

Domänensemantik-orientierte Integration heterogener konzeptueller Modelle

Janina Fengel, Michael Rebstock

Fachbereich Wirtschaft
Hochschule Darmstadt
Haardtring 100
64295 Darmstadt
vorname.nachname@h-da.de

Abstract: In Integrationsprojekten ist eine der Herausforderungen das Vergleichen und Verknüpfen heterogener Modelle. Dies erfordert nicht nur einen strukturellen, sondern insbesondere einen semantischen Abgleich. In diesem Beitrag stellen wir unsere dafür entwickelte Methode vor. Grundlage ist ein Reengineering von Modellen zu webbasierte Ontologien. Für die dafür erforderliche Modelldekomposition wurde eine unterstützende Brückenontologie entwickelt, die hier vorgestellt wird. Zur Unterstützung der anschließenden semantischen Integration nutzen wir Ontology-Matching-Verfahren zur automatisierten Ermittlung von Entsprechungen zwischen Modellen und ihren Elementen bezüglich der zu ihrer Bezeichnung verwendeten Domänensprache. Die prototypische Implementierung eines entsprechenden Systems zeigt die Machbarkeit unserer Methode.

1 Motivation

Eine der Kernaufgaben des Managements in Unternehmen ist die Gestaltung und laufende Anpassung des geschäftlichen Geschehens an aktuelle Erfordernisse und die Allokation der dafür erforderlichen Ressourcen. Grundlage der Gestaltung von Unternehmen und B2B-Kooperationen ist ihre umfassende, korrekte Beschreibung. Die Beschreibung geschäftlicher Abläufe und der dabei genutzten Informationen und Ressourcen erfolgt zumeist mit Hilfe von Modellen, insbesondere zur Darstellung der Unterstützung des geschäftlichen Geschehens durch IT. Die Modellierung von Organisationsstrukturen, Daten und Prozessen zur Beschreibung und Gestaltung betrieblichen Geschehens hat daher stark an Bedeutung gewonnen. Gegenwärtig sind verschiedene Modellierungssprachen in Verwendung. Zumeist sind sie fokussiert auf bestimmte Aspekte und daher anwendbar zur Erstellung unverbunden nutzbarer Modelle. Ansätze im Bereich der Unternehmensarchitektur thematisieren die Notwendigkeit der Erfassung von Zusammenhängen [La09; Fr04]. Architekturframeworks wie ARIS oder das Zachmann-Framework bieten einen Rahmen zur Integration von Modellen unterschiedlicher Sichten, wie beispielsweise für die Beschreibung der Beziehung zwischen Daten und Prozessen [SZ92; Sc96].

Allerdings wird dabei keine Möglichkeit geboten, Modelle derselben Sicht, ausgedrückt in verschiedenen Modellierungssprachen, miteinander und damit auch zu Modellen anderer Sichten in Bezug zu bringen. Indes ist zur Herstellung durchgängiger Prozessabwicklung eine konsistente Verknüpfung und Verwaltung von Modellen unabdingbarer Bestandteil. Eine der dabei anfallenden Aufgaben ist der Abgleich und die Zusammenführung bereits existierender, in Nutzung befindlicher Modelle. Dieser Bedarf entsteht in Geschäftsintegrationsprojekten wie im Falle innerbetrieblicher Geschäftsprozessoptimierungen oder Business Reengineering und in unternehmensübergreifenden Zusammenhängen bei Fusionen, der Umsetzung von Supply-Chain-Management oder der Gestaltung von B2B-Kooperationen. Für eine umfassende Geschäftsanalyse sind dabei in jedem dieser Fälle konzeptuelle Modelle in Bezug zu setzen, sowohl im Sinne einer vertikalen Integration von Modellen verschiedener Sichten als auch horizontaler Integration von modellsprachlich unterschiedlichen Modellen der gleichen Sicht. Zur Zusammenführung sind diese zu vergleichen bezüglich der Inhaltsbedeutung ihrer Elemente und ihrer Struktur, wobei erst der Abgleich der Domänensprache anschließend strukturelle Vergleiche erlaubt [SM07].

Das Vergleichen und Verknüpfen heterogener Modelle ist indes eine nicht-triviale Aufgabe. Prinzipiell explizieren Modelle aller Art unterschiedliche Sichten auf eine Organisation und damit Wissen. Allerdings unterscheiden sich häufig selbst Modelle gleichen Typs semantisch [BP08]. Dabei tritt semantische Heterogenität nicht nur im Bereich der Modellierungssprachen auf, sondern typischerweise bei der Auswahl der natürlichsprachlichen Fachbegriffe, die zur Benennung der Modellelemente verwendet werden [TF07]. Besondere die frei wählbare Fachterminologie behindert eine automatisierte Unterstützung semantischer Abgleiche und damit der zugrunde liegenden Daten und Prozesse. Die Modellelementbezeichnungen spiegeln die unternehmensspezifische Geschäftssprache wider. Existiert kein verbindlich definiertes Vokabular oder Regeln bezüglich deren Anwendung, können sich Modelle darin erheblich unterscheiden. Liegen gar Namenskonflikte vor, sind Modelle weder manuell noch automatisiert direkt vergleich- und damit integrierbar [BRS96; TF06]. Daher gilt es, semantische Ungleichheiten aufzulösen, um die Aussagen von Modellen inhaltlich in Bezug bringen und abgleichen zu können, besonders bei unterschiedlicher Herkunft der Modelle. Dabei handelt es sich um eine verknüpfende Modellintegration, die Modelle verbindet, ohne sie zu verschmelzen.

Zur Reduktion dieser Arbeitslast wird hier die Anwendung von Semantic-Web-Technologien vorgeschlagen und beschrieben. Es folgt die Vorstellung der Vorgehensweise zur Ontologisierung von Modellen und der dazu benötigten Brückenontologie UMCO in Abschnitt zwei. In Abschnitt drei wird die Nutzung von Ontology-Matching-Verfahren geschildert, sodass die Anwendung unserer Methode in Abschnitt vier gezeigt werden kann. Der Beitrag endet mit der Vorstellung und der Verbindung zu verwandten Arbeiten und einer kurzen Diskussion mit Ausblick auf zukünftige Arbeiten.

2 Ontologisierung von Modellen

Die Allgegenwärtigkeit des World-Wide-Web als globale Infrastruktur hat zur hohen Akzeptanz webbasierter Unterstützung elektronischer Geschäftsabwicklung beigetragen. Die Entwicklung der Idee des Semantic-Web und seiner spezifischen Technologien bietet nun weiterführend die Möglichkeit der Nutzung webbasierter Ontologien als Mittel zur Wissensstrukturierung und Herstellung semantischer Interoperabilität. Das Prinzip der Annotation von Information mit Metadaten erlaubt die Repräsentation von Wissen in strukturierter, maschinenzugänglicher Form basierend auf Internettechnologien, lesbar sowohl für Maschinen als auch von Menschen, nutzbar zum weiteren Ausbau der Informationsverarbeitung [SBH06].

