The Role of Metadata in the Longevity of Cultural Heritage Resources

Milena Dobreva
Centre for Digital Library Research,
University of Strathclyde
26 Richmond Street Glasgow, G1
1XH, Scotland, United Kingdom
milena.dobreva@strath.ac.uk

Nikola Ikonomov
Institute of Mathematics and
Informatics
bl. 8 Acad. G. Bonchev St.
1113 Sofia, Bulgaria
nikonomov@gmail.com

Abstract

Digital preservation is an integral part of the management of information and the institutions in the cultural heritage sector are seeking for ways to incorporate it into their everyday practice. While there are generic approaches to long-term preservation, further research and development work is needed to address any specifics of the digital objects in the cultural heritage domain. In this paper, we will take two case studies of recent projects and analyse to what extent the metadata accompanying digital objects contribute to guarantee longevity. We summarize this analysis in two scenarios for sustainability of resources produced by small projects because compared to big institutions their digital assets are facing a higher risk not to be preserved properly. We also identify processes where natural language technologies could be of help to make the preservation more efficient.

1 Introduction

An evaluation made in August 2008 in a recent Communication[1] of the EC states that "The absence of clear and comprehensive policies in many Member States was identified in the Recommendation as a threat to the survival of digi-

tised and born-digital material…" It also provides a good practice example from the UK, the Digital Preservation Coalition. Although this British experience is given as an example of a successful approach, the implementation of preservation in the current digital repositories in the UK is far from being satisfactory. The recent final report of the DRAI project (see Daisy Abbott 2008) on UK-based digital repositories identified that "Of the 3,707collections aggregated, only 737 (20%) had evidence of a preservation policy, and of these the vast majority were from within the Arts and Humanities Data Service (AHDS) Collections Repository which, since March 2008, is no longer funded, leaving only 6% of resources with ongoing policies for long term preservation."

Why the digital preservation which is recognised to be of high importance, is still not implemented widely? The Blue Ribbon Task Force Report (2008) summarized five major obstacles in building sustainable digital preservation programs: inadequacy of funding models to address long-term access and preservation needs; confusion and/or lack of alignment between stakeholders, roles, and responsibilities with respect to digital access and preservation; inadequate institutional, enterprise, and/or community incentives to support the collaboration needed to reinforce sustainable economic models; complacency that current practices are good enough; and fear that digital access and preservation is too big to take on. While these address mostly stakeholder-related issues, a recent report (DPE: 2007) reviewed the progress of digital preservation **research** world-wide. It concludes that "The analysis of the last 16 years of effort in the area of digital preservation… support our claim that while much work has been done on the periphery, the central problems of digital preservation

---

[1] Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions *Europe's cultural heritage at the click of a mouse. Progress on the digitisation and online accessibility of cultural material and digital preservation across the EU* [SEC(08) 2372] of 11 August 2008. Available: http://ec.europa.eu/information_society/activities/digital_libraries/doc/communications/progress/communication_en.pdf

have yet to be resolved." Automation in digital preservation is one of the ten areas named by this study as being in a need of accelerated research.

These two reports identify two key areas which need attention: the organisational aspects and the research. Any institution which currently faces the preservation challenge has to make a difficult choice under the circumstances of the rapidly changing technological environment and the lack of a generally recommended and effective solution.

However, the deeper reason why after two decades of efforts digital preservation still has not reached maturity, is the lack of consistent theoretical foundation. In 2001, James Cheney et al. articulated the need to develop a mathematical theory for the area of information preservation. Having such a theory should help to understand better the object models and the basic transformations which need to be supported in a preservation system; it will also be of great benefit to automation because it would be much more clear what processes can be automated. In addition, a coherent theory would help to identify any gaps in procedures or data flows and thus would contribute to the implementation of systems which behaviour could be consistently traced. However, since this first suggestion that a theory of preservation is needed, there has not been much progress in developing one. Even the identification of the basic elements of a theory of preservation is not a trivial task.

