

**HTW** Chur

Hochschule für Technik und Wirtschaft  
University of Applied Sciences

## Churer Schriften zur Informationswissenschaft

Herausgegeben von  
Wolfgang Semar und Brigitte Lutz

---

Arbeitsbereich  
Informationswissenschaft

**Schrift 58**

Die Georeferenzierung von Katalogdaten  
mit Hilfe von Linked Open Data

Patrick Wermelinger

---

Chur 2013

**Churer Schriften zur Informationswissenschaft**

Herausgegeben von Wolfgang Semar und Brigitte Lutz

Schrift 58

**Die Georeferenzierung von Katalogdaten**

**mit Hilfe von Linked Open Data**

Patrick Wermelinger

Diese Publikation entstand im Rahmen einer Bachelor Thesis zum Abschluss Bachelor of Science (BSc) FHO in Informationswissenschaft.

Referent: Prof. Dr. Niklaus Stettler

Korreferent: Prof. Dr. Rudolf Mumenthaler

**Verlag:** Arbeitsbereich Informationswissenschaft

**ISSN:** 1660-945X

**Chur,** April 2013

## Kurzfassung

Vorliegende Arbeit setzt sich mit der Georeferenzierung von Katalogdaten auseinander. Sie zeigt auf, wie sich mit Hilfe von Linked Open Data bibliografische Metadaten mit geografischen Koordinaten anreichern lassen. Die Georeferenzierung soll nicht von Hand, sondern automatisiert umgesetzt werden und sich nicht auf Bilder und Karten beschränken, sondern alle Arten von Dokumenten miteinbeziehen. Als mögliche Datenquellen für die automatisierte Georeferenzierung dienen DBpedia, GeoNames und LinkedGeoData, deren Daten als RDF-Tripel zu Verfügung stehen. Skizziert werden verschiedene Anwendungen, bei denen Geobrowsing und das Serendipity-Prinzip eine Rolle spielen. Neben der Punkt-suche (Point of Interest) ist dabei die geografische Flächensuche (Area of Interest) von grosser Bedeutung. Das Hauptaugenmerk liegt dabei auf den Schlagwörtern der Gemeinsamen Normdatei (GND), insbesondere der Geografika und der Ländercodes. Die geografische Themensuche zeigt, wie sich durch die Verwendung des Ländercodes die Suchergebnisse verbessern lassen, wenn ein sinnvolles Ranking implementiert ist.

**Schlagwörter:** Bibliothek, Geobrowsing, geografische Suche, Georeferenzierung, Kataloganreicherung, Linked Open Data, RDF

## Abstract

This document exposes with georeferencing of catalogue data. It points out how bibliographic metadata can be enriched with geographic coordinates. Georeferencing is to be realised automatized and not manually and includes not just pictures and maps but all types of documents. Potential sources for an automatized georeferencing are DBpedia, GeoNames, and LinkedGeoData whose data are placed at the disposal in form of RDF triples. Various applications are outlined whereby geobrowsing and the principle of serendipity are an issue. In addition to geographic point research (point of interest) is the geographic area research (area of interest) of importance. The focus is on the subject headings of the Gemeinsame Normdatei (GND), particularly the geographica and the country codes. The geographic subject research illustrates how the country codes can be used to improve search results if a reasonable ranking is implemented.

**Keywords:** library, geobrowsing, geographical research, georeferencing, catalogue enrichment, Linked Open Data, RDF

## **Vorwort**

Ohne die Unterstützung verschiedener Personen wäre die vorliegende Arbeit so nicht möglich gewesen. Ich danke meinen Referenten Prof. Dr. Niklaus Stettler und Prof. Dr. Rudolf Mumenthaler für ihre Unterstützung bei der Ausarbeitung des Themas. Ganz herzlich dan-ken möchte ich Franziska Geisser (Projektverantwortliche von BIBonMAP, ETH-Bibliothek), Nicolas Bugnon (wissenschaftlicher Mitarbeiter von Viaticalpes) und Guy Thomas (Projektleiter von BM Archives, Mission 21), dass sie sich Zeit genommen haben, mir ihre spannenden Projekte zu erläutern und jederzeit für Fragen zur Verfügung gestan-den sind. Bedanken möchte ich mich weiter bei Alice Spinnler (Leitung Sacherschliessung, UB Basel) für ihre Ausführungen zur GND und den verschiedenen Schlagwortsystemen des IDS Basel Bern sowie bei Tobias Viegner (Projektkoordination Swissbib, UB Basel) für seine Erläuterungen zu Linked Open Data. Mein Dank geht auch an Susanne Zollinger (Mitglied der Projektgruppe Kartenportal.CH, ETH-Bibliothek) und Markus Oehrli (Mit-glied der Projektgruppe Kartenportal.CH, ZB Zürich) für ihre wertvollen Anregungen. Linus danke ich für seine kritischen Fragen und konstruktiven Anmerkungen. Mein ganz besonderer Dank gilt Jenny. Für alles.

## Inhaltsverzeichnis

Abbildungsverzeichnis.....	7
Tabellenverzeichnis.....	7
Abkürzungsverzeichnis.....	9
1 Einleitung .....	11
1.1 Ausgangslage .....	11
1.2 Fragestellung .....	14
1.3 Angestrebtes Ziel .....	14
1.4 Vorgehen .....	14
2 Georeferenzierung in Bibliotheken und Archiven.....	17
2.1 Viaticalpes.....	17
2.2 BIBonMAP .....	19
2.3 BM Archives.....	21
2.4 Kartenportal.CH .....	22
2.5 Fazit .....	25
3 Die Datenlage .....	27
3.1 Katalogdaten: Die drei geografischen Bezüge von Dokumenten .....	27
3.1.1 Erscheinungsort .....	28
3.1.2 Orte der Urheber .....	29
3.1.3 Orte des Inhalts.....	32
Exkurs I: Der Ländercode der GND .....	37
3.1.4 Fazit .....	40
3.2 Linked Open Data .....	41
3.2.1 DBpedia .....	47
3.2.2 GeoNames.....	49
3.2.3 LinkedGeoData .....	51
3.2.4 Fazit .....	53
4 Die Georeferenzierung .....	55

4.1	Varianten der Anreicherung.....	55
4.1.1	Anreicherung der Katalogdaten .....	55
4.1.2	Anreicherung bei der Suchabfrage .....	56
4.1.3	Anreicherung des Index .....	57
4.2	Mögliche Anwendungen .....	57
4.2.1	Atlas der Basiliensia .....	58
4.2.2	Die geografische Suche alter Drucke .....	60
4.2.3	Die geografische Themensuche .....	61
	Exkurs II: Verbundübergreifende Kataloganreicherung mit SWD-Schlagwörtern .....	64
5	Empfehlungen .....	67
6	Schlussbemerkungen .....	69
7	Quellenverzeichnis .....	71
7.1	Mündliche Quellen .....	71
7.2	Schriftliche Quellen .....	71
8	Anhang.....	75

## Abbildungsverzeichnis

Abbildung 1: The Linking Open Data cloud diagram.....	13
Abbildung 2: Screenshot der Suchergebnisse für Basel in ViaticAlpes .....	18
Abbildung 3: Screenshot von Amersfoort op de kaart.....	20
Abbildung 4: Die räumliche Suche der geografischen Kartensuche.....	24
Abbildung 5: Screenshot User Interface der geografischen Kartensuche .....	25
Abbildung 6: Screenshot Titelaufnahme im MARC-Format.....	27
Abbildung 7: Autoritätsdatensatz eines Druckorts.....	29
Abbildung 8: Screenshot Normeintrag Autor.....	30
Abbildung 9: Autoritätsdatensatz einer Körperschaft .....	31
Abbildung 10: Autoritätsdatensatz einer Druckerei .....	32
Abbildung 11: GND-Normdatensatz eines Geografikum .....	34
Abbildung 12: Screenshot Index Kongress-/Ausstellungsort.....	36
Abbildung 13: Screenshot GND-Geografikum Schwäbische Alb .....	38
Abbildung 14: Aufbau eines Tripels .....	42
Abbildung 15: Illustration eines RDF-Tripels.....	42
Abbildung 16: Beispiel einer RDFS-Repräsentation in Turtle.....	43
Abbildung 17: Beispiel einer RDF/XML-Repräsentation.....	45
Abbildung 18: Beispiel einer FOAF-Repräsentation.....	45
Abbildung 19: Beispiel einer Geo-Repräsentation .....	46
Abbildung 20: Die RDF-Repräsentation der Stadt Embrun in GeoNames.....	50
Abbildung 21: Screenshot Bounding Box Schweiz. ....	53
Abbildung 22: Varianten der Anreicherung durch Linked Open Data.....	55
Abbildung 23: Beispiel eines Drilldown-Menüs für Druckereien .....	60
Abbildung 24: Bounding Boxes bei der Recherche.....	63
Abbildung 25: Drilldown-Facette aus HEIDI.....	75

## Tabellenverzeichnis

Tabelle 1: Ländercodes in der SWD.....	35
Tabelle 2: Ergebnisse von Testrecherchen im SWB.....	39

## Abkürzungsverzeichnis

ALEPH	Automated Library Expandable Program Hebrew University of Jerusalem
API	Application Programming Interface
AUT	Autoritätsdatensatz
BIB	Berufsverband Information Bibliothek
BIBO	Bibliographic Ontology
CSV	Comma-Separated Values
DDC	Dewey Decimal Classification
DMA-EST	Einheitssachtitel-Datei des Deutschen Musikarchivs
DNB	Deutsche Nationalbibliothek
DSV	Deutscheschweizer Verbund
ETH	Eidgenössische Technische Hochschule
FOAF	Friend Of A Friend
GKD	Gemeinsame Körperschaftsdatei
GND	Gemeinsame Normdatei
HeBIS	Hessisches Bibliotheksinformationssystem
HEIDI	Heidelberger Bibliotheks-Informationssystem
HLS	Historisches Lexikon der Schweiz
HTML	Hypertext Markup Language
IDS	Informationsverbund Deutschschweiz
JSON	JavaScript Object Notation
KIDS	Katalogisierungsregeln IDS
LCCN	Library of Congress Control Number
LGD	LinkedGeoData
LOD	Linked Open Data
MAB	Maschinelles Austauschformat für Bibliotheken
MARC	Machine-Readable Cataloging
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting

OPAC	Online Public Access Catalogue
OSM	OpenStreetMap
PND	Personennamendatei
POI	Point Of Interest
RDA	Resource Description and Access
RDF	Ressource Description Framework
RDFS	RDF Schema
RERO	Réseau Romand (Réseau des bibliothèques de Suisse occidentale)
RERO DOC	Digitale Bibliothek des RERO
REST	Representational State Transfer
RVK	Regensburger Verbundklassifikation
SIKART	Lexikon und Datenbank zur Kunst in der Schweiz und im Fürstentum Liechtenstein
SPARQL	SPARQL Protocol And RDF Query Language
SQL	Structured Query Language
SWB	Südwestdeutscher Bibliotheksverbund
SWD	Schlagwortnormdatei
Turtle	Terse RDF Triple Language
UB	Universitätsbibliothek
URI	Uniform Resource Identifier
URL	Uniform Resource Locater
URN	Uniform Resource Name
VIAF	Virtual International Authority File
W3C	World Wide Web Consortium
WGS84	World Geodetic System 1984
WWW	World Wide Web
XML	Extensible Markup Language

# 1 Einleitung

## 1.1 Ausgangslage

Online-Kartendienste wie Google Maps<sup>1</sup> oder OpenStreetMap<sup>2</sup> erfreuen sich grosser Beliebtheit. Bei der Recherche nach Orten, Adressen, Restaurants und anderen raumbezogenen Daten oder als Routenplaner sind die interaktiven Karten nicht mehr aus dem Alltag wegzudenken. Auch verschiedene Onlineversionen von Zeitungen oder Nachrichtenmagazinen haben den Nutzen dieser Technologie erkannt und angefangen, bestimmte Inhalte in georeferenzierter Form zur Verfügung zu stellen. So bietet eine interaktive Karte auf *Spiegel Online* im Ressort Reise unter dem Titel „Die Welt entdecken“<sup>3</sup>, Reisereportagen, Fotostrecken und Reisetipps an, die mittels verschiedenfarbiger Marker verortet sind. Eine fragwürdige Anwendung findet sich in der Online-Ausgabe der *Basler Zeitung*. Unter der reisserischen Überschrift „Die Basler Achse der Gewalt“<sup>4</sup> sind auf einer interaktiven Karte alle von der Staatsanwaltschaft publizierten Straftaten des Jahres 2011 verzeichnet, die sich lokalisieren liessen. Verschiedenfarbige Marker stellen verschiedene Deliktarten wie Gewaltverbrechen, Raubüberfälle, Sachbeschädigungen sowie Einbrüche und Diebstähle dar und sind (falls vorhanden) mit den entsprechenden Meldungen verknüpft.

Einzelne Bibliotheken und Archive haben damit begonnen, diese Technologie zu nutzen. Sie georeferenzieren ausgewählte Teilbestände mit räumlichem Bezug, um sie den Benutzern durch neue Suchmöglichkeiten zugänglich zu machen. Im Projekt Viaticalpes der Université de Lausanne sind Ansichten aus alpinen Reiseberichten von der Renaissance bis zum 19. Jahrhundert auf einer virtuellen Karte verortet und unter anderem nach geografischen Gesichtspunkten recherchierbar. Im Rahmen eines e-lib-Projekts wurde auf Kartenportal.CH eine geografische Suche umgesetzt, die die Recherche nach topografischen Karten erheblich vereinfacht. In beiden Fällen wird die klassische Katalog-Recherche um neue Suchmöglichkeiten erweitert. Die entwickelten Recherche-Tools bedienen sich unter anderem der Zoom- und Drag-Funktionen, wie sie die Online-Kartendienste verwenden und ermöglichen damit das Geobrowsing.

Diese Anwendungen erfüllen eine Forderung, die Wiesenmüller (2012, S. 22) in einem fünf Thesen umfassenden Aufsatz formuliert. Sie verlangt unter anderem, dass Kataloge Browsing und entdeckendes Suchen ermöglichen und verbessern müssen:

---

<sup>1</sup> <http://maps.google.ch/>. Daneben bieten weitere Unternehmen Kartendienste an: so z.B. Yahoo (<http://maps.yahoo.com/>) oder Microsoft mit Bing Maps (<http://www.bing.com/maps/>) [04.07.2012].

<sup>2</sup> <http://www.openstreetmap.org/> [04.07.2012].

<sup>3</sup> <http://www.spiegel.de/reise/karte/index.html> [02.07.2012].

<sup>4</sup> <http://bazonline.ch/basel/stadt/Die-Basler-Achse-der-Gewalt/story/29697970> [02.07.2012].

*„Im Sinne des „Serendipity“-Prinzips sollen Nutzer(innen) über den Katalog immer wieder auf Dinge stossen, die für sie interessant und relevant sind, ohne dass sie gezielt danach gesucht hätten“ (ebd.).*

Neben neuen Funktionen, wie beispielsweise dem Recommendersystem „BibTip“ oder der Anzeige von inhaltlich ähnlicher Literatur, regt sie an, das Semantic Web zu nutzen, um weiterführende Informationen in die Volltitelanzeige einzubinden (ebd., S. 22-23). An anderer Stelle schreibt Wiesenmüller (2011, S. 335f.), dass Linked Open Data eine wichtige Rolle spielen können, wenn es darum geht, den „Nutzwert der Daten“ zu erhöhen. Den grossen Vorteil dieser Daten sieht sie in der Tatsache, dass sie sich aufgrund ihrer Datenstruktur automatisiert nutzen lassen.

Die Idee des Semantic Web geht zurück auf Tim Berners-Lee, der 2001 in einem vielbeachteten Beitrag im amerikanischen Fachmagazin *Scientific American*, seine Vision einer zukünftigen Informationswelt skizziert. In dieser soll der Computer auf der Grundlage von strukturierten Daten Entscheidungen treffen können (Fürste 2011, S. 45). Berners-Lee et al. (2001, S. 37) beschreiben das Semantic Web im Verhältnis zum World Wide Web:

*„The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation.“*

Die Daten sollen so strukturiert werden, dass sie eine explizite Bedeutung haben und von Maschinen interpretiert werden können. Daten, die diese Voraussetzungen erfüllen, werden Linked Data genannt; sind sie frei verfügbar, handelt es sich um sogenannte Linked Open Data (LOD). Die Strukturierung der Daten beruht auf zwei Technologien:

*„Two important technologies for developing the Semantic Web are already in place: eXtensible Markup Language (XML) and the Resource Description Framework (RDF)“ (ebd., S. 38).*

RDF wird von Heath und Bizer (2011, S. 15) als „*lingua franca*“ des Semantic Web bezeichnet, die fähig ist, zwischen den verschiedenen Datenmodellen zu vermitteln, die im Web verwendet werden.

Seit einigen Jahren setzen sich Bibliotheken zunehmend mit dem Thema Linked Open Data auseinander. Am Bibliothekartag 2012 in Hamburg war Linked Open Data eines der Schwerpunktthemen<sup>5</sup>. Neben der Distribution von bibliothekarischen Daten ging es auch um die Nutzung der grossen Menge an Daten, die im RDF-Standard verfügbar sind. Gerade in den Bereichen Erschliessung und Datenkonsolidierung wird ein grosses Potenzial gesehen, das bereits in einigen Anwendungen zum Tragen kommt.

---

<sup>5</sup> Einen Überblick bietet der BIB Opus-Server, wo alle gehaltenen Vorträge als PDF verfügbar sind: <http://www.bib-info.de/verband/publikationen/opus/hamburg-2012.html> [02.07.2012].

Wie gross die Menge an veröffentlichten LOD-Datensets mittlerweile ist, zeigt Abbildung 1. (Stand September 2011). Die Daten aus dem bibliothekarischen Umfeld sind im grünen rechten Teil angesiedelt. So wurden beispielsweise die Katalogdaten des B3Kat<sup>6</sup>, des HeBIS-Verbundkatalogs<sup>7</sup> und der Universitätsbibliothek Mannheim<sup>8</sup> veröffentlicht. Die Deutsche Nationalbibliothek hat sowohl die Titeldaten als auch die GND<sup>9</sup> und die British Library die British National Bibliography<sup>10</sup> als Linked Open Data publiziert. Weiter sind die VIAF<sup>11</sup> sowie die Library of Congress Subject Headings<sup>12</sup> verfügbar.

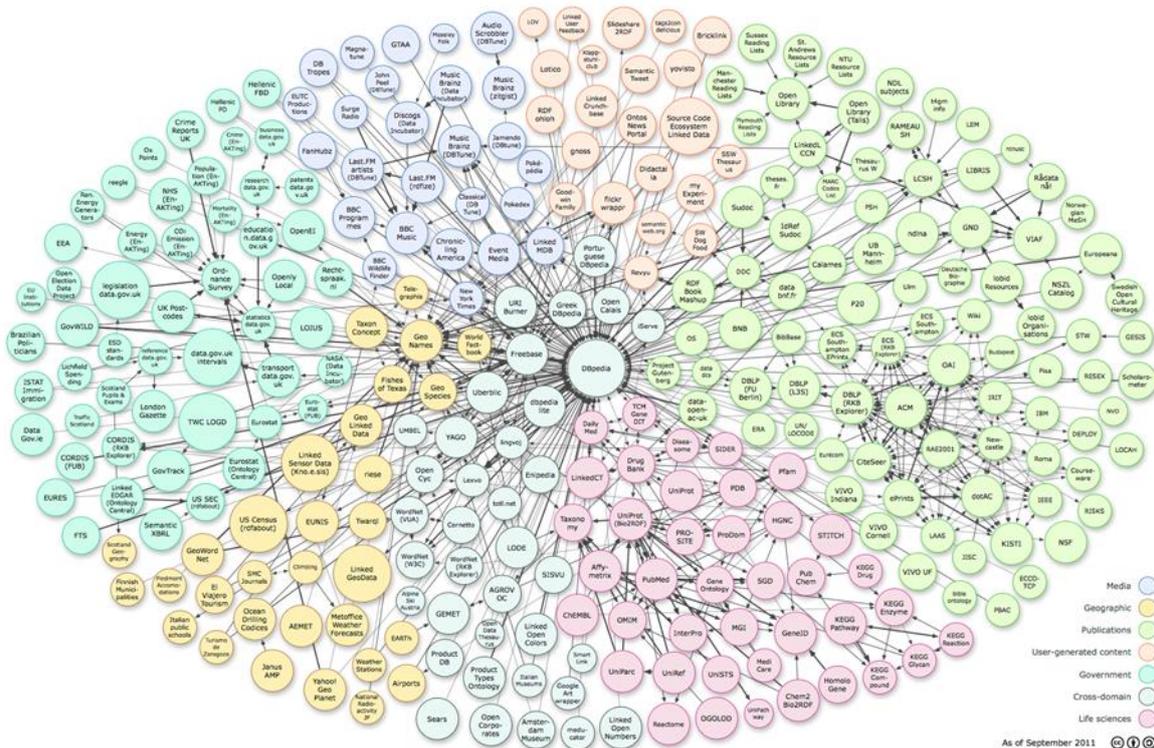


Abbildung 1: The Linking Open Data cloud diagram (Cygniak, 2011)

Mit Datenbanken wie beispielsweise DBpedia, GeoNames und LinkedGeoData liegen umfangreiche geografische Daten als Linked Open Data vor. Durch die Möglichkeit, diese automatisiert zu nutzen, bieten sie grosses Potenzial für die Georeferenzierung von Katalogdaten. Wie diese Daten für eine bessere Erschliessung von Bibliotheksbeständen genutzt werden können, ist Gegenstand der vorliegenden Arbeit.

<sup>6</sup> <http://lod.b3kat.de/doc> [25.07.2012].

<sup>7</sup> [http://www.hebis.de/de/1ueber\\_ uns/projekte/lod/lod\\_index.php](http://www.hebis.de/de/1ueber_ uns/projekte/lod/lod_index.php) [25.07.2012].

<sup>8</sup> <http://data.bib.uni-mannheim.de/> [25.07.2012].

<sup>9</sup> [http://www.dnb.de/DE/Service/DigitaleDienste/LinkedData/linkedata\\_node.html](http://www.dnb.de/DE/Service/DigitaleDienste/LinkedData/linkedata_node.html) [25.07.2012].

<sup>10</sup> <http://bnb.data.bl.uk> [25.07.2012].

<sup>11</sup> <http://viaf.org/viaf/data/> [25.07.2012].

<sup>12</sup> <http://id.loc.gov/> [25.07.2012].

## 1.2 Fragestellung

Konkret soll untersucht werden, wie die Georeferenzierung von ausgewählten Teilbeständen mit Hilfe von Linked Open Data aussehen kann. Im Gegensatz zu den bereits eingangs erwähnten Umsetzungen von Viaticalpes und Kartenportal.CH sollen die Dokumente nicht von Hand, sondern automatisiert georeferenziert werden.

Dafür sollen unter anderem Antworten auf folgend Fragen gefunden werden: Welche Datensets enthalten Koordinaten? Wie ist die Qualität der Daten? Wie soll die Anreicherung der Katalogdaten umgesetzt werden? Was sind die Stolpersteine, was die Grenzen der Georeferenzierung? Welche Anwendungen sind denkbar? Was ist der Mehrwert für die Benutzer, wenn Bestände geografisch erschlossen werden? Welche Empfehlungen an Bibliotheken lassen sich formulieren?

## 1.3 Angestrebtes Ziel

Vorliegende Arbeit verfolgt das Ziel, das Potenzial aufzuzeigen, das in der Georeferenzierung von Katalogdaten steckt. Es soll untersucht werden, wie sich Katalogdaten mit Hilfe von Linked Open Data mit Koordinaten anreichern lassen und welche Daten sich dafür eignen. Auf Grundlage dessen werden verschiedene Umsetzungsmöglichkeiten von Recherche-Tools für georeferenzierte Bestände skizziert. Der Fokus liegt dabei auf der Entwicklung von Konzepten für mögliche Anwendungen und nicht auf der technischen Realisierung. Gezeigt werden soll, wie bisher ungenutzte Informationen in den Katalogdaten für eine Verbesserung der Suchergebnisse genutzt werden können. Rechtliche Fragen bezüglich der Nutzung von Linked Open Data bleiben aussen vor.

Ein besonderer Fokus bei der Konzeptionierung von geografischen Recherche-Tools liegt dabei auf der Flächensuche. Bei dieser spielen Polygone, sogenannte Bounding Boxes, oder Minimum bounding rectangles eine wichtige Rolle. Bounding Boxes sind kleinste achsenparallel umschliessende Rechtecke, die durch zwei Koordinatenpaare ( $x_{min}$ ,  $y_{min}$ ) und ( $x_{max}$ ,  $y_{max}$ ) definiert sind. Aufgrund ihrer sehr kompakten Approximation von kompliziert abgegrenzten Objekten, lassen sie sich für geometrische Algorithmen sowie für Speicher- und Zugriffsmechanismen zur Performancebeschleunigung verwenden (Bill und Zehner 2001, S. 178).

## 1.4 Vorgehen

Auf Grundlage der Analyse verschiedener Georeferenzierungsprojekten und der vorhandenen Datenbasis, werden neue mögliche Anwendungen für die geografische Suche in Bibliotheksbeständen entworfen. In Kapitel 2 erfolgt eine Bestandsaufnahme der durch Viaticalpes, BIBonMAP, BM Archives und Kartenportal.CH verfolgten Lösungsansätze. Die

Ausführungen stützen sich zu einem wesentlichen Teil auf die mündliche Auskunft der projektverantwortlichen Personen. Danach werden in Kapitel 3.1 die Katalogdaten hinsichtlich ihrer Eignung für die Georeferenzierung analysiert. Die Datenbasis bildet der ALEPH-Katalog des IDS Basel Bern<sup>13</sup>, insbesondere der Bestand der UB Basel. Es handelt sich dabei um bibliografische Metadaten im MARC-Format. Kapitel 3.2 wendet sich den Linked Open Data zu. Nach einer allgemeinen Einführung, in der die wichtigsten Begriffe, Standards und Prinzipien dieser Technologie erläutert werden, sollen DBpedia, GeoNames und LinkedGeoData auf ihre Tauglichkeit als mögliche Datenlieferanten untersucht werden. Die Georeferenzierung ist Thema von Kapitel 4. Zuerst werden in Kapitel 4.1 drei Varianten der Anreicherung durch Koordinaten geprüft, bevor in Kapitel 4.2 die Konzepte von drei möglichen Anwendungen vorgestellt werden. Neben verschiedenen Suchfunktionalitäten spielt das Ranking eine wichtige Rolle. Ausführlich diskutiert werden zudem die Vor- und Nachteile der geografischen Flächensuche. Die gewonnenen Erkenntnisse münden in Empfehlungen an Bibliotheken, die zusammenfassen worauf bei Georeferenzierungsprojekten zu achten ist (Kapitel 5). Im abschliessenden Kapitel 6 werden die Ergebnisse der Arbeit thesenartig zusammengefasst und kritisch überprüft.

---

<sup>13</sup> <http://aleph.unibas.ch/menu.html> [310.07.2012].



