iPRES 2009

THE SIXTH INTERNATIONAL CONFERENCE ON THE PRESERVATION OF DIGITAL OBJECTS

Proceedings

October 5-6, 2009
Mission Bay Conference Center
San Francisco, California

California Digital Library

Significant Properties, Authenticity, Provenance, Representation Information and OAIS

David Giaretta[1], Brian Matthews[1], Juan Bicarregui[1], Simon Lambert[1], Mariella Guercio[2], Giovanni Michetti[2] and Donald Sawyer[3]

[1]Science & Technology Facilities Council, Rutherford Appleton Lab., Didcot, Oxon, UK
[2]University of Urbino, Italy
[3]VIE, Inc., USA

Abstract

The term "Significant Properties" has been given a variety of definitions and used in various ways over the past several years. The relationship between Significant Properties and the OAIS term Representation Information has been a puzzle.

This paper proposes a definition of Significant Properties which provides a way to clarify this relationship and indicates how the concept can be used in a coherent way. We believe that this approach is consistent with the actual use of the concept and does not invalidate the previous pieces of work but rather provides a clear and consistent view of the concept. It also links together Authenticity and Provenance which are also key concepts in digital preservation.

Introduction

The concept of Significant Properties is one which has been much discussed within the preservation community as a way of characterising, in some way, the essential features of a digital object which must be maintained over time. However, there is no consensus on a single definition of the concept of Significant Property or on how the Significant Properties are categorised and tested. OAIS [1], a widely recognised ISO standard concerned with systems which preserve digitally encoded information for the long-term, does not currently consider Significant Properties, but precisely defines a number of key concepts including Authenticity, Provenance and Representation Information and indeed one key function of OAIS has been to provide a consistent terminology. Clear and consistent terminology is a requirement for clear discussions about digital preservation. This leads to the question of how Significant Properties are related to these concepts and how Significant Properties can be adequately defined.

This paper attempts to draw together and provide a coherent view of Significant Properties, Authenticity, Provenance and Representation Information. This view clarifies the complementary roles of Significant Properties and Representation Information and their links to Authenticity. It also clarifies why the concept of Designated Community, which is so important in the OAIS, plays only a minor role, if any, in most discussions of Significant Properties. In addition this leads, in a consistent manner, to a way in which the concept of Significant Properties can be extended to scientific data. We also suggest why the digital library community stresses Significant Properties, while largely ignoring Representation Information, other than structure Representation Information (formats). We discuss the new proposed definitions as they exist in the draft revision of OAIS [2].

Significant Properties

The notion of Significant Properties has emerged as a key concept in preservation within the library community but has not been a concept that is much used in the context of the preservation of research data that is not normally viewed as a document. A number of definitions of Significant Properties have been proposed.

The CEDARS project [7] defined Significant Properties as

> *"those characteristics [both technical, intellectual, and aesthetic] agreed by the archive or by the collection manager to be the most important features to preserve over time".*

Sergeant [13] on the other hand proposed that

> *"Significant Properties are those attributes of an object that constitute the complete (for the intended Consumer) intellectual content of that object"*

However the example given of

> *Significant Properties for an e-thesis of the*
> - *complete text, including divisions into chapters and sections*
> - *the layout and style - particular fonts and spacing are essential*
> - *Diagrams*
> - *(perhaps web adverts are not Significant for our ejournals).*

does however seem more oriented to the rendering of the document in print or on screen, rather than its intended intellectual content. These could be consistent with OAIS concepts if the Designated Community (DC) had the appropriate knowledge base to understand the rendering, but would have problems if the knowledge base of the DC changes, for example if the language of the DC changes

from, say, English to, say, Chinese. There must be underlying Representation Information that supports the expression, or encoded value, of the Significant Properties listed for the e-thesis information object.

