Fundamental concepts in digital preservation

Laerte Pereira da Silva Júnior
University of Porto
Faculty of Letters
Via Panorâmica, 4150-564
Porto, Portugal
+351 226 077 100
laerte.psjunior@gmail.com

Maria Manuel Borges
University of Coimbra
Faculty of Letters
Largo da Porta Férrea, 3004-530
Coimbra, Portugal
+351 239 859 900
mmb@fl.uc.pt

ABSTRACT
The growing availability of digital information in repositories requires a new profile of information professional or curator, one who, among other aspects, master the fundamental concepts regarding digital preservation. In this paper, through a revision of the literature, we describe the bit preservation and functional preservation as fundamental concepts of the digital preservation problem. We recommend that public institutions should develop strategies to provide adequate training to raise this knowledge and awareness in order to assure that everyone who is involved in the production and curation of the digital preservation must have the profile needed to fulfil this function.

CCS Concepts
• **Social and professional topics~Information science education**

Keywords
Bit preservation; Functional preservation; Digital preservation.

1. INTRODUCTION
The Internet and the World Wide Web are the fundamental infrastructure to Open Science by allowing the sharing of research both with researchers and society at large. Projects like OpenAIRE, shows that this Open Science will continue to grow: more than 14.000.000 publications and 18.000 datasets from more than 5000 repositories and open access journals [1]. For Borgman [2], "strategic investments in knowledge infrastructures, such as data repositories, human resources with expertise in data management, better tools, and methods to provide credit for data sharing, may increase the release and use of data".

It is important to stress "open Science is an umbrella term encompassing a multitude of assumptions about the future of knowledge creation and dissemination" [3]. Regarding repositories, their diversity also mean that their commitment to digital preservation can assume distinct aspects: "backing up data in their submitted form can be expensive, but sustaining them over long periods of time requires much larger investments in migration to new technologies and formats as they appear. (...) The degree in to which they invest in verifying data content and structure, augmenting datasets with metadata and provenance documentation, or providing other value-added services varies considerably" [2].

Regardless of the greater or lesser degree of commitment that each repository assumes, the need to form digital preservation experts or curators have been referred to by several projects and initiatives (e.g. InterPARES, SCAPE, DigCurv, Digital Preservation Coalition, DPOE). In this sense, the categorization of basic theoretical aspects are essential to give an overview of the digital preservation issue for those who are willing to learn this field of knowledge, still in consolidation phase. This work aims to describe two basic theoretical concepts to apprehend digital preservation that any digital curator should master: bit preservation and functional preservation. The methodology consisted of a literature review from the concepts systematized in the Catalogue of Preservation Policy Elements – an SCAPE project's deliverable.

2. BIT PRESERVATION
PREMIS dictionary defines preserving the bit level as follows: Preservation strategy in which the sole objective is to ensure that a Digital Object remains fixed (unaltered) and viable (readable from media). No effort is made to ensure that the Digital Object remains renderable or interpretable by contemporary technology [4].

For a digital object remains unchanged, it is necessary to ensure its fixity[1] meaning that it does not suffer changes in a given time. In other words: the preservation ensures that the bit string of bits of a digital object remains intact and recoverable over time. However, the bit preservation is only a subset of aspects of digital preservation. In fact, it is the starting point of all activities of digital preservation. For this reason, it is necessary to define it so that they include various aspects of digital preservation by means of a holistic approach to maintenance bit, which is defined as "[...] an approach where the bit preservation is seen as something that should be recognized as part of a whole, where different circumstances can influence how the bit preservation should be performed" [6]. In this case, the "whole" are all aspects of digital preservation, as illustrated in figure 1.

TEEM'16, November 02-04, 2016, Salamanca, Spain
 ISBN 978-1-4503-4747-1/16/11...$15.00
DOI: http://dx.doi.org/10.1145/3012430.3012530

[1] To understand the concept of fixity, refer to the report *Checking your digital content: what is fixity, and when should I be checking it?* Retrieved 12 July, 2016 from http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA-Fixity-Guidance-Report-final100214.pdf.

