

Searchable Answer Generating Environment (SAGE): A Knowledge Management System for Searching for Experts in Florida

Dr. Irma Becerra-Fernandez
Florida International University
Decision Sciences and Information Systems
College of Business Administration
University Park Campus, BA 256A
Miami, FL 33199, becferi@fiu.edu
(305) 348-3278, (305) 348-3476 fax

Abstract

The need to acquire intellectual capital created the Knowledge Management (KM) movement, which aims to develop new practices and tools to capture knowledge. In-depth examination of current KM projects revealed the largest percent of KM projects attempting to create some kind of a knowledge repository. Current studies identify three types of knowledge repositories. One type of knowledge repository attempts to manage organizational knowledge by storing pointers to those who have specific knowledge within the organization. In this light, this paper discusses the development of the Searchable Answer Generating Environment (SAGE) application. The development of this application was funded through the NASA/Florida Minority Institution Entrepreneurial Partnership (FMIEP) grant. The purpose of this application is to create a repository of experts in the State of Florida (FL) State University System (SUS). This application creates one single web-enabled repository, which can be searched in a number of ways including research topic, investigator name, funding agency, or university. SAGE is a repository of Intellectual Capital within the state of FL SUS; helping locate FL SUS researchers for collaboration with industry and federal agencies, thus increasing the potential for research funding to the SUS. SAGE also enhances communication and allows more visibility for FL SUS experts, making universities more marketable while combining and unifying existing data from multiple sources into one user web-accessible interface. The SAGE system addresses an important KM problem: giving a user access to distributed knowledge, through a web-based Graphical User Interface.

Organizational Relevance of Knowledge Management

The need to acquire intellectual capital created the Knowledge Management (KM) movement, which aims to develop new practices and tools to elicit and capture organizational knowledge. Knowledge management in general tries to organize and make available important know-how, wherever and whenever it is needed. According to Dr. Peter Drucker (1997) "...this is the greatest social transformation in human history...the big challenge [we face] is how to make knowledge and knowledge workers productive". Dr. Drucker points out that new technology is an important part of these changes:

information technology allows to move information at a much lower cost than moving bodies through rapid transportation. But at the same time, the environment today renders new skills obsolete in a matter of years or months. These changes indicate that future success will depend on "a good deal of formal education and the ability to acquire and to apply theoretical and analytical knowledge". The basic economic resource is no longer capital, or natural resources, or labor, but is, and will be knowledge...today's enterprise richest resource is the knowledge and ideas residing in the minds of a particular organization's employees, customers, and vendors. Because of the increasing importance of Knowledge Management as a business movement, it has become increasingly critical to differentiate between three words often confused: Data, Information, and Knowledge. To distinguish between the three, we adhere to Zachman's (1997) definition of Data, Information and Knowledge:

1. Data is comprised of facts.
2. Information is data in context – its meaning depends on the usage.
3. Knowledge is information with direction or intent – thus, it facilitates a decision or action.

Several definitions for Knowledge Management abound in the literature today, among them (Becerra-Fernandez et.al. 1998a):

1. "Getting the knowledge from 'people who have done it' documented and available across the enterprise...as it was done by the team who did it" (Gundry & Metes, 1996).
2. An important new business movement that advocates the creation, sharing, and leveraging of knowledge within an organization to maximize business results (Milagro, 1997).
3. Creating the opportunities for private knowledge to be made public and tacit knowledge to be made explicit (Stewart, 1995).
4. Creation, acquisition and transfer of knowledge and modification of organizational behavior to reflect new knowledge and insights (Garvin, 1994).

Knowledge management in general tries to organize and make available important know-how, wherever and whenever it is needed (Becerra-Fernandez et. al. 1998b). This includes processes, procedures, patents, reference works, formulas, "best practices", forecasts, and fixes.

Technologically, Intranets, GroupWare, data warehouses, networks, bulletin boards, and video-conferencing are key tools for storing and distributing this intelligence (Maglitta, 1996).

It has been observed that KM systems underway at most organizations fall into three categories (Davenport et. al., 1997a):

1. Projects that created a knowledge repository, and whose objective was to take some form of knowledge and store it in a technical system for later access;
2. Projects which attempted to manage knowledge assets from a financial perspective (Becerra-Fernandez, 1997); and
3. Projects that attempted to improve the knowledge environment: either through the appreciation for knowledge, or by promoting knowledge sharing.

An in-depth examination of the first kind of KM systems and projects that created a knowledge repository, revealed three types of knowledge repositories (Davenport et. al., 1997a):

- a. In the first type of knowledge repository, organizational knowledge existed in some kind of explicit form, typically highly structured documents, for example systems to store marketing-oriented documents.
- b. The second type of knowledge repository consisted of developing less structured databases of employees' insights and observations. These projects are typically called "discussion databases" or "lessons-learned systems".
- c. Finally, the third type of knowledge repository attempts to manage organizational knowledge by storing pointers of those who have specific knowledge within the organization. For example, Microsoft uses this type of project to store software-systems development knowledge.

