

Scholars' Bank Development: May 2003 through September 2006

By Carol Hixson, Head, Metadata and Digital Library Services

This report is intended to provide some background for Scholars' Bank that may be of use to the new Steering Committee.

I. Brief History of Scholars' Bank

In January 2003, the Institutional Repository Group (IRG) initiative was established (<http://libweb.uoregon.edu/%7Ejqj/lib-ir/msg00000.html>). Co-chaired by JQ Johnson, the Director of the Center for Educational Technologies, and Carol Hixson, Head of Metadata and Digital Library Services, the group was charged to investigate the feasibility of an institutional repository for the University of Oregon and to move forward with an implementation if that seemed feasible and desirable. Other group members included Heather Briston, the University Archivist and Historian, Andrew Nicholson, the Geography/Map Librarian, and Barbara Jenkins, the Head of Reference and Research Services. Inspired by the work of SPARC and other similar groups, the IRG hoped to create an archive for faculty output, particularly peer-reviewed content. The expectation was that faculty would understand the advantages of making their work openly available and that this would ultimately lead to a change in the scholarly communication model, a model that librarians had come to consider irretrievably broken. As the difficulty of acquiring such peer-reviewed, formally-published content became clear, the IRG's vision for the archive changed to be far more inclusive. Scholars' Bank now includes scholarly output from any member (faculty, students, administrators) of the campus community, as well as materials supportive of the university's mission, such as newsletters, finding aids, and selected university records.

The IRG first met on February 12, 2003 and then met once or twice a month for the next year and a half. An archived discussion list and web site for maintaining a record of activities and progress were set up in January 2003 (<http://libweb.uoregon.edu/catdept/irg/>). Minutes and discussions can be reviewed there to trace how the thinking and strategies of the group evolved. Although the group has seldom met face-to-face in the past two years, it has continued to have discussion about strategies and policies, as needed, over the list (<http://libweb.uoregon.edu/%7Ejqj/lib-ir/>).

One of the first decisions the IRG made concerned the software to be used. There are several different institutional repository software packages; the most commonly used software packages at the time we started were DSpace and Eprints. We decided to use DSpace, an open-source system developed by MIT Libraries and Hewlett Packard. JQ Johnson did the initial setup, refitting an old PC that had been used previously as a server for Blackboard. The site came online in test in May 2003. JQ, Andrew Nicholson and Carol Hixson taught themselves how to use the system and began to load in test files.

Another early decision was about the wording of the license. The group called upon Heather Briston with her legal and archival training to come up with the initial wording. That wording was slightly modified once after a graduate student backed out of a submission because of uncertainty about what the license meant. In addition to modifying the wording slightly, we

developed web pages (<http://libweb.uoregon.edu/catdept/irg/SBlicense.html>) to explain the permission form. We gave people a chance to review the wording of the license file before getting into a submission, explained what the wording meant, and invited them to contact us if they had questions. The wording has served us well and has also been used as a model by other institutions setting up digital archives such as Scholars' Bank.

The first campus community in Scholars' Bank was the Economics Department's collection of Working Papers. Bill Harbaugh, the coordinator of the collection, had for some time been maintaining a collection of the Working Papers on a departmental web site, as well as contributing them to a disciplinary archive, RePEC. Although he was not originally convinced of the value of an institutional repository, he agreed to allow the library to harvest the papers from the two sites and contribute them to Scholars' Bank. This initial community and collection established two patterns that have been fundamental to the growth and success of Scholars' Bank. First, it set the pattern of the library (using MDLS staff) doing whatever it takes to get materials into Scholars' Bank. Second, it set the pattern of using one group's content as a model and marketing tool to acquire other content. Harbaugh later became one of our strongest faculty allies and has been responsible for bringing more relevant content into Scholars' Bank.

Another early decision was what to call the archive. The IRG put out a call to library faculty for suggestions. Paul Frantz suggested the name that was modified slightly by Deb Carver. The archive was officially named Scholars' Bank in January 2004.

Over time, the IRG's membership and mode of operating changed. Corey Harper joined the group in November 2003 after Andrew Nicholson left the UO; Corey eventually took over responsibility for maintaining and upgrading the software, as well as running reports, setting up statistics before the software had developed such tools, and even contributing code fixes that were later incorporated into subsequent DSpace releases.

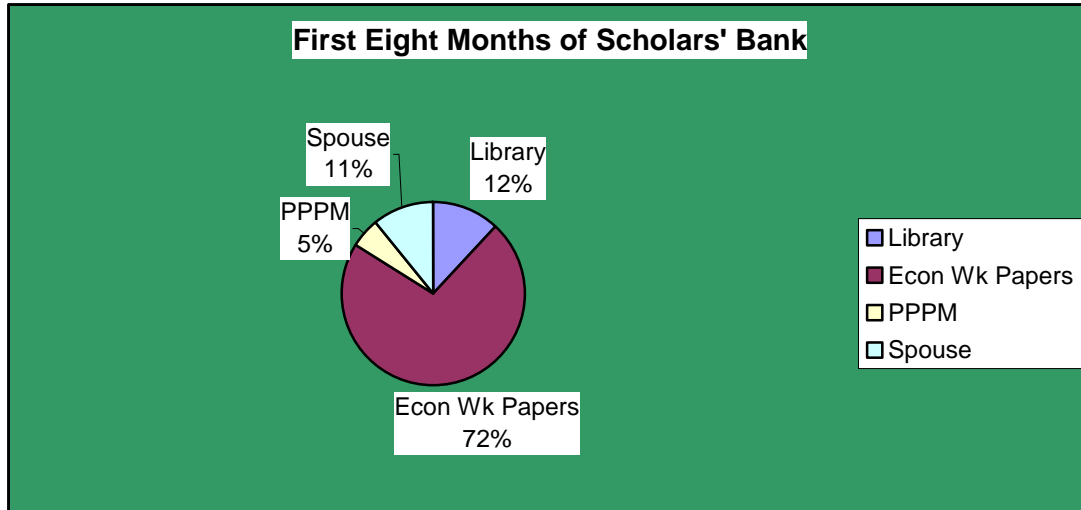
In January 2004, the Institutional Repository Group transitioned to serve as an advisory committee. Scholars' Bank administration was absorbed by Metadata and Digital Library Services (MDLS). Hixson became the unofficial coordinator: setting up communities and collections, training staff in submissions, developing and enhancing the contextual web pages for users of the archive, meeting with faculty, students, and other groups, and doing presentations to internal and external groups about the archive.

