A Pattern-based Approach for Semantic Retrieval of Information Resources in Enterprises
Application Within STMicroelectronics

Sara Bouzid[1], Corine Cauvet[1], Claudia Frydman[1] and Jacques Pinaton[2]
[1]*Laboratory for Information Sciences and Systems (LSIS), Aix-Marseille University, Marseilles, France*
[2]*STMicroelectronics, Rousset, France*

Keywords: Resource Retrieval, Semantic Search, Alignment Pattern, Goal-Oriented Mechanisms, Business Needs.

Abstract: Information-resource retrieval in enterprises is becoming a major concern nowadays because of the importance of business information in supporting the satisfaction of business objectives. To enhance resource retrieval in enterprises, this paper argues that it is necessary to coherently include the user need in the search process, in particular when this need is business-context dependent. A pattern-based approach is proposed for this purpose. The approach captures the business needs in a company using goal-oriented mechanisms and integrates them in a keyword search using alignment patterns. These patterns are used to both guide the search process and semantically fill the gap between the low-level description of information resources and the high-level needs of business actors. The approach has been applied for resource retrieval in the context of the manufacturing-process control within the STMicroelectronics Company.

1 INTRODUCTION

With the increase of information resources in enterprises, retrieving business information has become a daily challenge for the actors that have informational needs for achieving their business tasks. The complexity of the information systems in enterprises and the permanent evolution of the business needs make difficult the retrieval of the heterogeneous information resources that we can find in a company. Also, the growing use of commercial software platforms in enterprises for the extraction of data and the creation of business information does not facilitate the access to the resources produced, because existing commercial platforms do not often support the semantic management of information.

Indeed, in many recent research works, semantic techniques proved their ability to enhance information retrieval on the web and document retrieval (Chu, 2003). Semantic techniques rely on the use of ontologies and semantic web technologies, most of the time with a keyword search. Most of these techniques focus on the "what" of the user need and rarely on the "why" aspect. However, the users' needs in a company may be difficult to express with only keywords because such needs are generally complex and business-context dependent.

Basically, business information resources in enterprises may be related to the design process of products, the process control of a manufacturing activity, the service delivery of a company or either any information devoted to marketing purposes. To date, most semantic retrieval approaches are devoted to information retrieval on the web (Nunes, 2006). In addition, existing approaches do not consider the complex business problem of the user in the search process. The main used techniques to address a user need are commonly solution oriented. As a consequence, there is a lack of approaches suitable for the semantic retrieval of information resources used for business purposes in enterprises. The aim of this work is to enhance resource retrieval in enterprises considering, on one hand, the complex business needs of the users, and on the other hand, the distance that exists between a business need and the available information resources in a company.

We propose in this paper a pattern-based approach aiming at enhancing resource retrieval in enterprises by filling the gap between the low-level description of heterogeneous information resources and the high level needs of business actors. This approach relies on goal-oriented mechanisms carried by "alignment patterns" that we integrated in a keyword search.

Bouzid S., Cauvet C., Frydman C. and Pinaton J..
A Pattern-based Approach for Semantic Retrieval of Information Resources in Enterprises - Application Within STMicroelectronics.
DOI: 10.5220/0004886101930200
In *Proceedings of the 16th International Conference on Enterprise Information Systems* (ICEIS-2014), pages 193-200
ISBN: 978-989-758-027-7

The rest of the paper is organized as follows: section 2 tackles some recent works related to semantic resource retrieval in enterprises. Section 3 gives an overview of the pattern-based approach. Section 4 outlines the ontology structure. Section 5 presents the meta-model of the alignment patterns. Section 6 presents the pattern-based search and finally an example of implementation of the approach within STMicroelectronics is shown in Section 7.

2 RELATED WORKS

We focus here on the semantic retrieval approaches that include the business need of the user or its business context in the search process. Several techniques can be used to include the knowledge domain related to the user need in a search process (Yan 2010); (Nunes, 2006); (Chu, 2003); (Belkin and Croft, 1987).

