Supporting Knowledge Intensive Researches and Systematic Knowledge Organisation in Weakly Structured Business Processes

Karsten Böhm

Fachhochschule KufsteinTirol, University of Applied Sciences
Andreas-Hofer-Str. 7, A-6330 Kufstein, Austria
e-mail: karsten.boehm@fh-kufstein.ac.at

Paolo Salvatore

CiaoTech S.r.l.
Via Cassia 834 F, 00189 Roma, Italy
e-mail: p.salvatore@ciaotech.it

Abstract

The support of knowledge intensive tasks is a major concern of the research area of process oriented knowledge management. An interesting aspect is the support of weakly structured business processes with IT-tools. This contribution describes a software approach that combines the advantages of search engines that are accessible on the Internet and from professional information providers with a flexible organisation system for documents and conceptual structures (referred to as knowledge items) obtained from the research in a specific application domain. It aims at supporting systematic research activities that go beyond a single-step ad hoc retrieval and that request a systematic search strategy which includes iterative query refinement and the organisation of search results and documents. A domain specific information model will be used to organize all items (document and concepts) according to their domain and task specific relevance. During the research this domain model will be used to assist the retrieval process in the research activity.

The system is applied for the support of systematic researches in the domain of marketing and internationalization for the exploration of new markets for a company and should be used by SME, which have only limited time and expertise to carry our extensive information research activities or by consulting companies that carry out that research as a service for their clients.

1. Introduction

Systematic knowledge organisation becomes more important as the amount of information still increases rapidly and is relatively easy to access – compared to the possibilities that existed some years ago. A number of powerful search engines are available to search on Internet and Intranet resources and they provide the expert with powerful means to retrieve the needed information with a number of iterative searches.

Still, the appropriate search for users with less experience in the domain in question will be more difficult and the traceability of iterative searches is still a problem. Thus, *assistance during the individual search steps* would be desirable in order to query appropriate sources with the right questions (Assisted search). In order to be able to trace the subsequent (iterative) searches, a way of documenting the individual research steps would be useful to either *communicate* the course of a search process to persons not involved in the research (e.g. clients of a consultancy company) or to *document* and archive the search activities in order to reuse them later or to share them with colleagues.

After the retrieval process itself it is usually necessary to *store the relevant information* in order to capture the current state of the research for later reporting and to organize the information in some task-specific way, so that they form a dossier that files all the information relevant for the research carried out.

Finally, it should be possible to *document the knowledge* gained from the research in a more abstract way on a conceptual and system independent level, which should in turn be usable to organise the information pieces accordingly.

For this purpose, we have developed a software prototype within the project AMI-SME (Analysis of Marketing Information for Small And Medium-Sized Enterprises) that combines advantages of existing Internet search engines with modern text analysis functionalities and an intelligent ontology based storage system for documents and knowledge items. It has been envisaged as a system that can perform a variety of very complex information gathering and analysis tasks as well as the provision of a framework for the organization of useful information pieces on the basis of a conceptual structure. The solution aims at supporting systematic research activities that goes beyond a single-step ad hoc retrieval and that request for a systematic search strategy which includes iterative query-refinement and the organisation of search results and documents. A domain specific in-

formation model will be employed and used to organize all items (document and concepts) according to their domain and task specific relevance. During the research this domain model will be used to assist the retrieval process in the research activity.

A systematic research consists of a number of iterative retrieval steps which have to be carried out in a specific order with a dedicated task specific purpose. The results need to be filtered, sorted and stored for further processing or documentation. Current search capabilities do not support such a multi-step research activity and do also not support the organisation of the results in an application or domain dependent structure.

This contribution describes a practical application of semantic technologies in an application domain of marketing and internationalization for the exploration of new markets for a company. This domain is information intensive and needs a proper organization to enable re-use and transfer of the knowledge acquired during a research session. It employs a domain model that is used in a dual way: as a guiding structure during the research steps and as an organising structure for collecting the results of the research.

The rest of the article is structured as follows: after briefly introducing the characteristics of weakly structured business processes the application are of market research and its application within CiaoTech will be described in section 2. Afterwards, the use of ontologies as conceptual structures to support the two main objectives of the IT-solution is explained in section 3: assisted search and organisation of the research findings in a conceptual structure. The following section 4 will focus on the technical architecture and illustrate core functionalities using different aspects of the user interface. In section 5 we will discuss some related approaches und give an outlook in the concluding section 6.

2. Weakly Structured Business Processes and Knowledge Intensive Activities

The knowledge intensive activities within Business Processes are often only weakly structured regarding the execution sequence of the tasks. Moreover knowledge is often flowing between different business processes and acts independently of the sequence of actions within the core business processes, as indicated by the knowledge interactions in figure 1 below.

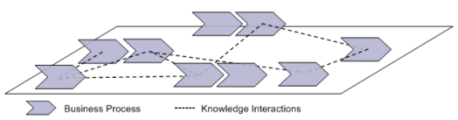

Figure 1: Knowledge Interactions spanning several Business Processes after [1]

The characteristics of weakly structured processes are usually specified as having a high variability on the executions of the single tasks and a low repetition rate. Knowledge intensive activities on the other hand are characterised by an information demand that could not be foreseen before the execution of the activities and that has to be satisfied within the task for successful completion. Thus, new knowledge must be acquired by the persons carrying out the process. Moreover the processes are often tightly connected to informal communication and information activities that will be carried out by a small number of actors which could span organizational boundaries. Those aspects together make these processes hard to manage and to support by classical approached to business process modelling and –support. A number of approaches have been suggested to support the individual activities of the knowledge worker in order to derive process patterns (process mining) or task structures (task-oriented knowledge management), see for example [2], [3]. These approaches try to identify the underlying structure, which is more task-oriented than strictly process oriented and try to model it using a bottom-up approach.

