

# Zur Zukunft der Digitalen Briefedition – kooperative Lösungen im kulturwissenschaftlichen Forschungsdatenmanagement

Workshop am 27.02.2018 auf der DHd2018 in Köln

Leitung: Thomas Bürger, Jochen Strobel

Vortragende: Patrick Sahle, Anne Bohnenkamp, Joachim Veit, Thomas Stäcker, Stefan Dumont, Beata Mache, Jochen Strobel

Stichwortprotokoll von Olivia Varwig

## 1a) Patrick Sahle: Offenheit und institutionelle Schließung

- a. Erschließung
  - b. Technik
  - c. Abschließbarkeit
  - d. Qualität
  - e. Nutzbarkeit
  - f. Nachhaltigkeit
- Briefedition als offener, inkrementeller Erschließungsprozess, Druckeditionen dagegen (ab)geschlossen
  - Briefedition: Identifikation, Beschreibung (Metadaten) und Repräsentation (Bilddigitalisat, Basistranskription + Editorische Regeln = Verarbeitete Transkription – dazu: Entitäten (Orte etc.), Textkommentar und Sachkommentar
  - Technische Dimension
    - 1) Drucke: Textverarbeitung (OCR)
    - 2) CMIF (Correspondence Metadata Interchange Format) für Metadatenstandardisierung, CMIF Creator zum Erstellen von Digitalen Briefverzeichnissen ([http://correspsearch.net/cmif-creator/index\\_de.html?l=en](http://correspsearch.net/cmif-creator/index_de.html?l=en))
    - 3) TEI
    - 4) TIFF (zur Archivierung), JPEG (zur Repräsentation) als Bildformate für Faksimiles
    - 5) JSON/RDF (Resource Description Framework) als Austausch-/Exportformate, um semantische Wissensbestände zu extrahieren
  - Problem: Heterogene Bearbeitungs-/Erschließungszustände (der Faksimiles, Transkriptionen, Entitäten, Kommentierung)
  - Lösung: Editionsstatus
    - 1) Projektintern?
    - 2) Projektübergreifend?
    - 3) Erschließungstiefe
    - 4) Zitation mit Statusangabe
    - 5) Autorschaft: Editornennung

## 1b) Diskussionsbeiträge zu Sahle

- Bohnenkamp: Ähnliche Probleme bei allen alle Digitalen Editionen? Was ist spezifisch für den Brief?
- Antwort von mehreren:
  - 1) Briefe sind paradigmatisch, gut identifizierbar

- 2) Wucherungen sind möglich, Moment des Unabschließbaren, Briefnetzwerke stets erweiterbar, projektübergreifender Kosmos möglich
- 3) Daher: Briefwechsel im Digitalen Medium besser erschließbar, Netzwerke darstellbar
- Briefeditionen sind noch stark projekt- und meist personenbezogen, die Nutzerperspektive dagegen fordert erweiterte Zugriffsmöglichkeiten
- Frage nach technischen Lösungen vs. Bearbeitersicht
- Statusfrage macht Vernetzung schwer, benötigt Systematik
- Stäcker: Generelle projektübergreifende Formulierungen sind unwahrscheinlich, aber: Editionen müssen besser dokumentieren, was genau am Einzelbrief bearbeitet wurde, damit die verschiedenen Projektstatus vergleichbar sind
- Forderungen nach Editionsrichtlinienveröffentlichung ist alt, gilt schon für Druckversionen
- Vernetzt codierbar? Vokabular schaffen, TEI
- Abtrennung von Text und Kommentar?
- Wir brauchen Mechanismen, um mit offenen und wandelbaren digitalen Ressourcen umzugehen

