A Model for Integration and Interlinking of Idea Management Systems

Adam Westerski, Carlos A. Iglesias, and Fernando Tapia Rico

Universiadad Politecnica de Madrid,
Escuela Tecnica Superior de Ingenieros de Telecomunicacion,
Avenida Complutense 30, Ciudad Universitaria,
28040 Madrid, Spain
westerski@dit.upm.es,cif@gsi.dit.upm.es,
fernando.tapia.rico@alumnos.upm.es

Abstract. This paper introduces the use of Semantic Web technologies for the Idea Management Systems as a gap closer between heterogeneous software and achieving interoperability. We present a model that proposes how and what kind of rich metadata annotations to apply in the domain of Idea Management Systems. In addition, as a part of our model, we present a Generic Idea and Innovation Management Ontology (GI2MO). The described model is backed by a set of use cases followed by evaluations that prove how Semantic Web can work as tool to create new opportunities and leverage the contemporary Idea Management legacy systems into the next level.

Keywords: idea; management; ontology; metadata; knowledge management; product development; process improvement

1 Introduction

Amongst other, one of the important reasons to develop Semantic Web [8] and rich metadata is to bring order to the current Web and harness the ever growing informational chaos. There have been a number of visions to achieve this at full in the Internet wide scale. Some initiatives took a global approach (e.g. Cyc [15]), while others claimed that building the new Web is through starting in small domains and interconnecting those islands slowly approaching the desired state (e.g. Linking Open Data community project [4]). However, in addition to those movements, as Semantic Web research grew in popularity, the industrial sector started to experiment with applying the technologies in closed environments, not to solve the inconveniences of the global Web but their own local problems with information overflow. In the following paper, we present one of such domains (Idea Management) and report on the first results of our research to bring Semantic Web technologies to this environment.

Idea Management is a promising industry sector [16] which produces software for collecting and organizing input from people regarding proposals for innovation of products and services. Furthermore, the goal of Idea Management

Systems is to provide tools that will enable to asses the collected ideas and select the best ones for implementation. However, as many Idea Management vendors admit [20], the biggest standing problem in the domain in how to quickly and efficiently cope with the sudden peaks of information that are submitted to the system and should be reacted upon.

As a corner store of our research we propose that the solution to this problem, ironically, comes through increasing the volume of data even more. We claim that applying Semantic Web technologies and rich metadata annotations to the assets of Idea Management Systems is the first step to interconnect those systems with other corporate systems and utilities as well as the global Web. Therefore, newly created connections can be used to pull additional data inside the Idea Management System and use this knowledge to automatically asses individual ideas and provide better aggregation, filtering and idea selection facilities.

In the following paper we focus on presenting the results of the first step of our work: to construct an ontology for Idea Management Systems and develop a model to fully describe the domain in pair with other existing ontologies. To justify our claims, we present examples that show how this research could be applied in practice to bring measurable goals (see Sec. 2). Further, we discuss the research process undertaken to model the ontology and show the connections with other vocabularies (see Sec. 3). Next, we present the results of the evaluation process during which we experiment with a number of different Idea Management Systems to check the coverage of the ontology (see Sec. 4.1). Finally, we give pointers to related work (see Sec. 5) and conclude the paper with pointing out the paths for future research (see Sec. 6).

2 Use Case Study

We present two use cases that can give an image how Semantic Web annotations can aid Idea Management in practice and improve the current systems. Firstly, we show how very basic Semantic Web technologies deliver interoperability between different systems (see Sec. 2.1). Next, we present a more sophisticated example where usage of common ontologies leads to discovering new useful data(see Sec. 2.2). The primary goal for both examples is to expose the benefits of interlinking Idea Management Systems with other systems. However, the description provided is only an overview of the situation, for more details please refer to homepage of the GI2MO project [3].

