



DIGITAL CURATION AND DIGITAL PRESERVATION: CONCEPTUAL INTERSECTIONS

CURADORIA DIGITAL E PRESERVAÇÃO DIGITAL: CRUZAMENTOS CONCEITUAIS CURADURÍA DIGITAL Y PREVERVASIÓN DIGITAL: CRUCES CONCEPTUALES

Thayse Natália Cantanhede Santos¹

ABSTRACT

Digital curation and digital preservation have as object of interest the maintenance of digital information over time. In the scientific literature, they have close definitions, however, digital curation is presented as a natural evolution of the preservation of digital resources management processes. Digital preservation can be understood as a method, process and model. Thus, it was found that there are points of intersection and differentiation elucidating and indicating why a concept appeared after the other and how each is presented in the face of the importance given to the management of large volumes of data, especially scientific data. After comparing the two concepts it was possible to conclude that there is more unity in the understanding of what curation is than there is in the preservation concepts. As a result, it is concluded that both concepts have fields of action that complement the management of digital information as a whole.

KEYWORDS: Digital Curation. Digital Preservation. Digital Information Management

RESUMO

A curadoria digital e a preservação digital têm como objeto de interesse a manutenção da informação digital ao longo do tempo. Tem na literatura científica definições próximas, contudo, a curadoria se apresenta como evolução natural dos processos de preservação de recursos digitais. A preservação digital pode ser vista como método, processo e modelo. Desse modo, foi constatado que existem pontos de interseção e de diferenciação que elucidam e indicam por que um conceito apareceu depois do outro e como cada um se apresenta em face à importância dada à gestão de grandes volumes de dados, especialmente os dados científicos. Após a comparação dos dois conceitos foi possível concluir que há mais unidade no entendimento do que é curadoria digital do que os conceitos de preservação e que as áreas de atuação de ambas complementam a gestão da informação digital como um todo.

PALAVRAS-CHAVE: Curadoria digital. Preservação digital. Gestão da Informação Digital.

RESUMEN

La curaduría y la preservación digital poseen el mismo objeto de interés: el mantenimiento de la información digital a través del tiempo. En la literatura científica, encontramos definiciones cercanas de los dos, sin embargo, curaduría se presenta como una evolución natural de la preservación de los procesos de recursos digitales. La preservación digital puede ser entendida como método, proceso y modelo. Por lo tanto, se encontró que existen puntos de intersección y diferenciación que elucidan por qué un concepto apareció después del otro y cómo cada uno se presenta a la vista de la importancia de la gestión de grandes volúmenes de datos, especialmente de datos científicos. Después de comparar los dos conceptos, se constata que hay más unidad en los conceptos de curación que en los conceptos de preservación encontrados. Como resultado, se concluye que ambos conceptos disponen de campos de acción que complementan la gestión de la información digital como un todo.

PALABRAS CLAVE: Curaduría Digital. Preservación Digital. Gestión de la Información Digital.

¹ Mestra em Ciência da Informação (UnB). Bibliotecária da Biblioteca Central (UnB). Brasília, DF. E-mail: thaysesantos@bce.unb.br. ORCID: <http://orcid.org/0000-0003-1052-9197>.

Submetido em: 03/08/2016 - Aceito em: 19/08/2016.

1 INTRODUCTION

Today the digital information works as a testament to the social, scientific and cultural memory, representative of the *zeitgeist* and the drivers of the possible forms of human experience in the advent of collectively built society.

Digital information resources are fragile differently to analog resources, largely because they are dynamic and complex. Their peculiarities, volume and variety, however, promoted the discussion of what should be preserved, when it should be preserved and how it should be preserved, requiring processes much more active than those demanded for analog resources.

The Herculean task of keeping what has been created or converted to digital featuring its context, its purpose and unique features led the discussion around what would 'digital preservation' designate. The term 'digital preservation' was first coined in 1990 in a joint research project of the Cornell University Library with Xerox to designate "using digital technologies to reformat analog media as part of the preservation process of those media " (HIRTLE, 2010, p. 124). According to Márdero Arellano (2008, p. 30) only in 1996, digital preservation came to be called as such.

The uncertainty of the environment, evolution of formats, technological supports to access, the corruptibility of data over time and the consequent risk of permanent loss are variables inherent to digital. The peculiarities of information in electronic media - especially after the rise of the Internet, and consequent *boom* of information available there, the variety and complexity of digital objects created - centralize the role of digital preservation. In Trevor Owens' words: "ensuring long term access and usability of complex digital objects is of critical importance to the future of nearly every area of arts, culture, the humanities and the sciences" (2015, p. 222).

