编者按： 近年来，数字资源长期保存领域经历了从基础理论研究、个体实验，到最佳实践的发展过程。继 2004 年北京、2005 年德国、2006 年北美成功举办之后，2007 年 10 月 11－12 日，数字资源长期保存国际会议（iPRES2007）再次回归北京，由国家科技图书文献中心主办，中国科学院国家科学图书馆组织承办。本次会议主题为“数字资源长期保存：项目进展和最佳实践”，结合具体项目和实践经验，分别从数字资源长期保存的战略计划与基础设施、相关管理问题、技术研究与实践、认证与评估、教育与培训 5 个方面进行了介绍，并就面临的问题和下一步发展进行了交流。

《现代图书情报技术》杂志作为此次会议协办媒体，组织部分参会论文形成专辑发表，集中反映国内外数字资源长期保存领域重大项目的研究进展以及实施路线，供广大研究人员参考、借鉴。

A Foundation for Developing Digital Preservation Policy: The InterPARES Policy Framework

SHERRY L. XIE

School of Library, Archival and Information Studies, University of British Columbia, Canada

sherryx@ interchange. ubc. ca

【Abstract】 The InterPARES 2 project (2001 – 2006) is an international research collaboration on long – term digital preservation and the world largest project of its kind. It brought together 100 researchers from 21 countries, who joined the project with their respective disciplinary expertise, and a large number of graduate research assistants, who contributed to various kinds of research activities. By investigating cases in complex digital environment, where digital information is generated and used to support the conduct of artistic, scientific, and governmental activities, the project has developed both theoretical and methodological knowledge essential to effective preservation of reliable, accurate, and authentic digital records.

The InterPARES policy framework, which contains two sets of complementary principles, is one of the core products of the project. It provides a sound foundation on which preservation policies and procedures can be formulated. Although it is produced based on research examining digital records, for which archival institutions are usually the legitimate preservers, some of them are applicable to other types of digital information such as digital publications, which libraries are charged to maintain. It is believed that the exchange of research findings between libraries and archival institutions – the two major players in the field of digital preservation – will facilitate the preservation of societal documentary heritage and ensure the collective accountability of the memory profession.

【Keywords】 Digital preservation　Preservation policy　InterPARES project　InterPARES policy framework

1　Introduction

In its simplest sense, digital preservation refers to the collection and storage of information created and used in digital format over a usually long time period. The demand or desire for long – term digital preservation implies provision of access to the preserved materials, as the National Library of Australia (NLA) maintains that digital preservation is “the processes involved in maintaining, and if necessary, recovering accessibility to digital information resources” (2002). Digital preservation[1] is a complex matter, requiring research attentions for many aspects it encompasses. Research on digital preservation is the response to the extensive use of digital technologies in creating and using information, the instability nature of such technologies, and the corresponding impact on memory institutions such as libraries and archives. A wide range of research projects started in the early 1990s and have been focusing on the nature of digital objects, preservation strategies[2], rights management, cost and sustainability, and the construction of standards and polices to guide preservation activities[3]. For any program

in organizations to start its preservation endeavor, a policy that provides a guiding framework, incorporates current developments in the field, and is endorsed by senior management should be the first establishment in order for the program to act effectively and grow smoothly. As ERPANET's digital preservation guidance puts it, "A digital preservation policy gives structure and general direction for specific actions … [and it] forms the pillar of a programme for digital preservation. From an external point of view, a written policy is a sign that the organization takes the responsibility to preserve digital material" (2003, p. 3). A well – thought and comprehensive high – level preservation policy should include at least components such as purposes and objectives, scope of the policy, rationales for the policy, nature of the collections, roles and responsibilities, and policy monitoring interval. Many research endeavors on digital preservation have generated recommendations or requirements for constructing preservation policies, which overlap in many areas of general concern but differ mostly in the scope of preservation[4]. The InterPARES project, an international and multidisciplinary undertaking specializing on digital records, is one of them.

Following a brief introduction to the project, this paper presents one of its core products, Policy Framework, first in the context of the project's intellectual organization, then in the course of analyzing individual principles contained within. It examines the applicability of these principles to born digital publications in the complex Web environment and offers its conclusions with regard to the impact of digital preservation on memory institutions such as archives and libraries and to the common goal of effectively preserving and providing access to trustworthy digital materials for as long as they are needed.

2 The InterPARES Project

InterPARES, or IP, is the short name for the research project entitled International Research on Permanent Authentic Records in Electronic Systems, which had run from 1999 to 2006. As its full title suggests, the project operates at an international scale and centers on long – term preservation of authentic electronic records. Although intellectually connected, the InterPARES project had physically undergone two phases[5]: IP1 (1999 – 2001), which primarily examined electronic records in textual format residing in document management system, and IP2 (2002 – 2006), which investigated electronic records that were multi – media enabled and were in interactive, experiential, and dynamitic computing environments. Marked as the world's largest research project of its kind, IP2 united 100 researchers from 21 countries, who joined the project with their distinctive expertise, and a large number of graduate research assistants from participating universities, who contributed to a variety of investigation activities. The tight connections between the two phases are demonstrated by the project's continued theoretical framework, multidisciplinary cooperation, and diversified research methodologies.