In addition to this methodological difficulty, the variety of domains where digital preservation is to be applied brings even more difficulties. It is still not clear whether a universal 'one-size-fits-all' solution could be used in all different domains. For example, in the cultural heritage domain which is considered in this paper there are various curatorial situations with regard to digital objects. Resources prepared by small projects are in more danger with regard to their longevity compared to resources of large institutions if we consider that they are created but the life-cycle for digital objects curation is not applied in full. In this paper it is not our aim to investigate what is the level of awareness and readiness of the whole range of cultural heritage institutions but we try to highlight what is the current state in the digital preservation field and what issues need to be addressed by projects in the cultural heritage domain.

In 2002, the reference model for an open archival information system (OAIS) was published and one year later adopted in the international standard ISO 14721:2003, see (OAIS: 2003). It provides a functional approach which is helpful to understand and describe the processes in a preservation system. However, the specification of the minimum information which should accompany a digital object in order to guarantee its accessibility, interpretability and usability in the future, is not addressed in OAIS.

There are different ongoing attempts to build a theory of preservation. Giorgos Flouris and Carlo Meghini in 2007 again suggested basing such a theory on firm mathematical foundations and such an effort will be made within the CASPAR[2] project. Paul Watry in 2007 presented the directions of future research of the SHAMAN[3] project which aims to study the context which guarantees that the essential properties of digital objects such as authenticity and integrity will be maintained over time. Reagan Moore and MacKenzie Smith suggested in 2007 a practical approach, which demonstrates how various requirements of the preservation systems can be implemented as sets of rules.

Attempts to find an approach to building a coherent theory of preservation should also reflect the specific features of various subject domains.

In Section 2 of this paper, we introduce the basic concepts of the ISO standard in the digital preservation domain OAIS[4]. Then in Section 3 we discuss issues related to the use of metadata for preservation in the cultural heritage domain. We take two examples of resources under preparation in the TEXTE project[5] and the KT-DigiCult-BG[6] project and suggest two scenarios for preservation of the products of small projects. In Section 4 we summarise the issues which need further development in the cultural heritage sector in order to address better the issue of longevity of the digital resources.

2 The Basic Standard in the Digital Preservation Domain: OAIS

In response to the increased need for the introduction of common practices in assuring long term digital preservation (DP) of digital objects, the International Standards Organisation (ISO) in the last decade developed a number of conceptual DP standards and also some technical guidelines. The most popular standard in the area is ISO 14721:2003 (Space data and information transfer systems – Open archival information system – Reference model), widely known as OAIS[7].

It is a conceptual framework which presents the main functional components and identifies the basic data flows within a digital preservation system. The development of OAIS[8] arose from the need for a model which would specify the basic components of a system for long-term preservation of digital objects and their relationships with the 'external world'.

In Fig. 1, an OAIS is represented as a box which is connected to other entities from the environment. These are **Producers**, **Consumers**, and **Management (**OAIS is not suggesting what roles could be defined within the archive). A special class of Consumers is the **Designated Community**. The Designated Community is the set of Consumers who should be able to understand the preserved information."[9]
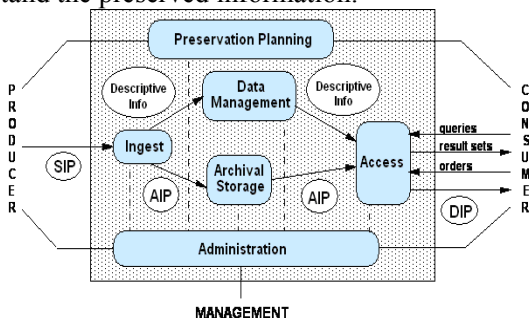

Figure 1: The OAIS functional entities and environment[10]

Within the cultural heritage domain, a special attention also should be paid to the concept of designated community which is essential for the understanding of resources in the future. Understanding the actual needs and the level of background knowledge of the designated community is not trivial in this domain, especially in the cases of creating scholarly resources.