## 2a) Anne Bohnenkamp: Kommentierung – ein Auslaufmodell?

- In Printeditionen derzeit Hochkonjunktur des Stellenkommentars
- Studienausgaben, z.B. Deutscher Klassikerverlag sowie Hist.-Krit. Ausgaben z.B. Erneuerung der Weimarer Goethe-Ausgabe, haben ausführlichen Stellenkommentar, bei den Goethe-Briefen teils 300 Seiten Text vs. 700 Seiten Kommentar
- Die historisch gewordenen Texte bereiten den heutigen Lesern oft Verständnisprobleme, d.h. der Kommentierungsbedarf steigt, historisch bedingte Verständnislücken (sekundäre Dunkelheit) müssen überbrückt werden Gleichzeitig sind viele Informationen jederzeit im Internet verfügbar: Kommentar überflüssig?
- Aufgaben des Briefkommentars:
  - 1) Klärung außertextlicher Zusammenhänge
  - 2) Erläuterung/Übersetzung sprachlicher Zusammenhänge
  - 3) Verweise auf Parallel- und Bezugsstellen
  - 4) Aufhellung von sekundärer Dunkelheit, die dadurch entsteht, dass der heutige Leser nicht dem eigentlichen Adressaten entspricht
  - 5) Abgrenzung zur Interpretation nicht primär (das gilt wohl aber für Werkausgaben)
- Kopf, Überlieferung, Textkritik, Beilagen, Register (bietet sich digital an), Erläuterungen, Verhältnis von Erläuterungen zum Register
- Kommentarfunktion
- Bezugs- und Antwortbriefe etc. als Verweis, Hyperlink
- Lemmata als Sprungmarken
- Stellenkommentar als kuratierter Hyperlink
- Möglichkeit von Doppel- und Mehrfachverweisen, Zwischenstationen, Mouseovertools, Kurzinformationen, die zu ausführlicheren Informationen führen
- Suchfunktion wichtig in Digitalen Editionen
- Wie behält der Nutzer in der Fülle der Informationen die Orientierung?
- Multimediales Medium nutzen, nicht nur Abbild vom Buch verfertigen
- Digitalen Editionen sind noch am Anfang, sie sind nicht nur ein ins digitale Medium verwandeltes Buch (bzgl. Visualisierungsmöglichkeiten, Organisation von Information)

## 2b) Diskussionsbeiträge zu Bohnenkamp

- Frage nach Digitaler Edition nach der Brentano-Printedition: Ausführliche Sachkommentierung wird im Print weggelassen, kommt stattdessen in die Digitale Edition

- Nutzerorientierung, Strukturierung, Visualisierung: Viele Informationen sind schon da, müssen aber noch visualisiert werden, Schnittstellenarchitektur muss entwickelt werden
- Bsp. Van Gogh-Ausgabe: speziell für diese Ausgabe entwickeltes, plausibles Informationstool
- Stäcker: Problem der Annotationsstabilität bei externen Links als Quellen: Wie hat man sich „kuratierte Links“ vorzustellen?
- Antwort: Ähnlich wie im Print, Brockhaus etc. muss auch mit Jahr angegeben werden, in Digitalen Editionen Versionierung
- Suche, Visualisierung: Briefe einer Epoche und über die Epochengrenzen hinaus
- Registerverknüpfungen müssen standardisiert sein
- Möglichkeit über einzelne Briefwechsel hinweg zu verknüpfen geht nur mit Standardisierung
- Briefkommentar funktioniert gut online, weil oft Verweise auf andere Briefe
- Briefe dialogisch angeregt, das ist die Spezifik des Briefs
- Adressierbare Ressourcen im Netz: Autorität, Qualitätssicherung?
- Problem, wenn sich der Kommentar von der Ausgabe löst und ins Netz bewegt
- Entweder eigene Ausgaben, oder Redaktionsteams: Wer ist autorisiert, eine Edition zu kommentieren? Welche Kommentare werden eingepflegt?
- Bsp. Wörterbuch der Editionswissenschaft, Kamzelak DLA: Crowd-Projekt mit Redaktionsteam: Vergleichbares für Editionen denkbar

### 3a) Joachim Veit: Versionierung /Zitation

- Analoge Welt: Auflage (mit Detailkorrekturen oder Aktualisierungen)
- Digitale Welt: Version (Erweiterung von Inhalten, Revision von Entscheidungen, neue Features)
- Versionen müssen identifizierbar sein, und zwar in Bezug auf die
  - 1) Einzelbriefe
  - 2) Gesamte Ausgabe
  - 3) Einzelnen Briefwechsel
- Van Gogh bietet engl. Übersetzungen der Briefe, wie ist das zitierbar?
- Beispiel Beethoven Edition: Briefe auch zum Hören:  
[https://www.beethoven.de/sixcms/detail.php?id=1505&template=einstieg\\_digitales\\_archiv\\_de&mid=Texte](https://www.beethoven.de/sixcms/detail.php?id=1505&template=einstieg_digitales_archiv_de&mid=Texte)
- Version im Zeitverlauf anzeigbar
- Noten, Detailreferenz: URL für vom Nutzer ausgewählte Detailzitate?
- Data first: Datenstrukturen zeigen
- Sollte man Binnen-IDs vom Briefobjekt vergeben oder Versionierung innerhalb der ID?
- Versionierung kann Metadaten/Brieftext/Kommentar betreffen
- Im Weber-Projekt: Downloads immer nur von der letzten Version möglich
- Frage der Länge von Zitatnachweisen

