Facilitating the Online Search of Experts at NASA using Expert Seeker People-Finder

Irma Becerra-Fernandez, Ph.D.
Assistant Professor
Decision Sciences and Information Systems
Florida International University, Miami, FL 33199
Tel: 305-348-3476, E-mail: becferi@fiu.edu

Abstract

People-Finder Systems are knowledge repositories that attempt to manage knowledge by holding pointers to experts who possess specific knowledge within an organization. This paper presents insights from the development of Expert Seeker, an organizational People-Finder KMS that will be used to locate experts at the National Aeronautics and Space Administration (NASA). This paper discusses insights and lessons learned from the development of this system, and the role of technology in automating the maintenance of the expert's profiles. Expert Seeker represents an important first step towards achieving our objective of automatically and intuitively discovering and identifying intellectual capital within the organization. While several systems in place today rely on self-assessment, we look at the potential of artificial intelligence (AI) technologies, in particular, data mining and clustering techniques, to uncover and map organizational expertise.

1 Introduction to Knowledge Management Systems

Knowledge Management Systems (KMS) have been defined as "an emerging line of systems [which] target professional and managerial activities by focusing on creating, gathering, organizing, and disseminating an organization's 'knowledge' as opposed to 'information' or 'data'" [Ala99]. KMS currently in use at most organizations, fall into three categories [Bec99A]:

1.Educational KMS: To elicit and catalog tacit knowledge, and at the same time serve as an educational tool.
2.Problem-solving KMS: Organizations with significant intellectual capital require eliciting and capturing knowledge for reuse in solving new problems as well as recurring old problems.
3.Knowledge repositories: Under the auspices of KM, tools historically used for singular, unrelated purposes are integrated to address the corporate memory problem. One type of knowledge repository is People-Finder Systems, also known as Knowledge Yellow Pages.

People-Finder Systems are knowledge repositories that attempt to manage knowledge by holding pointers to experts who possess specific knowledge within an organization. Several organizations in different business categories have identified the need to develop systems to help locate intellectual capital, or People-Finder KMS. The intent in developing these systems is to catalog knowledge competencies, including information not typically captured by Human Resources systems, in a way that could later be queried across the organization. A literary review and a table comparing the characteristics of hallmark People-Finder KMS in use in organizations today appears in [Bec00].

The paper presents insights from the development of Expert Seeker, an organizational People-Finder KMS that will be used to locate experts at the National Aeronautics and Space Administration (NASA). This paper discusses insights and lessons learned from the development of this system, and the role of technology in automating the maintenance of the expert's profiles.

2 Motivation: Developing an Organizational Knowledge Management Strategy

In order to assess the areas of Intellectual Capital for Kennedy Space Center, a Knowledge Management Assessment (KMA) was designed and implemented between the months of February and April 1998. The goals of this effort were:

Proc. of the Third Int. Conf. on Practical Aspects of Knowledge Management (PAKM2000)
Basel, Switzerland, 30-31 Oct. 2000, (U. Reimer, ed.)
http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-34/

1. To analyze the current types, sources and uses of knowledge in the organization;
2. To develop a detailed set of system specifications and implementation plan for future related activities;
3. To create a detailed plan for dealing with future needs; and
4. To gather data for the implementation of a prototype that will address some of KSC's Knowledge Management needs.

For this purpose, a series of assessment interviews were designed and implemented, with the cooperation of representatives of the majority of the functional groups at KSC. The goals of the interviews were to assist KSC in identifying key competencies and analyze the current knowledge architecture for the center. This step would ensure that the appropriate methodology is recommended at the end of this phase.

Following is a summary of the findings from the Knowledge Management Assessment of KSC's Knowledge Management (KM) needs and possible enhancements to the current KM Environment:

1. Expert Knowledge Elicitation and Virtual Mentoring Tools: Six of the eight interviewed groups identified the need for the implementation of tools to elicit, capture, and transfer the knowledge of experts acquired through years of experience at NASA.
2. An "Expert Seeker" Knowledge Management System: Six of the eight interviewed groups expressed the need for an application that holds pointers to experts with a particular background. This application would help locate Intellectual Capital within the center at all levels, from technicians to Ph.D.'s. The Expert Seeker application would store competencies available within the organization, including for all KSC employees completed past projects, patents, and their relevant expertise. An added benefit would be to include competencies outside NASA-KSC, for example those of subcontractors.
3. Collaborative Tools: Of the 8 technical groups interviewed, 6 expressed a need for Internet/Intranet based collaborative tools that capture knowledge as teams create it, integrated with an electronic document storage.
4. Decision Support and Expert Systems: The need for the implementation of KMS that would enhance decision making and would facilitate the decision process by incorporating knowledge factors from past projects that might prove useful and help make better, more educated decisions in the future.
5. Center-wide Lessons Learned Repository: Five of the eight interviewed groups expressed the need for a center-wide Lessons Learned Repository.