The approach taken in the AMI-SME project is similar to the task-oriented approaches but adds the methodological dimension as a structuring element – the goal is to support the indented application area as good as possible, in contrast of building a tool for all purposes. Specifically the following guidelines where used when designing the system:

- Focus on the underlying methodology that is used to carry out the process (support for marketing methodologies)
- Embrace the project instance, not only the process model (support for the market research project)
- Discriminate information sources at runtime (always provide all the available information sources)
- Support the communication and sharing of information from an individual perspective (organisation of information according to a domain specific model)

The information sources play a crucial role for knowledge intensive activities. It is more likely that such information sources specialised by topic provide more reliable and relevant information than general or multi-topic information sources. However, at least in the use cases of this research, such information sources cannot be identified in advance, only during the information search. It is most probably subject to the experience during this project to develop a stable set of information sources that provide the best information for the current demand.

2.1. The Case of Market Research

Especially small and medium-sized companies (SMEs) have to look for information with-

out investing too many resources in terms of time and money, in particular without being able to dedicate internal resources fully to this task. As valuable information about external business factors is readily available on the Web, what is needed is to explore the web resources properly. On the other hand the expertise of SMEs in using internet tools is rather restricted, especially if the domain that is researched for is only partly known. Therefore there is a need to provide tools that would simplify the Internet exploration process as a foundation for decision making for subsequent processes of an internationalisation project.

Increasing competition and globalisation trends are challenging companies to expand the target markets for their products and services in foreign countries. The process of internationalisation necessitates many decisions. Adequate information (e.g. relevant products and companies, or the market situation) about the specific industry niche is required to support decision making and ensure the successful implementation of the internationalisation strategy. The core questions of internationalising SMEe with respect to the target country are:

1. Who is a suitable partner?
2. What are functions that my product has to provide to maximize product success?
3. What regulations do have to be adhered?
4. Is there an interesting market, after all?

Such information is valid only for a particular geographical area and very niche specific. Therefore it is often not available in editorially proven commercial or public databases, but distributed over several homepages of companies, research and governmental institutions, media, and individuals. Moreover the research for the required information consists usually of several individual search steps and requires that one could keep track of the researched information at a later stage too (e.g. when the internationalization of the products is to be carried out).

The business case of Marketing Research and Internationalization for SME represents a challenging application setting for IT-based research tools, because of the following reasons: Firstly, these types of information researches are related with both the business perspective of the company (products, competitors, markets) as well as with the specific domain branch-related information (such as technologies, competitors, trends), thus spanning a heterogeneous field of research questions with a rather broad scope of potential useful information sources. Secondly, a marketing or internationalization research is carried out in an SME only occasionally, therefore research experts are not likely to be found within the enterprise; a proper guidance how to carry out the research is therefore desirable. Another approach for such a research would be the use of external consultants, but in this case the transparency of the research and the proper documentation to the SME that contracts the consultant and carries out the internationalization on the basis of the information of the research is crucial. The consultants, however, are familiar with the research methodology and need only a limited assistance for the search process itself (e.g. concerning domain specific issues), but are working for a number of client, maybe even simultaneously, and need therefore functionalities to organise the researches in a project like manner.

An analysis of several case studies suggested focusing to the following topics:

1. **Identification of products and services** with respect to competitive and complementary products. In both cases the identification of quality functions or customer needs often generated during (in-house) service or consultants. The identification of products that address such issues is an important task. In fact, competitive analysis is an information need identified in all business processes of the project partners. The search for complementary products is transferable, once the complementary functions are identified.
2. **Identification of companies:** Often SMEs look for local partners during internationalisation. Again the key issue is to identify the requirements imposed on companies, or in other words quality functions. The identification of quality functions for companies are again issues to (in-house) services or consultants. In the case of partners, competitors, suppliers, distributors, or customers the problem is transferable, once the key characteristics of the company is identified.
3. **Identification of experts:** The identification of experts in a given domain is a sensitive field. In many cases consultants offer their expertise for money. Scientists or authors often provide highly specialized knowledge, e.g. in their publications (at low cost).

Further important topics are the identification of macro- and micro-economic information, norms, regulations and laws that are in place, preferences due to culture, life-style, fashion, taste and the identification of property rights.

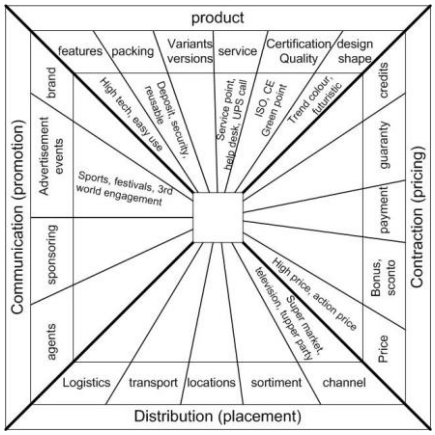

Figure 2: The Marketing Mix Schema, after [4]

Market researches can be carried out by employing a number of different methodologies. Two general methodologies that are supported in the AMI-SME project are the Marketing Mix [4] and the Competition Forces Model [5]. While a methodology like the Marketing Mix does not impose a dedicated process on the research it structures the research into several aspects (see figure 2) that have to be covered and that turn into a sequence of activities during the research project. This relation between a methodology and a process is illustrated in figure 3 for the Marketing Mix method, but can also be applied for the competition forces model or other methodologies. Thus a supportive tool must support the structural dimensions of the underlying methodology while maintaining the freedom for the user to arrange the certain research activities depending on the actual project which could obtain different perspectives on the research (e.g. a product-driven perspective, market-driven perspective or a strategy-driven perspective).

