DOCUMENT RESUME

ED 429 541 IR 019 449

AUTHOR	Payne, Deborah A.; Keating, Kelly A.; Myers, James D.
TITLE	Virtual Partnerships in Research and Education.
SPONS AGENCY	Department of Energy, Washington, DC. Office of Energy Research.
PUB DATE	1997-11-00
NOTE	7p.; In: WebNet 97 World Conference of the WWW, Internet & Intranet Proceedings (2nd, Toronto, Canada, November 1-5, 1997); see IR 019 434. Figure may not reproduce clearly.
CONTRACT	DE-AC06-76LO-1830
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Classification; Computer Interfaces; *Computer Mediated Communication; Computer Software; Computer Uses in Education; Cooperative Programs; Distance Education; *Educational Cooperation; *Institutional Cooperation; Research and Development; Science Laboratories; *Scientific Research; Shared Resources and Services; Teleconferencing; World Wide Web
IDENTIFIERS	*Collaborative Research

ABSTRACT

The William R. Wiley Environmental Molecular Sciences Laboratory (EMSL) at the Pacific Northwest National Laboratory (Washington) is a collaborative user facility with many unique scientific capabilities. The EMSL expects to support many of its remote users and collaborators by electronic means and is creating a collaborative environment for this purpose with capabilities ranging from chat and videoconferencing to shared applications, electronic notebooks, and remote-controlled instruments. This paper describes some of the particular capabilities required to support scientific collaborations, the status and direction of the EMSL tools, and several early uses of the EMSL software in both research and education collaborations. The first section presents a taxonomy of the types of research collaborations that currently exist (peer-to-peer, mentor-student, inter-disciplinary, producer-consumer) and evaluates the communications needs for each type. EMSL's real-time Collaborative Research Environment (CORE) is described in the next section, and the following capabilities/components are summarized: WebTour, file sharing, chat box, TeleViewer, Electronic Laboratory Notebook, on-line instruments, whiteboard, and audio/video conferencing. The third section describes the use of CORE in research and education settings. Together, these topics define a vision for natural, in-depth, virtual partnerships in research and education. A figure presents the CORE World Wide Web interface. (Author/AEF)

Virtual Partnerships in Research and Education

Deborah A. Payne, Kelly A. Keating, James D. Myers
Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory,
K8-91, P.O. Box 999, Richland, WA 99352, USA
da_payne@pnl.gov, kelly.keating@pnl.gov, jd_myers@pnl.gov

Abstract:
The William R. Wiley Environmental Molecular Sciences Laboratory (EMSL) at the Pacific Northwest National Laboratory is a collaborative user facility with many unique scientific capabilities. The EMSL expects to support many of its remote users and collaborators by electronic means and is creating a collaborative environment for this purpose with capabilities ranging from chat and video-conferencing, to shared applications, electronic notebooks, and remote-controlled instruments. This paper describes some of the particular capabilities required to support scientific collaborations, the status and direction of the EMSL tools, and several early uses of the EMSL software in both research and education collaborations. Together, these topics define a vision for natural, in-depth, virtual partnerships in research and education.

Introduction

The move toward virtual enterprises, seen in today's business world, is also occurring in the field of scientific research and education. National laboratories, such as the Pacific Northwest National Laboratory (PNNL), are making their data, instruments, and expertise, available to academic, industrial, and government collaborators, and conversely, are planning to make use of physical and intellectual resources at other institutions to supplement their own capabilities[WWW 97a][Kouzes et. al. 96]. Educators are looking to provide students with training in the latest techniques using state-of-the-art instrumentation as well as to motivate students' learning with cross-disciplinary examples of the use of science knowledge to solve real-world problems. PNNL's Environmental Molecular Sciences Laboratory (EMSL)[WWW 97b] is a new $230M facility for basic research in environmental and molecular sciences in support of the Department of Energy's mission to develop new technologies to clean up the nation's hazardous waste sites. The EMSL will house many unique facilities for basic scientific research, including the world's first commercial near-gigahertz Nuclear Magnetic Resonance (NMR) spectrometer, a scanning near field optical microscope, and the most powerful IBM parallel supercomputer yet built. Overall, the EMSL will house nearly 300 researchers with unique expertise, equipment, and software, seeking to understand the fundamental physical, chemical, and biological processes that underlie:

- the use of natural and engineered techniques to remediate and restore contaminated soils and groundwater,
- the processes and techniques used to extract and destroy chemical wastes, and to separate and safely store radioactive wastes,
- the development of a new generation of industrial processes that minimize or eliminate the use of toxic materials and the production of hazardous waste products, and
- the impact of toxic contaminants on the health of humans and the ecosystem.

While there are many research and education collaborations today involving EMSL researchers and their remote colleagues, the use of electronic collaboration tools promises to greatly enhance both the quantity and quality of such collaborations. However, simple videoconferencing tools are not sufficient to allow natural, in-depth collaboration, especially when the topic involves a complex scientific instrument, exploring multi-dimensional data, or synthesizing results from theory and experiment. Through interviews with EMSL researchers and their colleagues, and iterative feedback during software development, we've developed a simple taxonomy to describe the types of research collaborations that currently exist and have evaluated the communications needs in each type.

Some collaborations involve researchers in the same field sharing an instrument. The remote researcher might contribute to the design of a new detector and then use the instrument to study molecular systems of interest. In this **peer-to-peer** type of collaboration, the researchers share a common scientific vocabulary. The most important aspects of their collaborations are shared instruments and unanalyzed data, making remote instrument control and direct data file access important.

Other collaborations involve senior scientists and their more junior partners, such as students and postdoctoral fellows. In these collaborations, the mentor may use prepared materials and live demonstrations to teach data acquisition, analysis techniques, and scientific principles. The mentor must then observe as the student demonstrates mastery of the new concepts by using them appropriately. The necessary real-time interactions between mentor and student go far beyond standard conferencing: a mentor and student must be able to work collaboratively and interactively, sharing a view of an experiment in progress or the live output of a modeling/visualization package. In this **mentor-student** type of collaboration, real-time interactions are supplemented by asynchronous access to many types of archival information - data, notes, results, etc. This also allows the student to revisit the material as needed.

A third type of collaboration anticipated is between scientists doing complementary studies of the same molecular systems. For instance, a theorist may calculate structures of molecular clusters while an experimentalist uses laser spectroscopy to make an experimental measurement of the structure. Researchers in such **inter-disciplinary** collaborations share less of a common vocabulary and must often translate their results into each other's terms, alternating between the roles of mentor and student. Direct access to instruments or to raw data becomes less useful to the researchers, while access to summaries and analyses, perhaps recorded in an electronic notebook, and the ability to discuss unfamiliar concepts and to correct misunderstandings become more important.

A fourth type of collaboration, again involving researchers in different disciplines, involves one researcher, or research team, providing input for another. Examples of this type of collaboration include a mass spectroscopist determining the sequence of a protein or other biopolymer for a biologist, or a surface scientist providing reaction rate data to a geologist modeling the subsurface transport of hazardous wastes. Working with an analytical laboratory on a fee-per-service basis represents an extreme form of this **producer-consumer** type of collaboration. There is often a wider gap between the disciplines and motivations of researchers in such collaborations; a scientist may be interested in a new physical phenomenon while their collaborator, an engineer, is trying to reduce the cost of a clean up effort. They may have little chance for professional contact in their daily work or at conferences. Researchers in these types of relationships place the strongest emphasis on being able to receive a sample and information about it, and being able to transmit results back to the other party. However, new ideas and approaches can appear if these researchers communicate more closely. The EMSL and PNNL hold seminar series, workshops, and pizza dinner discussions, to foster this type of communication between basic and applied scientists. This suggests that if these researchers are provided with readily available tools for informal electronic discussions, their collaboration may become more complementary as they adjust their studies to incorporate new ideas from each other.

