The Research CYBERSTUDIO: Supporting Researchers as LifeLong Learners

proposal for a three year project

by Gerry Stahl

Center for LifeLong Learning & Design
and Institute of Cognitive Science
University of Colorado at Boulder

The Research CYBERSTUDIO Project addresses the problem of training and supporting learners at the most advanced end of the educational system to be skilled interdisciplinary researchers. It targets graduate students who have completed extensive classroom study within a discipline but who could benefit from practical research experience within a supportive context. The project's theoretical perspective of lifelong learning postulates that people in knowledge-intensive endeavors need to be continuously developing skills and constructing knowledge, and that this can be facilitated by *information delivery technologies within supportive collaborative contexts*. Accordingly, the proposed project approaches its goals from a technological and organizational approach: creating structured communities of learners (research studios), and providing adaptable computer-based support (the CYBERSTUDIO) for these communities. The technical approach builds on innovative software prototypes by the PI and collaborators; the organizational approach leverages substantial local opportunities. While the project is designed to assist novice researchers, it will develop computer software useful to interdisciplinary research communities generally.

The Problem of Training and Supporting Researchers

Research is an important aspect of contemporary universities like the University of Colorado (CU). Increasingly, much of this research is taking on an *interdisciplinary* character, spawning special groups like the Institute of Cognitive Science (ICS) and the Center for LifeLong Learning and Design (L^3D) at CU.

Despite a broad national effort to reform education from kindergarten through college and a significant attempt to develop computer support for education, little has been done to address the educational and computer needs of the most advanced students. The transition from an educated domain specialist to a skilled researcher is a lengthy and haphazard process, largely because the student is given little systematic support (Denning 1992). It is assumed that once students have completed their graduate course work they are capable of pursuing dissertation and post-doc research with minimal pedagogical support. However, experience within ICS and L^3D shows that fledgling researchers need to continue developing their skills in reading, writing, and mathematics just like students at any level. As they become involved in investigating problems that spill outside the discipline of their academic training, they need to learn to read broader professional literature, to prepare journal articles or conference presentations, and to master new methodologies (statistical evaluation, experimental design, computer modeling, discourse analysis, etc.). Whereas most professionals have specialized productivity software at their command, interdisciplinary researchers lack such tools.

The Research CYBERSTUDIO (RCS) Project adapts current constructivist educational theories to the problem of training researchers. In particular, the following pedagogical principles underlie the Project's approach:

1. Knowledge is constructed within communities of learners (Scardamalia & Bereiter 1994).
2. The approach of a design studio provides an effective setting for learning (Schön 1987, 1983).
3. Learning takes place through a person's increased participation in a community of practice (Lave & Wenger 1991).
4. Individual understanding can be fostered by appropriate computer-based systems (Papert 1993, 1980).

Based on these principles, the Project conceptualizes the problem of training novice researchers in the following terms:

1. The learner is viewed as a newcomer within a research community, as an apprentice who needs thoughtful mentoring.
2. Collaborative research activities are organized into a "research studio" structure in which individual and group projects are conducted and critiqued.
3. Learners are assisted in gradually participating more and more in their research community to acquire the tacit skills of their profession.
4. Special software acts as a "CYBERSTUDIO" in which community members communicate and contribute, work and learn.

The uniqueness of the RCS Project lies in the creation of computer support to promote lifelong learning and to manage organizational knowledge within a research community. The CYBERSTUDIO software will therefore be described first. Then its usage by the community will be discussed.

The CYBERSTUDIO Software

A central hypothesis of this project is that computer support can play an important role in developing interdisciplinary research skills as well as in accomplishing the research itself. The challenge of the project is to create new software adequate to the attainment of this potential.

The lack of computer support to deliver information on an as-needed basis is endemic to interdisciplinary research in general, not merely to novice researchers. The theories of situated learning and knowledge construction suggest that there is a significant untapped potential of computer support for building communities of learners and for capturing group memories to inform newcomers. Such knowledge sharing software could be particularly helpful in the interdisciplinary context, where community members speak different technical languages.

Project participants will design, prototype, explore, and evaluate CYBERSTUDIO's software to support learning, communication, and work within interdisciplinary research groups. This software will be designed to meet the information and collaboration needs of researchers, especially novices. The success of the community-of-learners approach requires a high level of communication and organization; CYBERSTUDIO will provide a medium in which this can take place. The software will also identify and deliver relevant ideas from the extensive and growing writings of the group and related published literature, allowing people to share ideas across time and space (Stahl et al. 1995a, 1995b).

Here is an illustration of how the Internet-based CYBERSTUDIO software can be used: Suppose that a graduate student drafts a thesis proposal for software to categorize the content of Internet sites by reading level. The proposal text is analyzed within CYBERSTUDIO. The software delivers a list of web links pointing to the most directly related excerpts from cognitive science papers, key terms in an interactive, multidisciplinary glossary, specific entries in threaded discussions within the research groups, email messages on the topic, and contact information for people in various disciplines who have done relevant work. The student can then review and respond to any of this information. For instance, the student might compile a set of notes with hypertext links to several of the retrieved sources, make annotations to the sources (for future users to read), send messages to referenced people. Then the student can revise the proposal draft and resubmit it to CYBERSTUDIO to obtain a refined list of relevant information. All of this is done within the CYBERSTUDIO system.