In the next sections, we look at the third type of Knowledge Management Systems in detail. The need for experts is not new but "the demand for specialized knowledge and experience is on the increase...restructuring and downsizing drive much of the increase, and the evolving challenges often call for management consultants across departments...technical areas demand new information and expertise" (Wheaton, 1997). To know who and where experts are is one of the most challenging activities that organizations face. Furthermore, "locating experts, specialists, and expert witnesses can take you all over the database and Internet map" (Schumacher & Dolan, 1995). Nowadays, when the Internet represents the biggest source of information, it is natural to suppose that it can be a good place for searching for experts. Currently, it is possible to go over the Internet and search for experts, but the search techniques and the existent databases are more useful if the searcher knows the name of the person he/she is looking for. Most of the time, this is not the case and what the researcher has is a group of parameters or requirements that need to be fulfilled. Furthermore, the most important

element in seeking for qualified experts is precisely that: "qualifying" the expert. Recognized indicators of expertise might include for example finding the most prolific writers, but being a prolific writer does not necessarily guarantee expertise, since authors may spend so much time writing that they don't spend much time "doing".

The Microsoft SpuD Project

Employees at the Microsoft Corporation need a high level of knowledge and competency due to the quick and evolving nature of their field. Therefore, Microsoft pays special attention to the hiring and promotion process in order to employ the most capable individuals. For this purpose, a pilot application was created for the Information Technology (IT) group within Microsoft. This project was named "Skills Planning und (and) development", known as "SpuD". The goal of this project was to develop a database containing job profiles available online. These profiles compile the knowledge and competency of each employee in the IT group. It is anticipated that this project will help match employee competency with jobs and work teams. The goals of this project are to develop a database containing job profiles available online, and to help match employee competency with jobs and work teams. The following are the five major components of the SPUD project (Davenport, 1997b):

1. Development of a structure of competency types and levels;
2. Defining the competencies required for particular jobs;
3. Rating the performance of individual employees in particular jobs based on their competencies;
4. Implementing the knowledge competencies in an online system;
5. Linkage of the competency models to learning offerings.

The objective of the employee categorizing process was to assemble a competency catalogue that could be used all across Microsoft. This way, a company employer seeking new personnel can search for an employee having the qualifications required for the position. As an example, Davenport (1997) suggests that the employer query the on-line system and ask, "Give me the top five candidates who have leadership skill levels on 80% of the competencies for this job and who are based in Microsoft's headquarters located in Washington". Using Microsoft Access, a prototype of the system was created based on a Structured Query Language (SQL) server as the database for MS Access, as well as a Web front end. Using a Web front end allowed for the publication of data around the world using the Internet and a web browser. To this day, the pilot project has had positive. In conclusion, the end results of a project like SpuD depend, to a large extent, on the acceptance of the individuals who are going use it. Furthermore, it will be important that employees and supervisors feel they have contributed to

the development of templates for the jobs, since it will be easier for them to adopt the Competency Model. Note that the validation of the data in this model rests with the manager, who essentially assigns the competency criteria to each of the employees under his/her supervision.

The Searchable Answer Generated Environment (SAGE)

In the same context of the Microsoft "SpuD" project, the NASA/Florida Minority Institution Entrepreneurial Partnership (FMIEP) grant is funding the development of the "Searchable Answer Generated Environment (SAGE) Knowledge Management System. The purpose of this KM System is to create a repository of experts in the State of Florida (FL) State University System (SUS). Previous studies have pointed out that there is a void in the ability to identify the capabilities in the SUS (Kotnour, 1998). Currently, each State University in Florida keeps a database of funded research, but these databases are disparate and dissimilar. The SAGE KM System creates one single repository by incorporating a quasi distributed database scheme, which can be searched by a variety of fields, including research topic, investigator name, funding agency or university. As NASA-KSC looks to develop new technologies necessary for the continuation of their space exploration missions, their need to partner with Florida SUS experts becomes evident.

The SAGE system combines the unified database by masking multiple databases as if they were one. One advantage of this method is that there is no need to reconfigure the data to fit it into one template. This methodology provides flexibility to the users and the database administrator, regardless of the type of program used to collect the information at the source (e.g. Dbase, FoxPro, Informix, Oracle, Sybase, MySQL, etc.). On the other hand if desired to conform the data under one system, it can be implemented as well. Although the project SAGE is specific in nature, what was desired was to develop tools and techniques that would make managing databases as seamlessly as possible.

