Agora: Enhancing Group Awareness and Collaboration in Floristic Digital Libraries

J. Alfredo Sánchez[1], Lourdes Fernández[1], John L. Schnase[2]
[1]*Laboratory of Interactive and Cooperative Technologies,*
Universidad de las Américas-Puebla
[2]*Center for Botanical Informatics, Missouri Botanical Garden*
{alfredo, lulu, jschnase}@cbi.mobot.org

Abstract.

Digital libraries can be regarded as virtual spaces in which collaborative scholarly research can be conducted. *Floristic digital libraries* provide such collaboration spaces for scientists working on solutions for Earth's biodiversity problems. However, group awareness and collaboration are not easily achieved in an increasingly distributed environment such as the virtual space in which digital library users (particularly botanists and biologists) do their work. We describe an environment that enables group awareness, communication, and collaboration among users in a globally accessible floristic digital library. This is achieved by extending existing library facilities with recommendation and alerting services, as well as various communication interfaces.

Keywords: Recommendation services, group awareness, agents, floristic digital libraries.

1. Introduction

Digital libraries comprise highly complex and dynamic information spaces on top of which a variety of services are provided for end users. These services are aimed at facilitating the use of available information resources for both individual and group knowledge-intensive activities. Inspired by visions of a worldwide digital library, researchers and practitioners around the world have started to develop a significant number of projects related to digital libraries, each aimed at addressing the needs of specialized domains and user populations [Fox and Marchionini 1998]. In this context, the Missouri Botanical Garden is leading the Floristic Digital Library Initiative (FDL), which provides virtual collaboration spaces for scientists and the general public utilizing collections of up-to-date botanical information.

A key aspect of digital libraries is their orientation to supporting collaborative work. As is common in many scientific disciplines, advancing botanical knowledge strongly relies on researchers being able to learn about related scholarly endeavors and, in general, on effective communication among peers. Indeed, botanical research is by nature conducted by scientists who are interested on inter-related problems and research questions but are physically distributed around the world. FDL provides an environment in which biodiversity studies can be shared among members of a global community, enabling and maintaining their ongoing discussion to reflect the state of the discipline. This paper describes *Agora,* a software component of the Floristic Digital Library that enhances the awareness of library patrons about users with common interests and generates communication and collaboration opportunities among them.

The remainder of the paper is organized as follows. Section 2 provides further details of the FDL context and some insight into botanical research practice and communication needs. Section 3 discusses the conceptual design of Agora, whereas Section 4 describes its current implementation. Finally, Section 5 presents the status of the project as well as the directions for ongoing and future developments.

2. The Floristic Digital Library Initiative

Harnessing the accumulating knowledge of Earth's biodiversity and the ecosystems that support it is one of the grand challenges for the 21st century. Plants are the source of food and oxygen for most living systems. They also provide other benefits to society such as raw material for shelter, clothing and medicines [Schnase *et al.* 1998]. Nowadays, plant species are disappearing at a faster rate than scientists are able to study them using traditional means of collecting and sharing information. Clearly, our best hope for solving this problem is to use information and communication technologies to support large-scale, distributed enterprises, change established work practices, and improve the efficiency and quality of information gathering and dissemination.

2.1 FDL Participants

The Floristic Digital Library Initiative (FDL) is an international enterprise that is taking advantage of developments such as the World Wide Web to reify a conceptual framework in which botanical research and collaboration can occur. Projects such as Flora of North America (FNA) and Flora of China (FOC), sponsored by the National Science Foundation, are participating in this effort.

Both FDL contributors (e.g. authors, editors) and patrons (e.g. biodiversity researchers, teachers, students, the general public) face a highly complex environment, but can expect a variety of services provided by the library to assist them in dealing with this complexity.

2.2 Environment for FDL Contributors

Building a floristic digital library requires a multidisciplinary and multi-institutional approach. Collaboration and communication in this context must account for specialized authoring and review processes as well as complex interactions among team members. The FNA project alone involves more than 800 scientists, including plant taxonomists in the US and other parts of the world, biologists in government agencies such as the US Forest Service, Bureau of Land Management, US Fish and Wildlife Service, and state conservation and biological survey offices. The FNA Editorial Committee consists of 34 plant taxonomists distributed throughout the United States and Canada. More than 30 institutions have committed staff time and facilities to the successful completion of the flora, which is expected to occur around 2006 [Morin *et al.* 1989; Schnase *et al.* 1997].

