

The Electronic Journal of Academic and Special Librarianship, 2009, Vol. 10, No. 2.

ISSN: 1704-8532

<http://www.icaap.org/>

<http://southernlibrarianship.icaap.org/>

[http://southernlibrarianship.icaap.org/content/v10n02/dotson\\_d01.html](http://southernlibrarianship.icaap.org/content/v10n02/dotson_d01.html)

© 2009 ICAAP

## Overhauling of a Science Library Web Area

**Daniel S. Dotson**

### Abstract

This article covers the revamp of a science library's web area. The library's web area contained outdated web pages, unused files, and an architecture that made it hard to maintain. All of these issues with the web area resulted in an examination and revamping of several aspects of all the files on the web area. Server statistics were used to determine the usage of the files in the library's web area. Files containing textual information were categorized according to their topic and the usage of each topic area was examined to determine the distribution of usage amongst the topic areas. Files were also categorized according to file type in order to determine the distribution of various file types. The architecture of the web area was improved in order to make the files categorized into proper topic folders and thus easier to maintain. Files were moved as needed in order to fit into the new architecture. Unused web pages were integrated into other web pages or deleted entirely. Redirects were created for users that have used bookmarks or links to the old location. Web pages were all examined and appropriately revised in order to provide concise, up to date, and accurate information and links to information. The article ends with the future plans for the web area, including a regular schedule of maintenance.

### Introduction

The Science and Engineering Library (SEL) at The Ohio State University (OSU) has a web area (<http://library.osu.edu/sites/sel/>) that has grown dramatically in the past several years. With 630 files in SEL's web area, there were many files to maintain. Yet many files on the server were outdated, unused, and even no longer needed to be there. The usage of SEL's web files was examined and the content and organization was assessed and updated. In addition to examination of server statistics and the content of the files, SEL's staff and subject specialists gave their opinions on perceived current and future needs for the web area.

### Literature Review

The University of Nevada-Las Vegas' library's web site redesign involved some of the same issues as SEL's redesign. These include content, maintenance, and policy evaluation and review. In addition, there was a desire for more consistency and better navigability (Vaughan 2001, 82). Statistics from this revamp experience showed that pages with information about electronic resources were the most commonly used pages (Vaughan 2001, 84). Vaughan (2001, 85) describes that the team working on this project found eight elements found on its web pages:

1. Electronic journals
2. Databases/indexes
3. Selected internet resources
4. Selected print resources
5. Style guides
6. New items

7. Contact information
8. Other links as determined by the content originator

With other possible elements, including online catalog subject searches, related local web sites, discussion lists, associations, etc.

Some other important issues in a web site include that the architecture of the site should be shallow and consistent (Xue 2004, 186), ease of use, perceived usefulness (Heinrichs et al 2007, 2326), not being too wordy, ease of getting the information, keeping the site simple (Adams and Dougherty 2002, 592), and maintenance of the web site (Breeding 2005, 26, 28-29). All of these issues apply to SEL's web page area. SEL had an inconsistent and unorganized architecture, some web pages were hard to use and contained bad links, some web pages were too wordy, and the web area was maintained irregularly. All of these issues pointed out by these authors were things that SEL's web area needing improvement and these improvements are discussed later in the paper.

### **SEL's Web Area**

The Science and Engineering Library primarily covers subjects in and serves the academic departments covering astronomy, chemistry, engineering (including computer science), mathematics, physics, and statistics. In 2004, architecture items moved from SEL to the newly opened Architecture Library. Several other libraries on campus serve other science subject areas.

SEL's web area, with 630 files, needed revamping. SEL's subject specialists determine the content of most of SEL's web area, with some pages' content being the responsibility of staff in the circulation department. However, the updating of the web pages fell in the past to whoever had the ability and time to keep the pages up to date. With turnover of subject specialists and staff, many pages have gone without being updated in some time (in some cases, over five years), while unused or unneeded files remained on the server without being deleted.

In addition, another problem found was that many files were sitting in the main area (i.e., not in a sub-folder), while some files were in topical folders. A file not in a topic area folder meant that not only would someone have to open it to determine what topic it about (especially files with ambiguous names), but it also made responsibility for its content unknown. When new subject specialists and staff are hired and others leave, it also makes it harder for them to know what files are in areas of their content responsibility.

