AAHSL

Institutional Repositories (IR)

Survey Summary and Analysis
2008 SURVEY
2005-2008 COMPARISON

This is a comparison of four short surveys on institutional repositories (IRs) that were distributed online to the AAHSL membership (regular members) on June 14, 2005; June 26, 2006; June 21, 2007; and June 24, 2008 (closed on July 25, 2008). The survey effort was supported by the AAHSL Charting the Future Committee and the Scholarly Communications Committee. Summary results of these surveys, observations, and conclusions follow.

Overall Survey Response Rates

	2005	2006	2007	2008
AAHSL Libraries	114	116	116	113
Respondents	51	54	55	66
Response Rate	45%	47%	47%	58%

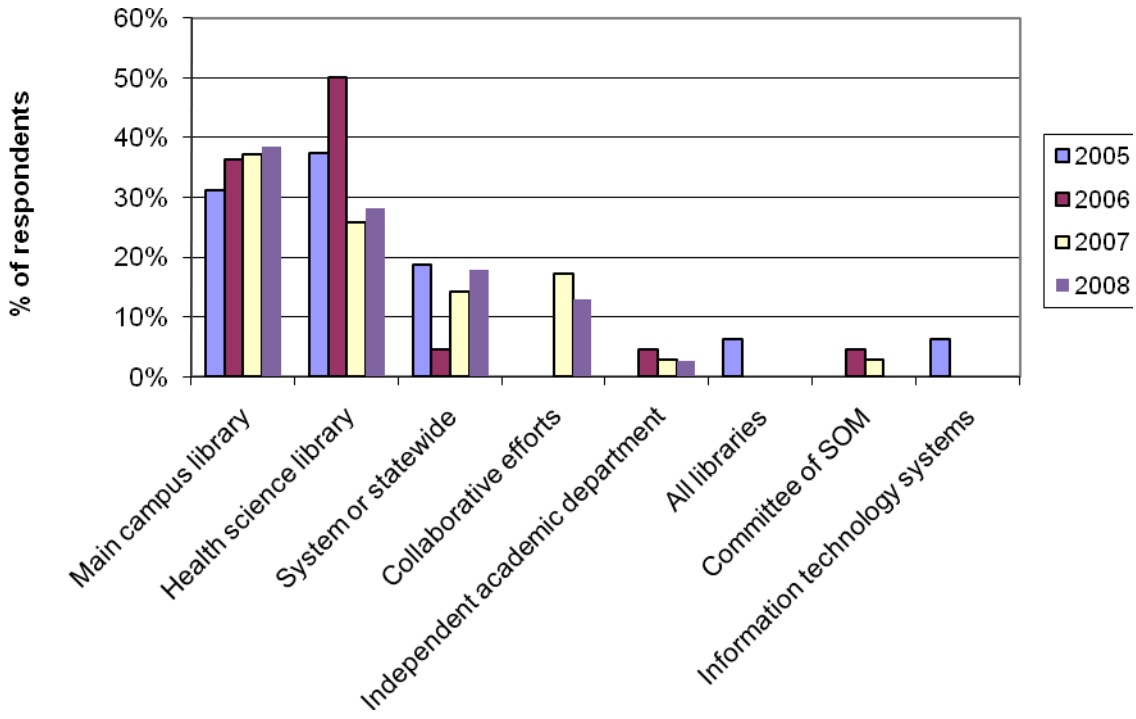
Of these respondents,

	2005	2006	2007	2008
No IR	35	32	27	34
Yes IR	16	22	27	32
No IR %	69%	59%	50%	52%
Yes IR %	31%	41%	50%	48%
% Increase		38%	23%	19%

Following are responses from those who reported "yes" to having an IR: *(note: not all respondents answered all questions)*

Who owns or promotes the IR?

