About an Architecture for Integrated Content-Based Enterprise Search

Manfred Grauer, Ulf Müller, Daniel Metz, Sachin S. Karadgi, Walter Schäfer
Information Systems Institute, University of Siegen, Germany
Email: {grauer, mueller, metz, karadgi, jonas}@fb5.uni-siegen.de

***Abstract* – The main goal of knowledge management is to improve the management of information and knowledge within and across enterprises. Enterprise knowledge is embedded in enterprise's data managed in a wide range of information systems (e.g., product data management, enterprise resource planning systems). These enterprise data can correspond to textual, numerical or multimedia. In the past, several search systems have been developed to provide access to primarily text-based enterprise content like web pages, data stored in database systems or emails. Similarly, dedicated search systems exist for searching information from multimedia data. As these search systems focus on particular enterprise content (e.g., textual data), they lack in providing a holistic view on available enterprise's knowledge. Therefore in the current contribution, architecture for integrated content-based enterprise search encompassing various enterprise data sources is elaborated. This architecture supports the exploitation of embedded enterprise knowledge. Further, the architecture is validated in an industrial scenario.**

Keywords* – *enterprise search; knowledge management; information retrieval; shape matching; image-based retrieval.

I. INTRODUCTION

Alaavi and Leidner define "data as raw numbers and facts, information as processed data, and knowledge as authenticated information" [1]. Knowledge is embedded into enterprise processes by enterprise members and described through various data types as illustrated in Fig. 1 (e.g., documents, 3D models, emails). This knowledge is enriched and enlarged as the enterprise process moves from upstream to downstream processes. Hence, necessitates for organizing and managing enterprise data in a manner that it can be identified quickly and assimilated by enterprise members for design and execution of enterprise processes.

In engineering, 3D computer-aided design (CAD) models, 2D drawings and design patterns have to be created during product development process [2]. Information and library sciences are managing text-based documents like books, journals and magazines. Order information and its corresponding offer documents, bill of material (BOM), email interaction and minutes of meeting have to be managed by sales department. Patient records have to be managed by hospitals. Due to the aforementioned diversity of applications, specialized IT-systems have been developed to manage different types of enterprise data (e.g., enterprise resource planning (ERP), customer relationship management (CRM), product data management (PDM) or clinical information systems (CIS)). Each of these IT-systems contains a large amount of enterprise data related to different enterprise entity types (e.g., products, customers, patients).

Many of the enterprises utilize aforesaid IT-systems in various combinations. Therefore, the IT landscape of an enterprise is often scattered and inadequately integrated i.e., enterprises often need to address data interoperability issues. As a consequence, enterprise members have to perform multiple searches for information and corresponding knowledge related to an enterprise entity in various IT-systems. In majority of the research, search system functionality is achieved by means of organising and retrieving structured and unstructured textual data from a certain IT-system [3]. Also, multimedia data (e.g., images, videos, sounds) is rarely considered, and only special research areas deal with organizing and retrieving of multimedia objects [4][5].

Overall, lack of integrated data from different IT-systems hinders establishing holistic view about the requested enterprise entity. To overcome these restrictions and drawbacks, enterprise search (ES) has emerged as a promising research area. ES is the practice of making enterprise content from numerous IT-systems, such as databases and intranets, searchable to certain stakeholders [3][6]. The vision is that an enterprise member is capable to request all necessary information and associated knowledge with minimum effort from various IT-systems, which is required to effectively design and execute enterprise processes. Therefore, an ES engine has to be provided to enterprise member incorporating enterprise data from different IT-systems.

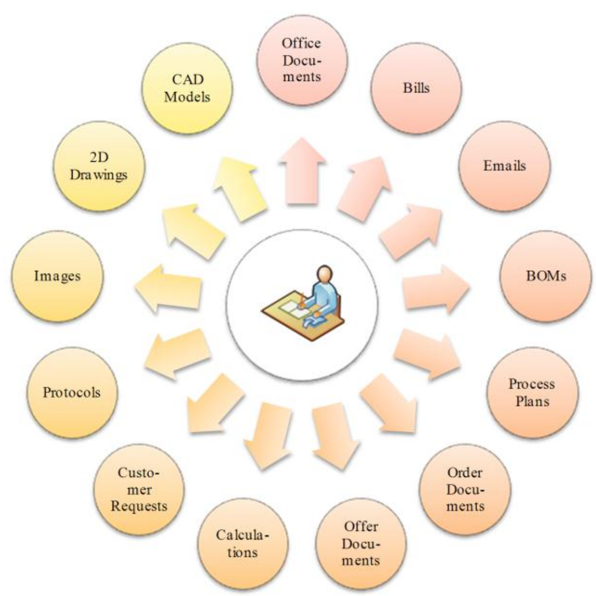

Figure 1. Different data sources and document types available to enterprise members

