Using xSearch for Accelerating Research—
Review of Deep Web Technologies
Federated Search Service

Abstract
Developed by Deep Web Technologies (DWT) (<http://www.deep-Webtech.com/>) in consultation with the Stanford University Libraries and Academic Information Resources, xSearch (<http://xsearch.stanford.edu>) is a locally named and customized version of DWT’s federated search service that can search 170+ resources at one time. (Explorit is the commercial version of Deep Web Technology’s federated search solution which Stanford has locally branded as xSearch.) In addition to the broad range of resources searched, xSearch provides an intuitive interface for searching, refining, and displaying results. The ability to choose the resources searched allows this tool to be used by researchers as well as undergraduates. Speed is enhanced by the fact that some results are displayed before the search is completed. Users are able to easily set up alerts from their search results. Both library staff and users are able to create custom search pages that can be a link on a Web page or embedded as a search box in a research or course guide. xSearch has provided a scalable, cost-effective, and robust solution for meeting a wide variety of information needs.

Pricing Options
Customized pricing is provided for each agreement. DWT’s pricing structure for Stanford contains: (1) Fixed and initial startup costs which include a software license for the Explorit Federated Search application, custom development, and support and maintenance fees for one Standalone Implementation, as well as initial connector development for resources. (2) Annual recurring fees to maintain connectors and for Explorit maintenance and support. (3) Capability to add new connectors, which will incur an additional one-time fee for setup and development, and a lower annual recurring fee to maintain the connector. We found DWT’s pricing to be competitive and, in some cases, significantly lower than similar products available on the market. All of the features developed for Stanford are available to other DWT customers. DWT also offers additional features not currently included in xSearch (e.g., multilingual searching, multimedia and image searching).

Explorit, DWT’s federated search engine, implemented at a mid-sized academic university with a population of up to 25,000 students and federated search services that contain 100 resources on a server hosted at DWT would cost about $20,000 a year. This price includes alerts services and the ability to build custom search engines.

Product Description
DWT has developed several federated searching sites that may be familiar to users. Sites include: Science.gov (<http://www.science.gov>), WorldWideScience.org (<http://worldwidescience.org>), SciTopia (<http://www.scitopia.org>), ScienceResearch.com (<http://www.scienceresearch.com>), Mednar (<http://mednar.com/mednar/>), and Biznar (<http://biznar.com/biznar>). xSearch is the name we decided to use for DWT’s federated search services that were customized for Stanford.

CONTENT
Because the intended audience for xSearch (<http://xsearch.stanford.edu>) includes researchers as well as undergraduates, the ability for subject specialists to select the resources to be searched via xSearch is a very important feature. It is worth noting that because content is queried in real time, not pre-indexed, the results include the latest information available from a resource.

Resources in xSearch provide depth and breadth for a wide array of subject areas and document types. In addition to abstracting and indexing services covering articles, conference proceedings, dissertations, patents, and technical reports, it was a high priority to include sites that have full-text searching of ebooks and ejournals in xSearch. Full-text search sites for newspapers and primary historical resources are also included. Other resources selected for xSearch include archival finding aids, government publications, and reference materials. Statistical resources searchable via bibliographic data are also included in xSearch. A few foreign language resources are included, but there are no special features implemented to assist in searching them.
A few of the resources that we wanted to include in xSearch would not work. There were several reasons: design of the search interface (e.g. must browse to select search terms), poor response time for a resource, and lack of persistent URLs for retrieved items. The native interface is being used for many queries but a few resources require using a Web Services API. As a result, some resources are accessible via xSearch that are not available from other federated search services.

**SEARCH INTERFACE**

End-user access to xSearch (<http://xsearch.stanford.edu>) requires only a Web browser and an internet connection. When working with DWT to customize the interface, great care was taken to make navigation and access ADA-compliant. A mobile version does not exist although many features work on an iPad. So that help information could be modified easily, a “companion” Web page was set up locally (<http://lib.stanford.edu/xsearch>) that contains a list of resources, a quick reference guide and tutorial, plus links to custom search engines created for subjects or courses.

The Quick Search page allows users to search Keywords in a fixed subset of 50 popular resources (see FIGURE 1). The Advanced Search page (see FIGURE 2) allows users to enter terms in Keywords, Title, and Author fields along with date ranges for all 170+ resources. Users may select resources they want to search by choosing individual titles or by choosing categories. Brief descriptions of each resource are displayed when a user hovers their computer mouse over the resource name. A separate search page for searching selected business resources by Company Name is in development.