2.1 Scenario 1: Extracting idea metrics via direct links to other systems

John is a working in a medium but rapidly growing enterprise. To wisely allocate the sudden influx of money his company invests in innovation. The enterprise has a large number of products and a huge client base, so John sets up Idea Management facilities that will help to gather the feedback from the clients.

Nevertheless, as John discovers, when new products are released, clients suddenly get very active and the feedback grows to incredible amounts nobody is capable to asses within reasonable time and effort costs. Furthermore, the metric generation capabilities embedded in the software are either insufficient or require a lot of effort to manually input business data for every idea to fully compare and judge client submissions. Therefore, John turns for help to emerging technologies and convinces his company management to invest in integration of systems with Semantic Web technologies. As time passes newly adapted technologies start to pay off. When a game changing product is released clients turn again to company website to submit their ideas with volume never seen before. However, this time John is prepared! The Idea Management platform is tightly interconnected with other development and corporate management systems that deliver a huge number of metrics and new capabilities, e.g.:

– Based on connections between Idea Management platform and project management system John can see which similar past ideas became successful and which failed in development. Therefore, he can asses the probability of success for new ideas.
– John discovers the true power of Semantic Web integration as he can see how past ideas have been causing problems during and after implementation. Although the Idea Management system has never been integrated with company bug tracking environment, it has been with the project management suit, which data in turn is semantically interlinked with ideas and via simple reasoning delivers desired metrics.
– John is not an engineer, neither in charge of the product production cycle - it is hard for him to judge accurately the production difficulties that might emerge, as well as time and cost implications. However, thanks to the integration with the PLM system, metrics for past similar ideas are automatically extracted and John can see how much time and effort it took to develop them.
– John prises the flexibility of the Semantic Web technologies. When, integrating the systems he did not have an idea what kind of metrics or connections between the systems he would need. However, as it became clear over time, the tight interlinking between ontologies for different systems and reasoning capabilities allow to quickly add new metrics without much effort.

2.2 Scenario 2: Discovering new ideas and assets through usage of common ontologies

The Idea Management Systems maintained by John turn out to work very good and provide a valuable supplement to the company innovation management process. However, John notices that he is missing a big amount of potentially good ideas that are submitted via other systems on the Web rather then his Idea Management facility (e.g. social portals, boards or blogs).

Fortunately, the Semantic Web technologies embedded in company facilities allow to discover and easily pull this new data inside the Idea Management

System. The connections between the assets in the Idea Management System and in other systems are discovered via user profiling described with common ontologies on all portals. In case of John's Idea Management System, the ontology favoured to describe users is FOAF [12]. Fortunately, it happens to be linked with a popular solution used to describe users across social spaces (SIOC ontology [11]). With little effort new idea mining system is deployed and starts to track connections via user profiles. In his growing happiness, while assessing his top contributors, John sees in an aggregated view that Mary also often publishes ideas using her blog. Thanks to using common technologies and ontologies that are interlinked, those ideas are pulled inside the Idea Management facility and can be assessed an analysed just like they would be posted normally via Idea Management System front-end.

3 Ontology Design and Implementation

3.1 Domain analysis and the research process description

To enable the presented use cases we have created an ontology to cover all the concepts described in the Idea Management Systems. As a preparation for that task we used a certain number of sources as a guide for modelling the data structure of this particular domain:

- analysis of publicly published data from operational Idea Management Systems (e.g. Dell IdeaStorm [2], myStarbucks [5] and othes)
- work with a sample commercial system (Atos PGI 2.0 [1])
- analysis of cases studies from the industry (presentations, publications, conference publications etc.)
- analysis of data based on a research done on definition of the Idea Life Cycle

As a result, we have defined a data model (see Sec. 3.2) that serves as a base for designing the ontology (see Sec. 3.3) and applying it to a number of heterogeneous systems (see Sec. 4).

