen, wie mit den unterschiedlichen Berufsbezeichnungen im Bereich der Filmmontage im Zusammenhang eines projektbezogenen und gleichermaßen infrastrukturell-nachhaltigen Forschungsdatenmanagements umzugehen ist. In der Datenbank des *WFPP* sind unter anderem die Varianten *cutter*, *editor* oder *scenario editor* zu finden, wobei die beiden letzteren damals nicht unbedingt die Tätigkeit des Schnitts involvierten, sondern vor allem dramaturgische Aspekte (Hatch 2013). Bezüglich der Datenbank des *WFPP* gälte es herauszuarbeiten, wann unterschiedliche Bezeichnungen auf dieselbe oder eine ähnliche Tätigkeit abzielten oder inwiefern verschiedene Tätigkeiten unter ein und derselben Bezeichnung versammelt wurden – und welche Implikationen damit jeweils einhergingen und -gehen (etwa soziale Erwartungen, Wertschätzung, berufliche Anerkennung, Bezahlung, Identität, Sichtbarkeit) (Hill 2016).

Neben kuratorisch-pragmatischen und forschungskonzeptionellen Fragen spielen nicht zuletzt ethische Aspekte eine Rolle in der Datenmodellierung. So ist zum einen herauszuarbeiten, wie wir Anschlussmöglichkeiten für andere Arbeitszusammenhänge gewährleisten oder Daten für konkrete Forschungsfragen aufbereiten können. Zum anderen ist aufzuzeigen, wie in diesem Prozess die Integrität von in Datenbanken erfassten Entitäten, zum Beispiel Personen und Objekte, aufrechterhalten werden kann. Klassifikationen stellen auch diesbezüglich eine entscheidende Aufgabe dar. Wie Flanders (2021) am Beispiel von Genderzuschreibungen und vor dem Hintergrund des von ihr initiierten *Women Writers Project*<sup>5</sup> argumentiert, können keine allgemeingültigen Klassifikationsregelungen geltend gemacht werden. Stattdessen ist, wie auch Flanders erklärt, von Fall zu Fall abzuwägen, auf welche Weise eine ethisch vertretbare und für das Anliegen sinnige Einordnung bestmöglich vorzunehmen ist. Für manche Zwecke mag eine binäre Gender-Klassifikation zielführend sein, etwa wenn wir die Repräsentation von Frauen in der britischen Filmindustrie diskutieren (Wreyford/Cobb 2017) oder die unzähligen Pionierinnen in der Stummfilmzeit in den Fokus rücken wol-

len. In anderen Projekten könnte die Kombination multipler Genderkategorien produktiv sein (Flanders 2021). Grundsätzlich sollten wir uns fragen, welche Prämissen mit einer Klassifizierung wie *Frau* verbunden sind (auch Garz/Rittiens 2023; Macgilchrist 2023). Hinsichtlich des *Women Film Pioneers Project* wäre beispielsweise zu erforschen, inwiefern es sich um eine mit Hilfe verschiedenen Quellenmaterials vorgenommene Fremd- oder Selbstzuschreibung handelt. Gehen wir von Wahrscheinlichkeiten oder eher Vermutungen aus? Betrachten wir einen Fakt oder ein Argument?

Bestenfalls werden in der Nutzung und Bearbeitung von Daten nicht nur die Provenienz und Funktion von Quellen festgehalten, sondern ebenso der Prozess des Klassifizierens und Kontextualisierens, sodass die Modulation für kritische Auseinandersetzungen und das Entwickeln von Folgeprojekten nachvollzogen werden kann (Flanders/Jannidis 2019b: 6). Denn wie wir Daten erheben, nutzen und aufbereiten, welche Klassifikationen wir anwenden, um Daten lesbar beziehungsweise interpretierbar zu machen, entscheidet darüber, welche Informationen wir von den Daten ableiten und letztendlich, welches Wissen wir daraus generieren können. Nicht immer ist nur das Ausgangsmaterial von Interesse, das mittels Metadaten angereichert und organisiert wird. Manchmal ist der kulturelle Diskurs, der sich um Artefakte entfacht, weitaus spannender, wie Gi-oele Barabucchi und Fabio Vitali darlegen (Digital History Berlin 2021). Aus diesem Grund sollten auch zeitliche, räumliche, personelle, intellektuelle und künstlerische Aspekte im Zuge von Datenmodellierungen reflektiert werden, um eine möglichst differenzierte (Re-)Kontextualisierung dieses Prozesses zu ermöglichen (ebd.). Welche theoretischen, aber auch praktisch umsetzbaren Modelle einen oder multiple Datenlebenszyklen umfänglich abzubilden vermögen, sodass ebenso widersprüchliche beziehungsweise mehrstimmige Interpretationen darin wiederzufinden sind, gehört daher in vielerlei Hinsicht zu den zentralen Fragen sowohl informatischer als auch geisteswissenschaftlicher Forschung an computerbasierter Verarbeitung und Darstellung von Informationen und Wissen.<sup>6</sup>

## Partikulare Einsichten und plurale Perspektiven

In der Wissenschaft bezieht sich o.J. auf eine Quelle, deren historischer Kontext nicht identifizierbar ist. Vage oder mehrdeutige Daten, also unsicheres Wissen oder eben blinde Flecken abzubilden, ist ein Hauptanliegen feministischer Filmhistoriographie (Gledhill/Knight 2015; Dall'Asta/Gaines 2015).