In (Li et al., 2007), the authors proposed a framework called EO-Search to semantically tag heterogeneous engineering resources (e.g., CAD drawings, design manuals, data sheets, etc.) and improve their retrieval during the design process in the automotive industry. The user queries consist of a list of keywords that are processed with a semantic disambiguation technique supported with a business ontology and a lexicon of terms. The authors in (Yao et al. 2009) proposed a retrieval framework for the retrieval of resources used in the processes of product design and manufacturing. The system maps the user queries (i.e. business keywords) with the concepts of two business ontologies developed by the authors. An intention feedback mechanism is proposed to the user at the results' display to refine them. In (Zaher et al., 2006), the authors proposed the HyperTopic approach for the collaborative description and retrieval of heterogeneous resources related to software projects conducted in two French companies. The concept of "point of view" was mainly used to link the resources to the user need. The retrieval process was based on navigation along the concepts of the resulting knowledge model (Guittard et al., 2005).

Overall, most of the existing works that tackles resource retrieval in enterprises focused on enhancing the keyword search with more concepts of the involved knowledge domain captured in ontologies. The user need is generally expressed with business key words or with simple goals. However, information resources in enterprises mainly address complex business needs. Up to now, the complex business needs of the users are not well included, neither in the resource description, nor in the search processes.

We propose with our approach to deal with the complex business needs of the users in the resource search process using goal-oriented mechanisms.

3 APPROACH OVERVIEW

The pattern-based approach aims at providing a business-need-oriented description and retrieval to information resources in the companies using *alignment patterns*. We define an alignment pattern as an artifact of business need that can be simple or complex. Each artifact constitutes a part of a description process of a resource that carries business information. In the alignment patterns, the business needs are expressed with goals and the patterns are implicitly linked to each other through goal decompositions.

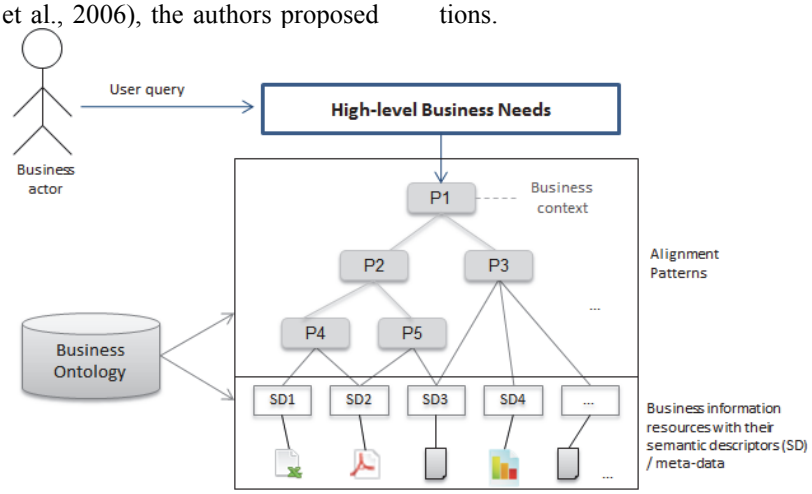

Figure 1: Role of alignment patterns in resource description and retrieval.

Also the goal decomposition may vary according to a context related to the user business activity. The pattern-based approach is supported with a business ontology that provides the necessary semantics for information-resource description and retrieval. This ontology is goal oriented.

Figure 1 depicts the role of the alignment patterns in filling the gap between business needs and the information resources that we can find in a company.

The alignment patterns carry the expression of business goals up to their satisfaction in a given business context. They provide another level of resource description close to the business needs of the users. Furthermore, we consider that these resources must have meta-data or semantic descriptors that provide their low level of description (business usage, business concepts, etc.), so to make the link with the alignment patterns. The business ontology is mainly used to consolidate this link. As a result, the alignment patterns enable somehow to "align" the low-level description of business information resources in a company with the high-level needs of the business actors.

In the proposed approach, the alignment patterns are integrated in a keyword search system, to provide a business-need-focused search of resources to the users. These patterns are progressively created, stored and reused during each search process required by the user. On the whole, the alignment patterns meet two purposes:

i. they constitute a source of knowledge related to business needs and how they must be addressed with resources, so to be shared and reused
ii. they guide the resource retrieval process, to increase the relevance of the search results to the user need.

4 THE BUSINESS ONTOLOGY

The business ontology proposed in this work gets a standard and abstract description of the knowledge domain enhanced with a goal orientation. This goal orientation enables to support the creation of the alignment patterns and the semantic retrieval of resources.