To address the aforementioned requirements of ES, architecture for integrated content-based enterprise search (CBES) is presented. The contribution is structured as follows. An overview of current research trends in ES and content-based retrieval is presented in section II. Section III motivates CBES, and defines challenges and requirements associated with CBES. The envisaged architecture and its components are elaborated in section IV. To validate the envisaged architecture, Section V describes an industrial case study. Finally, conclusion is presented in Section VI.

II. Related Work

The main task of ES is to improve information and knowledge management, and facilitate information access within and across enterprises [7]. ES allows enterprise members to search through documents, emails and other data sources available in an enterprise to identify information and associated knowledge [7]. Enterprise members spend one third of their work time searching for information necessary for designing and executing enterprise processes [8]. Therefore, it is not surprising that the investment in development of comprehensive ES during the last years has increased and major companies like Google, Microsoft, IBM or SAP have developed specific and proprietary ES-systems [9][10].

Data in an enterprise is primarily textual, in the form of (intranet) web pages, emails, orders details, bills, reports, and so forth (s. Fig. 1). Hence, majority of the ES implementations are focused on textual data [3][4][8]. An early approach of ES has been named as an enterprise intelligent system [11]. This system tries to integrate data from intranet servers, web servers and web search services. The usage of different search techniques (e.g., keyword search, semantic search) were analysed, and utilized in search of information from integrated vehicle health management data [12]. Similarly, ontologies were used to overcome the limitations of textual retrieval [13]. Further, an engineering information retrieval system has been elaborated which attempts to integrate data from various engineering data sources into a centralized dataset and make it accessible via ontology-based retrieval [14].

A search system for identifying information from huge collection of documents from digital libraries was presented [15]. The effectiveness of this search system was increased by exploiting techniques like full text search, collaborative filtering and multifaceted browsing. A search architecture was elaborated which employed the search functionalities of the source systems to locate relevant information [16]. Retrieved results are combined together in a global result list using case-based reasoning (CBR). Further, CBR was used for ranking the results in a result list.

Apart from textual data, nowadays multimedia data is widely used in enterprises. Multimedia data include images, photos, videos and sound recordings, among other. For instance, meeting involving enterprise members from multiple sites are recorded. Most of the available search systems have not integrated multimedia data along with the textual data. However, research has been carried out in isolation to search various multimedia data.

A survey of image retrieval techniques and systems were presented [17]. Image retrieval techniques were classified into two groups: text-based and content-based retrieval. Both groups deal with the same issue to overcome the gap between what is really experienced by the enterprise member when viewing the content on screen and what is really stored (e.g., in a database). This challenge is not only relevant for image retrieval but also for other kinds of multimedia data retrieval.

Several IT-systems have been developed for 3D model retrieval based on textual annotations, content, features and ontologies, and so forth [14][18][19]. All the presented approaches show that it is possible to search 3D content in large repositories. Another field in the context of multimedia data retrieval is face recognition. This field is relevant for identification of faces displayed on an image or in a video, and it has become more important from (public) security point of view. The state-of-the-art of face recognition techniques was carried out [20]. Besides images and photos, retrieval for other multimedia data is available (e.g., video, sound). Retrieval systems for these kinds of data have been developed and encouraging results have been achieved. A summary of video retrieval techniques can be found in [21] and survey of sound retrieval techniques is presented in [22].

The review of the related work has shown that effective techniques for content-based retrieval exist and dedicated systems for specific scenarios have been developed. However, an architecture which allows the incorporation of all aforementioned enterprise data types from different data sources could not be identified.

III. Challenges and Requirements for an Architecture for Integrated CBES

As already mentioned, most data in an enterprise is text-based like reports, emails, order details, and bills (s. Fig. 1). This obviously raises the question: why is it necessary to consider multimedia data and integrate it into an ES-strategy? The main reason is the increasing relevance for utilizing multimedia data during enterprise process design and execution. For example, design engineers and architects use 3D models to store information about design and construction of products or buildings. Respectively, images and drawings are used to identify and describe products, and voice recordings and videos are used to protocol meetings.

