

# Making Better out of Technologies: Responses of InterPares to Digital Records Management Challenges

Guanyan Fan<sup>1,2</sup>

<sup>1</sup>*School of Library, Archival and Information Studies, University of British Columbia, Vancouver, Canada*

<sup>2</sup>*School of Information Resource Management, Renmin University of China, Beijing, China*

**Keywords:** InterPARES Project, Digital Challenges, Digital Records, Technology and Information.

**Abstract:** The challenges presented to records and archival professionals by digital records have been extensively researched since the early 1990s. The InterPARES project is one of these early efforts. It progresses along with the evolution of information technologies in the past 20 years of its four phases, and continues to update, enhance and extend its professional knowledge and expertise today. This paper, by setting InterPARES as an example, offers a view on how traditional discipline response to and evolves with the rapidly changing technologies. The project's theoretical foundation and methodological design is first briefly introduced. The main part then elaborates on its research path and contributions via a thread that links the development of digital technologies, their impacts on the management of digital records, and the corresponding responses from the InterPARES project. The paper concludes with a set of meaningful insights drawn from the research experience of the InterPARES project.

## 1 INTRODUCTION

The International Research on Permanent Authentic Records in Electronic Systems (InterPARES) is a large, multinational collaborative research endeavor focusing on the preservation of authentic records created and/or maintained in digital forms. The project was launched in 1999 by Professor Luciana Duranti at the School of Library, Archival and Information Studies at the University of British Columbia. Up to now, the InterPARES Project have developed through four phases: phase 1 (1999-2001) focused on the long-term preservation of authentic records created and maintained in databases and document management systems; phase 2 (2002-2007) focused on records created in dynamic, experiential and interactive systems in the course of artistic, scientific and governmental activities; phase 3 (2007-2012) implemented findings of the first two phases in digital systems in small and medium-sized archival organizations; phase 4 (2013-2018), which is still underway, explores aspects of trust and online digital records.

Back to the late 1980s and early 1990s, records and archival profession, who deals with documented information as primary duty and has its discipline developed and matured in paperwork, was one of the

first communities that experienced the challenges imposed by the digital revolution. InterPARES was a pioneer in this community that had the expertise as well as resources to carry out comprehensive studies. So far, InterPARES has been running consecutively for 20 years with an enduring research interest in the issues of digital records as to their management and preservation. During its four research phases set up with different focuses and goals corresponding to challenges imposed by evolving technologies, InterPARES set a good example by showing how a non-technology community could react in a rapidly changing information age. This paper aims to sort out its research path and contributions via a thread that links the development of digital technologies, their impact on the management of digital information, in particular digital records, and the corresponding responses from the InterPARES project. It includes an introduction to the project's theoretical foundation and methodological design and interpretations of InterPARES findings against the technological background in its four research phases. The paper concludes on the implications drawn from the research experience of InterPARES project.

## 2 THEORETICAL FOUNDATION AND METHODOLOGIES

### 2.1 Theoretical Foundations

The InterPARES project was based on findings of a previous research project, titled “The Preservation of the Integrity of Electronic Records”, otherwise known as the UBC Project. The UBC project was undertaken by researchers including Luciana Duranti and Terry Eastwood, in collaboration with the United States Department of Defence, aimed at establishing standards for creating reliable electronic records and maintaining their authenticity during their active and semi-active life. One of its major findings is a two-phased records lifecycle model: one phase is to the control of the creation and maintenance of reliable and authentic active and semi-active records by the creator, who is supposed to be associated with records management, and the other phase is directed to the preservation of authentic inactive records by the preserver, who is supposed to be associated with archival administration. Yet, the division between the two stages focuses on the different states of the objects under care, the varied goals of ensuring their existence and persistence, and the distinct activities required to achieve the goals, rather than the working relationships between professions affiliated with the two types of management activities (Xie, 2013). The model later became one of the theoretical foundations for the whole InterPARES Project.

The other theoretical basis for the general premises concerning the nature of records and the conditions necessary for ensuring their trustworthiness are theories and methodologies of diplomatics and archival science. Diplomatics is a science originally developed in the 17th century for determining the authenticity and legal validity of individual documents (Duranti, 1999). Over the centuries it has evolved “into a very sophisticated system of ideas about the nature of records, their genesis and composition, their relationships with the actions and persons connected to them, and with their organizational, social, and legal context.” (Duranti et al., 2011) To establish the authenticity of a record, diplomatics analyses a record by breaking down it into various elements contributing directly and indirectly to the establishment of the record’s authenticity, then assesses those elements against the environment in which the record claims to be generated and kept. Whereas diplomatics studies records as individual entities, archival science

studies them as aggregations, which are records linked by “archival bonds” (Duranti, 2010). Archival bond is the network of relationships that each record has with the records belonging to the same records aggregation (Duranti, 1997). Both the professions of records and archives center their work on these aggregations and the organizations originating them (Xie, 2013). During the course of the research, the principles and concepts of diplomatics were integrated with those of archival science to reconstruct the definition of “record” within the framework of electronic systems. A record is defined as “a document made or received in the course of a practical activity as an instrument or a by-product of such activity, and set aside for action or reference. ” An electronic record is decomposed into elements that fall into four categories: documentary form, annotations, context, and medium (MacNeil, 2000). This conceptual analysis set the foundation for identification of records in digital environment.

