Health Information and Libraries Journal

Challenges to digital preservation of clinical information in the new info-communicational reality

Journal:	*Health Information and Libraries Journal*
Manuscript ID	Draft
Manuscript Type:	Regular Feature: Teaching and Learning in Action
Keywords:	Digital information resources, Medical records, Digital preservation, interdisciplinary research

Abstract: Current demands placed by users to health care institutions, regarding the access and management of their clinical information, introduced in this context a new info-communicational reality that involves the adoption of a strategic vision of the role of information management. Accordingly, health care institutions need to deepen their understanding of the information system in order to develop an information strategy to ensure the future preservation and (re)use of clinical information, allowing users an active role in managing their clinical information. With the objective of understanding the power that legitimized users to participate actively in managing their own health information, the present study, supported on the paradigms of complexity by Morin and technological by Castells, addresses the role of digital preservation in Health care as a critical factor in the future communication of the institution with its users, and discusses the different areas of knowledge mobilized in the research on this topic.

Challenges to digital preservation of clinical information in the new info-communicational reality

ABSTRACT
Current demands placed by users to health care institutions, regarding the access and management of their clinical information, introduced in this context a new info-communicational reality that involves the adoption of a strategic vision of the role of information management. Accordingly, health care institutions need to deepen their understanding of the information system in order to develop an information strategy to ensure the future preservation and (re)use of clinical information, allowing users an active role in managing their clinical information. With the objective of understanding the power that legitimized users to participate actively in managing their own health information, the present study, supported on the paradigms of complexity by Morin and technological by Castells, addresses the role of digital preservation in Health care as a critical factor in the future communication of the institution with its users, and discusses the different areas of knowledge mobilized in the research on this topic.

Keywords: info-communication; clinical information; digital preservation; interdisciplinary research

1. INTRODUCTION

Rapid technological developments in the recent years and the emergence of new forms of production and information management, as well as of the communication processes, force us to question not only how institutions evolved and continue to face the challenges but also how individuals position themselves and are prepared to meet the challenges of the Information Society. This question is not new and in 1992, addressing the impact of New Information Technologies (NIT) in society, David Lyon noted that old and reputable social institutions and lifestyles were undergoing a profound transformation (1). After more than 20 years, we recognize that, in many cases, these changes were insufficient or inadequate. This demands a reflection concerning the future challenges that we will face over the next 20 years.

The technological developments and the globalization process reflect on all sectors of activity, not only due to the increased movement of people, goods and images on a global scale, but also due to the existence of contradictory tensions between global trends, local and regional. The socio-economic and cultural contingencies in today's society led institutions to create new relationships with their customers, increasingly active and dynamic. In the health care context, the scientific and technological progress of the last decades of the twentieth century enabled health institutions to adopt latest technologies, to support the production and use of clinical information in the context of care, contributing to efficiency in the health care activities. However, concerning communication policies, the perspective of mandatory medical intermediation in the access to clinical information prevailed in Portugal until the publication of the Law 46/2007, following the EU Directive 2003/98. As a result, the focus on IT adoption was on the support of the activities performed by health

professionals, with emphasis on efficiency and productivity, and not the communicability of information.

The emergence of the "*informed patient"*, usually designated *ePatient*, as a result of the technological developments, introduced in the Health care context a new info-communicational reality. Current concerns with the access to clinical information by the patient / user and the adoption of technologies that allow citizens to play an active role in the access to their clinical information demand investment in communication strategies. Considering that the basis of that communication is the patients clinical records, we must focus on the production, use and preservation of clinical information currently produced and stored, as well as on the definition and implementation of digital repositories to ensure their preservation and safety. In this context, long-term preservation of clinical records must be recognized as a critical factor in the success of any health institution.

In the present work we refer to digital preservation not only as the ability to preserve information but also as the ability to reuse it, assuming that the digital preservation process is inseparable from the ability to access information.(2) In this sense, we use of the terms *digital preservation* and *digital continuity*, as well as the focus on reuse of information, considering that preserving digital information involves recreating it in an appropriate format for its use, either by human or a computer application. Considering the complexity and diversity of aspects involved in the preservation and reuse of digital information it has been the subject of research of many scientific domains, including not only Computer Science, other technology areas but also Information Science and Economics and Communication.

1.1. Motivation and objectives

The slow process of implementation of electronic patient records (EPR), widely discussed in the literature, and the necessary maintenance of paper-based records, along with the electronic patient record (EPR), has allowed institutions to preserve its clinical information. It is important to bear in mind that the current situation in health care is characterized by a hybrid context of paper and digital clinical records, with growing production of born-digital records scattered in various computer applications. Considering that most EPR applications have been adopted without a strategic vision of integrated information management, the question that arises is **which digital clinical information will we be able to preserve and reuse on the long term?** Although an operational approach may limit the discussion to the technological aspects, focusing on the definition of a digital repository, is to necessary to point out its implications on the context of production, use and reuse of the information. The lack of investment in information policies, considering the complexity of information systems as socio-technical systems, and in digital preservation or disaster recovery plans led us to rethink digital preservation based on the

complexity and technological paradigms by Morin and Castells, reflecting on the areas of knowledge involved. The objective of this work is to contribute to (i) understand what legitimized the power of the patients in adopting an active role in the management of their own clinical information, (ii) explore the implications of the new info-communicational reality on the current clinical information management strategy, particularly with regard to digital preservation and (iii) identify the main areas of expertise in addressing the topic of digital preservation.

1.2. Methodology

Regarding the methodological approach, the study is based on literature identified in searches performed on Google Scholar and Web of Science. The results from Google Scholar were the basis for the literature review presented and the Web of Science results were used for the analysis of scientific resources described in Section 3.

The search on Google Scholar using the terms "digital preservation" "digital continuity", retrieved a total of 158 results, 153 without patents or citations. With regard to language, only results in English and Portuguese were included. Bearing in mind the focus placed on the exploitation of scientific areas that contribute to the research on digital preservation results, sorted by relevance, were explored the basis of the abstract and the possibility of access to the full text, as well as the contribution to the work. In a second phase, a new search using the expression "digital preservation" OR "digital continuity" resulted in 16,700 results, and adding the term AND (healthcare OR "health care") we retrieved a total of 1 810 results. The search results (without patents and citations and only in English and Portuguese) were restricted to a temporal universe between 2010 – 2016 and the search resulted in 1 380 results, sorted by relevance and explored by applying the same criteria as the results obtained in the previous research. Section 3 presents a description of the analysis performed using the database Web of Science.

2. THE EMERGENCE OF THE “ePATIENT” AND THE IMPACT ON THE DOCTOR-PATIENT COMMUNICATION

The changes occurred in organizations, as a result of the technological innovation, must be understood transversely, bearing in mind the interactions between individuals in institutions. As Morin notes (3) in a company we cannot lay on one side the company's diagram and its production program and on the other side its other problems of human relations. The two processes are inseparable and interdependent. With regard to the health care sector, the impact of the Information Society led to profound changes in the relationship between doctor and patient. The phenomenon known in the literature as *patient empowerment* or *ePatient revolution*, as opposed to a context in which medical knowledge was owned by the health professional, has put health information available to the patient via the Internet. We refer not only to basic information about

health but also to specific scientific information about a particular disease or recent treatments in clinical trial phase. The ePatient, that on the present study we chose to designate as "informed patient", goes to the doctor already informed about diagnosis and therapy possibilities, as well as the rights of access to his / her clinical information. The ePatient wants to play an active role in his / her diagnostic and therapeutic decision making and demands to be informed of all clinical decisions.