Due to all of these factors, it was decided that a revamp of SEL's entire web areas was needed. Goals in the revamp of SEL's web area were:

1. Determine how much various types of pages are being used
2. Checking all pages for updated content
3. Make sure links go where they are supposed to go
4. Link print materials to the catalog so location will be up-to-date instead of listing a call number
5. Reduce overlapping pages (by combining pages or deleting duplicate pages)
6. Delete unused and unneeded files
7. Place files in folders with appropriate topic areas
8. Coming up with a regular maintenance routine

Many of the issues mentioned in the literature review were also inherent in SEL's web area. Particularly in need were a better architecture for the files in SEL's web area, content of the pages (including information on the pages, ease of use and non-wordiness), and maintenance of the web area.

The main purpose of SEL’s web area is to provide access to information in the previously mentioned subject areas. The goals of the revamp were to make SEL’s web area more up-to-date in terms of content, more accurate (links going to their correct destination, links taking users to the catalog so the current location of print materials will be available), easier to navigate, and easier to maintain. The revamp’s goals are thus meant to make SEL’s web area a better place for users to find accurate and up-to-date information.

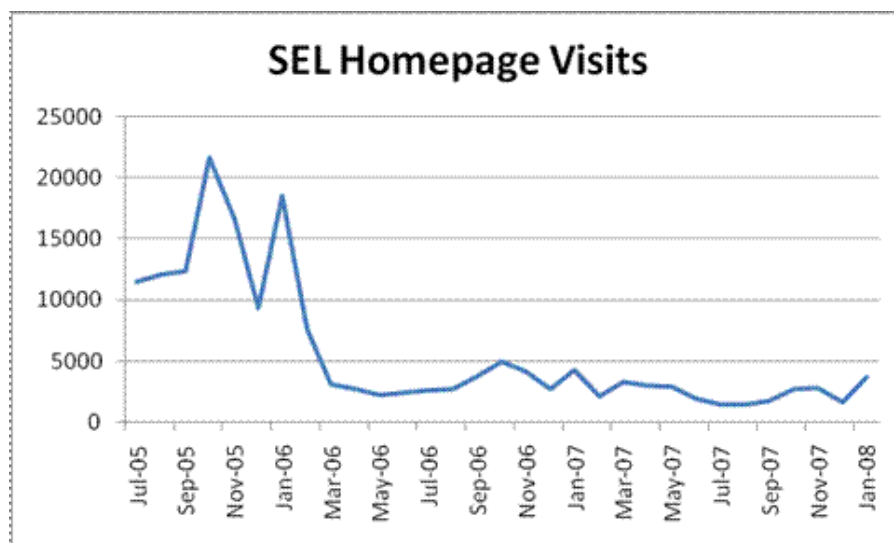
## SEL’s Web Area Usage

AWStats was used to gather usage statistics SEL’s web area. The information for SEL’s web area focused on all pages in the directory [library.osu.edu/sites/sel](http://library.osu.edu/sites/sel). Web statistics used for SEL’s web area were found under the “Navigation” area and the data used came from the “Viewed” column and viewed on a month-by-month bases. Figures for the overall OSU Libraries’ web site usage used in this article came from the “Summary” section.

One of the goals of the revamp of SEL’s web area was to determine the usage of all of SEL’s web page. SEL’s main web page (<http://library.osu.edu/sites/sel>) is the most used web page in SEL’s web area and is one of the most frequented web pages in the OSU Libraries’ web site. This page is the starting point that directs users to many of the other pages in SEL’s web area.

The first three months of 2006 saw a dramatic decrease in the number of visits. While a drop in usage could be part of the explanation, another factor comes into play to explain part of the dramatic drop seen. In 2006, SEL’s homepage was removed as the homepage for the public computers in the building. Since this change, the SEL homepage has seen a steady use, with drops during the summer sessions and breaks, when student traffic on campus in general is low (Chart 1). Examining the usage between May and June for the years 2004-2007, the usage of all of the OSU Libraries’ web pages dropped an average of 32%. The spring quarter typically ends before June 10th, indicating a drop in usage corresponding to the end of the quarter and the beginning of summer classes.

