Significance Is in the Eye of the Stakeholder

Angela Dappert and Adam Farquhar

The British Library
Wetherby, West Yorkshire LS23 7BQ
{Angela.Dappert,Adam.Farquhar}@bl.uk

Abstract. Custodians of digital content take action when the material that they are responsible for is threatened by, for example, obsolescence or deterioration. At first glance, ideal preservation actions retain every aspect of the original objects with the highest level of fidelity. Achieving this goal can, however, be costly, infeasible, and sometimes even undesirable. As a result, custodians must focus their attention on preserving the most significant characteristics of the content, even at the cost of sacrificing less important ones. The concept of significant characteristics has become prominent within the digital preservation community to capture this key goal. As is often the case in an emerging field, however, the term has become over-loaded and remains ill-defined. In this paper, we unpack the meaning that lies behind the phrase, analyze the domain, and introduce clear terminology.

Keywords: Digital preservation, properties, characteristics, significant properties, significant characteristics, applicable properties, requirements.

1 Introduction

Custodians of digital content take action when the material that they are responsible for is threatened by, for example, obsolescence or deterioration. At first glance, ideal preservation actions retain every aspect of the original objects with the highest level of fidelity. Unfortunately, achieving this goal can be costly, infeasible, and sometimes even undesirable. As a result, custodians must focus their attention on preserving the most significant characteristics of the content, even at the cost of sacrificing less important ones. Furthermore, we must verify that the preservation actions we apply actually preserve these characteristics. The concept of significant characteristics has become prominent within the digital preservation community to capture this key goal [9].

The term *significant characteristic* has become over-loaded and remains ill-defined. This has some unfortunate consequences. First, communication is hampered, because the term is used in substantially different ways by different authors. Second, based on an extensive analysis of policy and strategy documents related to digital preservation [7], the current definitions do not actually meet the needs of content custodians. Content custodians need to express priorities, as well as requirements that go beyond the significance of properties and values. Third, implementations based on existing definitions fail to meet the needs of content custodians because they focus too

tightly on characteristics of content and format, and do not take account of the context in which preservation actions take place.

1.1 Related Work

Chris Rusbridge [18] eloquently states why the quest for faithfulness to the original in all respects is both excessive and impractical in most preservation situations. Original work on significant characteristics comes out of the Cedars project [5], work at the Australian National Archives [14], the InSPECT project [12], PLANETS [1, 2, 7, 9, 19] and others. Surveys of related work are provided by Knight [13] and Wilson [21].

Terminology is used inconsistently and includes significant properties [e.g., 10, 12, 13], significant characteristics [1], essence [14], aspects [8], and others. Nonetheless, a widely accepted definition for significant properties is Andrew Wilson's [21]:

"The characteristics of digital objects that must be preserved over time in order to ensure the continued accessibility, usability, and meaning of the objects, and their capacity to be accepted as evidence of what they purport to record."

The term "characteristics", which describes what must be preserved in this definition, is interpreted in two conflicting ways. Some interpret it to refer to the abstract properties of file formats [e.g., 1, 12], whereas others interpret it to refer to the values of properties of specific digital objects [2].

We also find different interpretations of the term "digital objects", which describes whose characteristics need to be preserved. In 2002, an OCLC/RLG working group[16] stated that the properties of data objects need to be preserved; Brown [3] applies it to information objects as opposed to data objects in the OAIS sense of the terms [4]; Becker [1] applies it to the characteristics of specific file formats. Knight hints that the characteristics of the environments in which digital objects are rendered may also have to be preserved [12], but this idea is not fully articulated.

The need to clarify the difference between significant characteristics and representation information has repeatedly been voiced [e.g., 10, 13], but not yet addressed.

1.2 Contributions

In this paper, we probe into the meaning of Wilson's definition. The exploration has led us to shift focus from a priori significance of characteristics in files or file formats to a new model in which stakeholders state requirements expressing significance. In contrast with previous work, we

- distinguish "properties" and "characteristics" (Section 2.1);
- provide a conceptual model, identify the types of objects which may have properties and characteristics, and unify the treatment of properties and characteristics across preservation objects, preservation actions, and their environments (Section 2.2);
- clarify who and what determines significance (Section 3);
- list observations about practical uses of significant characteristics. They justify why we treat significant characteristics as first class concept that is a subtype of requirement (Section 3);
- clarify the difference between significant characteristics, applicable properties and representation information (Section 4).