Despite the above theoretical concepts and principles shared by the whole project, the fourth phase links record with another new concept – “trust”. Trust records are records that can be presumed authentic, reliable and accurate, relying on such factors that can be assessed as intellectual controls, protective measures, data partitioning and processing, legal compliance and risk management, identity and access management, service integrity, and endpoint integrity, etc (Duranti and Rogers, 2016). The relationship of “Trust” links records with activities in every sector and works as the core for the exploration of the fourth phase – InterPARES Trust.

### 2.2 Research Methodologies

The overall methodological principles the InterPARES Project follows are interdisciplinarity, transferability and multi-method design. Researchers with various disciplinary background have been convened as a whole team to better understand records generated in activities of different sectors. They contribute to shaping research processes, analysing data, and formulating final research products with their knowledge of concepts, principles and techniques in different disciplines. Yet as the Project is archival in nature, the work carried out in various disciplinary areas has to be constantly translated into archival terms and linked back to archival concepts as research findings. But once completed, the research outcomes will then be paraphrased against the unique background of each

discipline to make use of them. The whole research process highlights the interdisciplinarity and transferability feature of the project.

To take best advantage from this interdisciplinary intellectual team, the project has applied the principle of openness and flexibility, meaning that each task force, domain, national team, and project/study makes their own choices of methods and tools considered to be the most appropriate. Over the years, specific research methodologies include surveys, case studies, general studies, modelling, prototyping, diplomatic and archival analysis, and text analysis.

### **3 INTERPARES 1 (1999-2001)**

#### **3.1 Technological Background and Challenges**

In the late 1990s, many records that would have traditionally been created and preserved on paper were in electronic form. Databases were applied in organizations and governments to help manage large quantity of data. The proprietary nature of software applications, media obsolescence, and hybrid paper/digital environments were the main challenges faced by phase 1 when considering preserving authentic digital records in systems. Idiosyncratic software systems generated, managed, and stored digital information using proprietary technologies and media that were subject to the dynamism of the computer industry. Most digital information got lost in a self-perpetuating and expensive cycle of obsolescence and incompatibility. Even when luckily reserved, they still needed to experience one or more migrations with radical changes in the configuration and architecture of electronic systems (Duranti, 2005). Furthermore, organizations and individuals generated records in a variety of media and formats (Penn, 1994). In most modern offices it was already quite common for records related to a single matter to exist partly in traditional paper format, partly in an email box, word processing file, spreadsheet file, multiple database tables, or even hypertext linking. Lack of authenticity presented a problem as serious as lack of accessibility, which would render great problems when records were needed as evidence for legal compliance, typically in the court. As computer technology were still under development and staffs in organizations were just getting to be familiar with these “advanced tools”, records were largely exposed to the risk of inadvertent or intentional alteration, either in form or

content, and such alteration might not be readily perceptible. The preservation of records created in electronic systems was thus posing a critical challenge for the archival and records profession.

#### **3.2 Responses by IP1**

##### **3.2.1 Ways of Investigation by IP1**

The InterPARES 1 project, by contrast with the UBC project, took the perspective of the preserver in the two-phased lifecycle model, and sought to establish the means against the digital technological background for assessing and maintaining the authenticity of electronic records once they become inactive and are selected for permanent preservation (InterPARES 1 Project, 2011). The records examined were primarily textual documents produced and maintained in databases and document management systems. InterPARES 1 set up four task forces to investigate the domains of records authenticity, appraisal, preservation, and preservation strategy (Duranti, 2005). Domain 1 by Authenticity Task Force first established a theoretical framework represented by the Template for Analysis to identify the elements of electronic records that are necessary to maintain their authenticity over time. Researchers conducted twenty-two purposively selected, interpretive case studies of electronic systems that contained, or were deemed likely to contain, electronic records. Data gathered through these case studies were then used to test and extend the Template, as to lay the foundation for establishing conceptual requirements for preserving authentic electronic records over the long term. Domain 2 by the Appraisal Task Force set out to determine whether the theory and methodology of appraisal for electronic records differ from those for traditional records, and what role the activities of appraisal play in the long-term preservation of electronic records. It reviewed literature on electronic records appraisal and examined available documentation from archival institutions on their appraisal policies and procedures, as well as reports on actual appraisals of electronic records in practice. Domain 3, Methodologies for Preservation, aimed at identifying and developing the procedures and resources required for the implementation of the requirements and the criteria identified in the first two domains. It gathered empirical survey data about existing programs, plans, and technologies for preserving electronic records. Domain 4 by the Strategy Task Force sought to define the principles that should

guide the development of international strategies and standards for the long-term preservation of authentic electronic records, and the criteria for developing from them national and organizational policies and strategies.