Chart 1: SEL Homepage Visits



The web pages and pdf documents (with the exception of pages related to architecture, which were already slated for deletion) on the SEL server were divided into several topic areas,

depending on their main focus. SEL's main page, covering all topic areas, was not placed into any area and its usage is described above. A page's topic area was determined by its main topic. For example, a web page that mostly had information about library and web resources, but had short examples of catalog searches was labeled as Library & Web Resources, but not as Handouts and Instructional (Table 1).

**Table 1: Topic Areas of Pages and PDF Documents**

<b>Area</b>	<b>Description</b>
Displays	Many displays had corresponding web pages. The main SEL Displays page and all subpages for displays are included in this area.
Handouts & Instructional	Handouts (PDF or other format) about how to use resources and web pages about how to use resources are included in this area, included a SEL web page with a list links to handouts.
Library policies & Information	Information about the library (hours, staff, floor information, call number range explanations, etc.) is included in this area.
Library & Web Resources	Information about library and web resources, including free resources, links to the catalog for print resources, etc.
New books	Pages with information about new books.
News	News about the library, including PDF newsletters.
Personnel training	Training placed on the SEL server for SEL personnel.

Note these areas are similar to some of the elements as mentioned by Vaughan (2001, 85) in the literature review. His first four elements (Electronic journals, Databases/indexes, Selected internet resources, and Selected print resources) correspond to SEL's pages labeled as Library & Web Resources. His New Items corresponds to New books and also possibly News. Vaughan's Contact information would be included in SEL's Library policies & Information. His style guides would be part of SEL's Handouts & Instructional. His last element (Other links as determined by the content originator) could be included in any of the above, since the pages were assigned according to the majority of the information they hold.

AWStats software gave the number of times a file was viewed and these files were categorized as stated above. Server statistics were examined from the period November 2004 to January 2008 (Table 2).

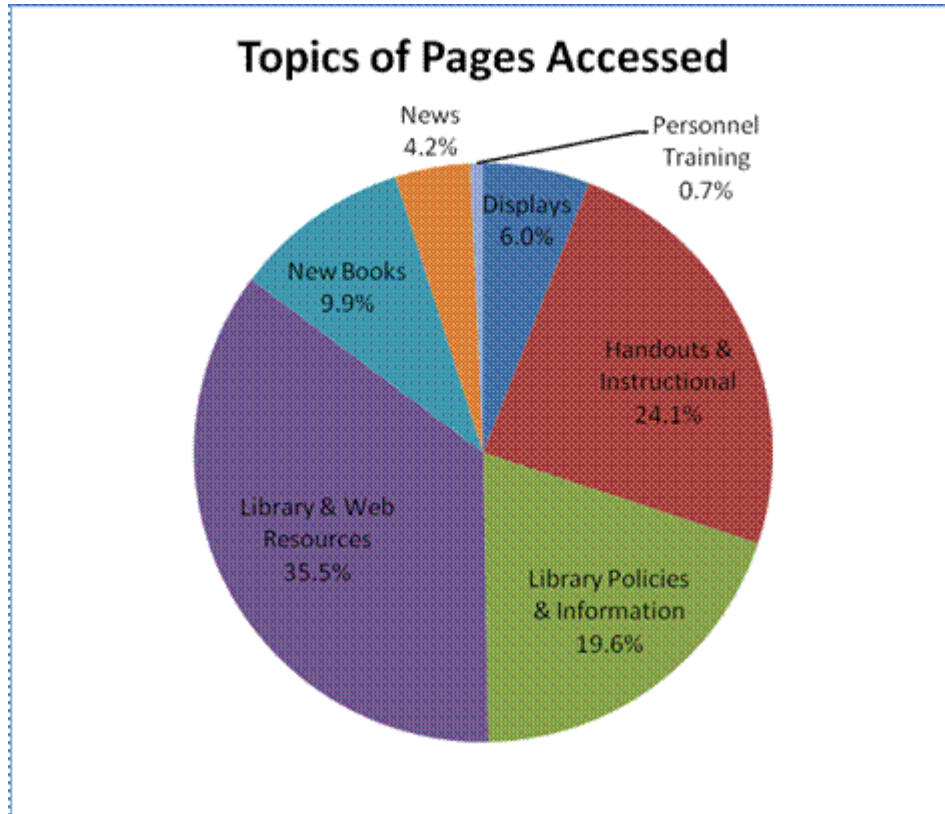
**Table 2: Overview of Usage of Web Pages and PDF Documents**