Figure 1: Holistic approach to the bit preservation Source: [6]

The holistic approach to bit preservation covers all bit preserving aspects and only some aspects of functional preservation. These two sets of features influence the choice of routes to ensure the preservation bit. It is still observed in figure 1 that this approach also includes other aspects of digital preservation that do not belong to the bit preservation or functional preservation (e.g. sustainability). The holistic approach is not another way to describe digital preservation, but a way to include aspects that influence the bit preservation.

At one point, the bit preservation should be ensured by a bit preservation solution. The holistic approach focuses on the requirements for such a solution when the bits of the preserved material and the solution found is part of a whole. These requirements may be equivalent to the requirements for an organizational subsystem chosen as a bit preservation solution. The subsystem is structured as a bit repository. There are several examples arising from "whole" to show that the bit preservation by the holistic approach is not restricted to the integrity and legibility of bits, such as: the aspects of information security - integrity, availability and confidentiality, according to ISO 27000; the cost requirements; the representation of aspects of digital material preserved at the bit level, among others.

The holistic approach to bit preservation aims to contribute to optimize a way to preserve the bit, based on three results Zierau research [6]: a model to define the bit preservation, separately from other aspects of digital preservation; a concept for digital material representations; a methodology to evaluate the choice of different bit preservation solutions. The type and importance of digital material for a given collection are decisive for the choice of conservation actions. Thus, one may adopt a digital preservation model levels, such as in table 1 below, to scale the bit preservation strategy associating it with other essential elements of the "whole".

Table 1 – levels of digital preservation

Storage and Geographic Location	
Level 1	Two complete copies that are not collocated; For data on heterogeneous media (optical discs, hard drives, etc.) get the content off the medium and into your storage system
Level 2	- At least free complete copies. - At least one copy in a different geographic location. - Document your storage system(s) and storage media and what you need to use them
Level 3	- At least one copy in a geographic location with a different disaster threat. - Obsolescence monitoring process for your storage system(s) and media
Level 4	- At least three copies in geographic locations with different disaster threats. - Have a comprehensive plan in place that will keep files and metadata on currently accessible media or systems.
File Fixity and Data Integrity	
Level 1	- Check file fixity on ingest if it has been provided with the content. - Create fixity info if it wasn't provided with content
Level 2	- Check fixity in all ingests. - Use write-blockers when working with original media. - Virus-check high risk content.
Level 3	- Check fixity of content at fixed intervals - Maintain logs of fixity info; supply audit on demand - Ability to detect corrupt data - Virus-check all content
Level 4	- Check fixity of all content in response to specific events or activities - Ability to replace/repair corrupted data - Ensure no one person has write access to all copies
Information Security	
Level 1	- Identify who has read, write, move and delete authorization to individual files - Restrict who has those authorizations to individual files
Level 2	- Document access restrictions for content.
Level 3	Maintain logs of who performed what actions on files, including deletions and preservation actions
Level 4	Perform audit of logs
Metadata	
Level 1	- Inventory of content and its storage location. - Ensure backup and non-collocation of inventory
Level 2	- Store administrative metadata. - Store transformative metadata and log events
Level 3	- Store standard technical and descriptive metadata
Level 4	- Store standard preservation metadata
File format	
Level 1	When you can give input into the creation of digital files encourage use of a limited set of known open formats and codecs
Level 2	- Inventory of file formats in use.
Level 3	- Monitor file format obsolescence issues.
Level 4	- Perform format migrations, emulation and similar activities as needed

This is a digital preservation model that excludes issues related to rights and/or the policies and restricted to technical issues. It was created by the National Digital Stewardship Alliance in order to help those interested in access to digital information in the long term to assess their practices to minimize the risk of loss, also by helping them to identify the next steps that should be taken in order to move all operations, or part of them, to the next level. These transactions are the activities along the rows levels. When there is a movement of the rows, from level 1 to level 4, such operations are moving from the basic need of bit preservation toward more comprehensive requirements in order to track the digital content and to make possible its accessibility for a long period of time.

