

# Digital Repositories and Possibilities of Their Integration into Higher Education

[doi:10.3991/ijet.v5s2.1212](https://doi.org/10.3991/ijet.v5s2.1212)

Radovan Vrana

University of Zagreb, Zagreb, Croatia

**Abstract**—Due to abundance of information available on the Internet, society seeks new types of digital information resources which would facilitate better delivery of information for different purposes. One such information resource type is digital institutional repository. Successful use of digital institutional repositories depends mostly on quality of content they offer, and on unrestricted access to that content. On the other side of the information chain, users i.e. students should be taught how to critically approach and evaluate content they find in digital institutional repositories in order to enrich their learning experience and make it more successful. To find out more about similarities and differences between top 10 largest open access digital information repositories, a comparison of their most important functions was carried out. The results of the comparison give us valuable indicators about possibilities of their use in learning process based on their properties and presence of the most important functions modern information resources should have. Generally speaking, as digital repositories gain more public attention, it can be expected that they will be used even more in future for various purposes, including education.

**Index Terms**—Digital repositories, e-learning, higher education

## I. INTRODUCTION

Information wealth available to us on a daily basis can create an illusion of freedom of access and good organization of the global information environment. In reality, Internet offers an immense quantity of human knowledge and a variety of ways for its collecting, organizing, storing and use. Internet users approach stored knowledge in conformity with their information needs and tasks they wish to complete. Their tasks vary greatly, from finding information necessary for the completion of a business task, scientific research or seminal paper at a university. All of them depend on quality information resources offering digital content and networked support for document delivery. One of the most interesting information resources at the present moment is digital institutional repository, a subtype of digital information repository. 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 [1]. 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 [2]. Diversity of content in digital institutional repositories suggests various possibilities of its inclusion in learning process at universities. In that sense, universities have an important role in the current information and knowledge society: as producers, transmitters, and disseminators of scientific knowledge and professional know-how, which makes them vital in the development of twenty-first century citizens [3].

Content stored in digital institutional repositories including learning materials can be integrated in the learning process as a part of the learning management systems in use at universities or as an external source of information. Prerequisites for this integration are access without restrictions i.e. open access and quality of digital content stored in the repository.

Open access to scientific and non-scientific literature means its free availability on the public Internet, permitting any user to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited [4]. More specifically its purpose is the world-wide electronic distribution of the peer-reviewed journal literature and completely free and unrestricted access to it by all scientists, scholars, teachers, students, and other curious minds. Removing access barriers to this literature will accelerate research, enrich education, share the learning of the rich with the poor and the poor with the rich, make this literature as useful as it can be, and lay the foundation for uniting humanity in a common intellectual conversation and quest for knowledge [4].

High quality of the content provided by digital institutional repositories is another important criterion for inclusion of it in the learning process at universities. The learning process at universities can employ various types of digital content for different purposes, but only those information resources which are able to guarantee high quality of its content are of interest to the teaching staff for selection and use. In science, the quality of published papers is guaranteed by the refereeing process; therefore, scientific papers are recognized as the high quality content appropriate for use in different environments including educational environments like universities. Another category, popular science represents a special category of content created by scientists for general public. Content of such articles and books is simplified when compared to

papers published in scientific journals so that everyone can understand it and can benefit from it. This type of articles can also be included in the learning process, especially in introductory university courses.

## II. ONLINE LEARNING AND ABUNDANCE OF INFORMATION

In addition to the availability of preferably unrestricted access to the scientific and ready made content for learning, modern information and communication technology (ICT) offers the scientists and teaching staff at universities its enhanced capabilities (broadband Internet connections, large computer storage, fast CPUs, mobile storage units etc.) for preparation of the content which can be included in the learning process at universities. For today's science which produces a significant quantity of all types of content (raw data as a result of a measurement, art works, published books and articles etc.), the ICT is very important. There is almost no scientific research today which is not supported or carried out with the help of computers, computer networks or other electronic equipment. This is the reason why science produce large quantities of data which are extremely difficult to collect, organize and store, and yet, these data can very useful for planning of a new research or for the students' learning. The ease with which digital information can be created, combined with the huge increase in computer power and network bandwidth, has led to the proliferation of a vast amount of 'born-digital' data, especially in science and engineering, where petabytes (10<sup>15</sup> bytes) of data are being generated by scientific instruments on a daily basis [2]. Due to the large quantities of online content, only a fraction of it is used in education properly or used at all. All available content is potentially useful especially for distance education where transfer of knowledge and data is essential for the success of the learning process. According to Anderson, learning theory must acknowledge the change from an era of shortage and restriction to an era of abundant content; content resources are now so large that filtering and reducing choice is as important as provision of sufficient content itself [5]. Pinto and Sales discuss the necessity of universities to teach students to deal critically with information; to analyze it, to select appropriate content, and to incorporate information that is relevant into a personal knowledge base. Furthermore, they suggest that students they speak about already have an increasingly detailed knowledge of technologies, they use new habits of communication, and they are interested in new forms of education and autonomous online learning but they have serious difficulties when it comes to dealing with the abundance of information [3].