November 2004 - January 2008						
Area	Individual Pages/ Documents Accessed	Total Views	Average Number of Views per Page/ Document	Median Number of Views per Page/ Document	Max Number of Views per Page/ Document	% of total Page/ Document Viewed
Displays	200	6901	34.5	22	188	6.0%
Handouts & Instructional	809	24141	29.8	16	389	24.1%
Library policies & information	659	37918	57.5	30	427	19.6%
Library & web resources	1191	63508	53.3	37	416	35.5%
New books	331	10878	32.9	16	324	9.9%
News	141	2085	14.8	8	93	4.2%
Personnel training	24	65	2.7	2	10	0.7%

AWStats gathers statistics by month. Note that the column “Total Pages/Documents Accessed” indicates how many unique pages or documents were accessed over the time period examined for every month examined. For example, if a page was examined each month during the period, that would account for thirty-nine unique pages viewed (one month included both html and php versions of a page, so possibly forty total views total if both versions were used that month). The column “Total Views” indicates the total number of views for all pages or documents in the “Individual Pages/Documents Accessed” column. Pages or documents without any views do not appear in AWStats statistics.

Library & Web Resources had the majority of the pages accessed, indicating that most of SEL’s web area usage comes from people needing to access information resources. Handouts and Instructional and Library Policies & Information make up the next largest chunks of usage (Chart 2).

**Chart 2: Topics of Pages Accessed**



However, some interesting information is apparent when looking at average usage of each category, which takes into account the total number of pages accessed each month. A few differences were noted when comparing the total number of hits versus the average number of hits (Table 3).

**Table 3: Comparison of Total Hits to Average Hits**

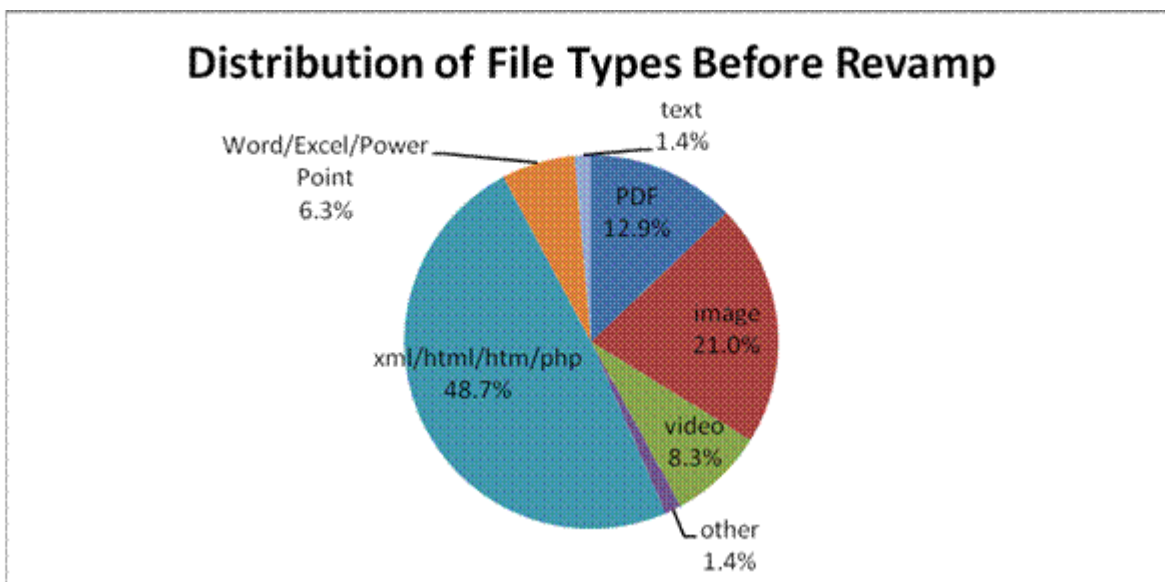
Rank by Total Hits (most to least)	Rank by Average Hits (most to least)
Library & Web Resources	Library Policies & Information
Library Policies & Information	Library & Web Resources
Handouts & Instructional	Displays
New books	New books
Displays	Handouts & Instructional
News	News
Personnel training	Personnel training

Library & Web Resources is clearly the area with the most files accessed. However, Library Policies & Information had more average hits. In addition, while Displays ranked fifth in the number of hits, it had the third highest average number of hits, indicating that users are accessing the Display pages. Similarly, Handouts & Instructional files dropped from third place in terms of total hits to fifth in average use. It is clear from the total and average usage of these pages that news-type information and personnel training both receive rather low usage.

