The Forest Portal: A Multidisciplinary Project

Marianne Koch[1], Lois Delcambre[2], Patricia Toccalino[3], Eric Landis[4], Fred Phillips[1], Tim Tolle[5], Len Shapiro[6], Nicole Steckler[1], David Maier[2], Mathew Weaver[2], Shawn Bowers[2], Balbinder Banga[1], Jason Brewster[6], Afrem Gutema[6], Sudarshan Murthy[2], Bill Howe[2], Rupa Tummala[2], Julia Norman[3], Kirsten Zillman[3], David Drake[4], Craig Palmer[7], and Ashley Burt[3]

[1]Department of Management in Science and Technology, [2]Computer Science Department, [3]Environmental and Biomolecular Systems Department, all within OGI School of Science and Engineering at OHSU, [4]Consultant, [5]USDA Forest Service, [6]Portland State University Computer Science Department, [7]University of Nevada Las Vegas

1   Introduction

This project partners the USDA Forest Service with researchers at several universities and federal agencies. The goal of the project is to improve the management of natural resources by providing easy access to documents. These documents may include decision notices, environmental impact statements, and specialist reports. They are produced and stored by natural resource managers at many different places across the Pacific Northwest in many differing forms, some of them not electronic (Delcambre, Tolle, et al., 2003).

The structure of this project is shown in Figure 1, with Drs. Timothy Tolle and Lois Delcambre serving as Project Co-Directors. On the right side of the figure, we see the various disciplinary specialists from the participant agencies. On the left side of the figure, we see the three teams of university researchers. In Section 2 of this paper we summarize the work of each discipline and the results of the work. Section 3 discusses the interdisciplinary interactions that occurred in this project.
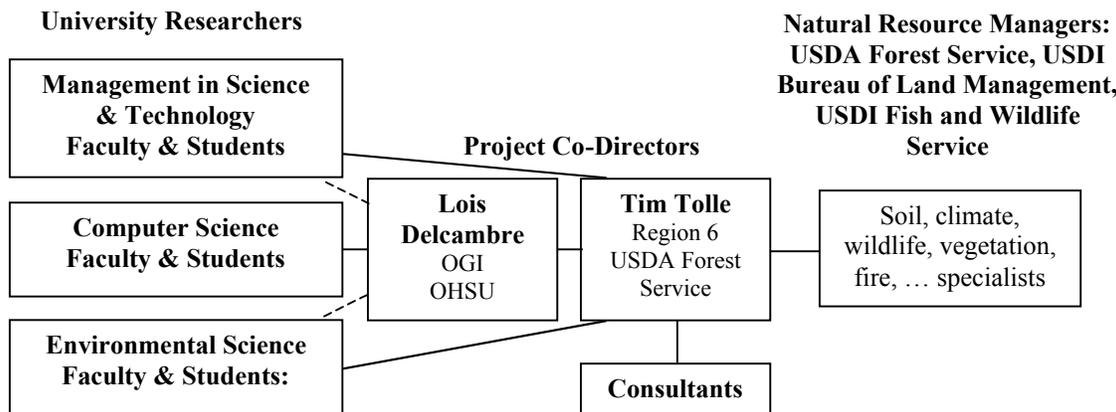

Figure 1: The Forest Team

2   The Component Parts of the Project

2.1   User Needs – Management Research

Our aim in conducting user-oriented interviews was to determine requirements to better enable users to achieve their goals; be it locating information, posting research findings, managing information, manipulating information, or providing legal defense. The target group was defined as local forest communities of the Pacific Northwest, specifically those on the management side of the public and private forestry sector.

Interviewees were asked to provide either responses to general inquiries regarding their information seeking and gathering behavior, or to articulate a case of a recent information inquiry they conducted (scenario). Several aspects of the users' information processes were targeted for discovery including:

- purpose of the information inquiry,
- desired content and structure (formats, scale, etc.),
- vocabularies and fields used for searching,
- methods for determining quality prior to retrieval,
- IM tools used during inquiry, retrieval and re-archiving process, and
- when/whether, in the inquiry process, person-to-person contact is required.

In all, approximately 65 interviews were conducted in person and by phone between April 2000 and December 2001. Highlighted user requirements include: a need for phone support throughout the inquiry process, a desire for expert (institutional or specialist) lists, availability of authors' biographies, availability of bibliographies of archived documents within metadata, ability to search by geographic names, the ability to easily archive retrieved documents for future use, and the ability to access ongoing research.

2.2   Controlled Vocabularies – Environmental Science Research

CVs are composed of terms that are commonly understood among targeted portal users. We explored standard terminology, keyword lists for describing database or WWW-site subject content in a controlled way, thesauri, classification schemes, and glossaries and other reasonable sources of terms in the domains of interest. We evaluated appropriate CVs to place in the automated system using the following five steps.