2.3 Environment for FDL Patrons

Similarly, end users of a floristic digital library are presented with the complexity of scientific data about plants (which include specialized taxonomic treatments and morphologic descriptions, distribution maps, plant illustrations and often dissenting taxonomic views). The richness of the digital medium also makes it possible for library patrons to interact with a very diverse user community, from the inexperienced, casual users looking for answers to basic questions, to the professional studying the ecological impact of a proposed construction development, to the botanical researcher extending the existing knowledge base about the flora of a given geographic region.

2.4 FDL Overall Architecture

Over the past three years we have developed a system architecture for FDL that addresses the needs for communication, collaboration and information management among the botanical community. In this architecture, library services are built on top of a distributed object repository comprising a rich mix of electronic documents such as textual descriptions, bibliographic references, maps, and illustrations, which together provide a comprehensive knowledge base of the current understanding of the plants of the world. Services are made available via interfaces that are provided for each type of user and for a variety of activities the library supports. We are particularly interested in exploring agent-based approaches to library services and user interfaces. The current architecture of FDL is an evolution of the system architecture presented by Sánchez *et al.* [1997], in which agent services play a key role. Among the services and interfaces we have developed are role-based views [Tomlinson *et al.* 1998], agent-assisted taxonomic treatment construction [Sánchez *et al.* 1998], agent-guided multi-taxonomic browsing [Flores 1997] and personalized information retrieval [Cabrera 1997].

2.5 The Need for Group Work Facilities

Users of the FDL are not expected to work in isolation. Much of the current interest in digital libraries springs from the opportunities for collaboration they enable [Lesk 1997]. In order for users to be able to take better advantage of the information and communication resources made available by the FDL, they should be able to communicate effectively with other library users. Users in the digital library need to know about the presence of other users, their level and areas of expertise, and the possible means to contact them. Experienced users in a floristic digital library may want to provide directions or recommendations to newcomers exploring a given taxonomic group. Inexperienced users, on the other hand, may benefit from expert recommendations and knowledge about authoritative opinions. The digital library should make this group awareness possible and should provide the necessary communication mechanisms, but at the same time it should protect the privacy of all users according to their individual preferences.

In the following sections, we describe *Agora*, a component of the FDL aimed at providing such group awareness and communication functionality.
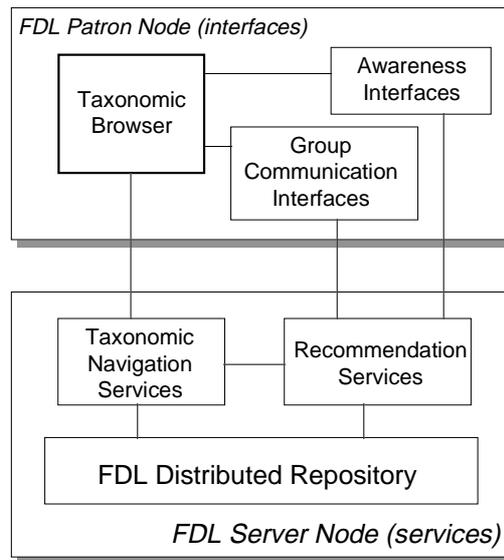
Figure 1. Components of the Agora environment in the Floristic Digital Library.

