

Semantische Prozessmodellierung – Konzeption und informationstechnische Unterstützung einer ontologiebasierten Repräsentation von Geschäftsprozessen

In dem Beitrag wird eine Erweiterung von Prozessmodellierungssprachen entworfen, mit deren Hilfe die in natürlicher Sprache formulierte Semantik der Bezeichner von Prozessmodellelementen durch formale Begriffe einer Ontologie repräsentiert werden kann. Durch diese Formalisierung der modellelementbezogenen Semantik können die mit der Verwendung der natürlichen Sprache einhergehenden Interpretationsspielräume eliminiert sowie die Suche in Modellierungswerkzeugen verbessert werden. Darüber hinaus werden durch die Verwendung von Regeln im Zusammenspiel mit Ontologien neue Möglichkeiten der Validierung von Prozessmodellen eröffnet.

DOI 10.1007/s11576-009-0201-y

Die Autoren

**Prof. Dr. Oliver Thomas
Michael Fellmann M.A.**

Universität Osnabrück
Institut für Informationsmanagement
und Unternehmensführung, Fach-
gebiet Informationsmanagement
und Wirtschaftsinformatik
Katharinenstraße 3
49074 Osnabrück
Deutschland
{oliver.thomas | michael.
fellmann}@uni-osnabrueck.de

Eingegangen: 2008-10-30
Angenommen: 2009-06-22
Angenommen nach zwei
Überarbeitungen durch Prof. Dr. Buhl.

This article is also available in English
via <http://www.springerlink.com> and
<http://www.bise-journal.org>; Thomas
O, Fellmann M (2009) Semantic
Process Modeling – Design and
Implementation of an Ontology-based
Representation of Business Processes.
Bus Inf Syst Eng. doi: 10.1007/s12599-
009-0078-8.

1 Einleitung

Zur Planung, Kontrolle und Steuerung von Geschäftsprozessen wird in Wissenschaft und Praxis eine Vielzahl an

Modellierungssprachen eingesetzt. Beispiele sind das Petri-Netz, die Ereignisgesteuerte Prozesskette (EPK), das UML-Aktivitätsdiagramm oder die Business Process Modeling Notation (BPMN). Bei den Modellierungssprachen haben sich semiformale, grafische Darstellungsformen durchgesetzt, welche einerseits eng an betriebswirtschaftliche Fachtermini angelehnt sind. Andererseits sind sie exakt genug, dass die Modelle als Ausgangspunkt zur Umsetzung von Anwendungssystemen dienen können. Auch wenn dies ein Grundgedanke der modellbasierten Systemgestaltung ist, so liegt in der genannten Verbindung zwischen natürlicher Sprache und grafischer Repräsentationsform eines der Hauptprobleme für den Modellierer: Unabhängig von seiner Entscheidung für eine Beschreibungssprache werden die Bezeichnungen der einzelnen Elemente eines Geschäftsprozessmodells, z. B. „Auftrag prüfen“ als Bezeichnung für eine EPK-Funktion oder eine BPMN-Aktivität, vom Modellierer in einer natürlichen Sprache hinzugefügt. Ein wesentlicher Teil der Semantik eines Prozessmodells ist somit immer an die natürliche Sprache gebunden, woraus zwei wesentliche Problemklassen für die Modellentwicklung und -anwendung resultieren.

Die Bezeichnung eines Modellelements mit Hilfe der natürlichen Sprache bringt *erstens* Interpretationsspielräume mit sich, die in der Literatur auch als Sprach- oder Begriffsdefekte bezeichnet werden; hierbei können Synonyme, Homonyme,

Äquipollenzen, Vagheiten und falsche Bezeichner unterschieden werden (Ortner 1997, S. 31 ff). Die genannten Sprachdefekte führen zu Problemen, welche die Verwendung der Modelle als Kommunikationsmedium und als Ausgangspunkt einer informationstechnischen Konzeption einschränken. Wird bspw. ein Modellelement in einem Modell mit „Handelsgut“ bezeichnet und ein weiteres Modellelement in einem anderen Modell mit „Handelsware“ und repräsentieren beide Artefakte dasselbe zugrunde liegende Objekt, so liegt eine Synonymie vor und damit verbunden das Risiko des Nichterkennens äquivalenter Konstrukte, die bei der Umsetzung in ein DV-Konzept zum mehrfachen Entwurf von Lösungen führen können (Rosemann 1996, S. 188).

Über die Sprachdefekte hinaus impliziert die Verwendung der natürlichen Sprache *zweitens* eine nicht maschinell verarbeitbare Semantik der Prozessmodelle. Dies führt sowohl bei der Suche in datenbankbasierten Modellierungswerkzeugen als auch bei der Validierung von Prozessmodellen zu Schwierigkeiten. Bezogen auf die Suche in Modellierungswerkzeugen besteht das Problem, dass eine Ermittlung der in Prozessmodellen enthaltenen Zusammenhänge, die nicht explizit vom Konstrukteur im Modell festgehalten wurden, jedoch mit Hilfe logischer Schlussfolgerungen ableitbar sind, nicht möglich ist. Ein Beispiel hierfür ist eine Funktion, die auf Ressourcen zugreift, die in einem Lager bevorratet werden. Sind diese Zusammenhänge

nicht maschinell verarbeitbar spezifiziert, so kann bspw. nicht gefolgert werden, dass die genannte Funktion den Lagerbestand reduziert. Bezogen auf die Validierung, verstanden als die Untersuchung, ob ein Modell den Zweck erfüllt, für den es konstruiert wurde, besteht das Problem, dass diese ohne eine maschinell verarbeitbare Semantik nicht möglich ist (Desel 2002, S. 24). Hierdurch können inadäquate Modelle nicht bereits zum Zeitpunkt der Modellerstellung erkannt werden.

Den beiden genannten Problemstellungen, die aus der Verwendung der natürlichen Sprache zur Bezeichnung von Prozessmodellelementen resultieren, wird in diesem Beitrag durch eine Erweiterung semiformaler Modellierungssprachen mit Hilfe von Ontologien begegnet. Die Grundidee dieser Erweiterung, die im weiteren Verlauf als *Semantische Prozessmodellierung* bezeichnet wird, ist die Zuordnung von in Ontologien formalisierten Begriffen (engl.: concepts) zu Prozessmodellelementen. Hierdurch kann die in natürlicher Sprache formulierte Semantik der Bezeichner von Prozessmodellelementen formal und maschinell verarbeitbar repräsentiert werden.

Zur Erreichung dieser Zielsetzung wird wie folgt vorgegangen: Zunächst werden verwandte Arbeiten gewürdigt, die semantische Problemstellungen im Geschäftsprozessmanagement untersuchen, wobei auch Beiträge aus angrenzenden Disziplinen herangezogen werden (Abschnitt 2). Um die erkannte Forschungslücke zu schließen, wird anschließend eine Methode entwickelt, die zur systematischen Herleitung der Semantischen Prozessmodellierung verwendet wird (Abschnitt 3). Diese Herleitung vollzieht sich nachfolgend in zwei Schritten. Im ersten Schritt wird das Konzept anhand ausgewählter Modellierungs- und Ontologiesprachen illustriert sowie mit Hilfe eines Informationsmodells generalisiert (Abschnitt 4). Anhand einer implementierten Testumgebung wird im zweiten Schritt die grundsätzliche Anwendbarkeit der entwickelten Konzeption demonstriert (Abschnitt 5). Der Beitrag schließt mit Nutzenpotenzialen des Ansatzes (Abschnitt 6) sowie Limitationen und einem Ausblick (Abschnitt 7).

2 Stand der Forschung und verwandte Arbeiten

Während im alltagssprachlichen Sinne mit dem Begriff der Semantik häufig auf die Bedeutung oder den Inhalt eines Wortes oder Satzes hingewiesen wird, stützt sich das wissenschaftliche Verständnis von Semantik auf die Begriffsbildung der Sprachwissenschaft (Linguistik). Innerhalb dieser Disziplin bezeichnet die Semantik (auch: Bedeutungslehre) dasjenige Teilgebiet, das sich mit dem Sinn und der Bedeutung von Sprache bzw. sprachlichen Zeichen befasst, m. a. W.: die Lehre von den Bedeutungen und von der Beziehung der Zeichen zum gemeinten Gegenstand. Überträgt man dieses Verständnis auf Prozessmodellierungssprachen, so kann unter der Semantik eines Prozessmodells – in einer ersten Annäherung – die Beziehung zwischen den Elementen des Modells (Zeichen) und einem existierenden oder neu zu schaffenden betrieblichen Geschäftsprozess (Gegenstandsbereich) verstanden werden. Dieses Verständnis von Semantik wird jedoch nicht von allen Autoren geteilt, insbesondere in frühen Arbeiten zur Prozessmodellierung (z. B. Keller et al. 1992) wird das Adjektiv „semantisch“ dazu verwendet, um die Bedeutung der Modellierung von Geschäftsprozessen aus einer betriebswirtschaftlichen, nichttechnischen Perspektive zu betonen.

Untersuchungen zur Semantik von Modellierungssprachen haben sich bislang hauptsächlich auf die *formale Semantik* der zur Verfügung stehenden Sprachelemente – im Folgenden auch Sprachkonstrukte genannt – konzentriert. Die formale Semantik ist u. a. in der Theoretischen Informatik und der Logik verankert und beschäftigt sich mit der exakten Bedeutung künstlicher (d. h. konstruierter) oder natürlicher Sprachen. Einen zentralen Stellenwert bei der Untersuchung der formalen Semantik von Modellierungssprachen nehmen mathematische Methoden ein (für die EPK vgl. stellvertretend Kindler 2006 und die dort zitierte Literatur). Die Semantik, die Modellelemente in Form von Modellelementbezeichnern hinzugefügt wird und gerade bei semiformalen Sprachen an die natürliche Sprache gebunden ist, wird in den Arbeiten nicht berücksichtigt.

Unabhängig von Modellierungssprachen wird die formale Repräsentation von Wissen im Forschungsgebiet der Künst-

lichen Intelligenz vorangetrieben. Zur Darstellung der komplexen Wissensbeziehungen werden hierbei häufig Ontologien verwendet, die aktuell durch die Bestrebungen, das World Wide Web zu einem Semantic Web zu erweitern (Fensel et al. 2003), an Bedeutung gewinnen. Unter einer Ontologie wird in der hier verwendeten informatiknahen Interpretation nach Gruber (1993, S. 199) eine explizite formale Spezifikation einer Konzeptualisierung (engl.: conceptualization) verstanden. Eine Konzeptualisierung ist dabei eine abstrakte, vereinfachte Sicht der Welt, die für bestimmte Zwecke repräsentiert werden soll (Gruber 1993, S. 199). Neuere Definitionsversuche betonen darüber hinaus die intersubjektive Gültigkeit der Konzeptualisierung, sodass Ontologien auch als ein von mehreren Individuen entwickeltes Vokabular aufgefasst werden können, das gemeinsam in einer Gruppe akzeptiert und genutzt wird (engl.: shared conceptualization) (Studer et al. 1998, S. 186; Gómez-Pérez et al. 2004, S. 8). Die Wirtschaftsinformatik macht sich die zum Teil umfangreichen Vorarbeiten der Ontologieforschung bspw. im Rahmen des Wissensmanagements (Zelewski et al. 2005), des Produktdatenmanagements (Hahn 2005) oder des *Information Retrieval* (Kuropka 2004) zunutze. In diesen Untersuchungen werden struktural organisierte Artefakte (z. B. Dokumente oder Produktmodelle) durch Ontologien repräsentiert. Die Nutzung dieser Repräsentationen zur semantischen Anreicherung von Geschäftsprozessmodellen ist – im Gegensatz zum vorliegenden Beitrag – nicht Gegenstand der Arbeiten.

