A knowledge reusing methodology in the product's lifecycle scenario: a semantic approach

Enrico Vezzetti

Dipartimento di Ingegneria Gestionale e della Produzione, Politecnico di Torino,

Corso Duca degli Abruzzi 24, 10129,Torino, Italy

Tel.: +39 011 5647294; Fax: +39 011 5647299;

e-mail: enrico.vezzetti@polito.it

Abstract. Considering that the market is always asking more and more complex and customized products, but that at the same time it is necessary to maintain under control cost and time, enterprises have to find appropriate methodologies and tools able to support their knowledge sharing for exploiting as best as possible all the experiences and to focus all the efforts on the development of the product innovative aspects. To reach this aim, it is necessary to introduce in the product life cycle management approaches appropriate methodologies for supporting knowledge formalizing during the product development process, in order to store in a comprehensible, accessible and simple way the company expertise and to support its retrieve by any other could be interested in. This paper deals with proposing a methodology for knowledge reuse during the product "concept" phase, working on customer requirements, product specification and semantic strategies. The proposed methodology has been experimented inside a student race team, supporting the development of the new version of the their vehicle.

Keywords: Knowledge Management, Knowledge Reuse, Product lifecycle Management, Product Concept

1 Introduction

The actual industrial scenario is driving many companies to be part of a globalised and distributed enterprise. Talking about of virtual enterprises means that companies are composed by many plants and located in different countries, but sometimes also means talking about joint venture between companies that work together on a same project, sharing competencies for reaching a common project and aim. In this scenario, product life cycle management [1] becomes always more and more strategic because it is necessary to provide to a lot of actors, involved in the product life cycle process, different kinds of data in real time. Moreover, the necessity to work in a mass customization scenario, every customer wants a low cost customized product, drives many companies not only to invest on product innovation, but also on process efficiency. For that reason, it is necessary to invest both in the management domain and in the information technology domain. In fact, if the actual use of the Product lifecycle solutions stress so much the efficiency in the use of product data, the new tendency seems to be working on a more aggregated level, that means dealing with knowledge. This means not only considering the simple data, as it is, but managing them with an integrated approach, where it is possible to store tacit and explicit knowledge. For this reason many companies start to invest in creating a knowledge management system for developing an efficient storage for a successive simple reuse.

Wiig [2] in fact states that most organizations operate in environments that they cannot control, they have a strong need to well manage knowledge. The intent with Knowledge Management is to manage knowledge practically and effectively to reach broad operational and strategic objectives. One doctrine of KM is the need to arrange our affairs to avoid rediscovering what earlier thinkers have created but maximize the reuse of valid knowledge and practices. Research on the nature of intellectual work will explicate how different kinds of knowledge is used, should be possessed, and accessed. Advanced information technology will increase abilities to supplant and support complex work tasks.

Bohn [3] observes that in dynamic environments and industries, knowledge about the process of product development is incomplete in the beginning and develops gradually over time, through various modes of learning. The process of design is characterized by complex deliberations about a series of interdependent decisions that lead to design solutions. Based on a study of concurrent product development activities, Ramesh and Sengupta [4] observe that knowledge about these deliberations is typically lost as it is never recorded. Davenport and Prusak [5] suggest that better knowledge of past, similar product development processes can lead to assessable efficiencies in product development and its consequent production. Such knowledge utilization is innately a collaborative process [6].

Dealing with new product development in particular, knowledge management system could be mainly focused on the product "concept" phase because of the significant effort implemented by designer in discounting a new efficient technical solution for customer requirements. This happens because in the concept phase, initial product ideas are developed into product specifications, that is, a product concept in the form of descriptions of the future product's major properties such as functionality, durability, cost, and so on.

Generally a number of strategic decisions are being made during this phase regarding such aspects as product features, target markets, competitive positioning, and so forth [7,8].

