

Vom Stellenkommentar zum Netzwerk und zurück: grosse Quellenkorpora und tief erschlossene Strukturdaten auf hallerNet

Stuber, Martin

martin.stuber@hist.unibe.ch
Universität Bern, Schweiz

Dängeli, Peter

p.daengeli@uni-koeln.de
Universität Bern, Schweiz; Cologne Center for eHumanities

Forney, Christian

christian.forney@hist.unibe.ch
Universität Bern, Schweiz

In Kooperation der Universität Bern und des *Cologne Center for eHumanities* CCeH entsteht seit August 2016 die Editions- und Forschungsplattform *hallerNet*, die im April 2019 veröffentlicht werden soll. Eine Spezialität von *hallerNet* sind die systematischen Verknüpfungen zwischen grossen Quellenkorpora und umfangreichen Metadaten, woraus sich ein grosses Analysepotenzial ergibt, das sich z.B. in explorativen Visualisierungen nutzen lässt.

Metadaten aus drei Jahrzehnten

In *hallerNet* integriert werden die umfangreichen prosopografischen und bibliografischen Strukturdaten, die in Form einer relationalen Verbunddatenbank seit anfangs der 1990er-Jahre im Rahmen der beiden SNF-Projekte zum Universalgelehrten Albrecht von Haller (1708-1777)¹ und zur aufgeklärten Reformsozietät *Oekonomische Gesellschaft Bern* (gegr. 1759)² erhoben worden waren. Ein mehrjähriges Transformationsprojekt (Kreditvolumen 1.18 Mio CHF) überführt diese untereinander vielfältig verknüpften Daten (vgl. Abb. 1) zu rund 48'000 Publikationen, 25'000 Personen, 20'000 Briefen, 3'000 Orten, 2'900 Pflanzenarten, 1'000 Versammlungen und 850 Institutionen in eine TEI-konforme XML-Datenstruktur (Recker-Hamm / Stuber 2015). Bemerkenswert ist dabei die grosse Tiefe der Metadatenerschliessung, die beispielsweise nicht nur biografische Eckdaten wie Geburts-/Sterbedatum und -ort oder die Hauptbeschäftigung

umfasst, sondern die Ausbildungsstationen, Reiseziele, Ämter und Mitgliedschaften usw. einbezieht und auch das Beziehungsnetz (Verwandte, Briefpartner) soweit möglich erfasst.³ Während die prosopografische Erhebung weitgehend auf biografischen Werken sowie auf edierten Universitätsmatrikeln und Mitgliederlisten von Akademien und Sozietäten basiert, stützen sich die bibliografischen Daten in erster Linie auf zwei autopsisch erarbeitete Grundlagenwerke (Monti 1983-1994, Steinke / Profos 2004).

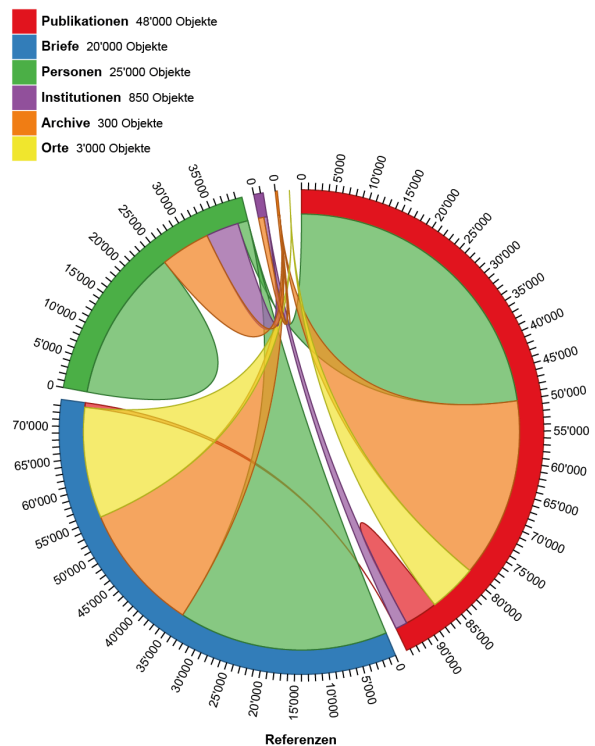


Abb. 1: Die rund 100'000 Metadatenobjekte sind vielfältig verknüpft, sowohl innerhalb des gleichen Entitätstyps als auch typenübergreifend.⁴