Hinsichtlich der ontologischen Analyse ist im internationalen Kontext der Information-Systems-Disziplin vor allem das Bunge-Wand-Weber-Modell hervorzuheben (Wand u. Weber 1995), das – vereinfacht gesprochen – als eine Theorie zur Beschreibung von Informationssystemen verstanden werden kann. Obwohl das Modell nicht geeignet ist, sämtliche Phänomene bei der Gestaltung von Informationssystemen zu erfassen, wird es in der Forschung vielfach angewendet, bspw. zur ontologischen Evaluation von Modellierungswerkzeugen (Green u. Rosemann 2000) und -methoden (Green 1996), von Interoperabilitätsstandards (Green et al. 2005) sowie zur Auswahl und Einführung von Standardsoftware (Soffer et al. 2001). Die hiermit verbundene Beurteilung der Eignung und der Ansprüche von Infor-

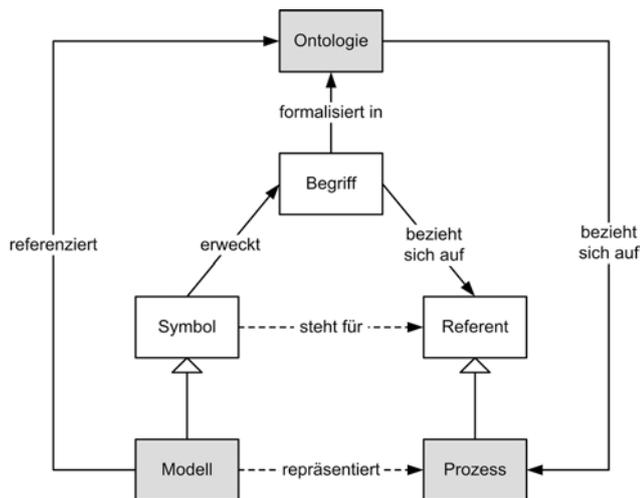


Abb. 1 Erweitertes Semiotisches Dreieck der Semantischen Prozessmodellierung

mationsmodellen und den zu ihrer Konstruktion verwendeten Modellierungssprachen sowie die Herleitung von Kriterien zur Bestimmung der Güte der Artefakte ist nicht Gegenstand des in diesem Beitrag vorgeschlagenen Ansatzes. Im Gegensatz zu der Anwendung von Ontologien zur Evaluation modellierungssprachlicher Artefakte baut die Semantische Prozessmodellierung auf einer Repräsentation der Artefakte in einer Ontologie auf.

Die Potenziale einer solchen Verknüpfung von Ontologien und Prozessmodellen werden seit geraumer Zeit in der Literatur erkannt (Hepp et al. 2005; Lin u. Strasunskas 2005; Ahlemann et al. 2006; Hepp u. Roman 2007). Die Autoren versuchen i. d. R. mit einer ontologiebasierten Attributierung von Prozessmodellen die Grundlage für eine automatisierte Verarbeitung der Ablaufmodelle zu schaffen. Die meisten Arbeiten sind hierbei sprachspezifisch ausgerichtet, d. h. sie behandeln ausschließlich die semantische Annotation von Prozessmodellen, die mit Hilfe einer bestimmten Sprache repräsentiert sind. Solche semantische Erweiterungen für Prozessbeschreibungssprachen existieren u. a. für das Petri-Netz (Koschmidler u. Ried 2005; Brockmans et al. 2006), die EPK (Thomas u. Fellmann 2007; Bögl et al. 2008), die BPMN (Abramowicz et al. 2007), die Demo Engineering Methodology for Organizations (DEMO) (Dietz 2006) und die Extended Enterprise Modeling Language (EEML) (Lin u. Ding 2005). Für das UML-Aktivitätsdiagramm liegt ein Konzept zur automatischen Synthese und Modifikation von Modellen nach Änderungen an Subprozessen vor (Lautenbacher u. Bauer 2006). Gelegentlich führt die semantische Annotation von Prozessmodellen auch zur Definition

neuer Sprachen, wie im Falle des Process Semantic Annotation Model (PSAM) von Lin (2008). Letztgenannte Autorin entwickelt darüber hinaus mit dem Process Semantic Annotation Tool die bislang umfassendste Untersuchung zur semantischen Annotation von Prozessmodellen. Im Unterschied zu den sprachspezifischen Ansätzen ist der in diesem Beitrag vorgeschlagene Ansatz zur Annotation von Geschäftsprozessmodellen sprachunabhängig ausgelegt.

Die mit der semantischen Annotation verbundene Vereinheitlichung der in Modellen und Modellelementen verwendeten Terminologie kann auch durch die Verwendung von Fachbegriffsmodellen erreicht werden (Ortner 1997; Rosemann u. Schwegmann 2002). Die Verwendung von Ontologien zur formalen Repräsentation einer Domäne besitzt jedoch den Vorteil einer maschinellen Interpretation. Hierdurch können insbesondere nicht explizit repräsentierte Fakten durch Verfahren des maschinellen Schließens automatisiert ergänzt werden, um eine vollständige Interpretation der Semantik bspw. bei der Suche oder Validierung von Modellen zu ermöglichen.

Diese maschinelle Verarbeitbarkeit ist zugleich der Grundgedanke von Forschungsbemühungen im Bereich der *Semantic Web Services* (Cabral et al. 2004; Cardoso u. Sheth 2005). Die semantische Beschreibung von Webservices intendiert eine verbesserte Auswahl und Orchestrierung der Dienste und ist somit auf die Spezifika einer serviceorientierten Softwareentwicklung zugeschnitten. Im Unterschied zu diesen Arbeiten befasst sich die vorliegende Konzeption ausschließlich mit Geschäftsprozessmodellen, die auf einer fachlichen Ebene beschrieben wer-

den. Dies trägt dem Umstand Rechnung, dass organisatorische Schwierigkeiten als Hindernis bei der Gestaltung und Veränderung von Geschäftsprozessen nach wie vor einen hohen Stellenwert einnehmen (AIIM 2007).

Ein entsprechender Ansatz zur automatisierten Planung von fachlichen Prozessmodellen wurde im Rahmen des Projekts SEMPRO entwickelt (Henneberger et al. 2008; Heinrich et al. 2008). Hierbei werden ausgehend von einzelnen Prozessaktionen oder -fragmenten sowie eines Start- und eines Zielzustands automatisiert Vorschläge für vollständige Prozessmodelle generiert. Die zu diesem Verfahren erforderliche semantische Beschreibung von Prozessaktionen durch ihre Eingabe- und Ausgabeparameter kann prinzipiell mit dem hier entwickelten Ansatz zur Verknüpfung von Ontologien und Prozessmodellen realisiert werden. Dieser ist jedoch nicht auf eine bestimmte Anwendung festgelegt, sondern bildet vielmehr das Fundament, um darauf aufbauend Ansätze zur Prozessplanung, zur Verbesserung der Anfrage an Prozessdatenbanken oder zur erweiterten Validierung von Prozessmodellen zu entwickeln.

3 Methode und Vorgehen

Der Kerngedanke der Semantischen Prozessmodellierung beruht auf einer Erweiterung des Semiotischen Dreiecks. Es besagt in seiner Grundform nach Ogden u. Richards (1923), dass Symbole keinen direkten und unmittelbaren Bezug zu den Referenten (Dingen) besitzen, für die sie stehen (inneres Dreieck, **Abb. 1**). Vielmehr wird ein Bezug erst indirekt durch Begriffe ermöglicht, die sich auf Referenten beziehen und durch das Auftreten von Symbolen aktiviert werden. Dies bedeutet, dass die Interpretation oder Wahrnehmung von Referenten durch ein Subjekt erst möglich wird, wenn dieses über einen mit dem Symbol assoziierten Begriff, Vorstellungsinhalt oder ein mentales Modell verfügt.

Analoge Beziehungen ergeben sich zwischen den Elementen des erweiterten Semiotischen Dreiecks „Modell“, „Ontologie“ und „Prozess“ (äußeres Dreieck, **Abb. 1**). Modelle sind als eine spezielle Form von Symbolen interpretierbar. Eine Beziehung zwischen Modell und Prozess ist nicht direkt und unmittelbar möglich, da das Modell das Ergebnis eines durch

Kreativität gekennzeichneten Konstruktionsprozesses ist. Die Bedeutung der Modellelemente hängt damit von der Vorstellung bzw. vom mentalen Modell der Konstrukteure des Modells ab. Dies trifft insb. auf die für Modellelemente vergebenen Bezeichner zu, die als Symbole für die während der Erstellung benutzten Begrifflichkeiten eingesetzt werden. Durch die Verknüpfung von Prozessmodellelementen mit Begriffen aus formalen Ontologien, in **Abb. 1** durch den Pfeil „referenziert“ angedeutet, kann die in Modellelementbezeichnern enthaltene, mit Hilfe der natürlichen Sprache formulierte Semantik explizit spezifiziert werden.

Wird bspw. ein Modellelement mit dem Bezeichner (Symbol) „Kundenauftrag“ vorgefunden, so kann zu dessen Interpretation die durch Domänenexperten formal vorgenommene explizite Konzeptualisierung in der Ontologie herangezogen werden. Diese stellt weitere Informationen über den bezeichneten Begriff zur Verfügung und setzt ihn in Beziehung zu anderen Begriffen. So kann in der Ontologie bspw. definiert werden, dass ein Kundenauftrag ein Datum besitzt und ein spezifischerer Begriff als „Auftrag“ ist, der wiederum ein spezifischerer Begriff als „Dokument“ ist. Diese Zusammenhänge können von einer Maschine verarbeitet werden, womit ein Kernaspekt der semantischen Erweiterung von Prozessmodellen mit Ontologien deutlich wird: Das in Ontologien repräsentierte Wissen kann zur maschinellen Interpretation der Prozessmodelle herangezogen werden. Durch Verfahren des maschinellen Schließens werden dabei neue Fakten generiert, die nicht explizit im ursprünglichen Modell enthalten sind. Die Suche nach einem Prozessmodell mit einem Element „Auftrag bearbeiten“ kann so als Ergebnis ein Prozessmodell mit einem Element „Kundenauftrag bearbeiten“ zurückliefern, da in der Ontologie in einer maschinell interpretierbaren Form die Information hinterlegt ist, dass die Kundenauftragsbearbeitung eine Spezialisierung der Auftragsbearbeitung ist.