All search pages support Boolean logic (AND, OR, NOT) plus nesting if users want to combine terms using AND and OR in the same query. Two trun-

---

**FIGURE 1. Quick Search**

**FIGURE 2. Advanced Search**
users refine results, clusters consisting of Topics, Authors, Sources, and Dates are generated dynamically. Terms in the Topics cluster are semantically generated. Results can be sorted by Rank, Date, Title, and Author. Users can also filter or limit the display of search results to a particular resource (see Figure 3). Session Preferences enable users to increase the number of items displayed per page from 15 up to 250.

Once search results are available, there are several features to help searchers identify and export records of interest. Stars are present to indicate relevancy ranking. Several sentences from an abstract or snippet showing hit term highlighting are also included. SFX links or full-text icons are present in the brief record display format. Full record displays for an item take a user into the native interface of a resource. Users are able to email or print entire answer sets from the main menu. If only some items are of interest in search results, users can check a box next to an item and then go to My Selections where they are able to directly export to RefWorks; export to EndNote Web; save in RIS format for importing into other reference manager software; print the selected citations; or email them.

A Results Summary link provides the total number of items found in a particular resource and the total number of items included in xSearch results (see Figure 4). A red X appears by the name of the resource if the connector failed to do a search and a stopwatch appears by the name of a resource if a search timed out due to slow response. This information allows users to verify that all the resources that they choose are included in the search results. The Results Summary also helps users identify specific resources they might want to search individually if they are doing in-depth research.

**ALERTS AND CUSTOM SEARCH PAGES**

Setting up alerts in xSearch is easy and intuitive. The search results page includes a link to “create an alert from this search.” Users can easily modify what resources are searched in their alert, select the frequency of alerts (weekly, monthly, quarterly, annually) and their format (email, RSS feed, Atom feed). Alert details (search strategy, resources searched, etc.) as well as results from several updates can be viewed in the xSearch interface by clicking on the My Alerts link.

Both library staff and users can Create a Custom Search (see Figure 5). This feature allows users to create a search page that contains only their favorite resources. This service also enables librarians to instantly create custom searches that are tailored to a particular

**FIGURE 3. Search Results**

![Search Results](image1.png)

**FIGURE 4. Results Summary**

![Results Summary](image2.png)

**FIGURE 5. Custom Search**

![Custom Search](image3.png)
class or research guide. Saved custom searches are accessible through the xSearch interface, as a link that can be added to a Web page, or as an embedded search box via widget code.

TECHNICAL INFRASTRUCTURE AND TECHNICAL SUPPORT

Stanford has worked with DWT for four years. Initially, DWT created three federated search prototypes for us: (1) a collection of library catalogs, (2) a collection of locally digitized materials, and (3) a collection of licensed resources that included databases and full-text ebook and ejournal sites. With shelving space limited on campus, before items were shipped off-site faculty wanted to have a more robust search interface for discovering library materials than was possible via Sirsi or DWT. Blacklight (<http://projectblacklight.org>), an open source discovery application, was selected and local development work followed. The new interface to the library catalog is called SearchWorks (<http://searchworks.stanford.edu>). Because SearchWorks also includes bibliographic access to locally digitized materials and because of the considerable investment the library makes in licensing resources that are not available via the open Web, we decided to focus our efforts with DWT on licensed resources. Our efforts included working with DWT in order to get a summary of search results, alerting services, a custom search engine, and support for open URL linking. DWT also customized the search interface for us in order to harmonize the user interface with other library Web applications. While there are no technical barriers to prevent the library catalog from being merged into xSearch, because development efforts have been underway for both interfaces, we have not yet integrated SearchWorks and xSearch.

DWT offers both hosted and locally installed versions of their federated search software; Stanford opted for local installation on its server farm. The initial production instance of xSearch was launched two years ago and included 28 resources running on a RedHat virtual machine. A year ago, when we decided to significantly expand the number of resources, to insures adequate bandwidth for production use and to provide a test environment for future developments, Explorit Search Engine and accompanying software were installed on dedicated dual CPU server with 4 GB of RAM.

