Data without meaning: Establishing the significant properties of digital research

Gareth Knight[1] and Maureen Pennock[2]

1. Centre for e-Research, Kings College London, 26 - 29 Drury Lane, London, WC2B 5RL
2. UKOLN, University of Bath, Bath, BA2 7AY
gareth.knight@kcl.ac.uk; m.pennock@ukoln.ac.uk

Abstract

It is well recognised that the time period in which digital research may remain accessible is likely to be short in comparison to the period in which it will have intellectual value. Although many digital preservation strategies are effective for simple resources, it is not always possible to confirm that all of the significant properties – the characteristics that contribute to the intended meaning – have been maintained when stored in different formats and software environments. The paper outlines methodologies being developed by InterPARES, PLANETS and other projects in the international research community to support the decision-making process and highlights the work of four recent JISC-funded studies to specify the significant properties of vector images, moving images, software and learning objects.

Introduction

In recent years, there has been a growing awareness of the need for digital preservation to maintain access to digital research. Unlike physical artefacts, it is considered to be infeasible to store digital data in its original form and expect it to be readable and usable over time [6]. Instead, there is an expectation that the environment in which digital records are accessed will change on an ongoing basis, e.g. as a result of updates to the computer hardware, operating system, or application software in use [24]. Institutions with a commitment to maintain digital research may adopt several digital preservation strategies, such as format conversion (normalisation, migration), emulation of the original hardware and software and, for certain types of data, re-implementation according to an existing specification. This paper will introduce the concept of significant properties and its role in maintaining the authenticity of research data across changing technological environments over time. It will highlight criteria for the evaluation of significant properties, through consideration of the requirements of those that have an investment in the availability and use of digital research. It will subsequently highlight work that has or is being performed to assist institutions with the task of understanding and evaluating significant properties. A final section provides a comparative analysis of the significant properties of vector images, moving images, software and learning objects that were identified by four recent JISC-funded studies.

Definitions of significant properties

The term 'significant properties'[1] was first used by the CEDARS Project [5] and has been interpreted using several different, but broadly consistent definitions [7]. For the purpose of this paper, significant properties are defined as the characteristics of an information object that must be maintained to ensure its continued access, use, and meaning over time as it is moved to new technologies [24]. The term is widely used in the archival community, where it is associated with authenticity (that it is what it purports to be) and integrity (that it has not been changed or corrupted in a manner that has caused the original meaning to be lost) [24, 9, 3]. Significant properties share some similarities with Representation Information and there is some crossover between the two concepts. In an OAIS, significant properties are the characteristics of the abstract information object (e.g. an image), while representation information indicates characteristics of the data object (e.g. format, encoding scheme, algorithm) [2].

Research on the topic of significant properties

The importance and position of significant properties in developing digital preservation strategies has been recognised by several parties over the past decade. The following list is not intended to be exhaustive, rather an illustration of the projects that have made an important contribution to the development of our understanding of significant properties:

- *CEDARS (Curl Exemplars in Digital ARchiveS)*: the JISC-funded CEDARS project (1998-2002) explored several digital preservation issues, including significant properties. The project defined the 'Underlying Abstract Form', an abstract model for preserving 'all the necessary properties of the data'[5].
- *Digital Preservation Testbed*: Complementary research took place in the Dutch Digital Preservation Testbed project (2000 – 2003) testing the viability of

[1] essence, essential characteristics, core features, properties of conceptual objects are other synonyms that are used in particular domains and institutions.

different preservation approaches for different types of government archival digital records. The research was based on the assumption that different types of records have different preservation and authenticity requirements [18]

- *National Archives of Australia*: The NAA developed the concept of the 'essence' as a formal mechanism to determine the characteristics that must be preserved and a 'Performance model' to demonstrate that digital records are not stable artefacts; instead they are a series of performances that change across time [14].
- *DELOS*: The preservation cluster in the EU-funded DELOS Network of Excellence in Digital Libraries built on the work of the Testbed project and developed a metric for testing and evaluating digital preservation strategies using utility analysis and an Objective Tree [19].
- *PLANETS*: PLANETS is an EU-funded project that is undertaking several projects that have relevance to the description of significant properties, including the continued development and integration of the DELOS Utility Analysis and Objective Tree into the PLATO Preservation Planning Tool and the creation of the eXtensible Characterisation Definition/Extraction Language (XCDL/XCEL) [17].
- JISC-funded Significant Properties projects: the JISC has funded four short projects to investigate the significant properties of vector graphics, moving images, learning objects and software that have produced some useful outputs [10].
- InSPECT Project: InSPECT is a JISC-funded two-year project performed by the Centre for e-Research at Kings College London and The National Archives. It is building on the work performed by the National Archives of Australia and Digital Preservation Testbed to develop a framework for the definition and description of significant properties, which will be integrated into the PRONOM format registry [12].