In current search environments, collections of multimedia data are managed by assigning textual annotations (i.e., textual metadata) which can be employed for retrieval [4]. Usually, enterprise members from different departments contribute during design and execution of enterprise processes. Different annotations are assigned to the same multimedia data depending on the executed process step and enterprise members' background. Therefore, knowledge managed using aforementioned metadata-based techniques places limitations during ES. In contrary to this metadata approach and to ensure that relevant product-related information and knowledge (e.g., embedded in a bill or 3D models) can be retrieved from different departments of an enterprise, it is mandatory to integrate different annotations, attributes from various IT-systems as well as content from

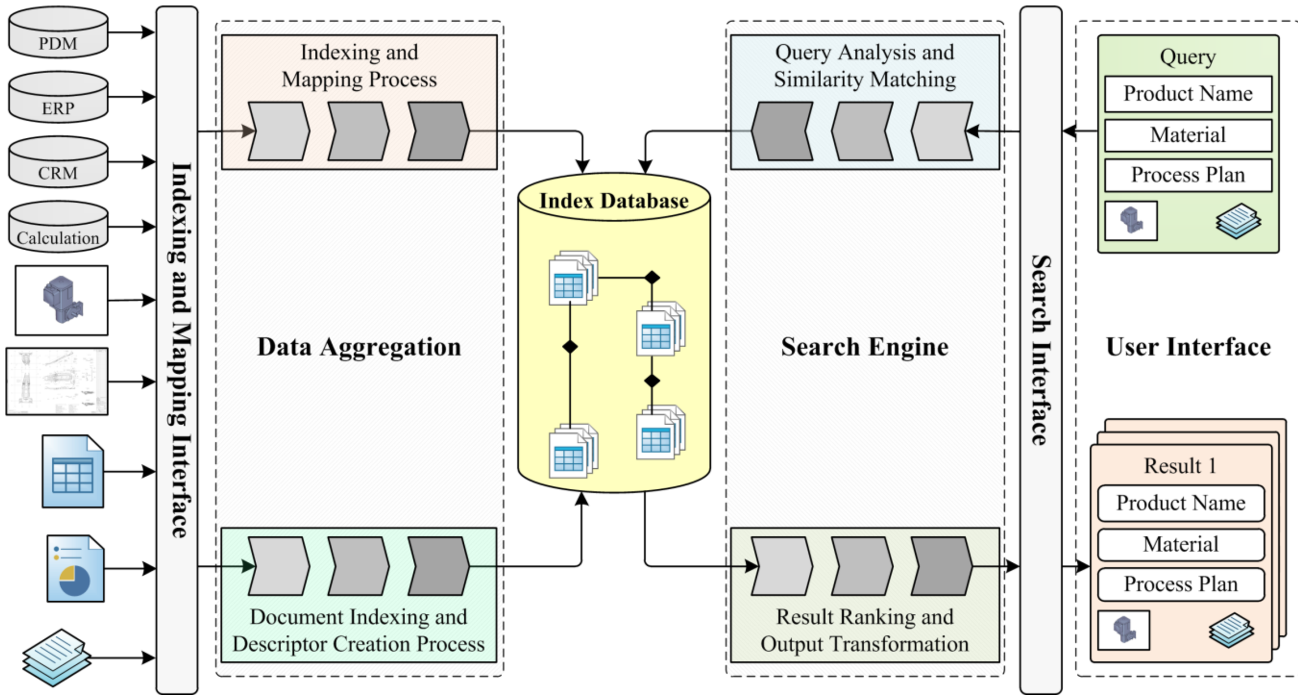

Figure 2. Architecture of integrated content-based enterprise search (CBES)

multimedia data. The overall goal is to create an integrated descriptor for retrieving relevant enterprise process entities (e.g., orders, bills, 3D models). In addition, the envisaged ES has to fulfill certain requirements which can be classified into four groups: (i) coverage, (ii) security, (iii) query support, and (iv) presentation of results.

Coverage in the context of a holistic ES means that all relevant IT-systems and data sources are included into the enterprise search. The challenge is to integrate information and knowledge which is scattered across an enterprise. Besides information in IT-systems like CRM or ERP systems, information and knowledge can also be available in emails, text documents, 3D models and images, and so forth.

