A software platform for semantics-based enterprise knowledge management

Francesco Taglino, Fabrizio Smith, Maurizio Proietti

National Research Council, IASI "Antonio Ruberti"
Viale Manzoni 30, 00185 Roma, Italy
{taglino,smith,proietti}@iasi.cnr.it

Abstract. In this paper we address the problem of knowledge management and interoperability in virtual enterprise environments where knowledge is often fragmented and heterogeneous. We propose a knowledge repository and management infrastructure, called Production and Innovation Knowledge Repository (PIKR), to support open innovation in virtual enterprises. The PIKR provides a set of reference ontologies to semantically describe enterprise knowledge resources, and semantics-based services for accessing and reasoning over such descriptions. We also give an overview of the implementation of the PIKR that is being carried on in the BIVEE European project.

Keywords: business innovation, ontologies, semantic services, virtual enterprises.

1 Introduction

In the era of the globalised market which has deeply transformed the world economy, and increased the international competition, SMEs are more and more pushed to form business alliances and work in virtual enterprises (VEs). Enterprise networks can be a means to reach the critical mass required by the expanding markets. However, heterogeneities of the network's members can generate interoperability problems, with the consequence of reducing the expected benefits. Here we focus on interoperability at knowledge and information level that impacts on the possibility of exchanging and accessing common knowledge resources within a VE, and in particular within and across the *production space* (where all the activities related to the core business take place), and the *innovation space* (mainly characterised by creative units and cooperative interactions) of the VE itself.

In a context of networked enterprises, this aspect is very crucial because it allows relevant resources to be shared, in order to assess how production activities are actually performed, and how performing they are, what kinds of resources (in terms of skills and expertise) the virtual enterprise can count on, what documental resources (e.g., market analyses, technical reports) have been produced or acquired by the VE.

It is also very important to know how innovation-related initiatives are carried on, e.g., the degree of participation of people to brainstorming activities, the number of relevant ideas collected in certain periods of time, the number of proposed ideas which have been concretely exploited, and so on.

To address interoperability problems, a good solution is the development of a semantics-driven common knowledge base, characterized by the adoption of shared reference ontologies. The purpose of this paper is to present a proposal of a knowledge-based infrastructure, named PIKR (Production and Innovation Knowledge Repository), to support knowledge interoperability and management in VE environments. This proposal is being conceived in the framework of the BIVEE European project[1], and adheres to the Linked Data approach [1].

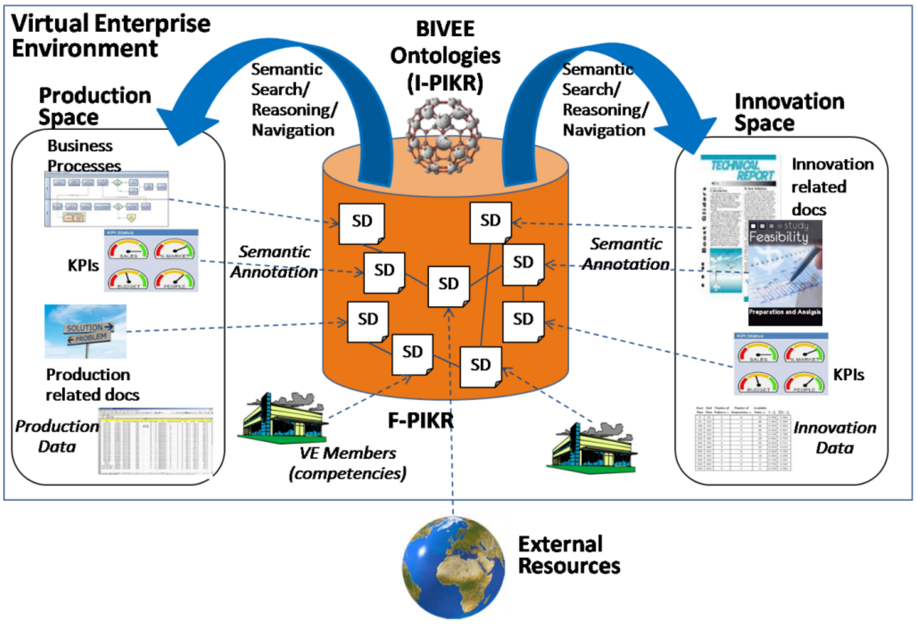

Fig. 1. The Production and Innovation Knowledge Repository overview

Following the Linked Data approach which recommends a set of best practices for exposing, sharing, and connecting pieces of data, information, and knowledge by using semantic web technologies, the PIKR provides, on the one hand a set of reference structures (i.e., ontologies) for the semantic description of enterprise knowledge resources, and on the other hand semantics-based services for accessing and reasoning over such descriptions.

