

# Compliance Check in Semantic Business Process Management

András Gábor<sup>1</sup>, Andrea Kő<sup>2</sup>, Ildikó Szabó<sup>2</sup>, Katalin Ternai<sup>2</sup>, and Krisztián Varga<sup>2</sup>

<sup>1</sup> Corvinno Technology Transfer Center Ltd  
Közraktár utca 12/a, Budapest, Hungary, 1093  
agabor@corvinno.hu

<sup>2</sup> Department of Information Systems,  
Corvinus University of Budapest  
Fővám tér 13-15. Budapest, Hungary, 1093  
{ko, iszabo, ternai, kvarga}@informatika.uni-corvinus.hu

**Abstract.** With a steady increase of requirements against business processes, support of compliance checking is a field having increased attention in information systems research and practice. Compliance check is vital for organizations to identify gaps, inconsistency and incompleteness in processes and sometimes it is mandatory because of legal, audit requirements. The paper gives an overview about our research and development activities in the field of compliance checking with the help of semantic business process management (SBPM). We propose a compliance checking approach and solution, illustrated with a use case from higher education domain.

**Keywords:** compliance checking, semantic business process management, ontology matching.

## 1 Introduction

The paper discusses a semantic technology-based approach for checking the compliance of business processes. Business process management (BPM) means modeling and managing processes which are interpreted as know-how platforms of the organization. BPM helps to maintain and manage the organizational processes and process attributes like input/output information, technology used and human resources needed to perform each activity. However current BPM approaches are not able to reflect to the fast changes in complex economic, social, and regulatory environment. Traditional and widely used BPM life cycles covers business process strategy formulation, process documentation, analysis and design, implementation and change management, operation, monitoring and controlling. A business process is seen as a sequence of activities, including the definition from the resource management perspective, which is responsible for the execution of each activity in terms of authority, responsibility (accountability) and competences (knowledge, skill, and attitude). Typically this information is stored in the RACI matrix (Responsible, Accountable, Consulted, Informed), which connects the organizational model to the process model. One of the

overall objectives of BPM is the articulation of the hidden knowledge in the organizational, production and service dimensions. An obstacle in achieving this objective is the contradiction between the static business process model and the dynamically changing organizational knowledge, that is aimed to be solved using agile BPM, interaction centric BPM or event-driven BPM. With a steady increase of requirements for business processes, automation support of compliance management is a field having growing attention in information systems research and practice. Several approaches have been developed to support compliance checking of process models [9] [15] [21] [23]. One major challenge in such approaches is their ability to handle different modeling techniques and compliance rules in order to enable widespread adoption and application in the organization. Compliance check is vital for organizations, because of the following reasons:

- to assess of processes by checking compliance with standards, regulation, benchmarks and reference models.
- to identify gaps, inconsistency and incompleteness of processes using open data as source
- to improve processes via recommendations and the collaboration mechanisms.

Compliance check of the processes and the embedded knowledge has to be proactive. Decision makers need to be supported via smart, semi-automated or automated compliance checks in their day-to-day work, which are currently not applied in practice, meanwhile open data; linked open data has the proper capacity to provide knowledge for the business process assessment. Becker, Defman, Hervig and Lis [2] draw attention to consider domain semantics in business process compliance-checking approaches, opening another important research field. Their review revealed the lack of real-world evaluation is a common problem in the field of business process compliance checking. The paper is structured as follows: A theoretical overview on semantic process management, ontology learning and ontology matching is presented along with its implications to the research approach. A short overview is given about compliance checking solution, illustrated by a use case from higher education domain. Finally, conclusion is drawn and future research directions are set.

## 2 Related Work

### 2.1 Semantic Process Management

Business processes have to perform well within dynamically changing organizational environments. The main challenge in BPM is the continuous translation between business requirements view on the process space and the actual process space. It can be expected that BPM will only come closer to its promises if it allows for a better automation of the two-way translation. SBPM is a new approach of increasing the level of automation in the translation between these two levels, and is currently driven by major players from BPM, and Semantic Web Services domain. The core paradigm of SBPM is to represent the two spheres using ontology languages and to employ machine reasoning for automated or semi-automated translation.

