An approach to Integrated Management System exploiting knowledge base to support business processes management

Roberto PAIANO, Andrea PANDURINO, Anna Lisa GUIDO

Department of Engineering for Innovation , University of Salento, Via per Monteroni, Lecce, 73100, Italy

roberto.paiano@unisalento.it, andrea.pandurino@unisalento.it, annalisa.guido@unisalento.it

Pierluigi RITROVATO

DIEM University of Salerno, via Giovanni Paolo II, 132, 84084 Fisciano (Sa)- Italy

pritrova@unisa.it

Ciro D'APICE, Giuseppe LARIA

Centre for Research in Pure and Applied Mathematics, c/o DIEM University of Salerno, via Giovanni Paolo II, 132, 84084 Fisciano (Sa)- Italy

dapice@crmpa.unisa.it, laria@crmpa.unisa.it

Abstract. Business organization are intrinsically complex because they have tackle constrains coming from both internal and external factors. Internal factors are related to production process that creates the value for the organization. External factors are related to the market contexts that also impose to respect several law and regulation impacting production processes. A company that wants to survive in their specific markets, needs of an integrated approach able to match the internal and external requirements and constrains. In this paper, we propose to extend traditional business process management systems with; a) a knowledge layer representing, through ontology, the application domain (internal factor) and the organizational environment (external factor); b) a customized version process management system based on BPMN (Business Process Management Notation) able to exploit, interact, and enrich the knowledge managed through the semantic technologies. The presented research is the main output of the HSEPGEST research project focused on the health, safety and environmental protection context.

Keywords: Integrated Management System, business process, knowledge base exploration, ontology

1 Introduction

Business organizations, whether public or private, operate in a context continuously changing and not just with respect to the competitive areas. These organizations, to survive in their specific markets (private sector), or to guarantee a competitive service level also with respect to ethical and social performance of public services (public sector), are required, among other things:

- to comply with a series of standards, constantly updated, which relate to occupational safety, environmental protection, quality management system (e.g. ISO, OHSAS, HACCP, EN, etc...)
- to verify the application of those standards while operating their daily activities.

The research presented in this paper arises from the observation that, to date, although there are numerous technological tools and methodologies for the practical application of regulations and procedures, they are poorly integrated for cross management and continuous monitoring of issues related to several sensitive areas (safety, health, environment, etc...). More in detail, the paper reports the results achieved in the frame of the HSEPGEST project funded by the Italian Ministry of Education, University, and Research (MIUR) in the context of the framework program 2007-2013. HSEPGEST addresses the challenge related to the Integrated Management Systems, following a synergic approach that takes into account different viewpoints (legal, technological, methodological, organizational) by introducing a set of tools and solutions that allow to align and to update the organizational information system with respect to a set of reference norms. In particular, this is possible by supporting the design and execution of organizational processes compliant with constraints introduced by the norms to be applied by the organizations in order to guarantee the required level of Quality, Environmental, and Safety management. This objective is achieved by improving the flexibility and adaptability of current business process management system through the integration of those systems with a knowledge management layer providing access to contextualized information according to the need of the design or of the execution of the business process.
The research is founded on two main pillars:

- the representation and management of the knowledge related to the application domain and to the organizational environment through the exploitation of semantic technologies. The knowledge model is able to provide a semantic representation of the Enterprise (including organizational structure, processes, information, resources) integrated with a representation of the regulatory context defined by the system of the norms implemented by the organizations.
- the design and development of a process management system based on BPMN notation able to exploit, interact, and enrich the knowledge managed through the semantic technologies. In detail, the presented BPMS uses the ontological model to support the design process providing the needed information. Moreover, In case of changes, (due to law changes, company reorganization, …), the associations

among chunks of knowledge, managed through the semantic technologies allow to redesign quickly the processes that need to be updated and modified guarantying the compliance with norms and regulations.

The results of the project have many innovative and distinctive features compared to current solutions of the market that are focused mainly on the design of the process. The presented solution is an integrated system that makes a powerful and effective use of a semantic layer to improve the quality of the process design.