Staff support from both DWT and Stanford is needed for xSearch. In addition to technical support provided by DWT staff for custom development, installation, and maintenance, technical support was also needed by Stanford for implementation (installation and configuration of the application, integration with University authentication systems, specification of open URL linking) and for maintenance (ongoing server management, coordination with DWT for upgrades and changes to connectors). Looking ahead, the goal is to have someone at Stanford assigned to xSearch as the Application Administrator. After an initial period of
training and knowledge transfer, it is estimated that this work will consume less than 10% FTE on an ongoing basis for routine monitoring and maintenance. The Application Administrator would be expected to increase their time commitment during upgrades.

**Usage Statistics**

Usage statistics far surpass what is in a typical COUNTER Report. Statistics that are provided include information such as the number of queries, response time for each resource, percent results that are relevant, and the number of click-throughs to full-text. Details for terms searched and particular articles viewed are also available. Daily, weekly, and monthly use statistics are available graphically and can be exported into Excel. Statistics to provide details about alerts were developed at Stanford and include total number of alerts and email addresses for those receiving alerts.

**Critical Evaluation**

Compared to other federated search products, Stanford found that DWT offered the most compelling package of performance, features, and design. While federated search engines’ performance is inherently limited by the performance of its target sites, DWT’s progressive delivery of results gives researchers near real-time response with the first set of responses while the application assembles a complete set of hits from all sources. When we chose DWT, other federated search services we considered (e.g., MetaLib) were much slower because they did not return any results until a search was completed. The recommended number of resources to search at one time through other services was much smaller too. DWT offered many options for customizing the interface, the only service that included alerts, and the only service that allowed us to create customized “search engines” locally.

DWT calculates relevance across search results on the fly, and clusters hits based on predefined criteria, giving searchers the ability to filter results to items from a particular resource is great. The ability to create custom search pages and set up alerts are features not offered by other federated search providers.

**Pricing:**

Deep Web Technologies federated search services are competitive with and in some cases considerably less expensive than competitor products.

**Contract Options:**

Contract language and options were reasonable.

---

**Contact Information**

Deep Web Technologies  
301 N Guadalupe Street, Suite 201  
Santa Fe, NM 87501  
Phone: 505.820.0301  
866.388.1407  
E-mail: <sales@deepWebtech.com>  
URL: http://www.deepWebtech.com
ery points than if they were searching Google Scholar or a publisher’s site. The search interface and facets for refining results are clear and easy to use. These facets make it much easier to refine results than in Google Scholar. The fact that xSearch also searches resources licensed by Stanford, and has good openURL integration, helps ensure that Stanford researchers have full access to the results they discover, and via a more direct path than can sometimes be the case in Google Scholar. Search results are returned while a search is still running so the lag time between starting a search and being able to begin reviewing results is very short. The number of items retrieved from a resource may vary, depending on the response time of that resource when the search was performed in xSearch. For items in xSearch results, the ability to filter the display of results to a particular resource is very helpful.

Relevancy ranking of the search engine is robust but results brought into xSearch from each resource are dependent on default sorting order for that resource. If the default sorting order for an individual resource is by date rather than by relevancy ranking, items retrieved and merged in xSearch results will not be the most relevant items from that individual resource. Search results include both ranked and unranked items with unranked items having only some of the search terms present. This means if results are sorted By Date then less relevant items may be displayed first.

The Brief Display includes the title, author, abstract, publication title, publication year, and the resource that an item was retrieved from, as well as a link to the full-text. Ideally, entries for articles would also include volumes and pages but this is a minor annoyance as the full citation is available either when the full record is displayed or when a citation is exported to a reference manager.

Contract Provisions

DWT is a commercial service that offers federated search solutions to a wide array of customers. Each contract is customized and is based on such factors as the type of organization, the number of resources included in the service, how the service is deployed (hosted vs. stand-alone), the amount of customization desired, support services needed, etc.

Because nearly all resources included in xSearch are currently licensed by Stanford, access to xSearch is limited to students, faculty, and staff with a current Stanford University Network ID. The DWT contract for xSearch defines “deliverables” that include features, functionality, and services. One key component of this license agreement is defining turn-around time for responding to service requests, for example, a prompt response if a connector is not returning results but batching requests for cosmetic changes to the interface. Statistics provided are not COUNTER Compliant but do offer a much richer array of data and reports that include data such as response time for each resource, number of queries, and number of full-text click-throughs.

Authentication

Anyone can view the search interface but only current Stanford users can run a search. Because most resources in xSearch are licensed by Stanford, IP filtering and EZProxy authentication using a Stanford University Network ID (SUNet ID) are invoked when a user presses the Search button, even if a user is located on-campus. SUNet ID entries are checked against the Stanford Directory to insure that users are current students, faculty, or staff (also called LDAP). Creating a custom search engine or setting up an alert also requires SUNet authentication.

