International Journal of Applied Engineering Research ISSN 0973-4562 Volume 18, Number 4 (2023) pp. 248-255

Reusing Knowledge in Design Engineering based on Semantic Approach

Achraf Ben Miled[1,2] and Aws Abu Eid[1]

[1]Computer Science Department, Science College, Northern Border University, Arar, Kingdom of Saudi Arabia
[2]Univ.Manouba, ENSI, LARIA UR22ES01, Campus Universitaire Manouba, 2010, Tunisia
**Corresponding Author: Achraf Ben Miled.*

Abstract
Today more than ever, it has become crucial to design faster, better and cheaper than competitors in an increasingly dense and complex market. An effective management of knowledge is thus seen essential to provide the right information at the right time and help the designers in their decision making. The main objective of this work is to propose a knowledge management system for knowledge reuse based on semantic organizational models. The system focuses on helping the user by providing the relevant knowledge needed to reduce the development time of the product, the costs of industrialization and improve its quality. Once the user formulates his need in the form of a natural language query, the system sends back the relevant knowledge ordered thanks to a new approach of similarity measure.

Index Terms— knowledge management, knowledge reuse, Semantic Search, Ontology, Similarity Measure.

Introduction

The design process is a strategic issue in the product's life cycle. In fact, the optimization of this process has a competitive advantage derives from the ability to build, at lower cost and faster than competitors.
Taking into consideration not only the quality and the cost but also the time of the design, we developed a reuse system that puts the relevant knowledge to the benefit of the user.
The main contribution of our work is the integration of a semantic model in the process of knowledge reuse for a quick and efficient optimization of a design project.
For the user, it is evident that there are concepts that are more related than others, but this is much less evident to formalize.
The present study is mainly concerned with the use of ontologies, i.e., the knowledge repositories for a conceptual representation of the information in the research process.
Using ontologies in information retrieval is driven by the need to return the documents that share the maximum of concepts rather than the maximum of keywords during a search query by similarity.
Ontologies are in fact reduced to concept hierarchies where semantic interpretation is based only on a similarity measure.
The reuse of knowledge is supported by a semantic search that explores the alloy ontology which is considered as the knowledge base and a new semantic similarity measure.
The remainder of this paper is organized as follows: Section II presents related works. Then, section III introduces the context and background. In section IV, we discuss the proposed approach. Finally, we conclude this paper in Section V.

Related works

Semantic search aims to improve search accuracy by understanding searcher intent and the contextual meaning of concepts as they appear in the searchable dataspace.
Avesani et al. [1] worked on the Tagsoratic project, which makes use of the tags and category labels in the project to enable users to find posts that are semantically connected, referring to the same topic being searched. The use of tags can help to achieve connectivity through all the contents on the semantic web-enabled structure. Tagsoratic project allows a user to find posts categorized by other users under labels that are semantically equivalent to a chosen one but it doesn’t analyze automatically the subject posted by the user and doesn’t take into account post evaluation while storing them.
Moller et al. [8] presented a prototype called SemiBlog which was created with the purpose of semantic analysis. This prototype allows the user to easily add semantic metadata but it doesn’t reuse it.
Karger and Quan [5] presented several examples of potential benefits for the integration of similarity in the Semantic Web. They built a tool called Haystack which provides a robust semantic blogging environment. Haystack enables users to input and publish information but doesn’t evaluate and reuse it.
In our approach, we analyze automatically text entered by the user and we reuse knowledge by proposing it to the user if they are similarity between entered text and exiting knowledge.

Background and Context

Our knowledge engineering approach is based on the mapping of knowledge which has been created, used and shared in the mechanical design process. This mapping enables us to define a model to structure our knowledge in the form of project memory [Monticolo and al, 2014]. The project memory model is a base to define a language for presentation and sharing knowledge, ie domain ontology which is used to represent knowledge to be capitalised and exploited. To understand the concept of ontology, we take Gruber’s definition [Gruber, 1993], "Ontology is an explicit specification of a conceptualization". We complement this definition with that of Uschold [Uschold and al, 1998]: "ontology may take different forms, but necessarily include a glossary of terms and a specification of their meaning. The latter includes definitions and an indication of how concepts are connected; links collectively impose a structure on the field and force the possible interpretations of terms". The ontology is a vocabulary that defines the meaning of concepts and relationships between those concepts. This vocabulary can be associated with a model that describes the content of a knowledge base, its properties, how it can be used and the syntax and constraints provided by the language of representation. The objective is to ensure the specification of explicit knowledge at the conceptual level using a formal language offering semantics to provide unambiguous use of domain knowledge. The domain ontology is a conceptualization of domain knowledge and a specification of their relationships, for reuse via artificial intelligence techniques.Our approach to build the design domain ontology, which we called OntoDesign, was carried out in six stages (Fig 1):