The distinction between the two phases is displayed by the shift on research subjects. As an observation of and a response to the advancement in the field of information and communication technology (ICT), IP2 invested a significant amount of its research power into the identification of records and their long – term preservation in working environments where ICT is heavily involved and business is frequently influenced by technological changes. Going through the two phases, the project had maintained its goal of developing both theoretical and methodological knowledge essential to effective long – term preservation of reliable, accurate, and authentic electronic records. The research findings of IP1 are distributed on the project's website[6] and also in a book entitled The Long – term Preservation of Authentic Electronic Records: Findings of the InterPARES Project[7]. The outcomes of IP2 are currently under consolidation; however, some of them (including the portion of research data permitted for public access) can be found on the same website.

3 The Research Outcomes of the Project

The IP2 research outcomes are closely associated with its intellectual organization, i. e., the conception and structure of its research interests. The project established task forces for three investigation domains, ranging from electronic records creation and maintenance, key concepts in managing

electronic records (i. e., reliability, accuracy, and authenticity), and to electronic records appraisal and long – term preservation. Each domain answered their research questions with empirical data collected from three types of focused activities, i. e., arts, science, and government. Four cross – domain task forces, who addressed dedicated inquiries in 1) electronic records and preservation terminology, 2) preservation legislation and regulations, 3) description (i. e., metadata creation) standards/tools, and 4) business/preservation activity analysis, answered research questions that are of relevance to all domains and focuses[8]. Outcomes and findings resulting from these domains, focuses, and cross – domains include full – length reports summarize and conclude their respective research activities, scholarly papers addressing specific issues emerging in research processes, and a set of core products drawn on collective wisdom and analytical results of empirical data and relevant literature. Where necessary, IP2 outcomes incorporate IP1 findings to represent a systematic and renewed view of the project's achievement, and references are made if IP1 findings stand out as independent products. The following table provides a visual representation of the relationships between the project's intellectual divisions and its core research products, with the latter emphasized in bold.

Table 1　IP2 Intellectual Organization and Research Outcomes

Digital Records	**Domain** 1 Creation & Maintenance	**Domain** 2 Management Key Concepts	**Domain** 3 Appraisal & Preservation
Focus 1 Artistic Activities **Focus** 2 Scientific Activities **Focus** 3 Governmental Activities	**Guidelines for Individual Records Creator** **Guidelines for Records Preserver** **Domain & Focus Reports** **Academic Papers**		
Terminology Cross Domain: **Terminology Database**			
Policy Cross Domain: **Policy Framework**			
Description Cross Domain: **MADRAS**			
Modelling Cross Domain: **Business and Preservation models**			

While the table above demonstrates the intention of mapping research products to specific research efforts put by particular task forces, these products should not be considered as discrete units; instead, they are interrelated to or sometimes interdependent on each other for full understanding, as a requirement demanded by the nature of the study subject, electronic records. The nature of electronic records require not only pertinent managerial measures to be conceived for each stage during their lifecycle, but also a systematic and coherent approach that takes into account relationships among stages and the influence that they have on each other. It also needs to be realized that the key concepts resulting from Terminology Cross – Domain are an attachment to all of the research findings and they are essential to interpretation and communication. The interrelatedness and interdependence of these research products will be more explicitly demonstrated in the following section, in which one of the core IP2 products, shortened as the Policy Framework, is the center of discussion.

4　The IP2 Policy Framework

The full title of the Policy Framework reads as A Framework of Principles for the Development of Policies, Strategies and Standards for the Long – term Preservation of Digital Records, and the full text of this document can be accessed at (http://www.interpares.org/ip2/display_file.cfm? doc = ip2(pub)policy_framework_document.pdf). Noticeable in the title is the use of the term "digital" rather than "electronic", which seems to be inconsistent with the wording in the title of the project[9]. The replacement of "electronic" with "digital" reflects a more accurate understanding of the nature of digital objects examined by the project, which are generated and maintained by digital technologies and represented by discrete values such as 1s and 0s. Information objects generated using electronic devices can be either digital or analogue (i. e., continuous signals), which makes "electronic" a more generic, umbrella – like term than "digital"[10]. While carried forward the title of IP1, IP2 has reached consensus among its researchers on this replacement. Thus, the term "digital record" will be used from now on in this paper in accordance with the principles in the Policy Framework.

