HAL open science

Vernetzung-Erzählung-Kollation. Digitale Methoden in der Biographieforschung

Anne Baillot, Anna Busch

To cite this version:

Anne Baillot, Anna Busch. Vernetzung-Erzählung-Kollation. Digitale Methoden in der Biographieforschung. BIOS – Zeitschrift für Biographieforschung, Oral History und Lebensverlaufsanalysen, 2019, 30 (1-2/2017), pp.22-29. 10.3224/bios.v30i1-2.01 . hal-02455253

HAL Id: hal-02455253

https://hal.science/hal-02455253v1

Submitted on 25 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Vernetzung - Erzählung - Kollation. Digitale Methoden in der Biographieforschung

Anne Baillot, Anna Busch

I.

Die grundsätzliche Veränderung im Umgang mit lebensgeschichtlichen Quellen im Zuge der Digitalisierung und die Möglichkeiten, die digitale Auswertungsszenarien mit sich bringen, haben zu zahlreichen neuen Fragestellungen und Herausforderungen – auch im Bereich der biographischen Forschung – geführt. Drängend ist die Beantwortung von Rechtsfragen hinsichtlich der Publikation, Verwendung und Weiterverarbeitung personenbezogener Daten, die im digitalen Rahmen neu zu stellen sind und zu Unsicherheiten führen. Etwa: Dürfen biographischen Daten im Netz verbreitet, verwendet, gespeichert, bearbeitet und weitergeben werden und wenn ja, unter welchen Bedingungen und wie ist die Quelle zu nennen?

Eine weitere Herausforderung, der sich Forschende und Nutzer_innen digital zugänglicher Informationen in Zukunft vermehrt werden stellen müssen, ist in der Heterogenität von Datenbeständen und der Verwendung dieser Datenbestände und dem Zugriff auf sie zu sehen. Gerade in Zeiten digitaler Goldgräberstimmung und unter der Maxime „alles muss digital werden“, kommt es entscheidend auf die Verwendung von Standards und Normen zur Erfassung von Daten an.

Die Frage nach dem Platz von biographischer Forschung in den Digital Humanities hängt aber nicht zuletzt mit dem grundsätzlichen Verständnis von Digital Humanities zusammen. Legt man den Fokus auf die algorithmische Analyse großer Datenmengen, ist biographische Forschung sicherlich – zumindest heute – nicht das naheliegendste Feld, mit dem sich die Digital Humanities beschäftigen. Das hat zum einen mit den oben genannten rechtlichen Einschränkungen, zum anderen auch mit der mangelnden Homogenität der Daten zu tun. Drei darüberhinausgehende Zusammenhänge können aber zu einer ergiebigen Verzahnung der beiden Felder führen. Erstens und selbst bei heterogener Datenlage können Digital Humanities-Tools es ermöglichen, insbesondere im Bereich der Netzwerkvisualisierung, Biographien ergiebiger in größere Kontext einzubetten und zu analysieren. Zweitens gibt es über die reine Visualisierung hinaus weitere digitale Methoden, die die biographische Forschung dezidiert und im Sinne einer Verbesserung der Datengrundlage fördern können. Das betrifft unterschiedliche Methoden in den Bereichen der Vernetzung und der Kollation. Drittens aber ist der epistemologische Beitrag des digitalen Wandels ein Punkt, der insbesondere in der

biographischen Forschung zukünftig zentral sein wird. Durch die digitalen Methoden sind Forscher_innen gezwungen, eigene Position gegenüber ihren Forschungsobjekten und denjenigen, die ebenfalls mit diesen Objekten arbeiten, zu überdenken. Diese Herausforderung, die es für Forscher_innen nötig macht, sich neu zu positionieren, ist eine, der sich auch die biographische Forschung stellen kann, um sich im Feld der Digital Humanities zu verorten.