### 3.2.2 Findings and Proposed Solutions by IP1

The Authenticity Task Force defined “authentic record” as a record that is what it purports to be and that is free from tampering or corruption. To assess the authenticity of an electronic record, the preserver must be able to establish its identity and demonstrate its integrity. The identity of a record refers to the attributes of a record that uniquely characterize it and distinguish it from other records. The integrity of a record refers to its wholeness and soundness: a record has integrity when the message that it is meant to communicate in order to achieve its purpose is unaltered, which implies that its physical integrity, such as the proper number of bit strings, may be compromised due to fragility of the media, the obsolescence of technology, and the idiosyncrasies of system. Both the identity and integrity can be demonstrated in metadata related to the record, or in one or more of its various contexts. Based on this conceptual finding, the task force developed the Benchmark Requirements that support the presumption of the authenticity of electronic records before they are transferred to the preserver’s custody when maintained by the creator, and the Baseline Requirements that support the production of authentic copies of electronic records followed by the preserver after the transfer.

The Appraisal Task Force confirmed the shared perception by archivists that electronic records must be appraised from the same theoretical and methodological standpoint as traditional records. They believed that monitoring change and determining its effects on selection decisions is nothing new in traditional appraisal but more pressing in the digital environment. It was the same with the need to appraise records early in their life. They developed a function model of selection using IDEF0, a U.S. Federal Information Processing Standard, in which selection is decomposed into four main activities: (1) managing the selection function, (2) appraising electronic records, (3) monitoring electronic records selected for preservation, and (4) carrying out the disposition of electronic records.

The Preservation Task Force assumed the existence of a paradigmatic shift in the concept of preservation of electronic record: it is not possible to

preserve an electronic record in terms of storage – it is only possible to preserve the ability to reproduce the record. They observed that much attention to the preservation of electronic records focused on the twin technological problems of the relatively short life expectancy of digital media and the rapid obsolescence of hardware and software in practice. But it was the archival and records management criteria instead of technology that determine the best preservation solution, with its appropriateness and adequacy as to the goal of preservation. A "Preserve Electronic Records" model following the IDEF method was designed to set a framework for organizations in developing solutions to the challenges of preserving electronic records. The model includes four main activities: (1) manage the preservation function, (2) bring in electronic records, (3) maintain electronic records and (4) output electronic records.

The main product of The Strategy Task Force was an intellectual framework through analysis and synthesis of results of work in the first three InterPARES domains. The framework contained a set of fourteen principles and corresponding criteria for the development of consistent policies, strategies, and standards adopted in contexts that are administratively, legally, and culturally diverse.

## 4 INTERPARES 2 & 3 (2002-2012)

### 4.1 Technological Background and Challenges

The challenges facing InterPARES in the second and third phase shifts from textual documents produced and maintained in database and static documentary system to records of multimedia form in Experiential, Interactive, Dynamic systems. This change owed to the rise of the technologies featured with “web 2.0”, which included broadband internet, better browsers, Ajax and JavaScript framework, dynamic HTML, Adobe Flash and the mass development of web-based widgets. They resulted in the great evolution of World Wide Web, from the static websites with proprietary applications to a more interactive web with various applications. It was characterized by its mass participation, flexibility and interoperability for end users. It allowed users to interact and collaborate with each other in a social media dialogue in a virtual community and provided users with information storage, creation, and dissemination capabilities that were not possible in the previous environment

(O'Reilly, 2018). These unique features contributed to the uprising of large numbers of complex systems which were active, experimental and dynamic. An interactive system is one in which each user entry or input from another system causes a response from or an action by the system. An experiential system is one that immerse the user in a sensory experience. A dynamic system describes flexible and adaptable approaches tailoring computing resources to demands. A system may be simply interactive, or both experiential and dynamic as usually an experiential or dynamic system is also interactive.