Eliz Breakstone joined the Steering Committee in February 2005 because of her strong interest and exemplary work making connections with various campus communities. Mary Greci joined the group in August 2005, serving as Hixson's administrative backup. The IRG ceased to meet on a regular basis and the archived list was also used less and less frequently, as maintaining Scholars' Bank was incorporated naturally as one of the core services provided by MDLS, in close cooperation with subject specialists and Library Administration.

II. Growth and Analysis of Scholars' Bank

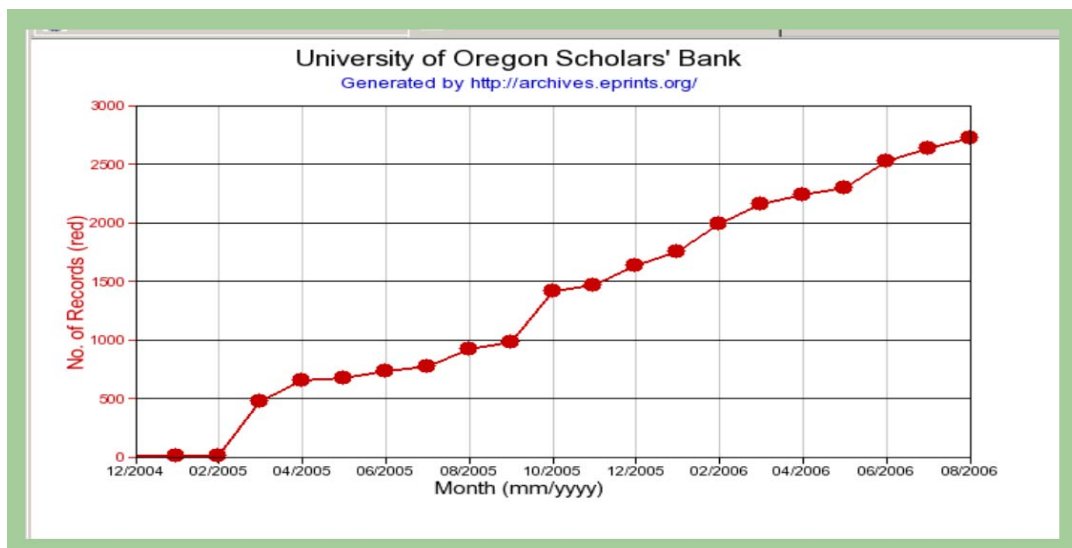
The growth of the archive was initially very slow. In the first 8 months that Scholars' Bank existed (May-December 2003), we succeeded in acquiring only 93 items. As the following chart

shows, 12% of those were from the library, in the form of library staff articles, presentations, or finding aids; 11% were items from a faculty member who was the spouse of one of the IR group members; 5% were faculty and student papers from the School of Planning, Public Policy, and Management. The majority (72%) were Economics Department Working Papers.

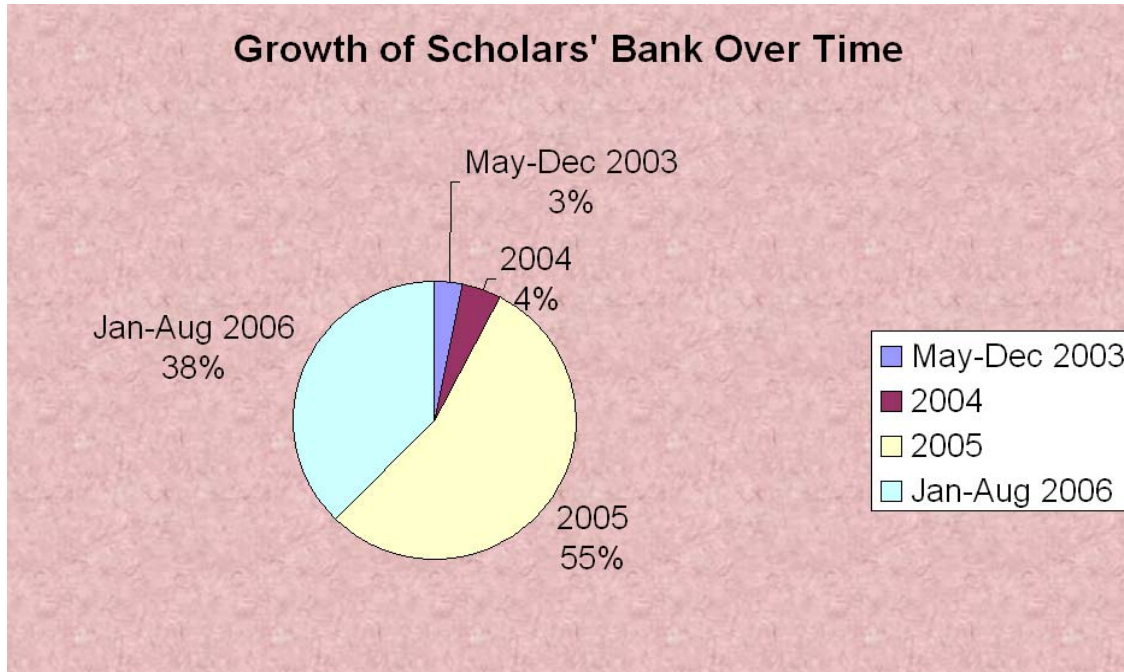


After 8 months of extensive work talking to individual faculty members trying to interest them in the archive and getting at best a lukewarm response and almost no follow through from those faculty who did express an initial interest, we realized that something had to change if we wanted to create a viable and useful repository. In general, we became more aggressive about marketing, we identified a broader type of content as being appropriate for the archive, we followed through quickly (within 24 to 48 hours) on every lead, and we developed a submission model whereby the majority of items in the archive have been submitted (and often digitized) by library staff.

The Registry of Open Access Repositories (ROAR) charts the growth of Scholars' Bank since December 2004 when we registered with them through August 2006, as follows:



The following chart shows the growth of Scholars' Bank from the very beginning through August 21, 2006.