Since much of the formal and informal knowledge along with the context associated with it is lost after the process is completed, development teams are unable to leverage knowledge actualized by earlier teams. This justifies the reason why it is necessary to have methodologies able to trace, store and retrieve knowledge.

While knowledge management is widely discussed in literature, different solutions emerge in specific areas and they also appear to be valid, though all these works are only addressed to solve ongoing problems and they lack for an approach focused on knowledge sharing and reuse, as much important issues, which also will increase.

Focusing the attention on this point, the next paragraph will explain the architecture and the main features composing the proposed methodology.

2 The proposed methodology

The structure that we set to establish must ensure not only a standard data organization but also their easy reusability. In companies often products are completely re-engineered also when in the organization's portfolio are available similar products or reusable parts.

Moreover, even upgrading a product is not easy, if you do not have at your disposal the starting designer, who owned the knowledge of his choices. Providing a tool that allows us to verify if the problem we are now facing has already been submitted in the past leads a completive advantage. Having the possibility to analyze projects similar to those we face allows us to make only a partial re-engineering of an existing product or to re-adopt its more compelling solutions.

Another important aspect is the fact that this instrument has to facilitate the learning process that every new member must perform within the company. The suggested method looks to reconstruct exhaustively the processes that are faced in the product "concept" phase.

Starting from the analysis of costumer/market requirements, because they indicate the problem to solve, with an efficient project/product, the designer has to find a set o technical specifications as answer.

Once implemented this step, it would be useful to have a tool able to discover similar projects, in term of requirements and technical specifications, in order to support the enterprise to identify a “physical principle” able to satisfy the costumers needs. If this analysis produces positive results and the matching score is high, it means that the actual project could be implemented employing the principle that has been already developed by the enterprise, for instance, it would be possible to adopt the same material, the same geometry, the same technological specifications, …. On the other hand, if the analysis result is at the same positive, but the matching score is lower, the previous project results could be employed only as starting point for the new project, for instance it is possible to adopt a geometry similar to the previous one, that could be modified for fitting the different scenario features.

For that reason, the proposed method starts from customers requirements, and for supporting enterprise and product lifecycle interoperability firstly associates them to a set of standard requirements coming from ISO10303[9]. In the second phase, once the requirements have been translated in a more univocal way, they are correlated with a set of technical specifications using a correlation matrix [10] and obtaining a set of "weighed" specifications.

The technical specification set will be composed with parameters coming from the ISO standards [9], in order to maintain an high level of interchangeability also with the technical specifications.

Finally, a project comparison between different projects, working on a similarity index coming from a correlation matrix between requirements and technical specification, is developed in order to extract the projects that tackle the same issues.

For a better understanding, the method has been graphically described, in all its steps, by the use of and IDEF [11] codification.

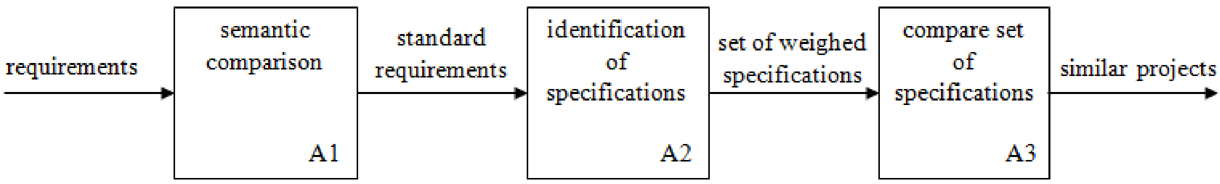

Fig. 1. Overview of the proposed methodology (IDEF)

2.1 Semantic Comparison

Considering the necessity to work in a collaborative environment and to support a more simple readability and understanding by all the stakeholders, the requirements have been matched with a standard set by the use of a semantic comparison approach, working on ISO10303 [9] standards.