2 PIKR Ontological Framework

The mission of the PIKR is to create a semantics-based unified view of the information and knowledge that flow within and across the Production Space and the Innova-

[1] Business Innovation in Virtual Enterprise Environments. http://bivee.eu

tion Space of VEs. In particular, these two spaces are seen through the following types of knowledge resources: *Processes,* which describe actual production activities; *Documents,* which are concrete footprints of all kinds of activities, both at production and innovation level; *Actors* and their competencies, which refer to the capabilities of the VE and its members; *Key Performance Indicators* (*KPIs*), for monitoring both the Production and the Innovation space.

Then, the PIKR is organized into two layers (Fig. 1): the *Intensional* PIKR (I-PIKR), which contains a federation of ontologies to describe the enterprise resources, and the *Factual* PIKR (F-PIKR), which contains the semantic representation (*Semantic Descriptors*) of the actual enterprise resources in terms of the above ontologies.

The ontologies in the I-PIKR are partitioned into *Knowledge Resource Ontologies* (KROs), and *Domain Specific Ontologies* (DSOs). KROs are independent of any application domain and declare what kind of information, links, constraints and business rules, for each type of knowledge resource (i.e., Processes, Documents, Actors, and KPIs), we intend to semantically represent (*Semantic Descriptors Skeleton*, *SDS*), while DSOs allow Semantic Descriptors to be enriched with domain specific contents (e.g., furniture domain).

According to this view, the I-PIKR contains four main KROs: *ProcOnto*, *DocOnto*, *ActorOnto*, and *KPIOnto*, for describing processes, documents, actors, and key performance indicators, respectively, and inter-connections between them. Consequently, the Semantic Descriptors, which describe actual knowledge resources (e.g., the technical report realized in a specific project, or the process for producing a certain product) will be instances of the KROs. The Semantic Descriptor Skeletons will be characterized by a common structure organized into the following sections:

- **Header**: collects information represented by traditional metadata like the ones proposed by the Dublin Core Vocabulary[2] (e.g., name, natural language description). This section also contains the link to the actual knowledge resources that is assumed to be stored in a proprietary system, e.g., a content management system.
- **Domain Specific Content**: collects information about what the content talks about in terms of the DSOs. For instance, in this section one can say that a given technical report is about the design of an innovative *contour chair in carbon fibre*.
- **Related Knowledge Resources**: collects links to related Semantic Descriptors, allowing the representation of semantic associations and dependencies among resources (e.g., the input document of an activity, an indicator occurring in a formula defining a KPI).
- **External Links**: links to resources external to the VE available on the internet (e.g., technical documentation, external policies or regulations, web-sites).
- **Extended Representation**: links to representations of the resource that will allow the enactment of specific reasoning facilities (e.g., a mathematical representation of a KPI, a machine processable representation of a business process).

Domain Specific Content and Related Knowledge Resources items can be enriched with business rules, i.e, constraints that characterize the semantic descriptor skeletons

[2] http://dublincore.org

with respect to the particular reality of a virtual enterprise. These constraints can depend, for instance, on the specific application domain, on the dimensions of the VE, or on the VE internal policies. For example a feasibility study for justifying the prototyping of a product needs financial information, if the expected cost is higher than a certain amount (e.g., 300 Keuro). This means that for feasibility studies, financial information is not always mandatory, even in the same VE.

For the definition of the KROs we are following different approaches with respect to the different kinds of resources. In the case of the DocOnto, we are considering both production and innovation related documents. For the production documents (e.g., invoices, bills of materials) there is plenty of literature and standards (e.g., UBL [2], RosettaNet [3]), which describe information items and dependencies among such documents. For the innovation documents (e.g., ideas, project proposals) we are mainly eliciting needed information through the interaction with end users of the BIVEE project. In particular, the BIVEE project has introduced a document-centric vision of innovation, based on four waves: *creativity*, *feasibility*, *prototyping* and *engineering*. We are then analysing how the BIVEE end users currently address innovation generation with respect to these waves, what kinds of documents they produce, which dependencies and constraints are among these documents.

