



# **InterPARES 2 Project**

**International Research on Permanent Authentic Records in Electronic Systems**

*International Research on Permanent Authentic  
Records in Electronic Systems (InterPARES) 2:  
Experiential, Interactive and Dynamic Records*

## **PART ONE**

**CASE AND GENERAL STUDIES IN THE  
ARTISTIC, SCIENTIFIC AND  
GOVERNMENTAL SECTORS**

### **Focus Task Force Report**

[including Appendices 3, 4, 5, 6, 7 and 8]

*by*

*Yvette Hackett, Library and Archives Canada*

*William Underwood, Georgia Tech Research Institute*

*Philip Eppard, University of Albany, State University of New York*

- Status:** Final (public)
- Version:** Electronic
- Submission Date:** May 2007
- Publication Date:** 2008
- Project Unit:** Focus Task Force
- URL:** [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_book\\_part\\_1\\_focus\\_task\\_force.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_book_part_1_focus_task_force.pdf)
- How to Cite:** Yvette Hackett, William Underwood and Philip Eppard, "Part One—Case and General Studies in the Artistic, Scientific and Governmental Sectors: Focus Task Force Report," [electronic version] in *International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records*, Luciana Duranti and Randy Preston, eds. (Padova, Italy: Associazione Nazionale Archivistica Italiana, 2008).  
<[http://www.interpares.org/display\\_file.cfm?doc=ip2\\_book\\_part\\_1\\_focus\\_task\\_force.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_book_part_1_focus_task_force.pdf)>

## Table of Contents

<b>Introduction.....</b>	<b>1</b>
Background and mandate.....	1
Research initiatives.....	1
<b>Research Methodology .....</b>	<b>3</b>
Case studies.....	3
Composition of teams .....	3
Research questions.....	3
Selection criteria .....	4
Additional documentation.....	5
Performances and viewings .....	5
General studies.....	6
Literature reviews .....	6
Analysis tools.....	7
Modeling.....	7
Diplomatic analysis.....	7
Terminology.....	8
Validation of case studies .....	8
Case study reporting framework.....	9
Prototyping solutions .....	9
Dissemination activities.....	10
<b>Focus 1 – the Arts.....</b>	<b>12</b>
Research team .....	12
Selection of Focus 1 case studies.....	13
Summary of Focus 1 case studies .....	14
Case study 01: Arbo Cyber, théâtre (?).....	14
Case study 02: Performance Artist Stelarc .....	15
Case study 03: <i>Horizon Zero/ZeroHorizon</i> Online Magazine.....	15
Case study 09: Digital Moving Images - Inputs, Processes and Outputs .....	15
Case study 10: <i>The Danube Exodus</i> .....	16
Case study 13: <i>Obsessed Again</i> .....	16
Case study 15: <i>Waking Dream</i> .....	17
Summary of Focus 1 general studies .....	17
General study 03: Preserving Interactive Digital Music—The MUSTICA Initiative .....	17
General study 04: Survey of Recordkeeping Practices of Composers .....	18
General study 07: Survey of the Recordkeeping Practices of Photographers Using Digital Technology .....	18
Findings specific to the artistic sector.....	18
Records creation and maintenance practices .....	19
Recordkeeping and preservation practices.....	20
The “work” versus the “performance”.....	21

Intellectual property rights.....	21
Authenticity and intent.....	22
The role of new technology .....	23
<b>Focus 2 – the Sciences.....</b>	<b>24</b>
Research team.....	24
Selection of Focus 2 case studies.....	25
Summary of Focus 2 case studies.....	26
Case study 06: CyberCartographic Atlas of Antarctica.....	26
Case study 08: Scientific Data Records from a NASA Spacecraft Mission.....	27
Case study 14: Archaeological Records in a Geographical Information System.....	28
Case study 19: Preservation and Authentication of Electronic Engineering and Manufacturing Records.....	29
Case study 26: MOST Satellite Mission - Preservation of Space Telescope Data.....	30
Summary of Focus 2 general studies .....	31
General study 01: Persistent Archives Based on Data Grids.....	31
General study 06: A Bayesian Belief Network: Supporting the Assessment of the Degree of Belief that a Recordkeeping System Maintains Authentic Digital Records .....	32
General study 09: Survey of the Digital Recordkeeping Practices of GIS Archaeologists Worldwide.....	33
General study 10: Preservation Practices of Scientific Data Portals .....	34
General study 12: Validation of the InterPARES 2 Project Chain of Preservation Model Using Case Study Data.....	34
Findings specific to the scientific sector.....	36
The concept of record .....	36
Authenticity, reliability and accuracy.....	37
Science data archives .....	38
<b>Focus 3 – Government.....</b>	<b>39</b>
Research team.....	39
Selection of Focus 3 case studies.....	40
Summary of Focus 3 case studies.....	41
Case study 05: Archives of Ontario Web Exhibits .....	41
Case study 12: Antarctic Treaty Searchable Database .....	42
Case study 17: New York State Department of Motor Vehicles On-line Services System.....	43
Case study 18: Computerization of Alsace-Moselle’s Land Registry .....	44
Case study 20: Revenue On-Line Service (ROS).....	45
Case study 21: Electronic Filing System (EFS) of the Supreme Court of Singapore.....	46
Case study 24: Preservation of the City of Vancouver GIS Database (VanMap) .....	47
Case study 25: Legacoop Bologna Web Site.....	48
General study conducted by Focus 3 .....	48
Findings specific to the governmental sector.....	48
<b>General Conclusions of the Focus Task Forces.....</b>	<b>50</b>

**References.....51**

**Bibliographies.....52**

**Appendices**

Appendix 3: InterPARES 2 Case Studies and General Studies: Researchers, Focus and  
Final Status.....53

Appendix 4: Genesis of the Case Study Research Questions.....60

Appendix 5: 23 Case Study Questions that the researchers should be able to answer at  
the completion of their investigation March 18, 2003 .....63

Appendix 6: Call for Case Study Proposals InterPARES 2 April 2002 .....65

Appendix 7: Diplomatic Analysis Template.....68

Appendix 8: InterPARES 2 Reporting Framework December 2003 .....72

## Introduction

The research of the Focus Task Forces constitutes the primary source of information upon which the findings and products of the Project's domains and cross-domains are based. Proper contextualization and understanding of the findings and products of the domains and cross-domains therefore requires a sufficient level awareness of the activities of the Focus Task Forces.

## Background and mandate

The organizational structure of the InterPARES 2 Project was developed to address the Project's guiding methodological principles of interdisciplinarity, transferability, open inquiry and multi-method design. To achieve a high level of multicultural and interdisciplinary collaboration, the research was structured into several intersecting areas of inquiry. The research teams responsible for each area were composed of investigators from a variety of disciplines and cultural backgrounds. The research concerning the various disciplines was divided into three focuses of inquiry, each of which was tasked with examining records created in the course of one type of activity as follows:

- the Focus 1 Task Force studied records created as part of artistic activities;
- the Focus 2 Task Force studied records created as part of scientific activities; and
- the Focus 3 Task Force studied records created as part of governmental activities.

## Research initiatives

As anticipated in the original research proposal, the focus task forces directed a good deal of their energy into case studies.

The primary role of Focus Task Forces is to gather and analyze case studies and other data of relevance to each type of activity across multiple domains of inquiry.<sup>1</sup>

Over the course of the Project, twenty-seven case studies were proposed and approved, addressing all three focus areas. Topics included performance art, moving images and music; archaeology, cybercartography and astronomy; and taxation, Supreme Court and land records, among many others. In all, twenty-three of the approved case studies were completed.

A small number of case studies were not completed for a variety of reasons. The number of organizations participating in InterPARES 2 changed over the course of the Project, as did that of individual researchers contributing to the work. One case study was simply re-classified from a case study to a general study when the distinction between these two types of research activities became more clearly defined.<sup>2</sup> One case study was proposed by an organization that left the Project early. Two others were proposed by researchers who were subsequently drafted into research initiatives in the Project's Domains or Cross-domains, eliminating the time they might have had available to conduct a case study. Finally, there were three instances where the subject of the case study withdrew from the process before data collection was finished. Of these three

---

<sup>1</sup> InterPARES 2 Project, "Milestones Report," October 2002, 2 (item 1.5). Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_milestones\(200210\).pdf](http://www.interpares.org/display_file.cfm?doc=ip2_milestones(200210).pdf)

<sup>2</sup> Case study 04, Persistent Archives Based on Data Grids, became general study 01 (under the same title).

instances, one case study was completed, one ended with an interim report and a third could not be completed.

Despite the five-year period covered by the Project,<sup>3</sup> it was not possible to conduct case studies into all aspects of records creation and maintenance in all the specific disciplines of the artistic, scientific and governmental sectors. In some cases, the necessary subject expertise was not represented among the researchers in the Project; in other cases, geographic obstacles would have been too expensive to overcome; and, overall, adequate human resources were frequently lacking, given the wide scope of inquiry being attempted by InterPARES 2.

In addition to case studies, the Focus Task Forces also conducted a number of general studies. Within the context of InterPARES 2, a *general study* was defined as an investigation carried out by any one of the Project's three Focus Task Forces, within its respective scope (i.e., arts, science or government), for the purpose of achieving the Focus Task Force's objectives, but which was *not related to a specific records creator*. This was in contrast to a case study, which, although also carried out by any one of the Project's three Focus Task Forces, was an investigation that focused on the records (or some portion of the records) and records management process *of a specific creator*. Thus, whereas the primary purpose of the case studies was to gather as comprehensive an understanding as possible of the *creator-specific* activities creating the records—including their purpose, their phases and the component actions, their by-products and their structure, their context, their technological environment and their use—the primary purpose of the general studies was to gather as comprehensive an understanding as possible of the above noted records-related issues, but in relation to a wider research context involving a more generalized, yet still bounded, aggregation of records creators within each Focus.

In effect, the general studies were undertaken to help the InterPARES researchers fill particular gaps in the coverage of each Focus area by the case studies. Some general studies allowed researchers to understand the degree to which a particular case study was representative of the work practices of a particular group. Thus, for example, while case study 13 was delving deeply into the work of one composer, Keith Hamel, and the technical details of one specific composition, *Obsessed Again...*, the Focus 1 researchers understood that this single case was not necessarily representative of the full range of adoption and use of digital technologies among composers and so introduced a general study consisting of a Web-based survey designed to examine the recordkeeping practices of a broader range of composers.<sup>4</sup> Other general study surveys looked at recordkeeping issues related to the use of GIS technology by archaeologists,<sup>5</sup> the use of digital technologies by photographers,<sup>6</sup> the functionality of government Web sites<sup>7</sup> and the development of data portals and repositories in the sciences.<sup>8</sup> Still other general studies adopted a more traditional research/report format, such as the study of Persistent Archives based

---

<sup>3</sup> The Project was granted a one-year extension in 2007 expressly for the purpose of disseminating its research findings, bringing to six the total number of years during which the Project was active.

<sup>4</sup> See Michael Longton (2004), "InterPARES 2 Project - General Study 04 Final Report: Survey of Recordkeeping Practices of Composers." Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_gs04\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_gs04_final_report.pdf).

<sup>5</sup> See Randy Preston (2006), "InterPARES 2 Project - General Study 09 Final Report: Digital Recordkeeping Practices of GIS Archaeologists Worldwide: Results of a Web-based Survey." Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_gs09\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_gs09_final_report.pdf).

<sup>6</sup> See Jessica Bushey and Marta Braun (2005), "InterPARES 2 Project - General Study 07 Final Report: Survey of Recordkeeping Practices of Photographers using Digital Technology." Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_gs07\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_gs07_final_report.pdf).

<sup>7</sup> See Mark Wolfe (2003), "InterPARES 2 Project - General Study 08 Final Report: Survey of Government Web Site Interactivity." Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_gs08\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_gs08_final_report.pdf).

<sup>8</sup> See Tracey P. Lauriault and Barbara L. Craig (2007), "InterPARES 2 Project - General Study 10 Final Report: Study of Science Data Archives/Repositories." Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_gs10\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_gs10_final_report.pdf).

on Data Grid Technology<sup>9</sup> and the study of the selection of digital file formats in support of digital preservation activities in a number of institutions.<sup>10</sup>

Of the eleven proposed and accepted general studies, nine produced final reports. The remaining two studies (general study 02, Survey and Analysis of Scientific Encoding Languages for Non-Textual Records and general study 05, An Examination of the Processes to Preserve and Manage Electronic Records: Round Three at the National Archives of Australia and WGBH) were retired prior to completion, although some preliminary reports exist. Appendix 3 contains a complete list of case and general studies and their participants, including completed, retired and irregularly classified studies.

Finally, each Focus Task Force conducted bibliographic work to identify and summarize articles from within its relevant disciplines in support of the Domain 2 Task Force's research into how the concepts of accuracy, reliability and authenticity were understood in the artistic, scientific and governmental environments.

## Research Methodology

### Case studies

#### Composition of teams

Each case study team included, at a minimum, a scholar of the activity under investigation, a technology specialist, an archival expert and a graduate research assistant. As the original Project proposal explained:

The singular nature of the research team, comprising leading scholars in archival science, the social and physical sciences, and the creative and performing arts, as well as representatives from archival institutions and other government bodies from twenty countries and five continents, will ensure that this research, so important for the preservation of our societal memory, will remain focused on 'records' rather than on all digital objects, on the preservation of their trustworthiness both as meaningful content and as records rather than on all issues related to digital preservation, and on the protection of their cultural character rather than on imposing uniform models and applications. This crucial perspective is unique.<sup>11</sup>

A shortage, within the Project, of both archivists and technical specialists meant that a full team was not always available. It must also be noted that a number of the graduate research assistants who joined case study and general study teams made significant contributions to the work, functioning as fully fledged researchers.

#### Research questions

One of the first instruments developed for use by the Focus Task Forces was a standardized set of case study research questions. The decision to use a standardized set of questions was based on the previous experience of the InterPARES 1 researchers who, having used a

---

<sup>9</sup> See Reagan W. Moore (2004). "InterPARES 2 Project - General Study 01 Final Report: Building Preservation Environments with Data Grid Technology." Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_gs01\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_gs01_final_report.pdf).

<sup>10</sup> See Evelyn Peters McLellan (2006), "InterPARES 2 Project - General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation." Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_gs11\\_final\\_report\\_english.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_english.pdf). French language version available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_gs11\\_final\\_report\\_french.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_french.pdf).

<sup>11</sup> Luciana Duranti (2001), "International Research on Permanent Authentic Records in Electronic Systems (InterPARES): Experiential, Interactive and Dynamic Records," SSHRC MCRI InterPARES 2 Project Proposal, 412-2001, 1.1-4–1.1-5. Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_detailed\\_proposal.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_detailed_proposal.pdf).



standardized set of questions for their case studies, found that the method greatly facilitated subsequent comparative analysis.

A standardized case study questionnaire was initially drawn up in quite general, non-discipline-specific language by a small group of researchers during the June 2002, InterPARES International Team meeting in Washington, D.C.<sup>12</sup> The questionnaire initially consisted of eighteen questions and was intended for use by all of the Focus Task Forces. As initially conceived, the questionnaire consisted of questions that were designed to be answered by the case study subjects themselves during interviews carried out by the InterPARES researchers. The set of questions grew from eighteen to twenty-two at the next InterPARES plenary workshop in Los Angeles.<sup>13</sup> The additional questions addressed the research concerns of the Policy, Description and Modeling Cross-domains.

The wider scope of inquiry of InterPARES 2, relative to InterPARES 1, was reflected in the questionnaire in two important ways. First, the questionnaire was designed to be easily adapted for use in artistic, scientific or governmental environments. Second, and even more specifically, it was designed to accommodate the known vocabulary and practices of the many specialized groups the Project would be investigating; be they composers or choreographers, biologists or geomatics experts, bureaucrats or information technology personnel. However, it was later decided that this approach might limit the ability of researchers to compare the results of the case studies, both across the focuses and even within each focus.

Consequently, the questionnaire was reformulated into a more controlled set of questions, using the more rigorous terminology of archival studies, designed to be answered by the case study researchers—with the support of a common, authoritative list of Project-related terms and concepts (i.e., the InterPARES 2 Terminology Database Glossary)—using information gathered by the researchers during the course of the case study through whatever means the researchers felt would be most effective and appropriate within the context of each case study. Thus, the goal of reformulating the questionnaire into a standardized set of questions to be answered by the case study researchers, rather than as a tool to elicit information from the case study subjects, was to, in effect, “normalize” the case study findings to facilitate subsequent comparative analysis of the case study data. From this process emerged a standardized set of twenty-three research questions that would eventually form the core of the case study reports.<sup>14</sup>

### **Selection criteria**

Selection criteria for case studies were kept as flexible as possible. The Call for Case Study Proposals<sup>15</sup> noted the lessons learned from the InterPARES 1 case studies and suggested that the most successful case studies addressed whole recordkeeping systems in the context of documented business procedures. Currently, any “whole” recordkeeping system would be a hybrid mix of analogue and digital formats, with the digital material including both entities created in digital form and digitized copies of analogue material. In judging the case study proposals, members of the International Team attempted to balance the amount of non-digital material included with the representativeness of the digital entities and the balance of digital to analogue.

---

<sup>12</sup> InterPARES 2 Project, “Proceedings of Workshop #2,” National Archives and Records Administration, College Park, Maryland, USA, 20-22 June 2002. For a summary of this workshop, see

[http://www.interpares.org/display\\_file.cfm?doc=ip2\\_wk02\\_summary.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_wk02_summary.pdf). The questionnaire is provided in Appendix 4.

<sup>13</sup> InterPARES 2 Project, “Proceedings of Workshop #3,” Crowne Plaza Beverly Hills, Los Angeles, CA, USA, 17-21 September 2002. For a summary of this workshop, see [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_wk03\\_summary.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_wk03_summary.pdf).

<sup>14</sup> See Appendix 5.

<sup>15</sup> See Appendix 6.

The most important selection criterion was the ability of the proposed case studies to answer the research questions formulated by each of the Project's research Domains and Cross-domains. As a result, the digital entities being proposed generally had to reside in interactive, experiential and/or dynamic systems. Although the archival interest of the InterPARES researchers was primarily in records, there was no attempt to exclude case studies that involved other digital entities that might more often be referred to as, for example, data, publications or works of art.

Each Focus Task Force also wanted to be as comprehensive as possible in studying the various disciplines in its subject area; be it the artistic, scientific or governmental sectors. At the same time, InterPARES 1 had demonstrated the usefulness of investigating similar systems, such as student record systems in universities or patent granting institutions in various jurisdictions. These parallel studies tend to highlight multiple approaches to similar recordkeeping and preservation situations.

Finally, the evaluation of case study proposals also factored in the logistics of the case study, including geographic location, the availability of archival, technical and subject specialists to form a team and the interest of the person or organization being studied in participating.

Despite this long list of considerations, selection of case studies was primarily guided by flexibility. In fact, regarding the process for accepting or rejecting case studies, it was decided during discussions at the second InterPARES 2 workshop "that there would be no specific selection criteria," and, hence, "no grounds for rejection of proposals, only for refinement of proposals."<sup>16</sup>

### **Additional documentation**

When available, case study investigators collected system documentation generated by the records creators. These ranged widely across the interests of the Domains and Cross-domains, including work procedures, policies governing creation and use, technical specifications and metadata standards.

### **Performances and viewings**

Usually, the work processes being examined in case studies are accessible to only the small group of researchers directly involved in the work. In InterPARES 2, a number of products that were the object of case studies could be more widely presented, however.

Among those works related to the artistic activities being investigated by Focus 1, CD and DVD products, such as the 3D reconstruction of the *House of Julius Polibius* (case study 09(1), Altair4 di Roma) and the documentation of Arbo Cyber, théâtre (?) (case study 01) were presented during plenary workshops. As well, a number of InterPARES 2 researchers were able to attend a performance of *Waking Dream* (case study 15) when it was mounted in Vancouver, and a "resurrected" version of the 1992 electroacoustic composition *Obsessed Again...* (case study 13)—a work that had fallen victim to technological obsolescence—was actually performed during the February 2006 plenary workshop in Vancouver.<sup>17</sup>

In the sciences, the *Antarctic Treaty Searchable Database* (case study 12) can be accessed via the Internet,<sup>18</sup> as can the *Archives of Ontario Web Exhibits* (case study 05) studied by Focus 3.<sup>19</sup>

<sup>16</sup> InterPARES 2 Project, "Proceedings of Workshop #2," op. cit., 6.

<sup>17</sup> Further details about the *Obsessed Again...* resurrection experiment are provided in the section titled "A strategy for preventing technological obsolescence of an artistic work" in the Domain 2 Task Force Report. Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_book\\_part\\_3\\_domain2\\_task\\_force.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_book_part_3_domain2_task_force.pdf).

<sup>18</sup> Available at <http://aspire.tierit.com/>. It is also noted that, although the first phase of the Cybercartographic Atlas of Antarctica (case study 06) is now accessible via the Internet (see <https://gcrc.carleton.ca/confluence/display/GCRCWEB/CAA+Project+Description>), the Atlas was only in a pre-Internet development phase during the course of the InterPARES 2 case study.

<sup>19</sup> Available at <http://www.archives.gov.on.ca/english/exhibits/>.

## General studies

The InterPARES 2 general studies utilized several different research methodologies. For example, the arts focus and science focus task forces conducted several large-scale, Web-based, questionnaire surveys.<sup>20</sup> Other general study research methodologies included literature reviews, large-scale, Web-based, researcher surveys (i.e., surveys in which the researchers collected data about the subjects or organizations being investigated from readily-available information sources, often without any direct input from the subjects or organizations themselves),<sup>21</sup> collaborative research and interviews,<sup>22</sup> and tool-building and experimentation.<sup>23</sup>

## Composition of teams

Each general study team included, at a minimum, a scholar of the general activity, discipline or community under investigation and a graduate research assistant. As noted earlier, a number of the graduate research assistants who joined the general study teams made significant contributions to the work, functioning as fully fledged researchers.

## Selection criteria

Aside from the absence of a formal Call for General Study Proposals and a requirement that the studies not focus on specific records creators, the selection criteria and selection process for the general studies closely mirrored that noted above for the case studies.

## Literature reviews

There was an early assumption that a review of existing literature about digital practices in the artistic, scientific and governmental sectors would also inform the work of the focus task forces. Researchers from each of the three focus task forces reviewed over 200 literature resources identified by the Project's graduate research assistants. The publications varied from product announcements, to detailed technical explanations of specific technologies, to high-level planning documents meant to precede the acquisition and use of any technology. The researchers soon discovered that, overall, very little was written about digital practices in the artistic, scientific and governmental sectors, whether from an archival perspective, or from a basic preservation perspective, or even from a practical perspective. As a result, this methodology was abandoned and researchers in each focus task force restricted their literature searches and reviews to works that discussed the concepts of authenticity, reliability and accuracy from the perspective of the researchers' individual specializations; be it in the artistic, scientific or governmental sectors. Any useful citations were forwarded to the Domain 2 Task Force, which was responsible for the analysis of these concepts.

---

<sup>20</sup> See Longton, "General Study 04 Final Report," op. cit.; Bushey and Braun, "General Study 07 Final Report," op. cit.; and Preston, "General Study 09 Final Report," op. cit.

<sup>21</sup> See Wolfe, "General Study 08 Final Report," op. cit.; Lauriault and Craig, "General Study 10 Final Report," op. cit.; and McLellan, "General Study 11 Final Report," op. cit.

<sup>22</sup> See Jennifer Douglas (2006), "InterPARES 2 Project - General Study 03 Final Report: Preserving Interactive Digital Music - The MUSTICA Initiative." Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_gs03\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_gs03_final_report.pdf).

<sup>23</sup> See William Underwood and Sheila Isbell (2007), "InterPARES 2 Project - General Study 06 Final Report: A Bayesian Belief Network: Supporting the Assessment of the Degree of Belief that a Recordkeeping System Maintains Authentic Digital Records." Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_gs06\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_gs06_final_report.pdf).

## Analysis tools

### Modeling

The final step in many of the case studies involved the creation of a “function model,”<sup>24</sup> using the IDEF modeling technique and IDEF0 graphical modeling language,<sup>25</sup> to provide a structured, graphical representation of the workflows and business practices documented in the case study final reports. The diagrams in these models illustrated the processes of creation, maintenance, use and occasionally even preservation that had been applied to the digital entities being studied.