Identify domains of interest. Twenty-eight domains of interest to natural-resource agencies, especially the USDA Forest Service, Bureau of Land Management, USDI Bureau of Land Management, USDI National Park Service, and USDI Fish and Wildlife Service were identified (Delcambre et.al., 2003). Domains that were evaluated included: air, aquatic biology, soils, vegetation, and wildlife.

Locate key-word lists, glossaries and indexes for each domain. Subject-matter experts familiar with each domain searched the web and other resources for appropriate CVs for each domain.

Evaluate the CVs. We used criteria reviewed and approved by the USDA Forest Service staff (Tolle et al, 2002).

Identify and implement relationships between and amongst terms. Examples of structured relationships include broader-term-narrower-term, synonym, and general functional relationships (Tolle et al, 2002). The next section explains how these are expressed in the automated system.

2.3   Metadata++ - Computer Science Research

The primary focus of the computer science team has been on the representation of the various controlled vocabularies in a uniform model called Metadata++ (Weaver et al, 2001). Metadata++ describes terms arranged in a broader-term/narrower-term hierarchy, with synonym and general association relationships among terms. The contribution of Metadata++ is in the uniform model where terms are presented in the hierarchy, terms are selected from the hierarchy to attach to a document or to express a search, and documents that match a search are presented directly with the term that they match in the hierarchy.

Other specific contributions from the computer science team include:

- We incorporated a GIS component into Metadata++ that allows us to (1) extract location CVs, such as hydrologic unit codes, to describe watersheds, from GIS datasets, (2) let users view maps and

select locations with the GIS, which are then transferred to Metadata++ and included in a compound search with non-location keywords, (3) determine spatial synonyms, and (4) display search results on a map (Shapiro et al, 2002). See Figure 2.

- • We developed the notion of a superimposed schematics browser (Bowers, Delcambre, Maier, 2002) and built a prototype system that allows the user to highlight specific items in an underlying document. Our motivating example used the superimposed schematic to highlight the issues, appellants, deciding officer, determination, and so forth for an appeals decision letter.

- • We developed a representation scheme called the Uni-Level Description that can represent information stored in various data models, such as XML, RDF, Topic Maps, and database data models, using a metamodel composed of basic structures such as struct, set, bag, and list (Bowers and Delcambre, 2000; Bowers and Delcambre, 2001; Bowers and Delcambre 2003).
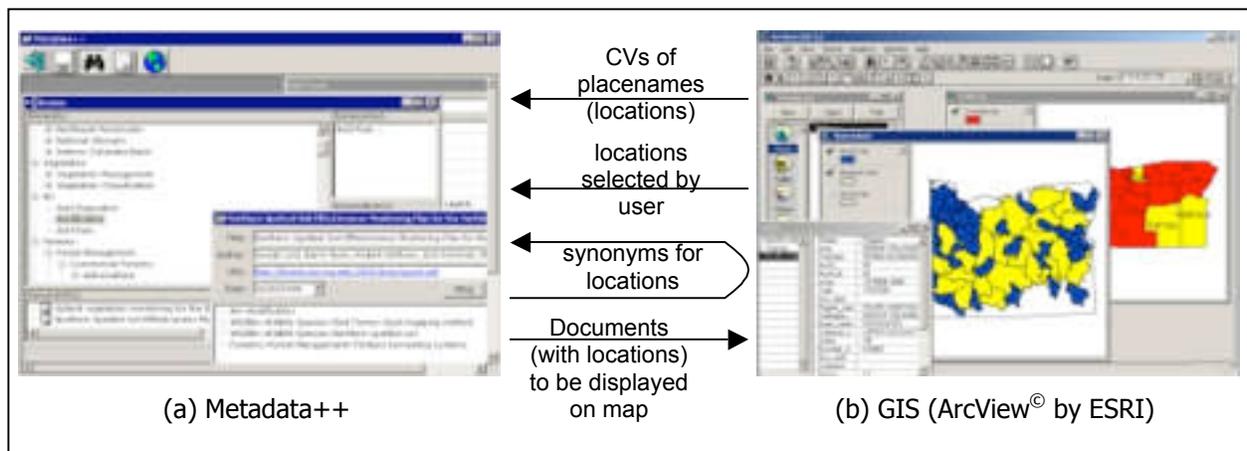

Figure 2: Functions supported by our integration

2.4   Facilitators and Inhibitors to Information Sharing – Social Science Research

Members of the social science team were funded to study facilitators and inhibitors of information sharing across federal agency boundaries. We interviewed a cross section of key participants in three federal agencies: USDA Forest Service, USDI Bureau of Land Management, and UDSI Department of Fish and Wildlife, and reviewed archival documents related to recent information sharing efforts.

