Preserving Authentic Electronic Art Over The Long-Term: The InterPARES 2 Project

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

Presented at the Electronic Media Group
Annual Meeting of the American Institute for Conservation of Historic and Artistic Works
Portland, Oregon
June 14, 2004

Abstract
InterPARES 2 is a project that was initiated in 2002 and is expected to be completed in 2006. It builds upon the findings of InterPARES 1, the purpose of which was to develop the theoretical and methodological knowledge essential to the long-term preservation of authentic records created and/or maintained in digital form. It focused on the preservation of the authenticity of records created and/or maintained in databases and document management system in the course of administrative activities. In addition to conceptual findings, it produced requirements for authenticity, methodologies of appraisal and preservation, and an intellectual framework for the development of policies, strategies and standards for the long-term preservation of the authenticity of electronic records. InterPARES 2 focuses on records produced in experiential, dynamic and interactive digital environments in the course of artistic, scientific and e-government activities. Digital art is therefore a primary focus of the on-going research, which involves specialists from the visual and performing arts, computer scientists, intellectual rights experts and other scholars from related fields.

1. The Problem

Ongoing technological change is causing widespread concern around the world regarding the preservation of the cultural heritage produced or stored using digital technologies. A portion of our society's documentary memory created and preserved digitally has already been compromised, and there are enormous costs associated with recovering electronic works that have become inaccessible. While the extent to which valuable digital material has been lost or has become retrievable only at great expense has yet to be adequately quantified, it is already apparent that the threat is real and widespread. Moreover, even if we could ensure the preservation of electronic entities and overcome media fragility and technological obsolescence, preserved works are of little value unless we can be sure they are authentic, that is, that their identity and their integrity have not been inadvertently or maliciously compromised, and they are what they purport to be, immune from corruption and tampering. For centuries, our presumption of authenticity has been premised on the presence or absence of visible formal elements and on an uninterrupted line of legitimate custody. The use of digital technology has not only reconfigured those formal elements, allowed for the bypassing of production controls, and made of physical custody an elusive concept, but, first and foremost, it has eliminated the original work, that is the first complete instantiation being communicated either across space (to persons other than the author) or time (saved for later access by the author or legitimate successors).

If electronic materials will ever be considered authentic as those on traditional media, the practices by which they are created, maintained, made accessible and used must be

analyzed, and strategies and standards for their authentic preservation must be developed. This is the mission of InterPARES (International research on Permanent Authentic Records in Electronic Systems), a research endeavor that aims to develop the theoretical and methodological knowledge essential to the permanent preservation of authentic materials generated and/or maintained electronically, and, on the basis of this knowledge, to formulate model policies, strategies and standards capable of ensuring that preservation. At the end of its first phase, which ran from 1999 to 2001, InterPARES issued, in addition to methods and activity models of selection and preservation, a series of authenticity requirements for materials that, although digital, were very similar to their analog counterparts, especially in that they had a fixed form.[1] Increasingly, however, organizations and individuals have been generating works of a dynamic, experiential, or interactive nature, which will need different, and perhaps work-type specific, authenticity requirements and selection and preservation strategies.

Dynamic materials depend for their content upon data extracted from databases, which may have variable instantiations. The challenge they present to those who generate and access them is their lack of fixity, but more serious issues are raised by experiential and interactive documents. Clifford Lynch describes experiential digital objects as objects whose essence goes beyond the bits constituting them to incorporate the behavior of the rendering system, or at least the interaction between the object and the rendering system. He also maintains that defining the authenticity of such objects is a much more complex problem than with raw data or traditional works, because it is dependent not on the ability to reproduce a copy of the object's original bit-stream, but on the ability to recreate the environment in which that object was experienced, an activity that involves issues of intellectual property, copyright, etc.[2]

An interactive system is one in which each user's entry causes a response from or an action by the system. To generate preservable works in such systems, we need to ascertain a) how user input affects the creation and form of electronic materials, and b) if and when the interactive system and its inherent functionality need to be preserved for those works to remain meaningful and authentic. Examples of interactive systems range from musical performances based on human-computer interaction to commercial video games.

