COLLABORATIVE FILTERING: POSSIBILITIES FOR LIBRARIES

Janet Webster
Associate Professor
Oregon State University Libraries
Marilyn Potts Guin Library
Oregon State University
Hatfield Marine Science Center
2030 Marine Science Drive
Newport, OR 97365
janet.webster@oregonstate.edu

Dr. Jon Herlocker
Assistant Professor
Oregon State University School of Electrical Engineering and Computer Science

Heather Pennington-Lehman
Masters student
University of Washington Information School

Seikyung Jung
Ph.D. student
Oregon State University School of Electrical Engineering and Computer Science

ABSTRACT: As librarians, we recognize that users seldom approach a research project with a clear search strategy thus we develop means to help them clarify their research objectives. However, our traditional methods do not translate well to the web environment, while at the same time demand for web access increases. Addressing this challenge, the Oregon State University Libraries and the Department of Computer Science are collaborating on the development of two test beds to investigate interfaces that will lead users to web resources that are valuable and useful by using collaborative filtering recommendation technology. Collaborative filtering (CF) allows an information portal to continually learn what resources are likely to be useful to which users (and which questions) by observing search patterns as well as explicit recommendations from users. The first test bed, the Tsunami Digital Library (TDL), is a portal to websites with relevant and authoritative information on tsunamis with an interface aimed at both researchers and the general public. The second test bed is a search and recommendation portal interface to the OSU Libraries' web resources. In this paper, we will explain collaborative filtering and its potential relevancy to libraries, describe the TDL project, and discuss future plans.

KEYWORDS: Digital libraries; collaborative filtering;

Introduction

As librarians, we recognize that users seldom approach a research project with a clear search strategy and thus we develop means to help them clarify their research objectives. We also know that many, if not most, users start with the web. Rather than be cut out of the information gathering process, how can we translate our methods and expertise to the web environment? As a computer scientist, my colleague, Jon Herlocker, is intrigued with how to develop more effective search engines. He sees inherent limitations to the evolution of search systems that rely on refining algorithms based on machine analysis of document content. Full-text keyword search is probably the content analysis we are most familiar with. Can we take a new look and apply collaborative filtering (CF) to improve the quality of information returned in a search while maintaining relevance and speed? Together, we are exploring ways to integrate librarians' expertise in evaluating resources and directing users towards appropriate information into a new approach to search interfaces using CF.

Collaborative Filtering and Libraries

Collaborative filtering, also referred to as social filtering or recommender systems, takes recommendations from many and applies them to the user's need. Consider how you find about good restaurants in a city you are visiting, or, what current movie may appeal to your unique taste. You ask friends, colleagues; you read reviews. You make choices from using all of this information. This is collaborative filtering. Collaborative filtering, developed and implemented in entertainment (e.g. MovieLens.org) and commercial settings (e.g. Amazon.com), incorporates the results of human analysis of content on a massive scale. Gladwell discusses the power of human input and opinion as applied to promoting small press books (1999). In this example, he refers to collaborative filtering as "an attempt to approximate ...insider knowledge" (p.50). Gladwell describes how the small bookstore owner with years of experience with books and customers can be compared to a sophisticated search system that compiles and synthesizes the recommendation and use patterns of people. Both can steer the user towards appropriate, relevant and perhaps surprising choices. Another example of the use of CF is a small experimental system within the education sector described by Relker and Walker (2003). The authors explore how CF can enhance a collection of educational web resources by involving the users of those resources. A small group of teachers and students describe, review and rate the resources. The system attempts to leverage the opinions and expertise of the individuals to help the entire group or community of users.

Several scientists, including the OSU Research Team, are exploring how to apply CF to improve search interfaces. In our prototype search interface integrating CF, users ask their question, and that question is matched against previous questions asked by other users. Then, the system recommends documents, pages, or other resources that other users found useful. The portal 'learns' what resources are valuable for which questions by observing the users' behavior and recording the recommendations. Every time somebody uses the search and recommendation system, the system becomes smarter.

Collaborative filtering could help librarians address three challenges we face:

- As electronic information increases in amount and value, how do we provide satisfactory access to it?
- As the definition of the digital library evolves, how to we continue to add value to our collections and services?
- As computer scientists and the commercial sector develop search interfaces for the electronic landscape, how can we integrate the expertise of librarians into the development process?

All three challenges can be answered with the development of improved search interfaces. We function in an electronic environment and are slowly being overwhelmed by the amount of information and the number of interfaces. At the same time, we recognize the increasing value of electronic information. Adapting our traditional methods into that environment meets with varying levels of success. We currently create subject guides on the web, catalog web sites, and constantly tinker with our web pages. Yet, we rely as much on Google as on the intricacies of our online catalogues or carefully selected electronic resources to find information for our users and ourselves. Our current approach to organization and collection of web resources is not keeping pace.