In the case of the ProcOnto, we refer to a logic-based language for representing and reasoning with process knowledge. We propose to adopt BPAL (Business Process Abstract Language), which is a process ontology, strongly inspired to the BPMN [4] notation. BPAL provides an explicit formalization of the meta-model and of the execution semantics thus allowing advanced BP querying facilities [5] that take into account both the structure (i.e., the workflow graph underlying the BPs) and the behavior (i.e., the possible executions) of BPs. Thanks to its grounding into logic programming, BPAL can be easily adopted in conjunction with rule-based ontology languages (e.g., OWL-RL [6]) for the annotation of BP schemas with respect to domain specific ontologies.

In the case of the KPIOnto, we refer to existing classifications like the one in [7] which categorizes KPIs into Operative, Administrative and Strategic, and in particular, to the Value Reference Model[3] (VRM) which provides a standard classification of KPIs both for production and innovation activities. Furthermore, we intend to address a formal representation of mathematical structures of KPIs in order to enable some forms of reasoning on them, such as the ability to check semantic correctness and redundancies of KPI definitions and the analysis of dependencies among KPIs [8].

3 PIKR Services

KROs, DSOs and Semantic Descriptors represent the Knowledge Repository of the PIKR, on top of which some semantics-enabled services are made available by the PIKR Reasoner. The services are here briefly described.

[3] http://www.value-chain.org/en/cms/1960

Search. This module provides keyword-based search services, following an interaction paradigm similar to traditional web information retrieval engines. The user request is expressed as an *ontology-based feature vector* describing the criteria for the selection of the resources of interest. The search engine returns a list of ranked results by applying semantic similarity techniques (e.g., the *SemSim* metric [9]) to compute the degree of matching between the concepts used to formulate the given request and the ones used to describe the available resources. For instance, suppose that the user is interested in finding all the *documents* which have been authored in the last two years and concerning the initial stages of the design of a piece of *furniture* equipped with an *electronic device*. The corresponding request should be then formulated by using terms defined in the I-PIKR, e.g., {*Resource:Document*, *Wave:Creativity*, *Content*:[*Domotics*, *Furniture*, *Electronic_Device*], *Year>2010*}. Rather than simply providing links from search results to the source documents in which the keywords are textually mentioned, the engine will retrieve semantically related resources, such as *Proposed_Idea* or *Project_Proposal* documents (which are assumed to be defined in the *DocOnto* as specific types of Creativity Wave documents) about a *Contour_Chair* with an embedded *Media_Player* (which are assumed to be defined in the DSO as kinds of piece of furniture and electronic device, respectively).

Query. This module provides services to retrieve pieces of knowledge which exhibit some given properties. Queries are posed in terms of the vocabulary and semantic relations provided by the PIKR ontologies, and the underlying reasoning engine returns a list of answers that satisfy all specified properties. These answers may consist of factual knowledge (semantic descriptors), conceptual knowledge (ontological terms), or references to resources. The most prominent standard for querying OWL/RDFS resources is the SPARQL (SPARQL Protocol and RDF Query Language) [10] standard, defined by the World Wide Web Consortium and widely accepted in the semantic web community. SPARQL is in fact designed to query RDF resources, that essentially are organized as directed and labeled graphs, by matching graph patterns over RDF graphs. We are currently developing a query language based on the SELECT-FROM-WHERE paradigm to extend the SPARQL language by providing additional primitives to be used specifically for querying particular resource kinds besides RDF models (e.g., BPs, KPIs). The Query service can be useful, for instance, in a scenario where there is the need of reengineering a process, as a consequence of an alert emerged by the KPI-driven monitoring. In this case we may want to retrieve all the documents related to the given process which have been defined in the Engineering Wave. The query engine will return as answer a list of (links to) documents specifying the procedures (e.g., *Quality_Protocol* or *Assembling_Protocol* documents) implemented by the process itself.