The importance of digital preservation as a primordial action for maintenance of scientific, cultural and human knowledge partially repeats the discussion of preservation of information recorded in more stable electronic supports. However, the volume of data and how quickly they are produced bring to light more dilemmas for the so-called memory institutions. As its tutors, endowed with the responsibility of maintaining accessible information, they directly face the challenges related to assessing the quality of electronic documents and media content, since it is up to them to estimate whether the documents are good enough to ensure their accessibility and long-term usability.

Digital information, whether produced intentionally or not, generates a series of data through views, clicks, purchases, searches and all user behavior being recorded for later use (EINAV, LEVIN, 2014). This huge volume of relatable data or Big Data add even more pressure to the issue of preservation of information for future use. According to Vieira *et al.* (2012), Big Data is characterized by: 1. Data on the order of tens or hundreds of terabytes; 2. Elastic horizontal growth power; 3. Easy distribution of data; 4. Various data types.

The intention of this work is to analyze how the definitions of preservation and digital curation converge through the intersection of the concepts. We start from the gradual understanding of preservation as a model, process or method and the adoption of the term 'curation' as a natural evolution of the efforts required to process the digital information in the face of the new challenges related to the management for the long term maintenance of large scientific datasets.

2 METHODOLOGY

For this qualitative research, the study focuses on the evolution of the concept of preservation and then addresses the concept of scientific data curation to build a comparative frame between two fields of study and also contribute to the broader understanding of both concepts, comprised within the existing collective set of efforts for the maintenance of digital information over time.

This study is structured as follows: after the introduction, we dissertate about preservation under different aspects, and then the digital concept of curation is displayed. In section 5, we will make a cross of the concepts in order to find common and differentiating points of both, in an attempt to understand them more deeply and contribute to the current discussion on the maintenance of digital information over long periods.

The principles of the concept of theory or Dalhberg "Analytical Theory of Concept" of (1978) will be used in the comparison of preservation and digital curation concepts, since "there is a consensus in the state that the definitions are indispensable assumptions in arguments and verbal communications and which are necessary elements in the construction of scientific systems "(p. 106). The Theory of Concept is intended to serve as a foundation for conceptual analysis of any initiative that relates to the study and standardization of terms. We start from the understanding that concepts of the terms discussed here are general.

According to that theory, the concepts are labeled syntheses of true statements about thought objects: these statements - assertions - lead to recognition or separation of characteristics of concepts that can also be considered as the elements of the concepts. Therefore, the contrast between the definitions found contributes to the broader understanding of the concepts under review.

3 DIGITAL PRESERVATION AS A METHOD, PROCESS AND MODEL

Neil Beagrie (2008) reports that "one of the biggest difficulties of any emerging discipline, such as digital preservation, is the lack of definitive taxonomy of terms" (p. 24). This is because the same term is used in different ways by different communities. The same is true when comparing definitions of key terms such as digital preservation and digital curation. Such confusion can be explained because the object of interest of many of the areas involved in the maintenance of digital information over time differs according to their goals.

This is easily verified when we confront the different digital preservation acceptance found in the literature. This phenomenon explains why sometimes it can be seen as a method, process or model. The views and different interests make it difficult to define the term clearly.

Initially, the first preservation efforts started from the need to find digital collections solutions especially in the context of libraries, which have to deal with digital information and question how these become accessible collections through digitization of documents.

3.1 Preservation for its purpose, use or object of interest

Conway opens the discussion by defining preservation from its purpose: "the purpose of preservation is to protect permanent value information for access to present and future generations"(1990). Four years later, the Report of the Task Force on Archiving of Digital Information is published and considered a theoretical framework on the requirements for long-term preservation. It called for a national system of digital archives that would have multiple functions and would be "collectively responsible for the accessibility of long-term social heritage, economic, cultural and intellectual in the nation's digital form" as well as an audit and certification system repositories. There is also extensive discussion regarding the construction of a national cyberinfrastructure for such undertaking and it involved individuals from various institutions (SANTOS, 2014). Despite being the cornerstone of the field, the document covers the subject without giving a clear definition of the term digital curation. Integrity, complexity, content, provenance and digital objects context and the institutional framework necessary to make them accessible are the highlighted items.

3.2 Preservation as a process or method

Preservation is commonly defined as a series of processes or tasks that are performed to guarantee the permanence of information over time. There is some fluctuation around what these processes would be. Thus, "digital preservation is the method of keeping digital material alive so that they remain usable as technological advances render original hardware and software specification obsolete" (PRYTHERCH, 2005). Though digital preservation is emerging as a reliable process, there is still much debate ongoing and skepticism abounds about the viability and even the meaning of this process. Given the nature of electronic

storage technologies and the ephemeral nature of Web pages, many doubt that digital preservation will ever become a reality. (JANTZ; GIARLO, 2005).