Before custom search engines were developed, we had users authenticate when they entered the URL or clicked on the link to use xSearch. After custom engines were developed, because these custom engines can be embedded as a search box, the authentication prompt was moved so that it appears when a user presses the search button.

Author’s References

DOI URL: <http://dx.doi.org/10.5260/chara.12.4.11>
DOI URL: http://dx.doi.org/10.5260/chara.12.1.5

About the Authors

Grace Baysinger is the Head Librarian and Bibliographer of the Swain Chemistry and Chemical Engineering Library at Stanford University. She was the project lead for implementing xSearch at Stanford. Grace’s professional interests include end-user searching, collection development and management, and scholarly communication issues. She serves on several advisory boards and committees for database producers and publishers in chemistry. After receiving her library degree in 1981 from the University of Michigan, Grace held professional positions at UM’s Natural Sciences and Chemistry Libraries. In 1989, she accepted the position as Head of the Swain Library. She also served as the Head of Stanford’s Science and Engineering Libraries from 1999-2005.

Tom Cramer is the Chief Technology Strategist and Associate Director of Digital Library Systems and Services for the Stanford University Libraries. In this role, he oversees the technical development and delivery of the full complement of Stanford’s digital library activities, including the digitization, description, discovery, delivery, preservation and management of digital resources that support teaching, learning and research. Tom is one of the founders of the Hydra Project <http://projecthydra.org/> and an active contributor to Blacklight <http://projectblacklight.org/>, open source projects rooted in higher education that deal with digital asset management and discovery. He currently serves as the director of PASIG, the Preservation and Archiving Special Interest Group, and as a steering group member for Open Repositories.
Heard on the Net

Fear & Self-Loathing in Libraryland

doi:10.5260/chara.13.2.61

By Jill Emery (Collections Librarian, Portland State University)

The library pundits are saying be afraid, be very afraid for your job, the future of librarianship is in peril, and hope is futile and librarians are not doing enough to maintain our professional relevance in the 21st century. These pundits speak loudly and carry catch phrases that get picked up by all our news channels and repeated over and over again. It started with the Taiga Forum in 2009 and has picked up momentum over the past three years with the latest provocative statements reeling like death knells: [http://taigaforumprovocativestatements.blogspot.com/].1 The echo of these pronouncements was picked up by ARL with their envisioning research library futures: [http://www.arl.org/rtl/plan/scenarios/index.shtml]2 and most recently the echo has been rebounding from the Education Advisory Board in their webinar entitled “Redefining the Academic Library” that is currently being presented to university provosts and administrators.3 Amidst the various concerns to be grappled with in regards to the librarian punditocracy stuck in this refrain, is that the majority of them are rhetoricians and not practitioners.

In a recent ReadWriteWeb post there was a report regarding the future Internet by UK’s national innovation agency, Technology Strategy Board. “The report defines the future Internet as ‘an evolving convergent Internet of things and services that is available anywhere, any time as part of an all-pervasive omnipresent socio-economic fabric, made up of converged services, shared data and an advanced wireless and fixed infrastructure linking people and machines to provide advanced services to business and citizens.”” [http://www.readwrite-web.com/archives/the_future_of_the_internet_is_converged_services.php]4

In Figure 1 illustrating the Internet development, the mix becomes machines, people and services and is labeled the Internet of People and Things. What a great space in which to be an information professional! The opportunity within this space for librarians and the library are numerous and here’s a small sampling of how librarians are already beginning to converge with their communities.

The Orbis Cascade Alliance in the Pacific Northwest is in many ways at the forefront of convergence with various initiatives currently underway.5 In addition to the member libraries participating in the Western Storage Trust (WEST) with the University of California and other Western U.S. institutions to create a shared print repository, a pilot project for shared demand driven e-book acquisitions is underway and so far looking like a successful way to partner with other academic institutions to leverage the buying power for patron requested material. The Alliance is also proceeding with a RFP for a shared ILS and has begun to share cataloging expertise of non-Indo-European language material starting with CJK languages (Chinese, Japanese, Korean). What the Alliance exemplifies is leveraged expertise can be capitalized upon to reward multiple institutions. [http://www.orbis-cascade.org/index/strategic-agenda>]

In 2010, the University of Calgary created a new campus unit entitled: Libraries and Cultural Resources bringing together the University Library, Archives, Special Collections, The Nickle Arts Museum, and the University of Calgary Press under one administrative unit.6 The purpose of this organizational change is to “provides a key strategic advantage to researchers and makes information - regardless of format - readily available.” [http://lcr.ucalgary.ca/] Instead of each individual component having to compete for budgetary resources, now there is a single management structure for supporting scholarship and learning at the University of Calgary.