In addition to helping the case study researchers better understand the records creation, maintenance and, where present, preservation activities of the case study subjects, this activity also resulted in case study models that could later be compared to the InterPARES 2 Chain of Preservation (COP) and the Business-driven Recordkeeping (BDR) models being developed by the Modeling Cross-domain. Such analysis could reveal deviations in practices by the case study subjects from the “ideal” models developed by the Modeling Cross-domain. These differences could suggest areas where the adoption of new procedures by the creator might ensure the authenticity of digital records and/or improve the likelihood of their long-term preservation. Alternatively, a consistent deviation in work-flows by the case study subjects in relation to either the COP or the BDR model might suggest a disconnect between archivists’ perception of implementable work practices and the reality faced by the records creators.

### Diplomatic analysis

Once each case study was completed and a draft report issued, graduate research assistants in the School of Library, Archival and Information Studies at the University of British Columbia performed a diplomatic analysis of each type of digital entity identified in each case study.<sup>26</sup> Diplomatics is a science that encompasses a set of principles and terminology that have been used to analyze records since the 17th century.<sup>27</sup> A significant amount of work was conducted during the InterPARES 1 Project to adapt the traditional tenets of diplomatics to the digital environment. The characteristics of a record are itemized in the InterPARES 1 Project’s *Template for Analysis*<sup>28</sup> and discussed in the Final Report of the Authenticity Task Force.<sup>29</sup>

---

<sup>24</sup> Defined as “a structured representation of the functions, activities or processes within [a] modeled system or subject area” (United States Secretary of Commerce, *Draft Federal Information Processing Standards Publication 183*, 21 December 1993. Available at <http://www.edef.com/pdf/idef0.pdf>).

<sup>25</sup> The IDEF (Integrated Definition for Function) modeling technique and the IDEF0 (Integrated Definition language 0) was used by both the Preservation Task Force of InterPARES 1 and the Modeling Cross-domain of InterPARES 2. A more detailed description of the modeling methodology and language within the context of InterPARES 2 research is provided in the Modeling Cross-domain Task Force Report. For a brief synopsis of the IDEF0 modeling process and language, see Randy Preston (2007), *InterPARES 2 Project - Integrated Definition Function Modeling (IDEF0): A Primer*.” Available at [http://www.interpares.org/display\\_file.cfm?doc=idef0\\_primer.pdf](http://www.interpares.org/display_file.cfm?doc=idef0_primer.pdf).

<sup>26</sup> The “Diplomatic Analysis” report for each case study is available on the InterPARES Web site (<http://www.interpares.org/>) and is also included on the DVD that accompanies this book.

<sup>27</sup> For a general introduction to diplomatics, see Luciana Duranti, *Diplomatics: New Uses for an Old Science* (Lanham, Maryland and London: The Scarecrow Press in association with the Society of American Archivists and the Association of Canadian Archivists, 1998). For a discussion of diplomatics in the context of the InterPARES 1 and 2 research, see Luciana Duranti and Kenneth Thibodeau (2006), “The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES,” *Archival Science* 6(1): 13–68, especially 15–21 and 52–55 (Note: a reprint of the Duranti/Thibodeau article is provided in Appendix 2. Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_book\\_appendix\\_02.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_book_appendix_02.pdf)).

<sup>28</sup> Authenticity Task Force (2000), “Appendix 1: Template for Analysis,” in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 192–203. Online reprint available at [http://www.interpares.org/book/interpares\\_book\\_j\\_app01.pdf](http://www.interpares.org/book/interpares_book_j_app01.pdf).

For the purposes of InterPARES 2, the diplomatic analysis was largely restricted to the testing of each type of digital object against the five necessary characteristics of a record to determine if each object could be considered a record, or whether an object was more appropriately identified as, for example, data, documents or publications. Non-records generally require a simpler preservation model because they exist autonomously from other documents and their purpose is, typically, limited to dissemination of information. Briefly, to be considered a record, a digital object must:

- possess a fixed form and stable content affixed to a stable medium;
- participate in an action;
- possess an archival bond, which is the relationship that links each record to the previous and subsequent record of the same action;
- involve at least three persons: the author, addressee and writer; in the digital environment, there are two more necessary persons: the creator and the originator; and
- possess an identifiable context (i.e., the framework in which the action in which the record participates takes place), including juridical-administrative, provenancial, procedural, documentary and technological contexts.<sup>30</sup>

The results of the diplomatic analyses carried out on the InterPARES 2 case studies are analyzed in the report of the Domain 1 Task Force, which focuses on the creation and maintenance of records.

### Terminology

To help ensure terminological consistency from one case study report to the next and to help keep the reports clear and accessible in the multi-disciplinary environment of the Project, each case study team regularly consulted the InterPARES 2 Project Glossary and Dictionary; two of three terminological instruments comprising the InterPARES 2 Terminology Database developed by the Terminology Cross-domain. As appropriate, subject-specific uses and narrow definitions of terms used in the case study reports were submitted to the Terminology Cross-domain Task Force for analysis and possible inclusion in the Terminology Database.<sup>31</sup>

### Validation of case studies

The findings of each case study were first tabled for discussion and validation by the members of the relevant focus task force. The primary goal of these discussions was two-fold: (1) to ensure the completeness and accuracy of the findings and (2) to identify, clarify and compare similarities and differences among discipline-specific practices within the broader artistic, scientific and governmental communities of practice. Cross-community comparisons were then carried out by presenting the completed and validated case studies to the plenary sessions held during the Project's bi-annual research workshops, which all Project researchers were expected to attend regardless of whether their research was concentrated in a focus, a domain or a cross-domain task force.

---

<sup>29</sup> Heather MacNeil et al., "Part One - Establishing and Maintaining Trust in Electronic Records: Authenticity Task Force Report," in Duranti, *Long-term Preservation*, op. cit., 19–65. Online reprint available at [http://www.interpares.org/book/interpares\\_book\\_d\\_part1.pdf](http://www.interpares.org/book/interpares_book_d_part1.pdf). Diplomats is discussed in several places in this report, especially at pp. 22–25, 33–37 and 52–56.

<sup>30</sup> For further elaboration on these characteristics, see the *Diplomatic Analysis Template* in Appendix 7. See also MacNeil et al., "Authenticity Task Force Report," op. cit.

<sup>31</sup> For further information about the Terminology Cross-domain and the InterPARES 2 Terminology Database, see the Terminology Cross-domain Task Force Report. Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_book\\_part\\_8\\_terminology\\_task\\_force.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_book_part_8_terminology_task_force.pdf).

An additional benefit of this approach was to provide members of the Domain and Cross-domain Task Forces with early indications of what each of the case study teams was finding. Another key benefit of these large-scale discussions was that they often led to a clearer and more refined understanding of the intent of the core twenty-three research questions. Finally, this validation and discussion process often highlighted potential issues, disciplines or communities of practice requiring additional research and/or particular attention from researchers involved in ongoing case studies as well as those researchers hoping to use the case study findings to support the work of the domains and cross-domains.

### **Case study reporting framework**

Given the great variation among the case studies undertaken in the artistic, scientific and governmental sectors, it was decided that a standardized reporting framework was necessary to “normalize” the findings and thus facilitate cross-community comparisons and use of the findings by the Project’s Domain and Cross-domain Task Forces. Among other things, the use of a standardized reporting framework would assist the InterPARES 2 researchers in navigating the findings and would help highlight the similarities and differences being uncovered in the practices of the various types of records creation environments. It was also felt that a standardized reporting framework would assist anyone who might choose to consult the case studies during future research.

A standardized reporting framework was initially developed by the Focus 1 Task Force and subsequently adopted by the other two Focus Task Forces.

The standardized reporting framework included the following eight sections:

- A. Overview
- B. Statement of methodology
- C. Description of context
- D. Narrative answers to the core research questions
- E. Narrative answers to applicable domain and cross-domain research questions
- F. Bibliography of relevant material, including articles about the methods and works of the subject(s)
- G. Glossary of terms
- H. Preliminary model

It was anticipated that all case studies would, at a minimum, address sections A through D, while contributions to sections E through H would vary, depending on the nature of the case study and its findings. Each of these sections is described in greater detail in Appendix 8.

### **Prototyping solutions**

Two early case studies offered the possibility to pursue the lessons learned in the case study beyond the final report. One such case study, VanMap (case study 24), was conducted by the government focus. The situation of the records creator in this study, which involved the use of a system (specifically, a Graphic Information System or GIS) that was configured and used in such a way that made it difficult, if not impossible, for the system to create records, was already familiar to the InterPARES 2 researchers who had also participated in InterPARES 1. Technical experts at the San Diego Supercomputer Center undertook an analysis of the system and proposed a system re-design that would ensure that the GIS could produce records and maintain them over the long term. The solution involves the preservation of historical data from the VanMap GIS in

an application environment that allows “point-in-time reconstruction” of VanMap views by re-assembling the data components that were in use at the time specified in a query.<sup>32</sup>

The second case study where a solution could be prototyped was *Obsessed Again...* (case study 13). Amid problems of hardware and software obsolescence, the essential issue to address was that the score for the electroacoustic composition that was the focus of this study did not fully express the composer’s intentions with regard to the performance of the work. An experiment in generating a better description of the work in performance, as an element of metadata, and then using those metadata to re-program the work onto a current technical platform, was attempted. The results of this experiment are described in an appendix to the case study’s final report.<sup>33</sup> The section of the Domain 2 Task Force Report titled “A strategy for preventing technological obsolescence of an artistic work” also summarizes the case study, the metadata and the subsequent attempts to re-perform the piece in a manner acceptable to the composer.

### Dissemination activities

The case studies, and some of the general studies, provided material of great interest to a number of specialized groups and associations whose users were involved in the same activities or who used the same digital technologies. These included musicians creating electroacoustic music, social scientists and government employees using Geographic Information Systems, and both producers and archivists working in the moving image and sound industries.

A sample of the target audiences to whom members of the Focus Task Forces gave presentations includes:

- American Historical Association (AHA)
- American Institute for Conservation of Historic & Artistic Works
- American Society for Information Science and Technology (ASIS&T)
- Archives Association of British Columbia
- Archives Association of Ontario
- Associacao Portuguesa de Bibliotecarios, Arquivistas, e Documentalistas (BAD)
- Association of Canadian Archivists
- Association of History and Computing UK
- Association of Italian Archivists (ANAI)
- Association of Library and Information Science Educators (ALISE)
- Association of Moving Image Archivists
- Association of Pacific Rim Universities (APRU)
- Association of Records Managers and Administrators (ARMA)
- Association for Computers and the Humanities (ACH)
- Association for Literary and Linguistic Computing (ALLC)
- Canadian Cartographic Association
- Centre de recherche interdisciplinaire en technologies émergentes (Canada)
- Centro de Estudos Judiciários do Conselho da Justiça Federal (Brazil)
- Committee on Data for Science and Technology (CODATA)

---

<sup>32</sup> For further elaboration concerning the details of this proposed preservation strategy, see Glenn Dingwall et al. (2008), “From Data to Records: Preserving the Geographic Information System of the City of Vancouver,” *Archivaria* 64 (Fall): 181-198.

<sup>33</sup> See J. Scott Amort (2004), “InterPARES 2 Project - Case Study 13 Final Report: *Obsessed Again...*” Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs13\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs13_final_report.pdf).

- Computer Applications and Quantitative Methods in Archaeology
- Daniel Langlois Foundation for Art, Science and Technology
- Digital Arts Network (DAN)
- European Commission
- Federal Institute of Access to Public Information (IFAI) (México)
- Health Canada
- Health Information Management Association of Australia
- Information Processing Society of Japan
- International Academy of Law and Mental Health
- International Association for Music, Libraries, Archives and Documentation Centres (IAML)
- International Association of Sound and Audiovisual Archives (IASA)
- International Congress on Archives (ICA)
- International Federation of Library Associations and Institutions (IFLA)
- International Musicological Society (IMS)
- Joint Committee on Antarctic Data Management (JCADM)
- Knowledge Management Africa
- le Laboratoire des nouvelles technologies de l'image, du son et de la scène at Université Laval
- le projet de Documentation et conservation du patrimoine des arts médiatiques
- Manitoba Library Association
- Manitoba Archival Association
- National Archives of Brazil
- National Archives of Mexico
- National Archives of the Netherlands
- National Archives of Singapore
- National Archives of Vietnam
- National Association of Government Archives and Records Administrators
- Preservation and Access for Electronic College and University Resources (ECURE)
- Records Management Association of Australasia
- Renaissance Society of America
- Scientific Committee on Antarctic Research
- Society for the History of Authorship, Reading and Publishing (SHARP)
- Society of American Archivists
- Special Libraries Association
- State Archives of Bologna (Italy)
- State Archives of Florence (Italy)
- State Archives of Genova (Italy)
- State Archives of Milan (Italy)
- State Archives of Torino (Italy)
- The Royal Society of Edinburgh
- UNESCO Memory of the World Program<sup>34</sup>

---

<sup>34</sup> A complete list of dissemination activities is available at [http://www.interpares.org/ip2/ip2\\_dissemination.cfm?proj=ip2](http://www.interpares.org/ip2/ip2_dissemination.cfm?proj=ip2).



## Focus 1 – the Arts

### Research team

The group of co-investigators conducting research on the records of artistic activities comprised experts in music, dance, photography, e-literature, theatre, film, multimedia, visual art and the design of multimedia software and hardware. Many have taken on multiple roles in their chosen discipline, including working as practicing creative artists, theorists, historians and archivists. This rich mix of knowledge and experience offered a balance between theory and practice, as well as illustrating the potential for tension between the interests of the artists creating the records and the archivists entrusted with their preservation.

The following is a list of researchers and research assistants who contributed to the work of the Focus 1 Task Force throughout the Project.<sup>35</sup>

#### *Chair:*

Yvette Hackett                      Jan 2001 - Dec 2006

#### *Researchers:*

Howard Besser	New York University, USA—Working Group 3.1
Marta Braun	Ryerson University, Canada—Working Group 2.1
Ann Butler	New York University, USA—Working Group 3.1
Martine Cardin	Université Laval, Canada—Working Group 1.1
Henry Daniel	Simon Fraser University, Canada—Working Group 1.1
Sidney Fels	The University of British Columbia, Canada
Yvette Hackett	Library and Archives Canada—Working Group 3.1
Keith Hamel	The University of British Columbia, Canada
Sally Hubbard	Getty Institute, USA—Working Group 3.1
Mary Ide	WGBH, USA—Working Group 3.1
Ian Lancashire	University of Toronto, Canada—Working Group 2.1
Brent Lee	Windsor University, Canada—Working Group 2.1
Michael Longton	University of Victoria, Canada—Working Group 1.1
Randal Luckow	Turner Broadcasting, USA—Working Group 3.1
Michael J. Murphy	Ryerson University, Canada—Working Group 1.1
Isabella Orefice	Associazione Nazionale Archivistica Italiana, Italy—Working Group 3.1
Jesse Read	The University of British Columbia, Canada
Andrew Rodger	Library and Archives Canada—Working Group 1.1
John Roeder	The University of British Columbia, Canada—Working Group 2.1
James Turner	Université de Montréal, Canada—Working Group 3.1

#### *Research Assistants:*

Melissa Adams	The University of British Columbia, Canada
Scott Amort	The University of British Columbia, Canada
Jessica Bushey	The University of British Columbia, Canada

<sup>35</sup> Researcher membership in Focus 1 changed somewhat over the five years of the Project. Among those who were interested in Focus 1 issues but were unable to participate for the full length of the Project are: Paolo Buonora, Archivio di Stato, Italy; Ben Howell-Davis, Davis International Associates, USA; Susan Kennard, Banff New Media Institute, Canada; Geoffrey Rockwell, McMaster University, Canada.

Natalie Catto	The University of British Columbia, Canada
Seth Dalby	The University of British Columbia, Canada
Heather Dean	The University of British Columbia, Canada
Glenn Dingwall	The University of British Columbia, Canada
Jennifer Douglas	The University of British Columbia, Canada
Coby Falconer	The University of British Columbia, Canada
Shanna Fraser	The University of British Columbia, Canada
Stephen Gage	The University of British Columbia, Canada
Nadine Hafner	The University of British Columbia, Canada
Peggy Heger	The University of British Columbia, Canada
Janine Johnstone	The University of British Columbia, Canada
Greg Kozak	The University of British Columbia, Canada
Tracey Krause	The University of British Columbia, Canada
David Litke	The University of British Columbia, Canada
Yvonne Loiselle	The University of British Columbia, Canada
Luke Meagher	The University of British Columbia, Canada
Cara Payne	The University of British Columbia, Canada
Philippe Perron	Université Laval, Canada
Carolyn Petrie	The University of British Columbia, Canada
Claudette Rocan	The University of British Columbia, Canada
Rebecca Russell	The University of British Columbia, Canada
Vincent Schillaci-Ventura	The University of British Columbia, Canada
Wendy Sokolon	The University of British Columbia, Canada
Adele Torrance	The University of British Columbia, Canada
Jill Teasley	The University of British Columbia, Canada
Melanie Wallace	The University of British Columbia, Canada
Keum Hee Yu	The University of British Columbia, Canada
Sherry Xie	The University of British Columbia, Canada

### Selection of Focus 1 case studies

The scope of inquiry for Focus 1 was limited to an examination of records generated in the course of artistic activities. Focus 1 profited considerably from the fact that issues related to music using digital technologies had already been introduced during the first phase of the InterPARES Project.<sup>36</sup> This advantage allowed Focus 1 to move through the case study process ahead of the scientific and government focus groups. As a result, Focus 1 influenced methodologies and reporting structures for the whole Project.

Most of the ten case studies completed by Focus 1 involved artistic creators; both individuals and organizations. Of particular note was the investigation into digital moving image production (case study 09, Digital Moving Images - Inputs, Processes and Outputs), which eventually studied four different animation and live-action production environments, including commercial

<sup>36</sup> See, for example, Brent Lee and John Roeder, "The Challenge of Digital Musical Archives: Research within the InterPARES Project," in *Digital Resources for the Humanities 2001-2002*, J. Anderson, A. Dunning and M. Fraser, eds. (London: Office for Humanities Communication [King's College], 2003), 171–180; and Brent Lee (2001), "The Growing Complexity of Music Preservation," in *Preserving Authentic Electronic Records: Preliminary Research Findings, an International Symposium held February 17, 2001 at the University of British Columbia, Vancouver, Canada*, Luigi Sarno, ed., 32–40. Available at [http://www.interpares.org/display\\_file.cfm?doc=ip1\\_symposium\\_2001.pdf](http://www.interpares.org/display_file.cfm?doc=ip1_symposium_2001.pdf).

and non-profit enterprises in Canada, Italy and the United States. Another fruitful study involved the pairing of a case study and a general study to investigate the use of digital technology by composers (case study 13, *Obsessed Again...* and general study 04, Survey of Recordkeeping Practices of Composers).

Even in Focus 1, which was the focus group with the largest membership, all sectors of the arts were not represented. Furthermore, it was not always possible to find artists willing to have their recordkeeping practices studied. In addition, some of the experts in the artistic disciplines who initially joined InterPARES 2 were unable to sustain their involvement over the five years of the Project, while others decided to concentrate their efforts on the cross-domain research areas of Policy, Description, Terminology and Modeling.

Of the two case studies that could not be completed by Focus 1, the first (case study 16, Model for Description and Preservation of Documents Created Using Unstable and Variable Artistic Techniques) focussed exclusively on metadata issues and was referred to the Description Cross-domain for possible follow-up. The second incomplete case study (case study 22, Electronic Café International: Aging Records from Technology-based Artistic Activities) was stopped when the interviewees had to withdraw from the process for personal reasons. In this case, an attempt was made to complete the final report with the information available; however, its findings could not be validated because no follow-up confirmation with the creators was possible.

### Summary of Focus 1 case studies

Contemporary art practices are increasingly interdisciplinary in nature. As a result, many of the case studies undertaken as part of Focus 1 cover more than one artistic discipline. The major artistic disciplines involved in the Focus 1 case studies can generally be categorized as follows:

- Performance art
- Experimental theatre
- Dance
- Music (composition and performance)
- Visual arts including installed works with digital moving image elements
- Experimental film/video art
- Computer-generated animation
- Documentary film production

The following brief descriptions of the Focus 1 case studies identify the digital entities that were examined, while also highlighting some of the more salient findings from each study. For a more detailed contextual analysis of these case studies in relation to the records creators and the nature of the activities resulting in document creation, see the “Characterization of the Case Studies” section in the Domain 1 Task Force Report. See Appendix 3 for a complete list of participants responsible for each case study in this focus.

#### Case study 01: Arbo Cyber, théâtre (?)<sup>37</sup>

The focus of this case study is a privately-owned theatre company, *Arbo Cyber, théâtre (?)* (Arbo), located in Québec City in Québec, Canada. Arbo’s artistic output involves performing arts, visual arts and media arts. The digital entity studied was Arbo’s Web site (the

---

<sup>37</sup> See Martine Cardin (2004), “InterPARES 2 Project - Case Study 01 Final Report: Arbo Cyber, théâtre (?)” Available in English at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs01\\_final\\_report\\_english.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs01_final_report_english.pdf), and in French at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs01\\_final\\_report\\_french.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs01_final_report_french.pdf).

*Ludosynthèse*),<sup>38</sup> which was developed to preserve the history and memory of the troupe—in part, through the digitization of Arbo’s past artistic works—while also allowing and encouraging continued audience interaction with Arbo’s past works. The case study researchers follow the processes of digitization and transformation of the creative materials originally created in analogue form by Arbo.

This case study offers an opportunity to observe how Arbo’s original creative materials have been appraised by the creators themselves, digitized, and integrated into Arbo’s “Web site of memory.” On a more general level, this case study explores and documents issues related to the digitization of analogue material as a strategy for long-term preservation by showing how selected archives can be modified by a process of electronic marking, or annotation, while still remaining linked to analogue documents.

### **Case study 02: Performance Artist Stelarc**<sup>39</sup>

This case study examines a performance artist, Stelarc, who combines robotics and technology with the human form, typically in collaboration with computer programmers, technicians and scientists. The digital entity studied as part of the case study was Stelarc’s Web site. Created for promotion, advertising and documentation of his performances, Stelarc also uses the Web site as an online archive of his work and performance documentation. One of the key issues examined in this case study is defining where record creation begins and ends within the context of performance art. In addition, issues related to the challenge of assessing the reliability and authenticity of the artist’s records, particularly in light of the “fragility” of the environments in which the works are created and performed, are examined.

### **Case study 03: *Horizon Zero/ZeroHorizon Online Magazine***<sup>40</sup>

This case study examines a media and visual arts institute situated within an institutional context in the city of Banff, Alberta in Canada. The digital entity studied as part of this case study is the bilingual, multimedia, online magazine, *HorizonZero*, and its related database, ZeroHorizon. In particular, the case study focuses on the digital entities that are created while producing an issue of *HorizonZero*—from conception to publication—including: administrative documents (i.e., those that facilitate communication and control of the production process); journalistic/artistic documents; and programming code that creates the presentation and interactive features of the online magazine’s Flash Web site.

### **Case study 09: Digital Moving Images - Inputs, Processes and Outputs**

This study involves four separate case studies, each of which examines a different moving image production environment. Given the complexity of the moving image production process, the analysis in each case study is limited to a sub-set of digital entities generated during the production process. In particular, each case study describes and explains the processes involved in creating digital moving image products, identifies the digital by-products created at each stage of the process (i.e., pre-production, production and post-production), models the structure of each digital entity and shows the relationships between the digital entities and their relationships to the overall production.

<sup>38</sup> See [http://www.lit.ulaval.ca/arbocyber/index\\_content.htm](http://www.lit.ulaval.ca/arbocyber/index_content.htm).

<sup>39</sup> Henry Daniel and Cara Payne (2004), “InterPARES 2 Project - Case Study 02 Final Report: Performance Artist Stelarc.” Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs02\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs02_final_report.pdf).

<sup>40</sup> See Brent Lee (2004), “InterPARES 2 Project - Case Study 03 Final Report: *HorizonZero/Zero Horizon Online Magazine and Media Database*.” Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs03\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs03_final_report.pdf).

The four production environments, and the basic production processes/digital entities examined, include:

1. a small commercial production company based in Europe (Altair4 di Roma) that produces computer-based 3D animation of historical and archaeological sites;<sup>41</sup>
2. a state-funded film and video production and distribution outlet (The National Film Board of Canada) and the work of its animation studio;<sup>42</sup>
3. a commercial animation studio and the procedures governing the production of image but not sound elements;<sup>43</sup> and
4. a large, U.S.-based public television production company (WGBH in Boston) and its transition from using a control system for analogue film footage to a digital asset management system.<sup>44</sup>

### Case study 10: *The Danube Exodus*<sup>45</sup>

This case study examines the processes, products and by-products involved in creating *The Danube Exodus: The Rippling Currents of the River*, an interactive, multimedia gallery installation by Hungarian filmmaker and artist Péter Forgács in collaboration with an art collective (the Getty Research Institute). The digital entity studied as part of the case study is the “work,” which was displayed at the Getty Research Institute in August and September 2002. As noted in the overview to the case study report, the work has a complicated history and provenance, and is (or was), effectively, a manifestation of several related pieces drawing on much of the same original material. Moreover, the work actually exists (or existed) in two forms: a gallery installation and a Web site. The gallery installation is no longer active but the Web site is still live.<sup>46</sup> This case study offers the opportunity to examine issues of reliability and authenticity within a virtual and hybrid context.