Table 1 provides the guidelines in five functional areas or categories that make up the core of digital preservation systems:

a) Storage and geographic location − The focus of this category is the storage of digital information. Each level change adds up new copies. Thus, the threats resulting from a degradation of bits or failures in the media or storage system are restricted. The diversification of geographic location in each level helps, in turn, to protect against threats to the storage system, either because of natural disasters or those caused by humans. On levels 2, 3 and 4, to enhance the longevity of the storage systems, the following requirements have been added: system documentation, obsolescence monitoring and accessibility plan. Thus simplifies the work, because the activities are performed in incremental steps.

b) File fixity and data integrity − The objective of this category is to provide a series of steps that will lead an organization to a stage where it is acting incisively to ensure fixity of its digital objects. At level 1 a fixity check[2] intake is required. If there is no fixity of information, it will be generate one. This process is essential to ensure that the preserved content is what really intended to preserve. The following levels lead to a continuous operation of fixity check, thus increasing confidence in the fixity of the content that has a high value to the preservationists of an organization.

c) Information Security − This category is intended to identify those who read, write, execute and delete a digital object, register 'logs' of manipulation's object and assign access restrictions. In this case, one can prevent, for example, the accidental deletion of a digital object. The levels are organized by increments of actions that meet the increasingly advanced and regulated requirements to minimize the risks that are also related to the perceived risks in other categories.

d) Metadata − In this category, additional layers of metadata are presented, which become the most protected content, identifiable and accessible. In most systems, almost all of these metadata (except descriptive) should be generated and processed exclusively by computer.

e) File format − Digital objects are subject to the structure and type of the file formats. Level 1 suggests that organizations seek to use open and popular file formats; Level 2 documents the use of formats; Level 3 monitors the obsolescence of formats; level 4 uses the migration, emulation of platform and scans other ways of ensuring that the contents will be preserved usable and accessible in the future.

It should be noted that the file formats are susceptible to bit errors. Different file formats will require different bit of preservation solutions in order to achieve the same maximization of risk prevention against loss of digital material [6]. The dictionary PREMIS recommends that the file format is determined on the repository intake process, since many preservation activities depend on detailed knowledge of the digital object format. The fixity check and virus check should also be made in intake, according to the guidelines of the fixity category file and complete data on the model of digital preservation levels. The model is in its first version, but serves to guide both those who are initiating actions to preserve their digital assets as the institutions that plan to improve their systems and workflows of digital preservation. It allows institutions to assess the level of preservation achieved for any given material in their custody; however, the model was not designed to evaluate the digital preservation programs in all its aspects, it does not address issues such as those related to conservation policy, personal or institutional support [7].

A digital object, as it is available on the Web, will be identifiable and accessible on a long-term basis, regardless of its Uniform Resource Locator (URL). A URL has the purpose of identifying a resource and describe its location, but may become inconsistent if the resource is moved to another location. For this reason, the use of persistent identifiers is considered the best solution to preserve access to digital resource, regardless of its URL, as the persistent identifier will be assigned to a new location when the resource is moved [8]. However, there are few strategies for the implementation of persistent identifiers and they depend on the technical, administrative and political institutions, their visions of the future and interest in interoperability with other systems. In particular, the strategies are as follows:

a) Redirection − is a minimum strategy, since it uses the standard features of the web server to redirect requests to the current position of the feature. This method is difficult to manage when it comes to large websites;

b) Installing a resolver supported by database – it assumes a link server software, running on a database and with the purpose to map the current location of the resource, i.e., the current URL. One option in this category is the PURL server software − Persistent URL − provided by the Online Computer Library Center (OCLC) (http://purl.oclc.org/) − (...);

c) Persistent identification system for hiring, offered by another organization − there are several persistent identification systems designed for use on the Internet, based on open standards, with different goals and approaches. For example: Digital Object Identifiers (DOI), Handle System and also PURL, since OCLC provides online identification service to third parties.[9].