Whether dynamic, experiential, and interactive digital objects are indeed to be preserved over the long-term depends of course on their relationship to the activity of their creator and on the value that society attributes to them. The entertainment and cultural industries have a long history of creating such objects, and clearly the professionals charged with the preservation of the archives containing them may have to face the concrete challenge of preserving views of dynamic systems, recreating the environment of experiential objects, and maintaining the functionality of interactive records. It is important both to know to what extent the requirements, methods and strategies

1. The requirements developed by InterPARES 1 can be found on the project's website at http://www.interpares.org/book/interpares_book_k_app02.pdf
2. Lynch, Clifford. "Experiential Documents and the Technologies of Remembrance." *I in the Sky: Visions of the Information Future*, edited by Alison Scammell. London: Library Association Publishing, 2000.

developed by the InterPARES 1 project to preserve authentic electronic material with a fixed form apply to these new situations, and to develop new ones where they do not.

These issues are further compounded when individual creators lack the knowledge and tools to generate electronic works that can be preserved over the long term. This point can be illustrated by reflecting on the challenge provided by some new habits that they have acquired. The ease with which digital works can be manipulated has in fact given those who generate them, particularly in the creative and research sectors, a new reason for keeping them: 'repurposing'. Makers and distributors of digital music and art works, as well as designers and architects, for example, often obscure the meaning and cultural value of their products by treating their form and content merely as digital data to be manipulated to generate new products, decontextualizing them and destroying their original identity. The potentially wide dissemination of repurposed materials threatens the authenticity of works of art, as well as their authors' moral rights.

For these reasons, it is necessary to develop an understanding of the new digital objects, not only in the later phases of their life cycle, but from the moment of their creation. In fact, it is probably necessary to revisit the concept of finished work itself, so that both the identification and the protection of these new types will be possible. We have to consider the possibility of substituting the characteristics of completeness, stability and fixity with the capacity of the system where the work resides to trace and preserve each change the digital object has undergone. And perhaps we may look at this new digital entity as existing in one of two modes, as an entity in becoming, when its process of creation is in course (even if such creation is ongoing), and as a fixed entity at any given time the work is viewed. There is no doubt that knowledge and strategies must be developed that are beneficial to both the creators and preservers of these complex new materials.

Technological obsolescence, which poses a continual challenge to the accessibility, readability and intelligibility of electronic objects, is of even more concern in the context of artistic activities than in that of administrative activities. Inadequate record-management practices have already precipitated the disappearance of many records pertaining to artworks that depended upon now obsolete software and hardware for their continued existence, including interactive musical materials, art works situated in virtual environments, and other performance works whose essential parameters were insufficiently documented to allow for their recreation. This has generated enormous difficulties for artists concerned with the long-term preservation of the unique and authoritative version of their work, requiring them to devote valuable time and resources to preservation efforts[3] and engendering an urgent demand for effective and tested strategies.

To meet these challenges requires an understanding of the nature of the new electronic objects and their creating processes in the creative and performing arts. For our society to fulfill the legal, social and cultural imperative of preserving these materials as

3. For example, internet artists Kit Galloway and Sherrie Rabinowitz have temporarily shut down their groundbreaking ecafe.com website for two years in an attempt to create an archive of their activities since they began the project in 1984. See <http://www.ecafe.com/about.html>.

authentic over the long term, research must be done into their characteristics and development, the requirements for their reliability, accuracy, and verifiable authenticity, and methods and strategies for their selection and preservation. To this end, the international team of researchers formed for InterPARES 1, together with additional researchers with discipline-specific knowledge, decided to initiate a second phase of its research, called InterPARES 2.