As introduced in the previous section that the Policy Framework is the outcome of the Policy Cross – Domain task

force which examined legislation and regulations with respect to long – term digital preservation and analyzed data collected by the three focus teams in relation to the interests of the three domains. As a result, in addition to its own research findings, it had in its research path received inputs from all other individual task forces, which allowed its development into one of the project's most significant products. These findings were concentrated and then prepared as guiding principles for the development of standards at national/international level and high – level preservation policies in institutions. The Framework contains two sets of principles, intended for two different groups of audience, i. e., records creators[11] and records preservers[12]. They are designed to complement each other and thereby recommended to be consulted in conjunction. Structuring these principles as corresponding requirements demonstrates the project's emphasis on the collaborative relationship between records creators and preservers, which is considered a vital enabler to the success of digital preservation.

4.1 Principles for Records Creators

Principles for records creators, or creator principles, specify thirteen requirements (labeled as C1 to C13, with C standing for creator) and are organized in the order of importance to the creator's operation. They address aspects of records creation and maintenance, the relationship between records and technologies, quality of records (i. e., accuracy, reliability and authenticity), changed working environment, and rights management. While maintaining a primary interest towards creators' business needs, all principles center on the concern of long – term digital preservation. This approach is supported by the now widely accepted recognition (and also confirmed by IP 1) that, for any preserver to possibly preserve digital records effectively, controls over the lifecycle of the records must start from their creation and remain in each and every stage. It is also justified by the fact that records creators always need to maintain a portion of their records – usually the most important and necessary ones to the continuation of their business operation – for a long even sometimes permanent time, which entails preservation challenges to be understood and preservation activities to be carried out by records creators.

4.1.1 *Records Creation*

Requirements regarding records creation are expressed in C1, C2, and C5. C1 stresses two characteristics essential to the identity of digital records: stable content and fixed documentary form. Requiring content to be stable and documentary form to be fixed is a developed response to the interactive and dynamic environments, such as database – driven websites, examined by the project. One typical feature of such environment is that contents are rapidly modified (including being added to, updated/overwritten, or deleted), and the same content can be easily re – structured either by human intentions or software functions, resulting in difficulties, if not impossibility, to re – access the same records over time. The requirement of stable content is comparatively straightforward and it simply refers to the quality that the intellectual content contained in digital records should remain unchanged or unchangeable after creation. The requirement of fixed documentary form, on the other hand, could include many variations[13]. Its basics, however, remain the same and refer to the presentation of the same appearance of digital records each time accessed – same in the sense that the appearance respects the rules governing records creation established by organizations according to legal and organizational stipulations. Both concepts are vital for qualifying information objects created in digital environment as records because only records with stable content and fixed documentary form are able to serve their legitimate functions of documenting faithfully what they have witnessed, i. e., of being evidencing facts and acts.

C2 introduces the concept of digital component in relation to the constructional nature of digital records. Unlike its traditional equivalent, which has all its parts affixed to a physical medium and thus presents itself as a coherent whole, a digital record consists of many separable components. The separability of components can either be derived from creation intentions (e. g., adding an attachment to an email), or be enabled by application functions (e. g., multimedia Web pages). Although they together offer the appearance on computer screen as a whole, digital components in one record may possess a great degree of technological variations, which may require distinctive strategies for man-

agement and preservation. Digital components of records thus need to be identified and it is desirable to do so at the stage of records creation. The identification of digital components and the design of corresponding strategy for their preservation will make their re – assembly in the future possible, which, at the same time, serve as a means of ensuring accuracy and authenticity.

The third requirement directly addressing records creation is C5, which stipulates that the design and management of the system – in which records are created – should be able to establish the reliability of records against regulatory requirements and best practices. The concept of reliability, along with the concepts of accuracy and authenticity, is the focus of principle C4, which deals with the quality of records in the course of creation, maintenance, and preservation. This principle provides definitions of the three concepts and clarifies their meanings in the context of records management and digital preservation. It also points out their relationships with the processes of record creation, maintenance[14], and preservation[15]. Reliability and accuracy are about the content of a record and are related to its creation processes. Reliability refers to "the trustworthiness of a record as a statement of fact" (InterPARES), and accuracy refers to "the degree to which data, information, documents or records are precise, correct, truthful, and free of error or distortion" (InterPARES). The existence of a record's reliability and accuracy can be verified through examining the completeness of the record's documentary form and the amount of managerial control exercised on the process of its creation. The system that is able to establish record reliability and accuracy is termed by the project as a trusted record – making system, which is defined as "a set of rules governing the making of records, and the tools and mechanisms used to implement these rules" (InterPARES). Such rules include integrating business and documentary procedures, building record metadata schemes (including classification scheme), analyzing records forms, establishing record – making access privileges, and acquiring technological abilities to implement the rules. Reliability is exclusively the responsibility of records creators; however, as it will be demonstrated in the section discussing principles for preservers, accuracy is the responsibility of both the creators and preservers.