- A RIOCK modelling of the domain to identify knowledge,
- The validation of domain knowledge by professional actors,
- The establishment of the knowledge typology and taxonomy,
- An analysis of existing domain ontologies for reuse,
- The specification of concepts, their attributes andrelationships

2.1 Step 1: Identify the knowledge using the formalism RIOCC

Our knowledge mapping is based on its identification resulting from an organizational model of the design process used during the projects. The model is built around the concepts of roles, interactions, skills and knowledge [Hilaire and al, 2000], [Monticolo and al, 2007]. An organization models the design process containing several sub-organizations modelling themselves on the phases and activities of the process. For each activity, we identify several roles played by professional actors, carrying out future tasks in a collaborative way. We define a role as a behavioural abstraction or a model of conduct, attached to a status, which may interact with other roles. In fact, a professional actor (human agent) instantiates an organization (roles and interactions) when it exhibits roledefined behaviour and when it interacts following interactions specified by the organization [Castelfranchi, 2000]. Within the process, the actors use and share their knowledge to carry out engineering activities. They develop their learning based on the Knowledge capitalization process. Working from experiments and observations realized in the company, we have defined for each organization (corresponding to a design process) several roles performed by professional actors. We assign to these roles the competences they use to accomplish tasks. Competence is defined at the individual level; "it is the ability of individuals to implement their knowledge and enhance their expertise in a professional context" [Le Bortef, 2002]. Thus, we add to the formalism RIO the concept 'competence' based on the work of Von Krogh [Krogh and Roos, 1995] which, through the concept of selfpoesies suggests that, in essence, knowledge can only be transmitted and recognized by interaction. Competence is an accumulation of knowledge from the learning process allowing the assimilation of knowledge. Professional competences are developed during the completion of professional activities, within which the dissemination and sharing of knowledge take place. Each competence is aggregated within a series of knowledge, which allows its characterizing. Within the organizational model, competence will be represented by a series of knowledge. We obtain a formal expansion of the RIO to RIOCK (Roles Interaction Organization Competence and Knowledge) by adding the concepts of 'competence' and 'knowledge'. In a RIOCK modelling, each competence is described by a series of knowledge. The competence defines by its name the action undertaken by the role. The interactions between multiple roles highlight the emergence of knowledge in the professional activities conducted through collaborative actions. Each activity is presented by organizational model as shown in Figure 1. In the activity (i.e. organization) 'to analyse the customer requirements' we observe two roles. The role 'Technical assistant' uses one of its competences; we read it like the capability to 'Formalize the customer requirements'. This competence requires one element of Knowledge which is used to satisfy the organization. In this organizational model, the type of knowledge is read like Knowledge on, for example the role 'Technical assistant' possesses the Knowledge on 'Customer requirements'.

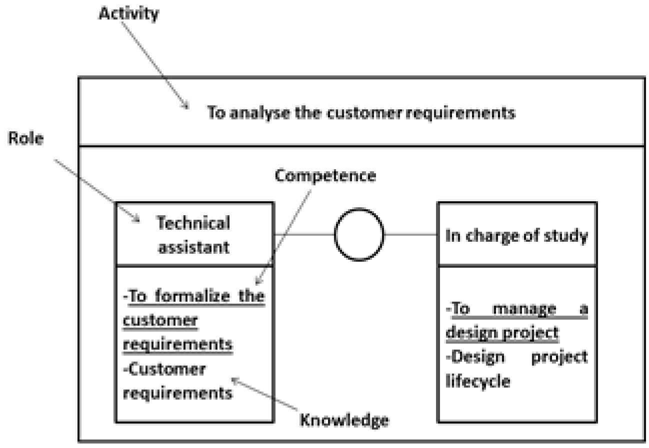


Figure 1. The organizational model for the activity "To analyze the customer requirements"

2.2 Step 2: Defining the knowledge to be capitalised upon by professional actors

Modelling using RIOCK allows us to obtain precise identification of knowledge used, shared and created during the design process. From this model we have established a series of knowledge that we have submitted to the professional actors so that they can define what knowledge can be capitalized upon in the project memory. We have drawn up a list of knowledge, from which the professional actors have identified knowledge which must be capitalized upon for each activity in each phase of the design process. The shared knowledge is that owned and used by the roles of professionalactors. The knowledge created is the knowledge that comes from the interaction between the professional actors.

