

Herausforderungen für Thementhesauri und Sachregister-Vokabularien zur Erschließung im Kontext des digitalen Editionsprojekts Cisleithanische Ministerratsprotokolle

Kurz, Stephan

stephan.kurz@oeaw.ac.at

Österreichische Akademie der Wissenschaften, Österreich

Zaytseva, Ksenia

ksenia.zaytseva@oeaw.ac.at

Österreichische Akademie der Wissenschaften, Österreich

Einführung

Gegenstand des umliegenden Projekts ist die Umstellung eines Langzeitvorhabens der Österreichischen Akademie der Wissenschaften (ÖAW) auf eine XML-basierte digitale Edition nach den Maßgaben der Vorschläge der Text Encoding Initiative (TEI).¹ Der Ministerrat war das zentrale Organ der Regierungstätigkeit in der Habsburgermonarchie beziehungsweise (nach 1867) in Österreich-Ungarn.² Seine Sitzungsprotokolle präsentieren alle Facetten staatlichen Lebens, von Fragen der Struktur und der Organisation des Staates bis zu gesellschaftlichen, wirtschaftlichen und technischen Entwicklungen sowie kulturellen und sozialen Problemen. Die Edition der Ministerratsprotokolle, die seit den 1970er Jahren unter wechselnder editorischer Verantwortung erscheint und in der ersten Serie 28 gedruckte Bände mit rund 18000 Druckseiten vorzuweisen hat, wird aktuell am Institut für Neuzeit- und Zeitgeschichtsforschung der ÖAW auf ein TEI-basiertes Editionsverfahren umgestellt, das die zu edierenden Bände in einer Print- und einer Online-Komponente aus einer Quelle speist.³

Die wissenschaftliche Edition der Ministerratsprotokolle zeichnet sich aus durch:

- den Umfang der zu edierenden Quellen
 - die große abzudeckende Zeitspanne von 1848 bis 1867 für die retrodigitalisierte erste Serie; von 1867 bis 1918 für die digitale Edition;
 - als Folge des umfassenden Zeitraums die Varianz etwa in den Schreibungen (mit Konsequenzen für die Textkritik, aber auch für die Arbeit mit Named Entities)
- die unterschiedliche Provenienz der Quellen
 - teilweise durch den Justizpalastbrand 1927 beschädigte Protokolle („Brandakten“)
 - Ergänzungsmaterial aus verschiedenen Archivbeständen
- vielfältige interne und externe Bezüge
 - zwischen Tagesordnungspunkten und Sitzungen, aber auch z.B. bandübergreifend
 - zwischen ediertem Text und wissenschaftlichem Kommentar
 - zu den Registern und Verzeichnissen, ergänzt durch Normdaten und Linked (Open) Data

Scope der Posterdokumentation

Das Poster präsentiert vor dem Hintergrund des Gesamtworkflows für die TEI-Edition eine Schlüsselstelle auf dem Weg zu einer digitalen Edition, die den mit dem Begriff verbundenen Anforderungen gerecht werden soll – die „Registerdaten“:

- Einerseits werden die Register der bereits edierten Bände rückerschlossen und ergänzt um neue Editionsdaten als Named Entities in einer relationalen Datenbank abgelegt, die das APIS-Projekt⁴ als biographische, prosopographische und mit Start- und Enddaten versehene Entitäten modelliert;
- andererseits werden die Sachregister in Form eines Thementhesaurus modelliert, der Politikbereiche sowie die administrative Gliederung der Habsburgermonarchie berücksichtigt.

Herausforderungen

Prüfsteine bestehen betreffend die Modellierung und Erstellung von notwendigen Auxiliardaten, wie etwa Named Entities mit diachroner Varianz (Beispiel: Namensänderung durch Personenstands- oder Standesänderungen; Umgang mit Mehrsprachigkeit bei Ortsnamen, wenn sie Zuordnungsveränderungen unterliegen), vor allem aber die Komplexität von Sachregister-Vokabularien und #Thesauri.

Methoden und Werkzeuge

Bei der Bearbeitung werden gängige Standards aus den Bereichen Semantic Web und Linked Open Data befolgt. Alle kontrollierten Vokabularien folgen dem

Datenmodell von Simple Knowledge Organization System (SKOS),⁵ um Wiederverwendbarkeit und Interoperabilität mit Daten in anderen Projekten zu gewährleisten. Die generierten Daten werden mit Dublin Core-Metadaten versehen und in ARCHE (A Resource Centre for the HumanitiEs),⁶ dem geisteswissenschaftlichen Repositorium der Österreichischen Akademie der Wissenschaften, archiviert.

Linked-Data-Anbindung

Folgende Sets von Linked Open Data-Datensätzen werden, so vorhanden, zur Anreicherung der Ministerratsprotokolltexte verwendet:

- Orte: GeoNames-IDs⁷
- Normdaten für Personen, Körperschaften, Ministerkonferenzen und Veranstaltungen, Geografische Entitäten, aus der Gemeinsamen Normdatei (GND)⁸
- Sachregister: kontrollierte Vokabularien, nach Möglichkeit Anbindung an den GEneral Multilingual Environmental Thesaurus (GEMET)⁹
- Prosopographische und biographische Daten: APIS-Datensatz zum Österreichischen Biographischen Lexikon (ab 2019 unter CC-Lizenz).

