

# Executive Summary

This document was created to highlight opportunities and drive discussion for the OCUL consortium in both the short term through the launch of a new Scholars Portal server in 2008, and in the long term by incorporating more 'social' means of sharing and organizing information within OCUL's Scholars Portal and the larger academic community that it serves. It was created by Scholars Portage librarians subgroup of the Scholars Portal Public Services Advisory Group.

With the implementation of Scholars Portal 2, a replacement for the existing online journal hosting service of Scholars Portal, comes a great opportunity to incorporate new features and new services to the OCUL community of students and researchers. This document was created with the intention of starting a discussion among OCUL librarians about what we would like to see be made available.

Our white paper begins with a brief survey of the landscape and a discussion of seven areas that are challenges to our current environment:

- User expectations are not being met
- Academic work is social
- Citations are hard work
- Academic work is not connected
- Ascendancy of Web 2.0 applications
- Our advanced features are not being used
- Metrics of scholarly authority are changing

Recognizing that many of these areas overlap, we suggest three main areas, with specific recommendations for each, where our institutions can help to alleviate these pressures.

Enhance and improve the user interface

- Enrich Scholars Portal content by bringing in metadata from sources outside the journal repository
- Explore the implementation of controlled vocabulary, thesauri and authority control
- Add user tagging functionality

Connect the citation network to user workflow

- Provide table of contents (TOC) RSS feeds with links that facilitate authentication. If it is possible, allow users to generate their own RSS feeds.
- Provide users of scholarly resources with social bookmarking services
- Consider services that support the whole of the user's research process and the development of online space for OCUL research communities.
- Seek means for Scholars Portal to be integrated into Learning Management Systems used by OCUL

Embrace standards and technologies that will allow present and future network discovery systems to make use of what we offer

- Provide both permalinks as well as COinS OpenURLs in the Scholars Portal server and to encourage OCUL libraries to adopt their own versions of LibX or promote other COinS readers
- Investigate how to take advantage of the attribute-based information that Shibboleth can provide
- Consider what semantic metadata could be provided through Scholars Portal

It is planned that a demo of the new server will be made available to OCUL librarians in October of 2007 and a beta version be developed for the spring of 2008. We heartily invite readers to comment.

# Scholar 2.0 : a white paper by Scholars Portage

## 1. Introduction

It is difficult for researchers who collaborate with others outside of their own institution to share and organize their work. We know this because we, the authors of this document, are from two different institutions and we ran into these problems ourselves.

This document was created by [Scholars Portage](#), librarians from the [Scholars Portal Public Services Advisory Group](#), to highlight opportunities for the [OCUL consortium](#) by incorporating more 'social' means of sharing and organizing information within OCUL's Scholars Portal and the larger academic community that it serves.

[Scholars Portal](#) was initiated in 2001 and since then it has developed into one of the world's largest collections of shared scholarly information resources and services. Scholars Portal hosts locally-loaded online journals on the Scholars Portal Server, offers cross-database searching through Scholars Portal Search, facilitates interlibrary loan service through RACER, and brings these resources together through its "GetIt!" OpenURL service and locally hosted RefWorks, an online personal bibliography service.

The existing Scholars Portal Server is set to retire and a new server will host OCUL's online journal content in 2008. With this new server comes a great opportunity to incorporate new features and new services to the OCUL community of students and researchers. This document was created with the intention of starting a discussion among OCUL librarians about what we would like to see be made available.

To get the conversation started, we begin our white with a brief a survey of the landscape to identify some areas where academic libraries are not meeting their users' needs and expectations. We then suggest three different areas where our institutions can improve the situation at hand by:

1. enriching Scholars Portal content by bringing in metadata from sources outside the journal repository
2. connecting the citation network to user workflow
3. embracing standards and technologies that will allow present and future network discovery systems to make use of what we offer

Or, as Emily Lynema of NCSU recently suggested in her presentation at the 2007 AALL Annual Meeting *Bringing the library to the user: the practice* ([ppt](#)) we suggest means by which we can improve:

1. Usability
2. Findability
3. Remixability

It would be deeply ironic if a group of librarians who wanted to incorporate more social means of sharing and organizing information did not reach out to their own community to help shape this vision. As such, we have publicly posted this document at <http://www.scholarsportal.info/commentpress> and heartily invite readers to comment and in doing so, let Scholars Portal better understand how to shape its development to meet present and future needs.

## 2. The landscape

There is little doubt that the library landscape is shifting under our feet. The past two years have seen: the adoption of new search and discovery tools (NCSU's Endeca-enhanced library catalogue, Georgia Public Library Service's open source PINES library catalogue, federated and metasearch products such

as Metalib, WebFeat and others); significant challenges to traditional bibliographic control (The Calhoun Report ([pdf](#)), The University of California's Bibliographic Services Task Force Report ([pdf](#)), and the formation of the Library of Congress [Working Group on the Future of Bibliographic Control](#) ); and the insertion of Web 2.0 technologies and terminologies into the library lexicon (blogs, wikis, chat reference, rss feeds, social software). All of these changes demand that we move beyond the confines of our systems to engage with our users in spaces that match their workflow and meet their needs. Our systems must become more user-centric if we are to have any traction in the new landscape.

Karen Coyle recently outlined what the user-centric library system requires:

...systems that are highly flexible that can be made to respond to observed patterns of use... personalization, but not only in terms of default searches or saved e-mail addresses but also in allowing the user to capture, link, or modify, to mash-up library data with other networked environments, such as taking their personal bookshelf into a virtual learning space. User-centricity also means providing services that do not end with the delivery of a book or article but that go forward into the use and re-use of resources. Just as a library provides services other than document delivery, a library system can provide virtual study space, group activity areas, reading groups and study aids. It can link to the classroom, virtual, or physical.<sup>1</sup>

In this document, we have tried to highlight some areas where this changing landscape is leaving the users of the academic library exposed to the elements.

This section is broken out into the following areas:

- 2.1 User expectations are not being met
- 2.2 Academic work is social
- 2.3 Citations are hard work
- 2.4 Academic work is not connected
- 2.5 Ascendancy of Web 2.0 applications
- 2.6 Our advanced features are not being used
- 2.7 Metrics of scholarly authority

## ***2.1 User expectations are not being met***

**Our current service environment is not meeting the needs of our users as they are choosing to do their research elsewhere.**

**Question: Should academic libraries make it a goal to mimic the qualities of popular online services such as Google Scholar or should we strive to differentiate ourselves?**

*"Stand on the shoulders of giants" - [Google Scholar](#)*

Students no longer put up with faint photocopies of a list and scribbled changes copied from the board. Most expect - and receive - booklists online, with up-to-date amendments following fast on the publication of new texts. Nor are they so willing to run to the library to grab the only copy of a key volume or wait until it is returned. Many will start their research with Google and expect booklists to include links that take them straight to the recommended text. [[Link](#)]

OCLC's 2005 survey [College Students' Perceptions of Libraries and Information Resources](#) found that 70% of college students start their research with a search engine as opposed to those 10% who started with an online library and the 14% who began their work in a physical library.

[April 2007 statistics from Neilson//NetRatings report](#) that Google currently holds a 55.2% market share of search engine traffic. What 'market share' Google has of 'scholarly' search engine traffic is not known but with their specialized services of Google Scholar and Google Books, it very well may be considerably larger than 55%.

