

Semantic Documentation in Project Management

Erick Casagrande Bastos, Monalessa Perini Barcellos, Ricardo de Almeida Falbo

Ontology and Conceptual Modeling Research Group (NEMO), Department of
Computer Science, Federal University of Espírito Santo– Vitória – ES – Brazil

{erickcasagrande, monalessa, falbo}@inf.ufes.br

***Abstract.** The existence of project management supporting tools did not eliminate the need of using desktop documents in this domain. However, retrieving information from documents is usually not trivial and depends on human effort. In this work we propose the use of semantic documentation in the project management domain as a way of getting useful information from data recorded in documents and spreadsheets. For this, we developed the Software Project Management Ontology and explored it to create features to support activities related to scope, time and cost management. The features were implemented in a specialization of the Infrastructure for Managing Semantic Documents, which was evaluated through an experimental study.*

1. Introduction

There are several tools to support project management, but they are not used by all organizations and often do not eliminate the need of using desktop documents [Villalobos *et al.* 2011]. Text documents and spreadsheets are frequently used in the context of projects, since they are instruments to record useful information to support communication among stakeholders and understanding about the project [Bruggemann *et al.* 2000]. However, one disadvantage of using documents is the difficulty of obtaining consolidated information from them, especially when information is spread in several documents. The access to document content typically depends on human intervention, since documents were originally created to be understood by humans and not by computers. Thus, retrieving and analyzing document content can be unproductive and sometimes inefficient. Besides, gathering relevant information from different documents can be so wearing that people may tend not to do [Arantes and Falbo 2010].

Semantic Documentation can be used to cope with this problem. Semantic Documentation consists in adding metadata into desktop document content so that it becomes available for computers interpretation. It allows relating annotated contents and use the relations to extract information from several documents, providing a general view that probably could not be gotten without the annotations [Arantes and Falbo 2010]. Ideally, metadata should be based on ontologies. Ontologies are ideal vehicles to describe metadata, since they represent a conceptualization and establish a common vocabulary to be shared [Sicilia 2006].

During a project, information regarding planning, execution, progress, monitoring and control is recorded in documents and spreadsheets (e.g., project plan and project monitoring and status reports). Talas *et al.* (2011) point out that, due to the importance of documents in a project, project members should share a common environment and be able to access the documents content in an easy and efficient way. Thus, if information is structured and annotated, computers can help deal with it.

Besides, annotating project documents helps store and retrieve knowledge acquired during a project and reuse it during other projects. Once documents are annotated, it is possible to build semantic repositories containing project documents and develop semantic documentation tools able to retrieve consolidated information from them.

Scope, time and cost are core areas in project management. The use of semantic documentation in this context can help extract and integrate data recorded in documents, providing an integrated view useful for project managers to make decisions.

The Infrastructure for Managing Semantic Documents (IMSD) [Arantes and Falbo 2010] is an environment able to manage desktop documents and store their content in a semantic repository. IMSD provides general semantic documentation features, such as document annotation and version control. Although these features are useful for managing documents produced in any domain, they are not enough to properly support project management activities.

Domain ontologies can be explored aiming at identification and implementation of features capable of supporting tasks in that domain [Falbo *et al.* 2014]. In this sense, explore concepts, relations and properties of a software project management ontology could provide a basis to semantic documentation features suitable for software project management. Thus, semantic documentation can allow retrieving data from several documents and accessing consolidated information regarding projects, helping project planning, execution, monitoring and control.

Considering that, the purpose of the work presented in this paper is to explore the use of semantic documentation to support software project management aspects, especially scope, time and costs management. We started by investigating the state-of-art through a systematic literature review that analyzed the use of semantic annotation in project management. After, aiming to establish the domain conceptualization to be used as basis to semantic annotations, we developed the Software Project Management Ontology (SPM_{Ont}), which addresses scope, time and costs planning and execution. Following, we analyzed IMSD and SPM_{Ont} to identify features to support project management aspects. Then, we extended IMSD to make it able to annotate spreadsheets and created the Infrastructure for Managing Semantic Documents for Software Project Management (IMSD-SPM), an IMSD specialization with specific features to support project management aspects. Finally, we conducted an experimental study to evaluate our proposal.

This paper is organized as follows: Section 2 talks briefly about semantic documentation and project management; Section 3 introduces the Software Project Management Ontology; Section 4 presents our proposal; Section 5 addresses the experimental study carried out to evaluate the proposal; Section 6 concerns related works; and Section 7 presents our final considerations.

2. Semantic Documentation and Project Management

Semantic Annotation was created to deal with limitations of extracting information from web pages content. It consists in adding metadata to the pages content to allow computers to retrieve data from them. This gave rise to *Web Semantic*, which uses ontology-based metadata to bring structure to the meaningful content of web pages, minimizing the lack of semantics and allowing both, humans and computers, interpret and retrieve data from web pages [Berners-Lee *et al.* 2001].

Problems related to extract information from text content also occur in the context of desktop documents, such as text documents and spreadsheets. Thus, Web Semantic principles can be applied to manage desktop documents content, given origin to *Semantic Documentation*, whose purpose is make desktop documents content interpretable for computers [Arantes and Falbo 2010]. Semantic documents can provide services such as advanced search, reasoning using document metadata, and knowledge management services, like document repositories and document management.

In the literature there are some tools to support Semantic Documentation. In the context of this work the Infrastructure for Managing Semantic Documents (IMSD) [Arantes and Falbo 2010] is particularly important. It provides: (i) a way to semantically annotate document templates; (ii) a mechanism for controlling versions of semantic content extracted from semantic document versions (and, therefore, a way for tracking the evolution of the data embedded inside a semantic document); and (iii) data visibility to end-users allowing for searches and data change notification subscription to aid developers to get up-to-date information about something they are interested in.

