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

APPENDIX 21

PRESERVER GUIDELINES
Preserving Digital Records:
Guidelines for Organizations

[including Appendices 21a, 21b and 21c]

by

Yvette Hackett
Library and Archives Canada

Status: Final (public)

Version: Electronic

Publication Date: 2008

Project Unit: Domain 3 Task Force

URL: http://www.interpares.org/display_file.cfm?doc=ip2_book_appendix_21.pdf

How to Cite: Yvette Hackett, Domain 3 Task Force, "Appendix 21: Preserver Guidelines – Preserving Digital Records: Guidelines for Organizations," [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_appendix_21.pdf>

PRESERVER GUIDELINES
Preserving Digital Records: Guidelines for Organizations[1]

Introduction

These guidelines have been developed to provide concrete advice to various groups that are responsible for the long-term preservation of digital records. They are not intended to be comprehensive but to highlight a number of areas that are particularly important to the preservation of authentic digital records and which experience has shown to be often overlooked in the rush to accept digital records into archival repositories.

As is widely recognized, digital records must be carefully managed throughout their entire existence to ensure that they are accessible and readable over time with their form, content and relationships intact to the extent necessary for their continuing trustworthiness as records. It is also widely recognized that management of digital records must proceed from a comprehensive understanding of all phases or stages of records' existence, from the time they are generated, through their maintenance by their creator, and during their appraisal, disposition and long-term preservation as authentic memorials of the actions and matters of which they are a part. From the perspective of long-term preservation, all the activities to manage records throughout their existence are linked, as in a chain, and interdependent. If a link in the chain fails, the chain cannot do its job. If certain activities and actions are not undertaken on records, their integrity (that is, their reliability and authenticity) and preservation are imperilled.

These guidelines focus on the preservation link in the chain of preservation and are organized according to the sequence of preservation activities presented in the InterPARES Chain of Preservation (COP) model,[2] which charts the many sequential steps in the creation, maintenance and preservation of authentic records. The alphanumeric number in parentheses following each section title in these Guidelines is a cross reference to the applicable preservation activity presented in the COP model.

The guidelines have been tailored to address the preservation needs of organizations or programs whose records must be retained and consulted for long periods and those of archival institutions that take on the responsibility for the long-term preservation of the records of others and for their continuing accessibility to the public they serve. In both these cases, human and financial resources as well as in-house technical expertise are frequently limited.

Institutions, organizations and programs with preservation responsibilities should also consult the *Framework of Principles for the Development of Policies, Strategies and Standards for the Long-term Preservation of Digital Records* (a.k.a., Policy Framework)[3] developed by the InterPARES 2 Policy Cross-domain, which complement these Guidelines. Many of the recommendations of these Guidelines may also be applicable to the preservation of digital objects other than records, such as documents, publications or data.

[1] These Guidelines have also been issued in an illustrated booklet form that is freely available at http://www.interpares.org/display_file.cfm?doc=ip2(pub)preserver_guidelines_booklet.pdf.
[2] Available at http://www.interpares.org/ip2/ip2_models.cfm.
[3] Available at http://www.interpares.org/public_documents/ip2(pub)policy_framework_document.pdf.

1. Manage Chain of Preservation

This aspect involves determining framework requirements, and designing, implementing and maintaining a chain of preservation framework. A *Chain of Preservation Framework* includes all the elements of policy, strategy, methodology and so on.

1.1. Establish scope and objectives

Preservers must define the scope and objectives of their digital preservation program. In the arts, for example, they may wish to preserve the recording of the performance(s) of a work, or they may choose to undertake the more complex preservation of the components of a work of art that support its reproduction or re-performance. In the sciences, preservers may wish to preserve only the final report of the results of an experiment, or hold raw data, normalized data and/or aggregated data to document the methodology used and the result obtained, as well as to ensure the availability of the data for future uses. Preservers should also consider who the eventual users of the archives will be. Technically sophisticated users generally require less assistance in accessing even technologically complex digital materials, while the general public might require extremely user-friendly access mechanisms and materials transformed into a few simple, but widely available, formats. The scope of the preservation program will help define which preservation strategies (see Section 4 and Appendix 21c, Section B) a preserver might need to support.

In defining the digital preservation program, preservers should build on previous efforts. To develop appropriate policies and strategies, preservers should consult the InterPARES 2 Policy Framework for guidance applicable at organizational, sectoral, national, international and supranational levels. For the functions of the preservation program, preservers should consult the ISO Open Archival Information System (OAIS) standard[4] and should follow the InterPARES 2 Chain of Preservation model for an adaptation of the OAIS standard specifically intended for digital records. Plans should also reflect the *Trustworthy Repositories Audit & Certification: Criteria and Checklist*, a revised and expanded version of the *Audit Checklist for Certifying Digital Repositories* originally developed by the NARA/RLG Digital Repository Task Force.[5]

1.2. Acquire resources

Digital preservation requires substantial resources in funding, technological capabilities and expertise. An organization responsible for digital preservation has several options, including: a) acquire new resources, b) reallocate existing resources and/or c) leverage other resources.

Regardless of the option(s) chosen, a fundamental requirement is that resources must be sustainable. One-time resources, such as grants, may be appropriate for specific finite tasks, such as establishing the preservation program or processing a given body of records, but a reliable source of sustained resources is a *sine qua non* for any preservation program.

Acquiring new financial resources will require a sound plan for the program and a matching communications plan to convince funding sources and stakeholders that preservers are likely to consult that the program should be funded. A viable strategy for a new program may be to start

[4] International Organization for Standardization, ISO 14721: 2003 - Space data and information transfer systems—Open archival information system—Reference model.

[5] See Online Computer Library Center, Center for Research Libraries (2007), "Trustworthy Repositories Audit & Certification: Criteria and Checklist," v. 1.0, February 2007. Available at http://www.crl.edu/PDF/trac.pdf.

small and plan on short-term successes to convince funding sources to incrementally increase resources for the program. An incremental strategy should evaluate whether funding sources are more likely to be influenced by short-term success in basic program objectives or in areas of more particular concern to the funding sources and stakeholders. For example, funders and stakeholders may be more swayed by demonstrations of technological capabilities than by a sound and comprehensive plan for appraising digital records.

For most organizations, reallocating resources to digital preservation is likely to entail painful choices. As with seeking new funds, an incremental approach may be best. Furthermore, ongoing adjustments can be made to the plan, based on the experience gained during each phase of implementation. If the digital preservation program is to be established in a larger institution, it would be helpful to address digital preservation as part of the overall strategic plan rather than as a special initiative.

Even when a preserver successfully acquires new resources or is able to reallocate existing resources to digital preservation, it is unlikely it will have sufficient resources to address all the challenges. Therefore, preservers should capitalize on opportunities for leveraging outside resources. There are a variety of paths for doing this. For example, rather than trying to hire technical experts on a permanent basis or training staff in all requisite technical knowledge and skills, preservers might engage outside experts on a consultative or task basis. They should not exclude options to contract for both basic and ad hoc services. On a basic level, preservers should evaluate the possibility of using a computer service provider rather than acquiring a dedicated preservation system. Ad hoc options include engaging specialized companies for tasks such as re-copying from obsolete digital media or converting rare formats. Another option is to participate—actively or passively—in open-source communities developing technologies needed for digital preservation (e.g., FEDORA,[6] Global Registry of Digital Formats[7]).