Who would have guessed that in June 2005 Google would account for over 56% of referrals to research articles in HighWire journals, while PubMed Central, a renowned life sciences repository, would account for less than 9%? If that stat isn't stunning enough, 72% of scholars surveyed for a report on self-archiving confessed to using Google to find scholarly literature on the web. [\[Link\]](#)

Google Scholar is in the top ten sources of incoming OpenURL links to library-licensed content for all the respective OCUL libraries, and as such, is a larger source of full-text article requests than many other expensive indexes.

There is also growing evidence of a disconnect between what library needs of faculty are and what librarians believe are library needs of faculty. In the summer and fall of 2006, a survey was commissioned to better understand the attitudes and perceptions of collection development librarians and faculty towards the transition to an increasingly online environment. The survey found:

In the future, faculty expect to be less dependent on the library and increasingly dependent on electronic materials. By contrast, librarians generally think their role will remain unchanged and their responsibilities will only grow in the future. Indeed, over four-fifths of librarians believe that the role of the library as the starting point or gateway for locating scholarly information will be very or extremely important in five years, a decided mismatch with faculty views.

Perceptions of a decline in dependence are probably unavoidable as services are increasingly being provided remotely, and in some ways, these shifting faculty attitudes can be viewed as a sign of the library's success. The mismatch in views on the gateway function is somewhat more troubling: if librarians view this function as critical but faculty in certain disciplines see it as declining in importance, how can libraries, individually or collectively, strategically realign the services that support the gateway function? [\[link\]](#)

## ***2.2 Academic work is social***

**The current online library environment can be seen as inhospitable to collaborative research.**

**Question: How can we support researchers as they work with increasing numbers of collaborators from increasingly far-flung locations and frequently across disciplines?**

*"If I have seen further it is by standing on ye sholders of Giants." - Issac Newton<sup>2</sup>*

"One study conducted by the Sante Fe Institute found that the average high-energy physicist now has around 173 collaborators. The same study found that the average number of authors per scientific paper has doubled and tripled in a number of fields. A growing number of papers have between two hundred and five hundred authors and the highest ranking paper in the study had an astonishing 1,681 authors." <sup>3</sup>

Collaboration is at the heart of scholarly pursuits for all disciplines. Collaboration can be defined as both building on the work of others and working directly with others. Why do libraries not make the

collection and sharing of research a seamless process? Why do academic libraries lag behind other non-academic environments (MySpace, Facebook) in offering collaborative online environments given that there is an arguable need to provide these services?

As part of a larger [Mellon-funded project on Academic Support](#), the University of Minnesota surveyed and interviewed faculty members and graduate students in sixteen departments in the humanities and social sciences. Of note, the survey found that:

- 68% - Faculty work collaboratively
- 52% - Collaborate with colleagues at other institutions
- 46% - Find that distance from colleagues is a collaboration obstacle

The American Historical Society devoted a recent issue of their journal to "[History and the Changing Landscape of Information](#)" and in one article about historical scholarship and [Zotero](#) (a free tool that allows for the collection, management, and citation of research sources in the Firefox web browser) the author suggests future possibilities for scholarly tools like Zotero:

The benefits of such networking and the emphasis on semantic entities like books, articles, and letters opens up new possibilities for scholarly communication. A group of historians interested in a topic lacking a chapter in the *Guide to Historical Literature* will be able to build a bibliography of important works in their field collaboratively, which then could be shared with students. Scholars will be able to track more easily publications of interest in their fields and hear of archival documents newly discovered or scanned by other Zotero users, based on tags, recommendations, and the holdings of personal collections. Historians from around the globe will be able to combine virtually to annotate a primary source that has just been digitized and placed online. Most intriguingly, this interaction of people, tools, and resources—what we might call an "ecology of scholarship" (which undoubtedly will include software other than Zotero)—perhaps will lead to the discovery of new knowledge by aggregating and analyzing our shared wisdom. [[link](#)]

Already, scholarly communities are forming around these new scholarly technologies:

Scientists involved in [OpenWetWare](#), an MIT project designed to share expertise, information, ideas in biology, are heralding the arrival of Science 2.0. Twenty labs at different institutions around the world already use the wiki-based site to swap data, standardize research protocols, and even share materials and equipment.<sup>3</sup>

[IISC](#), an organization established to support teaching and research through information technology funded by all the higher education institutions in the UK, is currently funding a project called [myExperiment](#), a "Virtual Research Environment which makes it easy for people to share experiments and discuss them". IBM's Collaborative Research Group has launched [Many Eyes](#) where users can share and discuss [data sets](#) and [visualizations](#).

## **2.3 Citations are hard work**

**We need to reduce the drudgery our users experience in the largely mechanical tasks of gathering, organizing and formatting citations and in doing so, strengthen the nodes of the citation network.**

**Question: To what extent should an academic library be responsible for the personal collection of citations of their users?**

*But citations and references are not only essential aids to scientists and scholars concerned to verify statements or data in the citing text or to retrieve further information. They also have not-so-latent symbolic functions. They maintain intellectual traditions and provide the peer recognition required for the effective working of science as a social activity. All this, one might say, is tucked away in the aphorism that Newton made his own in that famous letter to Hooke where he wrote: "If I have seen further, it is by standing on ye sholders of Giants." [pdf]*

Undergraduates are not the only user group of the academic library that run into difficulty in the management of their citations and the application of citation styles:

Wates and Campbell looked at copy editing changes carried out on a set of science, humanities and social science articles at Blackwell Publishing (as was) and reported that the biggest category of corrections by the publisher was concerned with the references (42.7% of all copy editing changes), the next biggest category (34.5%) was concerned with minor syntactical or grammatical changes and a small proportion (5.5%) of changes corrected author 'errors that might otherwise have led to misunderstanding or misinterpretation' [link].

As more 'digital natives' start creating and using 'natively digital' research, it can only be expected that the use of digital tools to handle online scholarship will grow.

Although a recent University of Minnesota survey found that only 8 percent of humanities scholars currently use reference managers like EndNote (compared to 36 percent in the social sciences), the usage of these tools will likely spread in the coming years. Graduate students are already over 50 percent more likely than existing faculty to use such software, and as historians of all ages become more comfortable with digital research, their use of digital tools to assist their studies will likewise expand. Moreover, with vast new historical resources rapidly moving online due to ambitious efforts such as Google's library digitization project, we will all need better ways of dealing with countless online documents without printing everything out or using note cards. [link]

## **2.4 Academic workflow is not connected**

**The integration of library tools and learning and teaching technologies should be a priority.**

**Question: Is the academic library's traditional role of knowledge gatekeeper holding back efforts to get into our users' workflow?**

*If I have not seen as far as others, it is because giants were standing on my shoulders.*

Many faculty want to build their own re-aggregated resources, using their own materials and mixing them with resources they have collected along the way. They are concerned about the significant inadequacy of the classroom technologies available to them. How to first manage the array of available resources, and then integrate them into teaching practice is a concern for those who are involved in supporting faculty pedagogies and developing useful technical tools. For faculty, there may be an array of tools available for collecting, developing, managing, and actually using resources, but the efficacy and interoperability of these tools for the immediate tasks that faculty need supported are questionable. A related issue is the integration of learning management systems (LMSs) with library resources and other course content. Current LMSs appear to have limited overall functionality, especially

since they may not allow easy integration with the diversity of digital resources that faculty want to use. [[link](#)]

The records contained in library catalogues cannot be bookmarked because OPACs insert a [session ID](#) number into each item's URL and this causes the link to 'expire' after a short time. Most articles in library-licensed databases cannot be easily bookmarked without an understanding of what a durable or permanent URL is as well as an understanding of their library's authentication technology -- which in the case of EZ-proxy, inserts an institution-specific prefix into the URL. Without readily available permalinks, users cannot easily share information about library resources via email, blog post, social bookmarking service such as [del.cio.us](#), or scholarly communication. In the current environment, it is surprisingly difficult to create something as simple as a reading list with links to each item.