Because this work was conducted with the STMicroelectronics Company, we developed a manufacturing process ontology (Bouzid et al., 2013b). The scope of this ontology relies on four views of description:

- *Organization View*: any business activity (function) related to the manufacturing process. We included in the organization view the business profiles that are related to the manufacturing activity (e.g. process engineering, maintenance, etc.)
- *Function View*: describes the manufacturing objectives of the company.
- *Data View*: gathers the data types involved in the manufacturing process (i.e. data about equipment, data about products, etc.)
- *Control View*: gathers a process control description of the manufacturing process. The process control regroups the set of techniques and methods used to control a manufacturing process to ensure its well achievement. Note that because of the vocabulary heterogeneity related to this view, we also proposed a process control dictionary that rationalizes the process control terminologies, but this part of work is out of scope of this paper (Bouzid et al., 2013c).

The proposed views refer in fact to the descriptive levels of the ARIS architecture (Ferdian, 2001). They were chosen as a starting point for the domain description because the ARIS approach seeks reducing the complexity of modeling industrial business processes using these levels

5 META-MODEL OF ALIGNMENT PATTERNS

The pattern notion was originally developed by Christopher Alexander (Alexander, 1979) to respond to recurrent problems in the architectural domain. This notion was reused and readapted by Gamma et al., (1995) in software engineering. In this paper, patterns are used to respond to recurrent business needs whose solutions are expressed with goals.

Figure 2 depicts the meta-model of an alignment pattern. An alignment pattern is generally composed of a **goal**, a **context** and a **solution**.

A **goal** expresses a business need, which can be complex or atomic. A complex goal can be refined into sub goals, whereas an atomic goal cannot be refined.

The **context** of a need requires the business profile or activity of the user. The business profiles and activities in a company are generally identified in the business ontology that describes its core business process.

The **solution** of an alignment pattern can be business goals or one or several semantic descriptors of resources, where a semantic descriptor refers to a resource. In fact, the solution for a pattern differs

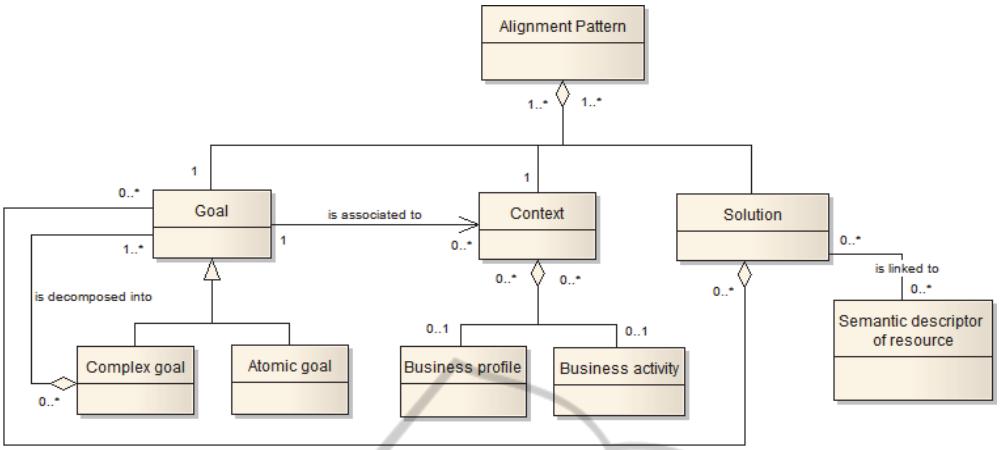

Figure 2: Meta-model of an alignment pattern.

according to the type of goal that composes it (i.e. complex or atomic). A solution of a pattern can be a set of goals when the goal of the pattern is complex. In this way, each pattern solution to a complex goal provides an "And" decomposition that enables to realize it.

Figure 3 shows a general example of an alignment pattern. *Variability reduction* is a complex goal in the example, its solution is the sub goals *Lot control* and *Equipment control* that contribute to its realization. The link between a pattern solution and another pattern is not defined in the content of the patterns, but rather deduced during the goal decomposition of the search process.

Goal = *Variability reduction*	**Context** = *Process engineering*
Solution = (*Lot control*) and (*Equipment control*)	

Figure 3: Example of an alignment pattern.

Each complex goal defined in a pattern can have alternatives of decomposition according to different contexts, because a same goal can be associated to different contexts for its achievement (Figure 2). The alternatives express in fact the "Or" decompositions, but in practice, the alternatives of decomposition are carried with different patterns.

Thus, two patterns P1 and P2 can have a same complex goal G1 in their definition but different "And" decompositions in their solution, depending on the context of achieving the goal G1.