Um die mit dem erweiterten Semiotischen Dreieck aufgezeigte semantische Erweiterung der Prozessmodellierung realisieren zu können, sind zunächst Sprachen zur Repräsentation von Ontologien (Ontologiesprache) und Prozessmodellen (Modellierungssprache) auszuwählen sowie eine Ontologie zu (re-)konstruieren.

Anschließend ist eine spezielle Annotation von Prozessmodellen auf Basis der ausgewählten Sprachen herzuleiten sowie sprachübergreifend (mit Hilfe eines Informationsmodells) zu generalisieren.

4 Konzeption der Semantischen Prozessmodellierung

4.1 Ontologien und Ontologiesprachen für das Prozessmanagement

Die Semantische Prozessmodellierung setzt das Vorhandensein von Ontologien und Ontologiesprachen voraus, die zur Repräsentation der Semantik von Modellelementen verwendet werden können. Zur Konstruktion einer Ontologie für die Semantische Prozessmodellierung kann auf bereits existierende Ontologien, wie die Enterprise Ontology (Uschold et al. 1998) oder TOVE (TOronto Virtual Enterprise) (Fox 1992), zurückgegriffen werden. Weitere Ontologien stehen durch die Übersetzung etablierter Industriestandards in Ontologiesprachen zur Verfügung, wie bspw. eClassOWL als Portierung des eCl@ss-Standards nach OWL (Hepp 2005), oder durch die Generierung der Ontologien aus bereits im Unternehmen vorhandenen semantischen Strukturen, wie bspw. Relationenmodelle (Gómez-Pérez u. Manzano-Macho 2003). Werden Begriffe aus mehreren Ontologien verwendet, so kann die daraus resultierende Ontologie durch eine übergreifende *UpperOntology* weiter strukturiert werden (vgl. die Übersicht bei Gómez-Pérez et al. 2004, S. 204 ff). Für die in diesem Beitrag zu vollziehende Ontologiekonstruktion wird SUMO (Suggested Upper Merged Ontology) (Niles u. Pease 2001) ausgewählt, da diese über ein umfangreiches, hierarchisches Kategoriensystem verfügt, das im Vergleich zu abstrakteren Upper Ontologies eine leichtere strukturelle Einordnung der zu definierenden Begriffe in die bestehende Ontologie ermöglicht.

Zur expliziten und formalen Repräsentation einer Ontologie existieren zahlreiche Ontologiesprachen (vgl. die Übersicht bei Gómez-Pérez et al. 2004, S. 204 ff). Insbesondere die Ontologiesprache OWL (Web Ontology Language) besitzt aufgrund ihrer weiten Verbreitung auch außerhalb der KI-Forschergemeinde und ihrer Standardisierung durch das W3C (Smith et al. 2004) eine beson-

dere Relevanz für die im Rahmen dieses Beitrags konzipierte Semantische Prozessmodellierung, im Sinne einer umfassenden Akzeptanz und Werkzeugunterstützung. Sie wird als Ontologiesprache ausgewählt. Da OWL in den drei Varianten Light, DL und Full auftritt, ist weiter eine geeignete Subsprache zu bestimmen. Aufgrund der mit der Semantischen Prozessmodellierung betonten maschinellen Verarbeitbarkeit von Semantik wird hier die an Beschreibungslogiken (Description Logics) angelehnte Variante OWL DL ausgewählt, nachfolgend vereinfacht mit „OWL“ bezeichnet. Diese baut auf der Beschreibungslogik SHOIN(D) auf und umfasst Verneinung, Disjunktion und eine eingeschränkte Form von All- und Existenzquantoren. Im Unterschied zu OWL Full existieren für OWL DL leistungsfähige Inferenzmaschinen, wie bspw. *Pellet* (<http://pellet.owldl.com>), *FACT++* (<http://owl.man.ac.uk/>) und *Racer* (<http://www.racer-systems.com/>). Zur kriterienbasierten Auswahl von Ontologiesprachen sei auf die Literatur verwiesen (Gómez-Pérez et al. 2004; Casely-Hayford 2005).

4.2 Ontologiebasierte Prozessrepräsentation

Die Beschreibung der Semantik von Prozessmodellen durch Begriffe einer Ontologie erfordert zunächst die Definition derjenigen Ontologiekonstrukte, welche die Konstrukte der verwendeten Prozessbeschreibungssprache (z. B. EPK-Funktion) repräsentieren (dieser Abschnitt). Anschließend werden diejenigen Ontologiekonstrukte definiert, welche die Bezeichner der Modellelemente (z. B. „Auftrag prüfen“) repräsentieren (Abschnitt 4.3).

Grundlegende Elemente, die von Ontologiesprachen unabhängig von den ihnen zugrunde liegenden Formalismen bereitgestellt werden, sind Begriffe und Relationen, bei OWL auch Klassen und Properties genannt. Die Ausgestaltung der Repräsentation ist einigen Freiheitsgraden unterworfen. So ist bspw. zu entscheiden, ob Beziehungen zwischen Sprachkonstrukten als Properties oder als Instanzen von Klassen in der Ontologie repräsentiert werden. In diesem Beitrag erfolgt eine graphentheoretische Interpretation von Prozessmodellen, d. h. ein Prozessmodell wird aufgefasst als ein gerichteter Graph mit Knoten und Kanten. Knotentypen, wie z. B. eine EPK-Funktion oder

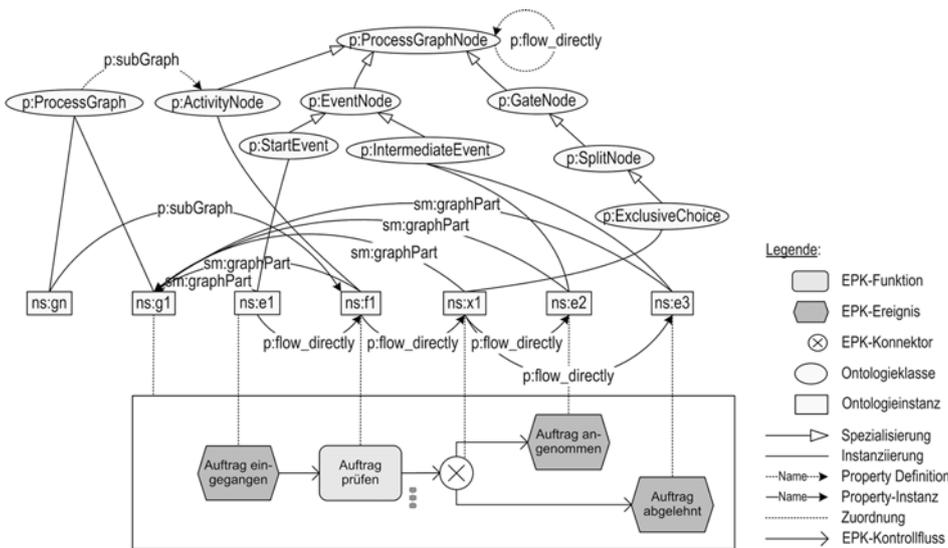


Abb. 2 Ontologiebasierte Repräsentation eines Prozessmodells am Beispiel der EPK

eine BPMN-Aktivität, werden als Klassen in der Ontologie definiert. Beziehungen zwischen diesen Knotentypen, z. B. in Form eines EPK-Kontroll- oder BPMN-Sequenzflusses, werden als Properties in der Ontologie erklärt. Dieses Vorgehen erlaubt eine einfache und intuitive Repräsentation von Prozessmodellen in Ontologien, deren Beschränkung liegt jedoch darin, dass die Möglichkeiten zur Definition von Beziehungen zwischen Sprachkonstrukten stark von den Möglichkeiten der gewählten Ontologiesprache abhängig sind.

Abb. 2 verdeutlicht die skizzierte ontologiebasierte Prozessrepräsentation an einem Prozessmodellfragment, das mit der EPK erstellt wurde. Die Ereignisse, die Funktion und der Konnektor des EPK-Modells werden in der Ontologie durch Instanzen von Klassen repräsentiert, die aufgrund der graphentheoretischen Interpretation aus der Klasse ProcessGraphNode abgeleitet sind. Es sei darauf hingewiesen, dass die den Bezeichnungen der Ontologieelemente vorangestellten Präfixe Namensräume kennzeichnen (Bray et al. 2006). Das Präfix „ns“ steht hierbei für einen beliebigen Namensraum, in dem die Klassen und Instanzen der Beispielsontologie liegen, „s“ für den Namensraum der SUMO-Ontologie, „p“ für den Namensraum der hier verwendeten SUMO-Erweiterung und „sm“ für den Namensraum der SUMO Mid Level Ontology (Niles u. Pease 2001). Im Folgenden wird zur besseren Lesbarkeit im Fließtext auf die Nennung der Namensraumpräfixe verzichtet. Die Ontologieklassen, die zur Repräsentation der EPK-Konnektoren verwendet werden, sind in Anlehnung an Arbeiten

zur sprachübergreifenden Beschreibung von Workflow Patterns definiert worden (van der Aalst et al. 2003). Der zwischen diesen Modellelementen bestehende zeitliche und sachlogische Zusammenhang, der bei der EPK Kontrollfluss genannt wird, ist in der Ontologie durch das Property flow_directly abgebildet.

4.3 Annotation von Prozessmodellen

Die Verknüpfung von in der Ontologie repräsentierten Prozessmodellen und Modellelementen mit weiteren Elementen der Ontologie wird als *semantische Annotation* bezeichnet. Das Attribut „semantisch“ unterstreicht hierbei – im Gegensatz zu einer Annotation bspw. in Form einer frei formulierten Textnotiz – den Anspruch der maschinellen Verarbeitbarkeit der durch die Annotation gegebenen Zusatzinformationen. Die folgenden Ausführungen beziehen sich auf die Annotation von Modellelementen, sie sind jedoch in gleicher Weise übertragbar auf die Annotation ganzer Modelle.

Ein wesentliches Gestaltungsmerkmal der Annotation ist die Vielfachheit (Kardinalität) der Verknüpfung von repräsentierten Modellelementen (im Folgenden auch als „Modellelementinstanzen“ bezeichnet) mit weiteren Instanzen der Ontologie (zur besseren Unterscheidung im Folgenden auch als „Domäneninstanzen“ bezeichnet). Eine 1:n-Beziehung besteht, wenn eine Modellelementinstanz, wie z. B. „Auftrag prüfen“, über geeignete Beziehungen mit mehreren Domäneninstanzen, wie z. B. „Vertrieb“ zur Dokumentation der ausführenden Organisationseinheit und „Produktkonfigurator“ zur Über-

prüfung der technischen Durchführbarkeit eines Auftrags, verknüpft wird. Die mehrfache Verknüpfung von Modellelementinstanzen mit Domäneninstanzen führt allerdings zu redundanten Annotationen, wenn semantisch gleichwertige Modellelemente in mehreren Modellen auftreten. Daher wird im Rahmen des hier entwickelten Ansatzes eine 1:1-Beziehung vorgeschlagen, d. h. die Annotation einer Modellelementinstanz mit nur *einer* sie semantisch vollständig repräsentierenden Domäneninstanz. Ein Beispiel hierfür ist eine Modellelementinstanz „Auftrag prüfen“, die mit einer Domäneninstanz „Auftragsprüfung“ in der Ontologie verknüpft wird. Durch dieses Verfahren der Annotation werden zwar prinzipiell mehr Domäneninstanzen benötigt (sind die Elemente der zu annotierenden Prozessmodelle semantisch vollständig disjunkt, so kann maximal für jedes zu annotierende Modellelement eine korrespondierende Domäneninstanz erforderlich sein). Die Anzahl der in der Ontologie vorzuhaltenden Domäneninstanzen kann jedoch verringert werden, indem die Annotationsrelation in Anlehnung an die Beschreibung der Beziehungen zwischen Elementen kontrollierter Vokabulare (Miles u. Bechhofer 2008) in eine Relation der semantischen Äquivalenz und eine Relation der höheren semantischen Spezifität aufgefächert wird.