Following from the recommendations presented at the conclusion of this study, the KSC Executive team decided to fund the implementation of the KSC Expert Seeker People-Finder.

3 Summary of previous research in People-Finder KMS: The Searchable Answer Generating Environment (SAGE)

The NASA/Florida Minority Institution Entrepreneurial Partnership (FMIEP) Grant is funding the development of the Searchable Answer Generated Environment (SAGE) which is in the category of People-Finder KMS [Bec99B]. The purpose of this KM System is to create a repository of experts in the State of Florida (FL) State University System (SUS). Currently, each State University in Florida keeps a database of funded research, but these databases are disparate and dissimilar. The SAGE KM System creates a single repository by incorporating a 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. Although the project SAGE is specific in nature, what was desired was to develop tools and techniques that would make managing these independent databases as seamless as possible**.** One of SAGE's advantages is that there is only one user point of entry at the web-enabled interface, allowing multiple occurrences of the interface and giving the end user deployment flexibility. The main interfaces developed 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.

SAGE has been online since August 16, 1999 at http://sage.fiu.edu. Future developments for SAGE include such projects as the development of algorithms that will facilitate the maintenance of SAGE in a more automatic fashion. This inter-organizational system will require coding developments at both the SAGE server

and at each of the university's servers. A complete description of SAGE, including implementation details and results, appears in [Bec00].

4 The Expert Seeker People-Finder System at Kennedy Space Center

The NASA Faculty Awards for Research (FAR) is funding the development of Expert Seeker, which is in the category of People-Finder KMS. Previous Knowledge Management studies at KSC affirm the need for a center wide repository, which will provide KSC with Intranet-based access to experts with specific backgrounds. Currently KSC is reorganizing from an operations center into a research and development center. Expert Seeker aims to help locate intellectual capital within NASA-KSC, and is this particular characteristic that differentiates Expert Seeker from SAGE (the latter a KMS to find experts within the Florida universities). Expert Seeker will be used to search for experts located at KSC, although its use is expected to expand to other NASA Centers. The Expert Seeker KMS will be accessed via KSC's Intranet. In contrast, the SAGE KMS, which is on the world-wide-web, is accessible through the Internet. Another important difference between SAGE and Expert Seeker is that the latter will enable the user to search for much more detailed information regarding the experts' achievements, including information such as intellectual property, skills and competencies, as well as the proeficiency level for each of the skills and competencies. The Expert Seeker KMS will provide access to competencies available within the organization, including items that are not typically captured by the typical Human Resource applications, such as completed past projects, patents, hobbies, and other relevant knowledge. This People-Finder KMS will be especially useful when organizing cross-functional teams.

The main interfaces on the query engine in Expert Seeker will use text fields to search the proposed data for keywords, fields of expertise, names or other applicable search fields. The application will process the end user's query and returns the pertinent information. The information will be collected from a conglomeration of multimedia databases, and the presented as queried. The purpose of the Expert Seeker KMS is to unify myriad data collections into web-enabled repository that could easily searched for relevant data. Prior to this project, there was no single point of entry into a unified repository that allowed identification of employees based on specific skills. Expert Seeker will allow KSC experts more visibility, and at the same time allow interested parties to identify available expertise within KSC. This People-Finder KMS will help to identify a researcher's expertise, within a discipline, and to facilitate communication with a point of contact.

5 Expert Seeker at Goddard Space Flight Center

To further create synergies between the efforts to develop Expert Seeker at Kennedy Space Flight Center (KSC), a similar effort to prototype Expert Seeker at Goddard Space Flight Center (GSFC) was funded by the Center of Excellence in Space Data and Information Sciences**.** Efforts related to this proposal will attempt to mirror, as funds allow, some of the efforts currently underway at KSC, including:

1. System specification and selection of the organizational groups to prototype the GSFC Expert Seeker People-Finder.
2. Development of the GSFC knowledge taxonomy.
3. Design and development of the GSFC-Expert Seeker.
4. Implementation of the system prototype.
5. Testing of the system prototype.
6. Rollout.