The requirements are in fact usually formalized in a linguistics form, therefore they are subject to a considerable variability of expressions. As a matter of fact, the same concept can be expressed with different words and grammatical constructions. In our case, we have the need to translate a requirement, expressed by customer with a sentence, with its corresponding parameters, within the standard set, that should be maximized or minimized. This last feature could be expressed by the use of a simple sign (-/+) associated to the standardized parameter.

To reach this aim we use a technique based on semantic comparison (Fig.2) between single words. This comparison is made possible by the examination of some lexical taxonomy, in which the words appear connected by their semantic relationships [12].

The similarity between two sentences is achieved by combining the similarity measures between individual words according to Lin method [13], that tackle the problem in depth, is adopted (Fig.3). This method involves first of all the processing of the two sentences to be compared, using chunking algorithms and shallow-parsing (parsers with restrictions on complexity). So, noun phrases, prepositional phrases and verbal phrases are extracted. The different kinds of phrases are compared each other and finally the measures of these comparisons are combined in order to obtain the overall similarity between the two sentences.

Prior to the calculation of semantic similarity each source sentence undergoes a transformation [14]. The first step is a stemming process, which removes very common words (also called stop words), such as articles and prepositions, and normalize other words: the names are taken in the singular form, the verbs in the infinitive form, etc..

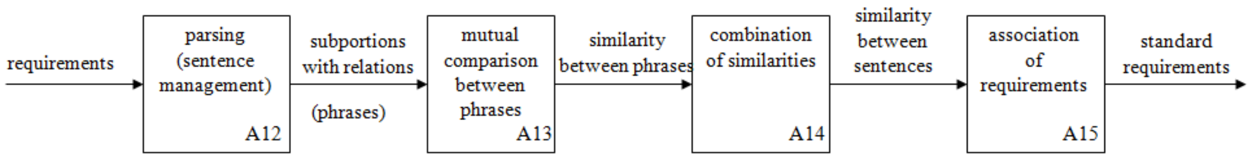

Fig. 2. Overview of A1 node proposed methodology: Semantic comparison (IDEF)

Afterwards there is the shallow parsing (Fig.4), a technique of simplified parsing whose task is to recover only a portion of syntactic information, and that is effective in the treatment of natural languages. This technique is usually made up of three different modules:

- **Part Of Speech Tagging**: given the context the algorithm performs the morpho syntactic analysis of the sentence, it associates to every word a tag indicating its type: noun, verb, adjective, adverb, article, preposition, etc..

- **Chunking**: words are now grouped in sub-portions, called chunks (noun, prepositional and verbal phrases).

- **Relation Finding**: it expresses the relationships that each phrase has with the main verb (subject, object, place, etc).

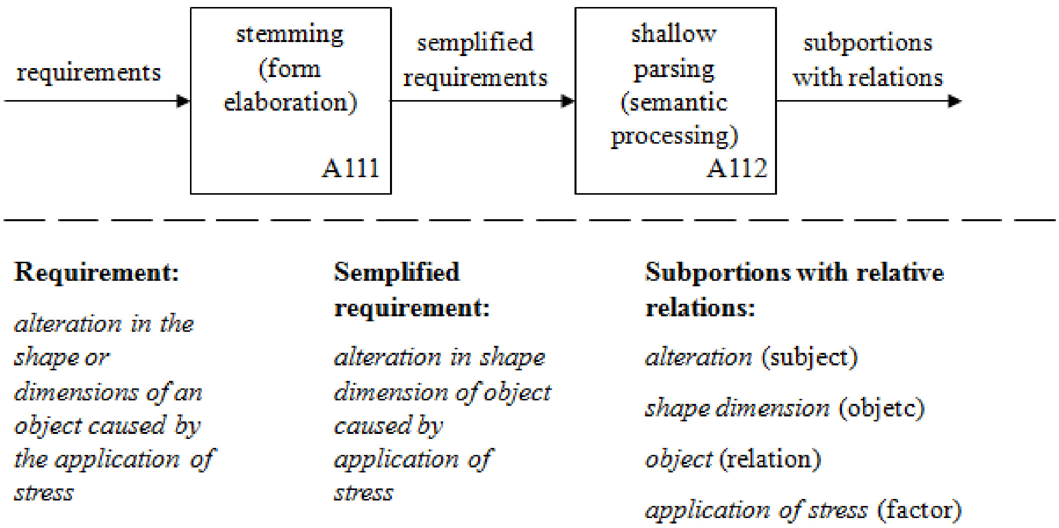

Fig. 3. Overview of A11 node proposed methodology: Parsing (IDEF)