In order to understand what contributed to the legitimate active participation of the patient in managing his clinical information it is important to note the American legislation HIPAA (Health Insurance Portability and Accountability Act), from 1996, and the European recommendations for protection of medical data (Recommendation No. R (97) 5), 1997. Both refer to the patients' right to play an active role with regard to their clinical record and to be informed of its content. In the European context it is important to note the Directive 2003/98 on the reuse of information in the public sector, which was transposed to the Portuguese law by Law 46/2007, recently substituted by the Law 26/2016 from 22 August. These changes defined the creation in all public institution of a person responsible for the access to information, with implications for the health sector, no longer maintaining medical intermediation as mandatory. Concerning the access to clinical information we highlight the work of the Northern Regional Health Administration in 2008 (4) and the various legal advices issued by the Commission Access to Administrative Documents[1], as well as specific legislation for health information, such as the Mental Health Act (Law Nº 36/98 of 24 July), the Law of Genetic and Health Information (Law Nº 12/2005 of 26 January and decree -law 131/2014 of 29 August) and the law on clinical research (Law Nº 21/2014 of 16 April). It is important to mention the contribution of the National Data Protection Commission, regarding the development and implementation of information management software in the context of health, especially concerning the case of the national recently created Health Data Platform[2]. In the macro context, it is important to stress the impact of the Patient Protection and Affordable Care Act (PPACA), commonly known as "Obamacare," on the ICT in Health care, contributing to boost the technological modernization in the health sector. In the near future it it is important to be aware of the changes that will be promoted in the European context due to the introduction of the Data Protection Regulation, Regulation (EU) 2016/679, on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, revoking the Directive 95/46/EC (General Data Protection Regulation).

In the Portuguese context, there are examples of initiatives of health institutions and of the Health Ministry designed to give patients an active role in managing their clinical information. We must refer the development the National Health Portal that allows patients to add information to their

[1] Available from http://cada.pt/

[2] Health Data Plataform available from http://www.portaldasaude.pt/portal. Authorization 940/2013. Retrived October, 28, 2014 from http://www.cnpd.pt/bin/decisoes/Aut/10_940_2013.pdf and Autorização 3742/2012 Retrived October, 28, 2014 from http://www.cnpd.pt/bin/decisoes/Aut/10_3742_2012.pdf.

clinical record, indicating allergies, therapeutic and monitoring data that may be relevant in a situation of care. Moreover, the portal also provides patients' access to the identification of health professionals who have accessed their clinical information via the Health Data Platform. This power, gradually given to patients by the increasing use of technologies in health care has been growing in the last years. The evolution that started with the use of the Internet to conduct research on health information (which can be analysed on google trends) has resulted in the creation of *online* portals that promote interaction between the user and the health institutions. The fact that nowadays the discussion focus on the implementation of functionalities that give patients the ability to allow or disallow the shared access to their health information using their personal mobile phone or have access to the access logs to their clinical record is a strong evidence of the ePatient revolution.

In the literature there are studies that address the patient doctor relationship in the new info-communicational reality, presenting a characterization of ePatient profile, which has changed over time. In 2004 Akerkar described the classic ePatient as young, the number of women exceeded the number of men, with a high level of education and more likely to belong to a group with high income. The study pointed to 63% of searches about a particular medical condition, 47% about treatment or medical procedures, 44% on nutrition and 36% on physical exercise. These are described as critical patients concerning their health problems (5). The active role of patients regarding their health information grew with the technological development and the focus of the current literature is on the emerging challenges faced by health institutions in promoting the involvement of the patient in health care decision making (6). The importance of the active patient is more than a simple technological infrastructure problem of information mediation, focusing on the impact of the patient power on the health care process and, consequently, reducing health care costs.

Making patients active players in the delivery of care is considered a possible solution to reduce health care costs: the more active and literate patient, able to engage in the management of his / her care might help reduce doctor shopping, the length of hospitalizations, and poor clinical outcomes (6).

Some studies concerning the patient access to clinical records suggest an improvement in doctor-patient communication and in patient education, helping patients to follow medical recommendations more carefully. Despite the possible difficulties in understanding parts of their clinical records studies indicate the patient satisfaction with the access experience (7,8). Although transparent access to the clinical record points out benefits concerning the clinical condition of the patient, contributing to increase the patient trust in the institution and its health professionals, it is important to mention recent news on the Portuguese Health Portal, pointing to a total of 833.000

users, two years after its implementation. The number of users is considered low, when compared with the initial number that aspired to 2 million users[3]. Explanations to this fact refer the users' digital literacy as a critical factor in IT investment to streamline communication with users. In this context, the term *digital literacy* is more than the technical capability to use digital devices, including cognitive skills used to perform tasks in digital environment. The contribution of Eshet (9) to the discussion of digital literacy points to the exercise of 6 different digital thinking skills that include: (i) photo-visual skills (understanding of picture messages); (ii) reproduction skills (use of digital reproductions to create new content with meaning from pre-existing content); (iii) branching skills (build knowledge from hypertext navigation); (iv) information skills (to critically assess the quality and validity of the information); (v) socio-emotional skills (understanding the dominant rules in cyberspace and apply their understanding in virtual communication); and (vi) real-time thinking (the ability to process huge volumes of simultaneous stimuli, video games or *online* learning) (9). The 6th competence, added by the author to the work of Eshet-Alkalai (10), focuses on current user behavior in the digital environment, bearing the impact of the technological developments in studies on digital literacy. In this sense, addressing patients' digital literacy in the National Health Service (NHS) context, based on the info-communicational paradigm, could benefit from research on digital literacy that has been carried out, providing a theoretical framework for various projects, contributing to maximize the benefit of IT investment in health care.

3. APPROACH TO DIGITAL PRESERVATION ACCORDING TO THE INFO-COMMUNICATIONAL PARADIGM

In health care institutions, the increasing adoption of IT that we have seen for years, along with the lack of a strategic vision of information management transversal to the organization, has contributed to the high complexity of the information system, understood as socio-technical system (11), i.e. including people organized in social systems (departments or teams) that use technology-based systems. To support the understanding of reality and the need to change and update the organizations, the present work is based on the paradigms of complexity of Edgar Morin (3) and on the technological paradigm of Manuel Castells, recognizing that technology has the ability to transform societies (12).

1.1. The contribution of the paradigms of complexity and technological

In the contemporary world the organizational management is increasingly geared towards improving performance and achieving results. Information management projects are designed to

[3] Gomes, Catarina. Only 838 thousand users registered on the Health Platform (Só 838 mil utentes se registaram na Plataforma de Dados da Saúde). Jornal Público, 28 set 2014. Retrieved October 26, 2014, from: http://www.publico.pt/sociedade/noticia/so-838-mil-portugueses-se-registaram-na-plataforma-de-dados-da-saude-1670996.

simplify reality through mapping of functional processes, focusing on oriented improvements and benefits that often lack an integrated perspective of the organization as an information system. This idea has echo in the perspective of Morin which states that the simplification of the reality that allowed advances in scientific knowledge since the seventeenth century entails harmful consequences that only began to be revealed in the twentieth century. The isolation of different scientific fields and the fragmentation of knowledge do not contribute to the understanding of reality since it is built on action and against the uncertain, a multiple set of interactions and feedbacks (3), and demanding complex knowledge. The organization of scientific knowledge reflects the principles that govern our world perspective, even if not consciously, resulting in paradigms that turn out to mutilate scientific knowledge, revealing useless to understand the complexity of reality. This is the basic idea from Morin that we used in the present work to understand the complexity paradigm that aims at integrating thinking.