2. InterPARES 2: Intellectual Framework

InterPARES 2 began in 2002 and its completion is scheduled for the end of 2006. It goal, objectives, structure and methodological principles have been articulated in an intellectual framework on which all co-investigators agreed.

2. 1 *Research goal*

The goal of InterPARES 2 is to ensure that the portion of society's recorded memory that is digitally produced in dynamic, experiential and interactive systems can be created in accurate and reliable form, and maintained and preserved in authentic form, both in the short and the long term, for the use of those who created it and of society at large, regardless of digital technology obsolescence and media fragility.

2.2 *Research objectives*

– To develop an understanding of dynamic, experiential and interactive systems and of the materials produced and maintained in them, of their process of creation, and of their present and potential use;
– to formulate methods for ensuring that these digital objects are generated and maintained by the creator in such a way that they can be trusted as to their content (that is, are accurate and reliable) and as works (that is, are authentic);
– to formulate methods for selecting among them those that have to be kept after they are no longer needed by the creator because of their social or cultural value;
– to develop methods and strategies for keeping the materials selected for continuing preservation in authentic form over the long term;
– to develop processes for analyzing and criteria for evaluating advanced technologies for the implementation of the methods listed above in ways that respect cultural diversity and pluralism; and
– to identify and/or develop specifications for policy, metadata, and automated tools necessary for the creation of an electronic infrastructure capable of supporting the creation of accurate and reliable, and the preservation of authentic digital objects.

2.3 *Guiding methodological principles*

2.3.1. Interdisciplinarity

The project is interdisciplinary in the measure in which its goal and objectives can only be achieved through the contribution of several disciplines and of all categories of stakeholders: individual creators, the information technology sector, the archival and conservation professions, etc. are involved in the formulation and selection of case studies, gathering of empirical evidence, and analysis. Such a mode of research ensures that the project's results will find ready acceptance within the targeted communities.

The scholars conducting the research come from the following fields: Archival Science, Chemistry, Computer Engineering, Computer Science, Dance, Diplomatics, Film, Geography, History, Information studies, Law, Library Science, Linguistics, Media Studies, Music, Performance Art, Photography and Theatre. The countries actively involved in the project are: Canada, United States, Australia, Belgium, China, France, Ireland, Italy, Japan, Netherlands, Portugal, Singapore, Spain and the United Kingdom. The Advisory Board also includes an archivist from South Africa.

2.3.2 Transferability

The ultimate goal of the project is archival in nature, in that it is concerned with the development of a trusted system for making and keeping digital entities[4] and of a preservation system that ensures the authenticity of the entities 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 digital entities 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 need to make use of them. In light of the above, all researchers are committed to learning the key archival concepts that are identified by the archival scholars in the team as constituting the core of the InterPARES 2 research, so that each discipline can identify the corresponding entities within its own body of knowledge.

2.3.3. Open inquiry

InterPARES 1 had its epistemological roots in the humanities, specifically in diplomatics and archival science. In contrast, InterPARES 2, while planning as one part of its research to test some of the outcomes of InterPARES 1 in a range of applied settings, espouses no epistemological perspective or intellectual definitions *a priori*. Instead, researchers in each working group identify the perspective(s), research design, and methods that they believe to be most appropriate to their inquiry. The reason for this openness is that InterPARES 2 is conceived to work as a "layered knowledge" environment, in the sense that some of the research work will build upon knowledge developed in the course of InterPARES 1, some will take knowledge of similar issues developed in other areas of endeavour and bring it to bear on creation and preservation of digital materials, some will reconcile knowledge about records and their attributes, elements, characteristics, behaviour and qualities existing in various disciplines and develop it for archival purposes, and some will explore new issues and study entities never examined before and develop entirely new knowledge.

2.3.4. Multi-method design

As stated, each research activity is carried out using the methodology and the tools that the dedicated investigating team considers the most appropriate for it. Examples of the

4. A trusted system comprises the whole of the rules that control the creation, maintenance, and use of the materials of the creator and that provide a circumstantial probability of the accuracy, reliability and authenticity of the digital objects within the system.

methods used are surveys, case studies, modeling, prototyping, diplomatic and archival analysis, and text analysis.