Finally, preservers in an organization lacking the required resources to support a digital preservation program should investigate the possibility of establishing collaborative partnerships or consortia to develop and finance a program that meets a minimum acceptable standard.

1.3. Focus on digital records

Preservers must ensure that digital preservation resources are primarily deployed to protect authoritative copies[8] of digital records, rather than to preserve digitized copies of surviving analogue records. The rationale for this is that most analogue records will survive without digitization, whereas digital records will be lost without a digital preservation program.

1.4. Offer advice

Because the chain of preservation of digital records begins at creation, preservers should provide advice on digital records creation and maintenance. Depending on the mandate of the preserver, this may be quite specifically targeted to, for example, employees in the preserver's organization or, as in the case of national archives, other government institutions. In other cases, the advice may be disseminated widely to special interest groups or to the general public, with

[6] See http://hul.harvard.edu/formatregistry/.
[7] See http://www.fedora.info/.
[8] Authoritative copy is defined as "The instantiation of a record that is considered by the creator to be its official record and is usually subject to procedural controls that are not required for other instantiations" (InterPARES 2 Terminology Database. Available at http://www.interpares.org/ip2/ip2_terminology_db.cfm).

the purpose of reaching the person(s) or organization(s) whose records fall under the mandate of the preserver.

1.5. Set a good example

Preservers must establish, within their own organization, a record-making and a recordkeeping environment such that their own control records produced in the course of their preservation function will be created and maintained in a way that satisfies the InterPARES 1 *Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records*.[9] Not only is this an essential requirement for any organization undertaking long-term preservation, but the development of this type of in-house environment will also provide:

- hands-on training to archivists in the technologies they are championing to records creators;
- an invaluable "user's eye view" of actual recordkeeping solutions and how they really work in a day-to-day operational environment;
- a testbed where upgrades and innovations can be introduced and evaluated; and
- a working prototype that can be used in demonstrations.

1.6. Develop procedures

Preservers must establish controls over records transfer, maintenance and reproduction, including the procedures and system(s) used to transfer records to their own organization or program within the organization; maintain them; and reproduce them in a way that satisfies the InterPARES 1 *Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records*.[10] These procedures must embody adequate and effective controls to guarantee the records' identity[11] and integrity,[12] and specifically that:

- unbroken custody of the records is maintained;
- security and control procedures are implemented and monitored;
- the content of the records and the required annotations and elements of documentary form remain unchanged after reproduction.

1.7. Implement maintenance strategies

Although much attention is paid to the development of complex long-term preservation strategies, such strategies are inapplicable if the records for which they are to be used are not properly maintained and protected in the recordkeeping and/or record preservation systems that contain them. A complete version of the eight primary maintenance strategies is available in Appendix 21c, Section A. Briefly, they include:

---

[9] See Authenticity Task Force (2002), "Appendix 2: Requirements for Assessing and Maintaining the Authenticity of Electronic Records," in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 204-219. Online reprint available at http://www.interpares.org/book/interpares_book_k_app02.pdf. See Appendix 21a for an abridged version.

[10] Ibid. See Appendix 21b for an abridged version.

[11] Identity is defined as "The whole of the characteristics of a document or a record that uniquely identify it and distinguish it from any other document or record. With integrity, a component of authenticity" (InterPARES 2 Terminology Database, op. cit.).

[12] Integrity is defined as "The quality of being complete and unaltered in all essential respects. With identity, a component of authenticity" (Ibid.).

A1. Clear allocation of responsibilities
A2. Provision of appropriate technical infrastructure
A3. Implementation of a plan for system maintenance, support and replacement
A4. Implementation of a plan for the transfer of records to new storage media on a regular basis
A5. Adherence to appropriate storage and handling conditions for storage media
A6. Redundancy and regular backup of the digital objects
A7. Establishment of system security
A8. Disaster planning

2. Appraise Records for Permanent Preservation (A4.2)

In cases where, as recommended in the InterPARES 2 Chain of Preservation model, retention scheduling is employed, decisions on the disposition of records will regularly be made as part of the management of a recordkeeping system. In some cases, appraisals may be conducted when it is determined that records in a longstanding system need to reach a disposition. Eight important aspects of the appraisal process are discussed below.

2.1. Appraise early

Given the technical difficulties involved in the preservation of digital records, the identification of what records need to be preserved for the long term should be carried out at the earliest possible opportunity. Performing appraisal, establishing transfer methods and even identifying potential preservation strategies with the records creator will improve the likelihood of success. This process may also provide the preserver with an opportunity to offer records creation and maintenance advice (see Section 1.4).

Professional preservers, such as archivists, are frequently encouraged to participate in the actual design of computer applications being developed by organizations with which they have a donor-preserver relationship. This approach will help integrate appropriate recordkeeping and preservation practices. Preservers who have joined system design teams have learned that it is an enormously time-consuming practice that requires a far more detailed understanding of the organization's internal workflows and procedures than an archivist normally acquires during an appraisal. Furthermore, system specifications are rarely an accurate depiction of the application that will eventually be implemented. An appraisal will still have to be conducted once the system is operational and is meeting organizational requirements. It may be more reasonable for archivists to contribute to system design as part of the advice function discussed in Section 1.4. Sharing high level strategies, principles and guidelines developed by the archival profession may prove to be a more realistic goal.[13]

[13] Many aspects relating to the creation of effective digital preservation programs have been studied in recent years. Among the Web sites containing useful information or examples are: the InterPARES Project at http://www.interpares.org; Model Requirements for the Management of Electronic Records (MoReq) at http://www.cornwell.co.uk/edrm/moreq.asp; the Metadata Encoding and Transmission Standard (METS) at http://www.loc.gov/standards/mets/; the Electronic Records from Office Systems (EROS) at the National Archives of the United Kingdom at http://www.nationalarchives.gov.uk/electronicrecords/advice/guidelines.htm; and the Australian DIRKS (Designing and Implementing Recordkeeping Systems) manual at http://www.records.nsw.gov.au/recordkeeping/dirks-manual_4226.asp.

2.2. Locate multiple owners

In cases where the intellectual components of a digital object have multiple owners, these owners must be identified during the appraisal process to assess the ramifications of this situation for long-term preservation. This can occur, for example, where institutions at various levels of government contribute, and share access to, data resources. Another example is illustrated by Web sites that access and use resources located outside their span of control. Although access agreements are frequently negotiated in these circumstances, they rarely include provisions for long-term preservation of all significant digital components.

2.3. Assess authenticity

The assessment of authenticity has always formed part of the traditional archival appraisal process. In the first instance, it has relied on confirming the existence of an unbroken chain of custody from the time of the records' creation to their transfer to the archival entity responsible for their long-term preservation. Periods when records were not subject to some form of protective measures by the records creator or by a successor institution with a vested interest in maintaining the accuracy and completeness of the records can cast significant doubt on the authenticity of the records.

The assessment of authenticity has also depended on the archivist's knowledge of recordkeeping practices, both historically and in relation to the record types and administrative procedures of a specific creator. The general framework for this assessment was originally codified in diplomatics.[14] A third, less frequently used method to confirm the identity and integrity of records is based on comparison. Records within a fonds are compared to copies forwarded to and held by external sources in the normal course of the creator's business.