Improving search interfaces is critical to manage and provide access to information effectively. Our OPACs and digital libraries rely on their user interfaces for their ultimate success (e.g. their use by us and others.) Yet, often change and innovation in those interfaces are tied to the expertise of our vendors and the conflicting demands they face in the marketplace. Change is also tied to developments in computer science. Some in the field believe that there are inherent limitations to the evolution of search systems that rely on refining algorithms based on machine analysis of document content. Recent search interface research is mostly incremental because researchers are reaching the limits of software's ability to autonomously *understand* content given today's computational limitations. The approach using CF circumvents this roadblock; it lets people perform the understanding of content, so we can use simpler, more dependable algorithms that learn from observing human reactions and responses. We already look to users for insight in some parts of library operations. For example, we incorporate usage patterns as we select which journals to purchase and which to cancel. Collaborative filtering suggests we can extend those observations of usage patterns to other parts of the collections as well as services.

Beyond providing improved access, librarians can also add value to digital information by incorporating our evaluation and collection expertise to gather high quality, 'good' information into manageable, searchable virtual collections. Currently, most digital libraries are electronic analogs of special collections or discrete collections of print material. As we expand the definition, a digital library becomes a gateway to both print and digital resources and services. This expansion adds value as it changes our perspective on what can be in our collections. Collaborative filtering leverages our collection efforts by validating our choices through user input. One goal of CF is to build digital libraries that improve with each user query, rather than just with each document added.

The final challenge, working with computer scientists, provides an excellent opportunity to exploit our expertise while expanding our knowledge base. Too often, the search interface designers neglect the human component preferring to create more complex code. In a recent interview, Wayne Rosing, Google's Vice President for Engineering observes that Google will eventually be need to be more than just a search engine that identifies items. It will have to "find you the good stuff. It will be an up-leveling of our ranking function...from what's the best document to what's the best, most well-formed knowledge on the subject" (Brown 2003, p. 20.) Librarians have a professional obligation to not only monitor this progression but help shape it. Shaping our search interfaces and exploring CF as a way to add human input is one possibility. A search system based on CF could potentially harness librarians' evaluation expertise and experience observing user needs with the expertise of users.

Collaborative Filtering and Libraries at Oregon State University

At Oregon State University (OSU), a Carnegie Doctoral Extensive University, the School of Electrical Engineering and Computer Science and the Libraries are working together to improve the effectiveness and accessibility of digital collections and web information portals using CF. The genesis of the project was threefold.

First, as a land/sea/space grant institution, the OSU Libraries has a tradition of trying to provide meaningful access to information to a wide variety of users. In recent years, we have been considering how to build a natural resources digital library that stores and serves information to not only the research community, but the general public as well. This entails developing different types of user interfaces or ways for people to get at the same data yet in ways that are understandable. Those of us in the libraries needed help with how to present complex data and information to an array of audiences.

The second impetus came when OSU garnered significant funding from the U.S. National Science Foundation's Network for Earthquake Engineering Simulation (NEES.) The goal of the program is to create an infrastructure for research and education, consisting of resources for experimentation, computation, model-based simulation, data management, and communication that are networked and geographically distributed. The recently dedicated Tsunami Wave Basin at the OSU O.H. Hinsdale Wave Research Laboratory is the first of 15 research installations nationwide. Integrated into the development of the basin is a system to capture, archive and deliver the data generated from experiments to scientists worldwide. Research is focused on the best way to do this and who are the possible audiences.

Finally, many of us work for institutions and agencies that encourage inter-disciplinary discourse. Often, it is merely talk. At OSU, various people recognized common interests in information access and starting looking at ways to address common issues. The university librarian initiated periodic meetings with the head of Northwest Alliance for Computation Science and Engineering (NACSE), the person working on the information technology system of the Tsunami Wave Tank. These two decided to explore the concept of a Tsunami Digital Library that provided a gateway to quality electronic information with a user appropriate interface. They identified staff with appropriate expertise and directed us to work on it. Dr. Jon Herlocker, a new professor in the School of Electrical Engineering and Computer Science, came with experience with the MovieLens project. Janet Webster, the marine science librarian at OSU, is familiar with the literature, the variety of users, and is interested in the problems of access to gray literature.