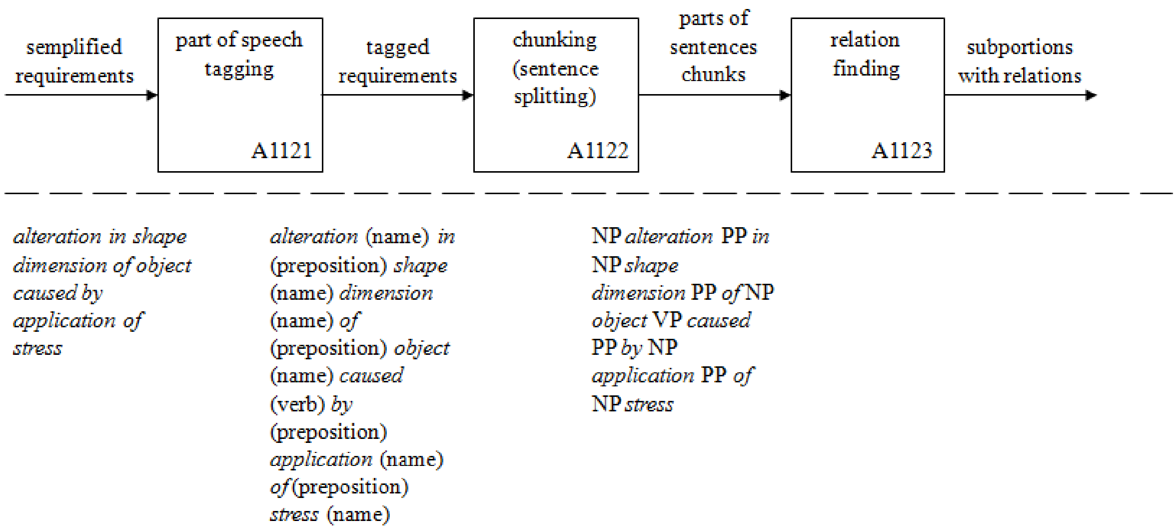

Fig. 4. Overview of A112 node proposed methodology: Parsing (IDEF)

The syntactic similarity between phrases is based on the similarity between two words. Adopting the Lin method [15,16,17] the similarity measure could be implemented, working with noun phrases *(np)*, verbal phrases *(vp)* and prepositional phrases *(pp)*, with the following formula:

$$Sim_{sent} = \gamma \cdot Sim_{np} + \delta \cdot Sim_{vp} + \eta \cdot Sim_{pp} \tag{1}$$

Where the factors γ, δ, η have [0.1] as range of variability and satisfies the condition:

$$\gamma + \delta + \eta = 1 \tag{2}$$

Taking noun phrases as an example, but the process is the same for the other phrases, it is possible to explain how the similarity is evaluated. First of all, for each of the two sentences that have to be compared, the different noun phrases contained are considered and used to form a vector (the noun phrase set, *np*), which contains the different words extracted from noun phrases. Then the union of the two noun phrase generates the noun phrase vector feature set, *(vfnp)*. Every word w_i of this vector is compared with its corresponding word in the noun phrase, the similarity value is multiplied by *i* and becomes the *i-th* element of the noun phrase vector.

