

BOBCATSSS 2010 @ Parma, Italy

Dates: Monday 25th, Tuesday 26th, Wednesday 27th January, 2010

Bridging the digital divide: libraries providing access for all?

"How user-friendly are user interfaces of open access digital repositories?"

Radovan Vrana

Department of Information Sciences, Faculty of Humanities and Social Sciences, University of Zagreb, Zagreb, Croatia

rvrana@ffzg.hr

Abstract

Digital information resources available on the Internet have become conditio sine qua non of modern research and teaching. In the last decade and a half the Internet and especially the Web had introduced many types of online information resources that emerged and vanished. Only those that were closely associated with important institutions in society (such as libraries and universities) and proved usable survived. Until recently, libraries have been places where university staff seeks quality information for research and teaching, and students for learning. With the proliferation of the Web, students have replaced libraries with search engines and are now using them as primary tools for discovery of information necessary for completion of their written exams, term papers, presentations etc. In their effort to provide researchers, teachers and students with quality content, universities and libraries started development of digital repositories, digital archives of the intellectual product created by the faculty, research staff, and students of an institution. Digital repositories contain research data, journal articles, preprints, technical reports, books, theses and dissertations and other material used in research and educational process. The diverse content of digital repositories represents rich resources for research and teaching. In addition to the diversity and quality of the content, another issue - user interfaces - attracted attention of computer specialists because user interfaces are means of successful use of digital repositories. The research of top 20 open access digital repositories showed that the biggest repositories share common characteristics which help

their users in their daily access to the content of repositories. Despite helpful similarities, some of these digital repositories should improve their design to become more attractive and attract younger generations of users seeking knowledge elsewhere on the Internet.

Keywords: digital repositories, science, teaching, learning

1. Introduction

Digital information resources available on the Internet have become conditio sine qua non of modern research and teaching. While scholars in all disciplines have recognized the importance of investing in physical infrastructure the bricks and mortar of libraries, museums, and archives for a long time, in the cyber age, collections of digital content and the software to interpret them have become the foundation for discovery; they have entered the realm of infrastructure (Arms and Larsen, 2007). The nature of digital information resources is changing fast with more information being migrated to digital format and more made available in digital format (digitally born material). It is therefore not surprising that information in digital format represents the prime and fastest growing collection in academic libraries today (Swanepoel, 2005). Such an important change in organization of scientific and educational content has preoccupied all the key stakeholders in the global information infrastructure. This change has introduced a new digital information resource especially relevant for research, teaching and learning - digital repository. Digital repository has gained popularity during this decade owing to the fast development of information and communication technology and need of scientists and students for access to quality material for research and learning. Despite the continuous growth of the global popularity of open access digital repositories, scientists are still not fully aware of the existence and impact of this type of digital information resource on their work. For instance, the 2005 survey of journal author behaviour and attitudes done by Rowlands and Nicholas showed that researchers' awareness (worldwide) about institutional repositories was very limited: only 9.7 per cent (of 5513 researchers) declared that they had "a little" or "a lot" of knowledge about repositories (Rowlands and Nicholas, 2005). Younger generations (students, among others) may be even less aware of the existence of digital repositories. They are particularly susceptible to digital technology and their demand for digital information has increased since most of new material is born digital and is readily accessible on the Internet. According to Mie and Nesta, the Internet has taken over the role of the main information provider among young generations, changing their perspective and attitude towards libraries greatly because of the inclination of users towards search engines instead of library resources (Mi and Nesta, 2006). As researchers and teachers, we want our students to have the best possible information resources available at their finger tips; however, we will have to make additional effort to present online information resources such as digital repositories in more appealing light. We should demonstrate the real value of digital peer reviewed content for their education and use it in a more innovative way.