The definition of the Preservation Handbook also interprets digital preservation as a series of activities to ensure continuous access to digital materials as long as it is needed.

Finally, digital preservation covers a wide range of activities designed to extend the life of computer files readable by machine and protect them from media failure, physical loss, and obsolescence (UNIVERSITY OF MINNESOTA LIBRARIES, 2016). The preservation is also considered in 'levels', as in the case of 2013 definition from the National Digital Stewardship Alliance (NDSA), a consortium of organizations committed to long-term preservation of digital information. The digital preservation levels are a hierarchical set of recommendations for how organizations should begin to build or improve their digital preservation activities.

Table 1. Version 1 of Digital Preservation Levels.

| | Level 1 (Protect your data) | Level 2 (Know your data) | Level 3 (Monitor your data) | Level 4 (Repair your data) |
|---------------------------------|---|---|---|--|
| Storage and Geographic Location | <ul style="list-style-type: none"> - 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 | <ul style="list-style-type: none"> - At least three 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 | <ul style="list-style-type: none"> - At least one copy in a geographic location with a different disaster threat - Obsolescence monitoring process for your storage system(s) and media | <ul style="list-style-type: none"> - 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 | <ul style="list-style-type: none"> - Check file fixity on ingest if it has been provided with the content - Create fixity info if it wasn't provided with the content | <ul style="list-style-type: none"> - Check fixity on all ingests - Use write-blockers when working with original media - Virus-check high risk content | <ul style="list-style-type: none"> - 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 | <ul style="list-style-type: none"> - 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 | <ul style="list-style-type: none"> - Identify who has read, write, move and delete authorization to individual files - Restrict who has those authorizations to individual files | <ul style="list-style-type: none"> - Document access restrictions for content | <ul style="list-style-type: none"> - Maintain logs of who performed what actions on files, including deletions and preservation actions | <ul style="list-style-type: none"> - Perform audit of logs |
| Metadata | <ul style="list-style-type: none"> - Inventory of content and its storage location - Ensure backup and non-collocation of inventory | <ul style="list-style-type: none"> - Store administrative metadata - Store transformative metadata and log events | <ul style="list-style-type: none"> - Store standard technical and descriptive metadata | <ul style="list-style-type: none"> - Store standard preservation metadata |
| File Formats | <ul style="list-style-type: none"> - When you can give input into the creation of digital files encourage use of a limited set of known open formats and codecs | <ul style="list-style-type: none"> - Inventory of file formats in use | <ul style="list-style-type: none"> - Monitor file format obsolescence issues | <ul style="list-style-type: none"> - Perform format migrations, emulation and similar activities as needed |

Source: NDSA Levels of Digital Preservation (NDSA, 2013, online)

The processing of information on levels is an interesting development as it brings closer the policies and practices for the management of digital information in a comprehensive treatment model for future access. The search for a representation of an information model that encompasses the creation, evaluation, processing and access points that the interest of preservation of information in digital media goes beyond the creation and use of systems that meet specific aspects of preservation. In this direction,

Although many software tools exist to carry out preservation tasks, their support status, quality, reliability, etc., are not under the control of the content owner and are too uncertain and unproven to be fully trusted and implemented in digital preservation workflows. (Fresa, A; Justrell, B; Prandoni, C., 2015, p. 192)

The leveled view of digital preservation is the closest to the preservationist approach that considers actions at all stages of the document life cycle.

3.3 *Digital Preservation throughout the life cycle of objects*

Digital preservation can also be interpreted as the management of the life cycle of a digital document, including storage activities of information in digital media, the maintenance of the collections, permanent access and distribution of digital documents. There is a consensus that such an undertaking, by its size and complexity, should be shared between all the actors directly involved in the creation and management of information, and the development of national and international cooperation activities. In 1998, Beagrie and Greenstein had already set some principles to be observed to reduce the risk of loss of digital material, which are:

- a) storage in a stable and controllable environment;
- b) development of strategies for preservation at the time that digital resources are created;
- c) ensure the continuity and integrity of the source, from an institutional policy that takes into account legal and cultural aspects such as licensing and copyright permissions;
- d) implementation of appropriate procedures for handling;
- e) transference of information to a standard storage media.

Briefly, the concepts and actions which were necessary to preserve have evolved to keep adding complexity to digital objects, in order to maintain their intrinsic characteristics and the creation context and future use as seen in Fig. 1.