### 3b) Diskussionsbeiträge zu Veit

- Lange URLs sind in der analogen und der digitalen Welt unschön: Möglichkeit von Alias-URLs

### 4a) Thomas Stäcker: Hemmnisse und Katalysatoren digitaler Brief-Infrastrukturen

- Hemmnisse:
  - 1) Was ist ein Brief?
  - 2) Was ist eine Briefsammlung?
  - 3) Wie viele Briefe gibt es?
  - 4) Wie lassen sich Briefe beschreiben?
  - 5) Kein klarer Zugriff auf Briefe

- Bibliothekskataloge
  - Archive
  - Kalliope
  - Einzelprojekte und Datenbanken, z.B. Arztbriefe der frühen Neuzeit
- 6) Druck vorherrschend
  - 7) Fehlende oder unzureichende Modellierung
  - 8) Keine Standardkategorien/-schema aller Elemente
  - 9) Unklare Rahmenbedingungen für Langzeitarchivierung
  - 10) Unzureichende technische und informationswissenschaftlicher Kenntnisse
  - 11) Unabgestimmte oder nicht existente (inter)nationale Förderpolitik
- Katalysatoren:
- 1) Communitybuilding, fachübergreifend, z.B. TEI
  - 2) Förderprogramme, z.B. COST
  - 3) Zentrale Nachweissysteme, z.B. CorrespSearch
  - 4) Datensynchronisationsverfahren
  - 5) Korpus- und Teilkorpusbestimmung
  - 6) Nationales Förderprogramm für den Brief (gedruckt und handschriftlich)
  - 7) Standards und definierbare Modelle
  - 8) Kuratierte Dateninfrastruktur
  - 9) Entwicklung von flexiblen Editions Konzepten und Workflows

#### 4b) Diskussionsbeiträge zu Stäcker

- Ermittlung des Korpus: fast unendlich, nicht abschließbare Tätigkeit, ob online oder offline
- Korpusdefinition: Was gibt es für Briefe? Inhaltliche Prioritäten?
- Für Zeitungen ähnlich, Masse vs. Klasse
- In Briefen steckt viel, Plädoyer für den Brief und erneute Forderung nach nationalem Förderprogramm speziell für Briefe
- Quellen = Bilder, Digitales Faksimile, Differenzierung
- Transkription ≠ Quelle, sondern Interpretation, Auswertung, immer eine intellektuelle Leistung, Fehleranfälligkeit
- Frage der Schöpfungshöhe: Ab wann beginnt die schöpferische Leistung?
- Forschungsfrage: Wie funktioniert die frühneuzeitliche Gelehrtenrepublik → Korpus
- Forschungsfragen werden gestellt, modellieren sich aber je nach Materialfund
- Forschung hat Instinkt für Material, Forschungsfrage ändert sich
- Unterschiedliche Faksimilequalitäten: Manchmal kann das Digitalisat selbst bereits eine Interpretation sein bzw. ist es unterschiedlich interpretierbar
- Frage der CC (Creative Commons) Lizenzierung, Rollen, Herausgeber, Autorschaft
- CC0 für Quellen, CC-BY für Editionen, Downloadbutton

#### 5a) Stefan Dumont: Schnelle Wege zu den Briefen

- CorrespSearch: Briefmetadaten
- 6 Thesen:
  - 1) Offenheit: rechtliche und technische Nachnutzbarkeit, muss aber zum Ursprungsangebot zurückverfolgbar sein
  - 2) Standards: Standardformate, TEI-Richtlinien, reduzieren Variationen und Möglichkeiten, Austauschformate, die zusätzlich zur Digitalen Edition bereitgestellt werden, schließen Graben zwischen notwendiger Spezialisierung und geforderter Standardisierung
  - 3) IDs aus Normdaten: muss vertieft werden, GND

- a. Problem: GND wird über Bibliotheken gepflegt, Forschungsprojekte sollten auch einen Beitrag leisten können. Das ist bereits angedacht, muss vereinfacht werden
  - b. Problem: Für Personen und Orte gibt es bereits gute Normdaten, für Werke gilt dies nur sehr eingeschränkt: nur aus bibliothekarischen Gesichtspunkten aufgenommen
- 4) Publikationslandschaft: vielfältig und dezentral, das wird vermutlich auch so bleiben (Föderalismus)
- 5) Retrodigitalisierung: Großteil der edierten Briefe hauptsächlich gedruckt, Forderung nach systematischer Retrodigitalisierung von Bibliotheken und Institutionen (inkl. Datenaufbereitung)
- 6) Akteure: Kooperation zwischen einzelnen Akteuren muss vertieft werden
  - a. Kalliope: archivalische Daten
  - b. Forschungseinrichtungen und Universitäten
  - c. Bibliotheken für retrodigitale Editionen
  - d. CorrespSearch: edierte und wissenschaftlich erschlossene Briefe