Auf der Grundlage der entstehenden Plattform startete anfangs 2018 ein vom *Schweizerischen Nationalfonds* SNF unterstütztes Editions- und Forschungsprojekt, das im geplanten Projektzeitraum von sechs Jahren zum einen die Gesamtedition der rund 9'000 Buchbesprechungen Albrecht von Hallers erarbeitet. Zum anderen soll eine begründete Auswahl von inhaltlich damit zusammenhängenden rund 8'000 Briefen ediert werden, dies als Zwischenetappe zur längerfristig angestrebten Gesamtedition von Hallers Korrespondenz, die insgesamt rund 17'000 Briefe umfasst.⁵ In der systematischen Verknüpfung zweier komplementärer Quellenkorpora der privaten (und tw. im kleinen Kreis öffentlichen) Kommunikation (Briefe) und des öffentlichen Diskurses (Rezensionen) liegt eine erste Innovation der Plattform.⁶

“Datenzentrierte” versus “textzentrierte” digitale Editionen

Auf *hallerNet* werden die v.a. quantitativen und formalen Bezüge der älteren Forschungsdatenbank mit edierten Textinhalten in Verbindung gesetzt. Die über Jahrzehnte systematisch erhobenen und homogenisierten Metadaten erleichtern die Editionsarbeit erheblich, namentlich die Referenzierung von Entitäten wie Akteure, Orte, Publikationen und Institutionen. In dieser Verbindung tief erschlossener Strukturdaten mit edierten Textinhalten liegt die zweite Innovation der entstehenden Plattform. Bisher sind “datenzentrierte” digitale Editionen die Ausnahme gegenüber den vielen existierenden und sich in Entwicklung befindlichen “textzentrierten” digitalen Editionen. Wo es sie gibt, verfolgen sie zumeist ganz spezifische Forschungsfragen und legen dazu ein entsprechend verengtes Datenmodell zugrunde.⁷ Daneben existieren biografische, prosopografische oder bibliografische Datenbanken und Forschungsportale, die sich einzig auf Metadatenerhebung und -analyse konzentrieren.⁸ Und schließlich gibt es digitale Editionsportale wie das *Nowa Panorama Literatary Polskiej*,⁹ die relativ dichte biografische Daten erheben, diese aber nicht in strukturierter Form bereitstellen.

Selbstredend lagen auch der Datenerfassung in den verschiedenen Projekten um Albrecht von Haller seit den 1990er-Jahren immer spezifische Forschungsinteressen zugrunde. So folgte auf eine erste Phase, welche die Gesamterschließung von Hallers Korrespondenz (Boschung et al. 2002) mit einer umfassenden prosopografischen Strukturanalyse (Stuber et al. 2005) kombinierte, eine zweite Phase, in der Netzwerkanalysen im Vordergrund standen und die in den Blick genommenen Akteure über Hallers Korrespondenznetz hinaus erweitert wurden (Dauser et al. 2008, Stuber / Krempel 2013). Die dieser Forschung zugrundeliegenden Daten wurden von Anfang an in der gleichen Datenbankanwendung (FAUST) und mit möglichst viel struktureller Homogenität gepflegt, die Datenpublikation erfolgte allerdings fast ausschliesslich in aggregierter Form und auf traditionelle Weise (Druck). Im Zuge der Datentransformation wird durch weitere Homogenisierung und die semantische Auszeichnung nach TEI-Richtlinien eine weitergehende Öffnung der Daten angestrebt, so dass diese auch in anderen Kontexten nutzbar werden.¹⁰ Die Metadaten werden zu diesem Zweck über spezifische Formate wie das CMI-Format bzw. *CorrespSearch* zugänglich gemacht, daneben aber auch in vollem Umfang parallel im FAIR-Repositorium *Zenodo* veröffentlicht werden.¹¹