This in fact opens a new essential issue related to the dynamics of resources: while the resources created by memory institutions such as libraries, archives and museums are 'static' because they represent the resources of the institutions, the resources created by projects tend to be dynamic, because they could be enriched and extended. These resources usually do not cover a fixed set of holdings but are oriented towards a specific theme and after a core set of resources is available, more materials could be added at later stages. Being familiar with the model will help cultural heritage professionals to organize better the digital objects' lifecycle and take informed decisions on the preservation aspects of their resources.

In addition to the definition of these functional entities, an OAIS information model explains the data flows between the environment and the archive, and also within the functional components of the archive.

Every act of submission of information to an OAIS by a Producer, as well as the dissemination of information to a Consumer, occurs either as a single discrete transmission, or as a series of transmissions. To describe the exchange of data, OAIS defines the concept of an **Information Package** as a "container" of two types of information: **Content Information (CI)** and **Preservation Description Information (PDI)**. The Content Information and PDI are viewed as being encapsulated and identifiable by the **Packaging Information (PI)**.

OAIS defines three specialised types of information packages (IP), namely:

- Archival Information Package (AIP): An Information Package, consisting of the Content Information and the associated PDI, which is preserved within an OAIS.

- Dissemination Information Package (DIP): The Information Package, derived from one or more AIPs, received by the Consumer in response to a request to the OAIS.

- Submission Information Package (SIP): An Information Package that is delivered by the Producer to the OAIS for use in the construction of one or more AIPs.

PDI is divided into four types of preservation information called *Provenance, Context, Refer-*

---

[7] A webpage which provides information on the basic stages of its development is http://nost.gsfc.nasa.gov/isoas/.
[8] For a record of the activities see
http://nssdc.gsfc.nasa.gov/nost/isoas/.
[9] Source: p.2.2-2.3, OAIS (2003).
[10] Source: Figure 4-1, page 4-1 (OAIS, 2003).

ence, and *Fixity.* Provenance describes the source of the Content Information; Context describes how the CI relates to other information outside the Information Package. Reference provides identifiers, by which the CI may be uniquely identified. Fixity provides a wrapper, which protects the CI from undocumented alteration. A key point to note is that OAIS does not suggest any specific metadata as instantiations of PDI; so that there is no guidance on what constitutes a minimum technical requirement in respect of representing and encoding PDI information within corresponding PDI data bitstreams. The definition of a minimum required set of data should be based on a study of what is required to assure a reliable, consistent, and measurable specification and implementation of a preservation system.

The consequences for the cultural heritage domain is that currently there is no single and uniformly accepted set of elements which guarantee the longevity of resources; but every single institution and project should consider how to structure the different types of archival packages if it establishes and digital archive and also decide what metadata need to be stored for preservation purposes.

In OAIS the closest analogue to the high-level notion of metadata can be found in the idea of *representation information* while on the implementation level PDI provides the closest notion to a particular instantiation of a set of metadata elements.

Having a reference framework for long-term digital preservation systems had an essential impact on the development of a common professional understanding and vocabulary. It has had a significant impact on the subsequent development of standards and as a guide to the setting up of preservation systems. Yet, the complexity of the area allows for multiple interpretations of how the framework should be implemented in real life applications. This arguably gives rise to a need to develop implementation guidelines and also to adjust any practical solution to policy frameworks.

3 Preservation-Related Metadata

3.1 Metadata for Preservation vs. Preservation of Existing Metadata

There are two key issues which need to be considered vis-à-vis metadata and preservation.
1. What metadata are needed for preservation purposes in the cultural heritage domain (besides assuring a reliable preservation process,

they should help the designated communities to understand the resources), and
2. How to preserve the metadata accompanying existing digital objects. Since it is often the case that cultural heritage resources have extensive metadata and there are multiple schemes used, this issue also touches upon understanding and use of present metadata schemes in the future; this issue is very similar to interoperability of metadata schemes but considered as interoperability between now and the future.

