

Significant Properties Report

Document type: InSPECT Work Package 2.2

Author: Andrew Wilson

Draft/Version: V2

Date of edition completion: 10/04/2007

Change history

Second version

Contents

Objective	2
1. Background	2
2 Digital Preservation	3
2.1. Preservation Approaches	3
3. Migration.....	3
4. Authenticity.....	4
5. Significant Properties (or Characteristics) and Essence	5
6. Initiatives.....	7
7. Project Concepts	8
References.....	9

Objective

This package will explore the current versions, views, and visions of the concept at the core of this proposal, the idea of 'significant properties', also referred to as 'significant characteristics' and 'essence' by collecting all discussions of the concept, analysing them and developing a clear, consistent articulation of the concept. The project will review work being undertaken within the DELOS project, and now continuing within the EU-funded PLANETS project (in which TNA is participating), within NARA's Electronic Records Archive project in the US, and at the National Archives of Australia.

1. Background

The nature of digital objects is very different from their physical analogues and preservation far from being a relatively straightforward technical process is fraught with difficulties and conceptual problems. What does preservation mean in the digital context? What are we actually doing when we 'preserve' a digital object over time? What is it that we are trying to preserve?

These questions have been of concern to those involved with digital preservation for at least the last decade. However, prior to 2000 there were few articulations of these concerns. Since then however, a number of important articles and papers have focussed on these issues of digital preservation at a conceptual level.

To digress briefly. Data curators, archivists, librarians, and other information managers, are very comfortable with operating in a physical world, where the objects that are the focus of preservation have a tangible reality and existence. In brief, preservation in this world aims to ensure that the physical object remains accessible and useable for as long as necessary or possible. Of course, the physical carrier of the data (eg a sheet of paper) can be separated from the data that it carries but for most purposes the carrier and the data are inextricable. To preserve the data in this context, it is necessary to preserve the physical carrier. But that is all, since neither exists independently as a data object without the other. Preserving the object results in preservation of the data.

Digital objects, on the other hand, although they are often functionally the same as their paper analogues, are inherently different. Perhaps the most obvious difference is that digital objects are mediated by technology. For the objects to be used, a user must have access to the right combination of hardware and software to enable the object to be re-created.¹ Certain non-trivial consequences flow from this unavoidable technological dependence of digital objects.

The major consequence is that it is not possible to leave the digital object alone and expect it to survive. In the past, a do nothing approach, frequently characterised as 'benign neglect', has often been enough to ensure the survival and useability of physical data. However, computer systems and software applications change so rapidly there is no guarantee that existing data sources will be accessible and useable on future computing platforms or software versions. Thus, in a digital world, the rate of technological obsolescence means that a 'do nothing' approach is dangerous and will result in the loss and/or destruction of the data. A second important consequence is that it is not enough to merely preserve the carrier medium. The best attention paid to ensuring the longevity of the digital carrier alone, be it CD, hard disk, DVD, digital tape or whatever, is not a recipe for success. Of course, the digital media must be looked after because no digital medium has anything approaching the archival longevity of paper. But this is not enough and will not enable the data itself to remain useable over time. There must, therefore, be 'active intervention' to make sure that the digital object can be located, accessed and used over time.

2 Digital Preservation

Digital preservation, then, can be viewed as covering all actions necessary to keep the digital object accessible and useable over time, and as ensuring that the object's contextual data and its informational content have not been compromised by anything that has been done to it in the preservation or access processes. More briefly, we may define the aim of digital preservation as ensuring that **authentic** digital objects remain accessible and useable over time.ⁱⁱ We will return to the issue of **authenticity** later in this paper.

It was observed in 2001 that

“traditionally, preserving things meant keeping them unchanged; however our digital environment has fundamentally changed our concept of preservation requirements. If we hold on to digital information without modifications, accessing the information will become increasingly difficult, if not impossible”.ⁱⁱⁱ

Thus, the fundamental challenge of digital preservation is to preserve the accessibility and authenticity of digital objects over time and domains, and across changing technical environments. This requires acceptance both of the inevitability of change, and of the inherent separation of the logical information object from its physical environment. Any successful preservation strategy must reconcile the requirement to maintain the fixity/integrity of that logical information object, with the inevitable transformation of the technical environment in which the object resides.