These complex systems, while providing more experimental and dynamic web experience for end users, posed great challenges for records management and preservation. First, what is a record? The interactive and dynamic environment composes various digital entities including but not limited to multi-media websites, online magazines, digital animation products, multimedia performance art pieces, Internet-based filing systems and untraditional databases, like GIS database (InterPARES 2 Project, 2007). There are not only "traditional" records as evidence of the result of action, but also digital documentation relating to the execution process, like original footage and footage logs, so the primary question is "what are necessary to be maintained and preserve". Issues of reliability and authenticity of records is another problem. Are there proper enforcement of access privileges and good control of system security? Are there effective measures of protection against loss and corruption of digital records? Has the creator capture adequate documentary evidence of the occurrence of the information exchange through the systems, as well as enough metadata to verify the identity and integrity of potential digital records? The virtual and hybrid context further exacerbate the problem. Use of traditional hardcopies coexist with the digital movements in office. Digitized copies of analogue material and born-digital entities mix together, manifesting themselves in various kinds of forms. With varied file formats generated by plenty of fancy software and applications, Media fragility and technology obsolescence are always at the core of access and preservation. Apart from those physical technological changes, much more is transforming in the structure and mode of organization transactions, as well as the mechanism of stakeholders and responsibilities. Divergences upon responsibilities and legal liabilities, authorship as well as intellectual property are sure to emerge.

## 4.2 Responses by IP2 & 3

### 4.2.1 Ways of Investigation by IP2

InterPARES 2: Experiential, Interactive, Dynamic Records focused on preservation of authentic records in the context of artistic, scientific and government activities (Duranti and Preston, 2008). The convened a multidisciplinary team with two thirds of its researchers from professions of art, government and government sectors to better understand the nature of the activities generating the records and their function and use in the context of those activities. Each focus was further divided into three domains of inquiry, including records creation & maintenance, reliability, accuracy & authenticity, and appraisal & preservation methods. Another four cross-domains addressed research questions common to all areas of inquiry, and they are: Terminology Cross-domain, Policy Cross-domain, Description Cross-domain and Modelling Cross-domain. All together three focus groups completed 22 specific case studies, that is, 10 in art, 8 in government, and 5 in science. The cases were required to describe in detail, through sets of templates designed by domain 1 and 3, the creator context, analyse the activities resulting in document creation, and examine all kinds of digital entities in complex systems using the diplomatic method. Each case included in its reports a bibliography and a glossary. The literature in the bibliographies were combed by domain 2 to find discussions of authenticity, reliability and accuracy, and of related but differently named concepts. The glossary defined the key terms used in the case study, both for purposes of possible inclusion in the IP2 glossary, and to allow the same definitions to be compared within different disciplines. The findings of the case studies provided much of the data necessary to answer the research questions of all domains and cross-domains. At the same time, each research unit carried out their own General studies, i.e., investigations within its scope for the purpose of achieving its objectives, but not related to specific cases. Topics of general studies included typology of interactive digital music compositions, recordkeeping practices of composers, photographers, GIS archaeologists, preservation practices of scientific data portals, survey of government web site interactivity, and digital formats for long-term preservation.

### 4.2.2 Findings and Proposed Solutions by IP2

The project, upon the observed realities drawn from

complex system cases of different sectors, refined the concept of electronic record as articulated by the first phase. The findings showed those complex systems contain documents which exhibit some variability in form and content, but can be considered records, as when the variability is due to technology rather than to the intention of the author or writer of the document. In addition, authors or writers can generate digital records that embed intentional variability, provided they are properly maintained and managed as intellectually interrelated parts of records aggregations. There were cases, most notably in the arts, but also in government and science of documents whose presentation or rendering always show some unique or spontaneous variation in content or form. This conclusion broadened, rather than contradicted, the finding of InterPARES 1 (Duranti and Thibodeau, 2005).

Domain 1 of Records Creation and Maintenance found that most case studies claimed to be—and were in fact—carrying out a new, non-traditional activity. However, through virtual equation of traditional corresponding activities, it was clear that there isn't such great differences between the traditional and electronic environments. The main change is an increase in the speed with which the process is accomplished and the inclusion of additional steps for verification or to take into account certain features or limitations of the technology used. Domain 2 of Records Reliability, Accuracy and Reliability found that artists, scientists and bureaucrats have very different ideas about the documents they create and reference, what needs to be kept and the features that are essential. However, despite their diversity, the cases shared common problems: technological obsolescence, lack of control over creation procedures, insufficient documentation and uncertainty about what digital objects needed to be saved. This finding resulted in the drafting of the Creator Guidelines, which outline a series of activities to carry out in practice to create and maintain authentic digital materials. The Guidelines were primarily for individuals, but may also be useful for small organizations or groups of individuals. Domain 3 of Records Appraisal and Preservation observed that too many records creators were still neglecting the long-term preservation of their digital files, whether they be static or dynamic, evidential or experiential, historically significant or interactive. They would rather adopt a put-it-on-the-Web preservation strategy, see digitization as a solution, outsource to the vendor or transfer the responsibility to somebody else. Preservation

problems posed by hardware dependencies (especially in arts) and customized or proprietary technologies were obvious as well. In this case, Domain 3 produced the Preserver Guidelines for Organizations. The guidelines are organized according to the sequence of preservation activities presented in the COP model (which will be referred to later) and reflected two perspectives: Actions that would have to be undertaken to avoid some of the situations encountered in most problematic case studies and actions that address the appraisal and preservation concerns.

