Semantic Technologies for Business and Information Systems Engineering:

Concepts and Applications

Stefan Smolnik
EBS Universität für Wirtschaft und Recht i.Gr, Germany

Frank Teuteberg
University of Osnabrueck, Germany

Oliver Thomas
University of Osnabrueck, Germany

BUSINESS SCIENCE Reference

Managing Director:	Lindsay Johnston
Book Production Manager:	Sean Woznicki
Development Manager:	Joel Gamon
Development Editor:	Michael Killian
Acquisitions Editor:	Erika Carter
Typesetters:	Jennifer Romanchak, Adrienne Freeland
Print Coordinator:	Jamie Snavely
Cover Design:	Nick Newcomer

Published in the United States of America by
Business Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

Semantic technologies for business and information systems engineering: concepts and applications / Stefan Smolnik, Frank Teuteberg and Oliver Thomas, editors.
p. cm.
Includes bibliographical references and index.
Summary: “This book investigates the application of semantic technologies to business and information systems engineering”-- Provided by publisher.
ISBN 978-1-60960-126-3 (hardcover) -- ISBN 978-1-60960-128-7 (ebook) 1. Semantic computing. 2. Semantic networks (Information theory) 3. Management information systems. I. Smolnik, Stefan, 1970- II. Teuteberg, Frank, 1970- III. Thomas, Oliver, 1971- IV. Title.
QA76.5913.S47 2012
005.7--dc23
2011031983

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this book is new, previously-unpublished material. The views expressed in this book are those of the authors, but not necessarily of the publisher.

Chapter 10
Stepwise Semantic Enrichment in Health-Related Public Management by Using Semantic Information Models

Hans-Georg Fill
University of Vienna, Austria & Stanford University, USA

Ilona Reischl
AGES PharmMed, Austria

ABSTRACT

The use of semantic technologies in practical scenarios requires carefully balancing the tradeoff between costs and benefits in order to gain acceptance. In this chapter we report on a project conducted together with the Austrian competent authority in regard to safety in healthcare. It is described how semantic technologies can be combined with conceptual models to support management executives in the distribution of knowledge and the analysis of compliance. The approach is based on a step-wise semantic enrichment of conceptual models with formal semantic schemata in order to support human analyses. It has been implemented on the ADONIS meta modeling platform and applied to a scenario dealing with the management of applications for clinical trials.

INTRODUCTION

Over the past years significant effort and money has been invested into research and industrial projects dealing with semantic technologies by funding organizations such as Darpa, the European Commission, and Asian funding organizations (Bussler, 2008), (Bertolo, 2005). As a result, a large number of technologies, standards, and research prototypes are today available which permit to easily implement semantic functionalities and apply them to concrete use cases. Thereby, the unique value propositions of semantic technologies in

DOI: 10.4018/978-1-60960-126-3.ch010

order to meet critical real world challenges can be directly assessed cf. (Cardoso, et al., 2007). In the area of information systems, the process to introduce semantic technologies to practical scenarios is governed by the principles of design science. By building and evaluating artifacts in the form of constructs, models, methods, and instantiations (March and Smith, 1995) it can be shown how technologies can be implemented in a working system, thus allowing for the assessment of the suitability of the artifacts for their intended purpose (Hevner et al., 2004). As a consequence, the relevance of the technology for the constituent community, i.e. the practitioners who deal with information systems and their technologies can be estimated (Hevner et al., 2004). With the following elaborations we report on a research project that has been undertaken by the University of Vienna in cooperation with AGES PharmMed as the Austrian Competent Authority in regard to safety in health care. The main objective of the project was to develop IT-based solutions for supporting executives in the area of health-related public management, in particular for the approval of clinical trial applications. The methodology that has been elaborated for this purpose is based on the concept of *semantic information models*. It allows for the combination of visual conceptual models and semantic technologies on the basis of a meta modeling approach and follows a design-oriented research objective. The particular benefit of the approach lies in the facilitation of human analyses of complex visual conceptual models and their alignment to internationally used semantic schemata by using a step-wise approach. The purposes are to support knowledge distribution and compliance management, as well as to establish a basis for performance and resource management in health-related public management. The chapter is structured as follows: The next subchapter will give a brief introduction to the foundations used for our approach and describe the linkages to existing work. Thereafter we will present the details of the approach and how it influenced the distribution of knowledge in the organization and the assurance of the performance and compliance of the processes. The chapter is concluded with an outlook on future research directions.

BACKGROUND

To provide the foundations for our approach we will briefly give some background information to outline the challenges in the domain of health-related public management, the use of conceptual modeling and meta modeling concepts to represent conceptual visual models and on the currently available semantic technologies in the area of health.

Health-Related Public Management

The development of new medical treatments and products at a high quality and an affordable price is today one of the central challenges in the area of health-related public management. In particular, the research for new drugs involves a large number of resources and imparts considerable risks that have to be taken on by pharmaceutical companies (Sauer and Sauer, 2007). Of every 5,000 molecules tested, about 250 substances enter preclinical testing, 10 enter clinical development and just one will be finally approved by the regulatory authorities and receive the marketing authorization (EFPIA, 2010). For the parties involved in this process it is therefore essential to ensure the efficient use of resources as well as the compliance to legal regulations. On the side of pharmaceutical companies, the findings from basic laboratory research have to be translated into new methods for diagnosis, therapy, and prevention. Despite the large number of tests that can already be conducted during this pre-clinical stage, the crucial data are accumulated during the clinical trial stage, where the substances are applied to humans. Depending on the phase of the clinical trial this involves either healthy volunteers or voluntary patients.

The compliance with established standards of manufacturing and laboratory practice and the consideration of relevant scientific guidelines is required to ensure the safety of substances before their first human application.

In all these stages *public authorities* play a major role. They are responsible for ensuring the completeness of the dossier, the verification of the use of state-of-the-art scientific methods and procedures, and the compliance with the legal framework. The application for clinical trials has to be approved by the authorities before they can be initiated. This involves the assessment of the dossier with respect to manufacturing, preclinical and clinical aspects. Parallel to the evaluation by the competent authorities, ethics committees have to assess the ethical impact of the trial, as well as pre-clinical and clinical aspects and the standard of care. The basis for all these activities is a number of national and international legal regulations and guidelines. This framework is continuously adapted to the advancement of science and technology and with a view of international harmonization.

In Europe, the approval of new medicinal substances or products has to be based on a licensing application that can either be submitted on a national level, in the course of a mutual recognition procedure (MRP) in case a substance has already been approved in another member state, by a decentralized procedure (DCP) to gain approval in several countries in parallel or, for a defined group of medicinal products, via a centralized procedure by the European Medicines Agency (EMA). Both during the clinical trial stage and at the time of an application for marketing authorization, the interaction between the applicant and the authorities, and amongst the authorities and the European Medicines Agency, relies on several IT systems and databases. In order to allow for an efficient and effective processing of applications for the clinical testing of new substances and their introduction to the market, the procedures executed by the regulatory authorities need to be constantly updated and adapted to new legal and scientific requirements. The effective communication of information about these updates and adaptations both to the staff of the authorities as well as to all external stakeholders participating in these activities is therefore a crucial task. In addition, the assurance of the legal compliance and the efficient use of resources requires a continuous analysis and optimization by the responsible executives.