This information is useful in determining information that users want to see on SEL's web area. Users clearly see policies and information about the library as important information. Library and web resources are also important pieces of information. It is therefore important that these types of pages have up-to-date information on them.

Most file types on the server before the revamp of the web area were web pages (html, php, or xml). Image files and PDF documents were the next most common file types. Video files, text files, various Microsoft Office files, and other miscellaneous file types made up just over eleven percent of the files on SEL's server area. With web page files and images associated (or at least at one point associated with) the web pages constituting nearly 70% of the total number of files, the files to be examined were mainly of those file types, with other files constituting around 30% of the total files in SEL's web area (Chart 3).

**Chart 3: Distribution of File Types Before Revamp**



### **Updating the SEL Web Area**

With over 630 files in the SEL web area, examining SEL's web area was a major undertaking. Some of the pages did not see much or any use, some files were duplicated in multiple locations, and some files were extremely outdated. It was therefore decided to examine all SEL web area files for purpose, need, content, usage, and topic. After examining each file, the process of reconstructing, updating, and streamlining SEL's web area began. When examining each of the files in SEL's web area, the following actions were taken:

- Every file in the web area was placed in a topic folder (with the exception of SEL's homepage, any system files required by it, and the site index page)
- A manually-produced site index page was created using the folder structure created as a basis
- Pages duplicated in other server areas (such as the architecture pages, see below) were eliminated
- Pages with a good deal of duplicate content were combined
- Removed duplicate files within SEL's web area where possible and point to a single file
- Redirects were created in order to send users to pages that have moved

- Files that have seen little or no use in the time period from November 2004 to January 2008 were examined for elimination or integration into other existing pages
- Images no longer in use and no seen need in the near future were eliminated

SEL personnel gathered in March 2008 to discuss the planned redesign of SEL's web area. The ultimate goals of the meeting were as follows:

- To convey the need to redesign the structure of SEL's web area
- To seek feedback from the SEL personnel about the web page's structure
- To seek ideas of SEL personnel about future web pages for the web area
- To see ideas about categories of web pages

Several ideas were gathered from this meeting, including having a password-protected area for training, possible topic areas for pages, and the need for certain pages. These were all taken into consideration when the web area was revamped. In particular, it was decided that all training materials be put in a folder which has content that will be eventually be made password-protected.

As indicated in the examination of the statistics, Library & Web Resources were the most used files. Therefore, it is vital that the information on those pages be current. Library Policies & Information received the highest average use, indicating users also find these pages important. When examining pages for updating, the usage of the various topic areas was taken into consideration when determining what to do about certain types of pages.

Two sets of files were removed totally. In 2004, the Architecture collection moved out of SEL into the new Architecture Library. Architecture files remained on SEL's server area even though the Architecture Library got its own area. These files were deleted and a CD of their contents was given to that library's staff in case they needed them. Redirects were put in place to direct users to the main Architecture Library page. Tutorials were created for students submitting honors theses to the institutional repository, the Knowledge Bank. These files were moved to an area outside SEL's web area. Redirects were put in place to these tutorials' new locations.

Examining the statistics of files in use, personnel training files were the lowest use by far. The files were student-training files that were no longer in use by the department that created them. These files were eventually put into a folder in SEL's web area that will later be made into a password-protected area. There was a stated need for making a student training area in the future and these can be used as a possible starting point.

In some cases, files were duplicated within SEL's web area. The most common were image files. Images still needed for the public pages were located into a single folder and unneeded and duplicate files were deleted.

The process of examining files, updating files, updating and/or combining pages, and deleting files as necessary then began. Ignoring system files and files in the folders mentioned above (training, architecture, and tutorials), twenty web page files were found not to be used. Of these, only one file was kept, which was updated and placed in a different folder. The other nineteen pages were redirected to a new page (six of which had their information integrated into that page).