## 2 Georeferenzierung in Bibliotheken und Archiven

Im Folgenden sollen vier Projekte vorgestellt werden, die mit selbstentwickelten Recherche-Tools georeferenzierte Dokumente suchbar machen. Viaticalpes und die geografische Kartensuche des Kartenportals.CH sind bereits realisiert, während sich BIBonMaps und BM Archives noch in Planung befinden. Im Zentrum der Untersuchung stehen Fragen wie: Welche Bestände wurden ausgewählt? Wie wurden die Dokumente georeferenziert? Auf welchen Technologien fussen die Anwendungen? Auf Grundlage der Analyse sollen neue Ideen für weitere Teilbestände entwickelt werden.

### 2.1 Viaticalpes

Viaticalpes<sup>14</sup> ist ein Projekt das die Faculté des Lettres der Universität Lausanne mit Unterstützung des Schweizerischen Nationalfonds realisiert hat. Aus einem Korpus von 850 illustrierten Reiseberichten des Alpenraums aus der Zeit des 16. bis 19. Jahrhundert wurden die 100 repräsentativsten ausgewählt. Die darin enthaltenen rund 3'000 Illustrationen wurden beschrieben, indexiert, georeferenziert und mit den zugehörigen Textstellen der Reiseberichte verlinkt. Die Bilder sowie die Berichte wurden von den beteiligten Bibliotheken digitalisiert (Université de Lausanne 2012a). Neben der Bibliothèque Cantonale et Universitaire Lausanne sind dies die Bibliothèque de Genève, die Bibliothèque Publique et Universitaire Neuchâtel, die Mediathek Wallis, die Schweizerische Nationalbibliothek und die Bürgerbibliothek in Bern (Université de Lausanne 2012b).

Die verwendete Datenbank Viatimages ist eine Entwicklung der Universität Lausanne in Zusammenarbeit mit der Universität Paris-Sorbonne. Für Hosting und Unterhalt ist die Universität Lausanne zuständig. Verwendung fand das als Open-Source-Anwendung verfügbare relationale Datenbankmanagementsystem MySQL. Mittels Google Maps API wurde Google Maps in die Website integriert (Mündliche Auskunft von Nicolas Bugnon, 06.06.2012).

Grundsätzlich bieten sich dem Benutzer zwei verschiedene Suchmöglichkeiten: eine Wortsuche sowie eine geografische Recherche. Die Wortsuche bietet die Auswahl zwischen der einfachen und der erweiterten Suche. Zusätzlich können über Dropdown-Listen Sachgebiete (domaines) oder die Werke angewählt werden. Darüber hinaus steht ein Sucheinstieg mittels Schlagwortwolke (tag cloud) zur Verfügung. Die geografische Suche stellt neben der freien Wortsuche mittels Toponymen auch die Möglichkeit per Dropdown-Listen Regionen, Städte, Seen, Gletscher sowie die einzelnen Reiseberichte auszuwählen. Die Suchergebnisse werden auf Google Maps angezeigt, auf der sich in gewohnter Weise navigieren lässt.

---

<sup>14</sup> <http://www.unil.ch/viaticalpes> [21.06.2012].

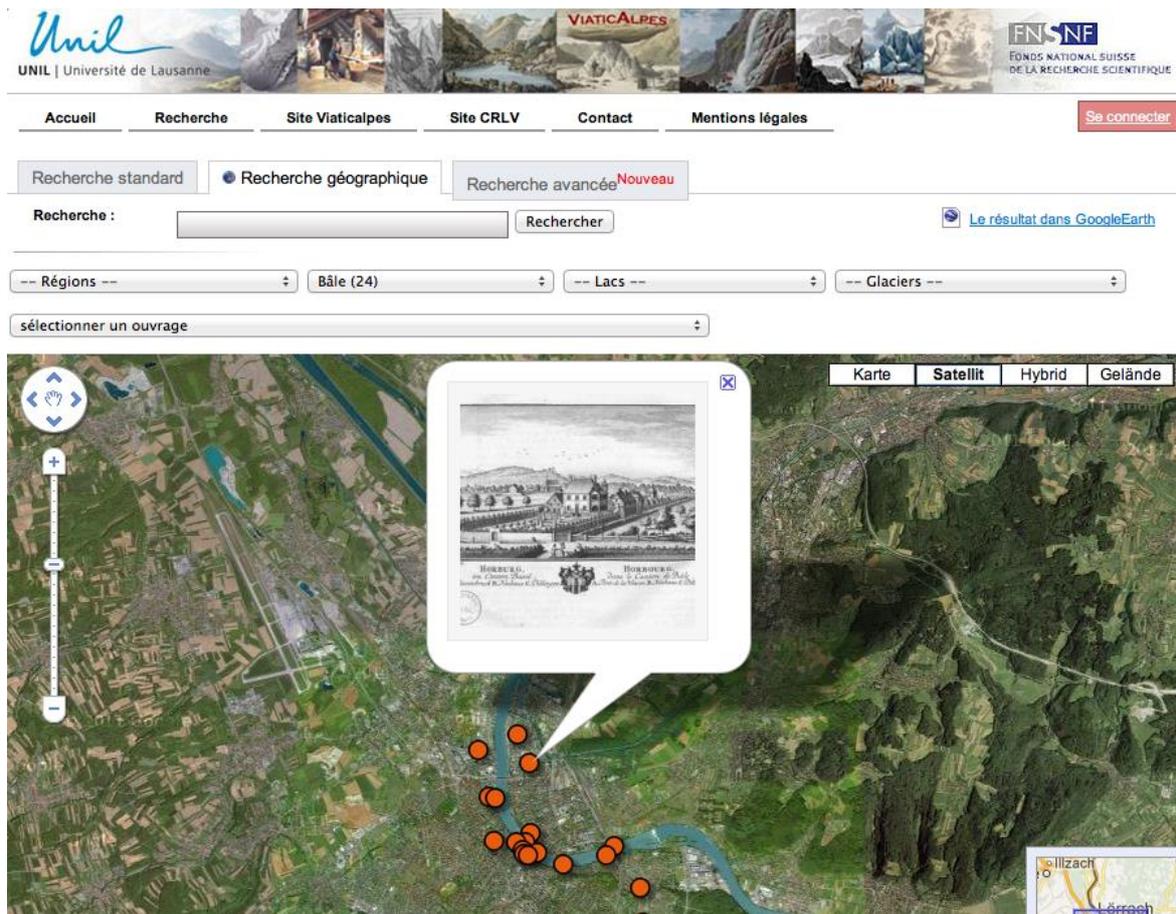


Abbildung 2: Screenshot der Suchergebnisse für Basel in Viaticalpes [22.06.2012]

Während bei der Wortsuche direkt die entsprechenden Bilder angezeigt werden, stellt sich die Trefferliste der geografischen Suche als orange Punkte auf der Karte dar. Der Benutzer sieht somit auf einen Blick, von welchen Orten Ansichten vorhanden sind. Durch Klicken auf einen der Punkte, erscheint ein Vorschaubild (siehe Abbildung 2). Wird dieses angeklickt, erhält man eine grössere, zoombare Version sowie die Metadaten: unter anderem Urheber und Titel des Bildes, Sachgebiete, Schlagworte, Titel und Autor des Werks, in dem das Illustration enthalten ist, die genaue Seitenzahl sowie ein Digitalisat des Volltextes, der sich auf das Bild bezieht. Zudem bestehen Links auf die Katalogaufnahmen sowie (falls vorhanden) auf die Digitalisate der vollständigen Reiseberichte z.B. in RERO DOC<sup>15</sup> (digitale Bibliothek des Westschweizer Bibliothekverbundes RERO), Gallica<sup>16</sup> (digitale Bibliothek der Französischen Nationalbibliothek) oder Google Books<sup>17</sup>. Verlinkt sind auch die Autoren der Reiseberichte sowie die Urheber der Bilder (Illustratoren, Stecher etc.) mit entsprechenden biografischen Artikeln in der Online-Ausgabe des Historischen Lexikon der Schweiz (HLS)<sup>18</sup>

<sup>15</sup> <http://doc.rero.ch/> [21.06.2012].

<sup>16</sup> <http://gallica.bnf.fr/> [21.06.2012].

<sup>17</sup> <http://books.google.ch/> [21.06.2012].

<sup>18</sup> <http://hls-dhs-dss.ch/> [21.06.2012].

oder in SIKART<sup>19</sup> (Lexikon und Datenbank zur Kunst in der Schweiz und im Fürstentum Liechtenstein).

Verschiedene Thesauri wurden eigenständig erstellt. So gibt es eine Liste von 30 Sachgebieten (domaines) wie beispielsweise „architecture et monuments“, „état et politique“, „géographie, cartographie et topographie“, „paysage“ und „transports et routes“. Weiter existiert ein geografischer Thesaurus, der in verschiedene, politisch und physische Klassen unterteilt ist und Deskriptoren für „continent“, „aire géographique“, „état“, „région/canton“, „toponyme/localité“ (für alle dargestellten Orte wie Gemeinde, Städte etc.), „montagne“, „vallée“, „lac“, „glacier“ und „cours d'eau“ enthält. Der thematische Thesaurus umfasst 300 Schlagworte, die je zur Hälfte aus Deskriptoren und Synonymen in französischer Sprache bestehen. Die Multilingualität ist geplant (Mündliche Auskunft von Nicolas Bugnon, 06.06.2012).

Die Georeferenzierung war eine sehr aufwändige Arbeit und wurde von Hand mit Hilfe von Google Maps API umgesetzt. In den meisten Fällen erfolgte die Lokalisierung motivbezogen, d.h. die Marker wurden mit dem Motiv verknüpft. Nur in Ausnahmefällen, namentlich bei Weitwinkelpanoramen, sind die Ansichten am Standpunkt verortet, von dem aus das Motiv aufgenommen bzw. abgebildet wurde. Insgesamt sind ca. 70-80% der Bilder georeferenziert. Bei den übrigen Bildern handelt es sich um mythologische Darstellungen, um Porträts oder Abbildungen von Trachten (ebd.).

## 2.2 BIBonMAP

Ähnlich wie das Projekt Viaticalpes wollte sich die ETH-Bibliothek in ihrem Projekt BIBonMAP (Arbeitstitel) ursprünglich auf Reiseberichte, in diesem Fall des 18. und 19. Jahrhunderts, fokussieren. Die Idee war es, virtuelle Reiserouten durch ausgewählte Schweizer Orte auf einer Karte darzustellen und über diesen ortsbezogenen Zugang neben den Beschreibungen der Reisenden weitere, auch zeitgenössische Bilder, Texte und sonstige Dokumente über diese Orte aus dem Bestand der ETH-Bibliothek einer breiten Öffentlichkeit zugänglich zu machen (Mündliche Auskunft von Franziska Geisser, 12.06.2012).

Im Juni 2011, neun Monate nach Projektstart, erfolgte eine Neuausrichtung des Projekts. An die Stelle der virtuellen Reiserouten trat die Idee einer Geobrowsing-Plattform, bei der auf einer interaktiven Karte prinzipiell alle Medien des eigenen Bestandes mit eindeutig geografischem Bezug lokalisiert werden können. Wichtige Voraussetzung dabei ist, dass Digitalisate der entsprechenden Informationsträger vorhanden sind. In einer ersten Phase

---

<sup>19</sup> <http://www.sikart.ch/> [21.06.2012].

wurden rund 1'000 Bilder mit Bezug zur Schweiz aus dem rund 170'000 Bilder umfassenden e-pics Bildarchiv<sup>20</sup> der ETH-Bibliothek ausgewählt (ebd.).

Mit der Georeferenzierung von Bildern widmet sich BIBonMAP nicht nur dem gleichen Medium wie Viaticalpes, die Anwendung basiert durch die Nutzung von Google Maps API zudem auf der gleichen Technologie. Eine weitere Gemeinsamkeit ist die geografische Verortung der Bilder, die wie im Lausanner Modell von Hand geschieht. Auch in Zürich hat sich gezeigt, dass die Georeferenzierung eine sehr zeitintensive Arbeit ist und es nicht immer ganz einfach ist, zu entscheiden, ob ein Bild mit dem Motiv oder dem Standort der Aufnahme verknüpft werden soll.

Doch im Projekt BIBonMAP haben sich aufgrund der Menge der Ansichten, die sich alle auf ein und dasselbe Objekt beziehen, Probleme mit der Übersichtlichkeit ergeben. So sind beispielsweise zahlreiche Ansichten des ETH-Hauptgebäudes vorhanden. Jedes einzelne Bild wurde möglichst genau georeferenziert, d.h. dem exakten Gebäudeteil zugeordnet. Das Ergebnis war ein Meer von roten Markern rund um das Gebäude, das unübersichtlich und alles andere als nutzerfreundlich in Handhabung und Orientierung war (ebd.).

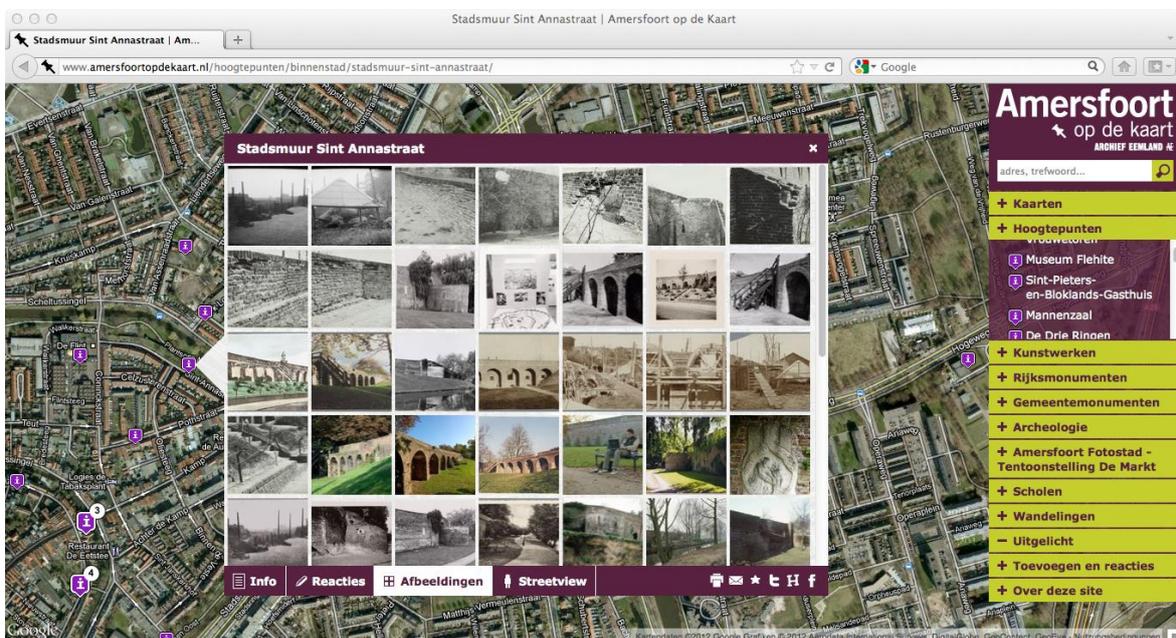


Abbildung 3: Screenshot von Amersfoort op de kaart [23.06.2012]

Auf der Suche nach einer Lösung wurde man bei Amersfoort op de Kaart<sup>21</sup> fündig, einem Projekt das im Jahr 2009 vom Archief Eemland anlässlich des 750. Geburtstags der Stadt Amersfoort verwirklicht worden war. Auf einer Google Map sind kulturhistorische Informationen und Ansichten der Stadt lokalisiert und mit den entsprechenden Orten (Gebäude, Monumente etc.) verknüpft (Archief Eemland 2012). Bei dieser Umsetzung wurde

<sup>20</sup> <http://ba.e-pics.ethz.ch/> [22.06.2012].

<sup>21</sup> <http://www.amersfoortopdekaart.nl/> [22.06.2012].

nicht jedes Bild einzeln georeferenziert, sondern Cluster gebildet. Alle Ansichten und Informationen, die zu einem bestimmten Gebäude, Denkmal etc. gehören, sind zusammengefasst und mit einem einzigen Marker auf der Karte lokalisiert.<sup>22</sup> Der Mouseover-Text bietet neben einem Vorschaubild den Namen des entsprechenden Objekts. Durch Klicken auf den Marker öffnet sich ein Layer, in dem der Benutzer unter anderem weitere Informationen und Abbildungen anzeigen lassen kann. Das Beispiel in Abbildung 3 zeigt einen Teil der Bilder die zur Stadtmauer Sint Annastraat verfügbar sind.

Trotz dieser eleganten Möglichkeit sollen bei BIBonMAP nicht alle vorhandenen Bilder zu einem bestimmten Ort, sondern nur eine Auswahl angezeigt werden. Die vollständige Sammlung digitalisierter Bilder zum jeweiligen Ort ist über einen Link ins Wissensportal erreichbar. Neben der Möglichkeit des geografischen Browsens sind verschiedene thematische Filter wie „Brücken“, „Berge“ oder „Gebäude“ vorgesehen sowie eine Zeitschiene, mit der sich der Zeitraum eingrenzen lässt (Mündliche Auskunft von Franziska Geisser, 12.06.2012).

Die historischen Reiserouten sind aber noch nicht vom Tisch, sondern sollen als speziell aufbereitete „Dossiers“ (z. Zt. ca. 100 Routen) auf der Karte dargestellt und gegebenenfalls mit historischen Bildern und Karten aus dem restlichen Bestand der ETH-Bibliothek angereichert werden (ebd.).

### 2.3 BM Archives

Ein Georeferenzierungsprojekt plant auch die Mission 21 in Basel im Rahmen ihres Online-Tools BM Archives<sup>23</sup>, das der „Präsentation und Vermittlung des historisch-kulturellen Erbes im Archiv der Basler Mission“ (Thomas 2012, S. 1) dienen soll. Vorgesehen ist die räumliche Verortung von Karten, Bildern und Missionsberichten.

Das Archiv der Basler Mission besitzt rund 9'000 historische Landkarten, Pläne und Skizzen in gedruckter und handschriftlicher Form, die zurzeit erschlossen werden. Eine Auswahl von ca. 6'300 Objekten soll digitalisiert werden. Neben einer vollständigen Sicherheitsverfilmung auf Makrofischen ist auch die Restaurierung von „ca. 100 besonders einzigartigen und repräsentativen Landkarten, Plänen und Skizzen“ (ebd., S. 2) vorgesehen. Das Bildarchiv der Mission 21 umfasst gut 30'000 Fotografien aus der Zeit zwischen 1850 und 1945, die in

---

<sup>22</sup> Eine ähnliche Lösung im Umgang mit vielen Markern auf kleinem Raum, hat die Mediathek Wallis umgesetzt. Das audiovisuelle Kulturgut des Kantons ist auf einer interaktiven Karte georeferenziert (<http://archives.memovs.ch/geoloc/> [22.06.2012]). Verwendet wurden verschiedenfarbige Marker mit unterschiedlichen Icons: roter Marker für Fotos, hellblauer Marker für Videos, dunkelblauer Marker für Tondokumente und grüne Marker, die als Sammelmarker (Cluster) dienen. Klickt man ein grünes Symbol an, wird in die Karte hineingezoomt und einzelne Marker anderer Farben werden angezeigt.

<sup>23</sup> <http://bmarchives.org/> [09.07.2012].

der Bildatenbank<sup>24</sup> der University of Southern California Libraries verfügbar sind (Mündliche Auskunft von Guy Thomas, 29.05.2012).

Im Zuge der Erschliessung wird ein Thesaurus erstellt, der sich an den Missionsfeldern der Basler Mission orientiert und ein Abbild der Organisationsstruktur sein soll. Der Thesaurus soll also die geografischen Schwerpunkte der Basler Mission widerspiegeln, die laut Thomas (2012, S. 1) „in Ghana, Kamerun, Südindien, Südchina und Kalimantan liegen. Er orientiert sich weniger an einer politischen Ordnung, sondern ist aus Begriffen der Berichtersteller aufgebaut. Eine Herausforderung dabei ist, die geografischen Begriffe zeitlich zu synonymisieren und die Mission Districts den modernen Toponymen zuzuordnen (Mündliche Auskunft von Guy Thomas, 29.05.2012).

Die Georeferenzierung der Bilder, Karten und (in einem späteren Schritt) auch der Missionsberichte soll auf Grundlage des Thesaurus erfolgen. Wie genau, ist allerdings noch unklar. Vorgesehen ist auch in diesem Fall ein Map Viewer, der auf Google Maps basiert und auf dem öffentlichen Google Maps API aufbaut (Ley 2009, S. 17).

## 2.4 Kartenportal.CH

Die geografische Kartensuche<sup>25</sup> ist Teil des e-lib.ch-Projekts Kartenportal.CH und wurde von der Firma Klokan Technologies mit einer Ausnahme auf der Grundlage von OpenSource-Produkten erarbeitet:

*„The search interface was written in JavaScript, using Google's Closure Library. The search map and the directory of places are integrated via the Google Maps API. [...] The server-side element of the geosearch consists of a FastCGI module, written in C/C++ that runs on a Linux system using a lighttpd web server. The web server supplies the data in JSON format, with the search parameters in line with the OpenSearch standard wherever possible. With the exception of the MapRank algorithm, open source components are used throughout“ (Oehrli et al. 2011).*

Der MapRank-Algorithmus ist eine Eigenentwicklung von Klokan Technologies und verwendet eine neue Art von Index. Er stützt sich nicht auf konventionelle Indizierverfahren mit Hilfe von R-Bäumen<sup>26</sup>, da die Performance bei diesen unbefriedigend ist.

---

<sup>24</sup> <http://www.bmpix.org> [09.07.2012].

<sup>25</sup> <http://kartenportal.mapranksearch.com/> [23.06.2012].

<sup>26</sup> „Der R-Baum ist eine räumliche Indexstruktur, die zwei- oder mehrdimensionale Rechtecke mit Hilfe überlappender Blockregionen auf dem Hintergrundspeicher organisiert“ (Brinkhoff 2008, S. 201). Dies erlaubt eine schnelle Suche nach mehrdimensionalen Objekten. In Geodatenbanken werden geografische Daten als Polygone gespeichert, die mittels Koordinaten effizient gefunden werden können. Ein R-Baum eignet sich gut für Rechtecksanfragen (ebd.).

*„MapRank [...] uses an index that already contains part of the precalculated ranking of the documents. The precalculation is carried out when the records are imported into the system. This therefore makes the system faster and more scalable, as the speed with which the list of results is generated is not significantly affected by the number of indexed records“ (ebd.).*

Als Datengrundlage dient der Geosuche der Metakatalog Swissbib<sup>27</sup>, ebenfalls ein Projekt von e-lib.ch. Swissbib vereinigt die Daten aus über 800 Schweizer Bibliotheken, Mediotheken und Archiven (Swissbib 2011) und beinhaltet zurzeit rund 130'000 Kartenkatalogisate (Oehrli et al. 2011). Da Swissbib die MARC21-Records harvestet und dedubliert, kann sich die Geosuche diese beiden Schritte sparen. Via OAI-PMH-Schnittstelle greift sie direkt auf die Daten zu. Die Suchresultate sind mit Swissbib verlinkt, d.h. ein Klick auf ein Treffer der Ergebnisliste führt direkt zu den Holdings der entsprechenden Bibliotheken.

Bei der räumlichen Suche der Geosuche handelt es sich nicht um eine Punkt- sondern um eine Flächensuche. Berücksichtigt werden weder die Eckkoordinaten, noch die Zentrumskoordinaten sondern die Kartenfläche<sup>28</sup>. Die Suchfunktion vergleicht zwei Gebiete:

*„the region covered by the cartographic material and the region covered by the user's search query. These two regions are typically linked by one of the following spatial relationships: equals, overlaps, contains, is contained by, disjointed“ (Oehrli et al. 2011).*

Abbildung 4 veranschaulicht, was das heisst: Das schwarze Rechteck entspricht der Suchanfrage, die roten Felder zeigen die Treffer, die gefunden werden, während die blaue Fläche nicht berücksichtigt wird.

<sup>27</sup> <http://www.swissbib.ch/> [23.07.2012].

<sup>28</sup> Ein Beispiel für ein Projekt, das Zentrums- und Eckkoordinaten verwendet, ist MapHappy der University of Minnesota Libraries (<http://www.lib.umn.edu/apps/maphappy/>) (Johnston et al. 2009).

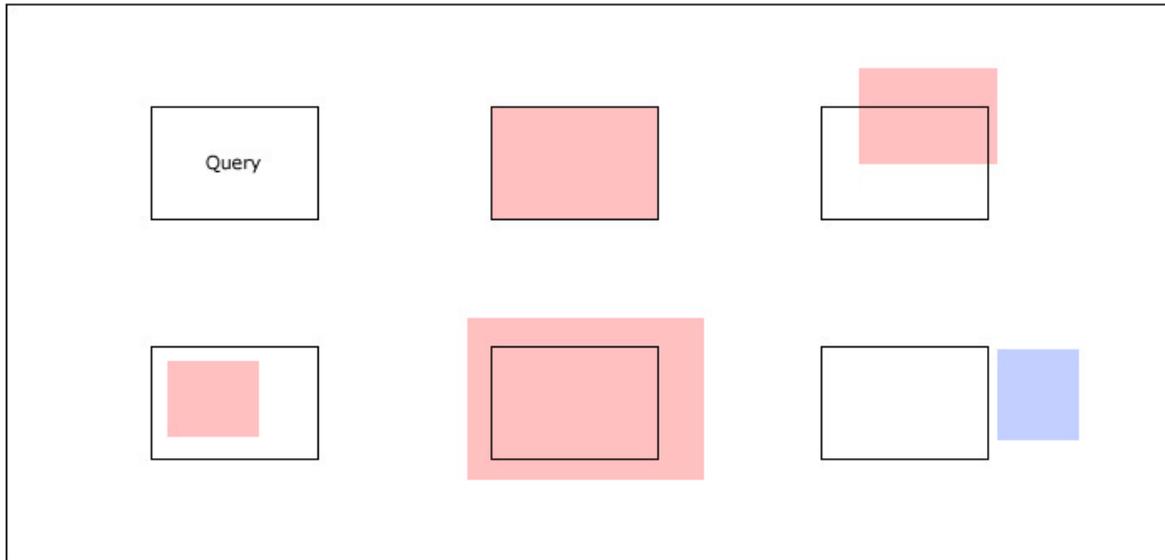


Abbildung 4: Die räumliche Suche der geografischen Kartensuche (Oehrli et al. 2011)

Die Suchoberfläche der Geosuche besteht aus drei Bereichen (Abbildung 5):

1. Mit den Zoom- und Drag-Funktionen von Google Maps kann der Benutzer nach dem gewünschten Kartengebiet suchen. Eine kleine Übersichtskarte in der rechten unteren Ecke zeigt die Suchanfrage (blau) sowie die am besten gerankte Karte (rot). Alternativ besteht die Möglichkeit, nach geografischen Namen zu suchen<sup>29</sup>.
2. Mit Zeitstrahl (Publikationsdatum) und Massstabsfilter kann die Suche weiter präzisiert werden.
3. In der rechten Spalte erscheint die Trefferliste mit den wichtigsten bibliografischen Angaben (Titel, Untertitel, Erscheinungsjahr, Verlag, Verlagsort, Massstab). Rot unterlegt ist der relevanteste Treffer, dessen entsprechendes Gebiet auf der Google-Karte rot eingefärbt ist. Wird der Cursor der Maus auf einen anderen Treffer geführt (Mouseover), erscheint das zugehörige Kartengebiet auf der Google-Karte rot eingefärbt. Das Anklicken eines Treffers führt zur Katalogaufnahme im Swissbib Katalog (Oehrli et al. 2011).