2 Foundations

2.1 Modelling Language - What Must Be Preserved?

In order to write with a reasonable level of precision, we need to introduce a basic vocabulary to talk about entities, properties, values, and so on. We use an object-oriented model with roots in [6]. The core terms in this vocabulary are:

Entity – Anything whatsoever.

Class – A class is a set of entities. Each of the entities in a class is said to be an instance of the class.

Individual – Entities that are not classes are referred to as individuals.

Property – A property is an individual that names a relationship.

Characteristic – A property / value pair associated with an entity. The value is an entity. This relationship is illustrated in Figure 1.

Facet – A facet is a property / value pair associated with a characteristic. The value is an entity.

Constraint – A Boolean condition involving expressions on entities.

Unless otherwise specified, a characteristic is directly associated with an entity. It is sometimes useful to associate a characteristic with all of the instances of a class. We refer to this as a *class characteristic*. Furthermore, we say that a property *applies* to a class if it can be meaningfully associated with some instances of the class.



Fig. 1. Properties and characteristics

We can use this language in the domain of digital objects and preservation. For example, *file* is a class; fl.txt is an instance of the class file; fileSize is a property; the property fileSize applies to file; the file fl.txt has the characteristic fileSize = 131342.

If every instance of *myDigitalSoundObject* has been virus-scanned, then it has the class characteristic *isVirusScanned* = "yes".

Important additional information about a characteristic, such as how a value is encoded, the unit of measure, or the algorithm or tool used to compute it can be specified using facets.

Under this terminology, it is clear that a characteristic (property / value pair) may be preserved by a preservation action, but that the abstract property cannot be. It is therefore not sensible to speak about preserving a "significant property."

2.2 Conceptual Model - Whose Characteristics Are Captured?

A key aspect of our model is that each of the classes *preservation object*, *environment*, and *preservation action* illustrated in Figure 2 may have properties and characteristics. It is important to distinguish the types of entity which are characterized. They play different roles during preservation processes and have different applicable

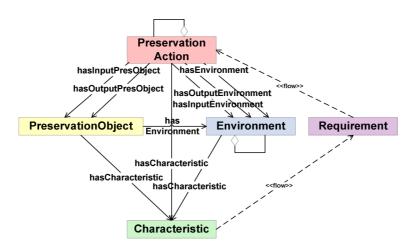


Fig. 2. Conceptual model: Characteristics of preservation objects, preservation actions and environments

properties. The labelled arrows summarise some of the properties that apply to the class' instances. This section discusses each of these concepts in more detail.

Preservation Object

The *preservation object* concept corresponds to those objects in need of preservation. It has subclasses on three tiers, as illustrated in Figure 3. The top two tiers are associated with specific physical representations of digital objects. The top tier comprises physical objects, such as bitstreams and its subclasses including *bytestreams* and *files*. The middle tier comprises *representations* of logical objects consisting of *representation bitstreams* that are needed to create a single rendition of a logical object (e.g., the set of *html* and *gif* files¹ needed to render the web version of a journal article). The bottom tier comprises logical objects such as *intellectual entities* and *components*.

These concepts are explained in detail in [8] and [9]. This presentation is somewhat modified to align terminology with PREMIS [15] and FRBR [11].

An *intellectual entity* is a distinct intellectual or artistic creation. PREMIS [15] defines it as a set of content that is considered a single intellectual unit for purposes of management and description. The *intellectual entity* can be extended in ways to meet the needs of stakeholders. For example, in the library setting, common subclasses include *collection*, *work*, and *expression*. In an archival setting, subclasses such as *fonds* and *series* are also relevant. Most repositories support discovery and delivery of *intellectual entities* such as books, videos, and articles. They may augment these with *work* and *expression* subclasses to capture useful FRBR distinctions [11]. *Intellectual entities* may also correspond to larger structures, such as *collections*, which may not be of interest to the end-user, but may be significant in preservation decisions.

¹ The formal definition of such a statement would of course contain a persistent unique identifier of the exact version of the file formats. For improved readability of examples we casually refer to file formats by their file extension.