## III. ISSUES AND POSSIBILITIES OF INTEGRATION OF DIGITAL INFORMATION RESOURCES INTO LEARNING PROCESS

There is no doubt that digital institutional repositories and content they deliver can be used in higher education for the purpose of achieving a better online learning. This situation introduced a dilemma to universities whether or not to employ online learning fully or just partially, combined with a conventional learning. In their article on theoretical framework that combines components of the online learning system and conventional learning system

Patel and Patel suggest that universities should provide education to students by combining components of the conventional learning system with components of the online learning system and that this will lead to success in creating superior student learning [6]. Such a task is not an easy one to complete. For instance, integration of quality information resource like digital library, another important information resource available on the Internet which may consist of several repositories into an e-learning environment requires considerable knowledge-building on the part of the organization involved [7]. According to Wang, digital library should [7]:

- Include all those learning resources which are relevant
- Classify the resources into logical categories
- Develop a knowledge vocabulary, including a thesaurus
- Create indexes and search mechanisms
- Constantly refine the classification categories.

Digital collections (in digital libraries) should respond to the users' needs and provide an appropriate balance between quality and demand [7]. Those who seek solutions for integration of digital institutional repositories into the learning process at their university could evaluate experiences of those universities which have done the same but with digital libraries.

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 [8]. They presented three categories of criteria for evaluation of online learning objects for inclusion in a digital repository [8]:

### 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

When a review of a learning object based on these criteria is positive; it can be included in the digital repository. Even learning objects which receive a negative review can be still included in a repository and have a status same as unrefereed scientific article have in science.

## IV. COMPARISON OF TOP 10 LARGEST OPEN ACCESS REPOSITORIES

Open access digital information repositories are among the most important and valuable information resources in the world. The expansion in number of available papers they store and improvement of their user interfaces are evident almost on a daily basis. To maintain quality of their organizational structure and content, they are frequently evaluated. In theory, a user interface of an online information resource should be designed with the end user in mind. Users' information needs are then mapped to task requirements and the information system built around these mappings [9]. The most important task is information seeking. Marchionini divided information seeking process into five functions [9]: define the information problem, select the source, articulate the problem, examine the results and extract information.

Online information resources such as digital institutional repositories can be evaluated by use of methods intended for evaluation of information retrieval systems. Most of these methods are oriented towards information retrieval and functions 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 [10]: database selection, query formulation, query reformulation, access to help function, organization and display of results and delivery of results.

Criteria suggested and used by Xie and Cool will be expanded for use in the digital repository comparison in this paper:

## Category 1. Access

- Browsing capabilities
- Searching capabilities
  - i. Simple
  - ii. Advanced

## Category 2. Query formulation

- Simple
- Complex (AND, OR, NOT operators)
- Reformulation

## Category 3. Help

- General
- Contextual

## 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. User interface language choice

The aim of this comparison is to show how similar user interfaces and functions of the largest open access digital repositories are. These similarities can help their integration into learning managements systems and can facilitate easier training of students about how to use these digital information repositories (in general) for research and learning 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 an online information resource can be very helpful when accessing the content of another and new information resource on the Internet.

Criteria divided in six categories will be applied to 10 largest open access digital repositories. According to Registry of Open Access Repositories (ROAR) at <http://roar.eprints.org> on January 26th 2009, the top 10 largest open access repositories (with the largest number of records) were:

- PubMed Central (<http://www.pubmedcentral.gov>),
- Networked Digital Library of Theses and Dissertations Union Catalog (<http://alcme.oclc.org/ndltd/index.html>),
- RePEc: Research Papers in Economics (<http://www.repec.org>),
- CiteSeer.PSU OAI (<http://citeseer.ist.psu.edu>),
- arXiv.org e-Print archive (physics, mathematics, related fields) (<http://arxiv.org>),
- Zentrales Verzeichnis Digitalisierter Drucke: Hauptseite (<http://www.digitalisierterdrucke.de>),
- Humanities Text Initiative (<http://www.hti.umich.edu>),
- Australian Research Repositories Online to the World: ARROW Discovery Service (<http://search.arrow.edu.au>),
- Library of Congress: American Memory (<http://memory.loc.gov>)
- National Institute of Informatics (<http://ju.nii.ac.jp>).

The research was carried out by accessing each digital repository separately and comparing its functions with the list of enumerated criteria in six categories. Each criterion will be applied to every repository on the list and results will be presented in tables.