The research is guided by detailed research questions that specifically address the records creation process in each of the examined areas of endeavour, and the characteristics, structure and interrelationships of the resulting materials; the issues related to the development of a chain of preservation for those materials that begins with creation and includes appraisal, description, and reproduction as authenticating procedures; the meaning of the concepts of accuracy, reliability and authenticity in the various artistic disciplines; the policies, strategies and standards in each area of activity covered by the research; the descriptive schemas necessary to the identification, use and preservation of the materials produced by each activity throughout their life-cycle; and the models that more appropriately represent the digital object that is investigated and the processes of its creation, maintenance, use, selection and preservation.

3. Research Progress

The need to concentrate the initial part of the research on gathering an understanding of the process of creation in dynamic, experiential and interactive digital environments has been especially encouraged and supported by the participant stakeholders. The researchers have carried out case studies and general studies. The case studies were identified according to the specific kind of activity that generated the material, and conducted by individual teams assembled in an interdisciplinary way for the purpose of investigating the entire life cycle of the digital objects that were examined. Each team comprised at least a scholar of the activity under investigation, a technology expert, an archival scientist, and a student research assistant. Depending on the complexity of the case study, additional experts and students might belong in the team. The general studies were developed to address issues relevant to each of the three types of activities producing records, but not specific to any given case. Examples of the case studies undertaken are:

– The work of Arbo Cyber, Théâtre (?), a company whose artistic output involves performing arts, visual arts and media arts. Arbo has created over 20 performances, numerous laboratories and school workshops between 1985 and 2001. In an effort to preserve their work, members of Arbo are now seeking to digitize their artistic works. The case study team is interested in following the processes of digitization and transformation of the creative materials created by Arbo in the course of realizing its original productions.

– The work of Stelarc, a performance artist who frequently collaborates with computer programmers, technicians and scientists. His art is exhibited or performed in diverse environments including galleries, aerial suspensions and the Internet. The case study team is interested in learning where record creation begins and ends with Stelarc's art. In addition, the fragility of the environments in which the works are created and performed raises questions relating to issues of reliability and authenticity.

– Obsessed Again…, a work for bassoon and interactive electronics written in 1992 by Canadian composer, Keith Hamel. The work was designed to use commercial hardware and software but the required equipment is quickly becoming obsolete. The commissioner of the work has expressed a wish to reconstitute the work. The case study objectives include identifying both digital and non-digital documents associated with the work, articulating the requirements for musical authenticity based upon the documents, building a performable, authentic realization of the work, and developing a method for the future storage, retrieval, migration and access of the work.

– The Electronic Cafe Intentional (ECI), a multimedia international network for showcasing creative, multi-cultural, multi-disciplinary, collaborative telecommunications. This case study deals with a wide variety of media types that now pose the problems of aging and obsolescent formats. ECI's activities took place from the mid-1970s to the present, and incorporated experiences that were dynamic and interactive, which is the focus of InterPARES 2 research. This study highlights the problems posed by interactive, experiential, and dynamic records many years after they were initially created.

– Three online exhibits developed by the Archives of Ontario as a means to enhance access to its holdings. The research team is investigating the experiential and interactive records that emerge from the creation and posting of these exhibits, and the present and potential uses of these records within the government sector.

Examples of the general studies undertaken are:

– A survey of the record-keeping practices of composers, in order to gather data relating to the types of records that they produce, the composers' assumptions regarding the future access of their records, and the nature and variety of digital music systems they use. This will also shed light on how composers use the records they create, what their expectations are in terms of their future accessibility and how they ensure accessibility of their records over the long-term.