Eine Relation der semantischen Äquivalenz, die als equivalentTo bezeichnet wird, besteht zwischen einer Domäneninstanz und einer Modellelementinstanz, wenn Erstere die intendierte Semantik der Modellelementinstanz (d. h. die Semantik des Objekts, das der Konstruktion des semiformalen Modells ursprünglich

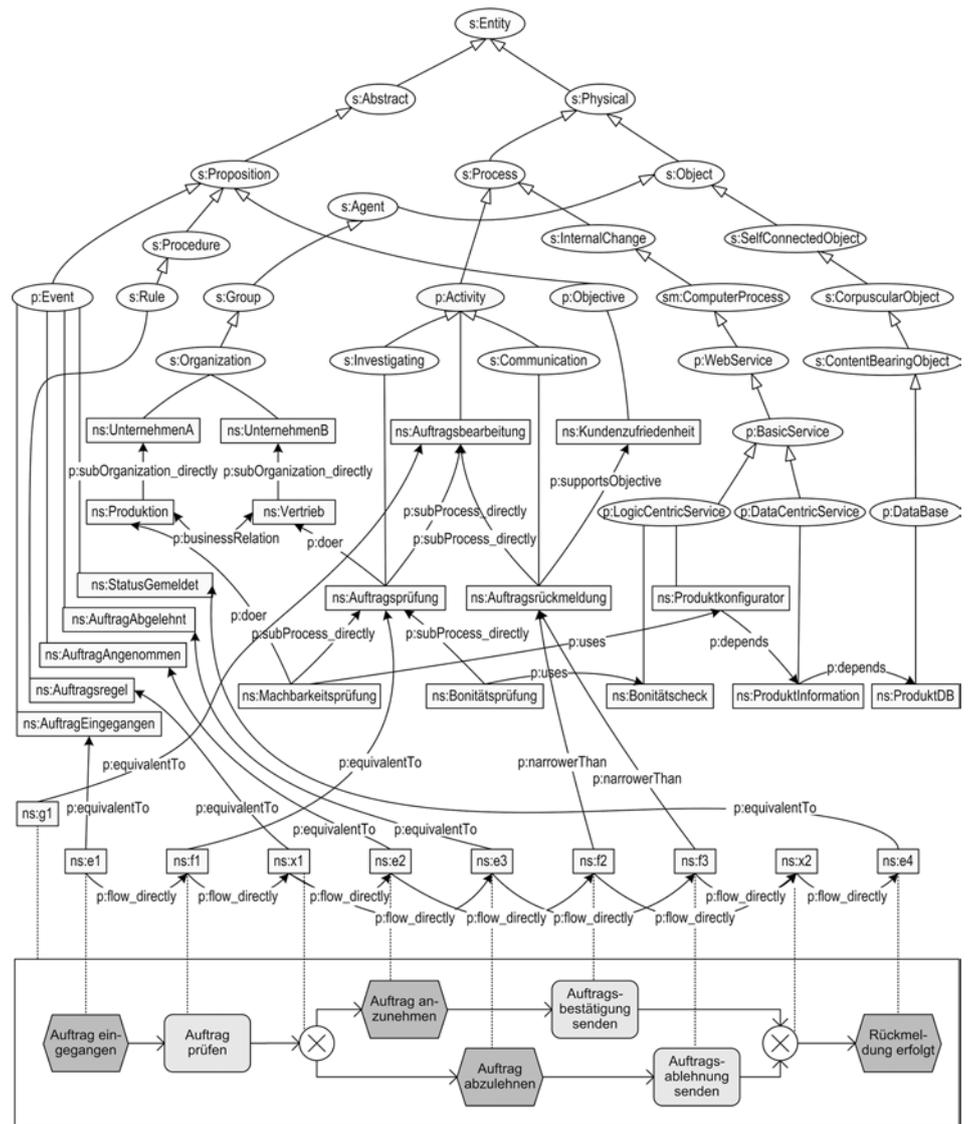


Abb. 3 Semantische Annotation eines EPK-Modells

zugrunde liegt) exakt widerspiegelt. Ein Beispiel hierfür ist in **Abb. 3** die Annotation einer Modellelementinstanz f1 („Auftrag prüfen“) mit der repräsentierenden Domäneninstanz Auftragsprüfung.

Eine Relation einer höheren semantischen Spezifität, die als narrowerThan bezeichnet wird, besteht zwischen einer Domänen- und einer Modellelementinstanz, wenn die Domäneninstanz die Modellelementinstanz semantisch enthält, die Modellelementinstanz jedoch in einigen Aspekten spezifischer oder „enger“ definiert ist. Ein Beispiel hierfür ist eine Domäneninstanz „Daten erheben“ und eine Modellelementinstanz „Kundendaten erheben“. Durch diese Relation können mehrere Modellelementinstanzen mit einer Domäneninstanz annotiert werden, womit eine Verringerung der benötigten Domäneninstanzen erreicht werden

kann. Darüber hinaus wird durch diese Relation eine Annotation auch dann möglich, wenn zum Zeitpunkt der Annotation noch keine äquivalente Domäneninstanz zu einer Modellelementinstanz vorhanden ist. In diesem Fall kann der Modellelementinstanz über narrowerThan eine unspezifischere Domäneninstanz zugeordnet werden. So werden im Beispiel von **Abb. 3** über narrowerThan die beiden Modellelementinstanzen f2 („Auftragsbestätigung senden“) und f3 („Auftragsablehnung senden“) der Domäneninstanz Auftragsrückmeldung zugeordnet.

Insgesamt zeigt **Abb. 3** beispielhaft die Annotation eines EPK-Modells unter Verwendung der beiden Relationen equivalentTo und narrowerThan. Die im oberen Teil der Abbildung verwendeten Ontologieklassen und deren Beziehungen bauen auf SUMO auf (vgl. Abschnitt 4.1).

4.4 Integration von Regeln

Mit der bisher beschriebenen Repräsentation und Annotation ist eine Verknüpfung von Prozessmodellen mit Ontologien möglich. Die in der Ontologie enthaltenen Definitionen können hierbei zur terminologischen Vereinheitlichung aber auch zur Anfrage und Validierung verwendet werden. Allerdings können nicht alle in einer Domäne relevanten Zusammenhänge durch die Definition von Begriffen und Relationen repräsentiert werden. Insbesondere zur Repräsentation und zur Schlussfolgerung bei komplexen Abhängigkeiten zwischen mehreren Begriffen bedarf es einer Erweiterung um Regeln.

Die im Rahmen der Semantischen Prozessmodellierung benötigten Regeln werden in Anlehnung an Boley et al. (2007, S. 273 ff) in deduktive und normative

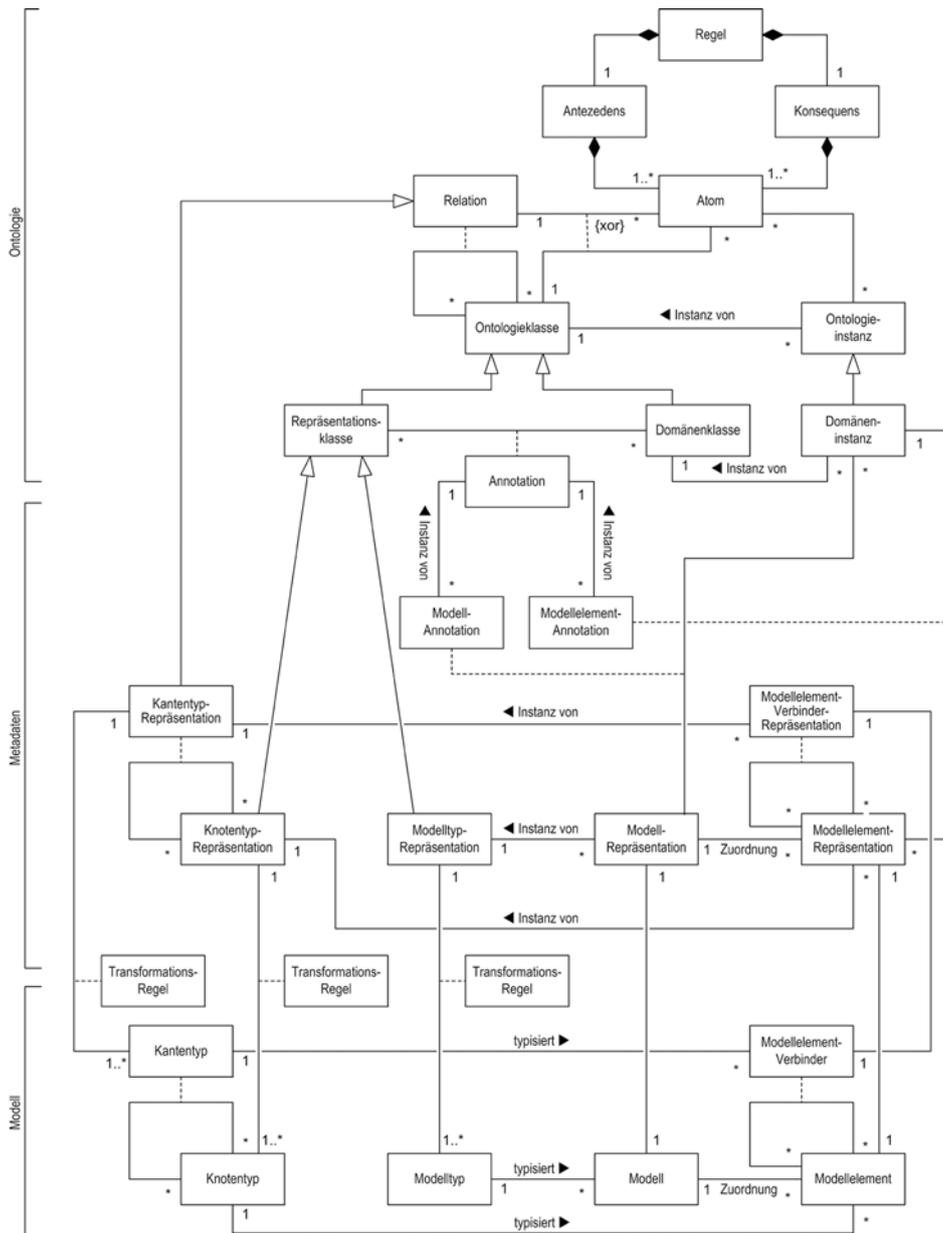


Abb. 4 Informationsmodell der Semantischen Prozessmodellierung

Regeln unterschieden. Deduktive Regeln (auch: Ableitungsregeln) dienen dazu, auf der Basis vorhandener Fakten durch die Nutzung logischer Schlussfolgerungen neue Fakten zu gewinnen. So kann bspw. bei der Interpretation des Prozesses aus **Abb. 3** in der gezeigten Ontologie unter der Prämisse, dass die von einem Sub-Prozess unterstützten Ziele gleichzeitig zur Menge der Ziele gehören, die vom übergeordneten (Super-)Prozess unterstützt werden, eine entsprechende Wenn-dann-Regel formuliert werden: „Wenn es zwei Aktivitäten x und y gibt und x ein Subprozess von y ist und ein Ziel z unterstützt, dann folgt daraus, dass y ebenfalls das Ziel z unterstützt“. Mit dieser Regel kann die Frage „Welche Ziele werden von der Akti-