The OCLC/RLG Working Group on Preservation Metadata [12] proposed the definition:

> *"Properties of the Content Data Object's rendered content which must be preserved or maintained during successive cycles of the preservation process"*

Hedstrom and Lee [6] defined Significant Properties as

> *"those properties of digital objects that affect their quality, usability, rendering, and behaviour".*

In that paper they have a number of links to the OAIS Reference Model, for example

> *"whether or not colour, for example, is a significant property of the given digital object or collection will depend on the extent to which colour features affect the quality and usability of the preserved object for a designated community",*

and

> *"decisions about which Significant Properties to maintain will depend on institutional priorities, anticipated use, knowledge of the designated community, the types of materials involved, and the financial and technical resources available to the repository"*

Within the InSPECT project, Wilson [4] defines Significant Properties in a similar fashion 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".*

He categorises these Significant Properties into Content, Context, Appearance, Structure, and Behaviour. Knight [3] built on Wilson's work and proposes a framework of description for Significant Properties which includes identifier, function, level of significance, optionally the designated community, and optionally notes of any property constraints. The project has applied this to a number of digital object types (structured documents, raster images, audio files, email messages). Four further studies considered Significant Properties of vector images [19], moving images [20], software [21], and learning objects [22]. It is notable that each of these studies took a different view on what constituted a Significant Property. Again, here we have notions of Significant Property which cover some aspects of meaning (content and behaviour), although it is not clear how these are to be supported. Thus it can be seen that there is a lack of agreement on the definition of what a Significant Property is, what its primary role is, or how they should be categorised, recorded and tested.

Limitations of Significant Properties

Clearly the uses of Significant Properties of necessity focus on those aspects of digital objects which can be evaluated in some way and checked as to whether they have been preserved. However, the meaning associated with a value of the Significant Property is nowhere defined. Therefore it must be the case that the Significant Properties, while useful, do not strictly contribute to understandability of the Information Object. For example a Significant Property might be that a text character must be red, however the meaning of that "redness" is not defined. That is, can the following type of questions be answered: "what is the meaning implied by the fact that this test is red, and what would be difference be if it were not red"?

OAIS, as described further in section 3, proposes that to ensure preservation, one must have enough Representation Information to allow the defined Designated Community to understand the information, given that Designated Community's knowledge base. This must include, for example, information used to express the value (i.e. how it is encoded into bits) of a Significant Property or the values used to derive it. This is consistent with the comment

> *"As with file formats, the Representation Information for a digital object should allow the recreation of all the significant properties of the original digital object."*

from the PARADIGM project [23].

It should be noted that even those studies of Significant Properties which include Designated Community only have it as an optional item and for example [3] states "By leaving the Designated Community value blank, the archive declares that the property is, as far as they are aware, important for all user communities". Thus the stress in this definition is on importance. It leaves open the issue as to whether the value of that Significant Property is understandable to that very broad Designated Community.

Comparing the various definitions only [4] includes "meaning" in its definition, and therefore seems somewhat out of step with the other definitions; [13] includes what might be interpreted as a more ambitious phrase "complete (for the intended Consumer) intellectual content"; [4] is the only one to include "accessibility". Both [6] and [4] include "usability" in their definitions which is plausible but hard to see, for example, with "redness". The terms "appearance" and "experienced" is used in [5] while [6] includes "rendering" and "behaviour"; [12] refers to "rendered content" and, as noted above, the example in [13] makes it fairly clear that the rendering is the main concern.

With such a diversity of definitions and a seeming clash with the OAIS definition of preservation, what is the real purpose of Significant Properties? In order to explore

this, we first discuss a number of important related concepts which are identified within OAIS.

OAIS

The Reference Model for an Open Archival Information System (OAIS) [1] is an ISO standard primarily concerned with systems which preserve digitally encoded information for the long-term. There are many models specified in OAIS; in terms of the conformance to the standard the most important of these is the Information Model, which introduces the concept of Representation Information. It is noteworthy that the 2002 version of OAIS did not mention Significant Properties; however, OAIS does define a number of concepts which capture some related, and essential, ideas.

