

Erweiterungspotenziale für die Modellierung mit „H2 for Reporting“ am Beispiel von Forschungsportalen

Ralf Knackstedt, Lukasz Lis

European Research Center for Information Systems (ERCIS)
University of Muenster

1 Fachkonzeptionelle Modellierung Data-Warehouse-basierter OLAP-Systeme

Für die Modellierung der fachlichen Anforderungen an Data-Warehouse-basierte OLAP-Systeme wurde eine Vielzahl unterschiedlicher Modellierungstechniken entwickelt (vgl. Böhnlein 2001): Klassische Ansätze lehnen sich eng an Modellierungssprachen an, die für die Fachkonzeption auf relationalen Datenbanken basierender, operativer Anwendungen etabliert sind, wie z. B. das Entity-Relationship-Modell (Chen 1976) oder die Objekttypenmethode (Wedekind 1981). Erweiterungen klassischer Datenmodellierungsansätze modifizieren die etablierten Datenmodellierungstechniken, wodurch insbesondere den rechentechnischen und sachlogischen Kennzahlenbeziehungen sowie der Modellierung von Bezugsobjekten auf Ausprägungsebene eine stärkere Beachtung geschenkt wird. Rein multidimensionale Modellierungsansätze verzichten auf eine Anlehnung an bestehende Datenmodellierungsansätze. Andere Modellierungsansätze entstammen dem Bereich der Scientific and Statistical Data Bases, der sich bereits seit den achtziger Jahren mit der Spezifikation multidimensionaler Datenräume befasst. Für die Spezifikation der bereitzustellenden Data-Warehouse-Daten werden darüber hinaus auch objektorientierte und erweiterte objektorientierte Modellierungsansätze vorgeschlagen. Einige dieser Ansätze zielen speziell auf eine Erweiterung der Unified Modeling Language (Booch et al. 1999) ab.

Angesichts der Vielzahl bereits bestehender Modellierungstechniken und vorhandener Standardisierungsbestrebungen erscheint es fraglich, ob die Weiterentwicklung einzelner Modellierungstechniken zukünftig eine gleich bleibend wichtige Aufgabenstellung für die Forschung im Bereich Business Intelligence darstellen wird. Anhand ausgewählter *Erweiterungspotenziale* plädiert dieser Beitrag dafür, die Entwicklung fachkonzeptioneller Modellierungstechniken für Data-Warehouse-basierte OLAP-Systeme nicht als abgeschlos-

sen anzusehen, sondern als ein interessantes Forschungsfeld, in dem sich weiterhin relevante Ergebnisse erzielen lassen.

Als Ausgangsbasis für die Weiterentwicklung wird die Modellierungstechnik H2 for Reporting gewählt, die im Folgenden kurz charakterisiert wird (vgl. Abschnitt 2). Die in diesem Beitrag vorgestellten Entwicklungspotenziale dieser Modellierungstechnik lassen sich auf andere Modellierungsansätze übertragen und kommen in unterschiedlichen Anwendungsbereichen Data-Warehouse-basierter OLAP-Systeme zum Tragen. Um die Entwicklungspotenziale trotz ihrer allgemeinen Anwendbarkeit konkretisieren zu können, wird ihre Bedeutung exemplarisch für die Fachkonzeption von Forschungsportalen erläutert. Die Auswahl dieses Anwendungsgebiets ist durch die Zielstellung des Workshops, in dessen Rahmen dieser Beitrag präsentiert wird, motiviert. Sie umfasst die Entwicklung einer „Forschungslandkarte für Business Intelligence“. Eine Forschungslandkarte kann als eine Ausprägung von Forschungsportalen angesehen werden. Im Anschluss an die Erörterung der Bedeutung von Business Intelligence für Forschungsportale (Abschnitt 3), werden konstruktionsorientierte Forschungsperspektiven aufgezeigt, die in einem engen bis weiteren Zusammenhang mit H2 for Reporting derzeit verfolgt werden bzw. werden sollen. Die Weiterentwicklungsperspektiven werden jeweils im Kontext des forschungsportalbasierten Controlling konkretisiert. Ein kurzes Fazit schließt den Beitrag ab (vgl. Abschnitt 5).

2 H2 for Reporting als Ausgangspunkt

Die Modellierungssprache H2 for Reporting ist in ihrer aktuellen Fassung aus einem mehrjährigen Entwicklungsprozess hervorgegangen (vgl. Holten 2003). Mit den anderen für die Fachkonzeption Data-Warehouse-basierter OLAP-Systeme vorgeschlagenen Modellierungstechniken weist sie wesentliche Gemeinsamkeiten, aber auch Besonderheiten auf.

Als gemeinsamer konzeptioneller Kern der Modellierungstechniken lassen sich Dimensionen und Kennzahlensysteme identifizieren, die zu Navigationsräumen zusammengefasst werden. Die Navigationsräume beschreiben die Datenbestände, die das OLAP-System auswertbar machen sollen. Die einzelnen Modellierungstechniken verwenden teilweise unterschiedliche Bezeichnungen für gleiche oder ähnliche Konstrukte. Darüber hinaus unterscheiden sie sich in weiteren Details, die darauf zurückzuführen sind, welche Zielsetzungen bei der Konstruktion der Modellierungstechniken schwerpunktmäßig verfolgt wurden.

H2 for Reporting zeichnet sich im Vergleich zu anderen Modellierungssprachen insbesondere dadurch aus, dass mit ihr neben den auszuwertenden Navigationsräumen auch die

Tabellenstrukturen beim Aufbau der OLAP-Berichte spezifiziert werden können. Darüber hinaus steht mit dem H2-Toolset ein Metamodellierungswerkzeug zur Verfügung, mit dem sich H2 for Reporting-Modelle erstellen und verwalten lassen.

Die inhaltliche Spezifikation der OLAP-Berichte erfolgt, indem über die Angabe von Zeilen- und Spalteninhalten eine zweidimensionale Projektion auf einen Navigationsraum festgelegt wird (vgl. Abbildung 1). Die Zeilen- und Spaltenstruktur legt den tabellarischen Aufbau des Berichts fest. Navigationsräume werden von Dimension(sausschnitt)en (z. B. Tage nach Jahren, Institutionen nach Ländern und Forschungsthemen nach Themengruppen) und Kennzahlensystemen (z. B. Anzahl der Forschungsergebnisse) aufgespannt. Der Navigationsraum macht auf diese Weise deutlich, welche Kennzahlen nach welchen inhaltlichen Kriterien analysierbar sein sollen.

