User-friendly Semantic Annotation in Business Process Modeling

Matthias Born Florian Dörr Ingo Weber

SAP Research, Karlsruhe, Germany
{Mat.Born, F.Doerr, Ingo.Weber}@sap.com

Abstract. Current problems in Business Process Management consist of terminology mismatches and unstructured and isolated knowledge representation in process models. Semantic Business Process Management aims at overcoming many of those weaknesses of Business Processes Management through the use of explicit semantic descriptions of process artifacts. However, this vision has a prerequisite: semantic annotations need to be added to the process models. In this paper, we present an approach that allows flexibly annotating semantics in a user-friendly way, by exposing ontological knowledge to the business user in appropriate forms and by employing matchmaking and filtering techniques to display options with high relevance only. By adding semantic information the precision of process models increases, ultimately supporting Web Service discovery and composition. As a proof-of-concept, the work has been implemented prototypically in a process modeling tool.

1 Introduction

A core aspect of Business Process Management (BPM) is creating models of business processes. These models are used in various contexts: communication, documentation, implementation, and automated execution. Companies try to establish a common basis of business terminologies using process repositories. However, this approach only helps to a certain degree and does not address the overall business process lifecycle. General issues within a business process lifecycle are that business consultants and IT experts do not speak the same language, do not share the same concepts of processes, or use the same tools. Semantically annotated process models could enable support for the modeler in various associated tasks: reusing parts of process models when creating new models; making process models executable; detecting cross-process relations; facilitating change management; and providing a structured basis for knowledge transfer. Semantic Web Service technology, like Web Service Modeling Ontology (WSMO) [16] or OWL-S [14, 2] provide methods and tools for creating these machine-accessible representations of knowledge. The Semantic BPM (SBPM) approach attempts to take BPM to the next level by integrating and utilizing semantics to improve the modeling and management of business processes.[7]

Within this context, our paper presents an approach for the integration of semantics in modeling tools to support the graphical modeling of business processes with information derived from domain ontologies. For this purpose, we identify suitable semantic information to specify business process models more precisely. We then present a concept for the integration of the identified semantic information in modeling tools utilizing ontological descriptions of the business process models and the domain world to augment and annotate process models. We exploit the particular nature of business process models, e.g., their control and data flow, and suggest a specific structure of the domain ontology. This structure defines business objects in the domain of discourse, alongside with their lifecycle. If the domain ontology is not expressed in this structure, our techniques for string-based matchmaking are still applicable, but the additional precision in filtering with regard to the objects and their states is lost.

Finally, we present matchmaking functionalities for supporting users in modeling semantically annotated process models. This is achieved by matching elements of the graphical business process model with elements of domain ontologies. We demonstrate a way how textual fragments are used to match semantic annotations to elements of a process model based on the Business Process Modeling Notation (BPMN)[1][2]. In short, this is achieved by comparing the context and any given textual descriptions to the applicable instances in a domain ontology.

Our approach enables that partial process fragments can be used to support the modeling tasks given an underlying ontological description of the domain. Further, the augmented business process models facilitate the discovery of appropriate Web Services. The semantic annotations can be used when querying an enterprise-wide process model repository, by allowing for more informed search techniques and fuzzy results.

As a proof-of-concept, the SAP Research modeling tool "Maestro for BPMN" has been extended with a prototypical implementation of this approach.

The remainder of the document is structured as follows. The requirements for our solution are described in Section 2. Based on this, we present an approach for user-friendly semantic annotation in Section 3, followed by a description of the related prototypical implementation in Section 4. Subsequently, related work is discussed in Section 5. Finally, Section 6 concludes.