Dabei besteht die Herausforderung darin, Lücken aufzuzeigen, ohne zu behaupten, dass, wenn wir uns nur genügend anstrengen, sich irgendwann ein vollständiges, finales Bild ergeben würde. Stattdessen gilt es, plurale Perspektiven darzulegen und Wissen in seiner Situiertheit zu reflektieren, ohne relativistisch zu sein (Haraway 1988). Dies trifft ebenso auf die Reflexion unserer Methoden, Ansätze und Praktiken zu: Wollen wir partikuläre Einsichten nachvollziehen beziehungsweise nachvollziehbar machen, gälte es im Grunde, die epistemologischen und methodologischen Bedingungen jeder einzelnen Anwendung in ihrem spezifischen Gebrauchskontext zu analysieren. Die im Feld der Digital Humanities zunehmende Forderung nach einer differenzierten Kritik digitaler Tools – bei der ebenso ökonomische und politische Bedingungen einbezogen werden – müsste bereits bei der Reflexion des Literaturverwaltungsprogramms oder anderer Forschungsdatenmanagementapplikationen oder eben der verwendeten Suchmaschine ansetzen.

Während Datenmodellierungen seit einiger Zeit im Zentrum datenbasierter Forschung stehen, wird alltäglichen Arbeitspraktiken in der Wissenschaft indes erstaunlich wenig Aufmerksamkeit geschenkt. Rechercheprozesse werden kaum systematisch und schon gar nicht öffentlich dokumentiert (Dang 2020: 132f.). Selten fließen die benutzten Werkzeuge oder Programme in eine Publikation ein.<sup>7</sup> Literatur, die bloß überflogen oder angelesen, aber doch nicht zitiert wird, bleibt in der Regel unerwähnt. Dabei wäre es im Sinne einer guten wissenschaftlichen Praxis durchaus aufschlussreich, wenn Leser:innen ebenfalls wüssten, welche Quellen von den Autor:innen *nicht* verwendet wurden oder zu welcher Frage sich *keine* weiteren Informationen finden ließen. Die interpretativen Bedingungen und Beschränkungen von Forschung könnten so verständlicher werden (ebd.). Mit der wachsenden Verbreitung digitaler Tools scheint sich das Bewusstsein für die eigenen Forschungspraktiken allerdings allmählich zu ändern. Dies zeigt auch ein Blick in aktuelle Konferenzprogramme (vDHD 2021). Grundsätzliche Fragen nach Gegenstand, Methodik und Erkenntnisinteresse und, wie sich diese in der Forschung gegenseitig bedingen, rücken in den Diskursen um die Potentiale und Grenzen der Digital Humanities wieder verstärkt in den Blick. Stellte Vonderau (2017) vor wenigen Jahren noch fest, dass trotz zahlreicher Einzelprojekte und internationaler Tagungen zu computergestützter Forschung „digitale Werkzeuge selbst bislang nicht zum Kanon neuer Forschungsfelder“ zählen, ist heute zu beobachten, dass immer häufiger die Forderung nach *Digital Literacy* oder *Data Literacy* zu vernehmen ist. *Digitale Quellenkritik*, *Toolkritik* oder gar

*Tool Science* (Wolff 2015) sind in diesem Zusammenhang Schlagworte, die auf die Reflexion neuer – sowie bisheriger – Arbeitspraktiken, Prämissen und Kontexte in den Geisteswissenschaften abzielen.