3. *Agora:* Conceptual Design

The main objective of Agora is to realize the potential of a digital library to become a *virtual gathering place* (hence the name) for a very diverse community. Naturally, virtual groups will emerge according to various criteria (including, for example, the users' interests or their level of expertise). Major design goals for Agora are the following:

- To enhance the group awareness of FDL users
- To improve the chances for communication among users with related interests
- To provide mechanisms for communication and collaboration among virtual groups emerging from regular use of the FDL

3.1 Components of Agora

As mentioned previously, Agora builds upon previous developments within the FDL architecture. Components of the architecture relevant to Agora are illustrated in Figure 1. For the sake of simplicity, various available FDL services and interfaces are not included in this diagram. *Taxonomic navigation* is one of the services provided by FDL that is required from number of different interfaces, including a *taxonomic browser*. Users of the botanical repository often get to the information they need by navigating through a taxonomic tree starting with a group of plants sharing a number of characteristics (a *taxon*) and moving from there to higher (more general) or lower (more specific) taxa. At each point in the tree, users have access to such data as taxonomic treatments, distribution maps, references to authoritative related publications and plant illustrations. Agora introduces three major components that work in conjunction with the taxonomic services and the taxonomic browser: recommendation services and communication and awareness interfaces.

3.2 Recommendation Services

First, on the FDL server side, Agora introduces the notion of *recommendation services.* Recommendation systems represent an emerging technology that promises to help users in utilizing otherwise unwieldy information spaces and connecting them with like-minded users [ACM 1997]. In Agora, recommendation services refer to an extensible collection of recommendation strategies available for use in client interfaces. For example, recommendations can be produced so users can be referred to new library elements which are similar to those the user has requested in the past. Similarly, users can be advised to examine library items other users with overlapping interests have requested. These two basic forms of recommendation have been termed *content-based* and *collaborative* recommendation, respectively [Balabanovic and Shoham 1997]. Recommendation services need to operate in close relationship with the taxonomic browser, since much of the success of recommendations depends on following the users' actions closely as they traverse the FDL. Agora allows users to define an interest profile that can be used to find other users in the library with similar interests. Experts in each taxonomic group can be identified and can provide and rank explicit recommendations for users browsing specific library areas.

3.3 Awareness Interfaces

The second component introduced by Agora is generically referred to in the diagram as *awareness interfaces.* As a user navigates through the FDL, he or she may access items for which a virtual group of users has already been formed, an expert or the author herself may be logged in at the time, or an explicit recommendation for related items has been made. Awareness interfaces bring these events to the users' attention and allow them to take immediate action to contact other users or groups of users or to access available recommendations.

3.4. Communication Interfaces

Indeed, a set of possible actions for the user in the events described above is enabled by the third component introduced by Agora, labeled *group communication interfaces* in Figure 1. FDL users are allowed to initiate or participate in synchronous or asynchronous individual or group communication. For example, if a scientist who is an expert in a given genus is present at the library while a student is examining that particular taxon, and the expert has specified in his profile that he is willing to take questions or engage in a live dialogue, then the student is presented with the option to contact the expert either synchronously (e.g. via an audio channel) or asynchronously (e.g. via email). Similarly, if a virtual group exists to discuss the item in question, the student is presented with the option to join the discussion, either synchronously (e.g. via a virtual room) or asynchronously (e.g. via a mailing distribution list).
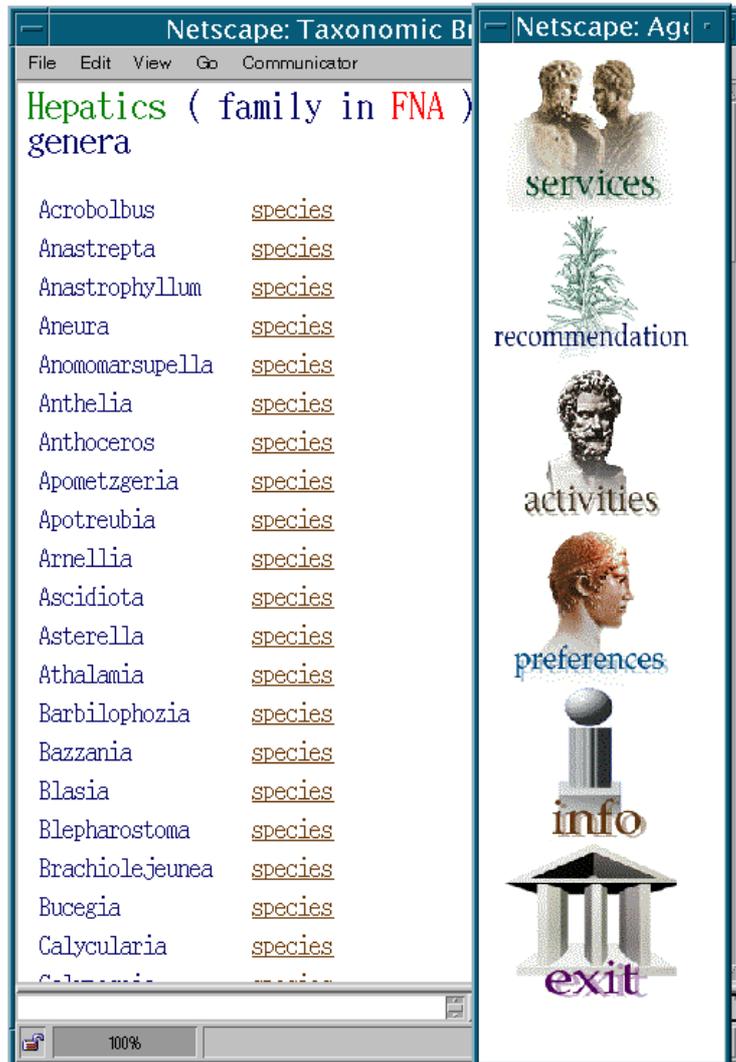
Figure 2. The Agora group awareness and collaboration environment.