Furthermore, our goal for the ontology was to make it available for others to apply to more then just the handful of systems that we experimented with. Therefore, one of the biggest challenges was to maintain the integrity with Semantic Web trends and standards yet keep the ontology simple and put impact on its usability and ease to appliance to encourage other developers. This resulted in a number of problems that can be generalized for every ontology design task but had to be resolved with our specific domain context in mind:

a. Modelling open data vs. closed data.
   A large number of data stored in Idea Management Systems is not published for users that generate ideas (internal metrics, assessments, internal reviews, business analysis etc.). Furthermore, often the main means of idea assessment are statistics that differer very much depending on Idea Management System implementation (e.g. number of posts by given user or complex business metrics). Such richness and diversity of information results in a situation,

were a big number of so-called Idea Management Systems are much more simple then the sophisticated ones. Therefore, a question rises whether the ontology should be extremely generic and simple or cover in detail the most sophisticated types of systems. Secondly, whether the ontology should be aimed for the sole goal of data publish and search (e.g. like SIOC ontology) or further data analysis and reasoning (e.g. for multimedia operations [19]).

b. Modelling for distributed publishing of Idea Management concepts vs. centralized model.
   In case of the distributed model (e.g. embedding RDFa on each page that represents different Idea Management concepts like Idea, Idea Comment etc.) the ontology grows double in size because of the necessity to implement inverse properties. Excluding such a possibility makes the ontology much easier to comprehend yet limits it's use (the evolution of SIOC ontology specification [9] is a very good example of problems that come with the distributed model and the needs to preserve data schema simplicity).
c. Usage of existing ontologies for modelling Idea Management concepts.
   Describing common concepts over many different systems on the entire Internet with same vocabularies brings many benefits and simplifies overall perspective of the Semantic Web. However, when narrowing down to a single domain, the necessity to comprehend all those vocabularies to model a single system becomes a problem for a potential developer. Therefore, a question that we had to face in our research was whether or not to model Idea Management Systems with the use of external ontologies and to what extent.

We addressed all of the above questions during the ontology design phase and applying the schema to operational Idea Management Systems (see Sec. 3.3).

3.2 Idea Management System Data Model

As an outcome of domain analysis and work with a number of different Idea Management Systems, we have listed all the data that is created or modified during each of the Idea Life Cycle phases. Based on that, we created a model for the Idea Management Systems that could be used as a reference to design the ontology. The basic concepts for each of the respectable phases are:

- Idea Generation data (idea title, summary, creation/modification dates, attachments, categorizations etc.)
- Idea Improvement data (comments, user ratings, idea versions etc.)
- Idea Selection data (internal reviews, metrics, analysis and assessments)
- Idea Implementation data (information related to development process of a product/service based on the selected idea)
- Idea Deployment metrics (most often business metrics such as Return Of Investment, total cost etc.)

Apart of the above there is a number of stable concepts that are present on each of the stages and can deliver useful information for idea selection and assessment, most interesting being:

– user data (Idea Management Systems can involve people working in a variety of roles that impact the innovation process in different ways[3])
– idea contest (a particular asset for Idea Management Systems - an themed event in time that initiates idea collection, e.g. "collecting ideas for the next product version release")
– idea status in the pipeline (can refer to the general Idea Life Cycle but also often has additional internal stages).

Figure 1 presents a simplified idea centric diagram of all the above concepts. For a more detailed model please refer to GI2MO project website [3].