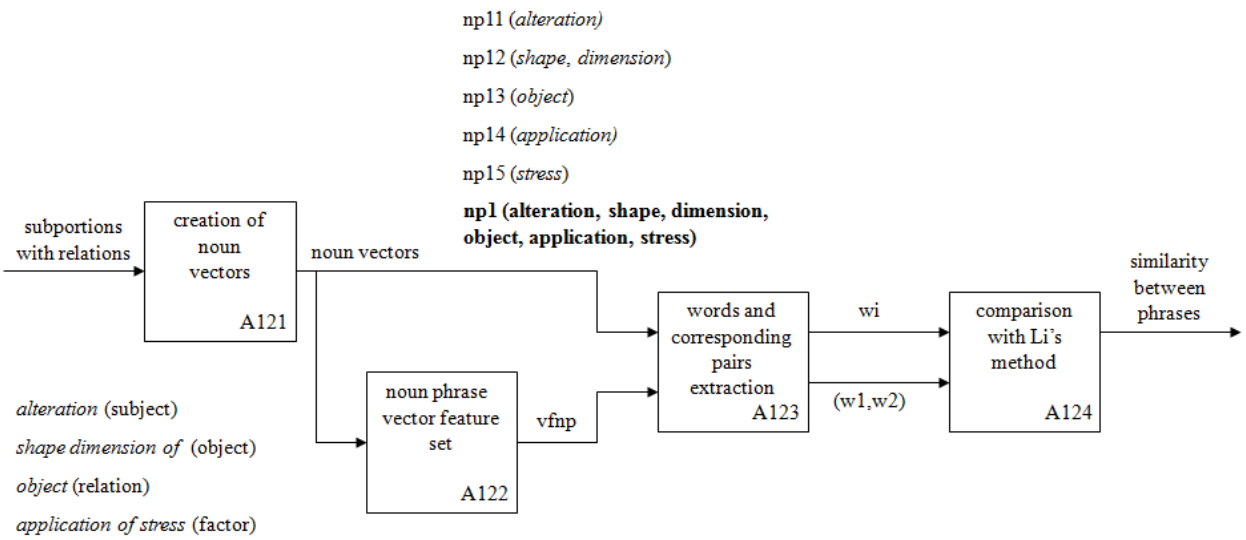

Fig. 5. Overview of A12 node proposed methodology (IDEF)

When for a word of the *vfnp* there is not a corresponding in the *np* the maximum value of the similarity between w_i and the entire set of words of the *np* is taken into account (Fig.5,6).

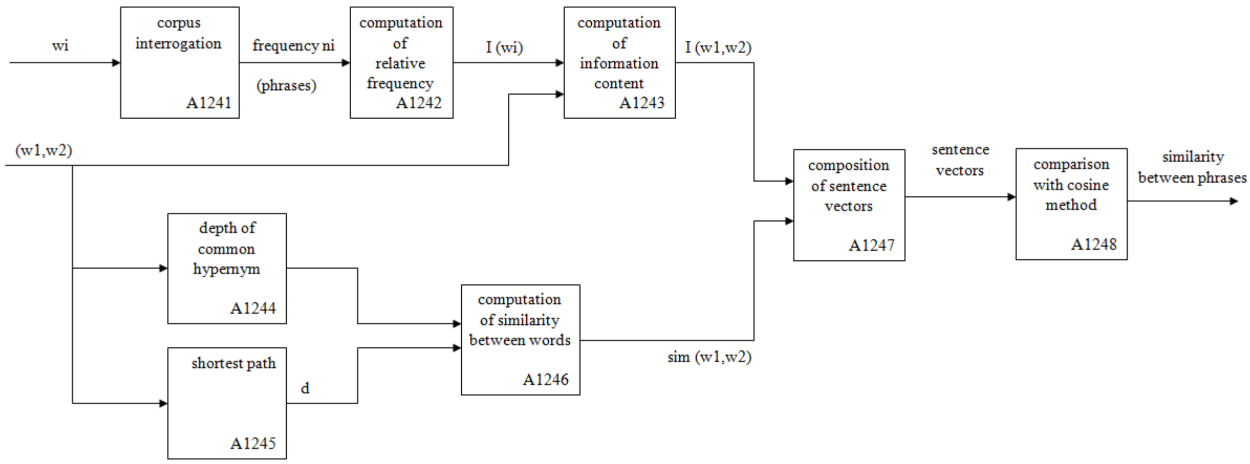

Fig. 6. Overview of A124 node proposed methodology: Li comparison method (IDEF)