<sup>29</sup> Der Suche nach Ortsnamen ist die Suche von Google Maps hinterlegt und funktioniert somit auch für Strassen, Gebäuden, Seen, Flüsse, Flurnamen, Täler und Berge. Die Mehrsprachigkeit ist gegeben und desweiteren führt beispielsweise die Abfrage „Leningrad“ zu St. Petersburg.

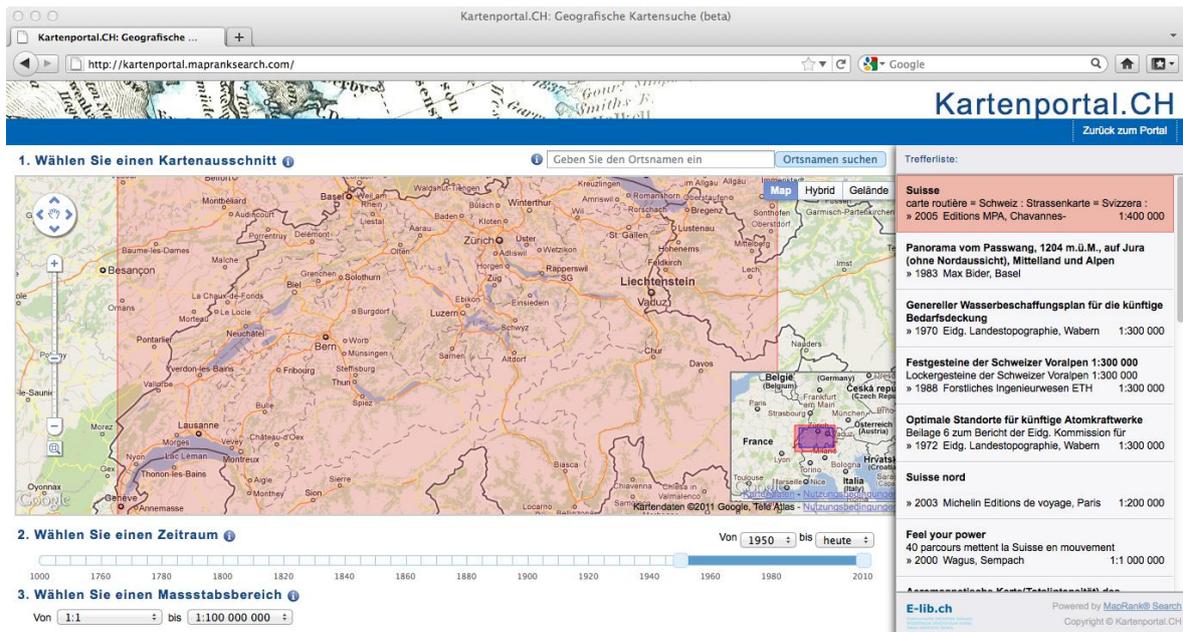


Abbildung 5: Screenshot User Interface der geografischen Kartensuche [23.06.2012]

Das Ranking der Geosuche funktioniert so, dass stets diejenige Karte zu oberst in der Trefferliste erscheint, deren Kartenfeld am genauesten mit dem Suchrechteck der Google-Karte übereinstimmt: „Der Suchalgorithmus findet alle Karten, die bezüglich Lage, Ausdehnung und Masstab am besten mit dem gewählten Kartenausschnitt übereinstimmen“ (Mouseover-Text Geosuche<sup>30</sup>). Filtert der Benutzer die Suchresultate mittels Masstabs- und/oder Zeitfilter, wird die Suche bzw. das Ranking entsprechend angepasst.

Die Geosuche bietet eine überaus benutzerfreundliche und intuitive handhabbare Suchmöglichkeit für Karten. Der Benutzer braucht weder Kenntnisse der Koordinaten, noch müssen ihm die Toponyme der gesuchten Orte oder Regionen bekannt sein. Es spielt keine Rolle, ob Städte oder Staaten im Laufe der Zeit ihre Namen geändert oder sich politische Grenzen verschoben haben. Auch braucht der Benutzer sich keine Gedanken über den Titel der Karten zu machen, um relevante Treffer zu erhalten. Die Recherche funktioniert sprachunabhängig. Grundlage für die Suche sind einzig und allein die Koordinaten, der Masstab sowie das Erscheinungsjahr.

## 2.5 Fazit

Gemeinsam ist allen vorgestellten Projekten, dass sie für die Benutzerschnittstelle die Technologie von Google Maps API verwenden und neben der Wortsuche das Geo-Browsing ermöglichen. Jede Anwendung bietet einen massgeschneiderten Zugriff auf einen ausgewählten Teilbestand, der von Hand georeferenziert worden ist, bzw. wird (über BM Archives lässt sich in dieser Hinsicht keine Aussage machen). Mit Ausnahme von BM

<sup>30</sup> Mouseover-Text des Info-Icon von „1. Wählen Sie einen Kartenausschnitt“ auf <http://kartenportal.mapranksearch.com/> (siehe Abbildung 5).

Archives, das vorsieht bei der Georeferenzierung auch Textdokumente zu berücksichtigen, widmen sich alle Geobrowsing-Plattformen jeweils lediglich einem einzigen Dokumenttyp: Während sich die Geosuche von Kartenportal.CH auf kartografische Erzeugnisse beschränkt, ist bei Viaticalpes und BIBonMAP ausschliesslich Bildmaterial recherchierbar. Zudem stehen bei allen Projekten, mit Ausnahme der Geosuche, digitalisierte Dokumente im Zentrum. Bei Viaticalpes und BIBonMAP sind die einzelnen Dokumente auf der Karte mittels Marker als Punkte, sogenannte Points of Interest (POI), lokalisiert, wohingegen Kartenportal.CH eine Flächensuche anbietet. In Entsprechung zu den Points of Interest könnten die Rechtecke der Kartenfelder als Areas of Interest bezeichnet werden.

Die vorgestellten Anwendungen bieten dem Benutzer einen klaren Mehrwert bei der Recherche nach Dokumenten mit geografischem Bezug. Im Fall von Viaticalpes sind die aufbereiteten Reiseberichte mit zusätzlichen Informationen aus anderen Quellen wie HLS und SIKART angereichert.

Die in Kapitel 4.2 entwickelten Konzepte von Geobrowsing-Anwendungen bauen auf den Technologien und Funktionalitäten der vorgestellten Plattformen auf. So sollen unter anderem die Flächensuche und der MapRank-Algorithmus eine entscheidende Rolle spielen. Im Gegensatz zu den betrachteten Beispielen liegt der Fokus dieser Arbeit auf der automatisierten Georeferenzierung, mit dem Ziel, den Aufwand so gering wie möglich zu halten. Zu diesem Zweck sollen Linked Open Data Verwendung finden, die geografische Koordinaten enthalten. Wichtig dabei ist es, sich nicht auf einzelne Dokumententypen zu beschränken, sondern Ansätze zu skizzieren, die unabhängig von der Art des Materials funktionieren. Analog zu den besprochenen Lösungen stehen (mehr oder weniger klar abgrenzbare) Teilbestände im Zentrum der Überlegungen. Doch bevor dargelegt wird, wie eine Georeferenzierung vonstattengehen kann, wird die Datenlage untersucht.

### 3 Die Datenlage

Grundlage für die Georeferenzierung sind die Daten. Im Folgenden soll untersucht werden, welche Daten verfügbar sind, wie deren Qualität ist und welche sich für eine Georeferenzierung eignen. Zuerst rücken die Katalogdaten in den Fokus, d.h. die bibliografischen Metadaten des IDS Basel Bern im MARC-Format, und mit ihnen die Frage, welche geografischen Informationen in Katalogaufnahmen vorhanden und verwertbar sind. Danach folgt eine kurze Einführung in die wichtigsten Standards und Grundfunktionen von Linked Open Data, bevor drei Datenbanken, namentlich DBpedia, Geonames und LinkedGeoData unter die Lupe genommen werden.

#### 3.1 Katalogdaten: Die drei geografischen Bezüge von Dokumenten

Ganz allgemein lässt sich festhalten, dass Medien drei geografische Bezüge aufweisen: Es sind dies der Erscheinungsort, die Orte der Urheber, d.h. Geburts-, Wirkungs- und Sterbeort von Autoren oder der Sitz von ortsbezogenen Körperschaften sowie die Orte des Inhalts, also inhaltliche geografische Bezüge.

FMT	MP
LDR	01819nem--22003974u-4500
007	aj-----
008	960122s1749---gw-----lat-d
019	ja Historischer Bestand. Exemplarspezifische Aufnahme, gesperrt für Veränderungen und das Anhängen von weiteren Exemplaren [5 30.12.2011/B400/mik
0341	ja a [b 850000 [d E0134400 [e E0195500 [f N0522500 [g N0491000
040	ja SzZulDS BS/BE B400
072	ja gg
245	ja Ducatus Silesiae tabula geographica generalis [h Kartenmaterial [b statui hodierno, ei nempe qui post pacem Dresdensem locum obtinet [c adaptata iustaque graduacione rectificata, per Tob. Maier [d <<Le>> duché de Silesie [b c.p.s.c.m.g.
246	[i Schreibweise des Titels auf der Karte [a Dycatus Silesiae tabula geographica generalis
255	ja [Ca. 1:850 000] [c E 13°44'-E 19°55'/N 52°25'-N 49°10'
260	ja Norimbergae [Nürnberg] [b impensis Homannianorum Heredum [c a. 1749
300	ja 1 Karte [b Kupferdruck [c 44 x 53 cm
500	ja Titeltartsche (lateinisch) oben rechts, französischer Titel oben, Legende rechts
500	ja Kartusche mit Erläuterungen und Massstabeleisten unten links
500	ja Nullmeridian: Ferro
500	ja Auch erschienen in: "Atlas Silesiae ..." hrsg. von den Homännischen Erben (Nürnberg, 1752)
590	ja Koloriert
651D	ja Schliesien [v Karte vor 1850
700	ja Mayer, Johann Tobias
752	[d Nürnberg
9022	ja Homann, Johann Baptist, Erben (Offizin, Nürnberg)
906	[c CM Karte = Carte
CAT	ja MK [b 40 [c 20111230 [i DSV01 [h 1604
CAT	ja MK [b 40 [c 20111230 [i DSV01 [h 1607
CAT	ja MK [b 40 [c 20111230 [i DSV01 [h 1617
CAT	ja MK [b 40 [c 20111230 [i DSV01 [h 1619
852	[b B400 [c 400M5 [h ZB Kart 1142 [p BM1538820 [4 Bern UB ZB [5 Magazin (U5)
PST	[0 Z30 [1 005714571000010 [b B400 [c 400M5 [o KARTE [e IP [f N [r DSV61-000000000 [h ZB Kart 1142 [3 Karte [4 Bern UB ZB [5 Magazin (U5) [7 In Bearbeitung
SBL	ja B400
LOC	[b B400 [c 400M5 [h ZB Kart 1142 [o KARTE [4 Bern UB ZB [5 Magazin (U5) [3 Karte
SYS	005714567

Abbildung 6: Screenshot Titelaufnahme im MARC-Format (IDS Basel Bern) [30.06.2012]

Als Beispiel wird eine topografische Karte mit dem Titel „Ducatus Silesiae tabula geographica generalis“ aus dem Jahr 1749 beigezogen, da an dieser Katalogaufnahme (siehe Abbildung 6) die meisten Felder mit geografischen Inhalt bzw. Bezug vorhanden sind und sich daran exemplarisch zeigen lässt, welche Daten sich in welcher Form für eine (automatisierte) Georeferenzierung eignen könnten.

### 3.1.1 Erscheinungsort

Der Erscheinungsort ist in unterschiedlicher Form in drei verschiedenen MARC-Feldern enthalten. Im **Feld 260 \$a** (siehe Abbildung 6) steht der Verlagsort. Dieser wird „in der Orthographie und im Kasus von der Vorlage übernommen“ (IDS 2011, 2: F-Kat, Teil I A, 2.4C1). Das bedeutet, dass in diesem Feld keine Normeinträge existieren. So kann beispielsweise Frankfurt am Main neben fremdsprachlichen Varianten unter anderem in den Formen „Frankfurt am Main“, „Frankfurt/Main“, „Frankfurt a. M.“ „Frankfurt, M.“ oder auch einfach als „Frankfurt“ erfasst sein. Wobei sich im letzten Fall nicht oder nicht ohne weiteres entscheiden lässt, ob es sich um besagtes Frankfurt in Hessen oder um Frankfurt an der Oder handelt. In einigen Fällen liesse sich aufgrund des Verlagsnamen, der im Feld 260 \$b erfasst ist, entscheiden, welches Frankfurt gemeint ist. Fehlt die Verlagsangabe, ist eine einwandfreie Identifizierung schwierig bis unmöglich. Zusätzliche Schwierigkeiten bei einer automatisierten Georeferenzierung bereiten folgende zwei Regeln der KIDS:

- „Wenn es zur Bestimmung als notwendig erachtet wird, kann eine andere Namensform der Stadt in eckigen Klammern beigefügt werden. Existiert eine Namensform in der Sprache der Katalogisierungsstelle, dann wird diese bevorzugt.

260 Christiania [Oslo]

260 \$a Lerpwl [Liverpool]“ (IDS 2011, 2: F-Kat, Teil I A, 2.4C2)<sup>31</sup> (dieser Fall liegt im bibliografischen Beispielsatz in Abbildung 6 vor).

- „Findet sich die Ortsangabe in der Vorlage nur in einer abgekürzten Variante, dann wird diese Schreibweise übernommen, die vollständige Form wird hinzugefügt oder der unvollständige Ortsname wird ergänzt.

260 \$a Mpls [i.e. Mineapolis ]

260 \$a Rio [de Janeiro]“ (IDS 2011, 2: F-Kat, Teil I A, 2.4C4).<sup>32</sup>

Zudem existieren derzeit etwas mehr 17'000 Aufnahmen, bei denen der Verlagsort nicht zu ermitteln war und daher „[S.I.]“ (lat. sine loco) erfasst wurde. Bei diesen Datensätzen ist eine geografische Zuordnung nicht möglich.

Wollte man dieses Feld für die Georeferenzierung verwenden, wäre ein grösserer Aufwand nötig, um die Toponyme eindeutig den entsprechenden Linked Open Data-Sätzen der Knowledgebases zuordnen zu können. Die Mehrsprachigkeit stellt kein generelles Problem dar, aber bei lateinischen und anderen „exotischen“ Sprachen dürfte Entsprechungen fehlen. Es ist fraglich ob sich der Aufwand lohnen würde, denn der Verlagsort dürfte in den meisten Fällen für den Benutzer nicht von Belang sein. Zwar kann er bei der Identifizierung einer

<sup>31</sup> Die eckigen Klammern sowie deren Inhalte sind Teil des Zitats.

<sup>32</sup> Die eckigen Klammern sowie deren Inhalte sind Teil des Zitats.

bestimmten Ausgabe eines Werkes wichtig sein, aber im Allgemeinen dürften Benutzer kaum nach Medien suchen, die an einem bestimmten Ort erschienen sind.

Im Gegensatz zum Verlagsort wird das Erscheinungsland in kontrollierter Form erfasst. **Feld 008/15-17** (siehe Abbildung 6) enthält einen Ländercode, der aus zwei bzw. drei Buchstaben besteht: z.B. sz für Schweiz, gw für Deutschland, fr für Frankreich, xxk für Grossbritannien (United Kingdom), xxu für die Vereinigten Staaten von Amerika und xx wenn kein Ort angegeben oder das Land unbekannt ist (Library of Congress 2008). Abgesehen von den 17'000 bereits erwähnten bibliographischen Datensätzen, bei denen der Ländercode xx steht, dürfte die Datenqualität in Feld 008/15-18 sehr gut sein.

```

008 010906---ad---ab--
040 la SzZuIDS BS/BE B400
152 ld Basel
452 ld Basilea
452 ld Basle
452 ld Bâle
452 ld Augusta Rauracorum
452 ld Augusta Munatiana
452 ld Basilea Rauracorum
452 ld Colonia Munatiana
452 ld Athena Helveticorum
452 ld Athena Rauracorum
667 la Varianten z.T. nach Benzing, 1982, S. 29 und VD16. Augusta Rauracorum wird im Druck Falk
3159a:22 von Konrad Waldkirch offenbar als Synonym für Basel verwendet
909A lc BBK
SYS 000136154

```

Abbildung 7: Autoritätsdatensatz eines Druckorts (IDS Basel Bern)

Bei Drucken, die vor 1800 erschienen sind, wird der Druckort im **Feld 752 \$d** erfasst. Für alle Eintragungen in diesem Feld werden Autoritätsdatensätze (AUT) erstellt. Das Beispiel in Abbildung 7 zeigt einen Ausschnitt der AUT des Druckorts Basel. Die Ansetzungsform erhält in der AUT das Feld 152 \$d, verworfene Formen erhalten je ein 452 \$d. Aufgrund des kontrollierten Vokabulars bietet sich dieses Feld für eine geografische Auswertung an.

### 3.1.2 Orte der Urheber

Autoren, Herausgeber, Fotografen, Kupferstecher etc. erhalten neben einer Erwähnung im Feld 245 \$c eine Haupteintragung im **Feld 100** oder eine Nebeneintragung in **Feld 700** (siehe Abbildung 6). Im Unterfeld \$a wird der Name erfasst. Durch die Unterfelder \$c (Berufsbezeichnung, Adelstitel etc.) und \$d (Lebensdaten) können die Personen individualisiert werden. Jedoch sind Autoren nur in den allerwenigsten Fällen individualisiert. Das führt dazu, dass beispielsweise unter dem Eintrag „Müller, Hans“ 268 Treffer gefunden werden (Stand 23.07.2012). Die Individualisierung von Autoren geschieht in AUTs und zwar aufgrund der Lebensdaten und optional zusätzlich über den Beruf. Allfällige

Namensvarianten sind ebenfalls Bestandteile der Normeinträge, nicht jedoch die Geburts- und Sterbeorte (siehe Abbildung 8).

Normeintrag - Autoren  
Müller, Hans, 1893-1971

[Zurück](#)

Klicken Sie auf ein unterstrichenes Feld in der Spalte *Liste*, um eine Indexliste zu sehen.  
Klicken Sie auf ein unterstrichenes Feld in der Spalte *Treffer*, um eine Trefferliste zu sehen.

Liste	Treffer
<u>Suche unter:</u>	<u>Müller, Hans, 1893-1971</u>
<u>Anmerkung</u>	Ingenieur, Schriftsteller
<u>Benutzt für:</u>	Müller-Schafir, Hans, 1893-1971

Abbildung 8: Screenshot Normeintrag Autor (IDS Basel Bern) [23.07.2012]

Aufgrund der fehlenden Individualisierung sind die Urheber nicht für eine Georeferenzierung geeignet, obwohl es durchaus interessant sein könnte, Autoren nach ihren Geburts- oder Wirkungsorte suchbar zu machen. Durch die Verortung von Autoren könnten Zeitgenossen, die in der gleichen Stadt gewirkt haben sowie Netzwerke dargestellt werden.

Ändern könnte sich die Situation, mit der Umstellung auf das Regelwerk RDA (Resource Description and Access). Dieses sieht vor, für formale und inhaltliche Erschließung nur noch eine Normdatei zu verwenden, was bedeutet, dass alle Personen individualisiert sein müssten. Damit wär zumindest eine Grundlage für die Georeferenzierung geschaffen. Allerdings ist fraglich, ob eine Bereinigung der Altdaten vorgenommen wird, wäre diese doch sehr aufwändig und kostspielig.

Einen möglichen Weg, wie die Disambiguierung von Personen bei Erfassung und Recherche möglich sein kann, haben Neubert und Borst (2012) in ihrem Vortrag am Bibliothekartag in Hamburg anhand von Autorinnen und Autoren der Wirtschaftswissenschaften aufgezeigt. Mit Hilfe der Gemeinsamen Normdatei GND, die durch die Deutsche Nationalbibliothek als Linked Open Data zur Verfügung gestellt wird<sup>33</sup>, sowie der Virtual International Authority File (VIAF), die die Normbestände von 22 Nationalbibliotheken mit anderen Datenanbietern aus 19 Ländern zusammenfasst (ebd., S. 25) und ebenfalls als Linked Data verfügbar sind<sup>34</sup>, soll die Identifikation von Personen erleichtert werden.

Bei Körperschaften sieht die Situation anders aus, denn für alle Einträge im **Feld 710** (Die Beispielaufnahme in Abbildung 6 enthält keine Körperschaft) werden AUTs erstellt. Abbildung 9 zeigt einen Ausschnitt der AUT für Kümmerly + Frey. Körperschaftsname oder Name der Gebietskörperschaft werden in Feld 110 \$a angesetzt, während die verworfenen Formen je ein Feld 410 \$a erhalten. Feld 510 enthält eine „Siehe-auch-Verweisung auf einen Körperschaftsnamen“ (IDS 2011, 1: Format, Feld 510), in diesem Fall Vorgänger- und Nachfolgerkörperschaften.

<sup>33</sup> [http://www.dnb.de/DE/Service/DigitaleDienste/LinkedData/linkedata\\_node.html](http://www.dnb.de/DE/Service/DigitaleDienste/LinkedData/linkedata_node.html) [24.07.2012].

<sup>34</sup> <http://viaf.org/viaf/data/> [24.07.2012].

```

008  901129---az---ab-n-----
040  la SzZuIDS BS/BE A100 lb ger
090  la 791565 lb DSV
099  la cr lb 40 lc 00000000 ll DST11 lh 0000
110  la Kümmerly + Frey (Bern)
410  la Kümmerly und Frey AG (Bern)
410  la Kümmerly + Frey (Zollikofen)
410  la Kümmerly + Frey Geographischer Verlag (Bern)
410  la Kartographische Anstalt Kümmerly und Frey (Bern)
410  la Geographisch-Artistische Anstalt Kümmerly und Frey (Bern)
410  la Geographische Anstalt Kümmerly und Frey (Bern)
410  la Verlag Kümmerly und Frey (Bern)
410  la Geographischer Karten-Verlag Kümmerly und Frey (Bern)
510  lw i la Gebrüder Kümmerly li Vor 1898 siehe
510  lw i la Hallwag Kümmerly + Frey (Urtenen) li Ab 2002 siehe
670  la LoCNA; Helveticat; Schweizerisches Handelsregister; Vorlage
680  la Kümmerly und Frey AG., Verlag und Graphische Betriebe in Bern/Zollikofen (1898-2001)
680  la Ab 2002 gemeinsam mit Hallwag in Urtenen
SYS  000020239

```

Abbildung 9: Autoritätsdatensatz einer Körperschaft (IDS Basel Bern)

Das geografische Unterschlagwort in Unterfeld 110 \$z findet keine Verwendung. Der Wirkungsort bzw. der Sitz der Körperschaft wird wie im Fall von Kümmerly + Frey in Klammer hinter dem Körperschaftsnamen angefügt, wenn er nicht Teil des Namens ist (z.B. Basler Botanische Gesellschaft). Hier zeigt sich ein Mangel an Granularität<sup>35</sup>. Die Information ist zwar vorhanden, aber nicht in einem eigenen Unterfeld untergebracht und so lässt sich das geografische Datum nicht ohne weiteres auswerten.

Bei Drucken, die vor 1800 erschienen sind, wird für Druckereien, die nicht Urheber und somit im Feld 260 \$b aufgeführt sind, das **Feld 9022 \$a** erfasst (siehe Abbildung 6). Für jede Druckerei wird ein AUT erfasst. Das Beispiel in Abbildung 10 zeigt den Ausschnitt eines Normeintrags für einen Basler Offizin. Als Ordnungshilfe enthält der AUT den Sitz der Druckerei in runden Klammern, ausser der Ort ist Bestandteil des Namens. Wenn aus dem Namen nicht ersichtlich ist, dass es sich um eine Körperschaft handelt, wird der Begriff „Offizin“ ergänzt (IDS 2011, 2: F-Kat. Teil I A, 2A12.2.2.) und „Druckerdynastien werden immer je Ort zu einer Körperschaft zusammengefasst“ (ebd.). Namensvarianten, Anmerkungen sowie die Wirkungszeit finden ebenfalls Eingang in die AUTs.

<sup>35</sup> Granularität bedeutet die Feinheit der Gliederung von Datenfeldern (Viegner 2012, S. 2). So ist beispielsweise auch die Ansetzung von Personennamen im MARC-Format nach dem Muster 100 \$a Nachname, Vorname problematisch, da keine Möglichkeit der Verfeinerung besteht. Da Vor- und Nachname nicht in separaten Feldern untergebracht sind, ist eine Verarbeitung „stets an einen zusätzlich programmierten Mehraufwand auf Applikationsebene gebunden“ (Fürste 2011, S. 27).

```

008 040518---ad---ab--
040 la SzZuIDS BS/BE A100
110 la Schröter, Johannes, Erben (Offizin, Basel)
410 la Schröter, Johann, Erben (Offizin, Basel)
410 la Schretter, Johann, Erben (Offizin, Basel)
410 la Schroeterus, Johannes, Erben (Offizin, Basel)
670 la Reske, 2007, S. 90 ; Benzing, 1982, S. 44
680 li Wirkungszeit: 1634-1635
680 la Johann Schröter aus Schleusingen (Thüringen) starb am 8.9.1634. Unter "Johann Schröter sel.
Erben", "Typis Haered. Joh. Schroeteri", vereinzelt auch "Johann Schröter sel. Wittib" firmierte die
Offizin bis 1635, als die Witwe, Johann Schröters zweite Frau Margaretha Zäsinger den vorher im
Betrieb arbeitenden Teilhaber Georg Decker heiratete (BBK, ergänzt aus Benzing und Reske)
909A lc BBK
SYS 000168234

```

Abbildung 10: Autoritätsdatensatz einer Druckerei (IDS Basel Bern)

Die Eintragungen in Feld 9022 \$a sind im Allgemeinen von sehr guter Qualität und aufgrund des aufgeführten Sitz der Druckerei ist ein Toponym vorhanden, das sich georeferenzieren liesse. Wie schon die Normdatensätzen der Körperschaften erweist sich auch dieses Feld als zu wenig granular. Interessant ist Feld 9022 \$a deshalb vor allem in Zusammenhang mit Feld 752 \$d, in dem der Druckort separat aufgeführt ist (vgl. 3.1.1 Erscheinungsort).

### 3.1.3 Orte des Inhalts

Inhaltliche geografische Bezüge von Katalogdaten treten in vier verschiedenen Feldern zu Tage. So können Toponyme Teil des Titels sein, der im **Feld 245** erfasst wird. (siehe Abbildung 6). Unterfeld \$a enthält den Haupttitel, der Untertitel wird ins \$b übernommen. Doch hier ist eine mögliche Auswertung noch schwieriger umzusetzen, als beim Erscheinungsort (260 \$a), denn in diesen Feldern treten geografische Bezeichnungen in Form von (flektierten) Nomen sowie als Adjektive in allen möglichen Sprachen auf. Neben den eigentlichen Toponymen sind hier auch Bauwerke (z.B. Basler Münster, Universität Basel, Fasnachts-Brunnen) zu finden, die einem bestimmten Ort zuordenbar sind. Grundsätzlich dürfte es möglich sein, die Titeldaten mit computerlinguistischen Verfahren nutzbar zu machen, jedoch wären Aufwand und Fehlerquote relativ hoch.