A competitive enterprise has to adapt core value-added processes with unprecedented speed, to act appropriately regardless of the situation. The focus of process designers is to make more sophisticated use of process architectures and continuous improvement of processes. Conceptual model captures the semantics of an application through the use of a formal notation, but the descriptions resulting from conceptual model are intended to be used by humans and not machines. The basic idea of SBPM is to combine Semantic Web Services frameworks, ontology representation, and BPM methodologies and tools, and to develop a consolidated technology [10] [23]. The SUPER project elaborated a semantic framework for compliance management. They presented five perspectives on compliance checking: design-time/run-time; forward/backward; active/passive; task checking/process checking or engine-based/query-based perspective [6].

## 2.2 Ontology Learning

Web 2.0 solutions, social media applications are characterized by the huge amount of content created, shared and consumed by the users. In most of the cases information exchange and knowledge sharing require semantic interoperability. Ontologies play a decisive role in supporting the interoperability, because they provide the necessary structure and framework. Manual ontology building is expensive, time consuming, error-prone, biased towards their developer, inflexible and specific to the purpose of construction [8][19]. Ontology learning is a research area, which deals with the challenges to turn facts and patterns from content into shareable high-level constructs or ontologies [21]. One of the first ontology learning surveys was published by the OntoWeb Consortium in 2003 [8]. They investigated 36 approaches of ontology learning from text. Shamsfard & Barforoush introduced a framework for comparing ontology learning approaches [20]. Buitelaar and his colleagues introduced the “ontology learning layer cake” phrase to describe the different subtasks of ontology learning process [3]. Zhou concluded that human involvement into ontology learning still remains necessary and desirable [24]. He highlighted the need for common benchmarks for evaluating ontologies. Discovery of fine-grained associations are open problem, too. Existing techniques should be customized for cross-domain text on a Web-scale, because expert-curate domain knowledge is no longer adequate. Researchers agreed that ontology learning requires more work in the following areas: 1) discovery of fine-grained relations between concepts; 2) developing common evaluation platforms for ontologies, and 3) knowledge acquisition should focus on Web - to make ontology learning operational on a Web-scale. It was also revealed from the surveys that fully automatic learning of ontologies may not be possible [22] [24].

## 2.3 Ontology Matching

Alasoud et al. [1] define ontology matching problem as follows: “given ontologies O1 and O2, each describing a collection of discrete entities such as classes, properties, individuals, etc., we want to identify semantic correspondences between the components of these entities.”

The general ontology mapping tools investigated by Noy [17] and Choi et al. [4], or the one developed by Protégé community use different kind of methods to identify the semantic correspondences between two ontologies. The most well-known methods are axiomatic correspondences (e.g. OWLDIFF [18], Compare Ontologies function in Protégé 4.X) or calculating similarity values. The latter takes probability distributions (e.g. Glue [5], OMEN [16]) or text similarity functions (e.g. LOM [14]) as a basis. Jung discussed the process specific methods [11] and used logical assertions and similarity measures to facilitate the interoperability among processes. Koschmider and Oberweis [12] used Petri nets „to obey an operational semantics that facilitates composition, simulation, and validation of business processes“.

### 3 Compliance Checking Solution Overview

The logic of the solution is summarized in the Figure 1:

1. Knowledge extraction and representation from open data (guidelines, standards, etc.) – Components populates the *reference ontology*.
2. Knowledge extracted from the business processes populates the *process ontology*.
3. Compliance check will be done with the reference and process ontology resulting the gap identification.
4. The identified gaps are interpreted and fed back to the process owner.