Conceptual Modeling and Meta Modeling

One way to deal with the complexity involved in the above outlined tasks is to use models that represent necessary parts of reality for the purposes of a human user cf. (Stachowiak, 1973). Modeling methods in general can be described according to a framework set up by (Karagiannis and Kühn, 2002). In this framework modeling methods are composed of a modeling technique and mechanisms and algorithms. The modeling technique is again composed of a modeling language and a modeling procedure. The modeling procedure defines the steps and results in the course of the application of the modeling language. The modeling language is composed of a syntax, semantics, and notation. In contrast to other approaches, e.g. as described in (Harel and Rumpe, 2000), the notation is separated from the syntax and stands for the visual representation of the elements of the modeling language cf. (Fill, 2009). This allows to modify the visual representation independently of the syntax. The semantics of the syntax elements is defined by applying a semantic mapping to a semantic schema. This semantic schema may either be formally defined or expressed in natural language. Based on the modeling language, mechanisms and algorithms can be implemented and used for the modeling procedure. The mechanisms and algorithms can either be generic, in the sense that they are independent of a specific modeling language. Or they can be specific, i.e. they are applicable only to one or a selected number of modeling languages or hybrid, i.e. they contain parameters that allow specific parts to be adapted to other modeling languages.

An example for a modeling paradigm that can serve as the basis for realizing modeling methods and that is capable of dealing with complex static and dynamic phenomena are conceptual models cf. (Wand and Weber, 2002). Modeling languages in this paradigm are characterized by a formal syntax that allows capturing the essential aspects of a given phenomenon and depicting them in the form of a visual notation (Fill, 2009). In contrast to knowledge representation approaches, conceptual modeling is directed towards the use by humans for purposes of understanding and communication and not towards the use by machines (Mylopoulos, 1998). Therefore, the semantics of conceptual modeling languages are usually not formally specified but given in natural language.

However, as has been shown by many applications of conceptual modeling techniques in science and practice, the formal definition of conceptual models may also be extended in order to allow for a technical implementation (Fill, 2004) or to support mathematical analyses such as verifications (Mendling and van der Aalst, 2007) and the application of algorithms, e.g. for the purpose of simulation cf. (Kühn and Junginger, 1999).

The implementation of conceptual model editors and processing tools requires a solid formal basis. This basis can either be realized by reverting to existing platforms and frameworks for implementing modeling languages such as EMF (McNeill, 2008) or can be performed from scratch in any kind of higher programming language. An established approach in this context is the use of meta modeling techniques (Karagiannis and Höfferer, 2006). Meta modeling techniques provide concepts to define the formal syntax of a modeling language together with their visual representation. Thereby, they enable the creation of mechanisms and algorithms working on the abstract and concrete syntax of the modeling language. They are available in the form of international standards such as the meta-object facility (OMG, 2006) or in proprietary formats (Junginger et al., 2000), (Gunderloy, 2004). In the terminology of meta modeling, a meta model defines the model of a modeling language cf. (Karagiannis and Höfferer, 2006). The meta model itself is again described in a particular meta modeling language that defines the syntax, semantics, and notation for the composition of meta models. The semantics of meta models is usually defined by mapping the elements of the meta model to natural language descriptions. Thereby, the semantics of the elements of the abstract syntax of the modeling language are described, which shall be denoted as *type semantics* cf. (Höfferer, 2007). This type semantics remains constant for all instances of the meta model and can therefore be directly used for implementing mechanisms and algorithms. However, a second kind of semantics is assigned to the instances of a particular meta model: To describe the content of a particular instance element and thus make it distinguishable from other instance elements, attributes such as a name are assigned to the instance. The content of these attributes is though not pre-determined but is assigned by a user during modeling. Therefore, the semantics described by this content – which we will denote according to (Höfferer, 2007) as *inherent semantics* – is not made explicit but described in natural language and therefore not directly accessible to machine processing.

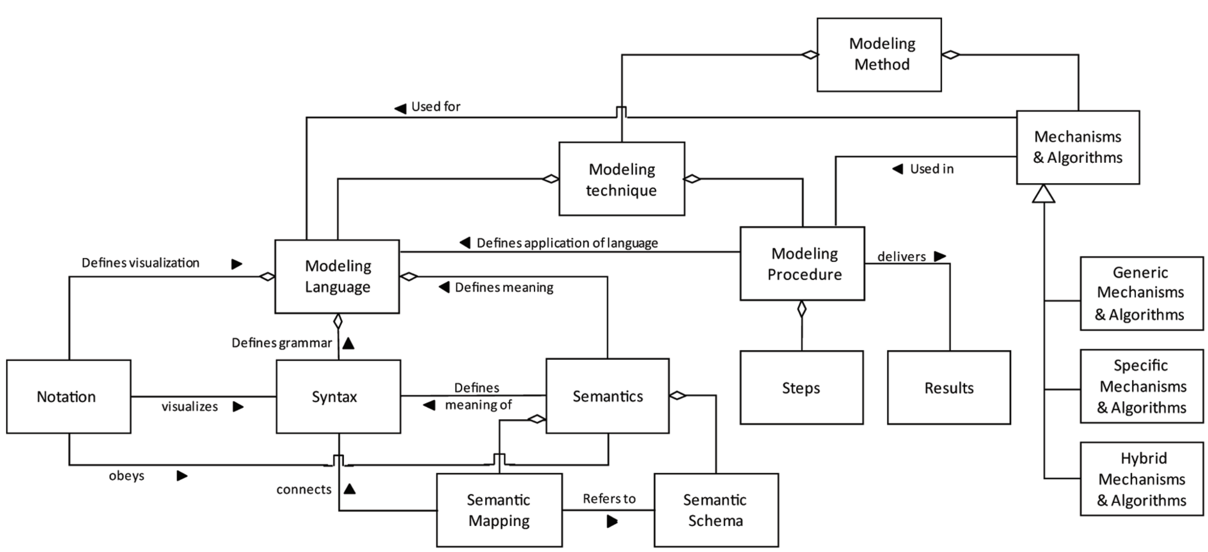
To process the inherent semantics of conceptual models, several proposals have been made in the area of business process management that can be subsumed under semantic business process modeling cf. (Thomas and Fellmann, 2009), (Höfferer, 2007), (Ehrig et al., 2007), (Lautenbacher et al., 2008). A common aspect of these approaches is the introduction of annotations for the instances of a business process modeling language in order to apply formal semantic techniques. An extension to these approaches for other model types in the context of business process management has been discussed recently in (Fill and Reischl, 2009). In this work, the approach of *semantic information models* has been described that stands for the combination of arbitrary types of conceptual models and formal semantic schemata. As will be shown in the main section of this chapter, this approach can be extended to achieve an alignment to shared semantic schemata and thus realize a step-wise semantic enrichment of conceptual models (Figure 1).

Semantics in Health-Related Public Management

As a foundation for expressing semantics in the area of health-related public management three areas have been identified that provide particular knowledge and technologies. These are: general biomedical ontologies, data interchange standards as issued by the Clinical Data Interchange Standards Consortium (CDISC), cf. (CDISC, 2010), and international harmonization initiatives by the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), cf. (ICH, 2010).