HathiTrust is a prime example of the convergence of librarians to leverage not only resources but expertise to provide shared collections and a digital framework for ongoing preservation of information resources.7 Their announcement last year regarding the tracking down of copyright permissions for orphan works speaks to the core values of our profession. [http://www.libraryjournal.com/fj/home/887388-264/hathitrusts_copyright_detectives.html.csp]8 However, the Authors Guild begs to differ and has now brought a copyright infringement lawsuit against HathiTrust for the “unauthorized use” of copyrighted works and even questions the security of the digital files.9 [http://chronicle.com/blogs/ticker/authors-guild-sues-hathitrust-5-universities-over-digitized-books/36178#top> This is definitely a development to watch to see which is the greater crime, the proposed breach of copyright or the “intolerable digital risk.”

LOCKSS/CLOCKSS is yet another example of how librarians have converged with scholarly publishers in order to provide collective management of electronic resources to maintain digital access to resources.10 [http://www.lockss.org/lockss/Home>]

Last but not least, the one rhetorical theme that has been overplayed and potentially does some of the greatest harm to our profession is the continued call for the reduction of professionals dedicated to cataloging library resources. Calling for any one arm of a profession to be diminished is a threat to the profession as a whole; noses and faces come to mind as a metaphor here. This theme is usually paired with the supposed “capability” to outsource the work these professionals perform. It is time the pundits learned that the reason why OCLC provides the largest catalog of resources is by and large due to the fact that our catalogers contributed the majority of the records found there. It can be argued that WorldCat is the single greatest achievement of convergence by librarians. WorldCat did not spring fully populated from the heads of the OCLC board of directors. Once this segment of our profession is minimized then the contributions for quality metadata also diminishes to all the initiatives listed above. OCLC does not hire enough personnel to perform the cataloging for every major academic research library in the United States of America. Furthermore, in their own words, OCLC laid off personnel in 2010 partly in response to the budgetary declines at the institutions utilizing OCLC.

[http://www.tmcmont.com/submit/2010/06/02/4822558.htm>]

The argument that vendor supplied records are “good enough” has been proved time and time again to be untrue. In fact, this is one area
where it has been shown that Google™ does not trump libraries. The access problems that have been outlined with Google™Books have shown this first hand [<http://chronicle.com/article/Googles-Book-Search-A/48245/>] and even more recently, Rick Anderson pointed out that a part of the implementation problems with the Espresso Book Machine at the University of Utah was due to bad metadata: “The real problem with search in EspressoNet isn’t the inflexibility of the interface, but the abominable quality of its metadata, much of which comes from Google Books.”


If MARC is to be replaced by a RDA-friendly container, our colleagues who are most proficient with metadata standards and controls will be needed more than ever.

Librarians and library administrators in the 21st century do need to make changes and it comes as a no shock to any practitioner that many of these changes are underway. The core values of our profession are as relevant as ever and there are many more opportunities to be explored than threats to our existence. Now if we can get the punditocracy to pay more attention to convergent efforts underway as opposed to lamenting the loss of our 20th century practices; the rhetoric can catch up with the practice.

Author’s Notes

10. LOCKSS. <http://lockss.org/lockss/Home>
You’re marooned at LaGuardia because of bad weather. You’re talking jobs with your fellow detainees. What would you tell them about ARTstor?

It’s a great mix of the old and the new – all the values of “old” institutions (pedagogy, libraries, museums) and all the things that the Web makes possible. I also enjoy the way that we are both completely mission-driven — in trying to support the people and places who are themselves mission-driven — and market-responsive. Knowing that places can either subscribe or not keeps us very focused on doing a good job and trying to be responsive to what our users really need. So, it’s a great project for people with a mix of skills.

How does ARTstor compare to such image providers as Google, Yahoo, or Bing? Is Art Project Powered by Google competition?