All other files were examined to determine if they needed to be retained and if their content was duplicated elsewhere. After deleting, moving, and combining various files, 304 files remained, down from 630. This resulted in a reduction of the number of web files by 52%.

## **Updating Pages for Content**

When updating pages for content, several goals were taken into consideration:



- Print items would link to the catalog's record for the item
- Links would be examined for linking to correct pages
- Where appropriate, new items and more recent editions would replace older items

SEL has undergone many changes within the past five years, which has resulted in print resources' locations not being a constant. Items can also, at any time, be placed on reserve, leaving the previous location, if listed on a web page, incorrect. Many print resources listed on web pages indicated their location and call number. In order to eliminate the need to update these whenever books' location changes, links to each item's catalog record were used instead, with an indication to the user that the item is print material (so that they know the link is not an electronic resource). As stated earlier, Library & Web Resources pages have the highest amount of use. Therefore, outdated information about a book's location would be of disservice to users. Linking to the catalog for the most up-to-date information will solve this problem.

Additionally, all pages were checked for the possibility of updated content. Newer print and electronic resources were added to pages. With the recent addition of thousands of ebooks, many web pages now had a wider variety of print and electronic resources available to users.

Links were manually examined for currency and going to the correct site. While a link checker may indicate that links work, not all links went to the intended page. One link in particular was now leading to a pornographic page (this link was, of course, removed). Where necessary, some links led to a page with a note as to what the user needed to click on to get to the needed resources, as some websites do not provide direct URLs to their web pages.

Wordiness having been mentioned in the literature as an issue, descriptions were kept short where needed and unnecessary descriptions were removed. In many cases, book titles and links to web pages were clear enough as to indicate what the users are clicking. Larger pages were given a linked directory at the top of the page in order for the user to be able to quickly get to the topic they need on the page.

In the case of one category, the statistics indicated that a possible problem was occurring. Some handouts received a lot of use - perhaps too much use. Some handouts showed hundreds or even thousands of uses since they were placed online. One would like to think that the handouts developed are being used, but when an extreme is noticed, one begins to wonder why so many uses are being seen. Handouts for JSTOR, CRC Handbook of Chemistry and Physics, and ASTM Standards received abnormally high number of hits. Examining hits via the OSU search box (which searches all OSU websites and appears on all public OSU pages), one sees that these resources handouts are among the top hits. The handouts for databases developed with SEL users in mind were then examined to see where they ranked in an OSU web search. The results (Table 4) indicated that the handouts often ranked near the top in the list of search results, indicating a very possible path for users to find the handouts (while they may be trying to find out how to access the resource itself).

**Table 4: Handout Hits and OSU Search Results Rank**

Search in OSU search box	Link # in search results	Handout hits	Time period
JSTOR	1	2548	March 06 – Jan 08
CRC	3	1151	Feb 07 – Jan 08
CRC handbook of chemistry and physics	4	1151	Feb 07 – Jan 08
ASTM Standards	1	747	Jan 07 – Jan 08
ISI citation indexes	1	153	Feb 07 – Jan 08
Web of Science	2	153	Feb 07 – Jan 08
MathSciNet	1	54	March 06 – Jan 08
Electronic Journal Center	1	54	April 07 – Jan 08
History of Science Technology and Medicine	3	28	April 07 – Jan 08

Due to the number of uses of several of the handouts and due to the fact that they appear near the top in results list, it was decided to add a link to the resource within any electronic resource handouts in SEL's server area. This enables people who find the handout in their search to be able to link to the resource needed. Since over 2,500 users have accessed the JSTOR handout within less than a two-year period and over 1,100 users have accessed the CRC handout within eleven months, it seems valuable to have links to the electronic resource within the handout so that users will be able to access it if they find the handout via the OSU web search. Such information would never have been known without the knowledge of the web statistics for SEL's web area and adding links to the resources within the handouts allows users who stumble upon the handouts via a web search to easily access the resource that they may be trying to find.

The web pages in SEL's web area were thus improved to provide more updated and accurate access to information in the subject areas SEL covers. All these changes made to SEL's web pages are key in having pages users will want to use.

### **Before and After Revamp**

As stated earlier, all files were determined to be of a specific topic area and placed in a topic area with that topic. New folders were created for new topic areas and folders no longer necessary were deleted (Table 5).