While the CYBERSTUDIO repository of information is primarily directed internally to the research group, it also includes external links to web sites globally and it allows outsiders to view many materials in the network. Thus, it provides a medium of communication and documentation within a local research community while participating in the broader discourse of the World Wide Web.

The CYBERSTUDIO software system envisioned to support working and learning by interdisciplinary researchers unifies three technologies that the PI and collaborators at L^3D and ICS have been exploring for many years:

1. Domain-oriented design environments (DODEs).
2. Dynamic web sites (DynaSites).
3. Latent semantic analysis (LSA).

These technologies will be integrated into CYBERSTUDIO network of research information services. CYBERSTUDIO captures knowledge as it is constructed within a research group and delivers items from this organizational memory when they are relevant to the new research of individuals, particularly newcomers.

1) Domain-oriented design environments. A DODE is a software application within which a professional conducts work. As the work progresses, the software responds by delivering domain-specific or community-historic information stored in its knowledge base that is relevant to informing the current state of the work (Fischer et al. 1993a, 1993b). By integrating working and information delivery, DODEs support lifelong learning or learning-on-demand.

2) Dynamic web sites. DynaSites are web sites that provide an interactive interface to a database of information shared by a group (Stahl 1997a). Based on intranet technology, they transform the World Wide Web from a generic broadcast medium to a group memory that allows collaborators to share their knowledge asynchronously. For instance, people in a research community can use their web browsers to find past discussions of ideas, glossaries of terms, and papers published by other members; as they generate new ideas, concepts, and essays using these resources, the new knowledge is added to the group memory interactively.

3) Latent semantic analysis. LSA is an automated technique for analyzing the semantic relations within a large corpus of text (Landauer & Dumais 1997). LSA can compare documents and rate the similarity of their technical content. When used properly, it can be effective for such tasks as evaluating the knowledge content of essays. Thus, it can be used by software to judge which of several student essays is most similar to a target

essay. Experiments have shown that LSA is approximately as reliable as people in grading SAT essays and in selecting readings that are most appropriate for a given reader based on an essay by that reader (Wolfe et al. 1997).

The CYBERSTUDIO software developed by this project will incorporate information sources relevant to the research of the interdisciplinary groups involved (e.g., L^3D and ICS). This includes both archival materials (published papers, technical reports, dissertations, seminar presentations) and process artifacts (on-going threaded discussions, email exchanges, meeting schedules or minutes, evolving glossaries of technical terms, annotated bibliographies, member information, etc.).

The project involves the development of techniques for capturing, structuring, evolving, retrieving, and presenting the information in CYBERSTUDIO. These techniques will include LSA applications (including the semi-automated production of a glossary of interdisciplinary technical terms in a corpus linked to key document excerpts), group perspectives (Stahl 1997a, 1995a, 1993a, 1993b), and a visual end-user language for querying and navigating the information base (Stahl 1993b, 1992a, 1991). Development of these techniques will be staged during the RCS Project period, with the glossary developed in Year I, perspectives in Year II, and the end-user language in Year III.

The Organizational Approach

The research studio is an approach to training through self-directed hands-on experience. Architecture students, for instance, spend a lot of their class time in studio classes, where they work on individual or group projects and receive critiques from peers and experienced designers. The project's research studios build on this model. Novice researchers will pursue their own dissertation research or participate in funded research within a community of learners, including both more and less experienced researchers. In addition to interacting informally and making formal presentations, people will share and co-construct ideas in settings such as reading groups, project meetings, and on-line discussion threads. Much of the communication associated with research studio activities will take place within the CYBERSTUDIO system and will be captured by it. Then, future newcomers can review the materials to learn relevant aspects of the group's intellectual history.

The project will investigate effective ways of structuring interdisciplinary research groups. This includes issues of physical office arrangements, meeting procedures, study groups, communication channels, and decision making. Like most professional workers, graduate students have too much to do; they must resolve conflicts of course work vs. research, individual projects vs. group efforts, meeting vs. working, learning vs. producing. Some of these conflicts can be ameliorated via institutional solutions such as adjusting requirements and reward structures. The framework of a research studio will be explored as a way of integrating research practice into the academic reward system, so involvement in group activities does not detract from personal achievement.

Apprenticeship or mentoring is important to the studio model of situated learning. A more experienced person provides systematic guidance or facilitation of student self-directed learning, and the student learns by working alongside old-timers. The mentoring relationship—just like the research studio—must be institutionally recognized in order to be effective. The project will investigate how this can be accomplished. It will start by formalizing apprenticeship relationships in the sense that they will be explicitly

recognized within the group. Both the mentor and the apprentice will receive recognition for their work together. An on-going dialog concerning benefits and problems of apprenticeship will evaluate this approach.

Because the RCS project is itself an interdisciplinary research effort involving reflective practitioners, the participants in the investigation will have the task of evaluating their own learning. They will incorporate evaluation methodologies from multiple disciplines (educational evaluation, psychological controlled studies, software engineering debugging, user testing, etc.). Assessment will itself be a topic of research—how to evaluate support for lifelong learning and interdisciplinary research in naturalistic settings.

