STRATEGIC AND TACTICAL USES OF THE WEB: THE STANFORD UNIVERSITY LIBRARIES PROGRAM

J W Coleman

Rapporteur: Dr Chris Phillips

Strategic and Tactical Uses of the Web: The Stanford University Libraries Program

Jim Coleman
Head, Academic Computing for the Humanities
Stanford University
Stanford, California

For the Stanford University Libraries/Academic Information Resources, an organization that combines both library and computing professionals, the Web offers a premier environment for developing a strategy for transforming old library paradigms into their newer, evolving counterparts. At the same time we view the Web as the premier tactical environment for satisfying the demands of our community, while fulfilling our mission and institutional obligations.

My presentation today will focus on the context of the change that is going on in the library at Stanford as it moves from "library old" to "library new", to examine some of the underlying strategic assumptions that accompany that change, to see how the Web and its uses at Stanford has become the principal level for change and examine some of those Web resources, and lastly to consider briefly the interactions between users, staff and technology, and how these interactions may themselves be evolving.

Library Old/Library New

In the past, the library has fulfilled a variety of roles and functions. Traditionally, the library has been seen as a repository of the intellectual and capital resources needed to support research and scholarship, as an archive of the achievements of the past (and, as such, an enduring place for those documents), as a place to study, to read, to contemplate and congregate. More recently, and particularly for libraries in university settings, libraries have become hubs of networked, information access for the entire research communities. The Stanford Libraries expect this trend to continue, and to become increasing the focal point of their activities in the electronic environment.

With the "old library" and its functions safely in tow, the "new" library will begin to take on new, strategic roles within the institution. Libraries will become both the creator of data and metadata resources and the broker of such resources to others. By metadata we mean not only the information about a library's holdings—or catalog in the traditional sense—we mean also information generated by the library or institution as a part of its creating its "electronic information space" for an institution. This includes not only the traditional catalog and its electronic records, but also a wide variety of other information about the collection and information space, such as electronic finding aids to the paper collections, digital objects (original or surrogates) within the collection, the Web space, and all of the other components of the digital collection.

For institutions such as Stanford that see the strategic possibilities in this new environment, the lines between libraries as "consumers" of information and publishers as "content providers" become markedly blurred. Libraries are increasingly content providers in their own right. By acknowledging and taking on some of the publishing roles, libraries such as Stanford become at the same time participants in the rapidly evolving publication environment, and are able as participants to shape that environment in ways in which a library as a consumer alone cannot.

Organizational Environment

The Stanford University Libraries/Academic Information Resources (SUL/AIR) is an organization that merges and melds academic computing (IT) and library functions into a single whole comprising 500 full-time employees equivalents.

The IT part of the organization includes programmers, system and application analysts, media specialists, and hard- and software technicians, while the library part includes curators, catalogers, reference staff, conservation professionals, paraprofessionals, and support staff. The advantages to such an organizational model are numerous: computing support for the library operations leads to a more robust processing model, library support of research and scholarship can rest on the best of both the Stanford collections and an electronic delivery environment tailored to those collections, and the traditional library prejudice towards service to users avoids the oft fatal splits between the "computing center" and "the user."

Stanford is equally blessed by a rich, quick and ubiquitous network: all of the libraries, virtually all of the offices, and more than 90% of the "beds" in University dormitories and housing[1] have a network connection. To this abundance of campus connectivity must also be added a keen and knowledgeable staff serving the administrative and networking technical infrastructure, which resides outside of SUL/AIR, but supports its activities. The network has long supported AFS with Kerberos authentication and encryption, and is currently transitioning to DCE environment. With this comes a very savvy staff, and likewise savvy users.

Stanford is likewise a participant in the current NFS Digital Library grant (with the Library as a co-participant), and the "Infobus" architecture developed out of that project is likely to have important and beneficial ramifications for the support and delivery of library services. The Silicon Valley connections are likewise important to us, and we find many opportunities for technical partnerships with mature companies such as Sun, IBM, and Hewlett-Packard, and the newer breed of startup.

---

[1] Approximately 95% of University undergraduates live on campus, and the number of graduate students likewise residing in networked University housing is more than 50%.

Technology Assumptions

In creating the "new" library within this operation environment, and given the continuing financial constraints on institutions of high learning in the US, SUL/AIR is basing its collections and technology efforts on the following principles:

1) Promote user self-sufficiency
   - Use technology where it can lead to more successful and flexible interactions with the user community
   - Create and leverage self-sufficiency thorough training and applications that work with users and/or are able to be tailored to individual needs
2) Use off-the-shelf applications
   - Integrate wherever possible
   - Avoid application development, particularly of "one time use" applications
3) Apply Open Systems principles
   - Avoid the proprietary, except where it has gained the force of a "standard" (i.e., word processing systems, spreadsheets, file formats like PDF, LaTex, and so on)
4) Deploy cross-platform, client-server products
   - Prefer application delivery environments that are not platform specific, and where thin/thick clients can be easily deployed, depending on need
5) Plan on short(er) equipment refreshment cycles
   - Replace most desktop equipment on a three year cycle; application development workstations on a two year cycle; servers on a twelve to eighteen month cycle
6) Bring power to the people -- outlets at every seat when building or renovating
7) Bring video/data into more classroom

Leveraging the Technology

From these principles, it is not hard to see that the organization recognized early on that the Web would offer opportunities to serve our users, move ahead on our strategic vision, and deploy new electronic resources to our users quickly and certainly more easily than the older technology.