These collaboration types suggest a range of useful tools, from email, voice and video, to shared computer displays, remote instruments, and electronic notebooks. Different collaboration types, and different tasks within them, will stress different communications channels. Other aspects of collaboration, and scientific collaboration in particular, affect the design of a shared electronic environment. During any collaboration, communication naturally switches between media as appropriate. An electronic collaboration environment should allow someone to talk, shrug, draw a graph, and point at new data from an instrument, all with minimal awareness of having switched to a new tool. Similarly, collaborations may move through different phases - acquiring data, analyzing results, writing papers - that require different communication tools. A collaboration environment must support easy transition between tools as required. Lastly, most scientific collaborators have intermittent contact. A relationship may lie dormant for weeks or months and then enter a period of high activity after a new capability is developed or new data is obtained. Any collaboration environment must support this use pattern.

It is important to note that while these classifications and examples all relate to scientific research, similar collaborations arise in education and business. Students may ask professors for help while working in teams of peers on projects. Workers might have peer-to-peer collaborations within their organization, and mentor-student or producer-consumer collaborations with suppliers and customers. Thus, software that is designed to support scientific collaborations will be applicable in other domains as well.

The EMSL Collaboratory Tools

EMSL's real-time Collaborative Research Environment (CORE) provides users with a single, simple way to access multiple electronic collaboration capabilities independent of their computer platform. CORE has a World Wide Web (WWW) main interface and provides cross platform capabilities to the user via both new software developed for CORE and via existing stand-alone tools, or combinations of compatible tools, that have been integrated. CORE hides the different syntax each tool has for launching and connecting to collaborators, helping to make collaboration more natural. Users start and join sessions using their names and a short topic description. Computer addresses, port numbers, and firewalls: all disappear from the user's view.

CORE relies on a central session manager and desktop executives that coordinate communications between participants and configure the various collaborative components. Use of the WWW paradigm makes the system easy for users to understand. The main interface of CORE is a WWW page that allows users to start or join collaborative sessions via a WWW form. This page, shown in [Fig. 1a], uses a common gateway interface (CGI) script to process user input. To start a new session, the user enters their name in the "User Name" text box and a brief topic description in the "Session Name" text box, and clicks on the "Start a New Session" push button. To join an existing session, the user enters their name and clicks the button showing the desired topic in the "Active Sessions" list.

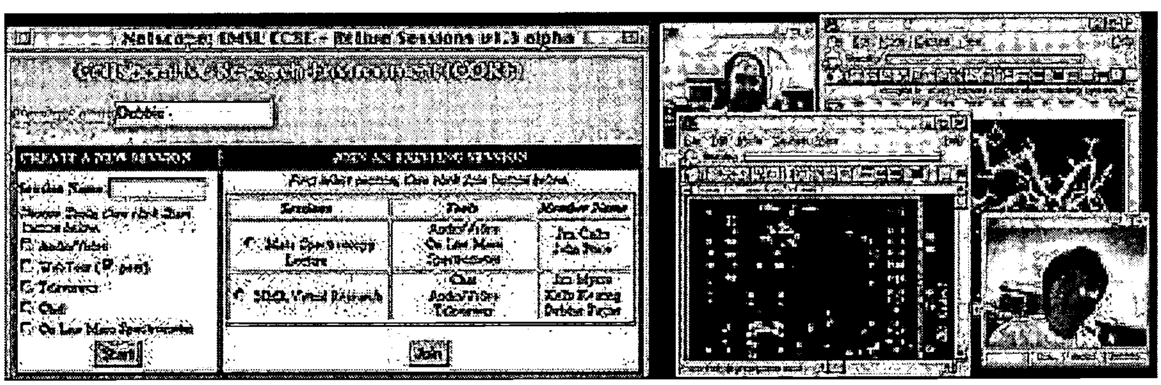

Figure 1. a) CORE's simple WWW user interface: select tools and start a session or click to join an existing one. b) Researchers discuss NMR data via CORE

When a new session is started, the user may select the tools desired for the given session. The session manager may start server processes for some of the tools, such as the EMSL TeleViewer described below. For other tools, such as videoconferencing, the user's IP address and platform type are used to determine the appropriate parameters for launching the client videoconferencing software. In our environment, we have implemented two third party options for audio/video conferencing. One is Cu-SeeMe, which we have implementing using a CU-SeeMe[WWW 97d] reflector bridge across PNNL's firewall. Macintosh and PC users use CU-SeeMe, connected to the appropriate end of the bridge to conference. The second option is use of multi-cast MBone[Eriksson 94] tools. Unix and PC users can launch MBone audio/video to either run independently or to connect to the reflector bridge. The session manager determines the appropriate parameters to launch software on each user's machine.