2.2 Similar projects identification: correlation matrixes comparison

The next step requires to compile a correlation matrix between costumer requirements and technical specifications.

The technical specification set will be composed with parameters coming from the ISO standards [9], in order to maintain an high level of interchangeability also with the technical specifications.

In order to obtain an additional layer of information onto the way the specification should take to fulfill the requirements, to each matrix correlation cell together with a value that explain which is the correlation level between the requirement and the technical specification is assigned a sign. This additional information could be positive, to indicate that it is necessary to maximize the technical specification to satisfy the requirement, or negative if it is necessary to minimize the technical specification value to fit the correlated requirement (Tab.1,2).

Table 1. Technical Specification and costumer requirement correlation level

Strong positive correlation	+9
Medium positive correlation	+3
Weack positive correlation	+1
Weack negative correlation	-1
Medium negative correlation	-3
Strong negative correlation	-9

Table 2. Correlation matrix structure

	Req. Relative weight	Tech. Specification 1	Tech. Specification 2	Tech. Specification 3	Tech. Specification 4	Tech. Specification 5	Tech. Specification 6
Requirement 1 [+]		-9					
Requirement 2 [-]			+3		+1	+1	
Requirement 3 [+]				-1			
Requirement 4 [+]						-3	+9

The comparison, implemented between the actual project and the already store ones, is developed working correlation matrixes.

Each project is characterized by a number of involved requirements, and it differs from another because of the importance that they assume (*pi*) (Tab.3)

Table 3. Customer requirements relative weight

Negligible	1
Preferable	2
Important	3
Very important	4
Fundamental	5

After assigning the importance to each requirement, it is easy to obtain the relative weight *(d_i)* that it has in the project, using the formula:

$$d_i = \frac{p_i}{\sum_{i=1}^{n} p_i} \quad (3)$$

Where *n* denotes the number of involved functional requirements. Then it is necessary to evaluate the importance *(w_j)* that each technical specification assumes in the specific project, using the precompiled correlation matrix. This importance is obtained by the sum of products between the importance of each requirement and the value of the relationship between the *j-th* specification with each requirement:

$$w_j = \sum_{i=1}^{n} d_i \cdot |r_{ij}| \quad (4)$$

where:

w_j = importance of the *j-th* feature, with $j = 1,2, ...,m$

d_i = relative importance of the *i-th* requirement, with $i = 1,2, ...,n$

r_{ij} = relationship between the *i-th* requirement and the and *j-th* specification

n = number of functional requirements involved

m = number of technical specifications

Finally the absolute importance level is transformed into the relative one w_j^*, expressed as a percentage:

$$w_j^* = \frac{w_j}{\sum_{j=1}^{m} w_j} \tag{5}$$

In this way a quantitative measurement of the importance for each technical specification within a given project could be obtained. The comparison between different projects is based on this vector of the relative importance for each technical specification (Fig. 7).