The Tsunami Digital Library: Background and Description

Digital collections of Internet resources place important information closer at hand. They facilitate the dissemination of new research faster, and create access for wider audiences. Rather than searching the entire Internet, a collection limits resources to those pertinent to a subject. Given the diversity of tsunami research and information, a digital collection could provide the means to search materials from all areas of tsunami research. The collection addresses language issues, data preservation, maintenance, and quality management. This last point, the assurance of high quality information, requires the diverse tsunami research community to be directly involved.

Collaboration is not unusual in tsunami research. Numerous national and international partnerships, including the Tsunami Hazard Mitigation Program, the Pacific Tsunami Warning Centre, the proposed Intra-Americas Sea Tsunami Hazards System, and the IOC International Tsunami Information Center share information. These partnerships have generated significant, high quality digital resources. However, the information is often located in disparate and hard to find sites, housed for an indeterminate amount of time, or not maintained for the most effective use. This lack of coordination is problematic as it does not adequately promote information sharing between the multi-faceted tsunami researchers, nor does it aid a public that is still mostly unaware of the basic facts of tsunami dangers.

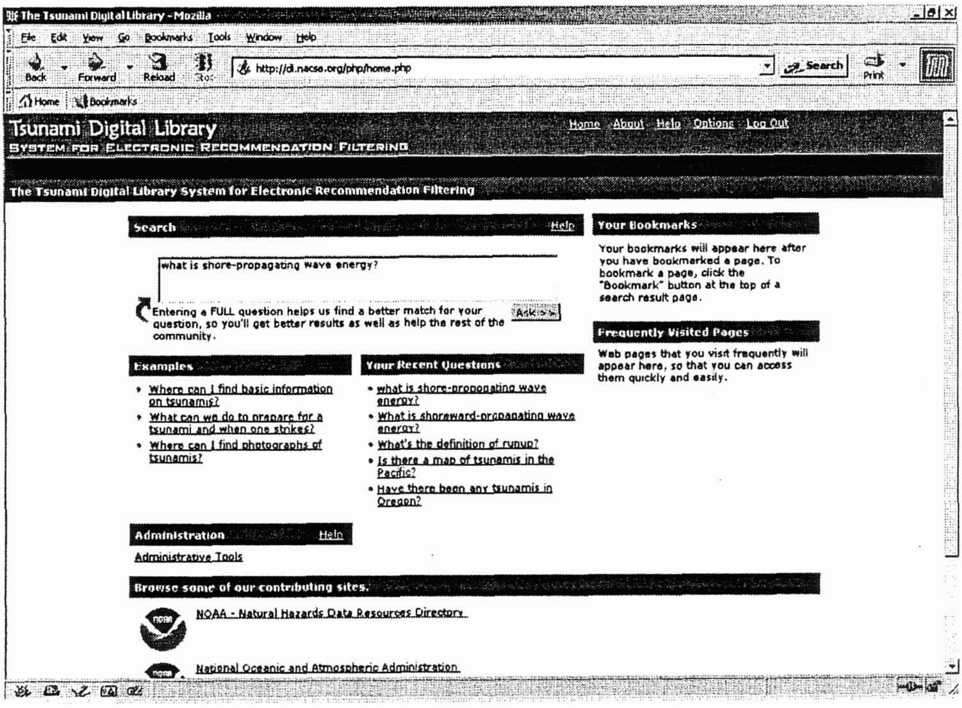
From a librarian's perspective, collecting tsunami information is messy yet seductive with its breadth and variety. Authority is multidisciplinary and originates from a wide variety of institutions, organizations and authors. It spans research data to school curriculum, evacuation maps to oral histories. Access to the information is variable with a few excellent web sites and many mediocre ones. Much historical info is in print and not well distributed. There is world-wide interest, although most focus is on the Pacific Basin. The variety of formats (e.g. real-time data, videos, PDFs) makes access a challenge. Multiple languages are another access challenge. Some of the web-based information is fragile with limited longevity. Maintaining accurate long-term sites is challenging and often beyond the scope of many researchers' work.

People in the field of tsunami research recognize the value of the existing information but mention problems with access and usability in particular. For example, Atwater's investigation of paleotsunami events along the Washington-Oregon coasts was controversial because of the lack of written documentation of historic events (1987). Ten years later, Satake et.al published an article that demonstrated the origin of the 1700 major tsunami in the Pacific Basin; they examined historic records housed in Japanese monasteries to do so (1996). These records were not readily available to others given language and physical barriers. This example, coupled with the earlier observations, demonstrates a need to improve access to and maintain the wealth of web based tsunami information.