The SAGE KM System combines and unifies existing data from multiple sources into one user accessible interface. Ideally the aggregated data would be accessible from one point of entry. For example, the data being accessed may consist of multiple data types that have been converted to a standard file format. SAGE consists of the typical university sponsored research data. One of SAGE's advantages is that there is only one point of entry or a web enabled interface, allowing multiple occurrences of the interface and giving the end user deployment flexibility. The main interfaces on the query engine use text fields to search the processed data for key words, fields of expertise, names, or other applicable search fields. The application processes the end user's query and returns the pertinent information (see Figure 1).

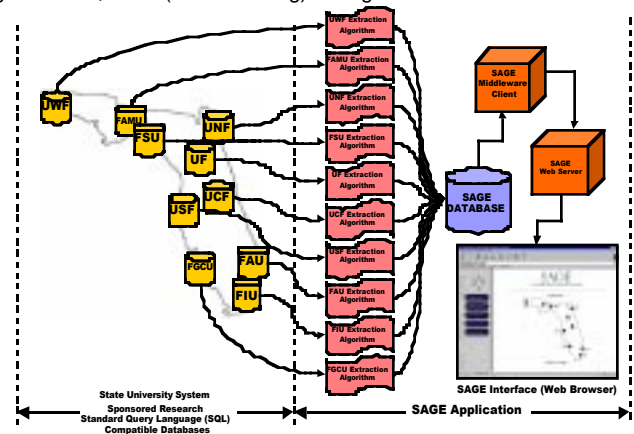


Figure 1: SAGE Architecture

The purpose of the SAGE KM System is to unify myriad data collections into one database collection that could easily be mined for relevant data. SAGE will give university researchers more visibility, and at the same time will allow interested parties to identify available expertise within the SUS. This application helps to identify a researcher's expertise within a discipline, and to facilitate communication or a point of contact. The benefits of SAGE are: (1) SAGE is a repository of Intellectual Capital within the state of FL SUS; (2) SAGE helps locate FL SUS researchers for collaboration with industry and federal agencies, thus increasing the potential for research funding to the SUS; (3) SAGE enhances communication and allow more visibility for FL SUS experts, making universities more marketable; and (4) SAGE combines and unifies existing data from multiple sources into one user web-accessible interface. The SAGE system addresses an important KM problem: giving a user access to distributed knowledge, through a web-based Graphical User Interface.

SAGE's strength also rests in its ability to search for experts by using a set of parameters instead of a proper name (even though this is a possibility also) and in the fact that it validates the data at the source, using the assumption that researchers who successfully obtain funded research grants are indeed experts in their fields. Other related current work is the development of "ProfNet, a database that lists public information officers for major universities, nonprofit organizations, and certain government facilities" (Ojala, 1995). A difference between SAGE and ProfNet is that the latter is not searchable by any individual over the Internet. Ojala explains that in Profnet public information officers who belong to the network query names of experts in a specific field. Then, the officers check in their institutions for the required expertise and return answers electronically. Thus, this is batch rather than interactive searching (1995). Another difference between SAGE and ProfNet is that SAGE is a more user friendly search engine that will give the user results immediately, instead of waiting for a response.

Another advantage that SAGE provides is that it's built upon a searching criteria that is recognized as a valid indicator of expertise, such as grants received and institution of employment. In most of the possibilities available on the Internet, the entity seeking for an expert has to use a combination of different tools in order to get this information. In short, SAGE is an inter-organizational system that provides users with the ability to locate experts.

The Technologies to Implement SAGE

SAGE is built upon the integration of the following technologies (see Figure 2):

1. Cold Fusion™— An off-the-shelf Rapid and Integrated Development Environment.
2. Open Database Connectivity (ODBC) allows middle-ware to interface with the database.
3. Verity's Search 97- enables document indexing, classification, search and retrieval personalized information dissemination. Search 97 allows the end-user to run sophisticated boolean searches, utilizing proximity operators, and relational operators. Verity 97 facilitates dynamic meta-data generation, along with incorporated relevance ranking in queries. Verity is used to perform the Keyword search, within the field that contains the grant abstract. In this way, you can use the Keyword "KM" to identify experts currently working in Knowledge Management, since the abstract that defines the project is expected to contain the word "Knowledge Management" as part of the research abstract.

The development of the SAGE database involved an initial design followed by incremental implementation phases. The initial design phase constituted a comprehensive survey of available tools and methodologies. An assessment was carried out to select the most efficient approach to data storage and data retrieval. The two methodologies considered for data access were Common Gateway Interface (CGI), and Middle-ware. Cold Fusion™ was chosen as the Middle-ware development environment because of its significant application strength and its demonstrated database interaction capabilities. Additionally, it provided the ability for secure transactions with Secure Socket Layer (SSL) technology, with the potential for strong encryption.

