

Carstens, Carola; Rittberger, Marc

## Ontologiebasierter Forschungsführer für die Bildungsforschung

formal überarbeitete Version der Originalveröffentlichung in:

formally revised edition of the original source in:

Hinneburg, Alexander [Hrsg.]: Lernen, Wissen, Adaption. Halle, Saale : Martin Luther Univ. Halle-Wittenberg 2007, S. 258-262



Bitte verwenden Sie in der Quellenangabe folgende URN oder DOI /

Please use the following URN or DOI for reference:

urn:nbn:de:0111-dipf-37995

10.25657/02:3799

<https://nbn-resolving.org/urn:nbn:de:0111-dipf-37995>

<https://doi.org/10.25657/02:3799>

### Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

### Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

### Kontakt / Contact:

DIPF | Leibniz-Institut für  
Bildungsforschung und Bildungsinformation  
Frankfurter Forschungsbibliothek  
publikationen@dipf.de  
www.dipfdocs.de

# Ontologiebasierter Forschungsführer für die Bildungsforschung

Carola Carstens, Marc Rittberger

Deutsches Institut für Internationale Pädagogische Forschung  
Informationszentrum Bildung  
D-60486 Frankfurt am Main  
carstens@dipf.de; rittberger@dipf.de

## Abstract

Dieser Beitrag beschreibt das Vorhaben, Semantic Web-Technologien für den Aufbau eines ontologiebasierten Forschungsführers einzusetzen. Heterogene, verteilte Datenquellen des Informationszentrums Bildung sollen auf diese Weise semantisch integriert, angereichert und über entsprechende Recherchemöglichkeiten zugänglich gemacht werden. Die einzelnen Schritte zur Umsetzung des geplanten Projekts werden vorgestellt, wobei insbesondere auf zu erwartende Mehrwerte gegenüber der aktuellen Datenquellennutzung eingegangen wird.

## 1 Einleitung

Semantic Web-Technologien werden vorzugsweise entwickelt, um die Maschinenlesbarkeit und die Interoperabilität von im Internet verteilten Informationen zu gewährleisten und darauf basierend Anbieter übergreifende semantische Suchfunktionalitäten zu entwickeln. Darüber hinaus können sie allerdings auch eingesetzt werden, um die Daten eines einzelnen Informationsanbieters semantisch zu integrieren, anzureichern und über ontologiebasierte Retrievalmethoden recherchierbar zu machen. Dies wird anhand eines konkreten Anwendungsvorhabens im vorliegenden Beitrag skizziert.

Die Konzeption eines ontologiebasierten Forschungsführeres soll verdeutlichen, dass durch den Einsatz von Semantic Web-Technologien auf der Datenbasis eines einzelnen Informationsanbieters bereits Mehrwerte für den Prozess der Informationssuche generiert werden können.

## 2 Projektbeschreibung

Das Projektziel besteht darin, einen ontologiebasierten Forschungsführer für den Bereich Bildungsforschung zu konzipieren, der Informationen über Forscher und Forschungsinstitutionen dieses Fachgebiets sowie ihre jeweiligen Forschungsbereiche auf der Basis semantischer Technologien in einem Portal präsentiert und recherchierbar macht.

Die für den Aufbau des Forschungsführers zur Verfügung stehenden Ausgangsdaten des Informationszentrums Bildung (IZB) liegen derzeit in verteilten heterogenen Datenquellen vor. Sie sollen mithilfe von Semantic Web-Technologien semantisch integriert und angereichert werden, sodass ein Mehrwert

gegenüber der aktuellen Nutzung der Datenquellen entsteht.

Die Anreicherung des Wissens erfolgt durch Inferenzregeln, die durch logische Schlussfolgerungen zu neuartigen Verknüpfungen des vorhandenen Wissens führen. Die auf diese Weise erweiterte Wissensbasis soll den Nutzern über eine Rechercheschnittstelle zugänglich gemacht werden. Die ontologiebasierte Datenintegration ermöglicht es, über Anfragen an die Ontologie Datenquellen übergreifende konzeptbasierte Suchen durchzuführen.