Given the nature of tsunami information, the impetus of the NEES funding and a willing ness to collaborate, we began developing the Tsunami Digital Library (TDL). The TDL is an intriguing test bed for the application of Collaborative Filtering. Maintaining wide and open access to timely and historic tsunami information is challenging as it is multi-disciplinary, multi-lingual, in a variety of formats, and its community of users highly varied. We can explore collecting digital documents, developing partnerships, maintaining quality tsunami information on the web, and evolving collaborative filtering in a circumscribed community. We can use this test bed to develop searching protocols, assess user needs, and propose working models for collaborative administration of web based information.

Basically, the TDL is an information portal designed to coordinate the access and distribution of Internet based tsunami related research. It does not house the tsunami information nor tsunami research sites. Rather, it is an intelligent interface that centralizes access to contributing partner sites within a digital library system. The TDL retrieves links to pertinent content and entire sites by "answering" queries posed by TDL patrons (see Figure 1). Research suggests that prompting users to submit full queries results in higher user satisfaction (Belkin 2003). We use a large query box to encourage entry of full questions.

Figure 1: Search screen from the Tsunami Digital Library


The Tsunami Digital Library - Mozilla
File Edit View Go Bookmarks Tools Window Help
Back Forward Reload Stop
http://dl.nacse.org/php/home.php
Search
Print
Home Bookmarks

Tsunami Digital Library

SYSTEM FOR ELECTRONIC RECOMMENDATION FILTERING

Home About Help Options Log Out

The Tsunami Digital Library System for Electronic Recommendation Filtering

Search Help

what is shore-propagating wave energy?

Entering a FULL question helps us find a better match for your question, so you'll get better results as well as help the rest of the community.

Ask >>

Your Bookmarks

Your bookmarks will appear here after you have bookmarked a page. To bookmark a page, click the "Bookmark" button at the top of a search result page.

Frequently Visited Pages

Web pages that you visit frequently will appear here, so that you can access them quickly and easily.

Examples

- Where can I find basic information on tsunamis?
- What can we do to prepare for a tsunami and when one strikes?
- Where can I find photographs of tsunamis?

Your Recent Questions

- what is shore-propogating wave energy?
- What is shoreward-propagating wave energy?
- What's the definition of runup?
- Is there a map of tsunamis in the Pacific?
- Have there been any tsunamis in Oregon?

Administration Help

Administrative Tools

Browse some of our contributing sites.

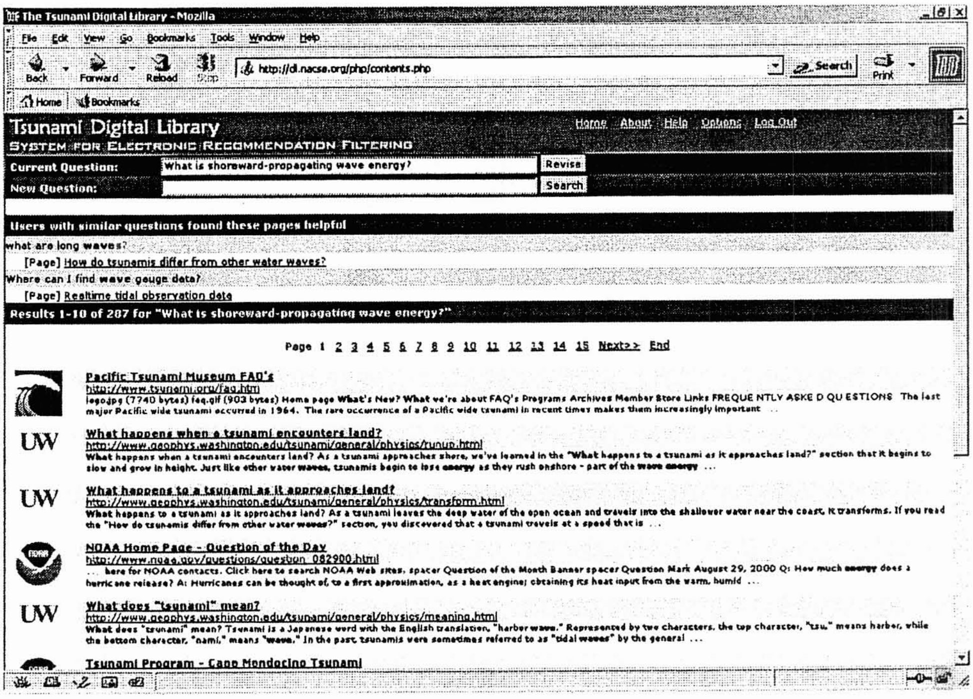
NOAA - Natural Hazards Data Resources Directory

National Oceanic and Atmospheric Administration

A TDL patron asking about "shoreward-propagating wave energy" will be provided with two sets of links (see Figure 2):

- a list of related questions with links to sites / pages
- a list of links to sites / pages that may contain information relevant to answering the question.

