

Institutional Repositories: Review and Knowledge Management Perspective

Maizatul Akmar Ismail, Mashkuri Yaacob & Sameem Abdul Kareem

Faculty of Computer Science and Information Technology
University of Malaya, 50603, Kuala Lumpur
Tel : 03-79676300, Fax : 03-79579249
E-mail : {maizatul, mashkuri, sameem}@um.edu.my

ABSTRACT

This paper provides a review on current literatures on the concept Institutional Repository (IR) and to propose the use of KM theory to provide understanding on the implementation of IR. We will discuss the issues relevant to IR from the KM view by providing examples which indicated that research can be build on the knowledge gained by KM researchers to augment the understanding of IR. This paper will contribute in bringing together recent research in IR and KM and how these two concepts can collaborate in resulting better understanding of IR implementation.

Keywords Institutional Repository (IR), Knowledge Management, Institutes of Higher Learning(IHL) , Knowledge Artifact

1.0 INTRODUCTION

Over the last two decades, the access to the Internet has tremendously changed the way of communicating the information to the end user. Scholars in particular, are among those users who benefits from the advent of the World Wide Web (WWW) where the process of assembling and disseminating the information became easier. Within this, an Institutional Repository (IR) became a free platform for knowledge sharing among the academia around the globe. IR is defined by Foster and Gibbons (2005) as “an electronic system that captures, preserves and provide access to the digital work products of a community”. IR is also demarcated as digital archive of intellectual product of university’s member i.e. the academicians and students and accessible to users within and beyond of the university (Poornima et. al, 2006). IR is also described by SPARC (2002) as a “digital collection that captures and preserves the intellectual output of a single university or multiple institutions”. Further, Poynder (2005) summed up the literature in IR by stating that “as a repository for a university’s research output, with the aim of increasing access to that research, and so enhancing its impact”. Albeit the varieties of definition given by IR’s expert, we are particularly interested with the working definition by Lynch (2003) whom defined IR as “a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members”.

Despite the difference of school of thoughts in defining IR, it could be deduced that IR are perceived as service and also an institution. As IR is regards as a new field of research, it seems to be lacking in a principle that could guide its formation. Thus, this paper aims to improve the establishment of IR by providing it with theoretical underpinning of Knowledge Management. This paper is organized as follows: First, it will give some overview on the research in IR. Next, we will highlight the issues and challenges in IR research. Subsequently, we will discuss about IR in KM perspectives. Finally, we will conclude the paper with discussion and future direction.

2.0 RESEARCH ON INSTITUTIONAL REPOSITORY

IR effort appeared to be an attempt to Open Access (OA) which is defined as “to read, download, copy, distribute, print, search or link to the full text of articles which are freely available on the Internet” (Willinsky, 2006) or in online repository supported by academic institution. With a properly plan IR, it will certainly have an impact to publishing behavior among academicians, researchers and students. Among the factors that influence towards encouraging electronic publishing is to get academic recognition, to increase the visibility of publication by letting the research work to be cited by others and to provide authors with central and permanent archive of their work (Zainab, 2006).

According to Ware (2004), IR is constituted by the following criteria:

- a Web-based database (repository) of scholarly material;
- institutionally defined (as opposed to a subject-based repository);
- cumulative and perpetual (a collection of record);
- open and interoperable (e.g. using Open Archive Initiative compliant software);
- collects, stores and disseminates scholarly material as part of the process of scholarly communication

Although the concept of IR is relatively new in Malaysia, two universities have started the initiative by implementing the so-called Knowledge Management

System (KMS- which concern with explicit knowledge) where the academicians are required to store their publication and merits are given based on the quantity and quality of the uploaded publication (Chua and Ismail, 2005). Amongst the factor that encourages academicians and staff to deposit their work in KMS is the reward system enforced by the university's authority. However, the resources are not freely available (limited to the university's member).

The existence of the first IR in Malaysia is tracked to the early year 2007 where the repository is filled with scholarly materials from the university members. It is a crucial step in preparing a platform where it can be utilized by students for research purposes. DSpace and E-Prints have been identified as a dominant package of IR software (Lynch and Lippincott, 2005). The structuring of IR content is identified as by document type and by subject area (Pickton and McKnight, 2006). Next step is to get the content into the institutional repositories (Westrienen, Lynch (2005)). Among the identified contents of IR are as shown in Table 1.0. From the table, the top most content is the theses, preprint/e-prints, conference proceeding and journal articles. In Malaysia however, the tendency to publish the "prepublished" or preprint articles are still low as perhaps because of the reasons posited by Crow (2002) : "they might fear plagiarism or anticipate copyright or other acceptance problems in the event they were to submit the work for formal publication. They might also fear the potential for criticism of work not yet benefiting from peer review and editing". Users are expected to deposit these material themselves yet the biggest hindrance of IR existence is the reluctance of the author to self-archive their work because it is time consuming and reluctance to accept new technology. In view towards populating the IR with the right content, another issue which is the interoperability arises especially to the content that is previously stored in the form of databases. Little have been said about IR software that could manage integration of multiple scholarly databases that are previously developed by journal publishers such as Malaysian Journal of Computer Science (MJCS) and Malaysian Journal of Library and Information Science (MJLIS) under the platform of Electronic Journal of University of Malaya (EJUM).