In the area of preservation metadata, the basic concern, as stated by Brian Lavoie and Richard Gartner, still seems to be the development of preservation elements' schemes (2005). However, presentation and interoperability issues are not the only concern: preservation metadata, as all other types of metadata are affected by the metadata bottleneck[11], a metaphor which indicates that the human efforts needed to create metadata can not cope with the pace of creation of new digital resources.

Another disquieting example comes from a recent evaluation of a German national digitisation programme which reveals "insufficient metadata practice, endangering the usage of the digital documents, not to speak of their preservation: 33% of the objects had no metadata at all, 33% bibliographic metadata only, 10% had both bibliographic and subject metadata (rest: no information). Less than a third of the metadata was digital."[12]

To suggest an OAIS-compliant preservation metadata solution, in 2002 OCLC (Online Computer Library Center, Inc.) and RLG (Research Libraries Group) created a working group to explore how a consensus between different projects and initiatives in the field of preservation metadata could be achieved. A body of recommended metadata for digital preservation was developed in the next years under the name PREMIS: PREservation Metadata Implementation Strategies (see PREMIS (2008)). The effort of OCLC and RLG to define preservation metadata is a major development in the digital preservation field. However, it analyses only preservation metadata related to the digital object itself. A number of issues likely to appear in real imple-

---

[11] The term metadata bottleneck was coined by E. Liddy in 2002.
[12] DELOS brainstorming on Metadata topics, Juan les Pins, 05/12/2005,
http://www.ukoln.ac.uk/ukoln/staff/t.koch/pres/Brainst2005 12-MDc.html

mentation scenarios are not considered, for example the preservation and re-use of descriptive and other metadata which might have been supplied with the digital object which is of special importance in the cultural heritage domain. Such metadata could be of help in generating preservation metadata, or could be later used for search.

Having all these in mind, preservation metadata area provides many challenges. What preservation metadata to use? What minimum set needs to be supplied in order to guarantee a reliable preservation process? How to automate the creation of preservation metadata? How to guarantee that the digital resources developed within a particular project are accompanied by sufficient preservation quality metadata? And how to guarantee interoperability between multiple existing schemes? It is not easy for any organisation or project to make decisions regarding the metadata in this situation.

In the last years there were several metadata schemes developed dedicated to long-term preservation. In the domain of cultural heritage institutions these new kind of metadata have to be combined with existing descriptive metadata based on the traditionally used catalogue systems. Amongst those we are The Metadata Encoding and Transmission Standard[13] (METS), a container format for metadata and content files maintained by the Library of Congress (USA); and LMER (Long-term preservation Metadata for Electronic Resources)[14] developed by the German National Library.

Currently, researchers come to the conclusion that the practical preservation activities require to combine several various metadata schemes to achieve a reliable preservation process. For example Angela Dappert and Markus Enders (2008) present an example of integrating METS, PREMIS, and MODS[15] metadata in an OAIS-compatible system which aims to ingest, store, and preserve eJournals in the British Library. They demonstrated that no single existing metadata schema accommodates the representation of descriptive, preservation and structural metadata necessary for such a system.

3.2 Two Examples

3.2.1. The TEXTE Project Digital Resources

TEXTE is an ongoing project coordinated by The Moore Institute in the National University of Ireland, Galway. It combines four tasks which aim to illustrate how the current computer-enhanced editing tools can be used to produce various types of high-quality scholarly editions (mediaeval manuscripts; correspondence; a set of journal issues and ephemera). The materials are digitised and then transcribed and encoded using the Guidelines of the Text Encoding Initiative[16]. The resulting resources are highly annotated scholarly editions, a combination of the original texts with the scholarly findings.