4.1.2 *Records Maintenance*

The maintenance of records after creation is addressed by C6, along with C4, for its portion on the concepts of accuracy and authenticity. C6 states that records should be maintained in a trusted record – keeping system in order to protect their accuracy and authenticity. Accuracy at the creation stage is pertinent to the processes of making or receiving records; with respect to maintenance, however, it changes to be relevant to preservation strategies such as media refreshing, file format conversion, and system migration. Because the currently available preservation approaches are unable to offer solutions that fully prevent data from loss when carrying out preservation activities, the content of records needs to be verified every time after, for example, system upgrade. Authenticity is the quality of a record being unaltered after creation, or of being consistent with its initial status (InterPARES). Authenticity does not receive any input from either accuracy or reliability and it is irrelevant to records content. In other words, a record that is reliable and accurate at its creation stage does not contribute at all to its authenticity later and thereby the establishment of authenticity of a record after its creation does not necessarily prove that the record was reliable and accurate. To infer a record's reliability and accuracy based on its authenticity entails a precondition: it was created in a trusted record – making system. On the other hand, to prove authenticity, a trusted record – keeping system in which the record resides needs to be in place. A trusted record – keeping system is a system equipped with "a set of rules governing the capture and storage of records and/or information about records and the tools and mechanisms used to implement these rules" (InterPARES). These rules include defining roles and responsibilities, identifying records, establishing access privileges, etc. and tools include record – keeping metadata scheme, retention schedule, registration system, and retrieval system.

4.1.3 *Long – term Preservation*

Two related principles, C7 and C8, provide guidance on organizational decisions on long – term digital preservation.

C7 informs creators who intend to preserve their digital records over a long time to take a proactive approach, i. e. , to embed preservation measures in all activities related to records creation and maintenance, and C8 recommends that a trusted custodian should be selected for such long – term preservation. A trusted custodian is a preserver who satisfies the requirements of 1) being a neutral third party, 2) possessing necessary preservation knowledge and skills, and 3) having a trusted preservation system. A trusted preservation system is a system consisting of preservation polices, procedures, and technical tools, capable of ensuring records accuracy and authenticity.

4. 1. 4 *Relationships Between Records and Technologies*

Three principles, i. e. , C3, C9, and C13, clarify the relationships between digital records and the technologies employed for their creation and maintenance. C3 offers guidance on the design of organizational information system. Based on case studies, legislation analysis, and industrial scanning, this principle instructs that policies on records making and keeping should focus on the formulation of records management requirements rather than specifying technological requirements. In a general sense, technologies are tools not goals, and in the case of digital records, requiring particular technologies will soon become irrelevant because technologies advance, and they advance rapidly. It is the responsibility of an action plan to design technological solutions that are capable of implementing records management requirements. Such action plan should be reviewed regularly, as this should be required by the policy, to accommodate or incorporate technological developments.

C9 recognizes the changed and changing working environment in which organizations are increasingly relying on centralized databases to carry out their business activities. One typical use of database is the procedure that more than one and sometimes very different business processes are allowed to save and/or retrieve records in one same database. This approach creates convenience for performing job duties but at the same time introduces a potential threat to records management. Records need to be clearly and precisely linked to the activity that brings them into existence so that they can provide evidence for that activity when requested. The project's case studies reveal that the lack of adequate documentations of the linkage between records and the business processes creating and/or using them has caused difficulties for managing such records (e. g. , how to classify and schedule them). It is the project's view that all business processes that contribute to the creation and/or use of the same records should be explicitly documented.

C13 introduces a new rule for records management, i. e. , the identification of original record, which has undergone a significant change in digital environment. An original record in the traditional world refers to its first manifestation that is deemed as perfect and complete and that is capable of reaching the effectiveness intended by its creation. While a digital record still conceptually complies with that connotation, its physicality disappears when it is being saved in the system after creation. Every subsequent re – assembly of the record is a copy of the original that now only exists as bitstream. In accordance with the laws and regulations (e. g. , evidence and electronic signature laws) examined by the Policy Task Force, C13 recommends that reproductions of a record made by the creator in its usual and ordinary course of business have the same effects as the first manifestation, and each is to be considered at any given time the record of the creator.

4. 1. 5 *Rights Management*

C10 – 12 are three principles address the management of rights, namely, intellectual property, privacy, and access, in record – making and record – keeping systems. Rights issues are identified as typical to the creation and use of records in many organizations, which either make or receive records containing third – party intellectual property and/or private information that are protected by legislation. Organizations that cross geographical and consequently juristic boundaries often find themselves in a situation where legislation on access to information varies. If it is not legally granted, reproduction of digital records as the means of preservation or distribution is a violation of such legislation. As legislative developments are always behind the progress of digital preservation, records creators are recommended to identify these rights at the creation stage and to establish handling measures accordingly (e. g. , soliciting permis-

sions from rights holders, or documenting legally – supportive justifications for providing or rejecting access) in order for use and maintenance activities to be smoothly carried out.