The cross-domains also generated rich findings and solutions. Policy Cross-domain developed the Framework of Principles Guiding the Development of Policies for records creating and preserving in organizations. The Description Cross-domain developed a Metadata Schema Registry to register relevant metadata schemes and sets and a Literary Warrant Database to identify existing literature requiring the creation and maintenance of archival description and other metadata supporting the preservation of authentic records. Terminology Cross-Domain developed an online terminology database that contains three instruments: Glossary, Dictionary and Ontologies. Modeling Cross-domain developed Unified Models (in IDEF0) of creating, managing and preserving digital objects, including the Chain of Preservation (COP) Model and the Business-Driven Recordkeeping (BDR) Model. The COP model, which depicts all the activities and the inputs and outputs that are needed to create, manage and preserve reliable and authentic digital records and consists of a series of diagrams depicting all the activities involved in the life-cycle management of digital records together with a glossary of all the terms appearing on the diagrams. The model distinguishes four main records activities: (1) managing the framework for the chain of preservation, (2) managing records creation, (3) managing records in a recordkeeping system and (4) preserving selected records. The team also modeled on the business processes of 14 case studies of the total 26 to better understand the environment in which the information (records) was created and used.

#### 4.2.3 Ways of Investigation by IP3

InterPARES 3 “Theoretical Elaborations into Archival Management (TEAM)” phase took the perspective of the preserver translated the theory and methods of digital preservation developed by InterPARES into concrete action plans for archives

and archival/records units within organizations endowed with limited resources (Duranti, 2007). The whole project was organized into 12 national teams, including Canada, Italy, Brazil, China and active participants from other countries. Each national team carried out specific case studies and general studies. Case studies were the investigations carried out by each team, in collaboration with test-bed partners. The 31 Cases initiated all together can be categorized into three types: records cases, recordkeeping cases and policy cases. For each case, context and research data concerning records program of the test-bed organization were first collected for later articulation of research questions and contextual and diplomatic analysis. Then all team members, in form of workshop, reflected on the data and collectively proposed possible solutions as from which action plans for each case would emerge. The plans of action included strategy, protocols, functional requirements, procedures and expected outcomes as needed. The development of action plan took iterative form with several tests before being put into implementation in test-beds (InterPARES 3 Project, 2012a). Compared with the action research in case studies, general studies were the investigations carried out by each Team concerning their own interest against their national background. They were carried out within its scope for the purpose of achieving their own research objectives, but not related to specific cases or test-bed partners (InterPARES 3 Project, 2012b).

#### **4.2.4 Findings and Proposed Solutions by IP3**

The investigation of 31 specific case studies produced tailored plans of action for test-bed partners. The general studies resulted in fruitful achievements as well. However, realizing the inability to present the detailed findings of all the studies, this chapter will only introduce the work of one of the leading teams, i.e., Team Canada as to reveal the research of the whole project. The studies Team Canada undertook can be categorized as preservation foundation, preservation mechanism, and preservation technological system (Xie, 2013).

Team Canada received 20 proposals from testbeds and researchers. It carried out two rounds of data collection and found that none of the digital objects proposed for preservation satisfied all the conditions for being qualified records set by the diplomatic analysis, which indicated a lack of systematic organizational management. The team realized that strong need to build a preservation

foundation, i.e., records management, for those organizations. They identified real records management issues and developed pertinent solutions in terms of action plans. The solutions generated include: organization-wide records management policies and procedures, activity-driven records classification systems, records retention schedules and retrospective records appraisal guidelines. The team went further with open-sourced records management software, e-mail management and records authenticity metadata application profiles. For test-bed partners that already held a better preservation foundations, improved preservation advises include: acquisition policies for university and community records incorporating appraisal guidance in InterPARES 1 that emphasizes authenticity and technological feasibility; documentation framework for acquiring and preserving digital-art works incorporating redefined concept of digital record in InterPARES 2; preservation policies and procedures for university records based on InterPARES 2 principles for the preserver; preservation of educational materials as community archives based on InterPARES 2 guidelines for individual records creators and preservation strategy for a website with identified technology and metadata based on the COP model of InterPARES 2. The team produced various reports with general studies as well, covering such diverse subjects as concrete strategies to preserve access to digitized and born digital records (emails, digital images, social media, websites), management concerning data warehouse and ERDMS Software, better recordkeeping system, and policies, procedures and concrete strategies for digital records preservation. Based on these studies, teaching modules were developed for in-house training programs, continuing education workshops, and academic curricula that provided professionals with the competence concerning preservation of authentic records.