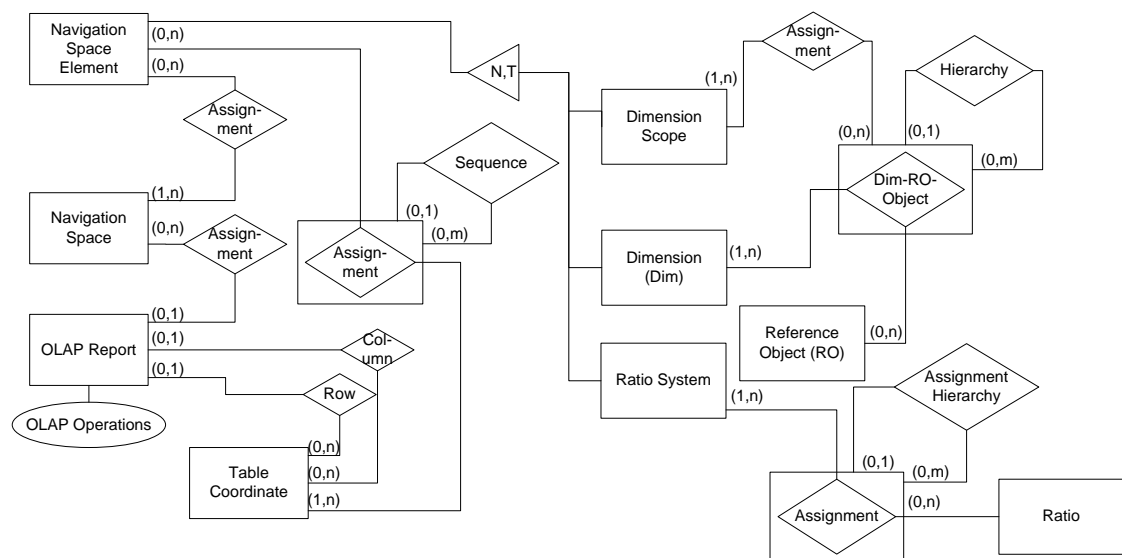


Abbildung 1: Metamodell von H2 for Reporting (Ausschnitt)

3 Fachkonzeption von Forschungsportalen als Gegenstand der BI-Forschung

Unter (internetbasierten) *Forschungsportalen* werden themenspezifische Websites verstanden, die Informationen, die an anderer Stelle in ausführlicher Form bereits vorhanden sind, spezifisch strukturiert zur Verfügung stellen, mit dem hauptsächlichen Ziel, die Zusammenarbeit in Forschungsprojekten zu fördern oder die Bekanntheit geplanter bzw. erzielter Forschungsergebnisse zu steigern. Um dies zu erreichen, dokumentiert die Website, wer Forschung betreibt und/oder welche Ziele in Forschungsprojekten verfolgt und/oder welche Ergebnisse in Forschungsprojekten erzielt wurden bzw. werden sollen. Forschung umfasst dabei sowohl akademische Forschung an Hochschulen und sonstigen

staatlichen Forschungseinrichtungen als auch Forschung in Unternehmen, öffentlichen Verwaltungen etc. Um die beschriebenen Ziele zu erreichen, stellt eine geographische Darstellung der Forschungsstandorte häufig eine wünschenswerte aber nicht immer notwendige Funktionalität von Forschungsportalen dar. Im Zusammenhang mit der geographischen Darstellung der Forschungsstandorte wird auch der Begriff Forschungslandkarte verwendet. Teilweise wird der Begriff der Forschungs(land)karte auch für die Visualisierung von Themengebieten verwendet, welche von dem Forschungsportal adressiert werden. Die thematische Eingrenzung spezialisiert den Zweck der Website. Beispiele sind die Förderung der Krebsforschung (<http://www.cancermap.org>) oder die Erforschung des Einsatzes von Windenergie (<http://forschungslankarte-windenergie.de>). Dabei sind auch geografische oder organisatorische Eingrenzungen des Betrachtungsgegenstands des Forschungsportals möglich. Ein Beispiel stellt die Beschränkung der Betrachtung der Stammzellenforschung auf diejenige im Bundesland Nordrhein-Westfalen dar (vgl. z. B. <http://stammzellen.nrw.de>). In Abhängigkeit von der themenspezifischen Zielstellung strukturiert die Website die zur Verfügung gestellten Informationen. Diese Struktur ist charakteristisch für die Website. Der Detaillierungsgrad und Facettenreichtum der Strukturierung ist dabei allerdings vom Qualitätsanspruch der Website und ihres Informationsumfangs abhängig.

Forschungsportale stellen eine Datenbasis zur Verfügung, deren Auswertung mit Business Intelligence (BI)-Systemen einen wichtigen Beitrag zum *Forschungs-Controlling* leisten kann. Ein forschungsportalbasiertes Controlling liefert dabei unterschiedlichen Anspruchsgruppen wichtige Informationen:

- *Einzelne Forscher* können z. B. analysieren, wer sich mit ähnlichen oder ergänzenden Themen beschäftigt, um Kooperationspartner für zukünftige Projekte zu identifizieren.
- *Forschungsabteilungen oder -institute* können mit entsprechenden Auswertungen z. B. dokumentieren, welchen Themen sie sich im Zeitverlauf mit welcher Intensität gewidmet haben und mit welchen anderen Forschern sie in welchem Umfang kooperiert haben.
- *Hochschulen und außeruniversitäre Forschungseinrichtungen* können z. B. die Forschungsleistungen ihrer Mitglieder in verschiedenen Themenbereichen untersuchen.
- *Staatliche und private Fördereinrichtungen* können z. B. untersuchen, zu welchen Forschungsbereichen bereits viele Ergebnisse vorliegen, um so gegebenen-

falls vorhandene Lücken zu identifizieren, für welche die Intensivierung der Förderung angezeigt wäre.