Biomedical ontologies are today used to search and query heterogeneous biomedical data, exchange data among applications, integrate information, process natural language, represent encyclopedic knowledge or apply computer reasoning to data (Rubin et al., 2008). They come in various types of formality ranging from controlled vocabularies, glossaries and thesauri to fully-fledged ontologies – for a good overview see (Rubin et al., 2008). Most of them are available for public use. The NCI BioPortal currently provides access to sixteen different terminologies including BiomedGT, the Gene Ontology, HL7, NCI Thesaurus, and SNOMED-CT (NCI, 2010). An ontology specialized on clinical research data has been recently presented by (Tu et al., 2009). A well known example for a controlled vocabulary in this area is the gene ontology (GO) that provides information about gene and protein roles in cells (Ashburner et al., 2000). It allows researchers to search for gene products that are either involved in particular biological processes, have certain molecular functions or that are located in a specific cellular component (Rubin et al., 2007). Furthermore, it can be used to gain insight into experimental results by annotating the results with

Figure 1. Components of modeling methods (Karagiannis and Kühn, 2002)



concepts from the GO and then executing analyses based on these annotations. Another example for a biomedical ontology is the NCI thesaurus developed by the National Cancer Institute. It is a biomedical terminology that contains about 80,000 concepts from several controlled vocabularies (De Coronado and Fragoso, 2009). Several formats are available for the NCI thesaurus including the common OWL DL ontology format.

CDISC is a global, open, multidisciplinary, and non-profit consortium that established standards to support the acquisition, exchange, submission and archive of clinical research data and metadata (CDISC, 2010). Its aim is to develop platform-independent data standards to enable the interoperability of information systems in medical research and related areas of healthcare. Currently, CDISC maintains primarily data standards in formats such as XML. However, also semantic definitions such as the CDISC glossary are available – but so far not in a technically processable format.

In the area of health-related public management an important organization in regard to the standardization of terminologies is the ICH. It is a joint initiative of regulators and research-based industry representatives of the European Union, the USA, and Japan (ICH, 2010). Currently the members of the ICH are the European Union and the European Federation of Pharmaceutical Industries and Associations (EFPIA), the Food and Drug Administration (FDA) and the Pharmaceutical Research and Manufacturers of America (PhRMA), as well as the Ministry of Health, Labour and Welfare (MHLW) in Japan and the Japan Pharmaceutical Manufacturers Association (JPMA). Its objective is to "increase international harmonization of technical requirements to ensure that safe, effective, and high quality medicines are developed and registered in the most efficient and cost-effective manner" (ICH, 2010). An important part of ICH is concerned with the development of common guidelines for the description of procedures of regulatory authorities. These guidelines are divided into the four major categories *quality, safety, efficacy, and multidisciplinary*. Although these guidelines follow an informal approach described in natural language they contain classifications of terms and detailed glossaries about the domain specific understanding of the terms and the procedures related to them. This makes them a valuable source of reference for the regulatory authorities that is continuously updated and approved by all member organizations.

AN APPROACH FOR A STEPWISE SEMANTIC ENRICHMENT OF CONCEPTUAL MODELS

As has been outlined above, the management tasks in the area of health-related public management are characterized by complex relationships between professional knowledge in the area of medical research, continuously changing compliance requirements and the need for an efficient use of resources. Additionally, several actors and IT systems are involved that need to communicate with each other and exchange information. To master these challenges, the use of conceptual modeling techniques has been found to be an appropriate solution (Fill and Reischl, 2009). In particular, conceptual models allow for the representation of the domain specific knowledge in regard to processes, organizational structures, and the involved IT systems. Several modeling languages are today available that support these tasks, e.g. event driven process chains (Keller et al. 1992), Adonis (Herbst and Karagiannis, 2000) or BPMN (Wohed et al. 2006). These modeling languages can be further extended and combined with other approaches, e.g. for integrating the modeling of IT architectures and services (Moser and Bayer, 2005) or for representing strategic relationships (Lichka, 2006).

With the increasing use of such modeling techniques, ranging up to several thousand models in some enterprises cf. (Rosemann, 2006), the use of (semi-)automatic analysis methods gains importance. This concerns not only the evaluation of traditional *process performance indicators* such as time, cost, and quality attributes but also the inspection in regard to *compliance requirements* or the *distribution of knowledge* to the different stakeholders. To support these analyses from a technical point of view, the conceptual models thus need to be extended. One way to enable such analyses of conceptual models is to use annotations with concepts from formal semantic schemata such as ontologies. In this way, semantic functionalities such as the provision of background knowledge for query expansion or rewriting, the facilitation of natural language processing, the retrieval and integration of information from different, distributed sources or the provision of navigation structures can be realized and directly employed for advanced analysis tasks cf. (Van Elst and Abecker, 2002).

The use of such semantic annotations for concrete practical scenarios has so far been described for some cases in the area of business process management e.g. (De Francisco, 2009). When bringing these approaches to practice, the tradeoff between costs and benefits of a particular approach has to be well argued in order to gain acceptance. Especially, when aiming for a high level of formality of semantic schemata, large efforts may be required including specially trained personnel and high costs for the time consuming, ambiguous and error-prone process of formalization (Van Elst and Abecker, 2002). Therefore the degree of formality has to be carefully chosen. In contrast, conceptual modeling approaches are characterized by the use of intuitive, visual elements that need not obey formal constraints. When the processing of the conceptual models becomes necessary, they can be gradually enriched with the formal concepts required for the processing.

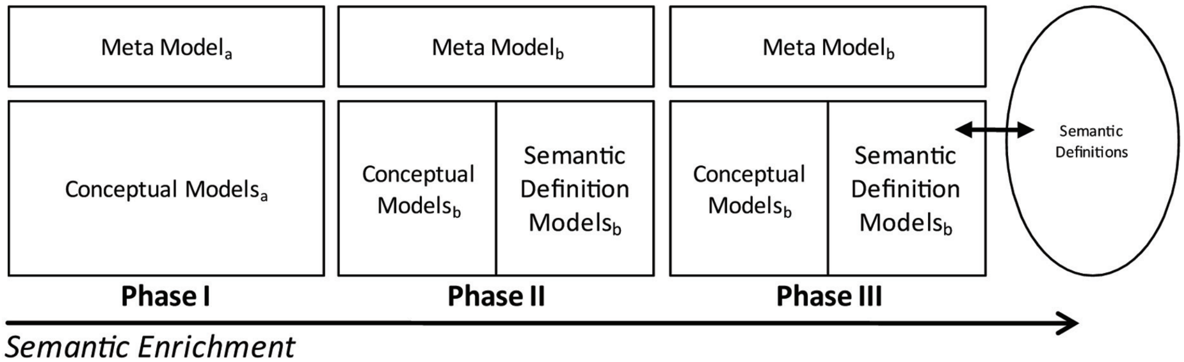
Based on these considerations, we developed an approach that is based on conceptual models and allows for a step-wise semantic enrichment of these models. This shall be described in the following sections.