Records created and maintained using digital technology present additional difficulties, and archivists have not yet developed standard practices to assess authenticity in this environment. Issues revolve around the fact that digital objects are easily duplicated, distributed, re-named, re-formatted or converted, as well as to the ease with which they can be falsified without leaving a trace. The following examples illustrate the extent of the loss to archivists, historians, lawyers and others who require authentic records in their work:

- The physical support on which digital documents are stored has largely lost its significance in confirming the date of a record or its place of manufacture. Anyone with access to functioning, obsolete equipment and storage media has the capability to copy digital files to, for example, 9-track tape or 5-1/4" diskettes.
- The date stamp on any digital file can be modified by adjusting the system clock.
- Few institutions understood what their employees would do once entrusted with word processing software. Standard document forms, such as memos and correspondence on letterhead, disappeared under the onslaught of new, individualized record forms, which rapidly included personalized colour, graphics and even sound effects, as well as the attribution of new meaning to capitalization, colour and the development of emoticons. The degree of erosion of standard records creation practices varied enormously across types and sizes of corporate and government organizations.

[14] See discussion of diplomatics in Luciana Duranti and Kenneth Thibodeau (2006), "The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES," *Archival Science* 6(1): 15-21.

- The introduction of e-mail networks allowed records to travel by many new routes among staff, rather than according to the well-established distribution routes of traditional office procedures.
- The severe reductions in records management personnel in most organizations, fuelled by an assumption that digital objects somehow did not need to be managed, played havoc with the holdings of the Records Office, which largely stopped receiving records created and transmitted in digital form.

When appraising records created in a digital environment, the assessment of the authenticity of records must become a more overt, visible process performed and documented by the preserver. Unbroken chain of custody, knowledge of recordkeeping practices, and verification may still offer some assurances of authenticity. To these must now be added the verification of compliance with each of the benchmark requirements for authenticity listed in Section 2.4.

2.4. Document the assessment of authenticity

The appraisal report should document the controls put in place by the creator to guarantee the identity and integrity of the records and thus the presumption of their authenticity. These controls include each of the benchmark requirements supporting the presumption of authenticity.[15] Briefly, these include:

A.1 Expression of Record Attributes and Linkage to Record (e.g., identity and integrity metadata)
A.2 Access Privileges
A.3 Protective Procedures against Loss and Corruption of Records
A.4 Protective Procedures against Media Deterioration and Technological Change
A.5 Establishment of Documentary Forms
A.6 Authentication of Records
A.7 Identification of Authoritative Record
A.8 Removal and Transfer of Relevant Documentation

2.5. Monitor records identified for long-term preservation

Once the appraisal is completed, the records identified for preservation must be monitored at regular intervals until such time as they will be transferred to the preserver. Monitoring involves confirming with the records creator that nothing has changed with regard to how classes of records identified for transfer are being created or maintained or, if changes have occurred, that they have not affected the nature and attributes of the records, their value, their authenticity or the feasibility of their preservation.

Many changes within an organization can affect the ongoing survival of digital records. The possibility that records will be destroyed in an instant is much higher than for traditional records. This danger is somewhat offset by the tendency to duplicate material in an uncontrolled fashion. Unfortunately, if the production of copies is uncontrolled, it is unlikely that anyone will realize when the last copy of a record is destroyed.

The simplest scenario may involve a system upgrade either to the hardware or to the software, which will affect the archives' ability to accept the records. An upgrade could also

[15] See Appendix 21a.

result in even minor system re-design that could remove the ability to separate temporary records from those that must be removed for transfer to the preserver.

A second scenario can involve changes in an organization's mandate or functions. This can easily lead to changes in how computer applications are used, and the nature and amount of data that they contain. People responsible for system re-design may not be aware of the requirement for transfer of the existing records to the designated preserver before the system can be modified. Without intervention, even documentation about the original application and backup tapes will move inexorably toward a scheduled destruction date.

Finally, the widespread collapse of proper records management practices in most organizations means that records are poorly identified and incorrectly stored in unsecured locations. Managers, and even records managers, may not understand the details of the technical infrastructure, while IT staff may be unfamiliar with either the history of an organization or the relative importance of older records in various data stores. Hard drives may be wiped, user accounts and all the files they contain may be deleted, tapes and discs may be recycled or destroyed, and obsolete playback technology may be disposed of to meet day-to-day operational requirements of speed and efficiency, with no understanding of the impact of such actions on an organization's records or on pre-existing transfer agreements designed to ensure their long-term preservation.

2.6. Update appraisals

Appraisals also need to be updated at regular intervals, though less frequently than records identified for transfer need to be monitored. Information gathered during a monitoring visit may provide the first indication that a new appraisal is required. Change within organizations and within their record-making and recordkeeping systems is inevitable. Organizational mandates and responsibilities may change, as well as the way those responsibilities are carried out, and data accumulated in existing systems may be put to new uses, which might increase their long-term value. At the simplest level, systems that did not initially contain records may be upgraded to do so. This is particularly true during this period of "hybrid" recordkeeping systems, where paper-based record systems continue to co-exist with the early stages of digital information, document or record systems.

2.7. Identify all digital components[16]

Paper records kept in traditional recordkeeping systems generally offer a tightly-wrapped package, where the content of the record is firmly attached to its paper support and the record itself is contextually filed with the related records. This seamless system began to break down with the introduction of technology when, for example, photographic negatives had to be processed to produce prints and moving images resulted from multiple layers of sound and images, combined and re-combined to produce the final composite print that is screened in cinemas or broadcast on television.

Digital technology has further dismantled the record into a series of components. To successfully extract digital records from the system in which they were created, or even from a

[16] A digital component is defined as "A digital object that is part of one or more digital documents, and the metadata necessary to order, structure or manifest its content and form, requiring a given preservation action" (InterPARES 2 Terminology Database, op. cit.).

secondary maintenance system, the preserver must ensure that all essential digital components are identified and that implicit relationships are made explicit in the metadata before the whole construct is transferred. One of the most common examples of a digital component is the library of fonts, any number of which can be selected by the creator to be used in the presentation of a word-processed document. In Windows, these are stored in '.dll' (or dynamic link library) files. For the preserver to be able to reproduce this record to reflect the creator's original intentions, both the digital component containing the text and the digital component containing the font must have been preserved, as well as the link between them established in such a way that the software attempting to display the content of the text file can find the appropriate font library.[17]

2.8. Determine the feasibility of preservation

Although not part of the assessment of the value of the records, the appraisal process must be completed by a careful investigation of the technical preservation requirements for preservation. Different preservation strategies (see Appendix 21c, Section B) can vary widely in cost and can produce very different results. A textual record stripped of all its formatting may be acceptable in a situation where the preserver is interested in carrying forward only the content of the record. However, where meaning is conveyed by the documentary form and the display characteristics of the record, a more complex preservation solution will be required.

A determination of the feasibility of preservation is essential if the preserving body is to clearly understand the cost of the acquisition and preservation to which it is committing. This is not a new activity; it is simply the extension to the digital realm of the identification of the resources needed to preserve, for example, paper records that are mouldy or moving image reels that are badly shrunken. The current state of digital preservation does mean, however, that preservation costs must be viewed as recurrent. Re-copying holdings from one physical carrier to another will be required as often as the selected format becomes obsolete. Conversion of file formats will be required when logical obsolescence threatens to make the content unreadable. In addition, the digital records considered for long-term preservation may require measures far too complex for the technological environment and the knowledge resources of the preserving organization, and this might imply a postponement of the transfer.

