

Data Feminism in DH: Hackathon und Netzwerktreffen

Lang, Sarah

sarah.lang@uni-graz.at
Universität Graz, Österreich

Borek, Luise

luise.borek@tu-darmstadt.de
Technische Universität Darmstadt, Deutschland

Probst, Nora

nora.probst@uni-koeln.de
Universität zu Köln, Deutschland

Einleitung

Das Ziel von *Data Feminism* ist es, ausgehend von Positionen des intersektionalen Feminismus die vorwiegend männlich dominierten Narrative in den Digital Humanities kritisch zu reflektieren und bislang marginalisierten Stimmen insbesondere von Frauen* mehr Sichtbarkeit zukommen zu lassen. In den Geisteswissenschaften ist häufig zu beobachten, dass sich der weiß, cis-männlich und hegemonial dominierte Bias aus den Archiven und Quellen in die erhobenen Daten und digitalen Technologien überträgt. Die vom *Data Feminism* angestrebte Balance im Sinne einer gendersensiblen Repräsentation ist ein wichtiger Baustein, um der Forderung nach einer grundlegenden Offenheit geisteswissenschaftlicher Forschung nachzukommen.

Data Feminism in den DH

Das 2020 von Catherine D'Ignazio und Lauren F. Klein veröffentlichte Buch *Data Feminism* (MIT Press) hat für die Digital Humanities eine hohe Relevanz, wendet sich allerdings eher an ein Publikum aus der Data Science. Zwar bieten Kleins Forschungen durchaus Schnittstellen zu Fragestellungen in den DH, doch lässt sich aus dem Buch keine unmittelbar anwendbare Anleitung für die Umsetzung von Data Feminism in den Digital Humanities generieren. Vielmehr handelt es sich um ein Manifest, das vor allem die Wichtig- und Dringlichkeit von intersektionalen, datenfeministischen Ansätzen deutlich macht und grundlegende Leitsätze zu deren Umsetzung formuliert. Diese Leitsätze auf konkrete Forschungsfragen oder Datenrepositorien aus den Digital Humanities zu übertragen, ist nach wie vor ein Desiderat.

Hier setzt unser Workshop an. Gemeinsam mit den Teilnehmenden soll eine Handreichung erarbeitet werden, die grundlegende Informationen und Leitlinien bezüglich folgender Fragen zusammenstellt:

1) Was lässt sich unter dem Begriff Data Feminism verstehen?

2) Wieso braucht man Data Feminism in den DH und inwiefern sollten dessen Ansätze in den bestehenden Diskursen der DH Berücksichtigung finden?

3) Wie können datenfeministische Projekte in den DH konkret aussehen? Inwiefern lassen sich Forschungsfelder definieren, die für Interessierte einen guten Ausgangspunkt bilden, um einige Ansätze des Data Feminism unmittelbar und praktisch umzusetzen?

Ein solche Handreichung sollte die Umsetzung datenfeministischer Ansätze in den DH insofern befördern, als sie die Einstiegshürden senkt und Interessierten diejenigen Informationen zur Verfügung stellt, die sie für einen effizienten, konkreten und praxisbezogenen Einstieg in Data Feminism benötigen.

Forschungsstand: Offene Fragen und Desiderate im Kontext des Data Feminism in den DH

Problemstellung 1: Data Gender Gap und das kulturelle Erbe

Beim Arbeiten mit Quellen zum kulturellen Erbe wird eine Vielzahl von Problemen offenbar, die bei vielen anderen Datensätzen (die beispielsweise auf statistischen Erhebungen und Befragungen beruhen) nicht bestehen und die sich genuin aus dem geisteswissenschaftlichen Forschungsgegenstand ergeben: So lassen sich Positionen des Data Feminism zum Teil gar nicht auf historische Daten anwenden, wenn beispielsweise Informationen über Frauen* oder andere Gruppen, die der hegemonialen Norm nicht entsprechen, gar nicht erst dokumentiert wurden (Mandell 2019, Lampe 2021, Rezai 2022). Nicht nur in den materiellen Beständen von vielen GLAM-Institutionen, sondern auch in den daraus hervorgegangenen Datenbeständen sind Frauen* unterrepräsentiert (Flanders 2018, Wiens et al. 2020). Diese Problematik eines Gender Data Gaps (Criado-Perez 2020) spiegelt sich auch in den verfügbaren Normdaten wider, da – etwa in der GND – zunächst vor allem publizierende Personen systematisch erfasst wurden.¹ Auch eine adäquate Verschlagwortung etwa in Bibliothekskatalogen stellt in diesem Zusammenhang eine große Herausforderung dar (vgl. Juen 2021).