Although each project has a distinct conceptual basis and methodology, the outputs of earlier work has contributed to the development of subsequent projects.

Criteria for evaluating significant properties

An implicit assumption in the use of terminology, such as 'significant' and 'essential' is the recognition that criteria is required against which the relative value of each property may be assessed. The Oxford English Dictionary defines 'value' as a noun to be 'a fair or adequate equivalent or return'. In diplomatics a distinction is made between 'intrinsic value' - that something has value 'in its own right'- and 'extrinsic value' - that value is derived from an external function. The InterPARES Authenticity Task Force has hypothesised that both intrinsic and extrinsic elements will play key roles in establishing the identity of a digital record [15]. For digital objects, value judgments made by an archivist or collection manager will determine the level of functionality that is retained in subsequent iterations of the object. It is therefore important to identify the potential stakeholders and understand the functions that will be required of the information object and the environment in which it will be used, as criteria for evaluating alternative preservation strategies [20, 5, 21]

The InSPECT project [12] has analysed several elements that may influence an institutions interpretation of value and, as a result the preservation activities that must be performed to maintain the various properties of the information object. These may be summarized into four categories:

1. Stakeholder requirements

The stakeholders represent the intended audience for the digital object. The consideration of the required functionality that an Information Object should provide must consider several stakeholders during its lifecycle. These may include:

1) The creator who produced the resource to fulfil specific aims and objectives in the short-term. For example, a paper written for publication.
2) Researchers in the designated community who wish to use the resource as the basis for further analysis and discussion, e.g. scientists, artists.
3) Tutors who wish to incorporate the resource into a learning object for use in teaching [1]

In addition a digital curator should be aware of their own requirements:

4) A curatorial institution that wishes to maintain an authentic copy of the resource for the purpose of curation and preservation.

The functionality required by each stakeholder may differ and change over time, influenced by aims and objectives directly defined by the stakeholder or imposed by business requirements (e.g. legal status, basis for funding, mandate, institutional policy of other stakeholders). Although a full analysis is required, it is reasonable to suggest that some or all stakeholders will require the digital object to be authentic. Each stakeholder will have different criteria for evaluating authenticity, which is influenced by the context of their work. For example, the InterPARES project [15] notes that the authenticity requirements for legal records are strict which requires the adoption of a risk-adverse strategy to preservation. In comparison, the authenticity requirements for a funding body may be much lower, limited to the requirement to maintain the intellectual content of the resource only [21]. A second function that may be required is the ability to use and modify content by the creator or a third-party, in addition to the ability to access it. For example, the ability to search and edit a spreadsheet, database, and word processing document have

been cited as potential useful functions that support the activities of financial institutions [19, 21].

2. Type of resource

The method in which a Creator first expresses an idea and renders it in a form that can be understood by others has an influence upon the properties that are considered to be significant. The creation process may be influenced by the design preferences of the Creator (e.g. an idea expressed as a page of text, a spider diagram, or audio recording), the software tools available, as well as consideration of the access method for a target audience. To illustrate the distinction between object types, a report may be written for communication in an email or a word processing document. Both will have common properties that are specific to the form of expression (words organised into paragraphs) and the method of embodiment (e.g. title may be indicated in subject line of an email or document body). However, an email will require additional properties to record details of the recipient.

3. Legal right

The copyright of digital research may be owned by one or more stakeholders. An institution with a commitment to curate and preserve the significant properties of a digital resource may be limited in its actions by the legal rights that have been assigned to it, which will limit the range of properties that it is capable of maintaining. For example, a research paper may contain text and images owned by the author that may be reproduced in a different format and typographical features owned by a publisher that cannot be reproduced [22].