The fact that the most-cited reasons for *not* using digital resources was that they simply do not mesh with faculty members' pedagogies is worth noting and has implications for those who wish to increase technology adoption by faculty. We should not expect faculty, who we can assume know more about teaching their subject than non-specialists, to shoehorn their approaches into a technical developer's ideas of what is valuable or the correct pedagogical approach. Tools and resources must be developed to support what faculty want to do. [[link](#)]

The OCUL consortium collectively subscribes to [RefWorks](#), an "online research management, writing and collaboration tool". RefWorks is accessible from within Microsoft Word through WriteNCite, allows for integration with the [Blackboard](#) course management system, and can generate publicly accessible reading lists through RefShare. But a number of OCUL institutions have course management systems other than Blackboard, have students and faculty who do not want to use Microsoft Word for their writing, and most unfortunately, RefShare does not generate OpenURL links that can be used outside of one's own institution alongside the citations it does make available. Fewer than 10% of OCUL faculty and students make regular use of RefWorks (although this is a respectable 30,000 users) and the take-up of RefShare has been surprisingly small.

The integration of RefWorks with the indexes and full-text databases that libraries subscribe to from commercial information providers is done on a case by case basis as these companies form strategic [linking partnerships](#). Sometimes these partnerships result in a poorer user experience as Mark Leggott, University Librarian of UPEI recently discovered:

I was surprised (and also not) recently when we asked the vendor ([ISI](#)) of one our oft-used databases, Web of Science, to remove a direct link to their citation product, [Endnote Web](#). The initial response to our request was "Yes, we can remove that for you and will shortly." A couple of weeks later we got a response back that said "The previous response was wrong, we cannot have the Endnote Web link removed."...

They not only make it extremely (sp) cumbersome to export citations to RefWorks, but they also embed a link to their own competitive (sp) product into their database interface. Now that they have resorted to questionable practices such as this (not even close to being in the best interest of their customers), how long will it be before people start looking at moving to alternative database options that do provide the desired flexibility (sp)? [[link](#)]

Ideally, our users should have a scholarly environment that does not require the acquisition of one or set of particular commercial software because its makes use of a proprietary format. What our online environment is missing is a robust [bibliographic backplane](#) or strong [bibliographic tissue](#).

## **2.5 Ascendancy of Web 2.0 applications**

### **Move from Scholars Portal 1.0 to Scholars Portal 2.0**



**Question: Should academic libraries develop social networking platforms or work towards integration with existing systems? Or is this a matter that will likely 'blow over' in time?**

*"I say with Didacus Stella, a dwarf standing on the shoulders of a giant may see farther than a giant himself." [\[link\]](#)*

There is a big unanswered question and that is how do Web 2.0 social networking models fit within the already well established social information networking that underlies scholarly activity? Peer review encompasses a lot more than just accepting journal articles for publication -- peer standing is the life blood of promotion and advancement. How does this kind of life and death social sharing translate to an environment where people are sharing vacation photos and movie reviews? The other dilemma is that students are often warned not to share information as they work on papers or assignments -- how do social networks fit within an environment with major concerns about copying and plagiarism? ([Alan Darnell](#), personal email to authors, April 29, 2007)

[According to Forrester Research](#) the number of 18 to 21 year old Americans who use social networking sites, such as Facebook or MySpace, is 70%. Peter Brantley, Executive Director for the [Digital Library Federation](#), a not-for-profit international association of libraries and allied institutions, made this observation as he noticed the graduate students attending Scholarly Communications Institute had all connected in Facebook for internal communication during the conference:

First, this is a fundamentally important shift generationally in what we expect from our software productivity tools. The grad students and young faculty using Facebook have used MySpace, and been Facebook members through their whole adolescent and adult school experiences. They are taking this experience with them into their work. The work of the people that I see most often is in research and teaching. But the lesson is broader: this generation will be working collaboratively in tools like Facebook. In schools, in corporations, in small non-profits, in community centers - people will collaborate and work together in social applications. And that is going to be as natural to them as email and text messaging. Second, regardless of the ultimate fate of Facebook, the set of characteristics that it has established - the sense of community; user control over the boundedness of openness; support for fine grained privacy controls; the ability to form ad-hoc groups with flexible administration; integration and linkage to external data resources and application spaces through a liberal and open API definition; socially promiscuous communication - these will be carried with us into future environments as expectations for online communities. Facebook is an empty wasteland for people who have not climbed over the hump of use. For those who have active community within it, it is this generation's Lotus 1-2-3. [\[link\]](#)

The commercial sector of the scholarly world is investing in social networks for academia:

[What is Nature Network?](#)

Connecting scientists at a global and local level.

Nature Network is the online meeting place for you and fellow scientists to gather, talk and find out about the latest scientific news and events. Science is an international endeavor and deserves a global stage for discussion. Scientists can also benefit from interactions at the local level. That's why you'll see an increasing number of local city hubs on Nature Network, starting with Boston and London....

Clearly there is a value in the communication of the underlying social networks of academia as Proquest/CSA sells this information as a product:

[COS Scholar Universe](#) provides direct access to more than 1.3 million active researchers working in over 200 disciplines and 9 countries. Searches in COS Scholar Universe provide unparalleled exposure to the community currently studying a topic of interest. Exploring this living research community expands awareness and access to cross-disciplinary research perspectives, publications, research methodologies, expertise, and collaborative opportunity. With COS Scholar Universe, your faculty and students are directly connected to the community of scholarly research.

*Benefits of COS Scholar Universe: Faculty*

- Find out who shares your research interests outside your discipline;
- Inform your teaching, research and writings with perspectives from Scholars from multiple fields to immediately make your work more inter-disciplinary;
- Easily find and contact active thought leaders in your new areas of interest;
- Find researchers with experience using the tools and materials you need to advance your research;
- Edit your profile to ensure the most accurate professional representation of yourself and your research.

## **2.6 Our advanced features aren't used**

**Academic librarians must seek and provide alternatives to advanced search screens and Boolean searching.**

**Question: How do we improve our systems and interfaces to account for the fact that users perform simple keyword searches and do not use advanced search functions or Boolean search techniques?**

*Pigmaei gigantum humeris impositi plusquam ipsi gigantes vident* [\[link\]](#)

Karen Markey recently reviewed 25 years of published research findings on end-user searching in online information retrieval systems and found most studies reported:

- 2 to 4 queries per session
- 2 to 3 terms used per query
- less than 15% of queries bear the Boolean AND operator
- less than 3% of queries bear the Boolean OR operator
- end users seldom use advanced search and when they use these features in their queries, they use them incorrectly about a third of the time
- the vast majority of end users are satisfied with the results of their searches<sup>4</sup>

In 2005, [The Pew Internet & American Life Project](#) found of the American adults who use the Internet, 63% used a search engine on an average day and that daily search engine use was slowly edging up on daily email use. In [a slightly earlier survey by the same group](#), it was found that 92% of those who use search engines say they are confident about their searching abilities, with over half of them, 52%, saying they're "very confident".

Even with a greatly improved interface for their Endeca-enhanced library catalogue, the librarians at NCSU have found that while the relative percentage of use of their classification and subject heading facets has risen relative to other facet selections (such as language, library, format, author, etc.) keyword searches with no additional faceted navigation still constitute 48% of searches [\[ppt\]](#).