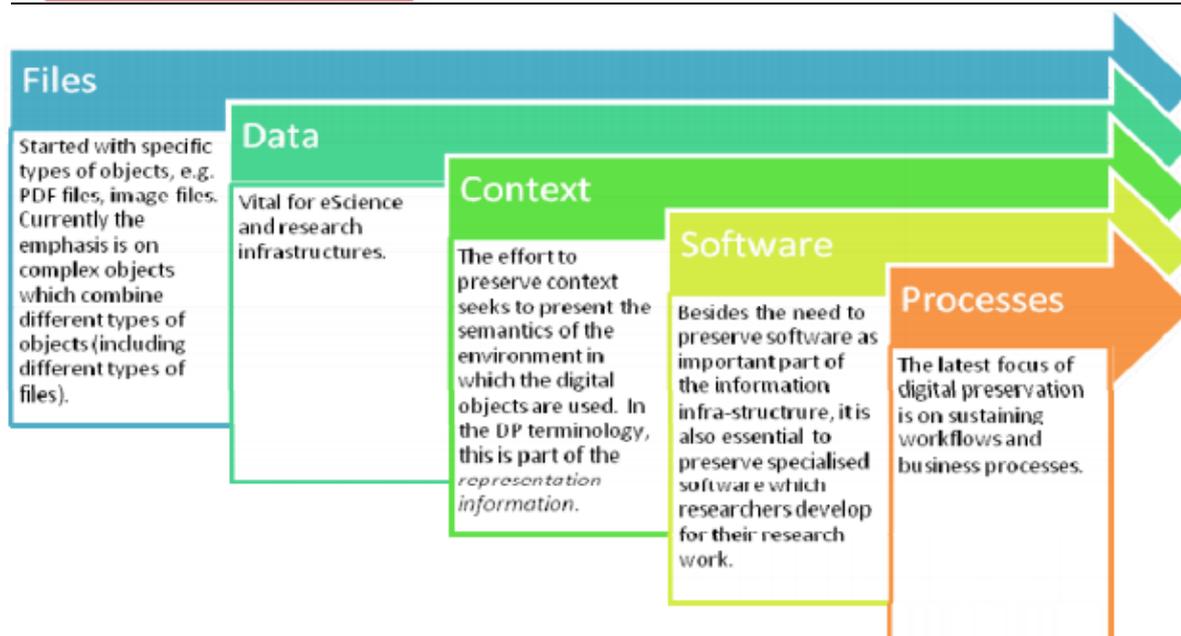


Figure 1. Evolution of digital objects addressed by digital preservation

Source: RUUSALEPP; DOBREVA (2013, p. 8).

The initial definitions understand digital preservation as a formal effort to ensure that digital information of permanent value will continue to be accessible and usable. These renderings are too generic and only circumscribe the boundaries of the area. Preservation also cannot be understood as a continuous sequence of events or transactions that, if executed with certain unity, guarantee the objectives of maintenance organizations' digital information. Digital preservation requires understanding the complexity that this enterprise comprises.

The authors start from different statements, considering characteristic elements or qualities attributed to digital preservation. The search for the definition of concepts is to "determine or set limits of a concept or idea." The authors do so when considering different species of complex traits (in terms of interest, quality and process) which combined result in own properties of digital preservation.

4 DIGITAL CURATION

The term curation can be defined as the "act or effect of healing, function, attribute, title, curator of power, trusteeship" (HOUAISS; VILLAR, 892 p., 2007). However, the term "curation" was transposed from museums and libraries to interactive media (Web and its new communication channels). The creation of distributed scientific data and the Big Data phenomenon (structured and unstructured data created in large scale), and the complex evolution of digital objects forged a favorable scenario of what would be called early 'data curation' or eScience. Huwe states (2013, p. 19)

the fact that even while we are moving full-speed into an era when crowdsourced, web-crawled, and tagged data are creating wholly new avenues for research, value still remains in ongoing data-acquisition programs. The emergence of Big Data research practices, which is revolutionizing how people parse datasets large and

small, can actually strengthen the impact of library discovery skills. As a result, information professionals stand to benefit not through digital curation and getting involved in Big Data analysis but also through the ongoing practice of reference and resource discovery. Because of this, I believe that it is important to promote our research and discovery acumen in the same manner that we are currently promoting the library as the “solution lab” for data curation. As admirable as that effort is, curation alone is, in my opinion, just half of the needed strategy. The crucial balance may be found by remembering that the skills inherent in reference work—discovery, pattern recognition, and analysis—offer a powerful means to convey our value proposition not only as data curators but also as information counselors with advanced data-acquisition skills.