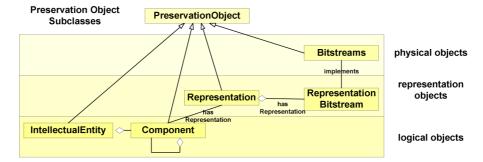


Fig. 3. Preservation Object Subclasses

During preservation, it is often necessary to consider fine-grained *components* of an *intellectual entity*. Examples include *table*, *image*, *title*, *substring*, or even an individual *character*. The *component* entity can be decomposed in several ways, such as by the type of content (e.g., *textComponent*, *imageComponent*), or by structure (e.g., *headerComponent* or *tableOfContentsComponent*). Values for characteristics of components can be measured from their associated representations (e.g. the *font* of a character component can be extracted from its representation bitstream.).

Properties can be applicable to objects in every tier. For example:

- fileSize or encoding are applicable to files.
- numberOfFilesInTheRepresentation, totalRepresentationSize, resolution, or preservationLevel are applicable to representations.
- pageCount or frameRate are properties applicable to intellectual entities such as a journal article or video. Alignment is a property applicable to a textComponent. SemanticInterpretation can be a characteristic of any component.

Environments

Preservation objects don't exist in isolation. A user or system interacts with an object in an *environment*. Therefore, every preservation object is associated with one or more environments that support different purposes or functions. Examples of environment purposes include delivery (remote or local), creation, ingest, and preservation. Examples of environment functions include rendering, editing, executing, and printing.

Every environment may be broken down into sub-environments that are needed for the interpretation and representation of the preservation object. Examples include hardware and software environments, the community, budgetary factors, the legal system, and other internal and external factors. They correspond to an extended notion of the environment description of *representation information* [4] and are enumerated in [8].

Environments have characteristics. For example:

- *memoryUsage* = "low" is a characteristic of a software tool environment that renders the preservation object.
- numberOfIntermediateCopies <= 3 and preservesColourDepth = "yes" are characteristics of a preservation service which is part of a preservation action's environment. They can be captured in a preservation services registry.

Preservation Actions

Custodians of digital content take actions to mitigate the risks that they identify. A preservation action event takes place when a preservation service is invoked. A preservation action is applied to an initial, or input, preservation object and environment. The result of the action is either a new output preservation object and/or a new environment. Together they mitigate the risk that the action addresses. For example, a Microsoft Word bytestream is migrated to a *pdf* bytestream in order to lock in the desired look-and-feel of the document. The output environment must support a *pdf* viewer. Characteristics of the output preservation object and the output environment are validated against significant characteristics in order to quantify the degree of compliance. This approach to describing preservation actions works for migration, emulation, hardware replacement, and other solutions.

Every preservation action is associated with the environment required for its own execution. The hardware on which the action is executed and the preservation service that is invoked are parts of this environment.

Preservation actions may have characteristics. For example, *numberOfIntermediateCopiesProduced* = 2 is a characteristic of a preservation action. This might be used to identify preservation actions that violate copyright regulations or license agreements that limit the number of intermediate copies created.

3 Observations about Significance in Digital Preservation

Observation 1

An idea, concept, act, or thing is not inherently significant. A stakeholder attributes significance to something, typically in a context relevant to some purpose or goal. In the digital preservation context, significance is determined by the stakeholders involved in the preservation process. These include the producer of the digital object, the custodian who holds it, and the consumer who will access it. The stakeholder's priorities may be captured as requirements ("business rules") by the custodian, who needs to ensure that preservation actions satisfy them. Requirements are an explicit statement of a stakeholder's values. These requirements influence the preservation process, and are often captured in preservation guiding documents, such as strategy or business documents. The conceptual model must have a requirement concept for capturing significance explicitly.

There is a notion that significant characteristics refer to the intellectual content—the essence of the digital object. In contrast, other characteristics are merely circumstantial, not significant, and can be ignored in preservation actions. Unfortunately, it is not possible to determine out of context which properties reflect content and which reflect circumstance. Consider a number that is formatted with the colour red. In some settings, the colour may be for a visual effect—simply pretty, circumstantial, and insignificant; in another setting, the colour may be to indicate that it is to be understood as negative and therefore has a significant semantic impact. This can only be determined by the stakeholder capturing significance explicitly.

Observation 2

Stakeholders specify constraints on both preservation objects and environments. Jeff Rothberg introduced widely used criteria to evaluate authenticity [17]: *content*,

context, appearance, structure, and behaviour. These are sometimes misinterpreted as exhaustive categories for significant characteristics [e.g. 12]. The consequence is to limit significant characteristics to "informational entities" - the logical preservation object itself - and exclude bytestreams, representations, or environments.