Meta Modeling as a Basis

As a basis for our approach we used the concepts of meta modeling to define the formal syntax of the involved modeling languages. The intuitive concepts of meta models proved particularly useful to communicate the concepts also to non-technical users and thus receive early feedback on the approach. By using meta modeling as the foundation for all involved modeling languages, particular advantages for the implementation of algorithms can be gained as the technical environment for the modeling languages remains the same and thus all information is available in a coherent format and structure. Therefore, application programming interfaces and tools supporting meta modeling can be easily re-used. To realize the step-wise semantic enrichment we defined three phases (see Figure 2):

The first phase only contains traditional conceptual models that are derived from a common meta model. The expression of semantics in this phase is conducted by creating conceptual visual models based on the modeling language as defined by the meta model. As the meta model itself is described by natural language, e.g. in the way of a manual explaining the concepts and relations, and the conceptual models do not contain formal semantics at this stage, formal analyses of the models can only be performed on the syntax and the type semantics of the modeling elements. The inherent semantics of the model elements cannot be processed at this stage. An example could be a conceptual model of a business process containing elements for activities, decisions, and relations that express a flow between these two elements. Thus, the type semantics of the elements is given, i.e. a creator of an algorithm could use the knowledge that the model can be used to express a flow of activities and decisions and implement her algorithm accordingly. However, the inherent semantics of the elements, e.g. the name of an activity, is given in natural language and thus not formally defined in semantic terms.

Figure 2. Phases of the semantic enrichment of conceptual models



In the second phase, semantic definition models are added. They are also defined based on a common meta model and can be used to express the inherent semantics of the conceptual models. The semantic definition models may contain any type of elements that serve the purpose of describing the inherent semantics of the conceptual models. Examples could be formal semantic schemata but also models of documents or other types of conceptual models taking the role of a semantic definition. By linking elements of the conceptual models to elements of the semantic definition models, semantic annotations can be created. Thereby, it becomes possible to assign, for example, to the name of an activity "Write approval letter" elements from the semantic definition model that detail the semantics of "Write" and "approval letter". Carrying on with the example, the semantic definition model may thus contain a controlled vocabulary with an entry for "approval letter" that defines in natural language what is understood by an approval letter. The purpose of semantic definition models in this phase is therefore to provide a common semantic reference to the conceptual models. The concrete composition and the modeling language of the semantic definition models are left open and may range from strictly formal to primarily syntactic forms.

For the third phase, the semantic definition models are synchronized with externally available semantic definitions. In contrast to the second phase, the semantic definitions in the third phase are not static, in the sense that they can only be modified within the framework of conceptual models and semantic definitions models. Rather, they can now be updated and aligned to external semantic definitions. An example could be a shared ontology or a wiki that is maintained as a web resource. From this resource, either parts or the complete resource can be integrated as semantic definitions. In case the resource is updated, corresponding mechanisms have to ensure the consistency with the conceptual models that have been annotated with semantic definitions from the resource. The advantage of integrating external semantic definitions is on the one hand the re-use of existing definitions and on the other the facilitation of semantic interoperability. A concrete instantiation of this three-phase framework will be described in the following section by reverting to the approach of *semantic information models*.

Instantiation Using Semantic Information Models

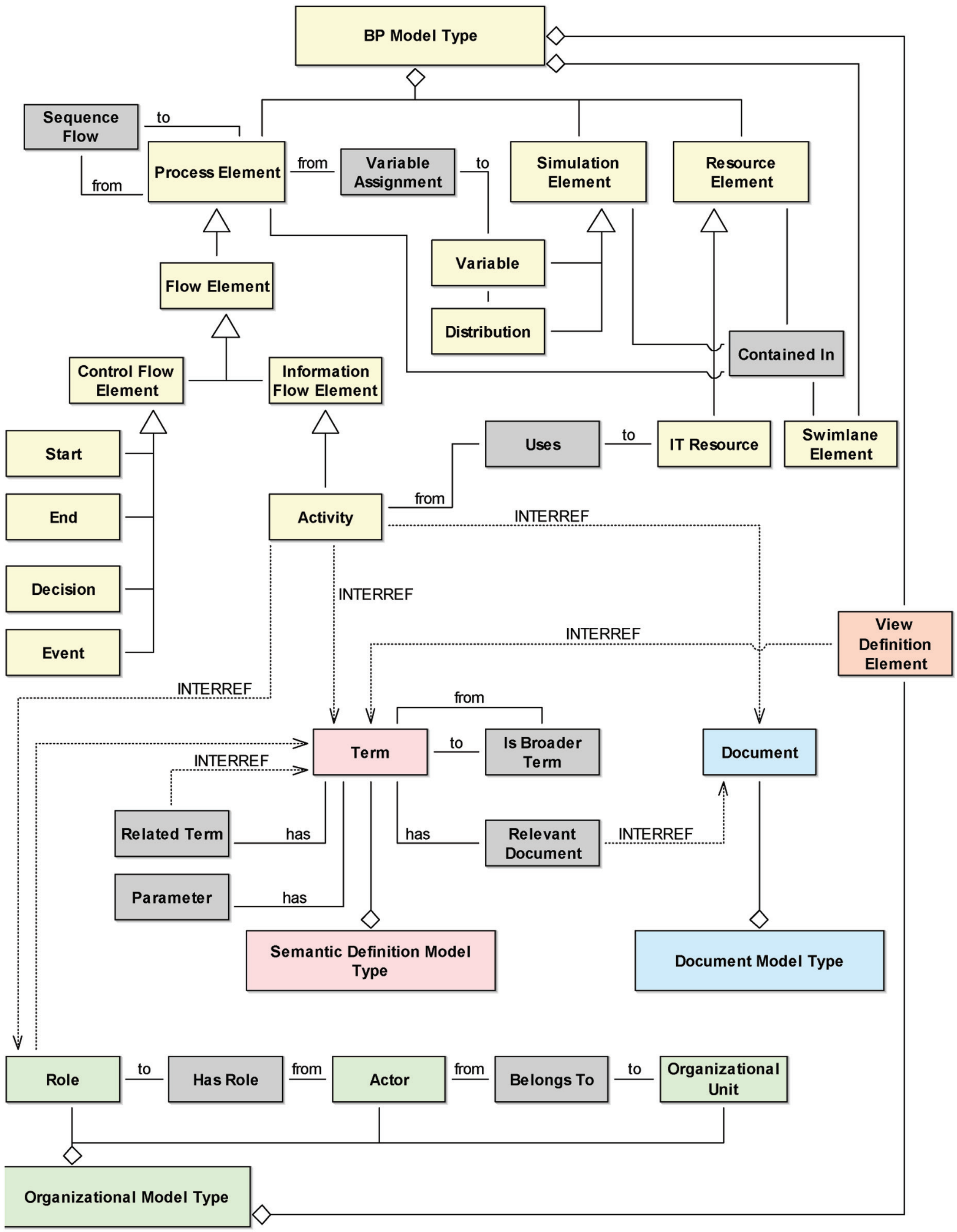
As has been described above, the conceptualization of semantic definition models can be performed in various ways, ranging from formal semantic schemata to primarily syntactic model definitions. The choice for the composition of the definitions depends on factors such as the goal of the annotation and the expected advantages of a formal definition, the familiarity of the users with certain semantic definition languages, and the technical support in regard to the available tools for processing semantic definitions.

