RDBCI | Revista Digital de Biblioteconomia e Ciência da Informação
Digital Journal of Library and Information Science

ISSN 1678-765X
DOI 10.20396/rdbci.v018i0.8659259/22488

ARTICLE

Digital Preservation
challenges, requirements, strategies and scientific output

Danilo Formenton [1] https://orcid.org/0000-0001-6531-4295
Luciana de Souza Gracioso [2] https://orcid.org/0000-0002-6320-4946

[1] Federal University of São Carlos, São Carlos, SP, Brazil. / e-mail: formenton.danilo@gmail.com
[2] Federal University of São Carlos, São Carlos, SP, Brazil. / e-mail: lugracioso@yahoo.com.br

ABSTRACT
The aim of this article is to provide a broad and reflective perspective on the main aspects of digital preservation, based on the challenges indicated, the recognized requirements and the strategies analyzed by the scientific community. The methodology adopts quantitative-qualitative and exploratory-descriptive research, with a review of the national and international literature of the last twenty-one years on digital preservation, in order to delineate the trends and policies on the theme as well as deepening the discussion on the needs for archiving and long-term preservation of digital content. Data is analyzed from the bibliographic survey of scientific publications indexed by Scopus and Web of Science from the last five years (2015-2019) that deal with the subject "digital preservation". It was found that among the themes discussed, budgets, costs and metadata for preserving and Web archiving are emerging and studies lacking in Brazilian Information Science. In the international scientific output, Brazil stands out for publication quantity, indicating a maturation of the theme, coinciding with the advance of national projects, such as the Cariniana Network. However, we have financial, human and technological demands that, together with the characteristics of strategies for digital preservation, highlight the usefulness of collaborations and of little-explored national topics.

KEYWORDS
Preservation of digital documents. Digital objects. Digital preservation strategies. Surveys. Information Sciences. Digital preservation.

Preservação Digital
desafios, requisitos, estratégias e produção científica

RESUMO
O objetivo do artigo foi prover uma visão ampla e reflexiva dos principais aspectos da preservação digital, a partir dos desafios indicados, dos requisitos reconhecidos e das estratégias analisadas pela comunidade científica. A metodologia adota pesquisa quanti-qualitativa e exploratória-descritiva, com revisão da literatura nacional e internacional dos últimos vinte um anos acerca da preservação digital, a fim de apoiar a apreensão de tendências e políticas do tema como aprofundar as discussões sobre as necessidades no arquivamento e preservação de conteúdos digitais a longo prazo. Analisa dados do levantamento bibliográfico de publicações científicas nas bases *Scopus* e *Web of Science* dos últimos cinco anos (2015-2019) que tratam do assunto "*digital preservation*". Observa que dos métodos discutidos, orçamentos, custos e metadados para preservação e arquivamento da *Web* são temas emergentes e carentes de estudos na Ciência da Informação brasileira. Na produção científica internacional, o Brasil destaca-se pelas quantias de publicações indicando uma maduração do tema e pactuando com o avanço dos projetos nacionais, como a Rede Cariniana. Porém, temos demandas financeiras, humanas e tecnológicas que, unido às características das estratégias para preservação digital, expõem a utilidade das colaborações e dos tópicos nacionais pouco explorados.

PALAVRAS-CHAVE
Preservação de documentos digitais. Objetos digitais. Estratégias de preservação digital. Levantamentos. Ciência da Informação. Preservação digital.

JITA: JH. Digital preservation.

1 INTRODUCTION

Modern societies are major producers and consumers of information, consisting of a vital asset for their cultural, economic, political and knowledge development. The use of Information and Communication Technologies (ICT) associated with the World Wide Web proposed by Tim Berners-Lee in 1989 and the Internet, allowed a worldwide information explosion with the rapid production, dissemination and acquisition of digital information resources, whether in the public or private spheres. However, the dynamism and ephemerality of digital environments impose rapid and definitive losses of important records available online, portraying one of the global challenges of the 21st century to ensure the preservation of and continuous access to a personal, corporate and cultural digital memory.

In recognition of these risks, organizations around the world - especially cultural heritage institutions and universities - have implemented techniques and tools for creating, managing and preserving accessible digital materials over time. In the Brazilian case, there are digital preservation policies planned by Higher Education Institutions (IES), such as at the State University of Campinas (2011) and the State University of Paulista (2017). Similarly, these universities are also part of the Brazilian Network of Digital Preservation Services (Cariniana[1]), which was created in 2013 by the Brazilian Institute of Information on Science and Technology (IBICT) in order to ensure continuous access to national electronic documents.

Digital preservation has been a theme of study in Information Science. It is a complex challenge, inevitable and current in national and international publications of the area, requiring analysis and inter/multidisciplinary solutions. As a definition of digital preservation, Grácio, Fadel and Valentim (2013, p. 113) interpret that digital preservation refers to "[...] an organizational management process that encompasses various activities necessary to ensure that a digital object can be accessed, retrieved and used in the future, from the ICT existing at the time and with guarantees of authenticity.", judging the concept of authenticity of a digital resource/object linked to safeguarding the original informational content of its production. From Márdero Arellano (2008) and Santos and Flores (2015), a digital object is any type of file in digital media, which is represented in bitstream and formed by logical structure, content and presentation structure.

Therefore, this work aims to provide a broad and reflective view of the main issues of digital preservation, which subsidizes the understanding of trends and policies on the subject as a competence required of professionals in information units (BOERES, 2017) and also deepen discussions on the needs of digital preservation to generate new knowledge and effective studies and future actions in this field. We consider that the understanding of current affairs and interests of the scientific community of digital preservation allows the synergy of strategic, institutional and technological contemporary approaches to the current and future problems of the management of preservation and long-term access to digital information.

For an ideal structure of the present work, we have the description of the methodology adopted, followed by the exhibition and debate of the results obtained, which include the problems, requirements and strategies of digital preservation described in classical and recent specialized literature, the data recovered in scientific databases, and the final considerations on the research developed.

[1] Available at: http://cariniana.ibict.br/index.php/inicio. Access on: Apr. 20, 2020.

2 METHODOLOGY

We have adopted a quantity-qualitative, exploratory-descriptive approach (SILVA; MENEZES, 2005), based on the review of specific national and international literature of the last twenty-one years on digital preservation; and on the collection and analysis of data, available in selected databases, related to the recent scientific production of this research theme. The discussions of the work provide a synthesis of the current conception of digital preservation issues, starting from the main difficulties discerned, the criteria of this process recognized and the strategic, political and technological solutions explored by the digital preservation community, portraying what has been articulated so far in the long-term archiving of digital contents.

Regarding the methodological procedures, the bibliographic method was adopted (MARCONI; LAKATOS, 2017; SEVERINO, 2016) in which scientific publications on the subject of 'digital preservation' were raised at Scopus (Elsevier[2]), the largest database of abstracts and citations from the literature with peer review, and the Web of Science Main Collection (Clarivate Analytics[3]), a database of publications of the highest quality and with the greatest impact in the world. Available via the Periodicals of the Coordination of Improvement of Higher Level Personnel (Capes), both databases with worldwide coverage, multidisciplinary and references with abstracts and citation data also offer bibliometric tools to monitor, analyze and/or visualize research.

As limiters in the bibliographic survey conducted on March 20, 2020, we use advanced research (Advanced) with the term "digital preservation" in the fields of title, abstract and keywords (Title, Abstract, Keywords/Topic), we define the filter options by conference proceedings/proceedings paper and journal articles (article), published in Portuguese, English and Spanish in the last five years (2015-2019) in order to obtain a more recent overview of publications on the subject through national and international academic sources with more intense pace of scientific production and validation. In the exploration and analysis of the results, through the Microsoft Excel 2016 program, we illustrate with figures of recent scientific production on digital preservation, highlighting the most productive researchers, in addition to institutions, countries and research areas with the highest number of publications in this period.

3 DIGITAL PRESERVATION: THE CHALLENGE FOR PRESENT AND FUTURE GENERATIONS

The main difficulties of digital preservation come from the specificities of the objects it seeks to safeguard over time. These digital objects, born digital or digitized, are susceptible to constant changes and the ephemerality of the means where they are created, transported or stored, as well as the high dependence on hardware, software and support technologies for their reproduction that quickly become obsolete or are physically damaged. Thus, these particularities result in reflecting the issues of reliability, authenticity and integrity of digital documents in their management, archiving and usable access over a long period. Through the National Archive (2016) and Barbedo, Corujo and Sant'Ana (2011), an authentic digital document constitutes what we prove to be and is exempt from unauthorized changes; in turn, the integrity of a digital document refers to the custody of completeness and fixity, where

[2] Available at: https://www.elsevier.com/pt-br/solutions/scopus. Access on: 20 Apr. 2020.
[3] Available at: https://clarivate.com/webofsciencegroup/solutions/web-of-science/. Access on: 20 Apr. 2020.

information to be recorded in metadata is crucial for identifying the provenance and context of creation and maintenance of the document over time. Santos and Flores (2014, p. 100) point out that the reliability of a document "[...] is related to its custody, so the reliable document must be able to achieve the same effects as when it was first manifested, to achieve the same effects reliability must incorporate quality of integrity and authenticity".