DIGITAL REPOSITORIES AND POSSIBILITIES OF THEIR INTEGRATION INTO HIGHER EDUCATION

TABLE I.  
CATEGORY: ACCESS

	Browsing	Searching	
		Simp.	Adv.
PubMed Central	Yes	Yes	Yes
Networked Digital Library of Theses and Dissertations Union Catalog	No	Yes	No
RePEc: Research Papers in Economics	Yes**	Yes**	Yes**
CiteSeer.PSU OAI	No	Yes	No
arXiv.org e-Print archive	Yes	Yes	Yes
Zentrales Verzeichnis Digitalisierter Drucke: Hauptseite	Yes	Yes	Yes
Humanities Text Initiative	Yes	Yes*	Yes*
Australian Research Repositories Online to the World: ARROW Discovery Service	Yes	Yes	Yes
Library of Congress: American Memory	Yes	Yes	No
National Institute of Informatics	Yes**	Yes**	Yes**

\* Depends on choice of a collection  
\*\* Depends on choice of a repository

TABLE III.  
HELP

	General	Contextual
PubMed Central	No	No
Networked Digital Library of Theses and Dissertations Union Catalog	No	No
RePEc: Research Papers in Economics	Yes*	No
CiteSeer.PSU OAI	No	No
arXiv.org e-Print archive	Yes	No
Zentrales Verzeichnis Digitalisierter Drucke: Hauptseite	No	No
Humanities Text Initiative	Yes	No
Australian Research Repositories Online to the World: ARROW Discovery Service	No	No
Library of Congress: American Memory	Yes	No
National Institute of Informatics	Yes**	No

\* Depends on choice of a database  
\*\* Depends on choice of a repository

TABLE II.  
QUERY FORMULATION

	Simple	Complex	Reformulation
PubMed Central	Yes	Yes*	Yes
Networked Digital Library of Theses and Dissertations Union Catalog	No	No	No
RePEc: Research Papers in Economics (IDEAS Web site)	Yes**	Yes**	Yes**
CiteSeer.PSU OAI	Yes	No	Yes
arXiv.org e-Print archive	Yes	Yes	Yes
Zentrales Verzeichnis Digitalisierter Drucke: Hauptseite	Yes	Yes	Yes
Humanities Text Initiative	No	Yes	Yes
Australian Research Repositories Online to the World: ARROW Discovery Service	Yes	Yes	No
Library of Congress: American Memory	Yes	No	No
National Institute of Informatics	Yes	Yes	Yes

\* Selective use of AND and OR in search for authors and AND, OR and NOT for article types  
\*\* Depends on choice of a database

TABLE IV.  
ORGANIZATION AND DISPLAY OF RESULTS

	Sorting	Limiting number of results
PubMed Central	Yes	Yes
Networked Digital Library of Theses and Dissertations Union Catalog	No	No
RePEc: Research Papers in Economics	Yes	Yes
CiteSeer.PSU OAI	Yes	No
arXiv.org e-Print archive	No	No
Zentrales Verzeichnis Digitalisierter Drucke: Hauptseite	Yes	No
Humanities Text Initiative	Yes*	Yes*
Australian Research Repositories Online to the World: ARROW Discovery Service	Yes	No
Library of Congress: American Memory	No	No
National Institute of Informatics	No	No

\* Depends on choice of a collection

TABLE V.  
DELIVERY OF LIST OF RESULTS

	File	Printer	Clipboard	E-mail
PubMed Central	Yes	Yes	Yes	Yes
Networked Digital Library of Theses and Dissertations Union Catalog	No	No	No	No
RePEc: Research Papers in Economics	No	No	No	No
CiteSeer.PSU OAI	No	No	No	No
arXiv.org e-Print archive	No	No	No	No
Zentrales Verzeichnis Digitalisierter Drucke: Hauptseite	No	No	No	No
Humanities Text Initiative	No	No	No	No
Australian Research Repositories Online to the World: ARROW Discovery Service	Yes	No	No	Yes
Library of Congress: American Memory	No	No	No	No
National Institute of Informatics	No	No	No	No

TABLE VI.  
USER INTERFACE LANGUAGE CHOICE

	Language Choice
PubMed Central	No
Networked Digital Library of Theses and Dissertations Union Catalog	No
RePEc: Research Papers in Economics	No
CiteSeer.PSU OAI	No
arXiv.org e-Print archive	No
Zentrales Verzeichnis Digitalisierter Drucke: Hauptseite	No
Humanities Text Initiative	No
Australian Research Repositories Online to the World: ARROW Discovery Service	No
Library of Congress: American Memory	No
National Institute of Informatics	Yes*

\* Depends on choice of a institutional repository / collection

## V. DISCUSSION

The results of the comparison show that researched digital repositories demonstrated similar approach in organization of their user interfaces and functions, differing only in the number of functions offered to a user. Most differences were found in category of organization and display of results ("Table IV."), where repositories offered different approaches to the manner of organization of screens presenting search results. Delivery of list of results ("Table V.") indicated that researched repositories are oriented more towards the direct delivery of the content itself rather than to the richness of functions for display and delivery of the list of search results. Two additional disadvantages that were discovered were lack

of multilanguage user interfaces in 9 out of 10 repositories, and lack of help, especially lack of contextual help ("Table VI."). Despite differences and drawbacks, all repositories were found to be easy to use.

