



Teilprojekt: AP 6 (UB) Realisierung orts- und zeitbezogener Anfragen auf Dokumentenservern

Antragsteller: Robert Zepf (UB, Direktor), Robert Stephan (UB, Digitale Bibliothek)

Bearbeiter: Karsten Labahn unter Mitwirkung von: Susanne Jürgensmann (wiss. Hilfskraft)

Themenstellung:

Ziel des Arbeitspaketes ist die Einbeziehung von bibliothekarischen Ressourcen in die Virtuelle Forschungsumgebung, um den Wissenschaftlern eine Recherche und Nutzung bereits vorhandener Publikationen und Datenbestände direkt bei der Arbeit in der Forschungsumgebung zu ermöglichen. Hierfür werden:

1. relevante Bibliotheksbestände bereitgestellt und aufbereitet (Anreicherung der Metadaten mit Geoinformationen) und
2. die technologische Infrastruktur der Digitalen Bibliothek der UB weiterentwickelt (Datenmodell, Schnittstellen, Austauschformate, Suchinterface).

Erreichte Ergebnisse:

1. Für die Einbindung in die Virtuelle Forschungsumgebung ist es notwendig, Raumbezüge, die in historisch-kulturwissenschaftlichen Ressourcen meist durch geographische Namen hergestellt sind, explizit zu machen, indem z. B. Ortsschlagworte in bibliographischen Metadaten mit geographischen Koordinaten angereichert werden. Der zunächst wichtigste Bestand zur Einbindung in VKLandLab ist die landesgeschichtliche Zeitschrift „Mecklenburgische Jahrbücher“, deren Jahrgänge 1835-1940 im Rahmen eines DFG-Projekts (2002-04) im Volltext (XML) digitalisiert wurden.¹

Im ersten Arbeitsschritt wurde ein weitgehend vollständiger Referenzdatensatz von Ortsnamen des heutigen und historischen (seit 1800) Mecklenburg-Vorpommern neu aufgebaut.² Zum Anderen wurden Algorithmen entwickelt, um Listen von Ortsnamen automatisch gegeneinander abzugleichen und Übereinstimmungen / Ähnlichkeiten zu ermitteln, wobei verschiedene Verfahren (Einschränkung des Prüfbereichs anhand von groben Lageinformationen, Normalisierung von Schreibungen, Ermittlung von ähnlichen Schreibweisen mittels „Levenshtein-Distanz“) verwendet wurden. Der Aufbau des Referenzdatensatzes mit zurzeit ca. 7.000 Ortsnamen mit Geokoordinaten wurde abgeschlossen.³ Anschließend sind ca. 2000 Ortsschlagworte zu den Aufsätzen der Mecklenburgischen Jahrbücher aus dem Referenzdatensatz mit Geokoordinaten angereichert

¹ <http://mvdok.lbm.v.de> .

² Eine eingehende Recherche zu bestehenden Ortsverzeichnissen und Geodatenbanken hatte zum Ergebnis, dass diese entweder nicht digital vorliegen, eine zu geringe Auflösung (kleine Ortschaften) haben und/oder keine historische Dimension enthalten.

³ Zum Aufbau des Grundbestandes an Ortsnamen und Ortslagen für den Referenzdatensatz HONV wurden folgende digital vorliegende Datenbestände genutzt: 1. ein aktueller, amtlicher Datensatz: Verzeichnis der Ortschaften Mecklenburg-Vorpommern 2011 (Landesamt für innere Verwaltung M-V); 2. ein historischer Datensatz um 1800: Gemeindeverzeichnis des „GIS zur Mecklenburg-Schweriner Volkszählung von 1867 (und vergleichend 1819)“ (Gyula Papay, http://histdem.de/sites/default/files/gis1867/StartGIS1819_67.html); 3. community-basierter Datensatz mit Abbildung der administrativen Zusammenhänge in ihren historischen Veränderungen: Genealogisches Ortsverzeichnis (Verein für Computergenealogie, <http://gov.genealogy.net/>).

worden. Auf dieselbe Weise werden im zweiten Projekthalbjahr weitere Datenbestände mit Geokoordinaten ausgestattet.⁴

2. Die UB Rostock betreibt auf Basis der OpenSource-Software MyCoRe, an deren Entwicklung sie maßgeblich beteiligt ist, unter anderem die Repositoriensysteme Rostocker Dokumentenserver (RosDok), Digitale Bibliothek Neubrandenburg und Dokumentenserver der LBMV (MVDok). Um geographische Koordinaten in den Metadaten digitaler Dokumente speichern zu können, wurde das MyCoRe-Metadatenmodell erweitert. Hierzu wurde das MODS-Metadatenformat⁵ verwendet, das mit den Elementen <subject>, <geographics> und <cartographics> die nötigen Voraussetzungen liefert.

Weiterhin wurde ein einfacher Webdienst für die Dokumentenserver entwickelt, der URL-kodierte Suchanfragen nach einem Koordinatenbereich verarbeitet und für diesen Suchbereich relevante Dokumente zurückliefert, wobei die Trefferliste im Metadatenformat Dublin Core als JSON-Datei ausgegeben wird. Damit sind für den Beispielbestand der Mecklenburgischen Jahrbücher die Voraussetzungen für eine Einbindung in VKLandLab geschaffen.

Zugleich wurde eine erste prototypische geographische Suche direkt auf dem Dokumentenserver MVDok entwickelt, wo auf der Basis einer Google-Karte eine Koordinatenbereich ausgewählt und Dokumente für dieses Gebiet gesucht werden können.⁶ Diese Erweiterung der üblichen Recherchemöglichkeiten in Dokumentenrepositorien wird perspektivisch in die MyCoRe-Software integriert.

⁴ Es wurde bereits mit mehreren Einrichtungen/Projekten eine Zusammenarbeit besprochen. Geplant ist die Einbeziehung folgender Datenbestände: Landesbibliographie M-V (Titelnachweise landeskundlicher Literatur), MVDok (weitere Digitalisate, z. B. Mecklenburgische Monatshefte 1925-1943), Dokumentenserver der Universität Rostock und der Hochschule Neubrandenburg (elektronische Publikationen, Hochschulschriften), Wossidlo Digital Archive (DFG-Projekt, volkskundliche Belegzettel), Rostocker Forschungsverbund Historische Demographie (Volkszählungsdaten des 19. Jahrhunderts).

⁵ <http://www.loc.gov/standards/mods/>

⁶ <http://mvdok.lbm.de/nav?path=left.search.geo>