The large volume of data, especially scientific data, demanded the creation of an infrastructure for its curation, “understanding the hyperonymic term designating the activities and strategies that comprise the administration of a part” (Santos, 2014). Specifically, digital curation is defined as the selection, preservation, maintenance, collection and archiving of digital assets.

Digital curation took off as a research field especially after the creation of the Digital Curation Centre (DCC) in 2004. Its implementation was scheduled in the Continuing Access and Digital Preservation Strategy for JISC launched in 2002 by JISC (Joint Information Systems Committee). Its function would be “to coordinate the development of new research, generic services and tools for digital curation that would support the next phase of the key elements of the JISC strategy” (Beagrie, 2004, p.7).

Beagrie speaks about the emergence of the term and also about its concept.

The term ‘digital curation’ is increasingly being used for the actions needed to maintain digital research data and other digital materials over their entire lifecycle and over time for current and future generations of users. Implicit in this definition are the processes of digital archiving and digital preservation, but it also includes all the processes needed for good data creation and management, and the capacity to add value to data to generate new sources of information and knowledge.

According to the DCC, digital curation involves maintaining, preserving and adding value to digital research data throughout its life cycle.

Rubridge et al. point out that the “DCC sees the digital curation as a continuum of activities, supporting the requirements for both current use and future” (2005, p. 1). Pennock (2006) relates the concept of curation with the concept of digital libraries, featuring the difficulties of doing curated repositories with information of different nature and formats. Hedges, Hasan and Blanke (2007) emphasize some of the limitations of digital preservation and indicate the curation as a set of actions to ensure the quality, integrity and audit of complex sets of information from actions performed during the life cycle of digital objects.

Uribe-Martinez and McDonald (2008, p. 276) argue that “all scientific data management activities boil down to the term data curation” and consider the DCC digital curation definition, which indicates that the terms data curation and digital curation are seen

as interchangeable, with no space to differentiation in the presented definition. Inge Angevaare states that "digital data requires specific care, the so-called 'curation', which includes preservation "to stand the test of time" (2009, p. 2).

Lee and Tibbo (2011, p. 124), state that digital curation is a term that came into use recently, brings together stream of disparate activities that have increasingly recognized to share a common set of challenges and opportunities, as well as reflecting the increasing convergence of communities before different. Ray (2012, p. 607) examines how the terms digital curation and cyberinfrastructure relate and describes the concepts associated as data to be managed, preserved, handled and made available for long-term use. For the author the "digital curation demands attention to the preservation and interoperability issues at the beginning of the data life cycle". Sayão and Sales (2012, p. 184) argue that digital curation is the result of accumulation of "knowledge and practices in preservation and access to digital resources that resulted in a set of strategies, technological approaches and activities that are now collectively known as "digital curation".

Weidner and Alemneh came up with an acceptance similar to the Digital Curation Centre:

Digital curation is the continuous activity of managing and enhancing the use of digital resources over their life-cycle and over time. Digital curation starts when an item is created (born-digital) or selected for digitization (analog) and continues through image processing, metadata capture, derivative creation, and preservation for long-term access.

All acceptations approximate what was established early by the DCC as digital curation. There is a certain accordance that digital curation is a more comprehensive term than digital preservation, to stress the need for handling of documents at all stages of their life cycle. Yakel (2007) indicates that curation is an umbrella term that encompasses digital preservation, data curation, management of electronic records and digital assets management. Despite thinking of the management of digital objects in a broader sense, and pointing out the importance of the planned intervention at all stages of their life cycles, digital curation should be understood as a hyperonymic term which is still in development.

In the concepts found, digital curation is seen as a continuum – an encompassing term that includes the necessary processes and activities for active data management: it comprehends the policies, processes, activities and techniques already present in digital preservation for the defined objective of managing information digital at all stages of their life cycle.

5 CONCEPTUAL INTERSECTIONS

In an attempt to cover the maintenance of digital information considering all its aspects and the large volume of data created for sharing and future use, either as a model, process or method, digital preservation has evolved from simplistic approaches (which were considered separate activities for treating the particularities of a digital document) to holistic approaches. The *Preservation Digital Reformatting Program*² from *Library of Congress* shows that:

because the term “digital preservation” can have multiple meanings, including ones as contrasting as the preservation of digital data and the use of digital technology to preserve analog data, “life-cycle management” is intentionally used in the Preservation Digital Reformatting Program to refer specifically to the progressive technology and workflow requirements needed to ensure long-term sustainability of and accessibility to digital objects and/or metadata.