4. Capability

Finally, the ability of the curator to perform preservation action for digital research may be influenced by the total money, time and resources available for the identification and evaluation of properties. The institution may have possess sufficient finances to purchase or develop a software tool to perform a data analysis; to allocate staff time to the identification of significant properties; and/or validate that they have been maintained in subsequent manifestations.

The creation of a definition of significance encompasses a range of qualitative requirements that may be unique to each institution. The PLANETS PLATO tool [17] may prove useful through the provision of a baseline set of characteristics that can be tailored to the requirements of each institution.

Framework for the evaluation of significant properties

The creation of a framework for the identification and analysis of significant properties has been a key area for research in recent years. The work of Rothenberg & Bikson [21], DELOS [19] and the InterPARES projects [15] has been particularly influential in this area. The following section provides a description of two frameworks, Digital Diplomatics and Utility Analysis that may assist curators to interpret the properties of digital research that must be maintained.

Digital Diplomatics

Digital diplomatics is the application of archival diplomatics to digital records, which was developed for use in the InterPARES1 project. The process emerged in the seventeenth century as a method for determining the authenticity of a physical record for legal purposes. On the basis of the examination, it may be possible to establish if the document was created at the time and place that is claimed. In comparison to other methodologies, their analytical method places a greater emphasis on the intended function (e.g. a legal document) that the record must perform as a basis for defining the significant properties. The InterPARES project indicates that many authenticity requirements are created and managed at an organizational level, and therefore cannot be entirely understood at the record-level. I To demonstrate the application of diplomatics to digital records, they indicate that properties may be organized into four categories:

1. *Documentary form*: The elements that establish its authority in an administrative or documentary context. These are separated into intrinsic and extrinsic elements. Intrinsic elements specify the context in which the record exists. For example, details of the creator, intended recipient, date of creation, and aspects that communicate the activity in which it participates. Extrinsic elements refer to the perceivable features that are instrumental in achieving an intended purpose. For example, the overall presentation of the intellectual content (text, image, sound), presentation features specific to the record (e.g. special layouts, hyperlinks, colours, sample rate), electronic signatures, digital time stamps and other 'special signs' (watermarks, an institution's logo).
2. *Annotations:* The aspects of the record that have been augmented after its creation. For example, additions made as part of: its execution (datetime that an email was transmitted, indication of attachments); its handling in relation to its intended use (comments embedded in the record that critique the work); and its handling for records management purposes (identifier, version number, cross reference to other records).
3. *Context:* the broader framework in which the record is created and managed. For example, judicial-administrative, documentary and technological context.
4. *Medium:* Diplomatic analysis specifies the medium on which information is stored as an essential element. However, the InterPARES Authenticity Task Force indicates that an analysis of the medium is transitory and may be an unnecessary consideration for many digital records.

The classification of different aspects of a digital object is a useful stage in the evaluation of the aspects that should be considered significant, in relation to one or more intended functions. However, the use of archival diplomatics as an analytical tool imposes certain well recognised limitations on the type of information that is considered to be significant. Specifically, there is an emphasis on textual elements of agents associated with the creation, augmentation and management process. The project has also noted the requirements for 'fixed form' records, which excludes certain types of dynamic data [15]. The approach taken by the InterPARES1 project in establishing the contextual basis for decisions at an organisational-level is useful, but further work is necessary, potentially based on less strict compliance with archival diplomatics analysis.

Utility Analysis

The preservation cluster in the EU-funded DELOS project built on the work of the Testbed project to develop a metric to test and evaluate digital preservation strategies, based on the conceptual Utility Analysis and Objective tree [19,17]. The metric may be used to define objectives and evaluate the results of preservation activities. The Utility Analysis model specifies eight stages (figure 2)

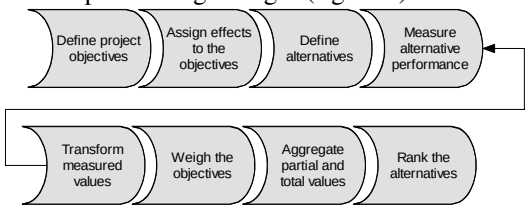

Figure 2: the eight steps of the DELOS Utility model

In the DELOS Utility Analysis and Objective tree, significant properties of digital objects are one of several factors that must be considered when defining and subsequently evaluating objectives. They may be divided into two major groups: 'file characteristics' that indicate the aspects of the digital object that must be maintained (e.g. horizontal and vertical dimensions of an image, frame rate of moving image) and 'process characteristics' that describes the objectives with which the resulting digital object must comply (e.g. authentic recreation of the significant properties, scalability, error-detection, usability, and others). The metrics developed in DELOS may be used to automatically weigh the performance of a given approach in preserving specific characteristics of records and the numerical evaluation of preservation strategies is consider to be a step towards the automation of the evaluation process.

