

Prosopographische Interoperabilität – Stand der Dinge

Vogeler, Georg

georg.vogeler@uni-graz.at
Zentrum für Informationsmodellierung, Universität Graz,
Österreich; Austrian Center for Digital Humanities,
Österreichische Akademie der Wissenschaften

Schlögl, Matthias

matthias.schloegl@oeaw.ac.at
Austrian Center for Digital Humanities, Österreichische
Akademie der Wissenschaften

Vasold, Gunter

gunter.vasold@uni-graz.at
Zentrum für Informationsmodellierung, Universität Graz,
Österreich

Prosopographical Interoperability – State of the Art

In einem wegweisenden Artikel zu Prosopographie aus dem Jahr 1971 schreibt Lawrence Stone (Stone, 1971) “The method employed is to establish a universe to be studied, and then to ask a set of uniform questions...”. Dieses *Erstellen des Universums* ist einer der Reize der Prosopographie, stellt das Feld aber auch vor besondere Herausforderungen. Ein *Universum* besteht gemeinhin aus Millionen von Objekten und noch mehr Relationen zwischen diesen Objekten. Projekte, die sich mit prosopographischer Forschung beschäftigen, können oder würden deshalb von weiterverwendbaren Daten besonders profitieren. Ihr Daten *universum* dreht sich nicht nur um für ihre Forschung zentrale Daten (die oft neu erstellt oder überprüft werden), sondern auch um angebundene Daten: Geburtsorte von Personen, Orte in denen sich Institutionen befinden, Lehrer von Personen in der Kerngruppe etc.

Eine logische Konsequenz daraus ist es, zumindest für periphere Daten der eigenen Prosopographie LOD-Daten nachzunutzen. Um dies ohne großen Aufwand tun zu können, bräuchte es kompatible Ontologien/Datenmodelle. In den letzten Jahren wurden mehrere Versuche unternommen, für das Feld der Prosopographie einheitliche Datenmodelle vorzuschlagen (Fokkens and ter Braake, 2018; Tuominen et al., 2017), ohne dass es schon zu einem Konsens gekommen wäre.

Das Poster wird über den Stand einer Initiative berichten, die seit vergangenem Jahr an der Definition einer RESTful API arbeitet, welche die Veröffentlichung von maschinenlesbaren prosopographischen Daten so

erleichtern soll, dass typische Anfragen performant und für Softwareentwickler einfach zu realisieren sind. Dieses “International Prosopography Interoperability Format” (IPIF) hat als Kern die Definition in einer RESTful API, die in OpenAPI beschrieben ist.¹

Das dort vorgeschlagene Datenmodell deckt zum einen die Notwendigkeit ab, zwischen Person, Quelle und Quelleninterpretation zu unterscheiden (“Factoid”-Modell, Bradley & Short 2005), zum anderen vereinfacht es den Zugriff auf Informationen über eine Person in klassischen Benutzungsszenarien der Prosopographie. In einem “Statement” über eine Person können verbale Beschreibungen oder Quellenzitate ebenso wie strukturierte Informationen enthalten sein. Um die prosopographische Benutzung zu erleichtern, lassen sich die strukturierten Informationen im Modell von IPIF als datierbare Ereignisse verstehen, wenn sie mit einer Property “date” versehen sind. Sie können aber für reine Identifikationszwecke auch einfache Eigenschaften (Name, Geschlecht) abbilden. Properties “relatesToPerson” und “isMemberOf” bedienen ein drittes zentrales Szenario der Nutzung von prosopographischen Daten, nämlich Beziehungen zu anderen Personen. Schließlich sind Ortsangaben zu einer Person mit der Property “place” möglich. Mit diesen Angaben ermöglicht IPIF Anzeigen wie die des DARIAH-Cosmotools² und Ego-Netzwerke wie z.B. in der Deutschen Biographie³.

Proof of concepts

Von Beginn an war eine Grundidee des Unterfangens, die Praxistauglichkeit des Datenmodells und der API Definition möglichst früh zu testen. Zu diesem Zweck wurden seit 2018 mehrere “Proof of Concept” Applikationen erstellt. Das Poster wird den aktuellen Entwicklungsstand dieser Proof of Concept Applikationen darstellen.

APIS (Austrian Prosopographical Information System)

APIS ist ein Web-basiertes System zur Arbeit an prosopographischen Daten (Schlögl/Lejtovicz 2018). Es bietet Webformulare, aber auch API-Schnittstellen zu den Daten. Aufbauend auf die schon vorhandenen APIs wurde ein Renderer erstellt, der vorhandene Daten in das IPIF Format überführt. Zusätzlich wurden als parallele API die IPIF endpoints implementiert und somit compliance level 1 laut API Definition (GET requests inklusive Filter) erreicht.