Da die Entwicklung einer Ontologie ein komplexes Unterfangen sein kann, schlagen wir zur Unterstützung die Nutzung semantischer Technologien vor, mittels derer sich die Inhaltsbedeutung von Information formal erfassen lässt zwecks Herstellung von Maschinenverarbeitbarkeit. Durch Erschließung und Wiederverwendung des in Modellen aller Art enthaltenen Domänenwissens lassen sich Ontologien der Modelle ableiten. Diesen Vorgang der Ermittlung und semantischen Formalisierung nennen wir Ontologisierung. Anschließend können diese Ontologien mit automatisierter Unterstützung semantisch integriert und dadurch verknüpft werden. Über die Zeit kann so eine zusammenhängende, terminologische Domänen- oder auch spezifische Unternehmensontologie aus bereits vorliegenden Modellen automatisiert unterstützt abgeleitet werden. Durch die dabei bottom-up erfolgende Erfassung der verwendeten Domänenterminologie kann diese wiederum zur Auflösung semantischer Ambiguität in Modellen eingesetzt werden.

2.1 Dekomposition anhand der Sprachräume

Eine Wiederverwendung von Modellen und Konvertierung in Ontologien erlaubt eine Verwertung existierender Modelle, während sie weiterhin unverändert aktiver Nutzung zur Verfügung stehen. Ein Reengineering dieser nicht-ontologischen Ressourcen bietet die Möglichkeit die Bedeutung der Modellaussage semantisch zu formalisieren. Dabei wird genutzt, dass bei der Erstellung von Modellen sowohl Kenntnis der Modellierungssprache als auch Kenntnis der Fachsprache erforderlich ist. Dies wurde bereits für die Modellierung von EPK gezeigt [TF06], lässt sich aber auf Modelle generell übertragen. Dabei enthalten Modelle Fakten aus zwei Wissensbereichen, wie in Abbildung 1 anhand eines Ausschnitts einer EPK namens „Travel Reservation“ aus der Domäne Touristik gezeigt.

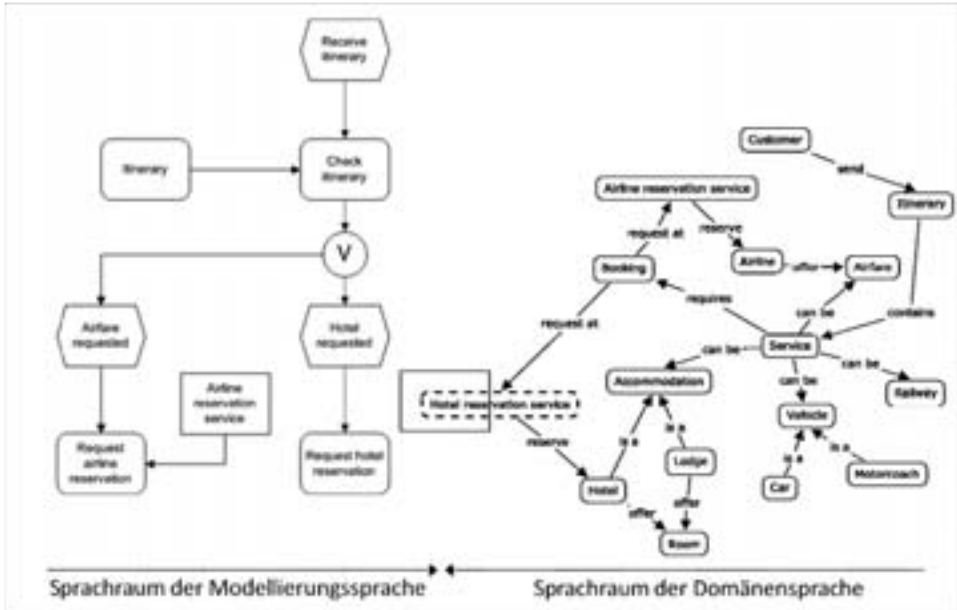


Abbildung 1: Sprachräume von Modellen

Aus dem Sprachraum der Domänensprache werden Konzepte zur Benennung von Modellen und ihren Elementen genutzt, während die Konzepte der Modellierungssprache zur Beschreibung im Sinne einer Typisierung und Anordnung dieser Konzepte herangezogen werden. In Umkehrung dieses Vorgangs lassen sich Modelle zerlegen, um die jeweils verwendeten Konzepte der Sprachräume zu extrahieren und in Form semantischer Modelle zu erfassen [FR09]. Modelle können so in eine Ontologie mit Modellname und Modelltyp und eine mit Modellelementnamen und Modellelementtypen zerlegt werden.

2.2 Konvertierung in Ontologien

Modelle repräsentieren in der Regel abgestimmtes Fachwissen. Durch Dekomposition und Überführung in Ontologien wird Maschinenzugang zu diesem Wissen hergestellt. Ontologien im informatiktechnischem Sinne sind Artefakte und können als konzeptuelle Schemata verstanden werden, ähnlich Datenbankschemata [AF05]. Die Elemente einer Ontologie sind Klassen oder Konzepte, die eine Menge individueller Objekte repräsentieren, mit ihren Instanzen, die diese individuellen Objekte repräsentieren sowie Relationen und Axiome zur Beschreibung der Beziehungen der enthaltenen Konzepte. Im Prinzip sind Ontologien Sammlungen von Definitionen von Elementen und ihren Beziehungen und enthalten ein abgestimmtes Vokabular [DOS03]. Bei der Entwicklung wird in der Regel Konsens hergestellt und eine eindeutige Inhaltsbedeutung der enthaltenen Begriffe definiert [SSS01].

Im Zuge einer Konvertierung von Modellen in Ontologien lässt sich der bei der Modellerstellung bereits erreichte Konsens wieder ableiten und für jedes Modell in Form einer spezifischen Ontologie des Modells formalisieren.