To demonstrate their approach the project carried out two case studies [19], indicating the requirements of a word processing document and an audio file. The analysis of the file characteristics in a word processing document identified a number of properties that must be maintained, including various aspects of the content (body text, embedded images, foot notes, page numbering), page layout (paragraphs, page margins, page breaks) and function of the creating application (Microsoft Word). The latter is surprising, but is supported by earlier work by Rothenberg & Bikson [21]. In terms of the process characteristics, the ability to track changes and search the document was considered to be significant. The criteria was subsequently used as a basis for evaluation of suitable file formats, indicating that the most suitable format to contain the 'file characteristics' and 'process characteristics' was another version of Microsoft Word. Whilst the high score may be due to fundamentally necessary compatibility between the source and target file formats, some would consider this an undesirable route in terms of format longevity. It is clear that any attribution of measured value can be subjective and is not necessarily transferable to other situations; different organisations with different baseline requirements will likely allocate different values to different properties and thus result in different final scores from the evaluation process.

The PLANETS project builds on the Utility analysis work by integrating it into the PLATO Preservation Planning Tool, a web-accessible system for measuring and evaluating the performance of preservation activities against stated requirements and goals. (http://www.ifs.tuwien.ac.at/dp/plato/) The project has defined four main groups of characteristics: object, record, process and costs. In recognition that requirements vary across settings, it is recommended that as many stakeholders as possible are involved in the definition of requirements, from producers, curators and consumers to IT staff, domain experts, managers, and lawyers. The tool is still in development and will eventually integrate with registries and services for file format identification, characterisation and preservation actions.

Analysis of significant properties studies

In recent years it has become increasingly evident that a renewed study on the topic of significant properties was necessary, to gain a better understanding of the significant properties of various object types that institutions must maintain. To address this need the JISC funded the InSPECT project and four studies that would investigate the significant properties of several object types, including vector images, moving images, learning objects and software. These projects have been informed by the 'Performance model' and associated methodology created by the National Archives of Australia [14], as well as related work that has been performed previously.

Although the various significant properties studies share a common objective, they each developed specific

methodologies for the identification and interpretation of significant properties, partially based on archival diplomatics, utility analysis, records management and other discipline specific standards (e.g. the SPeLOs [1] project was informed by web-based e-learning practices and the Significant Properties of Vector Images study [7] was influenced by the Computer Graphics Reference Model).

One of several recommendations identified during the course of a workshop on the topic of significant properties was that the outputs of these projects should be mapped onto a common model to identify similarities and differences [11]. The final section of this paper will provide a comparison of the significant properties identified by the four recent JISC-funded studies. This work will enable the recognition of common themes between different objects based on their complexity (e.g. a software package and a raster image) and content type (still images, moving image). In addition, the outputs of each study may be merged to correct shortfalls in the coverage of each study. For example, the analysis of composite objects, such as Learning objects may be informed by analysis at a lower level, through use of the outputs of the studies into moving images or sound [12].

To begin to analyse the significant properties of the objects a conceptual framework is required. The study on the Significant Properties of Software [16] recognised the FRBR (Functional Requirements for Bibliographic Records) as being potentially useful for analysing different layers of a resource. FRBR is a conceptual entity-relationship model that represents the 'products of intellectual or artistic endeavour' at four layers of analysis: Work, Expression, Manifestation, and Item. In practical use, these layers may be equated to a Record, version of the Record, a variant of the version (e.g. an moving image object saved as an AVI and MPEG2; two variants of software compiled for Microsoft Windows and Linux); and Object that represents a single example of the work (e.g. a AVI file located on a user computer). However, to use the FRBR model as a basis for analysing significant properties, we must introduce a fifth entity, Component that represents one or more constituent parts of an object (e.g. an audio bit-stream in a moving image; a file in a software or learning object package).