2.3 Step 3: Creating the typology and taxonomy of knowledge

Before drawing up a knowledge classification, we had created a combination of knowledge. From all the Knowledge which must be capitalized upon, we have six groups of Knowledge (table 1). Each group represents a type of Knowledge.

Table 1. Knowledge classification

Name of the knowledge type	*knowledge*
Context of the project (ProjectContext)	- Knowledge related to the project progress - Knowledge presenting the origin of the project - Knowledge describing the organization of the project
Evolution of the project (ProjectEvolution)	- Knowledge related to the history of the evolution of the project
Professional processes set up in the project (ProjectProcess)	- Knowledge presenting the activities carried out, the interventions of the professional actors and the information handled for each activity
Glossary of the projet (ProjectGlossary)	- Knowledge defining the vocabulary used during the project
Expertise in the project (ProjectExpertise)	- Knowledge related to the professional rules used to develop the product
Experience developed in project (ProjectExperience)	- Knowledge describing the errors, failures and difficulties in the project

In order to structure the knowledge, we have established a taxonomy. A taxonomy is a way to classify or categorize a series of information. It is classified as a hierarchy. This hierarchy of terms can be described as a tree, in that it consists of a root and branch, and each point between two branches is called a node. The Knowledge engineering defines taxonomy as a classification of information in the form of a hierarchy formalizing the suspected relationships between the entities from the real world it represents. Each node is a taxonomy of information that represents a real entity. Each link between two nodes represents a special relationship called subclassification. The taxonomy of professional knowledge has been realized in meeting with professional actors from the company.They formalized the knowledge they want to archive in a project.
We then compared this information with memory models from existing projects [Matta and al, 2000] and the knowledge/competences Knova model [SerraFero 04]. MemoDesign is based on a taxonomy of professional knowledge where there are high-level concepts based on existing models (Project Context, Project Process, Project Expertise and Project Vocabulary). However, all the knowledge presented in the classification which makes up MemoDesign is derived from a knowledge mapping based on RIOCK modeling and expertise of professional actors. Figure 3 shows the taxonomy structuring the project memory MemoDesign:

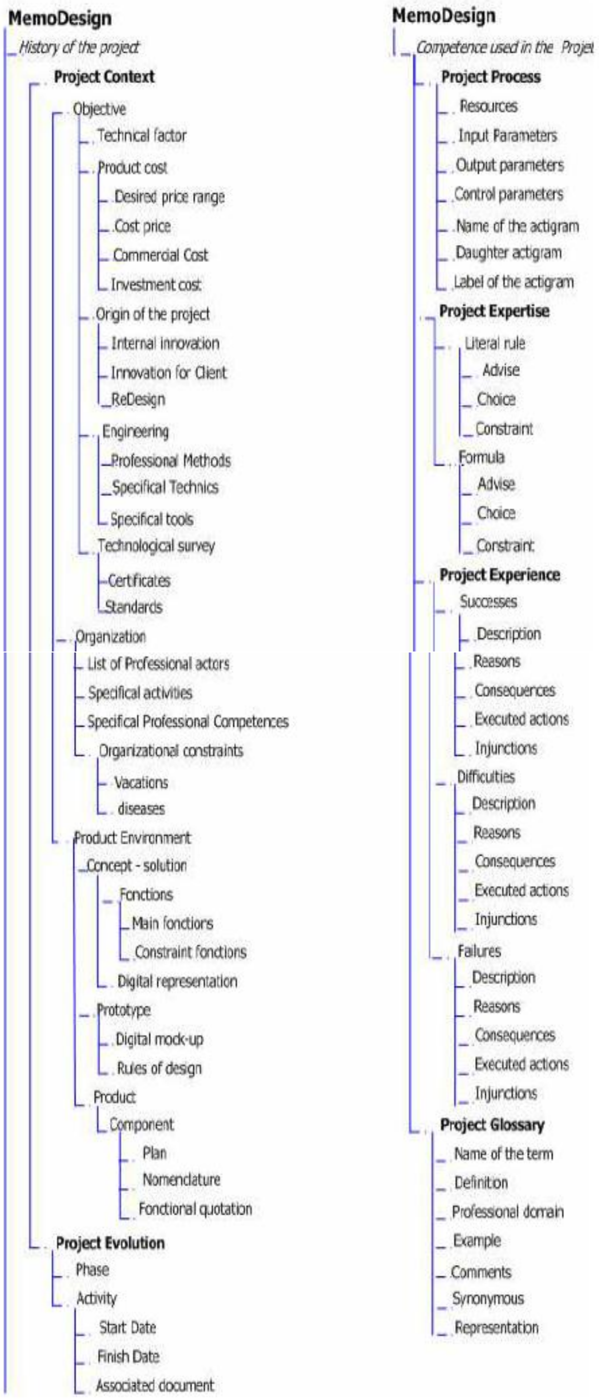


Figure 3. Taxonomy of knowledge and MemoDesign structure

2.4 Step 4: Analyzing the existing ontologies and conceptualizing

MemoDesign Our first task, after having identified which knowledge must be capitalized upon in the design project, was to analyze and study existing ontologies that we could reuse. Among the ontologies modeling the company we can cite 'Enterprise'
[Uschold and al, 1998], 'TOVE [Fox and Huang, 2005] and' O'COMMA [Gandon, 2002]. The first two ontologies are reusable at an informal level. Indeed some concepts can guide

us in the conceptualization, including the activity part ("activity") of 'Enterprise' and the product part ("Unit") of TOVE's ontology. The O'COMMA ontology is reusable because it was developed using the RDF schema, semantic web technology. This ontology is comprehensive enough to present a company memory but its concepts do not cover the terms used in product design projects. We must now conceptualize the MemoDesign memory project and define the attributes and relationships between these concepts, based on those used in the three domain ontologies studied.

2.5 Step 5: Specification of concepts, their attributes and relationships

This fifth step of the process of building the ontology OntoDesign is based on the methodology proposed in [Gandon, 2002], namely the use of tables to present concepts, attributes and relationships: The table of concepts (Table 2) has four columns, the name of the concept (term), the concept ID (ConceptID), the ID of the parent concept to which they are linked (Parent ID) and the definition of the concept in natural language.

The tables of relationships consists of five columns, the name of the relationship (relationship), the relationship ID(RelationID), the concepts involved in the relationship (Range,Domain) and a definition of the relationship in natural language.

Table 2: Extract from OntoDesign ontology concepts table

Terme	Concept ID	Parent ID	Définition au langage naturel
Driven action difficulty	Drivenactiondifficulty	Difficult	Guided action which has led to the difficult realisation of an activity
Driven action failure	Driven action failure	Failure	Guided action which has led to a failure during the realisation of an activity
Driven action success	ActionConsuite Succès	Success	Guided action which has led to carry out an activity which showed to be a success
Activity	Activity	ProjectEvolution	The whole specific action to be realized during a project
Specific activity	SpecificActivity	Organization	The whole action to be realized during a project
Functional analysis	FunctionalAnalysis	ProfessionalMethod	desciption and specification method of the product functions
Value analysis	ValueAnalysis	ProfessionalMethod	specification and estimatonmethod of the product functions costs
Success causes	SuccessCauses	Success	Action leading to the realization of the activity with remarkable results
Literal Rule Choice	LiteralRuleChoice	LiteralRule	Design rule proposed in literal form
Formula Rule Choice	FormulaRuleChoice	FormulaRule	Design rule proposed in formula form
Comment terme	CommentTerme	ProjectGlossary	Clarification of a term used in the project

Table 3: Extract from OntoDesign ontology relations table New approach and comparison

Relation	RelationID	Range	Domain	Definition
Compose Project team	ProjectTeam Composition	ProfessionalActor	ProjectTeam	Project team composition
Constraint function details	Constraintfunctiondetails	Functional analysis	ConstraintFunction	Constraint function specification
Principal function details	Principalfunctiondetails	Functional analysis	PrincipalFunction	Principal function specification
Estimate constraint function cost	Estimateconstraintfunctioncost	ValueAnalysis	ConstraintFunction	Cost estimation of constraint function
Estimate principal function cost	Estimateprincipalfunctioncost	ValueAnalysis	PrincipalFunction	Cost estimation of principal function
Carry the actigramme name	Actigramme Name	Activity	ActigrammeName	Activity name is the actigramme name
Considerate patents	ConsideratePatents	patent	SolutionSearch	Consideration of the existing patents
Considerate standards on the product	ConsideratestandardsonThe Product	Standard	LaboratoryTests	Consideration of the standards during the tests on the product
Considerate standards on the Prototype	Considerate standards on the prototype	Standard	PrototypeTests	Consideration of the standards during the tests on the prototype
Associated formula rule	AssociatedFormulaRule	FormulaRule	DesignRule	Formula rule used in prototype design