2 Related work

The research about the integration of management systems started with the publication of Environmental Management System (EMS) in 1996 [1] where a set of guidelines of integrating the EMS and Quality management system was proposed. The Integrated Management System (IMS) was defined by Beckmerhagen [2] as "putting together different function-specific management systems into a single and more effective integrated management system".
Ten years after, in the 2006 Jorgensen et al. [3] proposed a integration process based on three levels: correspondence, coordination and integration. On the basis of this work, Jorgensen [4] explored the relation between the sustainable management systems and process life cycle management. Zeng [5] classifies the internal and the external factors that are involved in the development of IMS. In the research of these factors, Wilkinson and Dale [6] are focused on two main approaches. Their first approach proposes to achieve the full integration using the common elements and similarities existing in the different standards. The second approach implements the integrated system through a total quality management approach. In the 2004, Labodova [7] proposes two methods of integration: the first is based on the introduction of individual systems followed by their incremental integration; while, the second one uses the risk analysis as guide to makes the integration. In the scientific panorama, other authors as Karapetrovic [8] Asif et al. [9, 10] and Lopez-Fresno [11] proposed a variety of methodologies.
Moreover, many organizations promote several standards designed to support integration: see, for example SAI Global [12], BSI [13] and ISO that published a guide for integration advice, methodology and examples [18].
On the basis of the studied literature, there are also several results in terms of semantic modeling of norms and business processes (e.g. [14], [15], [16], [17]) but they do not address the problem of the IMS as HSEPGEST proposes by exploiting semantics for providing an integrated support for the main aspects that the designer of the integration has to take in care: integration methodology, implementation strategy, level of integration and audit integration.

3 THE HSEPGEST SOLUTION

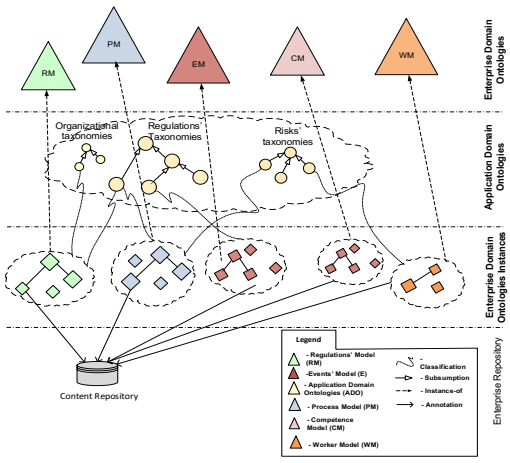
The HSEPGEST solution is based on:

- the use of semantic technologies for modeling and managing knowledge representing the domains of interests
- integration of workflow design and execution technologies with semantics

3.1 The modeling and management of HSEPGEST knowledge

The analysis of HSEPGEST application domain resulted in the identification of the following main entities to be modeled through ontologies to represent the HSEPGEST knowledge: processes, worker, regulations, working environments, risks, events, social networking, competences. In addition, the analysis of these entities allowed to identify a set of light weight ontologies (taxonomies) that specify and characterize the application domains. Example of these taxonomies are workers' roles, risks' types, regulations classification, etc. HSEPGEST ontologies allows to create filters defined to relate each other different entities of the HSEPGEST knowledge through a classification process based on taxonomies and specific constraints (e.g. introduced by the regulations). According to this view, the HSEPGEST knowledge model is based on two ontological levels:

- Enterprise Domain Ontologies (EDO): they model the entities (i.e. human resources, tasks, processes, objectives, etc.) that represent the organizational environment from a general perspective. It is worth mentioning the definition of two different models for competence (CM) and worker (WM), in fact competence is considered an asset of the enterprise that requires a specific model, to be referred by the worker model in the part that are relevant for workers. Because of their general purpose, these ontologies have been defined extending and combining ontologies (i.e. reference ontologies) already existing at the state of the art (e.g. FOAF[1], SCIOC[2], BPMN metamodel[23] oXPDL [17], etc.). This approach is widely applied in knowledge engineering and allows to preserve compliance with existing standards (de iure and/or de facto)