In order to meet this goals the AMI-SME system employs a *dialog structure supported approach* by guiding the user with dialogs that are aligned to the used methodologies (see section 4) and an *information structure supported approach* by providing tailored information models for supporting the information research (see section 3).

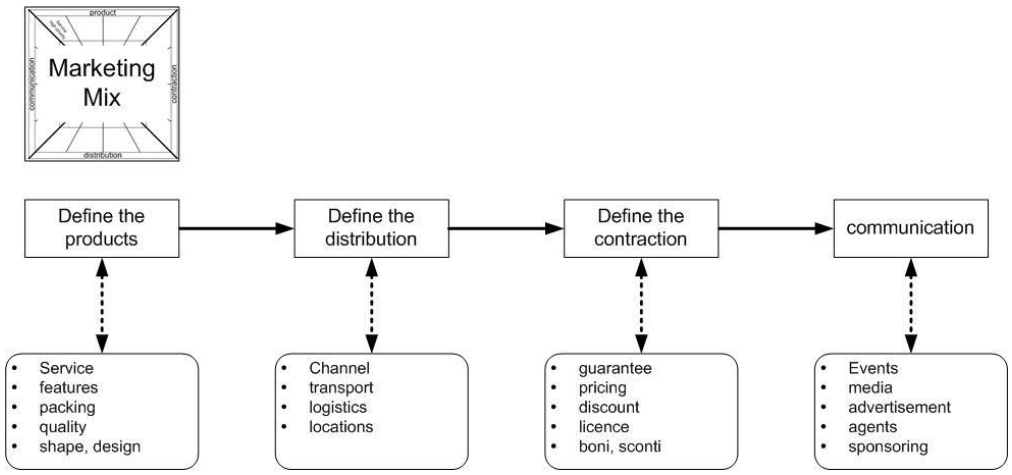

Figure 3: Sequence of tasks emerging from the application of the Marketing Mix Schema

The Business Process for the promotion of high-tech products in new markets at CiaoTech is usually divided into four main actions:

1. Preliminary Analysis
2. Detailed Market Analysis
3. Strategic Planning
4. Business Development

The Processes of Data collection and Analysis are realised during the first two phases of our Business Process (Preliminary Analysis and Detailed Market Analysis), which can be supported by the AMI-SME-tool. Table 1 synthesises the categories of data collected within the CiaoTech Business Process.

2.2. The CiaoTech Consulting Process for Internationalisation

CiaoTech is engaged in business development, fundraising, and technology transfer activities with European and US companies. CiaoTech's mission is that of increasing the competitiveness of high-tech companies by assisting them in Research and Innovation projects, Market Strategies and Planning, and Business Development activities. Here their main consulting processes for SMEs are described.

Information that CiaoTech is interested in for the consulting process, as an example of the application of AMI-SME:

- Information on local laws and regulations
- Information about local support
- Patent analysis
- Information about technical state of the art
- Information on infrastructure limitations
- Information about knowledge (awareness) on technologies and products in the market
- Information on socio-economic situations and trends
- Macro- and micro economic market data
- Information about size and age of substituted investments
- Information about customers, partners, distributors and competitors

It also need to be outlined, that in CiaoTechs experience of working with Italian and US SMEs, data collection and market researches have been always embedded in the overall consultancy process that is provided to access new or local markets.

While developing new products could require a relatively different process with respect to entering new markets with existing product, some data and analysis could be considered similar. In particular, some planning documents (such as Business Plan) which require similar information should be prepared, with evaluation of the new products (or of the new company) with respect to:

- technology State of the Art (including current patents)

- macro economic market data (business environment)
- micro economic market data (competitors, customers, distributors, system integrators)
- Return on Investment

Phase	Data	Description
Preliminary analysis	Local laws and regulations	Data on local laws and regulations in place in the desired region that affect technology, patents analysis
	Existing infrastructure limitations	Data on possible problems related to technology compliance and/or lack of infrastructures required by the product functionality
	Market data	1. Macro socio-economic data (current situation and trends on product end users) 2. Current data and trends on the infrastructures and application environment where the product will operate 3. Current data and trends on specific class of products
	Local knowledge of the new technology/ product and local presence of similar products	Data that could outline the level of knowledge about the product/technology in the local market. Such data could include data on local publications (magazines, papers produced by local Universities, local industries, etc.), local held seminars, etc.
	Size and age of the investments to be substituted by products like that to be promoted	Data extracted by the general data about investments realised in the product sector environment
Detailed market analysis	Analysis of the potential customers	Data collected on the potential customers, which include market dimension (number of potential clients, forecast of investments in the product categories, etc), lists of potential customers, contacts, etc.
	The competitors	Data collected on competitors, their value proposition, price, lists of competitors, etc
	The local support	Data collected on local support: list of distributors, national coverage, system integrators, etc