Den Einsatz derartiger semantischer Verfahren für die Prozesse der Datenintegration und des Retrievals gilt es abschließend zu evaluieren.

## 3 State of the Art

Der Vision des Semantic Web zufolge sollen Informationen im Internet in Zukunft derart semantisch ausgezeichnet sein, dass Agenten diese Informationen Anbieter übergreifend aggregieren und interpretieren können [Berners-Lee *et al.* 2001]. Eine wichtige Voraussetzung für die Realisierung dieser Vision stellen Ontologien dar, welche zur semantischen Annotation genutzt werden und den Anwendungen die Interpretation der semantischen Zusammenhänge zwischen den einzelnen Auszeichnungselementen erst ermöglichen.

Da es nicht realistisch ist, eine allumfassende Ontologie des gesamten Weltwissens zu modellieren, widmen sich verschiedene Semantic Web-Initiativen der Aufgabe, Standardontologien für abgegrenzte Anwendungsgebiete zu definieren. Im Zuge dieser Bestrebungen sind unter anderem die folgenden Ontologien entwickelt worden, welche für die Annotation der Inhalte des geplanten Forschungsführers relevant erscheinen: FOAF<sup>1</sup> („Friend-of-a-friend“) für die Darstellung von Personenprofilen und -netzwerken, SKOS<sup>2</sup> („Simple Knowledge Organisation System“) für die Abbildung von Thesaurusrelationen und SWRC („Semantic Web for Research Communities“) für die Modellierung von Forschungsgemeinschaften [Sure *et al.* 2005].

Semantic Web-Anwendungen sehen sich derzeit jedoch noch häufig mit dem Problem der mangelnden Verfügbarkeit von derart ausgezeichneten Daten im Internet konfrontiert. Sie müssen sich daher entweder auf die verfügbaren semantisch annotierten Daten beschränken, oder es wird der Ansatz verfolgt, herkömmliche Daten aus Webseiten zu extrahieren und in

<sup>1</sup> <http://www.foaf-project.org/>

<sup>2</sup> <http://www.w3.org/2004/02/skos/>

einen Semantic Web-Standard zu überführen, um sie für diese Anwendungen nutzbar zu machen.

Im Projekt DBpedia<sup>3</sup> wird beispielsweise das Ziel verfolgt, Informationen aus der Online-Enzyklopädie Wikipedia für das Semantic Web aufzubereiten. Zu diesem Zweck werden strukturierte Informationen, z.B. aus Wikipedia-Infoboxen, extrahiert und in RDF überführt. Auf der Grundlage dieser Datenbasis können über Anwendungen wie den DBpedia Relationship Finder<sup>4</sup> die semantischen Verknüpfungen zwischen Objekten der Datenbasis abgefragt werden [Auer und Lehmann 2007].

Im „Personal Publication Reader“-Projekt werden Informationen über Forscher und Publikationen mithilfe von Wrappern aus Webseiten extrahiert und in den RDF-Standard überführt. Über eine Ontologie werden diese Daten miteinander in Beziehung gesetzt, um regelbasiert personalisierte Sichten auf die ursprünglich verteilten Datenbestände generieren zu können [Baumgartner *et al.* 2005].

Das Projekt FLINK hingegen bezieht die Daten für die Darstellung von Forschernetzwerken in einem semantisch basierten Portal unter anderem aus verfügbaren FOAF-Profilen und mithilfe von Web Mining-Methoden [Mika 2005].

Das Projekt Ontoframe verwendet im Gegensatz dazu projektinterne Ressourcen für den Aufbau einer semantischen Informations- und Kollaborationsplattform. Zur Instanziierung der dem Portal zu Grunde liegenden Ontologie werden unter anderem die Datensätze einer Publikationsdatenbank in RDF-Tripel transformiert [Jung *et al.* 2007].