Eine Ontologie bietet die Möglichkeit Wissen eindeutig darzustellen und damit zu teilen und wiederzuverwenden, denn sie beschreibt formal eine vereinfachte abstrahierte Sicht auf einen bestimmten Weltausschnitt [Gr93]. Dabei können Ontologien unterschiedlich stark formalisiert sein. Zur Repräsentation webbasierter Ontologien gibt es verschiedene Sprachen. Für einfache, so genannte leichtgewichtige, schwach formalisierte Ontologien bietet sich die XML-basierte Beschreibungssprache RDF-Schema an. Für stärker formalisierte, schwergewichtigere Ontologien eignet sich die Web Ontology Language OWL. Diese vom World Wide Web Consortium publizierte formale Ontologiebeschreibungssprache kann in drei verschiedenen Varianten genutzt werden [SH05]. Eine ausreichend formale Beschreibung ermöglicht automatisierte Weiterverarbeitung, nicht nur durch Ontology-Matching, wie nachfolgend beschrieben, sondern auch durch logisches Schließen [BP06]. Daher wird hier die beschreibungslogikbasierte Sprachvariante OWL DL genutzt, da diese entscheidbar ist [SWM04]. Konzepte, Relationen und Instanzen können dadurch unabhängig voneinander beschrieben werden [Br04]. Bei Verwendung von OWL DL ist es weiterhin möglich, Inferenzregeln basierend auf den in der Ontologie enthaltenen Relationen zu formulieren [SK06]. Somit kann durch logisches Schließen aus vorhandenem Wissen neues abgeleitet [AH08] und eine durch Modellkonvertierung erstellte Modellontologie weiterentwickelt werden.

Eine Ontologie in OWL DL beinhaltet zwei Bereiche. Die TBox, die terminological box, enthält die Definitionen der Domänenkonzepte mit ihren Eigenschaften und Beziehungen. Die ABox, die assertional box, enthält die Fakten und Assertionen bezüglich der beschreibenden Eigenschaften von Konzepten und Instanzen und ihrer Beziehungen. Abbildung 2 zeigt den Bezug zwischen konzeptuellen Modellen und Ontologien in OWL DL über die verschiedenen Modellierungsebenen.

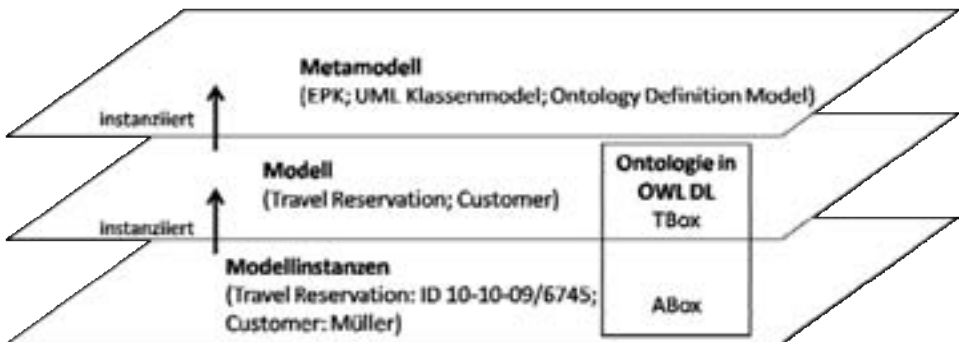


Abbildung 2: Modellebenen konzeptueller Modelle und Ontologien

Bei der Dekomposition von Modellen werden die Modellelemente in zwei Ontologien in OWL DL konvertiert, die Modellontologie und die Modelltypontologie. Zusammen beschreiben sie das Modell mit seinem Namen und Modelltyp sowie die Modellelemente mit ihren Namen und ihrem Modellelementtyp. Dabei werden sowohl die Benennungen als auch die Typen als Konzepte in die jeweilige Ontologie aufgenommen. Für den oben gezeigten Ausschnitt der EPK „Travel Reservation“ entstehen aus dem Ereignis „Receive itinerary“ zwei Ontologiekonzepte, nämlich das Konzept „Receive itinerary“ und das Konzept „Ereignis“. Auf diese Weise kann in die entstehende Modellontologie, die die Konzepte der Domänensprache aufnimmt, auch zukünftig anfallende Instanzinformation aufgenommen werden. Im genannten Beispiel könnten dies Kundendaten oder Informationen aus einem Workflowmanagementsystem sein. Bei der Konvertierung werden alle Modellnamen und Modellelementbezeichnungen ohne weitere Verarbeitung transferiert. Auf diese Weise können vollständige Ausdrücke zur Weiterverarbeitung übernommen werden, denn das Fachwissen bei der Modellierung zeigt sich oft erst in der Kombination von Worten zu häufig genutzten Formulierungen.

2.3 Entwicklung der Brückenontologie UMCO

Zur Abbildung der domänenspezifischen Aussage von Modellen werden diese wie oben beschrieben zerlegt. Dadurch werden die Konzepte der Domänensprache getrennt von den Konstrukten der Modellierungssprache, die sie typisieren. Die Verbindung einer solchen Typinformation mit dem jeweiligen Domänensprachenkonzept bleibt allerdings erhalten. Zur Darstellung dessen wird hier die Idee der Indexierung in Form von Topic Maps aus dem Bibliothekswesen übernommen. Das Domänenkonzept wird über die Beziehung „occurs as“ mit dem Typkonzept verbunden. Abbildung 3 zeigt das Metamodell eines dekomponierten Modells.

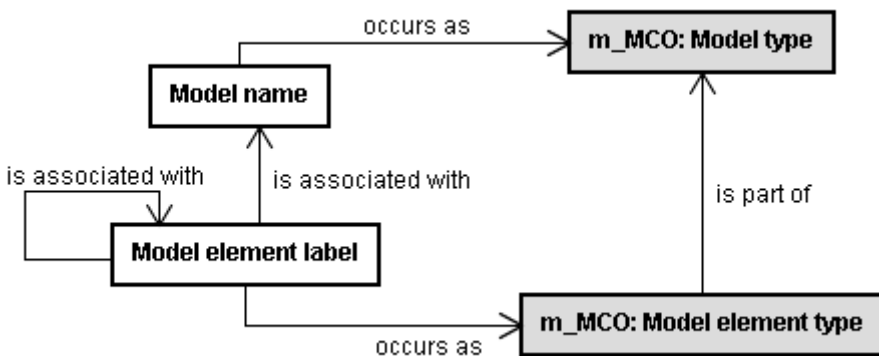


Abbildung 3: Metamodell eines dekomponierten Modells

Die Modellontologie, links im Bild in weiß gezeigt, umfasst das Domänenwissen ausgedrückt in den Elementbezeichnungen in natürlicher Sprache als Ontologiekonzepte und die bestehenden Beziehungen als Relationen, die je nach Bedarf mit Restriktionen versehen sind. Die Modellontologie importiert die Modelltypontologie, wie im Bild rechts in grau dargestellt. Für verschiedene Modellierungssprachen wurde von uns jeweils eine spezifische Modeling Concepts Ontology (MCO) entwickelt. Aktuell liegen MCOs vor für EPK, UML-Aktivitätsmodelle, BPMN, ERM, UML-Klassenmodelle und UML-Use-Case-Modelle, wobei für weitere beliebige Modellierungssprachen entsprechende MCOs entwickelbar sind. Jede MCO enthält als Konzepte die Modellierungssprachkonstrukte und damit Teile des jeweiligen Metamodells, die über Teil-Ganzes-Beziehungen verknüpft sind. Das Fachwissen, das in der Formulierung der logischen Beziehungen zwischen den Modellelementen im Modell enthalten ist, wird zusammen mit den Domänenkonzepten in die Modellontologie überführt. Bei dem gezeigten Beispiel wird entsprechend das Ereignis „Receive itinerary“ als Klasse der Modellontologie über die Relation „occurs as“ mit der Klasse „Event“ der EPC_MCO verbunden.