viät Auftragsbearbeitung unterstützt?“ mit „Kundenzufriedenheit“ beantwortet werden. Dies rührt daher, dass die Relation *supportsObjective*, die zwischen der Sub-Aktivität *Auftragsrückmeldung* und dem Ziel *Kundenzufriedenheit* besteht, auf die übergeordnete Aktivität der Auftragsbearbeitung übertragen wird (**Abb. 3**). Die Regel lässt sich wie folgt formulieren:
 $Activity(?x)$
 $\wedge Activity(?y)$
 $\wedge subProcess_directly(?x,?y)$
 $\wedge supportsObjective(?x,?z)$
 $\Rightarrow supportsObjective(?y,?z)$
 Der Wenn-Teil der Regel wird auch als *Antezedens* bezeichnet und beschreibt eine bestimmte Situation. Der Dann-Teil wird auch als *Konsequens* bezeichnet und

gibt eine geeignete Folgerung, Reaktion oder Entscheidung für die im Wenn-Teil beschriebene Situation an.

Normative Regeln (auch: strukturelle Regeln) werden verwendet, um Bedingungen für die von einer Anwendung verwendeten Daten oder der von ihr verwendeten Logik auszudrücken (Boley et al. 2007, S. 274). Sie können weiter in Konsistenz- und Integritätsregeln unterschieden werden. Als Konsistenzregeln werden Regeln bezeichnet, die zur Aufrechterhaltung der Widerspruchsfreiheit der Ontologie und der aus ihr ableitbaren Fakten eingesetzt werden. Im Rahmen der Semantischen Prozessmodellierung ist bspw. zu vermeiden, dass ein Prozessmodell und einzelne Elemente dieses Modells

über die Relation `equivalentTo` mit derselben Domäneninstanz annotiert werden. Als Integritätsregeln werden in Anlehnung an semantische Integritätsbedingungen im Datenbankumfeld Regeln verstanden, die zur Aufrechterhaltung der semantischen Korrektheit der Ontologie und der aus ihr ableitbaren Fakten eingesetzt werden. Ein Beispiel für eine Integritätsregel kann in Bezug auf die inhaltliche Validierung der in der Ontologie repräsentierten Prozessmodelle vor dem Hintergrund von **Abb. 3** angegeben werden. So kann in einem Unternehmen zur Verbesserung der Kommunikation mit dem Kunden die Richtlinie bestehen, dass nach der Auftragsprüfung in jedem Fall der Kunde über den Ausgang der Prüfung zu informieren ist.

Die Formulierung von Regeln ist im Sprachumfang zahlreicher nicht web-basierter Ontologiesprachen, wie bspw. OCML und Ontolingua, vorgesehen. Die hier verwendete OWL unterstützt die Formulierung von Regeln über die Erweiterungen der Semantic Web Rule Language (SWRL) (Horrocks et al. 2004).

4.5 Generalisierung des Ansatzes

Der besondere Anspruch der in diesem Abschnitt zu vollziehenden Modellkonstruktionen besteht darin, die zuvor entworfene spezielle Prozessmodellierung zu generalisieren, damit die erworbenen Erkenntnisse auf weitere Modellierungsmethoden und -sprachen sowie Situationen, in denen mit semantisch annotierten Modellen gearbeitet wird, übertragen werden können. Gegenüber der in den vorhergehenden Abschnitten an Beispielen orientierten Darstellung werden nun im Rahmen einer graphentheoretischen Interpretation die „Bausteine“ semantischer Geschäftsprozesse mit ihren Beziehungen betrachtet. Diese Betrachtung bezieht sich auf Aussagen, in denen die Elemente eines *generellen* Geschäftsprozesses, also ohne einen bestimmten Anwendungsbezug, erfasst werden. Als Modellierungssprache wird das UML-Klassendiagramm gewählt.

Aufgrund der graphbasierten Interpretation der Prozessrepräsentationen handelt es sich bei dem Informationsmodell der Semantischen Prozessmodellierung nicht um das Metamodell einer – wie auch immer gearteten – semantisch erweiterten Modellierungssprache, oder gar einer ontologiebasierten EPK, sondern um ein

objektsprachliches Klassenmodell, das unabhängig von einer bestimmten Sprache spezifiziert, wie Geschäftsprozesse semantisch annotiert und repräsentiert werden können (**Abb. 4**). Das Informationsmodell ist in die drei Ebenen „Modell“, „Metadaten“ und „Ontologie“ unterteilt. Die vertikale Schichtung dieser Ebenen im Datenmodell bedeutet jedoch keine Abstraktionsbeziehung. Sie ist vielmehr Ausdruck der verbindenden Funktion, die der mittlere Bereich der Metadaten zwischen Modell einerseits und Ontologie andererseits einnimmt. Als Metadaten werden hier Daten bezeichnet, die durch die Annotation eines Modells mit Elementen der Ontologie entstehen.

Auf der Modellebene werden konkrete Modelle durch die Klassenmodell, Modellelement und Modellelement-Verbinder repräsentiert. Jedes Modellelement (z. B. „Auftrag anlegen“ für eine EPK-Funktion oder eine BPMN-Aktivität) wird einem Knotentyp zugeordnet (z. B. EPK-Funktion oder BPMN-Aktivität) und ist mindestens einem Modell zugeordnet (z. B. „Auftrags erfassung“), das seinerseits durch einen Modelltyp (z. B. „EPK“ oder „BPMN“) typisiert wird. Es kann darüber hinaus mit anderen Modellelementen über einen Modellelement-Verbinder in Beziehung stehen (z. B. der EPK-Funktion „Auftrag anlegen“ folgt das EPK-Ereignis „Auftrag ist angelegt“), der durch einen Kanten typisiert wird (z. B. EPK-Flussbeziehung). Die Repräsentation der Modelle auf der Metadatenebene erfolgt mit Hilfe von Transformationsregeln, die angeben, wie Elemente der Modellebene in Elemente der Metadatenebene überführt werden.

Auf der Metadatenebene werden Modelle durch die Klassenmodell-Repräsentation, Modellelement-Repräsentation und Modellelement-Verbinder-Repräsentation dargestellt, die aus den Klassenmodelltyp-Repräsentation, Knotentyp-Repräsentation und Kanten typ-Repräsentation mittels der Transformationsregeln instanziiert werden. Eine Verbindung zur Ontologieebene wird über die Beziehungen mit den Assoziationsklassen Modell-Annotation und Modellelement-Annotation hergestellt.

Die zentrale Klasse Ontologieklassen repräsentiert alle Klassen, die in einer Ontologie enthalten sind. Ontologieklassen können miteinander in einer Beziehung stehen (Assoziationsklasse Relation), die wiederum Eigenschaften aufweisen

kann, wie bspw. Domäne, Wertebereich oder Bezeichnung. Ontologieklassen werden unterschieden in Klassen, die Konstrukte einer semiformalen Sprache repräsentieren, und in solche, die Konstrukte einer Domäne repräsentieren. Dieser Sachverhalt wird durch Spezialisierungsbeziehungen zwischen der Klasse Ontologieklassen und den entsprechenden Klassen Repräsentationsklasse bzw. Domänenklasse ausgedrückt.

Auf Basis der Ontologieklassen und -instanzen können weiter Regeln formuliert werden. Eine Regel besteht aus einem Antezedens und einem Konsequens. Beide können nicht sinnvoll außerhalb der sie beinhaltenden Regel existieren, dieser Sachverhalt wird durch eine Kompositionsbeziehung zwischen der Klasse Regel und den Klassen Antezedens und Konsequens verdeutlicht. In ähnlicher Weise bestehen der Antezedens und der Konsequens einer Regel aus einem Atom oder mehreren Atomen, die ebenfalls nicht sinnvoll außerhalb des jeweiligen Teils der Regel existieren können. Ein Atom kann sich entweder auf eine Relation oder auf eine Ontologieklassen beziehen. Dies wird durch zwei alternative Assoziationen repräsentiert, zwischen denen eine {xor}-Restriktion besteht. Darüber hinaus kann ein Atom sich durch die Verwendung von Variablen auch auf Ontologieinstanzen beziehen.

4.6 Übertragung auf andere Modellierungssprachen

Das gezeigte Informationsmodell ist sprachneutral ausgerichtet und kann auf weitere Modellierungssprachen übertragen werden. **Abb. 5** verdeutlicht dies am Beispiel der BPMN (OMG 2006). Im unteren Bereich ist ein BPMN-Modell abgebildet, das durch entsprechende Instanzen in der darüberliegenden Ontologie repräsentiert wird. Im linken Bereich werden Teile des Informationsmodells aus **Abb. 4** gezeigt. Die gestrichelten und gerichteten Kanten, die von den Klassen des Informationsmodells in der linken Hälfte zu den einzelnen Elementen in der rechten Hälfte verlaufen, verdeutlichen Instanzierungsbeziehungen. So sind alle in der Ontologie enthaltenen Klassen, die zur Repräsentation von Sprachkonstrukten von Prozessbeschreibungssprachen verwendet werden, aus der Klasse Knotentyp-Repräsentation des Informationsmodells instanziiert und die zwischen ihnen bestehenden Bezie-