Bevor also eine datenfeministische Analyse überhaupt beginnen kann, ist umfangreiche Erschließungsarbeit zu leisten und es ist je nach konkretem Anwendungsfall gegebenenfalls notwendig, bestehende Korpora um Quellen von oder über Frauen* zu erweitern. Ein Beispiel für solche Arbeiten, bei denen bestehende Datensätze mit bislang unterrepräsentierten Gruppen ergänzt werden, ist das Women Film Pioneers Project (WFPP) (<https://wfpp.columbia.edu/>, vgl. Wreyford/Cobb 2017, Dang 2020, Dickel et al. 2022, Gaines et al. 2022). In Themenfeldern, in denen Daten von Frauen* zwar vorhanden, aber noch tiefergehender erschlossen oder digitalisiert werden müssen, ist dies arbeitsaufwendig und wird erfordern, dass die DH auf Dauer ausreichend Energie

und Fördergelder in solche Tätigkeiten investieren. Wenn es also um Datenfeminismus und intersektionale Arbeit im Kontext der Digital Humanities geht, handelt es sich oftmals noch gar nicht um Datenanalysen, sondern vielmehr um erschließungsbezogene Tätigkeiten. Denn es gilt zunächst, die statistischen Ungleichgewichte der uns zur Verfügung stehenden Datensätze zum Thema zu machen und in diesem Sinne auf Lösungsstrategien hinzuwirken.

Eine weitere Herausforderung besteht hier ausgerechnet in der erstrebenswerten Offenheit von Daten, die inzwischen bei allen Förderinstitutionen und -richtlinien verlangt wird: Für marginalisierte und oftmals sensible Daten kann diese Offenheit aus Datenschutzgründen jedoch schnell zum Ausschlusskriterium werden. Hier müssen förderpolitisch Lösungen implementiert werden, die eine Erschließung und wissenschaftliche Nutzung ermöglichen, ohne den Datenschutz zu verletzen. Nicht zuletzt stellen Fälle, in denen Frauen* oder unterrepräsentierte Gruppen in Datensätzen unsichtbar, schwer sichtbar oder überhaupt nicht vorhanden sind, eine schwer zu lösende Aufgabe dar. Was ist mit Fällen, in denen es nicht möglich ist, diese Lücke durch neue Datenerhebungen zu füllen? Kann man unterrepräsentierte Daten aus bestehenden Datensätzen interpolieren?

Problemstellung 2: Modellierung, Kuratierung, Daten- und Korpuskritik

Positionen des Data Feminism halten bereits insofern Einzug in die DH, als in den vergangenen Jahren Hegemonie-kritische Forschungsthemen wie Gender-sensible Datenmodellierung, Korpuskritik, Decolonializing DH usw. stärkere Berücksichtigung finden (Risam 2015, Wernimont 2015, Koh/Stommel 2018, Losh/Wernimont 2018, Kim/Koh 2021, Guiliano/Heitmann 2019, Mandell 2019, Risam/Bordalejo 2019). Da ein feministischer und intersektional angelegter Umgang mit Daten in der Regel mit einer Analyse von Machtstrukturen beginnt, bieten die DH einen besonders geeigneten Nährboden für diesen Ansatz. Grundsätzlich gilt es nicht nur, dem Gender Data Gap durch die Erhebung von Daten zu Frauen* entgegenzuwirken, sondern es müssen auch die bestehenden Kategorien, Klassifizierungen und Datenmodelle auf Grundlage intersektionaler feministischer Perspektiven kritisch hinterfragt und ggf. komplexer gestaltet werden (Kyvernitou/Bikakis 2017) – dazu gehören Überlegungen zu Marginalisierungen an der Intersektion von Gender und rassistischen, klassistischen, ableistischen und weiteren Diskriminierungserfahrungen. Denn personenbezogenen Daten lassen sich in der Regel nicht durch Selbstbezeichnungen und/oder zusätzliche Befragungen ergänzen – Gender-bezogene Daten sind bei historischen Datensätzen in der Regel Zuschreibungen, die auf äußerlichen Merkmalen beruhen. In aller Regel erlauben diese Datensätze weder eine Unterscheidung zwischen biologischem Geschlecht und Gender, noch gehen sie über das binäre Geschlechtermodell hinaus. Data Feminism muss in den DH nicht nur den Data Gender Gap im Blick behalten und die Digitalisierung von bislang nicht digitalisierten Quellen forcieren, sondern auch bestehende Klassifizierungen, Datenmodelle und -kuratierungen kritisch hinterfragen.