Zur Verknüpfung der entstehenden Modellontologien können über die MCO Referenzen zwischen Modellen desselben Typs in derselben Modellierungssprache direkt hergestellt werden. Um nun auch Verknüpfungen mit Modellen anderen Typs oder anderer Art zu ermöglichen, wurde die Unifying Modeling Concepts Ontology (UMCO) entwickelt. Sie verknüpft die verschiedenen MCOs und überbrückt so verschiedene Sprachräume, denn sie enthält Konzepte, die Modellierungskonzepte verschiedener Modellierungssprachen mit ähnlicher Intention zusammenfassen. Beispielsweise enthalten Geschäftsprozessmodellierungssprachen das Konzept von Aktivitäten, genannt Aktivität, Funktion, Aufgabe oder Aktion, begonnen und beendet durch Ereignisse und verbunden über Kontrollflüsse, beschrieben mit Hilfe logischer Konnektoren [MK07]. Im Detail ist die Semantik dieser Modellierungssprachen nicht identisch, und Modelle können nicht ohne weitere Bearbeitung verlustfrei von einer Sprache in eine andere transformiert werden [KQC07]. Allerdings sind die grundlegenden Bedeutungen der enthaltenen Sprachkonstrukte vergleichbar. Diese Beobachtung gilt ebenso für die Sprachen zur Modellierung statischer Modelle. Konzeptuelle Datenmodelle können als UML Klassenmodelle, Entity-Relationship-Modelle oder direkt als semantische Modelle in OWL repräsentiert werden. Dabei entsprechen die Bedeutungen einer Klasse bzw. Entität den Klassen einer Ontologie, während die Attribute und Beziehungen den Relationen bzw. Properties in den meisten Ontologiesprachen entsprechen [UG04]. Daher sind UML Klassenmodelle auch direkt zur Ontologiemodellierung nutzbar [GDD04; BVE04].

Zum Zweck der semantischen Integration wurden nun auf diesen Überlegungen basierend in der UMCO generalisierende Modellierungskonzepte definiert und zu den entsprechenden Modellierungskonzepten der MCOs gleichgesetzt. Dies bietet die Möglichkeit logischen Schließens über verknüpfte Modelle hinweg. Abbildung 4 zeigt das Metamodell der UMCO in Erweiterung des in Abbildung 3 vorgestellten Metamodells.

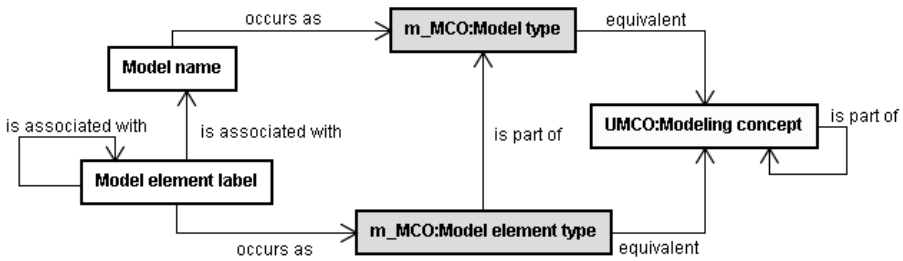


Abbildung 4: Metamodell eines dekomponierten Modells mit Beziehungen zur UMCO

Die UMCO enthält wie die MCOs neben den über Teil-Ganzes-Beziehungen verbundenen vereinigenden Modellierungskonzepte keine weiteren definierenden Beziehungen. Das Konzept der UMCO ist generisch und beliebig erweiterbar mit zusätzlichen MCOs. Die Entwicklung von MCOs für Prozess-, Daten-, Organisations- oder auch technischen Modellen erlaubt die Verknüpfung von Domänenwissen nicht nur innerhalb des Sprachraums einer einzelnen Modellierungssprache, sondern sprachraumübergreifend mittels der Konzepte der UMCO auch mit Domänenwissen ausgedrückt in anderen Modellierungssprachen. Beispielsweise sind die Konzepte einer EPK mit dem eines UML Aktivitätsmodells und eines Modells in BPMN gleichgesetzt zum Konzept `UMCO:Process`. `EPK_MCO:Function` in der EPK, `UML_AM_MCO>Action` im UML Aktivitätsmodell und `BPMN:Task` im BPMN-Modell entsprechen dem Konzept `UMCO:Activity` in der UMCO. Auf dieselbe Weise lassen sich ebenso Konzepte einer Modellontologie zu Konzepten in Modellontologien entstanden aus anderen Modellarten in Bezug bringen, z.B. lassen sich in Prozessmodellen enthaltene Konstrukte zur Beschreibung von Ressourcen wie Dokumente oder Teilnehmer zu den sie beschreibenden Daten- oder Organisationsmodellen in Bezug bringen. So erfüllt die UMCO eine Brückenfunktion zur Verbindung der Modellontologien anhand der Typinformation der eigentlichen Domänenkonzepte.

3 Semantische Modellintegration

Um nun die entstandenen Modellontologien, die die Domänensemantik der konvertierten Modelle enthalten, automatisiert miteinander in Bezug zu bringen, können Ontology-Matching-Verfahren angewendet werden. Für ansonsten manuell auszuführende Modellabgleiche kann so automatisierte Unterstützung geboten werden und die Modellelemente, die die Domänensemantik widerspiegeln, können so unabhängig von der ursprünglich genutzten Modellierungssprache verglichen werden.

3.1 Anwendung von Ontology-Matching-Verfahren

Ontology-Matching sind Verfahren zur Ermittlung semantischer Entsprechungen zwischen Elementen zweier Ontologien. Es werden Paare von Entitäten gesucht, die vermutlich dieselbe Inhaltsbedeutung haben [EMS08]. Die Ontologien werden dabei abgeglichen, um semantische Korrespondenzen zu entdecken. Diese können als semantische Mappings verstanden werden. Es sind Aussagen der Form "A in Ontologie X entspricht B in Ontologie Y". Die gefundenen Korrespondenzen beschreiben Äquivalenz oder Entsprechungen. Für weiterführende Nutzung können sie persistiert werden [ES07]. Dadurch können die verknüpften Ontologien bestehen bleiben, ohne zusammengeführt werden zu müssen. Dies ist besonders im Hinblick darauf nützlich, dass die zugrunde liegenden Modelle nicht ohne weiteres geändert werden können, sondern aktiv genutzt werden. Bewahrte Korrespondenzen bieten stattdessen die Möglichkeit einer virtuellen semantischen Integration.