2. Digital repositories

Digital repository is a broadest term for a digital archive storing scholarly output of an academic institution. Because of the fact that they are part of an institution, they are often called digital institutional repositories. According to Johnson, digital institutional repository (a digital information repository that is a part of university or other institution) is a digital archive of the intellectual product created by the faculty, research staff, and students of an institution and accessible to end users both within and outside of the institution, with few if any barriers to access (Johnson, 2002). It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation where appropriate, as well as organization and access or distribution (Swanepoel, 2005). For Sutradhar,

institutional repositories provide a web-based mechanism for researchers to deposit (self-archive) and access their research publications (Sutradhar, 2006). For Huwe, digital repositories are specialized environments which offer narrow slices of the information universe (Huwe, 2008). An institutional repository can contain e-prints of scientific papers, research data, but also e-learning materials and other forms of institutional intellectual outputs, which are generally not published or preserved elsewhere (Hockx-Yu, 2006). E-prints may also include: journal articles, preprints, technical reports, books, theses and dissertations (Warner, 2003).

Digital repositories emerged in first half of 1990s, and their number is constantly growing as academic community finds them to be important for storage and use of their research output. The history of digital repositories development can be divided into two groups. According to Jones, first repositories were little more than just databases and file systems, in the style of a standard web application, populated and maintained by librarians, guided principally by the ideal of Open Access, rising journal prices, and the desire to record institutional output for posterity and portfolio. The "next generation" of repositories represent the shift from relatively independent, stand-alone systems, to distributed, federated and highly integrated applications and services (Jones, 2007).

The development of institutional digital repositories can be perceived as a new strategy that allows universities to apply serious, systematic leverage to accelerate changes taking place in scholarship and scholarly communication, both moving beyond their historic relatively passive role of supporting established publishers in modernizing scholarly publishing through the licensing of digital content, and also scaling up beyond ad-hoc alliances, partnerships, and support arrangements with a few select faculty pioneers exploring more transformative new

uses of the digital medium (Lynch, 2003). Bravo and Diaz see the existence of digital repositories as the logical convergence of faculty-driven self-archiving initiatives, library dissatisfaction with the monopolistic effects of the traditional and still-pervasive journal publishing system, and availability of digital networks and publishing technologies (Bravo and Diaz, 2007). With digital repositories in place, scholarly communication increases because new opportunities are presented for faculty, staff, students, and even the university community as a whole, to share their work with a number of people without the reliance on traditional publishing venues (Graham, Skaggs and Stevens, 2005). The diverse content of digital repositories represents resources for research and teaching, which will play an important role in higher education for training students' research methods. As scientists in different fields gain experience and develop discipline-based methodologies for using largescale digitized content, special collection and new media collections, they will need to train students in the research methods (Waters, 2006) and how to use properly digital information resources such as digital repositories. As a result, digital repositories will help the advancement of knowledge and improvement of education. Furthermore, digital repositories will help libraries to cope with the important task of alleviating access to digital content for learners. According to Ellison, libraries are faced with the task of helping to provide distance learners with the resources and access to resources like never before in the history of profession. The major resources and access should be provided by the institution delivering the instruction (Ellison, 2000).

The existing books and articles on digital repositories mention more frequently the scientific side of use of digital repositories and less frequently describe their use in the learning process. With digital repositories, learners can find and access quality material. To support learning more directly, some digital repositories are oriented solely towards storing learning objects in

electronic format (e-learning objects). One such example is MERLOT (Multimedia Education Resource for Learning and Online Teaching) at http://www.merlot.org/merlot/index.htm. In their article on e-learning objects repositories, Schell and Burns state that it is a responsibility of faculty members to identify, select and assemble educational objects to provide a quality experience to students, and that teaching material must be academically enriching and provide challenges to the student beyond the standard lecture material of the class. They even presented three categories of criteria for evaluation of online learning objects for inclusion in a digital repository (Schell and Burns, 2002):

1.) Quality of Content:

- is clear and concise
- provides a complete demonstration of the concept
- is current
- is relevant in today's situation
- is self contained
- provides accurate information
- is flexible
- includes an adequate amount of material
- summarizes the concept well
- integrates the concept well

Overall quality

2.) Usability:

- is easy to use
- has very clear instructions

- is engaging
- is visually appealing
- is interactive
- is of high design quality

Overall usability

3.) Potential Effectiveness as a Teaching Tool:

- identifies learning objectives
- identifies prerequisite knowledge
- reinforces concepts progressively
- builds on prior concepts
- demonstrates relationships between concepts
- is easy to write assignments for
- is very efficient

Overall effectiveness

Prieto wrote about trust in digital repositories. Trust is necessary if we want to rely on particular digital information resource for research and teaching. Prieto pointed out how digital repositories can be trusted because they meet or exceed the expectations and needs of the user communities for which they are designed. They are designed to be usable and credible vehicles for disseminating information. Finally, he concluded that user communities are the most valuable component in ensuring a digital repository's trustworthiness (Prieto, 2009). We can conclude this part with thought that digital repositories will be trusted as much as they will be actually used by user categories for which they are built.