3. Acquire Selected Records for Permanent Preservation (A4.3)

The activity of the preserver acquiring selected records, and all the activities of preservation that follow from that, have as their goal the continued authenticity and accessibility of those records that are selected for continuing preservation. This movement of records from the creator's (or legitimate successor's) custody to the preserver's custody is a critical juncture in the chain of preservation and must be done with great care to ensure that nothing goes awry in the transfer process.

[17] A more detailed description of the "digital component," with additional examples illustrating the concept, is available in Preservation Task Force (2001), "Appendix 6: How to Preserve Authentic Electronic Records," in Duranti, *Long-term Preservation*, op. cit., 293–328. Online reprint available at http://www.interpares.org/book/interpares_book_o_app06.pdf.

3.1. Develop shared plan for transfer

A successful transfer from the current custodian of the records (be it original creator or legitimate successor) to the organization or program taking on responsibility for long-term preservation requires a plan agreed upon by both parties. Re-accessing obsolete systems or extracting inactive records from operational systems will definitely involve human resource costs for copying time and, potentially, for programming time. Special hardware and software may also be required. The logical and physical (or virtual) formats used for the transfer must be agreeable to both parties. As a general rule, the transfer plan should be developed when the technical feasibility of acquisition and preservation are undertaken. If the two parties cannot agree on a transfer process, the appraisal decision may have to be re-visited. Again, in this period of hybrid recordkeeping, paper and microfilm-based options may still exist. Alternatively, the preserver might encourage the records creator to adopt upgrades to the record system that will allow for easier regular transfers.

3.2. Enforce standardized procedures

The controls over the transfer of digital records from the creator's to the preserver's custody must include:

- establishing, implementing, and monitoring procedures for registering the records transfer;
- verifying the authority for transfer;
- examining the records to determine whether they correspond to the records that are designated for transfer; and
- accessioning the records.

As part of the transfer process, the authenticity of the creator's records, which was assessed as part of the appraisal process, should be verified. This includes verifying that the metadata relating to the records' identity and integrity have been transferred together with the related records and are linked to them, and that the records are accompanied by any relevant documentation of the technical and administrative environment in which they were created and maintained.

3.3. Keep the oldest available logical format

The logical format[18] in which the records were originally created, or in which they are held by the creator at the time of transfer, should, whenever feasible, be maintained by the preserver, in addition to any preservation or reference copies generated after the transfer. Should selected preservation strategies, such as a specific conversion path, fail over time, continued custody of the initial logical format will allow the preserver to essentially re-start the preservation process with the most authoritative copy of the records, by applying a different preservation strategy to the records. Over the long periods during which preservers hold records, experience may show that other preservation strategies are more stable over time or can more easily be carried forward

[18] Logical format is defined as "The organized arrangement of data on electronic media that ensures file and data control structures are recognizable and recoverable by the host computer operating system" (InterPARES 2 Terminology Database, op. cit.). Two common logical formats for files and directories are ISO 9660 for CD-ROMs, and Universal Disk Format (UDF) for DVDs.

over the long-term. Alternately, new methods of preservation may have been developed following the acquisition and initial processing of the records.

3.4. Avoid duplicates

Because of the ease of replication of digital records, the preserver must put in place procedures to ensure that digital records from a specific series are transferred by a specific creator to the preserver only once. Accurate identity information is an important first step in avoiding duplication of effort by the creator and the preserver. Also, if reference copies are provided by the preserver to the creator after the transfer of the records, they should be clearly identified and marked as such to prevent accidental re-transfer.

3.5. Document all processing

Initial processes applied during and immediately after transfer may or may not be related to preservation per se. Confirming the identity of the transferred material, checking for viruses and confirming completeness of files tend to leave the transferred file unchanged. File conversion, renaming digital objects and encapsulating files are more intrusive activities. In both cases, preservers must document all processing of digital records and the effects of processing while records are in their custody (see Appendix 21b, Requirement B.2). This documentation should include information such as:

- why certain processes were applied to the records;
- what records were processed;
- the date when the process was performed;
- the names of persons performing and documenting the various steps of the process(es);
- the impact of the process performed on the records' form, content, accessibility and use; and
- the description of any damage, loss or other problems encountered as a result of the processing, including any effect on the elements expressing the records' identity and integrity.

Should the preserver produce copies of the acquired records, it is important to remember that, as discussed in Section 1.5, these copies should be produced in an environment that satisfies the relevant requirements[19] from the InterPARES 1 Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records.

4. Preserve Accessioned Records (A4.4)

The designated records preserver is the entity responsible for taking physical and legal custody of, and preserving (i.e., protecting and ensuring continuous access to), a creator's records. Be it an outside organization or an in-house unit, the role of the designated preserver should be that of a trusted custodian[20] for a creator's records. The authentic copies of the creator's records are kept by the trusted custodian in a *trusted preservation system* (see Appendix

[19] Requirement A.5 (Establishment of Documentary Forms), where the creator establishes the documentary form of the record, would usually not apply to the preserver, except if the original documentary form of the record has been lost and the preserver must specify a substitute to permit access.

[20] A trusted custodian is defined in the InterPARES 2 Terminology Database as "A preserver who can demonstrate that it has no reason to alter the preserved records or allow others to alter them and is capable of implementing all of the requirements for the authentic preservation of records" (InterPARES 2 Terminology Database, op. cit.).

21c), which should include in its design a description and a retrieval system. This trusted preservation system must also have in place rules and procedures for the ongoing production of authentic copies as the existing system becomes obsolete and the technology is upgraded.

4.1. Describe the records

The information about the records and their contexts collected during the appraisal and processing stages should form part of the archival description of the fonds or series in which the records belong (see Appendix 21b, Requirement B.3). This should also include information about intellectual property rights or privacy concerns.

The archival description of the fonds or series containing the digital records should include—in addition to information about the records' juridical-administrative, provenancial, procedural, and documentary contexts—information about changes the digital records of the creator have undergone since they were first created. The description should also include an overview of the transfer and preservation processes based on the documentation discussed in Section 3.5 and the explanation of the relationships among digital components discussed in Section 2.7.

4.2. Identify legal ramifications of preservation actions

When a preservation strategy is selected, its legal implications should be reviewed. For example, format conversion out of a proprietary environment could involve the preserver in illegal actions. In the United States, the *Digital Millennium Copyright Act* has made it a criminal offence to produce tools that can circumvent copyright protection measures. Internationally, the World Intellectual Property Organization Copyright Treaty (WIPO WCT) contains provisions that include copyright protection for software as well as digital works and that introduce criminal penalties for infringement, which ranges from unauthorized copying of material placed on a Web site to the removal or alteration of rights management controls from digital works. Most software packages also include some type of similar restrictions, which users must agree to during the installation process.

4.3. Confirm the effectiveness of the selected preservation strategy

As discussed in Section 2.8, there are now a number of preservation strategies available. Ideally, the selected preservation strategy should be tested on the records prior to the formal transfer to the preserver, to ensure that it will perform as expected. Realistically, most preserving organizations or programs can only fund this type of testing on an exceptional basis. Just as traditional conservators carefully test proposed treatments before applying them wholesale to analogue records, digital preservers must be constantly alert to the impact that each preservation process may have on the records and ensure that it is the appropriate choice for preserving authentic records. Flaws in application software and variations in the functionality of versions over time can result in unexpected consequences when applied to a new group of records.