The creation of life cycles of digital information emerged from the need to manage content, whereas the formulation and workflows for specific types of content owners. In them, digital media information moves through stages, from its creation, to the continued preservation, management and access over time.

Data curation deals with the management activities required to keep research data so that they are available long-term for reuse and preservation. In science, data curation - also known eScience - may indicate the important information extraction process of scientific texts and research papers by experts, which will be converted into electronic format, as an input of an electronic database. In broad terms, digital curation means a range of activities and processes made to create, manage, maintain and validate a component. In an attempt to compare both concepts, Table 2 is presented:

Table 2 – Digital Curation and Digital Preservation compared

| | Digital Curation | Digital Preservation |
|-----------------|---|---|
| Goal | Ensure long-term data sustainability | Ensure long-term access information stored digitally |
| Approach | Involves the maintenance, preserving and adding value to digital research data throughout its lifecycle | Involves the retention of the informational object and its significance |

² Preservation Digital Reformatting Program. URL:
<http://www.loc.gov/preservation/about/prd/presdig/index.html>

| | | |
|-----------------------------|---|---|
| Necessary activities | Selection, preservation, upkeep, collection, archiving and digital asset revaluation. | Selection, upkeep, collection, archiving and digital asset revaluation. |
| Techniques | Creating policies, activities at each stage of the life cycle of the objects; considers the nature of each object | Creating policies, emulation, data migration and encapsulation; considers the nature of each object |
| Longevity | Considered a continuous activity | Considered a continuous activity |

Digital curation is a term that encompasses related definitions directed to the selection, enrichment, processing and preservation of information for access and future use, be it scientific, administrative or personnel. Both curation and preservation consider the nature of digital objects, their representation and concern for their future access. Digital curation is a broad term used to describe the activities required to manage information in larger digital means than the preservation does: it is about the conscious maintenance of information, oriented to preservation and access of digital assets resulting from scientific research.

6 RESULTS

The investigation showed that digital curation is a concept or expression that contains or describes the field of activity related to the management of digital information for long-term preservation and for future access. It can be seen that there is more closeness between the curatorial definitions found than the existing definitions of what would be considered preservation.

Digital preservation has evolved from a range of concepts and is still seen in many ways more narrowly (activities or processes) or broadly (general settings). Different points of interest give different importance to the objective or purpose of preservation and, therefore, define it. The diversity of representations of what digital preservation would be, although enriching the discussion, end up creating some confusion in understanding what its limits and practices are.

The creation of models that addressed all phases of the digital information life cycle has emerged as a natural process in an attempt to cover all the variables that should be considered when considering the treatment of digital information. Of course, the emergence of digital curation, a term even broader than digital preservation, indicates that digital preservation, even though it encompasses a broad spectrum of activities, would not respond alone to the new and more complete digital objects that have been created.

7 CONCLUSIONS

Digital curation was a term suggested in the early 2000s, at least a decade after the emergence of preservation as a solution to the issue of keeping of digital information. It was born in an age with the Internet already consolidated as a worldwide channel for creation and dissemination of information, and amid the growing realization that massive sets of collectively built data must be preserved for current and future use. Digital preservation already prevailed in the scientific community as a necessity inherent to the electronic medium. Digital curation can be interpreted as an evolution of understanding of preservation issues – not as a panacea, but as wide model for active management of digital information preservation for future use based in processes.

It did not come to replace the preservation actions but incorporate the best practices created and contribute to the preservation of cultural and scientific memory that is created collectively, in a distributed mode, which is massive and unstructured. Digital curation was born concerned with the management throughout the life cycle, and it is not useful to view it as a method or process for preservation, once in all the concepts found agree on the planned intervention on digital objects at all stages of their life cycle.

Digital curation can be understood as a general term: it exists so that digital preservation is possible, since its definition contains its concept and the actions associated with it.

REFERENCES

ANGEVAARE, Inge. Taking Care of Digital Collections and Data: "Curation" and Organisational Choices for Research Libraries. **Liber Quarterly: The Journal of European Research Libraries**, v. 19, n. 1, 2009. Disponível em: <<http://liber.library.uu.nl/index.php/lq/article/view/7948>>. Acesso em: 21 jun. 2016.

ASSOCIATION FOR LIBRARY COLLECTIONS & TECHNICAL SERVICES. **Definitions of digital preservation**: Prepared by the ALCTS Preservation and Reformatting Section, Working Group on Defining Digital Preservation. Washington: ALA, 2007. Disponível em: <<http://www.ala.org/alcts/resources/preserv/defdigpres0408>>. Acesso em: 25 maio 2016.