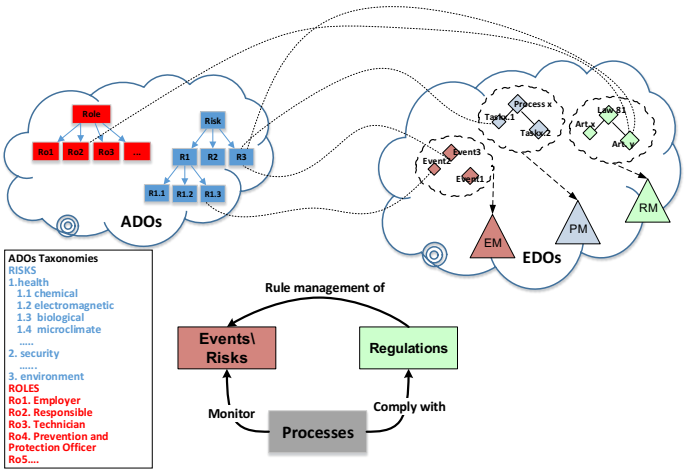


[1] http://www.foaf-project.org/

[2] http://www.w3.org/Submission/sioc-spec/

- Application Domain Ontologies (ADO): they model domain specific concepts and they support the classification of entities represented through the EDO. The regulations represent a typical example of ADO because they are application specific and they can be used to define normative taxonomies to classify the different entities representing the organization. Because of they are domain specific the ADO(s) are related to specific organizations.

The Figure above describes the roles and relations among the different entities involved in the structure of the HSEPGEST model. The bottom level of the structure shows the individuals, instance of the concepts part of the EDOs. These instances refer to/annotate organizational assets (e.g. human resources, documents, posts, facilities, etc.) that are stored in the information systems of the organization. A set of semantic services manage this structure by: i) Allowing the semi-automatic classification of the instances through the ADOs ii) Performing text analysis to extract concepts and support classification features iii) Providing search features through a link discovery approach. The availability of this infrastructure (i.e. services + knowledge model) can be exploited in several different scenarios related to the integrated management systems domain. In fact, a key issue, in this context, is the identification of relation among entities in order to make effective integrated management.



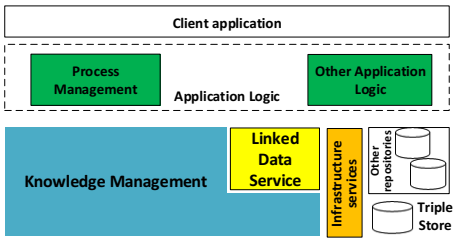
The following figure depicts a simple correlation scenario that, indeed, provides a powerful example about the potentiality of this infrastructure.
The figure highlights potential relations, defined by the EDO ontologies, among: processes, Events/Risks, and Regulation. The EDOs cloud show instances of these ontologies and the top level of the figure shows classification created by the semantic tools between EDO and ADO instances. The following situation could be imagined: i) process X monitors risks related to R3 risk type; ii) event 3 triggers a R3 risk type; iii) law 81 (art. Y) rules management of R3 risk type and requires involvement of Ro2.
In this situation the infrastructure can support in:

- Identifying processes to be executed to monitor risks arising from an unexpected event.

- Designing new processes complying with regulations ruling management of specific types of risk;

4 THE HSEPGEST SYSTEM ARCHITECTURE

The figure shows the HSEPGEST system high level architecture. It includes:

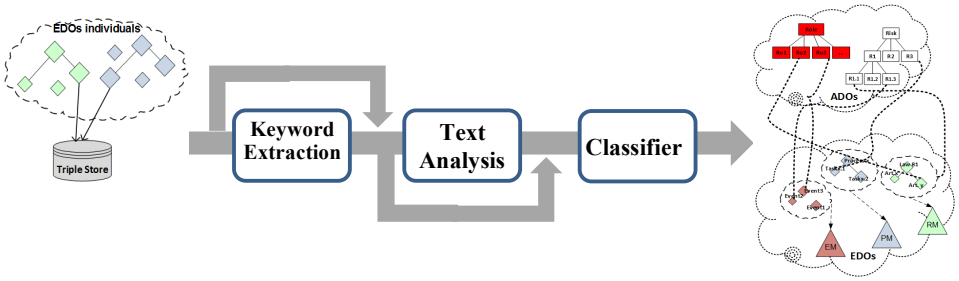


- The Knowledge management system: it is a middleware that provides access to knowledge through RESTfull and SOAP[3] based API[4]. The knowledge management system provides features for: concept extraction from text, text analysis, text classification based on concepts extraction and text analysis, search features based on the link discovery approach.
- Link Data Layer: provides an abstraction layer to access data stored in RDF[5] based repositories and aggregating data stored in traditional repositories.
- Infrastructure services: it includes basic services that could be used by different components of the architecture: connectors to access the available repositories; notification broker (NB) to support publish and subscribe based communication.
- Process Management System: it manages the design and execution of workflow exploiting features provided by the Knowledge Management System.

