

Workshop Historisch-biographisches Informationssystem

Veranstalter: Historische Kommission bei der Bayerischen Akademie der Wissenschaften; Bayerische Staatsbibliothek

Datum, Ort: 01.12.2014, München

Bericht von: Roxane Hagn/Maximilian Schrott, Historische Kommission bei der Bayerischen Akademie der Wissenschaften, München

Am 1. Dezember 2014 veranstalteten die Historische Kommission bei der Bayerischen Akademie der Wissenschaften (HiKo) und die Bayerische Staatsbibliothek (BSB) einen ganztägigen Workshop im Historischen Kolleg, den 70 Teilnehmerinnen und Teilnehmer aus ganz Deutschland besuchten. Anlass war der Abschluss des gemeinsam in den Jahren 2012 bis 2014 durchgeführten und von der DFG geförderten Projekts zur „Entwicklung eines historisch-biographischen Informationssystems für den deutschsprachigen Raum“.

Im Mittelpunkt der Veranstaltung stand die Präsentation des völlig neuen Online-Auftritts.¹ Statt bisher 130.000 können in dem historisch-biographischen Informationssystem „Deutsche Biographie“ (DB) – dessen Nukleus die historisch-biographischen Artikel der Allgemeinen Deutschen Biographie (ADB) und der Neuen Deutschen Biographie (NDB) sind – jetzt mehr als 260.000 Personen recherchiert werden. Durch die forcierte GND-(Gemeinsame Normdatei)Erschließung und die Kooperation mit namhaften nationalen Partnerinstitutionen sind nun zahlreiche für die wissenschaftliche Arbeit relevante Ressourcen direkt verbunden. Zu den Partnern zählen unter anderem die Nachlassdatenbank und das Bildarchiv des Bundesarchivs, die Archivschätze des Deutschen Literaturarchivs in Marbach, die Objektdatenbank des Germanischen Nationalmuseums in Nürnberg, der Bildbestand von Foto Marburg oder Sendenachweise zur Weimarer Republik aus dem Deutschen Rundfunkarchiv in Frankfurt am Main.

All dies lässt sich jetzt auch über kartenbasierte, stark erweiterte und facetiierte Suche recherchieren. Unter den neuen Features

finden sich erstmals auch Präsentationsformen, die durch computerlinguistische Analysen möglich werden, wie zum Beispiel die graphische Darstellung von Ego-Netzwerken.

Wie MALTE REHBEIN (Passau), der die digitale DB leitet, in seiner Einführung betonte, biete die DB in ihren biographischen Artikeln „zertifiziertes Wissen“, das verlässlich, redaktionell überprüft und zitierbar sei. Ziel des seit 2012 laufenden Projekts sei es gewesen, dieses Angebot signifikant zu erweitern und qualitativ zu optimieren. Ein wichtiger Teil dieses Vorhabens sei auch die Entwicklung des neuen Webauftritts gewesen, der parallel zum Workshop freigeschaltet wurde.² Durch zahlreiche neue Funktionen würden die Bestände der DB sowie die Angebote der kooperierenden Institutionen für die Benutzer leicht zugänglich gemacht.

Diese neuen Features wurden den Teilnehmern von DIRK SCHOLZ und MAXIMILIAN SCHROTT (beide München) vorgestellt. Insgesamt können jetzt über 260.000 Personen recherchiert werden. Das sind doppelt so viel wie vor dem Projektbeginn. Etwas mehr als 48.000 davon verfügen über einen Artikel in ADB oder NDB, der im Volltext online gelesen werden kann. Aber auch über die anderen Personen könnten die Benutzer Interessantes in Erfahrung bringen. Denn über Verlinkungen zu den Angeboten der kooperierenden Institutionen lassen sich zum Beispiel Quellen, Literatur, Werke oder Bilder von oder über bestimmte Persönlichkeiten leicht recherchieren.

Auch neue Suchfunktionen wurden vorgestellt: Features wie die automatische Vervollständigung der Sucheingaben und vielfältige Filtermöglichkeiten sollen den Benutzern künftig die Recherche erleichtern, zum Beispiel wenn sie nur Personen aus bestimmten Berufsfeldern finden möchten. Zusätzlich wurden für über 67.000 Personen die Orte, an denen sie geboren wurden, gelebt haben (Wirkungsort), gestorben sind oder begraben wurden, erfasst und geolokalisiert. Mit diesen Daten ist es nun möglich, sich auf einer Landkarte einen Überblick über die Lebens-

¹ vgl. Myriam Hönig, *Schöne neue Möglichkeiten*, in: Frankfurter Allgemeine Zeitung, 4.12.2014, S.8.