2.1. Preservation Approaches

Various approaches to preserving digital objects have been proposed and discussed in the literature of digital preservation. It is not intended to discuss each approach at length, instead we may summarise the various approaches briefly as:

- techno-centric: keeping and maintaining the object's original hardware and software;
- data-centric: maintaining objects in current data formats;
- process-centric: creating new processes to render the original object;
- post-hoc: digital archaeology or forensics

This paper is not the appropriate place to discuss the benefits or short comings of any or all of these digital preservation approaches. However, the emphasis in this paper will be on preservation approaches that are data-centric since it is in these methods that digital objects are changed over time. Emulation and techno-centric approaches depend on the underlying assumption that every single characteristic of the original object must be carried forward over time and, therefore, that digital objects must remain completely unchanged.^{iv}

Proponents of migration strategies (as we are)^v agree with the view of the digital preservation dilemma as articulated above and advocate the acceptance of change as a necessary aspect of digital preservation. This paper rests on the assumption that currently migration strategies offer the best hope for long-term preservation of digital objects.

3. Migration

The current consensus about digital preservation holds that approaches that are data-centric, ie. concerned about keeping the data object useable over time, offer better prospects for success than those which are process-centric, ie. concerned to keep original software and/or hardware environments operational over time. Three forms of migration have been considered in the literature and are variously implemented:

- migration at obsolescence;
- normalisation;
- migration on request.

Migration at obsolescence involves the ongoing migration of a digital object to new formats, or new versions of the original format, at the last possible moment before existing formats become obsolete. This approach seeks to take advantage of economies of scale and advances in migration techniques, and depends on access to reliable information to understand the lifecycle of file formats. The National Archives of the UK is an advocate for this approach.

Normalisation involves the migration of digital objects to a file format, or a very restricted range of acceptable formats, with enhanced preservation characteristics. In this approach migration into the preservation format(s) takes place as soon as possible to ensure ongoing access and use. This is the form of migration adopted in the National Archives of Australia's digital preservation operational unit.

Migration on request involves the retention of the digital object in its original format until is required for use. At this time it is migrated into a current format which is retained along with the original object. This approach requires the development of tools that can migrate the original object to the current format as necessary. This form of migration was conceptualised by the Cedars project in the UK, and some proof of concept applications developed. The LOCKSS (Lots of Copies Keeps Stuff Safe) Program for e-journal preservation also uses a migration on request approach to preservation.^{vi}

It is clear from this very brief and necessarily unsophisticated outline of migration approaches that the one thing they all have in common is that changes are made to the original format of the object. In other words, migration involves the alteration, and sometimes loss, of the original data. The original bitstream is altered in the migration process. Does this render the migrated object invalid or unauthentic?

4. Authenticity

Preservation of digital objects is closely tied to notions of authenticity.^{vii} Authenticity is commonly used in the context of digital preservation of a digital object (as above), as one of the defining factors of successful preservation. If the object does not remain authentic then the preservation system has not worked. But what is meant exactly by this term 'authenticity'?

Of course, in its broad meaning 'authenticity' carries with it all the connotations of that much overused word truth. Here we limit 'authenticity' to its archival meanings to do with what a record purports to be and how it was created.^{viii} Variations of this definition abound but the central core of the concept is fixed. The JISC definition, for example, reads: "An authentic digital resource is one that is what it purports to be, is free from corruption, and is intact in all essential respects". Here is introduced the idea of that authenticity incorporates what is often called fixity or integrity, ie. that there has been no corruption at the bit-level or deliberate alteration at the semantic level. For the UK National Archives assessing authenticity involves establishing the integrity and identity of the object – integrity here referring to the objects "wholeness and soundness", and identity referring to attributes such as context and provenance.^{ix} The idea of integrity arises from the area of authenticity that requires a digital object to be what it purports to be. If it has been changed at either the bit or semantic level, whether deliberately or not, it is not what it purports to be.

Why is authenticity important for digital objects? Does this mean that authenticity stands or falls on whether the objects is unchanged or not at the bit level? As the TNA document previously cited observes, "This does not mean that a record must be precisely the same as it first created for its integrity to exist and be demonstrated. A record is considered to be essentially complete and uncorrupted if the message meant to communicate in order to achieve its purpose is unaltered". In other words, digital objects do not need to remain in a state that is unchanged from their original state in order for them to be considered authentic.

