Long-term Accessibility and Re-use of Institutional Repository Contents of Some Selected Academic Institutions in Nigeria

Ezra Shiloba Gbaje
University Library, Federal University Lokoja, Kogi State, Nigeria
E-Mail address: shiloba@yahoo.com; ezra.gbaje@fulokoja.edu.ng

Monsurat Funmilola Mohammed
Department of Library and Information Science, Ahmadu Bello University, Zaria, Kaduna State, Nigeria
E-mail address: mohammedmonsurat38@gmail.com

Abstract:

The increase awareness of open access, growth in digital asserts, need for accessibility and the quest for global visibility by academic institutions in Nigeria has resulted in the rapid deployment of Institutional Repositories. These Institutional Repositories contains valuable institutional memories such as thesis, dissertations, pre-print and post-print scholarly articles, conferences papers and lots of other institutional memories that need to be preserved for the future generation. Purposive sampling was use to select five institutions that have deployed Institutional Repository with a view to examine the strategies they have in place for the long-term preservation of the contents. The findings of the study revealed that non-of these academic institutions has put in place any strategy to ensure the long-term re-use or accessibilities of the contents of their Institutional Repository. The paper concluded that back-up of Institutional Repository periodically was not adequate enough to ensure the long-term accessibility and re-use of the contents. Recommendation for adoption and modification of Open Archival Information System (OAIS) reference model as a strategy for digital preservation was made. This model defines the basic functional components of a system dedicated to the long-term preservation of information and could be applied to the long-term preservation of the content of an Institutional Repository. The proposed changes in the service models focused on four of the OAIS functional entities: the ingest of digital object, management of digital objects, storage of digital objects, and access, while the functions of the remaining two entities, Administration and Preservation Planning, can be performed by the staff of the Institutions.

Keywords: Long-term, accessibility, digital preservation, Institutional Repository, OAIS, Nigeria, Academic Libraries, re-use.

Introduction

Part of the core mandates of academic institutions is teaching and research, during these processes digital contents such as seminar papers, conference papers, technical reports, datasets, theses and dissertations, pre-print and post post-print journal articles, images, audio and video contents are created. These digital contents have grown tremendously with the increase use of computers and Internet in information production. In a bid to increase accessibility, global visibility and effective management of these digital contents academic institutions have deployed Institutional repositories. The investment in the deployment of institutional repository by academic institutions in Nigeria, for the purpose of managing their digital objects and ensuring accessibility for long-term re-use is threatened by the rapid and constant changes in computer software and hardware technologies. Digital objects regardless of whether they are ingested into an Institutional Repository or not are threatened by hardware and software obsolescence (Rosenthal, 2005). Supporting this assertion Hockx-Yu (2006), posited that the hardware and software used to store and access digital information are constantly upgraded and superseded. To achieve the long-term preservation of the digital contents deliberate efforts must be made to secure long-term access to these digital contents (Lawrence, 2000; Granger, 2002; Macdonald and Lord, 2003; Rosenthal, et al, 2005; and Gladney, 2007; Gbaje, 2011). Preservation in the digital environment means to maintain authenticity, reliability, and functionality of digital objects over long-term, so that they can still be accessed and understood through changes in technology. Li and Banach (2011) defined digital preservation as the act of physically and intellectually protecting and technically stabilizing the transmission of the content and context of electronic records across space and time, in order to produce copies of those records that people can reasonably judge to be authentic.

A great deal of research on digital preservation issues is currently being conducted around the world. *Berthon and Webb* (2000) assert that the past several years have witnessed a significant increase in collaborative international efforts addressing a variety of questions on digital preservation. The United Kingdom Joint Information Systems Committee (JISC) has been one of the leading institutions worldwide in undertaking research and development on the long-term preservation of digital materials through projects and services such as the Arts and Humanities Data Services, CURL exemplars in Digital Archives (Cedars) Project, Creative Archiving at Michigan & Leeds: Emulating the Old on the New (CAMiLEON) Project, and the JISC/NPO Preservation research studies (Beagrie, 2004). Digital preservation activities have also been the focus of collaborative initiatives, because no single agency can undertake the role of preserving all digital objects within its purview. These initiatives include national strategic alliances like the Digital Preservation Coalition (DPC) in the UK and Nestor (Network of Expertise in long-term STORage) in Germany, as well as initiatives like the National Digital Information Infrastructure and Preservation Program (NDIIPP) in the US.