Vernetzung - Erzählung - Kollation: Mit diesen drei Schlagworten sind Bereiche umrissen, in denen im Rahmen von Biographieforschung digitale Methoden zum Einsatz kommen können. Im Folgenden soll anhand dieser drei Felder exemplarisch angedeutet werden, wie geisteswissenschaftliche Fragestellungen anhand digitaler biographischer Daten beantwortet werden können. Zuvörderst werden biographische Metadaten auf ihr Potential für eine zukunftsweisende Vernetzungsstrategie hin überprüft. Im Anschluss werden Fragen der Entwicklung von biographischen Narrativen aufgeworfen. In einem dritten Schritt werden Kollationstools, wie sie bei der Analyse von Biographien verwendet werden, dargestellt, um das Potential aufzudecken, das auch in diesem Ansatz für biographische Forschung zu finden ist.

1. Vernetzung

Die Möglichkeit zur Vernetzung von Informationen stellt in mancherlei Hinsicht den größten Mehrwert ihrer digitalen Aufbereitung dar. Im Gegensatz zu ihrer rein visuellen Aufbereitung in Form eines digitalen Textes (ein Buch liest sich nach wie vor angenehmer als ein Text auf einem Bildschirm), ist der eigentliche Gewinn in der Aggregation von Daten zu sehen. Wenn es also nicht darum geht, linear zu lesen, sondern Informationen zusammenzuzutragen, hat sich der Einsatz von digital basierten Suchmaschinen bewährt. Die Auswertung einzelner Bücher, ihrer Inhaltsverzeichnisse, Personen-, Orts- und Werkregister ist ungleich mühseliger als eine gezielte digitale Suchanfrage.

Im Kontext biographischer Forschung stellt sich daher die Frage, was für Informationen sinnvollerweise miteinander in Verbindung gebracht werden sollen und in welcher Form dies geschehen kann, damit das Feld davon profitiert. An dieser Stelle liegt für Forscher_innen die erste Herausforderung, denn die Informationen, die vorrangig dazu geeignet sind, miteinander in Beziehung gesetzt zu werden, sind Metadaten. Wenn beispielsweise das Geburtsdatum einer bekannten Persönlichkeit überprüft werden soll, dann muss nach anderen Quellen für das Geburtsdatum dieser Person gesucht werden, dann wird verglichen und entsprechend begründet, warum dieses eine Geburtsdatum als das wahrscheinlichste zu gelten hat. Während

der Prozess der Festlegung und Begründung dieses Geburtsdatums eindeutig in das wissenschaftliche Arbeitsfeld der biographischen Forschung fällt, ist das Festhalten, Nachweisen oder Verzeichnen eines Geburtsdatums an sich nicht Aufgabe der Wissenschaft, sondern der Bibliotheken und Archive. Die Wichtigkeit der biographischen Metadaten in der digitalen Handhabung der biographischen Forschung macht eine Annäherung zwischen Forschung und Kulturerbeeinrichtungen auf einer Ebene und in einem Ausmaß notwendig, die es bislang so noch nicht gegeben hat. Das zwingt schon per se zu einem Umdenken.

Dies betrifft insbesondere die Zusammenarbeit der Digital Humanities mit der Gemeinsamen Normdatei (GND)[1], die schon seit mehreren Jahren immer wieder thematisiert wird. Digital Humanities-Projekte (z.B. Editionen, Datenbanken), die mit Blick auf eine größtmögliche Vernetzung ihre Forschungsarbeit konzipiert sind, bemühen sich um Interoperabilität, d.h. sie fügen in ihren Quellcode Schnittstellen zu anderen digitalen Ressourcen ein. Diese Schnittstellen lassen sich am ehesten durch die Verwendung etablierter Standards erstellen: ISO-Standards für Daten[2] sind eine davon, Authority Files[3] eine andere. Im deutschsprachigen Raum werden biographische Authority Files vorrangig von der Deutschen Nationalbibliothek verwaltet. Einige andere Kulturerbeeinrichtungen haben ebenfalls die Möglichkeit, Einträge anzulegen oder zu ergänzen, so beispielsweise das Deutsche Literaturarchiv Marbach (DLA), sowie immer mehr Forschungsprojekte, wenn sie eine einschlägige Schulung absolvieren. So werden biographische Einträge – meist zunächst mit minimalen Informationen – auf bibliothekarischer bzw. archivarischer Ebene angelegt, die dann in der Forschung zum Einsatz kommen. Das ist die Basis von der heutige Forschung ausgehen kann. Sie bringt sowohl Vernetzungspotenzial als auch Herausforderungen mit sich.