4.2 Applicability of Creator Principles to Digital Publications

Creators of digital publications can be anybody now due to the convenience and easiness for individuals to upload digital information into the cyberspace given by Web technologies. In the paper world, publications are conceptually different from records but can acquire a record status under certain circumstances (e. g., used in a business transaction), and this is still valid in the digital world. Publications are stand – alone, self – contained units, and they are produced with multiple copies for the purpose of massive distribution. Records, on the other hand, are byproducts or instruments of practical activities, and inherently possess a net of relationships with their creators, the activities that brought them into being, and other records relevant to them. Without the existence of such relationships, records cannot be understood fully and correctly. Both publications and records contain information and/or knowledge and are capable of serving the functions of education and entertainment. What fundamentally sets records apart from publications is their ability of evidencing the activities they participated. With respect to management, this characteristic has a decisive influence on the way records are organized, described, and preserved.

As its goal dictates, the InterPARES project's digital preservation principles are built around digital records. However, due to the fact that digital records and digital publications are now facing the same challenge of long – term preservation, some of the principles for records creators (the first seven, C1 – 7, to be precise), are considered instructive for producers of born – digital publications as well. Born – digital publications are those that only exist in digital format, as opposed to digitized ones that have equivalents in print format. While publications usually do not need to maintain a quality of being evidence in order to be admitted into court dispute, stable content (C1) should still be a necessity for carrying forward information and knowledge. Stable content enables general publications to be the memory of the society and its people and for scholarly publications to be the basis on which new knowledge can be built. Publishing on the Internet has been gaining popularity and the proliferation of websites has suggested that Web contents are becoming a portion of societal cultural heritage that should not be ignored. However, if contents are constantly updated, modified, or made disappeared, nothing will be left for preservers of websites such as libraries to come to perform their job tasks. Nevertheless, the requirement of fixed documentary form (C1) can be one at publishers' discretion. Depending on the purposes of creation and preservation, documentary form can be fixed if the "feel and look" of the publication in question is desired; or it can be open for changes if only contents matter.

The arguments for stable content go as well for the requirements of reliability, accuracy, and authenticity (C4, C5, and C6), which are traditionally not concerned (in the case of authenticity), or concerned differently (in the case of reliability and accuracy) with the publishing industry. In today's digital world, any publishers who wish to maintain the accessibility of their publications over time should take into considerations of these requirements because they indicate quality and value. They should be, however, mandatory for scholarly e – publications because satisfaction of these requirements provides ground on which the trustworthiness and integrity of scientific works can stand. Reliability and accuracy need to be considered when the content of e – journals and e – books is being transmitted within and between the processes of submission, edit, review, and typeset, and controls need to be established to prevent or correct both human and technology – related errors. Measures for protecting authenticity after the completion of publications need to be established as well. Data and information encoded in digital format and presented in the cyberspace can be easily modified and/or attacked. Authenticity considerations are especially vital for publishers who maintain e – publications in its own database and only distribute online access to libraries, who, then, do not possess local copies. Any corruptions to the copies in the publishers' database may cause the loss of their authenticity. In situations where multiple

copies do exist and verification against each other can be performed[16], each owner of the copies still needs to protect the authenticity of their own copies to allow effective verification.

The principle on digital component (C2) is useful in the sense that digital publications are taking increasingly complex composition forms, with features such as multi – media inclusion and/or cross – database search. While these features make publications more attractive and expressive, the preservation considerations they require are different from those for texts. Identifying digital components expresses the recognition that they need to be separately maintained in order for strategies to be designed according to their distinctive properties. Well – maintained components allow accurate and authentic re – assemblies in the future.

The general guidance on the relationship between records and technology (C3) and the lifecycle management approach (C7) can be adapted to digital publications as well. The utilization of advanced technologies for publications should not compromise the feasibility of long – term preservation, if such preservation is desired. In other words, preservation strategies should be developed in parallel with the technology involved in publishing. The applicability of the lifecycle management approach can be derived from publications' attachment to digital technologies. As the same with digital records, digital publications that are selected for long – term preservation require preservation considerations and activities to be associated with each and every step in both the publishing and maintaining processes.

Principles C8 – 13 are considered record – specific and not applicable to publications. For example, private information (C11) is not a concern for publications, neither is access crossing geographical boundaries (C12). Copyrights (C10) are usually held by publishers themselves, which frees the legal constrain on reproduction. It is worth pointing out, however, that publishers are records creators in doing their publishing business. While the above discussion only focused on the final products of their business activities, i. e., publications; records are generated in the course of carrying out these activities. In that regard, the creator principles could be fully applicable to the born – digital portion of them.

