End-User Issues Should Have a First Class Status

Terence R. Smith
University of California at Santa Barbara

The time has come to treat both end-users and knowledge about end-users as *first class entities* in the development of electronic information environments that support research and learning. First class status in this case implies that they are as much an object of research and development as the information technology itself.

The focus of the report *Revolutionizing Science and Engineering Through Cyberinfrastructure* (Atkins et.al., 2003) is the construction of a layer of "enabling hardware, algorithms, software, communications, institutions, and personnel". This cyberinfrastructure layer would lie above a hardware layer supporting computing, storage, and communications but below a layer of "software programs, services, instruments, data, information, knowledge, and social practices applicable to specific projects, disciplines, and communities."

As the report notes, "technology has crossed thresholds that now make possible a comprehensive cyberinfrastructure." Hence technology is probably sufficient for supporting the layer(s) above the cyberinfrastructure layer which would perhaps include the transparent scholarly communications environments denoted in the Background Paper for the workshop as Ubiquitous Knowledge Environments (UKEs.) Given Moore's Law and its extensions, technology will soon, if not already, be able to support analogous higher layers in which research and learning activities are not subject to the current heavy costs of using digital technology.

There is one and only one justification for the cost of developing the cyberinfrastructure necessary for UKEs: *its ultimate, net value to communities of end users.* The most fundamental measure for the net value of DL environments in supporting the discovery, learning, and application of knowledge derives from their *operational* use by communities of users. The Cyberinfrastructure Report is, however, naturally more focused on the technological aspects of cyberinfrastructure. It suggests that *operations in support of users* is only one of three sets of project activities although stating that "operations is a competition for users" and that a "primary point of post-evaluation should be the satisfaction of the users who are served, and to a lesser extent the number of users who are service, based on input from the user community."

The cognitive and physical activities of end-users, however, are both the most important and least understood component of any UKE. Yet we lack a deep and coherent body of applicable knowledge about the cognitive and physical activities of end-users in research and learning environments. The applicable knowledge that we possess about end-users is fragmented across a variety of disciplines and practices including, for example, Interface Design, Human Factors, Cognitive Psychology, User Evaluation, and Education.

A systematic and *applicable* understanding of how researchers and learners in any schol-

arly environment discover, learn, and apply information is surprisingly scarce, given the enormous literature on human perception, cognition, and behavior. For example, an understanding of the cognitive or collaborative activities by which scientists generate new *models* is clearly in the early stages of evolution (see, for example, Dunbar 1999) and has yet to see much practical application. A generally acceptable model of *cognitive representations* of scientific concepts is also at hypothetical stages of development (see, for example, Gardenfors 2000), despite the recent focus of effort on concept spaces in learning applications of DL technology (see, for example, Smith et. al. 2003.) Similarly, an understanding of the role of *multimedia presentations* in learning is in the early stages of being applicable, despite the importance that DL researchers have placed on multimedia. While Mayer (2001), for example, presents approximately a dozen empirically-established *heuristics* for making multimedia presentations effective (some of which are counter-intuitive) these hardly constitute a general theory adequate for the design of multimedia presentations in educational settings.

The lack of applicable and systematic knowledge has greatly affected the development of DLs and their natural extensions. End-users have been of significant, though largely second class, concern in a large proportion of projects funded under DLI-2, DLI-2, NSDL, DLESE, and PACI. An integrative and analytical study of how end-users and a knowledge of end-users have been integrated into these projects is not yet available. One may, however, informally categorize the large proportion of activities that have been focused on end user issues as involving: (1) proposal writing in which proposed end-users and support for their activities are given as the rationale for the project; (2) informally-obtained user characterizations and requirements; (3) integration of end users as experimental subjects in prototype testing; and (4) typically informal and non-comprehensive user evaluations.