Figure 2: Results screen from the Tsunami Digital Library


The Tsunami Digital Library - Mozilla
File Edit View Go Bookmarks Tools Window Help
Back
Forward
Reload
Stop
Search
Print
Home
Bookmarks

http://dl.nacse.org/php/contents.php

Tsunami Digital Library

SYSTEM FOR ELECTRONIC RECOMMENDATION FILTERING

Home About Help Options Log Out

Current Question: What is shoreward-propagating wave energy? Revise

New Question: Search

Users with similar questions found these pages helpful

what are long waves?

[Page] How do tsunamis differ from other water waves?

Where can I find wave gauge data?

[Page] Realtime tidal observation data

Results 1-10 of 287 for "What is shoreward-propagating wave energy?"

Page 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 Next>> End

Pacific Tsunami Museum FAQ's
http://www.tsunami.org/faq.htm
logo.jpg (7740 bytes) faq.gif (903 bytes) Home page What's New? What we're about FAQ's Programs Archives Member Store Links FREQUE NTLY ASKE D QU ESTIONS The last major Pacific wide tsunami occurred in 1964. The rare occurrence of a Pacific wide tsunami in recent times makes them increasingly important ...

UW What happens when a tsunami encounters land?
http://www.geophys.washington.edu/tsunami/general/physics/runup.html
What happens when a tsunami encounters land? As a tsunami approaches shore, we've learned in the "What happens to a tsunami as it approaches land?" section that it begins to slow and grow in height. Just like other water waves, tsunamis begin to lose energy as they rush onshore - part of the wave energy ...

UW What happens to a tsunami as it approaches land?
http://www.geophys.washington.edu/tsunami/general/physics/transform.html
What happens to a tsunami as it approaches land? As a tsunami leaves the deep water of the open ocean and travels into the shallower water near the coast, it transforms. If you read the "How do tsunamis differ from other water waves?" section, you discovered that a tsunami travels at a speed that is ...

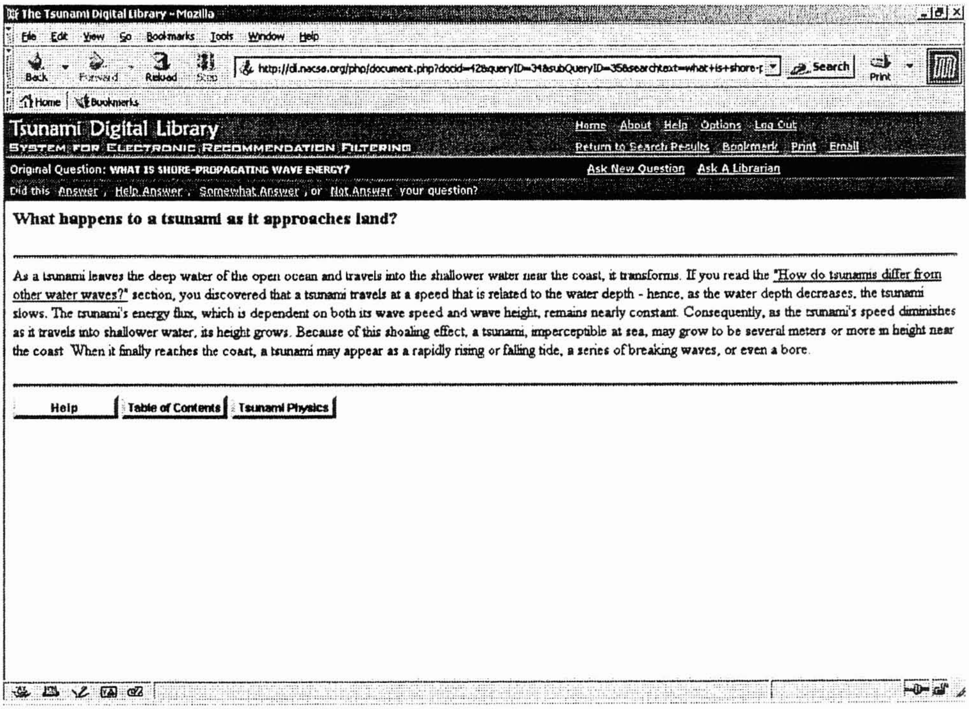
NOAA Home Page - Question of the Day
http://www.noaa.gov/questions/question_082900.html
... here for NOAA contacts. Click here to search NOAA Web sites. spacer Question of the Month Banner spacer Question Mark August 29, 2000 Q: How much energy does a hurricane release? A: Hurricanes can be thought of, to a first approximation, as a heat engine; obtaining its heat input from the warm, humid ...