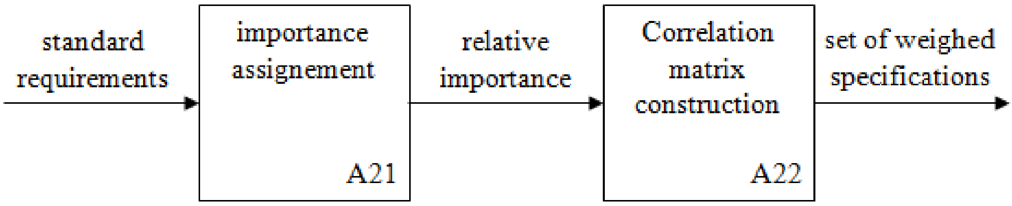

Fig. 7. Overview of A2 node proposed methodology: Technical Specification Identification (IDEF)

In order to maintain a certain significance, from a statistical point of view, in the selection only those technical specifications, with a relative importance greater than 10%, are extracted. Each vector element is compared with the corresponding one in all the other projects, considering a possible difference of about 10%. In this way it is possible to extract all the projects already implemented in which someone had faced a similar problem. This is very useful because it allows us to analyze the choices that we ourselves, or often someone else, had made in the past to meet demands (Fig.8).

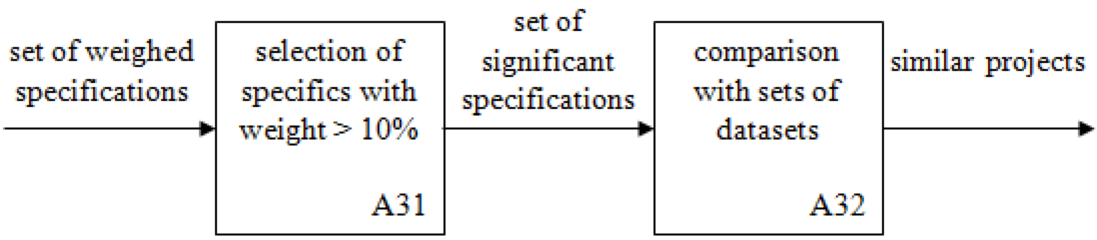

Fig. 8. Overview of A3 node proposed methodology: Technical Specification Comparison (IDEF)

3 The experimental Validation

In order to analyze the proposed methodology employed a case study has been employed. It has been decided to test the proposed methodology working together with a student race team. This choice has been justified by the significant turnover that characterize this kind of design group. Every year in fact some of its skilled members leave the group because of graduation, while new ones approach the design of the new vehicle needing to be skilled as soon as possible. For that reason, the availability of a methodology that support the storage and retrieve of the previous year vehicle development process represents a powerful tool.

Focusing the attention on the development of the vehicle structure, the designer group complete its work developing an efficient proposal for structure geometry, material and tolerances.

The present case study focuses its attention only on material, because on the other two aspects, the proposed method behavior works in the same way.

In order to have a sufficient number of already stored projects, on which testing the retrieve process, a first focus group, composed by new team member, has been created. They have to create the correlation matrixes related to previous vehicle structure projects. This has been necessary to have a sufficient knowledge base in order to test the retrieve approach of the proposed method.

Once obtained the set of previous project correlation matrixes the new team members have been involved in designing the new vehicle structure starting from the race rules imposed by the national organization about the new vehicle.

As described in the paragraph before the requirements set has been developed and successively matched with the standard set, as described in the previous paragraph.

The first requirements formulation is: *maximum permitted weight, maximum deformation for crash test, underbody stiffness to improve handling, competition number vehicle lifespan*.

The work has been completed translating the requirements using the standard set: *Weight, Deformation, Applicable Load, Fatigue Behavior*.

Creating the correlation matrix, that correlates technical requirements with standardized requirements, the project scenario has been completed working the following technical specification: *Density, Young Modulus, Shear Modulus, Yield stress, Ultimate tensile stress, Elongation, Fatigue limit* (Fig.9).

	importance	relative weight	Density	Young Modulus	Shear modulus	Yield stress	Ultimate Tensile Stess	Total elongation	Fatigue limit	
			ρ	Ey	G	ReH	Rm	A80	σD	
Weight [-]	5	38,5	-9							
Deformations [-]	3	23		9	3	3	3	1		
Applicable load [+]	3	23		3		9	3			
Fatigue behaviour [+]	2	15,5					3		9	
	14	100								
importance			346	277	69	277	154	23	138	1285
relative importance			27	21,5	5,5	22	12	2	10	100

Fig. 9. New vehicle structure project correlation matrix

Starting from the correlation matrix it is possible to extract the vector containing the relative importance of the technical specifications, describing the material, employable for developing the knowledge retrieve process.