#### 5b) Diskussionsbeiträge zu Dumont

- CorrespSearch als Basis, um Editionsprojekte zusammenzubringen
- GND-Öffnung birgt die Gefahr des Qualitätsverlusts der Daten, schon jetzt gibt es viele Doppelungen und verwaiste oder halbrichtige Datensätze, daher müsste die Redaktion der DNB gestärkt werden, diese kommt jedoch den Meldungen aktuell schon nicht hinterher
- GND nur wichtig bei editionsübergreifend wichtigen Personen/Werken, nicht beim Dorfpfarrer (der aber in einem bestimmten Briefwechsel eine Hauptrolle spielen kann)
- Forderung nach Einstiegsportalstruktur, die ALLE BRIEFE umfasst
- Möglichkeiten der Zusammenführung von Kalliope und CorrespSearch: schwierig, da Kalliope ein Verbund mit mehreren Partnern ist
- Kalliope und CorrespSearch als Tunnel, der von zwei Seiten gegraben wird und sich eines Tages (hoffentlich) treffen wird
- CorrespSearch und Kalliope zusammen über 3 Mio. Briefnachweise, davon ca. 3 % digitalisiert?
- Kalliope: nicht nur Briefe, sondern auch Werkhandschriften, aber geschätzt 80 % Briefe
- CorrespSearch: spezialisiert auf Korrespondenz

#### 6) Beata Mache: DARIAH-Services für Briefeditionen

- TextGrid, Rollengestaltung und -verwaltung
- Mindeststandards
- Basistranskription
- Definition
- Modulare Anbindung anderer Tools wie Wörterbücher möglich
- Aber: Arbeit im XML-Editor ist anspruchsvoll

#### 7a) Jochen Strobel: Akteure und Rollen

- Veränderte Rollenkonfiguration der Digitalen Edition
- traditionell: Rollenspektrum, Titelblattphilologie, Hierarchien
- Vermeidung von Anonymität: Angemessene Nennung aller Akteure und Institutionen inkl. rechtlicher Implikationen notwendig, Nennung aller Beteiligten im Impressum
- Nennung langfristig verantwortlicher Herausgeber (die mit den Projektverantwortlichen identisch sein können, aber nicht müssen) über Projektende hinaus, nach Förderungsende

- Karriererelevanz der Namensnennung, Abstufungen? Fachwissenschaftler, IT-Betreuer, Bibliothekare/Archivare
- TEI-Header als Ort dieser Zuschreibung
- Schöpfungshöhe
- Urheber kann nur natürliche Person sein, keine Institution (Anm. OV: Im Printbereich fungieren aber oft Institutionen als Herausgeber)

## 7b) Diskussionsbeiträge zu Strobel

- Wissenschaftliche, organisatorische vs. IT-technische Seite: Anteile der Akteure an Digitalen Editionen
- Neuroth: Wenn PraktikantInnen nach fachwissenschaftlichen Kriterien arbeiten, gibt es keinen Unterschied zu ProfessorInnen in der Datenautorschaft z.B. edierter Briefe und somit der Schöpfungshöhe
- Beispiel aus der Astrophysik: 100 Akteure mit Rollenzuweisung in Klammern alphabetisch genannt, ohne Wertigkeit
- Verweis auf die Richtlinien der Deutschen Forschungsgemeinschaft: diese legen genau fest, ab wann schöpferische Leistung beginnt
- Ehemalige Mitarbeiter müssten gefragt werden, ob man ihre Namen veröffentlichen darf (Datenschutz)
- Nachvollziehbarkeit, wer genau was gemacht hat, kann dies auch gefährlich sein?
- Namensnennung schafft Anreize für Projektmitarbeiter, sich für das Projekt zu engagieren
- Problem der Zitierbarkeit: 100 Beteiligte führt zu: Auernheimer et al.
- Vergleich: Filmabspann: ALLE Beteiligten werden genannt. Dagegen Generalherausgeber als Regisseur?
- Sollte man überhaupt einzelne Personen als Herausgeber nennen?
- Genaue Ausweisung, betrifft juristische Fragen, Lizenzen, Verträge
- Beispiel aus der Linguistik: Alle werden unmittelbar an der Stelle genannt, die sie bearbeitet haben mit Bearbeitungsschritt: Dient der Nachvollziehbarkeit, wer wieviel beigetragen hat
- Bsp. Schlegel-Projekt, SLUB Dresden: Es stellt sich die Frage, ob man auch die Restauratoren nennen sollte, die das Projekt durch die Ausbindung der Briefe vorbereitet haben?
- Jedes Fach hat andere Konventionen
- Wir haben ein nicht gut funktionierendes Zitationssystem/Reputationssystem, das sollte sich ändern, auch für 100+-Fälle
- Ressource, bei der mehrere Disziplinen (IT, Edition etc.) mitarbeiten, alle haben unterschiedliche Interessen, Anforderungen an die Nennung
- Auf der Ebene der Daten/Briefe: Transparenz herstellen
- Wie zitiert wird, entscheidet der Zitierende
- Frage, wer das Gesamtprojekt verantwortet hat, wer ist textverantwortlich?
- CC-BY: Projekt muss intern festlegen, wer in dieser Form juristisch genannt werden soll