Unlike the preservation of non-digital media, digital objects are easily able to dissociate into their individual elements making it difficult to maintain them entirely. One can, for example, maintain the content of an electronic document, but lose/adulterate its layout by continuous migrations, or even preserve the physical presence (the data archive) of an object, but no longer maintain its capacity for interpretation (BULLOCK, 1999; THOMAZ; SOARES, 2004). This being said, according to Márdero Arellano (2008, 2012), Baggio e Flores (2012) and Innarelli (2009, 2014), the preservation of digital and electronic documents requires specific efforts to maintain the original properties and the capacity to serve as a record and source of information, given their fragilities on complexity, costs, technological obsolescence and physical degradation.

In addition to these fundamental problems of digital preservation, there are other managerial, technical, legal, political, economic and social challenges, among which we highlight according to Box 1 below:

Table 1. Some challenges of digital preservation

Institution/Author	Year	Challenge
Barbedo; Corujo; Sant'ana	2011	Rapid and continuous technological obsolescence in hardware, software, formats and storage media, plus threats of physical damage to files and to hardware components and media.
National Museum of Australia	2012	The intellectual property rights and other legal obligations to be fulfilled, which interfere with the copying, storage, alteration and use of the content of digital resources for long-term preservation purposes
National Library of Australia	2013	The need for different preservation approaches at different scales; in addition to recurring cycles of maintaining continuous access to digital resources through the use of a variable set of tools.
Pennock	2013	The dynamic and ephemeral character of the Web combined with personalized browsing experiences, which bring dilemmas of authenticity and integrity validation in the rendering of large and complex websites archived on a large scale and over a long period.
Library and Archives Canada	2015	The wide variety of digital file formats and standards, making it essential to recognize and use the most appropriate and sustainable formats for long-term archiving of digital information
Digital Preservation Coalition	2015	The high demand for recruiting professionals with experience and practical skills continuously updated in digital preservation, in order to compose distributed or multidisciplinary teams in organizations.
The National Archives	2015	Cloud storage and service contracts require careful management to meet archiving, providing flexibility, low costs and secure, accessible data beyond the life of current technologies/providers.

National Archive (Brazil)	2016	The institutional commitment of not losing the significant properties of the digital object in custody and preserved for the long term, ensuring its recovery, intelligibility and authenticity, to serve as a source of evidence and information.
State University of São Paulo	2017	The continuous demand for financial resources for assiduous investments in organizational and technological infrastructure and in personnel training in order to keep digital objects accessible over time.

Source: The authors

In view of the challenges indicated, even if digital objects include dynamic content, multimedia, functionalities and advantages of transmission, replication and editing in digital environments, their complexities bring obstacles for long-term preservation and accessibility. From the aspects of digital preservation, the broad investments, legal requirements and guarantees of location, context, authenticity and integrity of digital content are evident. Therefore, it becomes vital to recognize the existing requirements, strategies and technological resources, described in the specialized literature, to be considered for effective and feasible efforts in this field.

4 THE REQUIREMENTS FOR DIGITAL PRESERVATION, BASED ON THE OAIS REFERENCE MODEL

Digital preservation requires compliance with a minimum set of functional and non-functional requirements in order to achieve the desired results. From the existing approaches, Bullock (1999), Formenton, Gracioso e Castro (2015), Thomaz (2004) and Thomaz e Soares (2004) identify nine requirements to be analyzed for the long-term preservation of digital objects, while Innarelli (2009, 2014) cites the principles of digital preservation in ten commandments that can be interpreted and applied according to the reality and structure of the organization; both proposals are supported by the Open Archival Information System (OAIS) reference model.

Based on the considerations pointed out by the mentioned authors, we have idealized a set of five basic requirements for digital preservation, which can be understood as follows:

- **Maintain a preservation policy** - elaboration, implementation and maintenance of guidelines, objectives and institutional methods for the archiving of digital collections, covering the clear delimitation of the types of information or which elements of the digital object will be selected, given the multimedia, hypertextual and dynamic nature of digital content.
- **Guaranteeing reliability, authenticity and integrity** - confidence that the digital object accessed is precisely the one that is sought, where the probable alterations or displacements resulting from the preservation measures for which it was submitted (the continuous migrations of media, formats and versions, for example), kept its identification and unequivocal location and its content with layout and original functionalities over time.
- **Keeping the context** - safeguarding the specific hardware and software dependencies that define the digital object, the reasons for its production, its distribution modes and the relationships with other objects. Preferably, open standards and formats, established in the

market and accepted by recognized national and international official bodies, should be used, allowing autonomy over developers, manufacturers or suppliers for the analysis of format specifications and access to information on file.

- **Maintain the origin** - identification of the origin or source of the digital object, its chain of custody and the detailing of the history of changes occurred, through the metadata for digital preservation, in order to prove or ensure the authenticity and integrity of the digital object and support its reconstitution, consistency and persistence over the long term. Regarding the description of electronic documents, we highlight the Extensible Markup Language (XML) format as an open standard for producing, storing and transferring documents electronically, independent of operational platforms and software manufacturers, understandable by various applications and self-explanatory. In digital preservation, XML language is considered a particular type of migration, which enriches information about structures and meaning, ensures the encapsulation of metadata and information required for the interpretation of the original digital objects and benefits the interoperability between resources from different areas (MÁRDERO ARELLANO, 2008).
- **Maintain recovery** - assiduous implementation and review of a backup policy or backup, which values the replication of the digital object (and its metadata) in place separate physical and the combined use of different types of storage technologies, for the purpose of ensuring reliable, integrated and secure access and restoration of data.

In summary, the five indicated requirements to be discussed and adapted by organizations committed to the long-term maintenance of digital information, aim primarily at safeguarding digital objects and the ability to continuously access and use their content, thus reflecting the assumptions of digital preservation strategies. Although each requirement has its own particularities, they are all intrinsically linked and have as key components the construction of preservation policy and the elaboration of metadata for preservation.

Given their general character, the described requirements are still used in the work as reference criteria for designing digital preservation. However, other aspects should be analyzed in the implementation of this process, incorporating them in a digital preservation policy, such as the identification of the informational needs of the user community; the provision of financial, human and technological resources; the ethical, moral and legal rights and duties of the various parties involved; and the application of methods and technologies for preserving the original properties of digital objects that support the validity of their reliability, authenticity and integrity. Furthermore, these requirements are also based on OAIS, one of the main reference models for standardization of digital preservation activities.

OAIS or Open Information Archiving System (OAIS[4]), a conceptual framework that standardizes a repository system for long-term preservation of and access to digital information, of the International Organization for Standardization (ISO) and the Consultative Committee for Space Data Systems (CCSDS), has defined a functional and information model, which specifies the operations to be performed by the system and the information recorded by metadata required for the representation of the materials maintained and long-term digital archiving.

To solve the current problems and challenges imposed by the digital environment to the long-term preservation of digital objects and, together, take advantage of scientific advances and the modern applications or technological tools that they make available, several strategies

[4] BRAZILIAN ASSOCIATION OF TECHNICAL STANDARDS. ABNT NBR 15472: spatial data and information systems: reference model for an open information archiving system (SAAI), Rio de Janeiro: ABNT, 2007.

have been proposed for digital preservation. It is necessary to understand them, exploring their capacities and limitations.

5 STRATEGIES: SOLUTIONS TO THE CHALLENGES OF DIGITAL PRESERVATION

Bullock (1999) brought together digital preservation strategies in two types: strategies to try to solve the problem of technological obsolescence, covering migration, transfer to analog supports, emulation and preservation of technology; and strategies to "take control", including the adoption of standards and guidelines, documentation and description of resources (metadata), building partnerships and establishing infrastructure. Pearson and Del Pozo (2009) still choose to order the strategies into primary preservation actions, which directly alter the digital materials to be preserved, such as migration; and secondary preservation actions, which change the way the material is accessed and how this access is preserved over time, including emulation and the collection and maintenance of a "technology museum". On the other hand, Long (2009) divides the strategies in terms of greater chances of providing at least partial preservation solutions, i.e., methodologies that are not sustainable in the long term, such as technological museums; and those that are sustainable in the long term, such as migration and emulation. Regardless of the preservation of technology and technological museums having different designations, in this work we will judge that both are the same strategy of digital preservation. Through Long (2009) and Pearson and Del Pozo (2009), a list of pros and cons can be obtained for consultation and research about the collection and maintenance dependent on the original technological context of creation and use of digital materials to be preserved in the short term.