Problemstellung 3: Anwendung außerhalb offensichtlicher Anwendungsfälle

Online finden sich eine ganze Reihe von Arbeiten, die sich des Data-Feminism-Begriffs bereits bedienen, doch handelt es sich häufig noch um Work-in-Progress, das noch nicht in Form von wissenschaftlichen Publikationen veröffentlicht ist (Bui/Gleißner/Kühn/Nenninger 2021, Keck 2021a, 2021b, Klein 2018, 2022). Zudem ergibt sich hier die Auseinandersetzung mit Data Feminism häufig unmittelbar aus einer thematischen Schwerpunktsetzung heraus. Wie aber könnte man mit Datensätzen oder Forschungsfragen umgehen, die nicht für Ansätze des Data Feminism prädestiniert scheinen, aber dennoch von diesen Ansätzen in hohem Maße profitieren würden? Um ein solches Paradigma zu etablieren, genügt es nicht, wenn sich lediglich einzelne Projekte für datenfeministische Perspektiven öffnen, bei denen der Nutzen des Data Feminism unmittelbar auf der Hand liegt. Um die Erkenntnisse der Intersektionalitäts- und Datenfeminismus-Forschung konsequent in den DH zu etablieren – wie unter anderem von Roopika Risam 2015 gefordert –, werden Leitlinien benötigt, die aufzeigen, wie diese auf eine größere Breite an Datensätzen und Forschungsfragen Anwendung finden können.

Problemstellung 4: Maschinelles Lernen als Bias-Verstärker oder Chance?

Zu fragen wäre schließlich, ob in diesem Kontext Algorithmen des maschinellen Lernens Abhilfe leisten könnten. Bisherige Publikationen zur Daten- und Korpuskritik in Bezug auf marginalisierte Gruppen halten sich diesbezüglich eher bedeckt und betonen, dass Algorithmen vor allem Bias der sie vornehmlich programmierenden (weißen cis-männlichen) Gruppen (Klinger/Svensson 2021) reproduzieren (GUILIANO/HEITMANN 2019).

Andere Publikationen verweisen allerdings darauf, dass die Aufgabe der DH in Bezug auf maschinelles Lernen gerade darin besteht, den Maschinen jenen "computer science gaze" abzutrainieren und stattdessen einen "curatorial gaze" zu etablieren (Bönisch 2021). Zu fragen wäre folglich, ob Algorithmen des maschinellen Lernens programmiert werden könnten, die trotz der in Problemstellung 1 und 2 skizzierten Herausforderungen einen "intersectional decolonialist gaze" auf Daten des kulturellen Erbes etablieren.

Aufbau des Workshops

Ziel dieses Workshops ist zum einen die Vernetzung aller Forschenden, die Positionen des Data Feminism bereits in ihren Forschungen berücksichtigen oder perspektivisch berücksichtigen wollen. Zum anderen geht es um die Erarbeitung konkreter Umsetzungsstrategien, wie sich die Forderungen des Data Feminism im Kontext der DH realisieren lassen. Der Workshop nennt sich Hackathon, um die kollektive Arbeitsform im Sinne einer gemeinsamen Daten-bezogenen Arbeitspraxis zu betonen. Zu Anfang werden zunächst die grundlegen-

den Konzepte des Data Feminism vorgestellt, um den Einstieg auch Personen ohne Vorerfahrungen zu ermöglichen. Im Anschluss können Teilnehmende eigene Arbeiten kurz vorstellen und ihren Zugriff auf Data Feminism im Kontext der DH erläutern. Aus diesen Inputs ergibt sich dann ein Austausch in Kleingruppen. Aufbauend auf den Erläuterungen und Diskussionen zu Beginn erarbeiten die Teilnehmenden in der Folge gemeinsam eine Handreichung, die auf wenigen Seiten grundlegende Informationen zum praktischen datenfeministischen Arbeiten in den Digital Humanities vermittelt und konkrete Leitlinien enthält. Auf diese Weise können sich Interessierte durch die Lektüre grundlegende Kenntnisse des Data Feminism selbstständig aneignen. Wir hoffen, dass damit die Einstiegshürde in datenfeministische DH-Forschung gesenkt und Data Feminism zukünftig in den DH einen breiteren Rückhalt finden wird. Die Handreichung soll idealerweise auf Deutsch und Englisch herausgegeben werden. Um all diese Aspekte (Einführung in grundlegende Konzepte, Kurzvorstellung eigener Arbeiten, Kleingruppendiskussion sowie Erstellung der Handreichung) zeitlich abdecken zu können, werden zwei Halbtage für den Workshop veranschlagt.