Zu den Dimensionen, nach denen im Rahmen des forschungportalbasierten Controllings beispielsweise ausgewertet werden kann, zählen: Forschungsergebnistyp, Realisationsgrad, Praxiseinsatz, Funktionsbereich, Organisationsklasse, Anwendungsbranchen und Standardisierung. Als Kennzahl wird häufig die Anzahl der im Portal beschriebenen Forschungsergebnisse ermittelt. Im Zuge einer leistungsabhängigen Mittelvergabe an Forschungseinrichtungen hat sich eine Vielzahl weiterer Kennzahlen entwickelt, deren Ermittlung von Forschungsportalen ebenfalls unterstützt werden kann. Viele der Bewertungsansätze ermitteln dabei nicht allein die Zahl der Publikationen, sondern bewerten diese entsprechend des Stellenwerts des jeweiligen Publikationsorgans. Beispielsweise wird von Garfield (1955) der Impact Factor (IF) propagiert, der im Rahmen des (Social) Science Citation Index bis heute aktiv berechnet wird. Dabei wird gemessen, wie oft Artikel einer Fachzeitschrift in anderen Fachzeitschriften zitiert werden und diese Zahl in Verhältnis zur Gesamtheit zitierbarer Artikel gesetzt. Perspektiv ist davon auszugehen, dass an Hochschulen und in Forschungsfördereinrichtungen über Forschungsportale Datenbanken zunehmend größeren Datenvolumens aufgebaut werden, in denen die Forschungsleistungen (Publikationen, Patente, Preise, Projekte) dokumentiert werden. Die Auswertung dieser Datenbasis zu unterschiedlichen Zwecken (strategische Ausrichtung der Forschungsbemühungen, Identifikation von Kooperationspartnern und Synergiepotenzialen, Mittelvergabe an Institute, Unterstützung von Gehaltsverhandlungen etc.) anhand jeweils verschiedener Dimensionen und Kennzahlen motivieren, die Datenbasis eines Forschungsportals in Anlehnung an die Data Warehouse-Architektur aufzubauen (vgl. Becker et al. 2008).

Mit yourResearchPortal.com wird ein softwaretechnisches System entwickelt, mit Hilfe dessen Forschungsportale automatisiert generiert, administriert und betrieben werden können. Bei der Ausgestaltung des Systems wurden zwei Hauptkomponenten unterscheiden. Die Datenpflegekomponente umfasst Funktionalitäten, die für das Erstellen, Betrachten, Aktualisieren und Löschen von konkreten Daten gebraucht werden. Typisch hierfür sind ein hoher Detaillierungsgrad der Daten und die Betrachtung von einzelnen Inhalten. Die Analysekomponente andererseits ermöglicht einen nur lesenden Zugriff auf die gesamte Wissensbasis eines Portals. Als Datenpflegekomponente fungiert das Open Source Content Management System Drupal (<http://www.drupal.org>), welches modular aufgebaut ist. Zusätzlich zum für die grundlegende Funktionalität zuständigen sogenannten Core können weitere Funktionalitäten durch die Installation verschiedener Module hinzugefügt werden. Die Analysekomponente wird in yourResearchPortal.com mit Hilfe des Open Source OLAP-Server Mondrian realisiert. Dieser ermöglicht sowohl die kom-

fortable Durchführung quantitativer Analysen als auch ein effizientes und interaktives Navigieren durch die mehrdimensionale Wissensbasis bis hin zur Ebene der einzelnen Inhalte. Die beiden Hauptkomponenten sind in einer webbasierten Benutzeroberfläche vollständig integriert. Sowohl die Datenpflege- als auch die Analysekomponente greifen auf die gleiche relationale MySQL-Datenbank (<http://www.mysql.com>) zu.

Obwohl die beiden Komponenten eine unterschiedliche Ausrichtung aufweisen (OLTP bei Drupal und OLAP bei Mondrian) wurde bei der DV-technischen Umsetzung darauf geachtet, ein relationales Schema zu entwickeln, das beiden Aufgaben gerecht wird. Wichtige Herausforderung war es hierbei, das System so zu gestalten, dass Änderungen in Drupal (z. B. das Hinzufügen einer neuen Dimension oder eines neuen Wertes) von Mondrian direkt übernommen werden, ohne dass eine zusätzliche Konfiguration der Analysekomponente nötig ist. Dies wäre im klassischen Star-Schema der Fall, wo jede Dimension durch eine eigene Tabelle repräsentiert wird (Kemper et al. 2006, S. 61-64). In Abbildung 2 wird ein Schemaausschnitt präsentiert, der den in yourResearchPortal.com verfolgten Ansatz schildert. Durch Rechtecke werden Tabellen samt ihren Namen und Spalten (Primärschlüssel unterstrichen in Fettschrift) dargestellt. Pfeile zeigen auf die von Fremdschlüsseln referenzierten Spalten. Neben der Faktentabelle Einordnung, in der die Zuweisung von Artefakten (Forschungsergebnissen) zu Dimensionsobjekten in Anhängigkeit vom Portal festgehalten wird, existieren drei Dimensionstabellen Portal, Dimensionsobjekt und Artefakt.

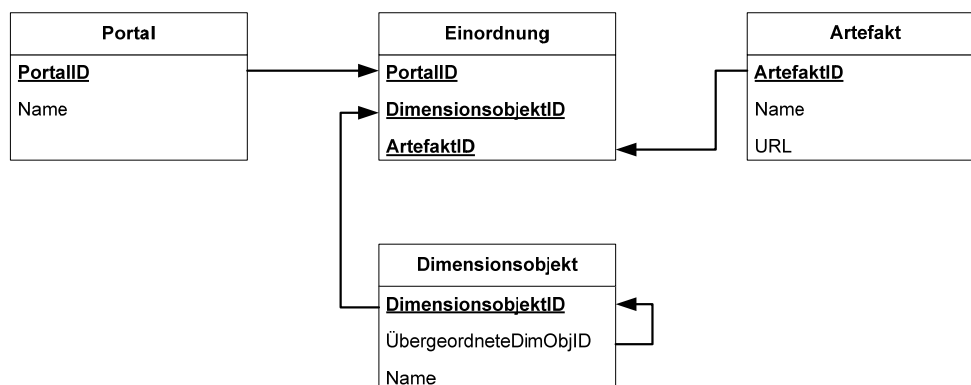


Abbildung 2: Generisches multidimensionales Analyseschema

Dimensionen werden durch Wurzel-Dimensionsobjekte abgebildet, die kein übergeordnetes Dimensionsobjekt besitzen. Diese werden wiederum von den Werten der ersten Hierarchieebene referenziert usw. Der OLAP-Server Mondrian ist so konfiguriert, dass den Benutzern eines Portals die entsprechenden Dimensionen samt ihren Werten und eventuellen untergeordneten Organisationen oder Projekten für Analysezwecke zur Verfügung stehen. Im Hintergrund wird ein Filter aktiviert, welcher nur Daten des aktuellen Portals

darstellen lässt. Hierfür werden die Ausprägungen der Analysedimension Portal auf die entsprechende ID eingeschränkt.