## **5 INTERPARES TRUST (2013-2018)**

### **5.1 Technological Background and Challenges**

Challenges faced in this new era is featured by the problems of trust residing upon issues of security, privacy, risk control and legal compliance against the technological background of cloud computing, big data and open data. Cloud computing is

generally recognized as a model of services delivered to multiple users through a connecting network, regardless of the location of the user and the provider's facilities, provisioned on demand and paid proportionally to usage (Duranti, 2012). Driven by lower costs, organizations today are increasingly moving their records into the cloud. However, a mass of challenges presents when control for records is relinquished to a third-party provider, only to name a few here (Franks, 2015): retained records when should have been destroyed, failed back-ups and unauthorized access by sub-contractors and hackers; proprietary issue when considering system and management metadata and legal compliance issues with unrecognizable jurisdictions. It may also be impossible prove the chain of custody as to verify the authenticity of the records (Duranti, 2015a); to ensure protection of legal privilege or trade secrets when using a third party; and to guarantee that the records that need to be permanently preserved are kept according to archival standards (Bushey et al., 2015). The ultimate essence of all these risks lie in one question: Can you trust the service provider and the records in the cloud (Duranti, 2015b)?

"Big Data" is also causing problems concerning trust. First defined by "three Vs" (volume, variety and velocity) in the IBM report, big data requires extensive data manipulation and mining through the intervention of various non-traditional technologies and tools (Zikopoulos and Eaton, 2011). Online governments, businesses and social media are amassing huge volumes of data to provide a host of services today. Big data thus fosters a range of democratic objectives, from promoting government transparency to supporting research to contributing to public-private sector goals and priorities – the reliability of data in big data is vital. The same is true with "open data". Open data is a term derives from the original concept of "open information" (McDonald and Léveillé, 2014). As one of the initiative that composes "open government", it is rooted in the objective to increase government transparency, generate public input and interest and stimulate social and economic development (Herly et al., 2016). The issues presented by this scenario of "big data and open data" are clear: Can the data be trusted? How and where are they stored? How secure are they? Will your privacy be protected (Duranti and Jansen, 2013)? With exponential growth of and reliance on Internet services today, and the waning level of public confidence in public and private organizations across, people are increasingly questioning how much they can trust digital information available on the Internet.

## 5.2 Responses by ITrust

### 5.2.1 Ways of Investigation

The fourth phase, InterPARES Trust (ITrust) starts from 2013 and related studies are still underway today. The project presumes the existence of an advanced technological infrastructure and a widespread use of complex technology embedded in the daily routines of people and organizations (InterPARES Trust, 2012). Centering around the topic of "trusted online records", the research is mainly carried out in five domains: infrastructure, control, security, access and legal domain. The infrastructure domain considers issues relating to system architecture and related infrastructure as they affect records held in online environments. The control domain focuses on classical records and archival issues concerning the management of digital material in online environments. The security domain considers online data security issues like security methods, data breaches, risks associated with shared servers, information assurance and risk assessment. The access domain researches open access and open data, the right to know and to be forgotten, privacy, accountability and transparency. The legal domain considers such issues as legal hold, chain of evidence and authentication of evidence offered at trial, etc. There are also five cross-domains, namely terminology, resources, education, policy and social domains. The whole project comprises 7 regional teams and each team investigates topics specified by those domains and cross-domains through individual studies involving trusted records in online environment.

### 5.2.2 Findings and Proposed Solutions

As some of the studies by InterPARES Trust are still in progress and not ready to be integrated to produce a complete book as InterPARES 1 and 2, this section will only give a brief summary of the ongoing investigations under different domains

Up to now, more than 90 studies have been launched with the efforts of more than 300 researchers and research assistants around the world, most of which go to the leading teams of North America and Europe (InterPARES Trust, 2018). Nearly half of the studies fall under the control domain, with a centralized interest in metadata, email management, government e-service and traditional records management issues like arrangement, description, retention and disposition. The 13 investigations carried out under the



infrastructure domain focused on the cloud service and covered topics of types of cloud and their reliability; cloud contractual agreements and their negotiation, cloud storage and repository and trusted certification. The access domain contains 13 investigations as well and researches topics on internet archives, open government (open data and information disclosure), and online public service. The security domain now has 6 studies underway around topics of security methods, risks associated with records management, forensic readiness and protection of authoritative records. The 4 studies initiated under legal domain discuss such specific issues concerning privacy and contracts of and legislations for cloud service. It seems that ITrust researchers show less interest in the cross-domains, but there are still some remarkable efforts to be named. The resource cross-domain has 5 investigations concerned with data sharing and archival services. In education, the North American Team incorporated the InterPARES Trust findings into a Mapping of Archival Competencies. Team Africa also investigated Curriculum Alignments at Institutions of Higher Learning in Africa for professionals to manage records created in networked environments. As for the terminology domain, there is a multinational vocabulary drawn from the emerging and evolving intersection of recordkeeping and information technology, which continues the work of previous InterPARES projects by exploring aspects of trust in cloud environments.