We’re really very focused on content, tools, and services that are directly useful for academic work. Other Web resources (be they image search services or aggregating services like Flickr) have awesome reach and are incredibly useful. But what we do is very focused. For example, we do something that is, essentially, old-fashioned: we bring very large image files together in one place. This requires work, but it allows us to provide access to details, to manage rights in a complex IP environment, and to be a consistent and dependable resource. Also, by bringing teachers and scholars into one shared environment, we can build on the knowledge revealed in their collective actions (to provide a “recommender” function, for example, based on which images other instructors have put in image groups with a given image). So, while those other image capacities of the Web are dazzling (and also useful for academics in many ways), we are really focused on providing coherent and high quality collections that respond to what users need and on building tools (like the one-button download of an image group into PowerPoint) that are really all about academic work.

Just as visual literacy became recognized in learning circles and catapulted by the Web into notoriety, funds fell away for innovative learning. How can ARTstor play a role in this learning/economic environment?

It’s great fun to be involved in improving the flow of images for academic work in a time when the appetite for visual evidence is shooting up. Art historians have been (for a hundred years) the leaders is using images for evidence — in other words, using the image as the primary source text for supporting an argument rather than as an illustration of some other piece of evidence. And this mode — of “reading” an image — is becoming important to a whole range of disciplines. With 80,000 images from Magnum Photos, ARTstor can help Political Scientists and Historians work with some of the most powerful primary source visual material of the 20th century.

When you ask what role we can play (besides providing content), I often think about how interesting it is to be working so closely with libraries at this point in time. Many of the librarians that we work with license ARTstor and think about how it can help in their evolving role in service provision – integrating the library more and more into the fabric of teaching and learning on campus. For them, we’re a great part of the natural experiment that is going on whereby service-oriented libraries not only have to reach out and do “marketing” of the resources, but also serve as creative partners as they help their users figure out what they can now do. While some libraries aren’t running full speed toward this future, those that are are great partners with experimenters like us.

Intellectual property rights and their management are key to ARTstor’s model. Can you briefly describe and explain ARTstor’s approach to this?

The copyright rules vary from country to country, and the doctrine of fair use (in the US) is crucial but not composed of black-and-white rules. Since we both build collections and distribute content worldwide, we have to be respectful of different legal structures. But what we’ve believed (in working on these issues for over a decade) is that the best long-term approach is one that respects and balances the needs of all involved — those who create works, those institutions and individuals who own and care for objects, and those who want to use images for educational purposes. Since our mission is so focused on educational non-commercial use, we’ve been able to work with this range of constituencies – for-profit organizations like Art Resource, Scala, Magnum Photos, and the private estates of artists — that all care about keeping the legacy of artists’ thriving. They are always relieved to hear that we care about the quality of the image and the data, and we respect their concerns about how the artist’s work is represented. We also care a great deal about the issues that are important to museums, archives, artists and photographers — about rights issues, about their need to generate revenue and thrive, about their concerns for how they are represented and that there be attributions for their work. The educational users that we serve benefit if we have positive, strong, and respectful relationships with all of those who will be able to make more and more content flow into users’ hands.

In your travels for ARTstor, what have you heard from faculty and students about visual resources and their use? How has this changed ARTstor?

The first thing that we always hear is, “more please.” Sometimes other people forget that primary source materials (and the hunger for them in academic work) are endless and insatiable; we never forget this. So, while we keep adding coherent and meaningful collections from all media, all cultures, and all time (we add about 150,000 images a year and we recently launched collections such as: Réunion des Musées Nationaux; SFMOMA; Yale Peabody Museum of Natural History; African field photography by Herbert Cole; Milton Rogovin’s archive of social documentary photographs; archaeological sites from Sites and Photos; Renzo Piano Building Workshop; Barnes Founda-
tion; Ringling Museum Circus Collection; and Warburg Institute to name a few), the real answer is to increase the reach of the network. We have been working with institutions on a shared cataloging and image management platform to increase the flow of images, and we work with the brilliant people at the Getty Research Institute and partners around the world on authority files and data standards that will help the flow of images. So, there are a range of approaches that we take to increasing the flow of images, but they are all in response to the understandable clamor for “more, more, more.”

In the past, users had issues with the offline viewer and using PowerPoint. Can you fill us in on your latest developments regarding image use off the Web site?

In the early days of ARTstor, when we constrained the export of images to very small postcard sizes because of rights concerns, we created our own proprietary offline viewer (OIV) which allowed for the use of much larger images offline, but in an encrypted software environment. For those who loved it, it was a great solution, but it was a whole new software (and thus a hassle for others). We worked for years towards greater compatibility with PowerPoint, and we came up with an approach that still respects the concerns of content providers (and keeps the source image safe) while providing an incredibly simple way to create PowerPoints (with images filling the screen and data being mapped into the Notes fields in PowerPoint).