Problems related to access and manage document content clearly occur in the project management context, since text documents and spreadsheets are frequently used as instruments for recording and sharing information among project members. In this sense, semantic documentation has potential use in this area.

Project management involves the application of knowledge, skills, tools and techniques to project activities aiming to meet project requirements [PMI 2013]. According to the PMBOK [PMI 2013], there are ten knowledge areas (KAs) to be managed in a project, namely: Integration, Scope, Stakeholder, Human Resource, Time, Cost, Risk, Quality, Communication, and Procurement.

Project management comprehends three main interactive phases [Pressman 2011]: planning, execution, and monitoring and control. During project planning it is established a plan to the project, including the project scope, team, schedule, and budget, among others. Execution consists of executing the project following the established plan. Monitoring and control aims to compare the plans with the execution, identify problems and present solutions. During this phase, performance indicators can help the project manager. Earned Value Analysis [Fleming and Koppelman 1999] and its indicators Schedule Progress Index (SPI) and Cost Progress Index (CPI) can be used to provide information about project progress and performance, as well as to calculate cost and time estimations for project conclusion.

Aiming at investigating the use of semantic annotation in the project management domain, we carried out a systematic literature review to analyze initiatives involving semantic annotation to support project management. Six digital libraries were searched, namely: Scopus, Engineering Village, ACM, IEEE Xplore, Springer, and ScienceDirect. Five initiatives were found: Inter-Organization Projects (CMIO) [Nakatsuka and Ishida 2006]; Collaboration in Public Policy Making, Implementation and Evaluation (CPPMIE) [Loukis 2007]; Semantic Annotation based on Software Knowledge Sharing Space (SKSS) [Lu et al. 2008]; Semex [Talaš et al. 2011] and Use of Semantic Wiki as a Capturing Tool for Lessons Learned (SMW) [Elkaffas and Wagih 2013]. Detailed information about the systematic literature review and its results can be found in [Bastos *et al.* 2015]. Next a brief summary of the results is presented.

By analyzing the selected publications, we noticed that only two proposals were conceived aiming to support project management. Although the other proposals support aspects related to project management, this is not their main concern. Four proposals use domain ontologies as a basis for annotating documents or web pages and all proposals adopt manual annotation. Spreadsheets are not annotated in any proposal. As for the project management aspects addressed, the proposals support some ones related to Scope, Integration, Communication and Stakeholder Management. Although six digital libraries have been searched, only five initiatives were identified and only two of them are truly devoted to the project management domain. This fact shows that the research topic is recent and has not been much explored. Since documents are still an important instrument to record and share information regarding projects, we believe that the use of semantic annotation in project management is a relevant topic and there are opportunities of research in this area. Thus, we propose to use semantic documentation as a way to support scope, time and cost management aspects in software projects. Since metadata to annotate documents should be based on domain ontologies, we developed the Software Project Management Ontology, which is presented in the next section.

3. The Software Project Management Ontology (SPMOnt)

SPMOnt was developed from the Software Process Ontology Pattern Language (SP-OPL) [Falbo *et al.* 2013] and addresses scope, time and costs planning and execution. Regarding costs, currently, only costs associated with human resources are considered. Figure 1 shows the SPMOnt conceptual model, represented by using OntoUML, a UML profile that enables modelers to make finer-grained modeling distinctions between different types of classes and relations according to ontological distinctions put forth by the Unified Foundational Ontology [Guizzardi 2005]. Axioms formalize constraints not captured by the model. Due to space limitations, they are not presented in this paper.

A *Software Organization* can be involved in *Projects*. In this case, the organization is committed to perform processes and activities defined to the projects. There are two types of processes defined to a Project: *General Project Process* and *Specific Project Process*. The first one is the global process defined to the Project. It is composed of specific process, allowing defining *sub-processes*. Specific Project Processes are composed of *Project Activities*, which can be *Simple Project Activities* or *Composite Project Activities*. Once a general project process is defined to a project, it is possible to plan duration, start and end dates, and cost of the process, their sub-processes and activities. The definition of duration, dates and cost to a Project Process gives rise, respectively, to *Process with Planned Duration*, *Scheduled Process* and *Process with Planned Cost*. Similarly, the planning of duration, dates and cost of a Project Activity gives rise to *Activity with Planned Duration*, *Scheduled Activity* and *Activity with Planned Cost*.

A *Human Resource Allocation* is the assignment of a Scheduled Activity to a *Human Resource* to perform a *Human Role*. The cost of a Human Resource Allocation is based on the cost of the allocated Human Resource, which is established in the *Employment* of that Human Resource.

A Project Activity can cause *Activity Occurrences*, which can be *Simple Activity Occurrences* or *Composite Activity Occurrences*. *Human Resource Participation* refers to the participation of a Human Resource in an Activity Occurrence.

A Project is performed to produce *Deliverables* that can be *Practical Results*, which are composed by other Deliverables, or *Work Packages*.