Capitalisation of terms in the text which follows normally indicates that the OAIS definitions are being used.

Information Model and Representation Information

The information model within OAIS provides the link between the encoded information and the Information Object itself. Representation Information is defined essentially as anything that is needed in order to understand the associated Data Object. OAIS provides a limited taxonomy of Representation Information made up of Structural, Semantic and Other Information.

In order for preservation to be testable the archive that is preserving the Data Object must define for whom it is being preserved. Without such a definition it would be impossible to test whether the preservation activities have been successful, as the archive could not establish sufficient criteria on the features required to be preserved relative to the knowledge of the community. This leads to the concept of the Designated Community and furthermore one sees that the amount of Representation Information needed is determined by the knowledge base of the Designated Community, another key concept introduced by OAIS. Digital preservation, in OAIS terms, requires that the Information Object, encoded as a Digital Object, must be understandable to the Designated Community. The Representation Information may need to be updated as the knowledge base of the Designated Community changes over time.

Authenticity

In addition to understandability and usability, good preservation also requires the concept of Authenticity. Much work has been done in the area of authenticity of digital objects, see for example within the InterPARES project [9]. OAIS (2002) [1] did not deal with Authenticity very well but the much improved working draft revision [2] defines Authenticity as:

> *"the degree to which a person (or system) may regard an object as what it is purported to be. The degree of Authenticity is judged on the basis of evidence".*

This definition is consistent with InterPARES and makes it clear that authenticity cannot be judged as a Boolean quantity, i.e. not simply true/false, and moreover it is not an absolute quantity. Instead it must be a question of evaluation by each user on the basis of evidence. For example even two bit-wise identical copies of a digital object may be judged to have different levels of authenticity if they have been in the custody of two different curators, handled with a different degree of accuracy and reliability.

The evidence on which the judgment of authenticity is made may consist of a mixture of technical and nontechnical elements. The former includes items such as digital digests, where the bit sequences are expected to have been unchanged (covered by the OAIS term Fixity). Examples of the latter type of evidence include such items as details of the curator of the data, which is an item of the Provenance of the Digital Object, supplemented by proof of identity of the curator (i.e. was the curator the person whom it is claimed to have been), plus perhaps evidence of the degree to which that person might be trusted.

A new situation arises if the bit sequences of the Content Data Object are changed (a Transformation in OAIS terms, a special type of Migration), in which case additional evidence must be supplied. Here the curator at the time of the Transformation is assumed to have taken steps, in particular performed some checks, to ensure that the new Content Information will not have lost important information and will continue to be judged to have the same level of authenticity as the object which was the input to the Transformation.

The question is the following: how did the curator come to that decision about the steps to take and what evidence can be supplied to future users to support authenticity? The following sections of this paper attempt to address those questions.

Authenticity Model

The complexity of the preservation function in the digital area requires the development of techniques able to ensure that the main elements and procedures relevant for the quality of the preservation are maintained, and that the authenticity of the preserved information objects can be evaluated. With reference to this crucial point, the CASPAR project [18] is documenting an Authenticity Model [14], and plans, as a reference solution, the implementation of an Authenticity Management Tool (AMT) with the ability to monitor and manage protocols and procedures across the custody chain. This tool will deliver, in the form of the documentation of the

preservation workflow, the benefits of authenticity into information systems, from the creation to the preservation phase. According to this approach the AMT has to identify mechanisms for ensuring the maintenance and verification of the authenticity in terms of identity and integrity of the digital objects. That is to provide content and contextual information relevant to the authenticity along the whole preservation process by capturing and making understandable over time all the required information.

The main components of are:

- the right attribution of authorship
- the identification of provenance in the life cycle of information objects
- the assurance of content integrity of the whole relevant digital components and their relevant contextual relationships
- the provision of mechanisms to allow future users to verify the authenticity of the preserved information objects or at least to provide the capability of evaluating their reliability in term of authenticity presumption.