### Case study 13: *Obsessed Again...*<sup>47</sup>

This case study examines *Obsessed Again...*, an electroacoustic work for solo bassoon involving the use of interactive electronics commissioned in 1992 by American bassoonist, Jesse Read, and created by Canadian composer, Keith Hamel. The digital entities studied are the instructions (score and software) required to perform the work. Although the commercial hardware and software elements used in the original composition were state-of-the-art, most have since become largely obsolete, making it difficult, if not impossible, to perform the piece today. The case study team set out to resurrect the work through a complex, controlled migration process that involved identifying both digital and non-digital documents associated with the work, articulating the requirements for musical authenticity based upon the documents, building

<sup>41</sup> See Isabella Orefice (2004), “InterPARES 2 Project - Case Study 09(1) Final Report: Digital Moving Images - Altair4 di Roma, A Multimedia Archaeological Project: *The House of Julius Polybius*.” Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs09-1\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs09-1_final_report.pdf).

<sup>42</sup> See Andrew Rodger (2006), “InterPARES 2 Project - Case Study 09(2) Final Report: Digital Moving Images - National Film Board of Canada.” Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs09-2\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs09-2_final_report.pdf).

<sup>43</sup> See James Turner et al. (2004), “InterPARES 2 Project - Case Study 09(3) Final Report: Digital Moving Images - Commercial Film Studio.” Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs09-3\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs09-3_final_report.pdf).

<sup>44</sup> See Mary Ide (2005), “InterPARES 2 Project - Case Study 09(4) Final Report: Digital Moving Images - WGBH Boston.” Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs09-4\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs09-4_final_report.pdf).

<sup>45</sup> See Sally Hubbard (2006), “InterPARES 2 Project - Case Study 10 Final Report: *The Danube Exodus*.” Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs10\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs10_final_report.pdf).

<sup>46</sup> See <http://www.danube-exodus.hu/en/>.

<sup>47</sup> See Amort, “Case Study 13 Final Report,” op. cit.

a performable, authentic realization of the work and developing a method for the future storage, retrieval, migration and access of the work.

To help assess the representativeness of the findings from this case study, Focus 1 researchers also conducted general study 04 (Survey of Recordkeeping Practices of Composers), which is summarized later in this chapter.

### **Case study 15: *Waking Dream***<sup>48</sup>

This case study examines *Waking Dream*, a multimedia, performance-based art piece involving dance/movement, soundtrack, live and pre-recorded video and remote controlled interactions between performers and various digital and analog technologies. The piece is the result of the collaboration between three partners at the Human Communications Technologies (HCT) Laboratory at the University of British Columbia. The digital entities studied as part of the case study are the Web site and the various digital components that participate in the performance of the work.

Similar to the situation noted by the researchers in case study 13, much of the hardware and software originally used to stage this piece are now obsolete. Unlike the situation in case study 13, however, no records of the original *Waking Dream* performance exist in any form,<sup>49</sup> making the task of staging an accurate and authentic reproduction of the piece all the more tenuous. The situation is further compounded by the fact that two of the partners disagree over the elements of the performance that determine its authenticity.

### **Summary of Focus 1 general studies**

Focus 1 did not limit its research to individual case studies. Three general studies were also conducted by members of the group.

### **General study 03: Preserving Interactive Digital Music—The MUSTICA Initiative**<sup>50</sup>

The first of the general studies in the arts Focus, MUSTICA, stands alone as a collaborative research project with external partners. MUSTICA was designed to fall halfway between a case study and a general study, with a short version of the case study questionnaire being administered to a large number of composers, technical assistants and other employees of two well-known French research institutes—the Institute de recherche et coordination acoustique/musique (IRCAM) and the Groupe de recherches musicales (GRM) of the Institut national de l'Audiovisuel (INA). The MUSTICA researchers worked to identify the various digital components generated during artistic creation and performance processes and to confirm which ones are necessary for long-term preservation and access. They also analyzed IRCAM's long-term experience with metadata, its creation and subsequent utility in providing access to digital files. Unfortunately, language issues, geographical distances and an inability to maintain continuity among personnel involved with the MUSTICA study forced a reduction in the extent of analysis that could be undertaken on the material acquired in the interviews.

---

<sup>48</sup> Sydney Fels and Seth Dalby (2004), "InterPARES 2 Project - Case Study 15 Final Report: *Waking Dream*." Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_c15\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_c15_final_report.pdf).

<sup>49</sup> As noted in the case study report, because "much of the performance takes place in infra-red light, and some parts occur in the dark among the audience, no video can capture the entire performance" (Ibid., 7).

<sup>50</sup> See Douglas, "General Study 03 Final Report," op. cit.

**General study 04: Survey of Recordkeeping Practices of Composers<sup>51</sup>**

This general study involved the development of a Web-based survey instrument to document the record-making and recordkeeping practices of composers using digital technology. The Composers' Survey gathered information about the use of digital technology, composers' intentions and strategies for maintaining digital records, and the forms that these records might take. As such, it served, among other things, to help assess the representativeness of the recordkeeping practices associated with Canadian composer, Keith Hamel, whose electroacoustic work, *Obsessed Again...*, was investigated in case study 13.

Five hundred composers were invited to participate in the survey. With a response rate of nearly 33% of those contacted, the results show a community already familiar with digital technology, with a small minority already beginning to create works intended to be performed in a Web-based environment.

**General study 07: Survey of the Recordkeeping Practices of Photographers Using Digital Technology<sup>52</sup>**

This general study adapted the composers' survey tool to document the recordkeeping practices of photographers. In this case, no detailed case study was underway among photographers, although the widespread use of digital cameras by professionals and amateurs suggested a much larger and more mature digital practice in the field. The photography community was invited to participate in the survey through a number of online professional forums. In total, 371 photographers responded to the survey. The pool of respondents reflected all three focus groups examined by InterPARES 2: the artistic, scientific and governmental communities. Findings of the survey indicate that:

...professional photographers have universally embraced the transition from analogue to digital photography. The majority of photographers who responded identified their practice as completely digital, allocating the use of analogue film to the occasional personal project. Even amongst those who identified their practice as a hybrid of digital and analogue, the bulk of their images were originally created in digital form; they made only a small percentage of analogue images and most of these were eventually digitized.<sup>53</sup>

**Findings specific to the artistic sector**

The findings of the Focus 1 case and general studies indicate that there are significant differences in attitude, procedures and concerns between individual creators or small groups of artists and the business and entertainment industry. For the most part, early adopters of specific digital technologies are not part of a mainstream industry, but are instead innovators who use, adapt or are inspired by particular technological tools, be they hardware or software. This is particularly true for individual artists who explore new forms of expression through technology. This more experimental type of work also seems to rely more heavily on specific types of hardware, which often results in more problems with short-term obsolescence than is the case for those whose work is more hardware-neutral or hardware-independent.

---

<sup>51</sup> See Longton, "General Study 04 Final Report," op. cit.

<sup>52</sup> See Bushey and Braun, "General Study 07 Final Report," op. cit.

<sup>53</sup> Ibid., 29.

## Records creation and maintenance practices<sup>54</sup>

### *Individuals and small organizations*

There will always be artists working alone and in small groups who adopt cutting-edge technology as a tool to pursue their artistic interests or, conversely, to explore the artistic possibilities in any new technology. It is likely that attempts by archivists and other preservers to recommend that artists restrict their use of technology only to widely-implemented technologies that are subject to international standards will be perceived by many artists—especially by those who incorporate cutting-edge technology into their work—as placing an undue constraint on the artistic process. As such, it is likely that such recommendations will, to a large degree, be ignored and, hence, unproductive. Instead, the findings suggest that a more effective preservation strategy for accommodating the needs and expectations of those artists who incorporate digital technology into their work will involve securing a stronger commitment from the information technology industry for the development of technology tools with *increased interoperability*. Software applications, in particular, must provide access to widely-supported interchange formats for files. An artist may choose to work within the proprietary confines of a specific application, but should always be given the means to export the resulting works into a less restrictive technological environment.

An additional danger for artists in this group is a continuing dependency on specialized hardware, such as synthesizers<sup>55</sup> or customized projectors,<sup>56</sup> to render their work. The rapid obsolescence of hardware components can leave an artist, or subsequent performers, unable to reproduce the work as it was initially conceived and presented. Although at times difficult to implement, it is better, from a preservation stand-point, to move the definition and control of this aspect of the work from hardware to software, as is currently happening, for example, with software tools used by composers of digital musical works.<sup>57</sup>

### *Corporations and government institutions*

As noted earlier, more formally established arts-related organizations are less likely to adopt experimental and cutting-edge technologies, preferring to wait for the endorsement of the marketplace before implementing new technology tools. Moreover, one of the key findings was that, in situations where technology is in rapid evolution and has failed to stabilize, corporations will often neglect or even abandon their original digital elements and focus instead on maintaining their analogue materials. This reflects the current belief among such corporations that, should the need ever arise, it will be cheaper to re-create the digital components on future digital platforms rather than attempt to perform multiple costly forward migrations through several generations of frequently incompatible technological environments.<sup>58</sup> Although coming from a different perspective than individual artists, this situation once again endorses the need for widespread interoperability among various file formats.

---

<sup>54</sup> This section provides a synopsis of some of the more salient records creation and maintenance practices in the arts based on the findings of the Focus 1 case and general studies. For more detailed discussion of these issues in relation to the Focus 1 case and general studies, see the relevant sections of the Domain 1 Task Force Report. Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_book\\_part\\_2\\_domain1\\_task\\_force.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_book_part_2_domain1_task_force.pdf).

<sup>55</sup> See, for example, case study 13, *Obsessed Again...*

<sup>56</sup> See, for example, case study 15, *Waking Dream*.

<sup>57</sup> The “resurrection” experiment in case study 13, *Obsessed Again...*, provides a good example of this strategy. See Appendix 1 in Amort, “Case Study 13 Final Report,” op. cit., especially the discussion in the section titled “Towards a Strategy” at pp. 28–29. See also the discussion in the section titled “Experience with a Possible Maintenance Strategy” in the Domain 1 Task Force Report, op. cit.

<sup>58</sup> Among the Focus 1 case studies, this view is most clearly expressed by the commercial film studio creator in case study 09(03). See Turner et al., “Case Study 09(03) Final Report,” op. cit., 4 (answer to research question three).



## Recordkeeping and preservation practices<sup>59</sup>

### *Individuals and small organizations*

Perhaps not too surprisingly, the findings indicate that individual artists and small non-profit and for-profit entities rarely maintain recordkeeping systems that adhere to accepted practices, such as the use of standardized descriptive schemas and metadata, persistent unique identifiers for digital files, documentation of actions/transactions regarding digital file use, or the adoption of standardized work procedures. Individual artists and the small arts organizations also do not employ the services of archivists. In a small number of cases, artists have made a conscious decision that their work should not survive beyond the “life” of the technological context in which it was created due, primarily, to concerns that efforts at long-term preservation could compromise certain characteristics that the artists consider essential to the essence of their artworks, such as the ephemerality and variability of works that change each time they are reproduced. More frequently, the absence of formal recordkeeping practices reflects, in part, a lack of money or a lack of knowledge as to how to best pursue proper maintenance of digital art objects. It also reflects the fact that, because many of the artists examined have yet to suffer a catastrophic loss of some or all of the components of any of their digital works, there has been little “incentive” for these artists to begin to consider the long-term preservation implications of their decision to adopt digital technology as part of their creative processes.

This is not to suggest, however, that the artists examined have not encountered problems related to their use of technology. In fact, all of the artists admitted having had trouble finding files. In response to this issue, almost all the case study and general study respondents acknowledged that they have, with varying levels of success, attempted to implement certain procedures to help manage the digital objects that they create, such as establishing a consistent directory structure and adopting some degree of consistency in naming the objects. These procedures are most evident, and most successfully implemented, in situations where some kind of partnership requires that collaborators other than the creator have access to the material. Of particular note is the fact that a collaborative environment is also most likely to result in the creation of multiple copies of each digital object. Although unwitting, this type of activity represents the adoption of the elementary component of a preservation strategy referred to as LOCKSS (Lots of Copies Keep Stuff Safe).<sup>60</sup>

### *Corporations and government institutions*

In contrast to the situation noted above for individual artists and small non-profit and for-profit entities, the findings indicate that large for-profit, non-profit and governmental organizations do utilize the services of records managers, recordkeeping systems and occasionally archivists (as was observed in two of the digital moving image case studies<sup>61</sup>). These larger organizations have more significant financial interests to protect in the digital assets produced and they also have the resources to implement this level of infrastructure support. Equally, the work created in these environments involves the input and contributions of many specialized individuals and departments. In this context, records management systems are often

---

<sup>59</sup> This section provides a synopsis of some of the more salient recordkeeping and record preservation practices in the arts based on the findings of the Focus 1 case and general studies. For more detailed discussion of these issues in relation to the Focus 1 case and general studies, see the relevant sections of the Domain 2 Task Force and Domain 3 Task Force reports. Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_book\\_part\\_3\\_domain2\\_task\\_force.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_book_part_3_domain2_task_force.pdf) and [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_book\\_part\\_4\\_domain3\\_task\\_force.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_book_part_4_domain3_task_force.pdf), respectively.

<sup>60</sup> LOCKSS is an international, non-profit, community initiative that provides open source software tools and support that are designed to help libraries collect, store, preserve and provide access to Web-published materials. For information on LOCKSS, see <http://www.lockss.org/lockss/Home>.

<sup>61</sup> Specifically, case studies 09(3) (Commercial Film Studio) and 09(4) (WGBH Boston).

seen as useful, if not necessary tools, to help facilitate collaboration by introducing workflow efficiencies and eliminating redundant practices. Such systems also provide a level of standardization for future access, use and re-purposing, as well as providing a level of documentation and contextualization for long-term preservation efforts.

The findings also indicate that corporate and governmental organizations tend to adopt digital technology only after it has been reasonably well-established and often for one or both of the following reasons: (1) to introduce technological efficiencies in terms of ready access to materials and (2) for the perceived financial benefits of ready access and asset re-purposing. Within these environments, the introduction and use of digital technology continues to co-exist with traditional records management practices that are already in place, such as printing to paper or accessing copies of previously recorded analogue audio or video recordings.

### **The “work” versus the “performance”**

Regarding an issue that is unique to the performance arts creators in Focus 1, the findings indicate that there is conflicting information from creators about the need to preserve the “means of production” versus the “documentation of a performance.” The first approach allows artists to re-use technological components of a work and, potentially, to continually alter, change and update the work, with or without the preservation of earlier versions. As well, some artists purposely choose to let a work “die,” keeping only the documentation of each new iteration of the work. The second approach, largely favoured by archivists, curators and conservators intent on capturing the artistic output of a creator, documents the “work” through various media.

The question of what constitutes the “work” varies widely across artistic disciplines of music, performance art, experimental theatre and contemporary fine art. It ranges from the score or script to the live performance to the documentation of an installed work. Archival institutions must be clear on the scope of their acquisition mandate since this central question has implications for what will need to be preserved and how. This issue illustrates that, in the arts, *standardized* preservation criteria and procedures may not be effective or appropriate and that preservation decisions may, in many instances, need to be addressed on a case-by-case basis. This situation mirrors the experiences of museum curators and conservators working with conceptual, installation-based and performative works with digital components, who have found a wide divergence among artists in their choices of what, if anything, needs to be preserved to fully represent the artistic intent of a work.

### **Intellectual property rights**

Some individual artists expressed little concern about establishing or protecting their intellectual property rights. This was most consistently expressed by artists working in highly experimental forms with limited commercial acceptance. Such artists are often part of small, although global, communities where their work is well-known and highly recognizable. Plagiarism in such contexts is unlikely, given the limited financial rewards and the likelihood that the deception will be uncovered and reported back to the creator quite quickly. The highly unique and idiosyncratic works of performance artist Stelarc (case study 02) offer the best example in support of this reasoning.<sup>62</sup>

---

<sup>62</sup> This is not to say, however, that issues related to copyright, intellectual property and patenting of technology are entirely ignored by Stelarc. In fact, as noted in the case study 02 final report, such issues “are of particular relevance to the digital entities examined” (Daniel and Payne, “Case Study 02 Final Report,” op. cit., 4). Nevertheless, explicit concern for such rights appears limited mostly, if not solely, to those elements of Stelarc’s works that are created by other members of his project development teams (i.e., engineers, computer scientists, visual artists, cognitive scientists, etc.), such that “copyright of the programs, codes

This “relaxed” approach to intellectual property rights exhibited by certain individual artists is obviously a major area of divergence from the corporate environment. In fact, where large corporate interests are at stake, copyright continues to be protected using traditional methods and mechanisms that have been tested and accepted by the legal system.

For individual artists, a number of current practices are, unwittingly perhaps, offering some measure of protection of their intellectual property rights. A number of individual artists indicated that although their files are usually stored on personal laptops, they also create backup copies on write-once CD-Rs to store in a second geographic location. For most artists, the practice of copying the files to external media is usually adopted to protect against loss of the original files due to theft of the laptop or failure of its hard drive, while storing the backup media in a second geographic location is seen as a means of protecting the files from small or large-scale disasters like fires and floods. However, the existence of “backup” CD-Rs containing multiple, dated copies of complete works, or versions of digital components developed during the artistic creation process, can also serve as evidence of ownership of the work.

On the other hand, although redundant storage practices can improve the chances of the long-term survival of the digital entities and, in some cases, also serve as evidence of their ownership, this practice can also create proof of ownership problems, particularly in relation to collaborative works. In fact, the ease with which artists involved in collaborative activities with other artists can exchange digital files often leads to the same files being stored on a number of different personal computers, thus making it difficult in some cases to prove exactly who originally created what.<sup>63</sup>

### **Authenticity and intent**

Artistic works are very much the creative output of the artists who create them and, as such, are intrinsically connected to the lives of the artists. Implicit in this understanding is the notion that the long-term preservation of authentic digital works is, in many ways, dependent on whether the artists who created the works are still alive when the works are transferred to the custody of a designated preserver. In many cases, digital art works are based on technological innovation and invention accompanied by little, if any, recordkeeping or documentation of the techniques or practices used; information that is often critical for preservation purposes. In such cases, knowledge and understanding of the exact techniques used often disappear with the artist, thus complicating or, in some cases, thwarting the efforts of others to preserve the works following the death of the artist. In general, it is the works of individual artists and small art organizations that tend to be the least well-documented and, therefore, the most vulnerable in terms of long-term preservation.

Many artists share a common assumption about the authenticity of the materials they create. While they are alive, most artists consider themselves to be the sole arbiter of the authenticity of their digital works. When asked “whether the creator thinks that the authenticity of the digital entities is assured, and if so, why?”<sup>64</sup> the majority of responses offer a surprising level of confidence that the authenticity of the digital entities is assured. For example, the performance artist Stelarc (case study 02) believes that the authenticity of the digital entities related to his

---

and likely the design of the technology itself belongs to the institution that designed and built the technology,” while “clearly...the copyright on the interaction of body and technology, as shown in the pictures and videos of the performance, belong to Stelarc” (Ibid.).

<sup>63</sup> For more detailed discussion of the issue of intellectual property rights in relation to the Focus 1 case and general studies, see the analysis of Domain 1 research question 3 in the Domain 1 Task Force Report, op. cit.

<sup>64</sup> Case study research question 11 (see Appendix 5).

performances is assured “primarily because of his own unique position at the centre of the entire [creation] process and the unique nature of the performance events.”<sup>65</sup>

Although this may be true when an individual is essential to the performance of the work, the ease with which most digital materials can be duplicated suggests that the authenticity of artistic works is, in many cases, less assured than most creators seem to believe or assume. This is particularly true in situations where artists, especially those who consider themselves to be the sole arbiters of the authenticity of their works, have not documented their intent in metadata that will remain inextricably linked to the digital components of the works, since it is these metadata that will be needed by others to reproduce the works in the future following the death of the artists.

The creators of *Danube Exodus* (case study 10) and *Waking Dream* (case study 15) expressed reservations about the ongoing preservation of the authenticity of their digital works. The works associated with these two case studies involve collaborations with two or more creators. In response to case study research question no. 11—which asked whether the creators thought that the authenticity of their digital entities was assured and, if so, why—the creator in case study 10 stated that “To the extent that either the works are published, or the stewardship discussed in the above question [question no. 10] continues, yes, but beyond that, for instance beyond their own life spans, no.”<sup>66</sup> In response to the same question, one of the co-authors (Sydney Fels) of the performance piece in case study 15 expressed concern that the digital entities “may be altered by future producers of *Waking Dream* to an extent that they will no longer be recognizable as his and his co-authors’ intellectual property.”<sup>67</sup> On the other hand, some individual artists and small collectives expressed greater concern with the preservation of their underlying artistic intent than with the preservation of the specific way in which they chose to manifest that intent in a particular work at a given time.

Regardless of the particular aspects of artistic works that artists consider to be central to the authenticity of those works, it is clear that, if there is to be any hope of preserving authentic copies of the works over the long term, additional metadata must be captured to more fully document the intent of the works from the living artists. Access to this metadata will greatly assist future long-term preservation efforts to ensure the longevity of a work beyond the life of the artist.<sup>68</sup>

### **The role of new technology**

On a more optimistic note, two case studies provide early indications of how the World Wide Web is encouraging the development of a new type of recordkeeping practice that meshes with the interests of performance artists and in the dissemination of their works.

---

<sup>65</sup> Daniel and Payne, “Case Study 02 Final Report,” op. cit., 11.

<sup>66</sup> Hubbard, “Case Study 10 Final Report,” op. cit., 8.

<sup>67</sup> Fels and Dalby, “Case Study 15 Final Report,” op. cit., 5. A variation of this sentiment was expressed by one of the interviewees in case study 9(2) (Digital Moving Images - National Film Board of Canada), who noted that, although the “final product is written out to film or to Betamax as a means of storage, [which] is seen as the basic means of assuring the continued existence of the animator’s work in the form he approved, [it is] nonetheless...broadcast on television, or distributed in VHS or DVD form and used on uncalibrated television sets [such that] “The degradation from what the film maker saw in their head to what they’re seeing at home on the tv or screen is so profound that ... some film makers just have nervous breakdowns”” (Rodger, “Case Study 09(2) Final Report,” op. cit., 10).

<sup>68</sup> For more detailed discussion of the concepts of authenticity, accuracy and reliability in relation to the arts in general, see the section titled “Conceptual Analysis: Authenticity, Accuracy and Reliability in the Literature of the Arts” in the Domain 2 Task Force Report, op. cit. For further discussion of these concepts in relation to the findings of the Focus 1 case and general studies, see the section titled “Authenticity, Accuracy and Reliability in the Arts Focus Case and General Studies” in the Domain 2 Task Force Report, *ibid*.

Performance art and documentation is the specific focus of the Arbo Cyber, théâtre (?) study (case study 01), which introduces the concept of the *Ludosynthèse*, an interactive Web site<sup>69</sup> that the artists in this study created to commemorate and disseminate their past performances and also to enable users of the Web site to continue to engage, in an interactive way, with the materials of those past performances. In this case, where the performance activity took place primarily from the mid-1980s to 2001, the *Ludosynthèse* contains digital records, the sources of which are both digital and analogue materials.

The performance artist Stelarc (case study 02) offers a second example. In this case, the artist's Web site<sup>70</sup> is essentially an online publication that provides the general public and collaborators with access to documentation about all aspects of the artist's existing body of work, which involves performance art, dance and technology. The Web site also serves as a digital repository<sup>71</sup> that Stelarc relies on for his own ongoing access to materials that will serve as a basis for his future works.<sup>72</sup>

## Focus 2 – the Sciences

### Research team

The group of co-investigators conducting research on the records of scientific activities comprised experts primarily from academia with knowledge in the disciplines of Geography (Geomatics), Computer Science, Physics and Archival Science. The following is a list of researchers and research assistants who contributed to the work of the Focus 2 Task Force at some point over the duration of the InterPARES 2 Project.