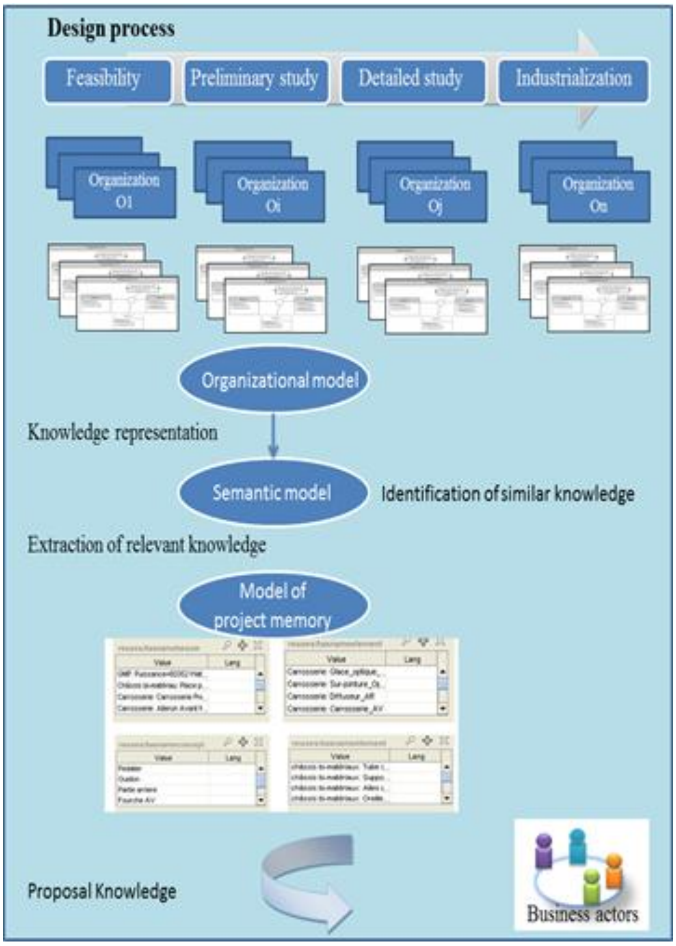
The degree of similarity measures in the ontologies are classified according to the relationships between the concepts, their content, or both [2] We are interested in the models that are based on the relations of concepts. Wu and Palmer used the length of the shortest path between concepts and their most specific common denominator in order to estimate their proximity [11]. However, their model reveals a weakness: in certain cases the value of similarity of two elements of an ontology contained in the neighborhood exceeds the similarity of two items in the same hierarchy.

A measure that takes into account all paths between two concepts and their weights is presented in the study of [10]. The disadvantage of distance-based approaches is that the similarity depends on the organization of concepts in the hierarchy.

The proposed system is based on the organizational model presented in the study of Monticolo et al. [9] which identifies the knowledge used and shared by stakeholders throughout the design process. The organizational model is responsible for highlighting the interactions between business actors and the knowledge that they will use to reach their goals.

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 18, Number 4 (2023) pp. 248-255

Our system is based on an approach that evaluates the degree of semantic similarity between the organizational model knowledge and the query posted by the user to deduce which information would be useful for the realization of its business.The figure 2 shows our approach to the reuse of knowledge.



Figuree 2. The process of knowledge reuse

As previously mentioned, a user formulates his needs in the form of a natural language query. With alloy semantic and organizational model, the system returns the most relevant knowledge.

Knowledge is presented by an ontology (fig 3) which is composed of concepts and relations. The instantiation of concepts composes the project memory which is a frame for the indexation of knowledge to be archived and reused during design project.