Apart from those national and international studies, preliminary findings related with the theme of the project are published as chapters of books or articles on refereed journals and conference proceedings. ITrust researchers have addressed more than 100 speeches on conference, workshops and seminars. In 2016, The Canadian Journal of Information and Library Science devoted a special issue for the Team North America to discuss data, records and archives in the cloud. The “Preservation as a Service for Trust” project, another effort by ITrust targeting preservation in the cloud, developed a set of requirements that establish a foundation for trusting the preservation of digital information.

## 6 CONCLUSIONS

With its 20 years exploration of the digital world, InterPARES as an example shows how it progresses along with the evolution of information technologies while persisting with its keen interest in preservation of authentic and trustworthy records. Despite the

concrete practices seemingly only valid to the records and archival field, InterPARES provides some significant enlightenments that can be shared by similar communities challenged by this digital trend.

- An adequate and sufficient understanding of the digital technologies. Understanding technologies relating to your discipline is extremely essential. As pioneers of digital records in its field, InterPARES has long realized this truth in its continuing efforts of four evolving phases. The team pointed out the insufficiency of records professionals in grasping and understanding of the digital technologies relating to digital records management tasks (Xie, 2011) and suggested an urgent need to know the new technologies in managing and preserving trusted digital records (Lemieux, 2016). What’s more, InterPARES team is trying to incorporate this insight into education of young professionals. A new interdisciplinary education program named “Digital records forensics” was initiated under its effort in 2011 and now well established as a new stream of study in the Master of Archival Studies (MAS) programme in the University of British Columbia, in partnership with the School of Information at the University of Washington. This program provides courses of digital forensics and information technology other than traditional knowledge of records management and archival science.
- Adaptation and evolution of traditional disciplinary theories. Based on classic thinking in diplomatic and archival science, InterPARES extracted the essence of its basic theories, and refined these traditional concepts, principles and methods by applying them in the digital environment. This indicates a wise and efficient way of thinking for non-technological community when handling with digital issues: to first reflect on traditional concepts and methods and then innovate on existing knowledge instead of a total reconstructing of the theory basis of the discipline, or surging of totally new concepts. This evolutionary progress demonstrates a profession’s capacity to continue enhancing and extending its existing professional knowledge and expertise.
- Interdisciplinary integration and cooperation. InterPARES was born as an interdisciplinary project at its very beginning. This distinctive

nature was particularly manifested in its second and fourth phases, with its most participating experts from professions other than records and archives, including Library, Computer, History, Law, Music, Film, Journalism, Geography, Engineering and Health Sector, to name only a few. This dependence and intersection among disciplines is not only emphasized by the InterPARES project, but also witnessing as a general trend by communities from various fields. Communications between professions shall be facilitated to encourage emerging areas of investigation, eliminate the duplication of theoretical efforts in different fields, and promote consistency of scientific knowledge.

Information and communication technologies have changed and will continue to change the world. Advanced information infrastructures and widespread use of complex technologies have already embedded in the daily routines of people and operations of organizations. Blockchain has become a buzzword, Artificial intelligence is already on its way. With more unpredictable new technologies to come, only those who grasp the key know how to react, survive and expand in this revolutionary information age.

## ACKNOWLEDGEMENTS

The author would like to express her sincere thanks to Professor Sherry L. Xie in Renmin University of China for her guidance and comments on the first draft of the paper.