#### *Chair:*

Kevin Glick                      May 2003 - December 2006

#### *Researchers:*

Paul Berkman	University of California, Santa Barbara, USA—Working Group 1.2
Su-Shing Chen	University of Florida, USA—Working Group 1.2
Barbara Craig	University of Toronto, Canada—Working Group 1.2
Michèle Cloonan	Simmons College, USA—Working Group 2.2
Fynnette Eaton	U.S. National Archives and Records Administration—Working Group 2.2
Kevin Glick	Yale University, USA—Working Group 1.2
Babak Hamidzadeh	Library of Congress, USA—Working Group 2.3
P.C. Hariharan	John Hopkins University, USA—Working Group 2.2
Ken Hawkins	U.S. National Archives and Records Administration
Holger Hoos	The University of British Columbia, Canada—Working Group 2.2
Yu Lijuan	Renmin University of China—Working Group 2.2

<sup>69</sup> See [http://www.lit.ulaval.ca/arbo cyber/index\\_content.htm](http://www.lit.ulaval.ca/arbo cyber/index_content.htm).

<sup>70</sup> See <http://www.stelarc.va.com.au/>.

<sup>71</sup> Interestingly, Stelarc considers the Internet to be part of his “presentation, performance and *recordkeeping system*,” and views his Web site as “the central unit in that system” (Daniel and Payne, “Case Study 02 Final Report,” op. cit., 7, 14, emphasis added).

<sup>72</sup> For more detailed discussion of the impact of technology in relation to the records management practices noted in the Focus 1 case and general studies, see the analysis of Domain 1 research question 7 in the Domain 1 Task Force Report, op. cit. See also the Domain 3 Task Force Report, op. cit., especially the “Findings” and “The Domain 3 Research Questions” sections.

Richard Lysakowski	Collaborative Electronic Notebook Systems Association, USA—Working Group 2.2
Reagan Moore	San Diego Super Computer Center, USA—Working Group 2.2
Eun Park	McGill University, Canada—Working Group 2.2
Xiaowei Qiu	State Archives Administration of China—Working Group 2.2
Fraser Taylor	Carleton University, Canada—Working Group 1.2
William Underwood	Georgia Tech Research Institute, USA—Working Group 2.2
Ronald Weiss	Arkival Technology Corporation, USA—Working Group 2.2
Mark Wolfe	University at Albany, State University of New York, USA—Working Group 2.2

*Research Assistants:*

Melissa Adams	The University of British Columbia, Canada
Bart Ballaux	The University of British Columbia, Canada
Natalie Catto	The University of British Columbia, Canada
Heather Dean	The University of British Columbia, Canada
Jennifer Douglas	The University of British Columbia, Canada
Coby Falconer	The University of British Columbia, Canada
Stephen Gage	The University of British Columbia, Canada
Erin Hanlon	The University of British Columbia, Canada
Joshua Hauck-Wheaton	University at Albany, State University of New York, USA
Janine Johnstone	The University of British Columbia, Canada
Eleanor Kleiber	The University of British Columbia, Canada
Tracey Krause	The University of British Columbia, Canada
Tracey Lauriault	Carleton University, Canada
Yvonne Loiselle	The University of British Columbia, Canada
Erin O'Meara	The University of British Columbia, Canada
Christina Miller	The University of British Columbia, Canada
Carolyn Petrie	The University of British Columbia, Canada
Randy Preston	The University of British Columbia, Canada
Wendy Sokolon	The University of British Columbia, Canada
Adele Torrance	The University of British Columbia, Canada
Brian Trembath	The University of British Columbia, Canada
Melanie Wallace	The University of British Columbia, Canada
Sherry Xie	The University of British Columbia, Canada
Keum Hee Yu	The University of British Columbia, Canada

**Selection of Focus 2 case studies**

The scope of inquiry for Focus 2 was limited to an examination of records generated in the course of scientific activities. The Focus 2 case studies were selected for their ability to generate the data required to answer the Project's core research questions regarding the creation, management and preservation of e-science data records. They were also selected for their ability to generate the data that could be used to address research issues in the Description, Policy and Modeling Cross-domains. Because of the small membership of the Focus 2 Task Force, as compared to the other two focus groups, it was decided that the Focus 2 researchers would concentrate on conducting case studies in their current areas of research or primary knowledge,

supplemented with a greater emphasis on general studies than was the case for the other two focus groups.

### Summary of Focus 2 case studies

The following brief descriptions of the Focus 2 case studies identify the digital entities that were examined, while also highlighting some of the more salient findings from each study, especially with respect to the concepts of authenticity, reliability and accuracy as these are understood in the sciences. For a more detailed contextual analysis of these case studies in relation to the records creators and the nature of the activities resulting in document creation, see the “Characterization of the Case Studies” section in the Domain 1 Task Force Report. See Appendix 3 for a complete list of participants responsible for each case study in this focus.

#### Case study 06: CyberCartographic Atlas of Antarctica<sup>73</sup>

The focus of this case study is the Cybercartographic Atlas of Antarctica (CAA), an online atlas that portrays, explores and communicates the complexities of the Antarctic continent for education, research and policy purposes. The Atlas is the key deliverable of a four-year (2003-2007) research project that was led by the Geomatics and Cartographic Research Centre (GCRC), an organized research unit in the Department of Geography and Environmental Studies at Carleton University, Ottawa, Canada.

The CAA is a complex system that is open source, interoperable, renders distributed data, includes scientific data visualization, is multisensory and incorporates multimedia data in a very wide array of formats from a number of international sources. In collaboration with experts from a number of disciplines—including Cartography, Geography, Psychology, International Studies, English, Cultural Mediation, Music Studies, Industrial Design and Computer Science—these data are used to develop theme-specific modules for use by the general public to facilitate knowledge sharing in multi-disciplinary science.

It was not the intention of the CAA Project to generate or collect substantive new data but to bring together selected existing datasets in a new multimedia form including experimental work with virtual reality and render these in a dynamic and engaging fashion. Distributed data are rendered “on the fly” in maps, charts, tables and text.

The CAA’s design incorporates recent specifications proposed by the Open Geospatial Consortium (OGC) and the International Organization for Standardization (ISO) that are enabling the development of the “spatial Web.” This spatial Web is making it possible for users to easily find, access and process digital geospatial data over the Internet. It is, of course, also creating new long-term preservation challenges.

The creators in this study rely on various processes and procedures to ensure the quality of the data in the CAA, including the professional practices and authority of the institutions from which the data are derived, implementation of access and security controls, and the use of peer review together with adherence to cartographic professional practices to choose the right level of data accuracy and to select cartographers for the right representation. These processes and procedures are very much reliant on metadata and professional practices.<sup>74</sup> Regarding authenticity, this is assessed by the creators of the CAA through analysis of what in Geography and Geomatics is typically referred to as “data lineage.” “Lineage,” which is one of at least seven

---

<sup>73</sup> See Lauriault and Hackett, “Case Study 06 Final Report,” *op. cit.*

<sup>74</sup> *Ibid.*, 14–15, 25–27, appendices P, T.

elements comprising “spatial data quality” and is a mandatory metadata element for any digital object that forms part of the CAA,<sup>75</sup> captures information about the chain of transmission of a dataset—from the moment the data in the dataset were originally recorded—that brought the dataset to the user. In other words, data lineage speaks to the *history* of a dataset; that is, its lifecycle from data collection to its many stages of compilations, corrections, conversions and transformations, including the generation of new interpreted products.<sup>76</sup>

Although clearly not a study of a recordkeeping or a preservation system, this case study raises the important and timely issue of how copies of digital records and related objects that an organization makes available over the Internet can be preserved authentic. As the findings of this study suggest, a conscious commitment to, and consistent reliance on, open source software, especially open source software that is highly interoperable, supported by a robust repository system that incorporates a classification system and “a multimedia metadata schema that meets both archival requirements and geospatial standards,”<sup>77</sup> is likely one of the most effective strategies for helping ensure the long-term sustainability and preservation of complex digital entities like the CAA.

### **Case study 08: Scientific Data Records from a NASA Spacecraft Mission**<sup>78</sup>

The focus of the case study is the data records of the Mars Global Surveyor (MGS) mission and the Planetary Data System (PDS). This case study was selected because it was also designed to collect data used in two general studies, general study 12 (Validation of the InterPARES 2 Project Chain of Preservation Model Using Case Study Data) and general study 06 (A Bayesian Belief Network: Supporting the Assessment of the Degree of Belief that a Recordkeeping System Maintains Authentic Digital Records).

Launched in 1996, the MGS was the first mission of the National Aeronautics and Space Administration (NASA) Mars Surveyor Program. The MGS spacecraft contains instruments that send observational data back to the Jet Propulsion Laboratory Spacecraft Control Center where they are stored as experimental data records in a Project Database. Spacecraft Operations Planning Computers at the Principal Investigator’s site and at seven discipline nodes—Atmospheres, Geosciences, Imaging, Navigation Ancillary Information Facility, Plasma Physics, Rings and Small Bodies—are used to create reduced data records.

The designers of the PDS adopted the use of self-describing data files—that is, files that describe the file format of the attached data as well as the context in which the data were created—as a preservation strategy. The use of self-describing files eliminates the need for converting scientific data to other formats to maintain their accessibility over time, which is one of the key elements supporting an experimental archival preservation strategy called Collection-based Persistent Object Preservation (POP).<sup>79</sup> Using this strategy, technology obsolescence requires only the migration of an interpreter, a data access library and a viewer, thereby

---

<sup>75</sup> Ibid., 14, 26.

<sup>76</sup> As is discussed in more detail in the Domain 2 Task Force Report, *op. cit.*, data lineage can also be characterized as “data provenance” and, thus, roughly analogous to an amalgamation of the archival concepts of provenance and chain of custody.

<sup>77</sup> Lauriault and Hackett, “Case Study 06 Final Report,” *op. cit.*, 29.

<sup>78</sup> See William Underwood (2005), “InterPARES 2 Project - Case Study 08 Final Report: Mars Global Surveyor Data Records in the Planetary Data System.” Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs08\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs08_final_report.pdf).

<sup>79</sup> For a brief description of POP within the context of this case study, see Underwood, “Case Study 08 final Report,” *op. cit.*, 34. For a more detailed discussion, see Reagan Moore et al. (2000), “Collection-Based Persistent Digital Archives - Part 1,” *D-Lib Magazine* 6(3). Available at <http://www.dlib.org/dlib/march00/moore/03moore-pt1.html>; and Reagan Moore et al. (2000), “Collection-Based Persistent Digital Archives - Part 2,” *D-Lib Magazine* 6(4). Available at <http://www.dlib.org/dlib/april00/moore/04moore-pt2.html>.



circumventing the need to continuously update or create new software for each of the file structures. The fact that the PDS has been operational since 1989, during which time it has not been necessary to update (convert or migrate) any of the self-described data products to other data formats, seems promising.

Concerns regarding the reliability and accuracy of the digital entities in this study are addressed in several ways. First, the creators in this study make use of various data processing plans, manuals, specifications and workbooks to guide processing, preparation and transfer of satellite data slated for long-term preservation in the Planetary Data System (PDS). Second, the data are subjected to a peer review process that is designed to check the digital entities for accuracy, reliability, suitability for archiving and to ensure that the PDS data processing, cataloguing and transmission standards have been followed. Finally, all data destined for the PDS are parsed by a computer program that validates the format and content of the “product label” metadata as well as the integrity of each data file using checksums.<sup>80</sup> Regarding authenticity, it is noted that “Project team members, PDS managers and engineers and other Planetary Scientists do not traditionally use the term authentic to characterize the data products that they create, maintain and use. They are concerned that the data records are complete, reliable, accurate, and that the integrity of the data record is assured.”<sup>81</sup> However, as the case study report further clarifies: “Given the definition of authentic digital record as “a digital record that is what it purports to be and is free from tampering or corruption,” one can conclude that, due to the emphasis on completeness and reliability of the planetary science data records, the peer review, role-based authentication of access to archived data products and data integrity checks, the scientific data records are maintained authentic. However, it is unlikely that the creators or maintainers of the records would customarily use that term to characterize their quality.”

#### **Case study 14: Archaeological Records in a Geographical Information System<sup>82</sup>**

This case study investigates archaeologists’ attitudes towards recordkeeping, especially in terms of what they consider authentic and reliable records and the value of these concepts within archaeology. The immediate goal of this case study was to gain insight into the creation processes of Geographic Information System (GIS)-related data and records created and used in the course of archaeological research, with an emphasis on better understanding how archaeologists themselves view and, subsequently treat, their data and the records resulting from the use of those data. The more general objectives of this study were to help answer questions about the nature of digital archaeological records in general; about how the increased reliance on GIS-supported research is impacting the archaeological community; and, especially, about how both of these issues are playing out in the recordkeeping habits of archaeologists.

Ultimately, this research was conducted with the objective of facilitating the development of an appropriate records management framework in which archaeologists and other scientists can approach the long-term preservation of the complex datasets and outputs that are commonly associated with a GIS. The long-term preservation of such datasets and outputs has become a critical issue with regard to enabling the kind of multidisciplinary research crucial to modern scientific knowledge.

---

<sup>80</sup> See Underwood, “Case Study 08 Final Report,” op. cit., 10–11, 23–24.

<sup>81</sup> Ibid., 24.

<sup>82</sup> See Richard Pearce-Moses, Erin O’Meara and Randy Preston (2004), “InterPARES 2 Project - Case Study 14 Final Report: Archaeological Records in a Geographical Information System: Research in the American Southwest.” Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_c14\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_c14_final_report.pdf).

The case study investigates the recordkeeping activities of those individuals responsible for the creation, maintenance and use of the Coalescent Communities Database and GIS that was created for the Center for Desert Archaeology in Tucson, Arizona, in support of a two-year U.S. National Science Foundation grant-funded research project. The database and related GIS consist of compilations of pre-recorded archaeological site data from multiple sources. These datasets exist in various formats, such as paper, spreadsheets and pre-existing databases.

As is true of virtually all applications of a GIS, the central function of the GIS in this case study is information management. In this particular instance, however, the process of incorporating large legacy datasets—many of which were collected long ago (more than 100 years ago in some cases) and originally recorded in analogue form and subsequently digitized—from diverse sources into a single database raises serious concerns about establishing and maintaining the authenticity, accuracy and reliability of the data and, hence, the records generated by the GIS from the data in the consolidated database.

Regarding the datasets used to populate the GIS database in this study, it is noted that an assessment of the accuracy and quality of each dataset may involve various processes or procedures. These may include preliminary “audits” of datasets both before and after they are ingested into the database; “routine” checks of the data in the database for data redundancies, errors and omissions by a volunteer who is a retired archaeology professor; and “spot” checks by the creator of data entered into the database by the volunteer to ensure that the data are being entered accurately. However, since there are no formal, documented data assessment rules, procedures or guidelines that control these activities,<sup>83</sup> the overall process is best characterized as idiosyncratic and ad hoc. The relative laxness of the data assessment approach used by the creator in this study is based, to a large degree, on his assumption that the reliability and accuracy of the data used to populate the database is determined by the reliability of the data *source*. In other words, the creator believes that if the source of the data can be assumed to be trustworthy, then the data acquired from that source can be assumed to be reliable and accurate. Moreover, this same argument also serves as an implicit means of “establishing” the authenticity of the data in that the creator assumes the data are authentic on the grounds that they are acquired from an official state repository and from fellow professional researchers whom the creator trusts to collect and maintain their research data using standard, professional archaeological procedures.<sup>84</sup>

To help assess the representativeness of the findings from this case study, Focus 2 researchers later conducted general study 09 (Survey of the Digital Recordkeeping Practices of GIS Archaeologists Worldwide), which is summarized later in this chapter.

### **Case study 19: Preservation and Authentication of Electronic Engineering and Manufacturing Records<sup>85</sup>**

This case study reports on an engineering experiment to develop an open-source preservation format for digital computer-aided design (CAD) records of solid models. The domain of the case study is science-based manufacturing of complex, high-assurance, high tolerance machined piece parts for the U.S. government. The business owner has an ongoing need to access and use the preserved records for business purposes over a long period of time (50+ years) with the assurance that the records remain accurate, reliable and authentic.

---

<sup>83</sup> See Pearse-Moses et al., “Case Study 14 Final Report,” op. cit., 7–8.

<sup>84</sup> Ibid., 2, 15, 24–26.

<sup>85</sup> See Kenneth Hawkins (2006), “InterPARES 2 Project - Case Study 19 Final Report: Preservation and Authentication of Electronic Engineering and Manufacturing Records.” Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs19\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs19_final_report.pdf).

The records involved contain information about complex geometric and topological measurements and relationships of various parts of three-dimensional objects. The case study involves analysis of an experiment that involved: (1) abstraction of this information from proprietary CAD formats; (2) expression of this information into enhanced logical forms that support reasoning about part shape and manufacturing actions; (3) rendering of the proprietary CAD formats into an open-source archival format; (4) transmission of the records in open-source format across a trusted network for ingestion into a persistent archives; and (5) retrieval of the archived records for verification of their continued authenticity, reliability and usability.

The intent of the experiment was to preserve not only the geometric specifications of the models but also their semantically encoded metadata, joined to make a “new logical preservation format” for archival purposes. By “logical preservation format,” the experiment partners in this case study meant a format encompassing not only the fixed form and content of information representing the models, but also instructions encoded within their metadata so that reasoning engines of the future can conduct “proofs” against the objects to authenticate them as fit to support the procedural action for which they were designed to be used.

Concerns regarding the authenticity, reliability and accuracy of the digital entities in this study are addressed in several ways. First, accuracy and reliability are in the hands of the designer and his/her adherence to design modeling standards such as the ANSI Y-14.5 tolerance standard, which provides for the assurance of measurements down to the millionth of an inch. Second, the CAD system used to create the initial CAD entities integrates rigorous model checks that can be used alert users whenever they attempt to create bad geometry. Finally, a quality guide is used to help control the digital entity production and manipulation processes. However, because of the extreme level of accuracy achieved by these processes—especially in relation to the “enhanced semantic knowledge about the geometric and topologic characteristics of the solid models”—the resulting design records cannot be accurately translated into the final archival format, at least not within the context of the engineering experiment.<sup>86</sup> The results of the engineering experiment show that, by the exacting standards of the engineering experiment, the final archival entities are unable to preserve all the advanced information contained in the “knowledge-enhanced” entities produced from the original CAD entities and, as such, “cannot be said to be authentic to a satisfactory degree.” In short, the experiment in this case study demonstrates that it is not possible, at present, “to preserve all elements and components of the entities in a persistent archives.”<sup>87</sup>

### **Case study 26: MOST Satellite Mission - Preservation of Space Telescope Data**

The focus of this case study is the preservation of astronomical data from the MOST (Microvariability & Oscillations of STars) Satellite Mission, Canada’s first space telescope. MOST is designed to monitor variations in the brightness of stars with unprecedented precision and time coverage. Launched in late June 2003, MOST is funded by the Canadian Space Agency (CSA) and operated jointly by Dynacon Inc., the University of Toronto Institute for Aerospace Studies (UTIAS) and the University of British Columbia (UBC).

This case study examines how the MOST researchers manage the data that are transmitted from the space telescope. In particular, the case study investigates how technological restraints complicate the creation, maintenance and preservation of the MOST satellite data and how future researchers can be assured that the preserved data are reliable, authentic and accurate. In so

---

<sup>86</sup> Hawkins, “Case Study 19 Final Report,” op. cit., 11, 16.

<sup>87</sup> Ibid., 17.

doing, this study explores the effect of the business process on the various phases of the life-cycle of the data and records. The study also examines how researchers interpret the concepts of accuracy, reliability and authenticity in the context of scientific research.

Perhaps not too surprisingly, similar procedures to those noted earlier for the NASA Spacecraft Mission (case study 08) are used to ensure the authenticity, reliability and accuracy of the digital entities in this study. The creators in this study have implemented a rigorous and systematic, two-stage data integrity process. The first stage involves an automated, computer-executed, checksum analysis. The second stage involves a hands-on “intellectual analysis” performed by the project’s instrument scientist. As well, some data are processed through a model that, among other things, helps the researchers spot and isolate any inaccuracies in the data that might require that the data be reprocessed.<sup>88</sup> Regarding authenticity, the report notes that “there is no reason for the creator to assume that the digital entities are not authentic.” This level of confidence is not specified is likely a consequence of the controlled and secure nature of the sources of the data used in this study,<sup>89</sup> together with the fact that there is only one researcher who is responsible for the creation, maintenance and preservation of the raw data files received from the space telescope and gathered from other sources, and one researcher who is responsible for processing (i.e., “reducing”) the data files during the interpretation phase of the research.<sup>90</sup>

## Summary of Focus 2 general studies

### General study 01: Persistent Archives Based on Data Grids<sup>91</sup>

This study focuses on the San Diego Supercomputer Centre’s project to develop a prototype for a “persistent archive”<sup>92</sup> based upon Data Grid technology for the U.S. National Archives and Records Administration (NARA). The study examines the minimal capabilities needed within grid technology for preservation of governmental records, focusing on activities related to the preservation of NARA’s selected digital holdings.

As discussed in this study, a Data Grid is

the set of abstractions that manage differences across storage repositories, information repositories, knowledge repositories and execution systems. Data Grids also provide abstraction mechanisms for interacting with the objects that are manipulated within the grid, including digital entities (logical namespace), processes (service characterizations or application specifications) and interaction environments (portals). The Data Grid approach can be defined as a set of services and the associated application programming interfaces (APIs) and protocols used to implement the services. The Data Grid is augmented with portals that are used to assemble integrated work environments to support specific applications or disciplines.<sup>93</sup>

<sup>88</sup> See Ballaux, “Case Study 26 Final Report,” op. cit., 11–12.

<sup>89</sup> In addition to the data received from a single space telescope, whose data collection functions are under the sole control of select project researchers, certain other data are received from a few other sources that are assumed to provide authentic and reliable data, such as the orbital information downloaded from the North American Aerospace Defense Command (NORAD) Web site (Ibid., 7).

<sup>90</sup> Ibid., 5, 13.

<sup>91</sup> See Moore, “General Study 01 Final Report,” op. cit.

<sup>92</sup> Persistent archives are “collections of digital entities that map from unique attribute values to a global, persistent identifier.” (Ibid., 5) which are distinct from “persistent storage systems “, where the emphasis is on providing “archival media that have a very long shelf life, such as heavy-ion beam encoded disk, film, etc.” (Ibid., 6).

<sup>93</sup> Ibid. For more information about Data Grids, see Fran Berman, Geoffrey Fox and Tony Hey (eds.), *Grid Computing: Making the Global Infrastructure a Reality* (Chichester, West Sussex: John Wiley & Sons, 2003).

Preservation environments for digital records are successful when they can separate the digital record from any dependence on the original creating infrastructure. Data Grid technology, which supports the management of distributed records, provides the software needed for infrastructure independence. Operating under the assumption that the capabilities of virtual data grids can be used to implement the traditional archival processes of appraisal, accession, arrangement, description, preservation and access, this study maps Data Grid capabilities to the archival processes required by the InterPARES 1 Model of the Preservation Function.<sup>94</sup>

### **General study 06: A Bayesian Belief Network: Supporting the Assessment of the Degree of Belief that a Recordkeeping System Maintains Authentic Digital Records**<sup>95</sup>

One of the research results of InterPARES 1 was the Authenticity Task Force's (ATF's) formulation of the *Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records*.<sup>96</sup> Quoting the ATF report: "A presumption of authenticity will be based upon the number of requirements that have been met and the degree to which each has been met. The requirements are, therefore, cumulative: the higher the number of satisfied requirements and the greater the degree to which an individual requirement has been satisfied, the stronger the presumption of authenticity." The purpose of this general study is to investigate a method of assessment for determining just how these requirements can actually be applied to assess the authenticity of a creator's digital records.

The approach used in the study is based on an interpretation of probability suggested by Bayesian theory, which holds that the concept of probability can be defined as the degree to which a person believes a proposition. Bayesian theory also suggests that Bayes' theorem can be used as a rule to infer or update the degree of belief in light of new information.

The researcher in this study conducted an experiment using a Bayesian Belief Network (BBN)<sup>97</sup> and data from case study 08 (Mars Global Surveyor Data Records in the Planetary Data System). A person, using data from the case study as evidence, was asked to express his degree of belief in the truth or falsity of the requirements at the terminal nodes of the BBN. As the degrees of belief were entered, the beliefs were propagated to the target hypothesis: "The Recordkeeping System is trusted and maintains authentic digital records." The result of this particular experiment was a very high degree of belief in the hypothesis.