UW What does "tsunami" mean?
http://www.geophys.washington.edu/tsunami/general/physics/meaning.html
What does "tsunami" mean? Tsunami is a Japanese word with the English translation, "harbor wave." Represented by two characters, the top character, "tsu," means harbor, while the bottom character, "nami," means "wave." In the past, tsunamis were sometimes referred to as "tidal waves" by the general ...

Tsunami Program - Cape Mendocino Tsunami

The TDL returns results on topic sensitive information by CF. Content can be and is intended to be rated for usability and accuracy by TDL patrons. Through this rating, subsequent patron queries can be answered with increased reliability and accuracy.

The user then can select various pages and vote on whether the selected resource answers her query. The vote is on a sliding scale: answered, helped answer, somewhat answered and did not answer. The vote is then tallied and used in the future when a similar query is submitted (see Figure 3).


The Tsunami Digital Library - Mozilla
File Edit View Go Bookmarks Tools Window Help
Back Forward Reload Stop
http://dl.nacse.org/php/document.php?docid=42&queryID=34&subQueryID=35&searchtext=what+is+shore-p
Search Print
Home Bookmarks

Tsunami Digital Library

SYSTEM FOR ELECTRONIC RECOMMENDATION FILTERING

Home About Help Options Log Out

Return to Search Results Bookmark Print Email

Original Question: WHAT IS SHORE-PROPAGATING WAVE ENERGY?

Ask New Question Ask A Librarian

Did this Answer , Help Answer , Somewhat Answer , or Not Answer your question?

What happens to a tsunami as it approaches land?

As a tsunami leaves the deep water of the open ocean and travels into the shallower water near the coast, it transforms. If you read the "How do tsunamis differ from other water waves?" section, you discovered that a tsunami travels at a speed that is related to the water depth - hence, as the water depth decreases, the tsunami slows. The tsunami's energy flux, which is dependent on both its wave speed and wave height, remains nearly constant. Consequently, as the tsunami's speed diminishes as it travels into shallower water, its height grows. Because of this shoaling effect, a tsunami, imperceptible at sea, may grow to be several meters or more in height near the coast When it finally reaches the coast, a tsunami may appear as a rapidly rising or falling tide, a series of breaking waves, or even a bore.

Figure 3: Voting page from the Tsunami Digital Library

In addition to rating content, users can suggest resources to add to the TDL. Site inclusion in the TDL aids the greater distribution of tsunami related information that might not otherwise be located easily. It also furthers the development of existing research networks. These functions, and the added ability of patrons to annotate information found through the TDL, facilitate information sharing within the tsunami research community.

A small research team consisting of Dr. Herlocker, Seikyung Jung (PhD candidate), four undergraduate students in computer science and the librarian developed the system. The librarian compiled an initial short list of the key tsunami sites to seed the collection. These sites were authored or published by authoritative sources and frequently referenced. The research team built a preliminary search interface using freely available software building on Dr. Herlocker's experience with the MovieLens Project. We did a preliminary controlled experiment to see how students responded to using the interface to answer tsunami questions. We derived the questions from two curricula aimed at high school students (Servicio Hidrográfico 2002; Goodrich & Atwell 2000). Results indicated that people liked the system and those with more recommendation data took less time and fewer clicks to find answers within the collection (Herlocker et al. 2003).

Even so, we knew had ongoing challenges to make the CF system work well. We identified three critical issues.

- Identifying when a document is accessed, and why.
  Identifying every time a document is accessed within a digital library is a challenge in a web based environment. When a user clicks on a hyperlink, the original site loses track of the user. The challenge is very apparent with the TDL because most of the content consists of web pages available at locations outside of Oregon State University's domain. Once we can track the usage adequately, we must record the context of the use. We need to know what information need was being pursued when that document was accessed.
- Collecting ratings from the users.
  Collaborative filtering relies on collecting ratings from users on the relevance and quality of documents. A mechanism must be provided to collect those ratings. Furthermore, since users may be unwilling to take the time to provide enough ratings explicitly, methods for observing implicit expressions of ratings must be developed. This includes monitoring the last document viewed and when a document is emailed or printed. We need to know what information satisfied the user's needs and what may satisfy future users.
- Personalizing information while addressing changing needs.
  Consumers who have used Amazon.com recommendations have experienced this challenge when they use Amazon to purchase a gift for another person. As a result of purchasing a baby book, they may receive recommendations for baby paraphernalia for months. We need to provide recommendations that are appropriate to the user's immediate need, and only use past history that is relevant to the current need.