4.3 Principles for Records Preservers

The set of principles for records preservers, or preserver principles, addresses the consideration for long – term preservation by first establishing the role of preservers as trusted custodians (P1), who should qualify the three conditions (explained in C8 in the above section) in order to perform preservation tasks in an effective manner. Seven principles subsequently address areas same to those in creator principles, including the concepts of reliability, accuracy, and authenticity (P2, C4), digital components (P4, C2), relationship between records and technologies (P6, C3), lifecycle management approach (P7, C7), third party intellectual property rights (P8, C10), privacy rights (P9, C11), and access to records in different jurisdictions (P13, C12), yet from the angle of building preservation policies in archival institutions or other entities responsible for preservation.

P3 and P5 together address authentic copies, a concept specific for records preservers. Authentic copies are faithful reproductions of creators' records and are made by preservers for the purpose of preservation. As designated custodians, preservers need to take creators' records into their physical custody, a step that starts archival functions. Since there are no physical original digital records existing in creators' record – keeping system, transferring originals from creators to preservers becomes producing copies and taking the copies into a preservation system. They are termed as copies rather than records because they are not reproduced by their creators in usual and ordinary business processes (C13). In other words, the use of "copy" instead of "record" is an explicit acknowledgement of the fact that originals no longer exist in the digital world, and such copies can only be treated as records when they are at their creators' hands. These copies are considered authentic because they are made by designated preservers, who are charged, usually legally, with preservation responsibilities (P3). Authentic copies made by preservers for preservation purpose hold the same effect as the creators' records do, i. e., they can be used as evidence. P5 emphasizes that, to achieve the goal of preserving memories and evidence, authentic copies can only be made from creators' records, i.

e. , digital objects that possess stable content and fixed documentary form (C1) .

P10 – 12 address archival functions. As a solution to complex working environment where more than one business processes contribute to the creation and use of records contained in centralized databases (C9) , P10 recommends identifying and analyzing these processes when performing archival appraisal. Recognizing the greater possibility of altering records in digital formats, P11 recommends archival appraisal to assess records authenticity through examining creators' record – keeping systems (C6) . As a unique function of archival administration, archival description is presented by P12 as a means of collectively authenticating records aggregations. The archival describing process is a process of analyzing and documenting the various relationships surrounding the aggregation, taking into consideration every event occurring in the time period during which it resides in the creator's record – keeping system (C6) . At the time of accessioning, archival description depicts a comprehensive historical overview of the aggregation and offers a conclusion regarding its collective authenticity. This description is maintained and monitored after accessioning, and updated when new events – for example, performing preservation activities to the aggregation – have taken place. Preservation strategies, such as converting file formats and migrating system platform, require detailed documentation for the purpose of authenticating the transformed aggregation. Such documentation should find its place in archival description.

4.4 Applicability of Preserver Principles to Digital Publications

Digital publications can be preserved by publishers and/or libraries, depending on how subscriptions to e – journal or purchase of e – books or databases are set down[17]. With respect to non – commercial Web publications, it is primarily libraries who, as cultural institutions, are currently assuming the responsibilities of preservation[18]. As introduced in section 4.2, publications conceptually differ from records, but both of them are now facing the same challenge of digital preservation.

One of the InterPARES' most significant findings is the establishment of the concept of record, which takes fully consideration records in digital format. Such establishment gives directions, both theoretically and methodologically, to the identification of digital records in complex technological environment. The identification serves as the first step in performing preservation activities and as a foundation on which policies and procedures governing the creation, maintenance, and preservation of records can be formulated. In the immediate context of the project, to identify records means to recognize them through the guidance of archival and diplomatic theories and with appreciation of their variations caused by various digital technologies. Considering digital preservation a national or even global mission, identifying records then facilitate the separation of different types of information that co – exist in the same environment and the application of pertinent preservation measures. To separate records from publications in the paper world has been straightforward due to the easiness of discerning their respective defining features. As a result, preservation responsibilities are clearly divided and respectively assumed by archival institutions and libraries[19]. However, in complex digital environments, although conceptually records still maintain their connection to business transactions and publications to massive distribution, it has become increasingly challenging to separate them from each other. In the case of organizations' use of websites for conducting business, records and publications are created in the same system and delivered to records addressees and information users through the same channel. A thorough understanding of the nature and characteristics of digital records could facilitate the separation and assist the assignment of preservation responsibilities among culture institutions.

Similar to creator principles, there are seven preserver principles, P1 – 7, which can be applied or adapted to digital publications requiring long – term preservation. The emphasis on preserver being a neutral third party to creators and users is a caution to the fact that digital information can be easily altered when access becomes available. Libraries have traditionally acted as neutral intermediates between information and information users. What is new to them is that libraries now need to take measures to protect authenticity of

digital publications if they locally own the copies of the publications and have access at administration level to them. It is self – evident that knowledge and skills for digital preservation and a trusted preservation system equipped with policies, procedures and tools are of necessity for libraries to effectively carry out preservation activities (P1), because it has become a reality that the number of digital publications in library collections has been consistently increasing. Justifications for the applicability of the concepts of reliability, accuracy, and authenticity (P2), authentic copies (P3), digital components (P4), stable content and fixed documentary form (P5), preservation purposes vs. availability of technology (P6), and lifecycle management approach (P7) to digital publications can be found the same in section 4.2.

