InterPARES

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

Abstract

This entry presents the International research on Permanent Authentic Records in Electronic Systems (InterPARES) project, an international multidisciplinary research endeavor that, since 1999, has sought solutions to the issue of long-term preservation of the authenticity of digital records. It first outlines the goal of the project, its membership, methodology, principles, activities, and products, and then it discusses its conceptual findings—specifically, the concepts of record and of authenticity—and its methodological findings regarding archival appraisal, preservation, and description. This entry concludes summing up the perspectives of the two phases of the project and providing a glimpse of its future direction.

INTRODUCTION

InterPARES stands for International research on Permanent Authentic Records in Electronic Systems. It is a multinational collaborative research project pursuing the goal of developing the theory and methods necessary to ensure that digital records produced in databases and office systems, as well as in dynamic, experiential, and interactive systems, in the course of artistic, scientific, and e-government activities can be created in accurate and reliable form and maintained and preserved in authentic form for the use of those who created them and of society at large, regardless of technology obsolescence and media fragility.

With the financial support of the Social Sciences and Humanities Research Council of Canada, the University of British Columbia (UBC), the United States National Historical Publications and Records Commission and National Sciences Foundation, UNESCO's Memory of the World Program, and several other agencies and organizations around the world, the InterPARES project began in 1999 under the direction of Luciana Duranti, establishing its headquarter in Vancouver, Canada, at the UBC School of Library, Archival and Information Studies. Over the years, it has involved investigators from numerous countries in five continents: Canada, United States, Australia, Italy, the Netherlands, Belgium, the United Kingdom, Ireland, Sweden, Norway, France, Spain, Turkey, China and Hong Kong, Japan, Singapore, Korea, Mexico, Brazil, Cuba, Argentina, Peru, and South Africa.

InterPARES Methodology

The research methodology has been based on the principles of "interdisciplinarity, transferability, open inquiry, and multimethod design." The project is interdisciplinary in the measure in which its goal and objectives can only be achieved through the contribution of several disciplines, among which are archival science, diplomatics and records management; computer science and engineering; jurisprudence and law; music theory, composition, and performance; film theory, production, and description; dance and theatre theory; geography; chemistry; archaeology; and several other hard and social sciences. In fact, in order to analyze the nature, characteristics, behavior, relationships, and process of creation of the records produced in the course of artistic, scientific, and e-government activities, it is not only necessary to gather a deep understanding of those activities, their purpose, their phases and the component actions, their by-products and their structure, and their context, but also their technological environment and their use. In addition, to analyze the results of the case studies, methodologies developed in the context of a variety of disciplines are needed. Among these are text analysis, diplomatic analysis, statistical analysis, etc.

The ultimate goal of InterPARES is archival in nature, in that the project is concerned with the development of trusted record making and keeping systems, and of preservation systems that ensure the authenticity of the records under examination over the long term. This implies that the work carried out throughout the project in the various disciplinary areas must be constantly translated in archival terms and linked to archival concepts, which are the foundation upon which the systems intended to protect the records are designed. However, upon completion of the research, the archival systems need to be made accessible and comprehensible to records creators, organizations and institutions, and disciplinary researchers. In other words, the research outcomes must be translated back into the language and concepts of each discipline that needs to make use of them. In light of the above, it was essential at the outset of the research to

examine the key archival concepts that are at the core of the InterPARES research, so that each discipline could identify the corresponding entities within its own body of knowledge.

The InterPARES research questions have epistemological roots in the humanities, specifically in diplomatics and archival science. However, the investigators involved in each research activity identify the perspective(s), research design, and methods that they believe to be most appropriate to their specific inquiry. The reason for this openness is that InterPARES is conceived to work as a "layered knowledge" environment, in the sense that some research work will build upon knowledge developed in a variety of other research projects on similar and different issues, and some will explore new issues, study entities never examined before and develop entirely new knowledge. Each case study, as well as each of the other research activities, is carried out using the methodology and the tools that the dedicated investigating team considers the most appropriate for it. The methods used are surveys, case studies, modeling, prototyping, diplomatic and archival analysis, and text analysis.