Security has gained considerable attention in the context of ES due to tightened legal requirements and enterprise's policies related to data privacy and confidentiality. Obviously, only privileged enterprise members should have access to enterprise data. An enterprise member should have access grants to relevant data, information and knowledge depending upon his position within a department and the activities he/she has to fulfill.

Apart from the scope of the ES and its security concepts, user-friendly and convenient search interfaces have to be provided to the enterprise members. The search interfaces have to respect an enterprise member's roles and privileges. For instance, an engineer will demand a dedicated search interface to search for 3D models where as sales department personnel will have a graphical user interface (GUI) to explore previous offers and orders. In either case, a search interface assists an enterprise member to input suitable number of search attributes. Based on these search attributes, appropriate similarity search techniques will be employed to construct the result list (e.g., containing 3D models or orders). In case of multimedia data, this can mean that functionality need to be provided where user can submit a template (e.g., 2D drawing of a product) to the search engine for requesting product related information and knowledge.

Last but not least, adequate presentation of the search result plays a prominent role. The retrieval system has to provide appropriate interfaces to display the result list, especially in case of multimedia data. For example, the ES-system provides a preview window for visualizing a 3D model of a selected result item. In addition, access to information and knowledge associated with the selected result item has to be implemented. Hence, the ES-system establishes links to the original data managed in one of the enterprise IT-systems.

IV. ARCHITECTURE FOR INTEGRATED CBES

As mentioned before, enterprise data is stored in different IT-systems (e.g., PDM system, ERP system). In addition, each data source has its own specialized data structures. For instance, a PDM system is used to manage product-related data created or modified along a product's life cycle [2]. The proposed process model for an integrated ES is composed of the following steps: (i) analysis of enterprise processes concerning their status and requirements regarding ES, (ii) create a centralized search index database with attributes and content-based descriptors revealed from enterprise data, (iii) engage search index to retrieve enterprise members' requested information, and (iv) evaluate, update and optimize the integrated ES system (i.e., especially the search index database).

Before implementing an ES system, it is mandatory to identify enterprise knowledge required by the enterprise members. Enterprise knowledge is used and shared by enterprise members along enterprise process execution. The knowledge flows within a certain enterprise process can be (re-)designed and analyzed employing the modeling and description language (KMDL) [23]. The identified characteristics of knowledge creation and assimilation assist the establishment of an appropriate search index database. In addition, the data sources required to execute enterprise processes have to be documented (e.g., with unified modeling language (UML)).

Based on the aforementioned (process) analysis, an appropriate index database has to be defined. This index database contains attributes and content-based descriptors related to enterprise process entities (e.g., orders, offers). Data mining techniques and structured interviews can be employed to determine relevant attributes and required content-based descriptors. The attributes and descriptors are derived from enterprise data which is managed in different enterprise IT-systems. Each entry in the index database points to the data sources.

Analysis of enterprise processes and setup of index database guides in definition of GUI and implementation of ES engine. GUI is utilized to submit search request to ES engine and display the search result list. Today’s enterprise processes are agile which requires periodical evaluation, refinement and optimization of index database and the corresponding ES-system. Therefore, enterprise process analysis has to be performed systematically to reflect the actual situation of an enterprise. Based on the envisaged process model for an integrated ES, architecture is elaborated in the following sub-sections.

The architectural overview of integrated CBES is illustrated in Fig. 2. The architecture is composed of four components: (i) data aggregation interfaces for indexing data sources and file types available in an enterprise, (ii) index database for textual and numerical attributes from the IT-systems and extract descriptors from multimedia data, (iii) ES engine to support integrated CBES, and (iv) GUI to the enterprise members based on their roles and privileges.

A. Data Aggregation

The data aggregation component provides functionalities to index and map enterprise data from different IT-systems, as depicted in Fig. 2. This component consists of two sub-components: (i) mapping and indexing of attributes from different IT-systems like ERP-, PDM- or CRM- systems to the columns of the predefined index database, and (ii) services for deriving content-based descriptors for multimedia data (e.g., images, 3D CAD models).

Mapping is used to establish a link between data fields of a source system (e.g., managed in a database or XML-structured files) to corresponding columns of the index database. Depending upon the data type of a source field (e.g., numerical, alphanumerical) dedicated indexing methods are employed to optimize the efficiency of the subsequent retrieval process. For instance, a text analysis and text aggregation component is available to extract the content from text documents. If a new value gets stored or an existing value gets updated in a certain source IT-system, the mapping functionality transfers this value to the appropriate column in the index database.