Table 1: Categories of data collected within the CiaoTech Business Process

Phase	Actions	Data Collection and Analysis
Planning Phase	Preparation of the comprehensive Business Plan	• Technology state of the art (including current patents) • Macro economic market data (business environment) • Micro economic market data (competitors, customers, distributors, system integrators, etc.) In such action sometime Start Ups are assisted by Incubators, Public Development Agencies, or Universities Technology Transfer Department.
Start up Phase	Operative Marketing Actions (contact potential customers, partners, distribution channels, etc.)	• Data on potential customers • Data on potential partners • Data on potential distributors • Data on competitors

Table 2: Data collected and Analysed for Start Ups

In particular with regard to new companies created to exploit potentialities of new products, our experience of working with Start-Ups (CiaoTech is certified tutor of BIC Lazio, and is currently assisting nine High-Tech start ups) is that in the start up process of new high-tech company could be supported by systematic research in the following phases. The AMI-SME tool can be helpful in the research steps listed in table 2.

In cases where new products are particularly innovative and require complex R&D action, the exploitation of the results is a complex matter. Some of the data indicated in the previous tables will be required in order to evaluate the market potential of possible applications of R&D results but also other activities have to be carried out that are beyond the scope of the AMI-SME project.

3. Conceptual Structures for Assisted Search and Document Organization

As mentioned in the introduction we use conceptual structures as a central model to support both the individual search processes as well as the organisation of the findings of the research, which we refer to as Documents

throughout the article. In order to realise this functionality we are using ontologies implement the central conceptual structure. An ontology is understood in this article as a common conceptualization to structure a domain.

The software application distinguishes system ontologies and project ontologies. System ontologies structure the domain of several projects, e.g. internationalisation. Project ontologies are specific for one project, e.g. selling simulation software in France.

If you define a new research project, you select the appropriate system ontology and all changes will be stored to its project specific copy. Thus, each project will have its own project ontology, extended and instantiated by modifications during the usage of the system. As illustrated in figure 1 below, the project ontology is the central element to support the core functionalities:

- **Structuring of relevant topics** is provided by the ability to label search results manually, which allows to find them within the ontology structure easily;
- **Suggestion of new search terms** is available as soon as enough results are labelled manually it is also used to suggest labels for search results automatically; this function will also assist in the definition of queries by topic suggestions related to the concepts and instances of the ontology, e.g. legal issues or distinct products and it also suggests keywords and synonyms as well as relevant Internet pages for concepts and instances.
- **Sorting and Finding of Search Results** allows saving and retrieving the individually extracted information about relevant knowledge items, e.g. about specific competitors or relevant regions;
- **Sorting and Finding of gained knowledge** will be supported by organising the researched information on a conceptual level (e.g. the competitors of a marketing research project that will be represented as instances in the ontology).
- Finally **the improvement and extension of the project ontology** itself will be considered as an important functionality for the user to tailor the conceptual structure according to the need of his or her information research.

Since all these activities relate to the same project specific ontology, wherever you add a new knowledge item (instance), a knowledge type (concept) or a relation, they are immediately available for the other purposes as well.

Currently one system ontology is already designed for and integrated in AMI-SME. It describes important concepts in internationalisation and is used to test the system. That ontology has four main top-level concepts, which are crucial to understand a given situation: product, company, target market and regional area. All of them are further specified with sub-concepts, relations and attributes, but also other concepts such as events and associations exist.

In order to obtain valuable search results it could be advisable to augment such industry independent system ontology with industry or company specific concepts attributes or relations.

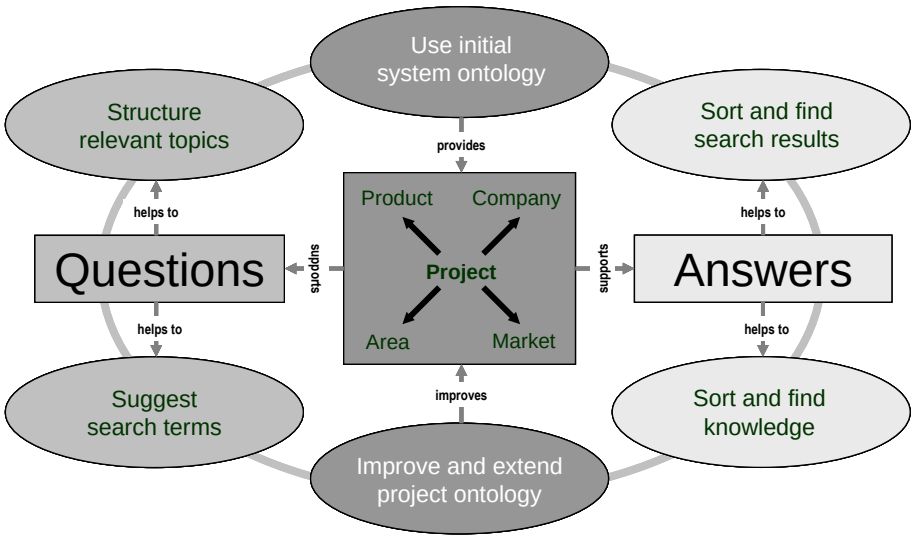

Figure 4: The Project ontology as central element for the main functions for supporting the formulation of search queries and the organisation of the retrieved relevant information.