Die Verfahren zum Ontology-Matching lassen sich in element- und strukturprüfende unterscheiden [ES07]. Elementprüfende Verfahren analysieren die Elemente einer Ontologie ohne Betrachtung der Beziehungen, die sie miteinander verknüpfen. Es werden Namen bzw. Zeichenketten und Mengen davon verglichen, wobei linguistische Ressourcen als Hintergrundinformation eingebunden werden können. Editierdistanz-basierte Methoden wie die Levenshtein-Distanz messen die Ähnlichkeit zweier Zeichenketten ausgedrückt als Anzahl der benötigten Änderungsoperationen zur Umwandlung des einen zum anderen Ausdruck. Berechnungsverfahren basierend auf der Jaro-Metrik messen die Anzahl und Anordnung gleicher Zeichen zweier Zeichenketten. Tokenbasierte Methoden führen diese Art von Messungen für Mengen von Zeichenketten aus [CRF03]. Strukturprüfende Verfahren dagegen analysieren die Struktur einer Ontologie und betrachten die Beziehungen der Elemente. In der Regel wird dazu das Vorliegen hierarchischer Beziehungen vorausgesetzt, da solche zumeist das Rückgrat einer Ontologie bilden [ES07].

Erste empirische Evaluationen von Matchings konvertierter Testmodelle wurden durchgeführt, wobei die Qualität der von verschiedenen Matchern automatisiert ermittelten Mappings durch Vergleich mit einem von Domänenexperten erstellten Referenzmapping gemessen wurde. Dabei zeigte sich, dass die Konvertierung von Modellen in Modellontologien wie beschrieben für Prozessmodelle zu Ontologien führt, die keine hierarchischen oder mereologischen Beziehungen enthalten, sodass nur elementprüfende Verfahren Ergebnisse liefern können. Weiterhin zeigte sich, dass die in Prozessmodellen übliche Benennung von Elementen mit mehreren Termen in einem Ausdruck durch Name-Matching-Verfahren allein zu qualitativ minderwertige Ergebnissen führt. Die Zerlegung solcher Multiterm-Ausdrücke und daran anschließendes tokenbasiertes Name-Matching führt dagegen zu höherwertigen Resultaten. Weitere Qualitätssteigerung ist durch zusätzliche Nutzung von WordNet als Hintergrundressource zu erreichen [Mi09]. Es bleibt zu eruieren, wie spezielle Glossare oder Lexika mit hinzugezogen werden können.

Die Konvertierung von Klassen- oder ER-Modellen bringt in den meisten Fällen Ontologien hervor, die auch mit strukturprüfenden Verfahren untersucht werden können. Unsere empirischen Evaluationen haben gezeigt, dass hier auf der Jaro-Winkler-Metrik basierende Vergleiche in Kombination mit Strukturvergleichen Ergebnisse erbringen, da Datenmodelle in der Regel Elementbezeichnungen bestehend aus ein bis zwei Termen enthalten.

3.2 Ermittlung von Korrespondenzen anhand der Domänensemantik

Da das Domänenwissen nicht als Instanz der jeweiligen Modellierungssprachkonstrukte konvertiert wird, sondern als Ontologieklassen, wird verhindert, dass Matching Entsprechungen der Modellierungssprachkonstrukte untereinander liefert. Diese sind dank der Nutzung der MCOs und der UMCO direkt verfügbar, ohne Aufwand für Matcher zu produzieren. Weiterhin werden durch das Nichtvorhandensein einer solchen unerwünschten Struktur die Modelle ausschließlich anhand der Domänensprache abgeglichen. Abbildung 5 zeigt beispielhaft die Verknüpfung von Modellen in Form einer Ontologie, dargestellt als semantisches Netz. Es handelt sich um den bereits bekannten Ausschnitt aus der EPK „Travel Reservation“ mit einem Ausschnitt aus einem UML Aktivitätsmodells namens „Travel Booking“, die abgeglichen und miteinander in Bezug gebracht wurden.

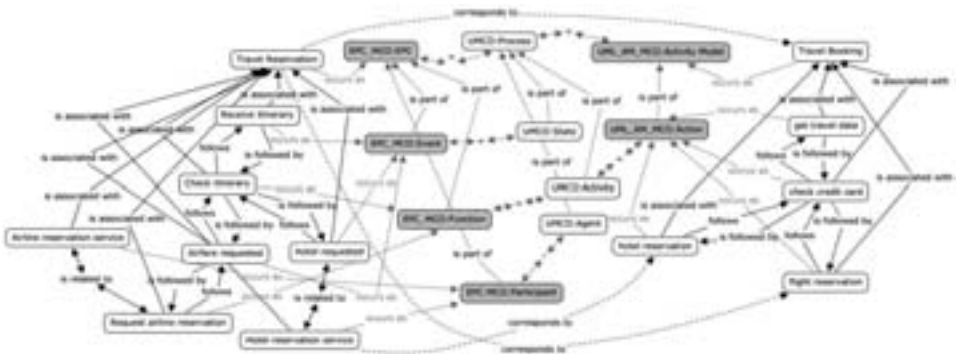


Abbildung 5: Semantisches Netz von Ausschnitten verknüpfter heterogener Modelle

Die konvertierten Modelle entstammen derselben Anwendungsdomäne, nutzen aber unterschiedliche Terminologie und Modellierungssprachen. Die Modellontologien, dargestellt in weiß, bilden die Domänensprache ab. Die Elemente referenzieren ihren Modellelementtyp, enthalten in der in grau dargestellten MCO. Die MCOs sind über die UMCO, hier in hell, verbunden durch die Definition der Entsprechungen der Elemente. Die durch Ontology-Matching ermittelten semantischen Korrespondenzen, im Bild durch gestrichelte Beziehungen dargestellt, verknüpfen die Modellontologien untereinander und damit die Konzepte der Domäne.