IR and Academic Institutions in Nigeria

According to the Nigerian Universities Commission website, there are currently 40 Federal, 44 State and 69 Private Universities. Scholars and researchers in these institutions generate large volumes of data and knowledge in the course of their job which is also criteria for their promotion. Scholarly publications of scholars and researchers in Nigeria are of great importance to solving local and regional problems as well as contributing to the body of knowledge (Gbaje, 2009). It can therefore be argued that over the years scholarly outputs produced within the academic institutions are digital-born while non-digital are being converted to digital format. The long-term preservation of these digital outputs is seen as responsibility for stewardship of intellectual and cultural heritage. To ensure wider dissemination and digital preservation of these scholarly outputs many universities in Nigeria have deployed Institutional repositories.

Statement of Problem

Academic institutions deploy institutional repository to manage and preserve their digital contents for long-term re-use, however, studies have revealed that because digital files have been ingested into the IR, does not mean that it can be accessed for re-use over a long-term (Robertson and Borchert, 2014). It was also observed that most academic institutions carry out back-up activities of their IR databases which are not enough to preserve the digital content for long-term because they are designed to be short-term solutions. However, ensuring the long-term re-use of the digital contents ingested into the repository requires preservation action. Hockx-Yu (2006) posited that the speed of changes in technology means that the timeframe during which preservation action must be taken is very much shorter than for paper. The researcher noted with concern the misconception amongst institutional repository mangers that digital contents will be preserved for long-term re-use simply because they are ingested into the Institutional repository. This study therefore seeks to identify the strategies put into place by some selected academic institutions in Nigeria to ensure the long-term access and re-use of the digital contents ingested into their IR. To achieve this the following research questions:

1. What strategies are put in place for the preservation of the contents of the selected Institutional Repository ?

2. What are the challenges encountered in the preservation of the contents of the selected Institutional Repository ?

Literature Review

Institutional repositories can be used to collect, preserve and provide free, unrestricted access to all types of institutional digital content. *Lynch (2003) defined* Institutional repository (IR) as a set of services that an institution offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members. It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation and facilitating access to digital asset. Most IRs are designed to accept and facilitate access to all forms of digital contents including text, images, video and audio files and ensure their availability to future generations in its original digital form. However, IRs are dependent upon software and hardware technologies, and Lynch (2003) asserted that a key part of the services that comprise an institutional repository is the management of these technological changes, and the migration of digital content from one set of technologies to the next as part of the organizational commitment to providing repository services. Lynch (2003) further argued that an effective institutional repository of necessity represents collaboration among librarians, information technologists, archives and records mangers, faculty, and university administrators and policymakers.

Li and Banach (2011) reported that even though some IR software are gradually integrating support for preservation, it was not sufficient to rely only on software since various facets have to be considered when preserving digital content. Supporting this assertion Fyffe et al.(2004) posited that long-term preservation consists of specific functions, such as ingesting digital objects and their metadata, storing these objects and retaining the link to their associated metadata, monitoring technological obsolescence, and migrating the objects over time. There is no doubt that preservation of digital contents of IR will require people, institutions, software applications, technological infrastructure, and procedures as well as control over what is ingested.

Digital Preservation Strategy and Policy

Over the last few years, a lot of efforts were put into defining, improving, and evaluating digital preservation strategies. Digital preservation strategies are methods for keeping stored digital objects permanently accessible for long-term re-use. Implementing a digital preservation strategy is a crucial part of managing the risk associated with rapid hardware and software obsolescence. Implementation strategies for digital preservation can be categorized into investment, short-term and long-term strategies. Digital preservation strategies are methods used for keeping stored digital objects permanently accessible for long-term re-use. The strategies adopted by organizations studied depend largely on the file format and infrastructures.