The following subsections describe more in detail the architecture and behavior of the Knowledge and Process Management systems.

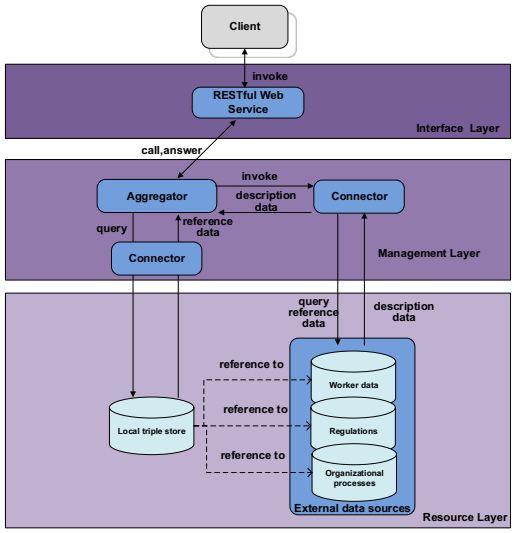
4.1 HSEPGEST – Knowledge Management system architecture

The knowledge management system implements a methodology based on the following elaboration pipeline:



[3] Simple Object Access Protocol
[4] Application Programming Interface
[5] Resource Description Framework

This pipeline takes as input the text provided as description of a new individual instantiated in the semantic repository (we use OWLIM accessed by mean of a SPARQL[6] endpoint). The first block (concept extraction) of the pipeline processes the text in order to extract a set of keywords to be used to perform the classification. The keyword extraction is performed applying two different approaches: i) An algorithm based on the multiple string patterns matching approach [20,21,22]; ii) The use of Wikipedia miner[7] for exploiting the rich semantics encoded in Wikipedia;
The results of the keyword extraction are used by the text analysis block for indexing the input text with respect to the concepts defined by the taxonomies of the ADOs. The indexing results, together with the extracted keywords, are used by the last block to classify the EDOs individual with respect to the ADOs taxonomy. For this purpose, ADOs taxonomy are defined using SKOS[8]. In particular, each concept of an ADO taxonomy is an instance of the `skos:Concept` class, and the `skos:preflabel` property indicates the concept label. In addition, each concept label is enriched using the `skos:hiddenlabel` property in order to add related words. The Classifier will search for `skos:preflabel` and `skos:hiddenlabel` matching the results of keyword extraction and text analysis. The matching `skos:concept` (s) are used to classify the EDO individual by instantiating a `HSEPGEST:hasCategory` property between the individual and the `skos:concept`, that will be enriched with additional `skos:hiddenlabel` properties modeling the keywords that were not included in the initial set. The ADO taxonomies together with the mechanism for creating relations between `SKOS:concept` and EDO individuals provide the foundation to link different resources belonging to the same class or to different classes of the EDO ontologies. The availability of these links allows to navigate the EDO individuals through the bridge generated by the ADO level of the HSEPGEST knowledge model.



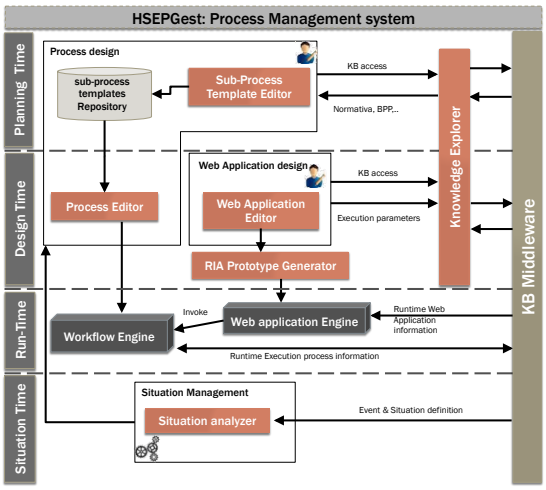
Furthermore, the navigation can be performed following a faceted browsing approach where the facets are represented by the classes of the EDO ontologies and the instances of the `SKOS:concept` part of the ADO taxonomies.
It is worth mentioning that the data layer has been designed taking into account the linked data model in order to reduce the impact of the integration of the semantics