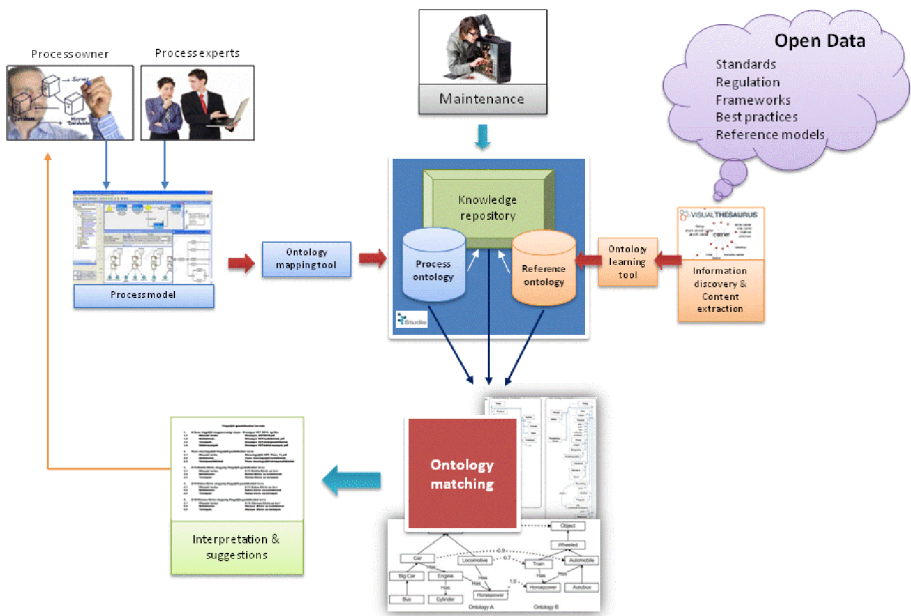


Fig. 1. Overview of compliance checking solution

The detailed description of the components will be given in the appropriate parts of the next section.

## 4 Use Case: Compliance Checking of the ‘Quality Assurance’ Process in a Higher Education Institution

Quality assurance is vital part of university governance, under the frame of Bologna process several directives give framework of quality assessment. One of the most accepted is the ESG [7].

### 4.1 Process Ontology

Adonis process modeling tool was applied as a business process modeling environment, in which “quality assurance” process was created. For the sake of extension and mapping the conceptual model to ontology model, XML export of the Adonis process was generated. In the export all objects of the business process model is an ‘instance’ in the XML structure, the attributes have the tag ‘attribute’, while the connected objects (such as the performer, the input/output data, etc.) have the tag ‘interref’. The next step is to establish links between model elements and ontology concepts. [13]. The „conceptual models - ontology models” converter maps the Adonis business process modeling elements to the appropriate ontology elements in meta-level. The model transformation aims at preserving the semantics of the business model. The general rule is to express each model element as a class in the ontology and its corresponding attributes as attributes of the class. This transformation is done by the means of XSLT script which performs the conversion. The converted OWL ontology in the structure of Protege/OWL XML format is imported into the editor of Protege 4.2.

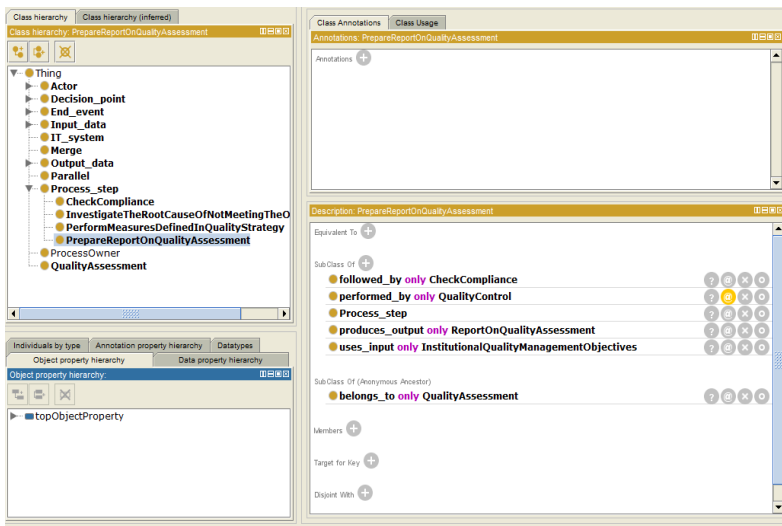


Fig. 2. The process ontology in Protege 4.2

The business model first must be represented within the ontology in order to specify the semantics of ADONIS model elements through relations to ontology concepts. In regard to the representation of the business model in the ontology, one can differentiate between a representation of ADONIS model language constructs and a representation of ADONIS model elements. ADONIS model language constructs such as “activity”, as well as the control flow are created in the ontology as classes and properties. Subsequently, the ADONIS model elements can be represented through the instantiation of these classes and properties in the ontology. The linkage of the ontology and the ADONIS model element instances is accomplished by the usage of properties. These properties specify the semantics of an ADONIS model element through a relation to an ontology instance with formal semantics defined by the ontology (Figure 2).