Comparing the vector just created with those available in the database it is possible to evaluate if a similar project scenario has been developed.

Starting from the results coming from the vectors comparison it has been verified that structure vehicle development project of two years back is matching the scenario of the actual project for 80%.

For that reason it is possible only to start from that project results for finding the solution for the new project, as an improvement of the previous one.

4 Conclusions

The availability of similar projects supports the adoption of benchmarking strategies. Working with a methodology that involve a correlation matrix between requirements and technical specifications, as proposed, it is possible to have direct and simple picture of the problem and of the possible solutions. In this way it is also possible to rationalize, in an efficient way, the budget available investing on those technical specification that are more important.

The resulting method here proposed is an efficient tool to organize the enterprise' knowledge portfolio. It allows a quick retrieve of the past experiences supporting an efficient enterprise innovation management of a new project design.

5 References

1. Vezzetti E., Moos S. Kretly S.: A product lifecycle management methodology for supporting knowledge reuse in the consumer packaged goods domain. In: Computer Aided Design Vol. 43 pp.1902 – 1911. Elsevier (2011)
2. Wiig K.,: Knowledge Management: An Emerging Discipline Rooted in a Long History. In Knowledge Management, Daniele Chauvel & Charles Despres (1999)
3. R.E. Bohn, Measuring and managing technological knowledge, Sloan Management Review 36 (1994)
4. B. Ramesh, K. Sengupta, Multimedia in a design rationale decision support system, Decision Support Systems 15 (1995).
5. T.H. Davenport, L. Prusak, Working Knowledge: How Organizations Manage what They Know, Harvard Business School Press, Boston (1998)
6. P.S. Adler, Technology and the Future of Work, Oxford Univ. Press, New York (1992).
7. Di Benedetto, C. A.: Identifying the key success factors in new product launch. Journal of Product Innovation Management, 16(6): 530-544 (1999).
8. Cooper, R.G.: Winning at new products: Accelerating the process from idea to launch (3rd ed.). Cambridge,MA: Perseus (2001).
9. ISO10303: Industrial automation systems and integration Product data representation and exchange, ISO/TS 10303-26 (2011)
10. Akao Y.: Quality Function Deployment: Integrating Customer Requirements Into Product, Productivity Press (2004)
11. Cheol H.: The complementary use of IDEF and UML modeling approaches: Computer in Industry Vol.50. Elsevier (2003)
12. Miller, G., "Wordnet: a lexical database for English", Communications of the ACM, Vol. 38, No. 11, 39–41, (1995)
13. Lin Li, Yiming Zhou, Boqiu Yuan, Jun Wang and Xia Hu, Sentence Similarity Measurement based on Shallow Parsing, In: Sixth International Conference on Fuzzy Systems and Knowledge Discovery, (2009)
14. James H., Miles O., Susan A., Walter D., "Introduction to Special Issue on Machine Learning Approaches to Shallow Parsing", Journal of Machine Learning Research, Vol.2
15. Liu, X., Zhou, Y. And Zheng, R., "Measuring Semantic Similarity in Wordnet", Proceeding of ICMLC2007 Conference, Hongkong, 2007
16. Aston, G. Burnard,L. BNC Handbook: Exploring theBritish National Corpus with SARA. Edinburgh, EdinburghUniversity Press, 1998.
17. Li, Y., McLean, D., Bandar, Z. A., O'Shea, J. D., and Crockett, K., "Sentence similarity based on semantic nets and corpus statistics", IEEE Transactions on Knowledge and Data Engineering Vol. 18, No. 8, 1138–1150, 2006