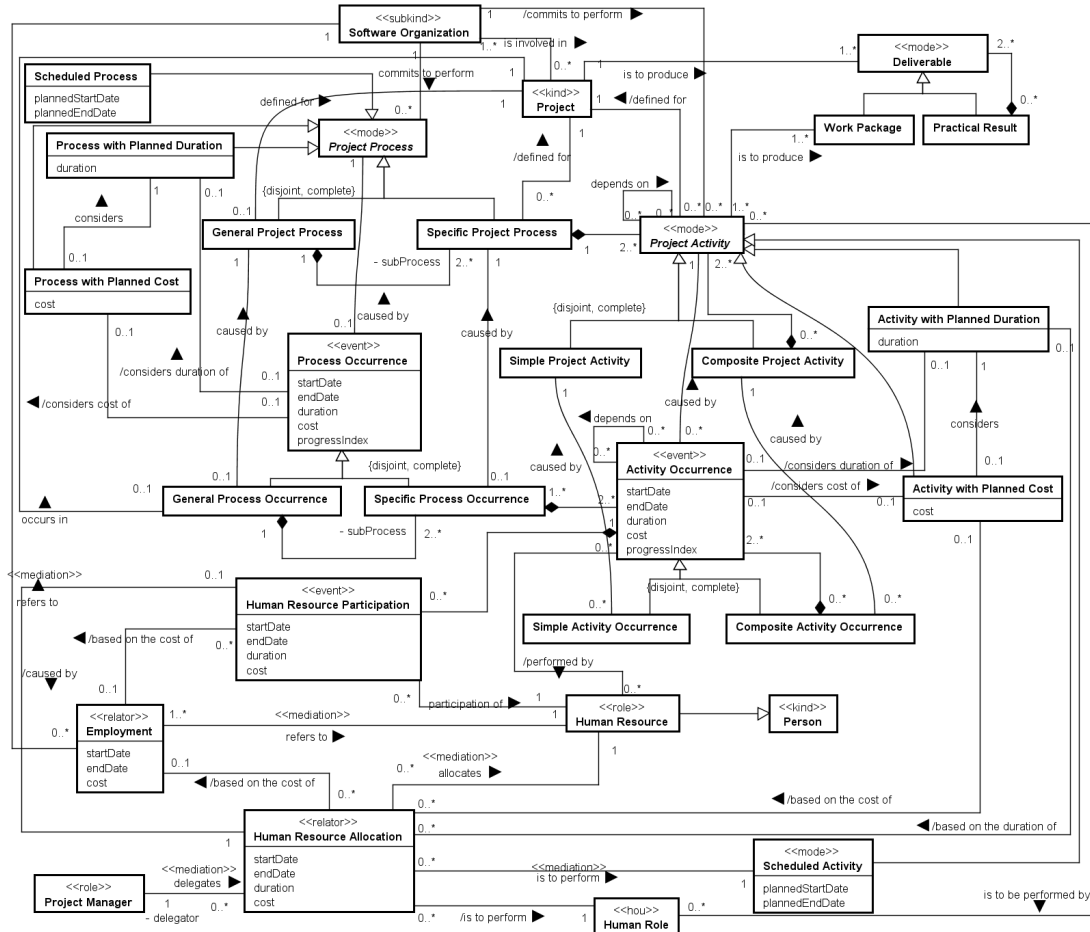


Figure 1 – The Software Project Management Ontology conceptual model.

4. Infrastructure for Managing Semantic Documents for Software Project Management (IMSD-SPM)

The Infrastructure for Managing Semantic Documents (IMSD) [Arantes and Falbo 2010] provides general features to support semantic documentation and it was developed in the same research group in which this work was carried out. Thus, we decided to extend it to provide features to support project management aspects. We analyzed IMSD considering two different perspectives: general, which concerns features useful for project management, but also relevant to support semantic documentation in any domain; and specific, which regards features devoted to support time, scope and cost management aspects.

Regarding the general perspective, we noticed that IMSD supported semantic documentation only in documents in text format. Since spreadsheets are very useful for recording data regarding projects (e.g., schedules and budgets), we identified as improvement opportunity (IO1) to extend IMSD to work with spreadsheets, expanding the scope of types of files that can be used as data sources.

Concerning the specific perspective, we considered the SPMOnt conceptualization to identify improvement opportunities to support project time, scope and cost management activities. The following improvement opportunities were identified:

(IO2) Planning project data can be recorded in different documents. Provide a consolidated view can help managers to have a global view about the project planning.

(IO3) It is important to know the dependencies between project activities and between them and project scope items. Thus, provide dependency matrices to represent these dependencies can ease visualizing dependencies (specially indirect and transitive dependencies) and analyzing impact of changes in the project.

(IO4) Project monitoring and control depend on information about progress and comparison between planned and performed values. Performance indicators can be useful in this context, because they quantitatively represent the project performance and also can be used to calculate estimates for the project conclusion. Thus, provide performance indicators can help managers to understand the project performance and identify situations that need corrective actions. Besides, estimations for project conclusion provide a view about the project performance implications and allow the manager to make decisions about project continuation or interruption.

(IO5) During a project, the amount of information to be manage increases. The exhibition of panels and graphics containing consolidated information about the project, including the project performance over time, provides an overview about the project and helps its management.

(IO6) Several projects are developed in a software organization. It is important to analyze them together in order to verify discrepancies and investigate their causes. Thus, present performance indicators of several projects provides a global view of the projects and allows comparisons among them. Besides, the performance indicators can be used to analyze the behavior of the project management process used.

(IO7) Identification of non-conformities in documents produced during a project can be eased by quality assurance features. Checklists containing domain-related criteria can be used to automatically verify documents quality.

The identified improvement opportunities were implemented by extending and specializing IMSD, as discussed in the next two sections.

4.1 Supporting Project Management with Semantic Annotations in Spreadsheets

In order to address IO1, we extended IMSD to work with spreadsheets and, then, we used SPMOnt as a basis to annotate spreadsheet and document templates related to the project management domain. The annotations are added into the templates resulting in *semantic templates* (or semantic models) that, when instantiated (i.e., when used to create documents and spreadsheets), give rise to *semantic documents*. Once annotated the templates, the documents produced using them are also annotated and can be used as data sources to IMSD. Spreadsheet templates were developed using the Open Document Format [Oasis 2015], since it is an open format, with great span. Annotations for cells were produced using Open Document Spreadsheet (ODS) in LibreOffice Calc.