Added to the above-exemplified systems the Uniform Resource Name and the Archival Resource Key. Persistent ID created for authors also has several deployment options, such as Author Claim, Scopus Author ID, Researcher ID, arXiv Author ID, ORCID. The working group Institutional Identifier of the National Information Standards Organization initiated the development of a pattern of persistent identification for the institutions. On the other

[2] Fixity check: a mechanism to verify that a digital object has not been altered in an undocumented manner; Checksums: message digests and digital signatures are examples of tools to run fixity checks; Fixity information: the information created by these fixity checks, provides evidence for the integrity and authenticity of the digital objects and are essential to enabling trust. Retrieved July 12, 2016 from http://www.digitalpreservation.gov/ndsa/ndsa-glossary.html.

hand, despite these initiatives demonstrate a growing awareness and interest in persistent identifiers, some difficulties continue to make persistent identification a complex problem, because different user communities can not guarantee the persistence of their identifier systems, specifically the resolvers.

For example, librarians, archivists, researchers, editors and funding agencies have different views and approaches to conceptualizing Persistent Identifiers and different business models, legal criteria, requirements and policies. Consequently, some identifier systems end up better address the needs of certain communities, but many of these particular solutions are widely used to meet specific requirements. This means that the discussion about Persistent Identifiers cannot be restricted to technical aspects to ensure persistent identification to digital resources, as it is to take into account the complexity of a range of responsibilities and requirements, which underlie the development and maintaining a system identifier. Each of these requirements involves the commitment of many stakeholders to maintain adequate infrastructure and ensure a consensus on policies, responsibilities, rights and duties. The difficulty in establishing the interoperability of these systems can reside on financial policy issues. How to create a globally unique identifier is far from being adopted, it remains the challenge of establishing a framework of interoperability between the known persistent identifiers solutions to enable persistent access, reuse and exchange of information through the use of existing identifiers and resources associated through different systems, locations and services. Therefore, no one would be dependent on a unique identifier system [6].

Sierman *et al.* recommend that an institution define what it understands about the bit preservation and relationship with functional preservation. This definition, stated in a digital preservation policy could include the actions described in table 1 according to the levels of preservation chosen for certain types of digital material. In order to ensure the performance of bit preservation, an institution should take some precautions in the intake process of this material: "ensure that the digital objects are free of viruses by examining the objects before ingest; ensure that the collection and the objects are complete; identify, characterise and validate formats" [10].

Likewise, it is important that the digital preservation policy includes the allocation of persistent identifiers to digital material − as they minimize the risk of losses caused by the lack of identification − Determine the number of copies of digital objects, the geographic distribution − this implies delegate the administration of the material distributed to another team − and have a contingency plan for disaster recovery.

3. FUNCTIONAL PRESERVATION

The functional preservation is the preservation of some or all functions of the original software environment. The preservation of bits will be useless if you cannot decode them and use the information. Therefore, the preservation of the functions of the original application will be the next level of conservation program. Still, it is not always necessary to preserve the entire application functionality that generated the information. For example, it will often be sufficient to preserve the ability to view and extract the information, but it need not be modified [12]. Functional preservation ensures the permanence of intelligibility and usefulness of the bits in accordance with the purposes of preservation, which are conditioned by institutional policy [6, 13]. The challenge of this type of preservation was already performing as a concern in the work of scholars, as Hedstrom [14], Kuny [15], Garrett and Waters [16]. Among the many reasons that contribute to the obsolescence of digital information, there are the rapid changes in the recoding media, the growing variety of files and integrated document features formats and running on different platforms, as evidenced by studies of Gladney [17], Rauch and Rauber [18] and Thibodeau [19]. However, several other aspects of functional preservation have been the subject of research over the past two decades, among which we highlight those related to the planning of preservation: "Preservation Planning, i.e. evaluating preservation strategies and choosing the most appropriate strategy, has turned into a crucial decision process, depending on both object characteristics as well as institutional requirements. The selection of the preservation strategy and tools is often the most difficult part in digital preservation endeavors; technical as well as process and financial aspects of a preservation strategy form the basis for the decision on which preservation strategy to adopt. The area of Preservation Planning has therefore attracted much interest in recent years" [20].