4. Prototypical Implementation

We have developed a prototypical implementation for the components of Agora described above. This prototype of Agora extends previously existing library browsing functionality [Flores 1997] and introduces various necessary extensions. FDL is currently implemented using the Illustra object-relational database management system as the basic storage manager for its data repository. Taxonomic browsing consists of a collection of functions that allow for the traversal of multiple taxonomic trees. These functions can be currently accessed from popular Web browsers. Newly developed functionality has been implemented as C programs that access the FDL database via the CGI interface, Java applets and JavaScript programs that display new information and communication options as independent windows containing textual and graphical elements.

Figure 3. Communication services in Agora

4.1 Group awareness extensions

Agora makes use of the existing infrastructure and presents its services to the user as an extension to the taxonomic browser. After initiating a session with the taxonomic browser, the user is presented with the Agora group awareness options on a separate panel, as shown in Figure 2. Each of these options is described next.

4.2 Communication services

Upon selection of the first option (*Services*) the user is given the options to contact experts on the taxon being examined or related taxa. As shown in Figure 3, one of the options allows the user to send a message to the group of taxonomic experts registered in FDL (registration is discussed below). This functionality is implemented as a mailing distribution list. If some of the experts are present at the time in the library, the user may contact them live by entering a virtual room. Currently, this is implemented as a real-time, multi-user, text-based communication tool, as displayed in Figure 4. An alerting window pops up for those experts who specify in their profile that they may be contacted synchronously while at the library, and each user decides whether or not to join the synchronous discussion. A third option in Figure 3 allows the user to obtain a listing of all registered FDL users and to contact them either synchronously or asynchronously.
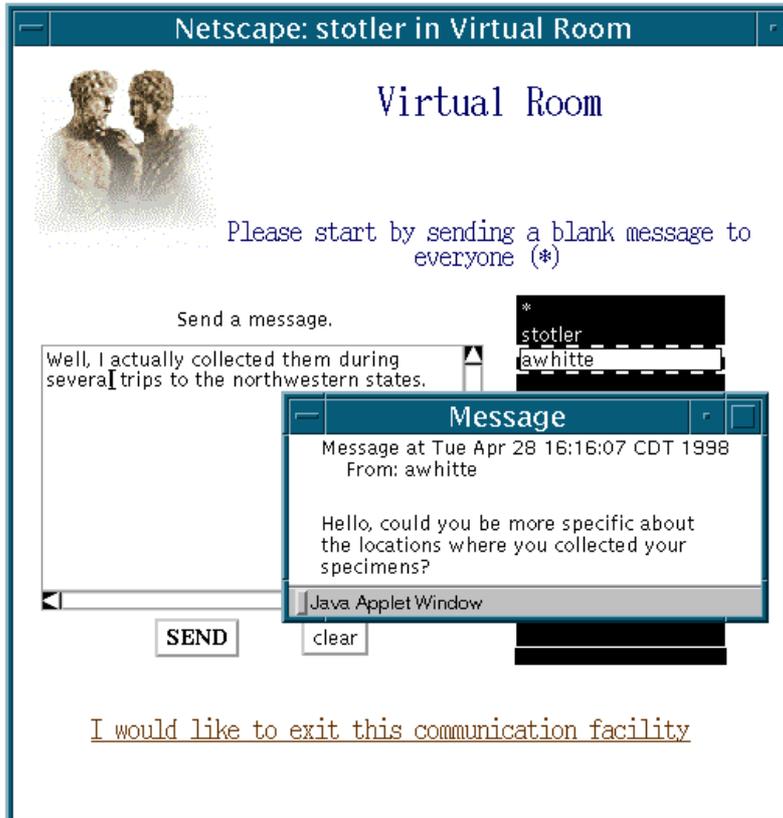
Figure 4. Agora's virtual room.