Herausforderungen stellen Datensatz-Doubletten dar. Bei einer ganzen Reihe von Personenaufnahmen existieren in der Gemeinsamen Normdatei zwei oder mehr biographische

---

[1] http://www.dnb.de/DE/Standardisierung/GND/gnd_node.html „Die Gemeinsame Normdatei (GND) ist eine Normdatei für Personen, Körperschaften, Konferenzen, Geografika, Sachschlagwörter und Werktitel, die vor allem zur Katalogisierung von Literatur in Bibliotheken dient, zunehmend aber auch von Archiven, Museen, Projekten und in Webanwendungen genutzt wird. Sie wird von der Deutschen Nationalbibliothek, allen deutschsprachigen Bibliotheksverbünden mit den angeschlossenen Bibliotheken, der Zeitschriftendatenbank (ZDB) und zahlreichen weiteren Einrichtungen gemeinschaftlich geführt. […] Normdaten erleichtern die Katalogisierung, bieten eindeutige Sucheinstiege und die Möglichkeit der Vernetzung unterschiedlicher Informationsressourcen. In den Normdatensätzen der GND sind nicht nur die Vorzugsbenennungen normiert, sondern neben abweichenden Benennungen auch Relationen zu anderen Normdatensätzen enthalten. Auf diese Weise ist ein Netz von miteinander in Beziehung stehenden Datensätzen entstanden, das sich besonders für die Nutzung im Web eignet, die Navigation innerhalb der Normdatei erlaubt und somit die Recherchemöglichkeiten für Nutzer verbessert. […] Normdatensätze der GND sind Bestandteil des Virtual International Authority File (VIAF) und werden dort mit den Daten anderer nationaler Normdateien zusammengeführt.“

[2] https://www.iso.org/iso-8601-date-and-time-format.html

[3] https://viaf.org/

Einträge, die in Konkurrenz zueinander stehen. Bei diesen Einträgen stellt sich die Frage nach Einschlägigkeit und Relevanz. Welche GND-Nummer ist heranzuziehen und auf welche ist zu verweisen? Ebenso – und das ist eine weitere Herausforderung – können sich in der Datenaufnahme Fehler finden oder es existiert mitunter zu einer bestimmen Person, die man auszeichnen möchte, gar kein Eintrag.

Bei der Handhabung von biographischen Metadaten kommt es zu einem Autoritätskonflikt, den es vor der digitalen Ära in dieser Form nicht gab. Der Austausch von biographischen Metadaten ließe sich durchaus effizienter gestalten, als er zu den Zeiten war, in denen erst die Rücksprache mit dem zuständigen Archivar möglicherweise eine Verbesserung oder Änderung der lokal vorliegenden Metadaten zur Folge hatte. Die Übernahme in einen regionalen oder gar in einen überregionalen Katalog war damit nicht notgedrungen verbunden. Eine zentrale Herausforderung im Umgang mit digitalen biographischen Forschungsdaten stellt diese Kette von Verantwortlichkeit für biographische Informationen dar: Ihre Erfassung, ihre Übermittlung und ihre Überprüfung müssen der Geschwindigkeit und Durchlässigkeit angepasst werden, die den wissenschaftlichen Alltag auszeichnen. Das ist unbefriedigend für alle Beteiligten: Die DNB ist personell nicht so umfangreich aufgestellt, dass jeder einzelne Eintrag überprüft werden kann, die Durchdringung der kompletten Systematik kann von Wissenschaftler_innen nicht verlangt werden und die bestandshaltenden Institutionen fühlen sich entmachtet. Für jedes Projekt, das mit biographischen Metadaten arbeitet, ist eine wohldurchdachte Strategie notwendig, um jedem Partner die Autorität und den Spielraum zu sichern, den er benötigt. Sonst sind mehr Doubletten, redundante Datenbänke, und letztendlich immer weniger zuverlässige biographische Metadaten die Folge. Hier ist auch die Deutsche Forschungsgemeinschaft in der Pflicht: Jedes DFG-geförderte Projekt, das biographische Daten in einer wissenschaftlichen Form zusammenstellt, könnte beispielsweise eine anteilige DNB-Pauschale erhalten, damit von der DNB aus die Koordination und der Informationsfluss in die GND hinein mit hinreichend Personal besetzt ist und jedes Projekt einzeln betreut werden kann. Eine solche Investition wäre im Vergleich zur heutigen Tendenz zu Redundanzen und Doubletten mehr als sinnvoll und auf längere Sicht zielführend.