For spreadsheets annotation, the syntax and instructions for annotating text defined in IMSD are used to capture the cell content. Instructions can be used to create instances, relations and properties based on the ontology. The syntax of the instance creation instruction is *instance (arg ,concept, accessVariable)*. This instruction creates the instance *arg* of the *concept* of SPMOnt. The SPMOnt was implemented in OWL and its URL is also informed in the *concept* field. The instruction result is a reference to the created instance and it is set on the *accessVariable* for later use. The syntax to create a relation is *property (arg1, prop, arg2)*. This instruction establishes a relation *prop* between the instances *arg1* and *arg2*. This instruction is also used to create properties and, in this case, it means that the value *arg2* is set as the property *prop* of the instance *arg1*.

For annotating spreadsheets templates, in the LibreOffice Calc, *Custom Properties* option is used to annotations recording and *Styles and Formatting* option is used to apply annotations to cells. The first thing to do when creating a semantic template is to create a custom property named *Semantic Document* and set its value to *True*. This way, IMSD can identify that the spreadsheet is a semantic document and searches for semantic annotations. Each annotation must be recorded in a new custom property whose value is the annotation instruction. For each annotation, a formatting style must be created and related to the custom property in which the annotation is recorded. Thus, when a formatting style is applied to a cell, the cell is annotated according to the annotation instruction recorded in the corresponding custom property.

Three templates related to project management were created and annotated: *Work Breakdown Structure* (WBS), a text document that presents the project WBS, describing the project deliverables; *Project Status Report* (PSR), a spreadsheet that contains information regarding planned and performed schedule and budget; and *Human Resources Costs* (HRC), a spreadsheet that contains information about costs of human resources. Semantic documents managed by IMSD must be created using the semantic templates. This allows partial automation of annotation and reduces the amount of work that users need to do to create semantic documents.

Figure 2 illustrates the template of the *Project Status Report* spreadsheet, which contains information about project activities, dependencies, human resources allocated and participants, deliverables, and planned and actual dates and duration. As examples, the annotations related to cells of Human Resource and Duration columns are shown. The first part of the human resource annotation creates instances of the *Human Resource* concept and stores in *hr* variable. The second part establishes the relationship *allocates* between instances of Human Resource and an instance of Activity, like in SPMOnt, in which the relation *allocates* connects a human resource to an activity, meaning that the human resource is allocated to perform the activity. The *break* tag means that one or various human resources can be related to one activity and they are separated by comma. In duration annotation, the tag *completeText* indicates that the instruction refers to the complete text stored in the cell. The instruction means that the cell content will be set as the property *Planned Duration* of an instance of Activity.

When the project manager uses the semantic templates to create a WBS, PSR or HRC, the created documents or spreadsheets become semantic documents. Then, they are submitted to IMSD, which extracts data from them and stores in OWL files, allowing searching and retrieval, and also performs version control. It is important to notice that, although we created semantic templates related to the project management

domain, spreadsheets annotation is a general feature, since by using it one can create semantic templates related to any domain.

```
[[completeText]];property($activity,
http://localhost/ontologies/SE/spmont
.owl#PlannedDuration,{content});
```

Project Status Report

Project: <<project name>> Date: <<Status Report date>> Responsible: <<person who created/updated the report>>

Id	Activity	Sub-activity	Planned						Executed					
			Start date	End date	Duration	Role	Human Resource	Activities from which it depends	Related Work Package	Start Date	End Date	Duration	Human Resource	% Done
1														
1.1														

```
[[break with ',' into 'var']];
instance({slice},http://localhost/ontologies/SE/
spmont.owl#HumanResource, $shr)
property($activity,http://localhost/ontologies/S
E/spmont.owl#Allocates,$shr)
```

Figure 2 – Project Status Report template (a semantic template).

4.2 Supporting Software Project Scope, Time and Costs Management

In order to address IO2 to IO7, we explored SPMont conceptualization in domain-specific features that were implemented to support scope, time and costs management activities. The features were implemented in an IMSD specialization named *Infrastructure for Managing Semantic Documents for Software Project Management* (IMSD-SPM). When a project starts, the manager creates the semantic documents related to project planning and submit them to IMSD-SPM. During the project, the manager updates the semantic documents with data related to project execution (or re-planning) and submit the updated documents to IMSD-SPM. Table 1 presents the implemented features and the IOs addressed by them. In each feature, some SPMont concepts, relations and properties were explored. After Table 1, some features are introduced and the use of SPMont in their context is discussed.

Table 1 – IMSD-SPM Features and Improvement Opportunities Addressed

Feature	Short Description	Addressed IO
Simplified Project Plan	Presents consolidated information about project scope, time and costs planning.	IO2
Dependency Matrices	Presents dependency matrices containing dependencies between activities and between activities and project deliverables.	IO3
Project Monitoring Panel	Presents consolidated information about project execution, pointing out the differences between planned and performed values.	IO4
Performance Indicators	Presents the current performance indicators of a project.	IO4
Estimates for Project Conclusion	Presents the optimistic, realistic and pessimistic estimates for project conclusion considering the current performance indicators.	IO4
Project Performance History	Presents project performance indicators and estimations over time.	IO5
Projects Comparison	Presents performance indicators and estimations for conclusion related to several projects.	IO6
Quality Assurance	Automatically evaluates documents and spreadsheets content quality considering a set of domain-related criteria.	IO7