In contrast, the characteristics of preservation actions constrain the context in which significant characteristics apply, but are not themselves significant.

Observation 3

Significant characteristics are not simple property/value pairs which a stakeholder declares to be significant. Our analysis of policy and strategy documents [8] shows that stakeholders need to state more complex requirements that can be expressed as constraints, using a constraint language such as OCL [20]. They often need to include specifications such as contexts, invariants, pre-conditions and post-conditions.

In many cases, for example, a stakeholder considers characteristics to be significant only when some additional conditions are met - that is, a context is specified. As a result, the language that we use to define significant characteristics must be expressive enough to include a context.

Sometimes the conditions involve preservation object or environment characteristics:

- If *componentType* = "text" then *fontSize* is significant.
- If *environmentType* = "preservation" then *resolution* is significant.

 At other times the conditions involve preservation action characteristics:
- If preservationActionType = "bitPreservation" then fileSize is significant.

Observation 4

Significance is not absolute and binary. We can not only choose which characteristics should be significant, but would like to specify an importance factor which is a measure of the relative significance of the characteristic for the stakeholders. I may consider each of two conflicting things significant and prioritise one as more significant than another. This prioritisation is essential for both decision making and planning.

Finally, requirements may tolerate some deviation or error. For example, an office document migration that produced a result with different hyphenation or pagination might be acceptable in many situations. We can allow for a tolerance factor which specifies to what degree deviation from the required value can be tolerated. During evaluation of a preservation action the importance and tolerance factors can be combined into a weighted measure of the significant characteristic.

Observation 5

In many cases, we wish to include the possibility of capturing improvements to an object. A common preservation action is normalization of digital objects upon ingest. This may be done to reduce the variety of formats held, but may also be done to improve characteristics in the original. For example, we might migrate files which are in formats that are susceptible to degradation to files in a more resilient format, or move static tables to spreadsheets which enable pivot tables. In this case the characteristics fileFormatResilience = "high" or enablesPivotTables = "yes" are significant characteristics which were not found in the original. Another preservation action which improves upon the original is the manual restoration of a file by a curator to the state it was presumed to have had before a corruption. Another common example can be found in CAD drawings or data sets. As technology improves, consumers desire to perform new functions on old data in ways that were previously not possible.

Observation 6

While characteristics capture values at a given moment in time, significant characteristics capture constraints on characteristics across time – before and after a preservation action.

As a result of Observations 2, 3, 4, 5, and 6, the language that we use to define significant characteristics must be able to express relationships other than the simple preservation of a value.

The above observations illustrate that significant characteristics are a subclass of preservation guiding requirements [8, 9]. Ideally, we would rename them to "significance requirements", but were reluctant to break too radically with current terminology. We recommend that significant characteristics which express requirements or business rules should in the general case be represented as explicit first class objects in a data model. Figure 2 introduces this separate concept.

We define significant characteristics as:

Requirements in a specific context, represented as constraints, expressing a combination of characteristics of preservation objects or environments that must be preserved or attained in order to ensure the continued accessibility, usability, and meaning of preservation objects, and their capacity to be accepted as evidence of what they purport to record.

4 Discussion

Using the conceptual model and the definition of significant characteristics, we can now investigate some implications of the definition and the relationship of significant characteristics to related digital preservation concepts.

4.1 Implications of the Conceptual Model

The conceptual model was motivated by our findings during the analysis of preservation policy and strategy documents [8]. It suggests the need for developing approaches that allow stakeholders to express constraints with prioritisation and tolerances.

It supports a wide array of preservation activities found in real organisations. Characteristics of different entities are used to express requirements for different preservation activities or purposes. For example, bit-preservation actions such as media refresh preserve characteristics at the file or representation level such as *fileSize*, *encoding*, or the *numberOfFilesInTheRepresentation*. In contrast, migration actions can be expected to change these characteristics.

Significant characteristics at the representation level can express requirements associated with the representations' different purposes, such as preservation versus access copies. *Resolution* = "high" and *preservationLevel* = "9" may be significant characteristics of a representation that is aimed at preserving archival quality.

A significant characteristic that is considered an inherent requirement of a logical component and does not vary from representation to representation should be captured on the logical level. These requirements need to be satisfied by all migration or emulation actions applied to this logical component. For example the requirement *sematicInterpretation* = "negative number" may be declared significant for all representations of a numberComponent. Different representations of the numberComponent

can satisfy it by rendering it as a red number, adding a minus sign or surrounding it by parenthesis, but the logical requirement must be satisfied for all of them.