In Santos and Flores (2017) the strategies for digital preservation are also organized by their priorities in the preservation of the three levels of digital objects: physical level (the integrity of hardware and support), integrating the refresh or recopy of content in more current media; logical level (the software and integrity of the original bit chain), encompassing the emulation, the preservation of technology and the encapsulation or gathering of everything required for access and interpretation of objects; and conceptual level (the visual representation of content interpreted by humans), including migration. According to Márdero Arellano (2008), Formenton, Gracioso e Castro (2015), Santos e Flores (2015, 2018), Thomaz (2004) and Thomaz e Soares (2004) the different strategies of digital preservation can usually be grouped into: structural strategies, which consist of the initial investments coming from institutions in order to build an appropriate environment for digital preservation; and operational strategies, which constitute the real measures of physical, logical or conceptual preservation of digital objects to be executed by the respective organizations.

Based on this generic categorization, we will discuss below some strategies for digital preservation that are currently more adopted, disseminated and/or reported in the specialized literature.

5.1 Structural strategies

a) Adoption of open standards

The adoption and compliance with open standards, whether for the creation or management of digital documents, enables the reduction of the effects of technological

obsolescence on digital preservation. Santos and Flores (2015, 2017) and Schäfer and Constante (2012), considering the migration of standards in case of obsolescence and the restriction of formats to store the data infer that the open and non-proprietary standards allow the reconstruction of interpreter software and unrestricted distribution, which ensures the non-dependence on updates from developers for access to information and reprogramming and adaptation of the software in consensus with the preservation policy. Examples of guidelines for identifying reliable digital file formats for long-term preservation include the "Guidelines on File Formats for Transferring Information Resources of Enduring Value" from Library and Archives Canada (2015); and the "Long-term File Formats" from the National Archives of Australia (c2020).

Open standards for digital preservation and archiving are set by official standards bodies and international consortia such as the National Digital Stewardship Alliance (NDSA[5]), the World Wide Web Consortium (W3C)[6], ISO[7] and the International Internet Preservation Consortium (IIPC)[8]. The NDSA, launched in 2010, is a consortium of organizations involved in the long-term preservation of digital information, which its activities occur through interest and work groups, such as Web Archiving Survey Working Group. W3C, founded in 1994, is an international community where member organizations (the Information and Coordination Nucleus of the BR Point - NIC.br of the Brazilian Internet Steering Committee, for example), a team and the public work to develop Web standards. Set up in 1947, ISO is a global network of national standards bodies from 164 countries (only one member per country), with Brazil participating in ABNT, which defines the international standard ISO 14721:2012 Space Data and Information Transfer Systems - OAIS. Formed in 2003 at the National Library of France, the IIPC has the participation of several organizations, such as the National Library of Chile and the Library of Congress of the United States, which are dedicated to creating standards and tools for web archiving.

b) Policy documents and institutional strategies

The set of documents created around the world, through the initiative of institutions involved with digital preservation and researchers on the subject, give guidelines for the right development and implementation of policies and strategies for managing digital materials. Examples include the "Recommendations for the Production of Digital Preservation Plans" of the Portuguese General Directorate of Archives (BARBEDO; CORUJO; SANT'ANA, 2011); and the "Digital Preservation Policy of the Permanent Program for the Preservation of and Access to Digital Archival Documents - AN Digital" of the National Archives (2016). A digital preservation policy, according to Grácio (2012) and Grácio, Fadel and Valentim (2013), is based on elements distributed in three categories: organizational, including management elements for setting and institutional stabilization on the policy and preservation measures; legal, covering institutional standards and legislation in force at national and international levels; and technical, involving elements on flows, processes and preservation measures. In addition, certain aspects should be analyzed in the creation of these policies and strategies, such as understanding and incorporating the organizational context that will exist (DIGITAL PRESERVATION COALITION, c2015).

[5] Available at: https://ndsa.org/about/. Access on: 20 Apr. 2020.
[6] Available at: https://www.w3.org/Consortium/. Access on: 20 Apr. 2020.
[7] Available at: https://www.iso.org/about-us.html. Access on: 20 Apr. 2020.
[8] Available at: http://netpreserve.org/about-us/. Access on: 20 Apr. 2020.

c) Budgets and costs of digital preservation

The calculation of the costs of digital preservation is a complex activity and is essential for the determination of profitable practices and the justification of investments of resources and expenses. Among the issues of impact for the costs to be analyzed are human resources (specialists and multidisciplinary team of professionals); the implementation, operation and maintenance of preservation measures; material resources; and the mission and institutional objectives, with the insertion of the type and volume of collections, the levels of preservation and access defined and the proposed deadline for actions (BARBEDO; CORUJO; SANT'ANA, 2011; DIGITAL PRESERVATION COALITION, c2015). There are still uncertainties about the costs of digital preservation, where most studies point to high expenses involved, such as Boeres (2017), and some indicate the low cost of this process compared to the preservation of traditional collections, such as Andrade, Borges and Jambeiro (2006).

d) Selection for digital preservation and legal compliance

The increasing volumes of information created in digital environments with their restrictions, dependencies and variable relevance, make it indispensable the selectivity of what access will be maintained for a certain end and period. According to Boeres and Márdero Arellano (2005) and Grácio (2012), the long-term preservation of everything and for all is unfeasible and null, so a policy of digital preservation should define selection criteria based on institutional objectives, academic and societal needs and in terms of cost-benefit of preservation. Another difficulty posed by the Digital Preservation Coalition (c2015), involves legality in the collection, preservation and access to managed and stored digital materials, given the legal duties on data protection/disposal, copyright and service contracts, in addition to the law that is often also lagging behind in terms of technological change and demands for digital preservation. In Brazil, we highlight Law no. 12,527, of 2011, which regulates the citizens' right of access to information produced or under the custody of public entities, such as the IES; Law no. 9,610, of 1998, which defines copyright in the country; Law no. 13,709, of 2018, which provides on the protection of personal data and their processing, including in digital media; Decree no. 10. 278, of 2020, which establishes the technique and requirements for the digitalization of public or private documents; the resolution of the National Council of Archives (CONARQ) no. 43, of 2015, which establishes guidelines for the implementation of reliable digital archival repositories; and Law no. 12,682, of 2012, which regulates the preparation and filing of documents in electromagnetic media.

e) Training and personnel development

As raised by Farias, Araújo and Evangelista (2018), the institutions of the Cariniana Network place the lack of human resources, specialized personnel and specific budget as preponderant factors in the adoption of digital preservation strategies. In this perspective, a program of staff training and continuous professional development is crucial, considering the different skills for preservation and the clear definition of the roles and responsibilities of the various stakeholders, with guarantees of their implementation, acceptance and improvement. These issues are notably discussed in Boeres (2017) which, through the indications of Brazilian and foreign specialists, identifies a set of skills required for teams of digital preservation professionals in information units, among which are: knowledge in long-term digital data management; and understanding about reliability, authenticity and integrity of records, trends and policies for digital preservation. Alternatives for training, development and learning of

teams include sharing information and exchanging personnel with similar organizations, as well as short courses or complete theoretical and practical programs about digital preservation, available in person or remotely and online (DIGITAL PRESERVATION COALITION, c2015).

f) Metadata for digital preservation

The effective adoption of standards or metadata schemes is one aspect of ensuring digital preservation in order to support the management of archiving and maintenance of continuous access of digital objects in informational environments such as repositories and cloud services. Examples of metadata standards for digital preservation are: the Metadata Object Description Schema (MODS) [9] as a descriptive metadata schema for library services; the Metadata Encoding and Transmission Standard (METS)[10] for encoding descriptive, administrative and structural metadata about objects in a digital library; and the PREservation Metadata: Implementation Strategies (PREMIS)[11] for encoding, managing and exchanging preservation metadata between repository systems for long-term digital preservation. According to Formenton et al. (2017) and Lavoie and Gartner (2013), these schemes will determine the identity, representation, cohesion and persistence of the object in the repository, with guarantees of reliability, authenticity and integrity, in addition to defining interoperability between systems. Interoperability is "[...] the ability of various systems with different hardware and software platforms, data structures and interfaces to exchange data with minimal loss of content and functionality. (NATIONAL INFORMATION STANDARDS ORGANIZATION, c2004, p. 2, our translation). On metadata standards of digital preservation in cloud services, Castro and Silveira (2018) infer that the subject is still little explored in the area of Information Science, either nationally or internationally, even with the growing use of these digital environments in the storage of digital objects as a strategy in the long-term preservation to contain technological obsolescence.

g) Investment and assembly of technological infrastructure

A factual digital preservation requires extensive investments in technological infrastructure to sustain the flows, processes and activities of digital material archiving. In view of the handling of increasing amounts of data, through Boeres and Márdero Arellano (2005) and Digital Preservation Coalition (c2015), some facilities should be assumed: cloud services and corporate storage systems for the replication of maintained data; reliable digital repository systems to store, manage and access the intellectual output of HEIs over time; and high-performance computing to manipulate large data, whether search or Web archives. In this infrastructure, two technological environments are suitable because, for Grácio (2012), Grácio, Fadel and Valentim (2013) and Universidade Estadual Paulista (2017), it should meet the preservation of the objects (and their metadata) and enable their ideal access, search and recovery.