Dass sich die in den beschriebenen Projekten angewandten Methoden der auf Semantic Web-Standards basierenden Informationsintegration, -anreicherung und -präsentation auch Mehrwert bringend für die Zusammenführung interner Datenquellen einsetzen lassen, soll im geplanten Projekt gezeigt werden.

Die Informationsintegration verfolgt hierbei das Ziel strukturelle und semantische Heterogenitäten zwischen den Ursprungsdatenquellen aufzulösen. Entitäten, die in den einzelnen Datenquellschemata unterschiedlich modelliert sind, sollen über die Ontologie zusammengeführt und semantisch kontextualisiert werden. Dass die Heterogenitätsbehandlung bei der Integration von Informationssystemen eine große Herausforderung darstellt, unterstreichen auch die Arbeiten von Baerisch und Stempfhuber, die Methoden zum Umgang mit semantischen und strukturellen Heterogenitäten bei der Implementierung von Datenquellen übergreifenden Rechercheschnittstellen beschreiben [Baerisch 2007; Stempfhuber 2003].

Im geplanten Projekt wird die Auflösung derartiger Heterogenitäten durch die ontologiebasierte Datenzusammenführung angestrebt.

## 4 Vorgehensweise

In diesem Abschnitt skizzieren wir die geplanten Schritte für die Realisierung des Forschungsführers. Hierzu zählen die Auswahl der Datenbasis, die Entwicklung eines Ontologieschemas für die semantische Datenintegration, die Entscheidung für eine Methode der

Datenquelleneinbindung, die Definition von Inferenzregeln für das Reasoning über die Ontologie, sowie die Konzeption einer Benutzerschnittstelle.

### 4.1 Auswahl der Datenquellen

Für den geplanten Forschungsführer sind alle verfügbaren Informationen über Forscher, Institutionen sowie Forschungsinhalte aus dem Bereich der Bildungsforschung von Interesse. An dieser Stelle wird daher ein kurzer Überblick darüber gegeben, welche Datenquellen des Informationszentrums Bildung relevante Daten enthalten, und wie sie derzeit genutzt werden.

#### FIS Bildung Literaturdatenbank

Die FIS Bildung Literaturdatenbank<sup>5</sup> enthält momentan über 600.000 Literaturnachweise zu pädagogischen und erziehungswissenschaftlichen Publikationen und ist über eine Rechercheschnittstelle des Fachportals Pädagogik<sup>6</sup> zugänglich [Bambey und Jornitz 2006].

Alle Literaturnachweise enthalten neben bibliografischen Angaben auch Schlagworte, die intellektuell nach vorgegebenen Standards vergeben werden. Die auf diese Weise entstehende Beziehung zwischen Autoren, ihren Publikationen sowie den vergebenen Schlagworten stellt eine wichtige Ressource für den Forschungsführer dar.

#### FIS Bildung Wörterbuch

Das Informationszentrum Bildung hat für die eigene Indexierungsarbeit ein Wörterbuch mit thesaurusähnlichen Strukturen wie Hyperonym-, Hyponym- und Synonymrelationen für den Bereich Bildung entwickelt. Darüber hinaus ist jedem Term eine Nummer zugewiesen, die ihn einer Kategorie der vom IZB genutzten Pädagogik-Systematik zuordnet. Das Wörterbuch ist nur intern zugänglich, wird in einer relationalen Datenbank verwaltet und für die intellektuelle Indexierung von Dokumenten der FIS Bildung Literaturdatenbank eingesetzt.

#### Institutionen- und Personendatenbank

Über das Fachportal Pädagogik ist zudem der Zugriff auf eine Institutionen- und Personendatenbank möglich<sup>7</sup>. Sie gibt Auskunft über die Forschungsgebiete, Adressen und Mitarbeiter von Institutionen und Personen aus dem Bereich der Bildungsforschung. Momentan ist der Zugriff auf diese Datenbank über ein hierarchisches Browsing sowie über eine Volltextsuche möglich.