2 Requirements

In the scope of our approach the additional information, i.e. the semantic annotations, sometimes referred to as tags or markups, for semantically enriched process models, should comply with the following requirements. First of all, the information should be definable by business experts during modeling time. Users of modeling tools should be able to understand the semantic information they deal with. Therefore, the semantic information may not be too IT specific or low-level. The ease of use is essential regarding user acceptance. Second, the additional information should facilitate the realization of the processes and support querying the process space. It should allow users to specify process models more

[1] OMG, BPMN Information, http://www.bpmn.org/

[2] Note that our conceptual solution is independent of BPMN as a concrete graphical business process modeling notations.

precisely, couple these process models tighter to existing domain concepts, and help to find and compose according Web Services for the activities of the process models (Web Service discovery and composition). Third, in order to support the semantic augmentation of the process models, information should be derived from appropriate domain ontologies. In order to realize this kind of support the following competency questions should be taken into account:

1. *What are the objects, states & actions of a domain and what are their names?*
2. *What are possible states of a certain object?*
3. *What are possible predecessor and successor states of a certain state?*
4. *Which actions are possibly relevant for a certain object?*
5. *Which objects are manipulated (state changes) by a certain action?*
6. *Which state changes (transitions) for which objects can be caused by which actions?*

Based on these requirements the next section shows the generic approach which we developed for user-friendly annotation of process models.

3 Solution Approach

To provide the information described above, different BPMN elements are extended and utilized. Figure 1 illustrates the main components of our semantic extension for business process modeling tools. Especially, BPMN Data Objects and Associations are used to describe the activities of a process model more precisely by defining associated objects and their state transitions: Data Objects identify the objects an activity deals with and Associations link the Data Objects to the corresponding activities in the process diagram. The user modeling the process may name and define the activities, may specify pre- and post conditions for the activities in natural language, and may define the objects as well as the objects' states before and after an associated activity has been executed within the graphical model. Through our extensions for business process modeling tools, business experts are supported in specifying this additional information during the graphical modeling phase of business processes. To help users annotating their process models, different matchmaking functionalities utilize parsed-in domain ontologies, which in turn describe available elements like objects and their states. The matchmaking functionalities thus help to link the model elements to available domain concepts.

The following subsections show how we used ontologies and describe the different matchmaking functionalities in more detail.

3.1 Ontology Engineering

We identified the following additional semantic information for business process models:

– The objects relevant for each activity in the process model can be specified and, if applicable, get directly linked to according objects of a domain.

– The states of these objects, before and after the according activity on the objects has been performed, can be specified.
– Natural language definitions of pre- and post conditions for the activities within a process model can be provided.

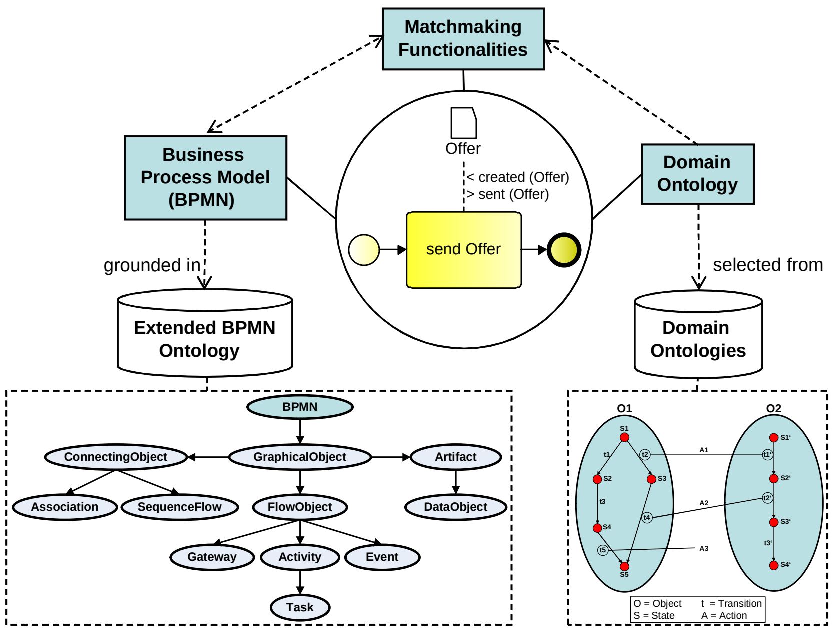

Fig. 1. Realization Overview.

All this information helps to ease the realization of the processes and supports querying the process space by answering questions regarding which objects are manipulated within which business process models.