InterPARES Principles

A few principles have guided InterPARES research through the years. The first is that technology cannot determine the solution to the reliable and accurate creation of digital records or to their authentic preservation over the long term: "organizational needs" define the problem and "archival principles" must establish the correctness and adequacy of each technical solution. A corollary of this principle is the idea that solutions to the digital records challenges are inherently "dynamic," because of the continuing change of technology and therefore of the issues that it raises, and "specific" to the cultural, disciplinary, administrative, and legal situations in which digital records are generated, because of the contingent reasons why records are produced and kept. The second principle is that preservation is a "continuous process that begins with records creation." Accordingly, the development of solutions to the issue of long-term preservation is dependent upon the development of guidance for the proper creation and maintenance of records that can be preserved. The third principle is that, given the fact that we cannot preserve digital records as physical entities, but only the ability to reproduce them, we cannot verify on the records at hand their authenticity, thus we must develop methods of records creation, maintenance, and preservation that allow us to presume it till proof to the contrary is established. Thus, the trustworthiness of digital records is always an inference made on the basis of the circumstances in which the records are made and kept over time.

InterPARES Activities

In order to achieve the goal of developing a body of theory and methods, InterPARES researchers have focused their efforts on the identification of "what constitutes a record" in each type of system being examined and in each context, and of "what record in each system and context has the force of an original." They have also studied large bodies of literature in a variety of disciplines for the purpose of defining "what a reliable, accurate, and authentic record" is in the arts, science, law, and administration. On this basis, they have developed the requirements for the design of "a trusted record making system," "a trusted recordkeeping system," and "a trusted record preservation system," using primarily modeling and the diplomatic analysis of the results of several case studies. Then, they have developed "methods and procedures for the creation, maintenance, appraisal, selection, and disposition, and long-term preservation of authentic digital records" and several products that can be autonomously used.

InterPARES Products

The products of InterPARES are: a *Framework of Principles Guiding the Development of Policies* for records creating and preserving organizations; *Guidelines for Making and Maintaining Digital Materials* for individuals and small communities of practice; *Guidelines for Digital Preservation* for archival institutions; *Authenticity Requirements* for records systems; a *Template for Analysis* of digital records; a *Metadata and Archival Description Registry and Analysis System* for the registration and analysis of metadata schemas; principles and criteria for adoption of *File Formats, Wrappers, and Encoding*; a *Model of the Chain of Preservation* (COP); and a *Terminology Database,* including a glossary, a dictionary, and ontologies. These products either are on the InterPARES Web site, www.interpares.org; in the book edited by the Project Director, Luciana Duranti, entitled *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*; in articles published in a dedicated space by the journal *Archivaria*, issues 64–67 (2007–2009); in articles published in a variety of venues, some of which are listed among the references and further readings, while others can be found or cited on the project Web site; and in a book edited by Luciana Duranti and Randy Preston,[1] entitled *InterPARES 2: Interactive, Dynamic and Experiential Records*, which is being published by the Associazione Nazionale Archivistica Italiana.

THE CONCEPTUAL FINDINGS OF InterPARES

The InterPARES project does not study digital preservation in general, but focuses its research on the long-term

preservation of the "authenticity" of digital "records." Thus, among its key findings are the concept of record in the digital environment and the concept of authenticity.

The Concept of Record

In the digital environment, an entity captured as a record presents the following characteristics: 1) a stable content; 2) a fixed form; 3) explicit linkages to other records within or outside the digital system, through a classification code or a naming convention; 4) an identifiable context; 5) five persons (i.e., an author—the physical or organizational person issuing the record; a writer—the person responsible for the record content; an originator—the person responsible for the electronic account or space where the record is made and/or saved; an addressee—the person for whom the record is intended; and a creator—the person in whose records accumulation the record exists) involved with its creation; 6) an action, in which the record participates or which the record supports either procedurally or as part of the decision-making process; and 7) formal elements, attributes, and digital components. Three of these characteristics require further explanation, i.e., stable content, fixed form, and the distinction among formal elements, attributes, and digital components.