It has been recognized that faculty and other 'domain knowledge experts' tend to employ search strategies that avoid subject searching altogether. In other words, many domain experts prefer to navigate the citation network<sup>5</sup>. This is a strategy that doesn't only benefit the expert user. In an article dedicated to the research processes of undergraduate students, Barbara Fister suggests that

Finding tools are not always the best route to good evidence. Our search strategies quite often describe the information-seeking process as one in which tools -- reference works, bibliographies, catalogs, indexes, are used successively and systematically to locate information is located through finding tools. In fact, students (and most other researchers) find the most direct and efficient route to sources through the citation network... If students find much of their material through the citation network and through serendipitous browsing of shelves, we should point those out as factors in the search strategy rather than emphasizing the use of privileged bibliographic tools as the correct way to locate information.<sup>6</sup>

## 2.7 *The new metrics of scholarly authority*

**We need to be more aware of the issues of metrics, scholarly communication and authority in online environments.**

**Question: How can academic libraries support the open exchange of research information and support fair and open metrics?**

*standing on the shoulders of giants / leaves me cold*  
- "King of Birds", R.E.M.

Michael Jensen, the director of strategic Web communications for the publisher [National Academies](#) recently weighed in on his thoughts on [the New Metrics of Scholarly Authority](#):

I hope it's clear that I'm not saying we're just around the corner from a revolutionary Web in which universities, scholarship, scholarly publishing, and even expertise are merely a function of swarm intelligences. That's a long way off. Many of the values of scholarship are not well served yet by the Web: contemplation, abstract synthesis, construction of argument. Traditional models of authority will probably hold sway in the scholarly arena for 10 to 15 years, while we work out the ways in which scholarly engagement and significance can be measured in new kinds of participatory spaces.

But make no mistake: The new metrics of authority will be on the rise. And 10 to 15 years isn't so very long in a scholarly career. Perhaps most important, if scholarly output is locked away behind fire walls, or on hard drives, or in print only, it risks becoming invisible to the automated Web crawlers, indexers, and authority-interpreters that are being developed. Scholarly invisibility is rarely the path to scholarly authority [[link](#)]

Michael Jensen contends that the reason why new tools to confer scholarly authority have to be expected is because scholarship is entering a period of unprecedented research abundance. Already there has been some development in both dealing with scholarly abundance and the granting of scholarly authority.

Recognizing that that physicians' daily routines afford them little time to search and review new medical information and that physicians admit that they lack the skills required to navigate literature databases and properly appraise medical literature, the Health Information Research Unit of McMaster University developed a system in 2005 called [MORE \(McMaster Online Rating of Evidence\)](#) that filters research articles through critical appraisal and a clinical relevance rating system and then channels these articles to physicians according to their practice disciplines [[link](#)]. This service is now incorporated into

[bmjupdates](#) a collaboration between the McMaster Health Information Research Unit and BMJ Publishing.

[BMJ Publishing Group](#) is also responsible for two products designed to deal with the problem of research abundance: [Faculty of 1000 Biology](#) and [Faculty of 1000 Medicine](#). These products are promoted as authoritative services in which, not 1000, but almost 5,000 researchers and clinicians "share their expert opinions by highlighting and evaluating the most important articles in biology and medicine". Of note, the Faculty of 1000 web site promotes the fact that "Recommendations and interpretations are based on the article's merits, not the journal's". This emphasis is especially interesting as The Faculty of 1000 is a paid-subscription service that is produced by BioMed Central, a private UK-based publisher of 179 peer-review open access journals. Their research is free but their authority is not.

Care must be taken to ensure that scholarly authority does not become an even more commercial product than it has already become:

For some scientists, it's the height of irony: The scientific method depends on being able to check what others have done. But the growing use of citation statistics and impact factors in academe goes against that basic principle, because researchers say they cannot assess the basic data used to produce those measurements, which are controlled by a private company, Thomson ISI. [[link](#)]

### 3.0 Unlocking Opportunity: Scholars Portal 2

*"In computer science, we stand on each other's feet" - Brian K. Reid*

The current environment of Scholars Portal makes the easy manipulation of data challenging and offers minimal scalability and minimal opportunity for the incorporation of Web 2.0 technologies. We have rich content and a rich network of relationships within that content. At 12 million articles from close to 20 publishers, tied together by over 100 million citation links between those articles, the Scholars Portal e-journal collection represents one of the largest aggregations of scholarly publishing ever assembled. But surfacing that content and making it usable is more difficult than it should be, in part due to the limitations of available technologies.

The ScienceServer system has been in operation since January 1997 when it was implemented by the University of Toronto, and it has served the Scholars Portal community well. For several years the Scholars Portal staff has been working with Endeavor (and subsequently with Ex Libris) to implement JOS (Journals OnSite) as a replacement for ScienceServer. Doubts about the future of this platform led to the commitment by the OCUL Directors to support the implementation of a new e-journal platform based on the Mark Logic Enterprise Server system. Subsequently, Ex Libris has announced support for ScienceServer / JOS will cease at the end of 2009...Mark Logic stores XML documents, an encoding format increasingly being used by our publishing partners, in native format. It builds indexes not only on the words in the documents but also on the XML elements and attributes that give structure and semantic meaning to those words. Using XQuery as its search language, Mark Logic supports searching of every element or attribute in a document. [Scholars Portal Annual Report 2007]

With the migration of Scholars Portal to Mark Logic, an important opportunity has been created to incorporate Web 2.0 technologies into the discovery layer of this service. With strong support for XML, XQuery, relevance-based searching, facet-based browsing, thesaurus expansion, language-based stemming and collations, automatic classification, web services, AJAX and other agile programming approaches, Mark Logic gives us the kinds of tools we need to incorporate the best of current Web technologies into a new Scholars Portal interface.

Which technologies, though, and how should these be presented?

In this section, we have tried to highlight some areas where Scholars Portal 2 offers opportunities to improve upon Scholars Portal recalling our main themes of **Usability**, **Findability** and **Remixability**.

### 3.1 Surfacing metadata and exposing content

- 3.1.1 XML and XQuery
- 3.1.2 Metadata enrichment
- 3.1.3 Controlled Vocabulary and Thesauri
- 3.1.4 Tagging

### 3.2 Connecting the citation network

- 3.2.1 RSS feeds
- 3.2.2 Social Bookmarking
- 3.2.3 Collaborative research spaces
- 3.2.4 Learning Management Systems

### 3.3 Embracing standards

- 3.3.1 Permanent URLs/COinS
- 3.3.2 Shibboleth
- 3.3.3 Semantic Scholars Portal

## 3.1 Surfacing metadata and exposing content

All too often libraries "hide" good metadata. Cataloguers create rich data in the form of MARC records, most of which is not utilised by most OPACS, publishers create content which is not used by libraries, and lots of useful content exists in other sources. Scholars Portal 2 offers the opportunity to expose more metadata, supplement internal data from external sources, collect and store user-generated content, and leverage relationships inherent in full text documents. This enables us to rethink our monolithic concepts of bibliographic data, allowing for multiple access points and multiple views, doing away with the notion of "order making" where such a concept is truly impossible. [[link](#)]

Metadata is not monolithic. Instead, it is helpful to think of metadata as multiple views that can be projected from a single information object. Such views can form the basis of customized information services, such as search engines. Multiple views -- different types of metadata associated with a Web resource -- can facilitate a "drill-down" search paradigm, whereby people start their searches at a high level and later narrow their focus using domain-specific search categories. [[link](#) ]

### 3.1.1 XML+XQuery=Data and content happiness

With ScienceServer and most traditional search engines, content is maintained separately from indexes. Content lives in file systems or relational databases and is harvested to pull out specific data elements that have been selected for indexing and retrieval. Most library OPACs work the same with, with many of the MARC fields not being indexed and therefore not available for retrieval. MarkLogic, in contrast, stores every element of data, indexing every element and all the structural relationships between those elements. Adding new access points does not require re-harvesting the data and re-indexing according to a new schema. All access points are always exposed in MarkLogic, accessible through a query language designed specifically for structured XML data. All content (records/citations and full text) is stored as [XML](#). [XQuery](#), a language specifically designed for accessing XML content, queries the XML content. This is where the happiness lives.

