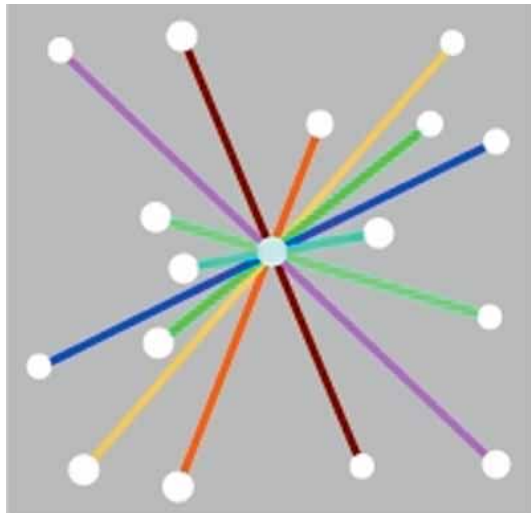


Guido Bee

Concepts in Context

Kölner Tagung zur Interoperabilität in Wissensorganisationssystemen und CrissCross-Abschlussworkshop



Wer Brücken errichtet, tut meist ein verdienstvolles Werk. Ist allerdings das Fundament brüchig, der Verwendungszweck unklar oder mangelt es an Zugängen zur übrigen Infrastruktur, kann sich das Vorhaben als Fehlplanung erweisen. In populären Darstellungen informationswissenschaftlicher Prozesse ist viel vom Brückenbauen die Rede. Und tatsächlich scheint diese Metapher gut geeignet, um sich die für Wissensvermittlung entscheidenden konzeptionellen Fragen vor Augen zu führen. Auch Strukturierung von Informationen ist eine Form des Brückenbaus, die aber nur gelingen kann, wenn die dabei verwendeten Dokumentationssprachen und Wissensorganisationssysteme ausdrucksstark, funktional und interoperabel gestaltet werden, um differenzierte Strategien der Informationssuche und der Erkundung neuer Wissensräume effizient zu unterstützen. Insbesondere der Interoperabilität bzw. Kompatibilität der einzelnen Systeme wird mehr und mehr eine entscheidende Rolle bei der Optimierung von Informationsprozessen zugeschrieben.

Interoperabilität dient der Optimierung von Informationsprozessen

Die Bedeutung der Interoperabilität von Wissensorganisationssystemen stand auch im Zentrum von »Concepts in Context«, einer gemeinsamen internationalen Fachtagung des Instituts für Informationsmanagement der Fachhochschule Köln und der Deutschen Nationalbibliothek (DNB), die vom 19. bis 20. Juli in Köln stattfand.¹⁾ Zu Beginn wies Klaus-Dirk Schmitz, Leiter des Instituts für Informationsmanagement, auf die vielfältigen Forschungsaktivitäten der Fachhochschule Köln und der ihr angeschlossenen Institute hin und beschrieb so ein Umfeld, in dem der Interoperabilität der einzelnen Bereiche gesteigerte Bedeutung zukommt. Ulrike Junger, Leiterin der Abteilung Inhaltserschließung an der DNB, stellte dann in ihrer Begrüßungsansprache die langjährige Zusammenarbeit heraus, die das Institut für Informationsmanagement und die DNB miteinander verbindet. Ein wichtiges Resultat war die von 2002 bis 2005 als gemeinsames Projekt durchgeführte Übersetzung der 22. Auflage der Dewey-Dezimalklassifikation (DDC) in die deutsche Sprache, die dem Einsatz der DDC in der Deutschen Nationalbibliografie vorausging.²⁾ Als Folgeprojekt schloss sich im Jahr 2006 das von der Deutschen Forschungsgemeinschaft (DFG) geförderte Projekt CrissCross an, das im Herbst dieses Jahres erfolgreich beendet wurde.³⁾ Ein Teil der Tagung fungierte daher als Abschlussworkshop, in dem die Ergebnisse des Projekts und die daraus resultierenden konkreten Einsatzmöglichkeiten vorgestellt wurden. Hauptziel von CrissCross war es, den Zugang zu heterogen erschlossenen Dokumenten für die Nutzer von Bibliotheken zu verbessern und an die heute gängigen Suchmethoden anzupassen. Dies wurde konkret dadurch ermöglicht, dass die Schlagwörter der Schlagwortnormdatei (SWD) mit der Dewey-Dezimalklassifikation Deutsch (DDC) als der weltweit meist verbreiteten Klassifikation verknüpft wurden. Das Resultat ist ein thesaurusbasiertes, benutzergerechtes Recherchevokabular, das die thematische Suche in heterogen erschlossenen Beständen unterstützt, den Anforderungen an