² Deutsche Biographie, <<http://www.deutsche-biographie.de>> (24.2.2015).

stationen einer Person zu machen oder aber direkt eine Suche über ein bestimmtes, eingegrenztes geographisches Gebiet durchzuführen, um am Ende zum Beispiel alle Personen angezeigt zu bekommen, für die die DB einen Lebensort in Württemberg aufweist.

Ebenfalls im Zuge des Projekts neu erfasst wurden die Beziehungen der biographierten Personen untereinander. Die mit Hilfe computerlinguistischer Methoden aus den Artikeln von NDB und ADB ermittelten Verknüpfungen können, wie gezeigt wurde, sowohl in Listenform, als auch in der Form eines Ego-Netzwerks visualisiert werden. Bei letzterem werden die Beziehungen einer bestimmten Person als ein Geflecht von Knoten und Kanten dargestellt, das sich um die Bekannten der Bekannten erweitern lässt. Wie Maximilian Schrott betonte, stünden diese Visualisierung, wie auch die geographischen Funktionen noch am Anfang. Es sei geplant, sie im Rahmen des im Oktober 2014 von der DFG bewilligten Folgeantrags mit weiteren Daten und Funktionen weiter zu optimieren.

Nach einer kurzen Pause setzte THOMAS BUSCH (München) den Workshop mit einem Vortrag über den Stand der GND-Erschließung fort. In dieser Gemeinsamen Normdatei soll, unter Federführung der Deutschen Nationalbibliothek (DNB) und maßgeblicher Mitarbeit der BSB, ein einheitlicher Katalog unter anderem für Personen, Geografika, Werktitel und Sachschlagwörter entstehen. Durch die Normierung der Daten werde es leichter möglich sein, Angebote im Internet miteinander zu verknüpfen.

Thomas Busch sieht große Fortschritte bei der GND-Erschließung. Im Oktober 2014 habe es bereits 10,8 Millionen überregionale Normdatensätze gegeben, und dank der etwa 300 teilnehmenden Institutionen wachse diese Zahl stetig weiter. Das gerade abgeschlossene Projekt habe dabei einen signifikanten Beitrag zu den personenbezogenen Daten innerhalb der GND geleistet. Durch Zusammenarbeit der HiKo, der BSB sowie der Partnerinstitutionen konnten umfangreiche Personendatenbestände neu in die GND aufgenommen werden. Gleichzeitig wurden schon vorhandene Datensätze mit genaueren Daten verfeinert oder individualisiert. Dafür war ein nicht unwesentlicher Aufwand nötig, da nur

ein Teil der Daten automatisiert erfasst werden konnte. In den meisten Fällen war hingegen intellektuelle Abgleicharbeit notwendig. Als Lohn für diese Mühen profitieren die kooperierenden Institutionen nun aber von der verbesserten Vernetzbarkeit ihrer Bestände.

An den Vortrag von Thomas Busch schloss sich eine Diskussionsrunde der Partnerinstitutionen der Projektphase 2012 - 2014 an. Hierzu stellten sich die Vertreter der Partnerinstitutionen zuerst mit einem kurzen Überblick über die eigene Einrichtung und deren Aufgabengebiete vor, um dann auf die bisherigen Erfolge mit dem GND-Matching und eigene zukünftige Pläne und Wünsche an die GND einzugehen.

Insgesamt lässt sich festhalten, dass die bisherige Verknüpfung mit der GND für die Partner erfolgreich war und auch zukünftig für neu hinzukommende Bestände in den Partnerinstitutionen ein Abgleich mit der GND stattfinden soll. Bei Erwähnung des erfolgreichen Matchings dankten die Partnerinstitutionen auch mehrfach den Vertretern der BSB und der HiKo für die erfolgreiche Unterstützung.

Neben dem Personenabgleich, so hieß es, werde die GND auch verstärkt für andere Teilbereiche genutzt. Eine Entwicklung auf die auch BRIGITTE WIECHMANN (Frankfurt am Main) aufmerksam machte und ein mögliches neues Webformular anregte, das besser für nichtbibliothekarische Anfragen geeignet sei.

Die Partnerinstitutionen machten darauf aufmerksam, dass vermehrt BEACON-Dateien³ zur Verfügung stünden. Institutionen können als Schnittstelle nach außen hin eine BEACON-Datei bereitstellen, in der die fest vergebene GND-Nummer mit einem Link auf das eigene Angebot zum jeweiligen Objekt verbunden ist. Auch das digitale Portal der DB nutzt die GND und realisiert so die Verknüpfungen zu ihren Partnerinstitutionen.