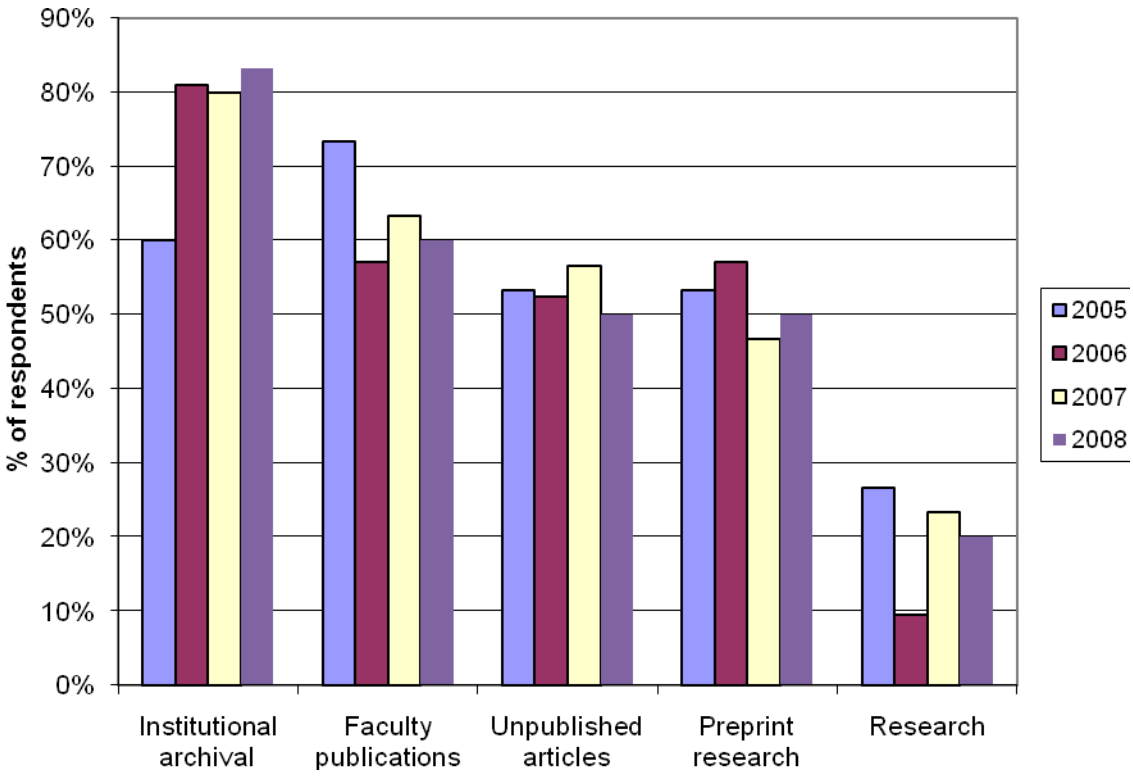
	2005	2006	2007	2008
Main campus library	5	8	13	15
Health science library	6	11	9	11
Collaborative efforts	0	0	6	5
System or statewide	3	1	5	7
Committee of SOM	0	1	1	0
Independent academic department	0	1	1	1
All libraries	1	0	0	0
Information technology systems	1	0	0	0



Observations: Main campus libraries continue to outnumber health sciences libraries in owning or promoting institutional repositories. The chart above clearly shows an upward trend in main campus library involvement over the past four years. We have eliminated a breakdown of entities in the "other category" here because it had yielded little meaningful information. Institutions identifying the "other" category in their responses cited collaborative efforts as owning or promoting their campus repositories. Collaborations were typically between the health sciences library and other libraries or academic units within the same institution.

Kinds of content maintained in the IR?