Two kinds of ontologies are used to enable the semantic support of modeling activities. First, the evolving *sBPMN* ontology [1] of the SUPER[3] project as a format for representing BPMN process models, featuring basic concepts and attributes for standard BPMN elements, has been extended. Our ontology provides possibilities to define states of a Data Objects before and after corresponding activities have been executed, to link objects, states, and activities to elements of domain ontologies describing them, and to capture natural language pre- and post conditions for activities. With these extensions, the *sBPMN* ontology can be used as an internal and external format for semantically augmented BPMN

[3] Integrated Project SUPER (Semantics Utilised for Process Management within and between Enterprises), http://www.ip-super.org/

process models in the scope of our approach. Second, we will define a possible structure of domain ontologies along with a short concrete example. Our domain ontology covers information concerning domain objects and states which helps to model business processes more precisely. This kind of domain ontology may then be used within process modeling tools to support the user in finding and defining appropriate activities, data objects, and states.

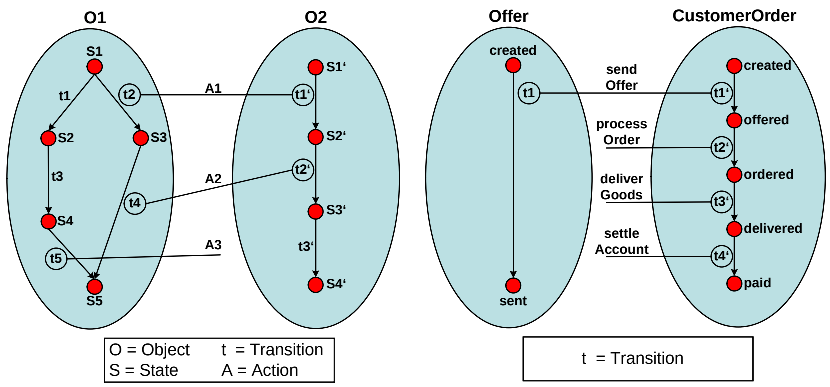

Fig. 2. Domain Ontology Visualization.

The left part of Figure 2 illustrates the information which our domain ontology provides. The ontology contains information about domain objects, states, transitions, and actions. For each domain object, possible states and state transitions are described which together form the lifecycle of a domain object. Actions represent activities in the domain and can cause multiple state transitions on different objects. States are described by more fine-granular definitions (e.g. constraints over attributes). The right part of Figure 2 shows an extract from an example of a domain ontology. Domain ontologies that provide this kind of information support our semantic business process modeling approach. During modeling, these domain ontologies are queried and utilized to help the user specifying model elements and states by proposing appropriate domain concepts or instances.

Not only states but also objects can be defined, described, and considered. The objects defined in the ontology represent the nouns used within this language, actions may be regarded as verbs and states virtually are adjectives. The structure relates objects (nouns), actions (verbs), and states (adjectives) to each other and thus defines a normalized modeling language in the scope of specific domains. This correlation and interrelation together with ontology-specific descriptions and definition potentials (e.g. hierarchies, inheritance, subsumption relationships, etc.), which may be exploited for reasoning tasks, is a major contribution and benefit of the approach developed within this paper.

3.2 Name-based matchmaking

Different matchmaking functionalities that are required to bridge the gap between the business process model and the domain world are incorporated into our semantic extension approach, in order to allow matching elements of the graphical business process model with elements of domain ontologies and to support the user in semantically specifying or refining the process. Utilizing appropriate domain ontologies, the matchmaking functionalities address the problem of deriving a list of proposals for a selected model element (Data Object, Activity, Association/State) that a user has chosen for semantic refinement. To solve this problem, we use a combination of different text and name matching methods and utilize process diagram context information as well as domain ontology knowledge. For name matching tasks, we use a combination of heuristic comparison methods on the strings of characters, well-known string distance metrics [4, 3], and matching methods considering synonyms and homonyms. The additional utilization of the diagram context information of selected model elements and domain knowledge to match model fragments with domain instances leads to even better results. Domain element proposals can be derived with the help of elements already specified in the process model and the information covered by domain ontologies.

3.3 Process context-based matching

Furthermore, the actual control and data flow in the process model can be leveraged in our approach.