Table 1: Types of IR Content (Source: Lynch and Lippincott, 2005)

Type of Content		
Theses/dissertations	Digitized institutional assets from library special collections	Campus Blogs
Preprints/e-prints	Digitized institutional assets from museum collections	Newsletters
Conference proceedings	University publications	Laboratory Protocol
Conference	University	Exhibition Guide

Presentations, e.g. PPT slides	electronic records	
Tech reports/working papers	Departmental materials or records	Book Manuscript
E-books	Digital images	Web Pages
Journals	Digital audio	Student papers other than theses or dissertations
Newspapers (born digital)	Digital moving images	E-portfolios
Data sets	Digitized musical scores	Course content, e.g. syllabi, lectures
Interview Transcripts	Exhibitions	Learning objects
Maps	Performances	
Plans/blueprints	Software	

Ware (2004) stated that IR possesses several benefits which include:

1. A solution to weaknesses of current local self-archiving;
2. Providing a long term solution for preserving scholarly content
3. Improvement of scholarly communication
4. Expanding the content of disciplinary repositories such as ArXiv and CogScience
5. Indirectly improving research and teaching

As the IR is sometimes referred to Digital Library (DL), it is wise to distinguish these two terms. The separation line between IR and DL are vague, as stated by Lynch and Lippincott (2005) as their study shows that there are "confusing relationships at many institutions among digital libraries, digital research collections and collections of materials in institutional repositories, and the ways in which all of these relate to the scholarly communications process". They further speculated that "a key distinguishing characteristic of digital collections and digital libraries is one of institutional rather than faculty-initiated accession and organizational efforts". We tried to draw more visible line between these two concepts as depicted in Table 2 below.

Table 2: Difference between Institutional Repository and Digital Library

Features	Institutional Repository	Digital Library
Content Management	Can be deposited by both user and system administrator.	Librarian will collect the content and deposit it on behalf of the user.
User Involvement	User can actively involve in managing the academic content and share it to others.	User act as passive reader and have no ability to manage the content
Collaborative Content	Allow collaboration between users – for e.g. reviewing pre-print academic articles	Does not allow users' collaboration
Types of resources	Academic and scholarly resources.	Can be academic and non-academic content. Support community based content such as old manuscript.

3.0 ISSUES AND CHALLENGE

Despite various literatures which govern the concept of IR, the fundamental theory that governs the formation of IR in the IHL seems to be lacking. We perceived that literatures and theories under the Knowledge Management umbrella is the most suitable body of knowledge that could assist further understanding of IR concepts which will lead to the successfulness of its formation. The challenge is to find the most suitable framework in KM that would make the description of IR clear and precise.

4.0 IR IN KNOWLEDGE MANAGEMENT PERSPECTIVES

The rapid growth of data and technologies trigger the transformation of data to useful information, known as knowledge. Nowadays, the academic society are more aware on the importance of knowledge and ways to acquire, recognize, capture, retrieve, use or measure, manage and share the knowledge in the form of knowledge artifacts as stated in Figure 1.0 above. The term 'Knowledge Management' (KM) is created for to resemble the process involve in managing the knowledge. Tiwana (2000) has stated that KM consists of explicit knowledge and tacit knowledge. KM is a new field especially in higher learning institutions. A survey by Hijazi and Kelly (2003) has indicated that higher learning institutions and the business world have some difficulty in understanding the implication of KM as a new model to support business process. From the research done, the higher learning institutions are no longer just providing knowledge to the students, but also have to manage and share the existing knowledge for future reference (Ismail and Chua, 2005).