This follows the approach in other domains, where generic ontologies exist, e.g. for organisational knowledge, that can be instantiated to a specific situation, e.g. to a specific company [6]. In addition it is possible to connect the system ontology with project specific ontologies; for example to specify the concepts “product” or “company”. Moreover, it is possible to define other system ontologies related to a different

domain, e.g. innovation management, and thus use the system for other purposes.

All changes to system ontologies have to be done outside the AMI-SME system, using ontology modelling tools. The sample internationalisation ontology was modelled with the graphically oriented software tool SemTalk [7]. The import of existing domain, industry or region specific ontologies and their use as system ontologies is also possible, as far as they follow some formal restrictions regarding the supported relations and constraints are taken into account.

Within the AMI-SME system, several independent search projects can exist in parallel or successively, e.g. in case of internationalisation for different products or towards different countries. Each user can access the projects that he defined, and where he gets assigned to by the project initiator or the system manager.

4. Description of the Technological Architecture

The core innovation in AMI-SME is the twofold usage of the ontology, in particular to provide means for organizing the local document and knowledge repositories and to support the query definition in order to access the information sources properly. The software-system therefore combines domain specific prestructuring, automatic analysis and manual annotation and structuring functionalities in a single application:

- A persistent storage for different search projects and their related queries and results allows working over a long time span on the same project with different users.
- A complex but intuitive and expandable ontology supports definition of queries and information sources, navigation in researched documents and organisation of the collected knowledge items.
- The integration of text analysis functionalities for clustering, abstracting labelling and filtering helps to keep an overview in the growing repository of information pieces.

The software development environment is based on the existing Web-application development framework ObjectLedge which is provided by Warsaw University and allows reusing of basic components and simplifying the development process [8]. The user interface screens are optimised in order to offer a wide range of intelligent functionalities while hiding the complexity of language processing and ontology manipulation for the user.

The high-level overview on the technical architecture (figure 5) shows the central significance of the ontology services, which is used as a basic service for the organisation of the information search and the found documents as well as the integration of the information analysis subsystem. It is also tailored to the intended application domain and supports the Marketing methods mentioned in section 2. The central position of the ontology as a dedicated information model resembles the information structure supported approach that was followed in the design of the AMI-SME-tool.

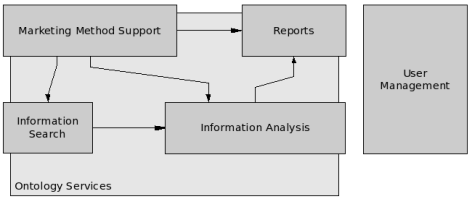

Figure 5: High level overview on the technical Architecture of the AMI-SME system.

As shown in figure 6, the tool is embracing the dialog structure supported approach by implementing several tabs that should guide the user from research projects to searches and from the organisation of results documents to the alignment of conceptual knowledge. In accordance to the conceptual design the interface is structured according to the three main project related views: the assistance of the search-steps during a marketing research, the storage of retrieved documents and the organisation of newly acquired knowledge-items. The following subsections will describe the main functionalities in these software components.

4.1. Search Assistance and Management

Individual searches can be defined for each research-project, they can be executed and saved for later reuse. Search results are automatically tracked and will be stored locally. Thus they are still available, even if the results delivered from the information sources are changing. There are different support levels for the definition of searches in AMI-SME:

- In ad-hoc search, only keywords and a search name have to be entered, and comments to the search are possible. This makes search simple as in widespread (meta) search engines, but saving and many result processing features are available (see figure 6).
- Assisted search gives easy access to the project ontology, which can be used to build more complex queries on the basis of selecting knowledge-items and knowledge-types and to direct the search to relevant Internet sites, e.g. of an association, a magazine or a company.
- Advanced search is based on the standard features of specific search settings in the web service interfaces of search engines (currently Yahoo, Google and A9-OpenSearch); this helps to direct and specify the query, e.g. for domain, language, document types or dates.

The query is translated to the standard corresponding to the selected search engines which are connected to AMI-SME using web services. The implementation of the interfaces to the information sources is done in a modular way, in order to enable the inclusion of other sources at a later stage. Search results received from various search engines are merged and ranked automatically according to relevance.