Significant characteristics of intellectual entities can model high level policy and strategy requirements, such as legal or fiscal requirements that must be satisfied after any preservation action.

Significant characteristics of environments make it possible to express requirements whose aim is preserving the look-and-feel of an information object, since the look-and-feel is determined by the combination of the data object and its environment. These significant requirements support emulation and migration activities equally. Environmental factors can also be external or internal policy factors which permit the expression of policy constraints.

4.2 File Formats and Properties

The basic consequence of this analysis is that significance is not inherent in or determined by the file formats of digital objects – but by the needs and requirements of stakeholders in the preservation activities. This enables us to make sense of common preservation activities, such as migration to less expressive file formats. For example, some stakeholders will be satisfied by migration from a *word* document to a simple text file when the original contains only simple text components (i.e., no formatting, headers, tables, and so on). A radio station might be satisfied by a migration that only preserves the audio stream of a video object. The analysis also shows why there can be disagreement about the significance of a property between stakeholders. Disagreement reflects different requirements and priorities among stakeholders. For example, the rotational frequency of a shape in a piece of online art may be significant to the artist, but not for many viewers.

The analysis also clarifies the role of archival subsets of file formats, such as *pdf/a*. The well-designed archival format profile will support properties that are of interest to a substantial community of stakeholders and appear in a substantial subset of content in the full file format.

The preservation community is establishing registries of file formats and properties that apply to them [12, 19]. These are registries of applicable properties² rather than of significant characteristics. A stakeholder may indicate that some of the applicable properties are not significant in certain contexts. This increases the set of preservation actions that are appropriate. Conversely, a stakeholder may indicate preconditions which rule out preservation actions that would have been appropriate considering only the file format's applicable properties.

4.3 Significant Characteristics and Representation Information

How do the significant characteristics of this conceptual model relate to representation information, as defined in OAIS [4, 16]? Representation information is "the

² There are also properties which describe a file format itself rather than the objects that are represented in files. They often appear in stakeholder requirements and enable stakeholders to choose formats that suit their business needs. For example, a custodian might require files to be represented in formats defined by an open standard, or in common use, or with high resilience to degradation damage.

information that maps a Data Object into more meaningful concepts. An example is the ASCII definition that describes how a sequence of bits (i.e., a Data Object) is mapped into a symbol."

Representation information is a set of characteristics describing the preservation object and its environment. Furthermore, representation information is specified for a specific context, namely for a given "designated community". It will vary for different designated communities. Additionally, the purpose of representation information is to guarantee the accessibility, usability, and meaning of preservation objects. All these characteristics of representation information agree with the definition of significant characteristics. It becomes obvious, that representation information is NOT a form of significant characteristics when we realize that it does not specify characteristics that need to be preserved or attained, nor does it specify requirements for preservation actions. Representation information is the set of important characteristics of a data object that are needed to make sense of it for a given designated community at a given time. It does not specify constraints for transformations over time, and it does not specify characteristics of an acceptable derived data object.

A piece of representation information, for example, may be the fact that a given data object requires a certain software package for its proper rendering. This does not imply that the corresponding information object after a migration must use this same software package.

Some pieces of representation information may, however, be declared to be significant for preservation purposes. For example, the semantic interpretation of a data object, such as the characteristic that a given *numberComponent* is to be interpreted as "body weight", is likely to be considered significant in most contexts.

5 Conclusion

This article has examined the concept of significance in digital preservation and presented a new model that places significance in the hands of stakeholders. The model has extended the domain of significant characteristics beyond digital objects to include environments. The model has consequences for implementations of preservation metadata dictionaries, property registries, and preservation services.

This work has been conducted within the larger context of defining a conceptual model and specific vocabulary for supporting preservation processes [8] within the PLANETS project. Significant characteristics can be considered one specific form of preservation guiding requirements which are discussed in [8].

This work has been presented within the digital preservation framework, but may apply to other transformation applications such as rendering accessible versions of digital objects for disabled users.

Acknowledgement

Work presented in this paper is partially supported by the European Community under the Information Society Technologies (IST) Programme of the 6th FP for RTD - Project IST-033789. The author is solely responsible for the content of this paper.