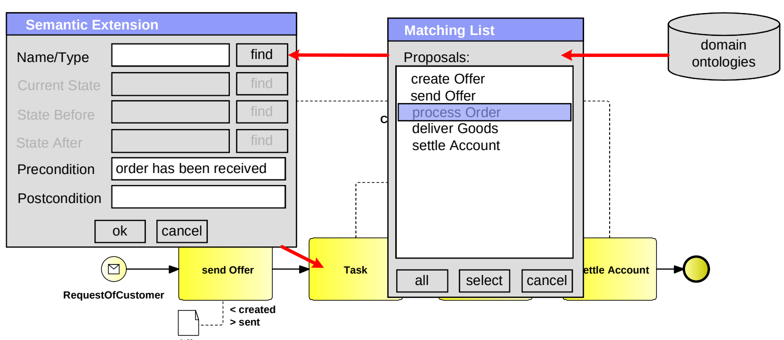

Fig. 3. Simple Ordering Example.

– With respect to the control flow, the lifecycle of a domain object can be used to suggest next activities during modeling. E.g., in Figure 3, the modeling tool can suggest "*process Order*" as the next activity in the process, because it follows "*send Offer*" in the CustomerOrder's lifecycle.

– Also, the lifecycle can be employed for making targeted suggestions for the semantic refinement of activities, e.g., by not suggesting an already used action again in a process model.

Analogously, the data flow in process models can be supported. Note that consistency checks between process control, data flow, and object lifecycles can be provided as well. Furthermore, the semantic business process models contribute to mechanize the realization of the processes. The additional information about how objects are changed by the activities of a process model can help to derive pre- and post conditions for goal descriptions. By considering the states – which may for example be defined in terms of object-attribute-constraints in domain ontologies – of all associated objects before and after a particular activity has been carried out, high-level goal definitions for the activities of a business process model can be derived. These goal definitions can then be used as an input to business process composition approaches [19] or directly to functional discovery engines, like the prototype integrated in Web Service Execution Environment (WSMX)[4]. By comparing goal descriptions, such discovery engines try to find appropriate Web Services to achieve the goals. On this way, Web Services implementing the desired functionality can automatically be discovered for the activities of a business process model. If no single Web Service satisfies the goal, a composition of Web Services can be searched, which together then achieve the goal.

4 Prototypical Implementation

The solution approach has been successfully implemented using the SAP Research modeling tool, namely "Maestro for BPMN". Our assumption of the implementation is that a domain ontology was created before a business experts starts modeling the process. The application is based on the Tensegrity Graph Framework[5], which provides basic functionalities like rendering, editing of diagram elements, event propagation mechanisms, a command stack, and a persistency service for diagrams. To give an impression how context information can be utilized for matchmaking tasks, Figure 4 sketches a simple example scenario concerning the definition of a Data Object.

The *Activity* "send Offer" has been linked to the *domain action* "send Offer". This action causes two transitions, one of them affecting the *domain object* "Offer" the other one affecting the *domain object* "Customer Order". When the user wants to define the *Data Object* more precisely, these two *domain objects* are proposed as possible elements because the *Data Object* is associated to the *Activity* "send Offer". If the *Activity* "send Offer" is not linked to the *domain action* "send Offer" the list of proposals would only contain the *domain object* "Offer", which can be found via name matching with "send Offer". Similar matchmaking capabilities regarding *Data Objects*, *States*, and *Activities* are also

[4] http://www.wsmx.org/

[5] http://www.tensegrity-software.com

incorporated into our extension approach. A whole business process diagram could be described in this manner.

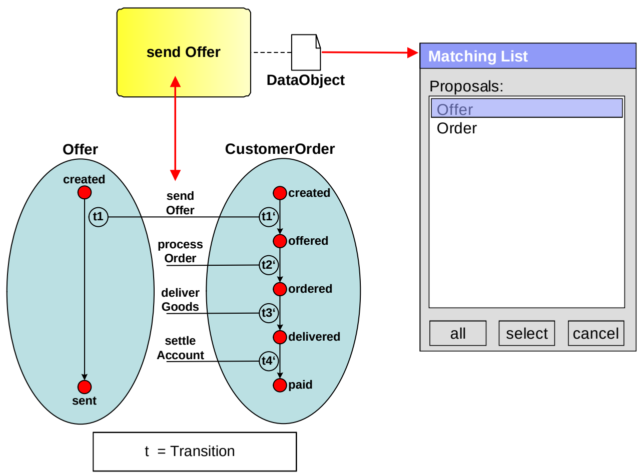

Fig. 4. Data Object Definition Example.