Durch die Koordinaten ist eine weitere inhaltliche geografische Referenz vorhanden. Im **Feld 255** (siehe Abbildung 6) werden Massstab (\$a) und geografische Koordinaten (\$c) für die Anzeige im OPAC jeweils im Sexagesimalsystem erfasst. **Feld 034** beruht auf diesem Feld und erhält die Daten in codierter Form in folgenden Unterfeldern (IDS 2011, 1: Format, Feld 034):

- \$a Art des Massstabs
- \$b Nenner des Quotienten, linearer Längenmassstab
- \$c Nenner des Quotienten, linearer Höhenmassstab
- \$d Koordinate – westlichste geografische Länge

- \$e Koordinate – östlichste geografische Länge
- \$f Koordinate – nördlichste geografische Breite
- \$g Koordinate – südlichste geografische Breite

Die Koordinaten stellen das abgebildete Gebiet als Rechteck dar und erlauben eine sehr genaue Georeferenzierung. Doch ist die Erfassung von Koordinaten gemäss KIDS ausschliesslich für Karten vorgesehen. Wie die Analyse der geografischen Suche des Kartenportals gezeigt hat (siehe Kapitel 2.4), werden diese Daten bereits in gelungener Weise genutzt.

Am deutlichsten ist der inhaltliche geografische Bezug in den Schlagwörtern ersichtlich. Seit Frühjahr 2012 wird im IDS Basel Bern die GND (Gemeinsame Normdatei) zur inhaltlichen Erschliessung verwendet.

*„In der Gemeinsamen Normdatei sind die bisher bestehenden Normdateien Gemeinsame Körperschaftsdatei (GKD), Personennamendatei (PND) und Schlagwortnormdatei (SWD) sowie die Einheitssachtitel-Datei des Deutschen Musikarchivs (DMA-EST-Datei) in einer gemeinsamen Normdatei (GND) zusammengeführt“ (Pfeifer 2012).*

Zuvor bestand in der IDS Basel Bern mit dem DSV12 ein Schlagwortsystem, das an der UB in Eigenentwicklung aufgebaut wurde und Schlagwortketten verwendete (Abbildung 6). Geplant ist, den DSV12 in die GND zu überführen. Doch nicht nur die Ketten bereiten Probleme, sondern auch die Relationen zwischen einzelnen Schlagwörtern. So figuriert beispielsweise „Metropole“ im DSV12 als verworfene Form von „Grossstadt“, während „Metropole“ in der GND als Unterbegriff von „Grossstadt“ auftritt. Bei Geografika ist dieser Sachverhalt weniger problematisch, da die Begriffe eindeutiger sind (Mündliche Auskunft von Alice Spinnler, 05.06.2012).

Neben der GND und dem DSV12 existieren zahlreiche lokale Schlagwortsysteme, die von Institutsbibliotheken und dem IDS Basel Bern angeschlossenen Institutionen gepflegt werden. Die Beispielaufnahme (siehe Abbildung 6) enthält im **Feld 651D** ein Lokalschlagwort der UB Bern. Neben den verschiedenen Systematiken bestehen im Katalog viele Altdaten, die überhaupt nicht inhaltlich erschlossen sind. Um die Zahl der beschlagworteten Aufnahmen zu erhöhen und eine Vereinheitlichung der Daten zu erzielen, plant die Schlagwortredaktion eine automatisierte Anreicherung, wie sie Pfeffer (2012) entwickelt hat (siehe Exkurs II). Die Methode, die sich auf die SWD bezieht, lässt sich auf die GND übertragen, da die SWD in der GND aufgegangen ist.

Die Ausführungen zu den Schlagwörtern sollen sich auf die GND beschränken und weitere Schlagwortsysteme aussen vor lassen. Die GND enthält unter anderem Normdaten zu Personen, Körperschaften, Veranstaltungen und Kongresse, Werktitel, Geografika sowie Sach-, Zeit- und Forms Schlagworte. Von besonderem Interesse sind in diesem Zusammenhang die Geografika, die nun etwas genauer betrachtet werden sollen. Abbildung

11 zeigt den Normdatensatz des Geografikum „Basel“ in einem Ausschnitt. Im Feld 151<sup>36</sup> ist die Hauptansetzungsform und in verschiedenen Feldern 451 sind die verworfenen Formen sowie die GND-ID aufgeführt. Die URI des Datensatz in der GND ist im Feld 0247 enthalten. Das Feld 043 enthält einen Ländercode.

```

001      (DE-588)4004617-5
005      20120420091902.0
008      880701nllazznnaabn-----l-ana----lc
0247     la http://d-nb.info/gnd/4004617-5 l2 uri
035      la (DE-101)040046176
035      la (DE-588)4004617-5
035      lz (DE-588c)4004617-5 l3 zg
040      la DE-101 lr DE-101 lb ger ld 1140 le rakwb lf rswk
043      lc XA-CH-BS
079      la g lb g lc l lq s lq f lu w lu z lv gik lv gif
08304    lz 2 la 49432 lt 2007-01-01 l2 22/ger
151      la Basel l2 gnd l1 (DE-588)4004617-5
451      la Kanton Basel lg Stadt l1 (DE-588)4004617-5
451      la Kanton Basel-Stadt l1 (DE-588)4004617-5
451      la Kanton Basel lg 1833- l1 (DE-588)4004617-5
451      la Basel l2 swd l1 (DE-588c)4004617-5
670      la B 2006, GKD, M
679      la Hauptort des Kantons Basel-Stadt, drittgrößte Stadt der Schweiz
913      lS swd li g la Basel l0 (DE-588c)4004617-5
COR      la Basel l2 swd-COR
SYS      000157403

```

Abbildung 11: GND-Normdatensatz eines Geografikum (IDS Basel Bern)

Die Ländercodes der SWD beruhen auf den Kürzeln für Staaten gemäss der DIN EN ISO 3166-2, die um zusätzliche Codes für Kontinente und Meere erweitert wurden. Auf einer dritten Ebene wurden auf Grundlage der DIN ISO 3166-2 für Deutschland, Österreich und die Schweiz weitere Verwaltungseinheiten (Bundesländer und Kantone) definiert (Wiesenmüller et al. 2011, S. 246). Tabelle 1 zeigt eine Auswahl verschiedener Ländercodes.

Neben Geografika erhalten diejenigen Schlagwörter Ländercodes, die einen geografischen Aspekt besitzen, wie

*„z.B. Ethnografika, Sprachen, Bauwerke, Körperschaften, Historische Ereignisse sowie Sachschlagwörter oder Werke mit einem eindeutigen geografischen Bezug. [...] Auch bei Personen werden Ländercodes vergeben, allerdings nur auf der zweiten Hierarchiestufe“ (ebd.).*

<sup>36</sup> Im Titeldatensatz erscheint das Geografikum im MARC-Feld 651.

<b>Ländercode</b>	<b>Gebiet</b>
XA	Europa
XA-CH	Schweiz
XA-CH-BS	Basel-Stadt
XA-CH-BL	Basel-Landschaft
XA-DE	Deutschland
XA-DE-BW	Baden-Württemberg
XA-DE-NW	Nordrhein-Westfalen
XA-AT	Österreich
XA-AT-1	Burgenland
XA-AT-2	Kärnten
XA-GB	Grossbritannien
XC	Afrika
XC-EG	Ägypten
XC-NG	Nigeria
XD	Amerika
XD-CA	Kanada
XK	Atlantischer Ozean
XK-GL	Grönland

Tabelle 1: Ländercodes in der SWD (Scheven, 2012)

Auf der Grundlage des Ländercodes haben Wiesenmüller et al. (2011) eine geografische Facette entwickelt, die in den Katalogen der UB Heidelberg und UB Mannheim implementiert worden ist. Die Einbeziehung der Ländercodes in die Recherche hat die Suchergebnisse zum Teil markant verbessert. Da dieser sehr interessante Ansatz wichtige Erkenntnisse für die Entwicklung von Geobrowsing-Anwendungen liefern kann, soll er weiter unten eingehender behandelt werden (siehe Exkurs I: Der Ländercode der GND).

Einen weiteren Ortsbezug gibt es neben den bereits erwähnten MARC-Feldern bei Ausstellungen und Kongressen. Diese erhalten Einträge in Feld 913, bei denen jeweils Jahr und Ort erfasst werden:

- 913 \$a Ausstellung = Exposition \$b 2009 \$c Basel
- 913 \$s Kongress = Congrès \$b 2011 \$c Yokohama

Der Ort wird in kontrollierter Form im Unterfeld \$c erfasst, d.h. die Orte werden bei Ausstellungen und Kongressen aus einer gemeinsamen Liste kopiert. Die Qualität der Daten ist somit über weite Strecken sehr gut. Allerdings gibt es keine Autoritätsdatensätze, die verschiedene Schreibweisen und Sprachvarianten zusammenführen. Dies hat zur Folge, dass zum Teil ein- und dieselben Orte in verschiedenen Sprachen angesetzt werden. Ein treffendes Beispiel ist das belgische Antwerpen, das unter der französischen Variante „Anvers“ angesetzt ist. Der Blick in den Index zeigt, dass 41 bzw. 42 Titel fälschlicherweise unter „Antwerpen“ indexiert sind (siehe Abbildung 12). Ähnlich präsentiert sich die Situation bei Den Haag, das unter dem Namen „’s-Gravenhage“ angesetzt ist.

**Blättern in einer Liste**

**Suchbegriff(e):**

**Wählen Sie eine Liste:**

**Liste: Kongress/Ausstellung**

Treffer	Eintrag
2	<a href="#">Antofagasta</a>
1	<a href="#">[Antwerpen]</a>
41	<a href="#">Antwerpen</a>
1	<a href="#">Anugraha</a>
50+	<a href="#">Anvers</a>
7	<a href="#">Anzio</a>
1	<a href="#">Anzola d'Ossola</a>
1	<a href="#">Aomori City</a>
50+	<a href="#">Aosta</a>
1	<a href="#">Aparecida, Brasil</a>

Abbildung 12: Screenshot Index Kongress-/Ausstellungsort (IDS Basel Bern) [10.07.2012]

Fehlende AUTs sind nicht nur für den Katalogisierenden etwas unglücklich, sondern vor allem für den Benutzer, denn er benötigt die Kenntnis, unter welchem Name ein Ort angesetzt ist – vorausgesetzt er verwendet den Index. Erhielte er mit der Indexabfrage „Antwerpen“ gar keinen Treffer, würde er vielleicht nach einem alternativen Toponym suchen, erhält er jedoch Treffer, merkt er nicht, dass der Recall bzw. die Treffermenge viel zu gering ist. Die Lage entschärft sich etwas bei einer Suche über alle Felder, da der Ausstellungs- bzw. der Kongressort oft im Notizfeld 500 \$a aufgeführt oder mittlerweile durch GND-Schlagwörter abgedeckt ist. In der GND werden Ausstellungen und Kongresse mittels Forms Schlagwort erfasst:

- 6557 \$a Ausstellung \$z Riehen (Basel, 2012)
- 6557 \$a Kongress \$z Yokohama (2011)

Bei der Georeferenzierung von Feld 913 \$c würden neben den korrekten auch die falschen Indexeinträge berücksichtigt werden und sich dadurch die Fehlerquote minimieren. Da Ausstellungen und Kongresse in der GND als Formschlagwörter auftreten, ist es fraglich, ob eine Georeferenzierung dieses Feldes nötig ist.

Allerdings dürfte eine geografische Suche nach Ausstellungen und Kongressen wenig sinnvoll sein, da die Verwendung des Ortes bei der Suche nach Kongressen und Ausstellungen in den meisten Fällen helfen soll, eine bestimmte Publikation zu finden. Die Suche nach allen Publikationen zu Kongressen und Ausstellungen in einem bestimmten Gebiet zu einer bestimmten Zeit dürfte eher die Ausnahme darstellen.

### **Exkurs I: Der Ländercode der GND**

Wiesenmüller et al. (2011, S. 245) bemängeln, dass die Datensätze der Schlagwortnormdatei „eine Fülle von Informationen [enthalten], die bisher nur zu einem erschreckend geringen Teil für die Recherche in Bibliothekskatalogen nutzbar gemacht werden“. Einerseits wird nur ein Teil der Thesaurusrelationen berücksichtigt. Ausgewertet werden normalerweise die Ansetzungsform („Geografikum“) und die Synonyme („Andere Namen“), nicht jedoch der Oberbegriff<sup>37</sup> (ebd.) (siehe Abbildung 13). Wird der Index jedoch um diese Relationen erweitert, werden deutlich mehr Treffer gefunden. Die Suche nach „Mittelgebirge“ findet unter Einbezug der Unterbegriffe auch die „Schwäbische Alb“.

*„Die Methode bringt allerdings einen gewissen Ballast in der Treffermenge mit sich und sollte deshalb mit einem Ranking kombiniert werden. [...] Ein Ranking würde Treffer zu Unterbegriffen [...] des Suchbegriffs erst weiter hinten in der Trefferliste anzeigen“ (ebd.).*

---

<sup>37</sup> Ungenützt bleiben oft auch der verwandte Begriff, der so genannte mehrgliedrige Oberbegriff und die DDC-Notation. Auf diese beiden Informationen soll hier jedoch nicht eingegangen werden.

<a href="#">GND</a>	
<b>Link zu diesem Datensatz</b>	<a href="http://d-nb.info/gnd/4053686-5">http://d-nb.info/gnd/4053686-5</a>
<b>Typ</b>	Naturräumliche Einheit (gin)
<b>Geografikum</b>	Schwäbische Alb
<b>Andere Namen</b>	Schwabenalb Schwäbischer Jura Rauhe Alb Jura / Schwaben
<b>Quelle</b>	B 1996, Geo-Du.
<b>Land</b>	Baden-Württemberg (XA-DE-BW); Bayern (XA-DE-BY)
<b>Oberbegriffe</b>	Beispiel für: <a href="#">Mittelgebirge</a>
<b>DDC-Notation</b>	T2--43473
<b>Systematik</b>	19.1b Physische Geografie
<b>Thema in</b>	<a href="#">310 Publikationen</a>  <ol style="list-style-type: none"> <li>1. <i>Hits für Kids Schwäbische Alb</i> Bayer, Antje. - München : Berg, 2012</li> <li>2. <i>Erlebnisreiche Radtouren</i> Eckstein, Eva. - Reutlingen : Oertel + Spörer, 2012</li> <li>3. ...</li> </ol>

Abbildung 13: Screenshot GND-Geografikum Schwäbische Alb (DNB) [30.07.2012]

Andererseits wird der Ländercode (siehe Tabelle 1, S. 35) nicht berücksichtigt. Wird dieser jedoch in die Suche einbezogen, lassen sich die Suchergebnisse markant verbessern. Für die Testinstallation wurde die Indexierung in der Weise abgeändert,

*„dass bei einer Recherche nach einem Ländercode nicht mehr die Schlagwörter mit dem betreffenden Ländercode ausgegeben wurden, sondern stattdessen die Titel, die mit diesen Schlagwörtern verknüpft sind“ (ebd., S. 246).*

Die Erwartungen in die Tests wurden „nicht nur erfüllt, sondern sogar noch übertroffen“ (ebd.). Aus diesem Grund sollen die Testergebnisse nun etwas genauer betrachtet werden (siehe Tabelle 2). Die Erhöhung des Recalls<sup>38</sup> ist in erster Linie vier Faktoren geschuldet (ebd., S. 248):

1. Über den Ländercode werden auch Teile des betrachteten Geografikum berücksichtigt, da sie denselben Ländercode tragen. Bei der Suche „Tourismus in Baden-Württemberg“, werden nicht nur die Titel gefunden, die sich mit dem ganzen Bundesland beschäftigen, sondern auch solche, die sich z.B. mit dem Tourismus in Heidelberg beschäftigen.
2. Bei einer trunkierten Suche werden auch diejenigen Geografika berücksichtigt, die auf einer untergeordneten Hierarchiestufe liegen. Bei der Suche „Klimaänderung in Afrika“ werden auch Titel zu den einzelnen Ländern gefunden.

<sup>38</sup> „Recall (Vollständigkeit) errechnet sich als Quotient aus der Anzahl der gefundenen relevanten Dokumentationseinheiten und der Gesamtzahl der relevanten Dokumente“ (Stock 2007, S. 63).

Suchinteresse	mit Geografikum	mit Ländercode	Erhöhung Recall
Wirtschaftskrise in den USA	145	173	19,31 %
Muslime in Österreich	23	33	43,48 %
Umweltschutz in der Schweiz	274	399	45,62 %
Klimaänderung in Afrika	22	52	136,36 %
Tibetischer Buddhismus in Indien	35	90	157,14 %
Tourismus in Baden-Württemberg	51	179	250,98 %
Sonnenenergie in Deutschland	21	74	252,38 %
Rechtsphilosophie in Deutschland	90	447	396,67 %
Höhlen in Baden-Württemberg	8	92	1'050,00 %

Tabelle 2: Ergebnisse von Testrecherchen im SWB (Wiesenmüller et al. 2011, S. 248)

3. Bei der Ländercode-Recherche werden auch Individualbegriffe (Instanzen) berücksichtigt, die im jeweiligen geografischen Raum liegen. Bei der Suche „Höhlen in Baden-Württemberg“ wird nicht nur die allgemeine Literatur gefunden, sondern auch Literatur zu einzelnen Höhlen, da bei den Individualbegriffen das Stichwort „Höhle“ im Oberbegriff enthalten ist.

4. Über den Ländercode werden auch Personen berücksichtigt. Bei der Suche „Rechtsphilosophie in Deutschland“ werden auch Treffer zu deutschen Rechtsphilosophen gefunden, da diese einen Ländercode erhalten. Allerdings funktioniert dies nur, wenn das entsprechende Sachschlagwort enthalten ist.

Durch die Nutzung des Ländercodes bei der Recherche konnte der Recall deutlich erhöht werden, ohne die Precision<sup>39</sup> zu verschlechtern. Zudem sind ungenutzte Informationen der Normdaten endlich beim Nutzer angekommen (ebd., S. 250).

Die guten Testergebnisse führten dazu, dass im HEIDI-Katalog der UB Heidelberg und im Primo-Katalog der UB Mannheim die geografische Facette implementiert wurde (ebd., S. 246). Im HEIDI-Katalog entschied man sich für eine mehrstufige Drilldown-Facette, die die hierarchische Relation der geografischen Codes abbildet (siehe Anhang, Abbildung 25). Es wurde zwischen den Titeln unterschieden, „die direkt mit einem Code der ersten oder zweiten Ebene verknüpft sind, und solchen, die diese Verknüpfung aus der Hierarchie erhalten“ (ebd., S. 251). Dies führt dazu, dass verschiedene Toponyme zweimal aufgeführt

<sup>39</sup> „Precision (Genauigkeit) ergibt sich als Quotient aus der Anzahl der gefundenen relevanten Dokumentationseinheiten und der Gesamtzahl der gefundenen Datensätzen“ (Stock 2007, S. 63).

sind. So sind beispielsweise unter „Deutschland (gesamt)“ die Titel indexiert, die sich auf das gesamte Deutschland beziehen und nicht auf eine Teilregion, während unter „Deutschland“ auch die einzelnen Bundesländer indexiert sind.

### 3.1.4 Fazit

Ein Feld wurde in den Ausführungen über die Katalogdaten bisher nicht berücksichtigt: das **Notizfeld 500 \$a**. In diesem können alle drei Arten geografischer Bezüge auftreten. So finden beispielsweise ursprüngliche oder alternative Erscheinungsorte (siehe Abbildung 6) oder Ausstellungen inkl. den Ausstellungsorten Erwähnung. Auch Nennungen von Geografika im Zusammenhang mit Autoren oder Körperschaften sind keinesfalls auszuschliessen. Die fehlende Eindeutigkeit im geografischen Bezug führt dazu, dass sich das Feld 500 \$a nicht für eine Georeferenzierung eignet.

Das Vorhandensein von Normeinträgen ist ein wichtiger Faktor bei der Beurteilung, ob sich ein bestimmtes Feld für die Georeferenzierung eignet oder nicht. Zwar liessen sich auch nicht-kontrollierte Datenfelder georeferenzieren, aber dafür wären computerlinguistische Verfahren nötig, die in der vorliegenden Arbeit nicht zur Sprache kommen. Wie die Untersuchung gezeigt hat, besteht ein Potenzial bei folgenden Feldern:

- 008: Erscheinungsland
- 0341 \$d-\$g Koordinaten
- 651: Geografikum der GND
- 752 \$d: Druckort
- 9022 \$a: Druckerei
- 913 \$c: Ausstellungs-/Kongressort

Exemplarisch sollen in drei Anwendungsbeispielen (siehe Kapitel 4.2) das Erscheinungsland, die Geografika der GND sowie der Druckort und damit verbunden die Druckereien berücksichtigt werden.

Die Koordinaten brauchen nicht weiter verfolgt zu werden, da das Kartenportal.CH bereits eine sehr gute Recherche-Plattform zur Verfügung stellt. Die Georeferenzierung von Ausstellungs- und Kongressorten wird ebenfalls ausgeschlossen, denn diese können durch die Referenzierung der GND abgedeckt werden. Zudem ist fraglich, ob die geografische Suche nach Ausstellungen und Kongressen gegenüber der konventionellen Katalogsuche einen wirklichen Mehrwert erzielen könnte.

Somit spielen zweimal Erscheinungsorte eine Rolle und einmal der inhaltliche geografische Bezug. Für die Orte der Autoren lässt sich aufgrund der Datenlage kein Beispiel ausarbeiten. Immerhin ist es möglich, die Druckereien indirekt zu referenzieren. Bevor die Daten dreier LOD-Anbieter hinsichtlich ihrer Eignung für die Georeferenzierung von Katalogdaten

untersucht werden, folgt nun eine Einführung in die wichtigsten Standards und Prinzipien des Semantic Web und der Linked Open Data.

### 3.2 Linked Open Data

Im World Wide Web liegen die Informationen in erster Linie in semi-strukturierten Dokumenten vor. Auszeichnungssprachen wie HTML dienen vor allem dazu, die Präsentation der Daten zu optimieren. Eine maschinell automatisierte Verarbeitung, d.h. Interpretation und Auswertung der Daten ist dem Computer nicht möglich, da er Relevanz und Bedeutung der Daten nicht erfassen kann. Dies ist dem Menschen vorbehalten (Fürste 2011, S. 43). Die Idee des Semantic Web besteht darin, Daten effektiv zu strukturieren und miteinander zu verlinken, damit sie maschineninterpretierbar sind und eine explizite Bedeutung besitzen (ebd., S. 45).

Pohl (2011)<sup>40</sup> weist darauf hin, dass Linked Open Data sowohl eine urheberrechtliche sowie eine technische Komponente umfassen: „Open-Data-Standards sorgen für die *rechtliche* Basis der Nutzung und Kombination verteilter Daten im Netz“, während „Linked-Data-Standards [...] die *technische* Kompatibilität“ unter den Daten gewährleisten. Fürste (2011, S. 45) fügt an, dass „offen“ nicht nur bedeutet, dass die Daten standardisiert und frei verfügbar vorliegen, sondern auch granular und über eine **URI** eindeutig identifiziert sind:

*„Ein Uniform Resource Identifier ermöglicht die eindeutige Benennung eines Informationsobjekts für das gesamte Internet und bildet den Überbegriff für URLs (ortsabhängige Bezeichner und URNs (ortsunabhängige Bezeichner)“ (ebd.).*

URIs können in verschiedenen Formen auftreten. HTTP-URIs beispielsweise sollten auflösbar (dereferenzierbar) sein und dienen nicht mehr nur der Identifizierung von Webressourcen, sondern identifizieren auch Objekte der realen Welt:

*„Any HTTP URI should be dereferenceable, meaning that HTTP clients can look up the URI using the HTTP protocol and retrieve a description of the resource that is identified by the URI. This applies to URIs that are used to identify classic HTML documents, as well as URIs that are used in the Linked Data context to identify real-world objects and abstract concepts“ (Heath und Bizer. 2011, S. 10).*

Eine über HTTP dereferenzierbare URI entspricht somit einer „syntaktisch identischen URL“. URIs müssen jedoch nicht zwangsläufig auch URLs darstellen, da sie „in erster Linie für die Identifizierung und Benennung von Entitäten verwendet werden“ (Fürste 2011, S. 46).

---

<sup>40</sup> Da die Paginierung des Dokuments fehlerhaft ist, muss auf die Angabe der Seitenzahl verzichtet werden.

Um zu erreichen, dass die Daten von Maschinen interpretierbar sind, hat das World Wide Web Consortium (W3C) 1999 ein neues Datenmodell vorgestellt (ebd.): Das Resource Description Framework (**RDF**) ist ein Modell mit dem einzelne Aussagen repräsentiert werden können (Pohl 2011) und das sich dabei bei einem grundlegenden Modell von Sprache bedient. Jede Aussage kann durch die Satzglieder Subjekt, Prädikat und Objekt ausgedrückt werden. Diese dreiteilige Struktur, in der ein Prädikat ein Subjekt mit einem Objekt verbindet, „wird als Tripel bezeichnet und bildet das Rückgrat der semantischen Idee“ (Fürste 2011, S. 46) (siehe Abbildung 14).

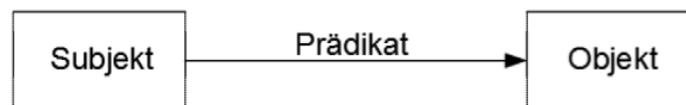


Abbildung 14: Aufbau eines Tripels (Fürste 2011, S. 46)

RDF-Aussagen bestehen somit aus drei Teilen und werden RDF-Tripel genannt. Pohl (2011) veranschaulicht das Ganze an einem Beispiel: Die Aussage „Platon ist der Autor von Sophistes“ besteht aus dem Subjekt „Platon“, dem Objekt „Sophistes“ und dem Prädikat „ist Autor von“. Mit dieser Information kann ein Computer – im Gegensatz zu einem Menschen, der die deutsche Sprache versteht und eine gewisses Mass an Bildung genossen hat – nichts anfangen. Damit der Computer die Aussage verarbeiten kann, sind Verknüpfungen mittels URIs notwendig. Die Aussage würde als RDF-Tripel wie folgt repräsentiert werden:

<[http://dbpedia.org/resource/Sophist\\_\(dialogue\)](http://dbpedia.org/resource/Sophist_(dialogue))>

<<http://purl.org/dc/elements/1.1/creator>>

<<http://dbpedia.org/resource/Plato>>.