4.3 Recommendation Services

Agora allows users to obtain recommendations or pointers to related materials in the library suggested by experts or by any other users in the library. If the users so specify in their individual preferences profile, when a given taxon is reached for which a recommendation exists, an alerting window will pop up providing links to related materials, as well as a ranking provided by the recommender for the suggested links. In the example illustrated in Figure 5, an expert has suggested to check the genus *Athalamia* and the genus *Conocephalum* when the user reached the morphologic description for the species *Bucegia romanica*. The expert has indicated, however, that the former is a much stronger recommendation than the latter. The user also can see the level of expertise of the person who is providing a recommendation.

For some users, windows popping up in an autonomous fashion may seem too intrusive. In this case, the user may specify in the corresponding preferences profile that unsolicited recommendations are not wanted. The user can still use the second option of the main panel (Figure 2) to explicitly request recommended library materials related to a given taxon.
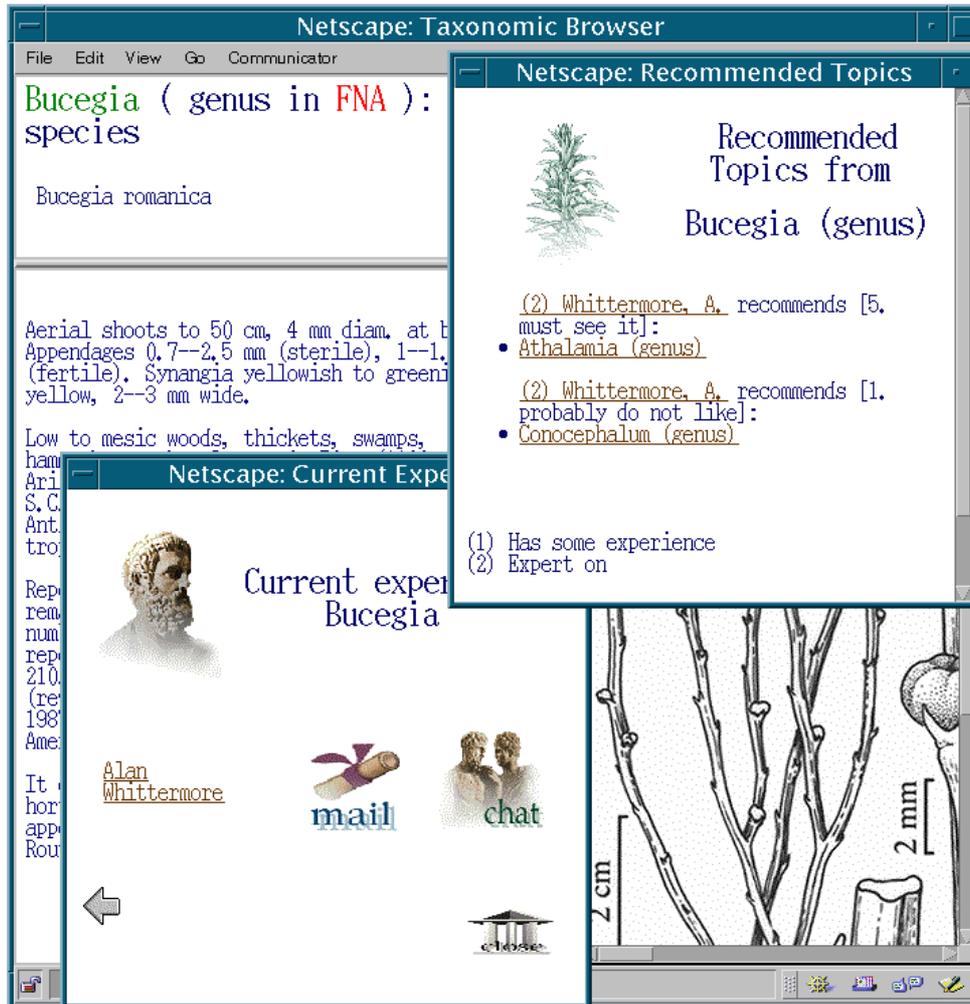
Figure 5. Recommendations provided by current Agora users.