Portalnutzer haben die Möglichkeit, ihr Profil formularbasiert in der Personendatenbank zu hinterlegen. Auf die gleiche Weise können sich auch Institutionen präsentieren.

### 4.2 Erstellung eines Ontologieschemas

Zur semantischen und strukturellen Integration dieser Daten muss ein Ontologieschema erstellt werden, das die Konzeptwelt des Bereichs Bildungsforschung abbilden kann. Es soll die oben aufgeführten Daten in der Form von Konzepten abbilden und miteinander über semantische Relationen in Beziehung setzen.

<sup>5</sup> [http://www.fachportal-paedagogik.de/fis\\_bildung/index.html](http://www.fachportal-paedagogik.de/fis_bildung/index.html)

<sup>6</sup> <http://www.fachportal-paedagogik.de/>

<sup>7</sup> <http://www.fachportal-paedagogik.de/branchenverzeichnis/index.html>

Da im Projekt der Anspruch verfolgt wird, semantisch ausdrucksstarke Relationen zu definieren, kann dieser Modellierungsprozess nur intellektuell erfolgen. Außerdem muss berücksichtigt werden, dass die Güte der späteren Anwendung maßgeblich von der Gestaltung des Ontologieschemas abhängt.

Bei der Definition dieses Schemas wird auf bereits bestehende Ontologien zurückgegriffen, um im Sinne der Semantic Web-Vision die spätere Interoperabilität des Systems mit externen Anwendungen zu unterstützen. Wie aus der folgenden Grafik (Abb.1) hervorgeht, wird das Vokabular der SWRC-Ontologie übernommen und um eigene Ausdrücke mit dem Präfix *vocab* erweitert. Darüber hinaus sollen auch die in Abschnitt 3 angesprochenen Standards FOAF und SKOS Anwendung finden.

Diese Grafik bildet einen Ausschnitt des auf diese Weise erstellten Ontologieschemas ab:

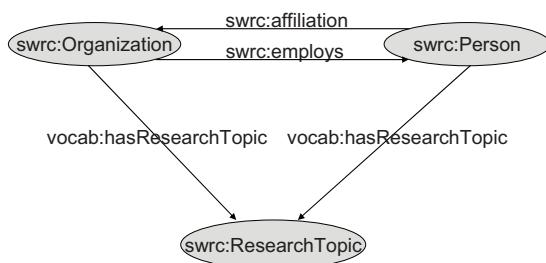


Abb.1: Ausschnitt aus dem Ontologieschema

In der Abbildung sind die Konzepte *swrc:Organization*, *swrc:Person* und *swrc:ResearchTopic* dargestellt, die über semantische Relationen miteinander verbunden sind.

Dass die Ontologie komplexe Modellierungsmöglichkeiten für die semantische Kontextualisierung der Daten bietet, veranschaulicht die folgende Grafik (Abb.2) beispielhaft anhand der Modellierung der Relation *swrc:email*. So lassen sich nicht nur Klassen hierarchisch organisieren, sondern auch Relationen.



Abb.2: Modellierung der Relation *swrc:email*

Aus der Abbildung wird ersichtlich, dass die aus der SWRC-Ontologie übernommene Relation *swrc:email* für die eigenen Zwecke differenzierter modelliert werden musste. Dies führte zur Erweiterung der Relation *swrc:email* um die Unterrelationen *vocab:privateEmail* und *vocab:businessEmail*.

### 4.3 Einbindung der Datenquellen

Die Herstellung der Beziehungen zwischen dem Ontologieschema und den in den Datenquellen enthaltenen Daten erfolgt über einen Datenimport.