Figure 5 shows an example for an enriched business process model. The Data Object states before the execution of the associated activity are indicated by a "less than" sign (<), the states afterwards by a "greater than" sign (>). The Data Objects and states identified with the help of domain ontologies could be linked to appropriate elements within these ontologies. Not all objects or states may be found in domain ontologies, there may also be unknown or new objects and states in the business process diagrams modeled. These unlinked objects and states may indicate that such new elements should be created and realized respectively within the domain described by the ontologies in order to implement the business process model.

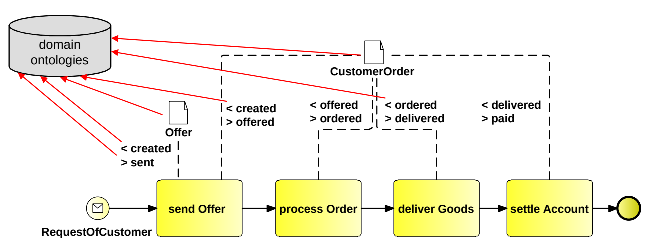

Fig. 5. Use Case Example.

The semantically augmented process models can be used to enhance querying the process space, by answering questions regarding which objects are manipulated within which business processes. If, for example, the EU enacts a new law concerning the expiration date of dairy products, the augmented process models and the knowledge of the domain ontologies can help to find out, which business processes are affected by the new law.

5 Related Work

Since this paper envisions an improvement regarding process modeling activities, the work is related to the field of business process modeling tools. There is a plethora of different BPMN modeling tools like for example Intalio Designer[6], the Eclipse STP BPMN Modeler[7], and ITP-Commerce's Process Modeler[8]. To our knowledge, no commercial tool enables the semantic annotation of business process models.

The working group concerned with semantic business process modeling and analysis tools within the SUPER project is currently working on a prototype of a Business Process Modeling Ontology (BPMO) Editor for modeling semantic business processes in SUPER. The first version of the BPMO modeling environment provides basic functionality for enriching existing process models with semantic annotations (e.g. by assigning a pre-defined semantic goal to an activity).[5, 18] In contrast, the approach introduced within this paper allows for a flexible and fine-granular semantic annotation of processes by directly accessing single ontology elements from the business domains of interest. Regarding process ontologies our work is closely related to the *sBPMN* ontology [1] developed in the SUPER project. The *sBPMN* ontology acts as the basis for our prototypical implementation. In a different context, Lin and Strasunskas [12, 13] present an approach of a General Process Ontology and discuss a semantic annotation framework.

Other related work can be found in [8–10], which is also based on semantic business process modeling. In [10], a class of Petri Nets, so-called Pt/T-Nets are modeled in OWL, and elements can be brought in relation to (ontology-based) data items. These relations are specified in terms of attributes and values, and how these are inserted or deleted as the process executes. There seems to be a clear focus on data instances, and there is a prototypical implementation in a Petri Net modeling tool. [8] presents, among others, a slightly updated model of the Petri Net ontology together with an approach to process decomposition, which is based on the "linguistic specificity" of the terms used in the labels of the elements in a process model. One of the goals is to normalize the terminology in a process model, but rather in terms of the abstraction level used in the labels. The focus in this work, in contrast, is on specifying the semantics of a process model by making use of a given domain ontology. Also, the data objects in our

[6] Intalio Designer, Available: http://www.intalio.com/products/designer/
[7] STP BPMN Modeler, Available: http://www.eclipse.org/stp/bpmn
[8] Process Modeler for Microsoft Visio, Available: http://www.itp-commerce.com/

domain ontology structure can be easily used for capturing business objects, not only data.