In the case of multimedia data, services are provided to extract relevant metadata from the corresponding file. The retrieved metadata or the calculated content-based descriptor is also stored in the index database. For each file type, a dedicated service is available which extracts the metadata and / or creates a content-based descriptor. For instance, a 3D model of a product is represented as a triangulated surface and stored as a stereolithography (STL) file. This file can be transformed into a content-based descriptor that describes the shape and dimensions of the product. This transformation process can be delineated in three steps (s. Fig. 3): (i) normalisation of the 3D shape essential to avoid problems resulting from translation, rotation and scaling invariance, (ii) creation of images from different perspectives of the normalized 3D model, and (iii) employ various algorithms on the taken images to create image-based descriptors (e.g., edge histogram).

B. Index Database

The index database contains an entry for each enterprise process entity of interest (e.g., order). The entries in the index database represent the enterprise process entities with a number of (alpha-) numerical attributes and if required, geometrical descriptors (s. Fig. 3). These attributes and descriptors are generated by the data aggregation component.

The stored data in the index database has to be indexed for efficient retrieval of enterprise process entities. Therefore, simple numerical and alphanumerical data types are indexed using the available indexing capabilities of relational databases. However, these indexing capabilities are not applicable for complex shape descriptors. Hence, feature space indexing structures (e.g., R*-tree) or metric indexing structures (e.g., M-tree) are available for indexing geometrical (shape-based) descriptors [24][25][26]. These indexing structures have been successfully used for retrieval of complex sheet metal components [27].

Obviously, index database and the available data sources

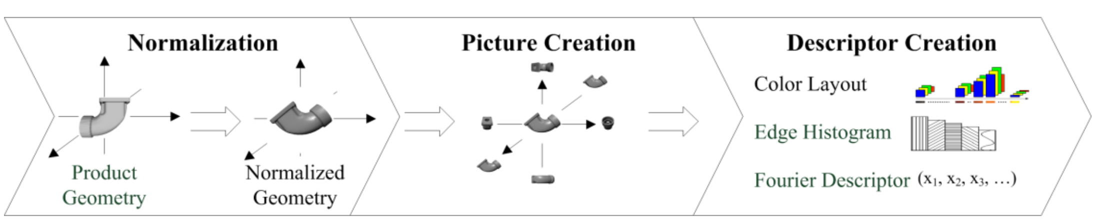

Figure 3. Illustration of image-based descriptor creation process for 3D models

ISBN: 978-1-61208-121-2

of the enterprise's IT-systems (e.g., PDM system) have to be synchronised. Different methodologies are applicable to guarantee that data sources and index database are in synchronous state. Based on the enterprise's IT strategy, these methodologies can be classified as following: (i) update the index database just-in-time, i.e., when the original data in an enterprise's IT-system has been modified, (ii) employ a batch process to update the index database (e.g., update every night), and (iii) combination of (i) and (ii).

C. *Search Engine*

The search engine of integrated CBES implements core functionality to compare and retrieve stored enterprise process entities from the index database. Hence, an enterprise member specifies a request by defining necessary (alpha-) numerical attributes. In addition to these attributes, the enterprise member might provide template objects like 3D models or 2D sketches to describe the desired enterprise process entity (e.g., product). In short, the enterprise member submits the aforementioned information to the search engine for retrieving most similar enterprise process entities.

For retrieving most similar enterprise process entities, different similarity and distance metrics for various data types and content-based descriptors have been developed. Levenshtein distance and Smith-Waterman similarity measures are used for comparing alphanumerical data [28]. Minkowski distance or its special form i.e., Euclidean distance are used to evaluate the distance between numerical data. In addition, Hamming distance can be used to estimate the distance between two input values of equal length for numerical, binary, or string. To assess structural similarities (e.g., BOM), a graph-based distance measure has been applied [29]. To retrieve multimedia items like 3D shapes or 2D drawings, special techniques are employed. For instance, to measure shape similarity various techniques are available which can be classified into feature-based, graph-based and geometry-/image-based [18]. For content-based image retrieval, techniques from the MPEG-7 standard are utilized to detect color; texture and 2D shape similarities [30].