Stable content and fixed form

Stable content is defined as the fact that the data and the message in the record are unchanged and unchangeable, meaning that data cannot be overwritten, altered, deleted, or added to. Fixed form means that the binary content is stored so that the message it conveys can be rendered with the same presentation it had on the screen when first saved (even if the digital presentation is different, as is the case with a record received in Word and saved in PDF). In complex dynamic and interactive systems, fixed form is a more articulated concept and involves a distinction between the "stored record" and the "manifested record," the latter being the one we see on the computer screen. If the same content can be presented on the screen in several different ways in a limited series of possibilities, we may have either a different view of "the same stored record" having stable content and fixed form (i.e., different documentary presentations, as in the case of statistical information rendered as a pie chart, a bar chart, or a table) or "several manifested records" with stable content and fixed form derived from the same stored record. In addition, the concept of fixed form may be linked to that of "bounded variability," which exists when there is no stored record, but only stored content data, form data, and composition data that are quite separate and only connected by a query when a user searches the system. In these cases, changes to the form of the manifested record are limited and controlled by fixed rules, so that the same query or user interaction always

generates the same result, and we have different views of different subsets of content, due to the intention of the author or to different operating systems or applications [2, pp. 47–52].

Formal elements, attributes, and digital components

A "formal element" is a constituent part of the record's documentary form as shown on its face. Formal elements may be extrinsic, when they determine the appearance of the record and influence its capability of reaching its purposes; or intrinsic, when they convey the action in which the record participates or to which it is linked, and its context. Extrinsic and intrinsic elements of form are listed and described in the *Template for Analysis* mentioned among the InterPARES products (see: http://www.interpares.org/book/interpares_book_j_app01.pdf).

Strictly related to the concept of formal elements is the concept of "records' attributes," which are the identifying characteristics of each given record or of a record element in it. An attribute may manifest itself as one or more elements of form. For example, the name of the author of a record is an attribute, which may be expressed as a letterhead or a signature, both of which are intrinsic elements of documentary form, i.e., record's elements. In addition to attributes that manifest themselves in the form of the record, i.e., on the face of the record, as record's elements, every record has attributes that are implicit in other parts of the record, such as the name of the creator or of the medium, but in digital records they are also expressed, albeit outside the documentary form. Because of this, they are mostly transparent to the user, and manifest themselves as metadata included in either a record profile, a topic map, or other digital entity linked to the record.

In addition to all the above, with electronic records, we also have to differentiate formal elements and attributes from the record's digital components. A "digital component" is a digital object that is part of one or more digital records, including any metadata necessary to order, structure, or manifest content, and that requires a given preservation action. For example, an e-mail that includes a picture and a digital signature will have at least four digital components (the header, the text, the picture, and the digital signature). Reports with attachments in different formats will consist of more than one digital component, whereas a report with its attachments saved in one PDF file will consist of only one digital component. Though digital components are each stored separately, each digital component exists in a specific relationship to the other digital components which make up the record. Each digital component requires one or more specific methods for decoding the bit stream and for presenting it for use over time. The bit stream can be altered, as a result of conversion for example, as long as it continues to be able to fulfill its original role in the reproduction of the

record. All digital components must be able to work together after they are altered, therefore all changes need to be assessed by the creator for the effects they may have on the record.

Record Trustworthiness

In the case of the concepts embedded in the record quality of trustworthiness, the definitions adopted by the Inter-PARES team were not derived from traditional archival theory. The team used the definitions developed by a previous research project, commonly known as the UBC-MAS project [3, pp. 23–30], in the context of which "reliability" is the trustworthiness of a record as a statement of fact, i.e., its ability to stand for the facts it is about, while "authenticity" is the trustworthiness of a record as a record, and refers to the fact that a record is what it purports to be and has not been tampered with or otherwise corrupted. Reliability is the exclusive responsibility of the record creator and is assessed on the basis of the completeness of the record, the authority and capacity of its author, and the degree of control exercised on the creation process. "Accuracy," a concern introduced later in InterPARES research by science researchers, can sometimes be subsumed under the concept of reliability, and refers to the exactness and correctness of content, mostly dependent on the competence of the author and the controls on the process by which data are recorded and transmitted through space (i.e., between persons, systems, or applications) and time (i.e., when stored off-line, or when the hardware or software used to process, communicate, or maintain it is upgraded or replaced). Differently from reliability, authenticity and accuracy are the responsibility of both the creator and the preserver as they depend on the controls exercised on the processes of transmission of the record across space or time.