Organisationsteam

Sarah Lang, Universität Graz, sarah.lang@uni-graz.at, <https://orcid.org/0000-0002-4618-9481>. Forschungsinteressen/Hintergrund: Digital Humanities PostDoc, Wissenschaftsgeschichte (Alchemie).

Luise Borek, Technische Universität Darmstadt, luise.borek@tu-darmstadt.de, 0000-0001-5849-374X. Forschungsinteressen/Hintergrund: PostDoc (TU Darmstadt) Vertretungsprofessur Digital Humanities (Universität Graz); germanistische Mediävistik. DFG-Netzwerk Offenes Mittelalter.

Nora Probst, Universität zu Köln, Deutschland, nora.probst@uni-koeln.de, 0000-0001-6932-0879. Forschungsinteressen/Hintergrund: Digitale Kulturwissenschaften, Vertretungsprofessorin der Professur "Kulturen der Digitalität".

Alle Einreichenden verbindet durch ihre Aktivitäten in der AG Empowerment das Interesse an Korpuskritik und Datenfeminismus.

Fußnoten

1. Criado-Perez' Buch *Invisible Women* hat in den letzten zwei Jahren viel Aufmerksamkeit für das Thema generiert. Wir distanzieren uns allerdings ausdrücklich von dem in der Studie vertretenen binären Geschlechtermodell, das Personen jenseits der Cis-Geschlechtlichkeit vollständig ignoriert, und vertreten unsererseits einen intersektionalen, queerfeministischen und inklusiven Ansatz nach D'Ignazio/Klein 2020.

Bibliographie

Aleksander, Karin. 2014. "Die Frau im Bibliothekskatalog." *LIBREAS. Library Ideas* 25 (2014). [https://li-](https://li-bre-as.eu/ausgabe25/02alexander/)

[breas.eu/ausgabe25/02alexander/](https://li-bre-as.eu/ausgabe25/02alexander/) (zuletzt zugegriffen 02.08.2022).

Bönisch, Dominik. 2021. "The Curator's Machine: Clustering of Museum Collection Data through Annotation of Hidden Connection Patterns between Artworks". *International Journal for Digital Art History* 5 (Mai 2021): 5.20-5.35. <https://journals.ub.uni-heidelberg.de/index.php/dah/article/view/75953> (zuletzt zugegriffen 02.08.2022).

Bui, Magdalena, Lea Gleißner, Fey Kühn und Amelie Nenninger. 2021. "Questioning Street Names Leipzig: wie Genderbias die Straßenbenennung in Leipzig beeinflusst und wie die DH helfen können Bias sichtbar zu machen." In: *Blog Public Humanities in den Digital Humanities*. 23.08.2021. <https://publicdh.hypotheses.org/282> (zuletzt zugegriffen 02.08.2022).

Criado-Perez, Caroline. 2020. *Invisible Women: Exposing Data Bias in a World Designed for Men*. London: Chatto & Windus.

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://www.digitalhumanities.org/dhq/vol/14/4/000528/000528.html>.

Dickel, Henri, Matija Miskovic, Kharazm Nori, Christian Schmidt, Atefeh Soltanifard, Sarah-Mai Dang und Thorsten Thormählen. *Women Film Pioneers Explorer*. <https://www.online.unimarburg.de/women-film-pioneers-explorer/> (zuletzt zugegriffen 02.08.2022).

D'Ignazio, Catherine und Lauren Klein. 2020. *Data Feminism*. Cambridge/MA: MIT Press. <https://datafeminism.io/> (zuletzt zugegriffen 02.08.2022).

Flanders, Julia. 2018. "Building Otherwise." In: *Bodies of Information: Intersectional Feminism and the Digital Humanities*, herausgegeben von Elizabeth Losh und Jacqueline Wernimont, 289-304. Minneapolis/MN: University of Minnesota Press. *Debates in the Digital Humanities*. DOI: <https://doi.org/10.5749/j.ctv9hj9r9> (zuletzt zugegriffen 02.08.2022).

Gaines, Jane, Radha Vatsal, and Monica Dall'Asta (eds.). 2022. *Women Film Pioneers Project*. New York, NY: Columbia University Libraries. <https://wfpp.columbia.edu/> (zuletzt zugegriffen 02.08.2022).

Guiliano, Jennifer und Carolyn Heitman (2019): „Difficult Heritage and the Complexities of Indigenous Data.“ *Journal of Cultural Analytics* 4/1: 1-25. doi:10.22148/16.044.

Juen, Sara. 2021. "Feminismus, Algorithmen, Gender-Data-Gap und was das alles mit Bibliotheks- und Informationswissenschaft zu tun hat." *LIBREAS. Library Ideas* 39 (2021). <https://doi.org/10.18452/23448> (zuletzt zugegriffen 02.08.2022).