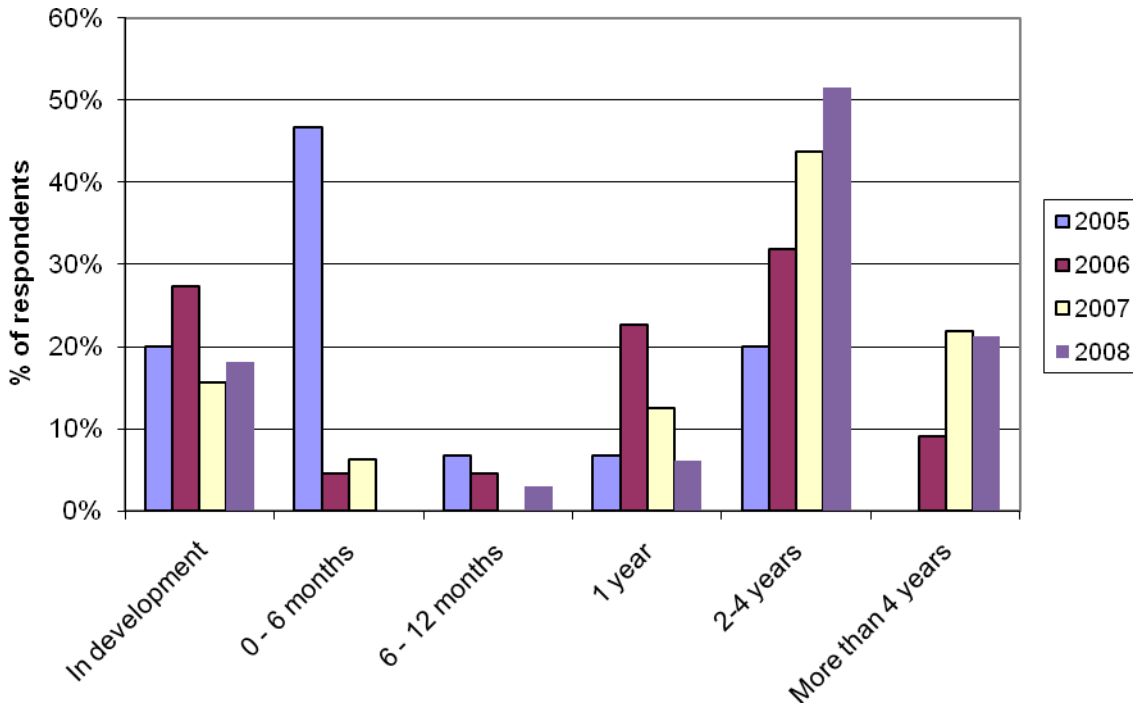
	2005	2006	2007	2008
Institutional archival	9	17	24	25
Faculty publications	11	12	19	18
Unpublished articles	8	11	17	15
Preprint research	8	12	14	15
Research	4	2	7	6
Institutions Responding	15	21	30	29



Observations: Institutional repositories continue to be used predominantly for archival purposes, but faculty publications, unpublished articles, and preprint research information also continue to be held by the majority of IRs in almost equal amounts. The use of IRs to house electronic theses and dissertations (ETDs) was not reported specifically for 2008 but was indicated in the comments section as increasing.

How long has the IR been in use?

	2005	2006	2007	2008
In development	3	6	5	6
0 - 6 months	7	1	2	0
6 - 12 months	1	1	0	1
1 year	1	5	4	2
2-4 years	3	7	14	17
More than 4 years	0	2	7	7



Observations: Like last year's observation, it appears that a new wave of institutional repositories is under development as the first couple of waves show a crest in the 2-4 year range. There is also the possibility that the IRs in development last year took longer than expected. Six IRs are still in development, but a few have been further developed and populated from the previous year, with the exception of those in the 2-4 year range. There is clearly staying power as IRs continue to grow on campuses, particularly those in use for 4 years or more that appear to have been developed in conjunction with the main campus library and/or university-wide system (e.g., California). New efforts appear to be under way, many of which are spearheaded by the health science library, or increasingly, in conjunction with the main campus library.

What is the number of unique digital objects in the IR?

	2005	2006	2007	2008
None	2	4	1	3
50 - 100	2	1	1	0
100 - 400	1	3	5	2
400 - 700	0	1	3	2
700 - 1,000	0	2	2	0
1,000 - 2,000	1	1	2	4
2,000 - 3,000	1	2	2	1
3,000 - 4,000	0	0	1	1
4,000 - 7,000	0	1	2	5
More than 7,000	1	0	3	3

Unknown	1	5	10	7

Observations: There continues to be a wide range in the number of objects contained within institutional repositories, and the situation continues to be related to whether the IR was developed system-wide or in conjunction with the main university library. Many of the initial efforts to populate IRs for individual health science libraries still range from just a few hundred discrete objects to thousands. Larger institutional efforts (i.e., main university library/system wide), as one expects, continue to have a greater number of objects in their IRs. Archival materials, when included, significantly add to the items in IRs. It is interesting that three institutions cited "none" and seven "unknown" for number of objects in their IR. Generally, the number of unique digital objects in the IR appears to be increasing slowly and has been relatively consistent across years.