Eine weitere Herausforderung der Arbeit mit biographischen Metadaten ist die soziokulturelle Beschränkung, die historisch von der DNB ausgeht. Da die Gemeinsame Normdatei ursprünglich auf der Grundlage von Bibliothekskatalogen entwickelt wurde, sind in den biographischen Einträgen Schriftsteller überproportional vertreten. Dies gilt sogar innerhalb des dort vorrangig vertretenen Bereichs der Buchproduktion: Auch Verleger oder Schriftstellerinnen sind im Vergleich zu Schriftstellern unterrepräsentiert. Aufgabe der

biographischen Forschung müsste es sein, einer solchen Verzerrung entgegenzuwirken, die im Prinzip kein Korrektiv zur (literarischen bzw. ideengeschichtlichen) Geschichtsschreibung ermöglicht, wie sie seit dem 19. Jahrhundert betrieben wird.

Mit qualitativ hochwertigen biographischen Metadaten wäre ein Schatz gewonnen, der biographische Forschung im Herzen der ideen- und literaturgeschichtlichen Forschung neu aufblühen lassen könnte. Das Potenzial, das mit diesen Möglichkeiten zusammenhängt, ist zu groß, um es nicht auszuschöpfen. Das Vernetzungspotential, das die GND-Nummern ermöglichen, lässt sich an einem einfachen Beispiel zeigen: Das BEACON[4] ist eine ganz einfache Plattform, die von Wikipedia initiiert wurde und an der die größten Kulturerbeeinrichtungen schon länger beteiligt sind. Das Prinzip ist einfach: Jeder (Forscher, Kulturerbeeinrichtung, Hobbygenealoge) speist über die GND-Nummern die Informationen über eine bestimmte Person ein, über die er verfügt. Diese Informationen sind im Anschluss von allen anderen Projekten, die ebenfalls mit GND-Nummern arbeiten und am BEACON teilnehmen, abrufbar. Das Wikipedia-BEACON ist aus zwei Gründen besonders relevant. Zum einen, weil es besonders einfach konzipiert ist und den Benutzer/die Benutzerin vor keine besonders hohen technischen Hürden stellt. Insofern ist es der beste Beweis für die Effizienz von Standards und ihrer Umsetzung im Sinne des Datenaustauschs. Zum anderen ist das BEACON deswegen reizvoll, weil die Informationen immer ihrer Quelle zuordenbar sind. So kann der Nutzer/die Nutzerin, der/die mit den biographischen Daten arbeitet oder sie mit den eigenen vergleicht, für sich entscheiden, ob er/sie lieber dem Wikipedia-Eintrag oder dem Deutsche Biographie-Eintrag folgen möchte. Dazu ist anzumerken, dass es oft gerade im Fall weniger bekannter Personen die Wikipedia-Einträge sind, die Informationen vorhalten. Grundsätzlich ist der BEACON-Ansatz vor allem deswegen sinnvoll, weil er den Autoritätskonflikt umgeht und die Zitierfähigkeit durch den Verweis auf die jeweilige online-Quelle elegant löst. Zwar muss man sich durch unterschiedliche Quellen klicken, um an die gewünschten Quellen und Informationen zu gelangen, aber im Prinzip verfügt er als Crowdsourcingverfahren über zuverlässige Autoritäts- und Qualitätskontrollemechanismen, die für jede Weiterentwicklung die bestmögliche Grundlage sind.