Infrastrukturen wie yourResearchPortal.com ermöglichen es, dass sich Betreiber von Forschungsportalen bei deren Realisierung auf die fachkonzeptionelle Ausgestaltung des Portals konzentrieren können. Die Spezifikation der Anforderungen an die Controllingauswertungen und damit die Festlegung der auch in der Datenpflegekomponente verwendeten Dimensionen steht dabei im Vordergrund. Die Diskussion und Dokumentation der gewünschten Ausgestaltung des Forschungsportals stellt daher ein Anwendungsgebiet für die fachkonzeptionelle Modellierung Data-Warehouse-basierter OLAP-Systeme dar.

4 Entwicklungsperspektiven für H2 for Reporting am Beispiel des forschungportalbasierten Controlling

Im Folgenden wird ein Überblick über weiterführende Entwicklungen gegeben, die für die Modellierungssprache H2 for Reporting untersucht wurden bzw. werden. Die vorgestellten Entwicklungspotenziale adressieren dabei jeweils konstruktionsorientierte Forschung. Im Sinne von Design Science beinhaltet diese die Entwicklung von Artefakten, die einen Problemlösungsbeitrag leisten, und die Evaluation dieses Problemlösungsbeitrags (vgl. z. B. Hevner et al. 2004). Zu den im Rahmen der Wirtschaftsinformatik besonders relevanten Artefakttypen zählen Sprachkonstrukte, Methoden, Modelle und Implementierungen. Die konstruktionsorientierten Forschungsfragen werden zunächst allgemein eingeführt und anschließend konkret auf das forschungportalbasierte Controlling angewandt (vgl. Tabelle 1).

Entwicklungsperspektive	Fragestellungen	Anwendungsbeispiele im forschungportalbasierten Controlling
Kollaborative Entwicklung von H2 for Reporting-Modellen	Welche Mechanismen unterstützen die Konsensbildung und wie lassen sich diese softwaretechnisch umsetzen?	Kollaborative Diskussion von Dimensionen, die zur Beschreibung von Forschungsergebnissen vom Forschungsportal genutzt werden

Entwicklungsperspektive	Fragestellungen	Anwendungsbeispiele im forschungportal- basierten Controlling
Integration von H2 for Reporting-Modellen mit weiteren Modelltypen	Wie lassen sich die Metamodelle unterschiedlicher Modellierungstechniken integrieren und wie kann die gestiegene Komplexität z. B. durch Sichtenkonzepte gehandhabt werden?	Integration von H2 for Reporting mit einer Technik zur Spezifikation von Websites
Referenzmodellierung zur Wiederverwendung von H2 for Reporting-Modell(teil)en	Mit welchen Mechanismen lässt sich die Wiederverwendung unterstützen und wie lassen sich diese softwaretechnisch umsetzen? Welche Berichtsspezifikationen eignen sich in bestimmten Anwendungsdomänen für eine Wiederverwendung?	Varianten von Fachkonzepten für Forschungsportalen inklusive ihrer Auswertungsfunktionalität und -inhalte spezifizieren und generieren
Optimierung einer Menge Berichtsspezifikationen z. B. durch Vergleich der Berichtsspezifikationen auf Ähnlichkeit	Wie lässt sich der Vergleich von Modell(teil)en auf Ähnlichkeit automatisieren z. B. durch die Adaption der Clusteranalyse?	Vergleich von Berichtsdefinitionen im Forschungs-Controlling zur Identifikation von ähnlichen Berichtsanforderungen
Automatisierung der Implementierung von Berichten auf der Basis H2 for Reporting-Modell(bestandteil)en	Welche Algorithmen sind geeignet, um OLAP-Systeme auf der Basis der fachkonzeptionellen Modelle automatisiert einzurichten?	Generierung von Auswertungen für Forschungsportale oder weiterer Bestandteile auf der Basis fachkonzeptioneller Modelle

Konstruktion von Reifegradmodellen zur Beurteilung von BI-Systemen bzw. BI-Komponenten enthaltenen Systemen	Welche Inhalte von H2 for Reporting (Referenz)modellen bzw. welche methodischen Aspekte der Modellierung mit H2 for Reporting eignen sich zur Beschreibung von Reifegraden?	Entwicklung eines Reifegradmodells für Forschungsportale
---	---	--

Tabelle 1: Erweiterungspotenziale von H2 for Reporting

Kollaborative Entwicklung: Fachkonzepte werden in aller Regel nicht durch eine einzelne Person, sondern unter Beteiligung mehrerer Anspruchsgruppen entwickelt, die möglichst einen Konsens herbeiführen. Eine erste Forschungsfrage, die über die reine Definition der OLAP-Spezifikationsprache hinausführt, aber zugleich auch eng mit ihr verbunden ist, stellt daher die Entwicklung einer Diskussionsumgebung dar, mit der sich fachkonzeptionelle Modelle kollaborativ erarbeiten lassen. Im Kontext der Prozessmodellierung sind derartige Ansätze schon relativ weit verbreitet. Die „Open Model Initiative“ (vgl. Frank, Strecker 2007) adressiert die gemeinschaftliche Erstellung beliebiger Modelltypen. Für den OLAP-Kontext ist zu untersuchen, welche Mechanismen sich zur Koordination der Diskussionen übernehmen und auf welche Modellbestandteile sich diese erfolgreich anwenden lassen. Für H2 for Reporting wurde eine Erweiterung des H2-Toolsets entwickelt, mit der sich Diskussionsprozesse unterstützen lassen, die in verschiedenen Anwendungskontexten evaluiert wurde. Im Rahmen des forschungportalbasierten Controlling wurde diese Erweiterung genutzt, um eine funktionale Dimension zu entwickeln, der später Forschungsergebnisse zugeordnet werden konnten. Die diskutierte Dimension wurde im Anschluss in einen funktionalen Ordnungsrahmen umgesetzt, der in der DIN PAS 1094 dokumentiert wird.