P8 – 13 address records – specific issues and archival functions, and are considered irrelevant to publications. However, as preservers of digital publications, libraries produce born digital records in the course of fulfilling its mission. Take the National Library of Australia as an example. Among the digital collections listed in its Digital Preservation Policy are two groups of digital records: The Library's corporate records in digital form and Metadata records of information resources[20]. In this regard, both sets of principles are applicable to the library's records management and digital preservation.

5 Conclusions

Digital preservation presents itself as an enormous challenge to both archival institutions and libraries. Committing to digital preservation entails a systematical approach, with policies, strategies, technologies, and competency in place. Building a preservation policy is the first vital element in this systematical approach. As a foundation for policy construction, the InterPARES Policy Framework reveals the most fundamental characteristics of digital objects requiring long – term preservation and addresses the most significant issues imposed by digital technologies. It offers both theoretical and methodological guidance for a variety of preservation considerations and at the same time leaves space for case – specific adoptions. It does not intend to be comprehensive or all – inclusive and management aspects for a preservation program such as cost and sustainability are not addressed. As a means of complementing, the project develops two sets of guidelines, one for individual creators[21] and the other for preservers[22], and recommends them as tools for implementing an established digital preservation policy.

The analysis on the applicability of the InterPARES principles to digital publications has demonstrated that, while traditionally differ in preservation theories[23], archival institutions and libraries are becoming to have more in common when it comes to digital preservation. Some archival concepts such as authenticity and trusted custodian are proven to be relevant to born digital publications. The analysis has also revealed that the identification of objects for preservation is essential, as it draws boundaries between records and publications and consequently facilitates the defining of preservation scopes. The recognition that various types of digital information co – exist in the Web environment and the gained understanding of their distinctive natures suggest that close collaborations among disciplines relevant to digital preservation should be established at regional, national, and international levels. Such collaborations include, but not limited to, exchanging research findings and best practices, investing joint preservation efforts, maintaining constant communication, and so on. The International Internet Preservation Consortium (IIPC) is an example of collaboration among (national) libraries[24], yet the UK Web Archiving Consortium includes the UK National Archives among its other library partners[25]. It is believed that collaboration will continue to emerge as a central theme in the realm of digital preservation and as a recognized means of effectively preserving society's digital documentary heritage and of ensuring the collective accountability of the memory profession.

Notes:

1 Definitions of digital preservation vary and are usually associated with preservation purposes and information needs of the targeted user community. For example, the definition provided by JISC (Joint Information Systems Committee) stresses the authenticity of the preserved digital materials, states that "Digital preservation is the series of actions and interventions required to ensure continued and reliable ac-

cess to authentic digital objects for as long as they are deemed to be of value," reflecting the demand for both authenticity and accessibility to fulfill the committee's mission of supporting education and research.

2 The term "preservation strategy" here is used to refer to the technological solutions for digital preservation such as refreshing, migration, normalization, and emulation. This usage is in accordance with many other sources on digital preservation, for example, the PADI (Preserving Access to Digital Information) project at http://www.nla.gov.au/padi/topics/18.html.

3 See, for example, the Cedars Project at http://www.leeds.ac.uk/cedars/, the CASPAR projects (cultural, Artistic and Scientific Knowledge for Preservation, Access and Retrieval) at http://www.casparpreserves.eu/caspar-project, The Planets (Digital Preservation Research and Technology) project at http://www.planets-project.eu/, the OCLC (Online Computer Library Center) research activities http://www.oclc.org/research/default.htm, the DPC (Digital Preservation Coalition) projects http://www.dpconline.org/graphics/join/projects.html, the Library of Congress - National Science Foundation Digital Preservation Projects of NDIIPP (National Digital Information Infrastructure and Preservation Program) at http://www.digitalpreservation.gov/partners/researchproj.html, and the NARA (National Archives and Records Administration) projects at http://www.sdsc.edu/NARA/Publications.html.

4 For a good source, see the PADI (Preserving Access to Digital Information) collection on preservation policies at http://www.nla.gov.au/padi/topics/172.html.

5 With a research fund being granted recently, the third phase of the InterPARES project, IP3, has started in September 2007 and will last for another five years.

6 InterPARES 2 project website is at www.interpares.org.

7 Luciana Duranti, ed.; San Miniato, Archilab, 2005.

8 For details about the project's intellectual organization, please see http://www.interpares.org/ip2/ip2_intellectual_organization.cfm and http://www.interpares.org/display_file.cfm?doc=ip2_overview_of_intellectual_framework(20030311).pdf.