1. Record

The Record is the top-level entity that equates to FRBR Work, The National Archives' concept of a Record [23], or software 'Package'. Several elements may be identified that indicate the significant properties for the Record entity in the studies on software [16], learning objects [1] and moving images [8] that describe the digital resources.

	Software	*Learning Objects*	*Moving Images*	*Vector Images*
Context	package name, keywords, purpose, Functional Requirement,	learning object classification, contextual, creator/Contributor, Description (Interactivity level, type, keywords) Educational Context, Metadata (catalogue type, references, subjects)	title	-
Context: Rights	provenance/ owner	Rights management		

Table 1: significant properties for the Record/Work entity

The information specified for the Record entity is informed by an archival diplomatics and records management methodology. The metadata is useful for establishing the chain of custody and provenance of the digital resource and may assist with its location and retrieval in a digital archive. However, it is provided for the purpose of completeness and is not considered to be relevant for the purpose of preservation to maintain access to the digital resource, in part or whole.

2. Expression / Version

A FRBR Expression is a realization of the intellectual work in a specific form. This may equate to different versions of an object containing updated or changed content (e.g. a learning object that is used for teaching in 2008 and later modified for the same course in 2009) or functionality (e.g. a software package that provides a new user interface, import/export option, or other features). Matthews et al (2008) identifies 17 entities that may be recorded for each software version. In addition, descriptive information in the Learning Objects and Moving Images study may be identified that are relevant for each version of an object.

	Software	*Learning Objects*	*Moving Images*	*Vector Images*
Context: descriptive	Version identifier, Functional description, Input format, output formats, Description of the algorithm used, API description, Software specification	LO classification, Educational context, Validator record, Author record, Creation date, Title, Learning Assembly	Title	-
Context:rights	Licence	Digital		

		Rights management		
Technical Environment	Software dependencies, Architectural dependencies, Hardware dependencies			

Table 2: significant properties for the Expression entity

The properties that are attributed to the Expression share a common theme, indicating specific contextual information that describes the function for which it has been created (e.g. a learning object for use in learning and teaching; a software tool for creation and processing of data) and its use by a Designated Community. The list of associated items specified in the study of software is not considered to be significant properties. However, the existence of documentation is a key component in understanding a software tool and recompiling or re-implementing it for a different environment[2].

3. Manifestation

A FRBR Manifestation is the embodiment of an expression in a particular medium or format. For example, the encoding of a moving image resource in the Apple Quicktime format or as a series of TIFF images, or the compilation of software code for Microsoft Windows or Linux systems. It is likely that Representation Information will be created for each manifestation, to interpret and render the digital resource in an appropriate technical environment. In the context of significant properties, the studies of Software and Learning Objects have identified several properties that may be categorised with the Manifestation entity:

	Software	*Learning Objects*	*Moving Images*	*Vector Images*
Context: description	Variant notes	Learning Unit classification, Digital object datatypes, reusability		-
Context: rights	licence			
Structure	software dependencies; configuration (software)	Delivery		-
Behaviour		Look and feel Delivery		
Technical Environment	platform (software); operating system (software).	interoperability		

[2] It is less common for researchers to create similar documentation for other types of digital object. Digital archives, such as the UK Data Archive and the Arts & Humanities Data Archive recommend that resource creators document the digital outputs that they produce.

Table 3: Significant properties for the manifestation entity

The Manifestation properties describe the technical composition of the digital resource. At this level of analysis, there is the potential for confusion between Representation Information and Significant Properties. Notably, the classification of environment properties is a matter for discussion, particularly in relation to software packages. However, other elements are simpler to interpret as a significant property. The 'Look and Feel' and 'Reusability' elements incorporate aspect of the technical composition, but use them as the basis for specifying the allowed usage of the digital resource.

4. Item

An FRBR Item is a single instance of a manifestation. For example, a learning object or software package that is stored in a digital repository or on a user's computer. It is equivalent to a software 'Download' or 'installation' [16]. A recipient may be provided with Representation Information to support its rendering and use or a description of significant properties to describe the content of the digital object. The majority of information provided with an item will have been created for each manifestation and, as a result will not require description at the item level. However, some object types may require the recording of information that indicate the digital rights and usage of the digital object in a specific environment (table 4).