Part of this process includes a constant awareness of the need to track the presence and the performance of all digital components. A change in one component may have unexpected results on a second component, or it may affect how the relationship functions between any two essential components of the record or affect these components' ability to interact. A different relationship that could be affected is that which exists among the members of a related group of records, such as a dossier or series, and the presentation of that aggregate in the correct order

(e.g., alphabetical, chronological or hierarchical). If the original order has been lost, corrective measures will have to be taken.

4.4. Maintain proper storage

It is a widely accepted archival preservation principle that maintaining an appropriate and consistent storage environment (temperature and relative humidity) for the material being stored is the most cost-effective contribution to the long-term preservation of records. Manufacturers of magnetic or optical storage media generally offer advice on optimum storage conditions. The environment must be monitored constantly and the readings checked on a regular basis. This recommendation is one of the eight mandatory maintenance strategies outlined in Section 1.7 and discussed in Appendix 21c, Section A.

5. Output Records (A4.5)

As noted earlier, continued accessibility (i.e., use) is an integral part of the archival process. Consequently, providing access to preserved records is an essential component in the chain of preservation. It should be managed by the preserver with the same sense of responsibility and degree of technical and professional competence imparted to records appraisal, acquisition/ transfer, description and storage.

5.1. Explain how the reference copies were made

The relationship between the records acquired from the creator and any copies produced by the preserver must be clearly described and readily accessible to users (see Appendix 21b, Requirement B.2.b). This should also include documenting how the reproduction process control measures that are in place were established and implemented and how they are monitored to ensure that the content of the reproduced records is not changed in the course of reproduction. Copies of records in the preserver's preservation system may not be designated authentic if the preserver has made them for purposes other than preservation; for example, a copy from which personal identifiers are removed may be made for access purposes.

Documenting the records reproduction process and its effects is an essential means of demonstrating that the reproduction process is transparent (i.e., free from pretence or deceit). Such transparency is necessary to the effective fulfillment of the preserver's role as a trusted custodian of the records. It also provides users of the records with a critical tool for assessing and interpreting the records by demonstrating the continuing authenticity of the records and by providing a complete history of the records, of which the history of reproduction is an essential part.

5.2. Explain the technical requirements for access

As mentioned in Section 1.1, different preservers provide reference services to different types of users. This will affect the reference formats and mechanisms adopted by the preserving organization or program, with simpler methods required for members of the general public who may not even own a computer or who may own a fairly simple machine with a few standard pieces of software. To meet the needs of these users, the preserver may have to undertake additional processing or create specialized tools to assist the researchers. More technologically

adept users, such as statisticians doing data analysis or forensic accountants conducting fraud investigations, are more likely to apply their own software tools to copies of the records.

Conclusion

This document has outlined a series of guidelines for institutions, organizations and programs with preservation responsibilities for digital records that can be presumed to be authentic and accurate while in the custody of the preserver. For individual preservers and small preservation organizations, the burden may seem great, but the alternative—loss of records or the emergence of corrupt and inauthentic records—would be an even greater problem in the long run. Small organizations will benefit by making a clear designation of the individual or individuals responsible for overseeing the preservation of the organization's digital records. Bear in mind, however, that not all recommendations presented in this document need to be implemented in each circumstance; each preserver should be able to select and adopt the measures that address its particular problems in the specific context in which it operates. There may also be cases in which additional measures are necessary because of legislative or regulatory requirements specific to the preserver's administrative or cultural jurisdiction. In such cases, consultation with legal experts may be required. Individuals, offices and small organizations responsible for preservation should not hesitate to contact such experts for advice on any issues relating to the preservation of the digital records in their custody and under their control.

Appendix 21a

Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records[1]

Preamble

The benchmark requirements are the conditions that serve as a basis for the preserver's assessment of the authenticity of the creator's electronic records. Satisfaction of these benchmark requirements will enable the preserver to infer a record's authenticity on the basis of the manner in which the records have been created, handled and maintained by the creator.

Within the benchmark requirements, Requirement A.1 identifies the core information about an electronic record—the immediate context of its creation and the manner in which it has been handled and maintained—that establishes the record's identity and lays a foundation for demonstrating its integrity. Requirements A.2–A.8 identify the kinds of procedural controls over the record's creation, handling and maintenance that support a presumption of the record's integrity.

Benchmark Requirements (Requirement Set A)

To support a presumption of authenticity the preserver must obtain evidence that:

REQUIREMENT A.1: Expression of Record Attributes and Linkage to Record	the value of the following attributes are explicitly expressed and inextricably linked to every record. These attributes can be distinguished into categories, the first concerning the identity of records, and the second concerning the integrity of records.

A.1.a Identity of the record:

A.1.a.i Names of the persons concurring in the formation of the record, that is:

- name of author[2]
- name of writer[3] (if different from the author)
- name of originator[4] (if different from name of author or writer)
- name of addressee[5]

A.1.a.ii Name of action or matter

A.1.a.iii Date(s) of creation and transmission, that is:

[1] Excerpted from: Authenticity Task Force (2002), "Appendix 2: Requirements for Assessing and Maintaining the Authenticity of Electronic Records," in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 204-219. Online reprint available at http://www.interpares.org/book/interpares_book_k_app02.pdf.

[2] The name of 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.

[3] The name of the physical or juridical person having the authority and capacity to articulate the content of the record.

[4] The name of the physical or juridical person assigned the electronic address in which the record has been generated and/or sent.

[5] The name of the physical or juridical person(s) to whom the record is directed or for whom the record is intended.

- chronological date[6]
- received date[7]
- archival date[8]
- transmission date(s)[9]

A.1.a.iv Expression of archival bond[10] (e.g., classification code, file identifier)

A.1.a.v Indication of attachments

A.1.b Integrity of the record:

A.1.b.i Name of handling office[11]

A.1.b.ii Name of office of primary responsibility[12] (if different from handling office)

A.1.b.iii Indication of types of annotations added to the record[13]

A.1.b.iv Indication of technical modifications;[14]

REQUIREMENT A.2: Access Privileges	the creator has defined and effectively implemented access privileges concerning the creation, modification, annotation, relocation, and destruction of records;

REQUIREMENT A.3: Protective Procedures: Loss and Corruption of Records	the creator has established and effectively implemented procedures to prevent, discover, and correct loss or corruption of records;

REQUIREMENT A.4: Protective Procedures: Media and Technology	the creator has established and effectively implemented procedures to guarantee the continuing identity and integrity of records against media deterioration and across technological change;

[6] The date, and possibly the time, of compilation of a record included in the record by the author or the electronic system on the author's behalf.
[7] The date, and possibly the time, when a record is received by the addressee.
[8] The date, and possibly the time, when a record is officially incorporated into the creator's records.
[9] The date and time when a record leaves the space in which it was generated.
[10] The archival bond is the relationship that links each record, incrementally, to the previous and subsequent ones and to all those [that] participate in the same activity. It is originary (i.e., it comes into existence when a record is made or received and set aside), necessary (i.e., it exists for every record), and determined (i.e., it is characterized by the purpose of the record).
[11] The office (or officer) formally competent for carrying out the action to which the record relates or for the matter to which the record pertains.
[12] The office (or officer) given the formal competence for maintaining the authoritative record, that is, the record considered by the creator to be its official record.
[13] Annotations are additions made to a record after it has been completed. Therefore, they are not considered elements of the record's documentary form.
[14] Technical modifications are any changes in the digital components of the record as defined by the Preservation Task Force. Such modifications would include any changes in the way any elements of the record are digitally encoded and changes in the methods (software) applied to reproduce the record from the stored digital components; that is, any changes that might raise questions as to whether the reproduced record is the same as it would have been before the technical modification. The indication of modifications might refer to additional documentation external to the record that explains in more detail the nature of those modifications.