Alternativ wäre es auch möglich gewesen, die Daten in den Ausgangsdatenquellen zu belassen und über das ontologische Schema lediglich miteinander in Beziehung zu setzen. Mithilfe der Sprache D2RQ können beispielsweise relationale Datenschemata auf ein Ontologieschema gemappt werden [Bizer und Seaborne 2004]. Dadurch wird es möglich, RDF-basierte SPARQL-

Abfragen an die Datenbanken abzusenden. Auf der Basis der in D2RQ definierten Mappings können diese Abfragen in SQL-Abfragen an die einzelnen Datenquellen transformiert werden, wobei das Ergebnis wiederum als RDF-Tripel zurückgeliefert wird.

Ein wichtiges Argument für den direkten Datenimport war jedoch die Tatsache, dass die prototypische Forschungsführer-Anwendung nicht in den laufenden Betrieb der Ursprungsdatenquellen eingreifen soll. Darüber hinaus bietet der skriptbasierte Datenimport hohe Flexibilität bei der Modellierung der ontologischen Wissensbasis. Als nachteilig stellen sich hingegen die redundante Datenhaltung und der Aufwand für eine regelmäßige Aktualisierung der Daten dar.

Die für den Datenimport erstellten Importskripte senden Abfragen an die Ursprungsdatenquellen und transformieren die zurückgelieferten Daten in Klassen, Instanzen und Relationen der Ontologie. Auf diese Weise wurden beispielsweise Daten aus der Personendatenbank ausgelesen und in der Ontologie als einzelne Personeninstanzen und ihnen zugeordnete Relationen wie *vocab:hasResearchTopic* und *swrc:email* abgebildet.

Die Kategorien der Pädagogik-Systematik wurden als Unterklassen der Klasse *swrc:ResearchTopic* angelegt, denen die einzelnen Wörterbuchterme als Instanzen zugeordnet werden. Beim Einlesen der in der Personendatenbank verzeichneten Forschungsgebiete der einzelnen Personen soll ein stringbasierter Abgleich mit diesen Forschungsgebieten stattfinden. Die daraus resultierende Verknüpfung zwischen Instanzen der Klassen *swrc:Person* und *swrc:ResearchTopic* integriert die Personendaten in einen größeren semantischen Kontext.

Die ontologiebasierte Integration von Daten aus zwei verschiedenen Quellen wird anhand von Abbildung 3 beispielhaft veranschaulicht.

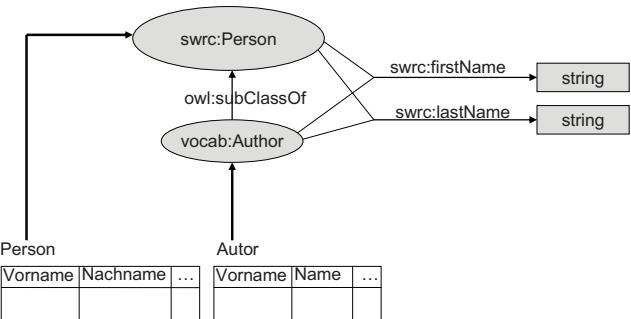


Abb. 3: Beispiel für die Überführung von Daten aus den Ausgangsdatenquellen in die Ontologie

Aus unterschiedlichen Quellen stammende Autoren- und Personendaten sollen über die Ontologie zusammengeführt werden. Dass Autoren eine Untergruppe der Klasse *swrc:Person* darstellen, wird in der Ontologie abgebildet, indem die Klasse *vocab:Author* als Unterklasse von *swrc:Person* modelliert wird. Letztere erbt daher die Relationen *swrc:firstName* und *swrc:lastName* der Oberklasse. Autorendaten und Personendaten aus verschiedenen Ursprungsdatenquellen werden auf diese Weise in einen semantischen Zusammenhang gebracht.

Auf der Basis der Ontologie sollen jedoch nicht nur strukturelle Heterogenitäten zwischen den einzelnen Datenquellen aufgelöst und eine semantische Integration

der Daten erzielt werden, sondern auch eine stärkere Verknüpfung der Daten wird angestrebt. Dies lässt sich durch den Einsatz von Reasonern realisieren.