Matchings können nach Bedarf schrittweise vorgenommen werden zwischen ontologisierten Modellen, unabhängig welchen Typs und in welcher Sprache. Über die Zeit wird mit zunehmender Zahl an Modellen die Tendenz zu häufig genutzten Begriffen sichtbar. Langfristig kann so eine erste einfache Unternehmensontologie aus den verknüpften Modellen abgeleitet werden. Diese kann genutzt werden, um Suchen nach Korrespondenz auszuführen. Es können Vergleiche zwischen Modellen und Abgleiche mit Referenzmodellen, Standards oder bereits vorliegenden Unternehmensglossaren vorgenommen werden. Eine generelle Suche über alle verknüpften Modelle nach Okkurrenz liefert Information zur Häufigkeit des Vorkommens eines Begriffs sowie dem Zusammenhang und Auftreten in bestimmten Modellen.

3.3 Vernetzung von Modellen

In Realisierung unserer Methode wird unser MODI-System entwickelt. Es ist in Java implementiert und über ein Web-Service-Interface nutzbar. Es besteht aus einem Kernsystem, an das über Adapter beliebige Ontology-Matching-Tools angekoppelt werden können. Zusätzlich ist ein Web-Service verfügbar, mit dem Modellkonvertierungen durchgeführt werden können. Viele Modellierungswerkzeuge bieten die Möglichkeit Modelle in XMI-Format zu exportieren, sodass eine Konvertierung mit Hilfe von XSLT vorgenommen werden kann. Allerdings hat sich gezeigt, dass aufgrund unterschiedlicher Verwendungen von XMI für jeden Modelltyp abhängig vom Werkzeug ein spezifisches XSLT-Sheet erforderlich ist.

Der Fokus der Entwicklungsarbeiten liegt auf der Realisierung von Matching und Mapping. Die Anwendung von Ontology-Matching-Verfahren liefert automatisiert eine initiale Menge semantischer Korrespondenzen. Diese dient als Ausgangsbasis für nutzergetriebene Weiterentwicklung. Nutzer können neue Korrespondenzen und neue Modelle hinzufügen. Unser System bietet verschiedene Abfragemöglichkeiten. Ein Nutzer kann gezielt zwei ausgewählte Modelle vergleichen oder generell nach vorhandenen Korrespondenzen zu einem bestimmten Begriff suchen. Das System liefert eine Liste mit Vorschlägen möglicher Entsprechungen, aus der der Nutzer auswählen kann. Obwohl die Anwendung von Matching-Werkzeugen die manuell zu erbringende Arbeit verringert, sind die Ergebnisse nicht perfekt, sondern können mehrdeutig oder inkorrekt sein. Zur Disambiguierung oder Korrektur ist menschliches Eingreifen erforderlich [Zh04]. Daher bietet unser System aktive Nutzerpartizipation und ermöglicht das Hinzufügen und Editieren von Korrespondenzen. Abbildung 6 zeigt die Vorschlagsliste für den Vergleich der beiden Beispielmodelle in textueller Form, der als Ausschnitt in Abbildung 5 zu sehen ist.

Search Correspondences New Correspondences View Model Ontology Manage Models

Filter by: Source Model Show Element Type Show Linking Type

Source Expression	Source Model	Destination Expression	Destination Model	Coefficient	Content	Terms
<input type="checkbox"/> TravelReservation	TravelReservation	TravelBooking	TravelBooking	10%	Travel	TravelReservation, Booking
<input type="checkbox"/> HotelCustomer	TravelReservation	HotelCustomer	TravelBooking	80%		HotelCustomer, Hotel
<input type="checkbox"/> TravelReservation	TravelReservation	RightReservation	TravelBooking	80%		TravelReservation, Right
<input type="checkbox"/> DataCreditCard	TravelReservation	CheckCreditCard	TravelBooking	80%		DataCreditCard, Check
<input type="checkbox"/> DataCreditCard	TravelReservation	HotelCreditCard	TravelBooking	70%		DataCreditCard, Credit
<input type="checkbox"/> HotelReservationSearch	TravelReservation	HotelReservation	TravelBooking	70%		HotelReservation, Search
<input type="checkbox"/> ReportHotelReservation	TravelReservation	HotelReservation	TravelBooking	50%		ReportHotelReservation

Found 7 results
12.06.2008

Abbildung 6: Vorschlagsliste semantischer Korrespondenzen eines Modellvergleichs

Die Kombination der Modellontologien mit ihren MCOs bietet weiterhin die Möglichkeit der Filterung anhand der Typinformation und damit der Suche nach oder in bestimmten Modelltypen, beispielweise nach allen EPKn, oder nach allen Modelltypen einer Art anhand des entsprechenden Konzepts in der UMCO, beispielsweise `UMCO:Process`. Umgekehrt liefert die Suche nach allen `UMCO:Activity` eine Liste aller Geschäftsaktivitäten als Ergebnis. So lässt sich sowohl modellsemantik- als auch domänensemantik-orientiert die entstehende Unternehmensontologie abfragen.

4 Verwandte Arbeiten

In der Praxis liegen häufig unterschiedliche unverbundene Modelle vor. Einige Arbeiten zur Modellintegration konzentrieren sich auf die Modellierungssprachen und Möglichkeiten der Migration oder Integration basierend auf der Übertragung der Modelle von einer Modellierungssprache in eine andere [Ge07; MK07; KQC07], dem Matching von Modellen über ihre Metamodelle [Ka05] oder dem Management von Modellen gleichen Typs [MRB03]. Dabei wird der Aspekt heterogen verwendeter Domänensprache nicht betrachtet. Die Erweiterung von Geschäftsprozessmodellen mit semantischen Annotationen wurde vorgeschlagen. Dabei werden die Domänenkonzepte zu Instanzen in ihrer Modellsprachenontologie [BCD07; TF09]. Andere Vorschläge zur Integration von Prozessmodellen setzen das Vorliegen eines separat erstellten Domänenmodells oder einer Domänenontologie voraus [BEK06; We07; EKO07; Hö07]. Die Nutzung einer manuell erstellten Domänenontologie bei der Prozessmodellierung wird in [HLD05; TF07] vorgeschlagen und für die Softwaremodellierung in [SK06]. Das Problem des damit verbundenen zeit- und kostenintensiven Erstellungsaufwands sowie die Problematik der korrekten Erhebung wird im Software Engineering bei der Erstellung von Domänenmodellen und formalen domänenspezifischen Programmiersprachen [Am08; Hi08] als auch im Knowledge Engineering für die Konstruktion von Ontologien betrachtet [AMS08]. Bestehende Vorschläge zur automatisierten Domänenmodellerstellung konzentrieren sich bisher auf die Wiederverwendung von UML-Klassenmodellen [GGD04].

Aktuell liegen keine Ansätze vor, die eine Lösung der Frage der semantischen Integration von Modellen unterschiedlicher Art und verschiedenen Typs sowohl für den Bereich der Modellierungssprache als auch für den Bereich der Domänensprache zusammen vorschlagen. Hier kann unsere Methode als Unterstützung und Ergänzung der vorliegenden Ansätze dienen.