Enterprise process entities in the index database are added to the search result list if their calculated similarity or distance fulfills a predefined similarity threshold. The threshold value is defined by the enterprise member. This threshold value specifies the accuracy of the retrieval process. The search result list is ordered by the similarity values and returned to the enterprise member. To improve the performance of the retrieval process, the search process starts with simple attributes and descriptors, and proceeds with complex ones. For example, a user searches for a product by providing product name, 3D model and product BOM. Initially, product name and 3D model are utilized to determine a relevant subset of the data available in index database. Next, the subset of the data is revised with the structure based (e.g., BOM) search.

D. *Graphical User Interfaces*

The graphical user interfaces (GUI) functionality is twofold: (i) provide necessary forms for specification and submission of search requests (s. Fig. 4) and (ii) visualization

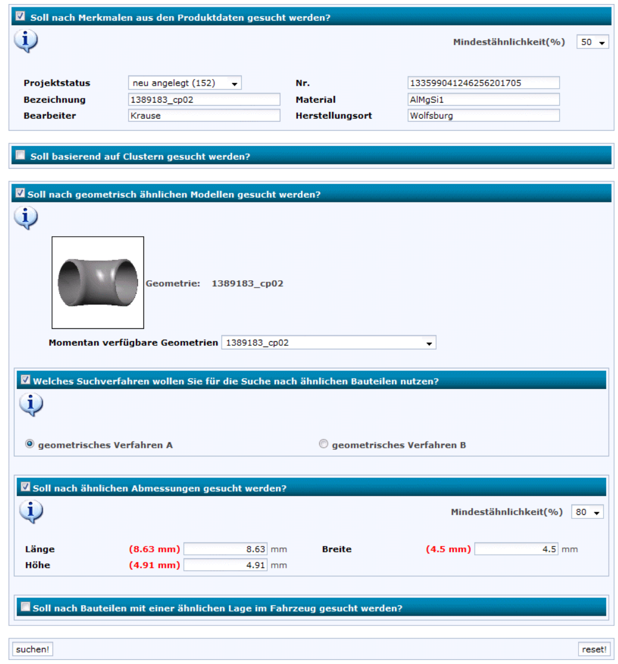

Figure 4. Illustration of an search interface for sales personnel searching for products using metadata and 3D model

of search results (s. Fig. 5). The aforementioned functionality of the GUI has to consider the roles and privileges of the enterprise members. Overall, this ensures the adherence of enterprise's IT policies and the filtering of the information and knowledge according to the enterprise members' needs.

V. CASE STUDY – INTEGRATED CBES FOR COST ESTIMATION IN AUTOMOTIVE SUPPLIER INDUSTRY

The proposed architecture has been validated in industrial scenarios with emphasis on cost estimation processes of automotive suppliers [31][32]. Usually, cost estimation process starts with a request by the customer for information from the supplier about cost, manufacturing techniques and delivery date of the requested component as illustrated in Fig. 6. At supplier's side, experts have to check the component feasibility and determine the requested information (e.g., production cost). Hence, knowledge about

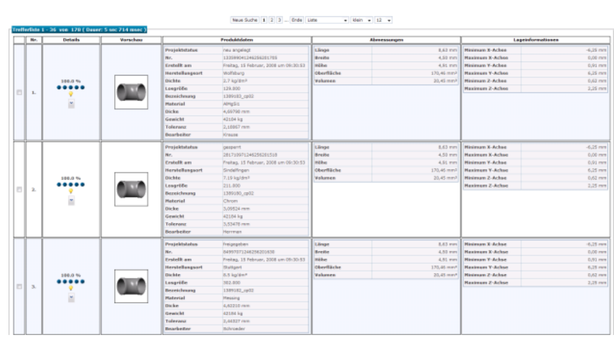

Figure 5. Result page for sales personnel showing similar products corresponding to search request submitted as shown in Fig. 4

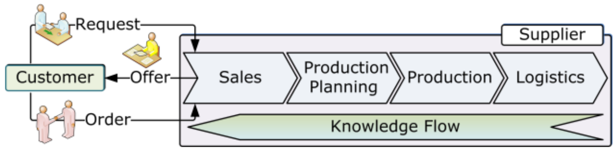

Figure 6. Typical customer supplier interaction in automotive industry

manufacturing techniques, materials and logistics are essential to perform the aforesaid tasks. Traditionally, experts from different departments support to determine the information. The data generated and used during the cost estimation process is not only textual but also multimedia-based. For example, a customer request often contains 2D drawings and / or 3D models of the requested component to be manufactured.