Rückblick auf andere Projekte

Projektziel – Zugang zu heterogen erschlossenen Datenbeständen verbessern

moderne Suchmethoden entspricht und darüber hinaus für weitere Forschungszwecke, insbesondere im Kontext des Semantic Web, zur Verfügung steht. Auf der Kölner Tagung gaben Jan-Helge Jacobs, Tina Mengel und Katrin Müller, die zusammen mit Teamleiterin Jessica Hubrich das Projektteam bildeten, zunächst eine Darstellung der konkreten Arbeitsweise und methodischen Fundierung der CrissCross-Verlinkungsarbeit. Ein wesentlicher Grundsatz des hier vollzogenen Mappings ist eine One-to-many-Strategie. Das impliziert, dass ein SWD-Begriff in den meisten Fällen mit mehr als einer DDC-Klasse verbunden wird. Es wird also keine eindeutige Zuordnung vorgenommen, sondern vorausgesetzt, dass SWD-Terme mit verschiedenen DDC-Klassen korrelieren können. Eine weitere Festlegung betrifft das so genannte Deep-Level-Mapping, das die Zuordnung eines SWD-Terms zur entsprechenden DDC-Klasse auf der tiefst möglichen Ebene vorsieht. Das entspricht einer feinklassifikatorischen Erschließung und macht – in vielen Fällen – die Bildung synthetischer Notationen erforderlich. Ein wesentlicher Teil der Arbeit ist schließlich die Einbeziehung eines Ranking-Systems, mit dem die Präzision der Mappings bewertet wird: Jede SWD/DDC-Verbindung wird durch einen von 1 bis 4 reichenden Determiniertheitsgrad qualifiziert, der Aufschluss über die Passgenauigkeit der jeweiligen Verlinkung gibt und für das Retrieval entsprechend ausgewertet werden kann. Die im Rahmen von CrissCross entwickelten methodischen Konzepte werden nach dem Ende des Projekts von den Fachreferenten der DNB übernommen und weitergeführt. In deren Händen liegt auch die Verlinkung der neuen SWD-Begriffe.

Über die konkrete Verlinkungspraxis hinaus wurden vom Kölner Projektteam auch Konzepte zur Integration der CrissCross-Mappings in Retrieval-szenarien entwickelt. So konnte gezeigt werden, dass die CrissCross-Daten auf vielfältige Weise in unterschiedliche Suchstrategien integriert werden können, indem sie zum einen die Einbeziehung lediglich klassifikatorisch erschlossener Dokumente in eine verbale Suche ermöglichen, darüber hinaus aber auch zu einer differenzierten Strukturierung von Treffersets beitragen können, wobei die Angabe der Determiniertheit als Steuerungsinstru-

ment fungiert. Weitere Perspektiven ergeben sich aus einem von Jessica Hubrich vorgestellten Stufenmodell semantischer Interoperabilität, in dem bestimmte Typen von Mappingrelationen mit spezifischen Retrievalfunktionalitäten in Beziehung gesetzt werden. Der Ansatz von Hubrich zielt insbesondere auf den Nachweis ab, dass die CrissCross-Daten zu einer Differenzierung des Relationsgefüges der SWD beitragen und so die semantische Struktur der SWD nachhaltig verbessern können.

Neben der Verknüpfung von SWD-Termen mit der DDC und der dadurch realisierten Verbindung von klassifikatorischer und verbaler Erschließung hatte CrissCross noch einen weiteren, multilingual ausgerichteten Arbeitsschwerpunkt, der in Kooperation mit der Schweizerischen Nationalbibliothek (NB) realisiert wurde und die Arbeiten des europäischen Projekts Multilingual access to subjects (MACS) fortführt.⁴⁾ Helga Karg, die an der DNB das CrissCross-Projekt geleitet hat, gab zusammen mit ihrer Kollegin Yvonne Jahns in ihrem Vortrag einen Überblick über die Durchführung dieses Projektteils und die mit ihm verbundenen Anwendungsszenarien. Im Rahmen der CrissCross-MACS-Kooperation wurden über 50.000 Links zwischen der deutschen SWD, den englischsprachigen Library of Congress Subject Headings (LCSH) und der französischen Schlagwortsprache Rameau (Répertoire d'autorité-matière encyclopédique et alphabétique unifié) erstellt. In den meisten Fällen erfolgte dies dadurch, dass ein bereits vorhandenes LCSH-Rameau-Tupel durch Hinzufügung eines SWD-Äquivalents zu einem Tripel ergänzt wurde. Der Hauptanteil der Arbeit wurde in der DNB Leipzig unter wesentlicher Beteiligung professioneller Übersetzerteams realisiert. Die Erstellung der Verknüpfungen wurde im Link Management Interface (LMI) der MACS-Datenbank durchgeführt. Die im Rahmen von MACS erstellten Daten werden auf vielfältige Weise Eingang in die Bibliothekswelt finden: Neben der bereits im Mai umgesetzten Veröffentlichung der MACS-Tripel als Linked Data durch die DNB ist eine Integration der Daten in das Portal von The European Library (TEL) sowie eine Nutzung innerhalb von Europeana geplant bzw. bereits prototypisch realisiert.⁵⁾ Das nun vorhandene Datenmaterial versetzt Bibliotheksnutzer künftig verstärkt in die Lage, mit der eigenen