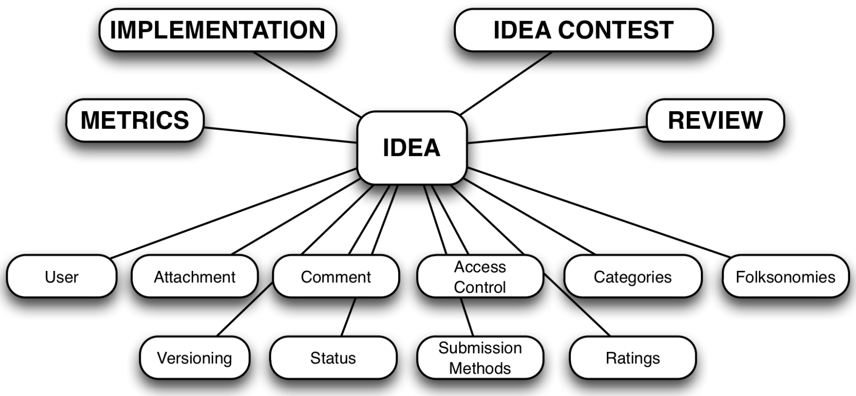

Fig. 1. Idea Management System - a simplified data model

3.3 Ontology Schema

Based on the presented earlier data schema we construed an ontology that aims to clip all the phases of idea management process together and allow to analyse the connections between (e.g. how idea input phase influences idea implementation etc.). Since we put most impact on connections between assets the natural choice was to base on research done in Semantic Web and support our work with achievements of technologies such as OWL, RDFS and research on other domain ontologies (e.g. GoodRelations [14]).

The most important concepts of Idea Management Systems that we wanted to interconnect and that have driven the design of the GI2MO ontology are:

– idea version control (track history of changes to e.g. see if ideas improved a lot provide better results in terms of different metrics such as revenue, cost etc.)

– idea pipeline modelling (building dependencies between phases and infrastructure o establish links)
– modelling dependencies with internal assets (other ideas) and external assets (ideas from other systems or other media resources)

The overview of classes included in the ontology is shown on Figure 2. For a detailed technical information with listing of all properties please refer to the ontology specification on GI2MO website [3].

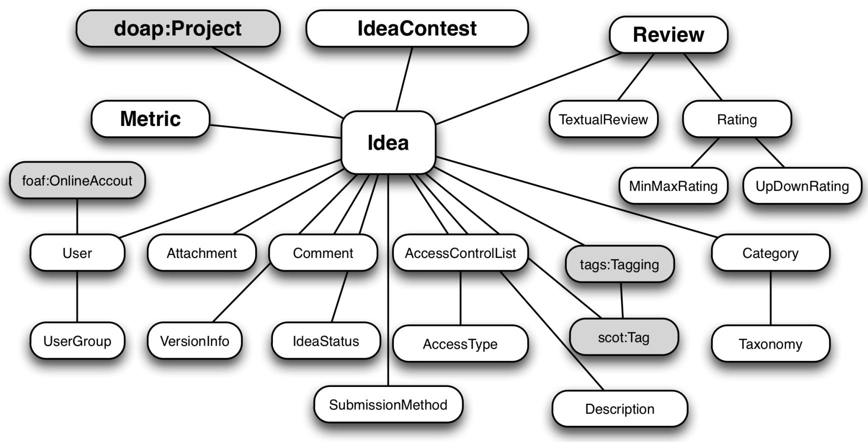

Fig. 2. GI2MO ontology class design overview

3.4 Connections with other ontologies

During the creation of the GI2MO ontology we decided to model a part of the Idea Management System using existing ontologies. To simplify the task we took a two step approach: first we modelled the entire data model with a single name space (GI2MO v0.1) and next we gradually started introducing other ontologies in certain areas (GI2MO v0.2). The final result can be seen in Table 1.

4 Ontology Evaluation

We have identified that the two most critical actions during applying an ontology to a domain specific system are: **data migration** and **connecting the new data layer with the application logic** to take advantage of the new capabilities. Therefore, we divided the evaluation activities into two separate phases respectable for those problems. Firstly, we wanted to test the coverage of

Table 1. GI2MO ontology imports

Ontology	Description of concepts modelled
Dcterms	Generic properties for many assets, e.g.: 'title', 'description' etc.
Doap	Idea implementation information
Foaf	Relation between User Account and personal data
Scot	Tags and tagging activities

the ontology on different data sets available and recognize any potential problems. Secondly, we aimed to evaluate the ontology in a development environment where data encoded with GI2MO vocabulary would be put into a particular use.