Assuming that action presupposes complexity, since it results from chance and unstable or unplanned interactions, we agree that the study of information, structured by human action, cannot be carried out based on a simplified or linear perspective because we cannot eliminate the random, the uncertain. This idea, conveyed by the complex thought of Morin, is also present in the technological paradigm from Castells, since action or innovation does not result from a previously programmed or simplistic environment but from change, disorder and unpredictability. Companies, as Morin stated (3), are orderly and organized but also a random phenomenon, since there is no certainty as to the sale of its products and services, operating in a market where there are possibilities, probabilities, plausibility. In order to keep up the organization must have the ability to regenerate and reorganize, by defining a strategy to take advantage of the disorder in which it functions. This notion of strategy is set by Morin, as opposed to the program concept, which is based on predetermined actions, removing the organization the necessary flexibility to keep up to date.

Considering complexity as a necessary factor to vitality and initiative, Castells discusses the transforming power of technology by analyzing the complexity of today's society. Although Morin states that technology does not determine society, the author recognizes that the interaction between technology and society results in scientific discovery, technological innovation and changes in social practices, pointing the individual culture of freedom, innovation of North American business initiative of the 60s as a cause of the technological progress that started in the 70s (12). Castells argues that society does not determine technology but notes that state intervention can play a decisive role in technological modernization, able to operate significant changes in society (economic power, military and social welfare). Bearing in mind that in the current information model the source of productivity lies in the production and processing of information technology, Castells adopted the designation of **technological paradigm based on information technology** (12). The structural principle of what Castells named *informationalism* is

characterized by a search for knowledge and information through technological development, which shapes social behavior.

Among the central aspects of the technological paradigm Castells points out (i) the information as raw material; (ii) the penetration capability of the effects of technology in all human activity and (iii) the logic of networks, as the driving force of innovation (12). The author explains the relationship between *informationalism* and capitalism, pointing out that the informational capitalism results from the influence of capitalist restructuring that took place since the 80s, on the technological paradigm, boosting social change. The fact that companies react differently to capitalist restructuring and the spread of *informationalism* taking into account its history, culture and institutions emphasized differences. Although technological advances and networks have contributed to significantly increase the ability to process information, we have seen a phenomenon of individual isolation that, according to Castells, forces us to reflect on the interaction between the Network and the Self.

The fact that Castells points out that the technological revolution does not focus on the centrality of information but on the use of information on the production of new knowledge and processing / communication devices, places the focus on the ability to preserve and reuse information, which is the topic of this study. In this sense, the ability of an health care institution to preserve and reuse its clinical information, promoting the production of knowledge about its patients, will be in the future a competitive advantage in the strategic positioning of health care institutions.

1.2. The power of domination - the "global information economy"

Analyzing the contribution of these authors Helena Santos (13) points out that both Morin as Castells put the emphasis on the need to adopt a broader vision that involves the relationship between the subject and the social, incorporating the relations between the social, physical and biological. In this regard, there is an approach of the social sciences to natural sciences to meet the challenges of social structuring processes of the contemporary society. The concept of complex thought by Morin anticipated the need for an integrated knowledge to understand the world, putting the emphasis on the difficulty of interconnecting various types of knowledge. Castells' vision is complementary to Morin as it seeks to understand the changes introduced by ICT, as a determining factor in contemporary social structure. The concept of *informationalism* presented by Castells, as the organizing principle of the network society made possible by ICT, including the influence of the increasingly global financial economy, carries the notion that social changes are shaped by technology, advocating the danger of certain determinism in the changes to be performed.

In the context of the current info-communicational paradigm, social sciences are required to assume a new perspective on the relationship between micro and macro, local and global (14).

According to Castells, the increased human capacity to process and communicate, made possible by technology, is what characterizes this info-communicational contemporary paradigm. The author uses the term "informational" to highlight a specific form of social organization in which the production, transmission and processing of information are source of productivity and power (12). The Network Society presents itself as a new development model, based on information technology, highlighting the ability to recombine or process information with a certain goal. The fact that the Network Society emerges in pre-existing contexts, which are not homogeneous, results in constraints and dangers of domination, which Castells stresses as a power of financial domination, using the term "global information economy."

This idea that the transformations in contemporary society are, to some extent, determined by technology is present in the reflection on the role of digital preservation in health care and the challenges posed to health care institutions. Bearing in mind that health care is based on the patients' trust on the health professional and / or the health care institution, then the ability of a health care institution to preserve and reuse clinical information of its patients will be a critical factor in the future performance of health institutions. Whereas the investment made in ICT in the public sector of health care has not been accompanied by investment in long-term management and digital preservation policies, we can say that in the future the ability of an institution to preserve and reuse the retrospective information of its patients will be a differentiating factor. It should also be noted that when we talk about digital preservation what is at stake is the need to invest in a long-term strategy, as opposed to the pressure that contemporary society and science itself currently suffer. The impact of the technological revolution, which recognizes information as a strategic resource in various activity sectors, is focused on interaction between production and (re)use of information, with an impact on the information science internal paradigm.

1.3. The Information Science paradigm

The need to adopt a scientific approach to information as an object of study and process led to the emergence of Information Science (IS), as part of the scientific paradigm also known as post-custodial, which focus on the properties of information as an object / phenomenon and process that is transformed by the dynamics of its properties (15). Researchers from the University of Porto have been working on the adoption of the scientific model in the information study, as a result of the work from Silva and Ribeiro (15), which focus on the need for an interdisciplinary approach to information as object of study (16). The IS approach rejects the concept of document as object of study, traditionally used by Documentation, considering that the technological advances and the emphasis on information access that promoted the Information Society overcome constraints of physical format or preservation location. The IS approach overcomes the record concept, as a place of document storage, considering the record/ archive as part of the information system, i.e.

actors that produce, use and manage information in a given context, in an integrated perspective. Moreover, the research framework of the informational paradigm, based on the scientific model, allows researchers to address information as a *phenomenon* and *process* (15,17). Considering the theory of complexity of Morin, the information managed by the organizations (object of study) analyzed according to the systemic theory, based on the idea that records/ archives are part of the information system (18), overcomes the disciplinary isolation and mutilation of knowledge, embracing the complexity of the system and relating inputs and outputs. The defence of IS based in a systemic approach contributed to the interdisciplinary approach, maximizing knowledge to apply to the complexity of organizational phenomena, considering the dynamism of its operation and the relationships and interdependencies with the environment, as an open system.

IS still carries a strong influence of the notion of document-information of the Archival and Librarianship disciplines from which it evolved, which is present in the definition of information:

Structured set of codified mental and emotional representations (signs and symbols), modelled with/by social interaction, and capable of being recorded on any material medium and, therefore, communicated in an asynchronous and multidirectional way (19).