Even projects whose activities were based on supporting large, active bases of "real" end-users have faced difficulties in incorporating knowledge of users into their design and development and in carrying out compelling, informative investigations of users. Such difficulties have typically arisen as a result of our lack of systematic knowledge about end-users and a lack of funding to support deep studies of end-user related issues. While the Alexandria Digital Library (ADL) Project, for example, crystallized around an operational Map and Imagery Library supporting approximately 40,000 client transactions/day, it began in 1994 (see www.alexandria.ucsb.edu) with a relatively informal modeling of user requirements as perceived by members of the library staff. Its more recent development of learning environments is based on similar based on relatively informal models of the learning requirements of undergraduate students (smith et. al., 2003.) Despite the the liberal use of workshops, consultants, and user-evaluation studies, many of ADL's design decisions were based on relatively informal knowledge concerning users and their responses.

Given how little we understand of such issues, it is reasonable to characterize the design and development of transparent information environments for researchers and learners as as informal, ad hoc, and occurring by trial-and-error. It is also reasonable to assert that a knowledge of end-users sufficient for the effective design and use of DLs, UKEs, and analogous environments *has not passed thresholds that are analogous to the thresholds of technological knowledge enabling the construction of cyberinfrastructure.* Acquiring such

knowledge, or even an understanding or what should constitute a body of such knowledge, is clearly a major undertaking.

It is suggested that an an investment of the same order of magnitude as that for cyberinfrastructure ($1B) is warranted to support an integrated program of research and development to create a coherent body of knowledge and practice about end-users. Such knowledge should be applicable and integrable in the development and use of advanced information environments that support research and learning in the sciences and the arts. A minimum of activity that *should* be supported is in developing an equivalent of the Blue Ribbon Advisory Panel Report for cyberinfrastructure. Such a report would provide a map to the knowledge of end-users that is currently fragmented across many disciplines and practices, as well as a proposed path for the research and development necessary to support rational, successful designs for cyberinfrastructure.

The goal of such a report would be to initiate a program analogous to, concurrent with, and integrated with a cyberinfrastructure development program. At the most general level, the foci for such a complementary program might include developing:

- an applicable understanding of the cognitive and collaborative activities of researchers and learners that are reasonably independent of underlying support technology and yet provide useful constraints on design and implementation (e.g., *how are new research ideas generated?*);
- an appropriate and applicable *theory* of how science and learning occur (e.g., *what are appropriate levels of granularity at which scientific knowledge is represented?*);
- an understanding of how researchers and learners respond to specific technologies in operational UKEs developed perhaps by monitoring, investigating, and evaluating their activities in *operational* environments (e.g., *how do researchers respond to machine learning procedures as tools for discovering new concepts?*);
- a knowledge of how to design DLs, UKEs, and other information environments in ways that maximize the efficacy of end-user activities (e.g., *what display environments, beyond current screen displays, are effective for learning?*);
- support for end-users, including educational programs in the utilization of services supported by any cyberinfrastructure (e.g., *are interface "agents" useful methods for providing education?*)

It may well be that an "end-user" project involves a significant component of large-scale experimental research in operational environments. By operational environments we mean sustainable sets of collections and services that while supporting the activities required in in research and learning do not detract significantly from end-user primary tasks of research and learning (i.e., they pass some threshold of usability for the the user.) The conundrum here is that we need operational information environments to understand what makes information environments operational.

Clearly the issues raised in trying to understand end-user issues, in ways that are "deep" and at the same time applicable, are immensely difficult. It is a project for which NSF is well-suited as a sponsor.

As a closing comment on such difficulties, the author notes that fifty years ago, the problem of rational procedures for designing effective medicines on the basis of a deep theory was considered to be a problem of immense difficulty and perhaps impossibility.

REFERENCES

Atkins, D. E., et. al., 2003. Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue Ribbon Advisory Panel on Cyberinfrastructure (http://www.communitytechnology.org/nsf_ci_report/.

Dunbar, K. (1999). The Scientist InVivo: How scientists think and reason in the laboratory. In Magnani, L., Nersessian, N.,& Thagard, P. Model-based reasoning in scientific discovery. Plenum Press

Gardenfors, P. (2000). Conceptual Spaces, Bradford Books, MIT Press

Mayer, R. E. (2001). Multimedia Learning. New York: Cambridge University Press.

Smith, T. R., et. al., 2003. The ADEPT Concept-based Digital Learning Environment. European Conference on Digital Libraries, Trondheim, Norway.