Consistency Checking. This module provides services for checking the compliance of the factual knowledge captured in the semantic descriptors with respect to business policies and internal regulations established for the whole VE or for individual enterprises. Such compliance requirements are represented in the I-PIKR in terms of business rules, i.e., statements that define or constrain some aspects of the business, specifying the structure of the domain entities (structural constraints) and influencing the way business operations are conducted (behavioral constraints). In our frame the compliance check takes place by verifying the consistency between the assertions contained in the F-PIKR and the axioms defined in the Knowledge Resource Ontologies formalizing the business rules. An example of structural constraint is “Each *Innovation_Report* needs to be composed by a *Project_Proposal* and a *Market_Report*”, while an example of behavioral constraint is “A *Monitoring_Sheet* cannot be produced unless a *Gantt_Chart* has been finalized before”.

KPI Reasoning. This module provides inference services for supporting KPI elicitation (i.e., the identification of the KPIs which are suitable for a given VE), by analyzing KPIs from different perspectives (e.g., organization and time dimensions). This module also supports the harmonization of the measures provided by VE members which are needed for the evaluation of KPIs. Indeed, since measures can be originated by different data sources (e.g., proprietary information systems) from different enterprises in the VE, they need to land on a reference representation compliant with the KPI formulas. Examples of heterogeneities between data definitions and required input for KPI evaluation could be in terms of terminology (e.g., *Customer_Requested_Date* vs. *Expected_Delivery_Date*), or granularity (e.g., aggregated vs. atomic data).

4 Architecture and Implementation

Figure 2 presents an overview of the architecture of the PIKR.

The ***PIKR GUI*** enables user interaction through a wiki-like environment, which provides the means to: *i)* semantically annotate knowledge resources (Semantic Annotation Widget); *ii)* navigate the resources by following semantic links expressed in the corresponding semantic descriptors (Resource Browsing Widget); *iii)* retrieve and process information through the facilities provided by the PIKR Reasoner (Reasoning Widget); and *iv)* supporting the editing of I-PIKR Ontologies, e.g., adding a new constraint or business rule (Ontology Building Widget).

The ***PIKR Reasoner*** implements the different semantic services described in the previous section, which are in the *Service Library* and made available through the *PIKR Service Manager* which is the entry point for the PIKR reasoning functionalities. The PIKR Service Manager is exposed as a web service through the *Service Integration Middleware*, in order to make available PIKR services to the other components of the BIVEE Platform (MCR and VIF). The core component of the PIKR Reasoner module is the *Inference Engine*, which is a rule-based reasoner implementing the various reasoning methods.

The ***PIKR Storage System*** is responsible for importing and maintaining the semantically enriched representation of the informative resources to be stored by the PIKR.