By providing this “export to PowerPoint,” we were conscious of the fact that doing so will lower the usage statistics regarding use of the ARTstor Digital Library (because users don’t have to come back to ARTstor every time they re-use a presentation or study for a quiz). But we just felt that it was the best way to serve our users well (and still respect the concerns of content owners). OIV is still used by a devoted crowd and we are working on updating it for those who are dedicated, but our guiding principle has been to recognize that different users have different presentation preferences and we should be as flexible and responsive as we can to those different preferences.

Is ARTstor related in any way to JSTOR?

We’re born from common roots (the Mellon Foundation) and from common principles (balancing community interests, developing mission-driven and sustainable business plans), and we have learned a ton from them, but we are organizationally independent of each other. Obviously our name was chosen to signal our admiration for JSTOR’s accomplishments and service to the community (they were already very successful when we were just starting in 2001). And we frequently talk and are always open to collaborations with them that would serve our shared users.

ARTstor used to host images supplied by subscribers. This has evolved into your Shared Shelf product. Give us a brief history of what this is all about, its value to users, and its cost to ARTstor members.

We began a hosting pilot program almost as soon as we launched the library as a live service in 2004. We did this because we knew that users needed to be able to mix remote (ARTstor), local (institutional), and personal content in their teaching and their research, and the rights environment that we were managing didn’t allow for a lot of exporting of content out of the ARTstor workspace. We began hosting image collections for individual institutions and for consortia and serving them back to end-users within the ARTstor workspace. It required a lot of work (on the part of local staff and our staff) to build and sporadically update these collections, but it turned out that end-users really liked having these different pools of content together in one place.

So, even when we started allowing the download of full-screen images that would allow users to mix content in softwares like PowerPoint, a number of colleges and universities asked us to work with them to build a cloud-based way of bringing together the various pockets of image content that lived in different softwares and in different data models around campus – in art history or architecture or classics or East Asian studies or the map collection. “We can’t share image collections on our own campus,” one library dean told us, “let alone share with ARTstor or with other campuses – and we want to.”

Shared Shelf has been designed in partnership with eight partner colleges and universities (Harvard, Yale, Cornell, NYU, Miami, Illinois, Middlebury, and Colby) and with the Society of Architectural Historians (with whom we built an early version of several components of the service). It provides a flexible enterprise-wide platform for managing image projects – efforts currently divided across campus among librarians, visual resources curators, instructional technologists, and individual faculty members. Projects can be very simple or very complex and can be published to a range of different targets – into the ARTstor work space, into an institutional repository, on to the open Web, for example.

It’s very early in the project, and the fees (which aim only at cost recovery and keeping the service robust) are designed to allow schools to begin using the system and pay more as they get more value out of it (to a capped limit). How useful can it be? Since the data model is flexible and it can publish content into an end-user friendly work environment, will it work for astronomy, geology, or medical teaching images? We’re still learning. Will it end up simple enough for individual faculty members to use it for their personal collections (so that they’ll be backed up, portable, and built with easy-to-use authority control)? Again, we’ll learn a lot in the next year or two of working with the first hundred places.

What’s in store for ARTstor mobile users?

Our first release of mobile (for iPhones and iPads) introduced a flashcard, quizzing capacity. There’s an endless amount that we could do (cool features, a thousand different mobile devices to support) and we’re still sorting through users’ priorities. Obviously, as a non-profit, we need to balance our financial constraints (and our desire to keep fees as low as possible) with the hunger of users for all the flexibility that mobile provides. We’re rarely in the best position to be inno-