Referring the registration material highlights the professional area of IS origin, which focus is on information representation and retrieval in a particular material medium. The definition also emphasizes communication perspective, highlighting the need to establish connections with communication areas. In this sense, the concept of information and the properties of information as an object of study, proposed by the authors, exceed the limitation of communication and statistical aspects from Shannon and Weaver information theory (3), focusing on the memory and the meaning of information. In this context, the integration of IS in the field of Information and Communication Sciences may be benefited from the complementary knowledge of both areas, regarding the production, storage, retrieval and dissemination of information. This perspective seems to be opposite to the Anglo-American influence of *information schools*, which associates IS with technology areas, separated from communication. Silva and Ramos (20) claim that the concerns with a strong professional fields of information management (archivists and librarians) and communication professionals have contributed to the separation of both areas in the professional and in the theoretical field.

From a scientific perspective it is important to understand which areas are focusing on digital preservation as a subject of study. Digital preservation implicates computer science and technological areas, in the construction of repositories to ensure the long-term preservation and security of digital information. Nevertheless, the fact that it involves the production, management, storage and retrieval and reuse of digital information emphasizes IS as a privileged knowledge area in the approach to digital preservation. Bearing in mind that archivists and librarians were

trained to work with computer applications designed as information repositories, and not to support the production of information in a particular activity working simultaneously as repository, **what may be the current contribution of IS for research in digital preservation?**

The fact that digital clinical information is produced and stored in health care applications, which were not designed as repositories but as work tools, causes constraints to its preservation and reuse. Whereas the definition of a digital preservation strategy requires, in certain circumstances, to decide between the preservation of a database and its application environment or the preservation of data in a readable and reusable standard format, it is important to consider information assessment and the costs involved in the decision key factors. Moreover, the fact that most applications are based on systems that store records in a database, displaying a pdf file used to export the information, creates on the user a false sense of preservation of a "document" in PDF format, similar to the physical paper-based document. Apart from authenticity problems, this aspect demands a more comprehensive approach on the definition of a digital preservation strategy. Since it involves the study of the socio-technical system, the assessment of information and the definition of data formats and technical requirements in the development of digital repositories, among other aspects, it is important to question the research areas that have been contributing to the production of scientific knowledge on digital preservation.

2. RESEARCH ON DIGITAL PRESERVATION

The dissemination of ICT and the development of a culture characterized by mass communication, where technology plays an important role in the mediation of information and knowledge, have impact in the scientific field (13). Helena Santos claims that the widespread belief in technology, as foolproof solution to the social and human problems, contributed to reduce scientific thought to the technical issues, resulting in subordination of social sciences to technology. The same author states that Edgar Morin anticipated the danger of scientific subordination, affirming that the physical sciences are not the pillar on which the other sciences are built, and that these are also human sciences, since they appear in human history and society (13). In this sense, the concepts of complexity and informationalism are essential to overcome disciplinary divisions, highlighting the need for an interdisciplinary perspective.

In the present work it is important to discuss the advantages or disadvantages of the contributions of different disciplines to research on digital preservation. Although the information production process and use summons, necessarily, a social science point of view, it is important to note critical aspects as the lack of qualification of social science researchers in understanding technological aspects. Moreover, the lack of recognition by researchers of technology areas towards the added value of social sciences methods seems to disrupt interdisciplinary projects. This highlights the danger of dispersion and fragmentation of knowledge, resulting in a purely technical approach

(Santos, 2014), bearing in mind the ambition of Morin complex thought to articulate between the disciplinary fields that are disrupted by the disjunctive thought (3). The communication difficulties between researchers of different fields, concerning the use of terminology, concepts and principles, as a result of their own internal paradigm, demand the establishment of commonly accepted concepts and methodologies for conducting interdisciplinary research. In addition, we must consider the point of view of different scientific areas on the issue of digital preservation, reflecting on three questions: (i) is technology alone a solution for digital preservation ?; (ii) may the introduction of changes in information behaviour of producers / users be part of the solution? and (iii) do the information producers know that digital information requires long term preservation? (21).

The questions above lead us to the current IS paradigm, which emerges as a science, integrating archival, librarianship and information technology, and the emerging info-communicational paradigm. The theoretical work of Silva and Ribeiro (17) highlight that the decision making process in information management must be approached in an integrated information perspective, from production to the preservation. In this sense, questioning the benefits of interdisciplinary research on digital preservation can only be understood considering the new info-communicational reality. Although IS is a recent area, reflecting strong influence of the professional perspective, it is important to note the existence of interdisciplinary research on digital preservation, integrating IS and Computer Science researchers. The challenge is to understand if carrying out interdisciplinary studies in these areas, mainly oriented to obtain results and problem solving, may result in future progress towards transdisciplinarity or, as Helena Santos alerts, if the weight of the extraordinary technological development we have witnessed is helping to reduce the scientific thought to technique or technological factors (13).

In an attempt to explore the areas of knowledge and views of the various researchers on digital preservation we assume the distinction discussed by Bernstein (22), based on other authors, stating that **interdisciplinarity** demands analysis, synthesis and links between disciplines into a coherent and coordinated whole, while **multidisciplinary** is merely a juxtaposition exercise of different disciplines. The hardest to achieve will be **transdisciplinarity**, i.e. go through the different disciplines and beyond each individual discipline, aiming to rebuild knowledge from scratch by recombining knowledge of various disciplines in order to recognize and analyse a problem from all possible perspectives and in its reality dimensions.

2.1. Identification of domains and research areas

Bearing in mind the need for an interdisciplinary approach on digital preservation we decided to question evidence in the published literature on digital preservation concerning the scientific areas of origin. This allowed us to explore the use scientific domains and research areas classification to

identify trends regarding interdisciplinary research. Searches were performed in Web of Science, and results were analysed concerning the research areas involved. In this context it is important to note the limitation of the classification used for each result, since it depends on the web of Science categories and research area considered for the publication in which it occurs . The firts 12 results are systematized in Table 1.

	Search expression	Timespan=2000-2016	Refined by Web of Science Categories	Results	
	TOPIC: ("DIGITAL PRESERVATION") OR TOPIC: ("DIGITAL CONTINUITY")	Timespan=2000-2016		570	
1			INFORMATION SCIENCE LIBRARY SCIENCE	311	54,56%
2			COMPUTER SCIENCE INFORMATION SYSTEMS	194	34,04%
3			COMPUTER SCIENCE THEORY METHODS	104	33,44%
4			IMAGING SCIENCE PHOTOGRAPHIC TECHNOLOGY	61	10,70%
5			COMPUTER SCIENCE ARTIFICIAL INTELLIGENCE	52	9,12%
6			COMPUTER SCIENCE SOFTWARE ENGINEERING	44	7,72%
7			ENGINEERING ELECTRICAL ELECTRONIC	41	7,19%
8			COMPUTER SCIENCE INTERDISCIPLINARY APPLICATIONS	30	5,26%
8			COMPUTER SCIENCE HARDWARE ARCHITECTURE	29	5,09%
9			ART	15	2,63%
10			TELECOMMUNICATIONS	10	1,75%
11			Refined by: WEB OF SCIENCE CATEGORIES: (INFORMATION SCIENCE LIBRARY SCIENCE) AND WEB OF SCIENCE CATEGORIES: (COMPUTER SCIENCE INFORMATION SYSTEMS)	108	18,95%
12			Refined by: WEB OF SCIENCE CATEGORIES: (INFORMATION SCIENCE LIBRARY SCIENCE) AND WEB OF SCIENCE CATEGORIES: (COMPUTER SCIENCE ARTIFICIAL INTELLIGENCE OR COMPUTER SCIENCE INFORMATION SYSTEMS OR COMPUTER SCIENCE THEORY METHODS OR COMPUTER SCIENCE HARDWARE ARCHITECTURE OR ENGINEERING ELECTRICAL ELECTRONIC)	120	21,05%

Table 1: Searches on Web of Science in 3 of September 2016 refined by Web of Science Categories.