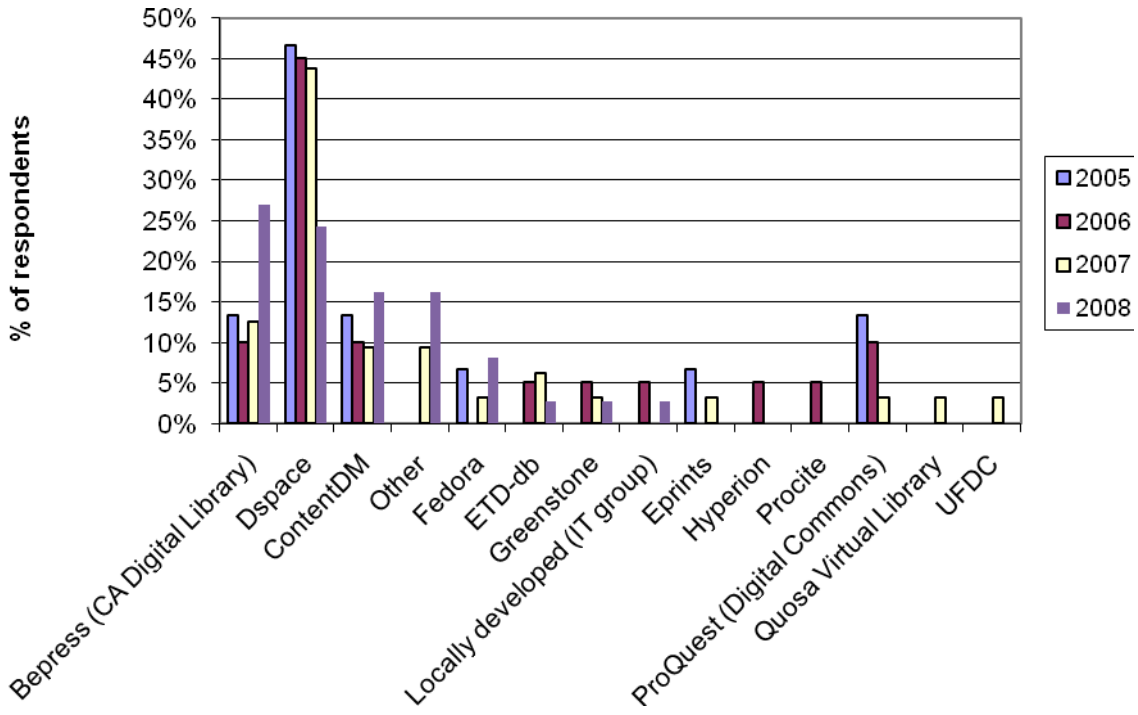
Percentage of faculty who contributed to the IR

	2005	**2007**	**2007**	**2008**
< 10%	12	20	29	28
> 10% and < 50%	0	0	1	2

Observations: Consistently across all years, faculty members are not contributing to their IRs, so others are doing the populating – most likely librarians. Populating the IR is an added chore for librarians if they are responsible for the IR.