The key concepts of knowledge management could be used to help understanding towards the formation of successful IR :

1). Tacit vs Explicit Knowledge – tacit knowledge refers to the knowledge that resides in human brain and difficult to be captured (Nonaka, 1991) such as teaching and research skills. Tacit knowledge also includes academia's experience, skills, "know-how", beliefs, values, perceptions and judgment or opinion. This knowledge is considered personal to each particular individual. Therefore, it is difficult to transferred or shared, expressed and communicated to the other party involved. As a result, the academia may face difficulties in formalizing this type of knowledge into formal documents. Their ideas and experiences may be captured in audio and graphic format for later used. On the other hand, explicit knowledge is the type of knowledge that exists in the form of documents, reports, design, blueprints, models, patterns, new rules, mathematics equations plan or formulate books, databases and text are known as explicit knowledge. This knowledge can be package, codified and transferred easily. Besides, it can be also be

communicated, shared and expressed easily with formal language.

2) Knowledge Artifact: Krupansky (2006) defined knowledge artifact as "*an artifact which represents an encoding of knowledge*". In the IR context, knowledge artifact is the output of research work done by a researcher in particular IHL. We emphasized that research output is explicit knowledge artifact inline with Holsapple (2003) which stated that it can be transferred to one person (or software agents) to another. The process of producing knowledge artifacts in IR involve the combination of the explicit knowledge in the form of documented data that gathered through survey or scientific experiment and tacit knowledge stored inside researchers' mind in the form of beliefs and research experience. For example, experience researcher would know that the best way to analyze a particular type of quantitative data is by using suitable statistical analysis techniques with the right analytical procedures. The term *knowledge assets* and *knowledge artifacts* will be use interchangeably in this paper.

3) *Community of Practice (CoP)* – Brown and Duguid (1991) posits that knowledge "*flows best through networks of people who may not be in the same part of the organization, but have the same work interests*". In IR, CoP includes academicians and scholars in various levels in IHL which involve in the business of manufacturing knowledge artifacts such as research articles in the form of journal, conference proceedings, thesis, technical reports, poster presentations etc

KM is a process where organizations have formulated ways in the attempt to recognize and archive knowledge assets within the organization that are derived from the employees of various departments or faculties and in some cases, even from other organizations that share the similar area of interests or specialization (Joseph, 2001) Besides, it is defined as the process of transforming information and intellectual assets into enduring value. It also connects people with the knowledge that they need to take action, when they need it (Kidwell et.al, 2000). In view to this, IR has become a valuable knowledge store which could be utilized by knowledge worker such as the researcher to produce new knowledge artifacts. Moreover, as KM is concerned with making the right knowledge available to the right processor such as human or computer, at the right time in the right presentations for the right cost (Holsapple and Joshi, 1999), IR software should be equipped with functional features that could make it utilization worthwhile. A survey and interview by Ismail et.al. (2007) concluded on the needs-and-wants of students as early staged researcher and the motivational factors that would encourage them to use the IR as part of the research process. They further stated that among the factors that will motivate students to search in IR are free to use, which are rated the highest by both groups, followed by efficient keyword search, user friendly interface and full-viewed display of the content. It is

interesting to observe that emerging research trend and expert identification are also marked by more than half of the respondents as motivating factors. All IHLs inherently store, access, and deliver knowledge in some manner. The question is what value is added to the products and services they deliver by the effective use of that knowledge capital and worker. Kidwell et.al (2001) explains that higher education institutions have “significant opportunities to apply knowledge management practices to support every part of their mission.” Other research have stated that some IHL have adapted to their changing role in a knowledge-based society (Metaxiotis and Psarras, 2003) and recognize the value of their intellectual capital to their continuing role in society (Rowley, 2000).

To assist in further understanding of how KM theory could be used to implement a successful IR, we utilized the framework proposed by Ismail and Chua (2005) as guidelines. This framework structure serves as the foundation and provides the fundamental of KM to be deployed in an environment specific to the Higher Learning Institution, and in our context, the IR implementation.

Due to the limitation of space, we will only focus on the middle and right section of the frame structure as the rest of the framework have been discuss in great detail in Ismail and Chua (2005). Firstly, the transformation of raw data into knowledge assets or artifacts. Yet, before data could be converted to knowledge, it will need to be converted to information. It will need to undergo the 5C’s Filter which consist the process of condensation, calculation, contextualization, correction and categorization (Davenport and Prusak, 1998), (Tiwana, 2002).