The implementation phase of SAGE involved the design of Cold Fusion modules, each with an assigned task. A robust server environment was set up along with a server operating system with remote user capabilities. Querying modules were used to provide search capabilities. Each query interface involved several modules that interacted with the database. Relational Databases were chosen because of their efficiency and flexibility with data. A sample database was obtained from the Florida International University Division of



Sponsored Research and Training (DSRT) to provide a prototype data storage structure in a relational database. After applying an extraction algorithm to the data, the SAGE database was populated with FIU data.

The next step was alpha testing the interface both locally and remotely. The evaluation phase of SAGE will include heavy end user testing and processing time optimization. The SAGE application was then web-enabled on the primary server for its current wave of beta testing.

Figure 2: The technologies to implement SAGE

Some of the technical challenges faced during the design and implementation of this project included the fact that databases from FL SUS DSRT offices were dissimilar in design and file format, and also in duplicate data occurrences, field names, and stored architecture. To resolve these problems PERL scripts were used to solve the database disparities, and a Data Dictionary was implemented to eliminate redundancies. One of the most important research contributions of SAGE is the merging of inter-organizational Database Systems through the use of correspondence tables, which function much like array pointers, and allow compliance to differing database formats. Through the correspondence tables we resolve finding similarities between database fields, no matter how different their original formats are.

The Future of SAGE

SAGE is a KM System in enterprise information management that is quite adaptable to the corporate environment. If used in corporate organization environments, SAGE can provide an organization with timely, pertinent, and cost effective data about internal strengths and capabilities, with a mature information base. This ability to catalog corporate and organizational assets is of prime concern to large organizations. Organizations have long desired to harness internal expertise, instead of relaying on consultants or third party contractors. Corporate prototype deployment would ideally involve

one group, preferably a unit that has Lightweight Directory Access Prototype (LDAP) directory enabled information network systems. This would drastically improve development time and overall usability and functionality. The directory "enabled" network will allow SAGE to keep a current listing for end user access, providing real-time responses to queries.

Acknowledgments

The Author wishes to acknowledge NASA-KSC and Bethune-Cookman College, for financial support to this project, entitled "Florida Minority Institution Entrepreneurial Partnership Grant", grant number NAGO-0220. Also the author wishes to acknowledge the contributions of the students who work in the FIU Knowledge Management Lab, particularly Marta March, Aracelys Rodriguez, Raheem Seecharan, Riyaad Seecharan, and Levi Stanley.

References

- I. Becerra-Fernandez (1997). "Methodology to Harvest Intellectual Capital at KSC". KSC Research & Technology 1997 Annual Report.
- Becerra-Fernandez, I., Riedel, J., & Lee, T. (1998a, February). Knowledge Management: Redefining Corporate Assets. Proceedings of the 7th International Conference on Management of Technology Conference. Orlando, Florida.
- Becerra-Fernandez, I. (1998b, April). Center for Innovation and Knowledge Management. Association for Computer Machinery SIGGROUP Bulletin, special Issue on Knowledge Management "Knowledge Management at Work", 19 (1), 46-51.
- Davenport, T. (1997a). Secrets of successful knowledge management. Knowledge Inc. 2 (2), 1-7.
- Davenport, T. (1997b) Knowledge Management Case Study: Knowledge Management at Microsoft. Published at <http://kman.bus.utexas.edu/kman/microsoft.htm>
- Drucker, P. (1997). [Excerpts from his speech at Ernst & Young's Knowledge Advantage Conference]. San Diego, CA.
- Garvin, D. A. (1994, January). Building a learning organization. Business Credit, pp. 19-28
- Gundry, J. & Metes, G. (1996). Team knowledge management: a computer-mediated approach. Published at <http://www.knowab.co.uk/wbwteam.html>.
- Kotnour, T. (1998, October) Partners in Education Conference, Cocoa Beach, Florida
- Maglitta, J. (1996, January). Know-how, Inc.: once a conference curiosity, knowledge management is catching on.
- Computerworld, pp. 73-75.
- Milagro (1997). [Excerpts from the "Austin Dialogue on Knowledge Management"]. Austin, Texas. co-sponsored by Milagro Systems.
- Ojala, M. (1995). People who need people: Extending the never-ending search for people. Online 19 (1), 36-40.
- Schumacher, J. E. & Dolan, D. R. (1995). Tracking down experts with online resources. Database 18(3), 14-20.
- Stewart, T. (1995, November). Getting Real About Brainpower. Fortune, pp. 201-203.
- Wheaton, Perry, In search of a few good experts. Electric Perspectives, 22 (6), 8-12.
- Zachman, John, Enterprise Architecture: the Issue of the Century. Database Programming and Design, February 1997