CrissCross-Daten verbessern die semantische Struktur der SWD

Multilingualer Arbeitsschwerpunkt

Übernahme neuer methodischer Konzepte nach Projektende

Schlagwortsprache Medien zu finden, die mit fremden Dokumentations-sprachen erschlossen wurden. Im Zusammenhang mit der Veröffentlichung der CrissCross- und MACS-Daten als Linked Data ging Antoine Isaac, Freie Universität Amsterdam, in seinem Beitrag auf diese immer weiter verbreitete Form der Datenveröffentlichung näher ein und diskutierte insbesondere den Nutzen von SKOS (Simple Knowledge Organization System) zur Darstellung und Vernetzung von Wissensorganisationssystemen. Er hob die Vorteile der mit SKOS gegebenen Ansätze zur Standardisierung von Mappingrelationen hervor und trat für eine möglichst einfache, vielen Anwendern geläufige Strukturierung des Linked-Data-Materials ein.

In den Kontext von CrissCross gehört auch ein von Claudia Effenberger und Julia Hauser vorgestelltes, innerhalb der DNB durchgeführtes Vorhaben, das im Kern jedoch nicht auf Aspekte des Mappings, sondern auf die Entwicklung der DDC bezogen ist. Ausgangspunkt ist die Tatsache, dass die DDC – wie jede Klassifikation – kein statisches System darstellt, sondern laufend aktualisiert wird. Dies hat – vor dem Hintergrund, dass eine Nachklassifikation aus arbeitsökonomischen Gründen nicht geleistet werden kann und dies auch künftig für die weltweit klassifizierten internationalen Datenbestände nicht zu erwarten ist – u. a. zur Folge, dass bestimmte an Dokumenten vergebene Notationen im Lauf der Zeit nicht mehr mit dem aktuellen Stand der DDC übereinstimmen. Eine Lösung des Problems könnte darin bestehen, jeder Version einer DDC-Notation einen eindeutigen Identifier zuzuweisen, um damit die tatsächlich verwendete Bedeutung einer DDC-Notation an einem Dokument identifizieren zu können. Sofern die Identifier der generischen URI-Syntax entsprechen, ist es möglich, die Beziehungen zwischen den DDC-Klassen, den Titeln und den Schlagwörtern der SWD mittels RDF (Resource Description Framework) als ein semantisches Netzwerk zu modellieren und nachfolgend zu untersuchen, ob hiermit das Retrieval verbessert werden kann.

Die genannten Vorträge machen deutlich, dass vom CrissCross-Projekt wichtige Impulse für den Umgang mit heterogen erschlossenen Dokumenten ausgehen. Das aber ist nur die eine Seite der Medaille. Winfried Gödert, in dessen Händen die

Projektleitung auf Kölner Seite lag, hob bereits zu Beginn der Kölner Tagung nachdrücklich hervor, dass das CrissCross-Projekt nicht nur Verbindungslinien geschaffen hat, sondern selbst Teil eines umfassenden Verbindungsnetzes und in vielfacher Weise mit anderen Projekten verbunden ist – MACS ist hierfür nur ein Beispiel. Im zweiten Teil der Veranstaltung standen daher andere Projekte im Vordergrund, die sich ähnlich wie CrissCross mit komplexen Fragestellungen bezüglich der Interoperabilitäts-Problematik auseinandersetzen.