Any successful preservation strategy must reconcile the perceived requirement to maintain the authenticity and integrity of the logical information object, with the inevitable transformation of the technical environment in which the object resides. A useful conceptual

model for understanding this is provided by the 'performance model' and the concept of 'essence' which were developed by the digital preservation project of the National Archives of Australia (NAA) in 2002.^x

The 'performance model' characterises a digital record^{xi} as the result of a mediation of technology and data. The data source (record, object, etc) needs to interact with a process in order to be understood by a user. That process is a combination of hardware (computer) and software (application + operating system). Without this mediation the digital object is meaningless since the data that makes it up exists independently only as a stream of bits, completely incomprehensible to the majority of humans. The NAA conceptualisation views the result of the combination of source and process as a performance, since it is created anew every time the record (or object) is viewed by a user. As the NAA document says, "A source may be mediated by many different software platforms, and each combination of source and specific process platform may produce a slightly different performance". What is rendered to a screen or printer, or any other output device, is the performance created when data and process interact. What this conceptualisation of digital object shows is that "neither the source nor the process need be retained in their original state for a future performance to be considered authentic. As long as the essential parts of the performance can be replicated over time, the source and process can be replaced".^{xii}

5. Significant Properties (or Characteristics) and Essence

The NAA developed the concept of the 'essence' of a digital record as "a way of providing a formal mechanism for determining the characteristics that must be preserved for the record to maintain its meaning over time. The performance model demonstrates that digital records are not stable artefacts; instead they are a series of performances across time. Each performance is a combination of characteristics, some of which are incidental and some of which are essential to the meaning of the performance". The 'essential characteristics' are what NAA calls the 'essence' of a record. In NAA terms the essence of a record is the set of characteristics of the record that are essential to the meaning of the performance.^{xiii}

Of course, the NAA researchers are not the only ones to talk about the 'essence' of a digital object, although the terms is theirs alone. Other projects and researchers have referred to 'significant properties' and 'significant characteristics', and even 'essential characteristics' (or properties). Indeed, the concept had been discussed even if obliquely, before the NAA articulation, although the NAA work remains the most well-thought out and articulated intellectual justification of the concept. The well-known Clifford Lynch discussed the idea of 'significant properties' in a 1999 D-Lib article, an early piece in the relatively short history of digital preservation writings.^{xiv} Lynch's view of "canonicalization" is reflected in the National Archives of Australia's more recent 'normalisation' approach to format migration.^{xv} In Lynch's paper there is an unargued assumption that 'reformatting' of the original digital object will need occur in order to keep the object useable. There is an early exploration of the idea of significant properties in this paper. Lynch notes that "We want to be able to guarantee that for a given object the reformatted version is equivalent to the original version with regard to some specific set of object characteristics". Lynch sees "canonicalization" as the answer to this need since this concept is the reformatting the digital object to a canonical form that "to some extent, captures the essential characteristics of that type of object in a highly determined fashion".

The same term reappeared in the course of a JISC funded research project at the Universities of Leeds, Oxford and Cambridge - the well-known CEDARS project, operating from April 1998 to March 2002. The aim of CEDARS was to explore issues of preservation of digital objects and to develop proof of concept implementations. The technical approach developed by the CEDARS work is summarised in the document "Cedars Guide to Digital Preservation Strategies" published on the CEDARS website in 2002.^{xvi} The CEDARS approach is a variant of migration, referred to as 'migration on request'. In this approach digital objects are kept in their original format until required for access. At that time they are migrated to a current format suitable for access purposes. For the CEDARS researchers, this migration must preserve all the 'significant properties of the original' [s. 11.]. CEDARS defined 'significant

properties' as "those components of a digital object deemed necessary for its long-term preservation", without further expanding on what those components are. Knowing the significant properties of the digital object is seen as the key to choosing the target format for migration of the obsolete original. Migration format options are assessed for their ability to preserve the significant properties and the most suitable chosen as the migration format. The CEDARS researchers realise that for complex formats there will be a lot of effort in "deducing significant properties" [s. 13].