Additional refinement and experiments are needed before the Bayesian approach outlined in this study can be considered a reliable and practical tool supporting the assessment of authenticity. However, among the results of this investigation is the identification of a number of conditional dependencies among the variables (requirements, attributes in record schema) that are not explicitly represented in the original statement of requirements. Additional consideration

---

<sup>94</sup> See Preservation Task Force (2002), "Appendix 5: A Model of the Preservation Function," version 6.0, in Duranti, *Long-term Preservation*, op. cit., 253–292. Online reprint available at [http://www.interpares.org/display\\_file.cfm?doc=ip1\\_ptf\\_model.pdf](http://www.interpares.org/display_file.cfm?doc=ip1_ptf_model.pdf). The InterPARES 1 Model of the Preservation Function was used in this general study because, at the time the study was conducted, the InterPARES 2 Chain of Preservation Model had not yet been developed.

<sup>95</sup> See Underwood and Isbell, "General Study 06 Final Report," op. cit.

<sup>96</sup> The *benchmark requirements* are the core information about a creator's digital records, and about the procedural controls exercised by the creator over the creation, handling and maintenance of its records, that set forth a basis for presuming or verifying the authenticity of the creator's digital records. Related to these are the *baseline requirements*, which are the minimum conditions necessary to enable the preserver to attest to the authenticity of copies of the creator's inactive digital records. For more information, see Authenticity Task Force, "Appendix 2: Requirements for Assessing and Maintaining the Authenticity of Electronic Records" in Duranti, *Long-term Preservation*, op. cit., 204–219. Online reprint available at [http://www.interpares.org/book/interpares\\_book\\_k\\_app02.pdf](http://www.interpares.org/book/interpares_book_k_app02.pdf). An abridged version of the benchmark requirements is provided in Appendix 21a. Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_book\\_appendix\\_21.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_book_appendix_21.pdf).

<sup>97</sup> A BBN is a graphical notation with an associated set of probability tables.

needs to be given to the relative contribution of requirements to the overall assessment of the authenticity of the records maintained in the recordkeeping system. Furthermore, there seem to be additional requirements (variables), particularly in the security area, that affect the authenticity of the maintained records but that were not included in this experiment.

### **General study 09: Survey of the Digital Recordkeeping Practices of GIS Archaeologists Worldwide<sup>98</sup>**

This general study summarizes the results of a forty-question, Web-based survey of the digital recordkeeping practices of GIS archaeologists worldwide. Invitations to participate in the survey were e-mailed to nearly 900 GIS archaeologists from sixty-nine countries worldwide. Additional invitations were posted to various professional archaeology- and GIS-related listservs and Internet discussion groups. The survey was available online for thirty days, during which 157 complete, or nearly complete, surveys were submitted.

Reliance on GIS technology has increased dramatically in archaeological research in the past two decades to the point where it is quickly becoming the de facto tool for facilitating data amalgamation, manipulation, synthesis, analysis and preservation in archaeological research. This fact raises concerns regarding the impact of GIS-facilitated research on archaeologists' ideas and attitudes towards digital recordkeeping practices, particularly in relation to their views towards creating and trusting digital records. The goal of the survey, therefore, was to gather and analyze baseline data about the existing digital recordkeeping knowledge and practices of GIS archaeologists worldwide to help gauge the current level of awareness and understanding within the global archaeological community about: (1) digital preservation issues, (2) digital recordkeeping practices and (3) the potential impact of such practices on the long-term preservation of accurate, reliable, authentic and accessible digital archaeological data and research records. This study was designed to complement the case study 14 (Archaeological Records in a Geographical Information System) research by providing more comprehensive and broad-based data on the recordkeeping practices of GIS archaeologists that could be used to help assess the representativeness of the findings in the case study.

The survey findings indicate that, despite the general perception of archaeologists as being meticulous and thorough in their documentation and recordkeeping practices, many GIS archaeologists currently engage in idiosyncratic or ad hoc file creation, management, preservation and/or documentation practices that have the potential to seriously compromise the accuracy, reliability, authenticity and accessibility of the files they create, especially over the long-term. One possible explanation for this apparent incongruity is that, until relatively recently, archaeologists typically have had little or no formal GIS training and have, instead, learned their GIS practices through a self-taught trial-and-error approach. These findings concur with the case study 14 findings.

Nevertheless, to the extent that the results of the survey can be considered representative of the habits of GIS archaeologists in general, it appears that, on the whole, there is a considerable and growing level of awareness among GIS archaeologists of the many technical, administrative, professional and societal issues surrounding the long-term preservation of their archaeological GIS data and research records.

---

<sup>98</sup> See Preston, "General Study 09 Final Report," op. cit.

### **General study 10: Preservation Practices of Scientific Data Portals<sup>99</sup>**

Geomatics and other science data are increasingly being discovered and accessed in data portals. Portals have a variety of names, such as data repositories, clearinghouses, catalogues, archives, geolibraries and directories. In this study, the term portal is used to encompass all of these. Portals can provide all or some of the following services: data search and retrieval, item descriptions, display services, data processing, platforms to share models and simulations, and data collection and maintenance. Much, but not all, of the data derived from portals are “raw” or unprocessed and require the user to interpret, analyze and/or manipulate them.

This general study was undertaken to collect structured information about the actual practices, standards and protocols currently used by existing science data portals in an attempt to understand how these types of entities address and demonstrate the lineage, reliability and authenticity of their data assemblages. To this end, a semi-structured survey of a random collection of thirty-two science data portals was conducted. The portals surveyed represented many different communities of practice in the sciences such as health, astronomy, biology, engineering, statistics, genetics, geosciences and ecology, to name a few.

Since the survey was undertaken for exploratory purposes, the sample size from each scientific discipline is small, thus limiting cross-disciplinary analysis. The study does, however, provide a deeper understanding of preservation practices in the natural and physical sciences in relation to data scientific portals. The report also includes an exploratory literature review that considers the importance of issues such as accuracy, reliability and authenticity in the management of scientific data exchanged through portals.

Among other things, this study demonstrates the heterogeneity of science data, the attention paid to scientific data quality, the complexity of the scientific methods that create and utilize these data and the myriad types and forms of scientific data that are created. In so doing, this study highlights the fact that the reliable and accurate creation of digital scientific records, and their authentic maintenance and preservation over the long term is a complex issue that needs to be addressed on several levels. This, in turn, emphasizes the need for a closer relationship between scientists and preservers to ensure that the digital records that scientists create and deem worthy of retention are created and managed in a manner that is conducive to their long-term preservation, and to ensure that preservers are able to preserve and provide access to the records in a manner that is sensitive to the often unique and varied requirements related to both the preservation and use of these types of records.

### **General study 12: Validation of the InterPARES 2 Project Chain of Preservation Model Using Case Study Data<sup>100</sup>**

This general study presents the results of a walkthrough of an early draft of the InterPARES 2 Project Chain of Preservation (COP) model using case study data. Walkthroughs using case study data are an effective way to test whether a model, design, program code or user interface achieve what is intended and to improve the quality of the product. A walkthrough is a peer group review of any information system product. A walkthrough of an activity model, such as the COP model, is concerned with the functionality of the system.

---

<sup>99</sup> See Lauriault and Craig, “General Study 10 Final Report,” op. cit. Additional findings and discussion stemming from this study can be found in Tracey P. Lauriault et al. (2007), “Today’s Data are Part of Tomorrow’s Research: Archival Issues in the Sciences,” *Archivaria* 64 (Fall): 123–179.

<sup>100</sup> William Underwood, Kevin Glick and Mark Wolfe (2007), “InterPARES 2 Project - General Study 12 Final Report: Validation of the InterPARES 2 Project Chain of Preservation Model Using Case Study Data.” Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_gs12\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_gs12_final_report.pdf).

The COP is a generic model of the processes involved in creating, maintaining, selecting and preserving authentic digital records. The COP model is prescriptive as well as descriptive. It prescribes criteria for determining whether digital records can be presumed to be authentic and it prescribes a method for applying these criteria. However, because the model was developed via functional decomposition rather than generalized from specific archival objects and information used to manage those objects, the following question arises: How can a user be sure that the model actually applies to the creation, maintenance, selection and preservation of digital records? The purpose of this study, therefore, is to demonstrate whether the COP model applies to specific record-making, recordkeeping and preservation systems for digital records and to refine and validate the model by conducting a walkthrough of the model using InterPARES 2 case study 08 (Scientific Data Records from a NASA Spacecraft Mission) data. As noted earlier, case study 08 was designed to collect information about the Planetary Data System (PDS) Space Science Data Archive.

NASA refers to the PDS as an active archive. Copies of the scientific datasets are transferred to the National Space Science Data Center (NSSDC) for long-term preservation. The NSSDC is referred to as a deep archives. The PDS activities of data preparation and management of datasets in the PDS Archive are similar to the COP activities “A2. Manage Records in a Record-making System”<sup>101</sup> and “A3. Manage Records in a Recordkeeping System.” The management at the NSSDC of scientific datasets from the PDS and from other space science disciplines appears to be similar to the COP activity “Select and Preserve Records.” However, further analysis of case study 08, together with the results of the walkthrough using the case study’s data, reveals that selection (appraisal) and preservation were central aspects of the PDS design and operation.

An interesting aspect of the walkthrough using the case study 08 data is that the description activity seems to take place during the record-making phase rather than during the preservation phase following the transfer of the records to the organization responsible for the long-term preservation of the records, as is prescribed in the COP model. Another interesting aspect of this case study is that parts of the appraisal and validation activities take place in the PDS *before* the recordkeeping phase, while in the COP model they take place *after* the recordkeeping phase. A possible explanation for this relates to the PDS management decision to actively involve scientists in the archiving process, which is based on the rationale that the scientists who create and use the data products are better able to describe and appraise them than are archivists (or scientists) far removed from the mission and data creation. Because of the expense of space science missions, the investment dictates early description, appraisal and validation of the datasets.

Of the sixty-eight lowest-level activities in the draft COP model, data from the case study were found to correspond to forty-six of those activities. There is no corresponding data in the case study for seven of the COP activities, while no data were collected for fifteen of the NSSDC activities that would correspond to the long-term preservation activities of the COP model.

These findings do not invalidate the COP model, rather they emphasize that it is a more general model of recordkeeping activities than the activities of scientific data recordkeeping in certain domains. In fact, the walkthrough of the COP model using the case study 08 data demonstrates that, overall, there is an interpretation of the record-making, recordkeeping and some of the preservation activities of the COP model in the domain of scientific data archives. That is to say, the COP model is satisfiable in this domain. A more thorough validation of the COP model would require further walkthroughs of the COP preservation activities using case study data with regard to the NSSDC preservation activities and walkthroughs using case study data from other scientific domains.

---

<sup>101</sup> This activity was called “Manage Records Creation” in the earlier draft version of the COP model used for the walkthrough.



## Findings specific to the scientific sector

The findings of the Focus 2 case and general studies indicate that, while there are certain obvious commonalities in the record-making, recordkeeping and record preservation concerns among creators in the sciences, there are also some significant differences in attitude, procedures and concerns between scientists in different disciplines, as well as among scientists within the same discipline in some cases. For the most part, the main focus in the sciences is on *data*, while the concept of record, at least in the archival sense of the term, receives far less attention or consideration. To some degree, the ambivalence of many scientists toward the archival concept of record is due to the fact that the concept of record as used and understood by most scientists is imbued with very specific and particular meaning that, for the most part, stands in stark contrast to the way the term is used and understood by archivists and other preservers. A somewhat similar situation is evident in relation to the concepts of authenticity, reliability and accuracy. Finally, despite the inherent experimental nature of scientific inquiry, evidence for the adoption and use among scientists of innovative, new or cutting-edge digital technologies is mixed. One indication of this is the use of customized and/or custom-built software applications, which was actually found to be higher among Focus 3 (government) creators than among Focus 2 creators. In fact, *all* of the creators investigated in the Focus 3 case studies were found to use substantially customized, commercial-off-the-shelf (COTS) software applications and/or custom-built applications (as well as unmodified COTS applications in most cases), while the same was not the case among the Focus 2 creators. One likely explanation for this is that, because scientists need to be able to easily share and exchange data, analysis results, etc., their activities are likely more sensitive to and influenced by the push to establish and follow industry standards regarding data creation and maintenance, as well as by the need to rely on applications and data formats with more robust interoperability functionalities and capabilities.<sup>102</sup>

## The concept of record

As was noted for the arts creators examined in Focus 1, the fundamental concept of what constitutes a “record” in a scientific context is open to wide and varied interpretation. This fact was made evident to Focus 2 Task Force members during both the review of the science literature and during the case and general studies research. The MIT Appraisal Project serves as a good example of the former. On the one hand, the MIT project, which focused on the nature of scientific activity and the scientific record, defines the scientific record as including experimental designs, documentation of instrumentation, experimental data records and analyses of experimental results; all entities that are in close alignment with the way the term record is interpreted in most archival contexts. On the other hand, the project also includes in its definition of the scientific record the *publication* of results in technical reports, conference proceedings and journal articles; all of which, because they are publications, are not considered by most archivists to be records—except in very specific and limited contexts (such as where the offprint of an author’s journal article is retained by the author as an evidentiary record of the act of publishing the article).<sup>103</sup> The process of publication was also seen by some of the creators in Focus 2 as

---

<sup>102</sup> For more detailed discussion of the impact of technology in relation to the records management practices noted in the Focus 2 case and general studies, see the analysis of Domain 1 research question 7 in the Domain 1 Task Force Report, *op. cit.* See also the Domain 3 Task Force Report, *op. cit.*, especially the “Findings” and “The Domain 3 Research Questions” sections.

<sup>103</sup> See Joan K. Hass, Helen W. Samuels and Barbara T. Simmons, *Appraising the Records of Modern Science and Technology: A Guide* (Cambridge, MA: Massachusetts Institute of Technology, 1985); and Joan K. Hass, Helen W. Samuels and Barbara T. Simmons (1986), “The MIT Appraisal Project and its Broader Implications,” *American Archivist* 49(3): 310–314.

instrumental in conferring legitimacy or “recordness” to the findings and outputs of their research.

Further complicating the situation in the sciences is the fact that, because of the inherently experimental and probatory nature of scientific inquiry, scientists view their research data, outputs and findings as *provisional* and continuously subject to revision by themselves and/or by other scientists. At a very fundamental level, it appears that this sense of the inherent impermanence of scientific research permeates and may, in many cases, even undermine the importance or relevance that scientists place on the characteristics of fixed form and stable content, both of which are central to the concept of recordness in an archival context. If so, then it is not too difficult to surmise how this attitude might shape, at a theoretical level, the basic concept of what constitutes a record in the eyes of many scientists, as well as, at a more practical level, how scientists perceive and carry out the creation, management, use and disposition of science records.<sup>104</sup>

### **Authenticity, reliability and accuracy**

Of the three concepts of authenticity, reliability and accuracy, it is the concept of accuracy, especially in relation to what is often collectively referred to as “data quality,” that receives the greatest attention in the science literature and in the Focus 2 case and general studies. This appears to be due in part to the fact that scientists typically focus more on data than on records, and in part to the centrality of data accuracy or quality to the validity of scientific research. Moreover, as the Focus 2 studies show, these concepts are interpreted and addressed in different ways by creators both between science disciplines and even, in some instances, within like disciplines; not unlike the situation noted earlier with creators in the arts. As the results of the Science Data Portals study clearly demonstrate, this situation is further compounded by the vast scope and breadth of scientific research, and the myriad forms and types of data records that this situation produces. To begin to successfully untangle and address the complex issues surrounding the reliable and accurate creation of digital scientific records and their authentic maintenance and preservation over the long term, preservers must, as a first step, develop a clearer understanding of the many nuanced ways that the fundamental archival concepts of authenticity, reliability and accuracy play out in the sciences.

Observations derived from the Focus 2 studies suggest that accuracy, which tends to be the most important data/records-related issue for scientists, is most commonly associated with, and addressed in relation to, the risk of having inaccurate data. Not surprisingly, it is also evident that the more critical the need for accuracy is, the more rigorous are the quality requirements and checks that tend to be established. Likewise, the more automated the data creation process is, the more technical the quality assessment procedures tend to be and, consequently, the more reliant the creators are on the technical systems in place, with human checks typically playing a secondary role in the assessment process.<sup>105</sup> On the other hand, in situations where very high data quality and accuracy levels are not as critical, attainable or practical, and where the data creation and quality assessment processes are less automated and more reliant on human

---

<sup>104</sup> For more detailed discussion of these findings on the concept of record in relation to the Focus 2 case and general studies, see the analysis of Domain 1 research question 4 in the Domain 1 Task Force Report, *op. cit.*

<sup>105</sup> See, for example, case studies 08 (NASA Spacecraft Mission), 19 (Electronic Engineering and Manufacturing Records) and 26 (MOST Satellite Mission).

intervention, concerns about professional practice and reliance on the trust associated with the integrity and authority of external data providers are given greater emphasis by the creators.<sup>106</sup>

The Focus 2 studies also reveal that, although, in many cases, scientists are unlikely to use the terms “authentic” or “authenticity” in relation to the data and records they generate or use in their research, the fundamental archival concepts underlying these terms are, nevertheless, often addressed. In some situations this is done implicitly, such as is noted in the Archaeology GIS and MOST Satellite Mission studies, where the authenticity of the datasets in both studies is assumed based on a presumption that the data are obtained from trustworthy sources. In other instances, the assessment of authenticity is more explicit, even if what is being assessed is not always identified using the terms authentic or authenticity. The Cybercartographic Atlas and NASA Spacecraft Mission studies provide good examples of this more explicit approach to authenticity issues. In the former study, the concept of authenticity (or at least its identity component<sup>107</sup>) is imbedded as an element (i.e., data lineage) within the more general concept of spatial data quality.<sup>108</sup> In the latter study, the concept of integrity—which, as noted, is one of the key components of authenticity—is explicitly identified as one of the characteristics of data records that is of primary concern to planetary scientists.<sup>109</sup>

### Science data archives

Scientists in many fields recognize the value of archiving digital scientific data, distributing the data to scientists in the field and preserving the data for future use. In fact, as the Scientific Data Portals study (general study 10) clearly demonstrates, there are scientific data “archives” in many fields of science, particularly in the areas where there is substantial government sponsorship; for example, the Space Sciences and High-Energy Physics.<sup>110</sup> Some of the scientific data archives include digital copies of scientific publications related to the scientific data. However, most institutional technical reports, conference records (proceedings) and journal

---

<sup>106</sup> See, for example, case studies 06 (Cybercartographic Atlas of Antarctica) and 14 (Archaeological Records in a GIS). It is noted, however, that the creator of the Atlas, unlike the creator of the Archaeological GIS, also relies heavily on good metadata provided by external data sources and on the metadata related to the Atlas modules themselves.

<sup>107</sup> As clarified by the InterPARES 1 Authenticity Task Force, authenticity is composed of two key elements: identity and integrity. “The *identity* of a record refers to the distinguishing character of a record, that is, the attributes of a record that uniquely characterize it and distinguish it from other records. From an archival diplomatic perspective, such attributes include: the names of the persons concurring in its formation (i.e., its author, addressee, writer, and originator); its date(s) of creation (i.e., the date it was made, received, and set aside) and its date(s) of transmission; an indication of the action or matter in which it participates; the expression of its archival bond, which links it to other records participating in the same action (e.g., a classification code or other unique identifier); as well as an indication of any attachment(s) since an attachment is considered an integral part of a record. The *integrity* of a record refers to its wholeness and soundness: a record has integrity when it is complete and uncorrupted in all its essential respects” (MacNeil et al., “Authenticity Task Force Report,” op. cit., 47).

<sup>108</sup> In general, data lineage speaks to the provenance or origin of a particular set of scientific data, which is essential to determining the accuracy, currency and validity of derived information and any assumptions, hypotheses or further work based on those data. Expressions of data lineage are found in the metadata that accompany the datasets or in other associated documentation such as peer-reviewed papers, reports, headers or notations. Data lineage or provenance is an area of considerable discussion in the sciences. See, for example, Yogesh L. Simmhan et al. (2005), “A Survey of Data Provenance in e-Science,” *Special Interest Group on Management of Data (SIGMOD) Record* 34(3): 31–36. Available at <http://www.sigmod.org/sigmod/record/issues/0509/p31-special-sw-section-5.pdf>.

<sup>109</sup> For more detailed discussion of the concepts of authenticity, accuracy and reliability in relation to the sciences in general, see the section titled “Conceptual Analysis: Authenticity, Accuracy and Reliability in the Literature of the Sciences” in the Domain 2 Task Force Report, op. cit. For further discussion of these concepts in relation to the findings of the Focus 2 case and general studies, see the section titled “Authenticity, Accuracy and Reliability in the Science Focus Case and General Studies” in the Domain 2 Task Force Report, *ibid*.

<sup>110</sup> Although, as the results of the general study 09 survey demonstrates, there are certain fields in the sciences, such as archaeology, that currently are conspicuously under-serviced by data repositories. See discussions of this issue on pp. 54 and 90–91 in the General Study 09 Final Report.

articles are preserved and made available to scientists and the public by science libraries, not by science archives. Redundant paper copies of scientific publications distributed throughout the world have reduced the risk of losing the publication part of the scientific record. However, with the advent of e-journals for scientific publications and the publication of technical e-reports and conference e-proceedings on Web sites, there is an increasing risk that without e-science record archives some of these elements of the scientific record will be lost or their authenticity will be in question.

## Focus 3 – Government

### Research team

The group of co-investigators conducting research on the records of e-government activities included a combination of researchers drawn from the public archives environment and from the university environment. Thus, the Focus 3 Task Force was an ideal blend of practitioners and researchers. During the course of five years, representation from the national archives participating in the Project changed as the primary work commitments of individual researchers shifted. The following list includes the names of all who contributed to the work of the Focus 3 Task Force at some point over the duration of the Project.

#### *Chair:*

Philip Eppard                      Jan 2002 - Dec 2006

#### *Researchers:*

Richard Blake	National Archives of the United Kingdom (withdrawn)—Working Group 3.3
Margaret Campbell	Archives of Nova Scotia (withdrawn)—Working Group 2.3
Filip Boudrez	City Archives of Antwerp, Belgium—Working Group 3.3
Hannelore Dekeyser	Katholieke Universiteit Leuven, Belgium—Working Group 2.3
Terry Eastwood	The University of British Columbia, Canada—Working Group 1.3
Fynnette Eaton	U.S. National Archives and Records Administration—Working Group 3.3
Philip Eppard	University at Albany, State University of New York, USA—Working Group 2.3
Elaine Goh	National Archives of Singapore—Working Group 3.3
Ken Hannigan	National Archives of Ireland—Working Group 1.3
Ken Hawkins	U.S. National Archives and Records Administration—Working Group 3.3
Hans Hofman	National Archives of the Netherlands—Working Group 3.3
Richard Marciano	San Diego Supercomputer Center, USA—Working Group 3.3
Terry Maxwell	University at Albany, State University of New York, USA—Working Group 1.3
John McDonough	National Archives of Ireland—Working Group 1.3
Sue McKemmish	Monash University, Australia—Working Group 3.3
Tom Quinlan	National Archives of Ireland—Working Group 1.3
Shelby Sanett	U.S. National Archives and Records Administration—Working Group 3.3
Jim Suderman	City of Toronto Archives, Canada—Working Group 3.3
Ken Thibodeau	U.S. National Archives and Records Administration—Working Group 3.3
Malcolm Todd	National Archives of the United Kingdom—Working Group 1.3

*Research Assistants:*

Melissa Adams	The University of British Columbia, Canada
Deidre Brocklehurst	The University of British Columbia, Canada
Natalie Catto	The University of British Columbia, Canada
Terra Dickson	The University of British Columbia, Canada
Jennifer Douglas	The University of British Columbia, Canada
Adam Farrell	The University of British Columbia, Canada
Dan Farrell	The University of British Columbia, Canada
Coby Falconer	The University of British Columbia, Canada
Fiorella Foscarini	The University of British Columbia, Canada
Joshua Hauck-Wheaton	University at Albany, State University of New York, USA
Peggy Heger	The University of British Columbia, Canada
Ted Hoppenstedt	University at Albany, State University of New York, USA
Janine Johnstone	The University of British Columbia, Canada
Eleanor Kleiber	The University of British Columbia, Canada
Tracey Krause	The University of British Columbia, Canada
Andrea Lam	The University of British Columbia, Canada
Yvonne Loiselle	The University of British Columbia, Canada
Rachel McMullin	University at Albany, State University of New York, USA
Brenda McPhail	University of Toronto, Canada
Catherine Miller	The University of British Columbia, Canada
Elisheba Muturi	The University of British Columbia, Canada
Jane Morrison	The University of British Columbia, Canada
Carolyn Petrie	The University of British Columbia, Canada
Peter Runge	University at Albany, State University of New York, USA
Vincent Schillaci-Ventura	The University of British Columbia, Canada
Geneviève Shepherd	The University of British Columbia, Canada
Wendy Sokolon	The University of British Columbia, Canada
Mary Beth Sullivan	University at Albany, State University of New York, USA
Adele Torrance	The University of British Columbia, Canada
Melanie Wallace	The University of British Columbia, Canada
Carol Ward	University at Albany, State University of New York, USA
Reginald White	University at Albany, State University of New York, USA
Mark Wolfe	University at Albany, State University of New York, USA
Jessica Zacher	University at Albany, State University of New York, USA

**Selection of Focus 3 case studies**

Defining the precise scope of Focus 3 activity was a question that occupied researchers in the first months of the Project. Although it was clear that the general focus was to be on the dynamic, interactive and experiential records in the governmental sector, there was some discussion as to whether Focus 3 should concentrate only narrowly on activities that come under the term of e-government, which is generally restricted to the use of information technology by government to improve the delivery of services to the citizenry. Alternatively, it was suggested that the work of Focus 3 might want to cover use of such technologies in the internal workings of government as well as in the delivery of external services. In the end, the Focus 3 researchers decided to lean more toward the traditional definition of e-government.