Goals that cannot be refined constitute the low level of the hierarchical structure of goals and are known as atomic goals. Thus, when an alignment pattern has an atomic goal, its solution is the set of semantic descriptors that satisfy this atomic goal according to the specified context.

Finally, during the search process, the alignment patterns are dynamically linked through the goal decomposition mechanism, which guides the resource retrieval for every user need.

6 THE PATTERN-BASED SEARCH

6.1 Context of Application

The basic problem facing any user searching for a resource in a business context is how to well capture his need and how to retrieve the resources relevant to his business objectives in a company. We consider that using semantic-based solutions to support information-resource retrieval in such contexts must consider not only the semantics of the resources but also the semantic alignment between the resources and the complex business needs of the users. This ascertainment has been deduced through the study of the information system of the STMicroelectronics Company, a semi-conductor manufacturer. We studied in this company a set of business information resources used to support the control of the manufacturing process. These resources must be shared and retrieved by company engineers of different business activities and profiles.

Basically, the STMicroelectronics' products are manufactured in lots where each lot is a silicon wafer used as support for the construction of electronic chips. The manufacturing process spread over several work areas (*Photolithography*, *Etching*, *defectivity*, etc) where each step and operation inside it can be remade several times. Controlling this complex

manufacturing process consists in monitoring and analysing manufacturing information extracted and processed with specific data management systems. In this context, the resulting resources carry information about the manufacturing activity.

A first part of this research work has been already done within the company to create semantic descriptors to information resources related to the context of work, using a bottom-up description approach. More details about this work can be found in (Bouzid et al., 2013b); (Bouzid et al., 2013c).

In a second part of this work, the pattern-based search is proposed to put the business needs of the users (i.e. business actors) in the heart of the resource retrieval process. The users formulate their query using *business keywords*, a *goal* and a *context*, and the system performs these queries and enriches at the same time the resource description through the alignment patterns. In fact, when a user formulates his query, two situations can happen:

a. the alignment pattern(s) do(es) not yet exist, the system builds pattern-based solutions starting from the user query
b. the alignment patterns that meet the user need exist, the system takes the solutions of the patterns that better meet the user query

The interesting asset of the proposed pattern-based search is that business need artefacts are progressively captured in alignment patterns and stored in a pattern base for reuse.

6.2 Creation of Alignment Patterns during the Search Process

Three goal-oriented mechanisms are used to capture and encapsulate business-need artifacts in alignment patterns: goal decomposition, goal-sibling decomposition and goal abstraction. These mechanisms are complementary and are repeatedly used as much time as they are needed for the construction of the necessary alignment patterns that respond to a given high-level business need. When some patterns already exist and others must be created during the search process, the created ones are dynamically linked to the existing patterns following the goal decomposition relations. The use of these goal-oriented mechanisms enables to progressively fill the gap between the business needs of the users and the information resources during each search process.

6.2.1 Goal Decomposition

The standard goal decomposition technique (Lamsweerde, 2001) is used here to refine complex business goals. The goal decomposition is expressed when, considering a complex goal $CG1$, all the sub-goals of $CG1$ must be achieved in order to achieve this goal. We can also say from a bottom up view that the goals *SG1.1* and *SG1.2* contribute to the realization of the goal *CG1*. In addition, $CG1$ can have alternatives of decomposition according to the given business context.

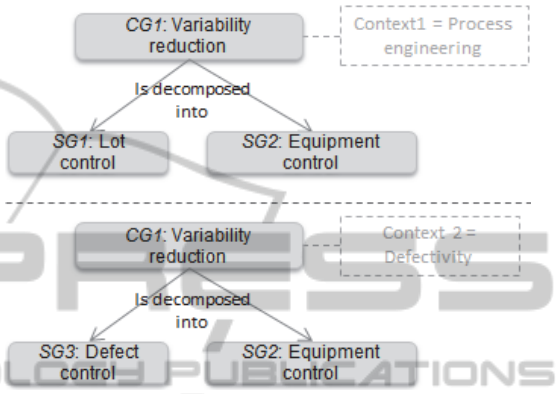

Figure 4: Examples of goal decomposition.

Figure 4 shows an example of a goal decomposition with two alternatives. *Variability reduction* is a goal that can be decomposed into *Lot control* and *Equipment control* in the business context *process engineering*, or into *Defect control* and *Equipment control* in the business context *Defectivity*.