Considering the current literature on digital preservation we used the search terms "digital preservation" or "digital continuity". Since the goal was to identify trends in the fields and current research areas results were restricted to the past 16 years. The analysis of the first ten categories on table 1 (searches 1 to 10) shows that from the 570 total results, 54.5% (311) of the results were classified as Information Science Library Science, and highlight the number of results identified in the following computer science and technology categories, presenting 194 results (34%) for Information Systems and 104 (18,2%) for Computer Science Theory Methods, 61 (10,7%) for Imaging Science Photographic Technology, in-between 52 (9,12%) and 29 (5%) we have 5 Computer Science and Electric Electronic Engineering categories, 15 (2,63%) for Art and 10 (1,75%) for Telecommunications. . The analysis of the distribution of results in these categories points out to the existence of results that are repeated in different categories, which may indicate the existence of trends in the interdisciplinary works published. This is visible in the results obtained when crossing the results in Information Science Library Science and the second category Computer Science Information Systems (line 11) . Considering that the total works in this topic classified as Computer Science Information Systems is 194, data shows that 55,6% of these works are also classified as Information Science Library Science. The same happens when

crossing results classified as Information Science Library Science and Computer Science and Engineering Categories (line 12) we get 161 results, representing a total of 28,5% of the works on the topic in these Computer Science categories (420). In order to overcome the fragilities of the web of Science categorization, which tends to be too broad in some categories, and to further explore the evidence of interdisciplinary work concerning the topic digital preservation in Information Science literature, the same results were refined by research area (Table 2). As in the categories the distribution of the results in research areas indicates that there are results that are repeated, especially the number of results in the areas of Information and Library Science (54,6%) and Computer Science (49,8%). From the results that can be identified as interdisciplinary 21,2% stand out as common to the areas of Information and Library Science and Computer Science.

	Search expression	Timespan	Search refined by Research Areas	Results 2000-2016	
1	TOPIC: ("DIGITAL PRESERVATION") OR TOPIC: ("DIGITAL CONTINUITY")	Timespan=2000-2016		570	
2			**Refined by: RESEARCH AREAS: (INFORMATION SCIENCE LIBRARY SCIENCE)**	**311**	**54,6%**
3			Refined by: RESEARCH AREAS: (COMPUTER SCIENCE)	284	49,8%
4			**Refined by: RESEARCH AREAS: (INFORMATION SCIENCE LIBRARY SCIENCE) AND Refined by: RESEARCH AREAS: (COMPUTER SCIENCE)**	**121**	**21,2%**
5			Refined by: RESEARCH AREAS: (ENGINEERING) AND Refined by: RESEARCH AREAS: (COMPUTER SCIENCE)	31	5,4%
6			**Refined by: RESEARCH AREAS: (ENGINEERING)**	**50**	**8,8%**
7			Refined by: RESEARCH AREAS: (INFORMATION SCIENCE LIBRARY SCIENCE) AND Refined by: RESEARCH AREAS: (ENGINEERING)	7	1,2%
8			**Refined by: RESEARCH AREAS: (ARTS HUMANITIES OTHER TOPICS)**	**6**	**1,1%**
9			Refined by: RESEARCH AREAS: (INFORMATION SCIENCE LIBRARY SCIENCE) AND RESEARCH AREAS: (ARTS HUMANITIES OTHER TOPICS)	1	0,2%
10			**Refined by: RESEARCH AREAS: (ART)**	**15**	**2,6%**
11			Refined by: RESEARCH AREAS: (INFORMATION SCIENCE LIBRARY SCIENCE) AND RESEARCH AREAS: (ART)	6	1,1%

Table 2: Searches on Web of Science in 3 of September 2016 refined by research areas.

Despite the limitations already mentioned, it is possible to identify key research areas concerning research on digital preservation. Although the results involving researchers from Information Science and researchers of Computer Science, possibly influenced by the iSchools model, are not surprising, **the question that arises is whether we are talking about multidisciplinary research with strong technological influence and subordination of Information Science area or if there is in fact an interdisciplinarity exercise.**

2.2. Knowledge mapping from the point of view of the authors

In order to overcome the limitations on the identification of research domains and areas, we decided to explore the results using the analysis of keywords assigned by the authors, which reflect more accurately the researcher's point of view on the work. This methodology was applied to the results of searches 2, 3 and 4 shown on table 2, classified as Information and Library Science and Computer Science. All results that had no keywords identified as chosen by the authors themselves were not considered. The results presented in Figures 1 and 2 below correspond to the first 70 most relevant keywords presented, visually identified according to the frequency of occurrence.

The analysis of the results classified as Computer Science (see Figure 1) highlights the term d*igital preservation* (22,8%), followed by terms that generically represent the perspective of the majority of the authors approach to digital preservation, namely: *digital libraries* (8,45%), *digital storage* (6,20%), archiving (2,25%) and archives management (2,25%). Besides these two last terms there is a high number of terms of medium occurrence that refer to an approach traditionally associated with the Library and Information Science, Archives or Records Management areas as academic archives (1,69%), libraries (1,69%), digital documents (1,69%), national libraries (1,41%) and records management (1,13%). Data shows a lower occurrence referring to technical aspects of digital preservation and standardization, as migration (0,56%), file format (0,56%), *oais* model (0,56%) and s*tandards* (0,56%)), showing the technical approach based on expertise.

When analysing the results classified as Library and Information Science (see Figure 2) the terms that stand out are digital preservation (31,52%), *digital libraries* (5,16%), *cultural heritage* (2,58%) and *digital storage* (2,58%). Terms that get medium occurrence in figure 1 also have significant occurrence in IS like *preservation planning* (2,58%) and *digital curation* (1,243%). Some of the terms that have a medium number of occurrence reflect IS inheritance from librarianship and archival disciplines, namely *digitization* (1,43%) and *heritage* (1,72%) and a significant number of terms that address technological aspects as *migration* (2,01%), *emulation* (1,15%) and cloud computing (1,15%). Besides the occurrence of terms that reflect the application of knowledge from technological areas to traditional IS working context, there is significant number of terms that relate to the research context, like *contextual information* (0,86%), *higher education* (0,86%), continuous function (0,57%) and knowledge management (0,57%), which may highlight a concern with the study of the research social context.