#### 4.4 Reasoning

Um die vorliegende Datenbasis durch die Generierung neuen Wissens anzureichern, können so genannte Reasoner eingesetzt werden. Sie greifen bei der Wissensgenerierung auf Inferenzregeln zurück, die in einer Regelsprache definiert werden müssen.

In unserem Anwendungskontext könnte eine solche Regel beispielsweise festlegen, dass eine Person, die ein Dokument zu einem bestimmten Thema publiziert hat, mit diesem Thema über die Relation *hasResearchTopic* verknüpft wird. Diese Regel würde beispielsweise in der „Semantic Web Rule Language“ (SWRL) folgenderweise formalisiert werden [O’Connor *et al.* 2005]:

$$\begin{aligned} & \text{hasPublished}(?x, ?y) \\ \wedge & \text{hasTopic}(?y, ?z) \\ \rightarrow & \text{hasResearchTopic}(?x, ?z) \end{aligned}$$

In der Ontologie würden nach der Aktivierung des Reasoners folglich sämtliche Verknüpfungen zwischen Autoren und Forschungsinhalten hergestellt werden, die aus der Analyse resultieren, mit welchen Themen sich die Publikationen der einzelnen Autoren beschäftigen.

Das Reasoning kann darüber hinaus auch für Zwecke der Datenbereinigung eingesetzt werden. Legt man beispielsweise fest, dass eine Emailadresse lediglich einer Person zugeordnet sein darf, so können zwei Objekte mit derselben Adresse als identisch identifiziert werden. Der Einsatz derartiger Regeln kann sich insbesondere bei der Zusammenführung von Daten aus unterschiedlichen Quellen als sehr hilfreich erweisen.

#### 4.5 Entwicklung einer Benutzerschnittstelle

Auf dieser ontologiebasierten integrierten Datenbasis soll der Forschungsführer als Informationsportal konzipiert werden, das sowohl eine ontologiebasierte Browsing-Möglichkeit als auch Retrievalkomponenten für bestimmte Suchszenarien bereitstellt. Diese Komponenten der Benutzerschnittstelle werden im Folgenden kurz beschrieben.

##### (1) Ontologiebasiertes Browsing

Informationen über Personen, Forschungsinhalte und Institutionen sollen auf Profilseiten dargestellt werden, die untereinander über ihre ontologischen Beziehungen verlinkt werden. Auf diese Weise wird das ontologiebasierte Browsing realisiert.

##### (2) Logikbasierte Suche

Die Retrievalkomponente soll eine logikbasierte Suche umfassen, welche es ermöglicht, anhand eines Pulldown-Menüs wie dem folgenden Anfragen an die Ontologie zu definieren:



Abb. 4: Logikbasierte Suche

#### (3) Darstellung ontologischer Verknüpfungen

Bei der Freitextsuche nach den Namen zweier Objekte der Ontologie soll es darüber hinaus möglich sein, die ontologischen Verknüpfungen dieser Objekte darzustellen. So würde die Eingabe zweier Personennamen A und B zu einer Ergebnispräsentation in der Form von Verknüpfungsketten wie der folgenden führen:

$$\begin{aligned} & A \text{ forscht zum Thema } P \\ & P \text{ ist Unterthema von } Q \\ & B \text{ forscht zum Thema } Q \end{aligned}$$

Auf diese Weise sollen Suchfunktionalitäten bereitgestellt werden, welche die ontologische Organisation der Daten nutzen, um dem Informationssuchenden einen Mehrwert gegenüber der Recherche in heterogenen Datenbankbeständen zu ermöglichen.

#### 5 Erwarteter Mehrwert

Durch die ontologiebasierte Datenintegration wird dem Informationssuchenden ein integrierter Zugriff auf die verteilten Datenbestände ermöglicht. Dies wird erreicht, indem Daten aus verschiedenen Datenbanken konzeptbasiert zusammengeführt werden. Durch die in Abschnitt 4.3 geschilderte Modellierung der Relation *swrc:email* wird es beispielsweise möglich, nach der Email einer Person zu suchen, ohne die Existenz aller verfügbaren Emailversionen abfragen zu müssen. Stattdessen kann direkt auf die subsumierende Relation *swrc:email* zugegriffen werden.