### 4.2 Reference Ontology

The name “reference” comes from the fact that the main sources of reference ontology development are standards, best practices, policies and laws. Several guidelines, standards, best practices, benchmarks in quality assurance are available; we don’t discuss the specific issues on discovery and collection of that. ESG as a source document was processed. The first step in text mining process executed by SPSS Modeler was to split the document into sentences. 1287 words were found, possessives, attributive structures in 1026 sentences. Based on these sentence elements 125 categories were defined (see in Figure 3).

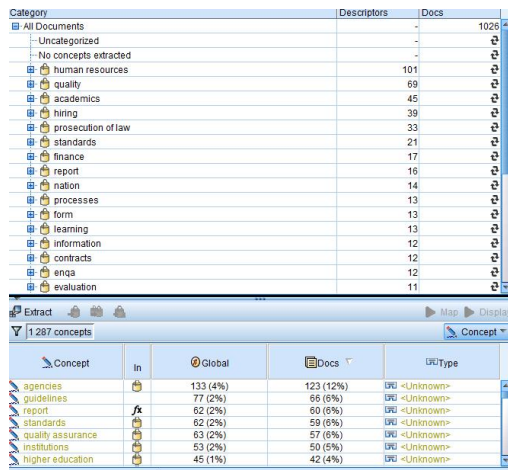


Fig. 3. Categories extracted by SPSS Text Mining

The categories did not follow the meta-structure given by the process model. In order to identify the main objects of the process model (Process step, Document, Actor) the restructuring of the existing categories were needed, new categories were inserted. The rationale behind of the text mining was to reveal sentences of the ESG document that relate to the process models. In above another text mining model was built in

RapidMiner in which ESG document was analyzed and resulted n-grams (4706 2-grams of tokens) and their occurrences. In the course of reference ontology the 2-grams of tokens provided by RapidMiner were taken into consideration. We dispensed the elements of these concomitances into the appropriate categories (see in Figure 4.).

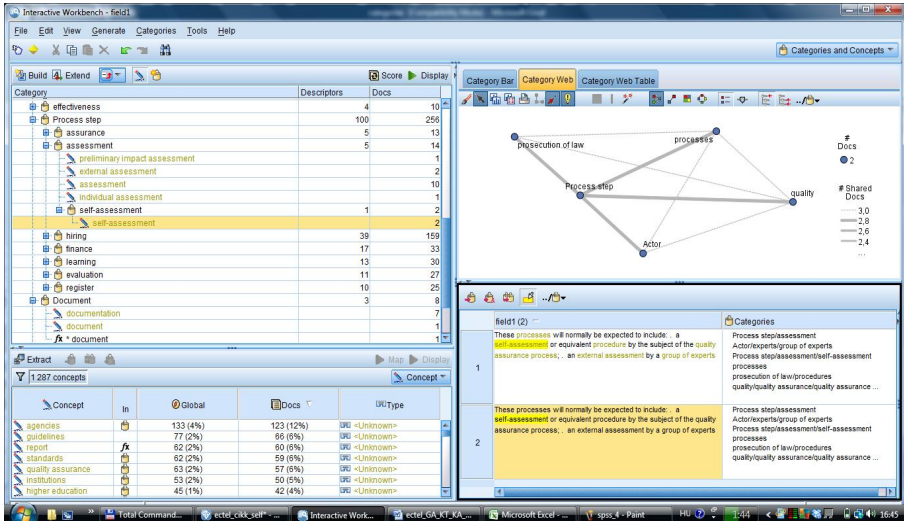


Fig. 4. The concomitances of categories

The next step of the text mining was to find the concomitances of these categories in sentences. Figure 4 shows the results for *self-assessment* category.

An algorithm written in Java searched the categories named by process model objects. The process model objects served as meta-classes in the process ontology, hence the sub-categories became sub-classes (Fig. 5).