API Wrapper

Eines der Anwendungsszenarien von IPIF ist die Suche schon vorhandener Identifier über mehrere Referenz Ressourcen hinweg (vgl. Vogeler et al 2020 forthcoming).

Um die Anwendbarkeit von IPIF für dieses Szenario zu testen, wurde eine einfache Applikation erstellt, die als Middle-Layer zwischen der IPIF API auf der einen Seite und dem Wikidata SPARQL Endpoint und der Lobid GND API auf der anderen Seite fungiert. Die Applikation übersetzt dabei Anfragen an die API in eine wikidata kompatible SPARQL query bzw. einen Lobid kompatiblen GET request und überführt die Antworten in das IPIF format. Die Applikation macht sich für diese Übersetzung die Django-Templating Engine zu Nutze und ist damit auch für andere APIs einfach konfigurierbar.

Papilotte

Auf der in OpenAPI veröffentlichten Spezifikation aufbauend lässt sich über Frameworks wie das von Zalando als Open Source Software bereitgestellte "Connexion" schnell ein Stand-Alone-Server in Python schreiben, der über flexible Konnektoren den Zugriff auf unterschiedlichste interne oder externe Datenquellen ermöglicht und diese IPIF-konform bereitstellt. Ein solcher Server wurde mit Beispieldaten aus dem Monasterium.net-Projekt erstellt.⁴

Papi-Cosmotool

Das DARIAH-Cosmotool bietet eine prototypische Oberfläche für eine prosopographische Datenbank an. Es enthält eine biographische ("Zeitleiste"), eine textuelle ("Ereignis-Detail") und eine geographische ("Kartendarstellung") Ansicht. Die Datenanzeige wird mit einem Quellenverweis ergänzt. Als Test für die Verwendbarkeit der API-Definition wird von Sebastian Stoff am Zentrum für Informationsmodellierung eine JavaScript basierte Anwendung erarbeitet, die vergleichbare Funktionalitäten bietet.⁵

JSON-LD

Schließlich arbeiten wir an einer Integration des JSON-Outputs der API-Definition in das Semantic Web. Dafür soll eine context.json-Datei bereitgestellt werden, die in den Resultsets der API-Anfragen gültige RDF-Aussagen identifiziert.⁶

Fußnoten

1. <https://github.com/GVogeler/prosopogrAPhI>
2. <https://cosmotool.de.dariah.eu/>
3. z.B. <https://www.deutsche-biographie.de/graph?id=sfz53095>
4. <https://github.com/gvasold/papilotte>, mit einer Beispielinstallation: <https://ginko.uni-graz.at/illurk/api/ui/>

5. <http://glossa.uni-graz.at/gamsdev/stoffse/erla/mapp/map/>, Code von Sebastian Stoff.

6. Ein erster Entwurf ist unter <https://github.com/GVogeler/prosopogrAPhI/blob/master/context.json> einsehbar.

Bibliographie

Bradley, John / Short, Harold (2005): "Texts into databases. The Evolving field of New-style Prosopography." in: *LLC* 20, suppl. 1: 3-24.

Fokkens, A. / ter Braake, S. (2018): Connecting People Across Borders: a Repository for Biographical Data Models, in: Proceedings of the Second Conference on Biographical Data in a Digital World 2017. Linz, Austria, November 6-7, 2017. CEUR Workshop Proceedings: 83-92.

Stone, Lawrence (1971): "Prosopography." in: *Daedalus* 100: 46-79.

Schlögl, Matthias / Katalin Lejtovicz . (2018). "A Prosopographical Information System (APIS)." In: Antske Fokkens, ter Braake, Serge, Sluijter, Ronald, Arthur, Paul, and Wandl-Vogt, Eveline (eds.). *BD-2017. Biographical Data in a Digital World 2017. Proceedings of the Second Conference on Biographical Data in a Digital World 2017* . Linz, Austria, November 6-7, 2017. Budapest: CEUR (CEUR Workshop Proceedings 2119): 53-58.

Tuominen, Jouni / Hyvönen, Eero / Leskinen, Petri (2018): "Bio CRM. A Data Model for Representing Biographical Data for Prosopographical Research." In: *BD-2017. Biographical Data in a Digital World 2017* , ed. by Antske Fokkens, Serge ter Braake, Ronald Sluijter, Paul Arthur, Eveline Wandl-Vogt, Budapest: CEUR (CEUR Workshop Proceedings 2119): 59-66.

Vogeler, Georg / Vasold, Gunter / Schlögl, Matthias (2020 forthcoming): Data exchange in practice: Towards a prosopographical API. In: *BD2019, Workshop on Biographical Data in Occasion of the RANLP 2019, Varna*, ed. by Antske Fokkens et al.