Referenzannotation im Zentrum

Die Tatsache, dass die Briefe je nach Qualität der vorliegenden Grundlagen auf drei unterschiedliche Arten

ediert werden – den Standards genügende Druckeditionen werden re-ediert, bei schlechteren Druckeditionen wird der Rohtext zur Arbeitserleichterung herangezogen, aber intensiv nachbearbeitet, und zuvor nicht edierte Dokumente werden ab dem Original neu ediert –, bedingte auch eine Reflektion über das Wesen der Fussnoten bzw. der Stellenkommentare beim Transfer vom Druck ins Digitale. Aus diesem Prozess resultierte ein Referenzsystem, das drei Annotationstypen vorsieht:

- Referenzannotationen: Referenzen auf Datenbankobjekte (`<rs key="person_00001">`), deren lokales Auftreten in einer `<note type="annotRef">` näher erläutert werden kann
- freie Annotationen zum Inhalt (`<note type="annotFree">`), d.h. inhaltliche Anmerkungen, die nicht in direktem Bezug zu einem bestimmten Datenbankobjekt stehen, sondern z.B. zu einem ganzen Absatz
- Annotationen zur Textkonstitution (`<note type="annotText">`)

Während für die letzten beiden Typen die digitale Umsetzung relativ nahe beim Druck bleibt (Anmerkungsnummer mit Tooltip bzw. Verweis zur Anmerkung), bieten die Referenzannotationen mehr Interaktivität und Zugangswege zu den Datenbankdaten. Wir erwarten, dass dieser Anmerkungstyp für künftige Editionen der *hallerNet*-Plattform noch an Bedeutung gewinnen wird. Es sind dann auch diese Referenzannotationen, die das Rückgrat der Entitätsauszeichnung bilden.

Doppelter Zweck der systematischen Entitätsauszeichnung

Innerhalb der Plattform verfolgt die Anbindung der digitalen Edition an die systematisierten Objekte der Datenbank einen doppelten Zweck: Zum einen macht sie die Quellentexte über plattformübergreifende Normdaten anschlussfähig zu anderen Ressourcen. Zum anderen lassen sich bei entsprechender Datenmodellierung aus solchen Referenzannotationen grosse Datenmengen für Netzwerkanalysen und räumliche Visualisierungen gewinnen. Diese Daten weisen erstens präzise Beziehungsqualitäten auf, so beispielsweise zum Verhältnis zwischen den Rezensenten und den Verfassern der rezensierten Publikationen. Zweitens sind die Nennungen der referenzierten Entitäten leicht den i.d.R. tagesdatierten Quellen Brief und Rezension zuzuordnen. Die generierten Forschungsdaten liefern damit Präzision sowohl in der raumzeitlichen Entwicklung als auch in der inhaltlichen Qualität der Netzbeziehungen und erlauben daher Antworten auf wichtige Postulate der aktuellen historischen Netzwerkforschung.¹² Dies möchten wir an