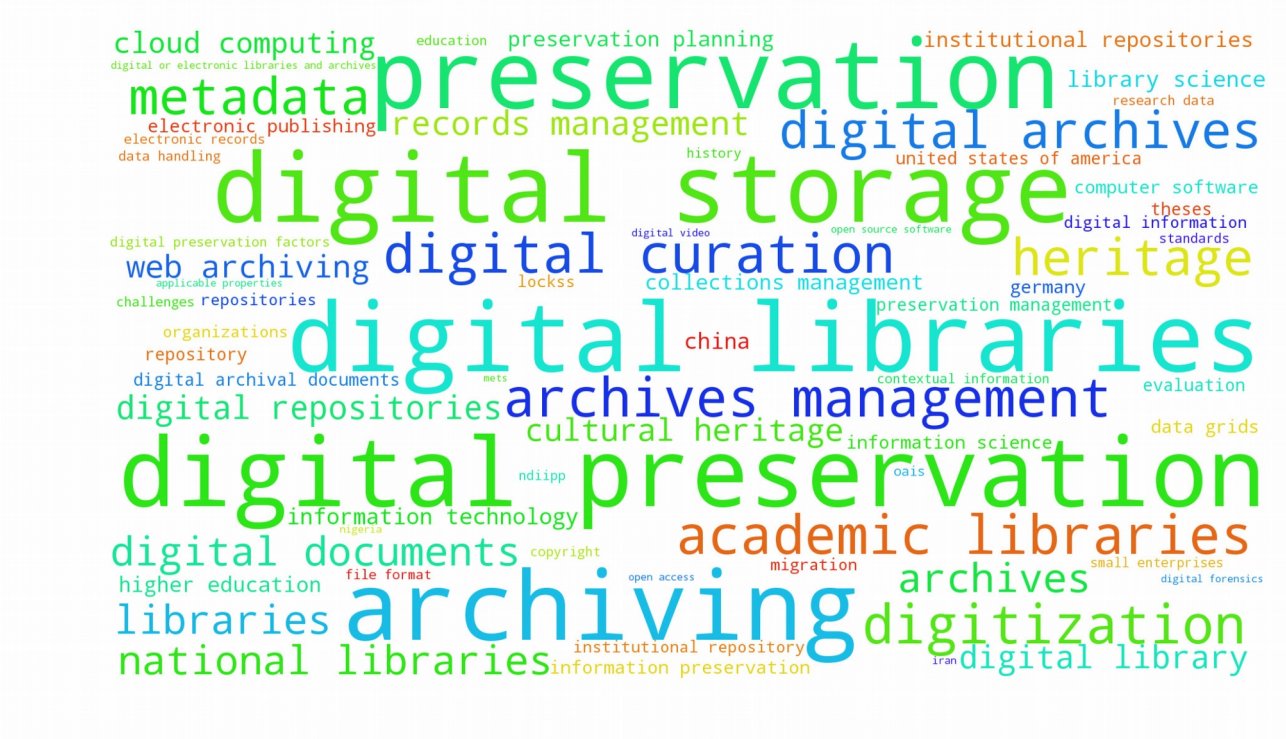

Figure 1: Author keywords used in works classified as Computer Science.

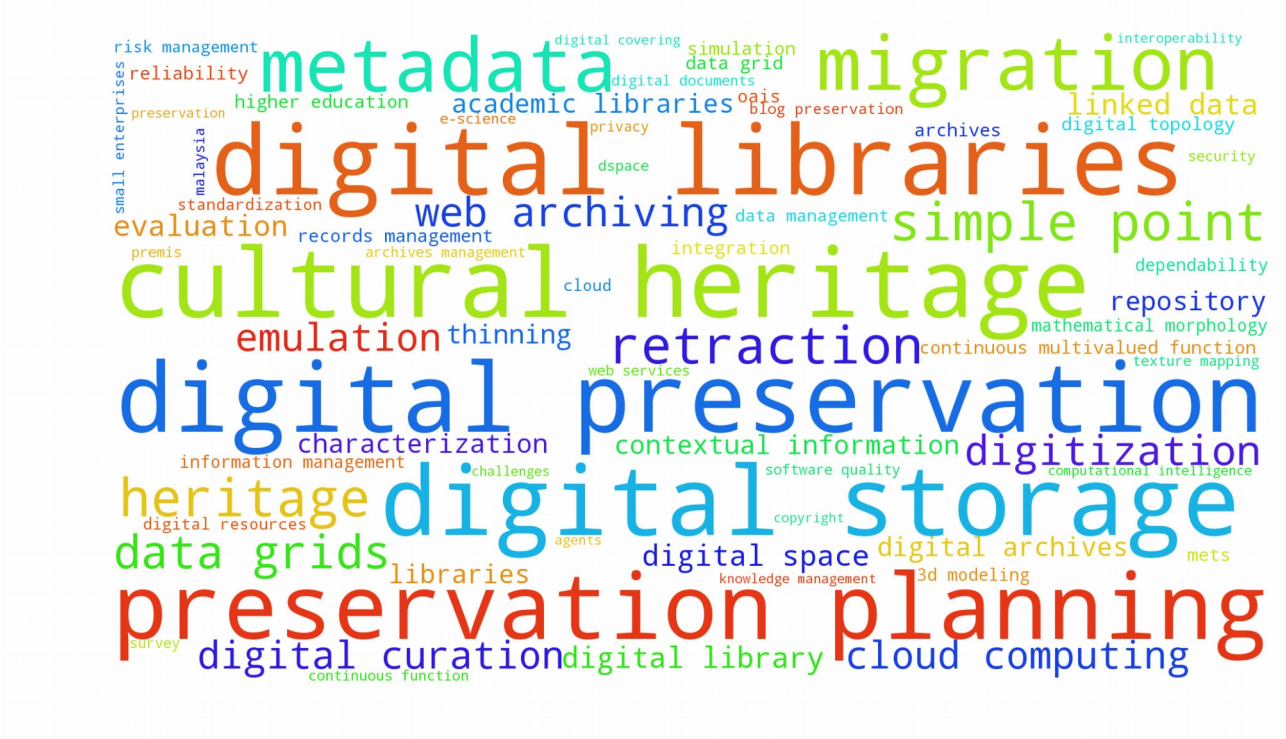

Figure 2: Author keywords used in works classified as Information and Library Science.

The intersection of keywords of both areas is presented in Figure 3. The highlight is on the terms *digital preservation* (23,16%) and *digital libraries* (7,91%), highlighting scientific production on

digital preservation concerning the development of repositories designed to store and retrieve information (digital libraries and repositories). The terms *digital storage* (5,08%), heritage (3,39%), academic libraries (1,69%) and digital archives (1,69%) stand out as a common central concern to researchers, showing participation of IS researchers in the approach to digital preservation with a focus on the production context involving business applications. Crossing both areas highlights that there is a significant occurrence of terms that reveal technological concerns, as well as terms that are traditional concerns of archival discipline, such as evaluation (1,13%), contextual information (1,13%), electronic publishing (1,13%) and records management (1,13%). Results show that although IS is a more recent area, in scientific consolidation, IS researchers interests in digital preservation share the specificity and variety of perspectives of Computer Science researchers.

Figure 3: Intersection of author keywords used in works classified as Computer Science and Information and Library Science.

The keywords analysis indicates that cooperation between the areas of Computer Science and Information and Library Science results in an enrichment of the research, focusing on information storage and highlighting the term *digital curation* that is currently referred as an emergent discipline in IS. There is specific knowledge of each of the areas, which is more evident in Computer Science. The absence of references to terms that focus on the core areas of IS, information representation and retrieval and information assessment, seems to indicate a disproportion in the results classified in both areas. The importance of research in IS, showing a strong influence of

archival and librarianship, points to a contribution of IS professionals to the topic. In this sense, it is positive to note the cooperation between IS researchers and computer science and technology researchers, integrating projects concerning preservation requirements in applications designed as repositories and the definition of production applications requirements of various business areas.

The geographical distribution of the authors' countries origin, taking into account the institutional affiliations, allows a geographical mapping of the scientific contribution of the different researchers. Apart from the possibility to show trends in certain geographical areas, the geographical distribution could allow a more elaborate analysis concerning the research institutions with higher reputation in these areas. The results of the distribution of articles by geographic area are systematized in Figures 4-6.