For the application in the area of health-related public management, we used the concept of semantic information models to create semantic definition models based on formal semantic schemata. As has been shown in previous work, semantic schemata that are suitable for this approach are for example controlled vocabularies and ontologies based on the web ontology language (OWL) specification (OWL, 2009), cf. (Fill and Reischl, 2009), (Fill and Burzynski, 2009). In the approach presented here, we developed an extension to the controlled vocabulary definitions in (Fill and Reischl, 2009). The conceptual model types are based on a common business process modeling language with extensions to allow for simulations of the business processes and to represent document models and organizational models. All model types have been specified in an integrated meta model as shown in Figure 3: The business process model type contains elements to define the control flow and the information flow of business processes. *Variable* elements can be used to further detail decisions taken in the process flow and thereby provide formal definitions of the *decisions* for executing the simulation. The organizational model type is based on role concepts. *Roles* can be assigned to actors for expressing their skills and qualifications. By further assigning *roles* to *activities* in the process flow, business process models and organizational models can be linked. In the same way, elements of the document model type can be linked to *activities*. To integrate semantic definitions, the semantic definition model type provides *term* elements. Term elements can be linked to each other by using *is broader term* relations. Additionally, relationships between terms can also be defined using the *related term* property. With this property three qualified relations can be expressed: *synonyms*, *comments*, and arbitrarily *related* terms. Thereby, basic semantic relations between terms of a controlled vocabulary can be defined which proved sufficient for providing solutions in the context of health-related public management so far. The *term* elements can be linked to elements in the business process model type, the document model type, and the organizational model type to express semantic annotations. For the creation of user-centric visualizations, a *view definition* element is available in the business process and the organizational model type. Through references to terms it can thus be defined, which elements have been annotated with a specific term. This is then used to derive a visualization that can visually highlight these elements.

In detail, these model definitions and the linkages to the semantic definition model allow for the following analyses to be conducted: By annotating *role* elements, *activity* elements and *document* elements with the *term* elements, analyses of the *compliance* to specific regulations can be supported. For example, this allows to easily identify activities that are related to a specific term and need to be updated due to a change of the regulations of this term as indicated through a link to the corresponding legal document. In the same way, not only a search for specific roles and thus actors who are responsible for certain activities can performed, but this search can also be done using semantic relationships, e.g. by using the *related term* or *synonym term* properties. This applies also to the search for documents that are relevant for the execution of activities in the process.

As all elements of the models are formally defined, the implementation of the meta model on a concrete technical platform can be directly performed. For our approach we used the ADONIS[1]

Figure 3. Excerpt of the meta model used for the scenario of health-related public management



meta modeling platform to implement the meta models and define the visual representations of the model elements. This also permits to easily distribute the knowledge contained in the models in other formats, e.g. for the distribution in the form of HTML pages that can be stored on a website and made accessible to all internal and external stakeholders. With the formal definitions based on a meta model mechanisms and algorithms working on these models can be immediately realized. In particular, algorithms for the simulation of the contained business processes can be applied to generate information about cycle times and the usage and requirements of resources for executing the processes. Furthermore, the annotation of the model elements can be used as input to advanced visualization functionalities (Fill, 2009). These can highlight for example particular parts of a complex model based on a set of given terms and semantically related terms to directly support the user in analyzing the structure of the model based on the view definitions.

Concerning the above described phases of semantic enrichment, the presented meta model is able to support all three phases. For the traditional use of conceptual modeling the semantic definition models can be left out and the use restricted to the business process model, organizational model and document model type. For the second phase, semantic definition models can be added based on user-defined semantic terms and their relationships. This ensures already a consistent terminology by linkages to the semantic terms throughout the various conceptual models. It also permits to conduct several semantic analyses. By integrating semantic definitions from external sources, the semantic capabilities of the approach can be even further extended. In this phase, even more powerful semantic definitions can be used for the analyses, including large ontologies and thesauri that can be mapped to the semantic definition models.

Application at AGES PharmMed

The above described approach has been applied in a practical scenario at AGES PharmMed. AGES PharmMed and the Federal Office for Safety in Health Care (BASG), the Austrian Competent Authority, went operative on January 2, 2006 following a reorganization and out-sourcing from the Federal Ministry of Health. Legal responsibilities of the BASG center on issues pertinent to drug development and licensing. The purpose of AGES PharmMed, which is fully owned by the Republic of Austria, is to support BASG by providing services, personnel, and location. The tasks of AGES PharmMed can be described by nine categories: (1) the approval of clinical trials with medicinal products and/or medical devices; (2) the approval of pharmaceuticals; (3) pharmacovigilance, i.e. the systematic logging of adverse reactions of medicinal products; and the vigilance of medical devices (4) the monitoring of the market of medicinal products; (5) the inspections of pharmaceutical companies; (6) haemovigilance, i.e. the monitoring of blood donations and transfusions; (7) the provision of scientific advice for pharmaceutical companies; (8) the official medicines control laboratory; and (9) the official international representation of Austria in European organizations such as the European Medicines Agency (EMA).

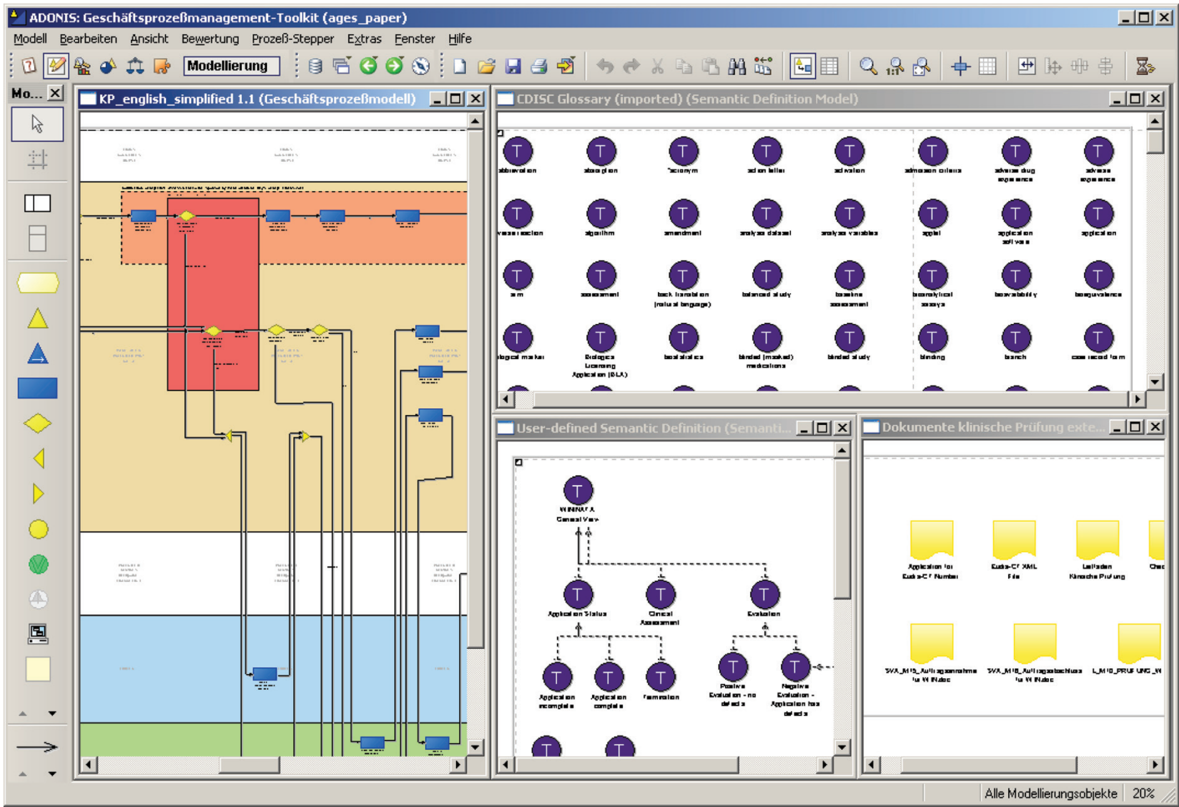
The focus of the scenario was the management of clinical trial applications with medicinal products. These applications are subject to European Directives that have been introduced to harmonize the different work practices in regard to clinical trials. European Directives have been translated into National Law and are applicable to all member countries (Heerspink et al., 2008). In particular, the clinical trials directive Dir/2001/20/EC clearly defines the role of central and local ethics committees and how they should interact in the procedures for the application of clinical trials cf. (European Parliament and the Council, 2001). For example, it is thus stated that an application for a clinical trial has to be submitted to a central ethics committee and the member state's competent authority in parallel and that both parties have to give an opinion on the trial application within 60 days from the receipt of the application. In addition, competent authorities have to provide mechanisms for the access of regulatory information by investigators and the general public and to evaluate and improve standards in clinical trials according to the recent scientific advancements. This constant adaptation to national and international procedures according to the legal framework requires a high degree of flexibility on the side of the competent authorities.