einigen konkreten Beispiel andeuten. Ausgangspunkt ist die auf *hallerNet* als Prototyp erstellte digitale Edition der Münchhausen-Haller-Korrespondenz von Otto Sonntag, die ein erweitertes Briefsample von rund 850 Briefen (inkl. Beilagen) umfasst (Sonntag 2019). Dabei werden die vorhandenen prosopografischen Forschungsdaten angereichert mit standardisierten Angaben über soziale Position und Aufenthaltsort sämtlicher in diesem Briefsample erwähnten rund 340 Personen (ca. 1'500 Nennungen) zum Zeitpunkt des entsprechenden Briefdatums. Damit wird diese Korrespondenz als zeitlich und sozial differenzierter Kommunikationsraum darstellbar. Diese Informationen lassen sich einerseits für einzelne edierte Dokumente nutzen, etwa um eine Brieftranskription um eine Akteursliste zu ergänzen, die auch die zum Briefdatum jeweils aktuelle Position der genannten Individuen aufführt. Sie sind aber andererseits vor allem auf aggregierter Ebene interessant, da sich mit verschiedenen Visualisierungsformen bestimmte Interaktionsmuster eruieren lassen.

Die DHd-Konferenz fällt zeitlich direkt in die letzte Phase vor dem Launch der Plattform. Nachdem die Encodings der Münchhausen-Briefe in den letzten Monaten erstellt und mit den bestehenden und neuen Datenbankobjekten referenziert wurden, bietet sich bis dahin die Gelegenheit für erste analytische Auswertungen auf der neuen Datengrundlage.

Den beschriebenen Visualisierungsformen kommen in vereinfachter Form auf *hallerNet* auch Navigations- und Filterfunktionen zu. Der Weg von der einzelnen Textstelle in der edierten Quelle zur Gesamtsicht in der Visualisierung ist also durchaus in beide Richtungen möglich.

Fußnoten

1. *Albrecht von Haller und die Gelehrtenrepublik des 18. Jahrhunderts* (1991–2003); Leitung: Urs Boschung (Institut für Medizingeschichte der Universität Bern); siehe allg. zu Albrecht von Haller: <http://www.albrecht-von-haller.ch>.
2. *Nützliche Wissenschaft, Naturaneignung und Politik. Die Oekonomische Gesellschaft Bern im europäischen Kontext (1750–1850)*; Leitung: André Holenstein, Christian Pfister (Historisches Institut der Universität Bern).
3. Siehe zum alten Haller-/OeG-Datenbankverbund: Flückiger / Stuber 2009, Steinke 2003.
4. Die Visualisierung zeigt die Verlinkung und die Anzahl der Referenzen zwischen sechs Entitätstypen. Die Daten wurden im Rahmen von Forney et al. (2018) berechnet und die Grafik mit *d3-chord* erstellt: <https://github.com/d3/d3-chord>.
5. Online-Edition der Rezensionen und Briefe Albrecht von Hallers. Expertise und Kommunikation in der entstehenden *Scientific community*. Leitung: André Holenstein (Historisches Institut), Hubert Steinke (Institut

für Medizingeschichte), Oliver Lubrich (Germanistisches Institut).

6. Siehe zur Grundidee: Stuber 2004.

7. Beispiele hierfür sind etwa die Projekte *Proceedings of the Old Bailey* (<https://www.oldbaileyonline.org/>), *buckhardtsource.org* (<http://burckhardtsource.org/>), *Intoxicants and Early Modernity* (<https://www.intoxicantsproject.org/>), *Jahrrechnungen der Stadt Basel 1535 bis 1610* (<https://gams.uni-graz.at/context:srbas>), *Urfehdebücher der Stadt Basel* (<https://gams.uni-graz.at/context:ufbas>), *Sound Toll Registers* (<http://soundtoll.nl>) oder das leider dem digitalen Zerfall anheimgefallene *Alcalá account book*.

8. Als Beispiele zur Gelehrtenwelt lassen sich hier *Amore Scientiae Facti sunt Exules* (<http://asfe.unibo.it>) oder das *Repertorium Academicum Germanicum* (<http://rag-online.org>) nennen. Bemerkenswert sind auch die der *Symogih*-Plattform (Système modulaire de gestion de l'information historique, <http://symogih.org>) angegliederten Projekte, die teilweise auch edierte Texte umfassen.

9. Nowa Panorama Literatury Polskiej / New Panorama of Polish Literature; <http://nplp.pl> und spezifischer <http://tei.nplp.pl>.