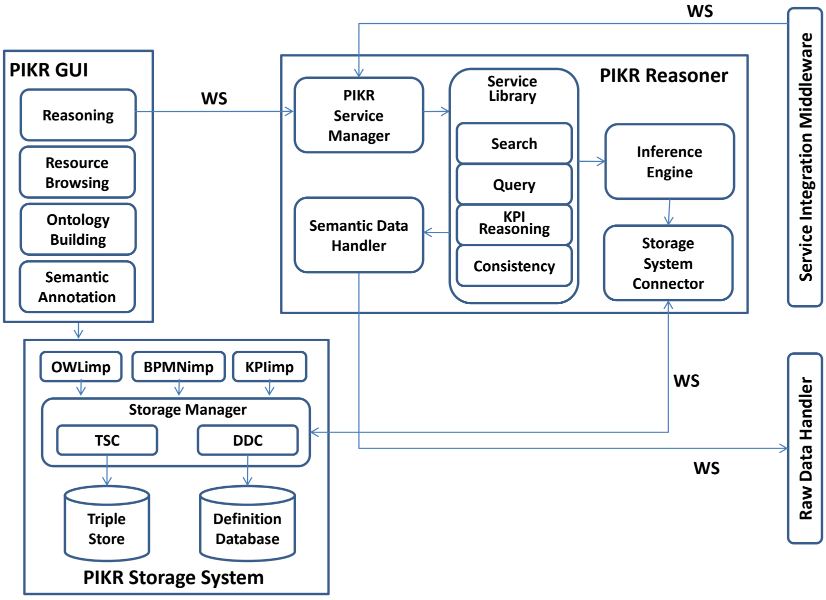

Fig. 2: PIKR Architecture

We are currently developing the PIKR on top of the Semantic Mediawiki plus[4] (SMW+) suite which provides a solid infrastructure for building powerful and flexible "collaborative knowledge-bases" upon a wiki, and also encompasses user-friendly environments for presenting and collecting both human-readable and machine-processable contents. SMW+ also allows the integration of different *triple stores* (we are currently using the Jena'[5] toolkit) providing basic storage and retrieval facilities for RDF data.
The Inference Engine, which is the core of the PIKR Reasoner, is being implemented as a Java application, interfaced with XSB Prolog [11], a research-oriented Logic Programming system. XSB extends the conventional Prolog systems with an operational semantics based on tabling, i.e., a mechanism for storing intermediate results and avoiding to prove sub-goals more than once. The PIKR Reasoner is exposed as a Web service.

5 Conclusions and Future Works

In this paper we presented a semantics-based infrastructure, called PIKR, aiming at providing a unified view of different kinds of knowledge resources that are present in a virtual enterprise context, for supporting knowledge management and interoperability in production and innovation related activities. This infrastructure is designed,

[4] http://www.smwplus. net
[5] http://incubator.apache.org/jena/index.html

according to the Linked Data approach, by describing knowledge resources and their semantic relations in terms of a federation of reference ontologies, which define production processes, documents, actors and key performance indicators. While the actual knowledge resources are stored at the premises of the respective owner companies in the virtual enterprise, the PIKR maintains resource images in the form of semantic descriptors that can be regarded as instances of the ontologies. On top of this descriptions, a set of semantic services is offered for easing the navigation and the retrieval of such resources, along with a set of facilities for reasoning over them.

While in this paper we give an overview of the PIKR infrastructure, in the next future we will address the following issues: the building of the specific reference ontologies, the full implementation of the PIKR.

Acknowledgments. This work has been partly funded by the European Commission through the ICT Project BIVEE: Business Innovation and Virtual Enterprise Environment (No. FoF-ICT-2011.7.3-285746). The authors wish to thank all BIVEE project partners for their contribution during the development of various ideas and concepts presented in this paper.

References

1. Bizer C., Heath T., Berners-Lee T.: Linked Data - The Story So Far. *International Journal on Semantic Web and Information Systems*. 5(3):1-22, 2009.
2. Universal business language v2.1. Technical report, OASIS, 2011, http://docs.oasis-open.org/ubl/UBL-2.1.pdf.
3. RosettaNet Program Office: Cluster, segments, and pips overview. Technical report, 2011, http://www.rosettanet.org.
4. OMG. Business Process Model and Notation. http://www.omg.org/spec/BPMN/2.0, 2010.
5. Missikoff, M., Proietti, M., Smith, F.: Querying Semantically Enriched Business Processes. Proc. DEXA, LNCS 6861, pp. 294-302, Springer, 2011.
6. Hitzler, P., Krotzsch, M., Parsia, B., Patel-Schneider, P.F., and Rudolph, S.: OWL 2 Web Ontology Language. W3C Recommendation, http://www.w3.org/TR/owl2-primer/, October 2009.
7. Eschenbaecher, J., Glaser, F.: Managing and optimizing innovation processes in collaborative and value creating networks. *International Journal of Innovation and Technology Management*. 8(3):373–391, 2011.
8. Diamantini C., Potena D. Thinking Structurally Helps Business Intelligence Design. Proc. of the itAIS Conference. Naples (Italy), Oct. 8-9, 2010.
9. Formica, A., Missikoff, M., Pourabbas, E., Taglino, F.: Semantic Search for Enterprises Competencies Management. Proc. KEOD 2010.
10. Prud'hommeaux, E., and Seaborne, A.: SPARQL Query Language for RDF. W3C Recommendation, http://www.w3.org/TR/2007/WD-rdf-sparql-query-20070326/, Jan. 2008.
11. Tennant, G., SIX SIGMA: SPC and TQM in Manufacturing and Services. Gower Publishing, Ltd.. ISBN 0-566-08374-4. 2011.