In summary, every relevant aspect should be described and documented at every stage in the life cycle with the aim of providing, at any time, a sort of 'Authenticity Card' for any object in the repository.

The Authenticity Model as here outlined is based on the fact that the preservation is a process, which requires – for the assessment of its quality in terms of the authenticity presumption of any object in the repository – the definition of specific procedures, referred to as Authenticity Protocols (AP), and their documentation (in the form of Authenticity Reports and their Authenticity Protocol History). The evolution of an AP may concern the addition, removal or modification of any step making up the AP, and the change of the sequence defining the workflow. In any case both the old and the new step and/or sequence must be retained for documentation purposes. An example of managing authenticity of data objects on the basis of CASPAR model has been successfully implemented by the IBM Research Laboratory in Haifa in conjunction with the University of Urbino with reference to the data transformation as part of the storage component [15].

Provenance

Provenance is defined [2] as:

"The information that documents the history of the Content Information. This information tells the origin or source of the Content Information, any changes that may have taken place since it was originated, and who has had custody of it since it was originated. The archive is responsible for creating and preserving Provenance Information from the point of Ingest, however earlier Provenance Information should be provided by the Producer. Provenance Information adds to the evidence to support Authenticity."

OAIS also remarks that Provenance is a special type of Context. It should be clear from the discussion of authenticity that while Provenance provides only one kind of the evidence required, nevertheless it is a very important type of evidence.

Authenticity and Significant Properties

Given the logic that authenticity requires evidence, some of this evidence is technical, for example Fixity, and some of the many types of Provenance are non-technical, for example that they tell one how trustworthy an individual is or was regarded to be.

As noted above, if the bit sequences are unchanged then there are well established mechanisms for checking this although, of course, issues arise over the long term as to, for example, the security of any particular message digest algorithm. If however the bit sequences of the digital object are changed then these mechanisms are ineffective. For example a Word file may be converted to a PDF; in that case the bit sequences will have been changed extensively. In such cases the curator presumably would have satisfied himself or herself that the object as transformed had not lost required information content and therefore was still being adequately preserved. The curator would see the new object has continuing to maintain authenticity.

This may have been done by, for example, checking that the words were the same, comparing the Word file and the PDF file; that the rendering of the pages was reasonably consistent between the two versions; that text which had been emphasised in the Word version by highlighting or by changing colour, was also emphasized in some appropriate way in the PDF version.

It will be recognised that for the Word to PDF conversion the curator checked and documented various properties that are often called out as Significant Properties.

Thus it will be seen that the function of Significant Properties, consistent with Wilson in [4], is the identification of "those characteristics [both technical, intellectual, and aesthetic] agreed by the archive or by the collection manager to be the most important features to preserve over time". Wilson presents a related argument in [10]. Also Rothenberg and Bikson [5] suggest, with respect to authenticity criteria:

"the intent of these criteria is to ensure that preserved records retain their original behavior, appearance, content, structure, and context, for all relevant intents and purposes"

which echoes Significant Properties.

However the important point to note is that their real significance is that Significant Properties provide some of the evidence about the Authenticity of the digital objects after Transformation (a point emphasized by Wilson), and are selected by the curator, who may or may not take the Designated Community into account, and moreover different curators may make different choices. Wilson considers the notion of Performance as a test of the authenticity of preservation with respect to Significant Properties. This is an important feature, which as we shall see, can be transferred into a science data context.

Significant Properties and Data

Scientific data has yet to be dealt with in studies of Significant Properties. However some clarification may be gained by considering another Transformation, this time of a FITS[1] file converted to a CDF[2] file. Again the bit sequences will have been changed extensively. In such a case it could be asked how a curator could have satisfied himself or herself that the object as transformed had not lost required information content and therefore was still being adequately preserved. This is the way in which the curator would see the new object has continued to maintain authenticity.