## 8) Abschlussdiskussion – Fazit – Ausblick

- Wie wünschen wir uns die Landschaft?
- CorrespSearch/Kalliope Zusammenführung?
- Frage: Suche ich eine Archivalie oder einen edierten Text, einen erschlossenen Brief?
- Ähnliche Probleme und Diskussionen gab es bei der Entwicklung des DFG-Viewer: Wunsch nach Standards, Austauschformaten vs. individuelle Anpassungsmöglichkeiten an projektspezifische Infrastruktur, Förderprogramme
- Man muss den kleinsten gemeinsamen Nenner identifizieren, was sind die Gemeinsamkeiten, Blick über den Tellerrand ermöglicht übergreifende Förderprogramme

- Community muss sich einig sein, was sie will, bevor sie Förderer anfragt
- Bamberg: Metadaten aus Editionen sollten automatisch in Archive und Datenbanken zurückfließen
- Medienfenster Brief in der DDB?
- Beispiel frz. Richtlinien für Digitale Editionen: <https://cahier.hypotheses.org/guide-correspondance>
- CorrespSearch als wichtigster Katalysator für die Zukunft, in die Breite und Tiefe weiterentwickeln
- Sahle/Stäcker: Wir müssen die Briefe von ihren Editionen befreien (Downloadbarkeit, Data first: Schöne Präsentationen sind gut, aber „arbeiten“ tut man am Ende doch nur mit den Daten, diese müssen barrierefrei zugänglich gemacht werden neben der schönen Darstellung, unabhängig von der Publikation): Wenn die größten Editionsprojekte das machen, ziehen die anderen nach (Vorbildfunktion)
- Identifikation: Werkidentifizier und Versionsidentifizier
- Frage der Kuratierung nach der Kuratierung: Was passiert nach Projektende?
- Nationale Forschungsinfrastruktur, Utopie?
- Institutionen haben dafür nicht die Ressourcen, für Einzelpersonen ist das unmöglich
- Stufenweiser Verlauf:
  - 1) Bibliotheken als Verantwortliche
  - 2) Zusammenarbeit der Bibliotheken untereinander
  - 3) Institutionen und Akademien
- Datenzentrumsdiskussion, dauerhafte Datenpfleger für Editionen
- Neugründung vs. bestehende Infrastrukturen/Ressourcen umwidmen?
- Gesellschaftlicher Konsens/Gesellschaftliche Aufgabe: Daten sollten bewahrt werden ebenso wie handfeste Archivalien
- Fachcommunity braucht klare Definitionen, was wie aufbewahrt werden soll
- Rohdaten, Qualitätskriterien, Durchsuchbarkeit (in der Fachcommunity klären)
- Informationsknappheit vs. Informationsfülle (Strukturierung)
- Aufbewahren von digitalem Wissen funktioniert nicht ohne aktive Arbeit: Welche Dateiformate kann man valide vorhalten, welche muss man gelegentlich konvertieren: nicht triviale Aufgabe! Dabei auch wichtig: Fachwissenschaftler, die die Daten kennen und verstehen, inhaltliche Überprüfung der LZA, dies ist dezentral nicht möglich
- Datenmenge wächst, technische Anforderungen werden immer größer
- Nachhaltigkeit, Infrastrukturen, Datenzentren, NFDI, Nationale Forschungsdateninfrastruktur
- Geisteswissenschaften müssen nachziehen, vgl. Naturwissenschaften
- Internetarchive (archive.org)
- Bewusstsein für Standardisierung wächst