4.2.1 The Simplified Project Plan and the Project Monitoring Panel

The *Simplified Project Plan* integrates data recorded in different semantic documents. Its first version is generated when the first versions of the WBS, PSR and HRC are submitted to IMSD-SPM. New versions can be generated when data regarding planning are updated and new versions of the documents are submitted. The Simplified Project Plan presents information extracted from semantic documents and also data generated by the IMSD-SPM. For instance, activities planned start and end dates are determined

by the IMSD-SPM considering activities planned duration, dependencies and the start date of the first activity. Also, IMSD-SPM calculates activities planned cost based on their planned durations and allocated human resources allocated. For this, the conceptualization provided by SPMont is explored. For example, the relation between the planned costs of activities and costs of the human resources allocated to them comes from the relation *based on the cost of* between *Activity with Planned Cost* and *Human Resource Allocation*, which determines that the planned cost of an activity is based on the costs of the human resources allocated to it, and from the relation *based on the cost of* between *Human Resource Allocation* and *Employment*, which determines that the cost of a human resource allocation is based on the cost of the allocated human resource. Besides, the composition relations between project activities and processes, and axioms that address planned costs in these relations (the planned cost of an activity is the sum of the planned costs of the activities that compose it, and the planned cost of a process is the sum of the planned costs of its activities) allow inferring that the total planned cost of the general process defined to the project is the sum of planned costs of all activities that compose the process. Finally, considering that, currently, only costs related to human resources are considered, it is possible to conclude that the project planned cost is equal to the planned cost of the general process defined to it. Figure 3 shows fragments of WBS, PSR and HRC semantic documents submitted to IMSD-SPM. Ellipses and arrows are used to highlight examples of related information.

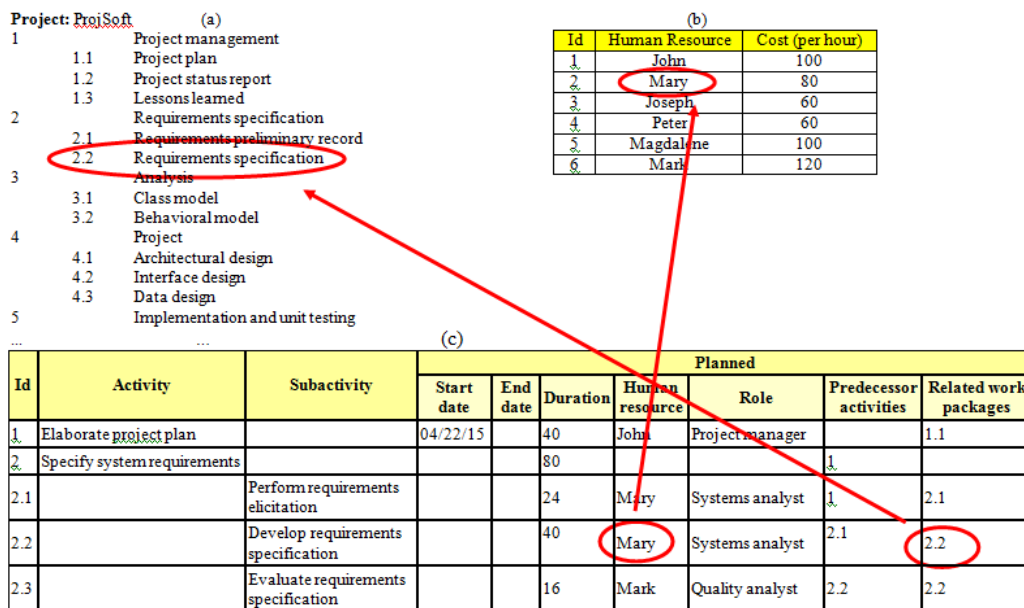


Figure 3 – Semantic documents fragments: (a) WBS, (b) HRC, (c) PSR.

A fragment of the Simplified Project Plan generated based on the documents shown in Figure 3 is depicted in Figure 4. Data inside solid ellipses are examples of data extracted from semantic documents. Data inside dotted ellipses are examples of data calculated by IMSD-SPM by exploring the SPMont conceptualization. To ease the plan visualization, a Gantt Graph is also provided (not shown in Figure 4).

Project		Planned							
Activity / Subactivity	Start date	End date	Duration	Human Resource	Role	Predecessor activities	WBS items related	Cost	
1 Elaborate project plan	04/22/15	04/28/15	40	John	Project manager	-	1.1	4000,00	
2 Specify system requirements	04/29/15	05/12/15	80	-	-	1	-	7040,00	
2.1 Perform requirements elicitation	04/29/15	05/01/15	24	Mary	Systems analyst	1	2.1	1920,00	
2.2 Develop requirements specification	05/04/15	05/08/15	40	Mary	Systems analyst	1, 2.1	2.2	3200,00	
2.3 Evaluate requirements specification	05/11/15	05/12/15	16	Mark	Quality analyst	1, 2.1, 2.2	2.2	1920,00	

Figure 4 – Simplified Project Plan fragment.

During the project, the manager submits new versions of the semantic documents containing data regarding project execution. Based on their content and considering SPMont conceptualization, IMSD-SPM provides a *Project Monitoring Panel* presenting planned and performed values and the difference between them. A Gantt Graph showing planned and performed values is also exhibited. Figure 5 presents a Project Monitoring Panel fragment. For sake of legibility some columns were omitted. The variations between planned and performed values are presented in the last columns.