Finally, "authentication" was defined as a declaration of authenticity, resulting either by the insertion or the addition of an element or a statement to a record. To make a distinction between authenticity and authentication was considered important, because governments have been legislating about the use of digital signatures and Au1 other similar devises as means of maintaining authenticity. The team wanted to emphasize the theoretical principle that authenticity is a property of the record that accompanies it for as long as it exists, while authentication is a means of proving that a record is what it purports to be at a given moment in time.

The researchers did not elaborate further the concept of reliability, but tried to elucidate the concept implied in the definition of authenticity by dividing it into two components: identity and integrity. "Identity" refers to the attributes of a record that distinguish it from other records. These attributes include the names of the persons concurring in its formation (i.e., author, addressee, writer, originator, and creator); its date(s) of compilation,

transmission and filing; the naming of the matter or action in which it participates; the expression of its relationships with other records; its digital and documentary presentations; and an indication of any attachment(s). These attributes may be explicitly expressed in a formal element of the record or in metadata related to the record, or may be implicit in its various contexts.

"Integrity" is the wholeness and soundness of a record. A record has integrity if it is intact and uncorrupted, i.e., if the message that it is meant to communicate in order to achieve its purpose is unaltered. This means that a record's physical integrity, such as the proper number of bit strings, may be compromised, provided that the articulation of the content and its required elements of form remain the same. Integrity may be demonstrated by evidence found on the face of the record, in metadata attesting to the responsibility for the record over time and to its technological changes, or in one or more of the record's contexts.

While in traditional archival theory, following jurisprudence, records that are relied upon by their creator in the usual and ordinary course of business are presumed authentic, with records in digital systems, the presumption of authenticity must be supported by evidence that a record has not been modified or corrupted in essential respects during transmission and maintenance. To assess the authenticity of a record, the preserver must be able to establish its identity and demonstrate its integrity by observing the existence of certain conditions. These conditions are authenticity requirements and are used by the prospective preserver to guide the assessment of the authenticity of the records during the process of appraisal. The Authenticity Requirements developed by Inter-PARES can be found on the InterPARES Web site (http://www.interpares.org/book/interpares_book_k_app02.pdf) and in the first InterPARES book [4, pp. 204–219].

THE METHODOLOGICAL FINDINGS OF InterPARES

It appeared quite clear to the InterPARES researchers that the team's understanding of the concepts of record and authenticity had a strong impact on the process of appraisal, which is inevitably influenced by the fact the objects to be selected for continuing preservation are not physical units but linked digital components whose continuing meaning is provided by attributes, and whose trustworthiness relies on metadata and on unbroken control both on their transmission through time and space and on the processes of their creation, maintenance, and preservation.

Appraisal

Contrary to the archival tradition, with digital records, "authenticity" has to become one of the values assessed

by the appraiser, because, in the absence of an original, the future users of digital records will have only the word of the record preserver as the basis for trusting the records they will use as sources. Thus, the role of the appraiser becomes that of a neutral third party who acts as the inspector first and the warrantor later of the authenticity of the records that will be preserved.

In addition, the researchers found that several activities extraneous to traditional appraisal methodology have to be introduced in the appraisal process. Although it has been accepted for decades that archivists and records creators have to participate jointly in records scheduling for retention and disposition and that, with digital records, such an endeavor must occur as soon as possible in the life of the record, insufficient emphasis has been put by the archival community on the necessity of "monitoring" the records identified for permanent preservation on a regular basis, in order to ensure that the inevitable ongoing changes of the technological environment of both the records and the creating office, and consequently of the business and documentary procedures of the creating office do not alter the records, their interrelationships, and their relationships with the business processes to the point that a new appraisal is warranted. Monitoring is a key activity also with respect to a new concern for the appraiser: "feasibility of preservation."

Feasibility of preservation was never a preoccupation with records on traditional media, but preservation of digital records is a very complex activity, requiring technological competence and resources that sometimes are not accessible to archival programs and institutions. Thus, the appraiser must assess the records in light of the present and future capability of the archival program or institution to preserve them. As a consequence, the appraiser may advice the creator to make certain technological choices that make the records preservable or may postpone their acquisition till such a time when preservation is feasible.