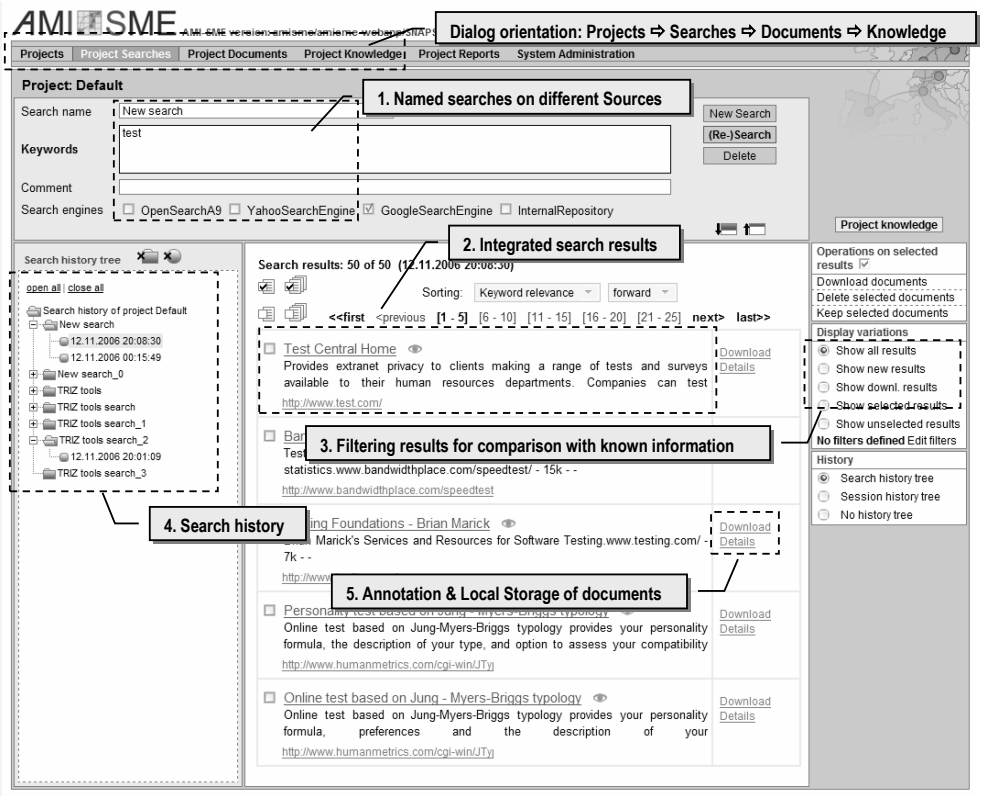

Figure 6: The Project Searches screen of the AMI-SME system, showing the ad-hoc search configuration with the search history tree, the aggregated search results and the access to the filter functionalities

For the results of all search modes, clustering structures the results and filtering searches in the results or its metadata. A search history allows to view and compare the results of existing queries, and to re-execute queries again.

A session history additionally offers a chronological overview of all searches that were executed or tacked in the current session while using the system.

4.2. The Organisation of Search Results in Project Documents

The search results will be stored in database stores as *Search Result Documents (SRD)*. A SRD represents a document, e.g. web page, PDF or MS Word file, etc. The same document only exists once in the system, even if it is found by several searches; this is realised by comparing URL and MD5sum. The uniqueness allows labelling and annotations that are always available, and saves storage volume as well. Moreover, the view on new search results can be limited to new documents.

The documents that the search results are referring to are not downloaded automatically, but manually only, in order to save storage resources and to avoid download time. If downloaded, they are accessible for more detailed analysis, such as automatic abstract generation, or text analysis algorithms including clustering, classification and name entity extraction.

A project document screen allows the user to concentrate on the evaluation of the results of all project related searches. Here all results for the whole project are available, but the view can be restricted to the results with a specific ontology label using the hierarchical view on the ontological information model.

For each result, a details pop-up screen (see figure 7) for annotation is available, as well from the search screen as from the project document screen. These document details contain partly system suggested values, e.g. by a summariser, language detector, or metadata extracted from the search engine (these results are displayed with a grey background).

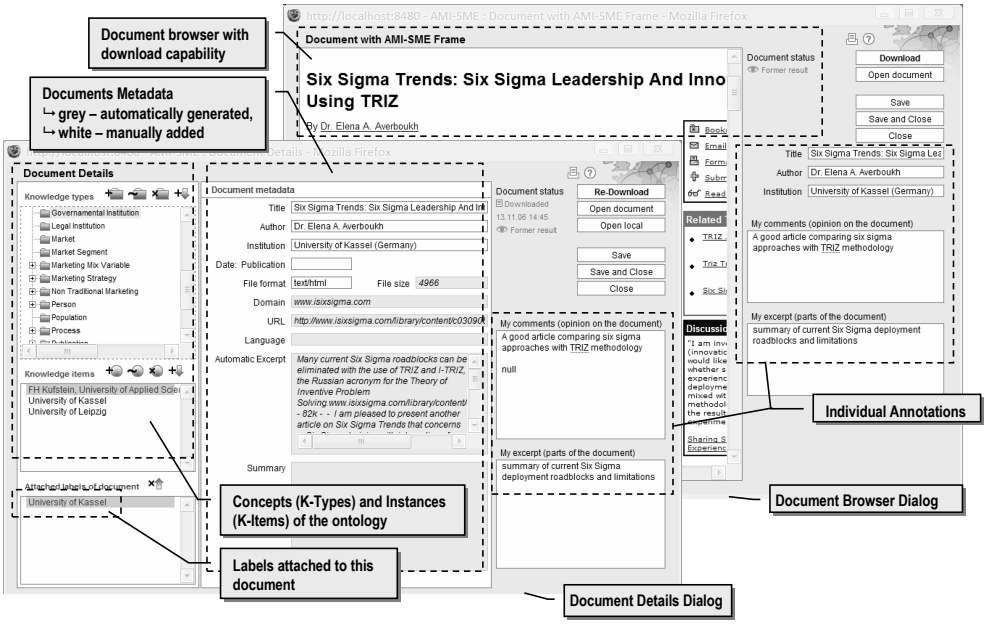

Figure 7: Screenshot of the Document details Dialog (left part in the foreground) for labelling the document and viewing all the meta data and the Document Browsing Dialog (right part in the background) for viewing a document from the search results within the context of the AMI-SME system with the ability to provide individual annotations.

The result document can also be labelled with concepts (Knowledge-Types) or instances (Knowledge-Items) of the project ontology; in addition the labelling of all selected documents is possible. A document that has been found during the research can also be opened in a Document Window, which enables viewing of the document content as it would appear on the web and provides also a limited amount of annotation functionalities (see the dialog in the background in figure 7).