An example of this interest is the work of Becker *et al.* [21], in which the authors distinguish the digital preservation policy concepts and a preservation plan to then describe the elements needed to make an informed and complete plan.

There are several methodologies[3] for digital preservation, but three of them are considered the main ones: migration, emulation and preservation technology. [4]

a) Migration - This methodology consists of the periodic transfer of digital materials from one hardware / software configuration to another configuration, or a generation of computer technology to another generation. The purpose of migration is to preserve the integrity of digital objects and allow users to retrieve it, display them and use them, even in the face of constant technological change [16]. As the bit string is changed after migration, we must also preserve the authenticity of the object so that the strategy is successful [6]. In the migration, an original file format is received and below is changed to another format. The migration strategy to be used will depend on each case. For example, some file types, in a given collection may be more appropriate for a migration intake while others may be more appropriate for a migration based on aspects of risk. As shown by studies of Pearson and Pozo [22] and Lee *et al.* [23], this methodology presents more disadvantages than advantages, however, is one of the most used as ensure Zierau [6] and Ferreira *et al.* [24].

b) Emulation - This methodology is the process of creating a virtual version of the original environment that was used to access a particular object. The virtualized environment is accessed through an executable emulator on a modern hardware and

[3] Other strategies are added in the entry *Digital Preservation* at: http://en.wikipedia.org/wiki/Digital_preservation. This entry is maintained by the Digital Preservation project on Wikipedia, which, in turn, is coordinated by the Working Group Standards and Practices of NDSA of Library of Congress.

[4] Pearson and Pozo [22] developed a practical approach to methods of digital preservation, structuring it according to the following topics: "What the methodology is, or what it purports to do; How it works; Its perceived advantages and disadvantages; and Different strategies for approaching or maintaining the methodology."

software platform. This allows access to original content is maintained through the emulator without the content undergo changes. The emulation does not keep the same shape and performance of the original hardware. This may have implications that depend on the planning of preservation. When the emulation, by itself, does not allow adequate access to the content of digital materials, it is necessary to associate it with another method, such as the renderers. In emulation (Figure 2), a software environment / original hardware is encapsulated in an emulated environment (emulated environment). This new environment is installed on a contemporary platform hardware and software. Once installed on the new platform, it uses the environment emulated to access the desired object.

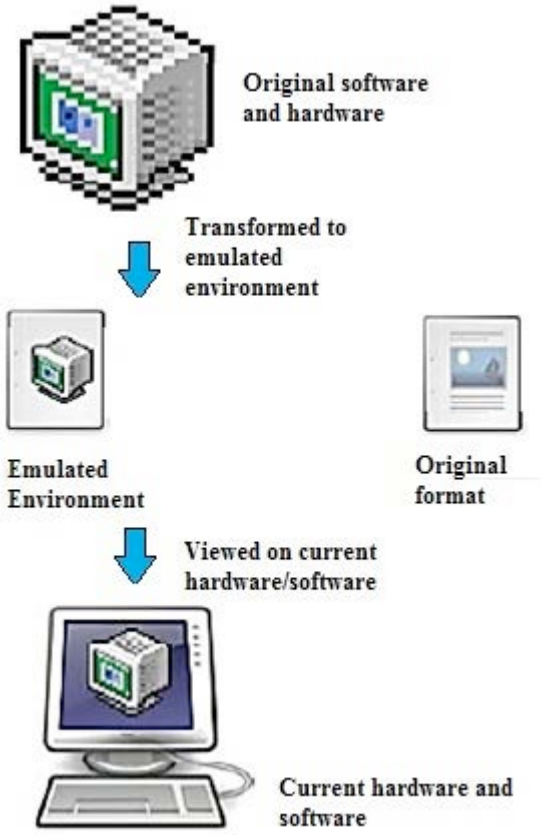

Figure 2: Emulation process. Adapted from [22]

c) Preservation of technology - this strategy is to preserve the software and hardware platform where the digital material was created or was accessed. According to Ferreira [25], it is the creation of a museum of technology, where the aim is to preserve the original digital object format. Proponents of this strategy consider it the only way to maintain the reliability of digital objects. The Handbook of Digital Preservation [11], Pearson and Pozo [22] list the advantages and disadvantages of this strategy. Among the advantages, we point out that it can be the only option to make reading digital materials. However, the hardware and software platforms are subject to extinction caused by technological obsolescence, which features the most significant disadvantage. As the Handbook of Digital Preservation [11], this strategy imposes the following requirements:

- Policies and guidelines on access.
- Hardware documentation and maintained software.
- Metadata required to maintain the hardware and software.