10. Zur Transformation ins TEI-Format vgl. Forney et al. 2018.

11. Diese Strategie verfolgt zwei Ziele: die langfristige Zugänglichkeit der Forschungsdaten, aber auch die Auffindbarkeit über spezialisierte Suchmaschinen wie *DataCite* oder die *Google Dataset Search*.

12. Siehe eine kritische Diskussion der Möglichkeiten und Grenzen der Netzwerkanalyse aus literaturwissenschaftlicher Perspektive (Baillot 2018).

Bibliographie

Baillot, Anne (2018): *Die Krux mit dem Netz. Verknüpfung und Visualisierung bei digitalen Briefeditionen*, in: **Bernhart, Toni et al. (eds.):** *Quantitative Ansätze in den Literatur- und Geisteswissenschaften. Systematische und historische Perspektiven*, Berlin/Boston: De Gruyter 355-370 <halshs-01278211>.

Boschung, Urs et al. (2002): *Repertorium zu Albrecht von Hallers Korrespondenz 1724–1777*. 2 Bde. Basel.

Dauser, Regina et al. (2008): *Wissen im Netz. Botanik und Pflanzentransfer in europäischen Korrespondenznetzen des 18. Jahrhundert*, Berlin.

Flückiger, Daniel / Stuber, Martin (2009): *Vom System zum Akteur. Personenorientierte Datenbanken für Archiv und Forschung*, in: **Kirchhofer, André et al. (Hrsg.):** *Nachhaltige Geschichte. Festschrift für Christian Pfister*, Zürich, 253-269.

Forney Christian et al. (2018): *Vom geschützt zugänglichen Datenbankverbund zur offenen Editions- und Forschungsplattform: kritischer Rückblick auf halber Strecke*, DHd 2018 (Poster).

Monti, Maria Teresa (1983-1994): *Catalogo del Fondo Haller della Biblioteca Nazionale Braidense di Milano, a cura di Maria Teresa Monti*, 13 Bde. Milano.

Recker-Hamm, Ute / Stuber, Martin (2015): *Haller Online – Konzept für den Umbau, Ausbau und die langfristige Sicherung der Haller-/ OeG-Datenbank*, 25 Seiten (8.6.2015).

Sonntag, Otto (2019): *The Albrecht von Haller-Gerlach Adolph von Münchhausen Correspondence*, hallerNet (in Vorbereitung).

Steinke, Hubert (2003): *Archive databases as advanced research tools: the Haller Project*, in: **Antonio Vallisneri**: *L'edizione del testo scientifico d'età moderna, a cura di Maria Teresa Monti*, Firenze 2003, 191-204.

Steinke, Hubert / Profos, Claudia (2004): *Bibliographia Halleriana. Verzeichnis der Schriften von und über Albrecht von Haller*, Basel.

Stuber, Martin (2004): *Journal and letter. The interaction between two communication media in the correspondence of Albrecht von Haller*, in: **Lüsebrink, Hans-Jürgen / Popkin, Jeremy (eds.): Enlightenment, Revolution and the periodical press (Studies on Voltaire and the Eighteenth Century)**, 114-141.

Stuber, Martin et al. (2005): *Hallers Netz. Ein europäischer Gelehrtenbriefwechsel zur Zeit der Aufklärung*, Basel.

Stuber, Martin / Krempel, Lothar (2013): *Las redes académicas de Albrecht von Haller y la Sociedad Económica: un análisis de redes a varios niveles*, in: REDES. Revista hispana para el análisis. De redes sociales, 24/1 (2013), 1-26: <https://doi.org/10.5565/rev/redes.450> / REDES Online English: *The Scholarly Networks of Albrecht von Haller and the Economic Society of Bern – a Multi-Level Network Analysis*: http://revista-redes.rediris.es/webredes/novedades/Stuber_Krempel_scholarly_networks.pdf.