As increasing amounts of information are stored, ex-changed, and presented using XML, the ability to intelligently query XML data sources becomes increasingly important. XQuery is designed to be a language in which queries are concise and easily understood. It is also flexible enough to query a broad spectrum of XML information sources, including both databases and documents. [SP Annual Report, 2007]

There is a great deal of flexibility in what elements can be searched in Scholars Portal 2. Beyond the traditional, highly structured data of the index records, the full-text content and content out in the web are available to be leveraged by XQuery. This massive amount of content can then be manipulated in some pretty interesting ways, easily scaling up or down depending on the need.

The sections below outline some of the possibilities.

### 3.1.2 Metadata enrichment

#### **Recommendation: enrich Scholars Portal content by bringing in metadata from sources outside the journal repository**

Scholars Portal 2 will utilize highly structured metadata already present in Scholars Portal, with the addition of offering more structured data through faceted classification and automated classification, while also adding the ability to collect enriched metadata from a wide variety of sources images, TOCs, information from external sources (Wikipedia, for example). This enriched metadata offers interesting opportunities in allowing researchers to better assess content and also move easily from an article, to related information.

Potential additions are:

- Reviews
- Table of contents
- Images
- Web content (Wikipedia)
- Article "kind": review article, peer review article, book review
- Enhanced author information
- Reading lists

### 3.1.3 Controlled Vocabulary & Thesauri

#### **Recommendation: explore the implementation of controlled vocabulary and thesauri.**

Because Scholars Portal is an e-journal repository which collects content across a wide swath of disciplines authority control and controlled vocabulary have always been problematic. Various thesauri (MeSH, LCSH) can be integrated into Scholars Portal 2, improving retrieval and while at the same time enabling clustering of like subjects and names. While most researchers may prefer to use keyword searching instead of structured browse searching [cite] the implementation of some form of controlled headings will allow for guided navigation and aid in providing better linkages between documents. While users may not choose to start searches with controlled vocabulary, surfacing terms in the "post search" environment may increase success. Or may not...

#### **More Indicators That Users Could Benefit From Direct System Intervention**

When researchers analyze end-users' failed searches, the number one problem is their initial choice of search terms (Debowski, [2001], p. 377; Hsieh-Yee, [1993], p. 169; Lucas & Topi, [2002], p. 105; Sewell & Teitelbaum, [1986], p. 241; Wildemuth & Moore, [1995], p. 299).



Instead of using a database's controlled vocabulary, users search for the first terms that come to mind. Failing to use the controlled vocabulary has an adverse effect on the precision of their searches and makes it impossible for users to enlist the vocabulary's special search features such as exploding terms, listing subheadings, and displaying term relationships. Researchers who compare end-user search performance to expert searchers remark on the various ways in which the latter enlist controlled vocabulary in their searches (Lancaster, Elzy, Zeter, Metzler, & Low, [1994]; Rudner, [2000]; Sutcliffe, Ennis, & Watkinson, [2000]). A longtime reference librarian assisting users at the Library of Congress, Mann ([2003]) observes "Even minimal experience working with readers at a reference desk - or, better, standing over their shoulders at a computer terminal - will demonstrate very clearly that most [us]ers, lacking either prior instruction or point-of-use assistance, are simply incapable of coming up with the best terms" (p. 53). [24]

The provision of controlled vocabularies may be of especial importance as Scholars Portal 2 will be able to access the full-text of documents, increasing the chances of large search results and potential false drops. However the usefulness of controlled vocabulary and its use in searching is being debated. [Mann, 2007 ([pdf](#))]

### 3.1.4 Tagging, Tag clouds and Folksonomies: Creating Meaning and Forging Connections

#### Recommendation: add user tagging functionality

A December 2006 survey has found that 28% of internet users have tagged or categorized content online such as photos, news stories or blog posts. On a typical day online, 7% of internet users say they tag or categorize online content. [[link](#)]

While tagging sites have flourished on the Internet since 2004 [17], helping users remember and organize email (GMail), photos (Flickr), web sites (del.icio.us), blogs (Technorati), research papers (CiteULike), books (LibraryThing), user tagging in scholarly or library research tools has been slow in coming.

#### Improving Access

Can tags improve access to content? If the observation holds that folksonomies generally contain more than 70% more terms than a formalized taxonomy, can this improve user searching and retrieval? [[link](#)] The museum community has noted in studies that often community taggers provide access to objects in ways which deviate from traditional indexing in the art world, attempting to capture what the art is "about", instead of limiting access to more easily quantifiable information like artist and date. [[link](#)] This hints at the fact that formalized systems of taxonomy may not provide for all the potential terms users may need while interacting with our library documents.

#### Tags Create Relationships

Tagging creates networks of meaning and helps researchers tap into the citation network. The museum community has been quick to take up tagging as a means to provide enhanced access to collections, calling on the community to assist each other. The Cleveland Museum of Art provides a link inviting the community to "Help others find this object". Would a similar invitation work in the research environment? [[link](#)] Can we envision an environment in which a professor "tags" papers of interest for students, encouraging the creation of expert knowledge? Or a system in which students follow tags to see collections of useful articles, as can be done in *Connotea*?

By analysing the data voluntarily supplied by users, *Connotea* is able to offer them a way of finding new and related content. In other words, *Connotea* is not only a place to store your bookmarks but is also a community-driven recommendation system. [[link](#)]

While traditional means of subject access, controlled vocabulary and descriptors, also create networks of meaning and help researchers find content, no system where subjects are predetermined can determine every possible use or relationship someone might have to an object. Tagging allows users to incorporate both personal meaning and assign headings that are either locally appropriate or are in more current terms.

### **The Motivation Issue**

Why would researchers tag within Scholars Portal 2? There are a number of papers which suggest that "Motivation to participate is likely to come from a user's personal connection with the museum or the content" [[link](#)] Therefore if we want to encourage user tagging the action of tagging must be connected to a need, some sort of active engagement with the content. This may mean that a Scholars Portal-based system for user tagging must be connected with a collaborative research space, a social bookmarking utility or a personal account.

A widely-held belief is that in such systems, tags are the primary means by which information producers organize their own bookmarks, and also how information consumers seek and discover new information that has been bookmarked by others on topics that interest them (Marlow et al., 2006; Udell, 2005). Through analyzing semi-structured interviews with twelve regular users of del.icio.us, we discovered that metadata reflecting the identity of the user who saved a web page, which is automatically associated with each bookmark when it is created, was more useful for consumers' information seeking than the user-generated tags. Producers created public bookmarks and tags for their own private reasons. While consumers received benefit from the web pages that were added to del.icio.us, the associated tags did not help their information seeking. In other words, we found that incentives are aligned for bookmarking actions, but not for tagging. [[pdf](#)]

The exception to the rule appears to be LibraryThing which collects more user-generated tags per day than the entire set of tags included in the PennTags library catalogue [Tim Spalding, Computers In Libraries Conference, 2007].