About James Shulman

James Shulman serves as the President of ARTstor, a nonprofit organization that provides over 1,300,000 images, software, and services to over 1,350 colleges, universities, museums, and schools. Prior to creating ARTstor in 2001 with colleagues at the Mellon Foundation, he had worked at the Foundation since 1994, writing about educational policy issues and the missions of not-for-profit institutions, and working in a range of research, administrative, and investment capacities. Two of his articles were published recently: in a collection on The Ethics of Philanthropy by Oxford University Press (on the creation of new nonprofits) and by the Stanford Social Innovation Review (on bringing innovation to innovation-resistant institutions).
The work that we do is driven by the characteristics of primary source material. These materials are endless – so people need endless amounts of it for their work. Managing such content is technologically and legally complex and we believe that managing such material should be thought of as a shared responsibility and as a shared need, and so it isn’t an issue that can be resolved locally. Everything that we do – whether it is aggregating market-responsive content in the ARTstor Digital Library, building tools that facilitate flexible sharing, or working with the community on shared cataloging thesauri -- is tied together by a need to build shared infrastructure – and to build trust. There are all sorts of barriers that need to be overcome: some have to do with individuals’ fears; some have to do with institutional pride or institutional politics. Many have to do with overcoming technical barriers. But no one will work toward shared solutions if they don’t trust those who are involved in the provision of the shared solution. We try to earn this trust by genuinely identifying with the community’s needs and delivering effective solutions.

It’s great fun!

---

Where do you see ARTstor in five years? Are you looking at media beyond still images? Is technology or content key in your evolution?

If the community wants us to do more work in video or other media, we certainly can, but we try to make sure that we start by doing well what the community has asked us to do. For us, the very good question of “is ARTstor about content or technology or data services?” doesn’t have one answer – we see ourselves as a value-adding node in an area (the academic use of primary source material) that must be networked and is best served by community-minded non-profit efforts.

---

18th and 19th Century Newspapers
As good as gold

From the birth of a nation to a nation in turmoil, get the perspectives on our country's first 150 years. Coverage of debates on finance, national debt, commerce, trade, the economy and the wars. How have things changed, and how have they stayed the same? From revolution to expansion to conflict to healing – follow contemporary coverage of these events as covered by:

- African American Newspapers: The 18th Century
- The Liberator (1831-1865)
- The Pennsylvania Gazette (1728-1800)
- The South Carolina Gazette (1732-177)
- The Virginia Gazette (1736-1780)

For a FREE TRIAL visit America's origins contact our exclusive sales and marketing agent:
iris.banney@unlimitedpriorities.com
Call: 239-549-2384

Accessible Archives
www.accessible.com
nal archiving program. This will make more than 500,000 journal articles from 200 journals accessible to anyone (about 6% of JSTOR total content). Kudos for opening up this valuable archive of historical content. <http://www.jstor.org/>

**Best Contract Options**

**E-book Library (EBL)** For several years EBL has led the e-book aggregator market in extremely flexible licensing options for e-books. Their combination of short term loans (defined by each customer) leading to a purchase (or not), combined with free previews of e-book content has lead the way in recent demand driven acquisitions programs. Other e-book aggregators are now emulating and catching up on EBLs lead. <http://www.eblib.com/>

**Best Customer Support**

**Serials Solutions** Serials Solutions has a strong suite of products and customers have found this company helpful and responsive. This is one of the many reasons why this company is at the head of the pack in its target market. <http://www.serialssolutions.com/>

**Best Effort**

**Pubget** This project is the brainchild of Ramy Arnaout (degrees from Harvard, Oxford and MIT) and represents one of the best efforts to create an end-user biomedical portal for PubMed, OAI, ArXiv and IEEE.

The interface will get the user directly to the full-text content (if no content is available your library’s link resolver can be invoked). Any non-profit organization or individual may use the service at no cost.

**Vaporware Awards**

**EBSCO’s Auto Repair Reference Center** has been awaiting a new platform for over a year and anxious users are still wondering when it will arrive. Of course, it’s better to announce and not launch than to prematurely launch a faulty product. <http://www.ebscohost.com/public/auto-repair-reference-center>

**Not Ready for Prime Time (Special One Time)**

**ProQuests New Interface** The new interface for the suite of ProQuest databases seems to have been prematurely launched. Users report many bugs and inconsistencies and the whole family of databases is not available under the new interface. But, the interface does look very powerful and promising. <http://www.proquest.com>

**Worst Pricing**

**Nature Publishing Group** Complaints about Nature’s high prices, unwillingness to work with consortia and annual price increases have upset librarians for some time. They have excellent content but someone just needs to say “no.” <http://www.nature.com/npg_/index_npg.html> ■

---

**Choice Reviews Online**

www.cro2.org

- Search all *Choice* content
- Create, save, and e-mail lists
- Ability to create a customized profile
- Access to over 162,000 *Choice* reviews
- Download, print or e-mail search results
- View Web exclusives

For more information visit

**Choice** at the Charleston Conference Vendor Showcase in Charleston, SC

November 2, 2011