Tatsächlich sind hinsichtlich biographischer Metadaten nicht nur rein technische Fragen relevant, sondern zentral ist der Aufbau eines Forschungsökosystems, das für die Zukunft der biographischen Forschung mit Sicherheit entscheidend sein wird.

---

[4] https://de.wikipedia.org/wiki/Wikipedia:BEACON

2. Das Narrativ

Über die bibliothekarisch-archivarischen Metadaten hinaus, die als Einstieg- und Vernetzungspunkt in der biographischen Forschung zum Einsatz kommen können, ist des Weiteren relevant, wie die Digital Humanities mit der Entwicklung von biographischen Narrativen umgehen.

Der Impuls, der in den ersten Jahren der Entwicklung der Digital Humanities von der Gemeinschaft der amerikanischen Klassischen Philologie kam, hat dazu geführt, dass sich biographische und prosopographische Datenbanken entwickelt haben, die allerdings den heutigen Standards nicht immer entsprechen.[5] Solche Vorhaben davon zu überzeugen, TEI-Standards zu verwenden, um interoperabel, d.h. mit anderen Datenbanken vernetzbar zu sein, ist nicht immer machbar. Eine größere institutionelle Unterstützung, die notwendig wäre, um sich in diese Zusammenhänge einzuarbeiten, fehlt mitunter. Nach wie vor wird die Arbeit an einer Datenbank als reine Zusatzleistung betrachtet, nicht zuletzt, weil die wissenschaftlichen Evaluationsmechanismen von Publikationen die Komplexität und den Gehalt der betroffenen Datenmodelle nicht im gebührenden Maße berücksichtigen. Für die prosopographische Forschung wäre eine Stellungnahme der einschlägigen Verbände und gelehrten Gesellschaften zu verbindlichen Standards und Normen ein wichtiges Zeichen, nicht zuletzt damit sich Nachwuchswissenschaftler_innen daran wagen, neue, interoperable, nachhaltige prosopographischen Datenmodelle zu entwickeln.

Das biographische Narrativ konstituiert sich jedoch nicht ausschließlich über prosopographische Arbeiten, sondern auch über andere Annäherungen an die Rekonstruktion von biographischen Einheiten. Auch da bietet das Digitale Raum zur Kreativität – einen Raum, der allerdings noch wenig genutzt wird. Um eine Einzelbiographie zu erarbeiten, ist es üblich, die betroffene Person in den Mittelpunkt zu stellen. So entstehen auch online beispielsweise ein Heinrich-Heine-Portal,[6] ein E.T.A. Hoffmann-Portal[7] etc. Dennoch würde es der Theorie der digitalen Methodik entsprechen, diese Form der biographischen Narrativbildung durch einen Netzwerkansatz zu ersetzen, d.h. Biographien von vornherein in ihrer gegenseitigen Verzahnung zu denken. Editionen und Portale haben es nicht nötig, figurenzentriert zu sein, sondern sie können sich ebenfalls um Konzepte einer Biographie ranken, die die Knotenpunkte

---

[5] Zu nennen wäre hier die Arbeit der Klassischen Philologie an der TUFTS University (vgl. bsp. http://ase.tufts.edu/classics/researchguide/index_files/Page603.htm).
[6] http://www.hhp.uni-trier.de/
[7] http://etahoffmann.staatsbibliothek-berlin.de/