At LibraryThing, people list their books. And, of course, we tag 'em up good. For example, Freakonomics has 993 unique tags (ignoring case differences), and 8,760 total tags. Now, tags are of course useful. But so are subject headings. So, Tim has come up with a clever way of deriving subject headings bottom up. He's introduced "tagmashes," which are (in essence) searches on two or more tags. So, you could ask to see all the books tagged "france" and "wwii." But the fact that you're asking for that particular conjunction of tags indicates that those tags go together, at least in your mind and at least at this moment. Library turns that tagmash into a page with a persistent URL. The page presents a de-duped list of the results, ordered by interestingness, and with other tagmashes suggested, all based on the magic of statistics. Over time, a large, relatively flat set of subject headings may emerge, which, subject to further analysis, could get clumpier and clumpier with meaning. [[link](#)]

### **Tag Clouds**

[Tag Clouds](#) display tags graphically, with the tag appearing larger the more items with that tag. Different graphical representations of content can be experimented with in Scholars Portal 2.

While user tagging cannot replace controlled vocabulary or descriptors, it can enhance usability and findability.

## **3.2 Connecting the citation network to workflow**

The superior XML capabilities of the new Scholars Portal server over the old should allow for much



greater ability for users to navigate the citation network. For example, unlike the current iteration of Scholars Portal, each author name can link to an automatic search of the server for that entity. The new Scholars Portal server will have the ability to provide for each article, lists of "Articles Cited in this Article", "Books Cited in this Article", "Articles Citing this Article", and Similar Articles.

As Scholars Portal is one of the largest repositories of scholarly information in the world, there is also an opportunity to develop modules that can make use of open measures of scholarly work such as the [Hirsch Index](#) or to create alternatives to ISI's impact factor or Google's [PageRank](#).

While the centralization of Scholars Portal makes the navigating the citation network less onerous, significant efforts still must be made to connect the citation network into workflow. For as Lorcan Dempsey has so succinctly put it ([ppt](#)):

**Then:** Users built workflow around the library  
**Now:** The library must build its services around user workflow

### 3.2.1 RSS feeds

**Recommendation: Provide table of contents (TOC) RSS feeds with links that facilitate authentication. If it is possible, allow users to generate their own RSS feeds.**

Many academics are likely to be interested in journal Table of Contents (TOC) alerts, but at present, this tends to involve either multiple registration at publisher's sites for email alerts, or a knowledge of RSS plus considerable effort required to find and select RSS feeds for journal TOCs from various sources - and then either insert them in a feedreader, Bloglines folder, or wishlist [[link](#)].

As the academic libraries of the OCUL consortium have most of their online journals hosted on a centralized server, Scholars Portal is in an enviable position to provide RSS feed services that are not readily available elsewhere. (IngentaConnect, a company that provides hosting, document delivery and other scholarly publishing services, provides five free TOCs per researcher but a purchased license is required for any more). Without their own centralized collection of scholarly material from various publishers, other library jurisdictions have had to seek funding to develop a centralized [TOC RSS service](#).

Currently, many users who make use of TOC alerts from publisher's sites find that they are only able to seamlessly access the articles of interest while on campus. When they are off-campus access they find that they are being denied access to articles by the publisher who does not recognize their institutional affiliation. Access to library-paid scholarly resources require being authenticated by their library first and for those libraries that make use of the EZ-proxy authentication service, users must use links that include a special prefix in order for the authentication process to be initiated. For these users, they must either access the articles once they are back on campus or go to the library's website to 're-find' these articles (or for those with a LibX Toolbar -- see 3.3.1 -- installed, they can simply re-load the page and LibX will add the ez-proxy prefix). To spare users this inconvenience, a centralized RSS service from Scholars Portal should incorporate a means to generate RSS feeds that allow for easier authentication, either by making use of the COiNS standard (see 3.3.1) or other method.

If it is possible, it would be ideal to allow users to generate their own RSS feeds based on search results for subject headings, author name, affiliations, and keyword matching. This would allow users to generate feeds that could be used in keeping up to date in a particular field, to add dynamic content to departmental web sites ("20 most recent publications by members of the U of T Chemistry Department"), and to create dynamic bibliographies. Once the hurdles of generating RSS feeds that facilitate authentication are met, it would not be difficult to then allow for the creation of Scholars Portal [widgets](#).

## 3.2.2 Social Bookmarking

### Recommendation: To provide for users of scholarly resource social bookmarking services

Collecting material for a bibliography is something which appeared to require an amazing amount of drudgery. All the existing options seemed to require more effort than strictly necessary to transfer the citation details for the article currently open in my web browser into some sort of permanent storage. I'm sure with a lot of practice I could have got the process down to twenty seconds or so, but that twenty seconds just presented enough unpleasantness of flipping between browsers and external applications, copying and pasting details, and opening downloaded "citation export" files that I was far less likely to actually do it. I'd need amazing amounts of self discipline to consistently bookmark everything I ever read on the off-chance that I might want it again. Unless, of course, it just involved clicking a button on the browser and having it all magically happen.

So, the obvious idea was that if I use a web browser to read articles, the most convenient way of storing them is by using a web browser too. This becomes even more interesting when you consider the process of jointly authoring a paper. There is a point where all the authors need to get together and get all the articles they wish to cite into the one place. If you do this process collaboratively on a web site, then it's easier.

The next obvious leap is that if all the references are available via a web interface on a central server, it would be really nice to see what your colleagues are reading and be able to show them what you're reading. It cuts down on the number of emails saying "have you seen this article?"

In fact, if enough users register on the system, you'll probably find people reading the same articles as you. That provides a great way of keeping on top of the literature - you simply share it with people who have common interests. [[CiteULike](#)]

Earlier in section 2.7 of this document it was suggested the abundance of the digital resources available to every user will necessitate new tools to manage these new resources. One such tool is the social bookmarking system.

Using [social bookmarking tools](#), users store links to web pages that they find useful in a communal online database rather than in their private web browser. While icons and menu options to easily 'share', 'save and tag', or 'add to' such bookmarking services are now commonplace on the Internet, including the websites of the major newspapers, these service links cannot be found in the scholarly resources provided by academic libraries. This is likely because of the generally unavailability of permanent links in library resources (see 2.4) and authentication issues (see 3.2.1), making it is very cumbersome to add library resources into social bookmarking services. Perhaps a module of the new Scholars Portal server could provide a permalink for each item or "dynamic service links" that would allow for the seamless import of the permalink to a select number of social bookmarking of services.

There are two social bookmarking tools of note that are geared towards capturing and sharing scholarly information: [Connotea](#) (developed by Nature Publishing) and [CiteULike](#) (developed by Richard Cameron). Unlike the more popular social bookmarking services such as [del.cio.us](#), these tools were designed to try to capture the unique and important metadata of the academic paper and to make use of services such as DOI (Digital Object Identifiers) and PMIDs (Pubmed Identifiers).

Alternatively, it would not be difficult for Scholars Portal to develop its own scholarly bookmarking service as the [source code of Connotea](#) has been released as open-source software as has Dan Chudnov's [unalog](#) program. There would be a number of benefits to doing so. First, it would provide a means for the faculty and students of OCUL to be informed of each other's research. It has been found that users of del.cio.us frequently browse others' bookmarking activities for a number of reasons: to support research on a particular topic, to follow a particular person's research interests and just for curiosity [[pdf](#)]. As well, as social bookmarking service based on the resources of Scholars Portal would

reduce the number of redundant links to scholarly research [that can be found currently in services like Connotea](#).