REQUIREMENT A.5: Establishment of Documentary Forms	the creator has established the documentary forms of records associated with each procedure either according to the requirements of the juridical system or those of the creator;

REQUIREMENT A.6: Authentication of Records	if authentication is required by the juridical system or the needs of the organization, the creator has established specific rules regarding which records must be authenticated, by whom, and the means of authentication;

REQUIREMENT A.7: Identification of Authoritative Record	if multiple copies of the same record exist, the creator has established procedures that identify which record is authoritative;

REQUIREMENT A.8: Removal and Transfer of Relevant Documentation	if there is a transition of records from active status to semi-active and inactive status, which involves the removal of records from the electronic system, the creator has established and effectively implemented procedures determining what documentation has to be removed and transferred to the preserver along with the records.

Commentary on the Benchmark Requirements Supporting the Presumption of Authenticity of Electronic Records

The assessment of the authenticity of the creator's records takes place as part of the appraisal process. That process and the role of the benchmark requirements within it are described in more detail in the "Appraisal Task Force Report." This assessment should be verified when the records are transferred to the preserver's custody.

A.1 Expression of Record Attributes and Linkage to Record

The presumption of a record's authenticity is strengthened by knowledge of certain basic facts about it. The attributes identified in this requirement embody those facts. The requirement that the attributes be expressed explicitly and linked inextricably[15] to the record during its life, and carried forward with it over time and space, reflects the task force's belief that such expression and linkage provide a strong foundation on which to establish a record's identity and demonstrate its integrity. The case studies undertaken as part of the work of the task force revealed very little consistency in the way the attributes that specifically establish the identity of a record are captured and expressed from one electronic system to another. In certain systems, some attributes were explicitly mentioned on the face of the record; in others they could be found in a wide range of metadata linked to the record or they were simply implicit in one or more of the record's contexts. In many cases, certain attributes (e.g., the expression of the archival bond)

[15] For the purposes of this requirement, *inextricable* means incapable of being disentangled or untied, and *link* means a connecting structure.

were not captured at all. The task force's concern is that, in the absence of a precise and explicit statement of the basic facts concerning a record's identity and integrity, it will be necessary for the preserver to acquire enormous, and otherwise unnecessary, quantities of data and documentation simply to establish those facts.

The link between the record and the attributes listed in Requirement A.1 is viewed by the task force as a *conceptual* rather than a *physical* one, and the requirement could be satisfied in different ways, depending on the nature of the electronic system in which the record resides. For example, in electronic records management systems, this requirement is usually met through the creation of a record profile.[16] In other types of systems, the requirement could be fulfilled through a topic map. A topic map expresses the characteristics (i.e., *topics*) of subjects (e.g., records or record attributes) and the relationships between and among them.

When a record is exported from the live system, migrated in a system update, or transferred to the preserver, the attributes should be linked to the record and available to the user. When pulling together the data prior to export, the creator should also ensure that the data captured are the right data. For example, in the case of distribution lists, the creator must ensure that if the recipients specified on "List A" were changed at some point in the active life of records, the accurate "List A: Version 1" is exported with the records associated with the first version, and that the second version is sent forward with those records sent to recipients on "List A: Version 2."

A.2 Access Privileges

Defining access privileges means assigning responsibility for the creation, modification, annotation, relocation, and destruction of records on the basis of competence, which is the authority and capacity to carry out an administrative action. Implementing access privileges means conferring exclusive capability to exercise such responsibility. In electronic systems, access privileges are usually articulated in tables of user profiles. Effective implementation of access privileges involves the monitoring of access through an audit trail that records every interaction that an officer has with each record (with the possible exception of viewing the record). If the access privileges are not embedded within the electronic system but are based on an external security system (such as the exclusive assignment of keys to a location), the effective implementation of access privileges will involve monitoring the security system.

A.3 Protective Procedures: Loss and Corruption of Records

Procedures to protect records against loss or corruption include: prescribing regular back-up copies of records and their attributes; maintaining a system back-up that includes system programs, operating system files, etc.; maintaining an audit trail of additions and changes to records since the last periodic back-up; ensuring that, following any system failure, the back-up and recovery procedures will automatically guarantee that all complete updates (records and any control information such as indexes required to access the records) contained in the audit trail are reflected in the rebuilt files and also guarantee that any incomplete operation is backed up. The capability should be provided to rebuild forward from any back-up copy, using the back-up copy and all subsequent audit trails.

[16] If the attribute values contained in the profile are also expressed independently as entries in a register of all records made or received by the creator, then, in addition to establishing the identity and supporting the inference of the integrity of the record, they would corroborate such identity and strengthen the inference of integrity.

A.4 Protective Procedures: Media and Technology

Procedures to counteract media fragility and technological obsolescence include: planning upgrades to the organization's technology base; ensuring the ability to retrieve, access, and use stored records when components of the electronic system are changed; refreshing the records by regularly moving them from one storage medium to another; and migrating records from an obsolescent technology to a new technology.

A.5 Establishment of Documentary Forms

The documentary form of a record may be determined in connection to a specific administrative procedure, or in connection to a specific phase(s) within a procedure. The documentary form may be prescribed by business process and work-flow control technology, where each step in an administrative procedure is identified by specific record forms. If a creator customizes a specific application, such as an electronic mail application, to carry certain fields, the customized form becomes, by default, the required documentary form. It is understood that the creator, acting either on the basis of its own needs or the requirements of the juridical system, not an individual officer, establishes the required documentary form(s) of records.

When the creator establishes the documentary form in connection to a procedure, or to specific phases of a procedure, it is understood that this includes the determination of the intrinsic and extrinsic elements of form[17] that will allow for the maintenance of the authenticity of the record. Because, generally speaking, that determination will vary from one form of a record to another, and from one creator to another, it is not possible to predetermine or generalize the relevance of specific intrinsic and extrinsic elements of documentary form in relation to authenticity.

A.6 Authentication of Records

In common usage, to authenticate means to prove or serve to prove the authenticity of something. More specifically, the term implies establishing genuineness by adducing legal or official documents or expert opinion. For the purposes of the benchmark requirements, authentication is understood to be a declaration of a record's authenticity at a specific point in time by a juridical person entrusted with the authority to make such declaration. It takes the form of an authoritative statement (which may be in the form of words or symbols) that is added to or inserted in the record attesting that the record is authentic.[18] The requirement may be met by linking the authentication of specific types of records to business procedures and assigning responsibility to a specific office or officer for authentication.

The authentication of copies differs from the validation of the process of reproduction of the digital components of the records. The latter process occurs every time the records of the creator are moved from one medium to another or migrated from one technology to another.