4.3. Organizing Conceptual Information in the Project Knowledge Repository

Relevant conceptual knowledge that is identified from the SRDs can be *directly* stored in the project specific knowledge base, which is identical to its project ontology: for example, competitors are instances of companies, and they may have values to attributes such as the name of the owner of a company. In this way, the user is able to extend the ontology in an easy and comfortable way by expressing the knowledge that he has gained during the market research project. The granularity of this information is different to that on the information level since it focuses on concepts (a certain competitor in a market), their relations (the products that the company might provide and the customers that use it) and some attributes (like contact information or annual turnover) that can be stored and retrieved in a more structured way.

A project knowledge screen (see figure 8) allows editing values to specific instances of the project ontology, e.g. the number of a company's employees, or adding relations, e.g. from a company to its products. The project knowledge dialog is directly available from the other main screens to view and edit information about identified knowledge items.

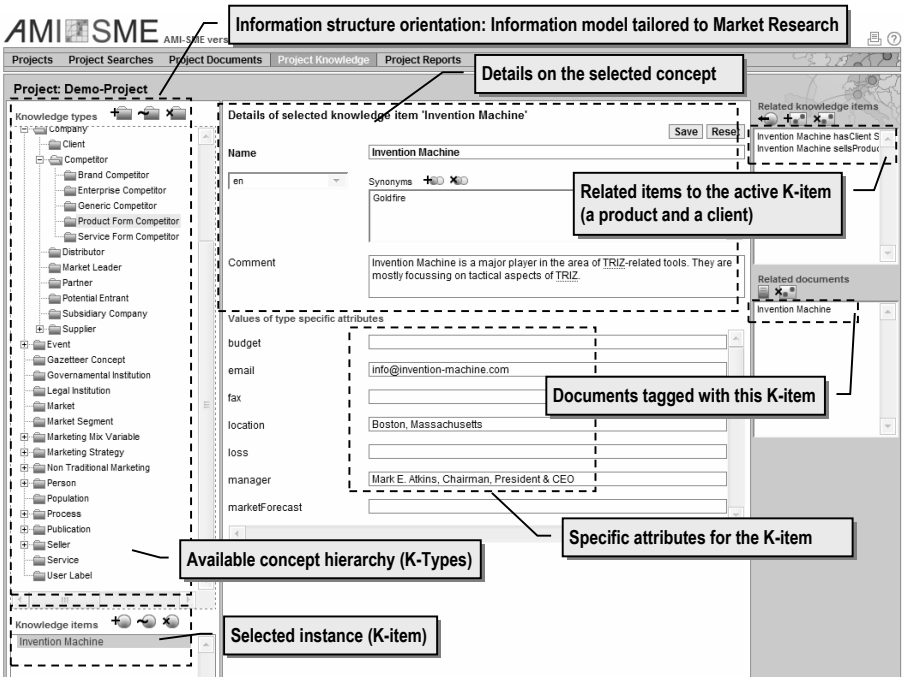

Figure 8: The project knowledge screen of the AMI-SME system enables the direct interaction with the underlying information model and allows the capturing of conceptual knowledge that has been gained during the research, which is expressed by Knowledge-types and Knowledge-items (left column), the values of their attributes (middle column) and their relation to other K-Items and documents (right column).

In the project knowledge screen, but also in project details and in the project documents, the concepts of the project specific ontology are displayed comparable to a folder structure in form of a tree in the left column, with the instances of the selected concept listed below. In the knowledge screen only (figure 8) also relations to other knowledge items and to labelled documents are presented in two separated lists in the right column; and also related documents are linked. This allows a well-known (hierarchical) navigation, despite of the more complex structure of the project ontology itself (which resembles a graph-structure).

5. Related work

Since the advent of Web the problem of exploratory search was addressed in many papers, albeit not sufficiently reflected in practical applications. Its growing meaning has been recently confirmed by a number of publications collected in [9], where the issue has been exhaustively revisited. Moreover, the focus of many approaches was on the retrieval side of exploratory search and not about organisation of findings which was often done in other IT-Systems such as local file-systems, DMS or CMS-System. We have attempted to implement AMI-SME in line with the idea expressed in [9] and paraphrasing the famous Hamming statement to the form "the purpose of exploratory search is insight, not data" [10]: "In intelligence analysis, as in other domains, that insight comes from the *process of exploration*, not just from its end result. We are interested in capturing and visually representing the iterative query processes and insights of the analyst to help them collect and compare information more effectively, as well as record and share the products of their analytic insights."

To an extent our approach refers to the idea of web farming systems, as defined by Hackathorn in [11] as "…the systematic refining of information resources on the Web for business intelligence", although the focus of our solution is not tailored to business intelligence, but the project specific information research and organisation to fulfil a dedicated information need. On the other hand it should provide means for quick exploration of new web areas. Until now, there were very few attempts to practically implement it, especially for the scale of SME needs. In this sense, AMI-SME is quite a unique system combining Web exploring functionality with advanced storage system and text analysis means. One of the main features of AMI-SME is building a repository by a sequence of consecutive queries. A similar approach has been implemented with the system SenseMaker, presented in [12] which already uses clustering techniques for visualizing the search results. Another approach that seems to be close to AMI-SME was INSYDER reported in [13]. The similarity resulted from the similar goals of gathering the information sources; however the proposed solutions referred mainly to using web agents, whereas in AMI-SME we expect to tap high quality information from the existing web search engines by a simultaneous search covering a number of existing engines. So instead of concentrating on building agents, we have put more attention to developing means for cumulating local repositories, on which advanced text

processing and text mining tools can be applied for knowledge extraction.