To apply the approach of semantic information models at AGES PharmMed the following steps were taken: In preliminary workshops the processes relevant for the administration of clinical trials were identified. These processes were then modeled in detail in ADONIS using mainly the business process model type together with the organizational model type and the document model type. As a next step, a controlled vocabulary was added using the semantic definition models. This controlled vocabulary contained terms that are central to the management of clinical trial applications and allowed to create first user-centric visualizations of the process models in order to highlight specific process parts. The last step so far comprised the integration of the CDISC glossary in the form of a semantic definition model. The CDISC glossary is freely available in a textual format and contains 400 terms from the area of clinical research. In contrast to most of the available biomedical ontologies it contained a large number of terms relevant for the management of clinical trial applications and did not require knowledge about formal semantic definitions. The textual format was parsed using a Java application and transformed into a model instance of the semantic definition model type - see Figure 4 for a screenshot showing the ADONIS meta modeling platform with a business process model on the left, the semantic definition model of the CDISC glossary on the right, another semantic definition model from the second phase at the bottom and a document model.

The terms contained in the CDISC glossary are explained in natural language. In addition, the glossary provides references to synonym terms, related terms, and relevant ICH and FDA guidelines. This information was also transferred to the semantic definition model and the links between the terms were established automatically in the model in case they were defined in the CDISC file. Finally, the elements of the conceptual models were partly annotated by hand using the terms from the CDISC glossary. Thereby it turned out that some of the terms and concepts were only applicable to a specific type of organization or even only to specific countries with particular regulations. Although some of the terms defined in the CDISC glossary are based on the ICH guidelines – which ensure a common understanding between all member countries – several terms are specific to the US way of handling certain procedures. For example the entry for "action letter" defines it as "…an official communication from FDA to an NDA sponsor…" (CDISC, 2010). Therefore, a direct annotation with these concepts was not possible as it would mislead a potential investigator of this term. However, it was found that the usage of terms from the CDISC glossary together with the conceptual models proved helpful in order to understand the origin of the meaning of a specific term and its embedding in a particular process.

Based on the annotations of the models with terms from the CDISC glossary, the user-centric visualizations of the process models could thus also be performed using an internationally established terminology standard - see Figure 5 for an example. In this screenshot, three activities in the process are shown that deal with the applications for "Substantial amendments" and "Non-substantial amendments" and the report-

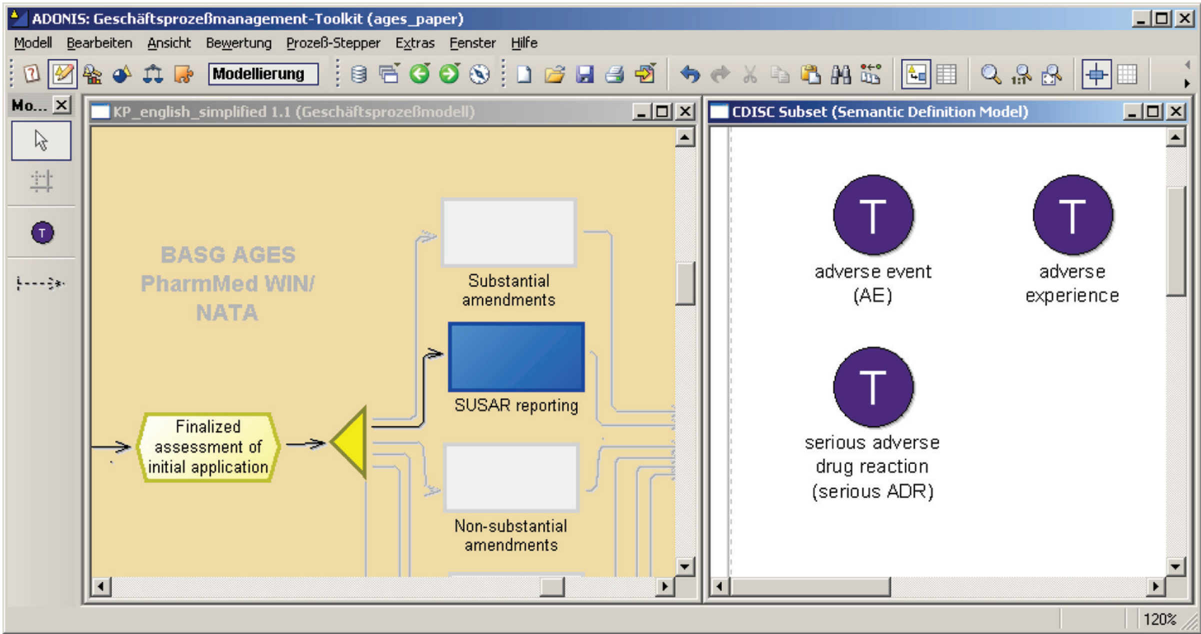
Figure 4. Screenshot of the implementation on the ADONIS meta modeling platform



ing of SUSARs (Suspected Unexpected Serious Adverse Reaction). The activity "SUSAR reporting" has been annotated with the CDISC term "serious adverse drug reaction (serious ADR)". In the CDISC glossary this term contains a relation to the term "adverse experience" via the *related* property. The term also has another relation to the term "adverse event (AE)", again with the *related* property. At the same time, a view definition element has been defined for the process model (not visible on the screenshot) that contains a reference to the term element "serious adverse drug reaction (serious ADR)" of the semantic definition model containing the CDISC glossary. It was therefore possible to adapt the visualization of the process model based on these relationships, so that in this view, only activities that have been annotated with the terms defined by the view definition element are highlighted and all others with lower contrast. By additionally using the semantic information of the *related* properties, it is also possible to expand the visualization semantically, in the way that any terms transitively related to "adverse event (AE)" can be found and used for highlighting correspondingly annotated activities.