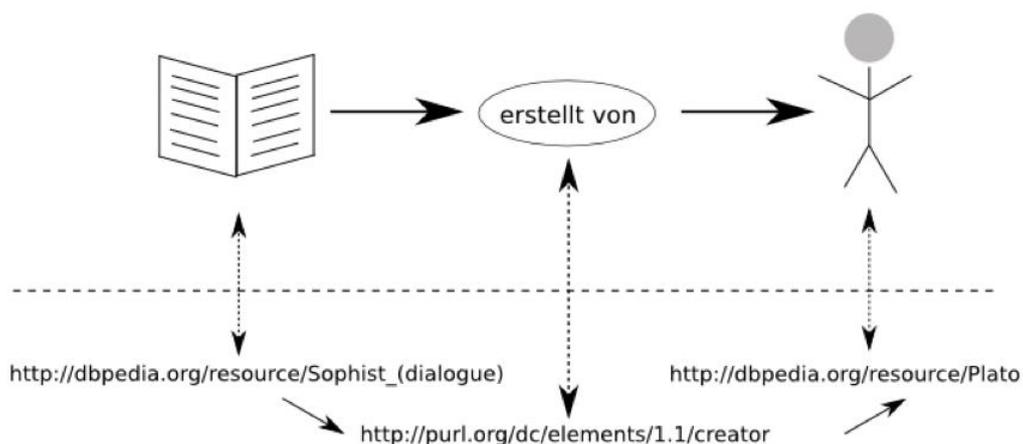


Abbildung 15: Illustration eines RDF-Tripels (Pohl 2011)

Die RDF-Aussage ist in Abbildung 15 illustriert. Die URIs für das Subjekt (Platon) und das Objekt (Sophistes) repräsentieren als eindeutige Identifier Entitäten der Knowledge Base DBpedia. In diesem Fall sind beide Links dereferenzierbare HTTP-URIs, die URLs entsprechen.

Beim Prädikat „handelt es sich um den URI für das Dublin-Core-Metadatenelement ‚creator““ (Pohl 2011). Im Unterschied zu den Links im WWW ist dieser typisiert, d.h. er verlinkt nicht nur auf Dokument, sondern hält zudem fest, *wie* die Art der Verlinkung ist. Dies ist eine grundlegende Eigenschaft des Semantic Web, denn hier ist jeder Link typisiert (ebd.). Im semantischen Kontext werden die Prädikate Properties oder Relationen genannt (Fürste 2011, S. 47).

Die Properties dienen dazu, semantische Entitäten – also Subjekte und Objekte – einzelnen Klassen zuzuordnen. Erst durch die Verknüpfung der Entitäten erhalten diese ihre Bedeutung und erst so wird eine sinnvolle Aussage ermöglicht. Dies verhält sich analog zur natürlichen Sprache. „Ich Pizza“ bekommt erst einen Sinn, wenn ein Prädikat hinzukommt: „Ich mag Pizza“ (ebd.).

Während Subjekt und Prädikat immer über einen URI repräsentiert werden, kann die Repräsentation eines Objekts durch einen URI oder eine Zeichenkette (Literal) geschehen. Literale spielen beispielsweise für Datumsangaben eine Rolle, da hier eine URI nicht sinnvoll wäre (ebd.).

Entitäten und Properties werden in Vokabularen definiert, die als Ontologien bezeichnet werden und ein mächtiges Werkzeug der Wissensrepräsentation darstellen. Klassen und Klassenhierarchien können jedoch nicht durch RDF ausgedrückt werden, da RDF lediglich das Datenmodell bereitstellt. Aus diesem Grund wurde das RDF Schema (**RDFS**) entwickelt, das ein Vokabular bietet, um Klassen von Entitäten sowie Properties, also die Relationen zwischen den Entitäten, zu modellieren (ebd., S. 48-49). RDFS verfügt mit *rdfs* über einen eigenen Namensraum und soll an einem Beispiel erläutert werden:

```
@prefix rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs:<http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd:<http://www.w3.org/2001/XMLSchema#> .

<http://example.org/Buch> rdf:type rdfs:Class .
<http://example.org/Buch> rdfs:subClassOf <http://example.org/Dokument> .

<http://example.org/hatISBN> rdf:type rdfs:Property .
<http://example.org/hatISBN> rdfs:subPropertyOf <http://example.org/hatIdentifier>

<http://example.org/hatISBN> rdfs:domain <http://example.org/Buch> .

<http://example.org/hatISBN> rdfs:range xsd:integer .
```

Abbildung 16: Beispiel einer RDFS-Repräsentation in Turtle (Fürste 2011, S. 49)

Abbildung 16 zeigt ein Beispiel einer RDFS-Präsentation, das im Folgenden Schrittweise nach Fürste (2011, S. 49-50) erläutert werden soll:

- Die ersten drei Zeilen binden die Vokabulare ein. Die Präfixe dienen als Platzhalter für den Namensraum der URI des jeweiligen Referenz-Vokabulars. Die grundlegenden Vokabulare von RDF und RDFS sind in den Vokabularen mit den Präfixen *rdf* und *rdfs* enthalten. In diesen Vokabularen sind die Properties definiert, die für die Zuweisung von Objekten zu den Subjekten benötigt werden.
- *rdfs:Class* typisiert die Klasse *Buch*.
- *rdfs:Property* typisiert die Property *hatISBN*.
- *rdfs:subClassOf* setzt die Klasse *Buch* mit der übergeordneten Klasse *Dokument* in Beziehung.
- *rdfs:subPropertyOf* setzt die Property *hatISBN* mit der übergeordneten Property *hatIdentifizier* in Beziehung.

Die hierarchischen Beziehungen der Klassen und Properties definieren, dass ein *Buch* immer auch ein *Dokument* ist und eine Property *hatISBN* immer auch eine Property *hatIdentifizier*. Durch diese Relationen gehen definierte Eigenschaften von Klassen und Properties (wie Titel eines Dokuments oder syntaktische Vorgaben eines Identifiers) auf die Unterklassen *Buch* und *hatISBN* über.

Durch *rdfs:range* und *rdfs:domain* werden Properties auf bestimmte Subjekt- und Objekttypen beschränkt. Konkret bedeutet dies:

- Die Property *hatISBN* darf nur bei der Klasse *Buch* vorkommen (und nicht bei der übergeordneten Klasse *Dokument*), sie wird also in ihrem Subjekt auf die Klasse *Buch* begrenzt (denn nur Bücher besitzen eine ISBN).
- Die Property *hatISBN* wird in ihrem Objekt durch *rdfs:range xsd:integer* auf einen ganzzahligen Wert begrenzt.

Das Beispiel in Abbildung 16 ist in der Turtle-Syntax notiert. Sie ist eine von mehreren Serialisierungsmöglichkeiten über die RDF verfügt. Das Turtle-Format ist wie auch die N3-Syntax in erster Linie auf die menschliche Rezeption ausgerichtet und somit einfacher zu lesen. Die RDF/XML-Syntax ist hingegen hinsichtlich einer maschinellen Verarbeitung von Bedeutung (Fürste 2011, S. 50). Die RDFS-Repräsentation in Abbildung 16 lässt sich in RDF/XML überführen (siehe Abbildung 17).

```

<?xml version="1.0"?>

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
         xmlns:xsd="http://www.w3.org/2001/XMLSchema#">

  <rdfs:Class rdf:about="http://example.org/Buch">
    <rdfs:subClassOf rdf:resource="http://example.org/Dokument" />
  </rdfs:Class>

  <rdfs:Property rdf:about="http://example.org/hatISBN">
    <rdfs:subPropertyOf rdf:resource="http://example.org/hatIdentifier" />
    <rdfs:domain rdf:resource="http://example.org/Buch" />
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#integer" />
  </rdfs:Property>

</rdf:RDF>

```

Abbildung 17: Beispiel einer RDF/XML-Repräsentation (Fürste 2011, S. 51)

Zusammenfassend lässt sich sagen: „Die Kernaufgaben einer Ontologie sind also die Definition von Klassen, möglichen Beziehungen und möglichen Schlüssen, die aus diesen Beziehungen hervorgehen“ (Pohl 2011).

Zwei populäre Vokabulare, mit denen Entitäten und Properties beschrieben und mit anderen Entitäten verlinkt werden können, sollen kurz erwähnt werden. **FOAF** (Friend Of A Friend)<sup>41</sup> ermöglicht es, personale Information zu beschreiben.

```

<foaf:Person rdf:about="#danbri" xmlns:foaf="http://xmlns.com/foaf/0.1/">
  <foaf:name>Dan Brickley</foaf:name>
  <foaf:homepage rdf:resource="http://danbri.org/" />
  <foaf:openid rdf:resource="http://danbri.org/" />
  <foaf:img rdf:resource="/images/me.jpg" />
</foaf:Person>

```

Abbildung 18: Beispiel einer FOAF-Repräsentation (Brickley und Miller 2010)

Das Beispiel in Abbildung 18 bedeutet: Es gibt eine *foaf:Person* mit einer *foaf:name* Property „Dan Brickley“. Die Person hat über *foaf:homepage* und *foaf:openid* eine Relation zu einer Entität mit Namen *http://danbri.org/* sowie mit *foaf:img* zu einem Objekt, das mit der relativen URI */images/me.jpg* referenziert ist (Brickley und Miller 2010).

**WGS84 Geo Positioning**<sup>42</sup> ist ein Vokabular für die Beschreibung von geografischen Informationen, wie z.B. Koordinaten (siehe Abbildung 19). Mit den Properties *geo:lat* und *geo:long* werden die geografische Breite (Latitude) und geografische Länge (Longitude) eines Punktes auf Grundlage des World Geodetic System 1984 (WGS84)<sup>43</sup> beschrieben.

<sup>41</sup> <http://www.foaf-project.org/> [29.07.2012].

<sup>42</sup> <http://www.w3.org/2003/01/geo/> [29.07.2012].

<sup>43</sup> Das World Geodetic System 1984 (WGS84) ist ein „international vereinbartes konsistentes geodätisches Bezugssystem“ (Bollmann und Koch 2001, S. 438).

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:geo="http://www.w3.org/2003/01/geo/wgs84_pos#">
  <geo:Point>
    <geo:lat>55.701</geo:lat>
    <geo:long>12.552</geo:long>
  </geo:Point>
</rdf:RDF>
```

Abbildung 19: Beispiel einer Geo-Repräsentation (Brickley 2004)

Für die Abfrage von Linked Data wurde die Sprache SPARQL entwickelt, die den Status einer W3C-Empfehlung<sup>44</sup> besitzt. Das rekursive Akronym steht für SPARQL Protocol And RDF Query Language und ist „syntaktisch stark an SQL, der Abfragesprache für relationale Datenbanken, orientiert“ (ebd., S. 53), mit dem Unterschied, dass Tripel und nicht Tupel abgefragt werden. Organisiert sind die Tripel in sogenannten Tripelstores, das sind „auf URIs und kurze Stringliterals optimierte Datenbanken“ (ebd.). SPARQL ermöglicht komplexe Abfragen über die Daten, die in grosser und komplexer Datenbasis vorliegen (Pohl 2011).

Für die Publikation von Daten hat Tim Berners-Lee (2009) vier Linked Data Prinzipien aufgestellt:

- „1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL)
4. Include links to other URIs. So that they can discover more things.“

Dadurch wird die Basis geschaffen, um einzelne LOD-Sammlungen zu verlinken. Das RDF-Datenmodell ermöglicht durch RDF-Links „Querverlinkungen zwischen Daten (sowohl Dokumenten als auch RDF-Repräsentationen“ (Fürste 2011, S. 58). „Ein RDF-Link besteht aus einem Tripel, in dem als Subjekt eine URI-Referenz im Namensraum A mit einer URI im Namensraum B über eine Property (s.B. *owl:sameAs* oder *rdfs:seeAlso*) verbunden ist“ (ebd., S. 58-59). Auf diese Weise können verschiedene Datenquellen miteinander verlinkt werden.

Auf den folgenden Seiten sollen drei dieser Datenquellen vorgestellt und unter die Lupe genommen werden. Unter anderem soll herausgefunden werden, wie die Daten strukturiert sind und ob neben Punktkoordinaten auch Koordinaten von Polygonen vorhanden sind.

<sup>44</sup> <http://www.w3.org/TR/2008/REC-rdf-sparql-query-20080115/> [29.07.2012].

### 3.2.1 DBpedia

Das Projekt DBpedia<sup>45</sup> wurde 2007 von der Universität Leipzig, der Freien Universität Berlin und der Firma OpenLink Software ins Leben gerufen (Isele 2012). Es verfolgt die Absicht, Wikipedia, in eine strukturierte Knowledge Base zu überführen und die Daten im RDF-Standard als Linked Data im Web zur Verfügung zu stellen:

*„DBpedia is a community effort to extract structured information from Wikipedia and to make this information available on the Web. DBpedia allows you to ask sophisticated queries against Wikipedia, and to link other data sets on the Web to Wikipedia data“ (Mendes 2011).*

Zurzeit beschreibt die Knowledge Base von DBpedia mehr als 3,64 Millionen Entitäten (things), wovon 1,83 Millionen in einer konsistenten Ontologie klassifiziert sind. Darin enthalten sind unter anderem 416'000 Personen, 526'000 Orte und 169'000 Organisationen. Mittlerweile besteht das Dataset aus rund 1 Milliarde RDF-Triples, von denen 385 Millionen aus der englischen Wikipedia-Version stammen und 665 Millionen aus Versionen anderer Sprachen, sowie aus Links zu externen Datasets (ebd.). Veröffentlicht wurden die Datensätze unter der Creative Commons Attribution-ShareAlike 3.0 License<sup>46</sup> und der GNU Free Documentation License<sup>47</sup> (ebd.). Verfügbar sind die Daten unter anderem als Raw Data in CSV und als RDF in den Serialisierungen N-Triples, N3/Turtle, JSON und XML. Die Datensets können als Dumps heruntergeladen<sup>48</sup> oder über den SPARQL Endpoint<sup>49</sup> abgefragt werden.

Zum grössten Teil bestehen Wikipedia-Artikel aus Freitext. Es sind jedoch auch verschiedene Formen strukturierter Informationen in Form von Wikitext (wiki markup), einer vereinfachten Auszeichnungssprache, vorhanden, wie beispielsweise Templates für die Infoboxen, Kategorien, Disambiguierungsseiten (Begriffserklärungen), Koordinaten, sowie Links zu externen Webseiten und Links zwischen verschiedenen Sprachversionen von Wikipedia (Bizer et al. 2009, S. 3). Diese strukturierten Daten werden extrahiert und mittels „Wikipedia live article update feed“ immer aktuell gehalten. Jede Entität weist einen zurückverfolgbaren Identifier auf (ebd., S. 2).

Wie sieht es mit der Datenqualität aus? Betrachtet man die Daten, die in diesem Zusammenhang von Bedeutung sind, dann zeichnet sich ein gutes Bild ab. Zumindest was die westliche Hemisphäre betrifft, dürfte die Abdeckung von Geografika ziemlich gross sein

---

<sup>45</sup> <http://dbpedia.org/> [30.06.2012]. Mit DBpedia Deutsch (<http://de.dbpedia.org/> [30.06.2012]) ist als Teil des internationalen Projekts eine Anwendung verfügbar, die sich auf die Daten der deutschen Wikipedia beschränkt.

<sup>46</sup> <http://creativecommons.org/licenses/by-sa/3.0/legalcode> [30.06.2012].

<sup>47</sup> <http://www.gnu.org/licenses/old-licenses/fdl-1.2.en.html> [30.06.2012].

<sup>48</sup> <http://wiki.dbpedia.org/Downloads37> [29.07.2012].

<sup>49</sup> <http://dbpedia.org/sparql> [29.07.2012].

und fast ausnahmslos sind die zugehörigen Koordinaten erfasst. Dies gilt nicht nur für Länder, Kantone, Bundesländer, Städte, Gemeinden, Gewässer, Berge und Täler, sondern auch für viele Bauwerke und Institutionen<sup>50</sup>. Existiert der Artikel jedoch nur in der deutschen Wikipedia, so findet er sich nicht in der DBpedia wieder. Dies ist beispielsweise mit dem Artikel, über das Bernoullianum der Fall (dessen deutschsprachiger Artikel Koordinaten aufweisen würde). Als Datenstamm fungiert die englischsprachige Version, die gleichsam durch die Daten der Artikel anderer Sprachen angereichert wird. Dadurch ist die Mehrsprachigkeit in den verschiedenen Namensvarianten gegeben, was für die Verwendung der Daten für die Georeferenzierung von Katalogdaten von grosser Wichtigkeit ist.

Dass so viele Artikel Koordinatenwerte aufweisen, dafür sorgt das WikiProjekt Georeferenzierung<sup>51</sup>. Das Projekt arbeitet unter anderem daran, möglichst viele Artikel mit WGS84-Koordinaten anzureichern und das georeferenzierte Objekt mit dem gleichen Ländercode, wie er in der GND verwendet wird, einem Staat oder einer Region zuzuordnen. (ISO-Standard ISO 3166 und ISO-3166-2). Die Codes für die Kontinente werden nicht erfasst. DBpedia enthält ausnahmslos Punktkoordinaten. Bei Ländern, Kantonen, Bundesländern, Gemeinden, Städte und Seen sind die Zentrumskoordinaten erfasst. Bei Flüssen wurden die Koordinaten der Mündung aufgenommen.

Am Beispiel der Entität Berlin (<http://dbpedia.org/resource/Berlin>) sollen ein paar Properties vorgestellt werden, die im Kontext dieser Arbeit von Belang sind, bzw. sein könnten:

- *rdf:type* definiert die Entität als zugehörig zu den Klassen *owl:Thing*, *dbpedia-owl:Place* und *dbpedia-owl:City*
- *rdfs:label* enthält den Namen der Entität in verschiedenen Sprachen, z.B. Berlin, Berliini, Berlino
- *geo:lat* enthält die geografische Breite als Gleitkommazahl (*xsd:float*), z.B. 52.500557
- *geo:long* enthält die geografische Länge als Gleitkommazahl (*xsd:float*), z.B. 13.398889
- *dbpprop:isoRegion*<sup>52</sup> enthält den ISO-Ländercode, z.B. DE-BE
- *owl:sameAs* enthält die Verknüpfung zur entsprechenden Entität eines anderen Vokabulars, z.B. <http://sws.geonames.org/2950159/>

---

<sup>50</sup> So sind beispielsweise die Artikel über die Universität Basel, die ETH Zürich, das Kunstmuseum Basel und die Kunsthalle Zürich georeferenziert, nicht jedoch die Universitätsbibliothek Basel im deutschsprachigen Artikel, da die Koordinaten jedoch im englischsprachigen Artikel vorhanden sind, finden sie sich auch in der DBpedia. Die HTW Chur ist weder georeferenziert, noch besitzt sie einen englischsprachigen Artikel.

<sup>51</sup> [http://de.wikipedia.org/wiki/Wikipedia:WikiProjekt\\_Georeferenzierung](http://de.wikipedia.org/wiki/Wikipedia:WikiProjekt_Georeferenzierung) [06.07.2012].

<sup>52</sup> Diese Property tritt auch in der Form *dbpprop:isoCodeRegion* auf.

### 3.2.2 GeoNames

GeoNames<sup>53</sup> ist eine geografische Datenbank, die unter der Creative Commons Attribution 3.0 Licence<sup>54</sup> verfügbar ist:

*„It contains over 10 million geographical names and consists of over 8 million unique features whereof 2.8 million populated places and 5.5 million alternate names. All features are categorized into one out of nine feature classes and further subcategorized into one out of 645 feature codes“ (GeoNames 2012).*

Die Datenbank enthält Toponyme in zahlreichen Sprachen sowie WGS84-Koordinaten, Höhe über Meer, Einwohnerzahlen, administrative Unterteilungen und Postleitzahlen. Die Daten stammen aus unterschiedlichen freien und öffentlichen Quellen<sup>55</sup> und können von eingeloggten Benutzern bearbeitet werden (ebd.). Das GeoNames User Manual<sup>56</sup> beschreibt, wie Daten editiert, falsch referenzierte Marker verschoben, neue Ortsnamen zur Datenbank hinzufügt oder alternative Bezeichnungen ergänzt werden können. Auf diese Weise ermöglicht GeoNames die Qualität der Daten laufend zu verbessern.<sup>57</sup>

---

<sup>53</sup> <http://www.geonames.org/> [07.07.2012].

<sup>54</sup> <http://creativecommons.org/licenses/by/3.0/> [30.07.2012].

<sup>55</sup> Unter den Datenquellen befindet sich neben Wikipedia unter anderem nationale geografische Institute wie Swisstopo oder National Geospatial-Intelligence Agency, statistische Ämter wie das Statistische Amt des Kantons Zürich oder das Amt für Statistik Berlin-Brandenburg sowie Online-Anbieter von Hotelzimmer. Die gesamte Liste ist unter <http://www.geonames.org/data-sources.html> verfügbar [30.07.2012].

<sup>56</sup> <http://www.geonames.org/manual.html> [07.07.2012].

<sup>57</sup> Das User Interface bietet den Benutzern eine einfache und erweiterte Suche sowie die Möglichkeiten durch Listen zu browsen oder auf einer Google Map zu navigieren. In Facetten werden die Klassen und die einzelnen Schlagwörter („feature codes“) angezeigt und können einzeln ausgewählt werden. Verschieden farbige Marker lokalisieren die Points of Interest (POI) auf der Karte. Unter dem Kartenausschnitt erscheint die Liste der angezeigten Treffer. Die Trefferliste lässt sich als Datei im CSV-Format herunterladen, die unter anderem GeoNameId, Name des Orts, Ländercode und Koordinaten enthält. Zudem kann die Trefferliste als Grafik ausgegeben werden, in der die einzelnen POI als rote Punkte auf einer einfachen Karte verzeichnet sind.

```

- <rdf:RDF>
- <gn:Feature rdf:about="http://sws.geonames.org/3020251/">
  <rdfs:isDefinedBy>http://sws.geonames.org/3020251/about.rdf</rdfs:isDefinedBy>
  <gn:name>Embrun</gn:name>
  <gn:alternateName xml:lang="la">Aebura</gn:alternateName>
  <gn:alternateName xml:lang="oc">Ambrun</gn:alternateName>
  <gn:alternateName xml:lang="la">Eburudunum</gn:alternateName>
  <gn:featureClass rdf:resource="http://www.geonames.org/ontology#P"/>
  <gn:featureCode rdf:resource="http://www.geonames.org/ontology#P.PPL"/>
  <gn:countryCode>FR</gn:countryCode>
  <gn:population>7069</gn:population>
  <gn:postalCode>05200</gn:postalCode>
  <gn:postalCode>05201 CEDEX</gn:postalCode>
  <gn:postalCode>05202 CEDEX</gn:postalCode>
  <gn:postalCode>05208 CEDEX</gn:postalCode>
  <gn:postalCode>05209 CEDEX</gn:postalCode>
  <wgs84_pos:lat>44.56387</wgs84_pos:lat>
  <wgs84_pos:long>6.49526</wgs84_pos:long>
  <gn:parentFeature rdf:resource="http://sws.geonames.org/6446638"/>
  <gn:parentCountry rdf:resource="http://sws.geonames.org/3017382"/>
  <gn:parentADM1 rdf:resource="http://sws.geonames.org/2985244"/>
  <gn:parentADM2 rdf:resource="http://sws.geonames.org/3013738"/>
  <gn:parentADM3 rdf:resource="http://sws.geonames.org/3016701"/>
  <gn:parentADM4 rdf:resource="http://sws.geonames.org/6446638"/>
  <gn:nearbyFeatures rdf:resource="http://sws.geonames.org/3020251/nearby.rdf"/>
  <gn:locationMap rdf:resource="http://www.geonames.org/3020251/embrun.html"/>
  <gn:wikipediaArticle rdf:resource="http://de.wikipedia.org/wiki/Embrun"/>
  <gn:wikipediaArticle rdf:resource="http://en.wikipedia.org/wiki/Embrun%2C_Hautes-Alpes"/>
  <rdfs:seeAlso rdf:resource="http://dbpedia.org/resource/Embrun%2C_Hautes-Alpes"/>
  <gn:wikipediaArticle rdf:resource="http://fr.wikipedia.org/wiki/Embrun_%28Hautes-Alpes%29"/>
  <gn:wikipediaArticle rdf:resource="http://it.wikipedia.org/wiki/Embrun"/>
  <gn:wikipediaArticle rdf:resource="http://nl.wikipedia.org/wiki/Embrun"/>
  <gn:wikipediaArticle rdf:resource="http://oc.wikipedia.org/wiki/Ambrun"/>
  <gn:wikipediaArticle rdf:resource="http://pl.wikipedia.org/wiki/Embrun"/>
  <gn:wikipediaArticle rdf:resource="http://vo.wikipedia.org/wiki/Embrun_%28Hautes-Alpes%29"/>
</gn:Feature>
- <foaf:Document rdf:about="http://sws.geonames.org/3020251/about.rdf">
  <foaf:primaryTopic rdf:resource="http://sws.geonames.org/3020251"/>
  <cc:license rdf:resource="http://creativecommons.org/licenses/by/3.0"/>
  <cc:attributionURL rdf:resource="http://sws.geonames.org/3020251"/>
  <cc:attributionName rdf:datatype="http://www.w3.org/2001/XMLSchema#string">GeoNames</cc:attributionName>
  <dcterms:created rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2006-01-15</dcterms:created>
  <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2012-03-30</dcterms:modified>
</foaf:Document>
</rdf:RDF>

```

Abbildung 20: Die RDF-Repräsentation der Stadt Embrun in GeoNames<sup>58</sup>

Die Objekte sind als RDF-Tripel verfügbar und können als Dumps heruntergeladen werden<sup>59</sup>.

Abbildung 20 zeigt die Repräsentation der Stadt Embrun. Die Entität wird durch den (dereferenzierbaren) URI <http://sws.geonames.org/3020251/>, der die 7-stellige GeoNameId enthält, identifiziert. Ein paar der Properties sollen hervorgehoben werden:

- *gn:name* enthält den Namen der Entität
- *gn:alternateName* enthält alternative Namensvarianten
- *gn:countryCode* enthält ein zweistelliges Länderkürzel
- *wgs84\_pos:lat* und *wgs84\_pos:long* beschreiben die Koordinaten
- *gn:parentCountry* stellt eine hierarchische Relation mit dem Objekt Frankreich her, das durch den URI <http://sws.geonames.org/3017382/> definiert ist.
- *gn:wikipediaArticle* verlinkt auf Wikipedia-Artikel über Embrun in verschiedenen Sprachen.

<sup>58</sup> <http://sws.geonames.org/3020251/> [20.07.2012].

<sup>59</sup> <http://download.geonames.org/export/dump/> [07.07.2012].

Die Qualität der Daten ist schwierig einzuschätzen, da sie aus sehr unterschiedlichen Quellen in die Datenbank gelangen. Es ist davon auszugehen, dass die Qualität stets verbessert wird, da die Benutzer daran arbeiten können. Polygone, also Bounding Boxes sind auch bei GeoNames nicht vorhanden. Jedoch bietet GeoNames über das GeoNames API einen Web Service<sup>60</sup>, der es erlaubt, diejenigen Wikipedia Artikel abzufragen, deren Koordinaten sich innerhalb einer definierten Bounding Box befinden. Das Ergebnis ist ein XML-Dokument.

### 3.2.3 LinkedGeoData

Stadler (2012) beschreibt das Projekt LinkedGeoData<sup>61</sup> (LGD) mit den Worten:

*„**LinkedGeoData** is an effort to add a spatial dimension to the Web of Data / Semantic Web. LinkedGeoData uses the information collected by the OpenStreetMap project and makes it available as an RDF knowledge base according to the Linked Data principles. It interlinks this data with other knowledge bases in the Linking Open Data initiative“ (Hervorhebung im Original).*

OpenStreetMap<sup>62</sup> (OSM) ist ein kollaboratives Projekt, das eine frei editierbare Weltkarte erstellt. Von Wikipedia inspiriert bietet es bekannte Wiki-Funktionalitäten wie Bearbeitungsfunktionen und Versionsgeschichte, mit dem Unterschied, dass nicht Artikel sondern geografische Entitäten editiert werden (Stadler et al. 2011, S. 2). Die wichtigsten Entitäten sind Knoten (nodes), Wege (ways) und Relationen (relations):

*„Nodes are the most primitive entities and represent geographic points with a latitude and longitude relative to the WGS84 reference system.*

*Ways are entities that have a list of at least two node references associated with them. Depending on whether the first reference equals the last one, a way is called closed or open, respectively.*

*Relations relate points, ways and potentially other relations to each other, thereby forming complex objects. Each entity participating in a relation plays a certain role in it. Multipolygons are modelled with relations“ (ebd., Hervorhebungen im Original).*

Jede Entität besitzt einen numerischen Identifier, genannt OSM ID. Die Knowledge Base von LinkedGeoData umfasst ca. 20 Milliarden RDF-Triples und besteht aus über einer Milliarde Nodes, etwa 100 Millionen Ways (Stadler 2012) und fast einer Million Relations (Stadler et

---

<sup>60</sup> <http://www.geonames.org/export/wikipedia-webservice.html#wikipediaBoundingBox> [30.07.2012].

<sup>61</sup> <http://linkedgeo.org/> [07.07.2012].

<sup>62</sup> <http://www.openstreetmap.org/> [07.07.2012].

al. 2011, S. 3). Lizenziert sind die Daten unter der Creative Commons Attribution-ShareAlike 2.0 Generic (CC BY-SA 2.0) License<sup>63</sup> (Stadler 2011).

LinkedGeoData ist mit DBpedia, GeoNames und der Food and Agriculture Organization of the United Nations (FOA), der Ernährungs- und Landwirtschaftsorganisation der Vereinten Nationen, verlinkt. Stadler et al. (2011, S. 9) beschreiben das Vorgehen bei der Verlinkung:

*„In all cases, we first manually aligned the classes of these knowledge bases with classes from LinkedGeoData on a best effort basis. The interlinking is then done on a per-class basis, where all instances of a set of classes of LGD are matched with all instances of a set of classes of another data source using labels and spatial information. As an example, cities in LGD and DBpedia are matched between all instances of *lgdo:City*, *lgdo:Town*, *lgdo:Village* and *lgdo:Hamlet* on one side and *dbo:Settlement* on the other“ (Hervorhebungen im Original).*

Im Gegensatz zu DBpedia weisen die Datensätze von LinkedGeoData nicht nur Punktkoordinaten auf, sondern wie bereits erwähnt auch Ways und Relations. Doch das Problem ist, dass nicht alle Daten in gleicher Weise verfügbar sind. So sind aus Gründen der Performance nicht alle OSM-Entitäten via SPARQL Endpoint<sup>64</sup> verfügbar (Stadler et al. 2011, S. 4). Welche Datensätze fehlen, wird nicht konkretisiert. Ein weiterer Zugang wird über REST API angeboten, doch die Abfragemöglichkeit ist eingeschränkt, denn gegenwärtig sind zwar alle Nodes und Ways, jedoch keine Relations als RDF-Daten suchbar (Stadler 2011, S. 4). Die kompletten Datensätze sind als Download verfügbar, in HTML<sup>65</sup> und als Dumps<sup>66</sup> oder als Linked Data via 303 Content Negotiation in den Formaten RDF/XML, Turtle, N-Triples oder HTML (ebd., S. 7-8).

Wie eine Relation aussehen kann soll am Beispiel der Staatsgrenze von Deutschland mit der OSM ID 51477<sup>67</sup> gezeigt werden. Insgesamt ist die deutsche Grenze aus über 1000 Ways zusammengesetzt, die alle ein bestimmtes Segment repräsentieren. So besteht etwa die Grenze mit Luxemburg aus rund 40 Ways. Diese Relation enthält über 30 Tags, die beispielsweise für die Zeitzone oder die Währung gebraucht werden. Dass es sich bei der Relation um eine Staatsgrenze handelt wird durch den Tag „boundary=administrative“ angegeben (ebd., S. 2f.).

<sup>63</sup> <http://creativecommons.org/licenses/by-sa/2.0/> [07.07.2012]. Gemäss Stadler et al. (2011, S. 3) ist geplant, zur Open Database License zu wechseln (<http://opendatacommons.org/licenses/odbl/>). [07.07.2012].

<sup>64</sup> Angeboten werden sowohl ein „static SPARQL endpoint“ (<http://linkedgeodata.org/sparql>) als auch ein „live SPARQL endpoint“ (<http://live.linkedgeodata.org/sparql>) [07.07.2012].

<sup>65</sup> <http://linkedgeodata.org/Datasets/> [07.07.2012].

<sup>66</sup> <http://downloads.linkedgeodata.org/> [07.07.2012].

<sup>67</sup> <http://www.openstreetmap.org/browse/relation/51477> [07.07.2012].

### 3.2.4 Fazit

Die drei untersuchten Datenbanken stellen alle eine grosse Menge an Koordinaten bereit. Stichproben haben gezeigt, dass die Datenlage für eine Georeferenzierung insgesamt sehr gut ist – zumindest was Punkte betrifft. Bei Flächen sieht es allerdings anders aus. In DBpedia und GeoNames sind alle Entitäten mit Punktkoordinaten georeferenziert, während LinkedGeoData komplexe Flächen wie administrative Grenzen in Form von Multipolygon-Relationen darstellt. Koordinaten für einfache Polygone in Form von Bounding Boxes waren nicht ausfindig zu machen. Ob die Multipolygone geeignet sind, um Katalogdaten zu georeferenzieren, müssten Tests zeigen. Es ist zu erwarten, dass sich die grosse Datenmenge negativ auf die Suchperformance auswirken könnte. Es sollte jedoch möglich sein, aus den LGD-Relationen die vier benötigten Maxima zu extrahieren.