**Table 5: Folders Before and After Revamp**

<b>Folders Before Revamp</b>	<b>Folders After Revamp</b>
<b>arch</b>	
<b>archtesting</b>	
	<i>astr</i>
chem	chem
collection	collection
comp	comp
ebooks	ebooks
eng	eng
ETD	ETD
	<i>genresources</i>
handouts	handouts
<b>KBtutorials</b>	
	<i>instruction</i>
	<i>Internal (to be made password protected)</i>
lists	lists
math	math
	<i>multidiscipline</i>
newbooks	newbooks
newsletters	newsletters
osu_navbar	osu_navbar
phy	phy
pictures	pictures
rss	rss
Scripts	Scripts
seldisplays	seldisplays
	<i>selinfo</i>
<b>student_training</b>	

Note: Folders in **bold** were deleted, folders in *italics* were created. Files related to student training were moved to the new internal folder.

With over half of the pages deleted, quite a few pages required redirects. Pages were redirected as follows:

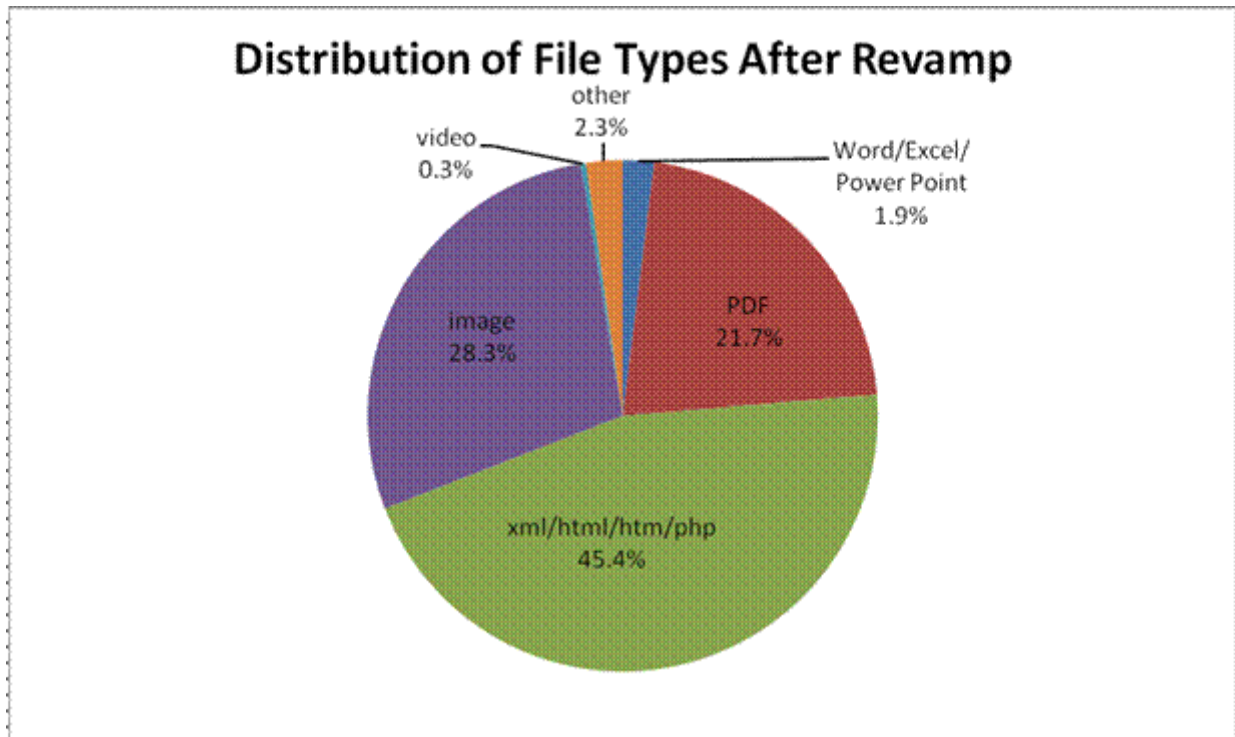
- All architecture pages were redirected to the main Architecture Library web page
- The video tutorials page and associated video were redirected to their corresponding page in their new area in a non-SEL server area
- Pages integrated into another page were deleted and redirected to that new page
- Pages deleted but not integrated into a new page were redirected to an appropriate page, in some cases SEL's main page

In order to provide better access to information on SEL's web area, site index was created to make access to SEL's web pages easier. One area left out was the Internal folder, containing in part the files used for student training. The folders listed in Table 5 were used as a basis for the site index, although the names were changed slightly as headers in order to make more sense to outside users.