Project		Planned				Performed				Variations				
Activity / Subactivity	Start date	End date	Duration	Human Resource	Cost	Start date	End date	Duration	Allocated HR	Cost	Cost	Start date (days)	End date (d)	Duration (d)
1 Elaborate project plan	04/22/15	04/28/15	40	John	4000,00	04/22/15	04/28/15	40	John	4000,00	0,00	0	0	0
2 Specify system requirements	04/29/15	05/12/15	80	-	7040,00	04/29/15	05/18/15	112	-	9920,00	2880,00	0	4	4
2.1 Perform requirements elicitation	04/29/15	05/01/15	24	Mary	1920,00	04/29/15	05/04/15	32	Mary	2560,00	640,00	0	1	1
2.2 Develop requirements specification	05/04/15	05/08/15	40	Mary	3200,00	05/05/15	05/13/15	56	Mary	4480,00	1280,00	1	3	2
2.3 Evaluate requirements specification	05/11/15	05/12/15	16	Mark	1920,00	05/14/15	05/18/15	24	Mark	2880,00	960,00	3	4	1

Figure 5 – Project Monitoring Panel fragment.

For calculating deviations between planned and performed values, the traceability between planned processes and its executions and between planned activities and its executions was explored in SPMont. The traceability is established by the relationships *caused by* between *Project Process* and *Process Occurrence*, and between *Project Activity* and *Activity Occurrence*. These relationships allow identifying start and end planned dates for processes and activities (*plannedStartDate* and *plannedEndDate* of *Scheduled Process* and *Scheduled Activity*); their planned duration and cost (*duration* and *cost* of *Process with Planned Duration*, *Process with Planned Cost*, *Activity with Planned Duration* and *Activity with Planned Cost*); and their actual duration and cost (*duration* and *cost* of *Process Occurrence* and *Activity Occurrence*).

Since activities planned and actual costs are respectively related to the activities human resources allocations and human resources participations, the *caused by* relationship between *Human Resource Participation* and *Human Resource Allocation* was also explored. *Caused by* allows relating a participation to the allocation that cause it and, then, comparing planned values (*duration* and *cost properties* of *Human Resource Allocation*) with actual values (*duration* and *cost* of *Human Resource Participation*).

4.2.2 Dependency Matrices

Dependency relations between activities and between them and scope items are important in the project management context and can be represented by means of dependency matrices. However, a lot of human effort is usually required to create and maintain these matrices.

In the semantic documentation context, relations can be established in semantic templates by using semantic annotations. Thus, data recorded in the semantic

documents are captured by IMSD-SPM and the changes made in the dependencies along the project are propagated when updated semantic documents are submitted to IMSD-SPM. Since the semantic global model of the semantic repository is generated by IMSD-SPM, the dependencies established in each semantic document, as well the ones between different documents, can be known.

Dependency relations established in semantic templates and in its instances (semantic documents) come from the relation *depends on* between *Project Activities* and the relation *is to produce* between *Project Activity* and *Work Package*. Figure 6 presents a fragment of a dependency matrix indicating dependencies between WBS items (first column) and project activities.

Dependencies between activities and project deliverables

WBS	1 - Elaborate project plan	2 - Specify system requirements	2.1-Perform requirements elicitation	2.2-Develop requirements specification	3-Evaluate requirements specification
1 - Project management	x				
1.1 - Project plan	x				
1.2 - Project status report					
1.3 - Lessons learned					
2 - Requirements specification			x	x	x
2.1 - Requirements preliminary record			x		
2.2 - Requirements specification				x	x

Figure 6 – Dependency Matrix fragment.

4.2.3 Performance Indicators, Estimates for Project Conclusion and Project Performance History

The Project Monitoring Panel allows managers to analyze in details differences between planned and performed values. However, it is also important to understand the project performance as a whole. Thus, IMSD-SPM calculates SPI and CPI, indicators that quantitatively represent the project performance. SPI and CPI are calculated from three variables: Planned Value (cost planned in the project baseline to the work to be done until a certain date), Earned Value (planned cost to the work done until that date), and Actual Cost (actual cost of the work done) [Fleming and Koppelman 1999]. Considering SPMont conceptualization, Planned Value is the sum of planned costs of activities (*cost* property of *Activity with Planned Cost*) whose *plannedEndDate* of *Scheduled Activity* is smaller than or equal to the considered date for calculating the indicators. Earned Value is the sum of planned costs of activities (*cost* property of *Activity with Planned Cost*) that caused the *Activity Occurrences* whose *endDate* is smaller than or equal to the considered date. Actual Cost, in turn, is the sum of actual costs of activities (*cost* properties of *Activity Occurrence*) whose *endDate* is smaller than or equal to the considered date. From SPI and CPI it is possible to determine time and cost estimates for the project conclusion. IMSD-SPM provides optimistic, realistic and pessimistic estimates [Fleming and Koppelman 1999]. Figure 7 illustrates fragments of IMSD-SPM screens showing values calculated to (a) SPI and CPI and (b) estimates.

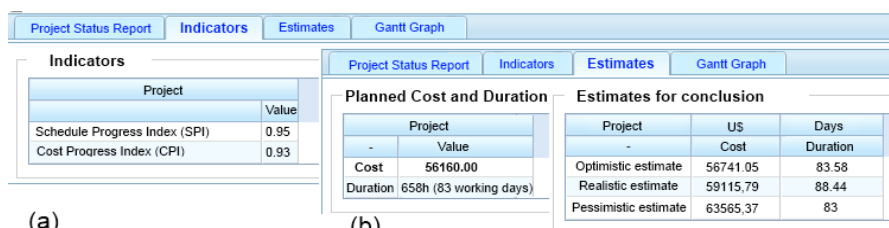


Figure 7 – Screens showing (a) SPI and CPI and (b) Estimates for Conclusion.