```
{
  "results" : [
    {
      "address_components" : [
        {
          "long_name" : "Schweiz",
          "short_name" : "CH",
          "types" : [ "country", "political" ]
        }
      ],
      "formatted_address" : "Schweiz",
      "geometry" : {
        "bounds" : {
          "northeast" : {
            "lat" : 47.80845460,
            "lng" : 10.49234010
          },
          "southwest" : {
            "lat" : 45.817920,
            "lng" : 5.956080
          }
        },
        "location" : {
          "lat" : 46.8181880,
          "lng" : 8.227511999999999
        },
        "location_type" : "APPROXIMATE",
        "viewport" : {
          "northeast" : {
            "lat" : 47.80845460,
            "lng" : 10.49234010
          },
          "southwest" : {
            "lat" : 45.817920,
            "lng" : 5.956080
          }
        }
      },
      "types" : [ "country", "political" ]
    }
  ],
  "status" : "OK"
}
```

Abbildung 21: Screenshot Bounding Box Schweiz (Google Geocoding API) [31.07.2012]

Eine Alternative für die Koordinaten von Flächen bietet das Google Geocoding API<sup>68</sup>, über das Bounding Boxes ausgegeben werden können. Geliefert werden die Koordinaten der südwestlichen und nordöstlichen Ecke des Rechtecks. Die Abfrage

<sup>68</sup> <https://developers.google.com/maps/documentation/geocoding/#ReverseGeocoding> [31.07.2012].

*<http://maps.googleapis.com/maps/api/geocode/json?address=Switzerland&sensor=false>* liefert die Bounding Box für die Schweiz (siehe Abbildung 21).

Für die Georeferenzierung von Katalogdaten dürften alle betrachteten Datenbanken in Frage kommen. Allenfalls könnten auch hier Testanreicherungen Aufschluss über die Qualität der Datenbasis geben. Die Verwendung von DBpedia bietet sich an, da sie gleichsam der Nukleus des Web of Data bildet. Sie ist mit zahlreichen anderen LOD-Sätzen verlinkt, insbesondere mit RDF-Tripel aus der Bibliothekswelt. Zudem sind mittlerweile zahlreiche Einträge von Personen mit Identifikatoren von Personennormdaten angereichert. Der Datensatz von Friedrich Nietzsche beispielsweise enthält neben der Library of Congress Control Number (LCCN) die Identifier der GND, der VIAF sowie der Web NDL Authorities, der Normdatei der Nationalen Parlamentsbibliothek in Japan. Interessant ist die Verlinkung zur GND vor allem deshalb, weil die Daten der GND auch als Linked Open Data publiziert sind.

## 4 Die Georeferenzierung

Nachdem die Analyse der Daten gezeigt hat, dass eine Georeferenzierung von bibliografischen Metadaten machbar ist, soll nun untersucht werden, auf welche Weise die Daten angereichert werden können. Im Anschluss daran werden mögliche Anwendungen diskutiert.

### 4.1 Varianten der Anreicherung

Grundsätzlich gibt es drei verschiedene Möglichkeiten, die Katalogdaten mit Hilfe von Linked Open Data zu georeferenzieren (siehe Abbildung 22). Denkbar ist, die Koordinaten direkt in Feld 034 in die bibliografischen Metadaten des Katalogs des IDS Basel Bern aufzunehmen (Variante 1). Eine andere Lösung sieht vor, die Daten ad hoc bei der Suchabfrage zu georeferenzieren (Variante 2). Die Anreicherung des Index stellt Variante 3 dar. Im Folgenden sollen die Vor- und Nachteile der drei Ansätze kurz diskutiert werden.

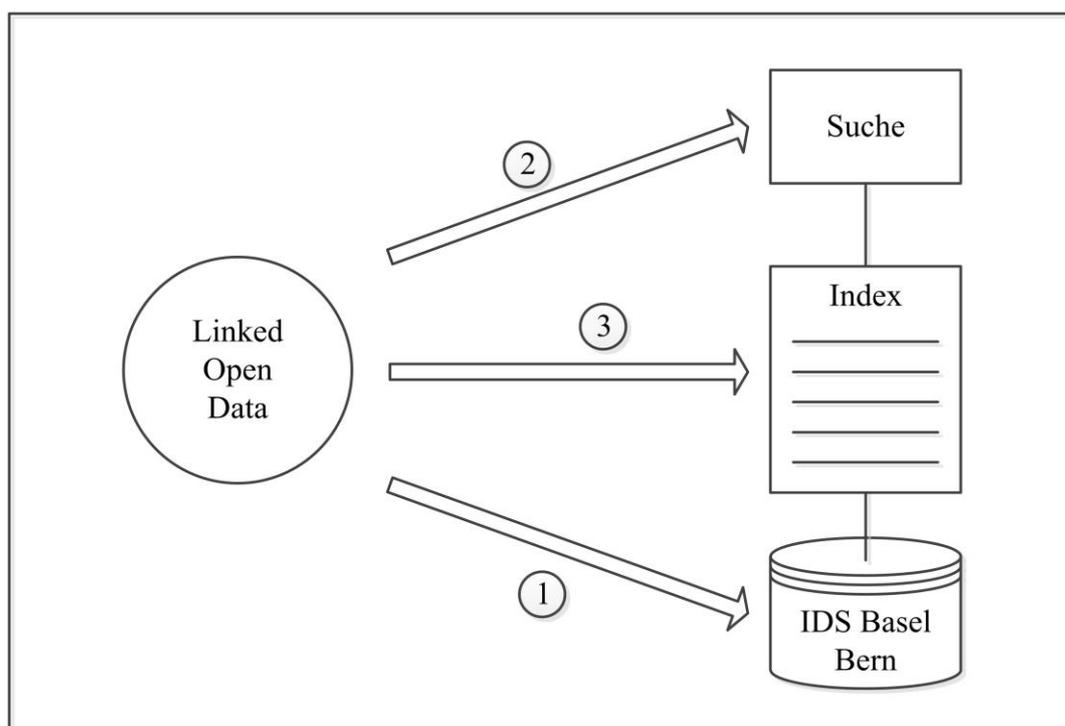


Abbildung 22: Varianten der Anreicherung durch Linked Open Data

#### 4.1.1 Anreicherung der Katalogdaten

Die Variante, die Katalogdaten mit Koordinaten anzureichern, mag die naheliegendste und auf den ersten Blick plausibelste sein. Gemäss KIDS ist die Erfassung von Koordinaten jedoch ausschliesslich für Kartenmaterial vorgesehen. Eine Regeländerung dürfte nicht oder nicht ohne weiteres durchzusetzen sein und soll ihm Rahmen dieser Arbeit auch nicht

angeregt werden. Selbst wenn das Regelwerk das Erfassen von Koordinaten erlauben würde, stellten sich ein paar grundlegende Fragen. Wie die Analyse der Katalogdaten gezeigt hat, bieten sich diverse Felder für die Georeferenzierung an. Wie wäre die Information unterzubringen, ob der Erscheinungsort, die Orte des Inhalts oder die Orte der Autoren georeferenziert worden sind? Und wie liessen sich die unterschiedlichen Typen von geografischen Bezügen kennzeichnen, wenn in einem einzigen bibliografischen Datensatz mehrere Felder georeferenziert würden? Diese Angaben wären nicht nur für die Maschine von Belang, damit sie den geografischen Bezug richtig zu interpretieren weiss, sondern auch für den Benutzer.

Nach derzeitigem Stand müsste nicht nur das Feld 034 sondern auch das Feld 255 erfasst werden und dieses wird im OPAC angezeigt. Dem Benutzer wäre verständlich zu machen, warum die Metadaten eines Buches mit Koordinaten versehen sind, bzw. welcher Aspekt des Dokuments georeferenziert wurde. Dieser Umstand würde die Katalogaufnahme im Fall von mehrfacher Referenzierung unnötig aufblähen und unübersichtlich machen.

Nicht zuletzt wäre die Anreicherung durch Koordinaten beim Metadatenaustausch mit anderen Bibliotheken problematisch. Von dieser Variante ist aus den genannten Gründen abzusehen.

#### **4.1.2 Anreicherung bei der Suchabfrage**

Geschieht die Georeferenzierung der Daten zum Zeitpunkt der Suchabfrage, hätte dies den Vorteil, dass einerseits den oben genannten Problemen aus dem Weg gegangen werden könnte und andererseits die verwendeten LOD immer auf dem neuesten Stand wären. Die Aktualität der Daten wäre ein klarer Vorteil gegenüber der ersten Lösung.

Der zu erwartende Knackpunkt bei diesem zweiten Ansatz ist die Performance. Die Ad-hoc-Georeferenzierung ist sehr rechenintensiv und dürfte zu langen Antwortzeiten führen, was alles andere als benutzerfreundlich wäre. Schliesslich sind sich die Benutzer durch die gebräuchlichen Suchmaschinen sehr schnelle Rechercheergebnisse gewohnt.

Um diese Form der Anreicherung überhaupt umsetzen zu können, wäre eine zusätzliche Hürde zu meistern: Die Katalogdaten müssten in RDF-Tripels umgewandelt werden. Nur wenn die kompletten bibliografischen Metadaten als Linked Data zur Verfügung stehen, ist eine Verlinkung mit den LOD anderer Datenanbietern möglich.

Die Distribution von Bibliotheksdaten als LOD wär ein eigenes Projekt und soll an dieser Stelle nicht weiter verfolgt werden. Der Ad-hoc-Anreicherung ist jedoch aus Gründen der unzureichenden Suchperformanz eine Absage zu erteilen. Bleibt also Variante 3.

### 4.1.3 Anreicherung des Index

Am besten eignet sich für die Anreicherung der Katalogdaten ein separater Suchindex, der auf die jeweilige Anwendung zugeschnitten ist. Die Daten könnten direkt aus dem IDS Basel Bern oder analog zur Geosuche aus Swissbib exportiert und in einer gesonderten Datenbank gespeichert werden. Dadurch bietet sich die Freiheit, die Daten so zu manipulieren, wie es die konkrete Anwendung erfordert. Dem Beispiel des Kartenportals folgend kann es sinnvoll sein, die Katalogaufnahmen bereits beim Aufbau des Index einem Ranking zu unterziehen, um die Suchperformanz zu optimieren.

Damit die verwendeten Daten stets aktuell sind, muss der Index regelmässig neu aufgebaut werden. Auf diese Weise ist garantiert, dass sowohl die Katalogdaten stets auf dem neuesten Stand als auch allfällige Korrekturen in den LOD-Ressourcen nachgeführt sind.

Die Anreicherung des Suchindex hat gegenüber den beiden genannten Varianten den Vorteil, dass keine Änderungen der Katalogisierungsregeln nötig wären, die Katalogdaten so belassen werden könnten, wie sie sind und die Suchperformanz nicht übermässig beansprucht wird. Zudem könnte man darauf verzichten, den gesamten Datenbestand in RDF zu überführen. Es würde reichen, einzelne Beziehungen zu anderen LOD-Datenbanken herzustellen.

Ein weiterer Vorteil gegenüber Variante 2 ist, dass ein Ort jeweils nur einmal georeferenziert werden muss. Der Parser könnte die Koordinaten bei bereits georeferenzierten Daten aus dem Index holen und so alle anderen Datensätzen, die diesen Ort enthalten, verarbeiten. Es wäre somit nicht nötig, die Daten in jedem Fall neu zu laden. Damit dies funktioniert, braucht es ein Format, bei dem die semantische Beziehung zwischen Toponym und Koordinaten erhalten bleibt.

## 4.2 Mögliche Anwendungen

Aufgrund der Analyse der Katalogdaten lohnt es sich, drei abgegrenzte Teilbestände zu betrachten, anhand derer mögliche Umsetzungen von Georeferenzierungen und den damit verbundenen Anwendungen diskutiert werden sollen. Dabei rücken unterschiedliche Aspekte von geografischen Bezügen in den Fokus der Verfahren: Der Atlas der Basiliensia wertet das Erscheinungsland aus; Grundlage für die geografische Suche nach alten Drucken ist die Georeferenzierung des Druckorts und bei der geografischen Themensuche werden Geografika und Ländercode der GND georeferenziert. Neben der Georeferenzierung spielen die User Interfaces mit ihren Recherche-Funktionen eine zentrale Rolle.

Alle Varianten besitzen ähnliche Benutzerschnittstellen, die je nach Anwendung unterschiedliche Funktionen aufweisen können. Im Wesentlichen sollen sie sich in Aufbau und Funktionalitäten am User Interface der geografischen Suche des Kartenportals

orientieren (siehe Abbildung 5). Zentrales Element ist die interaktive Karte, die neben den Zoom- und Drag-Funktionen auch den sehr nützlichen „drag zoom mode“ anbietet. Dieser Modus lässt sich über das Lupen-Symbol anwählen und erlaubt es, mit dem Cursor Bounding Boxes, also Suchrechtecke, aufzuziehen, um ausgewählte Gebiete gezielt zu vergrößern. Der Ausschnitt wird automatisch zentriert und in der Grösse an das Kartensuchfeld angepasst. Neben dem Geobrowsing stellt die Wortsuche nach Ortsnamen eine weitere Suchmöglichkeit zur Verfügung. Zusätzlich können je nach Anwendung weitere Suchschlitze angeboten werden. Auch der Zeitstrahl, mit dem sich der Erscheinungszeitraum einschränken lässt, ist hilfreich und überaus praktisch in der Handhabung. Allerdings wäre es denkbar, den abgedeckten Zeitraum und die Skalierung auf den jeweiligen Bestand anzupassen, so dass nur der Zeitraum erscheint, in dem auch tatsächlich Dokumente erschienen sind. Nicht fehlen darf die Ergebnisliste mit den Links in den Katalog. Gewisse Anwendungen erfordern zusätzlich die Möglichkeiten, die Suche durch Drilldown-Menüs oder Dropdown-Listen zu modifizieren. Aus Platzgründen können zusätzlichen Filter und Facetten in aufklappbaren Layern oder Fenstern untergebracht sein. Auf den Massstabsfilter kann verzichtet werden, da dieser nur bei der Kartensuche eine Rolle spielt.

Neben diesen Funktionen können entsprechend den Anforderungen einzelner Umsetzungen auch Marker eingesetzt werden, wenn es darum geht, Ergebnislisten als Point of Interest zu visualisieren. Nach diesen allgemeinen Vorbemerkungen sollen nun die angedachten Varianten erläutert werden.

#### 4.2.1 Atlas der Basiliensia

Interessant könnte die räumliche Auswertung des Erscheinunglandes, also des MARC-Ländercodes in Feld 008 sein, wenn es um Publikationen geht, die in der Basler Bibliografie verzeichnet sind:

*„Die Basler Bibliographie erfasst selbstständige und unselbstständige Publikationen aus allen Sachgebieten mit inhaltlichem Bezug zum Kanton Basel-Stadt und zur Region Basel. Berücksichtigt werden dabei auch Personen, deren Leben und Werk – unabhängig vom Bürgerrecht – mit diesem Raum verbunden sind“ (Universitätsbibliothek Basel 2012a).*

Begonnen wurde die Basler Bibliografie im Jahr 1919 und seit 2003 wird sie nur noch elektronisch erfasst. Interessant für die Georeferenzierung sind diese Dokumente, da sie einen klar abgrenzbarer Spezialbestand bilden: Alle Publikationen, die in die Bibliografie aufgenommen werden, erhalten einen Lokalcode im Feld 909A \$b. Der Code setzt sich zusammen aus dem Kürzel *basb* (für Basler Bibliografie) sowie zwei Ziffern für das Erscheinungsjahr, z.B. *basb12*. Mit der trunkierten Expertensuche *WLC=basb?* kann die Suche auf Publikationen der Bibliografie eingeschränkt werden.

Würden Ländercodes georeferenziert werden, wäre es möglich, die Werke gemäss ihrem Erscheinungsland auf einer Weltkarte zu verorten. In diesem Fall würden die Zentrumskoordinaten der Länder ausreichen. Für jedes Land, in dem Publikationen mit Bezug zu Basel erschienen sind, könnte ein Marker gesetzt werden, der die jeweiligen Werke zusammenfasst. Durch ein Klicken auf diesen Marker würde eine Ergebnisliste ausgegeben, die alle referenzierten Werke enthält. Mit einem Zeitstrahl könnte das Erscheinungsjahr eingeschränkt werden. Zudem könnten verschiedene Facetten eingebunden werden, die beispielsweise nach Schlagwörtern oder Autoren filtern lassen.

Bei dieser Anwendung soll jedoch weniger die Recherche, als vielmehr die Darstellung des Teilbestands im Vordergrund steht. So könnte auf einen Blick erfasst werden, in welchen Ländern Publikationen mit einem Bezug zu Basel erschienen sind (und sich im Bestand der UB Basel befinden). Interessant wäre es, die Grösse der Marker proportional zur Menge der Publikationen darzustellen, damit auch die quantitative Verteilung ersichtlich ist. Dies würde im konkreten Fall der Basler Bibliografie allerdings zu Schwierigkeiten führen, denn von den über 22'000 Titel sind gerade einmal 2'000 Titel ausserhalb von der Schweiz und Deutschland erschienen. Die Skalierung der Marker müsste so gewählt werden, dass diejenigen, die am meisten Titel in sich vereinen, nicht zu viel Platz einnehmen würden, aber dennoch die Relationen bei kleinen Mengen deutlich werden könnten.<sup>69</sup>

Die Georeferenzierung wäre in diesem Fall relativ einfach umzusetzen. Es würde reichen, den Ländercode, bzw. die toponymische Entsprechung, die in ALEPH hinterlegt ist, und die Systemnummer in RDF-Tripel zu überführen. Das Erscheinungsland könnte beispielsweise über die Property *ub:landOfPublication* repräsentiert sein. In einer Konkordanz müsste diese mit der Property *rdfs:label* in Beziehung gesetzt werden. Danach könnten die Objekte der Properties abgeglichen werden. Bei Übereinstimmung kann die entsprechende DBpedia-Entität mit dem bibliografischen Datensatz verlinkt und die Koordinaten (Werte der Properties *geo:lat* und *geo:long*) in den Index eingebunden werden. Das Ergebnis dürfte sehr gut sein, denn es ist davon auszugehen, dass in DBpedia alle Länder vorhanden und georeferenziert sind. Die Marker könnten mittels Google Maps API gesetzt werden.

---

<sup>69</sup> Im skizzierten Fall ist die Anwendung vielleicht etwas wenig aussagekräftig, da kein Auftrag besteht, der die UB Basel in ihrer Funktion als Kantonsbibliothek verpflichten würde, alle Basiliensia zu sammeln. Sie sammelt besagte Werke nur „sofern ein wissenschaftliches oder lokalhistorisches Interesse an ihnen besteht“ (Universitätsbibliothek Basel 2012b). Wirklich interessant könnte eine solche Anwendung beispielsweise für die Schweizerische Nationalbibliothek sein, die mit Artikel 3 des Bundesgesetzes vom 18. Dezember 1992 einen klaren Auftrag besitzt, (möglichst) alle Helvetica zu sammeln. Hier dürfte der Anteil an Publikationen, die ausserhalb der Schweiz und Deutschland erschienen sind, grösser sein.

## 4.2.2 Die geografische Suche alter Drucke

Während beim Atlas der Basiliensia die Darstellung im Vordergrund steht, soll bei der Georeferenzierung von alten Drucken neben der Darstellung die Recherche eine wichtige Rolle spielen. Auch dieser Teilbestand ist klar abgrenzbar: Alle Titel, die ein Feld 752 \$d aufweisen, könnten referenziert werden.

Die verschiedenen Druckereien einer Stadt würden in einem einzigen Marker zusammengefasst. Analog zum Kartenportal.CH könnten in der Ergebnisliste alle Drucke angezeigt werden, die in den Druckorten erschienen sind, die sich innerhalb des Kartenausschnitts befinden. Die einzelnen Druckereien eines Druckorts könnten in einem Pop-up-Fenster aufgeführt sein, das aufspringt, wenn der Cursor über einem Marker steht (Mouseover). Zusätzlich soll die Möglichkeit geboten werden über Facetten nach bestimmten Druckorten und Druckereien zu filtern. Diese könnten ähnlich wie bei der geografischen Facette als Drilldown-Menü konzipiert sein (siehe Abbildung 23). Zwischen Druckort und Druckereien könnte eine Art hierarchische Relation hergestellt werden, so dass die Offizine gleichsam als Unterbegriffe der Druckorte auftreten. Die Liste wird dynamisch aufgebaut und enthält nur diejenigen Orte (mit den zugehörigen Druckereien), die sich innerhalb des Suchfeldes befinden. Länder und Offizine können einzelnen ausgewählt oder ausgeschlossen werden. Als dritte hierarchische Ebene könnten die einzelnen Länder in die Facette mit einbezogen werden.

- Amsterdam (89)
  - Janssonius, Johannes (Offizin, Amsterdam) (47)
  - Mortier (Offizin, Amsterdam) (23)
  - Wit, Frederick de (Offizin, Amsterdam) (12)
  - Ohne Druckerei (7)
- Basel (103)
  - Officina Oporiniana (Basel) (56)
  - Ostein, Leonhard (Offizin, Basel) (34)
  - Schneider, Felix (Offizin Basel) (9)
  - Waldkirch, Konrad von (Offizin, Basel) (3)
  - Ohne Druckerei (1)
- Nürnberg (95)
  - Homann, Johann Baptist, Erben (Offizin, Nürnberg) (87)
  - Schneider und Weigel (Offizin, Nürnberg) (1)
  - Ohne Druckerei (7)

Abbildung 23: Beispiel eines Drilldown-Menüs für Druckereien

Bei diesem Lösungsansatz sind Darstellung und Recherche ungefähr gleichbedeutend. Der Benutzer sieht auf einen Blick, in welchen Städten Druckereien existierten. Er kann also gezielt nach Drucken aus einem bestimmten Gebiet suchen, ohne wissen zu müssen, wo überhaupt Druckereien ansässig waren. Eine weitere Vereinfachung ist der Umstand, dass der Benutzer es sich sparen kann, jeden einzelnen Ort einzutippen. Die Facetten bieten ihm die Möglichkeit, die Trefferliste nach einzelnen Orten oder Offizinen zu filtern und mittels Zeitleiste lässt sich der Zeitraum eingrenzen.

Wie bei der ersten Anwendung wäre auch in diesem Fall die Georeferenzierung relativ einfach zu bewerkstelligen. Wählt man GeoNames als Quelle, so müssten die verschiedenen Druckorte mit den Namen (Property *gn:name*) oder den alternativen Namen (Property *gn:alternateName*) der Entitäten nach Übereinstimmungen abgesucht werden. Auch hier könnte darauf verzichtet werden, die kompletten Katalogaufnahmen in RDF zu überführen. Es würde reichen den Druckort als Tripel zu codieren (z.B. mit der Property *ub:imprint*) und eine Konkordanz zwischen den beiden Namensäumen herzustellen. Für die Koordinaten sind die Properties *wgs84\_pos:lat* und *wgs84\_pos:long* ausschlaggebend.

### 4.2.3 Die geografische Themensuche

Im Gegensatz zu den beiden ersten Anwendungsbeispielen geht es bei der geografischen Themensuche nicht um die Anzeige von Trefferlisten, sondern einzig und allein um die Recherche. Aus diesem Grund sind bei dieser Lösung auch keine Marker vorgesehen. Um thematische Suchabfragen zu ermöglichen, braucht es neben der Möglichkeit nach Ortsnamen zu suchen einen Suchschlitz für die Eingabe von thematischen Begriffen.

Bei der geografischen Themensuche spielt die Flächensuche neben der Punktsuche eine entscheidende Rolle. Die Idee ist es, Orte (Dörfer und Städte) als Punkt zu georeferenzieren, Gebiete (Kantone, Bundesländer und Staaten) jedoch als Fläche in Form von Bounding Boxes. Diese Polygone könnten bei der Suche aufgrund von Lage und Ausdehnung in gleicher Art und Weise berücksichtigt werden, wie dies bei der Kartensuche der Fall ist. Das bedeutet, dass alle im Suchfeld enthaltenen oder von diesem angeschnittenen Polygone einen Treffer ergeben. Dies hat den Vorteil, dass man sich bei länder- oder kantonsübergreifender Suche nicht um die Namen der einzelnen Gebiete kümmern muss, ja noch nicht einmal die administrativen Einheiten kennen muss.

Ein Beispiel: Interessiert sich jemand für den Tourismus im Dreiländereck, also in Basel und der grenzüberschreitenden Umgebung, so reicht es, das gewünschte Gebiet ins Suchfeld zu rücken und gefunden werden alle Dokumente mit Bezug zu den Orten Basel, Saint-Louis, Lörrach, zu den Kantonen Basel-Stadt und Basel-Landschaft, zur Region Elsass und zum Bundesland Baden-Württemberg, sowie zu allen weiteren im Suchfeld enthaltenen und davon angeschnittenen Gebieten. Gefunden würden somit auch Dokumente über Tourismus

in der Schweiz, in Deutschland und Frankreich. Damit kann der Recall deutlich erhöht werden, allerdings dürfte auch einiger Ballast in der Ergebnisliste landen. Diesem Problem kann mit einem guten Ranking-Algorithmus, ähnlich demjenigen von MapRank, entgegen gewirkt werden. Ergebnisse müssten in der Trefferliste umso weiter nach hinten rücken, je kleiner das angeschnittene Gebiet im Verhältnis zu Gesamtgrösse ist, d.h. ein Kanton, der zu 60 % vom Suchfeld abgedeckt ist, erhält ein besseres Ranking, als ein Kanton, der gerade einmal zu 10 % abgedeckt ist. Ein gutes Ranking ist auch deshalb so wichtig, weil das Suchfeld so gut wie nie ausschliesslich das gewünschte Gebiet enthält. Die Vorteile und die Nachteile einer geografischen Suche und die Grenzen des Rankings lassen sich gut an einem weiteren Beispiel verdeutlichen (siehe Abbildung 24).