Etablierte medien- und filmwissenschaftliche Praktiken sowie Methoden bleiben im Zuge gegenwärtiger Transformationen weiterhin relevant, verändern sich jedoch und werden ergänzt. Wenn wir digitale Tools nicht bloß als Hilfsmittel begreifen, sondern als wesentlichen Bestandteil unserer wissenschaftlichen Arbeit, sollten wir diese sowohl zum Gegenstand unserer Reflexionen machen als auch zum Ziel. Auf diese Weise können wir – anstatt uns auf Google und Co zu verlassen – neue Anwendungen kreieren oder bestehende anpassen und weiterentwickeln, um diese gezielt und bewusst gemäß des spezifischen Forschungsinteresses einzusetzen, etwa um alternative Such- und Zugriffsmöglichkeiten auf Datenbanken zu ermöglichen. Hierzu ist es meines Erachtens nicht unbedingt erforderlich, nebst einem geisteswissenschaftlichen Studium programmieren zu lernen. Doch als Film- und Medienwissenschaftler:innen sollten wir uns mit den Herausforderungen digitaler Forschungspraktiken umfassend auseinandersetzen, etwa mit Datenmodellierungen, digitalen Infrastrukturen oder gar Statistik, um transdisziplinär und institutionenübergreifend kommunizieren und zusammenarbeiten zu können. Hinsichtlich der zunehmenden Produktion und Nutzung von Daten bedarf es schließlich einer kollaborativen Arbeitsweise, durch welche die verschiedensten Expertisen in einen produktiven Wissensaustausch gebracht werden, sodass bei der nächsten Suchanfrage nachvollziehbarer wird, nach welcher Logik Ergebnisse gelistet und Daten als Informationen aufbereitet werden. Wobei mich der unerwartete Verweis von Google auf O. J. Simpson nun tatsächlich dazu veranlasst hat, die preisgekrönte Serie endlich anzusehen.

## Anmerkungen

- 1 Dieser Text ist im Rahmen des Forschungsprojekts *Ästhetiken des Zugangs. Datenvisualisierungen in der digitalen Filmgeschichtsschreibung am Beispiel der Forschung zu Frauen im Frühen Kino (DAVIF)* entstanden, das vom BMBF zur theoretischen, methodischen und technischen Weiterentwicklung der digitalen Geisteswissenschaften gefördert wird.
- 2 Wie ich bereits an anderer Stelle in der Auseinandersetzung mit Forschungsdatenmanagement in der Film- und Medienwissenschaft umrissen habe (Dang 2020: 130f.).
- 3 URL: [wfpp.columbia.edu/](http://wfpp.columbia.edu/) [13.09.2021].
- 4 Eine umfassende Diskussion wesentlicher Aspekte sowie verschiedener Praktiken der Datenmodellierung bieten Flanders/Jannidis (2019b) in ihrem Sammelband *The Shape of Data in the Digital Humanities*.
- 5 URL: [wpp.northeastern.edu/](http://wpp.northeastern.edu/) [13.09.2021].
- 6 Dies bezugnehmend auf Veranstaltungen wie die internationale Konferenz *Data for History: Modelling Time, Space, Agents* (2021), URL: [d4h2020.sciencesconf.org/](http://d4h2020.sciencesconf.org/) [13.09.2021], auf der auch Barabucchi und Vitali ihr Paper zur Kontextualisierung von Datenmodellierungen präsentierten (Digital History Berlin 2021). Es gibt bereits eine Vielzahl an Datenmodellen, die sich dieser Herausforderung der Wissens- und Informationsintegration annehmen. Eine detaillierte Auseinandersetzung würde jedoch einen weiteren Beitrag erfordern.
- 7 Ein interessantes Ausnahmebeispiel stellt Peter Andorfers *DARIAH-DE* Working Paper dar, das sich unter anderem mit eben dieser Frage nach der Transparenz (Nachvollziehbarkeit) wissenschaftlicher Praktiken auseinandersetzt (Andorfer 2015).

## Referenzen

- Andorfer, Peter (2015). Forschungsdaten in den (digitalen) Geisteswissenschaften. Versuch einer Konkretisierung. *DARIAH-DE Working Papers*, 14.
- Burkhardt, Marcus (2015). *Digitale Datenbanken. Eine Medientheorie im Zeitalter von Big Data*. 1. Aufl. Berlin et al.: Transcript.
- Dall'Asta, Monica/Gaines, Jane M. (2015). Prologue. Constellations. Past Meets Present in Feminist Film History. In Christine Gledhill/Julia Knight (Hg.), *Doing Women's Film History: Reframing Cinemas, Past and Future*. Urbana et al.: University of Illinois Press, 13–25.
- Dang, Sarah-Mai (2018). Digital Tools & Big Data. Zu gegenwärtigen Herausforderungen für die Film- und Medienwissenschaft am Beispiel der feministischen Filmgeschichtsschreibung. *MEDIENWISSENSCHAFT: Rezensionen | Reviews*, 2–3: 142–156. <https://doi.org/10.17192/ep2018.2.3.7836>
- Dang, Sarah-Mai (2020). Forschungsdatenmanagement in der Filmwissenschaft. Daten, Praktiken und Erkenntnisprozesse. *montage/av*, 29(1), 119–40.
- Dang, Sarah-Mai/Hirsbrunner Simon D. (2019). Opening Research Data. Amplification and Reduction within Media Research Practices. *Open-Media-Studies-Blog*. URL: [kurzelinks.de/kb60](http://kurzelinks.de/kb60) [13.09.2021]
- Digital History Berlin (2021). *Gioele Barabucchi, Fabio Vitali: Context is all*. URL: [youtube.com/watch?v=rgn154jc8g](https://youtube.com/watch?v=rgn154jc8g) [13.09.2021]
- Drucker, Johanna (2016). „Graphical Approaches to the Digital Humanities“. In Susan Schreibman/ Raymond G. Siemens/John Unsworth (Hg.), *A New Companion to Digital Humanities*. Chichester et al.: Wiley/Blackwell, 238–250.