The CEDARS project overlapped with and informed another JISC funded project, the CAMiLEON project at the Universities of Michigan and Leeds, 1999-2003. The purpose of the CAMiLEON (Creative Archiving at Michigan and Leeds: Emulating the Old on the New) project was to develop and evaluate a range of technical strategies for the long term preservation of digital materials. CAMiLEON focussed on preservation of the 'look and feel' of digital objects and so concerned itself with emulation strategies. A number of the researchers involved in the CEDARS project were also involved with CAMiLEON so it is not surprising that significant properties were an important component of both projects. For both projects, significant properties are a function of the underlying abstract representation of a digital object – decide on the correct abstract representation and you have the significant properties determined.^{xvii} It is instructive that neither project developed or articulated an explicit definition of significant properties. Margaret Hedstrom, Project Director of CAMiLEON for the University of Michigan, further discussed the CAMiLEON approach in a paper delivered to the 2002 DLM Forum in Barcelona.^{xviii} In this paper is provided the project's definition of the term significant properties as used by the project, i.e. "those properties of digital objects that affect their quality, usability, rendering, and behaviour."^{xix}

Hedstrom states that "The purpose of our research is to identify the significant properties of digital objects that affect their quality, functionality, and look-and-feel so that custodians can select appropriate methods which preserve those significant properties of digital objects that are deemed important by designated user communities". This is a significant expansion of the project aims, as stated on the project website (see previous paragraph). However, although the term is used in a number of the project's papers, none of the published papers of the project actually set out, even at an abstract level, the significant properties of any digital objects.^{xx}

2002 was a busy year for digital preservation. Also in that year, the Council on Library and Information Resources (CLIR) published the proceedings of their Conference on "The State of Digital Preservation: An International Perspective".^{xxi} Included in those papers was one by Ken Thibodeau of the US National Archives and Records Administration, titled "Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years".^{xxii} In that important paper, Thibodeau accepts that it is not necessary to retain the object unchanged over time, arguing that the object only needs to be the same in "all essential aspects". He discusses 'essential properties' (sometimes 'characteristics') arguing that we need to be able to define these for a digital object in order to preserve it. Unfortunately, he does not define the term although it is implicit that essential properties or characteristics are what must be kept when you change an object for preservation purposes.

The terms 'significant properties', 'significant characteristics', 'essential properties', and 'essential characteristics' are used interchangeably in the literature. The National Archives of Australia is the only source for the use of the word 'essence', and the term does not appear elsewhere in the literature except in reference to the NAA work. Ken Thibodeau in his 2002 paper discusses "essential properties" of digital objects, a combination of the NAA and CEDARS terms that carries the same meaning as both.^{xxiii} The other terms are used frequently in the digital preservation literature, but there has been little discussion of significance in the literature since the work noted above. It seems to this observer that there has been a general acceptance of the idea that significant properties are important to the digital preservation mission, but most of those who use the term would be hard-pressed to define it or say why it is important. Nevertheless, some important initiatives are exploring the issue and its consequences for digital preservation.

6. Initiatives

The identification of those properties which are significant to the continued preservation and accessibility of authentic digital objects, across changing technical environments, is a fundamental task for successful digital preservation. Unless such properties can be defined in a rigorous and measurable manner, cultural memory institutions have no objective framework for identifying, implementing, and validating appropriate preservation strategies, nor for asserting the continued authenticity of their digital collections.

In recent years, the need to identify such properties has been highlighted within a number of notable digital preservation programmes. These include the National Archives of Australia, the Electronic Record Archives programme at the National Archives and Records Administration, The National Archives' Seamless Flow programme in the UK, and the EU-funded DELOS project. Some conceptual work on authenticity and object properties has been undertaken as part of the InterPARES 2 project at the University of British Columbia.^{xxiv} However, to date, little research has been undertaken on the practical application of the concept and approach. It is therefore widely recognised that there is a pressing need for practical research in this area, to develop a methodology, and begin identifying quantifiable sets of significant properties for specific classes of digital object.

National Archives of Australia (NAA)

<http://www.naa.gov.au/recordkeeping/preservation/digital/summary.html>

The NAA 'performance' approach has been implemented in a digital preservation operational unit within that institution. NAA uses a migration strategy based on 'normalisation', whereby digital records are migrated at time of ingest into a very small range of open, non-proprietary digital formats. For example, images are migrated to PNG format, audio is migrated to BWF and most other formats are migrated to an XML equivalent. The XML outputs for each ingest format type have corresponding XML schema which determines the structure and semantics of the output types. The XML outputs contain the 'essence' of the input objects, and the XML schema are an explicit statement of what NAA sees as the essence of the input object types. Recently, NAA has begun to document in textual form, essence statements for various ingest formats; unfortunately, these are not yet publicly available documents.