As a consequence of these functionalities, it is possible to support the distribution of knowledge in the way that commonly used terms of the CDISC glossary can be used to find the related information in the process model. Furthermore, the management of compliance to legal frameworks can be supported by identifying process parts that are related to terms that have undergone a change in scientific or legal procedures. An example could be a change in the handling of activities related to adverse events that require new reporting measures and that can be found via the above outlined mechanisms.

Figure 5. Screenshot of a user-centric visualization using the CDISC glossary



Lessons Learned

For the application in future projects we would like to give a short overview of the main lessons learned through applying the approach in practice. The representation of the business processes using the described modeling language together with the semantic annotations led to several benefits. Among them was the facilitation of the distribution of knowledge about the processes within the organization and the possibility of relating different information sources using a common terminology. This proved particularly helpful for the instruction of new staff members and their understanding of the complex dependencies between the various procedures and the according terminology standards. Furthermore, the separation of the organizational roles from concrete performers using the role concepts and their semantic annotation established a good basis for analyzing resource requirements by applying the described visualization technique. The insights gained by these analyses will be directly used in a subsequent project that will deal with performance management. What directly supported the application in an industry setting, was the fact that a stepwise approach could be taken for the formalization. Although formalization can offer a large range of additional processing functionalities, the effort for encoding just the basic semantic relationships of real-life business processes and the dependencies on a large number of legal regulations for ensuring compliance in formal terms can quickly outweigh its expected benefits. Especially the semantic technologies that have been reported in approaches for semantic enrichments of models so far still require a considerable amount of manual encoding that requires itself considerable knowledge about the formal foundations and the represented domain knowledge. Although the approach used here does not exempt domain users from dealing with formal issues, the stepwise semantic enrichment greatly smoothed the way towards it. Furthermore, the use of visual models helped to generate results rather quickly by the domain experts themselves.

FUTURE RESEARCH DIRECTIONS

Future work will include the evaluation of the approach in regard to the additional benefits of using semantic information models in the context of health-related public management. This concerns in particular the investigation of the mechanisms necessary to execute (semi-) automatic analyses of the compliance of the processes to legal regulations and the use of semantic schemata with a higher degree of formality such as ontologies. For this purpose it can be reverted both to a number of ongoing work in the area of compliance management of business processes (Karagiannis et al., 2008) and the application of semantic technologies for automated checking (Awad et al., 2008), (El Kharbili and Stein, 2008). However, for the concrete application field of health-related public management the most relevant semantic standards such as the ICH guidelines are currently not available as a formal specification. A fruitful approach to integrate also these guidelines could be to apply techniques of natural language processing in order to generate appropriate formal schemata for the integration as semantic definition models. Furthermore, it will have to be investigated, how the annotation of the conceptual models can be (semi-)automatically performed or at least how humans can be better supported during this task. Possible solutions for this goal might be found in the area of machine learning and artificial neural networks.

CONCLUSION

In this chapter we described an approach for the stepwise semantic enrichment of conceptual models. Our focus area was the field of health-related public management. To support management executives in this field we developed a meta model based approach for creating conceptual models and annotating them with elements from semantic definitions. The goal was to support knowledge distribution and compliance management, as well as to establish a basis for performance and resource management. By using a three-phased approach we showed how conceptual models can be semantically enhanced and related to externally defined semantic schemata. The approach benefits both staff and management in dealing with complex conceptual models. It has been applied and positively evaluated at the Austrian competent authority AGES PharmMed in a practical setting.

ACKNOWLEDGMENT

Parts of the work on this chapter have been funded by the Austrian Science Fund in the course of an Erwin-Schrödinger Fellowship Grant Number J3028-N23.

REFERENCES

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., & Cherry, J. M. (2000). Gene ontology: Tool for the unification of biology. *Nature Genetics*, *25*, 25–29. doi:10.1038/75556

Awad, A., Decker, G., & Weske, M. (2008). Efficient compliance checking using BPMN-Q and temporal logic. In M. Dumas, M.Reichert, & M.-C. Shan (Eds.), *Proceedings of the 6th International Conference on Business Process Management*, (pp. 326-341). Berlin.

Bertolo, S. (2005). *Funding strategies for the Semantic Web: Current activities and future trends in the European Union*. Retrieved January 15, 2010, from ftp:// ftp. cordis. Europa .eu/ pub/ ist/ docs/ directorate_e/kmcc/funding-strategies-for-the-semantic-web_en.pdf

Bussler, C. (2008). Is Semantic Web technology taking the wrong turn? *IEEE Internet Computing*, (January/February): 75–79. doi:10.1109/ MIC.2008.16

Cardoso, J., Hepp, M., & Lytras, M. (2007). *The Semantic Web - Real world applications from industry*. Springer.

CDISC (Clinical Data Interchange Standards Consortium). (2010). *Glossary - Request for comments on CDISC terms*. Retrieved January 15, 2010, from http://www.cdisc.org/glossary

De Coronado, S., & Fragoso, G. (2009). *NCI Edit Tab: Protege 3.4 plugin for editing and maintaining NCI thesaurus*. Paper presented at the 11th International Protégé Conference.

De Francisco, D., & Grenon, P. (2009). *Enhancing telecommunication business process representation and integration with ontologized industry standards.* Paper presented at the 4th International Workshop on Semantic Business Process Management (SBPM2009).

EFPIA (European Federation of Pharmaceutical Industries and Associations). (2010). *Industry in figures - A highly regulated industry*. Retrieved January 15, 2010, from http://www.efpia.eu/content/default.asp?PageID=361

Ehrig, M., Koschmider, A., & Oberweis, A. (2007). Measuring similarity between semantic business process models. In J. F. Roddick & A. Hinze (Eds.), *Proceedings of the Fourth Asia-Pacific Conference on Conceptual Modelling (APCCM 2007)* (vol. 67, pp. 71-80). ACM.

El Kharbili, M., & Stein, S. (2008). Policy-based semantic compliance checking for business process management. In P. Loos, et al. (Eds.), *Modellierung betrieblicher Informationssysteme (MobIS 2008), CEUR Workshop Proceedings 420*, Aachen, (pp. 165-177).

European Parliament and the Council. (2001). *Directive 2001/20/EC of the European Parliament and of the Council of 4 April 2001 on the approximation of the laws, regulations and administrative provisions of the member states relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use*. Retrieved January 15, 2010, from http://eur-lex. europa.eu/ LexUriServ/ LexUriServ. do?uri= CELEX: 32001L0020: EN: HTML

Fill, H.-G. (2004). UML statechart diagrams on the ADONIS metamodeling platform. *Electronic Notes in Theoretical Computer Science*, *127*(1), 27–36. doi:10.1016/j.entcs.2004.12.028

Fill, H.-G. (2009). *Visualization for semantic Information Systems*. Gabler. doi:10.1007/978-3-8349-9514-8

Fill, H.-G., & Burzynski, P. (2009). *Integrating ontology models and conceptual models using a meta modeling approach*. Paper presented at the 11th International Protégé Conference.