The types of files in SEL's web area changed dramatically after the revamp. Significantly, the percentage of Microsoft Office file types has decreased (from 6.3% of the files

to 1.9%) and since the video tutorial files were moved, the number of videos has dropped greatly (from 8.3% of the files to 0.3%). Images and PDF files increased their percentage of the files in SEL web areas, while the number of web page files showed a slight decrease, yet remained the majority of the pages on the server. It is important to note that after the revamp process began, several newsletters in PDF format were added, while few PDF files were deleted. This had a noted impact on the percentage increase in the number of PDF files on the server (Chart 4).

**Chart 4: Distribution of File Types After Revamp**



### **The Future of SEL's Web Area**

The process of examining server statistics and page content, deleting pages, moving pages to new folders, combining pages, updating links and content, and consulting with information technology staff and SEL subject specialists staff all resulted in dozens of hours of work spent on revamping SEL's web area. The work has resulted in a web area with a stronger architecture, elimination of duplication, and deletion of over half of the files.

In order to make sure that SEL's web area remains current, it was determined that a regular process of evaluation of web pages for content was needed. In addition, links need to be checked for accuracy of destination by a person (instead of a link checker) at least yearly. It was determined that the summer would be an ideal time to review pages, when the traffic in the building is lighter and allows more time for projects such as this.

Future pages created will also be placed in the appropriate area, not in the general area, in order to better keep track of content. In the past, many pages were not placed in topical folders, with many pages as a result being duplicated in content or even in their entirety in multiple locations. A better architecture will enable pages to be more easily controlled and monitored in the future.

Part of the reason for SEL's web area being overburdened with duplicate and unused files was lack of maintenance. As Breeding (2005, 29) suggested, maintenance is an important factor in keeping a web page up. Improving the architecture of SEL's web area will better enable files

to be controlled and monitored in the future.

This overhauling will be of great assistance in 2009 as the OSU Libraries make some dramatic changes to its web site, and the architecture will become a more topic-based architecture and its web pages will be created and managed via a content management system. SEL's revamp will result in making its transition to the new architecture and content management system easier.

Overhauling the web area was a major undertaking and took several dozen hours worth of time in order to evaluate and edit pages. All pages' links were updated to make sure they went to the appropriate page. Over half of the files in SEL's web area were deleted and the remaining files were moved into topic folders in order to provide for easier maintenance. Usage statistics indicate that while the bulk of the information in our pages, which falls into the Library & Web Resources category, the usage statistics indicate information about library policies and procedures and library displays to also be important to users. Regular maintenance and evaluation of page content and the improved and stricter architecture will prevent the need to do a major overhaul in the future.

## **Bibliography**

- Adams, Mignon. 2002. How useful is your homepage? A quick and practical approach to evaluating a library's Web site. *College & Research Libraries News* 63, no. 8:590-592. <http://vnweb.hwwilsonweb.com/hww> (accessed February 18, 2008).
- Breeding, Marshall. 2005. Analyzing Web Server Logs to Improve a Site's Usage. *Computers in Libraries* 25, no. 9:26-28-9. <http://vnweb.hwwilsonweb.com/hww> (accessed February 18, 2008).
- Heinrichs, John H., et al. 2007. Determining Factors of Academic Library Web Site Usage. *Journal of the American Society for Information Science and Technology* 58, no. 14:2325-2334. <http://journals.ohiolink.edu/ejc/> (accessed February 18, 2008).
- Vaughan, Jason B. 2001. Three iterations of an academic library Web site. *Information Technology and Libraries* 20, no. 2:81-92. <http://vnweb.hwwilsonweb.com/hww> (accessed February 18, 2008).
- Xue, Susan. 2004. Web usage statistics and Web site evaluation: a case study of a government publications library Web site. *Online Information Review* 28, no. 3:180-190. <http://journals.ohiolink.edu/ejc/> (accessed February 18, 2008).