In general, three scenarios of decompositions can be identified in a same context as follows:

- a complex goal CG is refined into complex sub goals SG_n. Each sub goal SG_i must be refined in turn until achieving atomic goals
- a complex goal CG is refined into complex goals SG_n and into atomic goals AG_m. The complex sub goals SG_n must be refined in turn until achieving atomic goals
- a complex goal CG is refined into atomic goals AG_m only. There is no more refinement to do in this case

6.2.2 Goal-sibling Decomposition

The goal-sibling decomposition is a special case of the goal decomposition. The decomposition of goals is identified starting from a sibling goal. Considering two goals *CG1.1* and *CG1.2* (Figure 5), if *CG1.1* is sibling of *CG1.2*, and *CG1.2* is used in the same context as *CG1.1*, thus *CG1.2* will be also captured in an alignment pattern and decomposed in turn if it is a complex goal.

Note that the goal decomposition scenarios mentioned before are the same for the sibling-goal decomposition.

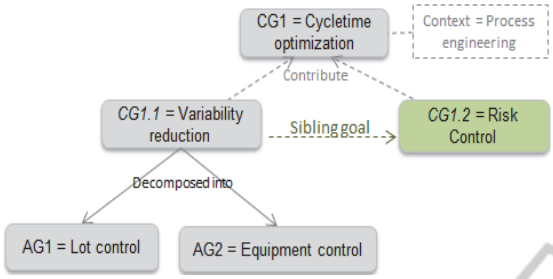

Figure 5: Example of goal-sibling decomposition.

6.2.3 Goal Abstraction

The goal abstraction is expressed when, considering two goals *SG1.1* and *SG1.2*, they contribute to the realization of a goal *CG1*. The abstraction of goals can also be associated to a given context.

In Figure 6, the goals *Cycle time optimization* and *Procedure optimization* contribute to the realization of the same goal *Cost optimization* in the same context.

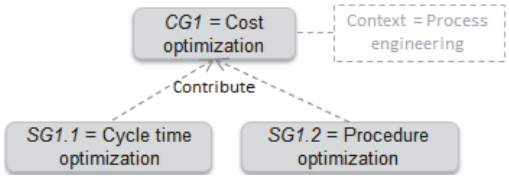

Figure 6: Example of goal abstraction.

6.2.4 Example

We consider the following user query as example from the STMicroelectronics' context:

Keywords: {'lot', 'out of control'}
Goal: {'cycle time optimization'}
Context: {'process engineering'}

The manufacturing process ontology of the company contains a description of its core business activity according to four views of description and each view describes a business aspect of the process, including the business objectives of the company related to each business activity (Bouzid et al., 2013c). Thus, in this ontology, the goal "C*ycle time optimization*" specified in the user query is in fact a complex goal. It will be then refined following its relations with other goals defined in the business ontology (Figure 7) and according to the specified business activity "P*rocess engineering*". Starting from this goal, a pattern P1 will be created (or identified in the pattern base). Figure 7 shows the obtained patterns for this example (to simplify the example, the patterns' identifiers like P1, P2, etc. are used to refer to the pattern solutions instead of using goals).

The pattern P1 with its goal and context is used as starting point to identify or create the patterns linked to P1 through the goal-oriented mechanisms. For example, the pattern P7 will be created through the goal abstraction, the patterns P8 and P3 through the goal-sibling decomposition and the patterns P4 and P5 through the goal decomposition. The resulting resources according to this example would be: R1, R2, R3, R7, R4, R6.