Fill, H.-G., & Reischl, I. (2009). *An approach for managing clinical trial applications using semantic information models*. Paper presented at the 3rd International Workshop on Process-oriented information systems in healthcare (ProHealth '09).

Gunderloy, M. (2004). Review: MetaEdit+. *Application Development Trends Magazine, 3*(12).

Harel, D., & Rumpe, B. (2000). *Modeling languages: Syntax, semantics and all that stuff - Part I: The basic stuff (No. MCS00-16)*. Rehovot, Israel: The Weizmann Institute of Science.

Heerspink, H. J., Dobre, D., Hillege, H. L., Grobbee, D. E., & De Zeeuw, D. (2008). Does the European Clinical Trials Directive really improve clinical trial approval time? *British Journal of Clinical Pharmacology*, *66*(4), 546–550. doi:10.1111/j.1365-2125.2008.03246.x

Herbst, J., & Karagiannis, D. (2000). Integrating machine learning and workflow management to support acquisition and adaptation of workflow models. *Intelligent Systems in Accounting . Financial Management*, *9*(2), 67–92.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in Information System research. *Management Information Systems Quarterly*, *28*(1), 75–105.

Höfferer, P. (2007). *Achieving business process model interoperability using metamodels and ontologies.* Paper presented at the 15th European Conference on Information Systems (ECIS2007), St. Gallen, Switzerland.

ICH. (2010). *ICH guidelines*. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Retrieved January 15, 2010, from http://www.ich.org/cache/compo/276-254-1.html

Junginger, S., Kühn, H., Strobl, R., & Karagiannis, D. (2000). Ein Geschäftsprozessmanagement-Werkzeug der nächsten Generation - ADONIS: Konzeption und Anwendungen (German: ADONIS: A next generation business process management tool - Concepts and Applications). *Wirtschaftsinformatik*, *42*(5), 392–401.

Karagiannis, D., & Höfferer, P. (2006). Metamodels in action: An overview. In J. Filipe, B. Shishkov & M. Helfert (Eds.), *ICSOFT 2006 - First International Conference on Software and Data Technologies* (pp. IS-27 - IS-36). Setúbal, Portugal: Insticc Press.

Karagiannis, D., & Kühn, H. (2002). Metamodeling platforms. In K. Bauknecht, A. Min Tjoa & G. Quirchmayer (Eds.), *Third International Conference EC-Web 2002 – Dexa 2002* (pp. 182). Aix-en-Provence, France: Springer.

Karagiannis, D., Mylopoulos, J., & Schwab, M. (2008). *Business process-based regulation compliance: The case of the Sarbanes-Oxley Act.* 15th IEEE International Requirements Engineering Conference (RE 2007). IEEE.

Keller, G., Nuettgens, M., & Scheer, A.-W. (1992). Semantic process modeling on the basis of event-driven process chains (in German). *Publications of the Institute of Business Informatics . Saarland University*, *89*, 29.

Kühn, H., & Junginger, S. (1999). An approach to use UML for business process modeling and simulation in ADONIS. In H. Szczerbicka (Ed.), *Proceedings of the 13th European Simulation Multiconference (ESM 99) - Modeling and Simulation: A Tool for the Next Millenium* (pp. 634-639). Warsaw, Poland.

Lautenbacher, F., Bauer, B., & Seitz, C. (2008). *Semantic business process modeling - Benefits and capability.* Paper presented at the AAAI Spring Symposium, Stanford University, California.

Lichka, C. (2006). *Business scorecarding as model-driven approach for strategy implementation.* PhD thesis, University of Vienna, Austria, Vienna.

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, *15*, 251–266. doi:10.1016/0167-9236(94)00041-2

McNeill, K. (2008). *Metamodeling with EMF: Generating concrete, reusable Java snippets*. Retrieved April 8, 2008, from http://www.ibm.com/developerworks/library/os-eclipse-emfmetamodel/index.html?S_TACT=105AGX44&S_CMP=EDU

Mendling, J., & Van der Aalst, W. M. P. (2007). Formalization and verification of EPCs with OR-joins based on state and context. In J. Krogstie, A. L. Opdahl & G. Sindre (Eds.), *Proc. of the the 19th International Conference on Advanced Information Systems Engineering* (Vol. LNCS 4495, pp. 439-453). Springer.

Moser, C., & Bayer, F. (2005). IT architecture management: A framework for IT services. In J. Desel & U. Frank (Eds.), *Proceedings of the Workshop on Enterprise Modelling and Information Systems Architectures*. Klagenfurt, Austria: Lecture Notes in Informatics - Gesellschaft für Informatik (GI).

Mylopoulos, J. (1998). Information modeling in the time of the revolution. *Information Systems*, *23*(3-4). doi:10.1016/S0306-4379(98)00005-2

NCI (National Cancer Institute). (2010). *BioPortal*. Retrieved January 15, 2010, from http://bioportal.nci.nih.gov/

Object Management Group OMG. (2006). *Meta object facility (MOF) specification 2.0*. Retrieved January 15, 2010, from http://www.omg.org/spec/MOF/2.0/PDF/

OWL. (W3C OWL Working Group). (2009, October 27). *OWL 2 Web ontology language document overview - W3C recommendation*. Retrieved January 15, 2010, from http://www.w3.org/TR/2009/REC-owl2-overview-20091027/

Rosemann, M. (2006). Potential pitfalls of process modeling: Part B. *Business Process Management Journal*, *12*(3), 377–384. doi:10.1108/14637150610668024

Rubin, D. L., Shah, N. H., & Noy, N. F. (2008). Biomedical ontologies: A functional perspective. *Briefings in Bioinformatics*, *9*(1), 75–90. doi:10.1093/bib/bbm059

Sauer, C., & Sauer, R. M. (2007). Is it possible to have cheaper drugs and preserve the incentive to innovate? The benefits of privatizing the drug approval process. *The Journal of Technology Transfer*, *32*, 509–524. doi:10.1007/s10961-007-9036-0

Stachowiak, H. (1973). *Allgemeine Modelltheorie (German: General Model Theory)*. Springer.

Thomas, O., & Fellmann, M. (2009). Semantic process modeling – Design and implementation of an ontology-based representation of business processes. *Business & Information Systems Engineering*, *1*(6), 438–451. doi:10.1007/s12599-009-0078-8

Tu, S., Carini, S., Rector, A., Maccallum, P., Toujilov, I., Harris, S., et al. (2009). *OCRe: An ontology of clinical research*. Paper presented at the 11th International Protégé Conference.

Van Elst, L., & Abecker, A. (2002). Ontologies for information management: Balancing formality, stability, and sharing scope. *Expert Systems with Applications*, *23*, 357–366. doi:10.1016/S0957-4174(02)00071-4

Wand, Y., & Weber, R. (2002). Research commentary: Information Systems and conceptual modeling - A research agenda. *Information Systems Research*, *13*(4), 363–376. doi:10.1287/isre.13.4.363.69

Wohed, P., Van der Aalst, W. M. P., Dumas, M., Ter Hofstede, A. H. M., & Russell, N. (2006). *On the suitability of BPMN for business process modelling.* Paper presented at the BPM 2006, Vienna, Austria.

ENDNOTE

1 ADONIS is a commercial product and trademark of BOC AG. A community edition is available at http://www.adonis-community.com/