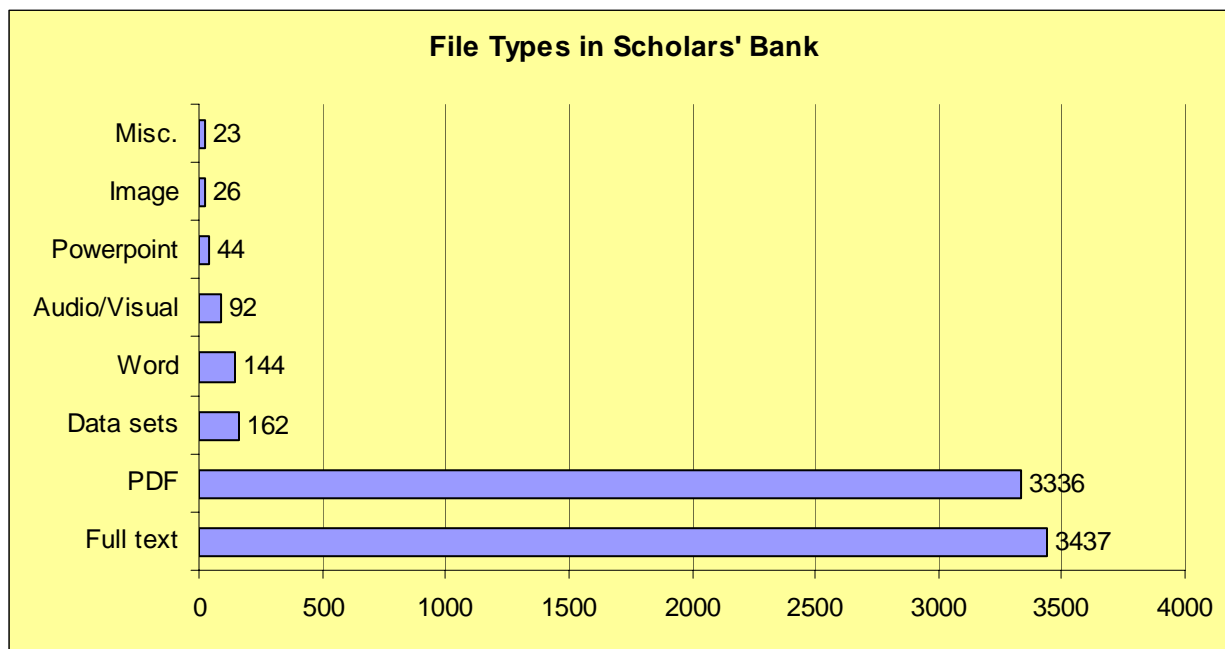


In the first eight months of the archive's existence (May to December 2003), we acquired 93 items. In all of 2004, we acquired only 124 more items. After concerted effort and a change of strategy, we achieved considerable success in 2005, with 1510 new items being added. That success has continued, with the first 8 months of 2006 seeing another 1049 items added to Scholars' Bank. What is perhaps even more significant is that a growing number of individuals and departments now approach us to request the establishment of communities and collections in Scholars' Bank. The archive has a high degree of name recognition on campus and a growing number of scholars are seeing the archive as the perfect place to deposit their intellectual legacy.

From May 2003 through August 2006, only 8.5% of the 3827 individually submitted (as opposed to system-generated) files in Scholars' Bank were submitted by the authors or content owners themselves. 7.5% of the total files have been submitted by Document Center staff and 84% have been submitted by MDLS staff on behalf of authors or content owners. The files that were submitted by library staff were often digitized from hard copy or harvested from web sites and stabilized as PDFs. A key to our success has been our willingness to handle the submissions on behalf of the content owners.

The file types currently in Scholars' Bank are primarily text-based, with a smattering of other types of content. The following chart shows the file types from the beginning in May 2003 through August 21, 2006. The full-text files shown on the chart are generated by the Media Filter program that is set up to run nightly to provide full-text searching capability. All other categories consist of files that have been individually submitted either by library staff or the content owners.

The miscellaneous category consists of html, css, xls, rtf, and indd files. The data sets are all SPSS data sets.



III. Definition

One of the primary factors enabling Scholars' Bank to grow and flourish has been the gradual redefinition of what it is. Our initial vision was fairly typical: a repository for peer-reviewed work created and self-submitted by faculty of the University of Oregon. The current definition of Scholars' Bank draws heavily on Clifford Lynch's (Coalition for Networked Information) definition of an institutional repository, with some local modifications. Scholars' Bank is:

- A set of digital collections that capture and preserve the intellectual output of the University of Oregon community. This community is defined in its broadest sense to include faculty, students, staff, and administrators of the University.
- A set of services offered to the members of the University community for the management and dissemination of digital materials created by the University and its community members.
 - "It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation where appropriate, as well as organization and access or distribution."—Clifford Lynch
 - Typical services includes harvesting or digitizing content, submitting it on behalf of the authors or content owners, and providing appropriate metadata to aid in its discovery and use.

- Perpetual. Scholars' Bank is an archive that promises long-term accessibility and stability.
- Openly accessible. The materials in Scholars' Bank are primarily open access.
 - With one or two exceptions, materials in Scholars' Bank are free of third-party copyright restrictions. Only a few files have restricted access and in those cases the metadata is openly available.
 - No content is available for purchase, unlike the model followed by some institutions, such as MIT
- Not just for formal, peer-reviewed publications, but also for working papers, presentations, class projects, newsletters, theses, etc. Scholars' Bank is complementary to the traditional publication process.

In short, Scholars' Bank contains materials that are themselves scholarly in nature or that support the University's scholarly mission – this has been very broadly defined. The materials in it must have some connection with the University of Oregon, although that connection is sometimes tenuous. This will be explained more in the policies and marketing sections of this document.