Document management software used for the IR

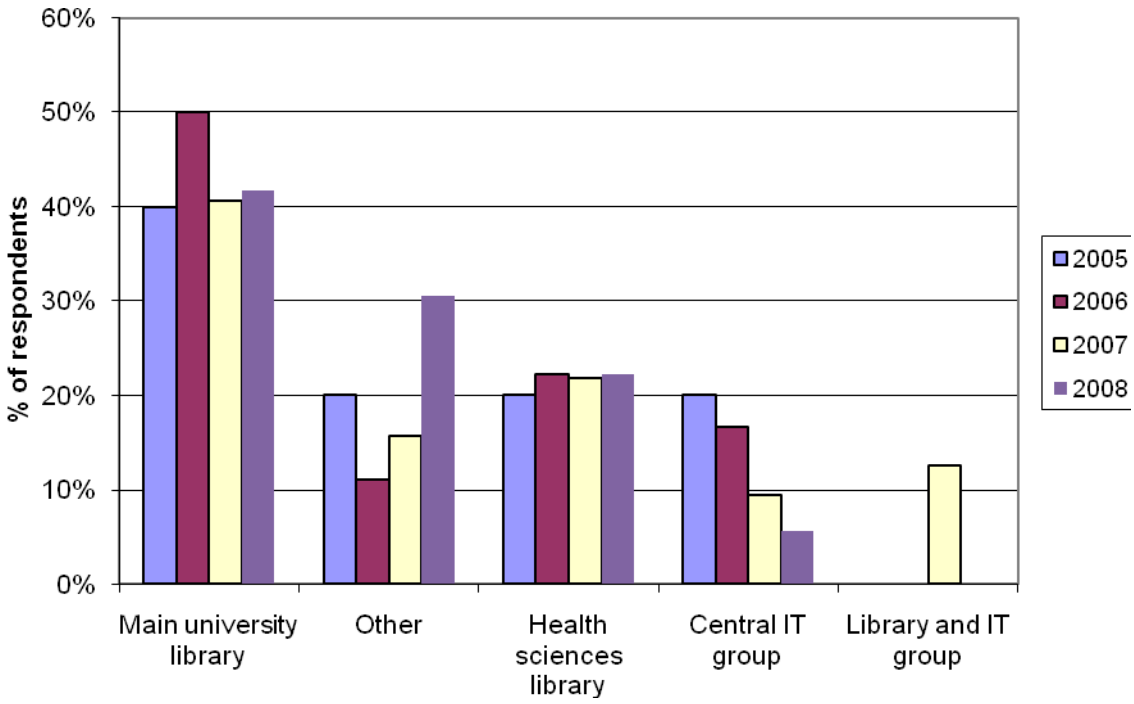
	2005	**2006**	**2007**	**2008**
Bepress (CA Digital Library)	2	2	4	10
Dspace	7	9	14	9
ContentDM	2	2	3	6
Other	0	0	3	6
Fedora	1	0	1	3
ETD-db	0	1	2	1
Greenstone	0	1	1	1
Locally developed (IT group)	0	1	0	1
Eprints	1	0	1	0
Hyperion	0	1	0	0
Procite	0	1	0	0
ProQuest (Digital Commons)	2	2	1	0
Quosa Virtual Library	0	0	1	0
UFDC	0	0	1	0
	15	20	32	37



Observations: The biggest change here is that DSpace no longer remains the dominant document management software in use as it dropped from 14 to 9 users. The Bepress software (e.g., University of California system) is now the dominant software used, with an increase from 4 to 10 users. ContentDM doubled in users from 3 to 6, as did the "other" category, and Fedora went from 1 to 3 users. The Eprints software is no longer cited as being used. New software names are listed from the previous survey. Virtually no electronic theses and dissertations (ETD) software was listed under the "other" category, although it was referred to in the comments section as being used.

Who provides technical support for your IR?

	2005	2006	2007	2008
Main university library	6	9	13	15
Other	3	2	5	11
Health sciences library	3	4	7	8
Central IT group	3	3	3	2
Library and IT group	0	0	4	0
	15	18	32	36



Observations: The findings here are consistent from the previous year, with the exception of the "other" category. Main university libraries remain dominant in providing both the impetus and central technical support for campus IRs. Efforts by health sciences libraries to support IRs remain strong, however. The campus IT group no longer provides technical support as often as the health science library, and apparently several medical libraries are on their own for developing and supporting the IR. The large state/system-wide efforts are unique. The support listed in the "other" category is varied, with support most often cited as being offered by Bepress or another software hosting company.

General Comments

The comments made by the AAHSL directors who responded to the survey are presented verbatim (with institutional names removed) for the year 2008:

2008

This is a university-wide repository. The total documents are from all the state campuses and in all fields.

The university library is just getting started and has invited the School of Medicine to participate; there is really nothing yet to report.