Sucht jemand Dokumente zum Thema Tourismus im Kanton Luzern und rückt das entsprechende Gebiet ins Zentrum des Suchfeld, kann nicht vermieden, dass auch Dokumente zu Bern gefunden werden. Wie die Grafik verdeutlicht wird die Bounding Box des Kantons Luzern (rot eingefärbt) fast vollständig von derjenigen des Kantons Bern (grau) abgedeckt. Das hat zur Folge, dass Dokumente zu Luzern nicht ohne Treffer zu Bern gefunden werden können. Zwar lässt sich, wie bereits erwähnt, das Problem durch ein gutes Ranking abschwächen. Im umgekehrten Fall hingegen hilft das Ranking nicht weiter. Die Suche nach Dokumente, die den Kanton Bern betreffen, schliessen (unter anderem) immer Dokumente über Luzern mit ein. Die nicht-relevanten Dokumente lassen sich nur über eine geografische Facette vollständig ausschliessen. Durch eine geeignete Georeferenzierung kann der Ballast zwar nicht vollständig ausgeschlossen werden, aber sie kann dazu dienen, neben dem Grad der Überlappung von Polygonen einen weiteren Faktor ins Ranking einzubringen.

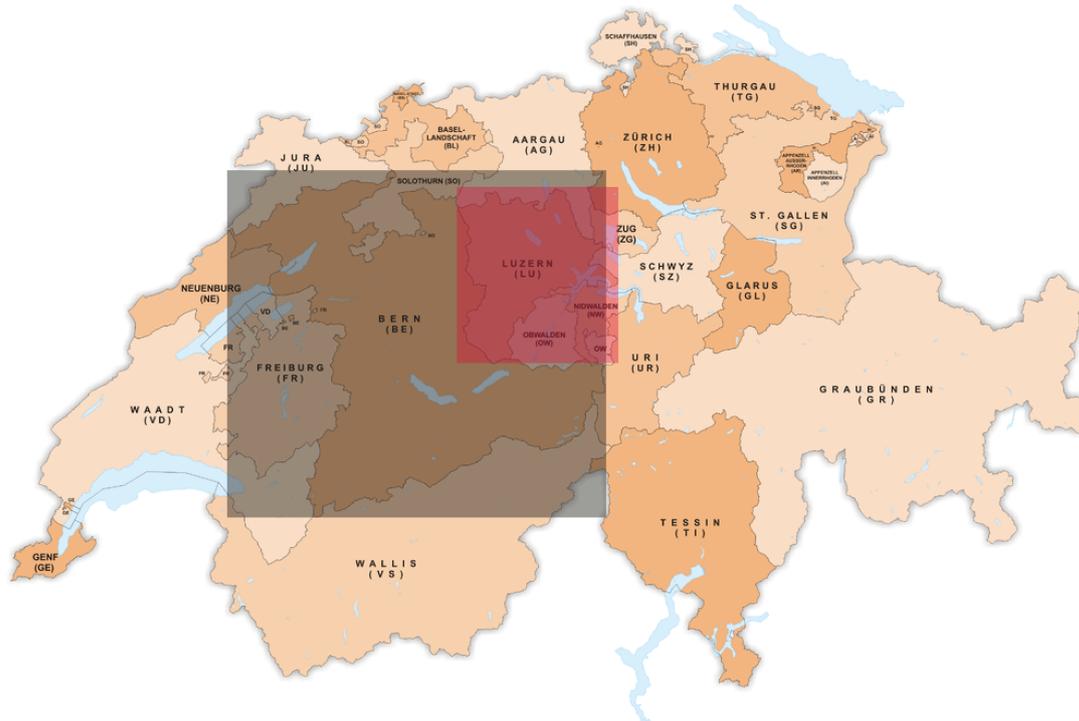


Abbildung 24: Bounding Boxes bei der Recherche (Tschubby 2011, Hervorhebung P.W.)

Die Georeferenzierung erfolgt bei dieser Anwendung einerseits über die Toponyme und andererseits über die Ländercodes der GND. Ein Geografikum wird also doppelt georeferenziert. Karlsruhe wird beispielsweise über das Toponym eine Punktkoordinate zugeordnet und über den Ländercode XA-DE-BW die Flächenkoordinaten für Baden-Württemberg. Auf diese Weise lässt sich ein ähnlicher Effekt erzielen, wie Wiesenmüller et al. (2011) ihn bei ihren Testrecherchen erzielt haben. Auch wenn sich die Punktkoordinaten von Karlsruhe nicht innerhalb des Suchfeldes befinden, werden trotzdem Dokumente mit diesem Schlagwort berücksichtigt, wenn die Fläche von Baden-Württemberg vom Suchfeld angeschnitten ist. Der Recall wird erhöht, das Ranking gewichtet den entsprechenden Titel jedoch weniger stark.

In der doppelten Georeferenzierung liegt eine Chance, das Ranking zu verbessern. Entgegen der eingangs gemachten Kategorisierung sollen neben den Toponymen von Orten auch die Toponyme von Gebieten (Bundesländer, Kantone und Staaten) über die Zentrumskoordinaten zusätzlich als Punkt georeferenziert werden. Der Ländercode wird in jedem Fall als Fläche georeferenziert. Die Zentrumskoordinate kann das Ranking verbessern, wenn berücksichtigt wird, wie weit die Zentrumskoordinate des vom Suchfeld angeschnittenen Gebiets von der Zentrumskoordinate des Suchfelds entfernt ist. Neben der Entfernung soll zudem eine Rolle spielen, ob die Zentrumskoordinaten des angeschnittenen Polygons innerhalb oder ausserhalb des Suchfeldes liegt.

Wie sich gezeigt hat, stellt LinkedGeoData durch Relationen modellierte Multipolygone zur Verfügung, die administrative Grenzen repräsentieren. Würden die sehr exakten Polygone zur Georeferenzierung verwendet, dürften sich die Suchergebnisse nur unwesentlich verbessern, da das Suchfeld ein Rechteck ist und dadurch immer eine Unschärfe bei der Suchabfrage bleibt. Aufgrund dessen lässt sich feststellen, dass einfache Bounding Boxes für die Georeferenzierung ausreichend sind. Die Präzision der Suche liesse sich allerdings steigern, wenn sich die Form des Suchrechtecks durch den Benutzer verändern liesse. Je nach Gestalt eines Gebiets liefert ein hochformatiges Suchrechteck gegenüber einem querformatigen die exakteren Treffer.

Eingerichtet werden könnte die geografische Themensuche beispielsweise innerhalb des Afrika Portals<sup>70</sup> des Zentrums für Afrikastudien Basel. Die African Studies sind eine der Schwerpunkte der Universität Basel wodurch diese „die Leading-House-Funktion unter den Schweizer Universitäten“ einnimmt (Universität Basel, 2007, S. 23). Auf dem Portal ist bereits eine Suchmaske<sup>71</sup> eingerichtet, die eine Recherche im IDS Basel Bern erlaubt, bei der nur Dokumente mit Bezug zu Afrika berücksichtigt werden. Die entsprechenden Aufnahmen sind im Feld 909A \$f mit verschiedenen Codes identifiziert, die alle mit *afbs* beginnen (z.B. *afbs-mono* für Monografien). Durch die geografische Themensuche liessen sich die Suchergebnisse deutlich verbessern, unter anderem deshalb, weil nicht nur codierte Dokumente berücksichtigt werden müssten, um einen geografischen Bezug zu Afrika herstellen zu können. Die Codes werden erst seit ein paar Jahre automatisch vergeben, ältere Bestände wurden nachträglich von Hand codiert.

Ein Nachteil einer solchen Anwendung ist der geringe Detailierungsgrad in den Ländercodes der afrikanischen Geografika. Der dreistufige Ländercode der GND wird wie bereits erwähnt nur für Deutschland, Österreich und die Schweiz verwendet. Die übrigen Länder sind zweistufig codiert.

Die Realisierung der geografischen Themensuche setzt allerdings eine grössere Abdeckung von GND-Schlagworten voraus, als dies derzeit der Fall ist. Um die Qualität der Suchergebnisse auf ein angemessenes Mass zu steigern, sollte erwogen werden, den Katalog nachträglich und automatisiert anzureichern. Eine mögliche Lösung wird im folgenden Exkurs aufgezeigt.

## **Exkurs II: Verbundübergreifende Kataloganreicherung mit SWD-Schlagwörtern**

Magnus Pfeffer (2012) zeigt in seinem am Bibliothekartag in Hamburg gehaltenen Referat auf, wie sich Katalogdaten automatisiert mit SWD-Schlagwörtern und RVK-Notationen

---

<sup>70</sup> <http://zasb.unibas.ch/de/portal/> [31.07.2012].

<sup>71</sup> <http://zasb.unibas.ch/de/portal/bibliotheken/> [31.07.2012].

(Regensburger Verbundklassifikation) anreichern lassen. Idee ist es, die Beschlagwortung und die Klassierung<sup>72</sup> aus Voraufgaben, Parallelausgaben und Übersetzungen zu übernehmen. In einem Projekt der Jahre 2010-2011 wurden die Verbunddatenbanken Südwestdeutscher Bibliotheksverbund (SWB) und Hessisches Bibliotheksinformationssystem (HeBIS) als Ausgangsdaten<sup>73</sup> verwendet (ebd., S. 8). Der Vergleich der Daten geschah auf der Grundlage von Einheitssachtitel, Titel und Untertitel, Autoren und Urheber sowie beteiligte Personen und Körperschaften. Ein Algorithmus wurde entwickelt, der Übereinstimmungen zwischen (Einheitssach-)Titel und mindestens einer Person bzw. Körperschaft aufspürt. Ist eine Übereinstimmung vorhanden, erfolgt das Matching der Titel und die SWD und RVK werden übernommen (ebd., S. 9f.).

Hier sollen nur die Ergebnisse für die SWD erwähnt werden, da die RVK in diesem Kontext nicht von Belang sind. Im SWB wurden über 5,8 Millionen Titel mit mindestens einem Match gefunden. Davon sind knapp 3,27 Millionen Titel ohne SWD. Eine Anreicherung war bei über 636'000 Titeln möglich, das sind fast 20 % der gematchten Titel ohne Schlagwörter. Im HeBIS sind die Ergebnisse noch besser ausgefallen. Von über 4,5 Millionen Titeln mit einem Match waren über 3 Millionen ohne SWD. Bei knapp 1,18 Millionen Titel war eine Anreicherung durch die Übernahme der Schlagwörter möglich, das entspricht über 38 % der gematchten Titel ohne Schlagwörter (ebd., S. 11-12).

Die Daten werden zur Nachnutzung sowohl als Download im Textformat (Titel-ID und gefundene Matches) als auch als Linked Open Data angeboten<sup>74</sup>. Die RDF-Tripel haben die Form ID>equalsForClassification-ID. Zudem wurden die Daten (Titel und gefundene SWD-IDs und RVK-Notationen) an die Verbundzentralen geliefert (ebd., S. 13).

Die Tests, die in den Testdatenbanken der Verbünde sowie im „Linked-Data Web“ durchgeführt worden sind, bestätigten die hohe Qualität der Ergebnisse (ebd., S. 14). Zudem hat sich gezeigt, dass umso mehr Titeldaten angereichert werden können, je mehr Quellen für die Übernahmen von SWD verwendet werden. Konkret heisst das: Wäre das Matching im SWB nur mit den Daten des eigenen Verbundes durchgeführt worden, wären die Anreicherung nicht bei den erwähnten 636'000 Titeln sondern nur bei etwas mehr als 450'000 Titeln möglich gewesen. Durch die Verwendung der Titeldaten aus HeBIS ist die Anreicherung um fast 30 % höher ausgefallen (ebd., S. 15).

---

<sup>72</sup> Pfeffer spricht fälschlicherweise von „Klassifizierung“, das gemäss Stock und Stock (2008, S. 192) „Bilden von Klassen“ bedeutet. In diesem Zusammenhang wäre „Klassieren“ der korrekte Begriff, da er „die Zuordnung von Klassen eines konkreten Klassifikationssystems zu einem gegebenen Dokument“ bezeichnet. Aus diesem Grund wird im vorliegenden Text „Klassieren“ anstelle von „Klassifizieren“ verwendet.

<sup>73</sup> Die Katalogdaten der beiden Verbünde liegen im MAB2-Format vor. Die Methode lässt sich jedoch problemlos auf MARC-Daten übertragen.

<sup>74</sup> <http://data.bib.uni-mannheim.de/>. Zurzeit ist der Linked Data Service nicht zugänglich, da eine neue Version entwickelt wird [26.07.2012].

Auf dieser Grundlage hat Kai Eckert von der UB Mannheim ein Software-Tool<sup>75</sup> entwickelt, das die Retroklassierung von grossen Beständen ermöglicht. Der praktische Einsatz hat gezeigt, dass je nach Bestand im Durchschnitt zwischen rund 320 und 430 Titel pro Stunde retroklassiert werden können (ebd., S. 27).

---

<sup>75</sup> <http://www.kaiec.org/reclassification.html> [26.07.2012].

## 5 Empfehlungen

Beabsichtigen Bibliotheken ihre Bestände zu georeferenzieren und mittels eines Geobrowsing-Tools zugänglich zu machen, sollten sie ein klares Projekt formulieren. Es gilt zu entscheiden, welcher Teilbestand georeferenziert werden soll. Von der Georeferenzierung des gesamten Bestandes ist eher abzusehen. Vor allem ist festzulegen, welcher der drei möglichen geografischen Bezüge (Erscheinungsort, Orte der Autoren oder Orte des Inhalts) relevant sein soll. Wichtig ist zudem die Definition einer Zielgruppe: An wen soll sich Anwendung richten? Welches Suchinteresse soll befriedigt werden?

Viel Wert ist auf die Suchfunktionalitäten zu legen: Welche Suchfelder und welche Filter sollen angeboten werden? Welche Facetten helfen, Hierarchien abzubilden und die Ergebnisliste sinnvoll einzuschränken? Eine gewichtige Rolle kommt dem Ranking zu. Gerade bei grossen Treffermengen oder bei der Flächensuche ist die Gewichtung der Treffer unverzichtbar.

Bevor komplexe Anwendungen umgesetzt werden, kann es sich lohnen mit der Implementierung einfacher Anwendungen Erfahrungen zu sammeln. Empfehlenswert sind in jedem Fall Test-Georeferenzierungen. Dadurch lassen sich Fehlerquellen und Qualitätsmängel eruieren. Die bereits vorhandenen Lösungsansätze Viaticapes und Kartenportal.CH, wie auch das geplante BIBonMAP zeigen, dass bereits ein beachtliches Mass an Sachkenntnis und Erfahrung vorhanden ist, was die Georeferenzierung und die Umsetzung von Recherche-Tools betrifft. Auf dieses Wissen sollte zurückgegriffen werden, denn es ist nicht nötig, das Rad immer aufs Neue zu erfinden.

Nicht zuletzt kann nur dafür plädiert werden, Linked Open Data für die Kataloganreicherung und die Erschliessung von Beständen zu nutzen. Die Georeferenzierung ermöglicht einen guten Anfang, aus dem riesigen Schatz an verfügbaren strukturierten Daten zu schöpfen und dadurch die Suchergebnisse zu verbessern.

In bestimmten Fällen kann es durchaus angebracht sein, die Qualität der Eintragungen in den ausgewerteten Datenfeldern zu erhöhen, indem Katalogaufnahmen durch den Einsatz von intellektueller Arbeit bereinigt und ergänzt werden. Die Konsolidierung der Daten mag zwar relativ zeitaufwändig und teuer sein, kann aber zu entscheidender Verbesserung führen, was Qualität und Quantität der einbezogenen Daten betrifft.



## 6 Schlussbemerkungen

Das Web of Data bietet eine immense Quelle an Daten, die für die Kataloganreicherung genutzt werden kann. Gerade im Bereich der Geodaten sind sehr umfangreiche strukturierte Daten in guter Qualität vorhanden. Der Vorteil von Linked Open Data ist, dass sie maschinenlesbar sind und eine automatisierte Anreicherung ermöglichen. Auf diese Weise lassen sich grössere Bibliotheksbestände georeferenzieren, was von Hand nur unter erheblichem Aufwand zu leisten wäre.

Die Georeferenzierung von Katalogdaten kann neue Zusammenhänge aufzeigen. Durch die Lokalisierung in einem räumlichen Kontext stehen die Dokumente in einem Bezug zu einander, der in herkömmlichen Bibliothekskatalogen nicht zu Tage tritt. Im besten Fall kann die Betonung des räumlichen Aspekts von Dokumenten zu neuen Erkenntnissen führen. Möglich ist, dass dies beim Atlas der Basiliensia der Fall sein kann, wenn ersichtlich wird, in welchen Teilen der Welt man sich mit Basel beschäftigt – und in welchen nicht.

Für ausgewählte Teilbestände bietet die Georeferenzierung die Chance der Visualisierung. Dokumente können einzeln oder in Clusters mittels Marker auf der Karte lokalisiert werden. So lässt sich der Bestand ausbreiten und durch das Geobrowsing werden Zufallstreffer ermöglicht. Die Verortung der Druckereien kann beispielsweise dazu beitragen, unbekannte Druckorte zu entdecken.

Auf der Grundlage von georeferenzierten bibliografischen Metadaten lassen sich Anwendungen entwickeln, die den Benutzern alternative Einstiege bei der Recherche bieten. Die skizzierten Anwendungen ermöglichen es, die Suchfunktionen auf die Eigenheiten eines Teilbestands anzupassen. Unterschiedliche Bestände erfordern unterschiedliche Recherchemöglichkeiten. Neben passgenauen Funktionalitäten kann ein ausgeklügeltes Ranking nötig sein, um grosse Treffermengen sinnvoll zu gewichten.

Ein gutes Ranking ist bei der Flächensuche, wie sie bei der geografischen Themensuche angedacht ist, unabdingbar. Mit dem MapRank-Algorithmus ist in dieser Hinsicht ein effizientes Werkzeug umgesetzt worden, das sich nicht nur bei der Kartensuche sondern allgemein bei der Suche nach Bounding Boxes einsetzen liesse. Allerdings scheint bei der Distribution von Koordinaten für umschliessende Rechtecke von Seiten der LOD-Anbieter noch Nachholbedarf zu bestehen.

Allen in vorliegender Arbeit skizzierten Anwendungsbeispielen ist gemeinsam, dass sie mit wenig Ontologie auskommen. Es ist nicht zwingend notwendig, alle Datenfelder in RDF-Triples zu überführen; oft reicht die Transformation von einzelnen Feldern aus oder es kann ganz darauf verzichtet werden. Nichtsdestotrotz ist zu erwägen, die gesamten bibliografischen Metadaten des betroffenen Teilbestandes in RDF zu überführen und zu

veröffentlichen. Schliesslich ist die Georeferenzierung nur eine von vielen Möglichkeiten der Nutzung von Linked Open Data und kann gleichsam als Einstieg in die Welt des Semantic Webs dienen. Es ist empfehlenswert, weitere Informationen miteinzubeziehen, sei es, um hierarchische Beziehungen wie beispielsweise zwischen Städten und Ländern für Facetten auszuwerten oder die Suchergebnisse mit Angaben zu den Offizinen anzureichern. Aus der GND, aus VIAF oder der DBpedia können Lebensdaten, Lebens-, Sterbe- und Wirkungsorte und Verbindungen zu anderen Personen eingebunden werden oder es lassen sich Angaben zu Titeln aus anderen Bibliothekskatalogen, wie etwa der Deutschen Nationalbibliothek ausgeben.

Vorliegende Arbeit versteht die Anwendungsbeispiele als möglichen Einstieg in die Nutzung der verfügbaren LOD-Datensets. Die skizzierten Anwendung sind relativ einfach realisierbar, beruhen jedoch auf einer Technologie, die zu wesentlich mehr in der Lage ist, als auf den vorangegangenen Seiten dargelegt worden ist. Es geht darum, einen Anfang zu machen, die Quelle anzapfen und erste Verlinkungen ins Web of Data zu setzen. Letztlich steht und fällt die Idee der Linked Open Data mit der Nutzung der Daten. Es reicht nicht aus, wenn Bibliotheken – und gerade Bibliotheken – nur ihre Daten veröffentlichen.

## Quellenverzeichnis

### 6.1 Mündliche Quellen

Bugnon, Nicolas, 06.06.2012

Geisser, Franziska, 12.06.2012

Spinnler, Alice, 05.06.2012

Thomas, Guy, 29.05.2012

### 6.2 Schriftliche Quellen

**Archief Eemland:** *Amersfoort op the kaart: Colofon*. Verfügbar unter:

<http://www.amersfoortopdekaart.nl/over-deze-site/colofon/> [22.06.2012].

**Berners-Lee, Tim (2009):** *Linked Data – Design Issues*. Verfügbar unter:

<http://www.w3.org/DesignIssues/LinkedData.html> [30.06.2012].

**Berners-Lee, Tim; Hendler, James; Lassila, Ora (2001):** *The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities*. Scientific American 284 (5), S. 34-43. Verfügbar unter:

<http://www.scientificamerican.com/article.cfm?id=the-semantic-web> [30.06.2012].

**Bill, Ralf; Zehner, Marco L. (2001):** *Lexikon der Geoinformatik*. Heidelberg: Herbert Wichmann.

**Bizer, Christian; Lehmann, Jens; Kobilarov, Georgi; Auer, Sören; Becker, Christian; Cyganiak, Richard; Hellmann, Sebastian (2009):** *DBpedia – A Crystallization Point for the Web of Data*. Verfügbar unter: <http://www.wiwiss.fu-berlin.de/en/institute/pwo/bizer/research/publications/Bizer-et-al-DBpedia-CrystallizationPoint-JWS-Preprint.pdf> [29.06.2012].

**Bollmann, Jürgen; Koch, Wolf Günther (Hrsg.) (2001):** *Lexikon der Kartographie und Geomatik: in zwei Bänden*. Heidelberg: Spektrum Akademischer Verlag.

**Brickley, Dan; Miller Libby (2010):** *FOAF Vocabulary Specification 0.98: Namespace Document 9 August 2010 - Marco Polo Edition*. Verfügbar unter:

<http://xmlns.com/foaf/spec/20100809.html> [29.07.2012].

**Brinkhoff, Thomas (2008):** *Geodatenbanksysteme in Theorie und Praxis: Einführung in objektrelationale Geodatenbanken unter besonderer Berücksichtigung von Oracle Spatial* (2., überarb. und erw. Aufl.). Heidelberg: Wichmann.

**Cygniak, Richard (2011):** *The Linking Open Data cloud diagram* (Version 10.09.2011).

Verfügbar unter: <http://richard.cyganiak.de/2007/10/lod/> [27.07.2012].

**ETH-Bibliothek (2012):** *BIBonMAP*. Verfügbar unter: <http://www.library.ethz.ch/de/Ueberuns/Projekte/BIBonMAP> [22.06.2012].

**Fürste, Fabian (2011):** *Linked Open Library Data: Bibliographische Daten und ihre Zugänglichkeit im Web der Daten*. B.I.T.online – Innovativ, hrsg. v. Rolf Fuhlrott, Ute Krauss-Leichert, Christoph-Hubert Schütte, Bd. 33. Wiesbaden: Dinges & Frick.

**GeoNames (2012):** *About GeoNames*. Verfügbar unter:

<http://www.geonames.org/about.html> [07.07.2012].

**Heath, Tom; Bizer, Christian (2011):** *Linked Data: Evolving the Web into a Global Data Space* (1st edition). Synthesis Lectures on the Semantic Web: Theory and Technology #1. Morgan & Claypool. Verfügbar unter: <http://linkeddatabook.com/editions/1.0/> [06.07.2012].

**IDS (2011):** *KIDS – Katalogisierungsregeln IDS* (Version 10.10.2011). Verfügbar unter:

[http://kids.informationsverbund.ch/kids\\_deutsch/a0kidsinh.pdf](http://kids.informationsverbund.ch/kids_deutsch/a0kidsinh.pdf) [05.07.2012].

**Isele, Robert (2012):** *DBpedia: Credits*. Verfügbar unter: <http://wiki.dbpedia.org/Team> [30.06.2012].

**Johnston, Lisa R.; Jensen, Kristi L. (2009):** *MapHappy: A User-Centered Interface to Library Map Collections Via a Google Maps „Mashup“*. *Journal of Map & Geography Libraries* 5, 2 (2009), S. 114-130. Verfügbar unter: <http://purl.umn.edu/92083> [23.06.2012].

**Ley, Andreas (2009):** *bm-maps.org Pflichtenheft*. Unveröffentlicht. Küsnacht ZH.

**Library of Congress (2008):** *MARC code list for countries* (2003 edition, last updated April 4, 2008). Verfügbar unter: [http://www.loc.gov/marc/countries/cou\\_home.html](http://www.loc.gov/marc/countries/cou_home.html) [28.06.2012].

**Mendes, Pablo (2011):** *DBpedia: About*. Verfügbar unter: <http://wiki.dbpedia.org/About> [30.06.2012].

**Neubert, Joachim; Borst, Timo (2012):** *Die Nutzung von LOD-Personennormdaten zur Disambiguierung bei Erfassung und Recherche*. *Foliensatz*. Vortrag gehalten am 101. Bibliothekartag in Hamburg. Verfügbar unter: <http://www.opus-bayern.de/bib-info/volltexte/2012/1240/> [24.07.2012].

**Oehrli, Markus; Přidal, Petr; Zollinger, Susanne; Siber, Rosi (2011):** *MapRank: Geographical Search for Cartographic Materials in Libraries*. In: *D-Lib Magazine*,

ember/October 2011, Vol. 17, Nr. 9/10. Verfügbar unter:

<http://dlib.org/dlib/september11/oehrli09oehrli.html> [23.06.2012].

**Pfeffer, Magnus (2012):** *Jetzt kommt zusammen, was zusammen gehört:*

*Verbundübergreifende Datenkonsolidierung und wie man damit grosse Bestände effizient retroklassifiziert. Foliensatz. Vortrag gehalten am 101. Bibliothekartag in Hamburg.*

Verfügbar unter: <http://www.opus-bayern.de/bib-info/volltexte//2012/1203/> [25.07.2012].

**Pfeifer, Barbara (2012):** *Gemeinsame Normdatei (GND).* Verfügbar unter:

[http://www.dnb.de/DE/Standardisierung/Normdaten/GND/gnd\\_node.html](http://www.dnb.de/DE/Standardisierung/Normdaten/GND/gnd_node.html) [25.07.2012].

**Pohl, Adrian (2011):** *Linked Data und die Bibliothekswelt. Preprint.* Verfügbar unter:

[http://www.hbz-nrw.de/dokumentencenter/produkte/lod/aktuell/pohl\\_2011\\_linked-data\\_ODOK.pdf](http://www.hbz-nrw.de/dokumentencenter/produkte/lod/aktuell/pohl_2011_linked-data_ODOK.pdf) [15.07.2012].

**Scheven, Esther (2012):** *Ländercodes in der SWD.* Verfügbar unter:

[http://www.dnb.de/DE/Standardisierung/Normdaten/SWD/swd\\_laendercodes.html](http://www.dnb.de/DE/Standardisierung/Normdaten/SWD/swd_laendercodes.html)  
[10.07.2012].

**Stadler, Claus (2011):** *LinkedGeoData Data Set.* Verfügbar unter:

<http://linkedgeodata.org/Datasets> [07.07.2012].

**Stadler, Claus (2012):** *LinkedGeoData: About.* Verfügbar unter:

<http://linkedgeodata.org/About> [07.07.2012].

**Stadler, Claus; Lehmann, Jens; Höffner, Konrad; Auer, Sören (2011):** *LinkedGeoData: A Core for a Web of Spatial Open Data.* Semantic Web Journal 2011. Verfügbar unter:

<http://www.semantic-web-journal.net/content/linkedgeodata-core-web-spatial-open-data>  