Keck, Jana. 2021a. "Text Mining America's German-Language Newspapers, 1830-1914: Processing Ger(wo)manness." Historikertag 2021, Peter Haber Preis für digitale Geschichte. <https://doi.org/10.5281/zenodo.5518019> (zuletzt zugegriffen 02.08.2022).

Keck, Jana. 2021b. "A Data Feminist Approach to Studying the C19 Social Network of German-Americans." *Global Digital Humanities Symposium 2021*. <https://www.youtube.com/watch?v=zrg1htGXRIA> (zuletzt zugegriffen 02.08.2022).

Klein, Lauren. 2018. "Data feminism: Community, allyship, and action in the digital humanities" (Keynote address). Digital Frontiers Annual Conference, Lawrence, KS, United States. <https://digital.library.txstate.edu/handle/10877/7839> (zuletzt zugegriffen 02.08.2022).

Klein, Lauren. 2022. "Data Feminism and Digital Humanities" (Vortrag). <https://www.youtube.com/watch?v=W0slTS8a6YY> (zuletzt zugegriffen 02.08.2022).

Klinger, Ulrike und Jakob Svensson. 2021. "The power of code: women and the making of the digital world." *Information, Communication & Society* 24 (14): 2075-2090. DOI: 10.1080/1369118X.2021.1962947.

Kyvernitou, Ionna und Antonis Bikakis. 2017. "An Ontology for Gendered Content Representation of Cultural Heritage Artefacts." *Digital Humanities Quarterly* 11/3 (2017). <http://www.digitalhumanities.org/dhq/vol/11/3/000316/000316.html> (zuletzt zugegriffen 02.08.2022).

Lampe, Moritz. 2021. *Diskriminierende Begriffe und Wissensordnungen im Bildarchiv. Eine postkoloniale Perspektive am Beispiel des 'Bildindex der Kunst und Architektur'*. Berliner Handreichung zur Bibliotheks- und Informationswissenschaft 481. Berlin. DOI: 10.18452/23766 (zuletzt zugegriffen 02.08.2022).

Losh, Elizabeth und Jacqueline Wernimont. 2018. *Bodies of Information: Intersectional Feminism and the Digital Humanities*. Minneapolis/MN: University of Minnesota Press. Debates in the Digital Humanities. DOI: <https://doi.org/10.5749/j.ctv9hj9r9> (zuletzt zugegriffen 02.08.2022).

Mandell, Laura. 2019. "Gender and Cultural Analytics: Finding or Making Stereotypes?" In: *Debates in the Digital Humanities* 2019, herausgegeben von Matthew K. Gold and Lauren F. Klein. University of Minnesota Press. DOI: <http://dx.doi.org/10.5749/j.ctvg251hk.4> (zuletzt zugegriffen 02.08.2022).

Rezai, Yasami. 2022. "Data Stories for/from All: Why Data Feminism is for Everyone." *Digital Humanities Quarterly* 16/2 (2022). <http://www.digitalhumanities.org/dhq/vol/16/2/000618/000618.html> (zuletzt zugegriffen 02.08.2022).

Risam, Roopika (2015). "Beyond the Margins: Intersectionality and the Digital Humanities." *Digital Humanities Quarterly* 9(2). <http://www.digitalhumanities.org/dhq/vol/9/2/000208/000208.html> (zuletzt zugegriffen 02.08.2022).

Risam, Roopika und Barbara Bordalejo (eds.). 2019. *Intersectionality in Digital Humanities*. Amsterdam: Amsterdam University Press. DOI: <https://doi.org/10.1017/9781641890519> (zuletzt zugegriffen 02.08.2022).

Wiens, Brianna, Stan Ruecker, Jennifer Roberts-Smith, Milena Radzikowska und Shana MacDonald. 2020. "Materializing Data: New Research Methods for Feminist Digital Humanities." *Digital Studies/le Champ Numérique* 10(1). DOI: <https://doi.org/10.16995/dscn.373> (zuletzt zugegriffen 02.08.2022).

Wernimont, Jacqueline (ed.). 2015. *Digital Humanities Quarterly* 9.2 (2015), Special Issue *Feminisms in Digital Humanities*. <http://www.digitalhumanities.org/dhq/vol/9/2/index.html> (zuletzt zugegriffen 02.08.2022).

Wreyford, Natalie, und Shelley Cobb. 2017. "Data and Responsibility: Toward a Feminist Methodology for Pro-

ducing Historical Data on Women in the Contemporary UK Film Industry". *Feminist Media Histories* 3/3 (2017): 107-132. doi:10.1525/fmh.2017.3.3.107.