[6] SPARQL Protocol and RDF Query Language
[7] http://wikipedia-miner.cms.waikato.ac.nz/
[8] http://www.w3.org/2004/02/skos/

with other data stored in existing repositories of the enterprise IT infrastructure. The figure depicts the architecture of the data layer. A central role is played by the aggregator service that acts as a broker between the local Triple store and external data source. The local triple store is the semantic repository and manages the ontologies (EDOs and ADOs) and their instances and, through the Connector hides all the details about the management of ontology languages (i.e. RDF[9] and OWL[10]) and related query languages (i.e. SPARQL). The semantic repository stores RDF triples representing individuals of the EDO ontologies and use "URI" to reference data stored in external source that represent actual value of the individual properties. Queries of external clients are resolved by the aggregator service that aggregates data from different data repositories according to the information stored in the Local Triple Store. From a technological viewpoint, the semantic repository is an OWLIM triple store with OpenRDF Sesame for its management. The connectors and the aggregator are Web Service developing REST and SOAP API. Furthermore the connector supports different response formats (e.g. XML, JSON, etc.).

4.2 HSEPGEST – PROCESS MANAGEMENT SYSTEM ARCHITECTURE

In despite of the standard BPMN suite, the business process design tool defined in the HSEPGEST project has as main goals not only to allow the user to design and to execute a business process but to improve the quality of the process design. The proposed architecture helps in all the phases of the management starting from the design of the business process until the development of the web application that support the execution of the business process.



The improvement of the process design is enabled by using the information stored in the knowledge base and through an increase of the process flexibility that could be rapidly modified. In the proposed tool during all the design process, the user can: i) interact with the knowledge base as described before ii) compose part of business process (called sub-process template) already defined in order to create a complete process able to satisfy the business need or to manage a specific company events such as sudden market changes. The process lifecycle proposed in HSEPGEST project is composed of four phases from the design to the busi-

[9] http://www.w3.org/RDF/
[10] http://www.w3.org/TR/owl2-overview/

ness process change management. The first two phases are related to the design process, the third phase provides the execution of the process, and the fourth phase allows to manage the process changes/upgrade. This four phases cover all the phases of the business process lifecycle from the understanding of the reference domain through the support of the knowledge base (Planning time) to the design of business process (Design time) and to the execution of the business process (Run time) (see figure). The architecture, also, manage particular unforeseen situations coming from changes in the application domain (Situation time) that may change the normal execution of the business process. In following we describe in detail each phase.

Planning Time: In the Planning Time phase, the process designer can select the concepts of the knowledge base that characterize the model and can model the sub-process template. The exploration of the knowledge base can be done using the Knowledge Explorer component available in the HSEPGEST design tool. The process designer is completely supported by the information of the knowledge base that is explored using the KB Middleware module. The Business Process Editor component allows to model the process using BPMN notation. Using this component the user can design the sub-process templates. Every designed elements of the process can be semantic characterized because the Business Process Editor allows to link the specific elements to the concepts/resources stored in the knowledge base. The resulting design of the business process is in BPMN notation and it is represented using a format suitable for the transmission to the knowledge base, and compatible with the characteristics required by the workflow engine. The interface of the BPMS and the knowledge base is defined through the KB middleware component in order to hide the complexity of the knowledge base making available API to access to the concept in the ontology.

Design Time: During the Design Time phase, the user models the process using the *Business Process Editor* component (the same of the Planning Time). During this design as happens for the Planning Time, the user can use the information stored in the knowledge base and can link (tag) the modelled element to specific concept of the KB. In this phase, the user has to model also the information than can assure the execution of the process including the execution parameters, the business logic and the User eXperience information. The execution parameters are the information elaborated by the task's business logic and they are provided by the knowledge base such as the number of employees or the name of the supervisor. Another type of information used in the design time is the User eXperience information. This type of information can assure the link between the process and the web application that execute the process. The User eXperience information allows the creation of the prototype of the process. The user can insert the business logic and the execution parameters through the *Business Process Editor* while the User eXperience need of a the *Web Application Editor* that use the IDM and Rich-IDM methodologies for the design of the prototype UX. The designers during the modelling of the UX can define integration points of information, which are dynamically retrieved from the knowledge base. The final model of UX, expressed in XMI format, is the main input for the module *RIA Proto-*

type Generator. The *RIA Prototype Generator* is responsible for the RIA prototype generation.

Run-Time: In the Run-Time phase, the process is executed. This task is completely assigned to the application server, which integrates the Workflow Engine and the Web Application Engine components. The Workflow Engine was based on the engine of the jBPM6.0 that has been extended with a more accurate monitoring module. The output of the process are sent to the knowledge base that allows to update the information in all the repository.