The FITS file might contain an image; the CDF file should contain a similar image. However just comparing the two images rendered on screens would be inadequate for scientific purposes. Instead the curator would need to be satisfied, for example, that the data values of the pixel elements were identical in the two images at corresponding points; that the co-ordinates associated with each pixel in the two images were identical, for example the same latitude and longitude; that the units associated with the numerical values were the same in both images.

Science data is largely numerical or documentary. In a transformation the way in which the numbers are encoded may change, for example from an IEEE real to a scaled integer. In such a case a number in the old and the new formats should be the same to within rounding errors or predefined accuracy. Alternatively, co-ordinate system transformations may also require changes to the numerical values, which however should be reversible. Thus the validity of the transformation in preserving these significant data values is testable.

Alternatively the curator might simply document the fact that the trusted application, which was widely believed to maintain these numerical values, had been used in the transformation and thus implicitly those important values would automatically be the same in the two versions. In that case details of the tool would need to be available and the adequacy of its preservation of significant values can be evaluated. Thus in these two cases, we can identify how the Performance of the transformed format can be evaluated to test the Authenticity of preservation.

By analogy one can see that (some) Significant Properties of the data in this case are the pixel data values, units and the co-ordinate values. However of course this would not provide enough information to use the image. For example - what frequencies of light were collected, what instrument was used, when was the data collected. All these, and more, would be needed to understand and use that data and, unless very specific definitions of the Designated Community were made which included this knowledge, all would therefore be required to be described in appropriate Representation Information.

Significant Properties and Representation Information

Rendered objects such as JPEG images or audio files tend to be accompanied only by structural information; in the OAIS terms this is equivalent to stating that the knowledge base of the designated community includes whatever is needed to interpret the contents of the JPEG image or audio file; as this can be anything, the designated community is not explicitly defined. This is analogous to normal library practice where the onus is on the reader to understand the printed document. Scientific Data on the other hand tend to be numerical; even in the simplest case, where the numbers are encoded in a document as text, although it may be acceptable to assume that for an implicit designated community with a general knowledge of standard Arabic numerals in decimal notation, they will be able to understand that the sequence of characters '1' '2' means twelve. However it is not reasonable to assume that the implicit designated community will understand what the twelve signifies, for example 12° Centigrade or 12 metres or 12 apples (or even eighteen in hexadecimal). In order to fill in this missing information some Semantic Representation Information must be provided.

The normal library practice of ignoring, by default, Semantic Representation Information, has allowed Significant Properties, as usually considered without attention to meaning of their values, to appear to play a more general role in preservation, to the detriment of the full use of the Representation Information concept. It is only when looking at a broader class of digital objects, including scientific data and software, and a broader definition of preservation, that their true significance may be seen.

---

[1] FITS (Flexible Image Transport System) is the most commonly used digital file format in astronomy.

[2] CDF (Common Data Format) is a data abstraction defined by NASA and used to handle multidimensional datasets.

For any Significant Property some aspect of the information object has been encoded in a way which is described by Representation Information (often structural). However to be useful to the designated community the meaning associated with this property's value must also be available in their knowledge base. If the knowledge base changes then appropriate additions should be made in the information object's Representation Information to again ensure understandability by the designated community.

On the other hand the Representation Information of an information object by itself does not provide much direct guidance as to what Transformation to apply. The transformation will usually alter the digital object and certainly new Representation Information must be provided. Clearly one could check that any new digital structure provided the capabilities needed to support the semantics of the information object. However Significant Properties provide a much simpler, albeit perhaps incomplete, way of choosing an appropriate transformation, consistent with their use in a number of testbeds [11]. For example for a Word document, if the only Significant Property consists of the characters in the text (including spaces and new lines) then an appropriate transformation from Word to ASCII or UNICODE if the text contains more than Latin characters. A database (ACCESS or Postgres) for which the embedded queries are Significant Properties, then an appropriate transformation could be to MySQL whereas it is clear that transformation to a CSV file would not be appropriate. A FITS file which contains an astronomical image with Significant Properties including a World Coordinate System, an appropriate transformation could be to NetCDF, but additional information about the World Coordinate System would have to be added somehow.