3. Digital repositories and e-science

Digital repositories are related to e-science. E-science is referring to scientific activities supported by high bandwidth computer-mediated telecommunications networks, and particularly to encompass the variety of such digital information-processing applications that are expected to be enabled by the grid i.e. the general purpose network technology which will serve to facilitate new, computationally intensive forms of scientific inquiry (David and Spence, 2003). E-science is based on the setting up of repositories and the development of infrastructures to permit analysis and sharing of information among researchers based in different places (Bravo and Diaz, 2007). These infrastructures are necessary for e-research, a collaborative activity that combines the abilities of distributed groups of researchers in order to achieve research goals that individual researchers or local groups could not hope to accomplish. Documents needed for research process must be discoverable and re-usable by others (Voss and Procter, 2009). Information tools that facilitate data being structured for efficient storage, search, retrieval, display and higher level analysis, and the codified and archived information resources that may readily located and reused in new combinations to generate further additions to the corpus of reliable scientific knowledge are or are becoming critical for modern science. The progress in this area has compressed the space and time in which data and information can be made available for analysis and use in further research. It has opened up the practical possibilities of integrating and transforming scientific and technical data into virtually unlimited configurations of information, knowledge, and discovery (David, 2004). Open access digital repositories can make this idea come true. As their number grows, it has become evident that digital institutional repositories are now clearly and broadly being recognized as essential infrastructure for scholarship in the digital world (Lynch, 2005). According to Prosser, there are many benefits to institutional digital repositories (Prosser, 2005):

• For the individual:

- o They provide a central archive of the researcher's work
- By being free and open they increase the dissemination and impact of the individual's research
- o They act as a full CV for the researcher

• For the institution:

- They increases the institution's visibility and prestige by bringing together the full range and extent of that institution's research interests
- They act as an advertisement for the institution to funding sources, potential new researchers and students, etc.

• For society:

- o They provide access to the world's research
- o They ensure long-term preservation of institutes' academic output
- They can accommodate increased volume of research output (no page limits, can accept large data-sets, 'null-results', etc.).

4. Evaluation of digital repositories user interfaces

Digital repositories across the world share some common characteristics. Huwe summarized positive characteristics of digital repositories that help their development: they conform to demanding standards for metadata and information architecture (ubiquitous Internet access); they often operate on open source platforms and are attached to research universities or non-profit outfits; they reinvigorate the best in long-term professional values and make them understandable for contemporary society (Huwe, 2008).

Online information resources such as digital institutional repositories can be evaluated by use of methods intended for evaluation of information retrieval systems. Xie and Cool selected six tasks which users have to achieve in order to accomplish their search tasks in online information retrieval systems and which are realized in an interface as functions (Xie and Cool, 2000): database selection, query formulation, query reformulation, access to help function, organization and display of results and delivery of results.

The original set of criteria suggested and used by Xie and Cool will be expanded for use in the digital repository comparison in this paper:

Category 1. Software used for repository operation

Category 2. Rate of freely accessible fulltext (estimate)

Category 3. Access

- Browsing capabilities
- Searching capabilities (Simple and advanced)

Category 4. Organization and display of results

- Sorting capabilities
- Limiting number of results

Category 5. Delivery of list of results

- File
- Print
- Clipboard
- E-mail

Category 6. Information about repository, Help. Guided tour, E-mail contact

- E-mail contact
- General help
- Guided tour
- Information about repository

Category 7. User interface language choice (multilingual user interfaces)

- Number of languages used
- Frequency of languages used

Category 8. Existence of policies and other documents that govern use of digital repositories

- Terms and Conditions (of use)
- Guide to Adding Resources
- Privacy policies
- Existence of peer review
- Citation databases

Category 9. Peer review information, citation in databases and content categorization

- Scientific papers
- Professional papers
- Educational material

To find out more about the user interfaces of the world largest digital information repositories, a comparison was carried out. Evaluation of user interfaces was done by comparing open access digital repositories' user interfaces elements against a list of criteria. The aim of the comparison is to find similarities among user interfaces of the largest open access digital repositories. These similarities can help their easier use, integration into

research and learning process at universities and can facilitate easier training of students about how to use these digital information repositories based on knowledge they already acquired using other available information retrieval systems (online databases etc.).