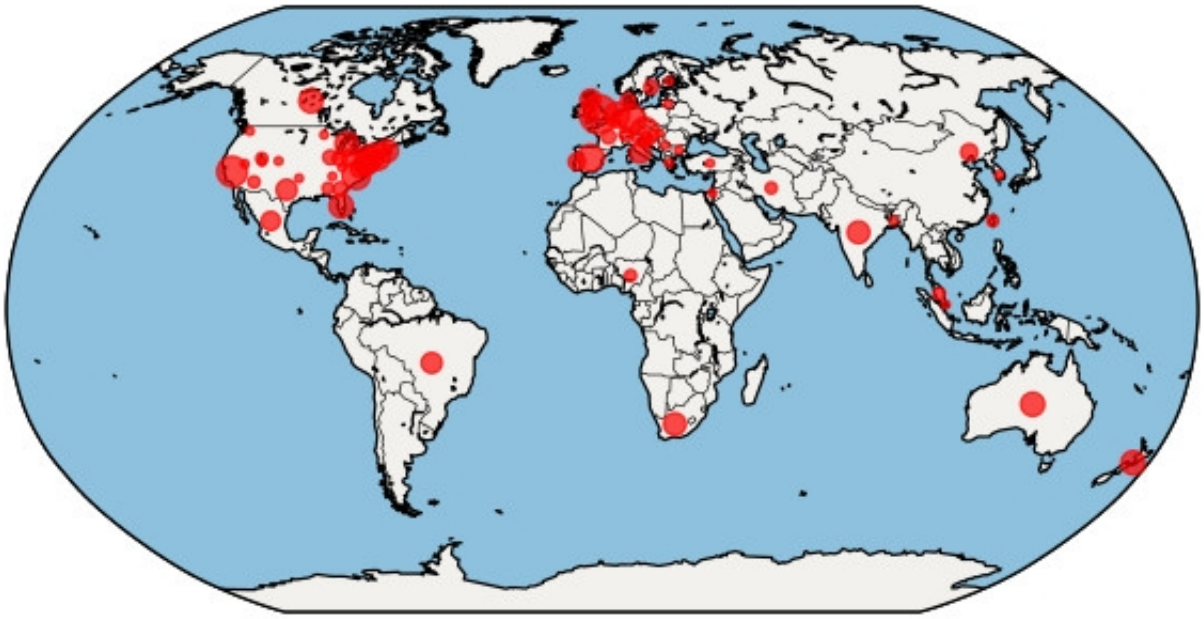

Figure 4: Geographical distribution by institution affiliation of the results classified as Computer Science.

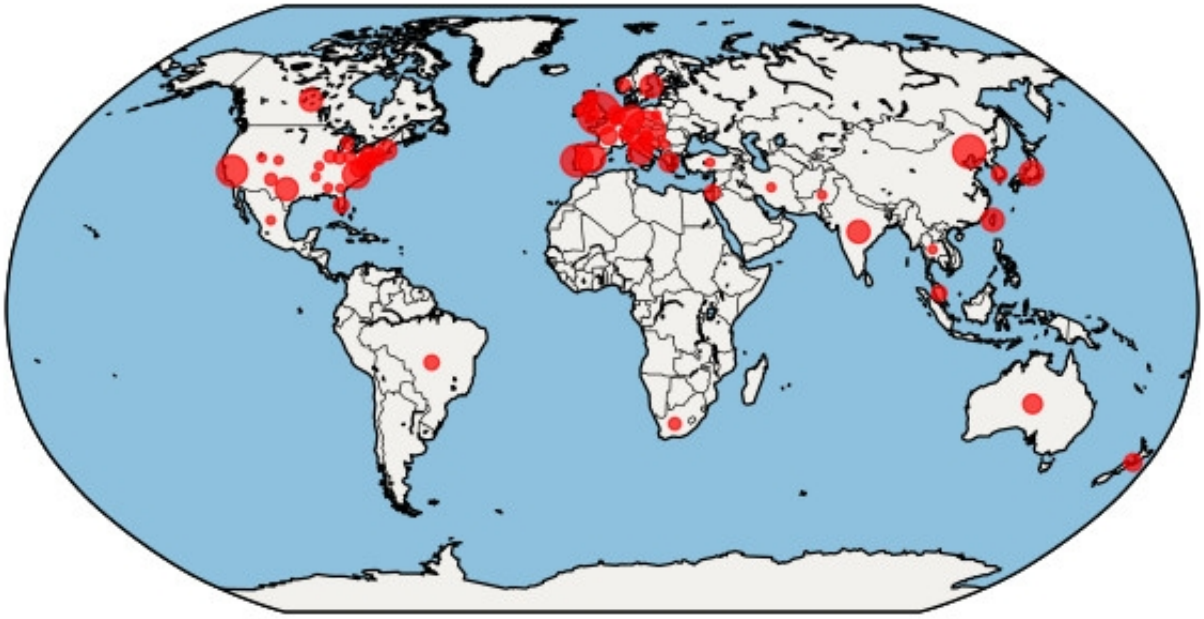

Figure 5: Geographical distribution by institution affiliation of the results classified as Information and Library Science.

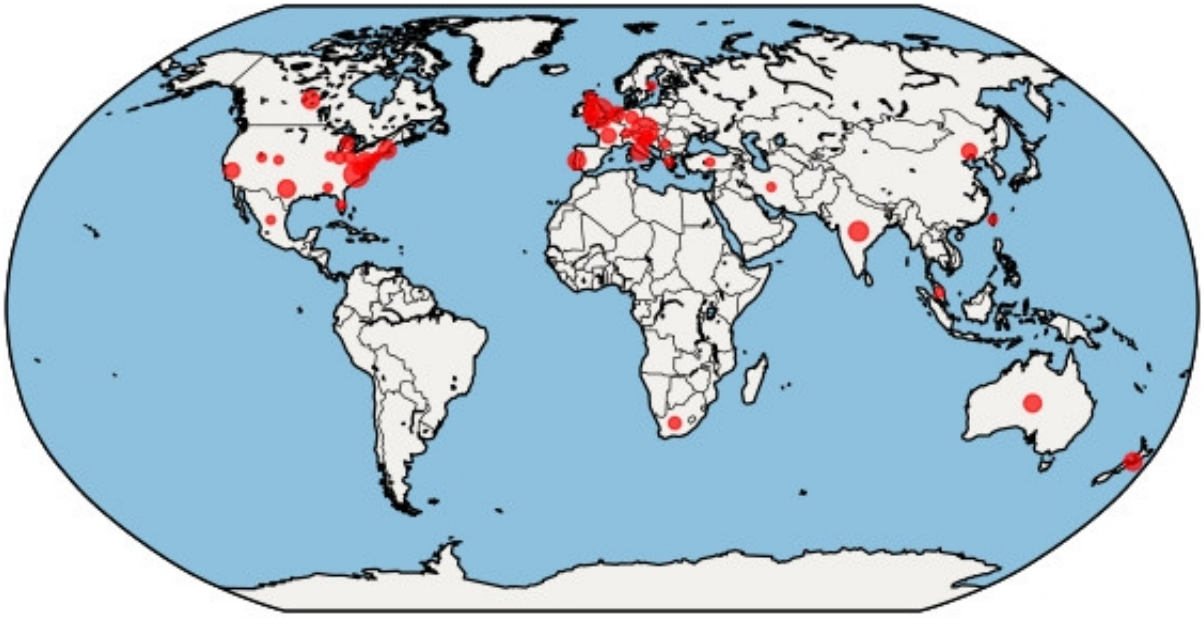

Figure 6: Geographical distribution by institution affiliation of the results classified as Computer Science and Information and Library Science.