All these resources are built using the same guidelines for text encoding. This common framework facilitates the development of the digital objects especially with regard to their online publication and visualisation and to the annotation of resources from the same historical and cultural background.

From digital preservation point of view, there are several issues which need to be taken into account:

The TEI P5: Guidelines for Electronic Text Encoding and Interchange[17] are not suggesting metadata elements to be used especially for preservation purposes. It is essential to have well-formed <TEI.Header> and to document the principles applied in the encoding of the text; these are prepared as a part of the project documentation. But the <TEI.Header> currently does not support elements which can help to trace the digital object's authenticity, integrity and chain of custody – these could be included as additions. But a difficulty for such projects is the lack of guidance on the significant properties of digital objects which need to be supported.

Projects such as TEXTE developing textual resources would benefit from clear guidance on this matter; currently the project faces the need to find its own solution. The final product of TEXT will be a set of web resources but their inclusion into repository is still under question; if this happens additional effort will be needed to define SIPs and to transform currently available XML files into SIPs.

[13] http://www.loc.gov/standards/mets/
[14] LMER description and LMER schema: http://www.d-nb.de/eng/standards/lmer/lmer.htm
[15] http://www.tei-c.org/index.xml

[16] Metadata Object Description Schema, MODS, http://www.loc.gov/standards/mods/
[17] http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index-toc.html

On the other metadata-related issue, the longevity of specific metadata schemes applied in the cultural heritage domain, the decision to use a widespread encoding framework might be seen as a guarantee of the future usability of resources. This is based on the expectation that the TEI will be continuously supported in the future.

3.2.2. The KT-DigiCult-BG Project Digital Resources

The KT-DigiCult-BG project funded by the Marie Curie programme of FP6 was recently finished (in 2008) and one of its outcomes was the creation of the Digitisation Centre in the Institute of Mathematics and Mathematics at the Bulgarian Academy of Sciences (IMI-BAS). This centre digitised a substantial number of various types of objects belonging to the Bulgarian cultural and scientific heritage (see Table 2). One of the issues the project faced was the organisation and the long-term preservation of these objects. These resources were not deeply annotated and are intended for the most general type of users – the citizens. In this case the metadata used in the mix of archival materials, old printed books, photographs and periodicals followed various encodings and were not encoded like the metadata of the TEXTE project using the same Guidelines.

This heterogeneity of approaches leads to a different situation in the produced resources. Currently the digitisation centre team is planning to place all scientific resources into a repository using DSpace[18] which will enable the use of its resources within the DRIVER[19] project repository.

Preservation-wise, these resources will be transformed according to this larger initiative requirements and the relevant SIP will be built. In this scenario the preservation task with regard to the digitised documents prepared for access in PDF format stays with the repository while the digitisation centre will take care for the physical copying of the master files which are stored in TIFF format. This is a mixed approach where various bodies take care for the preservation of clusters of digital resources.

Table 2. Digitised Resources in 2005-08 in IMI-BAS

Type of material	Size
Old printed books (National Library Ivan Vazov, Plovdiv)	17,000 pages
State Archives	8,000 pages
Research Archive of the Academy of Sciences	24 archival units, 1,000 pages
Old periodicals (newspapers)	1,200 pages
Musical periodicals	1,000 pages
Archive of the Higher Attestation Commission	8,500 pages
Archive of IMI-BAS	18,000 pages
Mathematical heritage (the heritage of Prof. Obreshkov)	2,000 pages
Audio archive of Bulgarian dialects	7 hours
Mathematical periodicals (1905 - 1991)	29,000 pages
Old photographs	200 photographs
TOTAL	85,900 units

3.3 Two Emerging Scenarios for Digital Preservation of Small Projects' Outputs

Two scenarios emerge from the analysis of the two case studies.

Scenario 1. Preparation of stand-alone resources. Under this scenario, the set of digital resources is prepared as a stand-alone collection and does not necessarily become a part of a digital repository.