Once all the connection information is determined, and appropriate servers are started, the CGI script sends a custom multipurpose internet mail extension (MIME) typed file to the user's browser. The CORE desktop executive is started as a viewer (helper application) for this custom MIME file, just as a video player is started to “view” a video/mpeg MIME type movie file. The helper application was developed in Java as is the session manager. The executive prepares its own communications, either opening a listening socket, or connecting to a listening executive, and then launches the requested collaborative tools. CORE provides a basic set of tools, some of which have been developed as part of the Collaboratory project and are highly integrated with the CORE executive, and others that are the product of other EMSL projects and third party efforts and use their own communications once launched. A brief description of each of the capabilities follows:

A. WebTour: WebTour provides the ability to synchronize WWW browsers, allowing users to hold lectures or discussions, using material on the WWW. WebTour can be run in either lecture mode (only the leader's browser is echoed) or peer-to-peer modes. The WebTour functionality is embedded in the CORE executive and uses its communications to the browser and to other executives.

B. File Sharing: The CORE provides file sharing as an extension to the WebTour. Any local files opened in the user's WWW browser are transmitted to collaborators and opened with their browsers. Because it uses the WWW's browser/viewer mechanism, it allows remote users to choose different applications to view transferred files, i.e. users may choose different word processors to view a rich text format (RTF) file.

C. Chat Box: A simple chat box is included in the executive as well. Messages are tagged with the user names given when starting the session. Proper serialization is guaranteed by sending all messages to the central executive (the one that started the session) which then redistributes them to all executives in the session.

D. TeleViewer: The EMSL TeleViewer[Keller 96] provides a cross platform shared computer display. Users may select a rectangle or window from their computer, or their entire display to share with collaborators. Using this tool, users can view any program running on the shared display, such as word processors, spreadsheets, instrument control software, and mathematical computations. The TeleViewer will soon provide annotation on top of the live image and eventually the ability to remotely control the shared application.

E. Electronic Notebook: The EMSL Electronic Laboratory Notebook (ELN)[Myers 96] provides users with a shared, interactive version of the traditional paper laboratory notebook. The current system allows users to create secure, dynamic, searchable, WWW pages, organized in notebooks, with text, links, images (files or screen capture), live views of the data with information about each file (instrument parameters that were used, the operator's name, the date, etc.), etc. The notebook is easily extended to handle additional data types. For instance, we recently added the capability to view protein structures stored in the protein data bank (pdb) format by incorporating a third party Java applet. Data from EMSL instruments can be sent directly to the ELN, where it is immediately available for viewing, download, comment, and analysis by all collaborators.

F. On-line Instruments: Other projects within the EMSL are developing on-line instruments that can be run remotely via the internet. CORE provides a mechanism to select and launch this software as part of a real-time session, while the notebook provides remote access to the acquired data and other information. One of the first of these instruments is a remote enabled radio frequency ion trap mass spectrometer. Commercial instruments, such as the EMSL's Varian Nuclear Magnetic Resonance (NMR) spectrometers, which already have remote capabilities, are being integrated with CORE and the electronic notebook.

G. Whiteboard: Whiteboards provide a shared space where users can write and draw, on a blank canvas or over a preexisting image.

H. Audio/video conferencing: Audio/video conferencing allows collaborators to see and hear each other, as well as to monitor instruments and laboratories. CORE currently launches CU-SeeMe or MBone's 'vic' and 'vat', depending on the user's preference and platform. As part of the Collaboratory project, PNNL set up a CU-SeeMe reflector bridge across our firewall that allows conferencing between EMSL researchers and external colleagues, while managing security.