A.7 Identification of Authoritative Record

An authoritative record is a record that is considered by the creator to be its official record and is usually subject to procedural controls that are not required for other copies. The

[17] The extrinsic and intrinsic elements of form are defined and explained in the InterPARES 1 *Template for Analysis* (see Authenticity Task Force (2000), "Appendix 1: Template for Analysis," in Duranti, *Long-term Preservation*, op. cit., 192–203. Online reprint available at http://www.interpares.org/book/interpares_book_j_app01.pdf).

[18] The meaning of authentication as it is used by the Authenticity Task Force in this report is broader than its meaning in public key infrastructure (PKI) applications. In such applications, authentication is restricted to proving identity and public key ownership over a communication network.

identification of authoritative records corresponds to the designation of an office of primary responsibility as one of the components of a record retention schedule. The Office of Primary Responsibility is the office given the formal competence for maintaining the authoritative (that is, official) records belonging to a given class within an integrated classification scheme and retention schedule. The purpose of designating an office of primary responsibility for each class of record is to reduce duplication and to designate accountability for records.

It is understood that in certain circumstances there may be multiple authoritative copies of records, depending on the purpose for which the record is created.

A.8 Removal and Transfer of Relevant Documentation

This requirement implies that the creator needs to carry forward with the removed records all the information that is necessary to establish the identity and demonstrate the integrity of those records, as well as the information necessary to place the records in their relevant contexts.

Appendix 21b

Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records[1]

Preamble

The baseline requirements outline the minimum conditions necessary to enable the preserver to attest to the authenticity of copies of inactive electronic records.

Baseline Requirements (Requirement Set B)

The preserver should be able to demonstrate that:

REQUIREMENT B.1: Controls over Records Transfer, Maintenance, and Reproduction	the procedures and system(s) used to transfer records to the archival institution or program; maintain them; and reproduce them embody adequate and effective controls to guarantee the records' identity and integrity, and specifically that	
	B.1.a	Unbroken custody of the records is maintained;
	B.1.b	Security and control procedures are implemented and monitored; and
	B.1.c	The content of the record and any required annotations and elements of documentary form remain unchanged after reproduction.

REQUIREMENT B.2: Documentation of Reproduction Process and its Effects	the activity of reproduction has been documented, and this documentation includes	
	B.2.a	The date of the records' reproduction and the name of the responsible person;
	B.2.b	The relationship between the records acquired from the creator and the copies produced by the preserver;
	B.2.c	The impact of the reproduction process on their form, content, accessibility and use; and
	B.2.d	In those cases where a copy of a record is known not to fully and faithfully reproduce the elements expressing its identity and integrity, such information has been documented by the preserver, and this documentation is readily accessible to the user;

[1] Excerpted from: Authenticity Task Force (2002), "Appendix 2: Requirements for Assessing and Maintaining the Authenticity of Electronic Records," in *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Luciana Duranti, ed. (San Miniato, Italy: Archilab, 2005), 204-219. Online reprint available at http://www.interpares.org/book/interpares_book_k_app02.pdf.

REQUIREMENT B.3: Archival Description	the archival description of the fonds containing the electronic records includes—in addition to information about the records' juridical-administrative, provenancial, procedural, and documentary contexts—information about changes the electronic records of the creator have undergone since they were first created.

Commentary on the Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records

The establishment and implementation of the baseline requirements take place as part of the function of managing preservation. The preservation function and the role of the baseline requirements within it are described in more detail in the "Preservation Task Force Report."

B.1 Controls over Records Transfer, Maintenance, and Reproduction

The controls over the transfer of electronic records to archival custody include establishing, implementing, and monitoring procedures for registering the records' transfer; verifying the authority for transfer; examining the records to determine whether they correspond to the records that are designated in the terms and conditions governing their transfer; and accessioning the records.

As part of the transfer process, the assessment of the authenticity of the creator's records, which has taken place as part of the appraisal process, should be verified. This includes verifying that the attributes relating to the records' identity and integrity have been carried forward with them (Requirement A.1), along with any relevant documentation (Requirement A.8).

The controls over the maintenance of electronic records once they have been transferred to archival custody are similar to several of the ones enumerated in the benchmark requirements. For example, the preserver should establish access privileges concerning the access, use, and reproduction of records (Requirement A.2); establish procedures to prevent, discover, and correct loss or corruption of records (Requirement A.3), as well as procedures to guarantee the continuing identity and integrity of records against media deterioration and across technological change (Requirement A.4). Once established, the privileges and procedures should be effectively implemented and regularly monitored. If authentication of the records is required, the preserver should establish specific rules regarding who is authorized to authenticate them and the means of authentication that will be used (Requirement A.6).

The controls over the reproduction of records include establishing, implementing, and monitoring reproduction procedures that are capable of ensuring that the content of the record is not changed in the course of reproduction.

B.2 Documentation of Reproduction Process and its Effects

Documenting the reproduction process and its effects is an essential means of demonstrating that the reproduction process is transparent (i.e., free from pretence or deceit). Such transparency is necessary to the effective fulfilment of the preserver's role as a trusted custodian of the records. Documenting the reproduction process and its effects is also important for the users of records since the history of reproduction is an essential part of the history of the record itself. Documentation of the process and its effects provides users of the records with a critical tool for assessing and interpreting the records.

B.3 Archival Description

Traditionally it has been a function of archival description to authenticate the records and perpetuate their administrative and documentary relationships. With electronic records, this function becomes critical. Once the records no longer exist except as authentic copies, the archival description is the primary source of information about the history of the record, that is, its various reproductions and the changes to the record that have resulted from them. Although it is true that the documentation of each reproduction of the record copies[2] may be preserved, the archival description summarizes the history of all the reproductions, thereby obviating the need to preserve all the documentation for each and every reproduction. In this respect, the description constitutes a collective attestation of the authenticity of the records and their relationships in the context of the fonds to which the records belong. This is different from a certificate of authenticity, which attests to the authenticity of individual records. The importance of this collective attestation is that it authenticates and perpetuates the relationships between and among records within the same fonds.

[2] Although, technically, every reproduction of a record that follows its acquisition by the preserver is an authentic copy, it is the only record that exists and, therefore, should normally be referred to as "the record" rather than as "the copy."

Appendix 21c

Digital Records Maintenance and Preservation Strategies[1]

This appendix includes a list of preservation strategies largely drawn from the UNESCO *Guidelines for the Preservation of Digital Heritage*,[2] which offers a framework for describing digital records preservation strategies that protect and maintain the accessibility of authentic copies of digital records throughout the chain of preservation.

The complete list of possible strategies adopted by InterPARES 2 is conceptually divided into two broad categories: a) maintenance strategies and b) preservation strategies.

A. Maintenance Strategies

Maintenance strategies[3] are the minimum necessary requirement to protect and maintain accessibility of authentic copies of digital records. There are eight primary maintenance strategies. All are necessary to ensure the records components will exist long enough for preservation strategies to be applied.

A1. Clear allocation of responsibilities

A person or office must be given unambiguous responsibility for managing records storage and protection. This is a technical responsibility that requires a specific skill set, dedicated resources, and an appropriate plan. This strategy can be undertaken by hiring a competent staff member devoted exclusively to this task or by assigning existing staff or an existing office a specific portion of time to carry out the responsibilities.

A2. Provision of the appropriate technical infrastructure

This includes all of the physical and administrative resources that enable the recordkeeping and/or maintenance processes (buildings, computer hardware, computer networks and the auxiliary staff necessary to maintain the same).

A3. System maintenance, support and replacement

The implementation of a plan for maintaining, updating and/or replacing hardware and software.