SUL/AIR had recognized the potential of the Web as an as an ubiquitous, platform-independent information delivery path early on. It also offered the ability to integrate disparate, heterogeneous information cleanly and efficiently; it was a good environment for rapid application development, and has since become even better with the advent of Web development suites from companies like Oracle, Netscape, NeXT, and Microsoft. As a client/server application environment, it was also relatively easy to update and maintain information on centralized servers rather than localized software applications on the desktop.

For SUL/AIR, Web accessibility and connectivity has become the current hallmark for application development and establishing digital library partnerships. In choosing a new, local library technical processing and cataloging system to replace the aging mainframe system, SUL/AIR saw the Unicorn system from SIRSI offering both the tools for integration and Web connectivity we were seeking. By being early adopters of relatively mature technologies, SUL/AIR was able to further these partnerships in ways that have had a positive influence on the development of the Unicorn system.

Using Oracle Web development tools in conjunction with the SIRSI system, SUL/AIR has been able to create a "next generation" library catalog. This system permits the interlinking and connection of larger, deeper collections of digital as well as traditional library resources. The catalog and system supporting it is becoming the environment in which SUL/AIR is able to deliver its comprehensive "view" of the information space—a view which contains the wisdom and intellectual capabilities of its staff and combines it with the rich electronic resources of Stanford has acquired through commercial sources. The SUL/AIR Web is on its way to becoming a comprehensive "information space" in which the Stanford community can view, react to, analyze, criticize, study, and interact with both the resources themselves, and the selectors and systems people responsible for administering them and making them available.

The digital resources present at Stanford have been further enhanced through the development of subject-specific discovery and retrieval engines, coupled with a customized Web interface. For example, the Academic Text Service offers full-text retrieval and customized analysis of a wide variety of SGML-based data resources. Researchers, instructors, and students at Stanford can search, for example, through all of Shakespeare's works, four translations of the Bible, and Kant's *Kritik der reinen Vernunft*. Others offerings include large files like the English Poetry Database by Chadwyck-Healey, British Philosophy, all of Joyce's novels, and multiple translations of Ovid's Metamorphoses, along with the original Latin. Similar services are offered by the Academic Data Service—a repository of social science data sets, and by a newly formed data resource unit in the natural sciences and technology.

Other programs combine these strategic technology goals with new and innovative approaches to faculty support. The Academic Technology Support Service (ATSS) was originally conceived by the Commission on Technology, Teaching and Learning, and is managed by the Stanford Libraries. This program places Academic Technology Specialists directly in nine departments who are skilled both in the discipline of the department and in technology. The Resources Specialists work with faculty to promote and support their use of technology in teaching and research. This pilot began in January, 1996, and is currently funded for three years. Another component of the ATSS is the Academic Technology Lab. Housed within Meyer Library, consultants work directly with instructors on the use of technology their classes, and assist in the coordination of Library facilities and resources. A series workshops for faculty and lecturers are offered each quarter on technology and its use in the classroom.

Tour of the SUL/AIR Web Environment

Perhaps the easiest way to get an idea of the richness of resources and the environment SUL/AIR is trying to create is by touring the SUL/AIR Web site[2].

- Home Page
  www-sul.stanford.edu

[2] Although all of these sites were demonstrated during the presentation, it should be noted that some of these resources are restricted to the Stanford environment, and may not be available to outside visitors.

— Main SUL/AIR Page/Top Level of Environment

- Stanford Web Catalog:
  jenson.stanford.edu/prod/owa/su_in

  — Stanford's newly created Web-based catalog. Site is currently limited to the Stanford community only, but it is expected that the catalog will be made world available

- Disciplinary/Topic-based Web Sites:
  www-sul.stanford.edu/depts/hasrg/index.html

  — Index of subject-based Web information spaces in the humanities. Many of these links point to extensive Web resources. See, for example, *Slavic and Eastern European Collections* (www-sul.stanford.edu/depts/hasrg/slavic/1slavic.html) or *Languages and Linguistics* (www-sul.stanford.edu/depts/ssrg/linguist/ling.html)
  www-sul.stanford.edu/collect/ejourns.html

  — Electronic journals available at Stanford

  www-sul.stanford.edu/geninfo/subdir.html

  — A complete listing of subject specialists, reference resources, linked to collections pages