Situation time: The situation time allows the management of process changes. In detail, the user can decide to change the process in every moment but in this phase we refer to the changes due of the changes of the information stored in the knowledge base. In detail, the knowledge base using a specific service provided by the KB middleware component notifies to the BPMS the specific events. This notification is characterized by a sub-set of concepts/resources stored in the knowledge base, and is intended to describe the event in detail. The Situation time has the task, based on notifications from the KB Middleware, to identify and suggests to the designer some business processes compatible with the event generated. The Situation analyser that identifies and suggests to the process designer some business process compatible with the event generated. The user is involved when there are more business process compatible with the event or when there is not business process compatible.

There are many technological choices in order to implement the HSEPGEST architecture. Three main aspects must be taken in consideration: i) The workflow engine; ii) The prototype architecture; iii) The access to the knowledge base. For the workflow engine an open source product has been selected after a technological scouting and a specific comparative study. We select jBPM6 platform that can be customized in order to satisfy all the HSEPGEST project requirements. jBPM6 is a complete suite that is in continuous updating thanks to a community of developers. The extension feature is the most important aspect for the HSEPGEST project. In HSEPGEST project in order to execute and to design the business process, we use both the workflow engine and the process editor of jBPM6. The editor has been customized to the design of sub-process template and to reference the knowledge base concept. In detail, we extend the editor in order to support these requirements and we relax the BPMN notation constraints check. We add a specific interface in order to acquire information from the knowledge base to design the business process too. As it regards the prototype architecture, we develop an eclipse-based design environment that allows to model web application using IDM [19] and Rich-IDM [23,24] methodologies. In order to obtain the web application model, the ATL [25] framework is used while the ACCELEO[11] framework allows the generation of the prototyped web application. Finally, the access to the knowledge base is taken through the KB middleware. The middleware is responsible of the communication between the tools of the business process management and the tools of the knowledge management. The access to the knowledge base has been taken using the RESTFUL API useful to read information stored in the knowledge base.

[11] http://www.eclipse.org/acceleo/

5 CONCLUSION

This paper presented an approach to the Integrated Management Systems and presents an architecture used in the HSEPGEST project useful to design and execute business process exploiting the knowledge base where the domain knowledge is stored. The added value of the proposed architecture is the close connection between the knowledge base (that contains the domain information) and the process design that facilitates the business expert in managing the complexity resulting from the application domain. The architecture allows managing correctly the correlation between the knowledge base and the design of the process due to the presence of a Knowledge Base middleware that hides the complexity of Knowledge Base. Very interesting is the possibility to define the sub-process template (already compliant with norms ands regulations) and to compose them in order to obtain a new process starting from the previous design. This is possible exploiting the structure of the HSEPGEST knowledge representation model that arranges the knowledge of the organization in two layers, ADO and EDO. This characteristic makes more effective searching and reuse of pieces of knowledge (in particular related to processes and regulations) thanks to the creation of interlinks that supports cross-entities browsing. It is worth noticing that the semantic model reuses and extends existing models (e.g. SKOS) and approaches (i.e. Linked Open Data principles) providing relevant benefits from the interoperability view and compliance with Semantic Web. The overall architecture has been developed in the HSEPGEST project and used in order to design and execute some business processes in the Health, Safety, Environmental protection field.

6 REFERENCES

1. S. Puri, Integrating environmental quality with ISO 9000 and TQM Portland, OR: Productivity Press, 1996.
2. I. Beckmerhagen, H. Berg, S. Karapetrovic, W. Willborn, “Integration of management systems: focus on safety in the nuclear industry”, International Journal of Quality & Reliability Management Vol. 20, No. 2, 2003, pp. 210-228.
3. T. Jorgensen “Towards more sustainable management systems: Through lide cycle managment and integration” Journal of cleaner production, Vol. 16, 2008, pp. 1071-1080.
4. T. Jorgensen, M. Remmen “Integrated Management Systems-three different levels of integration,” Journal of Cleaner Production, Vol. 15 No. 8, 2006, pp. 713-722
5. S.X Zeng., J.J Shi, G.X Lou,“A synergetic model for implementing an integrated management system: an empirical study in China” Journal of Cleaner Production, Vol. 15 No. 18, 2007, pp. 1760-1767.
6. G. Wilkinson, B.G. Dale, “An examination of the ISO 9001:2000 standard and its influence on the integration of management systems,” Production Planning and Control, Vol.13, No. 3, 2002, pp. 284-297.
7. A.Labodova “Implementing integrated management systems using a risk analysis based approach,”Journal of Cleaner Production, Vol. 12, No. 6, 2004, pp. 571-580.
8. S. Karapetrovic, J. Jonker “Integration of standardized management systems: searching for a recipe and ingredients” Total Quality Management & Business Excellence , 2003.