5 Schlussbetrachtung

Im vorliegenden Beitrag wurde eine Methode zur domänensemantik-orientierten Integration von Modellen mit Hilfe von Semantic-Web-Technologien vorgestellt und anhand von Beispielen gezeigt. Eine Vorgehensweise zur Wiederverwendung von Modellen durch Konvertierung in Ontologien wurde beschrieben. Die entstehenden Modellontologien lassen sich modellierungssprachlich-orientiert mit Hilfe der dafür entwickelten Brückenontologie UMCO mit ihren erweiternden MCOs verknüpfen. Die Modellontologien, die die Domänensemantik der Modelle enthalten, lassen sich weiterführend ohne manuelle Vorarbeiten zur Erstellung einer Domänenontologie abgleichen und in Bezug setzen. Dafür werden Ontology-Matching-Verfahren eingesetzt. Die auf diese Weise ermittelten semantischen Korrespondenzen können gesammelt werden, um langfristig die Ausgangsbasis für die Entstehung einer einfachen terminologischen Unternehmensontologie zu bilden. Gegeben durch die Ontologisierung der Modelle sind keine Vorarbeiten zur Erstellung einer Gesamtontologie erforderlich.

Durch die Einbindung von Nutzern kann die Qualität der automatisch ermittelten Korrespondenzen evaluiert und verbessert werden. Unser dafür entwickeltes MODI-System dient somit als Mediator und bietet Unterstützung zur Auflösung semantischer Heterogenität in Modellen. Das hier vorgestellte System wurde prototypisch implementiert und dient als Machbarkeitsnachweis der entwickelten Methode. Die entsprechende Evaluation ist noch zu Ende zu führen. Da das System die Kopplung externer Ontology-Matching-Werkzeug vorsieht, ist eine weiterführende Evaluation bezüglich deren Effizienz erforderlich, um geeignete Kombinationen zu ermitteln. Insgesamt hoffen wir, mit unserem Vorschlag die Nützlichkeit der Anwendung von Semantic-Web-Technologien zur Unterstützung in der Modellierung gezeigt zu haben, insbesondere, da die Brückenontologie beliebig erweiterbar ist mit weiteren MCOs, auch aus technischem oder anderen Anwendungsbereichen.

6 Danksagung

Diese Arbeiten wurden gefördert durch das BMBF im Rahmen des FHProfUnt-Programms im Projekt MODI, Förderkennzeichen 1728X07.

Literaturverzeichnis

- [AF05] Antoniou, G.; Franconi, E.; van Harmelen, F.: Introduction to Semantic Web Ontology Languages. In (Eisinger, N.; Maluszynski, J. Hrsg.): Reasoning Web. First International Summer School 2005, Msida, Malta, July 25-29, 2005, Revised Lectures. Springer, Berlin Heidelberg, 2005; S. 1–21.
- [AH08] Antoniou, G.; van Harmelen, F.: A semantic Web primer. MIT Press, Cambridge, Mass., 2008.
- [Am08] Ambler, S. W.: The object primer. Agile model-driven development with UML 2.0. Cambridge Univ. Press, Cambridge, 2008.
- [AMS08] d’Aquin, M. et al.: Toward a New Generation of Semantic Web Applications. In IEEE Intelligent Systems, 2008, 23; S. 20–28.
- [BCD07] Belecheanu, R. et al.: Business Process Ontology Framework, Super Project, 2007.
- [BEK06] Brockmans, S. et al.: Semantic Alignment of Business Processes. In (Manolopoulos, Y.; Filipe, J.; Constantopoulos, P.; Cordeiro, J. Hrsg.): Proc. of the Eighth Intern. Conf. on Enterprise Information Systems (ICEIS 2006). Information Systems Analysis and Specification. May 2006, Paphos, Cyprus. INSTICC, Setúbal, 2006; S. 197–203.
- [BP06] Blumauer, A.; Pellegrini, T.: Semantic Web und semantische Technologien: Zentrale Begriffe und Unterscheidungen. In (Pellegrini, T.; Blumauer, A. Hrsg.): Semantic Web Wege zur vernetzten Wissensgesellschaft. Springer, Berlin, 2006; S. 9–25.
- [BP08] Becker, J.; Pfeiffer, D.: Solving the Conflicts of Distributed Process Modelling – Towards an Integrated Approach. In (Golden, W.; Acton, T.; Conboy, K.; van der Heijden, H.; Tuunainen, V. K. Hrsg.): 16th European Conf. on Information Systems (ECIS 2008), 2008; S. 1555–1568.
- [Br04] Bruijn, J. de: Semantic Information Integration Inside and Across Organizational Boundaries, Deliverable. DERI, Innsbruck, 2004.
- [BRS96] Becker, J.; Rosemann, M.; Schütte, R.: Prozeßintegration zwischen Industrie- und Handelsunternehmen - eine inhaltlich-funktionale und methodische Analyse. In Wirtschaftsinformatik, 1996, 39; S. 309–316.
- [BVE04] Brockmans, S. et al.: Visual Modeling of OWL DL Ontologies Using UML. In (McIlraith, S. A. Hrsg.): The semantic Web - ISWC 2004. Third International Semantic Web Conference, Hiroshima, Japan, November 7 - 11, 2004. Springer, Berlin, 2004; S. 198–213.
- [CRF03] Cohen, W.; Ravikumar, P.; Fienberg, S.: A Comparison of String Distance Metrics for Name-Matching Tasks. In (Kambhampati, S.; Knoblock, C. A. Hrsg.): Proc. of IJCAI-03 Workshop on Information Integration on the Web (IIWeb-03), 2003; S. 73–78.
- [DOS03] Daconta, M. C.; Obrst L. J.; Smith K. T.: The Semantic Web. A Guide to the Future of XML, Web Services, and Knowledge Management. Wiley, Indianapolis, 2003.
- [EKO07] Ehrig, M.; Koschmider, A.; Oberweis, A.: Measuring Similarity between Business Process Models. In (Roddick, J.F.; Hinze, A. Hrsg.): Proc. of the Fourth Asia-