4.4 Domain Expert Options

Much of the success of Agora relies on the willingness of each of the library community members to provide help to other users and to participate in the various available forms of user communication. At any point while browsing the library, users may use the third option of the main panel of Agora (Figure 2) to keep their personal suggestions and rankings up-to-date. Of particular interest for all users are the recommendations and rankings from domain experts. Figure 6 shows the options available to the users for updating their recommendations. The user may register as an expert in the current taxon (indicating a level of expertise), update an existing record of expertise, specify a recommendation and ranking for materials related to the current taxon, or update existing recommendations. Authors of taxonomic treatments are automatically considered experts in the corresponding taxa.
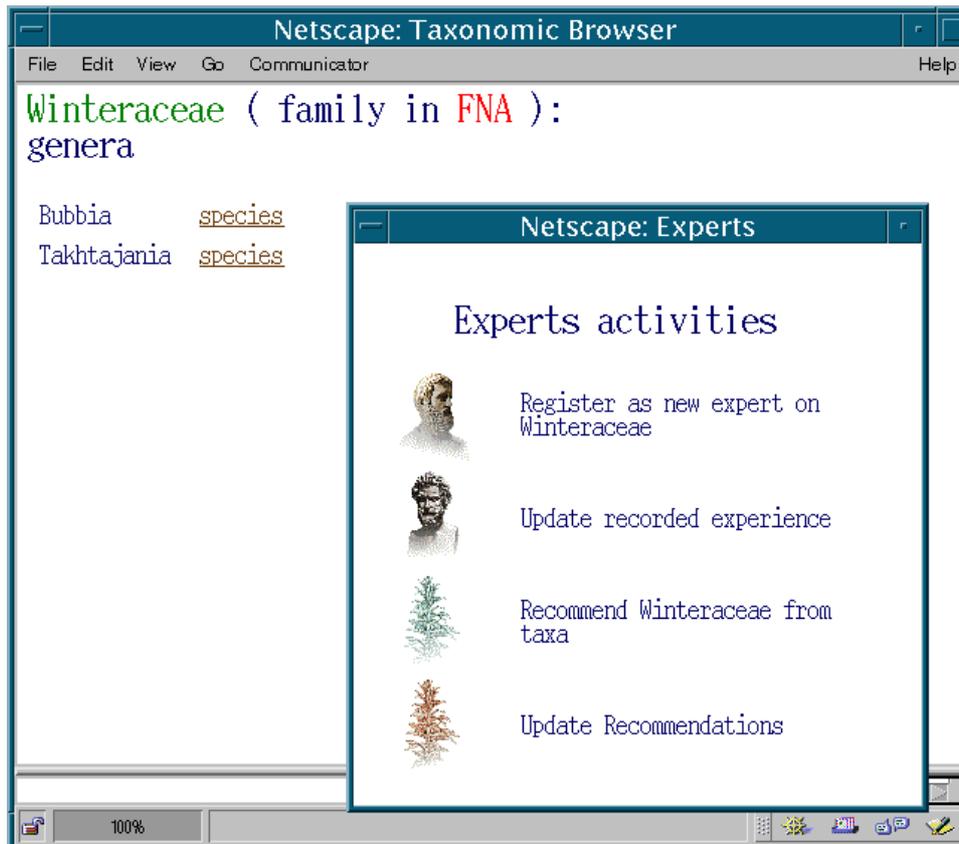
Figure 6. Recommending and keeping recommendations up-to-date.