File Format

The basic premise of preservation as it relates to digital information is the ability to provide meaningful access to file format of digital objects. Pearson and Webb (2008) observed that file format obsolescence is related to the development of new format encodings that take the place of ready existing format in the market place and changes in the availability and support of rendering software. As a result of prominence of file format in digital preservation the OCLC developed INFORM (INvestigation of FOrmats based on Risk Management), for investigating and measuring the risk factors of digital formats and providing guidelines for preservation action plans (Stanescu, 2004). The United Kingdom National Archives also built a repository called PRONOM which contains information about file formats, software products, and other technical components for long term access to digital objects of cultural, historical and business value (Reis and Lindley, 2007). The risk assessment will determine what environments are needed for the most reliable access to the content and what actions need to be taken on a regular basis to ensure preservation of the digital objects. In 2006, National Library of Australia in collaboration with the Australian National University (ANU) built the Automatic Obsolescence Notification System I (AONS I) prototype. The system was designed to analyse the digital repositories and determine whether any digital objects contained therein may be in danger of becoming obsolescent. It uses preservation information about file formats and the software which supports these formats to determine if the formats used by the digital objects are in danger (Curtis, 2006). In order to determine this, the AONS I system used information obtained from the PRONOM and Library of Congress Sustainability of Digital Formats (LCSDF) registries, which it periodically checked against the contents of the repository. When the repository was found to contain objects in danger of becoming obsolescent, a notification report was sent via email to the repository manager. At the conclusion of the AONS I Project, the software code was supported in a DSpace. Webb and Pearson (2008:96), assert that in 2007 the NLA and other Australian Partnership for Sustainable Repositories (APSR) partners collaborated in the AONS II software development project, to refine and expand the functionality of the prototype AONS I software.

Digital Preservation Policy

According to the Electronic Resource Preservation and Access Network Digital Preservation policy tool (2003), the primary aims of a digital preservation policy are to provide guidance and authorization on the preservation of digital materials and to ensure the authenticity, reliability and long-term accessibility of them. Digital preservation policy is directly connected to the aims and goal of the institution. It clearly states the principles and long-term direction that would guide preservation strategies and actions.

Digital preservation policy should explain how digital preservation can serve major needs of an institution and state some principles and rules on specific aspects which then lay the basis of

implementation. Beagrie et. al.(2008:3), proposed a model framework and guidance for the development of digital preservation policy and implementation clauses.

IR and OAIS Reference Model

Open Archival Information System (OAIS) Reference model was published 2002 by the Consultative Committee for Space Data Systems (CCSDS), to deal with long-term preservation. (Borghoff et al 2006). The term open refers to the fact that the reference model was developed and released in an open public forum, in which any interested party was encouraged to participate. An archival information system is "an organization of people and systems that have accepted the responsibility to preserve information and make it available for a Designated Community" (ISO 14721, 2003).The Reference model's main objectives are to clarify basic terms of long-term preservation and to identify the corresponding key processes. Although it originated from space research, the OAIS Reference Model is a general, theoretical model describing the organization of an archive. It examined the problem of long-term digital preservation from an information and process model point. The OAIS model also provides a comprehensive framework for all functions required for digital preservation including ingest, storage, retrieval, and long-term preservation of digital objects. At a very general level, it can be seen that IR provide a similar range of functionality as found in OAIS input and output, data management, and storage. The OAIS imposes more formality and discipline on these processes for the purpose of long-term preservation. It examined the problem of long-term digital preservation from information and process model viewpoints, as discussed below:

The Information Model

The information Model of the OAIS Reference Model provides a fundamental distinction between Data object and Information object. The digital object is the focus of preservation, along with metadata necessary to support its long-term preservation and access, bound into a single logical information package (International Organization for Standardisation 2003). Hence, the information model is built around the concept and structure of information as it moves into, though, and out of the archival system. Data turns into information only if the corresponding Knowledge and additional Representative Information are available. The knowledge which is required to understand a data object is called Knowledge Base in the OAIS Reference Model. Borghoff et al. (2006) assert that unavailability of the Knowledge Base can result in the inability to interpret the data. There is also the need for additional information to understand the data object, which is subsumed by the Representation Information in the OAIS Reference Model. Representation Information might include a description of the hardware and software environment needed to render the Content Data Object and/or access its contents (Lavoie 2004). The OAIS Reference Model models these processes and groups them according to the following:

1. Ingest Process: This process receives a Submission Information Package (SIP) and prepares it for storage and administration within the archive. The SIP is transformed into Archival Information Packages (AIPs) and Descriptive Information corresponding to the AIP created. Finally the AIP is passed onto the Archival Storage process and the corresponding Descriptive information to the Data Management process.

2. Storage Process: This process is responsible for the storage, management, maintenance of the bit-stream and the retrieval of the AIPs. The process ensures the periodic refreshing of storage media and guarantees the reconstruction of the AIPs in case of system failure.

3. Data Management Process: This process manages the Descriptive Information and the data necessary to run the system. The data are saved in a database and can be queried for editing and updating.
4. Administration Process: The Administration process handles routine work in the archive. The process checks whether the delivered SIPs meet the archive's standard. It is also responsible for the hardware and software architecture of the archive.
5. Preservation Planning Process: This process monitors the environment and provides recommendations to ensure the long-term accessibility of the stored information. This includes monitoring of the technology and designed community and evaluation of the archive and periodical recommendations on archival update for migration.
6. Descriptive Information is information that supports discovery and retrieval of Content Information by the Designated Community through the OAIS's finding aids. For example, Descriptive Information takes the form of a Dublin Core metadata record maintained by the OAIS to facilitate resource discovery by the consumer. The OAIS Reference Model is a conceptualization of the environment, functional components and information objects associated with a system designed to effect the long-term preservation of digital materials. Lavoie (2004) asserts that the OAIS Reference Model provides nothing about system architectures, storage or processing technologies, database design, computing platforms or any of the myriad technical details involved in setting up a functioning archival system.

The OAIS Reference Model only articulates the requirements to be met by archive and library systems (Borghoff et al. 2006). This makes the adoption of OAIS very flexible by any institution and in different ways.

Methodology
Data were collected from five out of ten Federal Universalities that have deployed Institution repositories as listed on opendoar.org. The five universities selected for this study were purposively selected because they have functional Institutional repository and have the necessary human capacity and technological infrastructure. The subjects of this study were university librarians or repository managers who were most likely individuals to be aware of and involved in IR activities in the universities. A survey method was used to collected data from the five universities using survey monkey.

Table 1: Sampled Academic Institutions

S/No	Name of University	IR URL
1.	Ahmadu Bello University Zaria	http://kubanni.abu.edu.ng:8080/jspui/
2.	University of Jos	http://irepos.unijos.edu.ng/jspui/
3.	University of Ibadan	http://ir.library.ui.edu.ng
4.	University of Nigeria Nsuka	http://repository.unn.edu.ng:8080/xmlui/
5.	University of Ilorin	http://uilspace.unilorin.edu.ng:8080/jspui/

Discussions and Results of finding
Strategies are put in place for the preservation of the contents of the selected Institutional Repository. The researchers sought to identify the motivation for the deployment of institutional repository in each of the institutions and the responses collected revealed that all the institutions sampled indicated that the main purpose of deploying IR was for the purpose of increasing the visibility of their digital contents and also to preserve their digital content for long-term re-use.

This is in agreement with Hockx-Yu (2006) who posited that preservation and access cannot be separated. The data collected further revealed that there was no digital preservation activities such as migration, emulation, and Universal Virtual Computer carried on the Institutional repository to ensure their long-term preservation and re-sure. The back-up of the IR databases carried out by the selected institutions is not in any way sufficient enough to ensure the long term-preservation as supported by McGath (2012).