[9] Available at: http://www.loc.gov/standards/mods/. Access on: 30 May 2020.
[10] Available at: http://www.loc.gov/standards/mets/. Access on: 30 May 2020.
[11] Available at: https://www.loc.gov/standards/premis/. Access on: 30 May 2020.

h) Formation of collaborative networks

Success in overcoming the challenges of digital preservation requires greater collaboration between organizations, teams of professionals and creators of digital objects to be maintained. For example, IBICT in the creation of the Cariniana Network joined the International LOCKSS Alliance for the preservation of national open access publications included in Open Journal Systems (OJS) and DSpace software, allowing the exchange of experiences with institutions around the world united in other collaborative networks (MÁRDERO ARELLANO, 2012). In fact, according to Grácio (2012) and Grácio, Fadel and Valentim (2013), collaborative initiatives include the attractions of the exchange of knowledge and experiences as well as the standardization of institutional strategies in support of interoperability of digital objects between systems. However, these measures impose flexibility within organizational structures and imply potential conflicts that can manifest themselves in the form of multiple agendas, deadlines or funding mechanisms. Thus, judging digital preservation as a global, vital and feasible effort, the key to building and maintaining collaborations lies in effective dialogue between stakeholders through the use of terms and language intelligible to all (DIGITAL PRESERVATION COALITION, c2015).

The first part of this section of the article was dedicated to the main structural strategies proposed. The second part will discuss current operational strategies for digital preservation.

5.2 Operational strategies

a) Definition of the storage medium

Digital preservation depends on the appropriate choice of data storage medium in the long term. In the last decades, the use of magnetic and optical media judged technical criteria for its continuous evaluation, such as the factors cited in Brown (2008) and Thomaz e Soares (2004): capacity (of storage, reading, etc.); obsolescence; standardization; feasibility (of retaining the probatory integrity, etc.); cost; and others. However, today the use of resilient local and/or cloud storage systems (DSpace)[12], the Flexible Extensible Digital Object Repository Architecture - Fedora[13] and the Lots of Copies Keep Stuff Safe - LOCKSS , for example) has become more common, supported by principles such as redundancy and diversity (the use of several independent copies of digital material stored in different geographic locations and the combined use of different types of storage technology, for example), fixation, monitoring and repair (the use of check sums to record and monitor the integrity of each copy of the material and the use of other copies to create a replacement in case of loss/corruption detection, for example) etc. , for increasing volumes of digital material to be accessed, preserved and recovered with ease, speed and greater accuracy (DIGITAL PRESERVATION COALITION, c2015; SANTOS; FLORES, 2017).

The DSpace project is an open source software developed by the Massachusetts Institute of Technology in the United States and widely used by academic organizations to create repositories for preservation and open access to all types of digital content. Launched in 1998 and developed by the Universities of Virginia and Cornell in the United States, the Fedora project is an open source software used in the construction of repository systems for

[12] Available at: https://duraspace.org/dspace/about/. Access on: 20 Apr. 2020.

management, dissemination, access and preservation of digital objects. As a digital preservation initiative in operation since 1999 under the auspices of Stanford University in the United States, the LOCKSS[14] Program offers reliable, open source repository software designed to preserve academic publications.

b) Migration

Migration is one of the most adopted strategies in long-term preservation. Through Baggio and Flores (2012), Barbedo, Corujo and Sant'Ana (2011), Hedstrom (2001) and Pearson and Del Pozo (2009) it consists in transferring digital objects from technological platforms in the process of obsolescence, physical degradation or discontinuity to newer, stable and standardized ones, ensuring the updating/conversion of versions, formats and supports, compatibility with current technologies and access to information. However, for the authors the migration variations intervene in the original structure and content of the object, requiring a continuous plan and the effective apprehension, control and documentation of the changes. In Santos and Flores (2017, 2018) and Schäfer and Constante (2012) constant migration may also reflect in data loss/addition, failures in faithful representation of complex objects and incompatibilities between source and destination formats, requiring other strategies. Taken as a type of migration (SCHÄFER; CONSTANT, 2012), the refresh strategy, according to Ferreira (2006, p. 33), relies on "[...] the transfer of information from one physical storage medium to another more current before the former deteriorates or becomes irreparably obsolete.", like a CD-ROM for the cloud. For unprotected information, this process implies a low degree of technical knowledge and investment in equipment but refreshing only solves the problems of physical degradation and obsolescence of the media (BAGGIO; FLORES, 2012).

c) Transfer to analogue media

Transfer to long-life analogue media is a method to be used as a last resort for the archiving of digital information. According to Bullock (1999), Hedstrom (2001) and Rothenberg (1999) the hard copy fixes the simple digital objects as a whole, keeps the content and, in a way, the layout, being a palliative, operative and applicable action in the absence of a technological infrastructure. Nevertheless, the authors infer that printing results in the loss of interactive or dynamic functionality and the original form of complex objects restricting the practice of this approach. Therefore, another option would be the hybrid strategy of using microfilm copies as archival substitutes and of generating digital copies because, according to Schäfer and Constante (2012) and Thomaz and Soares (2004), it allows the reformatting of documents originally produced on paper and the improvement of functionality and accessibility.

d) Emulation

Emulation proposes the creation and use of a modern emulator that replaces and reproduces the behavior of old and/or obsolete hardware and software technologies. As observed in Barbedo, Corujo and Sant'Ana (2011), Long (2009) and Rothenberg (1999), this method maintains the content and visualization of digital objects in their native format with layout and original functionality, being free from maintenance of specific platforms and systems and also useful when there is interest in preserving the original technological context

[13] Available at: https://duraspace.org/fedora/about/. Access on: 20 Apr. 2020.

[14] Available at: https://www.lockss.org/about/why-lockss. Access on: 20 Apr. 2020.

of the objects. However, the authors expose that emulation is susceptible to the risks of dependency and obsolescence of emulators and supposes limitations with time in the capacity of reliable representation of materials. Thus, according to Santos and Flores (2015), emulation is applied together with encapsulation and in exchange for the conservation of technology, having a short and medium-term function through high technical and financial resources. Ferreira (2006, p. 43) explains that the encapsulation is based "[...] on preserving, together with the digital object, all the necessary and sufficient information to allow the future development of converters, viewers or emulators". This strategy makes it possible to preserve the original format of the object, as long as it is accompanied by instructions that grant the interpretation of the document formats as the information content (BAGGIO; FLORES, 2012).

e) Conservation of technology

Technology conservation is an interim and short-term method. Based on Márdero Arellano (2008), Bullock (1999) and Santos and Flores (2017) involves maintaining the original hardware and software for creating or accessing digital objects in order to make them available for use. For the authors this strategy maintains the content and visualization of digital materials in their native format with layout and original functionality, but the creation of "museums" does not reduce the effects of technological obsolescence and requires conditions of cost, space and technical support difficult to operate, making its use useful for valuable objects in proprietary format and obsolete software. One of the possible roles of "computer museums" in digital preservation, according to Rothenberg (1999), may be to make heroic efforts to recover readable data from old media and to check emulators by comparing their behavior with that of saved obsolete machines. Although it is considered to be in decline, technology conservation is still used by some organizations. For example, the National Library of Australia has been active in collecting donations of obsolete hardware and software to make up its collection in order to support the recovery of valuable data in outdated digital formats with the application of emulators to create virtual environments that enable the use of outdated technologies (THORPE, 2015).

f) Digital archaeology

Digital archaeology is an expensive and partial method of preservation. As analyzed in Baggio and Flores (2012), Hedstrom (2001) and Schäfer and Constante (2012) consists in the rescue of inaccessible digital materials, either due to technological obsolescence and/or physical degradation of the support, which were not attended by other strategies or were lacking any preservation action. For the authors, digital archaeology is indicated only for situations in which the relevance of the information legitimates the high costs of the procedure, since there are no guarantees of recovery, restoration and interpretation of the fullness of the data that, in this way, compromise the definition of identity, integrity and context of the recovered material. By way of example, Galrão (2017) brings a practical case of digital data submitted to digital archaeology, where it was found that only part of the documentation could be interpreted.

g) Web archiving

Web archiving refers to the process of selecting and collecting, storing, retrieving, accessing and preserving long-term content on the Web. Through Costa, Gomes e Silva (2017) and Pennock (2013), the execution of these methods includes facing some difficulties, such as the large volume of information that is lost or becomes unavailable quickly in its original form

due to the dynamics of the Web and the conditions of legality and licenses/permissions of the content owner. To support the advancement and preservation of the Web, initiatives are being created in the world (for example, the Internet Archive[15] and the Arquivo.pt[16]) as tools and techniques are developed by consortia, stressing the W3C and IIPC. Created in 1996, the Internet Archive is a non-profit organization that provides free and universal access to a digital library of web pages, books, texts, videos, music, images, software, etc., through its official website and the Wayback Machine tool. The Portuguese Web Archive, started in 2008, consists of an infrastructure that provides search and access to Portuguese Web pages archived since 1996, aiming to preserve the information published on the Web for research purposes.