The Focus 3 case studies were selected to represent a range of governmental levels: national, state and provincial, and city. In addition, the range of case studies was expanded with the addition of two special cases: (1) an independently developed database of international treaties and related documents and (2) a Web site maintained by an enterprise association that is similar to government information and service Web sites. This brought to eight the total number of case studies completed in Focus 3. Two additional case studies were proposed and approved, but were later withdrawn (case study 11, Nova Scotia Business Registry Service and case study 23, UK Knowledge Network). In the first instance, the case study researcher was unable to continue participating in the Project. In the second instance, the researcher ultimately could not receive the necessary permissions to carry out the study.

### Summary of Focus 3 case studies

The following brief descriptions of the Focus 3 case studies identify the digital entities that were examined, while also highlighting some of the more salient findings from each study. For a more detailed contextual analysis of these case studies in relation to the records creators and the nature of the activities resulting in document creation, see the “Characterization of the Case Studies” section in the Domain 1 Task Force Report. See Appendix 3 for a complete list of participants responsible for each case study in this focus.

#### Case study 05: Archives of Ontario Web Exhibits<sup>111</sup>

This study focuses on the creation and posting of three Web exhibits within the operational contexts of two publicly funded archival institutions; the Archives of Ontario and the City of Toronto Archives. The goal of this study is to examine the processes by which Web exhibits in hopes of identifying indicators of authenticity, accuracy and reliability in experiential and interactive records. The records that the research team studied are, primarily, the final format of the exhibit records and, secondarily, the records created and used during the actual creation of the virtual exhibits. The types of records examined include scanned images, recorded sound and text files, all of which are combined into Web pages using HTML and accessed using standard Web browser applications. The basic process of creation involves decisions by Archives’ staff on the subject of an exhibit, on the choice of records used to include in the exhibit, and on how to present the chosen records.

Although the Web exhibits studied here are static documents from the creator’s perspective, the interaction of the rendering system (i.e., Web browser application) with the exhibits can cause considerable variance in the user’s perspective of the records. Corporate standards designed specifically to diminish or minimize the vagaries of this type of interaction for Web site development exist in the jurisdictions of both institutions participating in this study. However, this study reveals that the details of the preferred or “target” rendering system are only assumed by the creator and not actually conveyed to the Web exhibit users. As such, it is suggested that, insofar as experiential digital objects are “objects whose essence goes beyond the bits that constitute the object to incorporate the behaviour of the rendering system, or at least the interaction between the object and the rendering system,”<sup>112</sup> the Web exhibits in this study are experiential digital objects.

---

<sup>111</sup> See Jim Suderman et al. (2004), “InterPARES 2 Project - Case Study 05 Final Report: Archives of Ontario Web Exhibits.” Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs05\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs05_final_report.pdf).

<sup>112</sup> Duranti, InterPARES 2 Project Proposal, op. cit., 1.1-3.

Some of the key indicators found to impact the authenticity, accuracy and reliability of the digital objects examined in this study include:

- the presence of a nascent or “emerging” business process that is essentially “trust-based,” meaning that there is no established procedural context in terms of which individuals are to fulfil which roles in the creation process of the Web exhibits;
- a decentralized technological environment that is in a constant state of change;
- a limited or non-existent recordkeeping environment, particularly in relation to the digital records examined in this study, whose creation and maintenance was deemed to be ad hoc and at the discretion of the individuals responsible for the development and maintenance of the Web exhibits;
- legal and moral issues, particularly in relation to protecting copyright, providing accessibility to users with disabilities, upholding the requirements of donor agreements for the materials used in the exhibits, avoiding controversy in the selection of exhibit topics and source materials (censorship), and insuring the accuracy of the interpretive text;
- identification of the Web exhibits as records and not publications;
- the inability to determine the values the creating organizations place on the Web exhibit records due to the absence of any appraisals;
- spotty adherence to the InterPARES 1 benchmark requirements for supporting the presumption of the for authenticity of electronic records;
- the use of the same digital components in multiple records;<sup>113</sup> and
- the lack of feedback from Web exhibit users regarding their interactions with the exhibits.

### Case study 12: Antarctic Treaty Searchable Database<sup>114</sup>

The focus of this case study is the Antarctic Treaty Searchable Database (ATSD), an online, dynamic and interactive “policy archive” of digital records, and other supporting documents, resulting from the various measures that have been approved by the Antarctic Treaty Consultative Parties from 1959 to the present.<sup>115</sup> Now in its 8th edition,<sup>116</sup> the ATSD was originally produced for educational purposes in 1999 using materials from the U.S. Department of State as part of a two-year National Science Foundation project through the National Science, Technology, Engineering, and Mathematics Education Digital Library (NSDL). The ATSD is now used in many spheres outside of education and by many organizations worldwide, including by international governmental and non-governmental organizations, national government agencies, commercial institutions and industry. The objectives of this case study include determining the functional and conceptual challenges of modeling the creation and preservation of digital records and determining the technical and administrative challenges of re-purposing public domain records.

The ATSD interfaces with a proprietary search and data integration engine, the *Digital Integration System*<sup>TM</sup> (*DIGIN*®) from EvREsearch Ltd.,<sup>117</sup> which is designed to provide “technology independent” access to the records in the system. The *DIGIN*® engine automatically extracts and integrates information from digital records (and other digital objects)—all of which are maintained in ASCII format in the system—in response to user queries. Moreover, it does

---

<sup>113</sup> For a detailed discussion of the concept of digital components and the different relationships these objects may have with records, see the Domain 3 Task Force Report. See also Duranti and Thibodeau, “The Concept of Record,” op. cit.

<sup>114</sup> See Paul Arthur Berkman et al. (2005), “InterPARES 2 Project - Case Study 12 Final Report: Antarctic Treaty Searchable Database.” Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs12\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs12_final_report.pdf).

<sup>115</sup> See <http://aspire.tierit.com/>.

<sup>116</sup> The ATSD was in its 5th edition during the time that it was the focus of the InterPARES 2 case study.

<sup>117</sup> The *Digital Integration System*<sup>TM</sup> (*DIGIN*®) from EvREsearch Ltd. See <http://evresearch.com/Integration%20Engine.htm>.

this “without the use of markup or metadata that otherwise would create technology dependencies tied to “structured information” standards and legacy implementations.”<sup>118</sup>

Although not a true recordkeeping or preservation system, the ATSD nevertheless provides a number of identifiable records management functionalities, including: records registration, classification/organization, access control and retrieval. It also appears to incorporate certain functionalities designed to facilitate long-term preservation of the records in the database.<sup>119</sup> Like the Atlas in case study 06, the ATSD raises some important issues regarding long-term preservation and access of authentic, multi-source digital materials maintained in interactive, Web-based systems.

### **Case study 17: New York State Department of Motor Vehicles On-line Services System**<sup>120</sup>

The focus of this case study is the On-line Services System through which the New York State Department of Motor Vehicles (DMV) allows citizens to use a Web browser to conduct various online transactions related to activities such as registration renewal, driver’s license ordering and renewal, custom plate ordering and driver’s tests scheduling. Users conduct legal and financial transactions within the Web site, which generates records in a networked and online environment. In addition to these online transactions, the DMV’s Web site allows users to find information, download forms, read publications and statistics and send e-mail to the DMV. The DMV’s highly interactive online system features a complex set of interwoven electronic activities that complements the DMV’s internal electronic legacy system and the business rules that have applied to that system. Only a portion of the business transactions supported by the DMV can be accomplished via its Web site, and there are no services provided by the Web site that are exclusive to that medium. Of specific interest in this study are issues relating to records creation and maintenance within the online system, issues relating to ensuring accuracy, authenticity and reliability in the DMV’s recordkeeping system, and issues relating to technology.

The On-line Services System was designed in house, rather than using a commercial-off-the-shelf product; however, the DMV does use a mix of third-party and in-house products for software and security measures, including the implementation of a Public Key Infrastructure (PKI) that must conform to a strict, regulated legal framework. As a component of the New York state government, the DMV and its On-line Services System are required to comply with a number of state and federal laws, regulations and policies, including all applicable state and federal laws relating to freedom of information. Use and maintenance of the system is also subject to numerous internal DMV policies and standards.

The DMV maintains “core” records for each of its customers. A core record contains proof-of-identity information about a customer and must be created during an in-person office visit before the customer is eligible to conduct any transactions online. During an online transaction, the DMV system matches the identification information provided by the customer with the identification information in the core record file of that customer. If the two match, the information is assumed to be authentic. Because online customers only have access to transactions, not to the digital entities or the core records in the DMV’s mainframe database system, and because the system also limits what type of access each DMV employee has to the digital entities in the system, the creator is “99% sure” that its records are secure from unauthorized access by users both internal and external to the system.

<sup>118</sup> Berkman et al., “Case Study 12 Final Report,” op. cit., 41.

<sup>119</sup> See Ibid., 37–39 (specifically, the answer to case study research question no. 18).

<sup>120</sup> Philip Eppard and Mark Wolfe (2006), “InterPARES 2 Project - Case Study 17 Final Report: New York State Department of Motor Vehicles On-line Services System.” Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_c17\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_c17_final_report.pdf).



**Case study 18: Computerization of Alsace-Moselle's Land Registry**<sup>121</sup>

This case study examines the creation of a computerized land registry in Alsace-Moselle, a regional administrative entity in France. The primary focus of this case study is the long-term preservation of digital signatures within a dynamic information system designed to improve the efficiency of government-citizen relations in the context of the French civil law evidence system. To this end, the study documents the efforts of the Groupement pour l'Informatisation du Livre Foncier d'Alsace-Moselle (GILFAM), the administrative body that oversees development and maintenance of the computerized registry, and the Archives de France to agree on a method whereby relevant data may be transferred to the Archives while retaining their functionality. In so doing, the study also investigates the difference between authentication and authenticity, the design and implementation of a dynamic information system in the governmental sector, and how regulatory frameworks can better harmonize with archival requirements to help facilitate the long-term preservation of digital records.

The computerized Alsace-Moselle Land Registry consists of the transcriptions and scanned images of 40,000 existing paper land registry records within Alsace-Moselle, together with the inscriptions of new real estate transactions that have occurred since the system was implemented. Various procedural and state-of-the-art security technologies and methodologies are in place to ensure the authenticity, reliability and accuracy of each entry in the database. For example, each entry requires the signature of a judge, using a PKI that combines biometric access (fingerprint scan) and digital signatures. As well, there are sophisticated logging capabilities for recording all actions and transactions that take place in the system.

Of particular interest in this system is the use of a two-staged, software-controlled procedure that enables the use of digital signatures to provide continuous authentication services, rather than the more usual one-time authentication service provided by digital signatures in most systems. Nevertheless, the GILFAM, which has a legal responsibility to provide continued access to the land registry in a fashion that preserves its evidential value regardless of technological change, has not considered the problem of maintaining the digital entities, except through the mechanisms afforded by system vendors when upgrading the database management system. Nor has it considered the problem of maintaining the evidential value of digital signatures through technological evolution.

Apart from the preservation challenges stemming from the use of digital signatures, the eventual transfer of the data in the system to an archival institution is further complicated by the fact the land registry cannot be understood in the absence of the system's dynamic and interactive capabilities for organizing the records in the system and for providing access to the information that the records contain. For this reason, it is pointless to attempt to extract and transfer stand-alone documents from the system. Moreover, because of the complexity of the system and the costs involved in attempting to duplicate it, it is not possible for an archival institution to operate its own database management system, populated with inactive records of the land registry. The study discusses two possible solutions for overcoming these preservation obstacles. The first solution involves using an XML schema to serve as a translation device between the complex data model used by the land registry so that inscriptions can be imported into a relational database sufficiently simple to be maintained by an archival institution. The second solution is for the designated archival institution to grant to the GILFAM the permission

---

<sup>121</sup> Jean-François Blanchette, François Banat-Berger and Geneviève Shepherd (2004), "InterPARES 2 Project - Case Study 18 Final Report: Computerization of Alsace-Moselle's Land Registry." Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs18\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs18_final_report.pdf).

to act as the custodian of the land registry and transfer its inactive records to an archival database using the same software infrastructure.

### **Case study 20: Revenue On-Line Service (ROS)**<sup>122</sup>

This case study investigates Ireland's Revenue On-Line Service (ROS), a high profile e-government service offered to tax agents and customers by the Irish Revenue Commissioners. The ROS service enables the generation, maintenance, access and preservation of electronic-based tax and other records in a secure, networked, online environment. The case study focuses on the functionality and record creating and access properties of the ROS, especially in relation to the use, via a Revenue Public Key Infrastructure (PKI), of digital certificates within ROS and how these are managed. The claims that the system meets requirements for data authenticity, accuracy and integrity are also examined, particularly in the relation to the requirements of the *Irish E-Commerce Act 2000*.

The ROS system, which has been cited as a best practice in e-government products and has received several national and international awards, operates as both a stand-alone, offline and a networked, online application that links to a Web-based Revenue portal and allows authorized customers to access relevant tax information, complete and submit tax returns and, if necessary, make or arrange payments online.

At any given time, the ROS system only maintains a subset of the main body of a customer's records. Moreover, the records in this subset are record copies; the actual authentic copies of the digital records actually reside on Revenue's Integrated Taxation Processing (ITP) back-end database system. The ITP is the central component of Revenue's overall Integrated Taxation Services framework. The ROS system's architecture separates some of the functionality of the record-making and recordkeeping systems using business rules to dictate the records contained in the subset stored within the ROS database.

Record reliability and accuracy are, in part, addressed through the use of automated validation routines built into the ROS form templates to ensure that some of the data fields populated by users are accurate and complete. Authenticity is managed through the use of digital signatures administered via a PKI. The ROS system is very tightly integrated within the Revenue PKI environment, which is used to control access to the system and to ensure the authenticity, integrity, confidentiality and non-repudiation of all transactions between Revenue and its customers. However, although the Revenue PKI facilitates the transmission of authentic data into the ROS system, the case study researchers feel that this procedure cannot, at present, be used to continue to confer authenticity. This is due to the lack of "internal controls and procedures illustrating the mechanisms by which data elements are removed from the security wrapper, ingested into the ROS system, and processed," all of which, the researchers suggest, need to be better documented, articulated and maintained.<sup>123</sup> The "security wrapper" is a SOAP (Simple Object Access Protocol)<sup>124</sup> object that encompasses the entire transaction dataset received from the customer by the ROS system. Currently, Revenue retains the security wrapper to confer authenticity and non-repudiation over time; however, there is, at present, no defined policy regarding the retention and management of the security wrappers. Moreover, it appears that the security wrappers are retained only in the ROS system and not in the central ITP system.

---

<sup>122</sup> John McDonough, Ken Hannigan and Tom Quinlan (2005), "InterPARES 2 Project - Case Study 20 Final Report: Revenue On-Line Service (ROS)." Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs20\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs20_final_report.pdf).

<sup>123</sup> Ibid., 71.

<sup>124</sup> SOAP is an XML-based protocol that facilitates the exchange of information over HTTP between different applications running on different operating systems, with different technologies and programming languages.

Another key issue identified by the case study researchers is that, although creating agency maintains an “archive collection system” for certain PKI-related records such as public keys and certificates, it has not conducted a formal appraisal of the records in the ITP and its other back-end database systems, nor has it instituted a long-term archiving policy for data stored in either the ROS or the ITP back-end systems. This lack of formal appraisal and retention policies, especially with respect to the subset of ITP records and the security wrappers held in the ROS system, is identified as the most obvious deficiency in the current ROS system with respect to long-term preservation requirements.

### **Case study 21: Electronic Filing System (EFS) of the Supreme Court of Singapore**<sup>125</sup>

The focus of this case study is an online filing system of records of civil matters that is meant to facilitate the process of civil litigation through e-filing and electronic information services, as well as to implement the use of digital documents in electronic chambers or electronic courts in Singapore. Through the Electronic Filing System (EFS), law firms can electronically submit and sign their legal records to the courts online at all hours; they can also serve records to other firms electronically. The EFS also supports online searching of relevant case information and the extraction of records.

Due to the complexity of the business process and juridical context that governs the creation of various types of case files created within a single system in the EFS, this case study limits its examination to the filing of bankruptcy petitions and summons in chambers bankruptcy—records that the EFS of the Supreme Court of Singapore requires be submitted in digital form—focusing on an analysis of the authenticity and reliability of the records created within this judicial e-government context. More generally, the study examines the records creation process, appraisal and preservation of the records created within EFS, and the policies, procedures and standards taken to control the creation, modification and preservation of records within the system.

The EFS is comprised of standardized HTML style sheets, XML files, Visual Basic and PDF records. Essentially, the EFS is an integrated workflow application and document imaging system. All case files received at the Court’s end are stored onto WORM (Write Once Read Many) optical disks in a jukebox.

The reliability and accuracy of the records in the EFS are addressed through the application of documentary templates and strict procedural controls that govern the creation of the records. For example, if the information entered by the law firm is inaccurate, the Court’s registry staff will reject the record and request that the law firm correct the record and re-file its submission. Authenticity is addressed through the application of various procedures, protocols and technological systems, including: an “access control matrix” that identifies the names and various roles of the action officers as well as their access rights within the EFS (the EFS does not, for example, allow records filed by law firms to be changed or modified by the Court); a PKI, together with various software and hardware encryption technologies, to secure and authenticate the transmissions of records; and various “protective countermeasures” designed to protect the integrity of the system and its records, such as firewalls, anti-virus and intrusion-detection software and periodic, intentional hacking attempts by a government agency to expose any vulnerabilities with the system.

---

<sup>125</sup> Elaine Goh (2005), “InterPARES 2 Project - Case Study 21 Final Report: The Electronic Filing System (EFS) of the Supreme Court of Singapore.” Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs21\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs21_final_report.pdf).

**Case study 24: Preservation of the City of Vancouver GIS Database (VanMap)<sup>126</sup>**

The focus of this case study is VanMap, an extensive geographic information system (GIS) containing georeferenced data on the City of Vancouver. The system is maintained by the City's Information Technology Department and incorporates data that are supplied and updated on a regular basis by the Engineering, Planning, Social Planning, Permits and Licenses, Real Estate Services and other City departments and, to a much lesser extent, by external agencies such as the provincial government and crown agencies. The goals of the case study are to understand the types of records that are generated by VanMap and to explore issues relating to security, recordkeeping and long-term preservation in relation to the use of a GIS.

Essentially, VanMap is a Web-based reference tool created to provide City staff with ready access to a detailed visual realization of the City, in the form of interactive maps, as they carry out a variety of administrative tasks. Its fundamental purpose is to meet the needs of internal City users in providing zoning, permitting, licensing, emergency planning, utilities management, traffic control, street maintenance and numerous other services to Vancouver's citizens and businesses.

The data in VanMap are constantly being updated, with the frequency of the updates varying considerably. In most cases, the updating process physically overwrites (replaces) any existing data with the new data. For data that are overwritten, there is no way to track updates over time or to access previous instantiations of the data since copies of overwritten data are not routinely kept. This process presents significant conceptual and technical challenges relating to the need both to ensure that the City government can be held accountable for the way in which the data in the system at a specific point in time are used to support decision-making activities, and to preserve the authenticity of those data and the experience of accessing them in the form of interactive maps.

The reliability and accuracy of the data in VanMap are the responsibility of the originating departments and external agencies, not the VanMap Team. There are, however, no formal data quality agreements in place between the various data providers and the VanMap Team. Instead, the Team relies on informal agreements expressed verbally or through e-mail, supported by what is characterized as an "element of trust" that exists between the Team and the various City departments. In other words, the VanMap Team trusts that the data provided by each department are reliable and accurate since, as one Team member noted, "we all work for the City."<sup>127</sup> Furthermore, it is noted that only a limited number of well-trained staff within the departments are able to update the data using data entry formats that are highly automated and strictly controlled. With respect to data provided by external agencies (e.g., utility agencies), these are usually accompanied by disclaimers about the accuracy and reliability of the data. In such cases, the VanMap Team is careful to post the disclaimers in association with the relevant data.

The data provided to the VanMap Team by the City departments and external agencies are not altered in any way that would affect the authenticity, reliability or accuracy of the data. Moreover, once the data are input by specified staff they are not modified by the way they are used in VanMap, since VanMap is a read-only system. Thus, the VanMap Team can, in the end, only guarantee that the data that appear in VanMap are as authentic, reliable and accurate as the original source.

---

<sup>126</sup> Evelyn McLellan (2005), "InterPARES 2 Project - Case Study 24 Final Report: City of Vancouver Geographic Information System (VanMap)." Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs24\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs24_final_report.pdf).

<sup>127</sup> Ibid., 10.

### **Case study 25: Legacoop Bologna Web Site<sup>128</sup>**

The focus of this case study is the Web site of Legacoop Bologna, a business association cooperative operating in the Bologna area of Italy that promotes the development of cooperation, mutual aid and solidarity, stimulates economic relationships among member cooperatives and spreads the principles and values of the cooperation ideal. The Web site provides detailed information about the cooperative, along with facts about the economic and social situation within the local territory, to its members and the general public. It also offers specialized online interactive services to its members.

Although technically a support service provided by an association of private businesses, the Legacoop Web site could fall into the category of e-government as to its function if not as to its nature. Legacoop is indeed an institutional player in the local socio-economic system and its Web site, which presents both static and dynamic information, constitutes an essential instrument for maintaining what might be called the “cooperative network” of the association’s members. A restricted area of the Web site provides enhanced services to the association’s members, while an unrestricted area also provides basic information and services to the general public. Further, one of the main purposes of the Web site is to increase cooperation between Legacoop’s associate businesses and the city.

From a preservation viewpoint, this complex Web site presents significant challenges to preserving its authenticity over the long term. From a policy viewpoint, it provides a good starting point for conceptualizing the policy issues presented by the preservation of digital material of a private organization that behaves very much like a government. For these reasons, the case study addresses a number of interesting long-term preservation issues that are particularly relevant to the work of both the Policy Cross-domain and the Focus 3 Task Forces.

### **General study conducted by Focus 3**

At the beginning of the Project, the Focus 3 Task Force conducted a general study in an effort to gauge how extensively governments were using the World Wide Web to provide information and services to citizens. The Government Web Sites Survey (general study 08) used a typology developed by the National Archives of Australia to classify Web sites as static, static with form-based interactivity, Web sites based on dynamic data access, and dynamically generated Web sites. A total of 321 Web sites in six countries were analyzed, with 41% judged to be static, 34% static with forms, 19% dynamic data access and 6% dynamically generated.<sup>129</sup> This rudimentary survey gave researchers in the government focus a sense of the terrain in which their case study research would take place. It also provided a context in which researchers could understand the extent to which interactive and dynamic environments were shaping the delivery of e-government services.

### **Findings specific to the governmental sector**

Of the three focus areas in InterPARES 2, the governmental sector offered record creation and maintenance systems that were most similar to those studied by InterPARES 1. Nevertheless, because of the emphasis in InterPARES 2 on interactive, experiential and dynamic systems,

---

<sup>128</sup> Mariella Guercio (2004), “InterPARES 2 Project - Case Study 25 Final Report: Legacoop of Bologna Web Site.” Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_cs25\\_final\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_cs25_final_report.pdf).