Das im vorigen Abschnitt angeführte Beispiel zur Darstellung ontologischer Verknüpfungen veranschaulicht, wie die Ontologie dazu eingesetzt werden kann, dem Nutzer Zusammenhänge zwischen einzelnen Informationsobjekten zu präsentieren. Die Abfrage derartiger Zusammenhänge aus einer oder mehreren Datenbanken würde hingegen voraussetzen, dass der Nutzer bereits eine Vorstellung über die Art der Verknüpfung zwischen den Objekten besitzt. Durch die ontologiebasierte Informationsintegration wird es möglich, diese Zusammenhänge auch Datenquellen übergreifend darzustellen.

Der Einsatz eines Reasoners führt zudem dazu, dass Informationen abgefragt werden können, die in den Ursprungsdaten nicht explizit enthalten sind. Mithilfe der logikbasierten Suche könnte beispielsweise nach einer Person gesucht werden, die zu einem bestimmten Thema forscht. Während die Ursprungsdatenbank eventuell keinen Personendatensatz enthalten würde, welche den Suchterm im Feld „Forschungsgebiet“ aufführt, können über die neu ermittelten Verknüpfungen in der Ontologie auch diejenigen Personen ermittelt werden, die zu dem angegebenen Forschungsthema publiziert haben und für den Nutzer ebenfalls von Relevanz sind.

Das ontologiebasierte Browsing soll es dem Nutzer ermöglichen, in dem Forschungsführer frei zu navigieren. Der Nutzen dessen ontologischer Struktur lässt sich an einem Szenario wie dem folgenden illustrieren: der Nutzer betrachtet das Profil eines Forschers und klickt dessen Forschungsgebiet an, um sich über weitere Forscher aus diesem Bereich zu informieren. Gleichzeitig werden ihm auf der Basis der Ontologie auch verwandte Forschungsbereiche wie Teilgebiete und Obergebiete angezeigt. Auf diese Weise wird der Nutzer sowohl bei

der Verbreiterung als auch bei der Spezifizierung seiner Informationssuche unterstützt.

Derartige ontologische Verknüpfungen in der Form von Hyponymen, Hyperonymen und Synonymen sind auch für die Generierung von anfragesensitiven Suchvorschlägen einsetzbar. Integriert in eine Freitextsuchkomponente wird dem Nutzer damit die Möglichkeit geboten, seine Suchanfrage zu reformulieren oder zu skalieren.

Auch eine automatische Query Expansion-Funktionalität lässt sich durch die Nutzung der in der Ontologie hinterlegten Synonymrelationen in die Freitextsuche integrieren.

Durch die Nutzung von Standardvokabularen wird darüber hinaus die spätere Interoperabilität mit externen Semantic Web-Anwendungen angestrebt.

## 6 Evaluierung

Die abschließende Evaluierung verfolgt das Ziel, die Funktionalitäten des Forschungsführers anhand von Nutzungsszenarien zu bewerten. Einer internen Studie zufolge schreibt etwa ein Drittel der Fachportal-Nutzer den dort verfügbaren erziehungswissenschaftlichen Informationen über Experten und Institutionen ein hohes Gewicht zu<sup>8</sup>.

Im Rahmen von Nutzertests sollen daher praxisnahe Rechercheszenarien in Form von Use Cases eingesetzt werden, um das Retrieval auf den verteilten Ausgangsdatenquellen mit der Recherche im Forschungsführer-Portal zu vergleichen. Hierbei sind insbesondere die Aspekte der Informationsaufbereitung und –anreicherung sowie der Retrievalmöglichkeiten zu bewerten. Ziel ist es zu eruieren, ob der Einsatz semantischer Technologien im geschilderten Anwendungskontext einen Mehrwert für den Nutzer darstellt.