Cornell University risk-based approach to formats:

<http://www.clibr.org/pubs/abstract/pub93abst.html>

Cornell University in the US has developed an approach to managing digital formats over time that is based on a comprehensive risk assessment of migrating formats. Part of the risk assessment involves the comparison of the source format with the format that is the migration target. Although not made explicit in the paper this shares aspects of the current project and is essentially an assessment of the ability of the target format to retain the significant characteristics of the source format. The approach is explained in full in:

Lawrence, G. W., Kehoe, W. R., Rieger, O. Y., Walters, W. H., and Kenney, A. R.
(2000) *Risk Management of Digital Information: a File Format Investigation*,
Washington DC: Council on Library and Information Resources

UK National Archives (TNA) Seamless Flow Programme

http://www.nationalarchives.gov.uk/electronicrecords/seamless_flow/

TNA is developing an approach to modelling the significant properties of electronic records, primarily as a means of validating preservation actions such as migration. Work so far has focussed on developing a data model to support this, which allows the definition of extensible, type-specific properties for particular classes of electronic record. TNA is also developing a service-oriented architecture for deploying characterisation and preservation action tools, driven through the PRONOM technical registry, and is enhancing PRONOM to model property invariance for specific migration pathways.

PLANETS

<http://www.planets-project.eu/>

PLANETS is funded under the European Commission Information Science and Technologies Framework Programme 6 Call 5, which brings together European National Libraries and Archives, leading research institutions, and technology companies to address the challenge of preserving access to digital cultural and scientific knowledge. Understanding significant properties is an important element within PLANETS. The Preservation Planning sub-project is developing methodologies which use utility analysis for evaluating alternative preservation actions against standard criteria, which can include the capabilities of the preservation actions with respect to preserving significant properties. The Characterisation sub-project, led by TNA, is developing a generic framework to automatically deploy new and existing characterisation tools, and process their outputs in a standardised manner. These characterisation tools could be used to measure the significant properties of specific digital objects, and their outputs could therefore be used by preservation planning to assess the merits of preservation actions.

US National Archives and Records Administration (NARA) – Electronic Records Archive (ERA)
<http://www.archives.gov/era/>

NARA's ERA project is considering a template-based approach to describing the significant properties of particular types of electronic record. A given category of record, such as a particular form of report, can be described by a standard set, or template, of significant properties. Preservation action tools can then be categorised in terms of their capabilities measured against this template.

7. Project Concepts

This project will use the term 'significant properties' in preference to either 'significant characteristics' or 'essence', although we regard the terms as essentially interchangeable. For the purposes of this project we define 'significant properties' as:

the characteristics of digital objects that must be preserved over time in order to ensure the continued accessibility, usability, and meaning of the objects.

The significant properties of digital objects fall into 5 categories:

- content, eg. text, image, slides, etc.
- context, eg. who, when, why.
- appearance, eg. font and size, colour, layout, etc.
- structure, eg. embedded files, pagination, headings, etc.
- behaviour, eg. hypertext links, updating calculations, active links, etc.

For the object types under investigation the project will evaluate the characteristics of each type and assign them to the relevant category. This will form the basis for the assessment of the migration and representation outputs at later stages of the project.

Note that there are other types of property (e.g. technical/representation properties) which will be significant in terms of preservation planning, but which are quite distinct from the properties discussed here. It is not within the scope of the project to attempt to define and categorise every single property of digital objects.