4.1 Ontology Coverage Study

For the first evaluation task we used the data available on-line from the Idea Management facilities open to public use. The datasets where as follows:

- Dell IdeaStorm - based on SalesForce Ideas platform
- myStarBucks Ideas - based on SalesForce Ideas platform
- Adobe Acrobat Idea - based on BrightIdea platform
- Cisco i-Prize - based on Spigit platform

Data coming from the two instances of SalesForce Ideas platform were used to see the differences that can occur within a deployment of the same system but profiled for different companies. Next, we perused to mine data from two other systems and see further the variations that occur. In order to obtain data from the Idea Management Systems, we developed custom HTML scraping tools, while to encode the information in RDF we used D2RQ tool with specific mappings to GI2MO for each test case. To make the experiment more reliable we asked university students, being independent to the ontology creation process, to perform the mappings without our supervision. Furthermore, we repeated the experiment three times for different version of the ontology:

- GI2MO v0.1 - entire Idea Management System data model covered by GI2MO
- GI2MO v0.2 - same data model as v0.1 but introduces imported name spaces of other ontologies to describe a number of concepts (see Sec. 3.4)
- GI2MO v0.3 - version constructed after the main evaluation tasks to adjust the ontology to the preliminary results and include new concepts

The quantitative results of RDFization with those three different iterations of the GI2MO ontology are presented in Table 2.

Analysing the above we can make two interesting observations. First, the amount of ideas does not always has a direct impact on how data size and complexity scales. Although IdeaStorm and myStarBucks are based on the same

[1] Triples where the predicate is expressed with GI2MO ontology

Table 2. Quantitative results in RDF triples for ontology coverage experiment

Portal Name	# of Ideas	# of Triples	# of GI2MO Triples[1]		
			v0.1	v0.2	v0.3
Dell IdeaStorm	9851	520330	427248	250869	-
myStarBucks	10949	194086	153040	89638	-
Adobe Ideas	579	17859	13292	7499	8798
Cisco i-Prize	826	133413	94262	69628	81950

system with almost identical capabilities, a smaller amount of ideas in IdeaStorm produced a much larger number of connections between assets then in myStar-Bucks. As we found out, in this particular case the reason was huge user activity in IdeaStorm in terms of idea reviews. Furthermore, as we experimented with systems from other vendors, we noticed that the amount of descriptive data and interconnections produced can rise into very high numbers just because of the amount of metrics published. In those terms Cisco i-Prize was the most rich, while the SalesForce systems had least of such data. This state translated in a great way into relation between amount of ideas and triples: respectably 161 triples/idea in i-Prize and 17 triples/idea in myStarbucks.

Table 3. Property mapping results for ontology coverage experiment

Portal Name	Ontology	# of Properties Covered		
		v0.1	v0.2	v0.3
Dell IdeaStorm	GI2MO	21	11	-
	imported	0	10	-
	uncovered	0	0	-
	total	21	21	-
myStarBucks	GI2MO	14	8	-
	imported	0	6	-
	uncovered	0	0	-
	total	14	14	-
Adobe Ideas	GI2MO	24	13	18
	imported	0	11	14
	uncovered	8	8	0
	total	32	32	32
Cisco i-Prize	GI2MO	43	29	36
	imported	0	14	24
	uncovered	17	17	0
	total	60	60	60

Secondly, apart of the quantitative analysis more important for us was the coverage of the ontology versus the data mined from different Idea Management facilities. As we noticed, for the two first front-ends of the same vendor (IdeaStorm and myStarBucks), the data structure was quite similar. Therefore, as a result we got 100% coverage for GI2MO ontology for every iteration of the experiment with quite similar mappings reused in both cases. However, the last system (Cisco i-Prize) proved to be quite different and served us as a valuable lesson on how Idea Management Systems can variate depending on the vendor. In case of this platform, the coverage was around 71 % mostly due to rich metadata assigned to user profiles that we did not take into consideration before. Such evaluation made us rethink some of the elements of the ontology that shall be published in the next iteration of GI2MO specification. The full results in terms of coverage of GI2MO properties are shown in Table 3.