Neben dem Lob gab es auch Skepsis hinsichtlich des Ertrages einer zunehmenden Vernetzung. ANDREAS DAN (Frankfurt am Main) merkte an, dass seiner Erfahrung nach eine Verlinkung mit der GND nur zu einer ge-

³ Zu BEACON, siehe: <<https://de.wikipedia.org/wiki/Wikipedia:BEACON>> (24.02.2015)

ringen Erhöhung der Besucherzahlen führe. Abschließend lud Brigitte Wiechmann alle Institutionen ein, sich der GND anzuschließen.

Parallel zur Mittagspause wurden von einigen Mitarbeitern der HiKo und der BSB Poster zu technischen Hintergrundinformationen der Digitalisierung der Deutschen Biographie vorgestellt.

Im Anschluss wurden in einem Vortrag von BERNHARD EBNETH und MATTHIAS REINERT (beide München) die Ziele des von der DFG geförderten Fortsetzungsprojekts vorgestellt. Bernhard Ebnethe erklärte, dass durch die Aggregation personenbezogener Daten und Metadaten großer Gedächtnisinstitutionen ein zuverlässiges und stabiles biografisches Informationssystem aufgebaut werden solle. Dies geschehe durch die Beteiligung bedeutender Archive, Bibliotheken, Dokumentationseinrichtungen und Museen. Im DFG-Fortsetzungsantrag liege nunmehr der Schwerpunkt auf Akademieprojekten. Die neuen Partnerinstitutionen im historisch-biographischen Informationssystem verwendeten bereits die GND, durch deren Identifikatoren sich neue Vernetzungsoptionen ergäben.

Die Möglichkeiten im Semantic Web wurden anschließend von Matthias Reinert vorgestellt. Dabei handelt es sich um das Konzept, das Internet flächendeckend mit maschinenlesbaren Inhalten zu versorgen. Mit Hilfe des Resource-Description Framework-Standards (RDF) können so Informationen im Web verknüpfbar und automatisch auswertbar gemacht werden. Durch die Bereitstellung von RDF-Dateien für die biographischen Artikel und die Verlinkung über die GND sei die DB zusammen mit ihren Partnern Teil dieses semantischen Netzes.

Neben der Verknüpfung zwischen den Projekten solle auch eine kollaborative Arbeitsumgebung entstehen. Hierzu würden computerlinguistische Methoden, wie Entitäten-Erkennung und -Disambiguierung, weiterentwickelt, um so die Daten für das Ego-Netzwerk zu verbessern. Auch sollten weitere Metadatenquellen integriert und die Graphenvisualisierung mit Graphendatenbanken weiterentwickelt werden.

Unter der Leitung von Bernhard Ebnethe stellten sich schließlich die neuen Partnerin-

stitutionen vor. Wie bei den bisherigen Partnern präsentierten die Referenten zuerst ihre Einrichtung und deren Angebote und Ziele, um dann auf die Spezifika ihrer Datenbestände einzugehen. Dabei wurden neben den inhaltlichen Charakteristika und dem Umfang der Daten auch mögliche genutzte, angebotene oder geplante Schnittstellen beschrieben. Weiterhin wurde präzisiert, inwieweit die Bestände bereits elektronisch erschlossen und im Internet abrufbar sind.

Zuletzt gingen die Partner auf ihre Erwartungen an eine Vernetzung mit der GND und ihre Erweiterungsvorschläge ein. So erwarteten sich die neuen Kooperationspartner Hilfe bei schwer identifizierbaren Personen sowie eine Erhöhung des Grads der Verlinkung der Personen untereinander, zum Beispiel durch eine genauere Abbildung von Beziehungen. Auch sollen durch eine Vernetzung mit der DB noch nicht erfasste Personen in die GND aufgenommen werden oder es könne geklärt werden, inwieweit Personen aus bestimmten Datenbeständen überhaupt bedeutend und damit relevant sind. Des Weiteren wurde der Wunsch geäußert, nicht nur Verknüpfungen mit Personen, sondern auch mit Geografika, Institutionen, Bauwerken und ähnlichem zu erstellen.