The main hypothesis is that user interfaces and functions of modern digital repositories are very similar and that the previously acquired knowledge and experience of use of online information resources (in general) can be very helpful when accessing the content of another and new information resource on the Internet.

Enumerated comparison criteria (divided in categories) will be applied to the sample of 20 largest open access digital repositories. According to Registry of Open Access Repositories (ROAR) at http://roar.eprints.org on November 8th 2009 there were 1543 digital repositories registered. Although this number is high, the registry divided digital repositories in several categories (with number of repositories in parentheses): Other (219), Database/A&I Index (36), Demonstration (22), e-Journal/Publication (111), e-Theses (138), Learning and Teaching Objects (13), Research Cross-Institutional (144) and Research Institutional or Departmental (860).

E-journal/Publication category was chosen for this research as one of the most interesting and widely accessible. Many other categories, like institutional or departmental digital repositories sometimes have certain access restriction which can bee seen from the estimate of freely accessible full text. In case of institutional or departmental digital repositories, this estimate is sometimes very low: some repositories among top 20 in this category offer only 25% of freely accessible full text. Because of that, another category - e-journal/Publication - was chosen as it offers very high percentage of freely accessible full text and can be truly called open access.

Top 20 open access repositories from this registry in category e-journal/Publication (with the most records) on November 8th 2009 were: SciELO - Public Health - Brazil, Revistes Catalanes amb Accés Obert (RACO), BioMed Central (Journals), Hrčak - Portal of scientific journals of Croatia, NUMDAM - Numérisation de documents anciens mathématiques, I-Revues: Service d'édition électronique de l'INIST, Revues.org - Fédération de revues en ligne en sciences humaines et sociales, DSpace at Tartu University Library, WWW Conferences Archive, SciELO - Cuba, UPCommons - Revistes i congressos UPC, Universiti Teknologi Malaysia Institutional Repository (UTM-IR), SciELO - Chile, AMNH Scientific Publications, Érudit (Journals), Sistema Eletrônico de Editoração de Revistas, Sistema Eletrônico de Editoração de Revistas, Open Journal Systems, Vernadsky National Library of Ukraine.

Web pages of three digital repositories were unavailable at the time of research: WWW Conferences Archive, Vernadsky National Library of Ukraine and Bioline International EPrints Repository. Another three repositories were chosen as a replacement (next three repositories listed at http://roar.eprints.org): Temaria - Revistas digitales de biblioteconomía y documentación, SciELO – Argentina and SciELO – Peru.

The research was carried out by accessing each digital repository separately and comparing its user interfaces elements against the list of criteria. Each criterion was applied to every repository on the list and the results are presented in tables.

Category 1. Software used for repository operation:

Other softwares (various)	12
DSpace	5
Open Journal System	2

EPrints	1
Fedora	0

Category 2. Rate of freely accessible fulltext (estimate): average rate of freely accessible material in top 20 open access repositories: 76.25%

Category 3. Access

	Browse	Search simple	Search advanced
Yes	18	18	19
No	2	2	1

There were no major differences between digital repositories. All of them except very few, include browsing and searching capabilities. The only difference is the position of browsing and searching interface elements on the Web page (which could alleviate or make difficult use of these two functions). Some interface elements were easy to find while other required a few additional seconds for identification before they could be used. This remark is oriented towards the problem of difference in Web design of digital repositories which can make use of a particular digital repository difficult. Not all repositories included in this research use well known repository software such as E-prints or D-Space, so, differences in user interface design are likely to occur.

Category 4. Organization and display of results

	Sorting results	Limiting number of results
Yes	8	5
No	12	15

While searching capabilities are frequently present, additional manipulation of search results is not present in many repositories, and it is not crucial for the successful use of open access digital repositories.