BEAGRIE, Charles. The Digital Curation Centre. **Learned Publishing**, vol. 17, n.1, jan. 2004, p.7-9. BEAGRIE, Neil. **Preservation Management of Digital Materials: The Handbook**. Digital Preservation Coalition, nov. 2008. Disponível em: <http://www.dpconline.org/component/docman/doc_download/299-digital-preservation-handbook>. Acesso em: 04 jul. 2016.

BEAGRIE, Neil; GREENSTEIN, Daniel. **Strategic Policy Framework for Creating and Preserving Digital Collections**. London: King's College, 1998. Disponível em:

<http://www.ukoln.ac.uk/services/papers/bl/framework/framework.html>. Acesso em: 31 ago. 2016.

COMMISSION ON PRESERVATION AND ACCESS AND THE RESEARCH LIBRARIES GROUP. **Preserving Digital Information**: Report of the Task Force on Archiving of Digital Information. Maio 1996. Disponível em: <<http://www.clir.org/pubs/reports/pub63watersgarrett.pdf>>. Acesso em: 07 maio 2016.

CONWAY, Paul. Digitizing Preservation. **Library Journal**, v. 119, n. 2, p. 42–45, fev. 1994.

CONWAY, Paul. **Preservation in the Digital World**. Washington, D.C.: Commission on Preservation and Access, 1996.

CORNELL/XEROX/CPA JOINT STUDY IN DIGITAL PRESERVATION: Progress Report Number 2. **The Electronic Library**, v. 10 n. 3, p.155-163, 1992.

CURADORIA, In: HOUAISS, Antonio; VILLAR, Mauro de Salles. **Dicionário Houaiss da Língua Portuguesa**. Rio de Janeiro: Objetiva, 2007. p. 892.

DAHLBERG, Ingetraut. Teoria do conceito. **Ciência da Informação**, v. 7, n. 2, p. 102-108, 1978. Disponível em: <<http://revista.ibict.br/ciinf/article/view/115/115>>. Acesso em: 30 ago. 2016.

DAHLBERG, I. A referent-oriented, analytical concept theory for INTERCONCEPT”, **International Classification**, Vol. 5, pp. 142-51, 1978.

EINAV, Liran; LEVIN, Jonathan D. **The Data Revolution and Economic Analysis**. National Bureau of Economic Research, 2014. Disponível em: <<http://web.stanford.edu/~jdlevin/Papers/BigData.pdf>>. Acesso em: 09 jul. 2016.

FRESA, Antonella; JUSTRELL, Börje; PRANDONI, Claudio Digital curation and quality standards for memory institutions: PREFORMA research Project. **Archival Science**, v. 15, Issue 2, p. 191–216, jun. 2015.

HEDSTROM, Margaret. Digital preservation: a time bomb for digital libraries. **Computers and the Humanities**, v. 31, p. 189–202, 1998. Disponível em:<http://deepblue.lib.umich.edu/bitstream/handle/2027.42/42573/10579_2004_Article_153071.pdf;jsessionid=5395F9215FE2F1173FBD31FE3C55361A?sequence=1>. Acesso em: 30 jun. 2016.

HIRTLE, Peter. The history and current state of digital preservation in the United States. In: **Metadata and Digital Collections**: A Festschrift in Honor of Thomas P. Turner. Ithaca, NY: Cornell University Library: 2010.

HUWE, Terence K. Data Discovery and Data Curation Going Hand in Hand. **Computers in Libraries**, v. 33, n. 3, abr. 2013, p. 17-19. Disponível em: <<http://www.infotoday.com/cilmag/apr13/Huwe--Data-Discovery-and-Data-Curation-Going-Hand-in-Hand.shtml>>. Acesso em: 23 maio 2016.

JANTZ, Ronald; GIARLO, Michael J. Digital Preservation: Architecture and Technology for Trusted Digital Repositories. **D-Lib**, v. 11, n. 6, jun. 2005. Disponível em: <http://www.dlib.org/dlib/june05/jantz/06jantz.html>. Acesso em: 23 jul. 2016.

LIBRARY OF CONGRESS. **Preservation Digital Reformatting Program**. [S. d.]. Disponível em: <https://www.loc.gov/preservation/about/prd/presdig/preslifecycle.html>. Acesso: 17 jun. 2014.

MÁRDERO ARELLANO, Miguel A. **Critérios para a preservação digital da informação científica**. 2008. 356 f. Tese (Doutorado em Ciência da Informação) - Universidade de Brasília, Brasília, 2008. Disponível em: http://repositorio.bce.unb.br/bitstream/10482/1518/1/2008_MiguelAngelMarderoArella.no.pdf. Acesso em: 31 jul. 2016.