The two types of preservation discussed in this paper are related to other fundamental concepts such as authenticity, digital object, metadata, standards, access, organisation, and audit and certification. All this concepts, including rights, are defined in the guidelines of the Catalogue of Preservation Policy Elements [10], whose elements are the basis of Digital Preservation.

4. CONCLUSION

The growing availability of digital information in repositories, stimulated by mandates from public or private sector funding of research on national and European level (e.g. H2020 from the European Commission), asks for a new professional profile who must be aware of the problems that digital preservation brings. Open Science requires the availability of digital information to be stored in repositories in order to be shared and reused as a fuel for innovation processes. In this context, the discussion along the article tried to stress two fundamental aspects that any digital curator or information professional should master, bit preservation and functional preservation.

According the glossary of the Handbook of Digital Preservation [11], “Bit preservation is not digital preservation but it does provide one building block for the more complete set of digital preservation practices and processes that ensure the survival of digital content and also its usability, display, context and interpretation over time”. This “building block” has a certain complexity degree, as the holistic approach shows in the intersection with the bit preservation, functional preservation and digital preservation, that represents the ‘all’.

Bit preservation is not enough to turn the information of the digital object available at the long term. It is necessary to adopt other planned preservation actions that include the choice of the appropriated strategies.

We recommend that public institutions should develop strategies to provide adequate training in order to raise this knowledge and awaress to assure that everyone who is involved in the production and curation of the digital preservation must have the profile needed to fulfil this function.

5. ACKNOWLEDGMENTS

We thank the CAPES Foundation for granting full PhD scholarship abroad.

6. REFERENCES

[1] OpenAIRE. 2016. *Search*. Retrieved July 12, 2016 from https://www.openaire.eu/search/find.

[2] Borgman, L. C. 2015. *Big data, little data, no data: scholarship in the networked world*. The MIT Press.

[3] Bartling, S., and Friesike, S. 2014. Towards Another Scientific Revolution. In *Opening Science: The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing.* Sönke Bartling and Sascha Friesike (eds.). Springer Open. New York, NY, 3-16. DOI= dx.doi.org/10.1007/978-3-319-00026-8.

[4] PREMIS. 2015 *Data Dictionary for Preservation Metadata*. 3. ed. Washington, D. C. The Library of Congress. Retrieved July 12, 2016 from http://www.loc.gov/standards/premis/v2/premis-dd-2-0.pdf.

[5] Lavoie, B., Dempsey, L. 2004. Thirteen ways of looking at...digital preservation. *D-Lib Magazine*. 10, 7-8 (Jul./Aug. 2004). Retrieved 12 July, 2016 from http://www.dlib.org/dlib/july04/lavoie/07lavoie.html.

[6] Zierau, M. O. 2011. *A Holistic Approach to Bit Preservation*. Doctoral Thesis. Hvidrove : University of Copenhagen. Retrieved July 12, 2016 from http://www.diku.dk/forskning/phd-studiet/phd/thesis_20111215.pdf.

[7] Phillips, M., Bailey, J., Goethals, A., and Owens, T. 2013. *The NDSA Levels of Digital Preservation: An Explanation and Uses*. Retrieved July 11, 2016 from http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA_Levels_Archiving_2013.pdf.

[8] Bellini, E., Cirinnà, C., Lunghi, M., Bazzanella, B., and Bouquet, P. 2012. *Persistent Identifiers Interoperability Framework*. Technical Report, D 22.1. APARSEN. Retrieved July 12, 2016 from http://www.alliancepermanentaccess.org/wp-content/plugins/download-monitor/download.php?id=81.