in den Mittelpunkt stellen und gleichsam das Intertextuelle des biographischen Narrativs herausarbeiten. Entsprechendes wurde mit der Edition *Briefe und Texte aus dem intellektuellen Berlin um 1800* versucht.[8] Grundlegend – und hinsichtlich digitaler Editionen neu – war hier die Ausgangsidee, dass der Zugang zu literatur- und wissenschaftsgeschichtlichem Wissen im Netz nicht notgedrungen über einen bestimmten Autor oder ein einzelnes Werk zu erfolgen hat. Die einseitige Autorzentriertheit, die den gängigen Druckeditionen zu Grunde liegt, wurde zugunsten unterschiedlicher Annäherungsmöglichkeiten aufgebrochen. Diese konzeptionell entscheidenden Überlegungen ermöglichen dem Benutzer den Einstieg in die Edition – gleichwertig neben dem Zugang über den Autor – über eine Forschungsfrage, einen Zeitpunkt oder eine Zeitspanne, einen bestimmten Ort oder eine Textgattung.[9] An dieser Stelle müssen in Zukunft verstärkt auch theoretische Ansätze der Biographieforschung weitergedacht werden.

3. DH-Tools zur Analyse von Biographien

Durch digitale Untersuchungsmethoden verändert sich die Herangehensweise an biographische Forschung auch im Bereich der Textanalyse. Besonders interessant sind in diesem Zusammenhang Kollationstools, die es erlauben automatisierte Textvergleiche vorzunehmen. Es geht um maschinengestützte Möglichkeiten des Vergleichs von zwei oder mehreren sich unterscheidenden Versionen desselben Textes mit dem Ziel Unterschiede, aber – quasi als Nebenprodukt – auch Übereinstimmungen zu identifizieren. Es lassen sich mit solchen Tools Beziehungen bzw. Abhängigkeiten verschiedener Fassungen eines Textes rekonstruieren und je nach Tool auch visualisieren. Der Bereich, in dem solche Ansätze häufig genutzt werden, ist der der Plagiatsuntersuchung.[10] Zahlreiche Tools, die ähnlich konzipiert sind, werden aber auch im Bereich der Digital Humanities verwendet werden. Bereits in den 1960er Jahren gab es mit

---

[8] http://www.berliner-intellektuelle.eu/ „Diese Edition versammelt Texte verschiedener Autoren und unterschiedlicher Gattungen, die Eines gemeinsam haben: Sie beleuchten auf einprägsame Weise das intellektuelle Leben im Berlin des späten 18. und frühen 19. Jahrhunderts. Über die unterschiedlichen Einstiegsmöglichkeiten (Autoren, Gattungen, Themen, Zeitperioden, Suchfunktion) werden Einblicke in die Entstehungsgeschichte romantischer Literatur ermöglicht, zusammen mit Einsichten in Ideen- und Kulturtransfer, der sich in einer politisch und literarisch turbulenten Zeit in der preußischen Hauptstadt beobachten lässt.“ Grundlegend – und hinsichtlich digitaler Editionen neu – war die Ausgangsidee, dass der Zugang zu literatur- und wissenschaftsgeschichtlichem Wissen im Netz nicht notgedrungen über einen bestimmten Autor oder ein einzelnes Werk zu erfolgen hat. Die einseitige Autorzentriertheit, die den gängigen Druckeditionen zu Grunde liegt, wurde zugunsten unterschiedlicher Annäherungsmöglichkeiten aufgebrochen. Diese konzeptionell entscheidenden Überlegungen ermöglichen dem Benutzer den Einstieg in die Edition – gleichwertig neben dem Zugang über den Autor – über eine Forschungsfrage, einen Zeitpunkt oder eine Zeitspanne, einen bestimmten Ort oder eine Textgattung.

[9] Vgl. hierzu detaillierter: Anne Baillot, Anna Busch: „Berliner Intellektuelle um 1800“ als Programm. Über Potential und Grenzen digitalen Edierens, in: literaturkritik.de, Nr. 9, September 2014, Schwerpunkt: Romantik digital, unter: http://www.literaturkritik.de/public/rezension.php?rez_id=19678 (veröffentlicht: 1.9.2014)

[10] Beispielhaft seien genannt: CheckText.org, WriteCheck, PlagiarismFinder.