In the aforementioned scenario, it was determined that the existing search capabilities provided by the enterprises' IT-systems are insufficient to fulfill the requirements of an expert to identify information relevant to the cost estimation process (e.g., previously stored orders, process plans). To overcome these shortcomings, integrated CBES was developed which incorporates data from various IT-systems for previously executed offers, orders, 2D drawings and 3D models. Textual and numerical information along with multimedia information extracted from 3D models and 2D drawings were stored in the index database of the integrated CBES. This provides access to a search space containing data from various enterprise's IT-systems. To support different phases of the cost estimation process, several search interfaces have been designed. According to an enterprise member's roles and privileges, the interfaces grant access to different areas of the search space.

VI. Conclusions

In this contribution, architecture has been elaborated to provide access to enterprise's data sources through an integrated content-based enterprise search (CBES). The envisaged architecture bridges the gap between standard ES technology which is mostly focused on textual data and dedicated retrieval systems for multimedia data. The architecture consists of four components: (i) data aggregation contains techniques for indexing and descriptor creation, (ii) index database stores previously predefined descriptors, (iii) search engine implements numerous retrieval techniques, and (iv) GUI exists to submit search requests and view result lists. The developed architecture has been used in several enterprises. These enterprises are operating in different domains like automotive supplier industries and industrial chain manufacturers. The integrated CBES supports domain experts during design and execution of enterprise processes.

The effort to create an index database, implement required interfaces and integrate the required techniques for integrated CBES induces efforts (e.g., time, cost). However, the benefits of a quick and reliable access to information and associated knowledge contained in enterprises data are essential to retain an enterprise's competitive advantages. Overall, the integrated CBES improves the design and execution of enterprise processes. The throughput of enterprise processes can be increased significantly and at the same time, the quality of the enterprise processes' output can be strengthened.

Acknowledgment

The authors would like to thank their industrial partners LINDE + WIEMANN GmbH KG, Germany, Fischer & Kaufmann GmbH & Co. KG, Germany and KettenWulf Betriebs GmbH, Germany for the opportunity to implement and evaluate the elaborated architecture.

References

[1] M. Alaavi and D. Leidner, "Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues," MIS Quarterly, vol. 25, no. 1, 2005, pp. 107-136.

[2] J. Stark, Product Lifecycle Management: 21st Century Paradigm for Product Realisation, Springer, London, 2005.

[3] D. Hawking, "Challenges in enterprise search," Proc of the 15th Australasian Database Conference, vol. 27, 2004, pp. 15-24.

[4] D. V. Vranic, 3D Model Retrieval, Dissertation, University of Leipzig, 2004.

[5] J. M. Martinez, R. Koenen, and F. Pereira, "MPEG-7: the generic multimedia content description standard, part 1," IEEE Multimedia, vol. 9, no. 2, 2002, pp. 78-87.

[6] U. Crenze, S. Köhler, K. Hermsdorf, G. Brand, and S. Kluge, "Semantic descriptions in an enterprise search solution," in G. Antoniou et al. Eds., Reasoning Web 2007, LNCS 4636, Springer, pp. 334-337, 2007.

[7] M. Buttler, The Bussiness Value of Enterprise Search, Martin Butler Research, July 2009.

[8] Y. Fu, R. Xiang, M. Zhang, Y. Liu, and S. Ma, "A PDD-based searching approach for expert finding in intranet information management," in H.T. Ng et al. (eds.), AIRS 2006, LNCS 4182, Springer, 2006, pp. 43-53.

[9] P. Dmitriev, P. Serdyukov, and S. Chernov, "Enterprise and desktop search," Proc. of the 19th International Conference on World Wide Web, Raleigh, North Carolina, USA, 2010.

[10] L. Owens, The Forrester Wave: Enterprise Search, Q2 2008. Report, Forrester, 2008.

[11] E. K. Lee and W. Noah, "An enterprise intelligence system integrating WWW and intranet resource," Proc. of the Ninth International Workshop on Research Issues on Data Engineering: Information Technology for Virtual Enterprises (RIDE '99), March 1999, pp. 28-35.

[12] D. R. Throop, "Enterprise Search Tasks in IVHM Practice," Proc. of IEEE Conference on Aerospace, 2006.

[13] R. Winnemöller, "Semantic Enterprise Search (but no Web 2.0)," Proc. of 10th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing), Mexico City, Mexico, May 2009.