[9] Sayão, L. F. 2007. Digital libraries' interoperability: The role of persistent identifiers' systems - URN, PURL, DOI, Handle System, CrossRef and OpenURL. *Transinformação*. 19, 1 (Jan./Apr. 2007), 65-82. DOI=dx.doi.org/10.1590/S0103-37862007000100006.

[10] Sierman, B., Jones, C., and Elstrøm, G. 2014. *Catalogue of Preservation Policy Elements*. Technical Report, D. 13.2. SCAPE. Retrieved July 12, 2016 from http://scape-project.eu/deliverable/d13-2-catalogue-of-preservation-policy-elements.

[11] Digital Preservation Coalition. 2015. *Digital Preservation Handbook*. Retrieved July 20, 2016 from http://handbook.dpconline.org/.

[12] Waugh, A., Wilkinson, R., Hills, B., and Dell'oro, J. 2000. Preserving digital information forever. In *Proceedings of the fifth ACM conference on Digital libraries*. DL '00. ACM, New York, NY, 175-184. DOI= http://dx.doi.org/10.1145/336597.336659.

[13] Smith, M., Barton, M., Branschofsky, M., McClellan, G., Walker, and J. H., Bass, M. 2003. DSpace: An Open Source Dynamic Digital Repository. *D-Lib Magazine.* 9, 1 (January, 2003). DOI= http://dx.doi.org/10.1045/january2003-smith.

[14] Hedstrom, M. 1998. Digital Preservation: A Time Bomb for Digital Libraries. *Computers and the Humanities.* 31 (1998), 189-202.

[15] Kuny, T. 1998. The digital dark ages? Challenges in the preservation of electronic information. *International Preservation News.* 17 (May 1998), 9-18.

[16] Garrett, J., Waters, D. 1996. *Preserving Digital Information*. Report of the Task Force on Archiving of Digital Information. Retrieved July 19, 2016 from https://www.clir.org/pubs/reports/reports/pub63watersgarrett.pdf.

[17] Gladney, H. M. 2007. *Preserving Digital Information*. Springer-Verlag.

[18] Rauch, C., Rauber, A. 2004. Preserving Digital Media: Towards a Preservation Solution Evaluation Metric. In *Proceedings of the 7th International Conference on Asian Digital Libraries* (Shanghai, China, December 13-17, 2004). ICADL 2004. Springer, Heidelberg, 203-212.

[19] Thibodeau, K. 2002. Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years. In *Proceedings of The State of Digital Preservation: An International Perspective* (Washington, D.C., April 24-25, 2002). DAI, 2002. CLIR, Washington, D.C., 29-38.

[20] Strodl, S., Becker, C., Neumayer, R., Rauber, A. 2007. How to Choose a Digital Preservation Strategy: Evaluating a Preservation Planning Procedure. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries* (Vancouver, Canada, June 17-22, 2007). JCDL'07. ACM, Vancouver, 29-38.

[21] Becker, C., Kulovits, H., Guttenbrunner, M., Strodl, S., Rauber, A., and Hofman, H. 2009. Systematic planning for digital preservation. *Int. J. Digit. Libr.* 10, 4 (December, 2009), 133-157. DOI= http://dx.doi.org/10.1007/s00799-009-0057-1.

[22] Pearson, D., Pozo, N. 2009. Explaining Pres Actions. Retrieved July 19, 2016 from http://pt.slideshare.net/natlibraryofaustralia/explaining-pres-actions.

[23] Lee, K., Slattery, O., Lu, R., Tang, X., and Mccrary, V. 2002. The State of the Art and Practice in Digital Preservation. *J. Res. Natl. Inst. Stand. Technol.* 107, 1 (Jan./Feb. 2002). 93-106.

[24] Ferreira, M., Saraiva, R., Rodrigues, E. 2012. *Estado da Arte em Preservação Digital*. Technical Report. RCAAP. URI= http://hdl.handle.net/1822/17049.

[25] Ferreira, M. 2006. Introdução à preservação digital – Conceitos, estratégias e actuais consensos. School of Engineering of the University of Minho.