	Software	*Learning Objects*	*Moving Images*	*Vector Images*
Content	-	-	No. of streams	-
Context	Licensee, Conditions, Licence code		Creation date	-
Structure	File relationships	Relationship between constituent parts (files, metadata)	Relationship between constituent parts (bitstreams)	-
Technical Environ-ment	Environment variables, IP address, Hardware address			

Table 4: Significant properties for the Item entity

The Significant Properties of Software study has identified six properties that are distinct from those specified for the Expression or Manifestation entities. These indicate the licensee that is the user of the software; an individual licence tailored to the use of the particular item and user;

and hardware and software configurations that are distinct to the environment in which it will be used (e.g. the software can be used only if a specific IP or MAC address is defined). Similar requirements are not specified in the remaining three significant properties studies, though it is theoretically possible that a Learning Object, moving image, or vector image could be imprinted with a watermark or digital signature that is linked to a specific user. The location of the rights and environment properties is a matter for discussion. Although the study indicates that the properties are significant at the item-level, it may be better represented as a manifestation that has been tailored to the requirements of a specific user.

5. Component

A Component represents a unit of information that forms a logical group. The term is used by The National Archives [23], InSPECT [12] and Significant Properties of Software [16] projects to represent one or more sub-sections that, when aggregated and processed correctly will form the Item as a whole. It may be applied to several artefacts, including an audio bit-stream in a moving images file, a text paragraph in a HTML page and a shape in a vector graphics diagram. Significant properties that are defined for the component entity describe characteristics of the information content or the environment in which the content may be reproduced [13]. Each of the four studies identify information specific to the content type that they were responsible for analyzing:

	Software	*Learning Objects*	*Moving Images*	*Vector Images*
Content	-	text	duration	text
Context	functional description input format, Output format, Program language, Interface, Error handling			
Structure		-		
Behavi-our		-		
Render-ing	Algorithm	Text (format, character encoding, layout, fonts, colour) Animation (colours, frame rate, speed)	Gamut, Frame height, frame width, pixel aspect ratio, frame rate interlace	point, open path, closed path, , object, inline object, shape
Behav-iour		-		
Tech Environ-ment	hardware depend-encies, library depend-encies, package depend-encies	-	compressi on ratio, codec	

Table 5: Significant properties for the Component entity

The component entity is key for maintaining access to and use of the information object. The projects have recognized a range of technical properties that perform similar functions for each object type – recreation of the text, raster image and vector image of the object. However, it is questionable if elements classified under the Environment heading are properties of the information object or data object.

Conclusion

This paper has provided a definition of significant properties and outlined their role in a digital preservation strategy. It has highlighted criteria for their evaluation, through consideration of the requirements of those that have an investment in the availability and use of digital research, as well as work being performed in the international digital preservation community to assist institutions with the task of understanding and evaluating significant properties.The review of projects and institutions that have made some contribution to the development of digital preservation strategies suggests that there is a great interest in the identification, analysis and extraction of significant properties. However, the distinct methodologies adopted by each JISC project suggest that further work is necessary to encourage adoption of the Utility Analysis and Digital Diplomatics methodologies. The mapping of the significant properties to the FRBR entity-relationship model proved to be a useful exercise for understanding the disparate approaches taken by project and has highlighted similarities and differences between the properties for each object type. On the basis of the results obtained, it is evident that there remains some difference in the understanding of properties that may be categorized as significant for the information object and those that may be classified as Representation Information and that further work is necessary to map the significant properties of an information object onto a conceptual and practical model in a consistent manner.

We have yet to reach the stage where a researcher or academic in an institution is able to define the significant properties of their digital research without ambiguity. It is expected that ongoing work being performed by InSPECT, PLANETS and CASPAR and other projects will provide a common methodology and tools for understanding significant properties. In particular, work should be performed that maps the significant properties of an information object onto a conceptual and practical model in a consistent manner.