Collatio Codicum: An Exercise in COMIT Programming entsprechende Entwicklungen für automatisierte Textvergleiche. In Deutschland folgte dann in den 1970ern die Entwicklung von *TUSTEP*, dem Tübingen System von Textverarbeitungsprogrammen zur wissenschaftliche Textdatenverarbeitung innerhalb der Geisteswissenschaften. Besondere Verwendung hat *TUSTEP* im Bereich der Editionswissenschaft zur Erstellung textkritischer Ausgaben und dem Vergleich verschiedener Textfassungen zur automatischen Generierung der Apparate erfahren. Bei den zwei zurzeit regelmäßig verwendeten Open Source Tools, die besonders im Hinblick auf die Bedürfnisse von Geisteswissenschaftlern hin entwickelt worden sind, sind *Collate* (entwickelt von Peter Robinson) mit seiner Weiterentwicklung *CollateX*[11] sowie *Juxta*[12] zu nennen. Während *CollateX* zwei oder mehr kürzere Textversionen miteinander vergleichen kann, Variantengraphen (nach dem Modell von Schmidt 2009) generiert, parallele Ansicht der Textversionen (zur linear, tabellarisch, optional segmentiert) Verfügung stellt und (unter Umständen) Transpositionen erkennt, unterstützt *Juxta* den Vergleich von Textversionen als plain text sowie in beliebigem XML-Format (inkl. TEI-P5), visualisiert die Ergebnisse in verschiedenen Ansichten und ermöglicht HTML-Exporte der Ergebnisse als kritischer Apparat.

Mit den genannten Tools ist es auch im Bereich der Textanalyse möglich, der Textgenese von Biographien nachzuspüren. Biographien sind als Forschungsfeld in diesem Zusammenhang besonders ergiebig, da sie sich in der Regel aus unterschiedlichen Textzeugen zusammensetzen, bzw. unterschiedliches Quellenmaterial heranziehen und im Rahmen der Texterstellung verwerten. So fließen Briefe und Briefauszüge, Notizen, Tagebucheinträge, Mitteilungen von Freunden etc. in Biographietexte ein.[13] Die Identifikation und im Anschluss daran die Analyse von Fremdtextübernahmen in Biographien mithilfe von oben genannten Kollationstools verspricht besonders ertragreich zu sein. Als ein interessantes ergänzendes Beispiel ist ein Projekt der Internationalen Stiftung Mozarteum und der digitalen Mozartedition zu nennen, in dem unter der Federführung von Anja Morgenstern die Mozartbiographie von Georg Nikolaus Nissen auf Fremdtextüberahmen hin untersucht wird.[14] Hier werden sämtliche Quelltexte, die der Biographie zugrunde liegen, nachgewiesen. Dabei kommt der Arbeitsstelle entgegen, dass das Mozarteum sowohl über sämtliche relevanten Quelltexte in elektronisch erfasster Form verfügt als auch die Mozartbriefe digital ediert hat. Das erleichtert einen entsprechenden Vergleich entscheidend. Im Rahmen der Analyse der Nissen-Biographie ist an der dortigen

---

[11] http://collatex.huygens.knaw.nl/service/collate/console
[12] http://www.juxtasoftware.org/
[13] Vgl. hier Busch, Anna (2016): Briefbiographie um 1800. Legitimationsmechanismen einer Textsorte, in: Legitimationsmechanismen des Biographischen. Kontexte – Akteure – Techniken – Grenzen, hrsg. von Christian Klein und Falko Schnicke (Bern u.a.), S. 205–222.
[14] http://dme.mozarteum.at/DME/main/cms.php?tid=110&sec=briefe

Arbeitsstelle ein XML/TEI-P5-Schema zur Analyse von Biographien und Fremdtextübernahmen entwickelt worden, dass zukünftig bei GitHub zur freien Verfügung gestellt werden soll. Für die digitale Textanalyse von Biographien und vor allem die Weiterverwendung der erarbeiteten Analysedaten kann die Anwendung eines solchen Schemas zentral sein. Denn auch hier gilt: Je eher es zu einer Einigung auf einen Standard oder ein how-to kommt, desto eher und einfacher können die generierten Daten in anderen Forschungskontexten – auch zukünftig – weiterverwendet werden.