<sup>129</sup> See Wolfe, “General Study 08 Final Report,” op. cit.

relative to the primarily static systems examined in InterPARES 1, there were of course major differences in relation to the ways in which data and records were being generated and maintained by the creators studied in InterPARES 2. Governments carry out their core functions through the creation of records, usually under the constraints of fairly stringent guidelines or rules. The case studies carried out in Focus 3, therefore, show governments adapting their methods to these new digital environments to fulfill traditional roles with greater speed and accessibility.

These same case studies also demonstrate that there is a greater awareness among creators in the government focus of the formal elements of a document than there is among creators in the other two focuses. In fact, seven of the eight government focus case studies (the Antarctic Treaty Searchable Database being the exception) provide at least an implicit reference to the extrinsic and intrinsic elements of documentary form in relation to the records being examined, while four of the studies include explicit and often detailed discussions of the intrinsic and extrinsic elements. This stands in sharp contrast to most of the creators in the other two focuses who seem to exhibit little interest in, or awareness of, such matters. In fact, when asked to identify the key formal elements and attributes of their digital entities, creators in the arts and science focuses typically assumed the question was referring to issues such as media formats (e.g., VHS, DVDCam, etc.), file formats (e.g., .doc, .pdf, etc.) and/or the software applications that produce files in these formats, rather than to the intrinsic and extrinsic elements of documentary form that speak to the authenticity of the records.<sup>130</sup>

Most of the case studies in Focus 3 mainly depict governments doing traditional activities that are now carried out in digital formats. Thus, there is the question of adaptation of existing traditions to digital delivery and service mechanisms. In some cases, the new methods exist alongside older paper-based methods, acting as complements or alternatives to the older methods. On the whole, however, the purposes of creating documents have remained unchanged despite the switch or expansion to digital methods.<sup>131</sup>

New technologies do allow existing purposes to be augmented, however. When documents provide a service or information to citizens, the move to e-government allows for faster filing, quicker retrieval and improved access. Providing government services in an electronic environment reduces the costs of transactions while improving the speed and flexibility of the transactions. Additionally, digital security measures, such as the use of a Public Key Infrastructure,<sup>132</sup> allow for security and confidentiality to be safeguarded at every stage of the cycle in many cases. On the other hand, the switch to new technologies can, in some cases, create new security concerns.

As these changes represent traditional methods being carried out in a new fashion, existing methods of document creation remain largely unchanged. In most cases, only a few steps have been added to account for the new technologies used, such as digital security. In the few cases where new technologies have resulted in new methods of creation, it appears that the “newness” of the digital world obscured the fact that existing methods could still be carried out without large modifications.<sup>133</sup>

---

<sup>130</sup> For more detailed discussion of these findings in relation to the Focus 3 case studies, see the analysis of Domain 1 research question 3 in the Domain 1 Task Force Report, *op. cit.*

<sup>131</sup> For more detailed discussion of these findings in relation to the Focus 3 case studies, see the analysis of Domain 1 research question 1 in the Domain 1 Task Force Report, *ibid.*

<sup>132</sup> See, for example, the use of a PKI and digital signatures in case studies 17 (New York State DMV), 18 (Alsace-Moselle Land Registry), 20 (Revenue On-Line Service) and 21 (Supreme Court of Singapore).

<sup>133</sup> For more detailed discussion of these findings in relation to the Focus 3 case studies, see the analysis of Domain 1 research question 2 in the Domain 1 Task Force Report, *op. cit.*

Only about one half of the organizations examined in the Focus 3 case studies is actually creating records, at least according to the definition of a record as established by diplomatics and adopted by InterPARES 1.<sup>134</sup> In the remaining cases, some of the documents produced can be considered partial or potential records. Failure to meet record specifications frequently resulted from the use of hypermedia features, where the fluid and changing data produce a document with no fixed form that is never properly set aside. Other systems rely on providing services through dynamic databases, which are never fully fixed and set aside.<sup>135</sup>

An analysis of the e-government literature shows that there is awareness of and concern for the authenticity and reliability of records, although the terms are often used interchangeably. Accuracy was less of an issue, although one of the virtues of the use of digital forms for creating and entering data into systems is that a digital system can, through various automated and semi-automated controls, enforce a certain amount of accuracy that would not be easily achieved in a manual system. Authentication was a key concern in the literature and in some of the case studies, but in both cases it was sometimes taken as a guarantee of authenticity rather than a verification of authenticity at a certain point in time.<sup>136</sup>

The process of dealing with digital records should lead to a greater awareness of the need for records retention and preservation policies, especially because of the tendency of digital media to become inaccessible due to technological obsolescence. Ideally, this factor should encourage records creators to become more proactive in considering the retention needs of their organizations. This is, in fact, the case in some of the case studies. However, many others displayed a total lack of formalized preservation strategies. In some instances there were no long-term retention requirements on the records created—everything was considered part of the live system as long as it was needed. Even in these cases, however, the benchmark and baseline requirements set forth in InterPARES 1 should be adhered to if authenticity, accuracy and reliability of the records are to be demonstrated.<sup>137</sup>

## General Conclusions of the Focus Task Forces

All members of the Focus Task Forces also belonged to a Domain Task Force, where their experiences with their own case study and their knowledge of the case studies being done in their specific Focus could be shared with the other researchers and inform the work being done on Records Creation and Maintenance (Domain 1), Authenticity, Reliability and Accuracy (Domain 2) and Appraisal and Preservation (Domain 3). Questions addressing the key issues in each domain were included in the basic case study questionnaire.<sup>138</sup> The results of the analyses of the case studies that were conducted by the domains are available in the domain reports.

---

<sup>134</sup> For details on how “record” is defined by traditional diplomatics and by InterPARES 1, see Duranti, “Diplomatics,” op. cit.; and Duranti, *Long-term Preservation*, op. cit., especially Part 1 and Appendices 1 and 2. For a more concise summary, see the section titled “Definition of Record” in the Domain 1 Task Force Report, op. cit.

<sup>135</sup> For more detailed discussion of these findings in relation to the Focus 3 case studies, see the analysis of Domain 1 research question 4 in the Domain 1 Task Force Report, *ibid.*

<sup>136</sup> For more detailed discussion of the concepts of authenticity, accuracy and reliability in relation to e-government in general, see the section titled “Conceptual Analysis: Authenticity, Accuracy and Reliability in the Literature of e-Government” in the Domain 2 Task Force Report, op. cit. For further discussion of these concepts in relation to the findings of the Focus 3 case studies, see the section titled “Authenticity, Accuracy and Reliability in the Government Focus Case Studies” in the Domain 2 Task Force Report, *ibid.*

<sup>137</sup> For more detailed discussion of these findings in relation to the Focus 3 case studies, see the analysis of Domain 1 research question 7 in the Domain 1 Task Force Report, op. cit. See also the Domain 3 Task Force Report, op. cit., especially the “Findings” and “The Domain 3 Research Questions” sections.

<sup>138</sup> See Appendix 4.

Furthermore, each case study attempted to provide some input to the various cross-domains to supplement the research of the cross-domains. Representatives of each Focus Task Force were assigned as liaison to each of the cross-domains to ensure the free flow of relevant information among the researchers. The cross-domains were also invited to submit specific questions to be included in the case study questionnaire to learn about practices relevant to their area of study. These included questions 19 and 20 for the Policy Cross Domain<sup>139</sup> and questions 21 and 22 for the Description Cross-domain.<sup>140</sup> As a result, the Description Cross-domain received information about the existence of any in-house metadata schemas or about the adoption of existing metadata standards for use in records' control procedures by any of the case study subjects. Likewise, the Policy Cross-domain received information about legal or ethical issues underpinning current records creation practices along with pointers to other potential areas where standards were required and/or were being adopted.

Workflows were modeled for some case studies that particularly lent themselves to this form of study. The same *Integration Definition for Function Modeling (IDEF0)* methodology employed by the Modeling Cross-domain was adopted for the case studies.<sup>141</sup> This facilitated the comparison of actual records management practices in the field with the Project's Chain of Preservation model, an "ideal" records lifecycle model that integrates records creation, maintenance and preservation.<sup>142</sup> Finally, the terminology of all case studies was reviewed by the Terminology Cross-domain and selected terms were added to the dictionary section of the Project's Terminology Database to illustrate variations in the definitions of words among the many disciplines studied.

## References

To explore any individual case study in more details, please consult the information available on the InterPARES 2 Web site or on the DVD that accompanies this book. In many, although not all instances, the following documents have been prepared for each case study:

- case study proposal—includes a description of the case study subject, the rationale for choosing that case study subject, the research methodologies to be used, a description of the research team and their roles, and a timeline;
- areas that should be covered—contains a standardized set of information about the case study extracted from the whole of the documentation related to the case study; used to support the characterization of the study;
- case study characterization—contains information about: (1) the creator(s) of the digital entities under study, which allowed researchers to identify who produced the digital entities and for what reasons and (2) the creator's administrative and managerial framework and about the digital entities under study, which allowed researchers to characterize the entities and the types of activities related to their creation and management;

---

<sup>139</sup> Question 19: Have you had to make rules, or adopt standards to help you in your work? Do you find you have to update them regularly? and Question 20: Do any legal or ethical issues arise from your electronic work?

<sup>140</sup> Question 21: Did you create or adopt a standard list of information which you try to records about each file, or work? and Question 22: Where did you get it? Do you know if others use the same one?

<sup>141</sup> For more information regarding the IDEF0 modeling methodology, see the Modeling Cross-domain Task Force Report. Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_book\\_part\\_5\\_modeling\\_task\\_force.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_book_part_5_modeling_task_force.pdf).

<sup>142</sup> Ibid.



- case study overview—contains information extracted from the case study that is directly relevant to the work of all three Domains; used to facilitate analysis of case study data within each Domain;
- diplomatic analysis—contains an analysis of each type of digital object identified in the case study in relation to the five necessary characteristics of a record to determine if each object could be considered a record;
- Domain 1 analysis—contains information extracted from the case study to address the Domain 1 records creation and maintenance research questions;
- Domain 3 analysis—contains information extracted from the case study to address the Domain 1 records appraisal and preservation research questions; and
- final report.<sup>143</sup>

To locate conference presentations and articles addressing specific case studies, or Focus-specific analysis of case studies, please consult the list of participants, listed by case study, in Appendix 3. The contributions of the various investigators can be located in the extensive database where the Dissemination Activities of Project members are recorded. Again, this is available both on the InterPARES 2 Web site<sup>144</sup> and on the DVD that accompanies this book.

For example, the principal investigator for case study 12 (Antarctic Treaty Searchable Database) is Paul Berkman. A search of the Dissemination Activities database using the keyword “Berkman” will locate citations for two conference papers written by Berkman and citations for two refereed articles co-authored by Berkman, one of which is:

Berkman, Paul Arthur, George James Morgan III, Reagan Moore and Babak Hamidzadeh. “Automated Granularity to Integrate Digital Records: The “Antarctic Treaty Searchable Database” Case Study,” *Data Science Journal* 5 (June 2006): 84–89.

[http://www.jstage.jst.go.jp/article/dsj/5/0/84/\\_pdf](http://www.jstage.jst.go.jp/article/dsj/5/0/84/_pdf)

## Bibliographies

The bibliographies developed for each case study are available in their respective final reports. In addition, these bibliographies were merged with more extensive bibliographies (including, in some cases, annotated references) developed by the domains and cross-domains in their research. This integrated bibliography is available on the InterPARES 2 Web site<sup>145</sup> and on the DVD that accompanies this book.

---

<sup>143</sup> See Appendix 23 for a comprehensive listing of all major case and general studies-related documents organized by resource category. Available at [http://www.interpares.org/display\\_file.cfm?doc=ip2\\_book\\_appendix\\_23.pdf](http://www.interpares.org/display_file.cfm?doc=ip2_book_appendix_23.pdf).

<sup>144</sup> See [http://www.interpares.org/ip2/ip2\\_dissemination.cfm?proj=ip2](http://www.interpares.org/ip2/ip2_dissemination.cfm?proj=ip2).

<sup>145</sup> See [http://www.interpares.org/ip2/ip2\\_documents.cfm?cat=biblio](http://www.interpares.org/ip2/ip2_documents.cfm?cat=biblio).

## Appendix 3

### InterPARES 2 Case Studies and General Studies: Researchers, Focus and Final Status

#### Case Studies

##### **CS01 Arbo Cyber, théâtre (?)**

Focus: 1 (Arts)  
Status: Completed  
Principal investigator: Martine Cardin (archivist), Université Laval<sup>1</sup>  
Graduate research assistants: Philippe Perron, Université Laval  
Carolyn Petrie, The University of British Columbia

##### **CS02 Performance Artist Stelarc**

Focus: 1 (Arts)  
Status: Completed  
Principal investigator: Henry Daniel (dancer), Simon Fraser University  
Study team member: Luciana Duranti (archivist), The University of British Columbia  
Graduate research assistants: Peggy Heger, The University of British Columbia  
Cara Payne, The University of British Columbia

##### **CS03 *Horizon Zero/Zero Horizon* Online Magazine and Media Database**

Focus: 1 (Arts)  
Status: Completed  
Principal investigator: Brent Lee (musician), University of Windsor  
Study team member: Justine Bizzocchi (technologist), Banff Centre

##### **CS04 Persistent Archives Based on Data Grids**

Please see General Study 01.

##### **CS05 Archives of Ontario Web Exhibits**

Focus: 3 (Government)  
Status: Completed  
Principal investigator: Jim Suderman (archivist), Archives of Ontario  
Study team member: Marta Braun (art historian), Ryerson University  
Barbara Craig (archivist), University of Toronto  
Michael Murphy (technologist), Ryerson University  
Graduate research assistants: Deidre Brocklehurst, The University of British Columbia  
Terra Dickson, The University of British Columbia  
Peggy Heger, The University of British Columbia  
Brenda McPhail, University of Toronto

##### **CS06 Cybercartographic Atlas of Antarctica**

Focus: 2 (Science)  
Status: Completed  
Principal investigator: Fraser Taylor (geographer), Carleton University

---

<sup>1</sup> Institutional affiliations changed over the course of the Project. This appendix generally lists the researcher affiliations at the time the specific study was conducted.

Study team member: Yvette Hackett (archivist), Library and Archives Canada  
Graduate research assistant: Tracey Lauriault (case study lead), Carleton University

**CS08 Mars Global Surveyor Data Records in the Planetary Data System**

Focus: 2 (Science)  
Status: Completed  
Principal investigator: William Underwood (technologist), Georgia Tech Research Institute

**CS09 Digital Moving Images: Inputs, Processes and Outputs (a case study in four parts)**

Focus: 1 (Arts)  
Status: Completed  
Principal investigator: James Turner (information specialist), Université de Montréal  
Study team members: Marta Braun (art historian), Ryerson University  
Mary Ide (archivist), WGBH Boston  
Randal Luckow (archivist), Turner Broadcasting System  
Michael Murphy (technologist), Ryerson University  
Andrew Rodger (archivist), Library and Archives Canada

**CS09(1) Altair4 di Roma. A Multimedia Archaeological Project: The House of Julius Polybius**

Principal investigator: Isabella Orefice (archivist), Associazione Nazionale Archivistica Italiana

**CS09(2) National Film Board of Canada**

Principal investigator: Andrew Rodger (archivist), Library and Archives Canada

**CS09(3) Commercial Film Studio**

Co-Principal investigators: Randal Luckow (archivist), Turner Broadcasting System  
James Turner (information specialist), Université de Montréal

**CS09(4) WGBH Boston**

Principal investigator: Mary Ide (archivist), WGBH Boston

**CS10 The Danube Exodus: Interactive Multimedia Piece**

Focus: 1 (Arts)  
Status: Completed  
Principal investigator: Sally Hubbard (archivist), Getty Institute

**CS11 Nova Scotia Business Registry, Service Nova Scotia and Municipal Relations**

Focus: 3 (Government)  
Status: Retired  
Principal investigator: Margaret Campbell (archivist), Archives of Nova Scotia

**CS12 Antarctic Treaty Searchable Database**

Focus: 3 (Government)  
Status: Completed  
Principal investigator: Paul Berkman (environmental scientist), Ohio State University/EvREsearch, Ltd.  
Study team members: Babak Hamidzadeh (technologist), Library of Congress

Richard Marciano (technologist), San Diego  
 Supercomputer Center  
 Reagan Moore (technologist), San Diego Supercomputer  
 Center  
 George Morgan (technologist), EvREsearch, Ltd.  
 Jim Suderman (archivist), Archives of Ontario

**CS13 *Obsessed Again...***

Focus: 1 (Arts)  
 Status: Completed  
 Co-principal investigators: Keith Hamel (musician), The University of British  
 Columbia  
 Jesse Read (musician), The University of British Columbia  
 Study team member: Luciana Duranti (archivist), The University of British  
 Columbia  
 Graduate research assistant: J. Scott Amort, The University of British Columbia

**CS14 *Archaeological Records in a GIS: Research in the American Southwest***

Focus: 2 (Science)  
 Status: Completed  
 Principal investigator: Richard Pearce-Moses (archivist), Arizona State Library  
 Graduate research assistants: Erin O'Meara (case study lead), The University of British  
 Columbia  
 Randy Preston, The University of British Columbia

**CS15 *Waking Dream***

Focus: 1 (Arts)  
 Status: Completed  
 Principal investigator: Sidney Fels (technologist), The University of British  
 Columbia  
 Study team member: Luciana Duranti (archivist), The University of British  
 Columbia  
 Graduate research assistant: Seth Dalby, The University of British Columbia

**CS16 *Model for Description and Preservation of Documents Created Unstable and Variable Artistic Techniques***

Focus: 3 (Government)  
 Status: Completed

**CS17 *New York State Department of Motor Vehicles On-line Services System***

Focus: 3 (Government)  
 Status: Completed  
 Co-principal investigators: Phil Eppard (archivist), State University of New York,  
 Albany  
 Terry Maxwell (archivist), State University of New York,  
 Albany  
 Mark Wolfe (archivist), State University of New York,  
 Albany  
 Graduate research assistants: Joshua Hauck-Whealton, State University of New York,  
 Albany

Richard Hoppenstedt, State University of New York,  
Albany  
Rachel McMullin, State University of New York, Albany  
Peter Runge, State University of New York, Albany  
Reginald White, State University of New York, Albany  
Jessica Zacher, State University of New York, Albany

**CS18 Computerization of Alsace-Moselle's Land Registry**

Focus: 3 (Government)  
Status: Completed  
Principal investigator: Jean-François Blanchette, The University of British  
Columbia  
Graduate research assistant: Geneviève Shepherd (case study lead), The University of  
British Columbia

**CS19 Preservation and Authentication of Electronic Engineering and Manufacturing  
Records**

Focus: 2 (Science)  
Status: Completed  
Principal investigator: Kenneth Hawkins (archivist), National Archives and  
Records Administration

**CS20 Revenue On-Line Service (ROS)**

Focus: 3 (Government)  
Status: Completed  
Principal investigator: John McDonough (archivist), National Archives of Ireland  
Study team members: Ken Hannigan (archivist), National Archives of Ireland  
Tom Quinlan (archivist), National Archives of Ireland

**CS21 Electronic Filing System (EFS) of the Supreme Court of Singapore**

Focus: 3 (Government)  
Status: Completed  
Principal investigator: Elaine Goh (archivist), National Archives of Singapore

**CS22 Electronic Café International (ECI)**

Focus: 1 (Arts)  
Status: Retired (some reports available)  
Principal investigator: Howard Besser (moving images specialist), New York  
University  
Graduate research assistants: Nadine Hafner, The University of British Columbia  
Janine Johnston, The University of British Columbia  
Tracey Krause, The University of British Columbia  
Keum Hee Yu, The University of British Columbia

**CS23 UK Knowledge Network**

Focus: 3 (Government)  
Status: Retired (some reports available)

**CS24 City of Vancouver Geographic Information System (VanMap)**

Focus: 3 (Government)  
Status: Completed

Principal investigator: Evelyn Peters McLellan (archivist), City of Vancouver  
Study team members: Sue Bigelow (conservator), City of Vancouver Archives  
Glenn Dingwall (archivist), City of Vancouver Archives  
Luciana Duranti (archivist), The University of British Columbia  
Richard Marciano (technologist), San Diego Supercomputer Center  
Jonathan Mark (technologist), City of Vancouver, Information Technology Department  
Andrew Power (records manager), City of Vancouver  
Peter Van Garderen (archivist), President, Artefactual Systems, Inc.  
Reuben Ware (archivist), City of Vancouver Archives  
Liz Wright (archivist), City of Vancouver Archives  
Graduate research assistants: Eleanor Kleiber, The University of British Columbia  
Catherine Miller, The University of British Columbia

**CS25 Legacoop of Bologna Web Site**

Focus: 3 (Government)  
Status: Completed  
Principal investigator: Isabella Orefice (archivist), Associazione Nazionale Archivistica Italiana

**CS26 Microvariability & Oscillations of Stars (MOST) Satellite Mission: Preservation of Space Telescope Data**

Focus: 2 (Science)  
Status: Completed  
Principal investigator: Reagan Moore (technologist), San Diego Supercomputer Center  
Study team member: Luciana Duranti (archivist), The University of British Columbia  
Graduate research assistant: Bart Ballaux (case study lead), The University of British Columbia

**General Studies**

**GS01 Persistent Archives Based on Data Grids**

Focus/Area: Preservation  
Status: Completed  
Principal investigator: Reagan W. Moore (technologist), San Diego Supercomputer Centre

**GS02 Survey and Analysis of Scientific Encoding Languages for Non-Textual Records**

Focus/Area: Preservation  
Status: Retired (some reports available)  
Principal investigator: William Underwood (technologist), Georgia Tech Research Institute

**GS03 Preserving Interactive Digital Music - the MUSTICA Research Initiative**

Focus/Area: 1 (Arts)  
Status: Completed  
Co-principal investigators: Jean-François Blanchette, The University of British Columbia  
John Roeder (musician), The University of British Columbia  
Graduate research assistants: Jill Teasley (case study lead), The University of British Columbia  
Jennifer Douglas, The University of British Columbia  
Carolyn Petrie, The University of British Columbia

**GS04 Survey of Recordkeeping Practices of Composers**

Focus/Area: 1 (Arts)  
Status: Completed  
Principal investigator: Michael Longton (musician), University of Victoria  
Graduate research assistant: Vincent Schillaci-Ventura, The University of British Columbia

**GS05 An Examination of the Processes to Preserve and Manage Electronic Records: Round Three at The National Archives of Australia and WGBH**

Focus/Area: Preservation  
Status: Retired (some reports available)  
Co-principal investigators: Shelby Sanett (archivist), U.S. National Archives and Records Administration  
Michèle V. Cloonan (information specialist), Simmons College

**GS06 A Bayesian Belief Network: Supporting the Assessment of the Degree of Belief that a Recordkeeping System Maintains Authentic Digital Records**

Focus/Area: 2 (Science)/Appraisal  
Status: Completed  
Co-principal investigators: William Underwood (technologist), Georgia Tech Research Institute  
Sheila Isbell (technologist), Georgia Tech Research Institute

**GS07 Recordkeeping Practices of Photographers using Digital Technology**

Focus/Area: 1 (Arts), 2 (Science), 3 (Government)  
Status: Completed  
Principal investigator: Marta Braun (art historian), Ryerson University  
Study team member: Yvette Hackett (archivist), Library and Archives Canada  
Graduate research assistant: Jessica Bushey, The University of British Columbia

**GS08 Survey of Government Web Site Interactivity**

Focus/Area: 3 (Government)  
Status: Completed  
Principal investigator: Jim Suderman (archivist), Archives of Ontario  
Study team members: Yvette Hackett (archivist), Library and Archives Canada

Sue McKemmish (archivist), Monash University  
Mark Wolfe (archivist), State University of New York,  
Albany

**GS09 Digital Recordkeeping Practices of GIS Archaeologists Worldwide**

Focus/Area: 2 (Science)  
Status: Completed  
Principal investigator: Randy Preston (archivist), The University of British  
Columbia  
Study team member: Erin O'Meara (archivist), The University of British  
Columbia

**GS10 Preservation Practices of Scientific Data Portals**

Focus/Area: 2 (Science)  
Status: Completed  
Co-principal investigators: Tracey Lauriault (geographer), Carleton University  
Barbara Craig (archivist), University of Toronto  
Graduate research assistants: Heather Dean, The University of British Columbia  
Stephen Gage, The University of British Columbia  
Erin Hanlon, The University of British Columbia  
Christina Miller, The University of British Columbia  
Brian K. Trembath, The University of British Columbia  
Sherry Xie, The University of British Columbia

**GS11 Selecting Digital File Formats for Long-Term Preservation**

Focus/Area: Preservation  
Status: Completed  
Principal investigator: Evelyn Peters McLellan (archivist), Insurance Corporation  
of British Columbia  
Graduate research assistants: Tracey Krause, The University of British Columbia  
Yvonne Loiselle, The University of British Columbia

**GS12 Validation of the InterPARES 2 Project Chain of Preservation Model Using Case Study Data**

Focus/Area: Creation, maintenance, preservation  
Status: Completed  
Principal investigator: William Underwood (technologist), Georgia Tech Research  
Institute  
Study team members: Kevin Glick (archivist), Yale University  
Mark Wolfe (archivist), State University of New York,  
Albany



## Appendix 4

**Genesis of the Case Study Research Questions**

<b>Original Questions From Workshop #2 (21 June 2002, Wash., D.C.)</b>	<b>Re-worded Questions Distributed by Project Director (26 June 2002)</b>	<b>Additional Questions Added During Workshop #3 (17-21 Sept 2002, Los Angeles)</b>
	1. What activities of the creator are you investigating?	
	2. Which of these activities generate the digital entities that are the objects of your case study?	
1. What are the purposes of the information you record or create?	4. For what purpose(s) are the digital entities you are examining created?	
2. What information do you create to meet these purposes?		
3. What methods and paradigms inform your work?		
4. What forms do your information take?	3. What form do these digital entities take? (E.g. e-mail, CAD, database)	
5. What processes do you follow in creating information?	5. How are those digital entities generated?	
6. What are the key processes in creating the information?	6. From what precise process(es) or procedure(s), or part thereof, do the digital entities emerge?	
	7. To what other digital or non-digital entities are they connected? Is such connection documented or captured?	