Integration mit weiteren Modelltypen: OLAP-Berichte werden in bestimmten Kontexten verwendet. Beispielsweise lässt sich beschreiben, für welche betriebswirtschaftlichen Aufgaben OLAP-Berichte eingesetzt werden sollen und welche Organisationseinheiten, die Berechtigung zum Aufruf eines bestimmten Berichts erhalten sollen. Von besonderer Bedeutung für die Verbreitung fachkonzeptioneller Modelle erscheint daher die Integration der OLAP-Modelle mit anderen verbreiteten Modelltypen (vgl. Knackstedt 2006). Durch die Annotation von OLAP-Berichten an Funktionen in Prozessmodellen kann z. B. ausgedrückt werden, in welchen Arbeitskontexten OLAP-Berichte verwendet werden sollen. Insbesondere bei wiederkehrenden Aufgaben, wie z. B. der Sortimentsgestaltung im Handel, bietet sich eine Dokumentation in Prozessmodellen an. Die Integrationsauf-

gabe umfasst unter anderem die Konstruktion aufeinander abgestimmter Metamodelle mit geeigneten Sichtenkonzepten, welche die steigende Komplexität der integrierten Modelle handhabbar machen. Im Kontext des forschungsportalbasierten Controlling wurde die Integration der H2 for Reporting-Sprachkonstrukte in die Internetspezifikationsprache eW3DT nach Scharl (vgl. Scharl 1997) untersucht (vgl. Knackstedt et al. 2009). Dabei wurden OLAP-Berichtsdefinitionen als Inhalte einzelner Web-Seiten in die Spezifikationsprache integriert. Außerdem wurde erkannt, dass auch einzelne Dimensionen, Kennzahlensysteme und Navigationsräume zur Spezifikation von Bestandteilen spezifischer Web-Seitentypen, wie z. B. Index und Suche, herangezogen werden können (vgl. Abbildung 3). Die integrierte Modellierungssprache eignet sich dazu, Forschungsportale fachkonzeptionell zu spezifizieren und dabei sowohl die Seitenstrukturen der Forschungsportale abzubilden, als auch die OLAP-Berichte zur Auswertung der durch die Forschungsportale erfassten Daten zu spezifizieren.

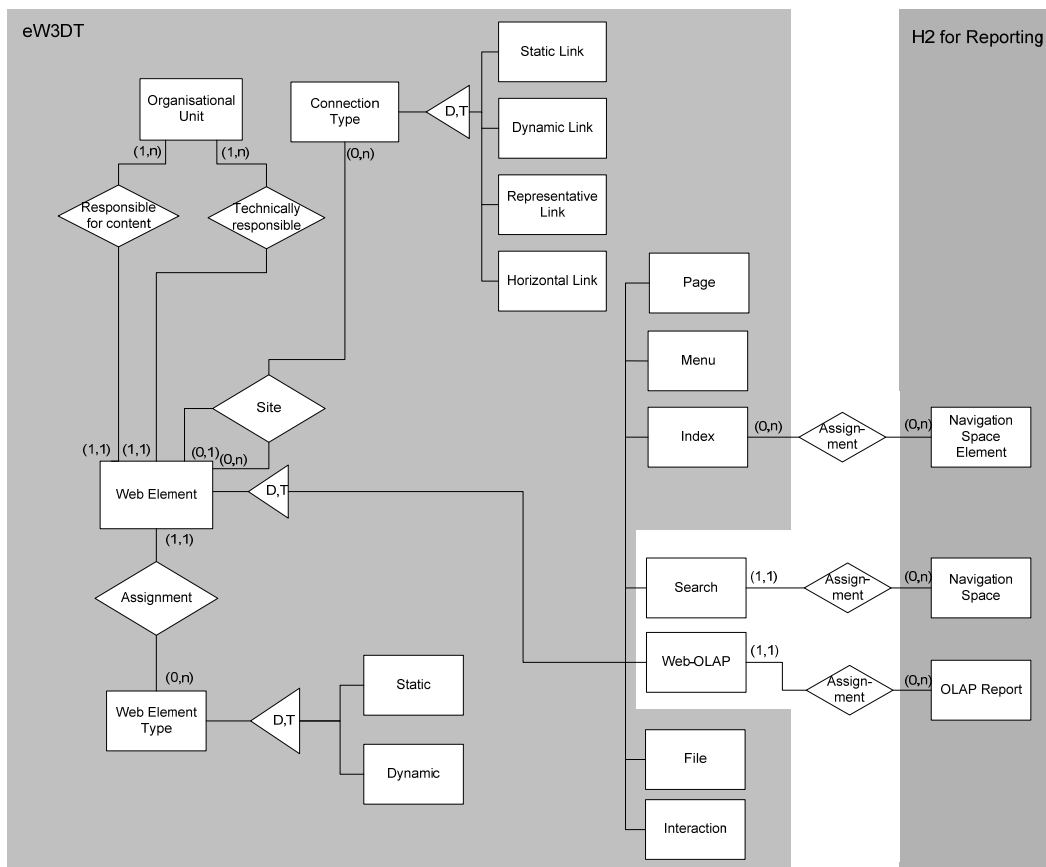


Abbildung 3: Integration von H2 for Reporting mit eW3DT (Ausschnitt)

Referenzmodellierung: Modellierungswerkzeuge profitieren in besonderem Maße davon, wenn sie nicht allein mit vordefinierten bzw. adaptierbaren Modellierungssprachen zur Verfügung gestellt werden, sondern auch mit anpassbaren, wiederverwendbaren Model-

len. Dieser Umstand und weitere Vorteile der Referenzmodellierung haben motiviert, zu untersuchen, welche Spracherweiterungen notwendig sind, um die Wiederverwendung von OLAP-Fachkonzepten zu unterstützen. Grundsätzlich lassen sich zur Unterstützung der Wiederverwendung Mechanismen unterscheiden, die auf Aggregation, Spezialisierung, Analogiebildung, Instanziierung und Konfiguration beruhen (vgl. Becker, Knackstedt 2004). Für H2 for Reporting wurde eine konfigurierbare Sprachversion entwickelt, die es ermöglicht, Regeln zu formulieren, die festlegen, wie ein fachkonzeptionelles Modell an bestimmte Anwendungskontexte anzupassen ist (vgl. Knackstedt, Seidel, Janiesch 2006). Im Kontext des forschungportalbasierten Controlling wird eine Referenzmodell für Forschungsportale entwickelt. Hierzu wurden bisher im Internet 66 Forschungsportale identifiziert und hinsichtlich ihrer Inhalte analysiert. Die im Sinne eines Common Sense üblichen Inhalte wurden mit einer Kombination aus H2 for Reporting und eW3DT modelliert (vgl. Abbildung 4). In weiteren Forschungsarbeiten soll die empirische Basis zur Entwicklung des Referenzmodells erweitert werden. Hierbei wird untersucht, welche typischen Varianten von Forschungsportalen existieren. Es wird z. B. vermutet, dass sich die Forschungsportale einzelner Forscher von denen von Forschungsnetzwerken durch spezifische Inhalte bzw. Funktionen unterscheiden. Des Weiteren sind Unterschiede zwischen Forschungsportalen, die von der Wirtschaft bzw. der Wissenschaft betrieben werden, wahrscheinlich. Die Arbeiten zur Spracherweiterung ermöglichen die adäquate Dokumentation der Varianten in einem konsistenten Gesamtmodell. Mittels Auswertung der Konfigurationsregeln lassen sich dann kontextspezifische Varianten aus dem Gesamtmodell ableiten.