References

[1] Ashley, K. Davis, R & Pinsent, E. 2008. Significant Properties of e-Learning Objects (SPeLOs), v1.0. http://www.jisc.ac.uk/whatwedo/programmes/programme_preservation/2008sigprops.aspx

[2] Brown, A. 2008. White Paper: Representation Information Registries. http://www.planets-project.eu/docs/reports/Planets_PC3-D7_RepInformationRegistries.pdf

[3] Bearman, D. & Trant, J. 1998. Authenticity of Digital Resources: Towards a Statement of Requirements in the Research Process, *D-Lib Magazine*, June 1998. http://www.dlib.org/dlib/june98/06bearman.html

[4] CASPAR Project. n.d. Cultural, Artistic and Scientific Knowledge for Preservation, Access and Retrieval. http://www.casparpreserves.eu/

[5] Cedars Project. 2002. Cedars Guide To : Digital Collection Management. http://www.leeds.ac.uk/cedars/guideto/collmanagement/

[6] Chen, Su-Shing. 2001. The Paradox of Digital Preservation. Computer, vol. 34, no. 3, pp. 24-28, Mar., 2001. http://doi.ieeecomputersociety.org/10.1109/2.910890

[7] Coyne, M et al. 2007. The Significant Properties of Vector Images. http://www.jisc.ac.uk/whatwedo/programmes/programme_preservation/2008sigprops.aspx

[8] Coyne, M. & Stapleton, M. 2008. The Significant Properties of Moving Images. http://www.jisc.ac.uk/whatwedo/programmes/programme_preservation/2008sigprops.aspx

[9] Digital Preservation Testbed. 2003. From digital volatility to digital permanence: Preserving text documents. http://www.digitaleduurzaamheid.nl/index.cfm?paginakeuze=185

[10] Grindley, N. June 2008. The Significant Properties of Digital Objects. http://www.jisc.ac.uk/whatwedo/programmes/programme_preservation/2008sigprops.aspx

[11] Hockx-Yu, H & Knight, G. 2008. What to Preserve?: Significant Properties of Digital Objects. *International Journal of Digital Curation*, Vol 3, No 1 (2008) http://www.ijdc.net/ijdc/article/view/70

[12] Knight, G. 2008a. Framework for the Definition of Significant Properties. http://www.significantproperties.org.uk/outputs.html

[13] Knight, G. 2008b. Significant Properties Data Dictionary. http://www.significantproperties.org.uk/outputs.html

[14] Heslop, H. Davis, S. Wilson, A. 2002. An Approach to the Preservation of Digital Records. http://www.naa.gov.au/Images/An-approach-Green-Paper_tcm2-888.pdf

[15] MacNeil, H. et al. Authenticity Task Force Report. http://www.interpares.org/book/interpares_book_d_part1.pdf

[16] Matthews, B. et al. 2008. The Significant Properties of Software: A Study, Version 5.7. http://www.jisc.ac.uk/whatwedo/programmes/programme_preservation/2008sigprops.aspx

[17] PLANETS Project. n.d. Preservation and Long-term Access through NETworked Services. http://www.planets-project.eu/

[18] Potter, M, 'Researching Long Term Digital Preservation Approaches in the Digital Preservation Testbed', RLG DigiNews (June 2002) - http://digitalarchive.oclc.org/da/ViewObjectMain.jsp?fileid=0000070519:000006287741&reqid=3550

[19] Rauch, C. Strodl, S. & Rauber, A. 2005. Deliverable 6.4.1: A Framework for Documenting the Behaviour and Functionality of Digital Objects and Preservation Strategies. http://www.dpc.delos.info/private/output/DELOS_WP6_d641_final__vienna.pdf

[20] RLG. 2002. Trusted Digital Repositories: Attributes and Responsibilities. http://www.oclc.org/programs/ourwork/past/trustedrep/repositories.pdf

[21] Rothenberg, J. & Bikson, 1999. T. Carrying Authentic, Understandable and Usable Digital Records Through Time: Report To the Dutch National Archives And Ministry of the Interior. http://www.digitaleduurzaamheid.nl/bibliotheek/docs/final-report_4.pdf

[22] The British Academy & The Publishers Association. 2008. Joint Guidelines on Copyright and Academic Research Guidelines for researchers and publishers in the Humanities and Social Sciences. http://www.britac.ac.uk/reports/copyright-guidelines/index.html

[23] The National archives, n.d. http://www.nationalarchives.gov.uk/electronicrecords/seamless_flow/default.htm

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