References

- ⁱ Adapted from Heslop, H., S. Davis, and A. Wilson, *An Approach to the Preservation of Digital Records*, National Archives of Australia, 2002.
- ⁱⁱ See H. Hofman, "Can Bits and Bytes be Authentic? Preserving the Authenticity of Digital Objects", *IFLA Conference*, Glasgow, 2002. Viewed at: http://eprints.erpanet.org/39/01/hofman_glasgow02.pdf. The OAI Reference Model defines 'long term preservation' as "the act of maintaining information, in a correct and Independently Understandable form, over the Long Term (p. 1-11). 'Independently Understandable information' is defined as information "that has sufficient documentation to allow the information to be understood and used by the Designated Community without having to resort to special resources not widely available..." (p. 1-10). Unfortunately for the purposes of this paper, 'correct' is not defined. Can we thus infer that 'correctness' is defined by the OAI or the Designated Community and is not a consideration for the Reference Model itself?
- ⁱⁱⁱ Su-Sing Chen, "The Paradox of Digital Preservation", *Computer*, March 2001, 2-6.
- ^{iv} This view is perhaps most strongly made by Jeff Rothenberg, a well-known advocate of emulation approaches. See for example Jeff Rothenberg, *Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation, A Report to the Council on Library and Information Resources*, CLIR, Washington DC, 1999. Available at <http://www.clir.org/pubs/reports/rothenberg/contents.html>.
- ^v The OAI Reference Model refers to these strategies as 'transformation'.
- ^{vi} DSH Rosenthal et al., "Transparent Format Migration of Preserved Web Content", *D-Lib Magazine*, vol. 11 no. 1, January 2005. At: <http://www.dlib.org/dlib/january05/rosenthal/01rosenthal.html>.
- ^{vii} In fact, maintenance of authenticity is one of the aims of management of physical resources as well, not just of digital ones. See Consultative Committee for Space Data Systems (2002) Reference Model for an Open Archival Information System (OAIS), CCSDS 650.0-R-1 – Blue Book, 2002, p. 2-1. Available at: <http://www.classic.ccsds.org/documents/pdf/CCSDS-650.0-B-1.pdf>
- ^{viii} See D. Bearman and J. Trant, "Authenticity of Digital Resources: Towards a Statement of Requirements in the Research Process", *D-Lib Magazine*, June 1998. This conceptualisation is given a more formal rendition in *ISO 15489 International Standard for Records Management, Part 1 – General*, 2002, section 7.2.2.
- ^{ix} TNA, *Defining the characteristics for authentic records*, 2002, p. 8. At: http://www.nationalarchives.gov.uk/documents/generic_reqs1.pdf.
- ^x H Heslop et al, *An Approach to the Preservation of Digital Records*, National Archives of Australia, 2002.
- ^{xi} The NAA Green Paper is focussed on digital records since that is the business of the National Archives of Australia, but the conceptualisation holds true for any digital object.
- ^{xii} H. Heslop et al., p. 11.
- ^{xiii} H. Heslop et al., p. 13.
- ^{xiv} Lynch, Clifford A, "Canonicalization: A Fundamental Tool to Facilitate Preservation and Management of Digital Information", *D-Lib Magazine*, September 1999. At: <http://www.dlib.org/dlib/september99/09lynch.html>. Viewed 26/02/2007
- ^{xv} Digital records are migrated to a very small range of open preservation formats, eg. XML and OpenDoc, which are assessed as having better preservation prospects, ie. their working life is significantly longer than the proprietary counterparts.
- ^{xvi} See: <http://www.leeds.ac.uk/cedars/guideto/dpstrategies/dpstrategies.html> [accessed 26 February 2007].
- ^{xvii} These concepts are discussed in a number of papers written by the project teams. See for example, P. Wheatley, "Migration – a CAMiLEON discussion paper", *Ariadne* 29, October 2001, at <http://www.ariadne.ac.uk/issue29/camileon#7>; and D. Holdsworth and P. Wheatley, "Emulation, Preservation and Abstraction", *RLG DigiNews*, vol. 5 no. 4, August 2001, at <http://www.rlg.org/preserv/diginews/diginews5-4.html#feature2>.
- ^{xviii} M. Hedstrom and C.A. Lee, "Significant properties of digital objects: definitions, applications, implications", *Proceedings of the DLM-Forum 2002*, at http://ec.europa.eu/transparency/archival_policy/dlm_forum/doc/dlm-proceed2002.pdf.
- ^{xix} Hedstrom and Lee, p. 218.
- ^{xx} The project papers are available from: <http://www.si.umich.edu/CAMILEON/reports/reports.html>.

^{xxi} Council on Library and Information Resources, *The State of Digital Preservation: An International Perspective. CONFERENCE PROCEEDINGS*, Washington 2002.

^{xxii} K. Thibodeau, "Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years" in *The State of Digital Preservation: An International Perspective*, CLIR Report 107, 2002. Available at: <http://www.clir.org/PUBS/reports/pub107/pub107.pdf>

^{xxiii} . See: <http://www.clir.org/pubs/reports/pub107/thibodeau.html>.

^{xxiv} InterPARES 2 Project website: http://www.interpares.org/ip2/ip2_index.cfm.