References

- Becker, C., et al.: A generic XML language for characterising objects to support digital preservation. In: SAC 2008: Proceedings of the 2008 ACM symposium on Applied computing, pp. 402–406 (2008)
- 2. Becker, C., et al.: Plato: A Service Oriented Decision Support System for Preservation Planning. In: JCDL 2008, Pittsburgh, Pennsylvania, USA (June 2008)
- 3. Brown, A.: White Paper: Representation Information Registries. PLANETS report PC/3-D7 (2008), http://www.planets-project.eu/docs/reports/Planets_PC3-D7_RepInformationRegistries.pdf
- 4. CCSDS. Reference Model for an Open Archival Information System (OAIS). CCSDS 650.0-B-1, Blue Book (the full ISO standard) (January 2002), http://public.ccsds.org/publications/archive/650x0b1.pdf
- 5. CEDARS Project, http://www.leeds.ac.uk/cedars/
- 6. Chaudhri, V., Farquhar, A., Fikes, R., Karp, P., Rice, J.: OKBC: A programmatic foundation for knowledge base interoperability. In: Proceedings of the 1998 National Conference on Artificial Intelligence (1998)
- 7. Clausen, L.: Opening Schroedinger's Library: Semi-automatic QA reduces uncertainty in object-transformation. In: Kovács, L., Fuhr, N., Meghini, C. (eds.) ECDL 2007. LNCS, vol. 4675, pp. 186–197. Springer, Heidelberg (2007)
- Dappert, A., Ballaux, B., Mayr, M., van Bussel, S.: Report on policy and strategy models for libraries, archives and data centres. PLANETS report PP2-D2 (2008), http://www.planets-project.eu/docs/reports/ Planets_PP2_D2_ReportOnPolicyAndStrategyModelsM24_Ext.pdf
- 9. Dappert, A., Farquhar, A.: Modelling Organisational Goals to Guide Preservation. In: iPRES 2008: The Fifth International Conference on Preservation of Digital Objects (2008), http://www.bl.uk/ipres2008/ipres2008-proceedings.pdf
- Hockx-Yu, H., Knight, G.: What to Preserve?: Significant Properties of Digital Objects.
 In: Report on the JISC/BL/DPC Workshop of April 7, 2008, British Library Conference Centre (2008); The International Journal of Digital Curation 3(1) (2008), http://www.ijdc.net/./ijdc/article/view/70/70
- 11. IFLA Study Group: Functional Requirements for Bibliographic Records (1998), http://www.ifla.org/VII/s13/frbr/frbr.pdf (Retrieved on December 1, 2007)
- 12. Knight, G.: Framework for the definition of significant properties (2008), http://www.significantproperties.org.uk/documents/wp33-propertiesreport-v1.pdf
- 13. Knight, G., Pennock, M.: Data Without Meaning: Establishing the Significant Properties of Digital Research. In: iPRES 2008: The Fifth International Conference on Preservation of Digital Objects (2008),
 - http://www.bl.uk/ipres2008/ipres2008-proceedings.pdf
- 14. National Archives of Australia: An Approach to the Preservation of Digital Records (December 2002), http://www.naa.gov.au/Images/ An-approach-Green-Paper_tcm2-888.pdf
- 15. PREMIS Editorial Committee: PREMIS Data Dictionary for Preservation Metadata, Version 2 (March 2008), http://www.loc.gov/standards/premis/v2/premis-2-0.pdf

- The OCLC/RLG Working Group on Preservation Metadata: Preservation Metadata and the OAIS Information Model. A Metadata Framework to Support the Preservation of Digital Objects (2002),
 - http://www.oclc.org/research/projects/pmwg/pm_framework.pdf
- 17. Rothenberg, J.: Preserving Authentic Digital Information. In: Authenticity in a Digital Environment. The Council on Library and Information Resources (2000), http://www.clir.org/pubs/reports/pub92/rothenberg.html
- Rusbridge, C.: Excuse me...some digital preservation fallacies? Ariadne, 46 (February 2006), http://www.ariadne.ac.uk/issue46/rusbridge/ ISSN 1361-3200
- 19. The National Archives: PRONOM, http://www.nationalarchives.gov.uk/pronom/
- 20. Warmer, A., Kleppe, A.: The Object Constraint Language. In: Getting Your Models Ready for MDA, Addison-Wesley Longman Publishing Co., Boston (2003)
- 21. Wilson, A.: Significant Properties Report, InSPECT Work Package 2.2, Draft/Version 2 (2007), http://www.significantproperties.org.uk/documents/wp22_significant_properties.pdf