## 7 Fazit

Der vorliegende Beitrag hat den geplanten Aufbau eines ontologiebasierten Forschungsführers für den Fachbereich Bildungsforschung geschildert. Es wurde dargestellt, wie Semantic Web-Technologien eingesetzt werden sollen, um mehrere heterogene Datenquellen des Informationszentrums Bildung semantisch zu integrieren und den Informationssuchenden darauf basierend Mehrwert bringende Recherchefunctionalitäten anbieten zu können.

## Literatur

[Auer und Lehmann 2007] S. Auer, J. Lehmann. What Have Innsbruck and Leipzig in Common? Extracting Semantics from Wiki Content. 4<sup>th</sup> European Semantic Web Conference (ESWC 2007), Innsbruck, Österreich.

[Baerisch 2007] S. Baerisch. Heterogenität in wissenschaftlichen Fachdatenportalen. In: A. Oßwald, M. Stempfhuber, C. Wolff: Open Innovation. Neue

Perspektiven im Kontext von Information und Wissen. Konstanz: UVK Verlagsgesellschaft, 509-518.

[Bambey und Jornitz 2006] D. Bambey, S. Jornitz. Fachportal Pädagogik – Recherche und mehr. Buch und Bibliothek, 57(4): 336-337, 2006.

[Baumgartner et al. 2005] R. Baumgartner, N. Henze, M. Herzog. The Personal Publication Reader: Illustrating Web Data Extraction, Personalization and Reasoning for the Semantic Web. 2<sup>nd</sup> European Semantic Web Conference (ESWC 2005), Heraklion, Griechenland.

[Berners-Lee et al. 2001] T. Berners-Lee, J. Hendler, O. Lassila. The Semantic Web. Scientific American, 284(5): 34-43, Mai 2001.

[Bizer und Seaborne 2004] C. Bizer, A. Seaborne. D2RQ - Treating Non-RDF Relational Databases as Virtual RDF Graphs. 3<sup>rd</sup> International Semantic Web Conference (ISWC 2004), Hiroshima, Japan.

[Jung et al. 2007] H. Jung, M. Lee, W. Sung, D. Park. Semantic Web-Based Services for Supporting Voluntary Collaboration among Researchers Using an Information Dissemination Platform. Data Science Journal, 6: 241-249, 2007.

[Mika 2005] P. Mika. Flink: Semantic Web technology for the extraction and analysis of social networks. Journal of Web Semantics, 3(2-3): 211-223.

[O'Connor et al. 2005] M. O'Connor, H. Knublauch, S. Tu, B. Grosof, M. Dean, W. Gross, M. Musen. Supporting Rule System Interoperability on the Semantic Web with SWRL. 4<sup>th</sup> International Semantic Web Conference (ISWC 2005), Galway, Irland.

[Stempfhuber 2003] M. Stempfhuber. Objektorientierte Dynamische Benutzungssoberfläche - ODIN. Bonn: GESIS Informationszentrum Sozialwissenschaften.

[Sure et al. 2005] Y. Sure, S. Bloehdorn, P. Haase, J. Hartmann, D. Oberle. The SWRC Ontology - Semantic Web for Research Communities. 12<sup>th</sup> Portuguese Conference on Artificial Intelligence - Progress in Artificial Intelligence (EPIA 2005), Covilha, Portugal, Dezember 2005.

<sup>8</sup> D. Bambey: Das Evaluationskonzept des Fachportal Pädagogik: Einsatz von Evaluationsmethoden und Erstellung eines Wiki-Leitfadens. Vortrag auf der ISI 2007 & IuK 2007 am 01.06.2007.  
[http://www.iuk2007.de/fileadmin/user\\_upload/iuk\\_abstracts.html#bambe](http://www.iuk2007.de/fileadmin/user_upload/iuk_abstracts.html#bambe) [zuletzt besucht am 22.08.2007].