7. How do you record and identify the information, the methods and the technologies you have followed?	8. What are the documentary and technological processes or procedures that the creator follows to identify, retrieve, and access the digital entities?	
8. How do you document the processes and procedures you use?	9. Are those processes and procedures documented? How? In what form?	
9. What are the key elements of the information you create?	10. What are the key elements, attributes, and digital components of the entities under examination? Should be asked of the interviewee (opinion question, subject to later analysis).	
10. What measures do you take to ensure the quality and reliability of the information you create or information sources that you use?	11. What measures does the creator take to ensure the quality, reliability and authenticity of the digital entities and their documentation? If no specific measure is taken, does the creator think that those qualities are to be assumed for its digital entities? ( <i>Note overlap with question 16.</i> )	
12. How are the changes made to your information and how are these recorded?	12. How are changes to the digital entities made and recorded?	
11. How do you use the information you create?	13. How does the creator use the digital entities under examination?	
13. How do others use the information you create?	14. Do external users have access to the digital entities in question? If so, how, and what kind of uses are made of the entities?	
14. Do others add to your information to create new information?		

<p>15. What do you conceive of as a record?</p>	<p>15. Among its digital entities, which ones does the creator consider to be records and why?</p>	
<p>16. What do you conceive of as an authentic record?</p>	<p>16. Does the creator think that the authenticity of his digital records is assured, and if so, why?</p>	
<p>17. How do you preserve this authentic record?</p>	<p>17. Does the creator keep the digital entities that are currently being examined? That is, are these digital entities part of a record keeping system? If so, what are its features?</p>	
<p>18. How do you preserve this through technological change?</p>	<p>18. How does the creator maintain its digital entities through technological change?</p>	
		<p>19. Have you had to make rules, or adopt standards to help you in your work? Do you find you have to update them regularly?</p>
		<p>20. Do any legal or ethical issues arise from your electronic work?</p>
		<p>21. Did you create or adopt a standard list of information which you try to record about each file, or work?</p>
		<p>22. Where did you get it? Do you know if others use the same one?</p>

## Appendix 5

### **23 Case Study Questions that the researchers should be able to answer at the completion of their investigation March 18, 2003**

1. What activities of the creator have you investigated?
2. Which of these activities generate the digital entities that are the objects of your case study?
3. For what purpose(s) are the digital entities you have examined created?
4. What form do these digital entities take? (e.g. e-mail, CAD, database)
  - 4a. What are the key formal elements, attributes, and behaviour (if any) of the digital entities?
  - 4b. What are the digital components of which they consist and their specifications?
  - 4c. What is the relationship between the intellectual aspects and the technical components?
  - 4d. How are the digital entities identified (e.g., is there a [persistent] unique identifier)?
  - 4e. In the organization of the digital entities, what kind of aggregation levels exist, if any?
  - 4f. What determines the way in which the digital entities are organized?
5. How are those digital entities created?
  - 5a. What is the nature of the system(s) with which they are created? (e.g. functionality, software, hardware, peripherals etc.)
  - 5b. Does the system manage the complete range of digital entities created in the identified activity or activities for the organization (or part of it) in which they operate?
6. From what precise process(es) or procedure(s), or part thereof, do the digital entities result?
7. To what other digital or non-digital entities are they connected in either a conceptual or a technical way? Is such connection documented or captured?
8. What are the documentary and technological processes or procedures that the creator follows to identify, retrieve, and access the digital entities?
9. Are those processes and procedures documented? How? In what form?
10. What measures does the creator take to ensure the quality, reliability and authenticity of the digital entities and their documentation?
11. Does the creator think that the authenticity of his digital entities is assured, and if so, why?
12. How does the creator use the digital entities under examination?
13. How are changes to the digital entities made and recorded?

14. Do external users have access to the digital entities in question? If so, how, and what kind of uses do they make of the entities?
15. Are there specific job competencies (or responsibilities) with respect to the creation, maintenance, and/or use of the digital entities? If yes, what are they?
16. Are the access rights (to objects and/or systems) connected to the job competence of the responsible person? If yes, what are they?
17. Among its digital entities, which ones does the creator consider to be records and why?
18. Does the creator keep the digital entities that are currently being examined? That is, are these digital entities part of a record keeping system? If so, what are its features?
  - 18a. Do the recordkeeping system(s) (or processes) routinely capture all digital entities within the scope of the activity it covers?
  - 18b. From what applications do the recordkeeping system(s) inherit or capture the digital entities and the related metadata (e.g. email, tracking systems, workflow systems, office systems, databases, etc.)?
  - 18c. Are the digital entities organized in a way that reflects the creation processes? What is the schema, if any, for organizing the digital entities?
  - 18d. Does the recordkeeping system provide ready access to all relevant digital entities and related metadata?
  - 18e. Does the recordkeeping system document all actions/transactions that take place in the system re: the digital entities? If so, what are the metadata captured?
19. How does the creator maintain its digital entities through technological change?
  - 19a. What preservation strategies and/or methods are implemented and how?
  - 19b. Are these strategies or methods determined by the type of digital entities (in a technical sense) or by other criteria? If the latter, what criteria?
20. To what extent do policies, procedures, and standards currently control records creation, maintenance, preservation and use in the context of the creator's activity? Do these policies, procedures, and standards need to be modified or augmented?
21. What legal, moral (e.g. control over artistic expression) or ethical obligations, concerns or issues exist regarding the creation, maintenance, preservation and use of the records in the context of the creator's activity?
22. What descriptive or other metadata schema or standards are currently being used in the creation, maintenance, use and preservation of the recordkeeping system or environment being studied?
23. What is the source of these descriptive or other metadata schema or standards (institutional convention, professional body, international standard, individual practice, etc.?)

## Appendix 6

# Call for Case Study Proposals InterPARES 2 April 2002

## Overview

Beginning in 2002, researchers in the InterPARES 2 Project will undertake a set of case studies in an effort to address in part the research questions assigned to each domain task force and cross-domain research team. Each case study will focus on the records (or some portion of the records) and records management process of a specific creator. Within the context of a case study, it may also be possible to test the appraisal and preservation models developed in InterPARES 1 and/or develop a prototypical system for the preservation of the records in question.

In March 2002, many InterPARES researchers participated in an exercise involving the sketching of records creation processes related to an activity with which they were familiar. This exercise has proven fruitful in identifying potential case study subjects and specific issues of concern. A number of case studies are currently in development based on these activity models, with more to follow in the coming months.

At the International Team Meeting in June 2002, the chairs of the various working groups and cross-domain teams will review the case study proposals that have been developed, offer suggestions to make the conduct of each case study more efficient and effective, and allocate resources as needed. Thus, the goal of the review as envisaged will not be to approve or reject case study proposals, but rather to coordinate the efforts of researchers, allow for the streamlining and sharing of research tools and offer advice and support to researchers conducting case studies.

This call is designed to assist researchers in the development of case study proposals. It is expected that researchers interested in proposing and/or leading a case study will be in contact with the chairs of their working groups as well as the InterPARES administration as their case study is developing. The International Team will review case study proposals as they are ready; those wishing their proposal to be reviewed at the meeting in June should submit their proposals by June 10 2002 to allow for distribution prior to the meeting.

## InterPARES 1 Case Studies

The Authenticity Task Force (ATF) of the InterPARES 1 Project undertook more than thirty case studies (in four rounds) of a variety of electronic systems. The initial goal of the ATF was to identify the formal elements shared by all electronic records, the elements that allowed for their differentiation by type, and the elements that allowed a record's authenticity to be verified over time. (It was further hypothesized that answers to these questions would lead to conclusions regarding the possibility of migrating electronic records from one system to another without compromising their authenticity.) Such analysis of a record's formal elements is *diplomatic* analysis; diplomacy as a science encompasses a set of principles and terminology that have been used to analyze records since the 17th century. Given that electronic records in most cases serve the same administrative functions that paper records have in the past, InterPARES 1 researchers were looking for possible parallels between paper and electronic records, specifically in their formal elements. To this end, the ATF developed a template for analysis of electronic

records, enumerating and describing a large set of formal elements that might be potentially found in an electronic record. Once the template had been finalized, a number of case studies of electronic systems were undertaken to empirically ascertain whether or not these elements manifested themselves in records contained within actual electronic systems.

To gather information about actual electronic systems and records, the ATF developed two tools: a Case Study Interview Protocol (CSIP) and a Template Element Data Gathering Instrument (TEDGI). The CSIP was a set of questions posed by a researcher to individuals familiar with the workings of the electronic system being studied. Based on the interviewees' responses to the CSIP questions, the researcher noted (in the TEDGI) the presence or absence of formal record elements. This two-step process was deemed necessary in light of the interviewees' unfamiliarity with diplomatic terminology. The ATF reported that the case study analysis did indeed confirm the presence or absence of certain formal elements and further indicated potential weaknesses in many of the records management system studied. It was also observed that many of the formal elements upon which a record's authenticity was presumed, which in the past had been visibly manifested on the face of the record, were supplanted in many electronic systems by procedural and technological controls.

In general the ATF concluded that the case study approach was very useful in addressing their assigned research questions, but had certain limitations and could be improved upon. The ATF recommended that in future case study research that 1) the record-keeping system be studied as a whole (including its paper elements), 2) the study begin with a careful analysis of the business procedures of the record creators in question in order to identify the actions in which records participate, 3) formulate a questionnaire (or revise and shorten the CSIP) with terminology familiar to the interviewees and 4) delve further into the technological context of the records under study. (Footnote 1: see the Authenticity Task Force Report on the InterPARES Web site <http://www.interpares.org/reports.htm>.)

It is hoped that in InterPARES 2 we can implement these methodological findings of InterPARES 1. Please keep these findings in mind when developing case studies and case study tools.

## **The Case Study Proposal**

Case study proposals will include a description of the case study subject, the rationale for choosing that case study subject, the research methodologies to be employed, a description of the research team and their roles, and a timeline.

- Description of the case study subject: briefly describe the case study subject in terms of their business mandate and business processes. If the records of interest to us are only a portion of the creator's records, contextualize the records within the creator's records as a whole.
- Rationale: why are this creator and these records of interest to InterPARES 2? Which research questions will be addressed in this case study?
- Research methodologies: how will the case study be conducted? What sorts of data will be gathered? How will the data be represented? What tools will be used in gathering the data? Below are listed a handful of possible methodologies and related issues:
  - a. Interviewing records creators: much information can be gathered by interviewing the creators of records. The questionnaire used in InterPARES 1 and the research questions for InterPARES 2 form a starting point for developing a set of questions; further questions may be appropriate for specific creators.

- b. Modeling records creation processes: collaboratively creating a model of the creator's business process may illuminate issues unanticipated in a questionnaire. Where applicable, the process models sketched in March, 2002 may serve as an outline. Also, a formal work-flow model may already exist for a given organization.
  - c. Application of IP1 Appraisal model: as much of the information that needs to be gathered is similar to the information typically gathered by archivists conducting appraisals, it may be a useful exercise to use the InterPARES 1 appraisal model to guide the study.
  - d. Replicating systems for hands-on study of records: though difficult with large systems, creators working on desk-top computers may be willing to make copies of all or a portion of their records which can be installed on an InterPARES computer for subsequent study. It may also be possible to replicate a larger system (such as a government Web site) and create a set of records for testing. This methodology would greatly facilitate diplomatic analysis of the records in questions, as they could be examined in their native environment.
  - e. Testing preservation strategies: it may be advantageous to transfer a sampling of records to an InterPARES computer to test various preservation strategies, or to walk through the InterPARES 1 preservation model. It may also be beneficial to prototype a preservation system for the records in question.
- Case study team: each case study will be conducted by a team of researchers. This team will comprise a lead investigator, other interested researchers, and research assistants as necessary.
  - Timeline: a timeline will outline the sequence and timing of activities in the conduct of the case study, as well as the researchers responsible for those activities (individually or collectively). Keep in mind that the team will have a chance to meet at the workshops in September 2002 and February 2003; further meetings may be conducted electronically. If further face-to-face meetings are necessary and require travel, include an estimate of the costs involved. In developing the timeline, consider the time and effort necessary to make arrangements with the case study subject, to develop a questionnaire, model or prototype, to seek human subjects approval, to represent and analyze the data, to follow up with the case study subject (pursuant questions, clarification of responses, validation of a process model, etc.) and to prepare a case study report.

## **Reporting Procedure**

Given the flexibility in the conduct of case studies, it is important that the reporting procedure be harmonized among case studies so as to facilitate comparison. The case study report will begin with information included in the proposal (description of subject, rationale, methodologies used), noting any modifications made in actually conducting the case study. The report will also include the case study data or summary of same, the observations and comments of the case study team and an evaluation of the effectiveness of the methodologies employed. Ultimately, case study reports will be reviewed by the appropriate task forces and will serve as the basis of much of their deliberations.



## Appendix 7

# Diplomatic Analysis Template

## Diplomatic Analysis

### CS[##] [Title of Case Study] Case Study

#### **INTRODUCTION**

Paragraph 1<sup>1</sup> - The InterPARES case study # ...

Paragraph 2<sup>2</sup> - The ... project/program/system/database ...

Paragraph 3<sup>3</sup> - The following text presents the results of the diplomatic analysis on the digital entity identified in the case study report. The purpose of the diplomatic analysis is to assess the status of the identified digital entity as record, and based on the analysis, Domain 3 of InterPARES 2 could propose applicable preservation strategies. The digital entity identified in the case study report is ... This diplomatic analysis therefore centers on the identification of the [digital entity] as record.

#### **IDENTIFICATION OF RECORD(S)**

The current version of the InterPARES glossary definition of record defines a record<sup>4</sup> as “a document made or received in the course of a practical activity as an instrument or a by-product of such activity, and set aside for action or reference.”<sup>5</sup> This definition implies that, to be considered as a record, a digital entity must comprise five indispensable elements: fixed content and form, embedded action, archival bond, persons and contexts, the establishment of which is based on diplomatic analysis, archival science and findings from InterPARES. The application of the definition on the [digital entity] is therefore analyzed accordingly:

- 1. TO BE IDENTIFIED AS A RECORD, THE DIGITAL ENTITY MUST POSSESS FIXED CONTENT AND FORM,<sup>6</sup> AND BE AFFIXED TO A STABLE MEDIUM (OR PHYSICAL CARRIER).**
  - The content of the [digital entity] is fixed or not and why
  - The documentary form<sup>7</sup> of the [digital entity] is fixed or not and why

---

<sup>1</sup> Introduction to the case study.

<sup>2</sup> Introduction to the case study subject.

<sup>3</sup> Introduction to the diplomatic analysis, including the indication of the subject of the diplomatic analysis, i.e., the digital entity in the report that needs to be analyzed.

<sup>4</sup> Current version of the definition of record as assessing criterion, which starts the process of analyzing the 5 components.

<sup>5</sup> Glossary definitions, in Terminology Database, accessible through the InterPARES website research restricted area.

<sup>6</sup> The InterPARES1 Authenticity Task Force has defined fixed form as the following: 1) binary content of the record, including indicators of documentary form, must be stored in a manner that ensures it remains complete and unaltered, and 2) technology must be maintained and procedures defined and enforced to ensure that the content is presented or rendered with the same documentary form it had when set aside. (See ATF Research Methodology Statement, available at: [http://www.interpares.org/documents/interpares\\_ResearchMethodologyStatement.pdf](http://www.interpares.org/documents/interpares_ResearchMethodologyStatement.pdf)).

<sup>7</sup> Definition of documentary form from Glossary Definitions, Terminology Database, InterPARES Web site: The rules of representation according to which the content of a record, its administrative and documentary context, and its authority are communicated. Documentary form possesses both extrinsic and intrinsic elements.

- The digital entity is fixed to a stable medium or not and why

**2. A RECORD MUST ALSO PARTICIPATE IN AN ACTION, DEFINED AS THE CONSCIOUS EXERCISE OF WILL BY THE AUTHOR OR BY AN EXTERNAL PERSON, AIMED TO CREATE, MAINTAIN, MODIFY OR EXTINGUISH SITUATIONS. A RECORD RESULTS AS A NATURAL BY-PRODUCT OF THE ACTION.**

Statement of the name of action that generated the [digital entity] and how it participates in the action.

**3. THIRD, A RECORD MUST POSSESS AN ARCHIVAL BOND, WHICH IS THE RELATIONSHIP THAT LINKS EACH RECORD TO THE PREVIOUS AND SUBSEQUENT RECORD OF THE SAME ACTION AND, INCREMENTALLY, TO ALL THE RECORDS WHICH PARTICIPATE IN THE SAME ACTIVITY. THE ARCHIVAL BOND IS ORIGINARY (I.E., IT COMES INTO EXISTENCE WHEN THE RECORD IS MADE OR RECEIVED AND SET ASIDE), NECESSARY (I.E., IT EXISTS FOR EVERY RECORD), AND DETERMINED (I.E., IT IS CHARACTERISED BY THE PURPOSE OF THE RECORD).**

The [digital entity] possesses (or does not possess) archival bond and why.

**4. FOURTH, RECORD CREATION MUST INVOLVE AT LEAST THREE PERSONS, WHETHER OR NOT THEY EXPLICITLY APPEAR IN THE RECORD ITSELF. THESE PERSONS ARE THE AUTHOR, ADDRESSEE AND WRITER; IN THE ELECTRONIC ENVIRONMENT, ONE MUST ALSO TAKE INTO ACCOUNT TWO ADDITIONAL NECESSARY PERSONS: THE CREATOR AND THE ORIGINATOR.**

- The record's author is the physical or juridical person having the authority and capacity to issue the record or in whose name or by whose command the record has been issued.

Name of the author and why.

- The writer is the physical or juridical person having the authority and capacity to articulate the content of the record.

Name of the writer and why.

- The addressee is the physical or juridical person(s) to whom the record is directed or for whom the record is intended.

Name of the address and why.

- The creator is the person in whose fonds the record exists.

Name of the creator and why.

- The originator is the person to whom the Internet account issuing or the server holding the record belongs.

Name of the originator and why.

**5. FINALLY, A RECORD MUST POSSESS AN IDENTIFIABLE CONTEXT, DEFINED AS THE FRAMEWORK IN WHICH THE ACTION IN WHICH THE RECORD PARTICIPATES TAKES PLACE. THE TYPES OF CONTEXT INCLUDE JURIDICAL-ADMINISTRATIVE, PROVENANCIAL, PROCEDURAL, DOCUMENTARY, AND TECHNOLOGICAL.**

- The juridical-administrative context is the legal and organizational system in which the creating body belongs.

Juridical context:

Administrative context:

- The provenancial context refers to the creating body, its mandate, structure and functions.

Provenancial context:

- The procedural context comprises of the business procedure in the course of which the record is created.

Procedures:

Diplomatic analysis of Procedural Phases in the Creation of the [digital entity]

- a) **Initiative:** The introductory phase of any procedure is “constituted by those acts, written and/ or oral, which start the mechanism of the procedure.”<sup>8</sup>
- b) **Inquiry:** This preliminary phase “is constituted by the collection of the elements necessary to evaluate the situation.”<sup>9</sup>
- c) **Consultation:** This phase is “constituted by the collection of opinions and advice after all the relevant data has been assembled.”<sup>10</sup>
- d) **Deliberation:** This phase is “constituted by the final decision-making.”<sup>11</sup>
- e) **Deliberation control:** This phase is “constituted by the control exercised by a physical or juridical person different from the author of the document embodying the transaction, on the substance of the deliberation and / or on its forms.”<sup>12</sup>
- f) **Execution:** “The documents created in this phase are the originals of those embodying the transactions.”<sup>13</sup> In other words, the execution phase results in the issuing of the first record capable of producing the consequences intended by its author.

---

<sup>8</sup> Luciana Duranti, *Diplomatics: New Uses for an Old Science* (Lanham, Maryland and London: The Scarecrow Press in association with the Society of American Archivists and the Association of Canadian Archivists, 1998), 115.

<sup>9</sup> Ibid.

<sup>10</sup> Ibid.

<sup>11</sup> Ibid.

<sup>12</sup> Ibid.

<sup>13</sup> Ibid., 116.

- The documentary context is defined as the archival fonds to which a record belongs and its internal structure.

Documentary context:

- The technological context is defined as the characteristics of the technological components of an electronic computing system in which records are created.

Technological context:

## **CONCLUSIONS**

Overall conclusions of the status of the digital entity under examination:

If it is not a record:

Summary of the digital entity as a publication:

Summary of the digital entity as a potential record:

## Appendix 8

### InterPARES 2 Reporting Framework December 2003

Case study reports should contain the following sections:

- A. Overview
  - B. Statement of Methodology
  - C. Description of Context
  - D. Narrative answers to the core research questions
  - E. Narrative answers to applicable domain and cross-domain research questions
  - F. Bibliography of relevant material, including articles about the methods and works of the subject(s)
  - G. Glossary of terms
  - H. Preliminary model
- 

- A. Give a brief overview of the subject and the nature of the case study.
- B. Briefly describe the data gathering and analysis methodology employed to achieve the research objective of answering the core twenty-three questions posed to researchers, and any applicable domain and cross-domain questions.

For most or all artistic focus case studies the primary information-gathering tool will be an interview or interviews that draws upon the lay restating and amplification of these questions ([ip2\\_possible%20cs\\_interview\\_questions.pdf](#)). Note that while recording and creating transcripts of interviews is highly recommended, transcripts should not be included in the case study report.

Other research methods might include document review, ethnographic analysis or participant observation, diplomatic analysis, Bayesian analysis, content analysis (of interview transcripts), etc.

- C. Describe the context of record or digital entity creation and management. InterPARES 1 defined five contexts.
  - a. Provenancial: the creating body, its mandate, structure, and functions (indicators include organizational charts, annual reports, the classification scheme, etc.).
  - b. Juridical-administrative: the legal and organizational system in which the creating body (indicated by laws, regulations, etc.) belongs.
  - c. Procedural: the business procedure in the course of which the digital entity is created (indicators include workflow rules, codes of administrative procedure, classification schemes, etc.).

- d. Documentary: the fonds to which the digital entity belongs and its internal structure (indicators include classification schemes, record inventories, indexes, registers, etc.)  
*Note: In some organizations, business procedures are integrated with documentary procedures.*
- e. Technological: the characteristics of the digital environment in which the record is created and maintained.

These specifications assume a legal and bureaucratic structure that might not be appropriate to all case studies. However, each of these contexts should be addressed at least to the extent that their inapplicability is documented, and the environment in which digital entities are created and managed, or the framework of action in which they participate, should be described.

- D. Narrative answers to the twenty-three core questions to researchers listed in ip2\_23\_questions.pdf.
- E. Narrative answers to any relevant domain and cross-domain research questions.
- F. The bibliography should draw from those bibliographies already created where appropriate, and additionally list any articles or monographs that are of particular relevance to the specific case study.
- G. The glossary should list and define the key terms used in the case study, both for purposes of possible inclusion in the IP2 glossary, and to allow definitions to be compared with those that already exist within the IP2 glossary.
- H. A preliminary case study model should be made following the guidelines provided by the modeling cross-domain group.