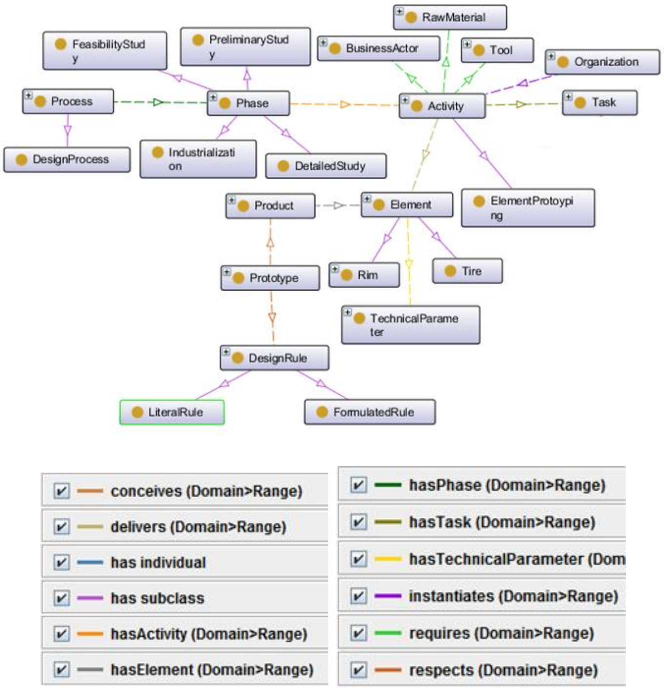


Figure 3. Ontology global overview

The extraction of this organizational knowledge is done by questioning the ontology via a SPARQL query.

When the user validates the result, the system loads to connect to the memory of the project and proposes the knowledge that is appropriate to his or her query.

The semantic search presented in our system is set in the framework of an attempt to optimize the conception process.

The semantic similarity measures play a pivotal role in the search for information; they have indeed become an abundantly explored path. Several methods have been proposed in this direction.

The aforementioned approaches for the calculation of the similarity between the concepts of an ontology either did not differ the links weight according to their type or did ignore some of the semantics that they represent.

Our proposed semantic search consists of comparing the user requests with the existing knowledge from a semantic perspective. We proceed by the lemmatization before calculating the degree of similarity between the query typed by the user and the organizational knowledge.

Non-meaningful words, called also stop words, are then omitted. This is a kind of filtering applied to information. For instance, after lemmatization, "lists of needs" becomes "list need". Notably, the step of lemmatization is applied to both the user query and the knowledge of the organizational model.

We chose to use the lemmatizer TreeTagger for its strength and efficiency.Once the lemmatization is complete, we pass to the calculation of similarity.We propose a new arcs-weigh based approach for measuring the semantic similarity.We assign a weight to each arc connecting two concepts. We adopt the coefficients of Sussna (Table 1). The lower the weight, the closer the concepts are semantically.

Table 1: Coefficients in relation to the type of relationship [10]

Relationship	Coefficient
Antonymy	2.5
Hypernymy, Hyponymy, Holonymie and Meronymy	2-(1 / nr(c1))
Synonymy	0

nr (c1): function counting the number of edge of type r from c1.We adopt the weighting arcs to choose the shortest path between two concepts. The semantic similarity between two concepts is calculated by the sum of the distances separating two consecutive concepts divided by all the concepts of the shortest path. We use Dijkstra algorithm to find the shortest path.

We normalize the semantic distance by dividing by 2 * n for a similarity between [0.. 2,5]. The similarity is given by the following formula:

$$sim(c_1, c_2) = [\sum_{i=1}^{n-1} w\,(c(i), c(i+1))]/(2 * n)$$

To be consistent with the stream of research that considers "the more weight between two concepts, the more similar they are", we change sim by Dsim where:

Dsim (c1,c2)=2.5-sim(c1,c2)

If the requested concept does not figure in the ontology, we assign the value 2.5 to Dsim.

We compared our approach with that of Wu and Palmer and Sussna following two criteria, namely, the response time and the relevance of the result.

Response time

The computation time of the similarity of two concepts in the ontology is very important. The faster the execution time is, the more relevant and favorable the measure is.

To check the validity of our measure, the speed of calculation should be tested with respect to the measure of Wu and Palmer which is considered the fastest in terms of generation time of similarity (Lin, 1998). The results show that our measure is at advantage compared to the approach of Wu and Palmer (Fig 4).