IV. Issues, challenges, and strategies

We have faced a number of challenges in building and promoting Scholars' Bank. These have been technical, practical, and cultural in nature. The least significant challenges have so far been the technical ones, for several reasons. One, the software is fairly robust and has a wide community of users who freely share their experiences and expertise; two, we have adapted our practices to fit the limitations of the software; three, we have had talented people in-house to help us overcome them, first JQ, then Corey, now several people in Systems as backup. Some of the technical issues are:

- Submission template
 - Although the current version of the software does provide the ability to set up templates with some metadata already supplied, the basic submission form cannot be modified for individual collections without modifying the code
- Metadata
 - Display labels are hardcoded with library jargon
 - There is lack of support for controlled terms and no easy way to do database cleanup; corrections are done item-by-item and require a high level of authorization
- Version control and revision
 - Inability of individual users to replace or update their files
- Inadequate report capabilities
 - Reports are available only by querying the underlying database
 - With the exception of some statistical data now available (which we have had to modify to get it to display in a more useful fashion and which we supplement through the use of Webalizer statistics for the server)
- Insufficient granularity of permissions
 - Too many functions require the highest level authorization, such as replacing files, modifying metadata, etc.
 - Some functions that should be available to delegated collection administrators are not
- User interface
 - Search interface is rudimentary
- Preservation

- Challenge faced for all of our digital content, not just that in Scholars' Bank

Some of the practical challenges are:

- Time
 - Lack of faculty time to think about or do anything new that doesn't lead to an immediate benefit
 - Lack of our time:
 - To pursue (multiple times) all the probable community and collection liaisons
 - To develop and keep up to date all the contextual pages
 - To develop searching guides
 - To pursue code development and tailor the archive to meet our users' needs more closely
 - To pursue a rigorous preservation program
- Money
 - Mostly translating into our staff support for mediating most of the submissions
- Identification of content and communities
 - Difficulty of finding out who is producing what and identifying the right people to talk to about it
- Acquisition of content
 - Simply getting it in – 93.3% of ours has been mediated by the library - of the 7264 bitstreams in the archive (bitstreams are the individual files and an item may consist of one or multiple bitstreams), only 487 (6.7%) have been submitted by the author or editor of the file

The area where we have so far devoted the greatest effort has been in trying to overcome the cultural challenges. The main ones have been:

- Intellectual property and copyright concerns
 - Will putting my material in the archive damage my relationships with my publishers?
 - How can I protect my content from being stolen or used inappropriately?
- Authenticity of submissions
 - From the Graduate School, how do we know that the students are submitting the same version of their dissertation that was approved by their review committee?
- Concern over quality
 - Campus administration was initially concerned over having an institutional stamp put on content that was not centrally approved
- Concern over control
 - Faculty were concerned that Campus administration would try to control what they put in the archive
- Terminology
 - Nobody understood repository, so we refer to it as an archive, which they do understand
 - Institutional conjured up negative connotations from Campus Administration and faculty so we avoid talking about that, as well
- Technophobia
 - Some people have limited experience with computer systems and are put off by anything that is new technologically

V. Policies

There are many policies in place for Scholars' Bank, although not all of them have been well articulated. There is a page where some policies have been collected and made publicly available at: http://libweb.uoregon.edu/catdept/irg/SB_Policies.html . Policies for Scholars' Bank roughly fall into the following categories, which I will elaborate on separately:

- Division of responsibilities and clarification of roles
- Structure of the archive
- Control of content
- Submission and withdrawal
- Metadata standards
- Institutional commitment
- Copyright, permission, and access
- Going it alone or multi-institutional

Division of responsibilities and clarification of roles

The division of responsibilities for Scholars' Bank (through August 2006) has been as follows:

- The Head of Metadata and Digital Library Services has functioned as the Coordinator of Scholars' Bank. This role has included: general oversight; assignment of project-based tasks for digitization and submission; training library staff in submission and other work; reviewing new submissions; setting up new communities and collections; meeting with faculty, students, and other campus groups to explain the archive; ongoing liaison work with community representatives; tracking of projects; writing occasional reports; developing new contextual pages; making public presentations to on and off-campus groups; bringing up policy issues to the Steering Committee and Library Administration.
- Subject specialists have identified communities or individuals whose materials would be appropriate for Scholars' Bank. This has been done in consultation with the Coordinator and may include meeting alone or in the company of the Coordinator with a particular group or individual.
- The Steering Committee (which up until August 2006 consisted of JQ Johnson, Barbara Jenkins, Corey Harper, Eliz Breakstone, Heather Briston, Carol Hixson, Mary Greci) has discussed questions and policies and shared relevant information, sometimes in person but often over the archived discussion list.
- Library Administration has made larger policy decisions for Scholars' Bank when they have been brought to its attention. These include decisions about new directions for the repository or communities desiring to join the archive that do not clearly fit into already established community or collection parameters. Some examples of decisions made by Library Administration have been: the name of the archive, not to accept fee-based content, to separate out government publications into a separate instance of the software, and others.
- Documents Center and MDLS staff have harvested, digitized, submitted, and reviewed content on behalf of various communities.

- Technical work (testing new versions of the software, backing up files and the database, setting up certificates, running reports against the database, modifying the code for local implementation (done in very few cases to date), solving problems when there are internal system errors, etc.) has been handled largely by Corey Harper, with recent backup from Library Systems staff.

Structure of the archive

The structure of the archive has been determined largely by the software being used, although the specific implementation at UO has been founded on implicit and explicit policies. DSpace supports four basic levels of hierarchy, which are defined in the software documentation available at: http://dspace.org/technology/system-docs/functional.html#data_model :

- Top level: Community
 - DSpace requires the establishment of communities, that its documentation says “typically correspond to a laboratory, research center or department.”
 - The policy statement for establishing communities in Scholars’ Bank found at http://libweb.uoregon.edu/catdept/irg/SB_New_Community.html indicates that: “To be a recognized community, a group must be affiliated with the University of Oregon. Any administrative or academic department, research institute, program, or publishing entity of the University of Oregon is recognized as a community of the University and may establish a community within Scholars’ Bank.”
 - The General Guidelines for Scholars’ Bank at: http://libweb.uoregon.edu/catdept/irg/Community_Guidelines.html provide further guidance.
 - A further policy decision is that all communities, sub-communities, and collections must be given an entirely unique name. This is to avoid confusion that would arise from system-generated alphabetical listings within the same hierarchical level.
- Second level: Sub-community
 - As of version 1.2 of DSpace, communities can have sub-communities. This level is optional. All other levels are required by the software. There doesn’t seem to be any limit on the number of sub-communities that can exist under a community. It is also possible to sub-divide sub-communities into smaller sub-communities, thus increasing the levels of hierarchy that can exist in DSpace.
 - Scholars’ Bank has no explicit policy statement on the establishment of sub-communities. In practice, we have followed the same basic policy for the establishment of a community and have established sub-communities that seemed to make sense and would allow a community to define its materials in useful ways. For examples of sub-communities in Scholars’ Bank, see:
 - <https://scholarsbank.uoregon.edu/dspace/handle/1794/1898>
 - <https://scholarsbank.uoregon.edu/dspace/handle/1794/2671>
 - A further policy decision is that all communities, sub-communities, and collections must be given an entirely unique name. This is to avoid confusion that would arise from system-generated alphabetical listings within the same hierarchical level.