4.2 Ontology Utilization Study

To test how the ontology would work in practice for an end user driven use case we turned for help to university students again and asked to implement a visualisation mechanism that would allow to categorize and view ideas from many heterogeneous sources. As a result, we got a web application (see Fig. 3) with the data back-end of the system entirely RDF driven and capable to work with the data mined earlier (see Sec. 4.1). During the experiment we did not observe any major issues related to comprehending the GI2MO ontology or the documentation delivered. However, the biggest standing problem that emerged was scalability - the application could not handle RDF dumps from the previous experiment at their full size.

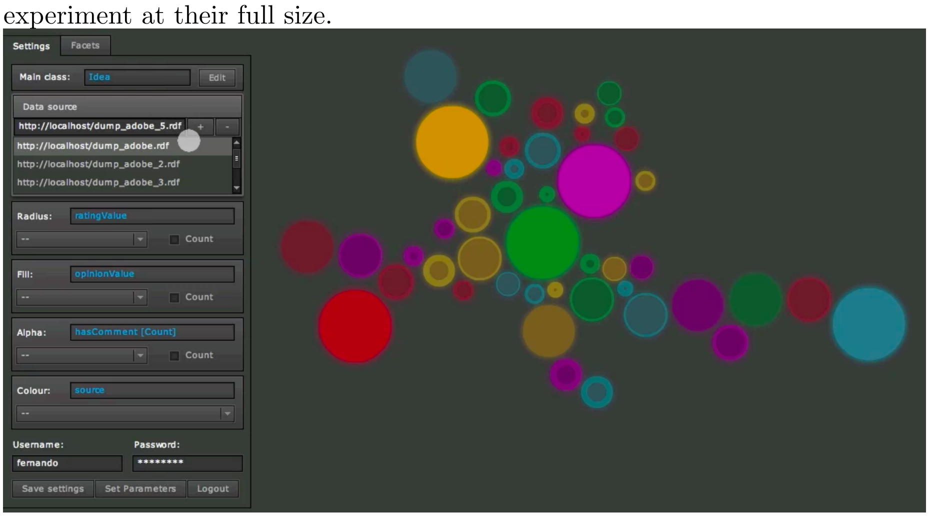

Fig. 3. Web application for faceted browsing of Ideas

5 Related Work

The topic of domain driven ontologies and their design has been investigated in numerous works for different areas and with different scope in mind e.g. BBC Music [17] Ontology, GoodRelations ontology[14] for e-commerce, SIOC ontology[11] and many others. In our work we tried to reuse the best practices from those attempts and employ them in our research. However, since Idea Management Systems are a rising technology there has not been much research done in terms of application of metadata and assets interlinking. To our knowledge Riedl et al. [18] are the only ones who present a similar attempt to ours. However, their Idea Ontology applies a different approach where less impact is put on interlinking and more on the sole goal of integration of idea repositories. As a result the two models are quite distinctive.

Furthermore, the Idea Management research has gone into a number of different directions trying to solve similar problems, e.g. data input user interfaces [7] or various knowledge management techniques often related to research on new collaborative methodologies that could aid idea ranking [10]. In relation to our primary goal of knowledge management there is a number of works refering to software supported innovation management. The most interesting, from our point of view, are by Adamides et al.[6] who tries to solve the data overflow problem by modelling innovation management as a problem solving case and Conn et al.[13] that shows the power of applying complex metrics for innovation assessment. Nevertheless both of those attempts do not refer to Semantic Web technologies in such extent as our research.

6 Conclusions and Future Work

In the paper we have shown how to model the data of Idea Management Systems using Semantic Web principles. By doing so we have described the research and difficulties that emerge in the process of designing a domain based ontology. However, we perceive the presented ontology only a first step to achieve our goals. The ontology lays foundations for knowledge management based on interlinking of enterprise systems and web assets to increase information awareness and help in innovation assessment. In terms of future work, we plan to experiment with interlinking Idea Management data with other specific systems and research on possibilities of automatic ranking and recommendation of ideas. Furthermore, the evaluation presented in this paper shows that experimenting with new systems can depict lacks of the ontology, therefore we shall continue its improvement to reflect Idea Management Systems data as best as possible.