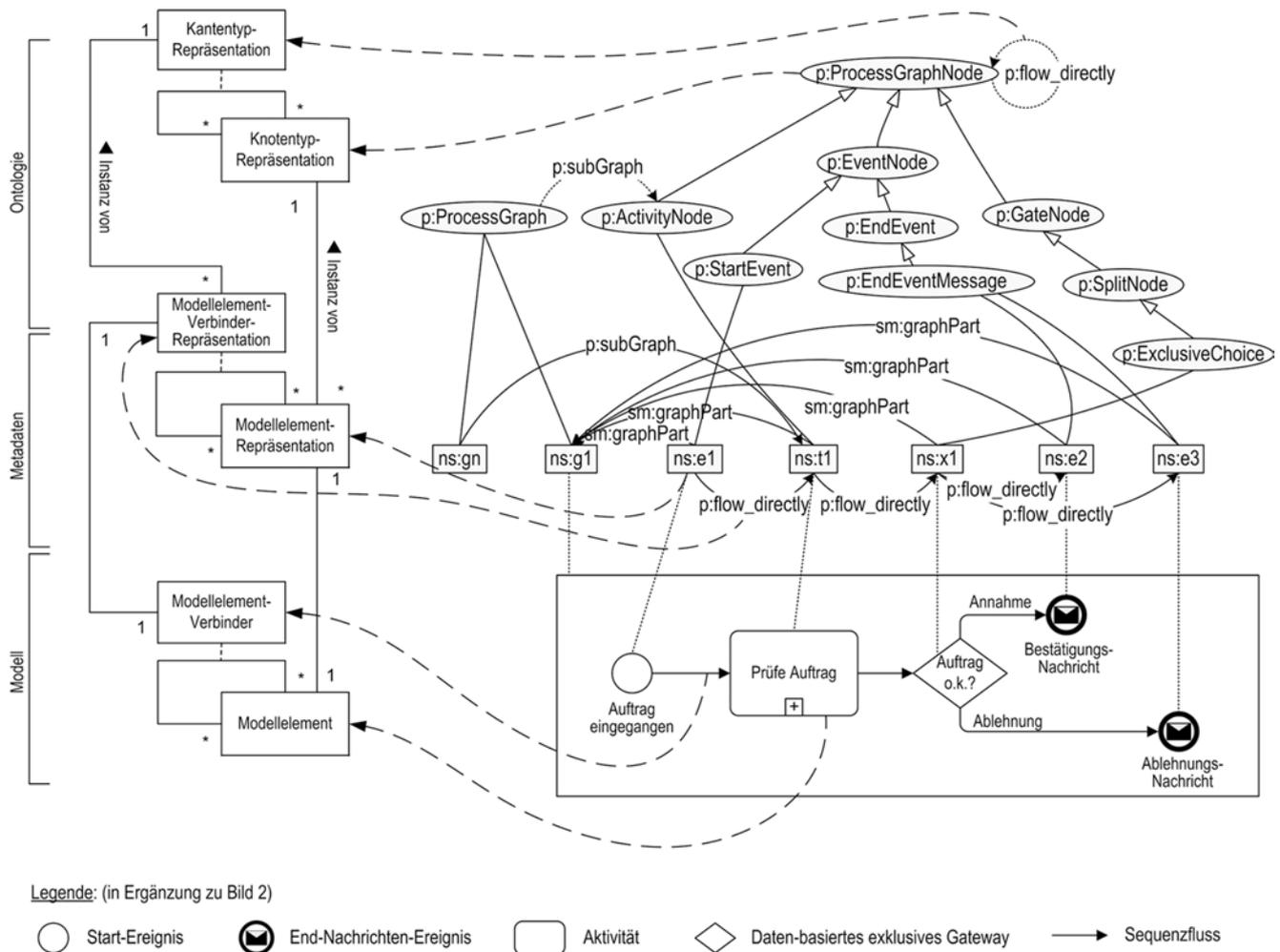


Abb. 5 Übertragung des Ansatzes auf die Sprache BPMN

hungen aus der Klasse Kantentyp-Repräsentation. Die Repräsentation konkreter Modellelemente und ihrer Relationen (Verbinde) in der Ontologie sind aus den Klassen Modellelement-Repräsentation und Modellelement-Verbinde-Repräsentation instanziiert. Die Elemente des zugrundeliegenden BPMN-Modells sind Instanzen der Klassen Modellelement und Modellelement-Verbinde.

Wie gegenüber **Abb. 2** erkennbar ist, sind zur ontologiebasierten Repräsentation eines BPMN-Modells nur geringfügige Erweiterungen in der Ontologie erforderlich. Diese beziehen sich auf zusätzliche Ereignistypen, wie bspw. das Endereignis vom Typ „Nachricht“, das die Sprache BPMN zur Verfügung stellt. Durch die graphentheoretische Interpretation erfordert die Übertragung des vorgestellten Ansatzes auf andere Prozessbeschreibungssprachen lediglich das Anlegen neuer Subklassen für entsprechende Knotentypen in der Ontologie.

In Bezug auf die Anwendung des vorgestellten Ansatzes kommt der IT-Unterstützung sowohl zur Modellierung und Annotation als auch zur maschinellen Verarbeitung der Semantik im Rahmen der Anfrage und Validierung von Prozessmodellen eine besondere Bedeutung zu. Im Folgenden werden daher zunächst eine Architektur und anschließend darauf aufbauend eine prototypische Implementierung vorgestellt.

5 IT-Unterstützung der Semantischen Prozessmodellierung

5.1 Systemarchitektur

Die wesentlichen Komponenten und deren Zusammenwirken zur IT-Unterstützung der Semantischen Prozessmodellierung zeigt die Systemarchitektur (**Abb. 6**).

Zur Erstellung und Anpassung der für die Annotation verwendeten Ontologie dient ein Editor, der mit dem zur Speicherung von Ontologien, Regeln und Modellen verwendeten Repository A verbunden ist. Den Kernbereich der Architektur bildet ein in Schichten unterteiltes sBPM-Rahmenwerk (sBPM = semantic Business Process Management). In der untersten Schicht stellt eine Komponente zur semantischen Datenverarbeitung für die darüber liegenden Komponenten grundlegende Funktionalitäten bereit, die das Laden der Daten, deren Interpretation mit einer Inferenzmaschine sowie deren Änderung und Speicherung betreffen. Die über dieser Schicht angeordneten Komponenten stellen die zur IT-Unterstützung der Semantischen Prozessmodellierung benötigten Kernfunktionalitäten bereit. Dies sind die von Modellierern während der Konstruktion und Anpassung von Modellen genutzten Funktionalitäten zur Anfrage, Validierung und Annota-

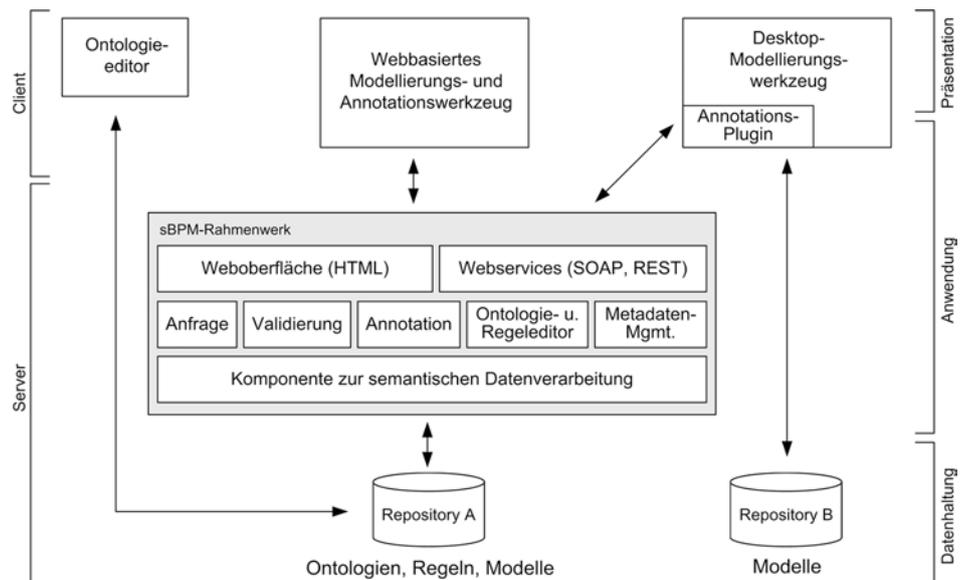


Abb. 6 Systemarchitektur eines Werkzeugs zur Unterstützung der Semantischen Prozessmodellierung

tion von Prozessmodellen. Die letztgenannte Komponente zur Annotation kann dem Modellierer hierbei auf der Basis von Abbildungen der natürlichen Sprache auf Ontologiebegriffe (Niles u. Pease 2003) Vorschläge zur Annotation unterbreiten. Weitere Elemente des sBPM-Rahmenwerks sind Komponenten zum Metadatenmanagement sowie ein integrierter Ontologie- und Regeleditor. Im Unterschied zum externen Ontologieeditor, der zur Erstellung der verwendeten Ontologie eingesetzt wird, dient dieser lediglich der Erweiterung der Ontologie, womit Modellkonstrukteure in die Lage versetzt werden, die Ontologie entsprechend ihrer Bedürfnisse während des Annotationsvorgangs, bspw. durch das Ableiten von Subklassen oder das Hinzufügen neuer Instanzen, zu erweitern.

Zur Nutzung der Funktionalitäten durch Clients über ein Netzwerk müssen geeignete Schnittstellen zum Zugriff existieren. Diese sind in der darüber liegenden Schicht mit einer Weboberfläche (Verwendung der Standards HTML bzw. XHTML) und einer Webservice-Schnittstelle (Verwendung der Standards SOAP/XML oder REST) gegeben.

Beim Zugriff auf die Funktionalitäten der Anwendungsebene durch einen Client lassen sich zwei Szenarien unterscheiden. Zum einen kann der Zugriff über ein webbasiertes Modellierungs- und Annotationswerkzeug erfolgen (Variante A). In diesem Fall können Ontologien, Regeln und Modelle in einem gemeinsamen Speicher vorgehalten werden. Semiformale Modelle können somit von einem einheitlichen Werkzeug erstellt, transfor-

miert und annotiert werden. Zum anderen kann der Zugriff über ein Plugin eines Desktop-Modellierungswerkzeugs erfolgen (Variante B). In diesem Fall können die Modelle weiterhin im eigenen Repository des jeweiligen Modellierungswerkzeugs gespeichert werden (Repository B), während deren ontologiebasierte Repräsentationen separat davon in Repository A gespeichert werden. Bei der Verwendung eines Desktop-Modellierungswerkzeugs mit Plugin ist davon auszugehen, dass nicht das gesamte Spektrum der Funktionalitäten zur Semantischen Prozessmodellierung durch das Plugin bereitgestellt werden kann, sodass im Ergebnis zwei Werkzeuge verwendet werden müssen.

5.2 Implementierung einer Testumgebung

Zur Demonstration der grundsätzlichen Anwendbarkeit der entwickelten Konzeption wurde auf Basis der Systemarchitektur die Testumgebung sBPMQuery entwickelt, die auf einem Modellierungswerkzeug zur Modellerstellung und einem Ontologieeditor zur semantischen Annotation aufbaut und darüber hinaus eine Serverkomponente mit Weboberfläche zur Anfrage und Validierung von Modellen enthält. Zur Implementierung wurden *EPC Tools* als Prozessmodelleditor, *Protégé* als Ontologieeditor und *Jena* als Semantic-Web-Rahmenwerk für die Serverkomponente verwendet. Letztere wurde zudem um die Inferenzmaschine *Pellet* erweitert. **Abb. 7** veranschaulicht die Beziehungen der Komponenten anhand der jeweiligen Benutzungsschnittstellen.