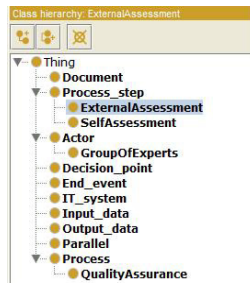


Fig. 5. Reference Ontology in Protégé

### 4.3 Ontology Matching

Protégé 4.2 and OWL DIFF [18] Protégé plugin was used as an ontology matching environment. *Compare Ontologies* is a built-in function in Protégé 4.X. It is capable of comparing ontologies with different name space while OWL DIFF isn't. The report created by *Compare Ontologies* is transparent; appropriate for ontology matching.

In the current stage of research it deals with building taxonomy from unstructured texts, without creating any non-taxonomic relationship between the classes; we used the matching solution to discover discrepancies among the subclasses of a given meta-class (Process step, Actor etc.). Using *Compare Ontologies* the following questions are answered:

- Q1: Are the same actors responsible for the execution of the processes?
- Q2: Do the processes consist of the same steps?
- Q3: Are the same documents used or created in the processes?

*Compare Ontologies* function compares two versions of the same ontology. In our case the Process Ontology was considered to be the original model, and, in order to assess the matching, the Reference Ontology was used as updated version. The result of the matching is exhibited in Figure 6. The *Created blocks* shows which ontology elements are in the Reference Ontology and are not in the Process Ontology. *Compare Ontologies* cannot compare ontology elements semantically. There are two possible outputs: 1) elements exist in Process Ontology, but with different names that are semantically similar; 2) elements do not exist in Process Ontology, therefore we have to insert them into the actual process. The *Deleted blocks* present which ontology elements are in the Process Ontology and are not in the Reference Ontology. There are two possible conclusions: 1) they are not identified in the text yet. In the future work, we have to elaborate a method to evaluate the efficiency of our ontology learning method, 2) they do not exist in the Reference Ontology, we have to eliminate them from Process Ontology. The ontology elements in the *Renamed and Modified block* are common and exist in the same taxonomic or non-taxonomic – axiomatic – relationship, but paired with a different element.

Aligned the lone unmatched siblings.		
Description	Baseline Axiom	New Axiom
Renamed	Class: QualityControl	Class: GroupOfExperts

Fig. 6. Part of the output of the Compare Ontologies ontology matching tool

In summary, the Created block can discover the new elements, the Deleted block can show the out-of-dated elements, and the Renamed and Modified block can reveal the structurally discrepancies between the process models. The first two blocks answer questions Q1 and Q3, the last blocks replies for Q2 question. The result of matching is the following: *Document process model* object appears in Process Ontology as *Input data* and *Output data*. In the course of ontology learning, we have to elaborate a rule to identify the role of a document in the process, based on its context. Decision maker should decide about the creation of a separate process for self-assessment and an external assessment, too. The tool discovered that a ‘group of experts’ is interested in the quality assurance process, instead of the ‘Quality Control actor’. The *interpretation component* gives a verbose explanation of the findings summarized in the above paragraph.

## 5 Conclusion and Future Research

This paper provided an overview about a compliance checking solution in SBPM. The main goal of our approach to demonstrate a way to assess business processes by checking compliance with standards, regulation, benchmarks and reference models. Higher education “quality assurance” process was detailed in our use case. Through the use case the suggested solution was justified as a proof of concept in terms of the potential to identify gaps, inconsistencies and incompleteness of processes. The solution heavily builds on open data as input. Our solution provides a semi-automatic support for compliance checking. The XSLT-transformation and the ontology matching can be executed automatically by using a Java API. The building of the Reference Ontology requires human intervention due to the interpretation of the ontology learning results. Future work includes the extension of the compliance checking solution with web mining, crawlers to support identification and collection the relevant sources and to continue testing the solution in other processes and domains. Another development direction is associated with the disadvantage of *Compare Ontologies* ontology matching tool. The tool does not compare semantically the name of ontology elements and as a consequence, the generated report may be less transparent in the case of complex process ontologies. Since it is open source software, we are going to improve it with a new algorithm working with similarity measure.