In the national sphere, according to Rockembach and Pavão (2018), the theme is current in the area of Information Science and there are still no systematized initiatives, implying the lack of a cultural memory of the Brazilian Web for present and future generations. Although there are no organized Web archiving initiatives in Brazil, in recent years research groups have been formed to investigate the topic, such as the Center for Research on Web Archiving and Digital Preservation (NUAWEB) at the Federal University of Rio Grande do Sul's Library and Communication Faculty (FABICO).

In view of the strategies discussed, we observe that there are no fully satisfactory and definitive solutions when applied in isolation, being necessary a combination of strategies for digital preservation. Nevertheless, operational strategies bring changes in the original properties of digital objects that may result in significant loss of data over time to end users, or even foresee restrictions on access, use and preservation of content linked to legal conditions and the availability of financial, human and technological resources; which contradict the principles of long-term preservation (in particular, ensuring reliability, authenticity and integrity, maintaining context and provenance). These issues are exacerbated by rapid technological obsolescence and the increasing complexity and interactivity of digital objects, reflecting the urgency of further studies, policies and standardized and collaborative technologies.

Thus, structural strategies (especially budgets and costs, preservation metadata and relationship networking) may mitigate the above-mentioned problems. The use of models for cost calculation, such as Boté, Fernandez-Feijoo and Ruiz (2013) and Willer et al. (2008), help in decisions about investments in preservation and the recording of changes by metadata provide assurance that the losses occurred did not affect the reliability, authenticity and integrity of materials (and their significant properties); but, they are methods lacking studies in national Information Science. Furthermore, the collaborative partnerships give answers to the lack of subsidies and the overcoming of obstacles where, through Márdero Arellano (2012) and Farias, Araújo and Evangelista (2018), can be evidenced by the expressive advance of digital preservation in the Cariniana Network with LOCKSS.

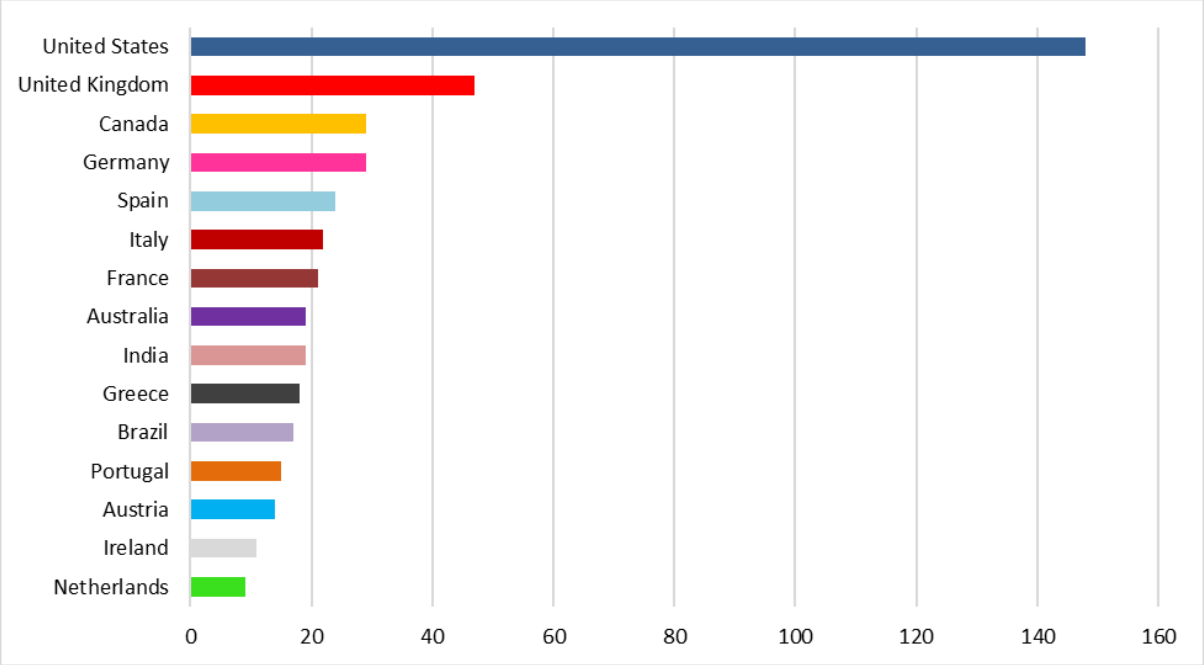
Considering the publications in the Scopus and Web of Science bases, below we present the sum of the scientific production in the last years on digital preservation, highlighting the most productive authors as the institutions, countries and research areas with the largest number of publications.

[15] Available at: https://archive.org/about/. Access on: 20 Apr. 2020.
[16] Available at: https://sobre.arquivo.pt/pt/ajuda/o-que-e-o-arquivo-pt/. Access on: 20 Apr. 2020.

6 ANALYSES OF PUBLICATIONS ON DIGITAL PRESERVATION AT SCOPUS AND WEB OF SCIENCE BASES

In the Scopus bibliographic survey, through the described methodological procedures, 274 papers (56.5%) and 211 conference proceedings (43.5%) were recovered, totaling 485 documents. The most productive authors were Efstratios Kontopoulos from Hellas Research and Technology Center[17] in Greece and Michael L. Nelson from Old Dominion University in the United States, with 9 publications each. The institutions with the highest scientific output are the Universities of Illinois at Urbana-Champaign and Old Dominion in the United States and the University of Toronto in Canada, in addition, the countries that published most were the United Kingdom and especially the United States as Chart 1.

Chart 1. Number of "digital preservation" publications per country (2015-2019) on Scopus database



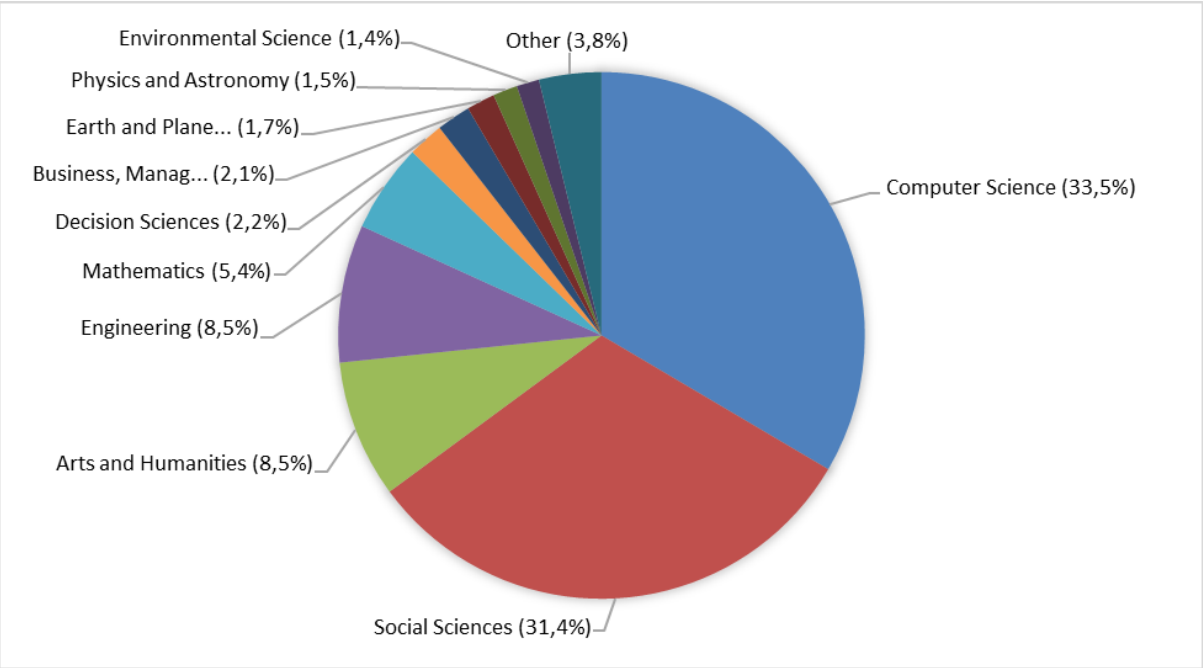
Source: The authors.

The University of Old Dominion, a member of the IIPC through the Computer Science Department, also coordinates with the Los Alamos National Laboratory the Memento Framework (RFC 7089[18]) and the Time Travel[19] tool for searching previous versions of websites stored in web archives, such as Arquivo.pt and Internet Archive. Also, part of the University of Toronto's publications are affiliated with the Faculty of Information, which has the Institute of Digital Curatorship[20], an interdisciplinary unit of researchers for research into the preservation of digital resources. Among the UK publications are the Universities of Glasgow, Edinburgh and Cambridge, members of the Digital Preservation Coalition (DPC)[21], a not-for-profit association founded in 2002 and dedicated to digital preservation.