Another advantage of hosting a scholarly social bookmarking service is that the information collected could provide useful information for other functions. For example, papers that were frequently bookmarked could be given added 'weight' in the search algorithm of Scholars Portal. Libraries do a poor job of capturing the wealth of domain expertise that exists on our campuses. Imagine what a wealth of knowledge a collection of faculty-suggested course readings could be!

It would be essential that a scholarly bookmarking service used by OCUL would be integrated with RefWorks and other citation management software. Social bookmarking services are better suited to the capturing citations during the reading and research workflow of scholars than RefWorks, even with its [RefGrabIt](#) capabilities.

### 3.2.3 Collaborative research spaces

**Recommendation: Consider services that support the whole of the user's research process and the development of online space for OCUL research communities.**

More than 70 percent of faculty said they maintain their own collections, although fewer of them (~35 percent) make their resources available to others on the Web. It was clear from our discussions and from comments on the surveys that many faculty want the ability to build their own collections, which are often composed of a variety of materials, including those that are copyright protected. How to manage this potpourri of resources and integrate them into teaching practice is the challenge. Although there may be an array of tools available to faculty for collecting, developing, and managing resources, the efficacy and interoperability of these tools for the immediate tasks that faculty need supported often fall short. [\[link\]](#)

[Through a Mellon Grant, the University of Minnesota developed a model for assessing support for scholarship and research on a large research campus.](#) The framework focused on three broad components: information resources, infrastructure services, and research behaviours and was built upon the results of a survey of faculty members and graduate students in sixteen departments in the humanities and social sciences. Of note, the survey found that:

- 37% - Have unique personal research collections
- 56% - Feel less than effective at disseminating research
- 87% - Draw on literature from multiple fields
- 49% - Feel less than effective at keeping up in their field
- 73% - Would use assistance in organizing/storing materials
- 39% - Have less than adequate methods for organizing materials

Each information need was placed in a model of four components: *Discover. Gather. Share. Create* and individual tool components were mapped onto meeting each need. Using this model as a foundation, the project then developed a prototype that demonstrates how the individual tool components could work together to form an online virtual research community [\[ppt\]](#).

Already researchers without formal spaces are beginning to share research and collaborate with others by re-purposing the other online tools that are on-hand:

Not long ago I observed — as did another wise librarian colleague in a previous life — that for all the work some libraries were doing with IRs, the faculty seemed aware of, and preferred to use.... well, Blackboard. I participated in a Blackboard focus group a few months back and was astonished to hear faculty talk about the joys of using it for sharing preprints

and other documents with their colleagues. It was easy to use. It was “in the flow” of their other activities. At least on that campus, they could share across and within disciplines. [[link](#)]

More dramatically, two researchers from the Netherlands has created a Mac OS X application called [Papers](#) which makes use of an iTunes-like application to store and share pdf files. Coincidentally, Dan Chudnov [in his keynote address to NASIG 2007](#) has outlined a potential roadmap how libraries could provide access to materials through a service like iTunes.

### 3.2.4 Working within Learning Management Systems (LMS)

**Recommendation: Seek means for Scholars Portal to be integrated into Learning Management Systems (otherwise known as Course Management Systems and Virtual Learning Environments) used by OCUL**

Because of the lack of persistent links that can easily authenticate (see 3.2.1) it is very difficult for users to add links to scholarly material into a LMS that can be accessed from off-campus. This is just one of the architectural challenges faced in the integration of library services into LMS and campus portals [[pdf](#)].

Efforts must be made to ensure that library-licensed scholarly material can be easily added to a reading list within an LMS as the LMS is gradually becoming where our students “are”. It is where they learn about their course readings, where they submit their assignments, and it is where they are encouraged to communicate with each other about what they are learning and researching. Providing appropriate library resources at the course-level may also prove an additional benefit to undergraduate students. After repeatedly observations in their library-web usability testing that undergraduates lack an understanding of an academic discipline and that this was hindering their ability to find appropriate library resources or library guides, the librarians at the University of Rochester are pursuing organizing library resources at the course level<sup>7</sup>.

Of note, the [IMS Global Learning Consortium](#), a worldwide non-profit organization representing the sectors of LMS sector, has already developed a standard called [IMS Resource List Interoperability Information Model for citations and citation lists](#).

## 3.3 Embracing standards

Libraries should be really good at technical standards for our data, the structures of our catalogues and digital projects are based on internationally recognized standards. We can share our bibliographic records very easily. However we need to be better at ensuring our data is open, usable and interoperable.

Sharing metadata and the resultant aggregations benefit users, particularly those users whose subject interest cuts across disciplinary boundaries. Not only do these aggregations minimize the time and effort expended on searching for all the resources on a particular topic, but they can yield higher quality resources in a variety of formats than would typically be found through an Internet search engine’s crawl of the Web. [[link](#)]

### 3.3.1 Permanent URLs and COInS

**Recommendation: to provide both permalinks as well as COInS OpenURLS in the Scholars Portal server and to encourage OCUL libraries to adopt their own versions of LibX or promote**

## other COinS readers

Many of the recommendations in this white paper that are related to connecting the citation network to user workflow require the availability of a permanent URL or an [OpenURL](#) link in the COinS format. COinS. "COinS is an acronym that stands for *Context Objects in Spans*, which represent a standardized way to embed citation metadata into a web page. COinS are actually included in the HTML code on the web page using OpenURLs. [This allows other processors—such as your web browser—to find the citation metadata and generate links to other resources that are accessible via OpenURLs.](#)"

COinS links can be found in [Open WorldCat](#) pages as well in [Wikipedia](#) – but the nature of COinS is that the OpenURL links are not immediately visible. In order to make COinS links visible, there must be functionality to read them in the browser that must be customized to reflect the user's library privileges. This may be done through a [bookmarklet](#) or [Greasemonkey script](#), [browser extension](#) or toolbar such as LibX. The [LibX Toolbar](#) is a multipurpose extension for the Firefox browser that is currently available for two OCUL institutions: The [University of Waterloo](#) and the [University of Windsor](#). One difference between how the two have been customized is that the University of Windsor LibX Toolbar allows users to send highlighted terms in the browser as a search to OCUL's Scholars Portal Search service.

### 3.3.2 Shibboleth

**Recommendation: Investigate how to take advantage of the attribute-based information that Shibboleth can provide**

A number of OCUL libraries are set up to take advantage of Shibboleth. "[Shibboleth](#) is standards-based, open source middleware software which provides Web Single SignOn (SSO) across or within organizational boundaries. It allows sites to make informed authorization decisions for individual access of protected online resources in a privacy-preserving manner." Shibboleth allows, for example, for users to sign on to an LMS system to then pass into library-licensed resources without having to enter another username and password to re-tell the system who they are.

As this software allows for some attribute-based authentication to be passed to Scholars Portal, there may be a potential to tailor Scholars Portal 2's interface or search results based on readership level.