The domain ontologies in this paper describe – among other information – data objects including their life-cycles. Hence, any suitable formalism may be used instead which is based on Finite State Machines [6]. However, our ontology-based approach allows capturing intentional human language and machine interpretable information jointly in a standardized language. With the help of these domain descriptions, object state information can be added to the process model. With a different focus, the idea of relating object state information to business process models is also presented in [17]. The approach taken there is to check the consistency of business process models and object lifecycles by adding object state information to specific points in a process model and comparing the process model's usage of the objects with the objects' lifecycles. The approach presented in this paper has a quite different target, as an object's lifecycle is not perceived as an authoritarian model. In [11], an approach for automatic "synthesis" of processes is presented, which means calculating the optimal combination of them. This work describes how annotated reference process models using UML syntax may be used for the synthesis task and describe synthesis algorithms.

The work concerning name-based matchmaking (see Section 3.2) is related to schema matching efforts. A survey of Rahm and Bernstein [15] presents a nice overview of this research area. However, the fundamental difference is that schema matching tries to map between two formalized schemas (e.g. defined in XML), whereas the matching problem addressed in this paper is to find appropriate domain ontology elements given a free-text name entered by the user and context information derived from the process model, concerning the selected element. Nevertheless, the name matching tasks are similar for the two approaches. A comparison of Cohen et al. [4] describes different string distance metrics, some of which are utilized for name matching tasks in this work. Therefore, these established, general approaches for name matching tasks are reused to some extent and are enhanced with new (context-related) matching functionalities.

6 Conclusion

In this paper, we presented an approach for the integration of semantics in modeling tools to support the graphical modeling of business processes with information derived from domain ontologies. The goal is to make these business process models more precise, to ease their (semi-automatic) realization, and to enable querying large process model repositories. For this purpose, we defined suitable semantic information, ontology-based descriptions for business process models and domains, and matchmaking functionalities to support users in modeling semantically annotated process models. Our current implementation assumes the existence of suitable domain ontologies which potentially are created by knowledge engineers. An example approach is to specify the concepts of such a domain ontology manually and derive the instance data from other (non-semantic) models, e.g. from MDD[9]/MDA[10] artifacts or other (structured) knowledge bases. An existing domain ontology does not replace a process model as the purpose of both is quite different. A domain ontology can be seen as a knowledge base for the actual process model, whereas a business process model is used by business experts to capture process knowledge graphically. A remaining open research question is whether a business expert should be able to change the structure of a domain ontology and, if so, what are the appropriate methods and techniques.

We also outlined how Web Service discovery and composition can be facilitated without forcing business process modelers to use unfamiliar, technical terminology, or even formal (ontology) languages. At the same time our approach is more flexible and generic then just attaching predefined goals to the activities of a process model. Besides the ontology-based matching functionalities supporting the process modeling activities with semantic techniques is a major benefit of our approach as opposed to the capabilities of current modeling tools. To prove our concept the SAP research modeling tool "Maestro for BPMN" has been used for a prototypical implementation. In addition, a smart technological concept is needed to utilize the semantically annotated process models and find appropriate services to implement these processes. The areas of potential future work comprise the utilization and enhancement of our approach for the implementation of Web Service discovery and for querying the process space as well as for composition activities in order to (semi-automatically) realize business process models and make them executable. We took first steps towards this goal, which may serve as a basis for further enhancements and new extensions.

References

1. Witold Abramowicz, Agata Filipowska, Monika Kaczmarek, and Tomasz Kaczmarek. Semantically enhanced business process modelling notation. In Martin Hepp, Knut Hinkelmann, Dimitris Karagiannis, Rüdiger Klein, and Nenad Stojanovic, editors, *Proceedings of the Workshop on Semantic Business Process and Product Lifecycle Management (SBPM 2007) in conjunction with the 3rd European Semantic Web Conference (ESWC 2007) Innsbruck, Austria, June 7, 2007*. CEUR Workshop Proceedings, Juni 2007.
2. Mark Burstein, Jerry Hobbs, Ora Lassila, Drew Mcdermott, Sheila Mcilraith, Srini Narayanan, Massimo Paolucci, Bijan Parsia, Terry Payne, Evren Sirin, Naveen Srinivasan, and Katia Sycara. OWL-S: Semantic Markup for Web Services. *W3C Member Submission*, November 2004.
3. Sam Chapman, Barry Norton, and Fabio Ciravegna. Armadillo: Integrating knowledge for the semantic web. In *Proceedings of the Dagstuhl Seminar in Machine Learning for the Semantic Web*, February 2005.
4. William W. Cohen, Pradeep Ravikumar, and Stephen E. Fienberg. A comparison of string distance metrics for name-matching tasks. In *Proceedings of the IJCAI-2003 Workshop on*, pages 73–78, Acapulco, Mexico, August 2003.