[17] Available at: em: http://dgp.cnpq.br/dgp/espelhogrupo/1769372358627653 Access on: 20 Apr. 2020.
[18] Available at:https://www.certh.gr/root.en.aspx. Access on: 20 Apr. 2020.
[19] Available at: https://tools.ietf.org/html/rfc7089. Access on: 20 Apr. 2020.
[20] Available at: http://timetravel.mementoweb.org/. Access on: 20 Apr. 2020.
[21] Available at: http://dci.ischool.utoronto.ca/about-the-dci/. Access on: 20 Apr. 2020.

Reinforcing Rockembach and Pavão (2018)'s research, the multi and interdisciplinary nature of digital preservation is also evident in the documents recovered, as shown in Graph 2.

Chart 2. Distribution of "digital preservation" publications by research area (2015-2019) in Scopus database

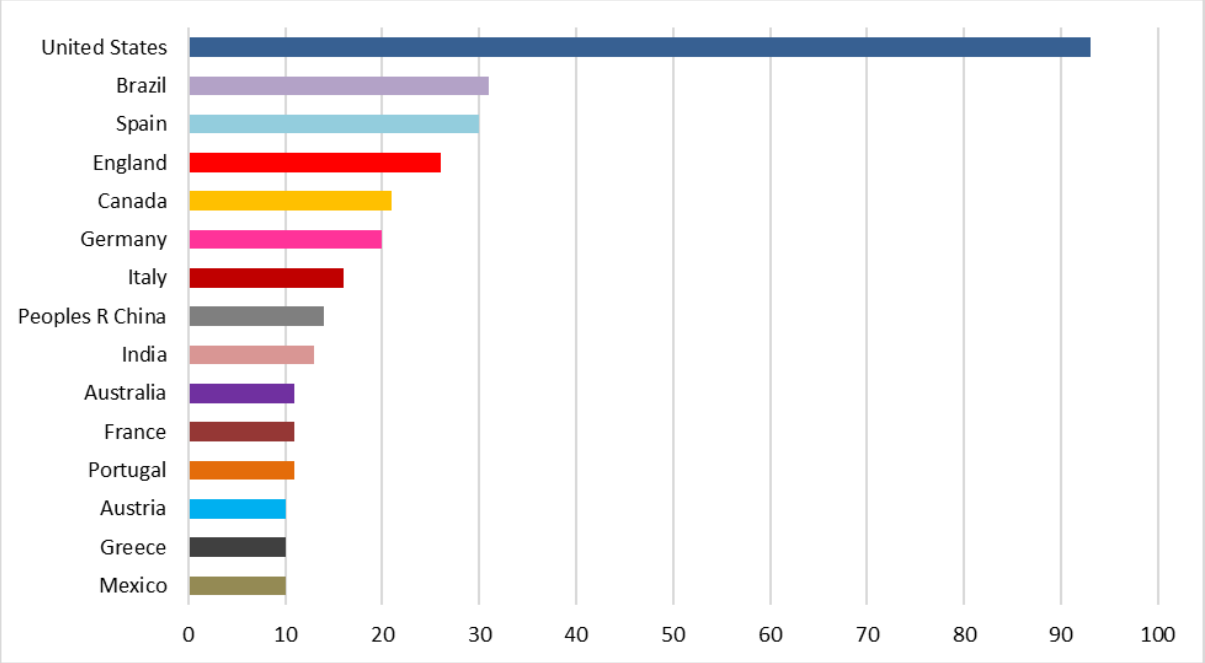


Source: The authors

In view of this, we find that the two major areas of research that study and publish on the subject of digital preservation are grouped in Computer Science and Social Sciences, together with Arts and Humanities, Engineering, Mathematics, Decision Sciences, Business, Management and Accounting, Earth and Planetary Sciences, Physics and Astronomy, Environmental Science and others. Of the terms used in the keywords of the 485 documents recovered, there were 241 records for "Digital Storage", 25 records for "Information Management", 23 records for "Metadata", 17 records for "Web Archiving", 15 records for "Digital Repository", 12 records for "OAIS" and 8 records for "Collaboration", reproducing part of the aspects of the subject treated in this work.

In turn, 363 documents were retrieved in the Web of Science base bibliographic survey, through the described methodology. The most productive authors were Michael L. Nelson and Michele C. Weigle, from the University of Old Dominion, with 8 and 6 publications each, respectively; and Daniel Flores, from the Federal University of Santa Maria (UFSM) in Brazil, with 7 publications. The two major research areas they publish on the subject focus on Library and Information Science and Computer Science. The institutions with the greatest scientific production are the University of Old Dominion and UFSM, and the countries that have published the most are Brazil and, mainly, the United States as Chart 3.

Chart 3. Number of publications in "digital preservation" per country (2015-2019) in the Web of Science base



Source: The authors

Publications in the United States are primarily affiliated with the University of Oklahoma, where through the Donald E. Pray Law Library, it participates in the Legal Information Preservation Alliance (LIPA)[22], a consortium of libraries created in 2003 and directed at projects for the preservation of printed and electronic legal information. Finally, in Brazilian publications, most of them are associated with UFSM and IBICT, accompanied by the University of Brasilia (UnB), University of São Paulo (USP), Federal University of Paraíba (UFPB), Federal University of Santa Catarina (UFSC), Federal University of Goiás (UFG), among others, institutions that are part of the integral partners of the Cariniana Network coordinated by IBICT[23] itself.

Given the results presented, we see that in recent years the scientific production on digital preservation has remained in general in developed countries, considering the limitations of the representation of journals, areas and countries of the databases dealt with in this work. Even so, countries such as Brazil, which hold the first publications on the subject in 2007 and 2003 on Scopus and the Web of Science, in this order, stands out for the remarkable amounts of studies in the Social Sciences (especially, Librarianship and Information Science). Reinforcing Boeres (2017), the data indicate a maturity of the theme (and its relevance) in Brazil in view of its greater solidity abroad, consistent with the growth of national initiatives where, for example, the use of LOCKSS gave rise to the application of theoretical knowledge in digital preservation (MÁRDERO ARELLANO, 2012).

[22] Available at: https://www.dpconline.org/about. Access on: 20 Apr. 2020.
[23] Available at: https://www.lipalliance.org/history-of-lipa. Access on: 20 Apr.2020
[24] Available at: http://cariniana.ibict.br/index.php/parceiros-da-rede/parceiros-integrais. Acesso em: 20 abr. 2020.

7 FINAL CONSIDERATIONS

By providing a broad and reflective overview of the main issues of digital preservation, the research conducted shows a current picture of challenges, experiences and opportunities among institutions and professionals involved in the production, maintenance and management of digital materials. Covering several types of digital objects - text, images, audio, videos, software, games, content in social media, web pages, etc. - The strategies for digital preservation, in defense of a possibility that future generations will have little or no record of the 21st century (held by Vint Cerf, one of the founders of the Internet, as a "digital dark age")[24], still present a relatively insufficient character of knowledge and practical essays in the scientific community.

Despite the prominence of the Carinian Network in publications obtained from the databases treated in the work, through the study of Farias, Araújo and Evangelista (2018) we observed that most of the partner institutions of the Network do not have an estimate of growth in activities or prospects of implementing strategies in the coming years due to the lack of specialized personnel, their own budgets and technological support. United to this reality, even if we are among the leading countries in the number of users in social media such as Twitter[25] and Facebook[26], or also have in 2019 approximately 149 million Internet[27] users, we found from Rockembach and Pavão (2018) that in Brazil the archiving of the Web is one of the most growing themes and lacking effective actions and studies in Information Science.

However, other themes related to digital preservation found in research are also little analyzed today by this large area of research in the national panorama, i.e.: budget planning and calculation of approximate costs for preservation, which bring an understanding about the feasibility of investments and effectiveness of cooperation or outsourcing; and the standards of preservation metadata, formal structures for describing resources, which record contextual information and provenance so that digital content can be interpreted in the present and future. Both topics are vital in large-scale digital preservation in web archives, systems, etc., given the high production of complex, heterogeneous and dependant objects.

In view of this, based on the discussions of this work, we propose that in addition to the urgency of further national studies of Information Science on digital preservation, new research could focus on the difficulties and little explored topics identified (especially, budgets, costs and metadata for digital preservation and Web archiving), foreseeing the expansion of our knowledge and experiences in this field. We must take care of the awareness of the importance of the process to those who produce, use and access digital objects and of the duty of the interested parties in the whole life cycle of the object, such as the definition of parameters to evaluate the effects of the preservation approaches in force (and the use of emulators, archiving technologies, etc.), in order not to distance ourselves from the principles of long-term digital preservation.

[25] Available at: https://www.bbc.com/news/science-environment-31450389. Access on: 15 Mar. 2020.

[26] Available at: https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/. Access on: 20 Apr. 2020.

[27] Available at: https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/. Access on: 20 Apr. 2020.

[28] Available at: https://www.statista.com/statistics/262966/number-of-internet-users-in-selected-countries/. Access on: 20 Apr. 2020.

REFERENCES

ANDRADE, Ricardo; BORGES, Jussara; JAMBEIRO, Othon. Digitalizando a memória de Salvador: nossos presente e passado têm futuro. **Perspect. ciênc. inf.**, Belo Horizonte, v. 11, n. 2, p. 243-254, maio/ago. 2006. Available at: http://www.scielo.br/pdf/pci/v11n2/v11n2a08.pdf. Access at: 20 abr. 2020.