Acknowledgements

This research has been partly funded by the Spanish Ministry of Industry, Tourism and Trade through the project RESULTA (TSI-020301-2009-31). We express our gratitude to Atos Origin R&D for their support and assistance as well as providing us access to their Idea Management Platform.

References

1. Atos pgi 2.0 idea management platform homepage, http://pgi2-en.atosorigin.es/
2. Dell ideastorm homepage, http://www.ideastorm.com/
3. Gi2mo project homepage, http://www.gi2mo.org/
4. Linking open data project website, http://esw.w3.org/SweoIG/TaskForces/CommunityProjects/LinkingOpenData
5. My starbucks idea homepage, http://mystarbucksidea.force.com/home/home.jsp
6. Adamides, E.D., Karacapilidis, N.: Information technology support for the knowledge and social processes of innovation management. Technovation 26(1), 50–59 (2006)
7. Baalsrud Hauge, J., Duin, H., Thoben, K.D.: Applying serious games for supporting idea generation in collborative innovation processes. In: Proceedings of the 14th International Conference on Concurrent Enterprising, ICE2008. Lisbon, Portugal (2008)
8. Berners-Lee, T.: Semantic web road map (1998)
9. Bojars, U., Breslin, J.G.: Sioc core ontology specification (2010), http://rdfs.org/sioc/spec/
10. Bothos, E., Apostolou, D., Mentzas, G.: Idem: A prediction market for idea management. designing e-business systems. markets, services, and networks. In: 7th Workshop on E-Business, WeB 2008. Paris, France (2008)
11. Breslin, J.G., Harth, A., Bojars, U., Decker, S.: Towards semantically-interlinked online. In: In Proceedings of the 2nd European Semantic Web Conference (ESWC '05). Heraklion, Greece (2005)
12. Brickley, D., Miller, L.: Foaf vocabulary specification. (2010), http://xmlns.com/foaf/spec/
13. Conn, S.J., Torkkeli, M.T., Bitran, I.: Assessing the management of innovation with software tools: an application of innovationenterprizer. International Journal of Technology Management 45(3/4), 323–336 (2009)
14. Hepp, M.: Goodrelations: An ontology for describing products and services offers on the web. In: Proceedings of the 16th International Conference on Knowledge Engineering and Knowledge Management (EKAW2008). vol. 5268, pp. 332–347. Springer LNCS, Acitrezza, Italy (2008)
15. Lenat, D.: Cyc: A large-scale investment in knowledge infrastructure. Communications of the ACM 38, 33–38 (1995)
16. Pettey, C., Stevens, H.: Gartner's 2009 hype cycle special report evaluates maturity of 1,650 technologies. Tech. rep. (2009)
17. Raimond, Y., Abdallah, S., Sandler, M., Giasson, F.: The music ontology. In: Proceedings of the International Conference on Music Information Retrieval (2007)
18. Riedl, C., May, N., Finzen, J., Stathel, S., Leidig, T., Kaufman, V., Belecheanu, R., Krcmar, H.: Managing service innovations with an idea ontology. In: Proceedings of XIX International RESER Conference, September 24th-26th 2009. Budapest, Hungary (2009)
19. Simou, N., Saathoff, C., Dasiopoulou, S., Spyrou, V., Voisine, N., Tzouvaras, V., Kompatsiaris, Y., Avrithis, Y., Staab, S.: An ontology infrastructure for multimedia reasoning. In: Proceedings of the International Workshop VLBV05. Sardinia, Italy (2005)
20. Turrell, M.: Collecting ideas from consumers & customers (2008), http://www.youtube.com/watch?v=QVZLrcOJR6Y