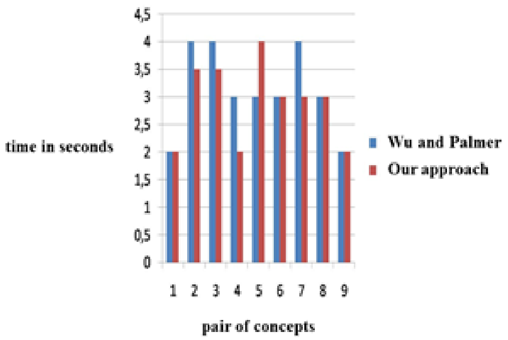


Figure 4: Comparative histogram of the speed of calculation provided by the measure of Wu and Palmer and ours

Relevance of the results

The measure of Wu and Palmer has the disadvantage of producing a similarity value of two close concepts which exceeds the value of two concepts in the same hierarchy. We therefore compared the relevance of similarity values obtained by our approach with that of Sussna. The experimental results show that our measure is more advantageous. They prove that the measure produced by our approach ensures both the speed of calculation and the relevance of values produced for the similarity of two concepts as shown in figure 5.

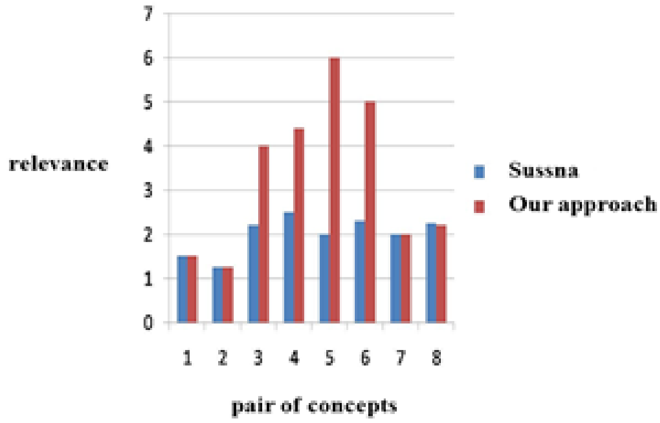


Figure 5. Comparative histogram of the efficiency provided by the measure of Sussna and ours

The results that helped drawing the histograms mentioned above are presented in the following table (Table 2). In the table 2, we chose a representation by pairs of concepts in the ontology.

Table 2. Experimental results comparing our measure with the measures of Wu and Palmer and Sussna

C1, C2	_Wu and Palmer Sim		_Sussna Sim		Our Sim	
	Values	Time	Values	Time	Values	Time
Constraints series, Lists rules	0.66	0.02s	0.95	0.4s	0.5	0.02s
Car, Vehicle	0.8	0.04s	0.3	0.8s	0.1	0.03s
Parameters techniques, Need	0.44	0.03s	1.2	0.7s	0.93	0.03s
Activities, Product	0.5	0.04s	1.32	0.3s	0.6	0.04s
Category, Four wheels	0.8	0.03s	0.3	0.3s	0.22	0.04s
Measure, Car	0.4	0.04s	0.8	0.4s	0.32	0.02s

The experimental results prove that the measure produced by our approach ensures both the speed of calculation and the relevance of values produced for the similarity of two concepts. Our new measure improves the average precision of the order of 5.5% compared to the conceptual measure proposed by Sussna.

The main objective is to match the query with a similar existing knowledge. The semantic search model integrate similarity measure that exploit an ontology hierarchical structure to calculate score between query and knowledge defined in the organizational model.
Once knowledge is identified, the system takes charge of the search for relevant information by querying the project memory and proposes to the user relevant knowledge related to the query.
Visualization of the relevant knowledge is done through a SPARQL query. For example: the professional actor enters the query "series needs" in the search field, it will have as a response to this application the knowledge "needs list." If valid, the system formalizes this SPARQL query:

```
"PREFIX reuses: <http://www.owl-ontologies.com/reuses.owl#>"
+ "SELECT ?hasnameRequirement"
+ " WHERE "
+ "{"
+ "?Requirements reuses:hasnameRequirement ?hasnameRequirement."
+ " } ";
```

The result of this SPARQL query is as follow:

```
{CustomerRequirements: Conform NF R30-020}
{CustomerRequirements: Standardization = 80% mini}
{CustomerRequirements: Autonomy GPS = 15 H}
{CustomerRequirements: Hygrometry = 0-100%}
{CustomerRequirements: Capacity of the battery = 8A/h}
{CustomerRequirements: Maximum mass of the VAE = 30 Kg}
{CustomerRequirements: Mass of the VAE + user's mass= 30 Kg}
```

Figure 4. Example of knowledge proposed to the professional actor.