Collaboratory Use

CORE and/or the electronic notebook are being used by several groups, most of which consist of an EMSL researcher or group working with their remote colleagues, though some groups have no EMSL research connection. The interests of the groups range from software development to quantum chemistry, mass spectroscopy, NMR spectroscopy, and reactive transport modeling. Some groups are strictly research oriented while others are using the collaborative tools to provide student research opportunities or to bring instruments and remote experts into classrooms. There have also been many non-science demonstration and trial uses of CORE, ranging from business meetings/presentations, remote training, and rapid response intelligence analysis.

The two groups described below demonstrate the use of CORE and the notebook in research and education settings.

An NMR Virtual Research Facility project is using the Collaboratory tools to let PNNL and Lawrence Berkeley National Laboratory (LBNL) structural biology researchers work closely together to determine the solution structure of proteins and DNA molecules. The NMR data will be collected at PNNL on a Varian 750 MHz NMR spectrometer, a resource not available at LBNL. Once the sample is inserted into the probe by a PNNL researcher, experiments can be run locally, or remotely and securely via the internet. In the initial joint experiment, between Dr. Kelly Keating of PNNL and Dr. Jeff Pelton of LBNL, preliminary work included sharing background references, known structures of similar molecules, and project plans through their electronic notebook. Dr. Pelton learned the specifics of PNNL's 750 MHz spectrometer control software by virtually sitting in on one of Dr. Keating's experiments, viewing the spectrometer console in real time via the Televiewer, while discussing the experiment parameters via videoconferencing, and recording notes in the electronic notebook. For the first, and several subsequent data acquisition sessions, Dr. Pelton controlled the experiment remotely with Dr. Keating observing. During experiment runs, which could last for two days for two dimensional (2D) NMR spectra, either or both collaborators would log directly into the spectrometer to check the progress, and/or use CORE to share the progress report and discuss the experiment. The notebook allowed similar, asynchronous, discussions, with Drs. Keating and Pelton viewing and commenting on current 2D data slices posted to their shared notebook. Once data were acquired, the collaborators continued to use CORE and the notebook as they began processing the data and assigning signals to specific atoms in the molecule. During analysis, the notebook again allows each researcher to link a copy of their results to the relevant ELN page with screen snapshots and comments to guide the other's work. Collaborative sessions, using videoconferencing and the TeleViewer allow joint analysis to complete difficult assignments. After perhaps several months of this cycle of NMR data collection, accessing the data, and data analysis, the collaborators will begin to jointly write their results into a paper for publication, exchanging documents and figures via the notebook and discussing changes on the fly using CORE.

The Collaboratory tools have also been used to provide a remote lecture to Professor Jim Callis' Chemistry 155 class at the University of Washington. The students were given a quick mass spectroscopy tutorial via videoconference and the WebTour by Dr. John Price at the EMSL, and then used his ion trap mass spectrometer remotely to complete a laboratory assignment, comparing the calculated and experimental spectra of a molecule containing three chlorine atoms. Their data was instantly available to all participants via the WWW based notebook. The ion trap mass spectrometer and CORE tools have also been used in student research collaborations with the University of Washington and Heritage College. In these cases, students were able to participate remotely, over a long term, in publishable research projects involving their local advisors and EMSL researchers.