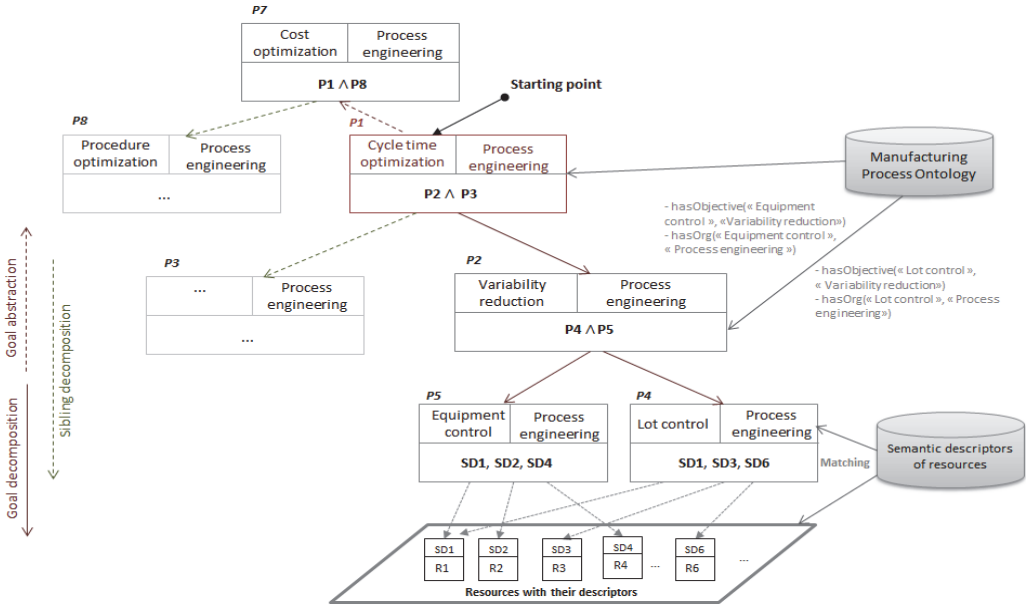

Figure 7: Example of resulting pattern architecture after query processing.

Finally, each new pattern resulting from a search process is stored in a dedicated base for reuse.

6.3 Combining the Alignment Patterns with a Keyword Search for Resource Retrieval

As mentioned before, the alignment patterns are integrated in a keyword search to enhance the resource retrieval process in business contexts.

Figure 8 outlines the search process after the required alignment patterns are created (or found in the pattern base). The alignment patterns with their solutions can be reused in several queries to guide the refinement of the user need until reaching the atomic goals and then the resources that meet them.

The keyword search is processed separately from the pattern-based search. The keyword search consists here in matching the keywords of the user query with the concepts of the semantic descriptors of resources.

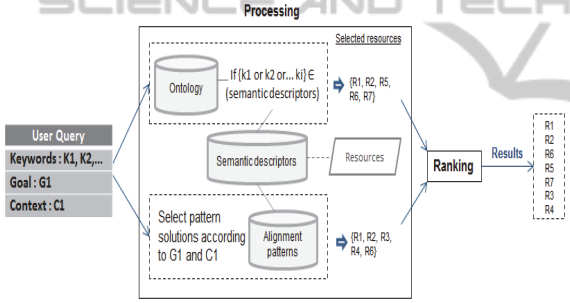

Figure 8: Processing of the user query by combining a keyword search and alignment patterns.

A standard string matching is applied, and this process is semantically supported with the business ontology and a process control dictionary. Indeed, each information resource has a semantic descriptor that provides its basic usage. In our context, the semantic descriptors provide a manufacturing description and a control description of each resource used to control a manufacturing process. These descriptors were created with a semantic mapping technique and capitalized in a dedicated base in a previous work (Bouzid et al., 2013a).

At the end of the search processing, the results obtained with the alignment patterns and with the keyword search are merged, filtered and ranked, so to only keep the corresponding resources to the user need. The ranking is also applied in order to display the results in a suitable way to the user.

Table 1 shows four categories identified for ranking the resources according to their matching with the user query. The display of the results to the user is done in the ascending order of these categories.

Table 1: Categories of ranking.

Goal	Context	Keyword(s)	Ranking categories
match(g_i)	match(c_i)	match(k_i)	1
match(g_i)	match(c_i)	match(k_i)	2
match(g_i)	match(c_i)	match(k_i)	2
match(g_i)	match(c_i)	match(k_i)	3
match(g_i)	match(c_i)	match(k_i)	4

According to the example in we consider that we have obtained the following resources with the keywords' matching: R1, R2, R5, R6, R7. Knowing that 5 resources were obtained with the alignment patterns, the resources that better satisfy the user need after ranking in this example would be: R1, R2, R6, R5, R7, R3, R4 and will be displayed in this order.

7 IMPLEMENTATION

The general retrieval approach was implemented within STMicroelectronics with a web prototype. Following the Information Technology policy of the company, the Php and C technologies were used for the programming logic and the XML technology was used to implement the business ontology.

An example of a pattern-based search is presented in the appendix. Basically, the system takes as user query one or a set of keywords, a goal and a context. A keyword here can be any concept of the manufacturing process description of the company. When the user specifies key concepts, the system automatically displays meta-information about each concept, so to help the user in choosing the right concepts to express his need. The semantics of the manufacturing process ontology is also used to assist the user in defining his business need with goals. The system processes the goal refinement task with the alignment patterns regardless of the level of abstraction of the selected goal by the user.