It is expected that implementing the GSFC version of Expert Seeker will be to a large extent a replication of the ongoing efforts at KSC, in order to minimize duplication of efforts and maximize the return-on-investment for NASA. The resources that will be provided by this grant will serve to ensure generic features for this innovative system. Furthermore, implementation of Expert Seeker at GSFC will further validate the effectiveness of this KMS and ensure the development of a system that could potentially be of value to all of NASA. On the other hand, it is expected that the Knowledge Taxonomy for GSFC will differ from the one for KSC. But this requirement does not pose a concern, as Expert Seeker could be developed so the software could be "configured" with customizable knowledge taxonomy.

6 The Technologies to Implement Expert Seeker

The development of Expert Seeker is being accomplish with the use of the following technologies:

1. Cold Fusion 4.0, Java Script, Active Server Pages ASP (Coding and Programming)
2. SQL Server 7.0 (Databases)
3. Verity (Search capabilities)
4. Adobe Phototshop 5.0 (GUI)
5. HTML and other web development tools

The development of Expert Seeker requires the utilization of existing data as much as possible. Expert Seeker will use the data in existing Human Resources databases for information such as employee's formal educational background, the X.500 Directory for the employee point-of-contact information, a Skills Database which profiles each employee's competency areas, GPES, an employee performance evaluation system, and PRMS a project resource management system. Figure 1 depicts the architecture of Expert Seeker. Furthermore, other related information deemed important in the generation of an expert profile which is not currently stored in an in-house database system can be user-supplied, such as employee's picture, project participation data, hobbies, and volunteer or civic activities Information regarding skills and competencies, as well as proficiency levels for the skills and competencies needs to be collected, to a large extent, through self-assessment. Recognizing that there are significant shortcomings of self-assessment, we propose to use an increased reliance in technology to update employees' profiles, and thus place less reliance on self-assessed data. For example, we are proposing the use of Global Performance Evaluation System (GPES), an in-house performance evaluation tool, to mine employees' accomplishments and automatically update their profiles. Typically, employees find it difficult to make time to keep their resumes updated. Performance evaluations, on the other hand, are without a doubt, part of everybody's job. We therefore seek to use this tool, augmented with appropriate queries, to inconspicuously keep the employees profiles up-to-date.

Finally, a data mining effort of the document repository will also contribute to update employees' profiles. Based on the assumption that authors of documents in the repository are subject matter experts, therefore, mining the electronic document repository will contribute to keeping employees' profiles up-to-date in an unobtrusive way. For this purpose, we are currently experimenting with the use of the Term Frequency Inverse Document Frequency (TFIDF) algorithm. The TFIDF algorithm is used as a measure of the uniqueness or relevance of a document within a collection of documents with respect to a specific keyword. TFIDF is calculated by the following formula:

$$w_{ij} = tf_{ij} * log(N/n)$$

where w is the TFIDF score for term i in document j, tf is the frequency of term i in document j, and the inverse document frequency or idf is calculated by the logarithm of the total number of documents divided by the number of documents term i appears at least once. A term that appears frequently in fewer documents will generate a higher TFIDF score for w_{ij} than a term that appears with comparatively high frequency but appears in many documents. Thus the TFIDF score is a measure of how relevant or unique a document is for a keyword in relation to a collection of documents. The resulting internal representation vector of the documents can then be searched by keyword. The TFIDF algorithm will be used within the Expert Seeker system to locate experts within the NASA Goddard Space Flight Center and NASA Kennedy Space Center by mining published documents within the Intranet of these organizations. This can be done periodically to keep the internal document representations up to date and to index new documents. The resulting TFIDF vector will be used for search queries. Documents that are returned as a query result will then be indexed by author name. The final result will rank authors according to those with the highest-ranking documents for that keyword and display these to the user as a subject-matter expert.

