

## **Germania Sacra Online – Das Forschungsportal für kirchliche Personen und Institutionen bis 1810**

Bärbel Kröger und Dr. Christian Popp, Germania Sacra

Wer im Netz nach wissenschaftlichen Informationen beispielsweise zu einem mittelalterlichen Kloster recherchiert, kann schon heute im Idealfall auf dem Forschungsportal der Germania Sacra ein ganzes Bündel von Informationen erhalten: Basisdaten zur Geschichte der Institution, kartographisch visualisierte Standortinformationen, Normdaten (GND, DBPedia, GeoNames), Links zu weiterführenden regionalen Online-Angeboten und bibliographische Informationen, Verknüpfungen zu dem in der Personendatenbank der Germania Sacra erfassten Klosterpersonal, die weitere fachübergreifende Verweise enthalten und damit den Weg zu neuen Erkenntnissen ermöglichen.

Die Germania Sacra hat in den vergangenen Jahren ein breites Portfolio digitaler Angebote zur Kirche des Alten Reiches erstellt. Hauptsäulen sind die digitalisierten Handbücher zur Geschichte kirchlicher Institutionen, die im Rahmen des Langzeitprojektes seit 1917 erarbeitet worden sind, ein umfangreiches digitales Personenregister zum kirchlichen Personal sowie die Datenbank zu Klöstern und Stiften des Alten Reiches. Alle digitalen Angebote der Germania Sacra sind work-in-progress: neue Bände werden nach 3 Jahren digital zur Verfügung gestellt, das Personenregister wird laufend erweitert (ca. 26.000 Einträge, Stand November 2014), die Klosterdatenbank befindet sich in der Aufbauphase (ca. 900 Einträge, Stand November 2014).

Die projektinterne Vernetzung der Daten wurde im Zuge der Integration in das Digitale Portal der Akademie der Wissenschaften zu Göttingen sichergestellt. Die hierfür verwendeten Technologien und Funktionalitäten werden im Rahmen des Vortrags präsentiert. Der Schwerpunkt des Vortrages wird jedoch auf der projektübergreifenden Vernetzung von Daten liegen. Ausführlich skizziert werden die hierfür bereits entwickelten Lösungsansätze sowie die zukunftsweisenden Kooperationen, die zu einer neuartigen digitalen Wissenslandschaft über die Kirche des Alten Reiches führen sollen.

Ein wichtiger Baustein für die Vernetzung ist die systematische Anreicherung der Datenbestände mit Normdaten. Für viele der durch die Forschung der Germania Sacra generierten Informationen kann auf bereits vorhandene Normdaten zurückgegriffen werden. Besonders relevant für unser Projekt ist der Datenbestand der Deutschen Nationalbibliothek mit den dort verwendeten Datensatznummern der Gemeinsamen Normdatei (GND). Für Personendaten wird üblicherweise das Beacon-Format verwendet, das das automatische Generieren von Links zu externen Datenquellen ermöglicht. Für andere Daten als Personen, etwa für Körperschaften, wird das Beacon-Format bisher kaum genutzt. Mit der Klosterdatenbank der Germania Sacra soll die Verwendung dieser Technik für Klöster und Stifte erprobt und eingeführt werden. Die Identifizierung der einzelnen Klöster und Stifte in der Gemeinsamen Normdatei der Deutschen Nationalbibliothek ist bereits vielfach erfolgt, fehlende Einträge in der GND werden durch die Germania Sacra ergänzt. So können in der Datenbank automatisiert direkte Links nicht nur zu externen Datenbanken, sondern auch zu relevanten Datensätzen in Bibliothekskatalogen, Bestandsübersichten von Archiven, Quelleneditionen, Bibliographien, Porträtsammlungen und weiteren Informationsangeboten bereitgehalten werden.

Um den Möglichkeiten zur semantischen Recherche einen Weg zu bereiten, werden die Inhalte der Datenbanken auf der Basis von Linked Data angereichert und im RDF-Format ausgegeben, dabei wird auf etablierte existierende Vokabulare zurückgegriffen. Für die Ausgabe der Datensätze im RDF-Format werden Normdaten für Orden, Bistümer, Personen wie auch Normdaten für Geografika (GeoNames) verwendet. Vorhandene Einträge in der Wikipedia werden referenziert. Das modellierte Schema bietet hohes Potential, das Informationsnetz zu den Beziehungen von Personen und geistlichen Institutionen für den Zeitraum des Mittelalters und der Frühen Neuzeit zu verdichten.

Während bei der Verknüpfung von Daten zu kirchlichen Institutionen aufgrund eindeutiger Identifikatoren (z.B. Name, Orden, Standortinformation) vergleichsweise gut automatisierte Lösungsansätze zu finden sind, ist die automatisierte Vernetzung von personengeschichtlichen Datenbanken aus dem Bereich der Mittelalter- und Frühneuzeitforschung nach wie vor ein ungelöstes Problem im Bereich der Digital Humanities. Die Zuordnung von Informationen aus unterschiedlichen Datenbanken zu einer bestimmten Person ist schwierig. Häufig liegen nicht genügend Daten vor, die eine sichere Identifizierung ermöglichen (Geburts- und Sterbedatum, Herkunftsort, Ämter und Amtsdaten). Erschwerend kommen die zum Teil erheblich abweichenden Namensvarianten, Übersetzungs- und Transkriptionsfehler, latinisierte Formen und der spät einsetzende Gebrauch von Zweitnamen hinzu.

Daher entwickelt die Germania Sacra in Kooperation mit dem Deutschen Historischen Institut in Rom (DHI) und dem Repertorium Academicum Germanicum (RAG) eine projektübergreifende Personenrecherche. Dabei gilt es geeignete technische Lösungen zu finden. Hier können beispielsweise Algorithmen, die phonetische und orthographische Varianten auffindbar machen, oder die Verwendung von Thesauri zur Erkennung latinisierter Namensformen hilfreich sein. Diese Metasuche soll nicht nur als ein datenbankübergreifendes Recherchetool fungieren, sondern zugleich unter Verwendung von Technologien des Crowdsourcing interaktive Verknüpfungsmöglichkeiten für wissenschaftliche Nutzer bieten, die so ihre Identifikationsvorschläge einzelner Personendatensätze in die Datenbanken zurückmelden können.

Gerade die Verknüpfung von Personendaten aus unterschiedlichen Forschungsprojekten und unterschiedlichen Quellenbeständen lässt eine Generierung neuen Wissens erwarten, die in dieser Form nur durch den Einsatz digitaler Werkzeuge möglich ist.

Die nachhaltige Nutzung der Forschungsdaten wird durch die Integration von Germania Sacra Online in das Digitale Portal der Akademie der Wissenschaften zu Göttingen gewährleistet, die die hierfür erforderliche Infrastruktur zur Verfügung stellt.