- Pacific Conf. on Conceptual Modelling (APCCM 2007), Australian Computer Science Communications, Ballarat, 2007; S. 71-80
- [EMS08] Euzenat, J.; Mocan, A.; Scharffe, F.: Ontology Alignments An ontology management perspective. In (Hepp, M.; Leenheer, P.; Moor, A.; Sure, Y. Hrsg.): *Ontology Management. Semantic Web, Semantic Web Services, and Business Applications*. Springer, Boston, MA, 2008; S. 177–206.
- [ES07] Euzenat, J.; Shvaiko, P.: *Ontology Matching*. Springer, Berlin, 2007.
- [FR09] Fengel, J.; Rebstock, M.: Model-Based Domain Ontology Engineering. In (Hepp, M.; Hinkelmann, K.; Stojanovic, N. Hrsg.): *Proc. of the 4th Intern. Workshop on Semantic Business Process Management (SBPM2009)*. ACM Press, New York, 2009.
- [Fr04] Frank, U.: Informationstechnologie und Organisation. In (Schreyögg, G., Werder, A. Hrsg.) *Handwörterbuch Unternehmensführung und Organisation HWO*. 4. Aufl., Schäffer-Poeschel, Stuttgart, 2004; S. 472-480.
- [GDD04] Gašević, D. et al.: Converting UML to OWL ontologies: The thirteenth International World Wide Web Conference. Alternate track papers & posters. ACM Press, New York, 2004; S. 488–489.
- [Ge07] Gehlert, A.: *Migration fachkonzeptueller Modelle*. Logos-Verl., Berlin, 2007.
- [Gr93] Gruber, T.: A translation approach to portable ontology specifications. In *Knowledge Acquisition*, 1993, 5; S. 199–220.
- [Hi08] Hillairet, G.: *Conception et utilisation d'ontologies pour la conception et l'exploitation des systèmes d'information*, Deliverable. 2008.
- [HLD05] Hepp, M. et al.: Semantic Business Process Management: A Vision Towards Using Semantic Web Services for Business Process Management. In (Lau, F. C. M.; Lei, H.; Meng, X.; Wang, M. Hrsg.): *Proc. of the IEEE Intern. Conf. on e-Business Engineering. ICEBE 2005*, October 2005, Beijing, China. IEEE, Los Alamitos, Calif., 2005; S. 535–540.
- [Hö07] Höfferer, P.: Achieving Business Process Model Interoperability Using Metamodels and Ontologies. In *Proc. of the 15th European Conference on Information Systems (ECIS 2007)*, 2007.
- [Ka05] Kappel, G. et al.: *ModelCVS - A Semantic Infrastructure for Model-based Tool Integration*, Technical Report, Wien, 2005.
- [KQC07] Kensche, D. et al.: GeRoMe: A Generic Role Based Metamodel for Model Management. In (Spaccapietra, S.; Atzeni, P.; Fages, F.; Hacid, M.-S.; Kifer, M.; Mylopoulos, J.; Pernici, B.; Shvaiko, P.; Trujillo, J.; Zaihrayeu, I. Hrsg.): *Journal on Data Semantics VIII*. Springer, Berlin, Heidelberg, 2007; S. 82–117.
- [La09] Lankhorst, M.: *Enterprise Architecture at Work. Modelling, Communication and Analysis*. Second Edition. Springer, Berlin, Heidelberg, 2009.
- [Mi09] Miller, G.A.: *WordNet - About Us*. WordNet. Princeton University. 2009. <http://wordnet.princeton.edu>, 12.5.2010.
- [MK07] Murzek, M.; Kramler, G.: The Model Morphing Approach – Horizontal Transformations between Business Process Models. In (Nummenmaa, J.;

- Söderström, E. Hrsg.): Proc. of the 6th Intern. Conf. on Perspectives in Business Information Research - BIR'2007, Tampere, Finland, 2007; S. 88–103.
- [MRB03] Melnik, S.; Rahm, E.; Bernstein, P. A.: Rondo: A Programming Platform for Generic Model Management. In (Halevy, A. Y. Hrsg.): Proc. of the ACM SIGMOD Intern. Conf. on Management of Data. 2003, San Diego. ACM Press, New York, 2003; S. 193–204.
- [SBH06] Shadbolt, N.; Berners-Lee, T.; Hall, W.: Semantic Web Revisited. In IEEE Intelligent Systems, 2006, 21; S. 96–101.
- [Sc96] Scheer, A.W.: ARIS – House of Business Engineering. Heft 133. Universität des Saarlandes, Institut für Wirtschaftsinformatik, Saarbrücken, 1996.
- [SH05] Stuckenschmidt, H.; Harmelen, F.: Information Sharing on the Semantic Web. Springer, Berlin Heidelberg, 2005.
- [SK06] Saeki, M.; Kaiya, H.: On Relationships among Models, Meta Models and Ontologies. In (Gray, J.; Tolvanen, J. P.; Sprinkle, J. Hrsg.): Proc. of the 6th OOPSLA Workshop on Domain-Specific Modeling, Jyväskylä, Finland, 2006.
- [SM07] Simon, C.; Mendling, J.: Integration of Conceptual Process Models by the Example of Event-driven Process Chains. In (Oberweis, A.; Weinhardt, C.; Gimpel, H. Hrsg.): eOrganisation. Service-, Prozess-, Market-Engineering. Bd. 1. 8. Intern. Tagung Wirtschaftsinformatik (WI 2007). Univ.-Verl, Karlsruhe, 2007; S. 677–694.
- [SWM04] Smith, M. K.; Welty, C.; McGuinness, D. L.: OWL Web Ontology Language Guide. <http://www.w3.org/TR/owl-guide/>, 27.12.2009.
- [SSS01] Staab, S. et al.: Knowledge Processes and Ontologies. In IEEE Intelligent Systems, 2001, 16; S. 26–34.
- [SZ92] Sowa, J.; Zachmann, J.: Extending and formalizing the framework for information systems architecture. In IBM Systems Journal, 1992, 31; 3, S. 590-616.
- [TF06] Thomas, O.; Fellmann, M.: Semantische Integration von Ontologien und Ereignisgesteuerten Prozessketten. In (Nüttgens, M.; Rump, F. J.; Mendling, J. Hrsg.): EPK 2006 - Geschäftsprozessmanagement mit Ereignisgesteuerten Prozessketten. CEUR-WS.org, Wien, 2006; S. 7–23.
- [TF07] Thomas, O.; Fellmann, M.: Semantic Business Process Management: Ontology-Based Process Modeling Using Event-Driven Process Chains. In IBIS Interoperability in Business Information Systems, 2007, 2; S. 29–44.
- [TF09] Thomas, O.; Fellmann, M.: Semantische Prozessmodellierung – Konzeption und informationstechnische Unterstützung einer ontologiebasierten Repräsentation von Geschäftsprozessen. In Wirtschaftsinformatik, 2009, 51; S. 506–518.
- [UG04] Uschold, M.; Gruninger, M.: Ontologies and Semantics for Seamless Connectivity. In ACM SIGMOD Record, 2004, 33; S. 58–64.
- [We07] Weske, M.: Business Process Management. Concepts, Languages, Architectures. Springer, Berlin Heidelberg, 2007.
- [Zh04] Zhdanova, A. et al.: Ontology Alignment Solution v2.0., Deliverable D1.4, Esperanto project. 2004.