ARQUIVO NACIONAL (Brasil). Programa Permanente de Preservação e Acesso a Documentos Arquivísticos Digitais – AN Digital. **Política de preservação digital**. Versão 2. [Rio de Janeiro, RJ], dez. 2016. 33 p. Available at: http://www.arquivonacional.gov.br/images/conteudo/servicos_ao_governo/Programas_e_Projetos/AND_Politica_Preservacao_Digital_v2.pdf. Access at: 20 abr. 2020.

BAGGIO, Claudia Carmem; FLORES, Daniel. Estratégias, critérios e políticas para preservação de documentos digitais em arquivos. **Ci. Inf.**, Brasília, DF, v. 41, n. 2/3, p. 58-71, maio/dez. 2012. Available at: http://revista.ibict.br/ciinf/article/view/1336/1515. Access at: 20 abr. 2020.

BARBEDO, Francisco; CORUJO, Luís; SANT'ANA, Mário. **Recomendações para a produção de planos de preservação digital**. Versão 2.1. Lisboa: Direção-Geral de Arquivos (DGARQ), nov. 2011. 111 p. Available at: http://arquivos.dglab.gov.pt/wp-content/uploads/sites/16/2014/02/Recomend_producao_PPD_V2.1.pdf. Access at: 20 abr. 2020.

BOERES, Sonia Araújo de Assis. **Competências necessárias para equipes de profissionais de preservação digital**. 2017. Tese (Doutorado em Ciência da Informação) - Faculdade de Ciência da Informação, Universidade de Brasília, Brasília, 2017. Available at: http://repositorio.unb.br/bitstream/10482/24354/1/2017_SoniaAraujodeAssisBoeres.pdf. Access at: 20 abr. 2020.

BOERES, Sonia Araújo de Assis; ARELLANO, Miguel Ángel Márdero. Políticas e estratégias de preservação de documentos digitais. *In*: ENCONTRO NACIONAL DE CIÊNCIA DA INFORMAÇÃO – CINFORM, 6., 2005, Salvador, BA. **Anais** [...] Salvador, BA: UFBA, 2005. p. 1-15. Available at: http://www.cinform-anteriores.ufba.br/vi_anais/docs/SoniaMiguelPreservacaoDigital.pdf. Access at: 20 abr. 2020.

BOTÉ, Juanjo; FERNANDEZ-FEIJOO, Belen; RUIZ, Silvia. Digital preservation cost: a cost accounting approach. **The Learning Organization**, [*s. l.*], v. 20, n. 6, p. 419-432, Sept. 2013. Available at: https://www.researchgate.net/publication/263478773_Digital_preservation_cost_A_cost_accounting_approach. Access at: 20 abr. 2020.

BROWN, Adrian. Selecting storage media for long-term preservation. **The National Archives Digital Preservation Guidance Note 2,** [London], Issue 2, p. 1-7, Aug. 2008. Available at: https://www.nationalarchives.gov.uk/documents/selecting-storage-media.pdf. Access at: 20 abr. 2020.

BULLOCK, Alison. Preservation of digital information: issues and current status. **Network Notes**, National Library of Canada, Ottawa, n. 60, Apr. 1999. Available at: http://epe.lac-bac.gc.ca/100/202/301/netnotes/netnotes-h/notes60.htm. Access at: 20 abr. 2020.

CASTRO, Fabiano Ferreira de; SILVEIRA, Júlio César Tauil. Mapeamento de padrões de metadados de preservação digital em cloud services. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB, 19., 2018, Londrina, PR. **Anais** [...] Londrina, PR: UEL, 2018. p. 5183-5204. Available at: http://enancib.marilia.unesp.br/index.php/XIX_ENANCIB/xixenancib/paper/view/1003/1697. Access at: 20 abr. 2020.

COSTA, Miguel; GOMES, Daniel; SILVA, Mário J. The evolution of web archiving. **International Journal on Digital Libraries**, [*s. l.*], v. 18, n. 3, p. 191-205, Sept. 2017. Available at: https://link.springer.com/article/10.1007/s00799-016-0171-9. Access at: 20 abr. 2020.

DIGITAL PRESERVATION COALITION. **Digital preservation handbook**. 2th ed. [Glasgow], c2015. Available at: https://www.dpconline.org/handbook. Access at: 20 abr. 2020.

FARIAS, Juliana Pinheiro; ARAÚJO, Luiza Martins de Santana; EVANGELISTA, Raimunda Lima. Percepções da importância da preservação digital. **RICI: R.Ibero-amer. Ci. Inf.**, Brasília, DF, v. 11, n. 1, p. 200-218, jan./abr. 2018. Available at: http://periodicos.unb.br/index.php/RICI/article/view/8475/7062. Access at: 20 abr. 2020.

FERREIRA, Miguel. **Introdução à preservação digital**: conceitos, estratégias e actuais consensos. Guimarães, Portugal: Escola de Engenharia da Universidade do Minho, 2006. 88 p. Available at: https://repositorium.sdum.uminho.pt/bitstream/1822/5820/1/livro.pdf. Access at: 28 maio 2020.

FORMENTON, Danilo. *et al*. Os padrões de metadados como recursos tecnológicos para a garantia da preservação digital. **Biblios**, Pittsburgh, n. 68, p. 82-95, 2017. Available at: http://www.scielo.org.pe/pdf/biblios/n68/a06n68.pdf. Access at: 20 abr. 2020.

FORMENTON, Danilo; GRACIOSO, Luciana de Souza; CASTRO, Fabiano Ferreira de. Revisitando a preservação digital na perspectiva da ciência da informação: aproximações conceituais. **Rev. digit. bibliotecon. cienc. inf.**, Campinas, SP, v. 13, n. 1, p. 170-191, jan./abr. 2015. Available at: https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/1587/pdf_91. Access at: 20 abr. 2020. doi:10.20396/rdbci.v13i1.1587

GALRÃO, Ana Filomena. Estudo de caso em arqueologia digital: o gabinete da área de Sines. **Páginas a&b**, Porto, n. especial, p. 99-114, 2017. Available at: http://ojs.letras.up.pt/index.php/paginasaeb/article/view/2658/2446. Access at: 20 abr. 2020.

GRÁCIO, José Carlos Abbud. **Preservação digital na gestão da informação**: um modelo processual para as instituições de ensino superior. São Paulo: Cultura Acadêmica, 2012. 214 p. Available at:

https://repositorio.unesp.br/bitstream/handle/11449/113727/ISBN9788579833335.pdf?sequence=1&isAllowed=y. Access at: 20 abr. 2020.

GRÁCIO, José Carlos Abbud; FADEL, Bárbara; VALENTIM, Marta Lígia Pomim. Preservação digital nas instituições de ensino superior: aspectos organizacionais, legais e técnicos. **Perspect. ciênc. inf.**, Belo Horizonte, v. 18, n. 3, p. 111-129, jul./set. 2013. Available at: http://www.scielo.br/pdf/pci/v18n3/08.pdf. Access at: 20 abr. 2020.

HEDSTROM, Margaret. Digital preservation: problems and prospects. **Digital Library Network (DLnet)**, [*s. l.*], n. 20, Mar. 2001. Available at: http://www.dl.slis.tsukuba.ac.jp/DLjournal/No_20/1-hedstrom/1-hedstrom.html. Access at: 20 abr. 2020.

INNARELLI, Humberto Celeste. Os dez mandamentos da preservação digital: uma brevíssima introdução. *In*: SEMINÁRIO SERVIÇOS DE INFORMAÇÃO EM MUSEUS, 2., 2012, São Paulo. **Anais** [...] São Paulo: Pinacoteca do Estado de São Paulo, 2014. p. 317-325. Available at: http://biblioteca.pinacoteca.org.br:9090/publicacoes/index.php/sim/article/view/57/54. Access at: 20 abr. 2020.

INNARELLI, Humberto Celeste. Preservação digital e seus dez mandamentos. *In*: SANTOS, Vanderlei Batista dos; INNARELLI, Humberto Celeste; SOUSA, Renato Tarciso Barbosa de. (org.). **Arquivística**: temas contemporâneos: classificação, preservação digital, gestão do conhecimento. 3. ed. Brasília, DF: Senac, 2009. p. 21-71.

LAVOIE, Brian; GARTNER, Richard. Preservation metadata. 2nd edition. **DPC Technology Watch Report**, [*s. l.*], v. 13, n. 3, p. 1-36, May 2013. Available at: https://www.dpconline.org/docs/technology-watch-reports/894-dpctw13-03/file. Access at: 20 abr. 2020.

LIBRARY AND ARCHIVES CANADA. **Guidelines on file formats for transferring information resources of enduring value**. [Ottawa], Feb. 2015. 25 p. Available at: https://www.bac-lac.gc.ca/eng/services/government-information-resources/guidelines/Documents/file-formats-irev.pdf. Access at: 28 ago. 2019.