1. Such projects need to pay special attention to proper documentation of the decisions taken on the encoding applied within the project – this is not only substantial for sustainability, but also for re-use of resources in the future as demonstrated by Claire Warwick et al. (2009). Documentation which refers to a general standard like TEI or XML is not sufficient because it does not guarantee future understanding and interoperability.

2. A set of preservation metadata needs to be supplied for every single digital object. This is currently not a trivial requirement because there is no common view on the structure of preservation metadata. Again, the internal project decisions need to be well-documented.

3. The overall responsibility to document significant properties of the digital objects which will allow checking authenticity and chain of custody stays with the project; it is unlikely to expect that anyone in the future will be able to fill in such gaps if they are present.

---

[18] http://www.dspace.org/
[19] DRIVER: Networking European Scientific Repositories, http://www.driver-repository.eu/.

Scenario 2. Preparation of resources which will be integrated into a larger digital repository. The following key considerations should be taken into account:

1. It is essential to be familiar with the processes of ingest of digital objects into the repository.
2. The structure of the SIPs for ingest should be discussed in advance.
3. The significant properties of digital objects need to be discussed with the digital repository; it should guarantee to retain authenticity and chain of custody related to the ingested objects.

The following Analysis Matrix (see Table 3) summarizes what issues needs to be taken into account in these scenarios.

Table 3. Summary of digital-preservation issues reflected in the two suggested scenarios

Issue	Scenario 1. (stand-alone resources)	Scenario 2. (resources to be integrated into a digital repository)
Digital object	The digital object is prepared completely and solely within the project.	The digital object is likely to be enriched in order to be ingested into a repository. This enrichment can be done by various parties according to the adopted procedures for ingest into the repository.
Significant properties of the digital object	Small projects usually do not consider significant properties but they are essential as future evidence of authenticity, integrity and chain of custody.	The values of the significant properties which guarantee authenticity, integrity and chain of custody need to be supplied jointly with the digital object when it is being ingested into the digital repository.
Preservation actions	The preservation actions e.g. copying of media or migration to newer file formats will be applied within the host institution; it needs to plan for these otherwise the danger is that the collection as a whole (or separate objects) will be lost because of media decay.	The preservation actions are applied within the repository. The creators of the collection do not have to plan for this but need to make sure what are the digital repositories policies on physical copying and format migration.
OAIS implementation	It is unlikely that small projects will implement the complete set of OAIS functional entities.	The repository hosting the project outputs may implement a specific subset of OAIS functional entities; this should be discussed in advance as a sustainability guarantee.
Possible application of natural language processing (NLP) technology	NLP can be applied for automated generation of metadata, including preservation metadata; we do not have evidence of such attempts so far.	An area of active research is the application of NLP for automated extraction of metadata from digital objects during ingest. This will help to overcome the metadata bottleneck.

We hope that this succinct analysis will help projects to understand better the preservation-related options they have.

4 Conclusions

Our presentation of digital preservation issues and the specifics of cultural heritage domain pinpoints several issues which are essential for future implementations:

1. The understanding of digital preservation standards and models is essential to organise a proper lifecycle management for the cultural heritage resources; there are examples of solutions applied in the large memory institutions[20] but smaller projects and initiatives need to address these issues when they develop their sustainability plans.
2. Widely used metadata standards in the cultural heritage domain, such as TEI, could offer better guidance on metadata elements subset designed especially to support long-term preservation. TEI provides mechanisms for follow-

---

[20] The German National Library and the Royal Library in the Netherlands are examples of institutions which apply successfully digital preservation solutions.

ing any changes in the digital object; it also could be used to encode information related to other copies made such as master or preservation copies and access or use copies; and possibly to record any processes applied to a resource, but this is not sufficient because it reflects the process of creation of an object but this is not sufficient to claim authenticity of this object. Currently there is no mechanism to check the integrity and the chain of custody of digital objects. However, TEI guidelines should suggest how correctly to represent these changes. For example Linda Cantara in 2005 suggested to form an AIP for the resources of the *Tibet Oral History Archive Project* especially for digital preservation purposes; more examples are needed in order to establish a good practices which could be incorporated in other cases.