7 Challenges in the Implementation of People-Finder KMS

Previous research [Bec00] conducted to establish the parameters to design Expert Seeker application has demonstrated that one of the challenges in developing People-Finder KMS is related to the inherent shortcoming of self-assessment. Most of the People-Finder KMS in place today, except for example SAGE People-Finder or Mitre's Expert Finder [Kot98], rely on each employee to complete a self-assessment of competency, which is later used when searching for specific knowledge areas. The disadvantage of self-assessment is that the results of self-assessment are subjective, based on each person's self-perception, the results could be hard to normalize, and employees' speculation about its possible use could 'skew' the results. For example, one particular organization conducted a skills self-assessment study during a period of downsizing. This resulted in employees' exaggeration of their competencies, for fear they might have been laid-off. On the other hand, another organization made it clear the self-assessment would be used to contact people with specific competencies to answer related questions. This self-assessment caused employees to be overly modest about their skill profiles, for they would be required to put to test their specific knowledge. Furthermore, one People-Finder in place at Microsoft [Bec00] required supervisors to ratify their subordinates' self-perceptions, and assign a quantifiable value to it, a requirement that many organizations would find this requirement too taxing on their supervisors.

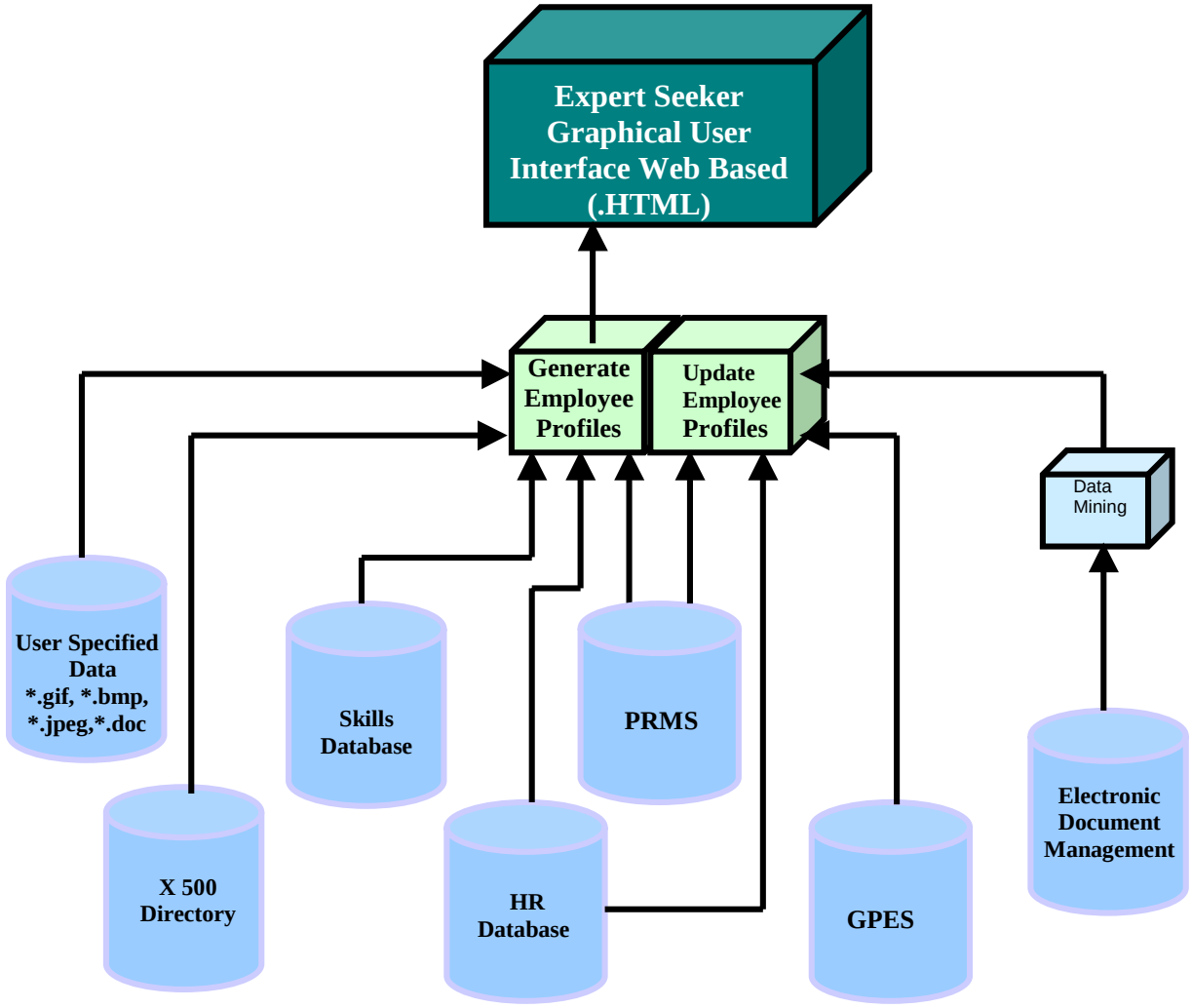

Figure 1:Expert Seeker Architecture

Another challenge in developing People-Finder KMS deals with the development of knowledge taxonomies. Taxonomy is the study of the general principles of scientific classification. Knowledge taxonomies allow organizing knowledge or competency areas in the organization. In the case of People-Finder systems, the taxonomy is used to describe, and catalog people's knowledge, an important design consideration. Furthermore, knowledge taxonomies could be critical in the People-Finder system's success [Bec00]. People-finder KMS in place have addressed this consideration keeping in mind that:

1. Taxonomies should easily describe a knowledge area.
2. Taxonomies should provide minimal descriptive text.
3. Taxonomies should facilitate browsing, not complicate them.
4. Taxonomies should have the appropriate level of granularity and abstraction. If the level is to high then it will be too complicated for the user, but if the level is too low it will not properly describe the knowledge areas.

HP's CONNEX [Bec00] has been one of the successful systems in place that has developed a fairly functional taxonomy to describe people's knowledge. According to