## VI. CONCLUSION

The abundance of available content on the Internet is a problem which cannot be solved easily. This content is not always organized in the same manner or organized at all, so a part of it is almost useless. Scientists and experts in information sciences, computer science and pedagogy are constantly discovering new and viable solutions for collecting, organizing, storing and usage of vast quantities of content (research data, articles, conference papers, books etc.). At the present moment, at least two solutions draw attention of the general public: digital libraries and digital (institutional) repositories (i.e. digital archives). Each of them offers different solutions to the already mentioned problems of collecting, organizing, storing and usage of digital content. Since all of the content in these digital information resources is accessible on the Internet (except in some cases in which repositories require authentication of its users either by use of an IP address or by use of a login and a password), they are usable at various levels of education depending on the complexity of the topic and quality of presentation of the content itself. As a result of the certain number of similarities found in digital repositories comparison, some of them can be easily integrated into learning management systems or learning process as quality and well organized external information resources. Students would be then given instructions how to use the available content in open access digital repositories carefully selected by the teaching staff. The principal criteria for their inclusion in the learning process at universities are: their quality, permanence (as a source of information), long term policies for preservation, open access to stored content and ease of use. Judging on these factors, it is not unrealistic to expect that digital information resources such as digital institutional repositories will be used even more in future for various purposes including education.

## REFERENCES

- [1] R. K. Johnson, "Institutional repositories: partnering with faculty to enhance scholarly communication", *D-Lib Magazine*, vol. 8, no. 11, November 2002. [Online]. Available: <http://www.dlib.org/dlib/november02/johnson/11johnson.html>. [Accessed Jan 20, 2009]
- [2] H. Hockx-Yu, "Digital preservation in the context of institutional repositories", *Program: electronic library and information systems*, vol. 40, no. 3, pp. 232-243, 2006. (doi:10.1108/00330330610681312)
- [3] M. Pinto and D. Sales, "Knowledge Transfer and Information Skills for Student-Centered Learning in Spain", *Libraries and the Academy*, vol. 8, no. 1, pp. 53-74, 2008.
- [4] "Budapest Open Access Initiative", Feb. 14, 2002. [Online]. Available: <http://www.soros.org/openaccess/read.shtml>. [Accessed Dec. 15, 2008]
- [5] T. Anderson, "Towards a theory of online learning" in *The Theory and Practice of Online Learning*, 2nd ed., Edmonton: AU Press, 2008, pp. 45-74.
- [6] C. Patel and T. Patel, "Exploring a Joint Model of Conventional and Online Learning Systems", *e-Service Journal*, vol. 4, no. 2, pp. 27-46, 2005. (doi:10.1353/esj.2006.0010)
- [7] M. Y. Wang, "The strategic role of digital libraries: issues in e-learning environments", *Library review*, vol. 52, no. 3, pp. 111-116, 2003. (doi:10.1108/00242530310465915)

- [8] G.P. Schell and M. Burns, "Merlot: A Repository of e-Learning Objects for Higher Education", e-Service Journal, vol. 1, no. 2, pp. 53-64, 2002.
- [9] G. Marchionini, "Interfaces for end-user information seeking", Journal of the American society for information science, vol. 43, no. 2, pp. 156-163, 1992. ([doi:10.1002/\(SICI\)1097-4571\(199203\)43:2<156::AID-ASIS3.0.CO;2-U](https://doi.org/10.1002/(SICI)1097-4571(199203)43:2<156::AID-ASIS3.0.CO;2-U))
- [10] H. Xie, and C. Cool, "Ease of use versus users control: an evaluation of Web and non-Web interface of online databases.", Online information review, vol. 24, no. 2., pp. 102-115, 2000. ([doi:10.1108/14684520010330265](https://doi.org/10.1108/14684520010330265))

AUTHORS

**R. Vrana** is with the Department for Information Sciences, Faculty of Humanities and Social Sciences in Zagreb, Ivana Lučića 3, 10000 Zagreb, Croatia (e-mail: [rvrana@ffzg.hr](mailto:rvrana@ffzg.hr)).

This article was modified from a presentation at The 32nd International Convention MIPRO on information and communication technology, electronics and microelectronics., in 25-29 May 2009, Opatija, Croatia. Manuscript received, 1<sup>st</sup> February 2010. Published as resubmitted by the authors 1st March 2010.