- Third level: Collection
 - Within DSpace, a collection can only exist under a community or sub-community. DSpace defines a collection as: “groupings of related content. A collection may appear in more than one community.”
 - In Scholars’ Bank, we have accepted any type of a collection that a community wants, within the parameters of what is considered appropriate for the archive. The explicit policy statements regarding collections can be found at: http://libweb.uoregon.edu/catdept/irg/SB_New_Community.html and note that:
 - Individual faculty may set up personal collections within the already-established [Faculty Research](#) community or they may encourage their department or other group to establish a new community.
 - Individual students or graduate students may submit only to established collections, such as the Graduate School's [Theses and Dissertations](#) collection. Students who are not members of already-established collections or communities must enlist the sponsorship of a recognized community before being able to submit their work to Scholars' Bank. [Click here](#) for a list of already-established student collections within Scholars' Bank.
 - To establish a collection, an individual or group of individuals must have the support of a recognized community of the University of Oregon.
 - Each collection within a community may have distinct membership, guidelines, or standards. Access to communities or collections may be open or restricted, according to the community's needs.
 - A further policy decision is that all communities, sub-communities, and collections must be given an entirely unique name. This is to avoid confusion that would arise from system-generated alphabetical listings within the same hierarchical level.
- Fourth level: Item
 - Within DSpace, an item is like a bibliographic entity. Like a bibliographic entity, an item can consist of a single or multiple parts – single or multiple files. The software documentation defines items as “the basic archival elements of the archive. Each item is owned by one collection. Additionally, an item may appear in additional collections; however every item has one and only one owning collection. Items are further subdivided into named *bundles* of *bitstreams*. Bitstreams are, as the name suggests, streams of bits, usually ordinary computer files. Bitstreams that are somehow closely related, for example HTML files and images that compose a single HTML document, are organized into bundles.”
 - In Scholars’ Bank, we have made a policy decision to accept any type of file. Out of practical consideration, files of greater than 1 Gigabyte must be pre-approved. That policy is explicitly stated at: http://libweb.uoregon.edu/catdept/irg/Community_Guidelines.html
 - Another policy decision is to allow items to be mapped to multiple collections, whenever it seems advisable. For instance, an individual faculty member’s works may be mapped to his personal collection, as well as to a collection established under his department’s or program’s community. For examples of items mapped to multiple collections, see the following:

- <https://scholarsbank.uoregon.edu/dspace/handle/1794/963>
 - <https://scholarsbank.uoregon.edu/dspace/handle/1794/2998>
 - <https://scholarsbank.uoregon.edu/dspace/handle/1794/3097>
- There are several other policy decisions based on practicalities that are not explicitly stated. These include:
 - Word documents are now converted to PDFs because Word files are not considered a good choice for long-term preservation
 - Streaming video content is stored on a separate server and reference files are loaded into Scholars' Bank. This is because of size and access considerations
 - Extremely large files are often broken up into several smaller files to ensure faster user access to individual files.

Control of content

Who controls the content of the archive?

- The community liaison or liaisons determine the appropriateness of materials for the archive. Working with the Scholars' Bank Coordinator, a community liaison defines the parameters of a specific collection. As it is explained in the [FAQ](http://libweb.uoregon.edu/catdept/irg/SB_FAQ.html) at: http://libweb.uoregon.edu/catdept/irg/SB_FAQ.html, "it is the responsibility of a community to set standards for the material that authors may contribute. In setting up the infrastructure for a community, the Libraries will consult with community representatives about specific policies and access restrictions. It is up to a community's liaison to communicate standards to individual authors." This aspect is further articulated in the General Guidelines where it says:
 - Some collections will require approval of submissions by a Community Liaison. Each designated representative reviews all items submitted to the Scholars' Bank Community that he or she oversees. If review of each item is not a realistic goal due to the size of the collection or due to time constraints, the representative should carefully review the first few items submitted by each member of the Community. The representative should continue periodically to review selected items to ensure that community policies are being followed and that appropriate metadata is being provided.
- Ultimately, it is the UO Libraries that control the content. There are several relevant policy statements regarding this aspect of Scholars' Bank that appear on the [General Guidelines](http://libweb.uoregon.edu/catdept/irg/Community_Guidelines.html) page at: http://libweb.uoregon.edu/catdept/irg/Community_Guidelines.html including:
 - Materials submitted should be the intellectual property of the author, the department, or the University and not under any copyright restrictions imposed by a third party.
 - When a new Scholars' Bank Community is established, a Metadata Librarian from MDLS will review the first few items submitted. Based on that review, the Metadata Librarian may recommend changes to the Scholars' Bank Liaison.

- The Steering Committee reserves the right to change these guidelines, to make formal policies, or to terminate the project. Should Scholars' Bank ever be terminated, materials will be returned to their authors or their communities.
- MDLS and the Steering Committee reserve the right to remove items from the Scholars' Bank if notified by a third party of potential copyright infringement. In such cases, authors and communities will be notified.

Submission, revision, and withdrawal

Many policies for submission, revision, and withdrawal of content have been explicitly stated, whereas others are only understood and applied by MDLS and the Scholars' Bank Coordinator.

- The general policy of Scholars' Bank is that any authorized user may submit to the collections where he or she has been authorized. Authors or content contributors may also designate a third-party (usually a library staff member) to submit materials on their behalf.
- Many of the policies on submission are articulated on the [Submitting to Scholars' Bank](http://libweb.uoregon.edu/catdept/irg/SB_Submit.html) page at: http://libweb.uoregon.edu/catdept/irg/SB_Submit.html and include:
 - Anyone who is a registered member of an established community in Scholars' Bank may submit appropriate materials to the community's designated collections. To determine which materials are appropriate, the user should consult with the Community Liaison for the designated collection.
 - Any user who needs assistance submitting materials to the archive, may consult with the Scholars' Bank Coordinator. Faculty may request that submissions be handled by library staff on their behalf. Graduate and undergraduate students are generally expected to submit their own materials to the appropriate collections, although assistance will be given by library staff, if needed.
 - Some collections require that new submissions be reviewed by a designated member of the sponsoring community. If that is the case, individual users will not see their submissions appear in Scholars' Bank until they have been reviewed and approved by the collection's reviewer.
- Another policy not explicitly stated but closely followed is that any collection that has open submission is always reviewed by a member of the library staff before it becomes publicly available. This is necessary to avoid inappropriate or bogus submissions, as well as to standardize or correct errors in author-supplied metadata.
 - In cases of doubt, the Scholars' Bank Coordinator checks with the community liaison to determine if a submission is legitimate.
- A policy decision is that Scholars' Bank is a permanent archive. The policies for withdrawal of content are articulated at the [Withdrawals from Scholars' Bank](http://libweb.uoregon.edu/catdept/irg/SB-Withdrawal.html) page at: <http://libweb.uoregon.edu/catdept/irg/SB-Withdrawal.html> and include the following statements:
 - Scholars' Bank has been established as a permanent archive. In exceptional cases, authors, editors, or community liaisons may request that works be removed from the archive. Reasons for withdrawal of content could be factual inaccuracy, copyright infringement, or plagiarism. Ordinarily, content will not be removed simply because the author has produced a revised version of a work.