The Local Context

This project addresses the problem of training and supporting researchers from within an exceedingly rich context of growing interdisciplinary research at the University of Colorado (CU). It will take advantage of considerable independent resources from federal, foundation, and university sources and focus them on the needs of interdisciplinary researchers.

Project level. The PI is currently directly involved in two interdisciplinary research projects: an effort to develop computer-based *organizational memories* and one to develop *educational software*. The first is sponsored by the Center for LifeLong Learning and Design (L^3D) and the second jointly by L^3D and the Institute of Cognitive Science (ICS). The organizational memory project integrates ideas from learning theory, anthropology, and organizational theory as well as various aspects of computer science and particular application domains. The educational software project involves issues of psychology and linguistics as well as computer science and education. During Year I, the RCS Project will encompass the teams of graduate students, post-docs, and visiting researchers working on these two projects.

Center level. ICS is an interdisciplinary institute by the nature of cognitive science; it expects to become an accredited interdisciplinary degree program in the next year. Within ICS, there is an active research group exploring latent semantic analysis (LSA). LSA is a statistical text analysis method with promising applications to practical problems in educational software as well as theoretical implications within cognitive science. The LSA research group includes cognitive psychologists, computational linguists, and computer scientists. L^3D is a center under both ICS and CU's Department of Computer Science, with strong involvements in education and environmental design. It encompasses projects developing conceptual frameworks and prototype software for applications in a variety of domains. L^3D and ICS members teach undergraduate and graduate courses in computer science, design, and cognitive science. Students in the research groups and courses within L^3D and ICS will provide the focus for Year II of the RCS Project.

University level. The CU administration is promoting the notion of a "total learning environment." As part of this commitment, L^3D is establishing a broader interdisciplinary initiative across many departments of CU—the Center for Interdisciplinary Research on LifeLong Learning (CIRLL)—likely to be funded as an NSF center next year. CIRLL will directly support seven graduate research assistants and four post-docs, as well as coordinating the work of many more novice and experienced

researchers across campus. In addition, L^{3}D has a growing network of industrial partners; students intern at the companies and company employees spend time at L^{3}D's research labs. In Year III, the RCS Project will expand to include novice interdisciplinary researchers in CIRLL and among the industrial interns.

Broader impact. It is anticipated that the lessons learned in the RCS project—pedagogical approaches, organizational supports, and computer software designs—will be disseminated beyond CU through research contacts at key centers like the Cognitive Studies of Interdisciplinary Communication program in the National Institute for Science Education at the University of Wisconsin, as well as through academic publications. Within CU the potential for dissemination is unlimited, with CU's focus on "total learning," its efforts to promote Internet support for teaching, its interest in distance learning, and its standing as a major research center.

Conclusion

The Research CYBERSTUDIO Project will explore organizational, pedagogical, and technological approaches to train advanced graduate students to be skilled interdisciplinary researchers. It will develop CYBERSTUDIO software to support the work of research groups. Gradually expanding its scope, the project will take advantage of a lively and growing community of interdisciplinary research at the University of Colorado. The approaches and software developed will be thoroughly evaluated, clearly documented, and broadly disseminated.

References

(See Biographical Sketch below for references to the PI's publications.)

Denning (1992) "Educating a New Engineer." *Communications of the ACM*. **35** (12). pp. 83-97.

Fischer, Nakakoji, Ostwald, Stahl & Sumner (1993) "Embedding Computer-Based Critics in the Contexts of Design." *Proceedings of InterCHI '93. Conference on Human Factors in Computing Systems*. pp. 157-164.

Fischer, Nakakoji, Ostwald, Stahl & Sumner (1993) "Embedding Critics in Design Environments." *The Knowledge Engineering Review*. **4** (8). pp. 285-307.

Greeno (1993) "For Research to Reform Education and Cognitive Science." In Penner, Batsche, Knoff, Nelson (Eds.) *The Challenge in Mathematics and Science Education*. APA. pp. 153-193.

Landauer & Dumais (1997) "A Solution to Plato's Problem: The Latent Semantic Analysis Theory of Acquisition, Induction and Representation of Knowledge." *Psychological Review* (forthcoming).

Lave & Wenger (1991) *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.

Papert (1993) *The Children's Machine: Rethinking School in the Age of the Computer*. Basic Books.

Papert (1980) *Mindstorms: Children, Computers and Powerful Ideas*. Basic Books.

Scardamalia & Bereiter (1994) "Computer Support for Knowledge-Building Communities." *Journal of the Learning Sciences*. **3** (3). pp. 265-283.

Schön (1987) *Educating the Reflective Practitioner*. Jossey-Bass Publishers.

Schön (1983) *The Reflective Practitioner: How Professionals Think in Action*. Basic Books.

Wolfe, Schreiner, Rehder, Laham, Foltz, Kintsch, Landauer (1997) "Learning from Text: Matching Readers and Text by Latent Semantic Analysis." Unpublished manuscript.