File Format for Ingest
The file format used to create digital objects for digital preservation plays a vital role in the future rendition of the digital object. Hence, the research sought to find out the file formats ingested into the IR for digital preservation. The responses obtained revealed that the original file format Adobe PDF file format were the most popular file format used to archive digital objects in their repository. The finding corroborates Borghoff et al's (2006) position that PDF and TIFF are the de facto standards for digital document. It can also be argued that the archived digital objects are products of the institution's digitization process, which results in PDF file formats in most cases. The use of original file format as file format for ingest into the institutional repository implies that same software requirement must be met to access the digital content some of which might be obsolete in the long-term thereby making the contents inaccessible for re-use.

The reasons for the choice of file format for ingest by the institution was also sought and the respondents revealed that facilities played a major role in the choice of file formats used for archiving the digital objects. It was also indicated that institutional policy, where available, played a role in the choice of file format for digital archive.

Strategies for Assessing Digital Objects for Preservation Action
Digital preservation action needs to be taken when the digital object has been assess to be in danger of either technology or software obsolescence. Data collected from the sampled institutions revealed that the only mechanism put in place was mechanism to check file format before ingest into the IR and no mechanism to monitor and provide warning on the need for preservation action before ingest of the digital objects. A mechanism therefore that will assess the risk of a digital object needs to be put in place.

Challenge of Preserver Digital Contents of IR
The study also sort to identify challenges associated with the preservation of IR contents from the sampled institutions and the responses collected indicated lack of institutional policy on digital preservation, lack of awareness of the challenges surrounding digital preservation and lack of technical know-how were the major challenges with the preservation of IR contents. This confirms Beagrie (2006) assertion that the development of institutional preservation policies is currently sporadic and digital preservation issues are rarely considered in key strategic plans. This therefore affects the level of commitment in terms of finance, staff, capacity building and digital preservation policy formulation and implementations.

Conclusion
Academic institutions in Nigeria deploy Intuitional Repository as a strategy of increasing their visibility and preservation of their digital objects. All the sampled institutions assumed that creation of backup of their IR databases was sufficient to ensure their long-term accessibility and re-use. Unfortunately, evolution in information technology has resulted to rapid hardware

and software obsolescence, hence Keeping digital resources for use by the future generations will require conscious efforts and continual investment. It is clearly established that preserving digital objects for long-term re-use should include policies and strategies for Assessing Digital Objects for Preservation Action. Such strategies could be imbibing actions for digital objects and technologies like the inclusion of a mechanism for risk assessment before they are ingested for long term preservation and accessibility. A working action for the development and implementation of a digital preservation policy should also be put in place.