Conclusions

The Collaboratory has been developed over the past two years, a time in which the internet and the WWW have changed greatly. In particular, the emergence of Java and distributed object frameworks, such as the Common Object Request Broker Architecture (CORBA), promise to revolutionize the development of dynamic WWW interfaces. As the current generation of CORE and the ELN move into productive use by EMSL researchers, we are also moving into a new round of development, in concert with other national laboratories, as part of the DOE's DOE2000 Collaboratory project[WWW 97d]. DOE2000 will result in a very usable set of tools as well as an extensible architecture for the development of more advanced and more domain specialized collaborative tools. (DOE2000 also includes two pilot collaborations with distributed academic, government, and industrial participants that will use the tools to enhance their research.) An iterative development approach will help ensure that the Collaboratory tools will meet the needs of collaborating researchers. To provide a similar experience in educational use of collaboration technologies to link national laboratories and academic sites, the EMSL and eight northwest academic institutions have formed the Collaboratory for Undergraduate Research and Education (CURE) group [Myers et. al. 97]. This NSF and DOE funded group is developing ways to maximize the benefit to students of exposure to the data, instruments, and expertise of the laboratories through combinations of remote lectures and laboratory experiments, student research projects, faculty development, etc. A major goal of the project is to encourage a web of collaboration between academic sites and the laboratory that will scale much better than a multitude of one-to-one collaborations.

Collaborative environments, such as the EMSL's Collaboratory suite, can provide users with a powerful array of collaborative capabilities to support distributed scientific research and education collaborations. By hiding the complexities of configuring individual tools, and providing cross-platform capabilities, collaborative environments reduce the barriers to communicating with remote colleagues. Extensions to the standard videoconferencing tools such as the Televiewer shared computer display, remote instruments, and electronic notebooks, allow collaborators to bring scientific resources directly into their discussions. Such environments hold the promise of making work with remote colleagues as simple, natural, and effective as working with people down the hall.

Acknowledgements

This work was supported by the U. S. Department of Energy through the DOE2000 program and the Distributed Collaboratory Experiment Environments (DCEE) program, both sponsored by the Mathematical, Information and Computational Sciences Division of the Office of Energy Research, and through the Laboratory Directed Research and Development program at Pacific Northwest National Laboratory. Pacific Northwest National Laboratory is a multiprogram national laboratory operated by Battelle Memorial Institute for the U.S. Department of Energy under Contract DE-AC06-76RLO 1830. We gratefully acknowledge the contribution of many individuals to this work (see [WWW 97a]).

References

[Eriksson 94] *MBONE, The Multicast Backbone*, H. Eriksson (1994). Comm. ACM, Aug. 1994, pp. 54-60, also see http://www.mbone.com/mbone/mbone.faq.html

[Keller 96] *The EMSL TeleViewer: A Collaborative Shared Computer Display*, P.E. Keller, J.D. Myers (1996). IEEE Fifth Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, June 19-21, 1996, Stanford University, Palo Alto, CA

http://www.emsl.pnl.gov:2080/docs/collab/papers/televiewer.wetice.html

[Kouzes et. al. 96] R.T. Kouzes, J.D. Myers, W.A. Wulf (1996). *Collaboratories: Doing Science on the Internet*, IEEE Computer, 29:8.

[Myers et. al. 96] *Electronic Laboratory Notebooks for Collaborative Research*, J.D. Myers, D. Le, J. Laird, C. Fox-Dobbs, D. Reich, T. Curtz (1996). IEEE Fifth Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, June 19-21, 1996, Stanford University, Palo Alto, CA http://www.emsl.pnl.gov:2080/docs/collab/papers/notebook.wetice96.html

[Myers et. al. 97] *Collaboratories: Bringing National Laboratories into the Undergraduate Classroom and Laboratory via the Internet*, Jim Myers, Norman Chonacky, Thom Dunning, Eric Leber (1997), Council on Undergraduate Research Quarterly, March 1997

[WWW 97a] *Collaboratory for Environmental Molecular Sciences*, Pacific Northwest National Laboratory http://www.emsl.pnl.gov:2080/docs/collab/

[WWW 97b] *W. R. Wiley Environmental Molecular Sciences Laboratory*, Pacific Northwest National Laboratory http://www.emsl.pnl.gov/

[WWW 97c] *CU-SeeMe Welcome Page*, Cornell University, http://cu-seeme.cornell.edu/

[WWW 97d] *DOE2000 Project Homepage*, http://www.mcs.anl.gov/DOE2000

NOTICE

REPRODUCTION BASIS

☒ This document is covered by a signed "Reproduction Release (Blanket) form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").