A4. Transfer of data to new storage media on a regular basis

The implementation of a plan for copying of data from one storage medium to another to avoid the impact of media decay. Such transfers should be undertaken in a systematic manner.

A5. Adherence to appropriate conditions for storage media

The rate of media decay may be dramatically reduced by adhering to appropriate environmental conditions. For instance, excessive heat, humidity and dust endanger storage media.

---

[1] Adapted from: Kevin Glick, "Electronic Records Preservation Strategies," (unpublished report, 2006).

[2] Colin Webb (2003), *Guidelines for the Preservation of Digital Heritage*. Prepared by the National Library of Australia for the Information Society Division, United Nations Educational, Scientific and Cultural Organization, report no. CI-2003/WS/3. Available at http://unesdoc.unesco.org/images/0013/001300/130071e.pdf.

[3] A maintenance strategy is defined as "A coherent set of objectives and methods for protecting (i.e., safeguarding authenticity and ensuring accessibility of) digital components and related information over time while still in active or semi-active use by the creator, and for reproducing the related authentic records and/or record aggregations" (InterPARES 2 Terminology Database, op. cit.).

A6. Redundancy and geographic location

The duplication of digital objects and the storage of the resulting multiple copies on different physical media protects them against media failure. Storage in different physical locations protects against poor environmental storage conditions, fire, flood, etc., at a particular storage site.

A7. System security

Controls should be implemented to ensure that digital components of records are exposed only to authorized users and/or processes. Such controls should include restricting physical access to places where computers are kept as well as restricting access to the digital records on the computers themselves. The latter can be accomplished through various means, including the use of passwords and/or biometric authentication to log on to the system.

A8. Disaster planning

The strategies listed above are designed to minimize accidental loss of data and maximize media longevity, but even with perfect storage conditions and excellent handling protocols, disasters may still happen. A disaster recovery plan should contain detailed procedures for restoring a damaged system and for guiding the effective recovery of recordkeeping and/ or preservation systems following a disaster.

B. Preservation Strategies

In addition to the maintenance strategies, every records preserver is responsible for establishing a trusted preservation system[4] for expressing one or more preservation strategies.[5] Twelve preservation strategies are listed below, in Section B, divided into four broadly defined groups. It is most likely that, in practice, a preserver will support two or more preservation strategies in addition to the eight maintenance strategies listed above in Section A.

B1. Use of standards

The use of widely available and supported standards increases the likelihood of stability and longer term support. Such standards may either be *de jure*,[6] if they have been formally agreed upon, or *de facto*,[7] if they have been widely adopted by industry. Standards can apply to many facets of a preservation system, including encoding methods, file formats, physical storage media, etc. Compliance with standards might also simplify the application and/or maximize the effectiveness of later preservation strategies. Standardization may be applied *prospectively*, by limiting the formats in which digital records may be transferred to the preserver; or *retrospectively*, by converting files received in other formats to standard ones.

[4] A trusted preservation system is defined as "The whole of the rules that control the preservation and use of the records of the creator and provide a circumstantial probability of the authenticity of the records, and the tools and mechanisms used to implement those rules" (Ibid.).

[5] A preservation strategy is defined as "A coherent set of objectives and methods for protecting (i.e., safeguarding authenticity and ensuring accessibility of) digital components and related information of inactive records over time, and for reproducing the related authentic records and/or archival aggregations" (Ibid.).

[6] A *de jure* standard is defined as a "Standard issued by an official standards-setting body, whether national (e.g., ANSI), multi-national (e.g., CEN) or international (e.g., ISO)" (Ibid.). For computer file formats, two recent *de jure* standards are PDF/A (PDF standard for archiving) and ODF (OASIS OpenDocument Format).

[7] A *de facto* standard is defined as a "Standard not issued by any official standards-setting body, but nevertheless widely used and recognized by its users as a standard" (Ibid.). Well known and widely used computer file formats that are considered *de facto* standards include PDF, TIFF, DOC and ZIP.

B1.1. Self-describing formats (persistent object preservation, tagging)

Analysis and tagging of records so that the functions, relationships and structure of specific elements can be described. The re-presentation of content can be liberated from specific software applications and can be achieved using different applications as technology changes.

B1.2. Encapsulation

Binding together a record and the means of providing access to it, normally in a *wrapper* that describes what it is in a way that can be understood by a wide range of technologies (such as an XML document). The wrapper often includes metadata that describe or link to the correct tools.

B1.3. Restricting the range of formats to be managed (normalization)

Storing records in a limited number of formats only.[8] The selection of acceptable formats may continue to include new proprietary formats or new generations of existing proprietary formats, or it may be restricted to non-proprietary formats, to carry standardization one step further. One example of this approach is referred to as *durable encoding*, which recommends encoding records to conform to well-known data processing standards down to the level of encoding bits as ASCII or Unicode UTF-8, and objects as XML.

B1.4. Conversion

Transferring digital records from one hardware or software generation to another. As distinct from *refreshing*, which copies the data stream from one carrier to another, conversion entails transforming the logical form of a digital object so that the conceptual object can continue to be correctly rendered or presented by the new hardware or software. The most commonly proposed conversion method involves permanently transforming one logical format into another in line with technological change, so that all converted objects can be presented with prevailing technology. It is also possible to propose a "conversion on demand" or "conversion at the point of access" model. This approach is discussed below in Section B2.4.

B2. Technology dependence

These strategies continue to rely on the original hardware and/or software without changing the records.

B2.1. Technology preservation

Maintaining the original software and hardware with which digital records were presented.

B2.2. Reliance on backward compatibility

Trusting the ability of some software to correctly interpret and present digital components of records created with previous versions of the same software. In this strategy, the presentation is limited to a temporary conversion for viewing or for non-archival copying purposes, whereas conversion permanently changes records into the format supported by the current version of the software.

[8] For a detailed analysis of current issues and trends in the selection of file, wrapper, tagging and encoding formats, together with recommendations for developing and implementing policies on selecting digital file formats for long-term preservation, see: Evelyn Peters McLellan (2006), "InterPARES 2 Project - General Study 11 Final Report: Selecting Digital File Formats for Long-Term Preservation." Available in English at http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_english.pdf, and in French at http://www.interpares.org/display_file.cfm?doc=ip2_gs11_final_report_french.pdf.

B2.3. Software re-engineering

Transforming software as technologies change. As such, it is similar to the transformation of record formats, discussed in sections B1.4.and B2.2. This may include anything from re-compiling source code for a new platform to re-coding the software from scratch in another programming language.

B2.4. Viewers and conversion at the point of access

The use of software tools or transformation methods that provide temporary accessibility when needed, using the original data stream.

B2.5. Emulation

Using software that makes one technology behave like another. In other words, making future technologies behave like the original environment of a preserved digital record, so that the original record could be presented in its original manifestation from the original, or converted, data streams.

B3. Non-digital approaches

Copying the digital records onto relatively stable analogue media, such as paper or microfilm; shifting the preservation burden to an analogue copy in place of the digital object. This approach destroys any functionality provided by the software, such as manipulability.

B4. Data restoration (digital archaeology)

Recovering records as bits from physical media followed by steps to restore the intelligibility of the recovered records. It is most often employed in the recovery of data from failed, damaged or degraded media, but methods to restore intelligibility have been used to rescue documents in obsolete formats.