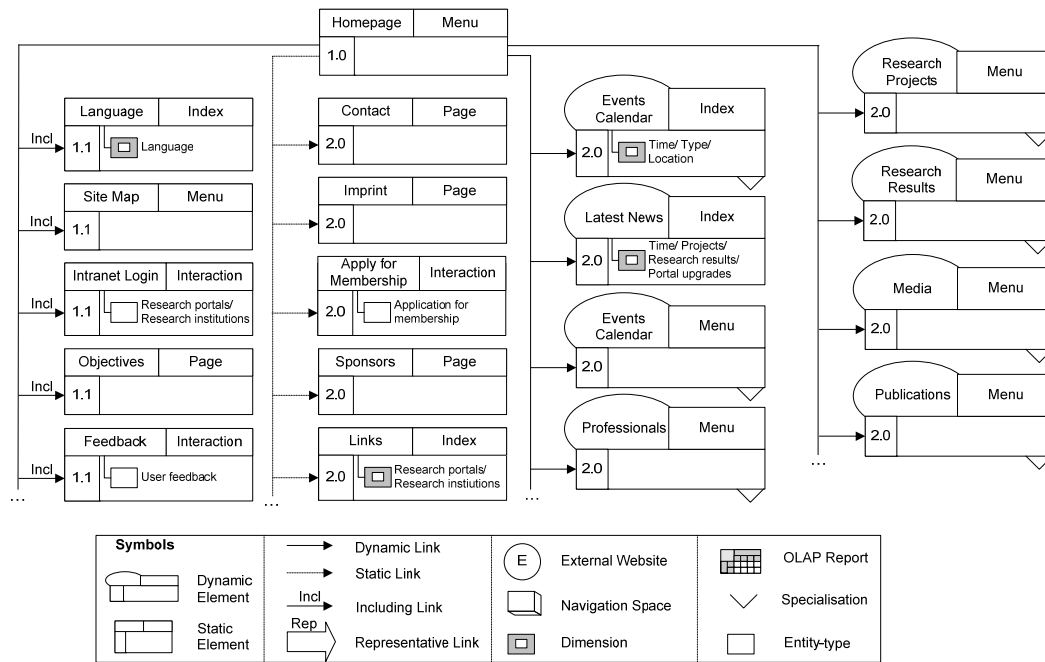


Abbildung 4: Forschungsportal-Referenzmodell (Ausschnitt)

Optimierung einer Menge OLAP-Berichtsspezifikationen: Der Abgleich eines bestehenden Berichtssystems mit einem Referenzmodell stellt einen möglichen methodischen Ansatz zu dessen inhaltlichen Verbesserung dar. Die Orientierung an den Begriffen eines Referenzmodells kann z. B. zu einer Vereinheitlichung des Sprachgebrauchs im Berichtswesen eines Unternehmens beitragen. Die Präzisierung und Vereinheitlichung der Kennzahldefinitionen, die bessere Abstimmung von Informationsangebot und Informationsnachfrage, die Verringerung der Berichtszahl sind weitere wesentliche Ansatzpunkte für die Verbesserung eines bestehenden Berichtswesens. Zur Ermittlung des Informationsbedarfs tragen auch die zahlreichen Ansätze zur Entwicklung von Soll-Modellen bei, die sich z. B. in datenangebots-, aufgaben- bzw. zielanalytische und datennachfrageorientierte Ansätze unterteilen lassen. Die Entwicklung eines konsistenten Sollkonzepts und die Konsolidierung eines historisch gewachsenen Berichtswesens lassen sich darüber hinaus durch statistische Ansätze unterstützen, die Verbesserungs- und Vereinfachungspotenziale in den OLAP-Berichten aufzeigen bzw. identifizieren helfen (vgl. Delfmann et al. 2008). Im Rahmen der konstruktionsorientierten Forschung ist zu untersuchen, welche statistischen Methoden sich auf dieses Problem anwenden lassen und welche Anpassungen gegebenenfalls an den Modellierungssprachen bzw. -editoren vorzunehmen sind. Für H2 for Reporting und das H2 Toolset wurde die Clusteranalyse adaptiert, um im Rahmen einer Konsolidierung umfangreicher Modellsysteme automatisiert Hinweise auf ähnliche Modelle generieren zu können (vgl. Knackstedt, Deinert, Becker 2009). Für die ähnlichen Modelle soll dann geprüft werden, ob sich die Berichtsspezifikationen durch geeignet

parametrisierbare Berichte ersetzen lassen, so dass die Komplexität des Berichtsystems reduziert werden kann. Im Rahmen der Entwicklung eines Referenzmodells für das forschungportalbasierte Controlling kann dieses Verfahren genutzt werden, um Ähnlichkeiten in den Berichten unterschiedlicher Forschungsportale zu identifizieren. Hierbei werden die in Forschungsportalen verwendeten Bereichsspezifikationen miteinander verglichen. Eine geeignete Datengrundlage wird sich in diesem Bereich derzeit vor allem im Forschungscontrolling von Hochschulen finden lassen. Neben den Berichtsspezifikationen lassen sich aber auch die Modelle der Forschungsportale selbst einer Clusteranalyse unterziehen. Auf diese Weise könnten Gruppen von Forschungsportalen gleicher bzw. ähnlicher Funktionalität identifiziert werden, die im Rahmen der Referenzmodellierung als Varianten ausgewiesen werden können.