In addition Significant Properties do provide hints on how the Designated Community has been defined (implicitly or explicitly) and what types of Representation Information must be present. In these ways the use of Significant Properties could supplement the role of Representation Information.

OAIS and Significant Properties

In the draft revision to OAIS [2] the term Significant Properties is not explicitly defined because it was felt that this would simply add to the already very disparate list of definitions. Instead a number of inter-linked definitions are provided, which are introduced here with some explanatory text. In particular a term, with a clear definition, **Transformational Informational Property** is introduced which provides a new term instead of Significant Property.

The Significant Properties concept, however loosely defined, leads one to think that there are "***Insignificant Properties***" i.e. properties which can be ignored from the preservation point of view. Therefore OAIS introduced the concept of an Information Property and its associated Information Property Description:

Information Property: That part of the Content Information as described by the Information Property Description. The detailed expression, or value, of that part of the information content is conveyed by the appropriate parts of the Content Data Object and its Representation Information.

and

Information Property Description: The description of the Information Property. It is a description of a part of the information content of a Content Information object that is highlighted for a particular purpose.

Having these definitions one can then go on to define the concept which the discussion earlier in this chapter suggests, namely something which comes into play when digital objects are transformed:

Transformational Information Property: An Information Property whose preservation is regarded as being necessary but not sufficient to verify that the Non-Reversible Transformation has adequately preserved information content. This could be important as contributing to evidence about Authenticity. Such Information Properties will need to be associated with specific Representation Information, including Semantic Information, to denote how they are encoded and what they mean. (Note that the term 'significant property', which has various definitions in the literature, is sometimes used in a way that is consistent with it being a Transformational Information Property).

Note that if the Transformation were reversible then it is reasonable to take it that no information is lost. It is for this reason that the above definition focuses on non-reversible transformations. For completeness the definitions of the reversible and non-reversible transformations are as follows:

Reversible Transformation: A Transformation in which the new representation defines a set (or a subset) of resulting entities that are equivalent to the resulting entities defined by the original representation. This means that there is a one-to-one mapping back to the original representation and its set of base entities.

Non-Reversible Transformation: A Transformation which cannot be guaranteed to be a Reversible Transformation. The important point is that the definition of *nonreversible* is drawn as broadly as possible.

Summary and Conclusions

Examining the various definitions of Significant Properties and comparing them with OAIS concepts, and then looking at scientific data, has allowed us to present what we believe is a clearer view of what underlies the various attempts at defining a Significant Property. We then put into context the new OAIS term for this

underlying concept, namely **Transformational Information Property**. This new term was introduced in order to avoid creating yet another definition of something called Significant Property.

The role of a collection of these in preservation, namely as partial evidence for Authenticity, and as guidance for the choice of Transformations, is then, we believe, consistent and clear. Their relationship to Representation Information has also been made clear.

It is hoped that this new term will be taken up by the digital preservation community in order to put the existing work on what has been termed Significant Properties into a consistent context, to see how such work might be extended to scientific data, to allow clearer exchange of ideas and better digital preservation activities.

References

[1] CCSDS. (2002), Reference model for an Open Archival Information System (OAIS). Retrieved from: http://public.ccsds.org/publications/archive/650x0b1.pdf

[2] OAIS update (at the time of writing under CCSDS review), http://public.ccsds.org/sites/cwe/rids/Lists/CCSDS%206500P11/Attachments/650x0p11.pdf

[3] Knight, G., 2008, Framework for the definition of significant properties. Retrieved from http://www.significantproperties.org.uk/documents/wp33-propertiesreport-v1.pdf

[4] Wilson, A., 2007, Significant Properties Report. Retrieved from http://www.significantproperties.org.uk/documents/wp22_significant_properties.pdf