– The MUSTICA Initiative. The study team seeks to develop a typology of interactive digital music compositions, that will support discussion and analysis of preservation needs by identifying the intellectual and physical components of the records of a variety of digital, interactive musical works created by composers at the Institute de Recherche et Coordination Acoustique/Musique (IRCAM) and Groupe de Recherches Musicales (GRM) of Institut National de l'Audiovisuel (INA). This research is partially funded by France's Centre National de la Recherche Scientifique (CNRS).

– Persistent Archives Based on Data Grids. This study focuses on the San Diego Supercomputer Centre's project to develop a prototype for a persistent archives based upon data grid technology for the National Archives and Records

Administration (NARA). This study examines the minimal capabilities needed within grid technology for preservation of digital records, focusing on activities related to the preservation of NARA's selected digital holdings.

– A survey of file formats and encoding languages that are used for non-textual materials. File formats and encoding languages are also analyzed to determine data, information and/or record structure and other properties related to the concepts of accuracy, reliability and authenticity of the digital objects in question. In addition, the study will determine equivalence classes of file formats and encoding languages and identify conversion tools that can be used for migration.

While waiting for the results of the case studies, the research unit responsible for investigating the key concepts of accuracy, reliability and authenticity in each discipline involved with the research has produced annotated bibliographies and reviews of literature; an analysis of those concepts as discussed in the reviewed literature; and a bibliographic database for managing references in bibliographies and literature reviews. The knowledge so acquired will guide the analysis of the findings of the case studies and the development of the related preservation strategies.

The Modeling research unit has developed an activity model of the management of the chain of preservation, the related entity model, and a methodology for walking through the model using case studies data; is testing the model by walkthroughs of selected completed case studies; has begun the development of an activity model of preservation from the creator viewpoint; has designed a protocol for representing in models the findings of case studies within the case study reporting framework; and is developing creation activity and entity models of the completed case studies.

The Policy research unit has conducted literature reviews of existing policies, strategies, guidelines, standards and legislation; has located international instruments that have been developed on freedom of expression, moral rights, etc., and examined how they have been or are being implemented individual countries; has developed a framework for the comparison of policies and has begun comparing those of different sample countries; and has located relevant legislation and ethical codes, analyzing them in relation to the domains research questions.

The Description research unit has conducted a literature review across all focus area in order to identify authorities addressing the accuracy, reliability, authenticity and preservation of digital materials by means of descriptive and other metadata schema and standards; has developed a database registering and describing salient features of relevant extant descriptive and other metadata schema and standards; has developed guidelines and provided training for researchers using the database; is populating the database; has established a metadata schema registry (developed specification, developed an XML-based DTD); has created the metadata schema registry database and populated with a pilot set and is beta testing it with existing standards; and is studying the extent to which all metadata schemas and standards identified in the database and through the completed case studies address the research questions related to description.

Several more research activities are conducted both by individual research units and by the team as a whole, the latter primarily aimed to the development of methodological research instruments and to the reconciliation of findings and drawing of conclusions, but this brief list is sufficient to provide the flavour of the InterPARES 2 work.

4. Conclusion

The InterPARES 2 Project has reached its mid-term. As shown by this summary, it has already produced a large quantity of the material on the basis of which it will develop the project's deliverables, that is, among other things, guidelines for records creators outlining methods for the reliable production and maintenance of digital entities that can be authentically preserved; prototypes of appraisal and preservation systems, and guidelines for records preservers; frameworks for developing policies, strategies and standards, and for the development of descriptive standards for the materials under examination; registries of metadata schemas; and literature and terminology databases. However, as Project Director, I have to recognize that, regardless of its final deliverables, the most desirable outcome of this project has already been achieved: the harmonious collaboration of scholars and professionals from such a large variety of disciplines, backgrounds and cultures towards the long-term preservation of their digital culture is the invaluable product of InterPARES that I watch in awe and cherish every day as the work progresses.[5]

5 For more information on the progress of the research, please consult periodically the www.interpares.org site, clicking on InterPARES 2.