

Die Herausforderung uneinheitlicher Daten

Um die digitale Suche nach historischen Personen oder Orten effizienter zu gestalten, arbeitet das Schweizer Projekt histHub an der Vernetzung und Normierung von Daten und ergänzt damit bestehende Infrastrukturen und Dienstleistungen im Bereich der Digital Humanities.

Das Grafengeschlecht von Werdenberg-Sargans war im 14. und 15. Jahrhundert im heutigen Kanton Graubünden äußerst einflussreich. So wurde etwa Hartmann IV. 1388 zum Bischof von Chur gewählt. Während das Datum der Wahl eindeutig belegt ist, sind deren Umstände umstritten. Umstritten ist auch das Geburtsjahr von Hartmann IV., während sein Todestag, der 6.9.1416, als gesichert gilt. Die Durchmischung eindeutiger mit unklaren oder gar widersprüchlichen Angaben ist typisch für historisch belegte Personen aus dem Spätmittelalter und der Frühen Neuzeit. Charakteristisch ist auch das Nebeneinander von unterschiedlichen Schreibvarianten, Bei- und Zunamen sowie von Ortsangaben und Funktionsbezeichnungen. So ist unser Hartmann IV. von Werdenberg-Sargans auch als Hartmann de Sanegans, oder – was die Verwechslungsgefahr deutlich erhöht – als Hartmann II. von Chur belegt.

Eindeutig und vernetzt

Dieses Beispiel weist auf ein typisches Problem bei historischen Personendaten hin: Die eindeutige Identifikation ist oft nur schwer möglich. Hinzu kommt, dass viele Personen bereits in verschiedenen Datenbanken erfasst sind. Dieses Nebeneinander von Datenbankeinträgen betrifft natürlich auch Daten zu Orten, Organisationen oder Schlagworten. Eine möglichst weit gefasste Verlin-

kung würde demnach Vergleiche der bestehenden Angaben ermöglichen und dabei die Sichtbarkeit der Datenbanken erhöhen.

Die Plattform histHub arbeitet gegenwärtig an einer Lösung für die eindeutige Zuordnung und Vernetzung von Daten. histHub verfolgt dabei zwei sich ergänzende Strategien. Erstens bietet das Portal Normdaten zu historischen Personen, Orten, Organisationen und Konzepten. Typologien und kontrollierte Vokabulare kommen dabei zum Einsatz. Die Entitäten werden so modelliert, dass feinste historische Unterschiede mit entsprechenden Belegen erfassbar sind. Durch die Vergabe einer permanenten ID sind die Daten auch für den Rechner eindeutig.

Von der Schweiz in die Welt

Die zweite Strategie von histHub besteht in der automatischen oder halbautomatischen Vernetzung bestehender Datenbanken. Konkordanzen werden gespeichert und dadurch Gleiches mit Gleichem vernetzt. Die Datenhoheit bleibt dabei vollständig beim Anbieter. Durch die eingesetzten, innovativen Technologien kann histHub eine große Anzahl Daten von Drittprovidern durchsuchen und verlinken. So generiert das histHub-Vernetzungstool für Ortsangaben zu Sargans aktuell Links von über 10 internationalen Providern. Auch wenn zurzeit der Fokus auf Daten zur Schweizer Geschichte liegt, sind

die bei histHub entwickelten Tools, Datenmodelle und Ontologien sicherlich auch für eine internationale Community von Interesse. ■

histHub ist ein Kooperationsprojekt der Diplomatischen Dokumente der Schweiz (Dodis), des Historischen Lexikons der Schweiz (HLS), des Schweizerischen Idiotikons mit dem Projekt ortsnamen.ch sowie der Rechtsquellenstiftung des Schweizerischen Juristenvereins und wird von swissuniversities mitfinanziert.

Links:

histhub.ch

dodis.ch

hls-dhs-dss.ch

idiotikon.ch

ortsnamen.ch

ssrq-sds-fds.ch

Alban Frei ist Historiker und Forschungsmanager. Er hat an der ETH Zürich promoviert und ist seit 2017 Geschäftsführer von histHub.