Dagobert Soergel wies in seinem Eröffnungsvortrag anschaulich auf das Grundproblem der Vielzahl konkurrierender Ontologien und Wissensorganisationssysteme hin, um diese Vielzahl dann gleich noch um ein eigenes Modell, das so genannte »Mapping through a Hub«, zu bereichern, das allerdings als übergeordnetes Steuerungssystem begriffen werden soll. Das Hub-Mapping orientiert sich an der aus dem Transportwesen bekannten Hub-and-Spoke-Architektur, einer sternförmigen Vernetzung aus einem Zentralknoten (Hub) und Endstellen, die über die Spokes (Speichen) miteinander verbunden sind. Die Hubs sollen dabei ein regelbasiertes Backbone-System repräsentieren, das mehrere Wissensorganisationssysteme verbindet und daraus eine allumfassende Wissensbasis herstellt.

Die von Soergel stark akzentuierte Notwendigkeit allgemein akzeptierter Standards für die Verknüpfungen verschiedener Vokabulare zog sich als Leitfaden durch den zweiten Teil der Tagung. So wies Stella Dextre Clarke – wie zuvor schon Isaac – auf die Bedeutung von SKOS für die Standardisierung der Wissensorganisation hin. Sie bedauerte, dass bei den meisten Mapping-Projekten eigene Wege der Festlegung und Umsetzung von Mappingtypen entwickelt wurden, die sich von den SKOS-Vorgaben unterscheiden, was die Interoperabilität der verschiedenen Projekte beeinträchtigt. Es liegt daher nahe, Standards für Mappingtypen zu entwickeln, um auf diese Weise alle weit verbreiteten Thesauri, Klassifikationssysteme und Schlagwortsysteme miteinander zu verknüpfen und universelle Interoperabilität zu ermöglichen. Ein solcher Standard sollte aber gewisse Vorgaben beachten: Er muss einfach sein, tatsächlichen Benutzerinteressen entsprechen, auf Regeln basieren, die von den Nutzern akzeptiert werden, mit gängiger Software kom-

patibel sein und kostenlos oder zumindest zu einem geringen Preis vertrieben werden. Das von Dextre Clarke geleitete Projekt ISO NP 25964 zur Aktualisierung und Revision internationaler Thesaurusstandards wird hierzu Anfang 2011 einen Entwurf vorlegen.⁶⁾

Die Schaffung kontrollierter Vokabulare und deren Standardisierung sind verdienstvoll, dürfen aber kein Selbstzweck sein. Philipp Mayr lenkte in seinem Beitrag daher den Blick auf das Retrieval als eigentliche Zielsetzung der Verlinkungsaktivitäten. Er konstatierte, dass zwar eine Vielzahl von Mapping-Projekten existiert, aber nur wenige dieser Projekte in Systeme implementiert sind, die tatsächlich genutzt werden können. Zudem mangelte es lange an differenzierten Retrievaltests. Auf dieses Desiderat reagierte das GESIS-Projekt (German Social Science Infrastructure Services) KOMOHE (Competence Center Modeling and Treatment of Semantic Heterogenity), in dem kontrollierte Vokabulare über Crosskonkordanzen verbunden und in einer abschließenden Retrieval-Evaluation in Bezug

auf ihre Effektivität in einer wirklichen Suchumgebung untersucht wurden.⁷⁾ Dabei zeigte sich, dass der Recall durch den Einsatz von Crosskonkordanzen deutlich verbessert werden kann. Eine Weiterführung der Resultate von KOMOHE stellt das Projekt IRM dar, das die Entwicklung und Erprobung von metadatenbasierten Mehrwertdiensten für Retrievalumgebungen mit mehreren Datenbanken zum Ziel hat.⁸⁾

Ein Arbeitsbereich, in dem der Lösung von Interoperabilitätsproblemen gesteigerte Bedeutung zukommt, ist die gemeinsame Modellierung der von der IFLA entwickelten Functional Requirements for Bibliographic Records (FRBR),⁹⁾ der Functional Requirements for Authority Data (FRAD)¹⁰⁾ und der Functional Requirements for Subject Authority Data (FRSAD)¹¹⁾. Obwohl die FRBR Review Group zu gegebener Zeit eine konsolidierte Version aller drei Modelle ausarbeiten soll, werden die RDF-Versionen der FRBR und der FRAD in unterschiedlichen Namensräumen erstellt, die durch eine eigenständige, auf der Web

Retrieval als Zielsetzung von Verlinkungsaktivitäten



BIS-C 2000
Archiv- und Bibliotheks-Informationssystem



DABIS.eu
Gesellschaft für Datenbank-Informationssysteme

DABIS.com - alle Aufgaben - ein Team

Synergien: Qualität und Kompetenz
Software: Innovation und Optimierung
Web - SSL - Warenkorb und Benutzeraccount
Web 2.0 und Catalogue enrichment
Lokalsystem zu Aleph-Verbänden