Carrozza [e-mail 1999] HP's knowledge taxonomy is based on standards such as the U.S. Library of Congress Classifications (available online at http://lcweb.loc.gov) and the INSPEC Index (available online at http://www.iee.org.uk/publish/inspec), but is customized to their business area. Other firms have followed this model and have created their own taxonomies, as in the case for Microsoft's SpuD and for BA&H's Knowledge On-Line. According to Remeikis [phone interview, Sept. 3, 1999], this effort was successful. In contrast, the National Security Agency used a taxonomy based on a standard from the Department of Labor (O*Net – available online at http://www.doleta.gov/programs/onet). Finally, according to Timothy Horst, Vice-President and Manager of Construction Resources and Technologies, Bechtel Construction Operations Incorporated also developed a taxonomy for their Knowledge Bank. It is based on standards developed by the National Center for Construction Education Research (NCCER – available online at http://www.nccer.org), but is only being used to catalog skills of manual workers [Phone interview, August 17 and 26, 1999]. While a number of work classification standards have been developed, that could be used to organize knowledge areas, we have not been able to apply any of these standards directly, without some thought and further development of the taxonomy. A deep analysis of the People-Finder KMS in place reveals that many attempts to create knowledge taxonomies are unsuccessful [Remeikis phone interview 1999] or sub-optimal [Carrozza phone interview and follow-up e-mail, 1999].

8 Practical Applicability of Expert Seeker

The Expert Seeker application has completed it's first year development. The need for such a system to locate experts in an organization of more than 10,000 individuals, all with exellent qualifications, in order to reduce the time and effort spent in resolving issues pertaining to the R&D conducted at the different NASA centers. A preliminary usability study by NASA officials [Naus phone interview and follow-up e-mail, 2000] revealed minor weaknesses in the graphical-user interface that since have been corrected. Suggestions also focused on methods to access information, such as the capability of the system to allow combined searches.

Expert Seeker, when completely implemented, is expected to become an effective tool in the management of knowledge required for new product and process development at NASA. Chris Carlson of NASA-Kennedy Space Center, envisions how Expert Seeker will be used in the future [Carlson phone interview and follow-up email, 1999], as he describes a possible scenario*:*