Category 5. Delivery of list of results

	Export - File	Export - Print	Export - Clipboard	Export E-mail
Yes	15	14	11	12
No	5	6	9	8

Another additional capability is export of search results. Surprisingly, some forms of export are present in significant number of repositories in this research sample while sorting and limiting of number of results are present less frequently.

Category 6. Information about repository, Help. Guided tour, E-mail contact

	About	Help/Support	Guided tour	E-mail contact
Yes	20	14	2	13
No	10	6	18	7

Usually, development of digital repositories is well documented as they are important parts of universities. Digital repositories in this sample offer basic and often short introduction about their history, content and use. Help is present in its most basic form usually directing repository users to browsing and searching functions. E-mail addresses for contact with institutions behind repositories are sometimes present on homepages of repositories, but are, in some cases, very difficult to discover. For repositories that share common software platform, e-mail addresses were easily discovered (due to similarities in site design). Though significant number of digital repositories offers help, only two repositories offer a guided tour to their users.

Category 7. User interface language choice (multilingual user interfaces)

	Language choice
Yes	15
No	5

Multilingual user interface is a prerequisite for repositories if their creators want them to be used by users from different sides of the world. Fifteen digital repositories offer user interfaces in two or three languages, while five repositories offer only one language. User interfaces of digital repositories offer use of several major languages. In all cases, first language offered is the language of the country where digital repository is located.

Number of languages	1	2	3	4 and more
	5	7	8	0

Eight digital repositories offer their user interfaces in three languages which is excellent, and seven repositories offer user interfaces in two languages. There were no digital repositories that offered more than three languages, but all repositories that offered more than one language chose major languages spoken by great number of the Internet users. This result indicates that the more languages are offered the more users will be able to find and retrieve content of digital repositories.

	English	Spanish	Portuguese	French	Catala	Croatia	Germa
					n	n	n
Languages and dialects	19	10	5	4	3	1	1

English, Spanish and Portuguese are three most frequently used languages used for use interfaces of digital repositories. The choice of language depends on the location of the repository server and target population. The results from this sample indicate that almost all repositories offer English language as common choice for use of digital repository user interface.

Category 8. Existence of policies and other documents that govern use of digital repositories

	Terms and conditions	Guide for adding resources	Privacy policy
Yes	5	3	3
No	15	17	17

Very few open access digital repositories offer important documents such as terms and conditions, guide for adding resources and privacy policy. These documents should be made available in all digital repositories to inform repositories users about critical issues about participation in repository development and its use.

Category 9. Peer review information, citation in databases and content categorization

	Peer review	Citation databases	Content categorization
Yes	2	1	0
No	18	19	20

Information about peer review was found in only two repositories, leaving this responsibility to journals and publishers. In only one digital repository was presented data about databases in which content of that repository was cited. There were no data about any type of content categorization in digital repositories (presentations, professional papers, scientific papers etc.).

5. Conclusion

As the Internet has taken over the role of the main information provider among young generations, libraries and universities offer digital repositories as an answer to search engine domination. Since the number of digital repositories continues to grow, they are going to become more visible to different Internet user communities, not just to researchers, teachers and students. They have grown from simple document management systems to distributed, federated and highly networked information resources that are systematically developed and offer possibilities for the advancement of science and education. Open access digital repositories are especially important because they facilitate access to knowledge without restrictions for the widest possible audience on the Internet. The open access initiative will help in broadening knowledge horizons to the Internet users of all ages.

Although user interfaces of the compared open access digital repositories share common functions and elements, they still look too formal to attract students seeking knowledge elsewhere on the Internet. Multilingual user interfaces indicate international orientation of digital repositories. Each additional language attracts new users from different corners of the world. In cases of all repositories, similarities in user interface elements use of repositories'

content easier. This proves the main hypothesis that user interfaces and functions of modern digital repositories are very similar and that the previously acquired knowledge and experience of use of online information resources (in general) can be very helpful when accessing the content of another and new information resource on the Internet.