## References

1. Alasoud, A., Haarslev, V., Shiri, N.: An Effective Ontology Matching Technique. In: An, A., Matwin, S., Raś, Z.W., Ślęzak, D., et al. (eds.) ISMIS 2008. LNCS (LNAI), vol. 4994, pp. 585–590. Springer, Heidelberg (2008)
2. Becker, J., et al.: Generalizability and Applicability of Model-Based Business Process Compliance-Checking Approaches – A State-of-the-Art Analysis and Research Roadmap, vol. 5(2), pp. 221–247. German Academic Association for Business Research, VHB (November 2012)
3. Buitelaar, P., et al.: Ontology Learning from Text: An Overview. IOS Press (2003)
4. Choi, N., et al.: A survey on ontology mapping. SIGMOD Rec. 35, 34–41 (2006)

5. Doan, A.H., et al.: *Ontology matching: A machine learning approach*. In: Staab, S., Studer, R. (eds.) *Handbook on Ontologies*, pp. 385–516 (2004)
6. El Kharbili, M., et al.: *Towards a framework for semantic business process compliance management*. In: *Proceedings of the Workshop on Governance, Risk and Compliance for Information Systems*, pp. 1–15 (2008)
7. ESGStand: *European Standards and Guidelines*, <http://www.eqar.eu/application/requirements/european-standards-and-guidelines.htm> (accessed April 9, 2013)
8. Haller, A.: *Ontologies and semantic Business Process Management 2007* Digital Enterprise Research Institute (2007), <http://www.soberit.hut.fi/T-86/T-86.5161/2007/Ontologies%20and%20semantic%20BPM.pdf>
9. Hepp, M., Cardoso, J., Lytras, M.D.: *The Semantic Web: Real-World Applications from Industry*. Springer (2007) ISBN: 0387485309
10. Gómez-Pérez, A., Manzano-Macho, D.: *A survey of ontology learning methods and techniques*. In: *OntoWeb Consortium* (2003)
11. Jung, J.J.: *Semantic business process integration based on ontology alignment*. *Expert Systems with Applications* 36(8), 11013–11020 (2009)
12. Koschmider, A., Oberweis, A.: *Ontology based business process description*. In: *Proceedings of the CAiSE*, pp. 321–333 (2005)
13. Kramler, G., Murzek, M.: *Business Process Model Transformation Issues* (2006)
14. Li, J.: *LOM: a lexicon-based ontology mapping tool*. In: *Information Interpretation and Integration Conference, I3CON* (2004)
15. Liu, A.Y., Müller, S., Xu, K.: *A Static Compliance-Checking Framework for Business Process Models*. *IBM Systems Journal* 46(2), 335–361 (2007)
16. Mitra, P., Noy, N., Jaiswal, A.R.: *OMEN: A Probabilistic Ontology Mapping Tool*. In: Gil, Y., Motta, E., Benjamins, V.R., Musen, M.A. (eds.) *ISWC 2005*. LNCS, vol. 3729, pp. 537–547. Springer, Heidelberg (2005)
17. Noy, N.F.: *Semantic integration: a survey of ontology-based approaches*. *SIGMOD Rec.* 33, 65–70 (2004)
18. OWLDiff: *OWL Diff Documentation* (2008), <http://krizik.felk.cvut.cz/km/owlldiff/documentation.html>
19. Sabou, M., et al.: *Learning domain ontologies for Web service descriptions: an experiment in bioinformatics*. In: *Proceedings of the 14th International Conference on World Wide Web*, pp. 190–198. ACM, New York (2005)
20. Shamsfard, M., Barforoush, A.: *The state of the art in ontology learning: a framework for comparison*. *The Knowledge Engineering Review* 18(4), 293–316 (2003)
21. Schmidt, R., Bartsch, C., Oberhauser, R.: *Ontology-based representation of compliance requirements for service processes*. In: *Proceedings of the Workshop on Semantic Business Process and Product Lifecycle Management, SBPM 2007* (2007)
22. Wong, W., et al.: *Ontology learning from text: A look back and into the future*. *ACM Comput. Surv.* 44(4), 20:1–20:36 (2012)
23. Wetzstein, B.; Zhilei Ma; Filipowska, A.; Kaczmarek, M.; Bhiri, S.; Losada, S.; Lopez-Cobo, H.M.; Cicurel L.: *Semantic Business Process Management: A Lifecycle Based Requirements Analysis*. In: *Proceedings of the Workshop SBPM* (2007), <http://ceur-ws.org/Vol-251/paper1.pdf>
24. Zhou, L.: *Ontology learning: state of the art and open issues*. *Information Technology and Management* 8(3), 241–252 (2007)