Zum Abschluss des Workshops übergab Malte Rehbein das Wort an MARTEN DÜRING (Luxemburg) für einen Impulsvortrag über Historische Netzwerkforschung und das Semantic Web. Netzwerkforschung, so Düring, sei ein interessantes Thema für die Geschichtswissenschaften, vor allem wenn es darum gehe, soziale Beziehungen aufzuschlüsseln. Er selbst warne aber vor übertriebenen Erwartungen und davor, Netzwerkforschung nur zu betreiben, weil das Konzept gerade in Mode gekommen sei. Selbst wenn die nötigen Daten vorhanden seien, gelinge das Erstellen eines aussagekräftigen Netzwerks nicht automatisch. Ein methodologisch überlegtes Vorgehen und genaue Reflexion der zu erforschenden Zusammenhänge sei unabdingbar. Düring zog in diesem Zusammenhang den Vergleich zu einer Straßenkarte. Genau wie diese nicht jedes Detail des Geländes erfasse und dafür andere besonders hervorhebe, müsse auch bei der Netzwerkepräsentation die Wirklichkeit gezielt abstrahiert

und vereinfacht werden. Auch bei richtiger Anwendung könne Netzwerkforschung keine Lösungen frei Haus präsentieren, aber ein hervorragendes Werkzeug sein, um komplexe Beziehungsstrukturen sichtbar zu machen und dann zu analysieren. Um diese Methoden in Zukunft zugänglicher zu machen, kündigte Marten Düring für das kommende Jahr an, zusammen mit Kollegen einen Sammelband zu veröffentlichen, der die Prinzipien der Netzwerkanalyse speziell für Historiker darlegt.⁴

Im zweiten Teil seines Vortrags ging Düring dann auf die Möglichkeiten des Semantic Webs ein. Bereits bei vielen Ressourcen, wie der Wikipedia oder dem Linked Open Data-Projekt werde der Gedanke vom intelligenten Internet umgesetzt. Aber trotz des raschen Fortschritts gäbe es auch noch immer Zweifel, wie erfolgreich dieses Vorhaben wirklich sein könne. Denn Datenlücken oder unsaubere Daten könnten die Nutzbarkeit des Semantic Webs gefährden. Aber wie Düring zeigen konnte, gibt es bereits jetzt erfolgreiche historische Projekte auf Basis des Semantic Web. An ihnen sei erkennbar, welche Chancen ein Internet, an das hochflexible und intelligente Anfragen gestellt werden können, für die Geisteswissenschaften biete.

Insgesamt war der Workshop vor allem vom Gedanken der semantischen Verknüpfung geprägt. Von ihr erwarten sich die teilnehmenden Organisationen ganz überwiegend neue Möglichkeiten für die historischen Wissenschaften im Internet. Immer wieder wurde dabei die Bedeutung der GND als Schlüsselinstrument für diese Perspektive herausgestellt. Die DB will mit ihrem neuen Internetauftritt ein zentraler Knotenpunkt für die Fächergrenzen überschreitende Beschäftigung mit historischen Persönlichkeiten aus allen Bereichen und Epochen des deutschsprachigen Raumes innerhalb dieses Netzwerkes werden. Dieses Ziel soll, mit weiteren Funktionen und der Einbindung der neuen Partner kontinuierlich fortgesetzt werden.

Konferenzübersicht:

Helmut Neuhaus/Rolf Griebel (München), Begrüßung

Malte Rehbein (Passau), Einführung

Dirk Scholz/Maximilian Schrott (München), Vorstellung des erweiterten Internetauftritts der Deutschen Biographie

Thomas Busch (München), Bericht über die GND-Erschließung

Diskussion mit den Partnerinstitutionen der Projektphase 2012-2014

Moderation: Matthias Reinert (München)

Es stellen sich vor:

Angela Kailus (Marburg)

Karin Schmidgall (Marbach)

Rainer Jacobs (Koblenz)

Andreas Dan (Frankfurt am Main)

Birgit Jooss (Nürnberg)

Georg Hohmann (München)

Brigitte Wiechmann (Frankfurt am Main)

Poster-Session

Sophia Stotz/Valentina Stuß (München), Relationsextraktion

Roxane Hagn/Dolores Sarancic (München), Personen- und Ortsabgleich

Christian Kaufhold/Maximilian Schrott (München), Suchindizes

Bernhard Ebner/Matthias Reinert (München), Ziele des Verlängerungsantrags

Präsentation der neuen Partner

Moderation: Bernhard Ebner (München)

Es stellen sich vor:

Matti Stöhr (Berlin)

Yannik Weber (Mainz)

Martin Lies (Mainz)

Eva-Maria Dickhaut (Marburg)

Gerhard Müller (Berlin)

Georg Eckes (Frankfurt am Main)

Jens Bove (Dresden)

Daniel Fährle (Stuttgart)

Dorothee Goetze (Bonn)

Caroline Sternberg (München)

Marten Düring (Luxemburg), Vortrag: Historische Netzwerkforschung und Semantic Web

Tagungsbericht *Workshop Historisch-biographisches Informationssystem*. 01.12.2014, München, in: H-Soz-Kult 04.03.2015.

⁴ Marten Düring / Ulrich Eumann / Linda von Keyserlingk / Martin Stark (Hrsg.), Handbuch Historische Netzwerkforschung, Berlin vsl. 2015.