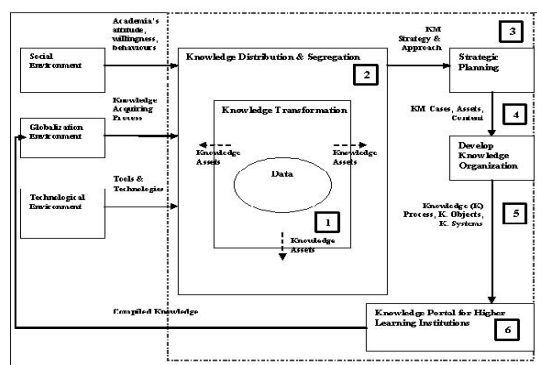


Figure 1: Overview of KM Framework in IHL (Source : Ismail & Chua, 2005)

This information will later be transformed into knowledge. Knowledge worker, in the context of IR will be the academician and students whom doing a research work for carrier advancement and merits of academic achievement will make extensive uses of knowledge assets in the form of scholarly content to assist the research work. Graziano and Raulin (2000) stated that

there are seven stages of research procedures which need to be closely follow in order to produce high quality research output. The research output in the form of knowledge artifacts or assets will then be distributed and segregated to other researchers around the globe. In the IR milieu, the knowledge segregation and distribution means publishing the researchers work (academia and students) to the international platform for other scholars to appreciate and indirectly gaining recognition for the IHL where they are working with. Such respect is useful as it determine the prestigious of one institution in which the research output acts as a performance indicator. This knowledge possessed by particular IHL also represents the strategic resources that will create competitive advantages (King and Marks, 2006). With a repository well in place, the IHL seems to be well equip to be in a “World class University” provided that it follows the right KM strategy and approaches. Boyer’s (1990) view in determining academic excellence focus on four aspects of scholarship i.e. discovery, teaching, application, and integration of knowledge. Weiser (1996) pinpoints three basic forms of scholarship (as stated in Virginia Polytechnic Institute and State University’s Website, 2007), namely: 1) *Discovery* of new knowledge; 2) *Development* of new technologies, methods, materials, or uses; and 3) *Integration* of knowledge leading to new understanding.

Next, IHLs need to embark on a Strategic Planning, which is a method to determine the vision, scope and objective of the KM initiative and identify the best methods to achieve their desired goals. Thousands of IRs have been implemented since the value of research artifacts been realizes as valuable assets that could open another door of new knowledge discovery. Open archives repositories such as ArXiv and CogScience are examples of famous disciplinary-based repositories which allow scientist and researchers from the same discipline to have some sort of knowledge sharing activity among their members. The fact that some IHL in Malaysia realize the importance of IR and implement one of their own proves that the IHL management has envisaged the need for such repository. Subsequently, the generated results that consist of KM cases, knowledge content and knowledge assets are used as input to “Develop Knowledge Organization” strategy for the development of KM system structure. The ideal knowledge organization as posited by Awad and Ghaziri (2004) is one where “people exchange knowledge across the functional areas of business by using technology and established processes”. In the context of IR, the KM assets are in the form of scholarly content as depicted in Table 1.0. These databases will be used to populate the resources for the development of knowledge organization. In this sense, IHL which have chosen to develop their own IR have taken the initiative in developing the knowledge organization as a way to capture, preserve and provide access to the digital work products of a community (Foster and Gibbons, 2005). The knowledge artifacts resulted from the formation of knowledge organization will be use as an input to the

knowledge portal where it could be aligned to the existing infrastructure of the IR system itself. The most important criteria that became a motivational factor for the utilization of KM system is "Perceived Usefulness" (Clay et.al, 2005). Their research also posited that the loyal used of KM system will increase with the improvement in factors such as System Quality, Ease of Use and Relevant Content.

5.0 DISCUSSION AND CONCLUSION

IR is newly crafted term which are lacking on theoretical underpinning that could govern its implementation. This paper tried to complement theory and research in KM which can be use to assist further understanding on IR. Scholarly content became the most valuable knowledge artifacts in IHL. What is required perhaps in the implementation of IR which seldom be taken into account is the need-and-want of another major stakeholder i.e. the students as this is the group of people that appear to be a reader and contributor to the IR content (Pickton and McKnight, 2006). The phrase "build it and people will come" is no longer valid. Further research is to map the architecture of KM system to suit the ideal formation of IR system.