9 The use of electronic records in the title of the project is inferred, in which the phrase "records in electronic systems" is used. It is explicit, however, in the title of the book presenting IP1 findings, which reads, The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project, see Note 7.

10 See the definitions of "electronic records" and "digital records" developed by the Terminology Task Force at http://www.interpares.org/ip2/ip2_terminology_db.cfm.

11 Defined as the physical or juridical person who makes, receives, or accumulates records by reason of its mandate/mission, functions or activities.

12 Defined as the entity responsible for managing the permanent preservation of records.

13 See variations of fixed documentary form in Luciana Duranti and Kenneth Thibodeau, "The Concept of Record in Interactive, Experiential and Dynamic Environments: The View of InterPARES", Archival Science (2006) 6:13 - 68.

14 Addressed in C6.

15 Addressed in the first principle for records preserver, i.e., P1.

16 Using multiple copies for verification can be said a LOCKSS (Lots of Copies Keep Stuff Safe, http://www.lockss.org/lockss/Home) idea. LOCKSS is a digital preservation project led by the University of Stanford and primarily focused on providing local and long-term access to e-journals, a solution to prevent loss of access to back issues in the event of cancelling subscription or publishers' cease of business.

17 Discussions on subscription to digital publications deserve another paper. Various factors contribute to the result whether libraries get to keep a set of copies of the digital publications they subscribe. Deposit libraries, of course, are exceptions.

18 See, for example, the US Library of Congress's preservation project, MINERVA, at http://lcweb2.loc.gov/cocoon/minerva/html/minerva-home.html; National Library of Australia, PANDORA project, at http://pandora.nla.gov.au/apps/PandasDelivery/WebObjects/PandasDelivery.woa; National Library of Sweden, Kulturarw project, at http://www.kb.se/kw3/eng/.

19 It is worth noting that in reality, the division of collecting and preserving responsibilities between archival institutions and libraries is not always clear. Many libraries in North America have manuscripts departments, which collect mostly private records. However, the point of discussion here is that the central missions of the two cultural institutions are obviously distinguishable.

20 NLA (National Library of Australia), Digital Preservation Policy.

21 Guidelines for Individual Creators at http://www.interpares.org/ip2/display_file.cfm?doc=ip2(pub)creator_guidelines_booklet.pdf.

22 Guidelines for Preservers at http://www.interpares.org/ip2/display_file.cfm?doc=ip2(pub)preserver_guidelines_booklet.pdf.

23 Preservation theories here refer to the conceptual framework designed to guide preservation activities that are based on the distinctive natures of records and publications, such as arrangement and descriptions. They do not include, however, the principles and methodologies for taking physical care of media that hold the intellectual content.

24 The members for this consortium include the national libraries of Australia, Canada, Denmark, Finland, France, Iceland, Italy,

Norway, Sweden, the UK and the USA; see its website at http://netpreserve.org/about/index.php. The National Library of Canada, however, now is a part of the Library and Archives of Canada, which is the result of the incorporation of the former two national cultural institutions in 2004.

25 This consortium includes the British Library, the National Archives, the National Library of Wales, the National Library of Scotland, the JISC and the Wellcome Trust; see its website at http://www.webarchive.org.uk/.

Reference:

Duranti, L., ed. (2005). The Long – term Preservation of Authentic Electronic Records: Findings of the InterPARES Project. San Miniato, Archilab, 2005.

Duranti, L., & Thibodeau, K. (2006). The Concept of Record in Interactive, Experiential and Dynamic Environments: The View of InterPARES. *Archival Science*, 6, 13 – 68.

ERPANET (Electronic Resources Preservation and Access Network). (2003). Digital Preservation Policy, Retrieved August 20, 2007, from http://www.nla.gov.au/policy/digpres.html.

InterPARES. (2006). Terminology. Retrieved August 20, 2007, from http://www.interpares.org/ip2/ip2_terminology_db.cfm.

InterPARES 2. Home. Retrieved August 20, 2007, from http://www.interpares.org/ip2/ip2_index.cfm.

InterPARES 2. (2006). A Framework of Principles for the Development of Policies, Strategies and Standards for the Long – term Preservation of Digital Records. Retrieved August 20, 2007, from http://www.interpares.org/ip2/display_file.cfm? doc = ip2(pub)policy_framework_document.pdf.

InterPARES 2. (2002). Intellectual Organization. Retrieved August 20, 2007, from http://www.interpares.org/ip2/ip2_intellectual_organization.cfm and Overview of Intellectual Organization. Retrieved August 20, 2007, from http://www.interpares.org/display_file.cfm? doc = ip2_overview_of_intellectual_framework(20030311).pdf.

NLA (National Library of Australia). (2002). Digital Preservation Policy. Retrieved August 20, 2007, from http://www.nla.gov.au/policy/digpres.html.