[9] Model Driven Development
[10] Model Driven Architecture

5. Marin Dimitrov, Alex Simov, Sebastian Stein, and Mihail Konstantinov. A BPMO based semantic business process modelling environment. In Martin Hepp, Knut Hinkelmann, Dimitris Karagiannis, Rüdiger Klein, and Nenad Stojanovic, editors, *Proceedings of the Workshop on Semantic Business Process and Product Lifecycle Management (SBPM 2007) in conjunction with the 3rd European Semantic Web Conference (ESWC 2007) Innsbruck, Austria, June 7, 2007*. CEUR Workshop Proceedings, Juni 2007.
6. Arthur Gill. *Introduction to Theory of Finite-state Machines*. McGraw-Hill Education, 1962.
7. Martin Hepp, Frank Leymann, John Domingue, Alexander Wahler, and Dieter Fensel. Semantic Business Process Management: A Vision Towards Using Semantic Web Services for Business Process Management. pages 535–540, 2005.
8. Agnes Koschmider and Emmanuel Blanchard. User Assistance for Business Process Model Decomposition. In *Proc. 1st IEEE Intl. Conf. on Research Challenges in Information Science*, pages 445–454, Ouarzazate, Marokko, April 2007.
9. Agnes Koschmider and Andreas Oberweis. Ontology based Business Process Description. In J. Castro and E. Teniente, editors, *Proceedings of the CAiSE05 WORKSHOPS, no. 2*, pages 321–333, Porto, Portugal, June 2005.
10. Agnes Koschmider and Daniel Ried. Semantische Annotation von Petri-Netzen. In *Proceedings des 12. Workshops Algorithmen und Werkzeuge fr Petrinetze (AWPN'05)*, pages 66–71, Berlin, Germany, September 2005.
11. Florian Lautenbacher and Bernhard Bauer. Semantic reference and business process modeling enables an automatic synthesis. In *Proceedings of Semantics for Business Process Management Workshop (SBPM) of the European Semantic Web Conference (ESWC), 2006, Budva, Montenegro*, 2006.
12. Yun Lin and Darijus Strasunskas. Ontology-based semantic annotation of process templates for reuse. In *Proc. of 10th Intl. workshop EMMSAD'05*, Porto, Portugal, June 2005.
13. Yun Lin, Darijus Strasunskas, Sari Hakkarainen, John Krogstie, and Arne Slvberg. Semantic annotation framework to manage semantic heterogeneity of process models. In Eric Dubois and Klaus Pohl, editors, *CAiSE*, volume 4001 of *Lecture Notes in Computer Science*, pages 433–446. Springer, 2006.
14. Deborah L. Mcguinness and Frank van Harmelen. OWL Web Ontology Language Overview. February 2004.
15. Erhard Rahm and Philip A. Bernstein. A survey of approaches to automatic schema matching. *The VLDB Journal*, 10(4):334–350, December 2001.
16. Dumitru Roman, Uwe Keller, Holger Lausen, Jos de Bruijn, Rubn Lara, Michael Stollberg, Axel Polleres, Cristina Feier, Cristoph Bussler, and Dieter Fensel. WSMO - Web Service Modeling Ontology. *Applied Ontology*, 1(1):77–106, 2005.
17. Ksenia Ryndina, Jochen Malte Küster, and Harald Gall. Consistency of business process models and object life cycles. 4364:80–90, 2006.
18. Michael Stollberg and Barry Norton. A refined goal model for semantic web services. In *Proc. of 2nd International Conference on Internet and Web Applications and Services (ICIW 2007)*, May 2007.
19. Ingo Weber, Ivan Markovic, and Christian Drumm. A Conceptual Framework for Composition in Business Process Management. In *BIS-07: Proceedings of the 10th International Conference on Business Information Systems*, Poznan, Poland, April 2007.