Archiv	Bibliothek
singleUser	multiUser
Lokalsystem	Verbund
multiDatenbank	multiServer
multiProcessing	multiThreading
skalierbar	stufenlos
Unicode	multiLingual
Normdaten	redundanzfrei
multiMedia	Integration

Software - State of the art - Open Source

Über 20 Jahre Erfahrung und Wissen	Sicherheit
Leistung	Offenheit
Standards	Verlässlichkeit
Stabilität	Adaptierung
Generierung	Erfahrenheit
Service	Support
Outsourcing	Zufriedenheit
Dienstleistungen	
GUI - Web - Wap - XML - Z 39.50	

Portale

<http://landesbibliotheken.eu>
<http://VThK.eu> <http://bmlf.at>
<http://VolksLiedWerk.org> <http://bmwfj.at>
<http://behoerdenweb.net> <http://wkweb.at>

DABIS.com
 Heiligenstädter Straße 213, 1190 - Wien, Austria
 Tel. +43-1-318 9777-10 * Fax +43-1-318 9777-15
 eMail: office@dabis.com * <http://www.dabis.com>

Zweigstellen: 61350 - Bad Homburg vdH, Germany / 1147 - Budapest, Hungary / 39042 - Brixen, Italy

Ihr Partner für Archiv-, Bibliotheks- und DokumentationsSysteme

Ontology Language (OWL) basierten Ontologie miteinander verbunden werden. Mit den daraus resultierenden Problemen setzten sich gleich drei Tagungsbeiträge auseinander. Gordon Dunsire unternahm eine detaillierte Darstellung des ontologischen Konzepts mit den dazugehörigen Definitionen von Klassen und Eigenschaften und thematisierte zudem Fragestellungen, die sich aus der RDF-Repräsentation der RDA-Metadatenelemente ergeben. Maja Žumer gab einen Überblick über die Entwicklung des FRSAD-Modells und skizzierte dessen Auswirkungen auf die künftige Ausgestaltung von Inhaltserschließungsdaten. Felix Boteram stellte ein auf der Basis von Ergebnissen des RESEDA-Projekts (Repräsentationsmodelle für semanti-

sche Daten) der FH Köln entwickeltes Modell vor, das sich mit der Interoperabilität der FRSAD auseinandersetzt.¹²⁾

Das Ende der gleichermaßen anspruchsvollen wie anregenden Tagung machte deutlich, dass mit der Standardisierung der Wissensorganisationssysteme große Vorteile verbunden sind, die Realisierung dieses Ziels jedoch noch viel Zeit in Anspruch nehmen wird. Eine weitere große Herausforderung liegt darin, die Verknüpfungen zwischen den Systemen in konkrete Suchsysteme einzubauen und so in stärkerem Maße für Nutzer fruchtbar zu machen. CrissCross hat hier eine Diskussion beflügelt, die auch nach dem Ende des Projekts weiter anhalten wird.

Ausblick

Anmerkungen

1 Nähere Informationen zu der Tagung inkl. Abstracts und Präsentationen finden Sie unter: <<http://linux2.fbi.fh-koeln.de/cisko2010>>

2 Informationen zum Projekt DDC Deutsch und zum Einsatz der DDC in der DNB finden Sie unter: <<http://www.ddc-deutsch.de>>

3 Projekthomepage: <<http://linux2.fbi.fh-koeln.de/crisscross>>

4 Projekthomepage von MACS: <<https://macs.hoppie.nl/pub>>

5 vgl. zu TEL: <<http://search.theeuropeanlibrary.org/portal/en>> zu Europeana: <<http://www.europeana.eu/portal>>

6 vgl. <<http://www.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2007/papers/dextre.pdf>>

7 KOMOHE: <<http://www.gesis.org/en/research/programs-and-projects/knowledge-technologies/project-overview/komohe>>

8 IRM: <<http://www.ib.hu-berlin.de/~mayr/arbeiten/IRM-ISK09.pdf>>

9 FRBR: <<http://www.ifla.org/en/publications/functional-requirements-for-bibliographic-records>>

10 Zu FRAD vgl. <<http://archive.ifla.org/VII/d4/wg-franar.htm>>.

11 Zu FRSAD vgl. <<http://nkos.slis.kent.edu/FRSAR/report090623.pdf>>

12 RESEDA: <<http://linux2.fbi.fh-koeln.de/reseda>>