MÁRDERO ARELLANO, Miguel. Preservação de documentos digitais. **Ciência da Informação**, v. 33, n.2, p.15-27, maio/ago. 2004. Disponível em: <https://core.ac.uk/download/files/418/11878016.pdf>. Acesso em: 31 jul. 2016.

NATIONAL DIGITAL STEWARDSHIP ALLIANCE. The NDSA Levels of Digital Preservation: An Explanation and Uses. In: **Proceedings of the Archiving (IS&T) Conference**, April 2013, Washington, DC Disponível em: http://ndsa.org/documents/NDSA_Levels_Archiving_2013.pdf. Acesso em 31 jun. 2016.

NORA, Pierre. Entre memória e história: a problemática dos lugares. **Revista Projeto História**. São Paulo, v. 10, p.7-28, 1993.

OWENS, Trevor, Preserving Complex Objects, Janet Delve, David Anderson (Eds.). Facet Publishing, Greenwich, CT (2014), **Journal of Academic Librarianship**, v. 41, n. 2, p. 222, mar. 2015.

PENNOCK, Maureen. Digital curation and the management of digital library cultural heritage resources. **Local Studies Librarian**, v. 25, n. 2, p. 3-7, 2006. Disponível em: www.ukoln.ac.uk/ukoln/staff/m.pennock/publications/docs/lsl-curation_mep.pdf. Acesso em: 23 jul. 2016.

PRYTHERC, Ray. **Harrod's Librarians' Glossary and Reference Book**. 10th ed. Ashgate: Aldershot and Burlington, VT, 2005.

RUUSALEPP, Raivo; DOBREVA, Milena. **State of the art in digital preservation and multi-agent systems**. INNOVATIVE DIGITAL PRESERVATION USING SOCIAL SEARCH IN AGENT ENVIRONMENTS, 2013. Disponível em: http://durafire.eu/wp-content/uploads/2014/08/D1-1_State-of-the-Art.pdf. Acesso em: 19 jul. 2016.

RUSBRIDGE et al. The Digital Curation Centre: a vision for digital curation. In: **Local to Global Data Interoperability** : challenges and technologies .24-25 jun. Sardenha, Italia. [S.l.]: IEEE, 2005, p. 31-41. Disponível em: eprints.erpanet.org/82q01/DCC_Vision.pdf. Acesso em: 12 jul. 2016.

SANTOS, Thayse Natália Cantanhede. **Curadoria digital: o conceito no período de 2000 a 2013**. 2014. 165 f., il. Dissertação (Mestrado em Ciência da Informação)—Universidade de Brasília, Brasília, 2014.

SAYÃO, Luis Fernando; SALES, Luana Farias. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. **Informação & Sociedade: Estudos**, v.22, n.3, p. 179191, set./dez. 2012. Disponível em: <http://www.ies.ufpb.br/ojs/index.php/ies/article/view/12224/8586>. Acesso em: 23 fev. 2013.

STATE LIBRARY OF NORTH CAROLINA. **Digital Preservation Best Practices and Guidelines**, 2013. Disponível em: <http://digitalpreservation.ncdcr.gov/>. Acesso em: 17 maio 2016.

UNIVERSITY OF MINNESOTA LIBRARIES. **Digital Preservation at the University of Minnesota Libraries: what is Digital Preservation?** Disponível em: <https://www.lib.umn.edu/dp>. Acesso em: 05 jun. 2016.

URIBE-MARTÍNEZ, Luis; MACDONALD, Stuart. Un nuevo cometido para los bibliotecarios académicos: data curation. **El profesional de la información**, vol. 17, n. 3, p. 273-280, 2008. Disponível em: www.elprofesionaldeinformación.com/contenidos/2008/mayo/03.html. Acesso em: 21 jul. 2016.

VIEIRA, Marcos Rodrigues. et. al. Bancos de dados no SQL: conceitos, ferramentas, linguagens e estudos de casos no contexto de Big Data. In: Simpósio Brasileiro de Bancos de Dados, 2012, Mato Grosso. **Minicurso...** Mato Grosso: Universidade Federal de Mato Grosso, 2012. Disponível em: http://data.ime.usp.br/sbbd2012/artigos/pdfs/sbbd_min_01.pdf Acesso em: 29 jul. 2016.

WEIDNER, Andrew James; ALEMNEH, Daniel Gelaw. Workflow Tools for Digital Curation. **Code4Lib Journal**, n. 20, 2013.

YAKEL, Elizabeth. Digital Curation. **OCLC Systems & Services**, v.23, n.4, 2007, p. 335-340.