Automatisierung der Implementierung: Ein wesentlicher Beitrag zur Reduktion des Aufwandes der Einführung von Data-Warehouse-gestützten OLAP-Systemen der Fachkonzeption stellt die automatische Überführung der fachkonzeptionellen Modelle in techniknähere (Teil-)spezifikationen bzw. Implementierungen dar, wie sie z. B. von MDA (Model Driven Architecture) adressiert wird. Im Kontext von OLAP-Modellierungsmethoden wurde diese Zielsetzung bereits frühzeitig aufgegriffen. Auch im Rahmen der Entwicklung von H2 for Reporting wurden Algorithmen zur Automatisierung der Implementierung entwickelt (vgl. Holten 2003). YourResearchPortal.com ermöglicht die Generierung von Forschungsportalen. Das System erfragt hierzu wesentliche Auswertungsdimensionen vom zukünftigen Administrator des zu generierenden Forschungsportals. In Übereinstimmung mit den Strukturen der H2 for Reporting-Sprache kann der Nutzer über das Internet interaktiv die Auswertungsdimensionen anlegen. Im Anschluss wird ein spezifisches Forschungsportal mit angepassten OLAP-Berichten generiert. Auch die übrige Funktionalität des generierten Forschungsportals lässt sich über die Administrationskomponente anpassen. Ein erstes Forschungsportal, das mit diesem Werkzeug generiert wurde, stellt die „Forschungslandkarte zur hybriden Wertschöpfung“ (<http://www.forschungslandkarte-hybridewertschoepfung.de>) dar (vgl. Abbildung 5). Die Dimensionsspezifikationen der zu generierenden Forschungsportale ließen sich prinzipiell auch aus H2 for Reporting-Modellen auslesen, was derzeit allerdings noch nicht unterstützt wird. Mit den Arbeiten zur Referenzmodellierung von Forschungsportalen sind weitere wichtige Grundlagen zur Realisierung einer modellgetriebenen Generierung von Forschungsportalen gelegt. Das Modell, das der Generierung zu Grunde liegt, könnte dabei zunächst durch die Adaption eines Forschungsportal-Referenzmodells erzeugt werden.

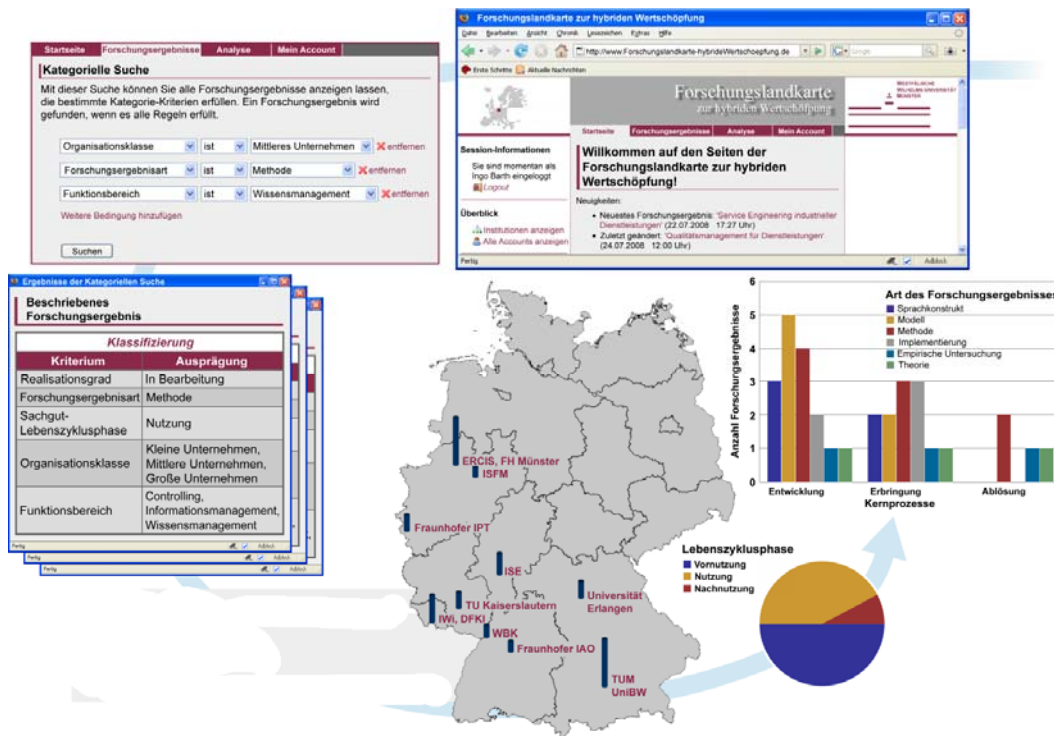


Abbildung 5: Mittels yourResearchPortal.com generierte „Forschungslandkarte zur hybriden Wertschöpfung“

Konstruktion von Reifegradmodellen: Zur Beurteilung des Entwicklungsstands des BI-Systems eines Unternehmens werden verschiedene Reifegradmodelle vorgeschlagen (vgl. Chamoni, Gluchowski 2004). In der Regel wird in Reifegradmodellen die Beurteilung nach unterschiedlichen Sichten vorgenommen. Die Verwendung von Modellierungstechniken zur Fachkonzeption des BI-Systems geht in die organisatorische Beurteilung der Prozesse der Entwicklung und des Betriebs eines BI-Systems ein. Die technische Dimension beurteilt die von den Auswertungssystemen zur Verfügung gestellte Funktionalität und die Infrastruktur zur Vereinheitlichung und Verteilung der Daten. Die inhaltliche Beurteilungsdimension untersucht, ob die für das Unternehmen relevanten Informationen bereitgestellt werden. Neben der allgemeinen Beurteilung des BI-Systems einer Unternehmung als ganzes, können Reifegradmodelle auch spezifische Controllingbereiche, wie z. B. das IT-Controlling, adressieren (vgl. Becker, Knackstedt, Pöppelbuß 2009). Je spezifischer die Controlling-Bereiche gewählt werden, für die ein Reifegradmodell entwickelt wird, desto konkreter können die bereitgestellten Inhalte von BI-Systemen beurteilt werden. Für das forschungsportalspezifische Controlling wird ein Reifegradmodell entwickelt, das neben forschungsportalspezifischer Funktionalität insbesondere auch die Abdeckung relevanter Inhaltsbereiche, wie z. B. Forschungsprojekte, -organisationen, -ziele, -ergebnisse und -publikationen untersucht. Für höhere Reifegrade ab Stufe 3

sieht das Reifegradmodell vor, dass das Forschungsportal die Analyse der Datenbasis mittels OLAP-Funktionalität unterstützt. Dementsprechend ergibt sich eine Überschneidung der Beurteilungskriterien des Reifegradmodells für Forschungsportale mit den BI-spezifischen Reifegradmodellen.