While there are a number of solutions available for assisting during the individual search steps there are not many for approaches for organising the research findings using a conceptual structure that was also applied to assist the search process. There are a number of solutions to organise web resources in bookmark like systems, e.g. BlueOrganizer [14], some even follow a collaborative approach such as del.icio.us, (see [15]) but most of them only keep track of the link to the resource, which is not sufficient for documenting research findings, as the original resource might disappear or become unavailable. New application allow to clip parts of Web-Resources (see ClipMarks, [16]), but again this Web application is implemented as a public collaborative service, which is not appropriate for the business professional to support the addressed business case of Marketing Research and internationalization. Finally the alignment of the search result organisation with a powerful conceptual knowledge representation has not been used in existing systems is a unique feature of the AMI-SME system.

6. Conclusion and Outlook

The realised concept of ontology based search and storage improves the interface between search engines and applied Knowledge Management: it makes Internet search more intelligent and integrates it closely to the working context of the users in an information research project. Especially for tasks that companies seldom conduct, like internationalising, such a guiding structure helps not to forget important issues while conducting the research. This is provided by tailoring the internal information model to the needs of the underlying research methodology. In case of weakly structured information, as for individual industrial niches, that is even more important since the reader has the effort to extract relevant content on his own and is now supported to store results in a proven way.

The general idea can be transferred to other topics than internationalisation, e.g. new product development or innovation management, which is currently under evaluation for another research project [17]. Also the automatic extraction of available content from the Internet to the knowledge base is a research activity that would additionally extend the usage of the software, either by integration of relevant external RDF sources (e.g. the World Factbook) which contains a wide range of useful information about countries, see [18] or with specific wrappers, like in the PiggyBank-project [19]. During the exploitation in other domains, the project partners will offer additional services to the users of the software, such as ontology modelling, system integration, design adaptation, training as well as support in the interpretation of the search results.

Acknowledgements

The described IT-solution is being developed within an EU-supported CRAFT project "AMI-SME: Analysis of Marketing Information for Small And Medium-Sized Enterprises" (Contract Nr. 017566) by six RTD partners and seven SME partners from five European countries. More information on the project can be found here: http://www.ami-sme.org.

References

[1] Strohmaier, M.: B-KIDE: A Framework and a Tool for Business Process Oriented Knowledge Infrastructure Development, Shaker, 2005.

[2] Uwe V. Riss, U. V.; Rickayzen, V.; Maus, H.; van der Aalst, Wil M. P.: Challenges for Business Process and Task ManagementJournal of Universal Knowledge Management, vol. 0, no. 2 (2005), 77-100

[3] Harald Holz, H.; Maus, H.; Bernardi, A.; Rostanin, O.: From Lightweight, Proactive Information Delivery to Business Process-Oriented Knowledge Management. Journal of Universal Knowledge Management, vol. 0, no. 2 (2005), 101-127

[4] Kotler, P.: Marketing Management, Prentice Hall, 2005

[5] Porter, E. Competitive Advantage. Free Press, 2004.

[6] Gualteri A., Ruffolo M. (2005) "An Ontology-Based Framework for Representing Organizational Knowledge", In: Proceedings of I-Know 05, Graz, Austria, June 29- Juli 1, 2005, pages 71-78.

[7] Fillies, C.; Weichhardt, F.; Smith, B. (2005) "Semantically correct Visio Drawings", Proceedings of the 2nd European Semantic Web Conference, Heraklion, Crete, May 29 – June 1, 2005.

[8] Caltha (2006) Objectledge, [online], http://objectledge.org

[9] Gersh j.;Lewis, B.; Montemayor, J.; Piatko, Ch.; Turner, R.: Supporting exploratory search: Supporting insight-based information exploration in intelligence analysis, Communications of the ACM, Vol. 49(4), 2006

[10] Hamming, R.: The Art of Doing Science and Engineering: Learning to Learn. CRC Press., 1997

[11] Hackathorn, R. : Web Farming for the Data Warehouse, Morgan Kaufman, 1999.

[12] Wang, M. Q., Baldonado, Winograd, T.: "SenseMaker: An Information-Exploration Interface Supporting the Contextual Evolution of a User's Interests", Proc. of the SIGCHI Conf. on Human Factors in Computing Systems, 1997.

[13] Reiterer, H., Müller, G., Mann, T. M., Handschuh, S.: INSYDER - An Information Assistant for Business Intelligence", Proc. of the 23rd Annual Int'l ACM SIGIR Conference, 2000

[14] adaptiveblue: BlueOrganizer [online] http://adaptiveblue.com/about.html

[15] Golder, S.; Huberman, B. A.: The Structure of Collaborative Tagging Systems, [online] http://arxiv.org/abs/cs.DL/0508082

[16] Clipmarks [online] http://clipmarks.com/how-to-clip

[17] KNOW-IT: Supporting Innovation Process Development [online] http://www.know-it-project.eu

[18] CIA "The World Factbook" [online] http://www.cia.gov/cia/publications/factbook

[19] Huynh, D., Mazzocchi, S., Karger. D (2005) Piggy Bank: Experience the Semantic Web Inside Your Web Browser. Proc. of the International Semantic Web Conference 2005, 2005.