Mit Hilfe der Testumgebung kann ein Ausgangsmodell (1) zunächst automatisiert durch ein XSLT-Skript in dessen ontologiebasierte Repräsentation transformiert werden, die in den Ontologieeditor (2) importiert wird. Nachdem im Ontologieeditor die Annotation des Modells manuell erfolgt ist, werden die daraus resultierenden semantischen Metadaten aus dem Ontologieeditor in den Speicher der Serverkomponente (3) importiert. Über eine Anfrage-Oberfläche (4), die als Client in einem Browser abläuft, kann mit der Serverkomponente kommuniziert werden. Hierbei können die durch die semantischen Metadaten repräsentierten Modelle sowohl validiert als auch mit der Anfragesprache SPARQL (Prud'hommeaux u. Seaborne 2005) durchsucht werden. Bei beiden Funktionalitäten kann zusätzlich eine Inferenzmaschine oder Regelausführungseinheit hinzu geschaltet werden, sodass eine Interpretation der Daten auf einer semantischen Ebene möglich wird und nicht explizit im Modell enthaltene Fakten zum Anfragezeitpunkt geschlossen werden können. So ist etwa das in **Abb. 7** auf der rechten Seite zu sehende Ergebnis der Beispielanfrage um das automatisiert geschlossene Faktum erweitert, dass der Webservice zur Produktkonfiguration *ProductConfiguratorService* von der Resource *ProductDatabase* abhängig ist (im Kontext der Testumgebung wurden die englischsprachigen Bezeichnungen der Ontologiebegriffe beibehalten). Dies kann durch eine transitive Relation *depends* geschlossen werden, die in dem zugrunde liegenden Beispiel aus **Abb. 3** zwischen

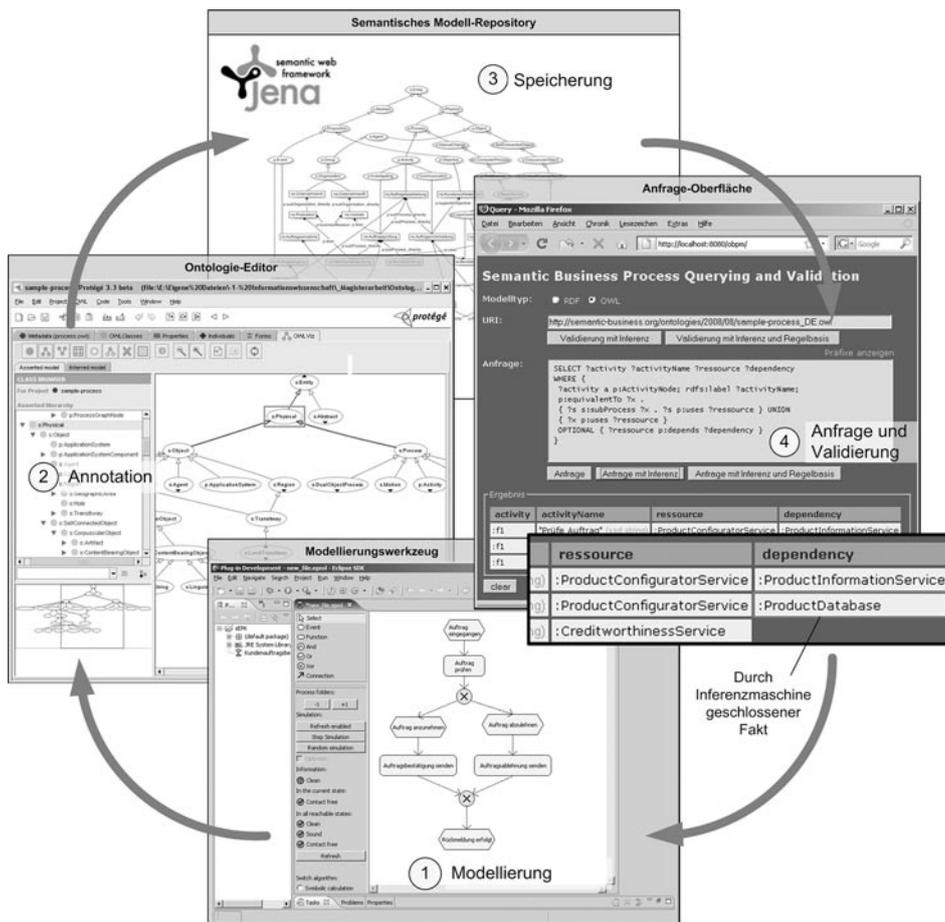


Abb. 7 Anfragen an semantisch erweiterte Modellierungswerkzeuge

den Webservices Produktkonfigurator, Produktinformation und der Datenbank ProduktDB besteht. Die hierbei gewonnenen Erkenntnisse können vom Modellierer zu weiteren Änderungskonstruktionen am Modell (1) verwendet werden.

6 Nutzenpotenziale der Semantischen Prozessmodellierung

Mit Hilfe von Ontologien kann die in Prozessmodellen enthaltene Semantik formal repräsentiert werden. Hieraus resultieren für den in diesem Beitrag entwickelten Ansatz der Semantischen Prozessmodellierung folgende Nutzenpotenziale:

- Verteilte Modellierung:** Mit Hilfe der Semantischen Prozessmodellierung kann das gegenseitige Verständnis der an der Analyse und Gestaltung prozessorientierter Informationssysteme beteiligten Akteure gefördert werden, da die natürlichsprachlichen Bezeichnungen der Modellelemente mit den in einer Ontologie von einer Experten-gruppe formal definierten Begriffen einer Domäne hinterlegt werden. Insbesondere in der arbeitsteiligen und

verteilten Modellentwicklung kann dadurch eine einheitliche Interpretation der Modelle erreicht werden.

- Modellmanagement:** Eine Modellverwaltung, die auf einer semantischen Annotation von Prozessmodellen aufbaut, ist für das Modellmanagement von hohem Nutzen. Sie systematisiert und erleichtert den Zugang zu den Modellen und ist zur Unterstützung der Suche und Selektion von Prozessmodellen geeignet. Zentrale Argumente sind hierbei *erstens* die durch die maschinell verarbeitbare Semantik verbesserte „Passgenauigkeit“ eines Modells in Bezug auf die Kriterien einer Suchanfrage und *zweitens* die Vollständigkeit des Ergebnisses einer Suchanfrage, wobei durch Verfahren des maschinellen Schließens nicht alle zur Beantwortung einer Suchanfrage erforderlichen Informationen durch einen Nutzer explizit angegeben werden müssen. Dies reduziert den Aufwand zur Adaption der insbesondere von Softwarehäusern und Beratungsunternehmen eingesetzten Referenzmodelle.

- IT-Business-Alignment:** Im Rahmen des Entwurfs und der Realisierung von Anwendungssystemen stehen Unternehmen vor der zentralen Herausforderung, sowohl fachliche als auch technologische Aspekte berücksichtigen zu müssen. Problematisch ist hierbei, dass die fachliche und die technische Modellierung i. d. R. nicht miteinander gekoppelt sind. Aufgrund dieser bestehenden Lücke ist keine konsistente Überführung fachlicher Anforderungen in unterstützende IT-Systeme gewährleistet. Die hierzu notwendige Abstimmung zwischen IT- und Fachabteilungen wird durch die Verwendung eines gemeinsam akzeptierten Vokabulars, das Bestandteil der Semantischen Prozessmodellierung ist, erheblich vereinfacht. Damit kann letztlich ein höherer Geschäftswertbeitrag durch IT im Sinne eines IT-Business-Alignment erbracht werden.
- Compliance:** Durch die Verwendung von Regeln, die auf einer Ontologie aufbauen, eröffnen sich neue Möglichkeiten der Validierung von Modellen. Hierdurch können inadäquate Modelle bereits zum Zeitpunkt der Modeller-

stellung erkannt werden. Fehlerhafte Repräsentationen können sich bspw. durch die Nichterfüllung von Anforderungen ergeben, die in Form von Gesetzen und Richtlinien (z. B. Sarbanes-Oxley Act (SOX) zur Unternehmensberichterstattung, Normen ISO/IEC 27001:2005 zur Informationssicherheit und DIN ISO 154891 zur Information und Dokumentation) aber auch als unternehmensinterne Maßgaben existieren (Compliance). Durch die mit der Semantischen Prozessmodellierung erreichbare inhaltliche Prüfung kann die Weitergabe fehlerhafter Modelle an die der Modellkonstruktion nachgelagerte Phasen vermieden werden.

7 Limitationen und Ausblick

Der Aufwand zur Erweiterung von Modellierungssprachen zwecks Unterstützung semantischer Technologien kann aus konzeptioneller Hinsicht sehr hoch sein und den Nutzen, der durch eine Modellsuche oder -validierung im Rahmen von Modellierungsprojekten erzielbar ist, durchaus überkompensieren. Die Forschung muss sich in diesem Zusammenhang zukünftig Evaluationsfragen widmen, bspw. unter Zuhilfenahme von Kriterien zur ökonomischen Bewertung des Einsatzes von Ontologien oder semantisch annotierter Modelle.

Darauf aufbauend ist zu evaluieren, inwiefern durch die Nutzung der Testumgebung sBPMQuery eine Verbesserung der Modellentwicklung und -anwendung in der Unternehmenspraxis erreicht werden kann. Hierbei ist zu beachten, dass das zentrale Artefakt dieses Beitrags – das Informationsmodell der Semantischen Prozessmodellierung – nicht das Implementierungsmodell des gezeigten Prototyps ist, folglich wird durch die Umsetzung der Testumgebung keine Evaluation des Informationsmodells erreicht. Auch dies kann Gegenstand weiterer Forschungsarbeiten sein.

Ergänzend zu einer Modellevaluation ist die Einbettung der Semantischen Prozessmodellierung in vorhandene Organisationsstrukturen zu erforschen. So ist bspw. zu klären, wie eine semantische Annotation organisatorisch in der Unternehmenspraxis verankert werden kann, d. h. welche Mitarbeiter die Annotation durchführen, wer die Ontologie pflegt und anpasst sowie durch welche Mitarbeiter eine Syn-

Zusammenfassung / Abstract

Oliver Thomas, Michael Fellmann M.A.

Semantische Prozessmodellierung – Konzeption und informationstechnische Unterstützung einer ontologiebasierten Repräsentation von Geschäftsprozessen

Eine Erweiterung von Prozessmodellierungssprachen wird entworfen, mit deren Hilfe die in natürlicher Sprache formulierte Semantik der Bezeichner von Prozessmodellelementen durch formale Begriffe einer Ontologie repräsentiert werden kann. Diese Verknüpfung von semiformalen Modellen einerseits und formalen Ontologien andererseits wird als Semantische Prozessmodellierung charakterisiert. Das Konzept wird beispielhaft anhand der Sprachen EPK (Ereignisgesteuerte Prozesskette), BPMN (Business Process Modeling Notation) und OWL (Web Ontology Language) eingeführt und in Form eines Informationsmodells generalisiert. Durch die Formalisierung der modellelementbezogenen Semantik im Zusammenspiel mit der Nutzung von Inferenzmaschinen können die Suche in Modellen verbessert und neue Möglichkeiten der inhaltlichen Validierung erschlossen werden. Mithilfe einer Systemarchitektur und eines Prototyps wird die Einbettung des Ansatzes in bestehende IT-gestützte Arbeitsumgebungen von Modellierern demonstriert. Es wird deutlich, dass durch Modellierungs-Tools, die um semantische Technologien angereichert sind, Vorteile in den Bereichen verteilte Modellierung, Modellmanagement, IT-Business-Alignment und Compliance erzielbar sind.