A general analysis of the results concerning the intersection of both areas in Figure 6 shows relevant results in England and Central Europe, the US East Coast and Portugal. The three maps highlight England as a leader in the data analyzed, with 7,22% for Computer Science, 9,06% for Information and Library Science and 10,26% concerning the intersection of both areas. Next, it should be noted that the United States of America / California stand out in the areas of Computer

Science, with 4,47% results and 4,71% in Information and Library Science, with a low score of 2,56% results in the intersection of both areas. Austria also stands out with 4,12% results in Computer Science, 8,33% in Library and Information Science and 8,55% in the intersection of two areas. Data show relevance in Central Europe in research on digital preservation. With regard to Portugal and Spain, Portugal has 1,37% results in Computer Science, 5,07% in Information Science and Library and 3,42% in the intersection of both, while Spain has 3,44% results for computer Science, 4,35% for Information and Library Science and not showing any results in the intersection of both areas.

Regarding the USA the intersection of both areas highlights geographic areas of the East Coast showing North Carolina (6,84%), District of Columbia (Washington) (4,27%), Texas (3,42%) and Massachusetts (3,42%). On the Weast Coast It is important to notice California which has the highest occurrences in the US for Computer Science (4,47%) and Information Science (4,71%), although it has a lower occurrence in the intersection of both areas (2,56%) when compared with other states. It is curious to note an incidence in the same geographical areas when analysing a map of iSchools in the USA, revealing its influence in the scientific production.

3. FINAL REMARKS

Data indicates that research in digital preservation has been dominated by Computer Science, in a dominant technological perspective. There is a contribution of Information and Library Science concerning digital preservation research, focusing on research repositories specifically designed for storage and / or information retrieval. Although there is a benefit in the articulation of knowledge of both areas in research on digital preservation, it is important to note the need to include in digital preservation research projects an analysis of the production context and functional circuits, considering the socio-technical information system. Regarding the contribution of archival discipline, as a part of IS, it is important to note that the fact that we are talking about clinical information of living patients, which is dynamic, involves the creation and management of descriptive metadata with different characteristics from the traditional archival description commonly used for archives.

In the Health care context, it is important to have in mind that the digital informational universe to be considered for preservation and future reuse is stored in business applications, designed as working tools, and not applications designed for repository or storage, with the exception of PACS systems for medical image management. This reinforces the need to combine different knowledge regarding the research on digital preservation, integrating health professionals as producers / users. At the same time, considering that the investment to be made and the information assessment will be key factors in the definition of a digital preservation strategy, it poses challenges beyond the areas of Computer Science, Information Science or the Health care,

implicating Communication, Law, Economics and Organizational Management areas, among others. The fact that research on digital preservation, according to the info-communicational paradigm, requires an interdisciplinary approach can encourage the production of new scientific knowledge, in a future perspective of transdisciplinarity, released from the dominant technological determinism.

In the current informational paradigm the consequences of not adopting an interdisciplinary approach to digital preservation research will result in the institutions inability to preserve and reuse information. Bearing in mind that the clinical records are a reflection of the whole process of health care of a patient we can say that in the context of Health care, like in other sectors of activity, it is crucial. Ensuring the access and re-use of clinical information, in response to the challenges of the information society, is much more than a technological issue, imposing itself as a guarantee of citizens' rights and democracy.

REFERENCES

1. Lyon D. A sociedade da informação: questões e ilusões. 1992.
2. Thibodeau K. Overview of technological approaches to digital preservation and challenges in coming years. state Digit Preserv an Int Perspect. 2002;4–31.
3. Morin E. Introdução ao pensamento complexo. 4ª ed. Porto Alegre: Sulina; 2011.
4. ARSN. Relatório do Grupo de Trabalho ad hoc sobre o Direito de Acesso à Informação de Saúd [Internet]. Porto; 2008. Available from: http://portal.arsnorte.min-saude.pt/portal/page/portal/ARSNorte
5. Akerkar SM, Bichile LS, others. Doctor patient relationship: changing dynamics in the information age. J Postgrad Med. Medknow; 2004;50(2):120.
6. Barello S, Graffigna G, Vegni E, Bosio AC. The challenges of conceptualizing patient engagement in healthcare: a lexicographic literature review. J Particip Med. 2014;6(11).
7. Ross SE, Lin C-T. The effects of promoting patient access to medical records: a review. J Am Med Informatics Assoc. The Oxford University Press; 2003;10(2):129–38.
8. Ferreira A, Correia A, Silva A, Corte A, Pinto A, Saavedra A, et al. Why facilitate patient access to medical records. Stud Health Technol Inform. IOS Press; 1999; 2007;127:77.
9. Eshet Y. Thinking in the digital era: A revised model for digital literacy. Issues Informing Sci Inf Technol. 2012;9(2):267–76.
10. Eshet-Alkalai Y. Digital literacy: A conceptual framework for survival skills in the digital era. J Educ Multimed Hypermedia. Association for the Advancement of Computing in Education; 2004;13(1):93.
11. Whitworth B. Handbook of research on socio-technical design and social networking

systems. IGI Global; 2009.

12. Castells M. A Sociedade em Rede, Volume I A Era da Informação: Economia, Sociedade. Lisboa: Fundação Calouste Gulbenkian; 2002.

13. Santos H. Complexidade e informacionalismo: as contribuições de Edgar Morin e Manuel Castells. In: e-Infocomunicação: estratégias e aplicações. São Paulo: Senac; 2014. p. 25–48.

14. Santos H. Complexidade e informacionalismo: as contribuições de Edgar Morin e Manuel Castells. In: Passarelli B.; Silva AM da , Ramos F (org. editors). e-Infocomunicação: estratégias e aplicaçõeses. São Paulo: Senac; 2014. p. 25–48.

15. Malheiro AS, Ribeiro F. Das Ciências Documentais à Ciência da Informação. Lisboa, Afrontamento. 2002;

16. da Silva AM, Ribeiro F, Martins F. Information science and cognitive psychology: A theoretical approach. In: International Symposium on Information Management in a Changing World. 2010. p. 189–99.

17. da Silva AM, Ribeiro F. Information science and philosophy of information: Approaches and differences. In: Luciano Floridi's Philosophy of Technology. Springer; 2012. p. 169–87.

18. da Silva AM, Ribeiro F, Ramos J, Real M. Arquiv{í}stica - Teoria e pr{á}tica de uma ci{ê}ncia da informação. Edições Afrontamento; 1998. 23-44 p.

19. Silva AM da. A informação - da compreensão do fenómeno e construção do objeto científico. Porto: Edições Afrontamento; 2006. 1 p.

20. Silva AM da, Ramos F. As ciências da comunicação e da informação: casos e desafios de uma interdisciplina. In: Passarelli B., Silva AM da,Ramos F (org., editors). e-Infocomunicação: estrat´egias e aplicaç~oes. São Paulo: Senac; 2014. p. 49–77.

21. West B. The Digital Dark Ages: Preserving History in the Era of Electronic Records. 2014;

22. Bernstein JH. Disciplinarity and Transdisciplinarity in the Study of Knowledge. Informing Sci Int J an Emerg Transdiscipl. 2014;17.