Based on this exploratory study we developed a conceptual framework (Steckler and Koch, 2003) describing how the dynamics between three subcultures found within government organizations (a scientist culture, a politician culture, and a bureaucrat culture) create barriers to sharing information between and within agencies.

Two cases of inter-agency collaboration among federal agencies charged with managing and regulating natural resources were used to illustrate the framework and show how differences between the three cultures lead to difficulties with both technology-mediated and interpersonal communication and information-sharing (as presented in Koch, Steckler and Drake 2003). The conceptual framework can be used to improve implementation of data- and knowledge-sharing technologies toward the goal of better public resource management.

3   How it All Comes Together

The Forest Project is an example of interdisciplinary research. On the academic side, researchers from multiple disciplines worked together on various parts of the project, in various combinations. On

the government partner side, there was extensive collaboration and coordination with scientists from different fields within and across federal government agencies. The project has touched many disciplines beyond the management, environmental, and computer science fields: we've published papers in computer science, knowledge management (Delcambre, Koch, et al, 2003), software development, data envelopment analysis, and management conferences and journals, and touched on other fields like geology (a team member once said the geologists were ahead of us in the use of controlled vocabularies), marketing, and user interface design.

On the agency side, this project has contributed a rich set of controlled vocabularies appropriate for natural resource management that has drawn attention from a wide audience. Users in several agencies are currently evaluating the Metadata++ system and documents from other digital libraries are being automatically loaded into the Metadata++ system. This will allow a broader range of users to evaluate the system directly. The details regarding deployment of this system are under development.

References

S. Bowers, L. Delcambre, D. Maier, "Superimposed Schematics: Introducing E-R Structure for *In-Situ* Information Selections," Proc. of the 21st Int'l Conf. on Conceptual Modeling (ER 2002), Lecture Notes in Computer Science 2503, S. Spaccapietra, S. March, Y. Kambayashi, (eds.), Springer-Verlag, 2002, pp. 90-104.

S. Bowers and L. Delcambre, "Representing and Transforming Model-Based Information," Proc. of the Semantic Web Workshop 2000, held in conjunction with the European Conference on Digital Libraries (ECDL) 2000, Lisbon, Portugal, September 2000, pp. 5-18.

S. Bowers, L. Delcambre, "On modeling conformance for flexible transformation over data models," To appear in *Knowledge Transformation for the Semantic Web*, IOS Press, Amsterdam, 2003.

S. Bowers, L. Delcambre, "A Generic Representation for Exploiting Model-Based Information," ETAI Journal, 2001, Linköping Electronic Articles in Computer and Information Science, ISSN 1401-9841, Vol. 5, Section D (2001), pp. 1-34. http://www.ida.liu.se/ext/epa/ej/etai/2001/019/epapage.html

L. Delcambre, M. Koch, E. Landis, D. Maier, F. Phillips, N. Steckler, P. Toccalino, T. Tolle, M. Weaver, and B Banga, Knowledge Management: A Re-assessment and Case, chapter to be published in book on Global Knowledge Management, edited by Hamid R. Nemati, Ph.D., ISOM Department, Bryan School, The University of North Carolina at Greensboro.

L. Delcambre, T. Tolle, D. Maier, F. Phillips, P. Toccalino, N. Steckler, M. Koch, L. Shapiro, E. Landis, B. Banga, S. Bowers, J. Brewster, A. Gutema, W. Howe, S. Murthy, J. Norman, R. Tummala, M. Weaver, and K. Zillman. 2003. Harvesting Information to Sustain our Forests. Communications of the ACM. 46(1):38-39.

M. Koch, N. Steckler and D. Drake, 2003. "Implementing Knowledge Systems: Avoiding Common Pitfalls," presented at Portland State University Engineering and Technology Management seminar.

L. Shapiro, L. Delcambre, T. Tolle, M. Weaver, D. Maier, D. Guenther, J. Brewster, A. Gutema, "G-Metadata++: Rich Keyword Search Enhanced with a GIS", Proceedings of GIScience 2002, Boulder CO, Sept. 2002.

N. A. Steckler and M. J. Koch, "Science, Politics, and Bureaucracy: Cultural Barriers to Information-Sharing in Government Agencies." Unpublished manuscript, under review

T. Tolle, Toccalino, P., Delcambre, L., Norman, J., Phillips, F., and D. Maier. 2002. Using controlled vocabularies as a knowledge base for natural resource managers. Paper prepared for dg.o 2002.

M. Weaver, L. Delcambre, D. Maier, "A Superimposed Architecture for Enhanced Metadata", DELOS Workshop on Interoperability in Digital Libraries, held in conjunction with European Conference on Digital Libraries (ECDL 2001), Darmstadt, Germany, Sept. 2001.