Schlüsselwörter: Prozessmanagement, Prozessmodellierung, Modellierungssprachen, Semantik, Ontologien, Ontologiesprachen

Semantic Process Modeling – Design and Implementation of an Ontology-based Representation of Business Processes

An extension of process modeling languages is designed which allows representing the semantics of model element labels which are formulated in natural language by using concepts of a formal ontology. This combination of semiformal models with formal ontologies will be characterized as semantic process modeling. The approach is exemplarily applied to the languages EPC (Event-driven Process Chain), BPMN (Business Process Modeling Notation) and OWL (Web Ontology Language) and is generalized by means of an information model. The proposed formalization of the semantics of individual model elements in conjunction with the usage of inference engines allows the improvement of query functionalities in modeling tools and enables new possibilities of model validation. The integration of the approach in the IT-based work environments of modelers is demonstrated by a system architecture and a prototypical implementation. Evidently, advantages in the areas of modeling, model management, business/IT alignment, and compliance can be achieved by the application of modeling tools augmented with semantic technologies.

Keywords: Process management, Process modeling, Modeling languages, Semantics, Ontologies, Ontology languages

chronisation annotierter Modelle vorzunehmen ist, wenn sich Elemente im Geschäftsprozess oder Begriffe in der Ontologie verändern.

Literatur

- Abramowicz W, Filipowska A, Kaczmarek M, Kaczmarek T (2007) Semantically enhanced business process modelling notation. In: Hepp M et al (Hrsg) Proceedings of SBPM 2007. Innsbruck, S 88–91
- Ahlemann F, Teuteberg F, Brune G (2006) Ontologie-basierte Attributierung von Informationsmodellen: Grundlagen und Anwendungsbereiche. In: Teuteberg F, Ahlemann F (Hrsg) ISPRI-Arbeitsbericht, Nr. 01/2006, Universität Osnabrück
- AIIM (Hrsg) (2007) BPM: Not just workflow anymore. Silver Spring
- Bögl A, Schrefl M, Pomberger G, Weber N (2008) Semantic annotation of EPC models in engineering domains by employing semantic patterns. In: Cordeiro J, Filipe J (Hrsg) Proceedings of ICEIS 2008. Barcelona, S 106–115
- Boley H, Kifer M, Patranjan PL, Polleres A (2007) Rule interchange on the web. In: Antoniou G et al (Hrsg) Reasoning web. 3rd international summer school 2007. Springer, Heidelberg, S 269–309
- Bray T, Hollander D, Layman A, Tobin R (Hrsg) (2006) Namespaces in XML 1.1. W3C
- Brockmans S, Ehrig M, Koschmider A, Oberweis A, Studer R (2006) Semantic alignment of business processes. In: Manolopoulos Y et al (Hrsg) Proceedings of ICEIS 2006. Paphos, S 191–196
- Cabral L, Domingue J, Motta E, Payne TR, Hakimpour F (2004) Approaches to semantic web services: An overview and comparisons. In: Bussler C et al (Hrsg) The semantic web: Research and applications: Proceedings of ESWWS 2004. Springer, Berlin, S 225–239
- Cardoso J, Sheth AP (2005) Introduction to semantic web services and web process composition. In: Semantic web services and web process composition. 1st International Workshop, SWSWPC 2004, San Diego, S 1–13
- Casely-Hayford L (2005) A comparative analysis of methodologies, tools and languages used for building ontologies. CCLRC, Swindon
- Desel J (2002) Model validation – A theoretical issue? In: Esparza J, Lakos C (Hrsg) Application and theory of Petri nets. Springer, Heidelberg, S 23–43
- Dietz JLG (2006) The deep structure of business processes. *Comm. ACM* 49(5):58–64
- Fensel D, Hendler J, Lieberman H, Wahlster W (Hrsg) (2003) Spinning the semantic web. MIT Press, Cambridge
- Fox MS (1992) The TOVE project: A common-sense model of the enterprise. In: Belli F, Radermacher FJ (Hrsg) Industrial and engineering applications of artificial intelligence and expert systems. Springer, London, S 25–34
- Gómez-Pérez A, Fernández-López M, Corcho O (2004) Ontological engineering. Springer, London
- Gómez-Pérez A, Manzano-Macho D (Hrsg) (2003) A survey of ontology learning methods and techniques. Universidad Politécnica de Madrid
- Green PF (1996) An ontological analysis of information systems analysis and design (ISAD) grammars in upper case tools. PhD Thesis, University of Queensland, Brisbane
- Green PF, Rosemann M (2000) Integrated process modeling: An ontological evaluation. *Information Systems* 25(2):73–87
- Green PF, Rosemann M, Indulska M (2005) Ontological evaluation of enterprise systems interoperability using ebXML. *IEEE Transactions on Knowledge and Data Engineering* 17(5):713–725
- Gruber TR (1993) A translation approach to portable ontology specifications. *Knowledge Acquisition* 5(2):199–220
- Hahn A (2005) Integration verteilter Produktmodelle durch Semantisch-Web-Technologien. *WIRTSCHAFTSINFORMATIK* 47(4):278–284
- Heinrich B, Bewernik M-A, Henneberger M, Kramer A, Lautenbacher F (2008) SEMPMA – Ein Ansatz des Semantischen Prozessmanagements zur Planung von Prozessmodellen. *WIRTSCHAFTSINFORMATIK* 50(6):445–460
- Henneberger M, Heinrich B, Lautenbacher F, Bauer B (2008) Semantic-based planning of process models. In: Bichler M et al (Hrsg) Multikonferenz Wirtschaftsinformatik 2008. GITO, Berlin, S 1677–1689
- Hepp M (2005) eClassOWL: A fully-fledged products and services ontology in OWL. In: Poster Proceedings of ISWC2005. Galway
- Hepp M, Leymann F, Domingue J, Wahler A, Fensel D (2005) Semantic business process management: A vision towards using semantic web services for business process management. In: Proceedings of IEEE ICEBE 2005. Beijing, S 535–540
- Hepp M, Roman D (2007) An ontology framework for semantic business process management. In: Oberweis A et al (Hrsg) eOrganisation: Service-, Prozess-, Market-Engineering: 8. Internationale Tagung Wirtschaftsinformatik, Band 1. Universitäts-Verlag, Karlsruhe, S 423–440
- Horrocks I, Patel-Schneider PF, Boley H, Tabet S, Groszof B, Dean M (2004) SWRL: A semantic web rule language. W3C
- Keller G, Nüttgens M, Scheer A-W (1992) Semantische Prozeßmodellierung auf der Grundlage „Ereignisgesteuerter Prozeßketten (EPK)“. In: Scheer AW (Hrsg) Veröffentlichungen des Instituts für Wirtschaftsinformatik, Nr. 89. Universität des Saarlandes, Saarbrücken
- Kindler E (2006) On the semantics of EPCs: Resolving the vicious circle. *Data & Knowledge Engineering* 56(1):23–40
- Koschmider A, Ried D (2005) Semantische Annotation von Petri-Netzen. In: Workshop für Algorithmen und Werkzeuge für Petri-Netze (AWPN'05). Humboldt-Universität Berlin, S 66–71
- Kuropka D (2004) Modelle zur Repräsentation natürlicher sprachlicher Dokumente. Logos, Berlin
- Lautenbacher F, Bauer B (2006) Semantic reference- and business process modeling enables an automatic synthesis. In: Hinkelmann K et al (Hrsg) Proceedings of SBPM 2006. Budva, S 89–100
- Lin Y (2008) Semantic annotation for process models: Facilitating process knowledge management via semantic interoperability. PhD Thesis, Norwegian University of Science and Technology, Trondheim
- Lin Y, Ding H (2005) Ontology-based semantic annotation for semantic interoperability of process models. In: Proceedings of CIMCA-IAWTIC, 06, vol 1. IEEE, Washington, DC, S 162–167
- Lin Y, Strasunskas D (2005) Ontology-based semantic annotation of process templates for reuse. In: Proceedings of 10th CAISE/IFIP8.1/EU-NO international workshop on evaluation of modeling methods in system analysis and design (EMMSAD05). Porto, S 162–167
- Miles A, Bechhofer S (2008) SKOS simple knowledge organization system reference. W3C
- Niles I, Pease A (2001) Towards a standard upper ontology. In: Welty C, Smyth B (Hrsg) Proceedings of the 2nd international conference on formal ontology in information systems (FOIS-2001). Ogunquit, 2001, S 2–9
- Niles I, Pease A (2003) Linking lexicons and ontologies: Mapping WordNet to the suggested upper merged ontology. In: Proceedings of the 2003 international conference on information and knowledge engineering (IKE 03). Las Vegas, S 23–26
- OMG (Hrsg) (2006) Business process modeling notation specification. OMG, Needham
- Ogden CK, Richards IA (1923) The meaning of meaning, 1. Aufl. Kegan Paul, London
- Ortner E (1997) Methodenneutraler Fachentwurf. Teubner, Stuttgart
- Prud'hommeaux E, Seaborne A (Hrsg) (2005) SPARQL query language for RDF. W3C
- Rosemann M (1996) Komplexitätsmanagement in Prozessmodellen. Gabler, Wiesbaden
- Rosemann M, Schwegmann A (2002) Vorbereitung der Prozessmodellierung. In: Becker J et al (Hrsg) Prozessmanagement, 3. Aufl. Springer, Heidelberg, S 47–94
- Smith MK, Welty C, McGuinness DL (Hrsg) (2004) OWL web ontology language guide. W3C
- Soffer P, Golany B, Dori D, Wand Y (2001) Modelling off-the-shelf information systems requirements: An ontological approach. *Requirements Engineering* 6(3):183–199
- Studer R, Benjamins VR, Fensel D (1998) Knowledge engineering: Principles and methods. *Data & Knowledge Engineering* 25(1–2):161–197
- Thomas O, Fellmann M (2007) Semantic business process management: Ontology-based process modeling using event-driven process chains. *International Journal of Interoperability in Business Information Systems* 2(1):29–44
- Uschold M, King M, Moralee S, Zorgios Y (1998) The enterprise ontology. *The Knowledge Engineering Review* 13(1):31–89
- van der Aalst WMP, ter Hofstede AHM, Kiepuszewski B, Barros AP (2003) Workflow patterns. *Distributed and Parallel Databases* 14(1):5–51
- Wand Y, Weber R (1995) On the deep structure of information systems. *Information Systems Journal* 5(3):203–223
- Zelewski S, Alan Y, Alparslan A, Dittmann L, Weichelt T (Hrsg) (2005) Ontologiebasierte Kompetenzmanagementsysteme. Logos, Berlin