REFERENCES

- Boyer, E. (1990). *Scholarship Reconsidered: Priorities of the Professoriate*. Princeton, NJ: The Carnegie Foundation for the Advancement of Teaching.
- Brown, J.S., and Duguid, P. (1991), Organizational Learning and communities of practice: toward a unified view of working, learning, and innovation. *Organization Science*, 2, 1 40–57.
- Clay, P.F., Dennis A.R., Ko, D.G, (2005), Factors Affecting the Loyal Use of Knowledge Management Systems, Proceedings of the 38th Hawaii International Conference on System Sciences, IEEE pp 10
- Davenport TH., and Prusak, L (1998), Working Knowledge: How organizations Manage What They know, Harvard Business School Press, Boston
- Foster, N.F., Gibbons, S. (2005), "Understanding Faculty to improve content recruitment for Institutional repositories", *D-Lib Magazine*, Vol. 11 No.1 at url: www.dlib.org/dlib/january05/foster/01foster.htm
- Gerard van Westrienen, Clifford A. Lynch (2005) Academic Institutional Repositories, D-Lib Magazine Volume 11 Number 9 ISSN 1082-9873
- Graziano A.M, and Raulin M.L, (2000), Research Methods: A Process of Inquiry, 4th Edition, Allyn and Bacon, Boston
- Hijazi, S. & Kelly, L. (2003). Knowledge Creation in Higher Education Institutions: A Conceptual Model. *Florida Keys Community College William Seeker Campus* at url: <http://fits.depauw.edu/ascue/Proceedings/2003/p78.pdf>
- Holsapple, C.W, (2003), Handbook of Knowledge Management, Springer, ISBN 3540435271 pg 306
- Ismail M.A, Chua L.Y, Implication of Knowledge Management in Higher Learning Institution, *International Conference on Knowledge Management (ICKM'05)*, 7th – 9th July 2005, Kuala Lumpur, Malaysia (in CD ROM)
- Ismail M. A, Mashkuri Yaakob, Sameem Abdul Kareem (2007), Semantic Search Engine in Institutional Repository : An Ontological Approach, Proceedings of the International Conference on Libraries, Information and Society ICoLIS 2007 Kuala Lumpur, Malaysia. pp.55-63. ISSN 97898343491-0-3
- Joseph M. F (2001). Key Issues in Knowledge Management. *Knowledge and Innovation: Journal of the KMCI*. At url: <http://www.kmci.org/media/firestoneissueskiv1n3.pdf>
- Kidwell, J., Linde, V., Karen M., and Johnson S.L, (2000) "Applying Corporate Knowledge Management Practices in Higher Education." In Bernbom, Gerald, editor, *Information Alchemy: The Art and Science of Knowledge Management. EDUCAUSE Leadership Series #3. San Francisco: Jossey-Bass*. 1-24.
- Krupansky, J (2006), Foundation of Software Agent Technology at url: http://www.agtivity.com/def/knowledge_artifact.htm
- Lynch C. (2003) "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age." *ARL Bimonthly Report*, No. 226, at url: <http://www.arl.org/newsltr/226/ir.html>
- Lynch, C and Lippincott, J.K, (2005), Institutional Repository Deployment in the United States as of Early 2005, DLib Magazine Volume 11 Number 9, ISSN 1082-9873 <http://dlib.org/dlib/september05/lynch/09lynch.html#Lynch>
- Nonaka I. The knowledge-creating company. *Harvard Business Review*, 79, 6 (1991),96–104.
- Pickton, M and McKnight, C., (2006), Research Students and the Loughborough Institutional Repository, *Journal of Librarianship and Information Science*, 38 (4), pp 203- 219
- Poornima. N, B.S. Biradar, and I.R.N. Goudar, (2006) Institutional Repositories in India: A Case Study of National Aerospace Laboratories. *ICADL 2006*: pp 533-534
- Poynder, R. (2005), "Institutional repositories", *American Scientist Open Access Forum*, Vol. 23
- Raym Crow, (2002) *The Case for Institutional Repositories: A SPARC Position Paper*, available at <http://www.arl.org/sparc/IR/ir.html>

- Tiwana, A. (2002). *The Knowledge Management Toolkit: Orchestrating IT, Strategy, and Knowledge Platforms* (2nd Edition). Prentice Hall. ISBN: 013009224X.
- Ware, M. (2004) Institutional repositories and scholarly publishing Learned Publishing, Volume 17, Number 2, 1 April 2004 , pp. 115-124(10)
- William R. King and Peter V. Marks, Jr (2006), Motivating Knowledge Sharing Through a Knowledge Management System, *International Journal of Management Science*, Volume 36, Issue 1, February 2008, pp 131-146
- Weiser, C. J. (1996). The value of a university-- Rethinking scholarship. Oregon State University at <http://www.adec.edu/clemson/papers/weiser.html>
- Willinsky J. 2006. The access principle: the case for open access to research and scholarship. Cambridge, Mass.: MIT Press.
- Zainab Awang Ngah. (2006). *Scholarly Skywriting: E-Print Archives and E-Journals, Panacea or Problem?*, Inaugural Lecture, Faculty of Computer Science and Information Technology, University of Malaya.