Recommendation

In the context of this work the use of computer hardware and Dspace Institutional Repository Software is recommended for inclusion in the OAIS reference model. OAIS Reference Model provides nothing about system architectures, storage or processing technologies, database design, computing platforms or any of the myriad technical details involved in setting up a functioning archival system. The Models only articulate the requirements to be met by digital archive and library systems (Borghoff et al., 2006). The proposed framework is depicted in figure below:
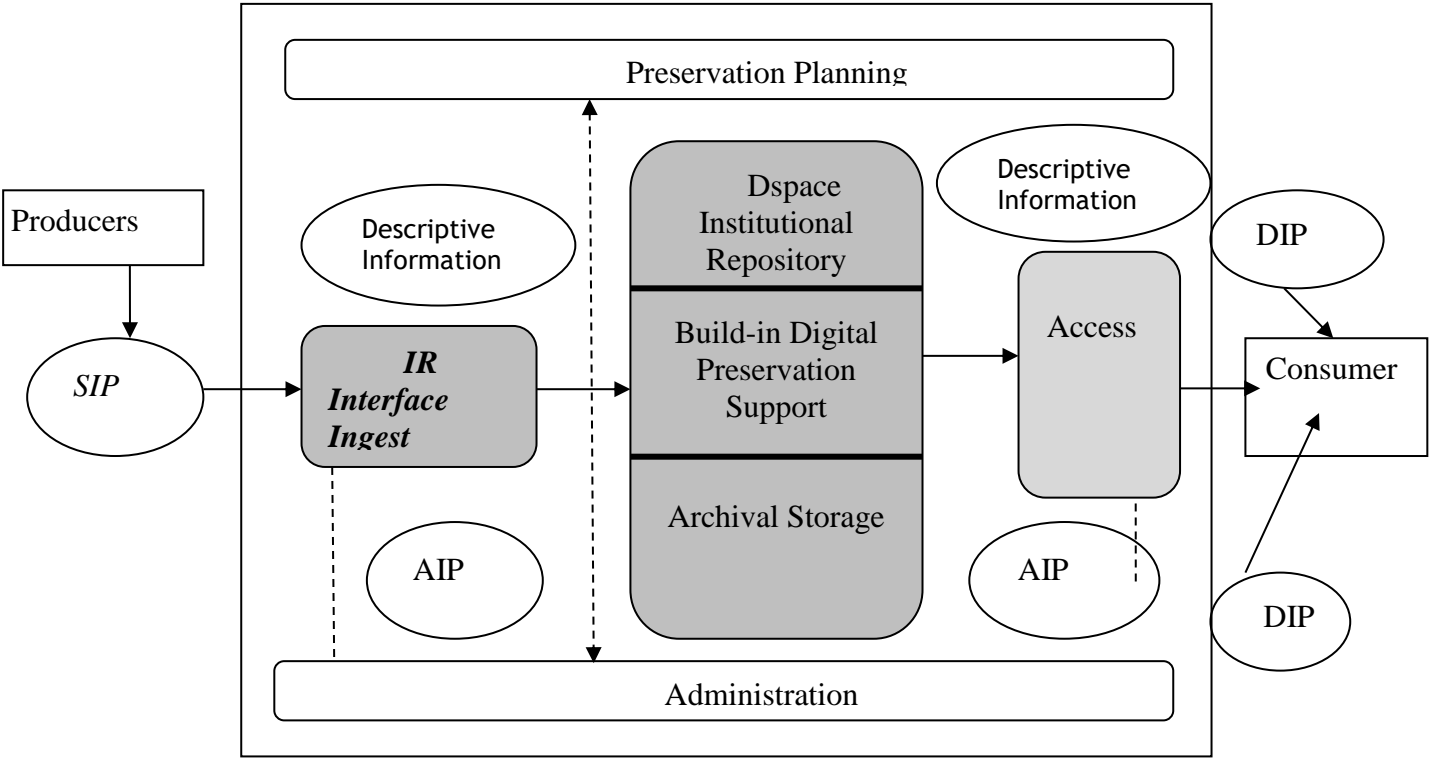

Figure 1: Proposed Academic Institutions Digital Preservation Strategy

The changes in the service models are shown in the dark-grey rectangle with the highlighted text and are all focused on four of the OAIS functional entities: the ingest of digital object, management of digital objects, storage of digital objects, and access. The relationships are shown by the connecting arrows. The functions of the remaining two entities, Administration and Preservation Planning are performed by the staff of institutions library and will also be responsible for the development and implementation of digital preservation policy. The activities of preservation planning are to monitor the hardware and software markets, and to check whether all preserved documents can be rendered on newly developed systems.. The Administrative entity goes further to decide whether to update or to migrate the Institutional Repository contents If they cannot it must decide which strategy (eg migration or emulation) will guarantee best future accessibility of the document under the new circumstances.

References

Beagrie et al.(2008). Digital Preservation Policies Study. A study funded by JISC. Available at http://www.jisc.ac.uk/Home/publications/documents/jiscpolicyfinalreport.aspx (accessed on Retrieved on 8/7/2017)

Borghoff et al (2006). Long-Term Preservation of Digital Document: Principles and Practices. Springer-Verlag Berlin Heidelberg.