The project performance can change along time. Thus, a performance history showing performance indicators over time helps the manager to understand the project performance behavior, identify performance improvements or deteriorations, investigate causes, and make decisions. IMSD-SPM provides a graph of SPI and CPI values over time and also displays the history of optimistic, realistic and pessimistic estimates. Figure 8 shows an example of graph depicting a project performance history provided by IMSD-SPM.

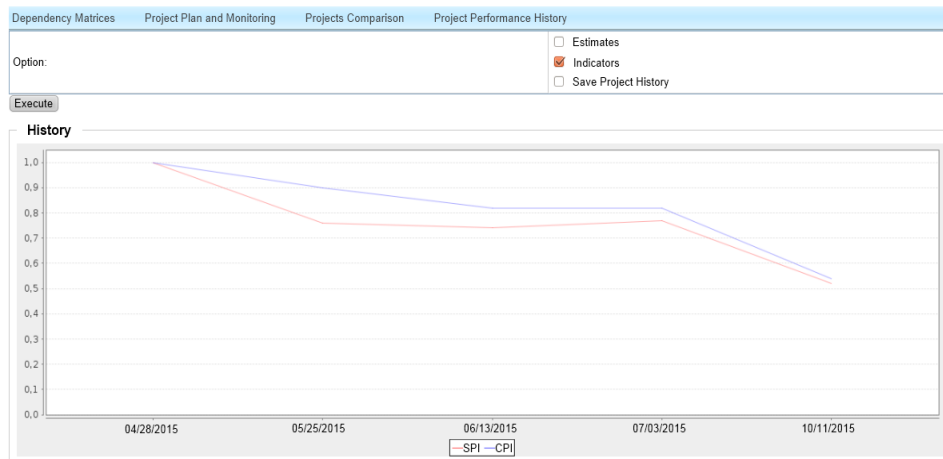


Figure 8 – Project Performance History.

5. Proposal Evaluation

To evaluate the proposal, we conducted an experimental study. The purpose of the study was to verify if IMSD-SPM is able to properly support scope, time and costs management activities. For this, two indicators were considered: (a) templates adequacy and (b) features utility. Seven graduate students and one undergraduate student participated in the study. All of them have theoretical knowledge of project management and practical experience in projects. During the study, the researcher presented information related to a project. Based on that, each participant planned the project scope, schedule and cost and created WBS, PSR and HRC semantic documents from semantic templates. Then, some situations concerning the project execution were introduced by the researcher (e.g., someone in the project got sick and had to be replaced, the supplier did not deliver material on time impacting on project activities, some activities took more than the planned time, etc.). The participants updated the semantic documents with data about project execution. The produced semantic documents were submitted them to IMSD-SPM and the participants used the IMSD-SPM features.

For each semantic template, the participants were asked to evaluate its adequacy. All the templates were evaluated as *adequate* for most of the participants. Some of them considered the templates *partially adequate*. In the case of PSR e HRC, the comments presented by the participants who considered the templates partially adequate were outside the scope of this work (e.g., record information related to human resources competences and abilities) or referred to operational improvements (e.g., automatic data generation). As for the WBS template, the reasons put by the participants who considered the template partially adequate were related to lack of graphic representation.

For each feature provided by IMSD-SPM the participants were asked to evaluate its utility. 5 out of 8 were evaluated as *very useful* or *useful* by 100% of participants. The other 3 were considered *useful* or *very useful* by at least 75% of participants.

The participants were also asked to compare the benefits of using ISDM-SPM versus the ones of using traditional documents. 62,5% answered that ISDM-SPM provides *much more* benefits and 37,5% answered that it provides *more* benefits.

The study results can be understood as initial evidences that the use of IMSD-SPM as a way to support project time, scope and cost management is feasible. However, due to study limitations (e.g., the small number of participants, the sample homogeneity and the use of a controlled environment), the results cannot be generalized. Therefore, the preliminary results obtained in the study must be confirmed (or not) in additional studies, involving more participants, with different profiles. In addition, other studies types (e.g. case studies) can also contribute to a better evaluation.

6. Related Works

As discussed in Section 2, there are some initiatives involving semantic annotation that support project management aspects. There are some similarities between our work and the proposals we found in the systematic review. However, there are also differences.

As for similarities, like IMSD-SPM, four proposals use domain ontologies as a basis to annotations and provide general features for managing semantic content (annotation, storage, indexing and retrieving). Regarding technologies used, IMSD-SPM was built using Java programming language, Postgres database and Subversion version control system. Ontology design and implementation were made using Web Ontology Language (OWL) and the SPARQL language was used to perform queries on ontologies and semantic repositories. Some proposals found in the SLR use similar technologies, such as Subversion to versioning control and Resource Description Framework (RDF), which is similar to OWL, to data description and modelling.

The main differences between our proposal and the ones found in the systematic review concern the types of annotated files, the annotation type and the project management knowledge areas supported. Regarding types of files, the proposals annotate web pages, electronic forums, pdf and text documents. IMSD-SPM also annotates text documents, but it is the only one to annotate spreadsheets. As for annotation type, all initiatives adopt manual annotation. IMSD-SPM, in contrast, uses semantic templates to automatically annotate documents and spreadsheets created by users. By doing this, users do not need to annotate files, because annotated instances of documents are generated from annotated templates. Regarding knowledge areas supported, the proposals support aspects related to Scope, Integration, Communication and Stakeholder Management. IMSD-SPM, in turn, deals with aspects related to Scope, Time and Costs Management.