## REFERENCES

- Bushey, J., Demoulin, M. and McLelland, R., 2015. Cloud Service Contracts: An Issue of Trust. *Canadian Journal of Information and Library Science*, 39(2), pp.128-153.
- Duranti, L., 1997. The archival bond. *Archives and Museum Informatics*, 11(3-4), pp.213-218.
- Duranti, L., 1999. Concepts and principles for the management of electronic records, or records management theory is archival diplomatics. *Records Management Journal*, 9(3), 149-171.
- Duranti, L., 2005. The long-term preservation of authentic electronic records: Findings of the InterPARES Project. [San Miniato, Italy]: Archilab. Accessible at <http://www.interpares.org/book/index.cfm>.
- Duranti, L. and Thibodeau, K., 2005. The concept of record in interactive, experiential and dynamic environments: The view of InterPARES. *Archival Science*, 5(2-4), pp.13-68.
- Duranti, L., 2007. An Overview of InterPARES 3 (2007-2012). *Archives & Social Studies*, 1(1), 577-603.
- Duranti, L. and Preston, R., 2008. International research on permanent authentic records in electronic systems (InterPARES) 2. Padova: CLEUP. Available at: <http://www.interpares.org/ip2/book.cfm>.
- Duranti, L., 2010. Concepts and principles for the management of electronic records, or records management theory is archival diplomatics. *Records Management Journal*, 20(1), pp.78-95.
- Duranti, L., Eastwood, T. and MacNeil, H., 2011. *Preservation of the integrity of electronic records*. Dordrecht: Springer.
- Duranti, L., 2012. Preserving Digital Records: InterPARES Findings and Developments. In: P. Bantin, ed., *Building trustworthy digital repositories: theory and implementation*. Lanham, MD: Rowman & Littlefield Publishing Group.
- Duranti, L. And Jansen, A., 2013. The InterPARES Trust Project—Trust and Digital Records in an Increasingly Networked Society. In *Future2013: information governance*. Zagreb: Department of Information and Communication Sciences, Faculty of Humanities and Social Sciences, University of Zagreb, pp.63-68.
- Duranti, L., 2015a. Preface, *Canadian Journal of Information and Library Science*, 39(2), pp.91-96.
- Duranti, L., 2015b. Digital Records and Archives in the Commercial Cloud.” In CS. Yoo and Blanchette, J. eds., *Regulating the Cloud: Policy for Computing Infrastructure*. Cambridge, Massachusetts and London, England: The MIT Press, pp. 197-214.
- Duranti, L., & Rogers, C., 2016. Trust in records and data online. Integrity in Government through Records Management. Essays in Honour of Anne Thurston, pp. 203-214.
- Franks, P.C., 2015. New Technologies, New Challenges: Records Retention and Disposition in a Cloud Environment/Nouvelles technologies. *Canadian Journal of Information and Library Science*, 39(2), pp.191-209.
- Herly, G., Léveillé, V., McDonald, J., Rovegno, K., Suderman, J. and Timms, K., 2016. The Implications of Open Government, Open Data, and Big Data on the Management of Digital Records in an Online Environment. [online] Available at: [https://interpares.trust.org/assets/public/dissemination/IPT\\_NA08\\_Final\\_Report\\_1Oct2016\\_fordistribution\\_.pdf](https://interpares.trust.org/assets/public/dissemination/IPT_NA08_Final_Report_1Oct2016_fordistribution_.pdf) [Accessed 10 Jun. 2018].
- InterPARES 1 Project., 2002. Project Summary. [online] Available at: [http://www.interpares.org/ip1/ip1\\_index.cfm](http://www.interpares.org/ip1/ip1_index.cfm) [Accessed 9 Jun. 2018].
- InterPARES 2 Project., 2007. Case Studies. [online] Available at: [http://interpares.org/ip2/ip2\\_case\\_studies.cfm](http://interpares.org/ip2/ip2_case_studies.cfm) [Accessed 9 Jun. 2018].
- InterPARES 3 Project., 2012a. Case Studies. [online] Available at: [http://www.interpares.org/ip3/ip3\\_ca](http://www.interpares.org/ip3/ip3_ca)

- se\_studies.cfm [Accessed 10 Jun. 2018].
- InterPARES 3 Project., 2012b. General Studies. [online] Available at: [http://www.interpares.org/ip3/ip3\\_general\\_studies.cfm](http://www.interpares.org/ip3/ip3_general_studies.cfm) [Accessed 10 Jun. 2018].
- InterPARES Trust., 2013. Formal Partnerships. [online] Available at: [https://interparestrust.org/trust/about\\_research/partnership](https://interparestrust.org/trust/about_research/partnership) [Accessed 10 Jun. 2018].
- InterPARES Trust, 2018. Research Dissemination. [online] Available at: [https://interparestrust.org/trust/research\\_dissemination](https://interparestrust.org/trust/research_dissemination) [Accessed 10 Jun. 2018].
- Lemieux, V. L., 2016. Trusting records: is Blockchain technology the answer?. *Records Management Journal*, 26(2), pp.110-139.
- MacNeil, H., 2010. Providing Grounds for Trust: Developing Conceptual Requirements for the Long-Term Preservation of Authentic Electronic Records. *Archivaria*, 50, pp.56-67.
- McDonald, J. and Léveillé, V., 2014. Whither the retention schedule in the era of big data and open data? *Records Management Journal*, 24(2), pp.99-121.
- O'Reilly, T., 2018. *What Is Web 2.0*. [online] Oreilly.com. Available at: <https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html?> [Accessed 10 Jun. 2018].
- Penn, I., 1994. *Records Management Handbook*. 2nd ed. [S.l.]: Routledge.
- Xie, S., 2011. Building foundations for digital records forensics: A comparative study of the concept of reproduction in digital records management and digital forensics. *The American Archivist*, 74(2), pp.576-599.
- Xie, S., 2013. Preserving Digital Records: InterPARES Findings and Developments. In: V. Lemieux, ed., *Financial Analysis and Risk Management*. Heidelberg: Springer, pp.187-206.
- Zikopoulos, P. and Eaton, C., 2011. Understanding big data: Analytics for enterprise class hadoop and streaming data. *McGraw-Hill Osborne Media*.