The pattern-based approach is still under test today within the STMicroelectronics Company. The first experimentations were done using a set of scenarios of needs defined with their solutions by the business experts of the company. 12 scenarios were tested on the keyword search alone and on the pattern-based search (i.e. the combination of the keyword search and the alignment patterns). The approach showed good results on the latter with 96%

of precision and 87% of recall comparing to the keyword search alone (78% of precision and 85% of recall).

8 CONCLUSION

This paper presented the pattern-based approach, devoted to enhance resource retrieval in enterprises. The proposed approach emphasizes a business-need focused search based on alignment patterns in a standard keyword search. The alignment patterns are created dynamically when a user searches for information resources in a business context. These patterns progressively capture business need artefacts and support their reuse for addressing recurrent needs. This original search approach presents two main interesting assets: (i) it enriches resource description with a high-level business semantics maintaining in this way the link between the resources and the needs of business actors in a company, and (ii) it progressively capitalizes on business know-how by storing the semantics of the information resources used in an activity domain.

In future work, more experimentations of the approach are planned with the business experts of the STMicroelectronics Company. A specific string-similarity measure is also planned to be tested on the matching between the keywords and the resource descriptors. This technique was previously used in our work where it proved its efficiency (Bouzid et al., 2013a). We consider that it would be worth to reuse it to foster the keyword matching during the search process.

REFERENCES

Alexander, C., 1979. *The Timeless Way of Building*, New York, USA: Oxford University.

Belkin, N.J. & Croft, W.B., 1987. Retrieval Techniques. In M. E. Williams, ed. *Annual Review of Information Science and Technoloy (ARIST)*. Elsevier Science Publishers B.V., pp. 109 – 145.

Bouzid, S. et al., 2013a. A Bottom up Search Technique of Manufacturing Indicators. In *24th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC 2013)*. Saratoga Springs, New York: IEEE.

Bouzid, S. et al., 2013b. A Semantic Mapping Approach to Retrieve Manufacturing Information Resources. In *The IEEE IFAC-MIM Conference*. Saint-Petersburg, Russia: IEEE Xplore.

Bouzid, S. et al., 2013c. A Semantic Support to Improve the Collaborative Control of Manufacturing Processes in Industries. In *17th IEEE International Conference on Computer Supported Cooperative Work in Design (CSCWD 2013)*. Whistler, BC, Canada: IEEE.

Chu, H., 2003. Information Representation and Retrieval : An Overview. In I. Information Today, ed. *Information Representation and Retrieval: An Overview*. pp. 1–25.

Ferdian, 2001. *A Comparison of Event-driven Process Chains and UML Activity Diagram for Denoting Business Processes*,

Gamma, E. et al., 1995. *Design patterns: Elements of Reusable Object-Oriented Software*, Addison-Wesley.

Guittard, C., Zaher, L.H. & Cahier, J., 2005. Experimentation of a socially constructed " Topic Map " by the OSS community. In *In proceedings of the IJCAI-05 workshop on Knowledge Management and Ontology Management (KMOM)*. Edimbourg.

Lamsweerde, A. Van, 2001. Goal-Oriented Requirements Engineering : A Guided Tour. In *Proceedings Fifth IEEE International Symposium on requirements Engineering*. Toronto: IEEE, pp. 249–263.

Li, Z., Raskin, V. & Ramani, K., 2007. A Methodology of Engineering Ontology Development for Information Retrieval. In *International Conference on Engineering Design, ICED'07*. Paris, France, pp. 1–12.

Nunes, S., 2006. *State of the Art in Web Information Retrieval*, Porto, Portugal.

Yan, W., 2010. Component Retrieval Based on Ontology and Graph. *Journal of Information & Computational Science*, 4, pp.893–900.

Yao, Y., Lin, L. & Dong, J., 2009. Research on Ontology-Based Multi-source Engineering Information Retrieval in Integrated Environment of Enterprise. In *International Conference on Interoperability for Enterprise Software and Applications*. China: Ieee, pp. 277–282.

Zaher, L. H. et al., 2006. The Agoræ / Hypertopic approach. In *International Workshop IKHS - Indexing and knowledge in Human Sciences (Sdc)*. Nantes, France.

APPENDIX