[07.07.2012].

**Stock, Wolfgang G. (2007):** *Information Retrieval: Informationen suchen und finden.*

München etc.: Oldenbourg.

**Stock, Wolfgang G.; Stock, Mechthild (2008):** *Wissensrepräsentation: Informationen auswerten und bereitstellen.* München etc.: Oldenbourg.

**Thomas, Guy (2012):** *www.barchives.org: Online-tool zur Präsentation und Vermittlung des historisch-kulturellen Erbes im Archiv der Basler Mission. Unveröffentlicht.* Basel.

**Tschubby (2011):** *Karte Kantone der Schweiz farbig 2011.* In Wikipedia. Verfügbar unter:

[http://de.wikipedia.org/w/index.php?title=Datei:Karte\\_Kantone\\_der\\_Schweiz\\_farbig\\_2011.png&filetimestamp=20110116171128](http://de.wikipedia.org/w/index.php?title=Datei:Karte_Kantone_der_Schweiz_farbig_2011.png&filetimestamp=20110116171128) [27.07.2012].

**Universität Basel (Hrsg.) (2007):** *Strategie 2007: Zur Entwicklung der Universität Basel 2007-2013.* Basel. Verfügbar unter:

<http://www.unibas.ch/index.cfm?3F207D103005C8DEA3E318DD51878884> [07.07.2012].

**Universitätsbibliothek Basel (2012a):** *Basler Bibliographie: Informationen*. Verfügbar unter: <http://www.ub.unibas.ch/ub-hauptbibliothek/recherche/kataloge/spezialkataloge/basler-bibliographie/informationen/> [31.07.2012].

**Universitätsbibliothek Basel (2012b):** *Sammelschwerpunkte der Universitätsbibliothek Basel*. Verfügbar unter: <http://www.ub.unibas.ch/ub-hauptbibliothek/wir-ueber-uns/weiteres/sammelschwerpunkte/> [03.07.2012].

**Université de Lausanne (2012a):** *Viaticalpes UNIL: Le Projet FNS*. Verfügbar unter: <http://www.unil.ch/viaticalpes/page53767.html> [21.06.2012].

**Université de Lausanne (2012b):** *Viaticalpes UNIL: Notre réseau*. Verfügbar unter: <http://www.unil.ch/viaticalpes/page53769.html> [21.06.2012].

**Viegener, Tobias (2012):** *Formate im Bibliothekswesen. Unveröffentlichtes Skript*. Chur.

**Wiesenmüller Heidrun (2011):** *Die Zukunft der Katalogisierung: Den Kern erhalten – Qualität an der richtigen Stelle*. In: Ulrich Hohoff und Daniela Lülfig (Hrsg.): *Bibliotheken für die Zukunft – Zukunft für die Bibliotheken: 100. Deutscher Bibliothekartag in Berlin 2011* (S. 327-337). Hildesheim etc.: Georg Olms.

**Wiesenmüller, Heidrun (2012):** *Zwischen Wunsch und Wirklichkeit: Bibliotheksdaten und Bibliothekskataloge*. VDB-Mitteilungen, Heft 1, S. 20-24. Verfügbar unter: <http://www.vdb-online.org/publikationen/vdb-mitteilungen/vdb-mitteilungen.php?jahrgang=2012&heft=1> [23.07.2012].

**Wiesenmüller, Heidrun; Meylein, Leonhard; Pfeffer, Magnus (2011):** *Mehr aus der Schlagwortnormdatei herausholen: Implementierung einer geographischen Facette in den Online-Katalogen der UB Heidelberg und der UB Mannheim*. B.I.Tonline 14 (2011), Nr. 3, S. 245-252. Verfügbar unter: <http://www.ub.uni-heidelberg.de/archiv/12555/> [10.07.2011].

## 7 Anhang



Abbildung 25: Drilldown-Facette aus HEIDI (Wiesenmüller et al. 2011, S. 252)

## Bisher erschienene Schriften

Ergebnisse von Forschungsprojekten erscheinen jeweils in Form von Arbeitsberichten in Reihen.  
Sonstige Publikationen erscheinen in Form von alleinstehenden Schriften.

Derzeit gibt es in den Churer Schriften zur Informationswissenschaft folgende Reihen:  
Reihe Berufsmarktforschung

Churer Schriften zur Informationswissenschaft – Schrift 1  
Herausgegeben von Josef Herget und Sonja Hierl  
Reihe Berufsmarktforschung – Arbeitsbericht 1:  
Josef Herget  
Thomas Seeger  
Zum Stand der Berufsmarktforschung in der Informationswissenschaft  
in deutschsprachigen Ländern  
Chur, 2007 (im Druck)  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 2  
Herausgegeben von Josef Herget und Sonja Hierl  
Reihe Berufsmarktforschung – Arbeitsbericht 2:  
Josef Herget  
Norbert Lang  
Berufsmarktforschung in Archiv, Bibliothek, Dokumentation  
und in der Informationswirtschaft: Methodisches Konzept  
Chur, 2007 (im Druck)  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 3  
Herausgegeben von Josef Herget und Sonja Hierl  
Reihe Berufsmarktforschung – Arbeitsbericht 3:  
Josef Herget  
Norbert Lang  
Gegenwärtige und zukünftige Arbeitsfelder für Informationsspezialisten  
in privatwirtschaftlichen Unternehmen und öffentlich-rechtlichen Institutionen  
Chur, 2004  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 4  
Herausgegeben von Josef Herget und Sonja Hierl  
Sonja Hierl  
Die Eignung des Einsatzes von Topic Maps für e-Learning  
Vorgehensmodell und Konzeption einer e-Learning-Einheit unter Verwendung von Topic Maps  
Chur, 2005  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 5  
Herausgegeben von Josef Herget und Sonja Hierl  
Nina Braschler  
Realisierungsmöglichkeiten einer Zertifizierungsstelle für digitale Zertifikate in der Schweiz  
Chur, 2005  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 6  
Herausgegeben von Josef Herget und Sonja Hierl  
Reihe Berufsmarktforschung – Arbeitsbericht 4:  
Ivo Macek  
Urs Naegeli  
Postgraduiertenausbildung in der Informationswissenschaft in der Schweiz:  
Konzept – Evaluation – Perspektiven  
Chur, 2005  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 7  
Herausgegeben von Josef Herget und Sonja Hierl  
Caroline Ruosch  
Die Fraktale Bibliothek:  
Diskussion und Umsetzung des Konzepts in der deutschsprachigen Schweiz.  
Chur, 2005  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 8  
Herausgegeben von Josef Herget und Sonja Hierl  
Esther Bättig  
Information Literacy an Hochschulen  
Entwicklungen in den USA, in Deutschland und der Schweiz  
Chur, 2005  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 9  
Herausgegeben von Josef Herget und Sonja Hierl  
Franziska Höfliger  
Konzept zur Schaffung einer Integrationsbibliothek in der Pestalozzi-Bibliothek Zürich  
Chur, 2005  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 10  
Herausgegeben von Josef Herget und Sonja Hierl  
Myriam Kamphues  
Geoinformationen der Schweiz im Internet:  
Beurteilung von Benutzeroberflächen und Abfrageoptionen für Endnutzer  
Chur, 2006  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 11  
Herausgegeben von Josef Herget und Sonja Hierl  
Luigi Ciullo  
Stand von Records Management in der chemisch-pharmazeutischen Branche  
Chur, 2006  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 12  
Herausgegeben von Josef Herget und Sonja Hierl  
Martin Braschler, Josef Herget, Joachim Pfister, Peter Schäuble, Markus Steinbach, Jürg Stuker  
Evaluation der Suchfunktion von Schweizer Unternehmens-Websites  
Chur, 2006  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 13  
Herausgegeben von Josef Herget und Sonja Hierl  
Adina Lieske  
Bibliotheksspezifische Marketingstrategien zur Gewinnung von Nutzergruppen:  
Die Winterthurer Bibliotheken  
Chur, 2007  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 14  
Herausgegeben von Josef Herget und Sonja Hierl  
Christina Bieber, Josef Herget  
Stand der Digitalisierung im Museumsbereich in der Schweiz  
Internationale Referenzprojekte und Handlungsempfehlungen  
Chur, 2007  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 15  
Herausgegeben von Josef Herget und Sonja Hierl  
Sabina Löhner  
Kataloganreicherung in Hochschulbibliotheken  
State of the Art Überblick und Aussichten für die Schweiz  
Chur, 2007  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 16  
Herausgegeben von Josef Herget und Sonja Hierl  
Heidi Stieger  
Fachblogs von und für BibliothekarInnen – Nutzen, Tendenzen  
Mit Fokus auf den deutschsprachigen Raum  
Chur, 2007  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 17  
Herausgegeben von Josef Herget und Sonja Hierl  
Nadja Kehl  
Aggregation und visuelle Aufbereitung von Unternehmensstrategien  
mithilfe von Recherche-Codes  
Chur, 2007  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 18  
Herausgegeben von Josef Herget und Sonja Hierl  
Rafaela Pichler  
Annäherung an die Bildsprache – Ontologien als Hilfsmittel für Bilderschliessung  
und Bildrecherche in Kunstbilddatenbanken  
Chur, 2007  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 19  
Herausgegeben von Josef Herget und Sonja Hierl  
Jürgen Büchel  
Identifikation von Marktnischen – Die Eignung verschiedener Informationsquellen  
zur Auffindung von Marktnischen  
Chur, 2007  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 20  
Herausgegeben von Josef Herget und Sonja Hierl  
Andreas Eisenring  
Trends im Bereich der Bibliothekssoftware  
Chur, 2007  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 21  
Herausgegeben von Josef Herget und Sonja Hierl  
Lilian Brändli  
Gesucht – gefunden? Optimierung der Informationssuche von Studierenden  
in wissenschaftlichen Bibliotheken  
Chur, 2007  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 22  
Herausgegeben von Josef Herget und Sonja Hierl  
Beatrice Bürgi  
Open Access an Schweizer Hochschulen – Ein praxisorientierter Massnahmenkatalog für  
Hochschulbibliotheken zur Planung und Errichtung von Institutional Repositories  
Chur, 2007  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 23  
Herausgegeben von Josef Herget und Sonja Hierl  
Darja Dimitrijewitsch, Cécile Schneeberger  
Optimierung der Usability des Webauftritts  
der Stadt- und Universitätsbibliothek Bern  
Chur, 2007  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 24  
Herausgegeben von Nadja Böller, Josef Herget und Sonja Hierl  
Brigitte Brüderlin  
Stakeholder-Beziehungen als Basis einer Angebotsoptimierung  
Chur, 2008  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 25  
Herausgegeben von Robert Barth, Nadja Böller, Sonja Hierl und Hans-Dieter Zimmermann  
Jonas Rebmann  
Web 2.0 im Tourismus, Soziale Webanwendungen im Bereich der Destinationen  
Chur, 2008  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 26  
Herausgegeben von Robert Barth, Nadja Böller, Sonja Hierl und Hans-Dieter Zimmermann  
Isabelle Walther  
Idea Stores, ein erfolgreiches Bibliothekskonzept aus England – auf für die Schweiz?  
Chur, 2008  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 27, im Druck  
Herausgegeben von Robert Barth, Nadja Böller, Sonja Hierl und Hans-Dieter Zimmermann  
Scherer Auberson Kirsten  
Evaluation von Informationskompetenz: Lässt sich ein Informationskompetenzzuwachs messen?  
Eine systematische Evaluation von Messverfahren  
Chur, 2009  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 28  
Herausgegeben von Robert Barth, Nadja Böller, Sonja Hierl und Hans-Dieter Zimmermann  
Nadine Wallaschek  
Datensicherung in Bibliotheksverbänden.  
Empfehlungen für die Entwicklung von Sicherheits- und Datensicherungskonzepten  
in Bibliotheksverbänden  
Chur, 2009  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 29  
Herausgegeben von Robert Barth, Nadja Böller, Sonja Hierl und Hans-Dieter Zimmermann  
Laura Tobler  
Recherchestrategien im Internet  
Systematische Vorgehensweisen bei der Suche im Internet  
dargestellt anhand ausgewählter Fallstudien  
Chur, 2009  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 30  
Herausgegeben von Robert Barth, Nadja Böller, Sonja Hierl und Hans-Dieter Zimmermann  
Bibliotheken und Dokumentationszentren als Unternehmen:  
Antworten von Bibliotheken und Dokumentationszentren  
auf die Herausforderungen der digitalen Gesellschaft  
Chur, 2009  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 31  
Herausgegeben von Robert Barth, Nadja Böller, Sonja Hierl und Hans-Dieter Zimmermann  
Karin Garbely, Marita Kieser  
Mystery Shopping als Bewertungsmethode der Dienstleistungsqualität  
von wissenschaftlichen Bibliotheken  
Chur, 2009  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 32  
Herausgegeben von Robert Barth, Nadja Böller, Sonja Hierl und Hans-Dieter Zimmermann  
Tristan Triponez  
E-Mail Records Management  
Die Aufbewahrung von E-Mails in Schweizer Organisationen als technische,  
rechtliche und organisatorische Herausforderung  
Chur, 2009  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 33  
Herausgegeben von Robert Barth, Nadja Böller, Urs Dahinden, Sonja Hierl  
und Hans-Dieter Zimmermann  
Die Lernende Bibliothek 2009  
Aktuelle Herausforderungen für die Bibliothek und ihre Partner im Prozess  
des wissenschaftlichen Arbeitens  
Chur, 2009  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 34  
Herausgegeben von Robert Barth, Nadja Böller, Sonja Hierl und Hans-Dieter Zimmermann  
Rene Frei  
Die Informationswissenschaft aus Sicht des Radikalen Konstruktivismus  
Chur, 2009  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 35  
Herausgegeben von Robert Barth, Nadja Böller, Sonja Hierl und Hans-Dieter Zimmermann  
Lydia Bauer, Nadja Böller, Sonja Hierl  
DIAMOND Didactical Approach for Multiple Competence Development  
Chur, 2009  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 36  
Herausgegeben von Robert Barth, Nadja Böller, Sonja Hierl und Wolfgang Semar  
Michaela Spiess  
Einsatz von Competitive Intelligence in Schweizer Spitälern  
Chur, 2009  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 37  
Herausgegeben von Robert Barth, Nadja Böller, Sonja Hierl und Wolfgang Semar  
Jasmine Milz  
Informationskompetenz-Vermittlung an Deutschschweizer Fachhochschulen:  
eine quantitative Inhaltsanalyse der Curricula  
Chur, 2010  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 38  
Herausgegeben von Robert Barth, Nadja Böller, Sonja Hierl und Wolfgang Semar  
Corinne Keller  
RFID in Schweizer Bibliotheken – eine Übersicht  
Chur, 2010  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 39  
Herausgegeben von Robert Barth, Nadja Böller, Sonja Hierl und Wolfgang Semar  
Herausgegeben von Robert Barth und Iris Kuppelwieser  
Bibliotheksbau in der Schweiz 1985 – 2010  
Planung – Nutzung – Ästhetik  
Chur, 2010  
ISSN1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 40  
Herausgegeben von Robert Barth, Nadja Böller, Sonja Hierl und Wolfgang Semar  
Stephan Becker  
Klassifikationsraster zur Relevanzanalyse aktueller Themenanfragen  
an einer Mediendokumentationsstelle in der Schweiz  
Chur, 2010  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 41  
Herausgegeben von Robert Barth, Nadja Böller, Sonja Hierl und Wolfgang Semar  
Reihe Berufsmarktforschung – Arbeitsbericht 5:  
Iris Capatt, Urs Dahinden  
Absolventenbefragung 2010  
Bachelorstudiengang Informationswissenschaft und Diplomstudiengang Information und  
Dokumentation der HTW Chur  
Chur, 2010  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 42  
Herausgegeben von Robert Barth, Nadja Böller, Sonja Hierl und Wolfgang Semar  
Saro Adamo Pepe Fischer  
Bestandserhaltung im Film-/Videoarchiv des Schweizer Fernsehens  
Chur, 2010  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 43  
Herausgegeben von Robert Barth, Iris Capatt, Sonja Hierl und Wolfgang Semar  
Patricia Düring  
Ökonomischer Mehrwert von Bibliotheken, aufgezeigt anhand ausgewählter Dienste der Zentral-  
und Hochschulbibliothek Luzern  
Chur, 2011  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 44  
Herausgegeben von Robert Barth, Iris Capatt, Sonja Hierl und Wolfgang Semar  
Pia Baier Benninger  
Model Requirements for the Management of Electronic Records (MoReq2).  
Anleitung zur Umsetzung  
Chur, 2011  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 45  
Herausgegeben von Robert Barth, Iris Capatt, Sonja Hierl und Wolfgang Semar  
Martina Thomi  
Überblick und Bewertung von Musiksuchmaschinen  
Chur, 2011  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 46  
Herausgegeben von Robert Barth, Iris Capatt und Wolfgang Semar  
Regula Trachsler  
Angebote für Senioren in Deutschschweizer Bibliotheken  
Chur, 2011  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 47  
Herausgegeben von Robert Barth, Iris Capatt und Wolfgang Semar  
Wolfgang Semar (Hrsg.)  
Arge Alp Tagung 23.-24. September 2010, Chur  
Informationsgesellschaft und Infrastrukturpolitik im Alpenraum  
Chur, 2011  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 48  
Herausgegeben von Robert Barth, Lydia Bauer, Iris Capatt und Wolfgang Semar  
Heinz Mathys  
Jungs lesen weniger als Mädchen.  
Was können Bibliotheken gemeinsam mit den Schulen tun, um dies zu ändern?  
Chur, 2011  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 49  
Herausgegeben von Robert Barth, Lydia Bauer, Iris Capatt und Wolfgang Semar  
Anina Baumann  
Stärken und Schwächen von Discovery Diensten am Beispiel des EBSCO Discovery Service  
Chur, 2011  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 50  
Herausgegeben von Robert Barth, Lydia Bauer, Iris Capatt und Wolfgang Semar  
Reihe Berufsmarktforschung – Arbeitsbericht 6:  
Iris Capatt, Urs Dahinden  
Absolventenbefragung 2011  
Hochschule für Technik und Wirtschaft HTW Chur Weiterbildungsstudiengänge  
Informationswissenschaft.  
Externer Bericht.  
Chur, 2011  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 51  
Herausgegeben von Robert Barth, Lydia Bauer, Iris Capatt und Wolfgang Semar  
Reihe Berufsmarktforschung – Arbeitsbericht 7:  
Iris Capatt, Urs Dahinden  
Absolventenbefragung 2011  
Hochschule für Technik und Wirtschaft HTW Chur Weiterbildungsstudiengänge Management.  
Externer Bericht.  
Chur, 2011  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 52  
Herausgegeben von Robert Barth, Lydia Bauer, Iris Capatt und Wolfgang Semar  
Salome Arnold  
Auf den Spuren der Barrieren für ein barrierefreies Webdesign  
Chur, 2011  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 53  
Herausgegeben von Robert Barth, Lydia Bauer, Iris Capatt und Wolfgang Semar  
Laura Stadler  
Die Gläserne Decke in Schweizer Bibliotheken  
Chur, 2012  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 54  
Herausgegeben von Robert Barth, Lydia Bauer, Brigitte Lutz und Wolfgang Semar  
Ruth Süess  
Evaluation von Web Monitoring Tools zur softwaregestützten Informationsbeschaffung  
am Beispiel ausgewählter Open Source Web Monitoring Tools  
Chur, 2012  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 55  
Herausgegeben von Robert Barth, Lydia Bauer, Brigitte Lutz und Wolfgang Semar  
Michael Hunziker

Approval Plans und andere Outsourcing-Formen im Bestandesaufbau an den  
Wissenschaftlichen Bibliotheken der Deutschschweiz  
Chur, 2012  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 56  
Herausgegeben von Wolfgang Semar und Brigitte Lutz  
Urs Dahinden, Michael Aschwanden und Lydia Bauer  
Verpasste Chancen? Altersspezifische digitale Ungleichheiten bei der Nutzung von  
Mobilkommunikation und Internet  
Chur, 2012  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 57  
Herausgegeben von Wolfgang Semar und Brigitte Lutz  
Grégoire Savary  
Eine Konservierungsstrategie für das Archiv der Siedlungsgenossenschaft Freidorf bei Muttenz.  
Eine Hilfestellung für kleine Archive mit gemischten Beständen  
Chur, 2013  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 58  
Herausgegeben von Wolfgang Semar und Brigitte Lutz  
Patrick Wermelinger  
Die Georeferenzierung von Katalogdaten mit Hilfe von Linked Open Data  
Chur, 2013  
ISSN 1660-945X

---

## Über die Informationswissenschaft der HTW Chur

Die Informationswissenschaft ist in der Schweiz noch ein junger Lehr- und Forschungsbereich. International weist diese Disziplin aber vor allem im anglo-amerikanischen Bereich eine jahrzehntelange Tradition auf. Die klassischen Bezeichnungen dort sind Information Science, Library Science oder Information Studies. Die Grundfragestellung der Informationswissenschaft liegt in der Betrachtung der Rolle und des Umgangs mit Information in allen ihren Ausprägungen und Medien sowohl in Wirtschaft und Gesellschaft. Die Informationswissenschaft wird in Chur integriert betrachtet.

Diese Sicht umfasst die Teildisziplinen Bibliothekswissenschaft, Archivwissenschaft und Dokumentationswissenschaft. Auch neue Entwicklungen im Bereich Medienwirtschaft und Informationsmanagement werden gezielt aufgegriffen und im Lehr- und Forschungsprogramm berücksichtigt.

Der Studiengang Informationswissenschaft wird seit 1998 als Vollzeitstudiengang in Chur angeboten und seit 2002 als Teilzeit-Studiengang in Zürich. Künftig wird ein berufsbegleitender Masterstudiengang das Lehrangebot abrunden.

Der Arbeitsbereich Informationswissenschaft vereinigt Cluster von Forschungs-, Entwicklungs- und Dienstleistungspotentialen in unterschiedlichen Kompetenzzentren:

- Information Management & Competitive Intelligence
- Records Management
- Library Consulting
- Information Laboratory

Diese Kompetenzzentren werden im **Swiss Institute for Information Research** zusammengefasst.

## IMPRESSUM

### Verlag & Anschrift

#### Arbeitsbereich Informationswissenschaft

HTW - Hochschule für Technik und Wirtschaft  
University of Applied Sciences  
Ringstrasse 37  
CH-7000 Chur  
www.informationswissenschaft.ch  
www.htwchur.ch

**ISSN 1660-945X**

### Institutsleitung

Prof. Dr. Niklaus Stettler  
Telefon: +41 81 286 24 61

Email: [niklaus.stettler@htwchur.ch](mailto:niklaus.stettler@htwchur.ch)

### Sekretariat

Telefon : +41 81 286 24 24  
Fax : +41 81 286 24 00  
Email: [clarita.decurtins@htwchur.ch](mailto:clarita.decurtins@htwchur.ch)

---