Our IR is in the pilot stages. We hope to go live with the IR in the next 6 months.

The main library is in the process of recruiting a Scholarly Communications Librarian to take charge of the project. The system and systems support is in place, and some material -- theses -- uploaded. The Health Sciences Library will be working on the CIHR (Canadian equivalent of NIH) grant output requirements for open source access to research output.

Whoever created this survey ignored the type of content in our IR, the most popular type in the world, electronic theses. That is why I did not check any of the survey content categories.

There are plans to begin looking at an institutional repository at the University Libraries level.

Our library will begin depositing library-related materials during the summer/fall of 2008. We hope to begin a pilot project with a department in the School of Public Health to post materials for them as well.

The President has recently indicated that the 4 primary Libraries: University, Medical, Law and Theology, will have primary responsibility for development of the University's Institutional Repository, so it is just now in the early planning stage of development.

The University Library has an archival digital depository, and will support other projects on request. It's not a true institutional depository yet. No significant change from last year.

We have a database of our dissertations and theses available through Proquest Digital Dissertations and Theses.

Note: Health Sciences Libraries collaborate with the University Library on developing and recruiting content for the repository. The server itself is maintained by the University's Office of Information Technology.

Our IR solely consists of ETDs.

Our repository houses electronic these and dissertations using the software developed at our Institution.

Fairly low activity at this point.

We are excited about implementing an institutional repository for scholarly works and are in the
process of creation with expected live date sometime summer 2008

A development person for digital collections; IR part of digital collections.

The main university library has asked us to work with them in developing such an institutional repository. Perhaps by next year, there will be something underway.

We are beginning to house dissertations in the repository. We also house learning objects developed by faculty and others. We are looking at using for case based learning. Repository has been slow to catch on.

Most of the content is digital theses & dissertations which are now required for all students. In addition there are some image collections and sound recordings (oral histories).

DSpace, in fact the concept of repositories, has been slow to catch on here. The Program in

Audiology and Communication Science is actively and enthusiastically providing documents for deposit. They use the system in their recruiting efforts. We are in the process of a reintroduction of DSpace to the professional library staff and is re-examining appropriate uses and document types for inclusion. We will be developing a marketing campaign aimed at a defined audience within the school to help us judge the value of the repository to various programs.

Our main campus library is beginning to investigate potential platforms for an IR for the entire State University campus. I have indicated my interest in working with them on it.

We continue to move slowly into the IR realm. The software we chose was based on selection by the university for data management. Documentum is a powerful tool for input but is woefully inadequate as a search engine. We are exploring open source options as well. Eventually we will add much content but for now it is still in its infancy stage.

Both the Library and Museum have promoted the need for the Institution to establish a repository, but so far no action.

We are inching closer as there is an interest not so much in publications are there is in teaching objects and the archival retention of them.

We hope to begin participating soon by contributing some archival materials.

Observations: The comments again appear quite varied. There appears to be more collaborative relationships, with some institutions slowly continuing to develop IRs. The perceptions of respondents are interesting, and new software is apparently being used to develop new IRs or refine existing ones. Particularly interesting is how slowly across the last couple of years the number of items in institutional repositories has grown. Institutions and libraries appear to be looking for collaborative relationships for developing IRs, and the support necessary to develop, maintain, and grow an IR. There appear to be a fair number of AAHSL libraries still interested in developing an IR for their institution or being part of the effort to do so.

Conclusions

There is not a significant amount of new or changed information from past surveys, although there are some interesting indicators of change. There is a greater response rate from the previous survey (55 to 66 respondents), with more respondents indicating that they have an IR (27 to 32), but more also indicating that they did not have one (27 to 34), which means more respondents indicated "no IR" than on previous surveys. The first IR survey was completed in 2005, and AAHSL libraries continue to be interested in IR efforts, but except for the large university-wide efforts, the development of IRs in most health science libraries still appears to be fledging and slowly growing. Development of IRs (with the exception of the large collaborative players) has been slow. Most of the successful, concerted, large efforts continue to be aligned with main university libraries or systems (e.g., California Digital Library), and individual health science library efforts continue to slowly increase and even more slowly expand. The same appears true for large private university systems. The larger university systems appear to have provided

the impetus and support for early IR development and therefore have leapfrogged ahead in terms of number of objects and staying power, and they continue to have the largest holdings. DSpace is no longer the dominant document software, and has just been slightly passed by Bepress (e.g., California Digital Library) with the largest number of users.