3. OAIS does not suggest any specific metadata as instantiations of preservation description information and in particular in the cultural heritage domain there is no single and uniformly accepted set of elements which guarantee the longevity of resources; therefore it is necessary to define what metadata need to be stored for preservation purposes. We can not offer a general solution to this problem because it requires consensus of the communities of professionals.

We also have presented two scenarios on digital preservation derived from case studies on two EC-funded projects. This work shows that there is definitely a current need to suggest good practices for such endeavours, because while the projects tackle the preservation issue, it is very difficult to find individual good models to guarantee the longevity of resources.

Reference

Angela Dappert and Markus Enders. 2008. Using METS, PREMIS and MODS for Archiving eJournals. In: *D-Lib Magazine* September/October 2008, Vol. 14, Number 9/10, ISSN 1082-9873 Available:
http://www.dlib.org/dlib/september08/dappert/09dappert.html

Blue Ribbon. 2008. The Blue Ribbon Task Force on Sustainable Digital Preservation and Access (2008). Interim Report. Available:
http://brtf.sdsc.edu/biblio/BRTF_Interim_Report.pdf

Brian Lavoie and Richard Gartner. 2005. Preservation metadata, DPC Technology Watch Series Report 05-01. Available:
http://www.dpconline.org/docs/reports/dpctw05-01.pdf

Claire Warwick, Isabel Galina, Jon Rimmer, Melissa Terras, Ann Blandford, Jeremy Gow, George Buchanan (2009) Documentation and the Users of Digital Resources in the Humanities. Journal of Documentation, Vol. 65 (1), pp. 33-57.

Daisy Abbott. 2008. JISC Final Report – Digital Repositories and Archives Inventory Project, HATII, University of Glasgow, July 2008.

DPE. 2007. Digital Preservation Europe (DPE) (2007). Access to and preservation of cultural and scientific resources: research roadmap. http://www.digitalpreservationeurope.eu/publications/reports/dpe_research_roadmap_D72.pdf

Giorgos Flouris and Carlo Meghini. 2007. Some Preliminary Ideas Towards a Theory of Digital Preservation, 2007, Proceedings of the 1st International Workshop on Digital Libraries Foundations (DLF1-07). Available:
http://www.ite.gr/ics/isl/publications/paperlink/DLF107.pdf

James Cheney, Carl Lagoze and Peter Botticelli. 2001. Towards a Theory of Information Preservation . In: Constantopoulos & I.T. Sølvberg (Eds.): ECDL 2001, LNCS 2163: 340–351, Springer-Verlag Berlin Heidelberg.

Linda Cantara. 2005. The *Tibet Oral History Archive Project* and Digital Preservation. ACH-ALLC 2005 conference. Available:
http://mustard.tapor.uvic.ca/cocoon/ach_abstracts/xq/xhtml.xq?id=141

OAIS. 2003. ISO 14721:2003 (Space data and information transfer systems – Open archival information system – Reference model). Available:
http://public.ccsds.org/publications/archive/650x0b1.pdf

Paul Watry. (2007). Digital Preservation Theory and Application: Transcontinental Persistent Archives Testbed Activity. International Journal of Digital Curation 2(2). Available:
http://www.ijdc.net/ijdc/article/view/43/0

PREMIS. (2008) Data Dictionary for Preservation Metadata: PREMIS version 2.0, Available:
http://www.loc.gov/standards/premis/v2/premis-2-0.pdf.

Reagan Moore and MacKenzie Smith. (2007). Automated Validation of Trusted Digital Repository Assessment Criteria. Journal of Digital Information 8(2). Available:
http://journals.tdl.org/jodi/article/view/198/181.