- Flanders, Julia (2021). Gender in the Machine. Representing Gender in Digital Publication Frameworks (Internationale Tagung), Friedrich-Schiller-Universität Jena. URL: [gw.uni-jena.de/digitalgender-history](http://gw.uni-jena.de/digitalgender-history) [13.09.2021]
- Flanders, Julia/Jannidis, Fotis (2019a). Data Modeling in a Digital Humanities Context. An Introduction. In Dies. (Hg.), *The Shape of Data in the Digital Humanities. Modeling Texts and Text-based Resources*. London et al.: Routledge, 3–25.
- Flanders, Julia/Jannidis, Fotis (Hg.) (2019b). *The Shape of Data in the Digital Humanities. Modeling Texts and Text-based Resources*. London et al.: Routledge.
- Garz, Jona T./Riettiens, Lilli (2023). –. In Sandra Hofhues/Konstanze Schütze (Hg.), *Doing Research*. Bielefeld: Transcript, 12–17.
- Gitelman, Lisa (Hg.) (2013). *Raw Data Is an Oxymoron*. Cambridge et al.: The MIT Press.
- Gledhill, Christine/Knight, Julia (2015). Introduction. In Dies. (Hg.), *Doing Women's Film History: Reframing Cinemas, Past and Future*. Urbana et al.: University of Illinois Press, 1–12.
- Haraway, Donna (1988). Situated Knowledges. The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist studies*, 14(3), 575–599.
- Hatch, Kristen (2013). Cutting Women. Margaret Booth and Hollywood's Pioneering Female Film Editors. In Jane M. Gaines/Radha Vatsal/Monica Dall'Asta (Hg.), *Women Film Pioneers Project*. New York: Columbia University Libraries.
- Hill, Erin (2016). *Never Done. A History of Women's Work in Media Production*. New Brunswick et al.: Rutgers University Press.
- Moeller, Katrin (2019). Standards für die Geschichtswissenschaft! Zu differenzierten Funktionen von Normdaten, Standards und Klassifikationen für die Geisteswissenschaften am Beispiel von Berufsklassifikationen. In Jana Kittelmann/Anne Purschwitz (Hg.), *Aufklärungsforschung digital Konzepte, Methoden, Perspektiven*. Halle: Mitteldeutscher Verlag, 17–43.
- Macgilchrist, Felicitas (2023). o.D. In Sandra Hofhues/Konstanze Schütze (Hg.), *Doing Research*. Bielefeld: Transcript, 312–319.
- Schäfer, Mirko T./van Es, Karin (Hg.) (2017). *The Datafied Society: Studying Culture through Data*. Amsterdam: Amsterdam University Press.
- Schäffer, Burkhard (2023). Abs. In Sandra Hofhues/Konstanze Schütze (Hg.), *Doing Research*. Bielefeld: Transcript, 72–81.
- Underwood, Ted (2014). Theorizing Research Practices We Forgot to Theorize Twenty Years Ago. *Representation*, 127(1), 64–72.
- vDHd 2021 *Experimente* (2021). Hypotheses. URL: [vdhd2021.hypotheses.org](http://vdhd2021.hypotheses.org) [13.09.2021]
- Vonderau, Patrick (2017). Quantitative Werkzeuge. In Malte Hagen/Volker Pantenburg (Hg.), *Handbuch Filmanalyse*. Wiesbaden: Springer, 1–15.
- Wilkinson, Leland (2005). *The Grammar of Graphics*. 2. Aufl. New York: Springer.
- Wolff, Christian (2015). The Case for Teaching “Tool Science”. Taking Software Engineering and Software Engineering Education beyond the Confinements of Traditional Software Development Contexts. In IEEE EDUCON (Hg.), *Proceedings of 2015 IEEE Global Engineering Education Conference*. Tallinn: IEEE, 932–938.
- Wreyford, Natalie/Cobb, Shelley (2017). Data and Responsibility-Toward a Feminist Methodology for Producing Historical Data on Women in the Contemporary UK Film Industry. *Feminist Media Histories*, 3(3), 107–132.