Chandra S and Gokhale A.P(2012). Implementing Open Archival Information System Model for Digital Preservation at Indian Institute of Geomagnetism. *DESIDOC J. Lib. Inf. Technol.*, 2012, 32(4)

Curtis, J. (2006). *AONS system documentation revision 169*. 2006-09-29. Canberra, Australia: Australian National University. Available on http://www.apsr.edu.au/publications/aons_report.pdf (accessed 2/6/2017)

Electronic Resource Preservation and Access Network(2003). Digital Preservation Policy Tool. Available at *www.erpanet.org/guidance/docs/ERPANETPolicyTool.pdf (accessed on* 2/06/2017).

Fyffe R, Ludwig D., Roach M., Schulte B., and Warner B.F, (2004). *Preservation Planning for Digital Information: Final Report of the HVC2 Digital Preservation Task Force* (Lawrence: University of Kansas), available at https://kuscholarworks.ku.edu/bitstream/handle/1808/166/Preservation%20Planning%20for%20Digital%20Information.pdf?sequence=1&isAllowed=y , accessed 25/07/2017.

Gbaje E.S. (2009). Global Visibility of Scholarly Journal Published in Nigeria. *Nigerian Libraries: Journal of Nigerian Library Association.* Vol. 42. 32-44

Gbaje E.S (2011). Digital Preservation Strategies: A Case Study of Nigerian National Information Centres. ***International Federation of Library Association and Institutions (IFLA) Journal*** Vol.37(3). 218-227.Available at http://www.ifla.org/en/news/out-now-october-2011-issue-of-ifla-journal. Accessed on 31/10/2011

Hitchcock S., Brody T., Hey J.M.N. and Carr L.(2005) Preservation for Institutional Repositories: practical and invisible At *Ensuring Long-term Preservation and Adding Value to Scientific and Technical Data (PV2005)*. *21 - 23 Nov 2005*. 9 pp. Available at http://www.ukoln.ac.uk/events/pv-2005/pv-2005-final-papers/033.pdf (accessed 24/08/2017)

Hockx-Yu H. (2006) "Digital preservation in the context of institutional repositories", Program, Vol. 40 Issue: 3, pp.232-243, https://doi.org/10.1108/00330330610681312

ISO 14721(2003). *Space data and information transfer systems Open archival information systems– Reference model (ISO 14721:2003).*

Lavoie. F.B.(2004). The Open Archival Information System Reference Model: Introductory Guide. *OCLC Online Computer Library Center, Inc. and Digital Preservation Coalition. Available from http://www.dpconline.org/docs/lavoie_OAIS.pdf (accessed on 5/7/2017)*

1.1.1   Li Y. and Banach M. (2011). Institutional Repositories and Digital Preservation: Assessing Current Practices at Research Libraries. D-Lib Magazine Vol. 17(5). Available at http://www.dlib.org/dlib/may11/yuanli/05yuanli.html. (accessed 9/7/17)

Lynch, C. A (2003). "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age" ARL, no. 226 (February 2003): 1-7. Available on http://www.arl.org/resources/pubs/br/br226/br226ir.shtml (Accessed on 6/06/2017)

McMillan, G. (2004). Digital preservation of theses and dissertations through collaboration. Resources Sharing & Information Networks, 17 (1-2), 159-174

Pearson D. and Webb C. (2008). Defining File Format Obsolescence: A Risk Journey. The International Journal of Digital Curation. Vol. 3(1). Pp91

Reis M. and Lindley A. (2006). Digital Long-term Preservation. Research Studios Austria-Studio Digital Memory. Available at http://sat.researchstudio.at/werkstatt/wp-content/uploads/2007/08/reis_lindley_digital_long_term_preservation.pdf (accessed 14/06/2017)

Robertson C.W. and Borchert A.C.(2014). Preserving Content from Your Institutional Repository. Serial Librarian Vo. 66:278-288. Available at http://ir.uiowa.edu/cgi/viewcontent.cgi?article=1251&context=lib_pubs (accessed 9/7/17)

Stanescu, A.(2004). Assessing the Durability of Formats in a Digital Preservation Environment: The INFORM Methodology. D-Lib Magazine Vol. 10(11). Available at http://www.dlib.org/dlib/november04/stanescu/11stanescu.html#6 **(accessed 9/6/17)**