- When the Scholars' Bank Coordinator removes content, the description (metadata) for the work will remain as a placeholder in the archive, with its affiliated handle, or URL. A brief statement explaining the reason for withdrawal of the content will be added to the descriptive metadata.
- No materials will be removed without the authors' knowledge. The Scholars' Bank Steering Committee reserves the right to remove items from the Scholars' Bank if notified by a third party of potential copyright infringement. In such cases, authors and communities will be notified.
- There is a separate page where the policies regarding revision of content have been articulated at http://libweb.uoregon.edu/catdept/irg/SB_Revision.html : Those policies are:
 - Authors or editors of works should ensure that the works they submit are factually accurate and that they have the right to make them publicly available. If the status of a work is uncertain, it should not be submitted to the archive.
 - In cases where authors or editors have substantially reworked the content and wish to make a newer version available, they are encouraged to submit the new or revised version as a new item. The Scholars' Bank Coordinator will provide assistance in linking the earlier version with the later version and will, if the author requests it, provide information indicating which version is preferred. However, it is the policy of Scholars' Bank also to retain the original version of works submitted to the archive. In exceptional cases, an earlier version may be completely removed from the archive, at the author's or editor's request.
 - If an author discovers after submitting a work to the archive that the file(s) contained errors, the Scholars' Bank Coordinator will work with the author or editor to replace the earlier file with the corrected version.

Metadata standards

DSpace supports Dublin Core, simple and qualified. However, there are limitations to the display and searching of some metadata elements that render the use of some qualified Dublin Core elements moot. The general policy decisions regarding metadata for items in Scholars' Bank have not generally been documented.

- Metadata application have followed these principles:
 - Do whatever is necessary to provide for a logical and clear display. This will vary depending on the nature of the collection and what the community hopes to accomplish with the collection. This aspect has been covered extensively in Hixson's presentations in 2005 and 2006 to NASIG that are available in Scholars' Bank at: <https://scholarsbank.uoregon.edu/dspace/handle/1794/2484> or <https://scholarsbank.uoregon.edu/dspace/handle/1794/843>
 - Provision of metadata is not the same as cataloging. Metadata within Scholars' Bank represents the minimum necessary for retrieval and accurate identification of an item. When the files that comprise an item can be searched as full text, we have not generally supplied keywords in the metadata.
 - Metadata has not been applied consistently across collections. Similar types of information have been supplied in different fields across collections.

- There is not a one-to-one correspondence between MARC and Dublin Core. Therefore, multiple pieces of information that would appear in separate MARC fields are often grouped together into a single Dublin Core field.
 - Individual communities and collections may apply distinct standards for the terms and other descriptive metadata. This reflects the basic structure and philosophy of the software which expects author self-submission.
 - There are no controlled vocabularies within the current implementation of Scholars' Bank. The current version in production provides a simple way of checking for consistency in names of authors and the Scholars' Bank Coordinator has made an attempt to keep the names file reasonably clean and consistent.
 - The standards for names and subjects in Scholars' Bank are articulated at: http://libweb.uoregon.edu/catdept/meta/subj_access_background_Scholars_Bank.pdf
- Additionally, an unwritten policy is that metadata will be revised by the Scholars' Bank Coordinator (or designated staff) to reflect changes to the content that may have occurred. One example of such revision of metadata can be found at: <https://scholarsbank.uoregon.edu/dspace/handle/1794/633>

Institutional commitment

Scholars' Bank was conceived as a digital repository to which the University of Oregon Libraries has made a long-term commitment. Policy statements about that commitment can be found at: http://libweb.uoregon.edu/catdept/irg/SB_Preserve.html , http://libweb.uoregon.edu/catdept/irg/SB_Overview.html , and http://libweb.uoregon.edu/catdept/irg/SB_FAQ.html . Some of the specific commitments include:

- Nightly production of ASCII text backup, when possible, for those files that are text-based
- Nightly incremental backup of files and metadata to the Libraries' mass storage unit
- Bi-weekly backup from the mass storage unit to tape
- Bi-monthly export of the metadata and directory structure as a tar file to another system storage of magnetic tapes outside of the main library building
- Conversion of some file formats upon deposit to more secure file formats
- Periodic checking of file integrity to avoid data corruption
- Monitoring the technological environment to prepare for file migration as file formats become obsolete or the software needed to run them becomes unavailable

We have always attempted to be cautious in what we state explicitly, which is why the policy statement also says:

- Because the technological environment is changing rapidly, it may not be possible to preserve the usability of every file format that is deposited in Scholars' Bank. The University of Oregon Libraries will make a good-faith effort to maintain the viability of the materials deposited in Scholars' Bank or return them to their authors if unable to do so.

Long-term preservation of digital content is a challenge and is a serious issue with Scholars' Bank. We promote Scholars' Bank as an archive where people can turn over responsibility for the long-term preservation of their content. There are several serious flaws in our current procedures, however, that should be addressed:

- We need to have off-site storage of our tape backups. Tapes are currently stored in the Science Library. We should pursue the possibility of swapping backup tapes with one of our sister libraries, such as Oregon State University Libraries. The Head of UO Librray Systems was going to investigate this but there have been no reports back on those discussions, if they ever happened. (This is an issue not just for Scholars' Bank but for our online catalog and any other machine-readable files or databases.)
- We need to undertake regular checks of the integrity of the files. Although DSpace software collects a checksum at the point of ingest, we do not go back and check for loss of integrity in any of the files. A process should be put in place that would check for file corruption or deterioration. There have been discussions about this with Systems staff at different times but, to the best of my knowlodge, there has not been time for them to pursue this aggressively.
- We need to be monitoring the file types accepted and periodically checking to see if files need to be refreshed or migrated to newer versions of the same file type or to an entirely different file type.