The knowledge semantic search limits the search space and helps users to find relevant knowledge. To achieve our aim we have installed a prototype which helps professional actors in all the phases of the product design (fig 6).

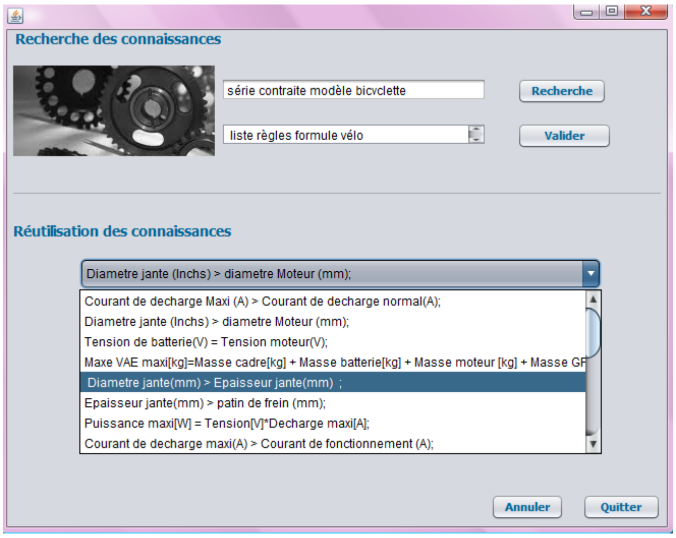


Figure 6: knowledge reuse system

Conclusion

In this article, we propose a knowledge reuse approach based on the semantic search.
The first step of this approach is based on the knowledge identification. We use semantic search to match the query with a similar existing knowledge. The semantic search tries to identify similar knowledge using a new approach of similarity measure.
Once knowledge is identified, the system takes charge of the search for relevant information by querying the project memory and proposes to the user relevant knowledge related to the query. This work facilitates knowledge management and help users to find relevant knowledge which is a time consuming task. Our future work will focus on evaluating the approach with using an ontology composed of an important number of concepts.

References

[1] Avesani P., Cova M., Hayes C. and Massa P. (2005). Learning Contextualised Weblog Topics. In E. Adar, N. Glance and M. Hurst (eds), Proceedings of WWW 2005 2nd Annual Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics.

[2] Budanitsky, A. et Hirst, G. (1999). Semantic distance in WordNet: An experimental, application-oriented evaluation of five measures. In Proceedings of the Workshop on WordNet and Other Lexical Resources. ACL.

[3] Evbuomwan, N.F., Sivaloganathan, S., Jebb, J. (1996) 'A survey of Design Philosophies, Models, Methods and Systems', Journal of Engineering Manufacture, Vol. 210, pp.301-320.

[4] Gomes, S., Sagot, J.C. (2000) 'A concurrent engineering experience based on a cooperative and object oriented design methodology', 3rd International Conference on

Integrated Design and Manufacturing in Mechanical Engineering, IDMME 2000, Montréal.

[5] Karger D.R. and Quan.D. (2004). What would it mean to blog on the semantic web? In S.A. McIlraith, D. Plexousakis and F. van Harmelen (eds), Third International Semantic Web Conference (ISWC-04), Hiroshima, Japan, Vol. 3298 of Lecture Notes in Computer Science, Springer.

[6] Le Bortef,G. (2002), 'L'ingenierie des competences', editions d'organisation.

[7] Lin, D. (1998). An information theoretic definition of similarity. In Proceedings of the 15th ICML

[8] Möller K., Breslin J. and Decker S.. (2005). SemiBlog – Semantic Publishing of Desktop Data. Proceedings of the 14th Conference on Information Systems Development.

[9] Monticolo D., Hilaire V., Koukam A., Gomes S. (2007). OntoDesign; A domain ontology for building and exploiting project memories in product design projects. International Conference in Knowledge Management in Organizations, Lecce Italia, 6p.

[10] Sussna, M. (1993). Word Sense Disambiguation for Free-text Indexing Using a Masive Semantic Network. In Bhargava. Of the 2nd International Conference on Information and Knowledge Management (CIKM). Washington. DC. USA. ACM, pp. 67-74.

[11] Wu, Z. et Palmer, M. (1994). Verb semantics and lexical selection. In 32nd. Annual Meeting of the Association for Computational Linguistics. New Mexico State University. Las Cruces. New Mexico, pp. 133-1