In addition, in digital systems, a record may be created for one purpose, and then subsequently used for different purposes and by different persons. Any appraisal decision should consider all uses of the record, simultaneous and sequential, by different users and by the same users, and be aware of the business processes behind them. This is necessary in order to make an informed decision about what to preserve, as well as to be able to dispose effectively of all possible copies of the records that have not been selected for preservation.

The use of records or information within records by different business processes may be desirable from the creator's standpoint in terms of providing a degree of interoperability among the creator's information and record systems. In such situations, the preserver should advise the creator that metadata attached to records used by many business processes must identify each relevant business process. This is critical for the creator because it ensures the authenticity of the records by establishing their identity and integrity in each context. It is also critical for the preserver who must understand all contexts in which the records were used in order to effectively undertake appraisal and also to meet the requirements for maintaining authenticity for any records acquired into the preservation system.

Preservation

In light of the principle that it is not possible to preserve an electronic record, but only the ability to reproduce it, InterPARES has established that preservation of electronic records involves creating authentic copies of the records of the creator. The authenticity of these copies is guaranteed by: 1) a controlled process of migration of the records acquired from the creator to the archives technological environment; 2) the accurate documentation of any change that the records undergo during such process and every time that the archives technological environment is upgraded; 3) the implementation and monitoring of privileges concerning the access, use, and reproduction of the records within the archives; 4) the establishment of procedures to prevent, discover, and correct loss or corruption of records; 5) the establishment of procedures to guarantee the continuing identity and integrity of the records (i.e., their authenticity) against media deterioration and across technological changes; and 6) if authentication of individual records is required, the existence of rules determining responsibility for and means of authentication.

The person responsible for all these activities, the designated records preserver, has to take physical and legal custody of the records of the creators, protect them, and ensure continuous access to them. Be it an outside organization or an in-house unit, the role of the designated preserver should be that of a "trusted custodian" for a creator's records. To be considered as a trusted custodian, the preserver must

- Act as a neutral third party, i.e., demonstrate that it has no stake in the content of the records and no reason to alter records under its custody, and that it will not allow anybody to alter the records either accidentally or on purpose,
- Be equipped with the knowledge and skills necessary to fulfill its responsibilities, which should be acquired through formal education in records and archives administration, and
- Establish a trusted preservation system that is capable of ensuring that accurate and authentic copies of the creator's records are acquired and preserved.

The authentic copies of the creator's records are kept by the trusted custodian in a trusted preservation system, which should include in its design a descriptive and

a retrieval system. This trusted preservation system must also have in place rules and procedures for the ongoing production of authentic copies as the existing system becomes obsolete and the technology is upgraded. It should be noted that, the simple fact of reproducing records in the preserver's preservation system does not make the result an authentic copy. Such designation must be provided by the preserver's authority.

A sustainable preservation strategy requires close collaboration between a records creator and its designated trusted custodian. It is the preserver's responsibility to take the initiative in collaborating with the creator to establish acquisition and preservation procedures and in advising the creator in any records management activities essential to the preserver's acquisition and preservation activities.

Archival Description

The InterPARES researchers have agreed that, no matter how careful the preservation procedures and how trusted the records custodian, ultimately, to all future users, the most important source of the authenticity of the records is "archival description."

It has always been the function, either explicit or implicit of archival description to authenticate the records by perpetuating their administrative and documentary relationships but, with digital records, this function has moved to the forefront. In fact, as original digital records disappear and an interminable chain of nonidentical reproductions follows them, the researchers looking at the last of those reproductions cannot find in it any information regarding provenance, authority, context, or authenticity.

The authentication function of archival description is different from that of a certificate of authenticity, because it is not simply an attestation of the authenticity of individual records, but a collective attestation of the authenticity of the records of an archival "fonds" (defined as the whole of the documents made or received by one creator and accumulated for action or reference) and of all their interrelationships as made explicit by their administrative, custodial, and technological history, the scope and content, and the hierarchical representation of the records aggregates. And, it is different both from the identity and integrity metadata attached to individual records, which are part of the record itself and are reproduced time after time with it, and from the additional metadata attached to records aggregations (e.g., file, series) identifying them and documenting their technological transformation. The unique function of archival description is to provide an historical view of the records and of their becoming while at the same time presenting them as a universality in which each member's individuality is subject to the bond of a common provenance and destination. Never before

archival description has had such a key function in the preservation of records.