The Research Team felt we had passed the first test well, but were unsure where to head given limited funds. We decided to proceed on two fronts. We recognized that the search system needed to be tested within a larger test bed. So, we approached the OSU Libraries for permission to further develop our nascent system using the Libraries large and varied web site as a test bed. At the same time, we felt an assessment of the information needs of tsunami researchers would bolster our arguments for further development of the TDL, or show that the concept was not viable for this community. Coupling the needs assessment with further development of the search system kept our interests stimulated.

Initial Needs Assessment of Tsunami Researchers

We hired a library school student to assist with the assessment as well as collection development of the TDL. Initially, a small group of expert users was identified in consultation with Dr. Harry Yeh, the Edwards Chair for Engineering at OSU, who has been involved in the project from the beginning. The list has now been expanded to include the contact people at the various sites that are currently in our collection as well as those who are cited frequently. We selected five people for the pilot project, developed the questions, and have completed the first round of interviews. These were initiated by email and the completed by phone when possible. Having a library school student to work on this was essential and we recommend making use of the talent of these students.

The initial responses indicate our questions are relevant, but the order could be revised to improve the flow of the conversation (Appendix A.) Our pilot group, as anticipates, uses the Internet for their work and research. While extensive users, they are not expert searchers. Timeliness of information is the most mentioned reason for why a site is useful. Most do not immediately mention specific sites unless prompted. All describe common problems including difficulty navigating, dead links, language differences, and repetitious information. They also mention bad site design, lack of useful indexing, and the slowness of working through logic trees. Identification of these problems is particularly helpful as we are interested in alleviating many of these navigation challenges with the TDL and its CF search interface.

As we are promoting a digital library, we are curious as to people's perceptions of the concept. Although these are intense Internet users, most are not sure how to describe a digital library. They mention lists of links (such as some of their own sites) and access to electronic resources through their academic institutions. Librarians use the term somewhat casually and even many of us would have a hard time coming up with a succinct definition. We know it when we see it. Is it important that we define digital libraries so we can promote them? Or, is the concept still evolving and it does not matter at this stage? We prefer to keep the concept loose and monitor users' perceptions.

In general, this pilot group was intrigued with the TDL. They would use it if it was easy to navigate and contained relevant information. Relevant information includes real time data, current research, raw data, observations, model results, pictures, videos, bibliographical references, contacts, and maps. They indicate an interest in being involved if that involvement is not a time sink. These people are already involved in maintaining web sites, so they recognize the possible benefit of a collaborative effort.

The TDL's Future

After completing the survey, we anticipate using the results to move forward with formalizing partnerships, moving the TDL from a development to a production platform, and then opening it for use by the tsunami community and others in late 2003. We are also working on a collection policy that is challenging as we are focusing on electronic documents. How do we ensure access? What do we include? Should we 'collect' resources that may disappear?

This last question suggests exploring how we can help manage sites included in the TDL. People are interested in what we can do to help administer sites. Yet, we recognize this as opening Pandora's box. The TDL offers site statistics and tracking for partner site administrators. They can monitor site use, better manage dead links, and review the queries that are being answered by specific site content. Assistance with the management of critical sites is a possibility in the future Do we want to mirror or even capture certain sites? Are we willing to take over the "Ask for help?" function?

As with many experimental projects, the TDL funding is cobbled together from a variety of sources. These include the OSU Libraries Gray Chair for Innovative Technologies, the Northwest Alliance for Computational Scientific Engineering funding for undergraduate research and the Georgia Pacific HMSC internship. Our attempt to garner major funding through the U.S. National Science Foundation's International Digital Library Initiative failed when the program was cut. The grant writing process, through, was an excellent collaborative effort and all involved gain a better understanding of the TDL and its possibilities. Currently, we are seeking funding from NSF and other foundations.

The lack of specific funding led us to consider other applications of the CF system. Consequently, we are broadening our scope by implementing a search and recommendation interface for the OSU Libraries' extensive web site in a way that enlists users to recommend pages, databases and e-journals to others asking similar questions of the site. Developing and testing the merits of a recommendation system in this more diverse setting present challenges including resolving issues of integration with existing library systems and library tradition; dealing with noisy and untrustworthy data; computation, display and explanation of recommendations; and inferring recommendations from user behavior.

Conclusion

Collaborative filtering offers possibilities for translating traditional approaches to reference services and resource discovery into the web environment. It has potential to improve the efficiency of our evolving digital libraries. It suggests ways to improve our search interfaces making systems learn through human interaction rather than the tweaking algorithms.

It can be an avenue to use librarians' expertise and help us effectively realize the potential of the electronic information environment. Collaborating between librarians and computer scientists stimulates both to think differently and expansively about mutual problems. Just as CF needs humans to make the evaluations and recommendations, computer scientists need librarians as users and collaborators.