> *You are working in a project to build a new cryogenic handling storage facility. You encounter a problem, where upon testing, a valve fails. There is a design problem. You have two choices:*
>
> - *The first choice is to go back through the same process with the same company and NASA engineers working the problem*
>
> *or*
>
> - *The second choice is to use Expert Seeker to organize the Rapid Answer Collaborative Knowledge Expert Team (RACKET). Using the expertise keyword 'cryogenics' Expert Seeker finds the following experts:*
>
> 1. *A collection of scientists from the University of Arizona for cryogenics studies;*
> 2. *A valve manufacturing expert from a plant in Detroit;*
> 3. *A cryogenic expert that worked on problems during shuttle that transferred to Marshall Space Flight Center.*
> 4. *In addition, the Expert Seeker uncovers a collection of technical white papers and lessons learned that NASA has published from similar projects.*
>
> *The RACKET collaborates by video teleconference and the Internet to pinpoint the design problem, identify a feasible solution, and fixes the design problem in two days.*

9 Conclusions and Future Work

Results from the KSC Knowledge Management Assessment revealed the importance of a system to identify experts within the organization. Expert Seeker represents an important first step towards achieving our objective of automatically and intuitively discovering and identifying intellectual capital within the organization. While several systems in place today rely on self-assessment, we look at the potential of artificial intelligence (AI) technologies, in particular data mining and clustering techniques, to uncover and map organizational expertise.

Data mining technologies could contribute to updating employees' profiles. Based on the assumption that authors of documents in the repository are subject matter experts, mining the electronic document repository could contribute to keeping employees' profiles up-to-date in an unobtrusive way. Furthermore, clustering techniques

could be instrumental in clustering similar data objects together [Meh99]. In this case clusters of expertise, could reveal expertise areas that may not be currently defined. The use of clustering techniques provides the potential of creating a domain dictionary of "pseudo-keywords" that could serve to increase the semantic domain of the keywords, and which could be used to identify relationships that may not be necessarily obvious. Another application of this clustering notion is the development of a "super" concept, which would allow to group experts together, developing a group-level of expertise. Given the individual areas of expertise, these could be clustered together into groups of expertise or virtual "centers of excellence". In the case of Expert Seeker, grouping of experts within KSC or GSFC with complementing expertise areas could result in virtual "centers of excellence". This effort could reveal areas of strength that could otherwise go unnoticed in the organization.

Acknowledgements

The author wishes to acknowledge NASA-Kennedy Space Center and NASA-Headquarters under the Faculty Awards for Research (FAR-99), grant number NAG10-0259, as well as NASA-Goddard Space Flight Center and the CESDIS, contract number NAS5-32337 and subcontract number 5555-97-74, for financial support for the development of Expert Seeker. The author also wishes to acknowledge James Jennings, Shannon Roberts, Gregg Buckingham, Milt Halem, Susan Hobbard, Nancy Laubethal, and Jim Green, all of NASA who championed this initiative at KSC and GSFC. Finally, the author wishes to acknowledge the contributions of the students who work in the FIU Knowledge Management Lab and who collaborated in this research, specifically Claudia Alarcon, Greg Copeland, Hector Hartmann, Nelson Ospina, Haniel Pulido, Gabriela Rizzo, Juan Rodriguez, Hernan Santiesteban, and Luis Felipe Villegas.

References

[Ala99] Alavi, M., Leidner, D. *Knowledge Management Systems: Issues, challenges, and benefits*. Communications of the Association for Information Systems, 1, article 7 (1999).

[Bec99A] Becerra-Fernandez, I. *This is rocket science: The Knowledge Repository of Kennedy Space Center.* Presentation at the Delphi's International Knowledge Management Summit (IKMS'99), San Diego, California (March, 1999).

[Bec99B] Becerra-Fernandez, I. *Searchable Answer Generating Environment (SAGE): A Knowledge Management System for Searching for Experts in Florida.* Proceedings of the 12th Annual International Florida Artificial Intelligence Research Symposium (FLAIRS) - Knowledge Management Track, Orlando, Florida (May, 1999).

[Bec00] Becerra-Fernandez, I., *The Role of Artificial Intelligence Technologies in the Implementation of People-Finder Knowledge Management Systems*, Knowledge Based Systems special issue on Artificial Intelligence in Knowledge Management, forthcoming in 2000.

[Kot98] Kotnour, T. *Presentation at the Partners in Education Conference*, Cocoa Beach, Florida (October, 1998).

[Meh99] Mehrotra, M., Alvarado, S., and Wainwright, W. *Laying a Foundation for Software Engineering of Knowledge Bases in Spacecraft Ground Systems.* Proceedings of the 12th Annual International Florida Artificial Intelligence Research Symposium (FLAIRS), Orlando, Florida (May, 1999).