5. Ongoing and Future Work

The prototype described in Section 4 demonstrates all major functions included in the design of Agora. The system (as part of an evolving digital library) is currently undergoing various tests in actual library use. Based on initial positive feedback and the success of related efforts [ACM 1997], we expect that Agora will contribute significantly to assist FDL users in using more fully the library resources (data, software and colleagues), discovering otherwise inaccessible information and relationships among library elements, and dealing with a highly complex and dynamic environment.

Work is still needed to realize the full potential of Agora as an enhancement for communication and collaboration in a digital library. As mentioned in Section 3, a number of recommendation strategies are amenable to inclusion as part of Agora's recommendation services. The current prototype allows for the operation of a social protocol in which users explicitly provide recommendations and rankings, as well as indication of their willingness to participate in discussions and collaborative activities. We are currently working on the implementation of strategies based on observed user behavior and similarity measures among user profiles to improve the quality and diversity of recommendations while reducing the amount of explicit input from users. Other areas of interest for future developments include

exploring audio channels and graphical representation for virtual rooms [Dourish and Bly 1992; Roseman and Greenberg 1996] as alternatives to the current text-based implementation.

Acknowledgments

References

ACM. 1997. Special issue on recommender systems. *Commun. ACM 40,* 3 (March).

Balabanovic, M., and Soham, Y. 1997. Fab: Content-based, collaborative recommendation. *Commun. ACM 40,* 3 (March), 66-72.

Cabrera, J. 1997. Integración de servicios y agentes de usuario en la recuperación de información en una biblioteca digital. Tech. Rep. ICT-97-2. Laboratory of Interactive and Cooperative Technologies, Universidad de las Américas-Puebla, Cholula, Puebla, México. (http://ict.udlap.mx).

Dourish, P., and Bly, S. 1992. Portholes: Supporting awareness in a distributed work group. In *Proceedings of CSCW'92*, 809-815.

Flores, C. A. 1997. Agentes de usuario como guías en bibliotecas digitales. Tech. Rep. ICT-97-3. Laboratory of Interactive and Cooperative Technologies, Universidad de las Américas-Puebla, Cholula, Puebla, México. (http://ict.udlap.mx).

Fox, E., and Marchionini, G. 1998. Toward a worldwide digital library. *Commun. ACM 41,* 4 (April), 29-32.

Lesk, M. 1997. *Practical Digital Libraries.* Morgan Kaufmann.

Morin, N. R., Whetstone, R. D., and Tomlinson, K. L. (Eds.). 1989. Floristics for the 21st Century, Monographs in Systematic Botany from the Missouri Botanical Garden. Vol. 28, Missouri Botanical Garden, St. Louis, MO.

Roseman, M., and Greenberg, S. 1996. TeamRooms: Network places for collaboration. In *Proceedings of CSCW'96,* 325-333.

Sánchez, J. A., López, C. A., and Schnase, J. L. 1998. An agent-based approach to the construction of floristic digital libraries. To appear in *Proceedings of the 3rd ACM International Conference on Digital Libraries* (DL'98, Pittsburgh, PA, July), *forthcoming.*

Sánchez, J. A., Leggett, J. J., and Schnase, J. L. 1997. AGS: Introducing agents as services provided by digital libraries. In *Proceedings of the 2nd ACM International Conference on Digital Libraries* (DL'97, Philadelphia, PA, July), 75-82.

Schnase, J. L., Lane, M. A., Bowker, G. C., Star, S. L., and Silberschatz, A. 1998. Building the next generation biological information infrastructure. In *Proceedings of the National Academy of Sciences National Research Council Forum on Biodiversity and Human Society: The Quest for a Sustainable World*. Washington, DC. (forthcoming).

Schnase, J. L., Kama, D. L., Tomlinson, K. L., Sánchez, J. A., Cunnius, E. L., and Morin, N. R. 1997. The Flora of North America digital library: A case study in biodiversity database publishing. *Journal of Networks and Computer Applications 21*, 1, 87-103.