Librarians can do more to shape the future information landscape. The current chaos causes much hand wringing. New approaches are needed, can be intellectually challenging and are fun to develop. We should be thinking beyond the efficient search to the future that Rosing described when we figure out how to deliver the best knowledge and not just the best document (Brown 2003). Collaborative filtering may be part of that future.

References:

Amazon.com, Inc. [Online]. Available: http://www.amazon.com. [Accessed: 17 October, 2003].

Atwater, Brian F. 1987. Evidence of great holocene earthquakes along the outer coast of Washington State. *Science* 236(4804): 942-944.

Belkin, N.J., Cool, C., Kelly, D., Kim, G., Lee, J-Y., Muresan, G., Tang, M-C., & Yuan, X-J. 2003. Query length in interactive information retrieval. *Proceedings of the 26th Annual International ACM SIGIR*, pp. 203-212.

Brown, David J. 2003. A conversation with Wayne Rosing. *Queue* 1(6): 12-20. [Online]. Available: http://doi.acm.org/10.1145/945131.945162 [Accessed: 17 October, 2003].

Gladwell, Malcolm. 1999. Annals of Marketing: The silence of the sleeper: How the Information Age could blow away the blockbuster. *New Yorker* 75(29): 48-55. [Online]. Available: http://www.gladwell.com/1999/1999_10_04_a_sleeper.htm [Accessed: 17 October, 2003.]

Goodrich, Micheal & Atwill, Teresa. 2000. *Oregon Earthquake and Tsunami Curriculum: Grades Seven Through Twelve, Revised edition.* National Tsunami Hazards Mitigation Program, Oregon Department of Geology and Mineral Industries.

GroupLens Research Project. Movielens; Helping you find the right movie. [Online]. Available: http://movielens.org [Accessed: 17 October, 2003].

Herlocker, Jon, Jung, Siekyung & Webster, Janet. 2003. Collaborative Filtering for Digital Libraries. Technical Report 03-40-01, Department of Computer Science, Oregon State University. [Online]. Available: http://eecs.oregonstate.edu/library/files/2003-17/jcdl2003.pdf [Accessed: 17 October, 2003].

OSU Libraries Digital Library Team. Oregon State University Libraries System for Electronic Recommendation Filtering. [Online]. Available: http://dl.nacse.org/osu [Accessed: 17 October 2003].

Recker, Mimi M. & Walker, Andrew. 2003. Supporting "word-of-mouth" social networks through collaborative information filtering. *Journal of Interactive Learning Research* 14(1): 79-98.

Satake, K., Shimazaki, K., Tsuiji, Y. & Ueda, K. 1996. Time and size of a giant earthquake in Cascadia inferred from Japanese tsunami records of January 1700. *Nature* 379 (6562): 246-259.

Servicio Hidrográfico y Oceanográfico de la Armada de Chile, Departamento de Oceanografía, Programa de Geofísica Marina, Intergovernmental Oceanographic Commission & International Tsunami Information Center. 2002. *Earthquakes and Tsunamis: High School Teacher's Guidebook.* [Online]. Available: http://www.igf.fuw.edu.pl/kolo_naukowe/czytaj_pliki/tsunami/hsteacher.pdf [Accessed: December 5, 2003].

Tsunami Digital Library Team. The Tsunami Digital Library System for Electronic Recommendation Filtering. [Online]. Available: http://dl.nacse.org/ [Accessed: 17 October, 2003].

Appendix A: Survey of the Tsunami Researchers and Their Information Seeking Behavior.

1. Do you use Internet resources for your work or research? What types of tsunami information do you use/look for through the Internet? How do you use that information?

2. Are there specific tsunami related sites that you visit often for your work or research?

3. Do you use these sites because the information is timely, specific to your needs, etc? Please describe what makes that information useful.

4. What types of problems have you encountered accessing electronic information?

5. Do you know what a digital library is? Have you ever used a digital library before?

6. If you had access to a digital library dedicated to tsunami research would you use it?

7. What collections, information, resources would make a tsunami digital library most useful to you?

8. The Tsunami Digital Library (TDL) is a collaborative project. Its emphasis is fast, efficient access to quality information developed by a community of users and constantly reviewed by those users. Although it might take some of your time, being a partner in the TDL would allow the following opportunities:

    - the opportunity to help assure the quality of tsunami information within the digital library
    - the ability to recommend information, web sites, and agency materials for the tsunami digital library
    - access to site usage and function statistics

    Is this something that you find interesting?

9. Do you have any other comments about tsunami information, digital information in general, or access issues to digital information?