[5] J. Rothenberg and T. Bikson, 1999, 'Carrying Authentic, Understandable and Usable Digital Records Through Time' report to the Dutch National Archives and Ministry of the Interior. Retrieved from http://www.digitaleduurzaamheid.nl/bibliotheek/docs/final-report_4.pdf

[6] M. Hedstrom and C.A. Lee, "Significant properties of digital objects: definitions, applications, implications", Proceedings of the DLM-Forum 2002. Retrieved from http://ec.europa.eu/transparency/archival_policy/dlm_forum/doc/dlm-proceed2002.pdf

[7] Cedars project, http://www.leeds.ac.uk/cedars/

[8] Investigating the Significant Properties of Electronic Content over time (InSPECT) http://www.significantproperties.org.uk/

[9] The InterPARES project, http://www.interpares.org/

[10] Wilson, A., 2008, Significant Properties of Digital Objects, presented at "What to preserve? Significant Properties of Digital Objects". Retrieved from http://www.dpconline.org/docs/events/080407sigpropsWilson.pdf

[11] DELOS Digital Preservation Testbed. Retrieved from http://www.ifs.tuwien.ac.at/dp/testbed.html

[12] OCLC/RLG Working Group on Preservation Metadata, 2002, Preservation Metadata and the OAIS Information Model, A Metadata Framework to Support the Preservation of Digital Objects. Retrieved from http://www.oclc.org/research/projects/pmwg/pm_framework.pdf

[13] Derek Sergeant, 2002, Interpretation of the OAIS Model. Retrieved from http://www.erpanet.org/events/2002/copenhagen/presentations/dmserpanet.ppt

[14] CASPAR Access Model (especially section 2) http://www.casparpreserves.eu/Members/cclrc/Deliverables/report-on-oais-access-model/at_download/file

[15] Michael Factor, Ealan Henis, Dalit Naor, Simona Rabinovici-Cohen, Petra Reshef, Shahar Ronen, IBM Research Lab in Haifa, Israel and Giovanni Michetti, Maria Guercio, University of Urbino, Authenticity and Provenance in Long Term Digital Preservation: Modelling and Implementation in Preservation Aware Storage, TaPP '09. First Workshop on the Theory and Practice of Provenance. San Francisco, 23 February 2009, http://www.usenix.org/event/tapp09/tech/full_papers/factor/factor.pdf

[16] CASPAR Conceptual Model, http://www.casparpreserves.eu/Members/cclrc/Deliverables/caspar-conceptual-model-phase-1-1/at_download/file

[17] Giaretta, D., 2007, The CASPAR Approach to Digital Preservation, The International Journal of Digital Curation, Issue 1, Volume 2, http://www.ijdc.net/index.php/ijdc/article/viewFile/29/18

[18] CASPAR – Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval. See http://www.casparpreserves.eu

[19] Mike Coyne, David Duce, Bob Hopgood, George Mallen, Mike Stapleton. The Significant Properties of Vector Images. JISC report, 27 November 2007. http://www.jisc.ac.uk/media/documents/programmes/preservation/vector_images.pdf

[20] Mike Coyne, Mike Stapleton. The Significant Properties of Moving Images. JISC report, 26 March 2008. http://www.jisc.ac.uk/media/documents/programmes/preservation/spmovimages_report.pdf

[21] Brian Matthews, Brian McIlwrath, David Giaretta, Esther Conway. The Significant Properties of Software: A Study. JISC report, March 2008 http://www.jisc.ac.uk/media/documents/programmes/preservation/spsoftware_report_redacted.pdf

[22] Kevin Ashley, Richard Davis, Ed Pinsent. Significant Properties of E-learning Objects. JISC report, March 2008. http://www.jisc.ac.uk/media/documents/programmes/preservation/spelos_report.pdf

[23] PARADIGM project, Workbook on Digital Private Papers. http://www.paradigm.ac.uk/workbook/preservationstrategies/