[14] Y. Yao, L. Lin, and J. Dong, "Research on ontology-based multi-source engineering information retrieval in integrated environment of enterprise," Proc. of International Conference on Interoperability for Enterprise Software and Applications China (IESA '09), April 2009, pp. 277-282.

[15] S. R. Kruk, S. Grzonkowski, A. Gzella, and M. Cygan, "DigiMe - Ubiquitous Search and Browsing for Digital Libraries," Proc. of 7th International Conference on Mobile Data Management (MDM'06), 2006.

[16] J. Bahrs, B. Meuthrath, and K. Peters, "Information Retrieval Services for heterogeneous Information Spaces," Proc. of I-KNOW '08 and I-MEDIA '08 Graz, Austria, 2008.

[17] Y. Alemu, J. Koh, M. Ikram, and D. Kim, "Image retrieval in multimedia databases: A survey," Proc. of 5th International

Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2009.

[18] J. W. H. Tangelder, and R. C. Veltkamp, "A survey of content based 3D shape retrieval methods," Multimedia Tools and Applications, vol. 39, 2008, pp. 441-471.

[19] J. Pu, Y. Kalyanaraman, S. Jayanti, K. Ramani, and Z. Pizlo, "Navigation and discovery of 3D models in a CAD Repository," Computer Graphics and Applications, IEEE , vol. 27, no. 4, July-Aug. 2007, pp. 38-47.

[20] A. Ruifrok, A. Scheenstra, and R. C. Veltkamp, "A survey of 3D face recognition methods," Audio- and Video-based Biometric Person Authentication (AVBPA 2005), LNCS 3546, 2005, pp. 891-899.

[21] P. Geetha and V. Narayanan, "A survey of content-based video retrieval," J. of Computer Science, vol. 4, no. 6, 2008, pp. 474-486.

[22] M. A. Casey, R. Veltkamp, M. Goto, M. Leman, C. Rhodes, and M. Slaney, "Content-Based Music Information Retrieval: Current Directions and Future Challenges," Proc. of the IEEE Advances in Multimedia Information Retrieval, vol. 96, no. 4, 2008, pp. 668-696.

[23] N. Gronau, "Modelling knowledge intensive engineering processes with the knowledge modeler description language KMDL," In: F. Weber, S. Kulwant, K. D. Thoben, (eds.) Enterprise Engineering in the Networked Economy, University of Nottingham, 2003, pp. 195-202.

[24] N. Beckmann, H. Kriegel, R. Schneider, and B. Seeger, "The R*-tree: An efficient and robust access method for points and rectangles," SIGMOD Conference, 1990, pp. 322-331.

[25] P. Ciaccia, M. Patella, and P. Zezula, "M-tree: An Efficient access method for similarity search in metric spaces," Proc. of 23rd International Conference on Very Large Data Bases, 1997, pp. 426-435.

[26] C. Böhm, S. Berchtold, and D. A. Keim, "Searching in high-dimensional spaces - index structures for improving the performance of multimedia databases," ACM Computing Surveys, vol. 33, no. 3, 2001, pp. 322-373.

[27] U. Müller, T. Barth, and B. Seeger, "Accelerating the Retrieval of 3D Shapes in Geometrical Similarity Search using M-tree-based Indexing," Proc. of Workshop of Uncertainty, Knowledge Discovery and Similarity in Case-base Reasoning (UKDS'09), 2009.

[28] D. Gusfield, Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology. Cambridge University Press, New York, 1997.

[29] C. J. Romanowski and R. Nagi, "On comparing bills of materials: A similarity/distance measure for unordered trees," IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans vol 35, no. 2, pp. 249-260, 2005.

[30] B.S. Manjunath, P. Salembier, and T. Sikora, Introduction to MPEG-7: Multimedia Content Description Interface, Wiley & Sons, March 2003.

[31] C. Lütke Entrup, T. Barth, and W. Schäfer, "Towards a process model for identifying knowledge-related structures in product data," Proc. of the 6th Int. Conf. on Practical Aspects of Knowledge Management, LNCAI, vol. 4333, 2006, pp. 189-200.

[32] S. Karadgi, U. Müller, D. Metz, W. Schäfer, and M. Grauer, "Cost estimation of automotive sheet metal components using knowledge-based engineering and case-based reasoning," IEEE Int. Conf. on Ind. Eng. and Eng. Management, Hong Kong, 2009, pp. 1518-1522.

 ISBN: 978-1-61208-121-2