Still, some of digital repositories in this comparison need new design of Web pages to become not only more attractive (less important) but to become more functional (more important). However, this is highly subjective viewpoint. Judging from the number of total records in each repository compared, open access digital repositories are very popular and many journals found their interest in publishing their content in them. This fact helps digital repositories in becoming a world wide publishing platform for even more journals. Unfortunately, the establishment of digital repositories requires human, technical and financial resources which are not always easily available. The number of 1543 existing open access repositories registered at http://roar.eprints.org indicates that digital repositories have already become part of the global information infrastructure.

References

Arms, W. and R. Larsen. (2007). The Future of Scholarly Communication: Building the Infrastructure for Cyberscholarship. Retrieved from http://www.sis.pitt.edu/~repwkshop/SIS-NSFReport2.pdf

Bravo, B. R. and L. A. Diaz. (2007). E-science and open access repositories in Spain. OCLC Systems & Services: International digital library perspectives, 23(4), 363-371.

David, P. A. (2004). Towards a cyberinfrastructure for enhanced scientific collaboration:

Providing its 'soft' foundations may be the hardest part, Stanford Institute for Economic

Policy Research, Stanford, Retrieved from http://129.3.20.41/eps/le/papers/0502/0502002.pdf

David, P. A. and M. Spence. (2003). Towards institutional infrastructures for e-Science: the scope of the challenge. Oxford: The University of Oxford.

Ellison, J.W. (2000). Distance learning for today's librarian. Library Review, 49(5), 240-242.

Graham, J.B., Skaggs, B. L. and K. W. Stevens. (2005). Digitizing a gap: a state-wide institutional repository project. Reference Services Review, 33(3), 337-345.

Hockx-Yu, H. (2006). Digital preservation in the context of institutional repositories. Program: electronic library and information systems, 40(3), 234-245.

Huwe, T. K. (2008). The Surprising Impact of Digital Repositories. Computers in Libraries, 28(9), 42-44.

Johnson, R. K. (2002). Institutional repositories: partnering with faculty to enhance scholarly communication. D-Lib Magazine, 8(11), Retrieved from http://www.dlib.org/dlib/november02/johnson/11johnson.html

Jones, R. (2007). Giving birth to next generation repositories. International Journal of Information Management, 27(3), 154-158.

Lynch, C. A. (2003). Institutional Repositories: Essential Infrastructure For Scholarship In The Digital Age. Libraries and the Academy, 3(2), 327–336.

Lynch, C. A. and J. K. Lippincott (2005). Institutional Repository Deployment in the United States as of Early 2005. D-Lib Magazine, 11(9). Retrieved from http://www.dlib.org/dlib/september05/lynch/09lynch.html

Mi, J. and Nesta, F. (2006). Marketing library services to the Net Generation. Library Management, 27(6/7), 411-422.

Prieto, Adolfo G. (2009). From conceptual to perceptual reality: trust in digital repositories. Library Review, 58(8), 593-606.

Prosser, D. C. (2005). The Next Information Revolution - How Open Access will Transform Scholarly Communications. In International Yearbook of Library and Information

Management 2004-2005: Scholarly Publishing in an Electronic Era (pp. 99-117). London:

Facet Publishing

Rowlands, I. and D. Nicholas. (2005). Scholarly communication in the digital environment: the 2005 survey of journal author behavior and attitudes. Aslib Proceedings: New Information Perspectives, 57(6), 481-497.

Schell, G.P. and M. Burns. (2002). Merlot: A Repository of e-Learning Objects for Higher Education. e-Service Journal, 1(2), 53-64.

Sutradhar, B. (2006). Design and development of an institutional repository at the Indian Institute of Technology Kharagpur. Program: electronic library and information systems, 40(3), 244-255.

Swanepoel, M. (2005). Digital repositories: all hype and no substance?. New Review of Information Networking, 11(1), 13-25.

Voss, A. and R. Procter. (2009). Virtual research environments in scholarly work and communications. Library Hi Tech, 27(2), 174-190.

Warner, S. (2003). E-prints and the Open Archives Initiative. Library Hi Tech, 21(2), 151-158.

Waters, D. (2006). Managing digital assets in higher education: an overview of strategic Issues. ARL Bimonthly Report, 244. Retrieved from http://www.arl.org/arldocs/events/fallforum/forum05/waters.pdf

Xie, H. and C. Cool. (2000). Ease of use versus users control: an evaluation of Web and non-Web interface of online databases. Online information review, 24(2), 102-115.