- Academic Text Service:
  www-sul.stanford.edu/depts/hasrg/ats/ats.html

- Academic Data Service:
  http://www-sul.stanford.edu/depts/ssrg/ads/ads1.html

- Information Resource Specialists
  www-leland.stanford.edu/dept/SUL/irs/

- HighWire Press, the Internet imprint of SUL/AIR:
  highwire.stanford.edu

Building on SUL/AIR's Human Resources

What I have described here highlights the positive aspects of technology and information delivery at Stanford, this is an arena where change and the force of the new and unestablished has a sizeable impact, both on staff and users. We have found that having a very active education and instruction program for faculty, staff, and students is critical. The level of technology and systems in place in faculty offices and in student dormitories is less high than is preferable, and the technical abilities of faculty and students is often challenged.

While many of the IT problems have workable solutions, these solutions can stress the infrastructure, the organization, and those of us who participate in it. This means that even though the core values of librarians remain, the culture and the arena in which they perform and interact changes, often in ways that make those values difficult to maintain. For

SUL/AIR, this means a continual organizing and reorganizing the fundamentals of the workplace—a flatter organizational structure, and more points of interaction and intersection between divisions that share common users and common technology problems. If we want to ensure that the transmission of scholarly information, our productivity, and potential continues at a speed that in pace with the possibilities, we need to leverage against our greatest resources: the information and computing professionals on which the organization rests. We are confident that SUL/AIR as an organization and the professionals within it will continue to meet these challenges and provide Stanford with the resources it needs to compete and demonstrate continuing excellence.

DISCUSSION

Rapporteur: Dr Chris Phillips

Lecture One

The fairly slow response during a demonstration of the Stanford Library Catalog provided the opportunity for a number of questions.

Professor Shepherd commented that although Dr Coleman had earlier claimed to use off-the-shelf components, the reality was somewhat different. Dr Coleman replied that although the goal is to use off-the-shelf components, in practice few provide exactly the required functionality and so some modification is necessary.

Mr Kay's question about whether the system was available to the general public was answered in the affirmative, and response time is normally good on campus.

The demonstration included a search for occurrences of the word "blood" in the works of Shakespeare. In a response to an enquiry by Professor Capriz, Dr Coleman stated that various versions of Shakespeare were available. Professor Capriz continued by saying that there is a legacy issue in that people like to see an image of a book page. Dr Coleman responded by saying that data is stored in various formats; in the form of transcriptions, but also as page images.

A number of observations were made on the problems of maintenance, that in ten years time we will have moved on from SGML, that the medium itself will have changed. Dr Coleman felt that the transition path should be smooth.

Professor Shepherd enquired as to the additional benefits provided by the Catalog, which could be viewed as just another search engine. Dr Coleman replied that such a service is not available elsewhere. Ten or twenty years ago it was possible to complete a dissertation on work that now takes a few minutes. In response to other sceptical comments Professor Randell commented that exactly the same things used to be said about the computer itself. Dr Coleman said that paper-writing had improved because hypotheses could be tested more rigorously. Dr Weber commented that the mere fact that we are now asking 'so what?' is an illustration of the rate of technological change.

Professor Kopetz enquired as to what extent sensitive information is made available. Dr Coleman replied that it was policy to provide information that people need; access will be restricted if this is deemed reasonable.

There was some debate as to the main potential users of the system. Professor Farber stated that at Pennsylvania the least use of the Web for research is made by Computer Science students. Professor Shepherd claimed that this was not the case at Lancaster. Professor Randell feared that students would become blinkered, that they would only use what is available on the Web. Dr Coleman argreed that the Web is not yet a satisfactory single source of information.

The discussion moved onto access, duplication and persistence issues. Given that clients are encouraged to use resources of other units and vice versa, Mr Kay argued that in the long term the kind of system being demonstrated should not be associated with a single institution.

It was suggested that creating such a large catalogue is beyond the resources of a single institution. It is essential to ensure there is no duplication of effort, that collaboration ensures that we do not all provide separate systems for analysing the works of Shakespeare. Dr Coleman said that it is important to check what others are doing, and to produce preservation models. Unfortunately the very nature of the Web means that information stored is transient,

and therefore difficult to quote. Dr Weber reported that there exists a company involved in archiving Web pages. Dr Coleman responded that this is not providing a complete archive of all sites. Professor Farber was concerned that if the Stanford system became a national resource it could be prone to natural disasters, and that some off-site duplication would be necessary.

Dr Coleman agreed, although digital archiving is likely to be expensive. Professor Randell wondered how the growth of digitized material was following that of information recorded on microfilm. Professor Farber remarked that considerable funds were available for the former as opposed to the latter. Professor Capriz enquired whether there exists a system analogous to the Library of Congress, whereby anything published is lodged in an archive. Dr Coleman replied in the negative, although it is a fact that, for example, in some US states all e-mail messages are copied and stored before delivery.