5 Fazit

Angesichts der Vielzahl an Modellierungstechniken, die für die fachkonzeptionelle Spezifikation Data-Warehouse-basierter OLAP-Systeme bereits vorgeschlagen worden sind, erscheint es nahe liegend, daran zu zweifeln, dass die (Weiter)entwicklung dieser Modellierungstechniken noch immer eine viel versprechende Forschungsperspektive darstellt. Die fachkonzeptionelle Modellierung ist eine aufwändige Aufgabe. In der Praxis wird diese häufig vernachlässigt, um vermeintlich Aufwand und Zeit zu sparen. Die Einsicht in den Nutzen der fachkonzeptionellen Modellierung steigt, wenn die Modelle in mehreren Kontexten verwendet werden. Am Beispiel des forschungportalbasierten Controlling wurden sechs ausgewählte Themenbereiche vorgestellt, die zeigen, dass die fachkonzeptionelle Modellierung durchaus weiterhin interessante Perspektiven für konstruktionsorientierte Forschung bietet. Die diskutierten Themenbereiche erheben dabei keinen Anspruch auf Vollständigkeit, sondern repräsentieren priorisierte Arbeitsbereiche, die im Rahmen der Fortsetzung der Entwicklung der Modellierungssprache H2 for Reporting gewählt wurden.

Danksagung

Dieser Beitrag wurde ermöglicht durch die Förderung des Projektes "Flexible Informationssystemarchitekturen für hybride Wertschöpfungsnetzwerke" (FlexNet) durch das BMBF (Förderkennzeichen 01FD0629) im Rahmen des Programms "Innovationen mit Dienstleistungen". Wir danken an dieser Stelle auch dem Projektträger Deutsches Zentrum für Luft- und Raumfahrt (DLR) für die Betreuung.

6 Literatur

Becker, J.; Beverungen, D.; Knackstedt, R.; Glauner, C.; Stypmann, M.; Rosenkranz, C.; Schmitt, R.; Hatfield, S.; Schmitz, G.; Eberhardt, S.; Dietz, M.; Thomas, O.; Walter, P.; Lönngrén, H.; Leimeister, J.M. (2008): Ein Plädoyer für die Entwicklung eines multidimensionalen Ordnungsrahmens zur hybriden Wertschöpfung. In: Arbeitsberichte des Instituts für Wirtschaftsinformatik der Universität Münster, Nr. 119. Editors: Becker, J.; Grob, H.L.; Klein, S.; Kuchen, H.; Müller-Funk, U.; Vossen, G. 2008.

- Becker, J.; Knackstedt, R. (2004): Referenzmodellierung im Data-Warehousing – State-of-the-Art und konfigurative Ansätze für die Fachkonzeption. In: *Wirtschaftsinformatik*, 46 (2004) 1, S. 39-49.
- Becker, J.; Knackstedt, R.; Pöppelbuß, J. (2009): Entwicklung von Reifegradmodellen für das IT-Management – Vorgehensmodell und praktische Anwendung. In: *Wirtschaftsinformatik*, 51 (2009) 3, S. 249-260.
- Böhnlein, M. (2001): *Konstruktion semantischer Data-Warehouse-Schemata*, 1. Auflage, Deutscher Universitäts-Verlag, Wiesbaden, 2001.
- Booch, G.; Rumbaugh, J.; Jacobson, I. (1999): *UML – The Unified Modeling Language Reference Manual*. Reading 1999.
- Chamoni, P.; Gluchowski, P. (2004): Integrationstrends bei Business-Intelligence-Systemen. *Wirtschaftsinformatik*, 46 (2004) 2, S. 119-128.
- Chen, P.P.-S. (1976): The Entity-Relationship-Model – Toward a Unified View of Data. *ACM Transactions on Database Systems*, 1 (1976) 1, S. 9-36.
- Delfmann, P.; Knackstedt, R.; Niehaves, B.; Winkelmann, A. (2008): Berichtswesenvereinfachung: Perspektiven für eine gestaltungsorientierte Wirtschaftsinformatikforschung. In: vom Brocke, J.; Becker, J. (Hrsg.): *Einfachheit in Wirtschaftsinformatik und Controlling*. Festschrift für Heinz Lothar Grob. München 2008. S. 189-206.
- Frank, U., Strecker, S. (2007). Open Reference Models – Community-driven Collaboration to Promote Development and Dissemination of Reference Models. *Enterprise Modelling and Information Systems Architectures* 2 (2), pp. 32-41.
- Garfield, E. (1955): Citation indexes for science: a new dimension in documentation through association of ideas. *Science* Vol. 122, S. 108–111.
- Hevner, A. R.; March, S. T.; Park, J.; Ram, S.: Design science in information systems research. In: *MIS Quarterly*, 28 (2004) 1, S. 75-105.
- Holten, R. (2003): Specification of management views in information warehouse projects. *Information Systems*, 28 (2003) 7, S. 709-751.
- Kemper H-G, Mehanna W, Unger C (2006) *Business Intelligence – Grundlagen und Praktische Anwendungen: Eine Einführung in die IT-basierte Managementunterstützung*. 2. Aufl. Wiesbaden.
- Knackstedt, R. (2006): *Fachkonzeptionelle Referenzmodellierung einer Managementunterstützung mit quantitativen und qualitativen Daten. Methodische Konzepte zur Konstruktion und Anwendung*. Berlin 2006.
- Knackstedt, R.; Deinert, M.-O.; Becker, J. (2009): Modellvergleich mittels Clusteranalyse am Beispiel einer automatisierten Ähnlichkeitsanalyse für OLAP-Berichtsspezifikationen. In: Hansen, H. R.; Karagiannis, D.; Fill, H.-G. (Hrsg.): *Business Services: Konzepte, Technologien, Anwendungen*. 9. Internationale Tagung Wirtschaftsinformatik. 25. – 27. Februar 2009, Wien. Band 2, Wien, S. 339-348.
- Knackstedt, R.; List, L.; Stein, A.; Becker, J.; Barth, I. (2009): Towards A Reference Model for Online Research Maps. In: *Proceedings of the 17th European Conference on Information Systems (ECIS 2009)*. Verona, Italy, 2009.
- Knackstedt, R.; Seidel, S.; Janiesch, C. (2006): Konfigurative Referenzmodellierung zur Fachkonzeption von Data-Warehouse-Systemen mit dem H2-Toolset. In: J. Schelp, R. Winter, U. Frank, B. Rieger, K. Turowski (Hrsg.): *Integration, Informationslogistik und Architektur*. DW2006, 21.-22. Sept. 2006, Friedrichshafen. *Lecture Notes in Informatics*. Bonn 2006, S. 61-81.

Scharl, A. (1997): Referenzmodellierung kommerzieller Masseninformati­onssysteme. Idealtypische Gestaltung von Informationsangeboten im World Wide Web am Beispiel der Branche Informationstechnik, Frankfurt am Main 1997.

Wedeking, H. (1981): Datenbanksysteme I – Eine konstruktive Einführung in die Datenverarbeitung in Wirtschaft und Verwaltung. 2. Aufl., Mannheim u. a. 1981.