CONCLUSION

The InterPARES project began in 1999 to find a solution to the problem of the long-term preservation of the authenticity of digital records. In its first phase it focused on the preservation function, but the findings of its research identified the need for a holistic approach that looked at preservation as an integral part of records creation, maintenance, use, and dissemination through time. For this reason, the second phase of the project began with the analysis of the entities being created in complex systems, while still live and often incomplete, in order to understand their nature, characteristics, behavior, and use. Among the results of this analysis there is a new diplomatic theory for interactive and dynamic records and a new archival theory of preservation, along with methodologies supporting their application.

However, this second phase of the research has made clear that, although the body of concepts, principles, and methods developed through scientific research constitutes the essential foundation and framework of best practices, any solution to digital preservation problems is situation specific, and must be devised by preservers taking into account: 1) the cultural, administrative, legal, and functional context in which they operate; 2) the nature and characteristics of the organizations producing the digital material to be preserved; 3) the typology of the material produced and its documentary and technological features; 4) the limitations imposed by the available financial and human resources; 5) the organizational culture of both the producer of the material and the preserver; and 6) access to educated professionals or educational programs and resources. Furthermore, while the conceptual and methodological findings of InterPARES are equally applicable to larger and smaller organizations and programs, archives with limited resources, which often have the greatest need for assistance, will find the outcomes of the research difficult to apply without specific directions on how to move forward.

Thus, a third phase of InterPARES has begun in 2007. It intends to translate the theory and methods of digital preservation drawn from research to date into concrete action plans for existing bodies of records that are to be kept over the long term by archives—and archival/records units within organization—endowed with limited resources. In the process, detailed knowledge can be developed on: 1) how general theory and methods can be implemented in small- and medium-sized archives and units and become effective practices; 2) what factors determine the type of implementation that is appropriate for each body of records in each context; and 3) what skills professionals will require to conduct such operations.

On this basis it is possible to build teaching modules for in-house training programs, continuing education work-shops, and academic curricula that will provide archives with professionals who are competent not only to preserve over the long term our documentary heritage in digital form, but also to ensure the accountability of organi-zations and institutions through the protection of the ac-curacy and authenticity of the digital information they produce.

REFERENCES

1. Duranti, L.; Preston, R., Eds. *InterPARES 2: Interactive, Dynamic and Experiential Records*; ANAI: Roma, 2008.
2. Duranti, L.; Thibodeau, K. The concept of record in inter-active, experiential and dynamic environments: The view of InterPARES. Arch. Sci. **2006**, *6* (1), 13–68. (online: http://dx.doi.org/10.1007/s10502-006-9021-7).
3. Duranti, L.; Eastwood, T.; MacNeil, H. *Preservation of the Integrity of Electronic Records*; Kluwer Academic Pub-lishers Group: Dordrecht, 2002.
4. Duranti, L., Ed. *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*; Archilab: San Miniato, 2005.

FURTHER READINGS

Duranti, L.; MacNeil, H. The preservation of the integrity of electronic records: An overview of the UBC-MAS research project. Archivaria **1997**, *42*, 46–67.
MacNeil, H. Providing grounds for trust: Developing conceptual requirements for the long-term preservation of authentic electronic records. Archivaria **2000**, *50*, 52–78.
MacNeil, H. Providing grounds for trust II: The findings of the authenticity task force of InterPARES. Archivaria **2002**, *54*, 24–58.
MacNeil, H. Contemporary archival diplomatics as a method of inquiry: Lessons learned from two research projects. Arch. Sci. **2004**, *4* (3–4), 199–232.

Author Query Form

**Book: Encyclopedia of Library and Information Sciences, Third Edition
Article No: 120043455**

Dear Author,

During the preparation of your manuscript for typesetting some questions have arisen. These are listed below. Please check your typeset proof carefully and mark any corrections in the margin of the proof or compile them as a separate list.

Query	Details Required	Author's response
AU1	Please check whether "devises" should be changed to "devices" in the sentence: "To make... authenticity."	