IMSD-SPM differs from the proposals investigated in [Bastos *et al.* 2015] mainly due to the features to support project management activities, obtained by exploring the SPMOnt conceptualization in functionalities that help managers to plan, monitor and control projects. Although the proposals support some project management aspects, the domain ontologies used by them do not address aspects that allow comparing project planning and execution. Also, none proposal provides indicators or estimates to help project managers to monitor projects performance. Summarizing, by

exploring the SPMOnt conceptualization, domain-specific features are provided by IMSD-SPM, better supporting project management activities.

7. Final Considerations

Text documents and spreadsheets are often used in the context of project management. They are used to record information about projects and, usually, information is distributed in several documents. Thus, obtain consolidated information about the project and its progress demands effort and is prone to fails.

Considering this scenario, in this paper we presented a proposal that applies semantic documentation to software project management, supporting scope, time and cost management aspects. The proposal was implemented as an extension of the IMSD [Arantes and Falbo 2010] that enables semantic annotation of spreadsheets, and as a specialization of IMSD, the IMSD-SPM, which provides domain-related features defined based on the Software Project Management Ontology.

The main contributions of this work are: (i) the IMSD extension and the IMSD-SPM, (ii) the Software Project Management Ontology, and (iii) the systematic literature review that investigated the use of semantic annotation in the project management domain [Bastos *et al.* 2015]. It is worth noting that although IMSD-SPM is proposed to the software project management domain (SMPOnt was developed from a software process ontology), it can also be useful for project management in other domains.

It is important to point out that problems to retrieve information from documents and spreadsheets could be also addressed by approaches such as corporative architectures and information systems. However, although these approaches are capable of dealing with information recording and extraction, they imply in modifying the way organizations and people perform activities, since they replace documents by information systems. Semantic documentation can be considered an advantageous approach because it accesses information recorded in documents, allowing organizations and people to keep the way they perform activities, i.e., recording information in documents.

As future works, we plan to carry out other studies to evaluate the proposal, to extend SPMOnt and use semantic documentation to support other project management areas. Also, we intend to enable IMSD to annotate images. As a result, it will be possible, for example, to represent WBS graphically and annotate it. Currently, we are working on integrating IMSD-SPM and project management tools (DotProject and MS-Project). By doing this, organizations that use these tools can also benefit from IMSD-SPM features that are not available from other tools used by them (e.g., performance history, estimates and quality assurance).

Acknowledgment

This research is funded by the Brazilian Research Funding Agency CNPq (Processes 485368/2013-7 and 461777/2014-2).

References

ARANTES, L. O. and FALBO, R. A. (2010) "An infrastructure for managing semantic documents", In: Joint 5th VORTE and MOST Workshop, p. 235-244.

- BASTOS, E. C., BARCELLOS, M. P. and FALBO, R. A. (2015) "Exploring Ontologies for Semantic Documentation in Project Management", In: ONTOBRAS 2015, São Paulo, Brasil, p. 34 - 45.
- BERNERS-LEE, T., HENDLER, J. and LASSILA, O. The semantic web. Scientific American: 34-43 p. 2001.
- BRUGGEMANN, B. M., HOLZ, K.-P. and MOLKENTHIN, F. (2000) "Semantic documentation in engineering", In: Eighth International Conference on Computing in Civil and Building Engineering, California, USA, p. 828-835.
- ELKAFFAS, S. M. and WAGIH, A. S. (2013) "Use of Semantic Wiki as a Capturing Tool for Lessons Learned in Project Management", In: Science and Information Conference, London, UK, October 7-9, 2013.
- FALBO, R. A., BARCELLOS, M. P., NARDI, J. C., GUIZZARDI, G. *et al.* (2013) "Organizing Ontology Design Patterns as Ontology Pattern Languages", In: 10th European Semantic Web Conference – ESWC 2013, France, p. 61-75.
- FLEMING, Q. W. and KOPPELMAN, J. M. (1999), Earned value Project Management, Project Management Institute, 2nd edition.
- GUIZZARDI, G. (2005) "Ontological Foundations for Structural Conceptual Models", ISBN 90-75176-81-3, Universal Press, The Netherlands, 2005.
- LOUKIS, E. N. (2007) "An ontology for G2G collaboration in public policy making, implementation and evaluation", Artificial Intelligence and Law, v. 15, n. 1, p. 19-48.
- LU, Q., CHEN, M. and WANG, Z. (2008) "A semantic annotation based software knowledge sharing space", In: IFIP International Conference on Network and Parallel Computing (NPC), China, p. 504 - 509.
- NAKATSUKA, K. and ISHIDA, T. (2006) "Content management for inter-organizational projects using e-mail metaphor", In: International Symposium on Applications and the Internet (SAINT), Phoenix, Arizona, USA, p. 202-205.
- OASIS. Open Document Format for Office Applications. Visited in: 09 jul. 2015, https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=office.
- PMI (2013), A guide to the Project Management Body of Knowledge (PMBok), Project Management Institute, 5th edition.
- PRESSMAN, R. S. (2011), Engenharia de Software, McGraw Hill, 7th edition.
- SICILIA, M. (2006) "Metadata, semantics and ontology: providing meaning to information resources", International Journal of Metadata, Semantics and Ontologies, v. 1, n. 1, p. 83-86.
- TALAŠ, J., GREGAR, T. and PITNER, T. (2011) "Semantic wiki in environmental project management", IFIP Advances in Information and Communication Technology, 359 AICT, Brno, Czech Republic, p. 437-444.
- VILLALOBOS, J., SANABRIA, S. and CACERES, R. (2011) "Activity scheduling through gantt charts in an ms excel spreadsheet", Revista Facultad de Ingeniería, n. 61, p. 132-145.