Responsible, long-term digital preservation entails much more than simple backups. We do regular backups, but even those procedures need to be improved upon to meet even minimal standards. The Digital Content Coordinators has been working for more than two years to develop policies and procedures that would ensure the long-term viability of our digital content. However, the efforts of that group are too diffuse and unfocused and little has been accomplished.

Copyright, permission, and access

Our policies regarding copyright, permission, and access are more often understood than explicitly stated. We promote Scholars' Bank as an open-access archive, meaning that the materials in it are freely available to anyone with Web access and the necessary software for reading the files. Some specific policy statements can be found at:

http://libweb.uoregon.edu/catdept/irg/SB_Copyright.html,
<http://libweb.uoregon.edu/catdept/irg/SBlicense.html>,
and http://libweb.uoregon.edu/catdept/irg/SB_FAQ.html

Some of the policies regarding copyright, permission, and access are:

- Authors who submit their work to Scholars' Bank retain the copyright to their work, unless they explicitly give it away to a third party.
- The University of Oregon Libraries do not seek nor claim copyright on any of the works submitted to Scholars' Bank. The Libraries ask all authors to agree to a non-exclusive distribution license (<http://libweb.uoregon.edu/catdept/irg/SBlicense.html>), which means

that authors may make other copies of their work available on other web sites or through other means without obtaining permission from the Libraries. They may also formally publish their work, in the same form or in a revised form, without obtaining permission from the Libraries.

- All of the metadata and almost all of the files are freely available to the world at large.
- Authors may set conditions on the re-use of their materials in Scholars' Bank by affixing permission and copyright statements to the items. Two examples of author restrictions can be found on the items at: <https://scholarsbank.uoregon.edu/dspace/handle/1794/2572> and <https://scholarsbank.uoregon.edu/dspace/handle/1794/3075>
- Access to materials is via a Web interface and materials are searchable through Scholars' Bank or standard Web search engines, such as Google
- Indexing is mapped to international standards so that the materials in the archive can be routinely found by international search engines and harvesters, such as [OAIster](#)

Permission in Scholars' Bank is complex because the software expects and was designed to handle author self-submission. The mechanism it employs is to attach the email identifier of the person submitting the files to every item. Since we know that only 8.5% of the files in the archive have been submitted by the content owners, it was necessary to devise a mechanism for tracking the author's permission that would be separate from the submitted files. Ultimately, Hixson devised three online permission forms that send emails to a specified address. That email contains the basic text of the permission, as well as the email and other contact information of the author or content owner. That email is converted to a simple txt file, is added as a new bitstream to the item, and is then coded as a license file within the archive (it is not viewable in the public interface but can be viewed in the administrative interface.) The policies inherent in this practice are:

- When authors request the library to submit files on their behalf, we direct them to the appropriate permission form and ask them to fill it out and submit it
- Permissions are stored in the archive as part of the bitstreams affiliated with an item

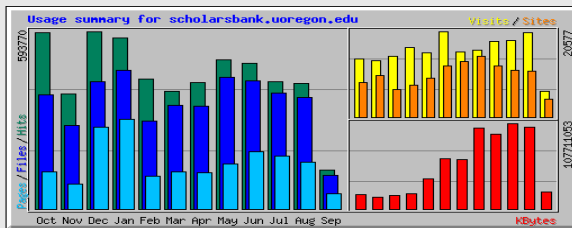
More information on permission and access issues for Scholars' Bank can be found in Hixson's presentation to the to the ALCTS Networked Resources and Metadata Interest Group Meeting at the American Library Association Annual Conference in June 2006 available at: <https://scholarsbank.uoregon.edu/dspace/handle/1794/2900> .

Going it alone or multi-institutional

Early on in the work of the Institutional Repository Group we decided to establish an archive for the University of Oregon community. However, we also made a commitment to have an archive that would comply with the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). In this way, we have been able to have an archive that met our specific needs while at the same time sharing information and access to the materials in our archive. As a matter of policy, Scholars' Bank:

- Complies with OAI-PMH
- Is registered with numerous registries such as ROAR, OAIster, and Google Scholar.
These registries harvest the metadata from Scholars' Bank on a periodic basis and help

get the word out about Scholars' Bank. Thanks to such registries, Scholars' Bank is very heavily used, as the following charts from Webalizer show:



Summary by Month										
Month	Daily Avg				Monthly Totals					
	Hits	Files	Pages	Visits	Sites	KBytes	Visits	Pages	Files	Hits
Sep 2006	11764	10336	4682	557	4401	21706227	6134	51506	113699	129410
Aug 2006	13523	12029	5011	652	11002	102233437	20239	155354	372914	419225
Jul 2006	13758	12537	5705	590	11333	107711053	18290	176857	388662	426513
Jun 2006	16252	14244	6393	602	12264	93864950	18077	191798	427329	487574
May 2006	16124	14237	4881	519	14513	101645067	16118	151322	441374	499872
Apr 2006	14078	11429	4018	519	13346	62427251	15582	120568	342891	422361
Mar 2006	12654	11165	4046	663	12308	63721063	20577	125433	346120	392278
Feb 2006	15528	10458	3915	548	9412	38296863	15360	109641	292832	434806
Jan 2006	18420	14985	9706	537	7609	19281054	16672	300886	464542	571033
Dec 2005	19153	13749	8789	467	6557	17081847	14494	272479	426246	593770
Nov 2005	12849	9283	2747	450	10008	15280500	13523	82435	278496	385492
Oct 2005	18966	12334	3966	453	8302	17822344	14050	122949	382364	587950
Totals						661071656	189116	1861228	4277469	5350284

It would not be difficult to establish a registry that would harvest metadata from institutional repositories of any target group (such as GWLA libraries), as long as they complied with OAI-PMH. The difficulty of such an effort comes when the member institutions try to reach agreement on the type of content to be archived and the descriptive metadata standards. The strength of registries such as ROAR and OAIster is that they accept whatever metadata institutions supply, with no attempt to review or regularize input standards for different data elements, and they also ignore an individual institution's selection criteria and simply provide access to whatever content the registered institutions choose to supply.