There continues to appear to be an increasing effort by main campus libraries to participate in the IR effort. However, individual autonomous health science libraries are still getting in the game, and several plan to do so in the next year or so. Several health science libraries clearly intend to develop an IR and are in a planning and development stage. This past year we saw a slight increase in the number of libraries mounting IR efforts, and perhaps a few being dormant.

Although there appears to be little consensus overall from AAHSL institutions regarding the importance, need, or wish to develop an IR via their library for their institution, there continues to be some effort by a few health science libraries to explore developing an IR on their own (we saw an increase from 27 to 32 libraries that indicated "yes," they have an IR or one in development). But IRs can be added work for a library. Few faculty members use the IR to archive their own materials, leaving the tasks of identifying content and populating the IR to librarians. This is perhaps one reason why the number of discrete objects appears to increase so slowly. There is a wide range of objects in IRs. This range is apparently related to whether the IR is developed system-wide or in conjunction with the main university library, how long it has been in existence, and what content items are identified as important to populate the IR. Many of the initial IR population efforts for individual health science libraries still include just a few hundred discrete objects (or unknown), and archival materials leads the list, followed by faculty publications and research. The institutions that have the largest number of objects in their IRs continue to be system-wide or main university led IRs, and the number of objects continues to grow in these large repositories, but slowly in other smaller IRs.

Technical support for the IR document software is related to the type of software used for the IR. DSpace is no longer the dominant document management software in use – surpassed by Bepress (CA Digital Library), but they are virtually tied in place, followed by ContentDM, Fedora, and the "other" category. The Eprints software is no longer cited as being used. Several new IR software names are listed from the previous survey.

What happened to DSpace that saw its market share decline? Could it be that the other document management software offers greater "one-stop" technical support, or is there some other reason? Most often the main university library, when included with the health science library, provides the technical support for the IR. This is not surprising, given their large staff sizes and greater resource support. But the technical support provided by the health science libraries as discrete units appears still in development or growing.

There has been a realization that the development of an IR takes considerable work, time, and energy, and the effort to develop an IR in a health science library is often an institutional issue. If the institution determines that as part of its strategic plan it wants an IR, then the library will participate; often this participation takes place as a university-wide effort, or in collaboration with others.

Summary

This survey has tracked IR development for the past 4 years, and there have not been as many changes as one might have anticipated from 2005. Certainly, the number of health science libraries promoting and/or maintaining institutional repositories continues to slowly increase, but the development has been slow, except in the case of large university-wide efforts or in conjunction with main university libraries. Although the IR effort is increasing, development has not proceeded as quickly as some respondents previously indicated. Based on individual comments, it appears that additional resources are needed to adequately support a health science library in developing an IR, and most often the impetus is to protect the archives of the institution, followed closely by faculty publications and research information. The most successful efforts (as determined by number of objects populating the IR) continue to appear to be via the main campus library and statewide higher education efforts. As more and more health sciences libraries internalize information technology practices, we may see more development and a changing face in the IR in the health sciences world. However, more often than not the academic health science library IR development is part of a larger university or collaborative library effort.

finis

Prepared by:

Paul Schoening, MS, MBA
Associate Dean and Director
The Bernard Becker Medical Library
Washington University School of Medicine
660 S. Euclid Ave., Box 8132
St. Louis, MO 63110
voice: 314.362.3119
fax: 314.747.4416
email: paul.schoening@wustl.edu

Thomas Singarella, Ph.D.
Professor, Chairman, and Director
Health Sciences Library and Biocommunications Center
The University of Tennessee Health Science Center
877 Madison Avenue
Memphis, TN 38163 USA

voice: 901.448.5694
fax: 901.448.6855
email: tsingarella@utmem.edu

Revised: August 6, 2008