### 3.33 The Semantic Scholars Portal

**Recommendation: Consider what semantic metadata could be provided through Scholars Portal**

We believe that this easy-to-use research environment and the way in which it takes advantage of semantic bibliographic information on the Web instantly makes Zotero an attractive choice for managing one's research. Hundreds of sites already seamlessly integrate with Zotero, with more being added each week. Some research sites, like Copac, the union catalog of major libraries in the United Kingdom and Ireland, have automatically become compatible when they have added semantic formats to their site. [\[link\]](#)

Providing structured data that can be recognized by Semantic Web technologies allows for the possibility for the data to be repurposed and remixed by our users. Examples that illustrate potentially how scholarly data could be 'remixed' can be found at MIT's [Simile Projects](#) which have been developed to "improve access, management and reuse among digital assets". One such project is [Exhibit](#) which allows users to create web pages that allow for sorting and filtering and that appear to be generated by a database but are actually driven by Javascript that takes advantage of semantic metadata. Exhibit-created websites can then be further be mashed or re-mixed with other structured data using Simile's [Potluck](#).

As Google Maps has shown us, if you allow users to mash up your data they will create uses that are as individual as they are. Educators have already begun to explore the potential of "mashed-up" scholarly resources. Tony Hirst of The Open University [has documented](#) his ongoing experiments with [Yahoo!Pipes](#):

Among Hirst's products is a "[deliSearch Pipe](#)," which allows users to run a constrained Internet search over domains that have been flagged by a user of the del.icio.us social bookmarking service. Another Pipe that Hirst has created is the "[OpenLearn Unit Outlinks Search Hub Pipe](#)," which extracts "all the outgoing links from a course unit, then feeds these into a Yahoo Search pipe, which uses the domains as search limits for the search." In other words, this Pipe can create a filtered search of trusted domains that are relevant to a particular course, and the filtered search will adjust automatically as new links are added to the course materials. Of course, this added functionality requires open content and a reusable data format in order to work properly. [[link](#)]

There has been some early work on creating a [Bibliographic Ontology](#) that seeks to re-use existing semantic web ontologies including

- [FRBR](#): as the basement of the ontology
- [FOAF](#): as the way to describe authors
- [SIOC](#): as a way to describe everything related to the social software world
- [MO](#): as a way to describe everything related to musical things
- [DC](#): Dublin Core
- [Event](#): as a way to describe some events like workshops, conferences, etc.
- [Timeline](#): as a way to describe complex temporal frameworks

In the meantime, it may be beneficial to investigate the use of [microformats](#) in Scholars Portal to create semantic structures that can be used to express relationships and information about people, licenses, and social networks.

## 4.0 Conclusion

The goal of this white paper is to generate discussion about which potential services we would like to see the new Scholars Portal Server provide. It is planned that a demo of the new server will be made available to OCUL librarians in October of 2007 and a beta version be developed for the Spring of 2008.

The comments generated by this white paper are one means by which development of the new server will be guided. In the next months it is hoped that input on this matter will be also be collected from OCUL students and researchers, by means of either focus groups, surveys, or user panel.

The success of this work is dependent on you. Please let us know what you think.



## **5.0 Histories / Bios**

### **OCUL**

OCUL is a consortium of twenty university libraries in the province of Ontario. The member libraries cooperate to enhance information services through resource sharing, collective purchasing, document delivery and many other similar activities.

OCUL's vision is to be a recognized leader in provincial, national and international post-secondary communities for the collaborative development and delivery of outstanding and innovative library services that are critical to the success of Ontario's universities. Established in 1967, OCUL's twenty member institutions continue to enhance their libraries through activities such as resource sharing, collective purchasing, and, joint creation of the digital library Scholars Portal.

### **Scholars Portal**

Scholars Portal was established with four primary objectives:

1. To provide for the long term, secure archiving of resources to ensure continued availability
2. To ensure rapid and reliable response time for information services and resources
3. To provide an environment that fosters additional innovation in response to the needs of the users
4. To create a network of intellectual resources by linking ideas, materials, documents and resources

Currently Scholars Portal serves 360,000 researchers and students that are a part of OCUL member institutions. There are currently 12 million full text articles in Scholars Portal and 150 million article abstracts in Scholars Portal Search.

### **Scholars Portal Public Services Advisory Group**

The Scholars Portal Public Services Advisory Group relates to all the user services of Scholars Portal (for example Scholars Portal Search, SFX, RefWorks, MultiSearch, RACER, Verde). Reporting to the Scholars Portal Operations & Development Committee (SPODC), the Scholars Portal Public Services Advisory Group will:

1. Solicit feedback and ideas about Scholars Portal from public services and systems staff at OCUL institutions, and facilitate user-driven evaluation of services.
2. Alert SPODC and Scholars Portal staff to user issues on an ongoing basis, and provide advice and feedback on strategic directions.
3. Liaise with OCUL-IR and Scholars Portal staff on matters relating to Scholars Portal content and database configuration.
4. Track developments in the presentation of Scholars Portal at OCUL institutions and its promotion and instruction, and facilitate the sharing of best practices.
5. Contribute to the development of content for a web portal for staff at OCUL institutions, to enhance communications regarding all aspects of Scholars Portal.
6. With SPODC and Scholars Portal staff, develop the programs for Scholars Portal meetings of staff from OCUL institutions.
7. Report regularly to SPODC.

## **Membership**

Six public services librarians representing a range of disciplines and OCUL institution sizes, serving two year terms. One Scholars Portal staff member, ex officio. A member of SPODC chairs the group.

Members of the group for 2007-2008:

- Stacy Allison-Cassin, York
- Jim Brett, Guelph
- Valerie Critchley, Ottawa
- Kate Davis, SPOT
- Heather Matheson, Carleton
- Dana Thomas, Ryerson
- Martha Whitehead, Queen's, (Chair)
- Mita Williams, Windsor

## **Scholars Portage**

Scholar's Portage is a sub-committee of the Scholars Portal Public Services Advisory Group. It does collaborative work between OCUL librarians in pursuit of incorporating more 'social' means of sharing and organizing information within OCUL's Scholars Portal and the larger academic community.

**Current members:**

### **Stacy Allison-Cassin, York University**

Stacy holds Masters degrees in Music Performance (Duquesne University, 2002) and Information Studies (University of Toronto, 2004) and is beginning a PhD in Humanities at York U. in the fall of 2007. Stacy is the music cataloguer at York University, a position she has held since 2005. Her first professional position was as National Librarian at the Canadian Music Centre, a non-profit library.

### **Mita Williams, University of Windsor**

After graduating from McMaster University with an Honours B.Sc in Geography and Environmental Science and acquiring a M.L.I.S. from McGill University, Mita was employed in a number of contract positions in public, corporate, non-profit, government, and academic libraries before joining the Leddy Library, University of Windsor as Science Librarian in 1999.



## References

1. Coyle, K. (2007). The Library Catalog: Some Possible Futures, *The Journal of Academic Librarianship*, 33(3), 415
2. Letter to Hooke, 5 February 1675/6. In H. W. Turnbull (ed.), *The Correspondence of Isaac Newton*, 1, 1661–1675 (1959), Vol. 1, pp.416
3. Tapscott D, Williams AD. *Wikinomics : How mass collaboration changes everything*. New York, N.Y.: Portfolio; 2006. 324 p.
4. Markey, K. (2007). Twenty-five years of end-user searching, part 1: Research findings. *Journal of the American Society for Information Science and Technology*, 58(8), 1071-1081.
5. Markey, K. (2007). Twenty-five years of end-user searching, part 2: Future Research Directions. *Journal of the American Society for Information Science and Technology*, 58(8), 1123-1130.
6. Fister, B. (1992). The research processes of undergraduate students. *Journal of Academic Librarianship*, 18, 163.
7. Reeb, B., & Gibbons, Susan (Susan L.) 1970-. (2004). Students, librarians, and subject guides: Improving a poor rate of return. *portal: Libraries and the Academy*, 4(1), 123-130.