VI. Marketing

When we market Scholars' Bank, we emphasize a number of key points. Those key selling points (taken from the *What's in it for me?* page for faculty at http://libweb.uoregon.edu/catdept/irg/Faculty_Reasons.html) include:

- Scholars' Bank can collect in one place the results of faculty research, whether it be a conference presentation, working paper, peer-reviewed article, class lecture, book review, data set, or other type of work. For examples, see <https://scholarsbank.uoregon.edu/dspace/handle/1794/691>

- Items submitted in Scholars' Bank can appear in more than one collection, such as a personal collection and a departmental collection. For an example of such an item, see <https://scholarsbank.uoregon.edu/dspace/handle/1794/2215>
- Scholars' Bank provides a unique - and stable - URL for each work that can be cited and linked to.
- Every work is archived and preserved at no cost to faculty or to their home departments.
- Submissions are indexed in Scholars' Bank and available on the Web immediately after approval by the collection's reviewers. (*If a collection has no reviewer, the work becomes available without any delay.*)
- Text-based materials are indexed automatically so that the full text is searchable.
- Google, Google Scholar and other similar services routinely harvest information from Scholars' Bank and add links to their indexes so that works submitted here may also be found through searches on the open Web. (*Be aware that such indexing sometimes takes a few weeks.*)
- Submissions are immediately date-stamped, marking the content as having been created by the faculty member on the date of submission. (*Earlier dates of availability or publication can also be added.*)
- Citations can be added to the records in Scholars' Bank to link to formal publications or to versions that have been previously made accessible through other venues.
- Faculty collections can be established in such a way that they mirror the peer-review process, if desired.
- Faculty can select their own keywords or terms to describe the subject of their work, if desired.
- Faculty can handle the submissions themselves without any library involvement (*once the library sets up a collection for them*) or they can arrange to have the library do all the work for them, free of charge

For students, there is one additional marketing point taken from the *What's in it for me?* page for students at http://libweb.uoregon.edu/catdept/irg/Student_Reasons.html :

- With the support of a faculty member or academic department, Scholars' Bank provides a forum for student work that would otherwise be hard to locate or would be completely unavailable to the general public.

Marketing entails getting the word out anywhere and everywhere. This has been discussed numerous times in presentations and articles by Hixson, Breakstone, Briston, and Jenkins – all of which can be found in Scholars' Bank. Some of the strategies we have employed to market Scholars' Bank, locally and beyond include:

- Developing local context
 - Provide a general overview without jargon
 - Personalize the information
 - Answer questions of local interest
 - Build in redundancy so that the same information is available in multiple places
 - Provide varying levels of detail allowing people with deeper needs or interest to drill down to more detail
- Explaining the broader context

- Appeal to their professional self-interest
- Link to worldwide movement
- Educate on issues
- Obtaining a critical mass

Our definition has expanded, enabling us to go after a broader range of content, much of which would not otherwise be archived – or not effectively and reliably archived. Some of the content we have acquired falls into the following categories:

 - Faculty output
 - Individual works and scholarly journals
 - Library collections
 - Electronic records, finding aids, historical materials, etc
 - Campus publications and presentations
 - Newsletters, planning documents, historical publications, streaming video, etc.
 - Student works
 - Theses and dissertations, honors projects, etc.
 - Miscellaneous
- Obtaining interesting and attractive content
 - Oregon Quarterly, Dissociation, etc.
- Being open to anything (almost)
 - Doing whatever it takes to get content in and searchable (digitize, submit for authors, apply metadata creatively to improve searching and display)
- Making connections wherever possible

Another of our key strategies is to make appropriate connections, bi-directionally whenever possible:

 - One way is to register the IR in appropriate registries, such as OAIster and the Institutional Archives Registry
 - Provide links to subject or keyword searches in our local catalog
 - Make connections between collections (Office of President in CONTENTdm and related collection in Scholars' Bank)
 - Catalog collections in Scholars' Bank and link back from the library's online catalog
 - UO Channel site and Scholars' Bank – bi-directional link
- Publicity – being creative and shameless
 - Promote anywhere and everywhere
 - Inside Oregon, Oregon Quarterly and other articles about Scholars' Bank
 - Use subject specialists in the effort
 - Make connections with communities and assist with outreach
 - Share success stories
 - Hilary Hart's testimonial
 - Work with known allies
 - Economics Department is good example
 - Never miss an opportunity
 - Getting winners of Undergraduate Library Research Awards to speak about the archive
 - Doing presentations at every opportunity, on and off-campus
 - Writing articles about our work
- Fast turnaround and follow-up
 - Set things up, at least in the test site, within 48 hours of getting the go-ahead from a community
 - Lift contextual information from departmental and other web sites, submit some sample files with proposed metadata, send them URL to react to

- All new communities and collections are tracked on a spreadsheet, with names and emails of contact people
- Improving the suite of services
 - Full-text searching, statistics, mapping items to multiple collections
 - Web-based permission forms
 - Digitization of alumni theses and dissertations
- Borrowing ideas from other archives

VII. Future challenges

There are many challenges that face Scholars' Bank. Continuing to identify and acquire appropriate content is still paramount. Beyond that, the top challenges that I would identify for the future are:

- Preservation – getting serious about it and investing some institutional resources to support it
- Metadata
 - standardizing it more and documenting the standards
 - adding metadata to identify peer-reviewed versus non-peer-reviewed content; faculty versus student; etc.
- Improving the search interface and developing user guides for searching
- Overcoming the resistance of a few individuals in the Graduate School who are blocking the easy acquisition of some of the most important content created at this university
- Investigating new software options and possibly migrating Scholars' Bank to another software platform
- Integrating Scholars' Bank (and other UO-created digital content) in with other UO owned or licensed content. Currently, this content has been marginalized and has been made difficult to find within the library context. Given the use of these resources worldwide, it is ironic just how difficult finding them within the library has been made. Before the library web site redesign, we had succeeded in getting a place for digital content, including Scholars' Bank, on the top page. However, because a handful of students in a focus group said they didn't know what "digital collections" meant, all the important digital content that the Libraries spend so much time and effort creating has been hidden. Ironically, the world knows where Scholars' Bank and the other digital content is and uses the sites heavily (5,364,947 hits against Scholars' Bank in the past 12 months, with files being opened 30% of the time). But our own users are denied easy access because of the Libraries' web site design. (Author's note: this situation was corrected regarding Scholars' Bank, but not the other digital content, on September 13, 2006.)