LONG, Andrew Stawowczyk. **Long-term preservation of web archives**: experimenting with emulation and migration methodologies. [*S. l.*]: International Internet Preservation Consortium (IIPC), Dec. 2009. 54 p. Available at: https://www.ltu.se/cms_fs/1.67312!/file/LongtermPresOfWebArchivesOsv.pdf. Access at: 20 abr. 2020.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Fundamentos de metodologia científica**. 8. ed. atual. São Paulo: Atlas, 2017. 368 p.

MÁRDERO ARELLANO, Miguel Ángel. Cariniana: uma rede nacional de preservação digital. **Ci. Inf.**, Brasília, DF, v. 41, n. 1, p. 83-91, jan./abr. 2012. Available at: http://revista.ibict.br/ciinf/article/view/1354/1533. Access at: 20 abr. 2020.

MÁRDERO ARELLANO, Miguel Ángel. **Critérios para a preservação digital da informação científica**. 2008. Tese (Doutorado em Ciência da Informação) - Departamento de Ciência da Informação e Documentação, Universidade de Brasília, Brasília, 2008. Available at: http://repositorio.unb.br/bitstream/10482/1518/1/2008_MiguelAngelMarderoArellano.pdf. Access at: 20 abr. 2020.

NATIONAL ARCHIVES OF AUSTRALIA. **Long-term file formats**. Canberra, [c2020]. Available at: https://www.naa.gov.au/information-management/storing-and-preserving-information/preserving-information/preserving-digital-information/long-term-file-formats. Access at: 20 abr. 2020.

NATIONAL INFORMATION STANDARDS ORGANIZATION. **Understanding metadata**. Bethesda: NISO Press, c2004. 16 p. Available at: https://www.lter.uaf.edu/metadata_files/UnderstandingMetadata.pdf. Access at: 20 abr. 2020.

NATIONAL LIBRARY OF AUSTRALIA. **Digital preservation policy**. 4th ed. [Canberra], Feb. 2013. Available at: https://www.nla.gov.au/policy-and-planning/digital-preservation-policy. Access at: 20 abr. 2020.

NATIONAL MUSEUM OF AUSTRALIA. **Digital preservation and digitisation policy**. Version 2.2. Canberra, Aug. 2012. 11 p. Available at: https://www.nma.gov.au/about/corporate/plans-policies/policies/digital-preservation-and-digitisation. Access at: 20 abr. 2020.

PEARSON, David; DEL POZO, Nick. **Explaining pres actions**: a working document. [Canberra]: National Library of Australia, Nov. 2009. 42 p. Available at: https://www.nla.gov.au/content/explaining-pres-actions-a-working-document. Access at: 20 abr. 2020.

PENNOCK, Maureen. Web-Archiving. **DPC Technology Watch Report**, [*s. l.*], v. 13, n. 1, p. 1-45, Mar. 2013. Available at: https://www.dpconline.org/docs/technology-watch-reports/865-dpctw13-01-pdf/file. Access at: 20 abr. 2020.

ROCKEMBACH, Moises; PAVÃO, Caterina Marta Groposo. Políticas e tecnologias de preservação digital no arquivamento da web. **RICI: R.Ibero-amer. Ci. Inf.**, Brasília, DF, v. 11, n. 1, p. 168-182, jan./abr. 2018. Available at: http://periodicos.unb.br/index.php/RICI/article/view/8473/7060. Access at: 20 abr. 2020.

ROTHENBERG, Jeff. **Avoiding technological quicksand**: finding a viable technical foundation for digital preservation. A Report to the Council on Library and Information Resources. Washington, DC: Commission on Preservation and Access and Council on Library and Information Resources, Jan. 1999. 35 p. Available at: https://clir.wordpress.clir.org/wp-content/uploads/sites/6/pub77.pdf. Access at: 20 abr. 2020.

SANTOS, Henrique Machado dos; FLORES, Daniel. As estratégias de emulação como fundamento para a preservação de objetos digitais interativos: a garantia de acesso fidedigno em longo prazo. **Informação Arquivística**, Rio de Janeiro, v. 3, n. 1, p. 95-116, jan./jun. 2014. Available at:

http://www.aaerj.org.br/ojs/index.php/informacaoarquivistica/article/view/79/34. Access at: 29 maio 2020.

SANTOS, Henrique Machado dos; FLORES, Daniel. Os impactos da obsolescência tecnológica frente à preservação de documentos digitais. **Brazilian Journal of Information Science**: research trends, Marília, SP, v. 11, n. 2, p. 28-37, jun. 2017. Available at: http://www2.marilia.unesp.br/revistas/index.php/bjis/article/view/5550/4511. Access at: 20 abr. 2020.

SANTOS, Henrique Machado dos; FLORES, Daniel. Preservação de documentos arquivísticos digitais: reflexões sobre as estratégias de emulação. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, Florianópolis, SC, v. 20, n. 43, p. 3-19, maio/ago. 2015. Available at: https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2015v20n43p3/30007. Access at: 20 abr. 2020.

SANTOS, Henrique Machado dos; FLORES, Daniel. Preservação de documentos arquivísticos digitais: reflexões sobre as estratégias de migração. **PRISMA.COM**, Porto, n. 37, p. 42-54, 2018. Available at: http://ojs.letras.up.pt/index.php/prismacom/article/view/4707/4395. Access at: 20 abr. 2020.

SCHÄFER, Murilo Billig; CONSTANTE, Sônia Elisabete. Políticas e estratégias para a preservação da informação digital. **Ponto de Acesso**, Salvador, v. 6, n. 3, p. 108-140, dez. 2012. Available at: https://portalseer.ufba.br/index.php/revistaici/article/view/6449/4817. Access at: 20 abr. 2020.

SEVERINO, Antônio Joaquim. **Metodologia do trabalho científico**. 24. ed. rev. e atual. São Paulo: Cortez, 2016. 320 p.

SILVA, Edna Lúcia da; MENEZES, Estera Muszkat. **Metodologia da pesquisa e elaboração de dissertação**. 4. ed. rev. e atual. Florianópolis, SC: Universidade Federal de Santa Catarina (UFSC), 2005. 139 p. Available at: https://projetos.inf.ufsc.br/arquivos/Metodologia_de_pesquisa_e_elaboracao_de_teses_e_dissertacoes_4ed.pdf. Access at: 20 abr. 2020.

THE NATIONAL ARCHIVES. **Guidance on cloud storage and digital preservation**. 2th ed. Surrey, Mar. c2015. 39 p. Available at: http://www.nationalarchives.gov.uk/documents/archives/Preserving-Digital-CloudStorage-Guidance_March-2015.pdf. Access at: 20 abr. 2020.

THOMAZ, Katia de Padua. **A preservação de documentos eletrônicos de caráter arquivístico**: novos desafios, velhos problemas. 2004. Tese (Doutorado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2004. Available at: https://repositorio.ufmg.br/handle/1843/VALA-68ZRKF. Access at: 20 abr. 2020.

THOMAZ, Katia de Padua; SOARES, Antonio José. A preservação digital e o modelo de referência Open Archival Information System (OAIS). **DataGramaZero - Revista de Ciência da Informação**, Rio de Janeiro, RJ, v. 5, n. 1, fev. 2004. Available at: http://www.brapci.inf.br/index.php/res/download/45229. Access at: 20 abr. 2020.

THORPE, Clarissa. Trash to treasure: retro computer, software collection helps national library access digital pieces. **Australian Broadcasting Corporation** (**ABC) News**, Sydney, 19 June 2015. Available at: https://www.abc.net.au/news/2015-06-20/collecting-retro-computer-technology-to-save-digital-treasures/6560494. Access at: 20 abr. 2020.

UNIVERSIDADE ESTADUAL DE CAMPINAS. Procuradoria Geral. **Resolução GR-017/2011, de 29 de junho de 2011**. Estabelece diretrizes e define procedimentos para a gestão, a preservação e o acesso contínuo aos documentos arquivísticos digitais da Universidade Estadual de Campinas. Campinas, SP, 1 jul. 2011. Available at: https://www.pg.unicamp.br/mostra_norma.php?id_norma=3057. Access at: 20 abr. 2020.

UNIVERSIDADE ESTADUAL PAULISTA. **Política de preservação digital para documentos de arquivo da Unesp**. Versão 1.0. São Paulo, dez. 2017. 14 p. Available at: https://www2.unesp.br/Home/cppd/politica-de-preservacao-digital--para-documentos-de-arquivo-da-unesp---v-1.0.pdf. Access at: 20 abr. 2020.

WILLER, Mirna. *et al*. Selective archiving of web resources: a study of processing costs. **Program:** electronic library and information systems, [*s. l.*], v. 42, n. 4, p. 341-364, Sept. 2008. Available at: https://www.emerald.com/insight/content/doi/10.1108/00330330810912043/full/html. Access at: 20 abr. 2020.

Article submitted to the similarity system

Submitted: 22/04/2020 – Accept: 22/05/2020 – Published: 09/06/2020