9. M. Asif, E. Bruijn, O. Fisscher, C. Searcy, H. Steenhuis, “Process embedded design of integrated management systems.” International Journal of Quality & Reliability Management Vol. 26. No. 3, 2009, pp. 261-282.
10. M. Asif, E. Bruijn, O. Fisscher, C. Searcy. “Meta-management of integration of management systems”. The TQM Journal Vol.22, No. 6, 2010, pp. 570-582.
11. P. Lopez-Fresno. “Implementation of an integrated management system in an airline: a case study”. The TQM Journal Vol. 22, No, 6, 2010, pp. 629-647.
12. SAI Global. AS/NZS 4581 Management System Integration e Guidance to Business. government and community organisations, Sydney, Australia, 1999.
13. BSI. PAS 99 Specification of Common Management System Requirements as a Framework for Integration. British Standards Institution, London, UK, 2005.
14. Henry M. Kim & Mark S.Fox, “using enterprise reference models for automated ISO 9000 compliance evaluation” Proceedings of the 35th Hawaii International conference on System Sciences – 2002
15. Amjad Abou Assali, Dominique Lenne, Bruno Debray, "KoMIS: An Ontology-Based Knowledge Management System for Industrial Safety" - 18th International Conference on Database and Expert Systems Applications (DEXA 2007)
16. Iman Poernomo, Timur Umarov, "Normative Ontologies for Data-Centric Business Process Management" - Enterprise Distributed Object Computing Conference Workshops, 2008 12th.
17. Armin Haller, Walid Gaaloul, Mateusz Marmolowski, "Towards an XPDL Compliant Process Ontology", 2013 IEEE Ninth World Congress on Services 2008, pp. 83-86, doi:10.1109/SERVICES-1.2008.71
18. ISO, 2008a. The Integrated Use of Management System Standards. International Organisation for Standardization, Geneva, Switzerland, 2008.
19. Bolchini, D., Paolini, P., Interactive Dialogue Model. 2006. A Design Technique for Multi-Channel Applications. In IEEE Transactions on Multimedia, 8, 3. (2006) 529-541.
20. Aho, Alfred V. & Corasick, Margaret J. (1975). Efficient string matching: an aid to bibliographic search. Communications of the ACM, 18, 333-340.
21. Boyer, R. S. & Moore, J. S. (1977). A fast string searching algorithm. Communications of the ACM, 20, 762-772.
22. Horspool, N. (1980). Practical fast searching in strings. Software: Practice and Experience, 10, 501-506.
23. Pandurino A., Mainetti L., Paiano R., “Rich-IDM: Transforming the User Experience of Legacy Applications”. Proceedings of 13th IEEE International Symposium on Web Systems Evolution, WSE 2011, September, 2011, Williamsburg, VA, USA. pp. 81-86. ISBN: 978-1- 4577-0698-1
24. Pandurino A., Bolchini D., Mainetti L., Paiano R., “Rich-IDM: Extending IDM to Model Rich Internet Applications”. Proceedings of 12th International Conference on Information Integration and Web Based Applications & Services iiWAS2010, 8-10 November, 2010, Paris, France. Pp. 145-152. OCG ISBN 978-3-85403-272-4 (print) - ISBN 978-3-85403-272-4 (print) - ISBN 978-1-4503-0421-4 (electronic)
25. Eclipse Foundation (2013) ATL: Atlas Transformation Language, sito Web: http://www.eclipse.org/atl/. Eclipse Foundation (2013),.
26. Paiano R., Guido A., Pandurino A. DESIGNING COMPLEX WEB INFORMATION SYSTEMS: INTEGRATING EVOLUTIONARY PROCESS ENGINEERING Publisher: Information Science Reference (igi-global). Publication date April 2009, Pages 319 ISBN 978-1-60566-300-5 (cap. 4).