Dokumentiert und im Poster zur Diskussion gestellt wird der zum Zeitpunkt der DHd 2019 aktuelle Stand der Arbeiten am Editionsprojekt, inklusive der bisherigen Recherchen zu parallel gelagerten Editionsprojekten (u.a.: Die Protokolle des Bayerischen Ministerrats 1945–1962 Online)¹⁰ und einer Einladung zur Nachnutzung der entstehenden Daten.

Fußnoten

1. <https://www.tei-c.org/> . [XML: eXtended Markup Language]
2. Vgl. den Einleitungsband zur Serie 1, insbes. darin Rumpler, Helmut: *Ministerrat und Ministerratsprotokolle 1848 bis 1867*. Wien: ÖBV 1970, S. 11-108.
3. Zum aktuellen Stand bei der Edition vgl. <https://www.oeaw.ac.at/inz/forschungsbereiche/kulturelles-erbe/forschung/ministerratsprotokolle-habsburgermonarchie/> .
4. Vgl. <http://apis.acdh.oeaw.ac.at/> ; APIS steht für Austrian Prosopographical Information System; das 2019 auslaufende Projekt wurde in Kooperation zwischen dem Austrian Centre for Digital Humanities (ACDH), dem Institut für Stadt- und Regionalforschung (ISR) und dem Institut für Neuzeit- und Zeitgeschichtsforschung (INZ) – alle an der Österreichischen Akademie der Wissenschaften – entwickelt und bildete in einem ersten Schritt prosopographische Daten aus dem Österreichischen Biographischen Lexikon (ÖBL) als Set von Relationen zwischen persons, places,

institutions, events und works ab. Modelliert als relationale Datenbank, ermöglicht APIS die Darstellung der Entitäten und Relationen in mehreren Exportformaten, die Linked-Open-Data-fähig sind, darunter neben diversen Serialisierungen auch RDF und [TEI]-XML.

5. Vgl. Miles/Bechhofer 2009.

6. <https://www.oeaw.ac.at/acdh/tools/arche/> ; das Repositorium wird betrieben von ÖAW-ACDH und dem Rechenzentrum der Österreichischen Akademie der Wissenschaften (ARZ) und basiert auf Fedora Commons.

7. Vgl. <https://www.geonames.org/> . Vorgesehen ist auch eine Verknüpfung mit einem weiteren am ÖAW-ACDH angesiedelten Projekt, das geografische Informationsdaten mit einer historischen Tiefendimension erweitert:

HistoGIS, siehe <http://histogis.acdh.oeaw.ac.at/> .

8. Die von vielen überregionalen Bibliotheksverbünden des deutschsprachigen Raums betriebene Datei (verwaltet von der Deutschen Nationalbibliothek, <http://portal.d-nb.info/>) bietet einerseits Ressourcen, andererseits eindeutige Identifikatoren zur Disambiguierung. Für den deutschsprachigen Raum des 19. Jahrhunderts sind in der Mehrzahl der Fälle auch für die zweite Reihe der administrativen Rangordnung bereits zumindest Rumpfdaten vorhanden, und auch das Virtual International Authority File (VIAF) der Library of Congress (ein Superset von GND) hätte diesen gegenüber keine Vorteile zu bieten. Eine Online-Abfrage bietet auch das Bibliotheksservicezentrum Baden-Württemberg, siehe <http://swb.bsz-bw.de/DB=2.104/> .

9. Online-Zugang siehe <http://www.eionet.europa.eu/gemet> , Vorteile von GEMET sind die Verbreitung und die Verfügbarkeit (auch als SKOS/RDF).

10. Siehe <http://www.bayerischer-ministerrat.de/> .

Bibliographie

Viele Herausgeber und Bearbeiter (1970–2015): *Die Protokolle des österreichischen Ministerrates 1848–1867*. Wien: mehrere Verlage. [<https://hw.oeaw.ac.at/ministerrat/>].

Viele HerausgeberInnen und BearbeiterInnen (1973–2018): *Die Habsburgermonarchie 1848–1918*, 12 Bände, Wien: Verlag der Österreichischen Akademie der Wissenschaften.

Bernád, Ágoston Zénó / Gruber, Christine / Kaiser, Maximilian (2017): *Europa baut auf Biographien*. Wien: new academic press.

Fokkens, Antske / ter Braake, Serge / Ockeloen, Niels / Vossen, Piek / Legêne, Susan / Schreiber, Guus / de Boer, Victor (2017): „*BiographyNet: Extracting Relations Between People and Events*“, in: *Europa baut auf Biographien*, 193–223.

Heath, Tom / Bizer, Christian (2011): *Linked Data: Evolving the Web into a Global Data Space* (1st edition). Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. San Rafael: Morgan & Claypool

Miles, Alistair / Bechhofer, Sean (eds.) (2009):
SKOS Simple Knowledge Organization System, W3C
Recommendation, 18 August 2009, [http://www.w3.org/](http://www.w3.org/TR/skos-reference)
TR/skos-reference .

Eine ausführliche historiographische Literaturliste kann
bei Bedarf zur Verfügung gestellt werden.