

**Data Dictionary section
from
PREMIS Data Dictionary
for Preservation Metadata**

version 2.0

This is an excerpt from the PREMIS version 2.0 document. It includes only the Data Dictionary section. The Introduction, Special Topics, Methodology, and Glossary are in a separate excerpt. The full document and both excerpts are available online from: <http://www.loc.gov/standards/premis/>

PREMIS Editorial Committee

March 2008

CONTENTS

Acknowledgments.....	ii
PREMIS Web Sites and E-Mail.....	iv
Introduction	1
<i>See separate document</i>	
The PREMIS Data Dictionary Version 2.0	22
Limits to the scope of the Data Dictionary.....	23
Object Entity	25
Event Entity	130
Agent Entity	151
Rights Entity	157
Special Topics.....	195
<i>See separate document</i>	
Methodology	208
<i>See separate document</i>	
Glossary.....	209
<i>See separate document</i>	

ACKNOWLEDGMENTS

PREMIS Editorial Committee members

Rebecca Guenther, Library of Congress, *Chair*
Steve Bordwell, General Register Office for Scotland
Olaf Brandt, Koninklijke Bibliotheek, Netherlands
Priscilla Caplan, Florida Center for Library Automation
Gerard Clifton, National Library of Australia
Angela Dappert, British Library
Markus Enders, Staats- und Universitätsbibliothek Göttingen/British Library
Brian Lavoie, OCLC
Bill Leonard, Library and Archives Canada
Zhiwu Xie, Los Alamos National Laboratory

Special thanks

These individuals contributed their expertise as former members of the PREMIS Editorial Committee:

Rory McLeod, British Library
Yaniv Levi, ExLibris

These individuals were the original *Preservation Metadata: Implementation Strategies* (PREMIS) Working Group that developed version 1 of the Data Dictionary:

Priscilla Caplan, Florida Center for Library Automation, *co-chair*
Rebecca Guenther, Library of Congress, *co-chair*
Robin Dale, *RLG liaison*
Brian Lavoie, *OCLC liaison*
George Barnum, U.S. Government Printing Office
Charles Blair, University of Chicago
Olaf Brandt, Göttingen State and University Library
Mikki Carpenter, Museum of Modern Art
Adam Farquhar, British Library
David Gewirtz, Yale University
Keith Glavash, MIT/DSpace
Andrea Goethals, Florida Center for Library Automation
Cathy Hartman, University of North Texas
Helen Hodgart, British Library
Nancy Hoebelheinrich, Stanford University
Roger Howard, J. Paul Getty Museum

Sally Hubbard, Getty Research Institute
Mela Kircher, OCLC
John Kunze, California Digital Library
Vicky McCargar, Los Angeles Times
Jerome McDonough, New York University/METS
Evan Owens, Ithaka-Electronic Archiving Initiative
Erin Rhodes, U.S. National Archives and Records Administration
Madi Solomon, Walt Disney Corporation
Angela Spinazze, ATSPIN Consulting
Stefan Strathmann, Göttingen State and University Library
Günter Waibel, RLG
Lisa Weber, U.S. National Archives and Records Administration
Robin Wendler, Harvard University
Hilde van Wijngaarden, National Library of the Netherlands
Andrew Wilson, National Archives of Australia and British Library
Deborah Woodyard-Robinson, British Library and Woodyard-Robinson Holdings Ltd.

PREMIS WEB SITES AND E-MAIL

PREMIS maintenance activity Web site: www.loc.gov/standards/premis/.

PREMIS Implementers' Group discussion list: pig@loc.gov. To subscribe, send email to listserv@loc.gov with the message, "subscribe pig [your name]"

Please send comments and questions to premis@loc.gov.

INTRODUCTION

The PREMIS Introduction (pages 1-21) has been removed from this excerpt; it is available in a separate excerpt and in the full document. All the PREMIS documents are available online at: <http://www.loc.gov/standards/premis/>

THE PREMIS DATA DICTIONARY VERSION 2.0

The PREMIS Data Dictionary includes semantic units for Objects, Events, Agents, and Rights. The fifth entity in the model, the Intellectual Entity, is considered out of scope because it is well served by descriptive metadata. The template for each entry includes a place for notes about how to create or use the semantic unit. In some cases the group felt additional information, such as the reason for a semantic unit's definition or issues that arose in the group's deliberations, would be useful; for these details, see "[Special Topics](#)" page [195](#).

A semantic component always inherits the applicability of the containing semantic unit. That is, if the containing semantic unit specifies that it is applicable to files but not to representations, each of its semantic components is applicable to files and not to representations. Repeatability and obligation, however, may vary.

Each entry in the Data Dictionary offers these attributes of a semantic unit:

- **Name of the semantic unit:** Names were devised to be descriptive and unique within the Data Dictionary. Using these names for the exchange of metadata among preservation repositories will aid interoperability. These names need not be used internally within any individual preservation repository.
- **Semantic components:** The semantic components each have their own entries later in the Data Dictionary. A semantic unit that has semantic components does not have any value of its own. Only semantic units at the lowest level have values.
- **Definition:** The meaning of the semantic unit.
- **Rationale:** Why the semantic unit is needed, if this is not self-evident from the definition.
- **Data constraint:** How the value of the semantic unit should be encoded. Some common data constraints are:

Container – The semantic unit is an umbrella for two or more semantic components and has no value of its own.

None – The semantic unit can take any form of value.

Value should be taken from a controlled vocabulary – The preservation repository should establish an authority list of values that are useful and meaningful to the repository. The PREMIS Data Dictionary does not specify what this authority list should be, and it is assumed that different repositories will use different vocabularies. In general, when a value is taken from a controlled vocabulary, the source of the vocabulary should be recorded. A mechanism to record the source is provided in the PREMIS XML schemas.

- **Object category:** Whether the unit applies to a representation, file, or bitstream Object. Semantic units that apply to files also apply to filestreams (see page 7).
- **Applicability:** A scope of "applicable" means it applies to that category of Object.

- **Examples:** One or more examples of values the semantic unit may take. Examples are intended to be illustrative.

An example of an actual value is set in normal text. Text in brackets presents a description of the value rather than the value itself. For example, “SHA-1 message digest” reflects the actual value of the semantic unit, while “[SHA-1 message digest]” means the value of the semantic unit is an SHA-1 message digest such as “7c9b35da4f2ebd436f1cf88e5a39b3a257edf4a22be3c955ac49da2e2107b67a1924419563”

- **Repeatability:** A semantic unit designated as “Repeatable” can take multiple values. It does not mean that a repository must record multiple instances of the semantic unit.
- **Obligation:** Whether a value for the semantic unit is mandatory (if applicable) or optional.

A mandatory semantic unit is something that the preservation repository needs to know, independent of how or whether the repository records it. The repository might not explicitly record a value for the semantic unit if it is known by some other means (e.g., by the repository’s business rules). “Mandatory” actually means “mandatory if applicable.” For example, an identifier for a bitstream is mandatory only if the repository manages data at the bitstream level. When exchanging PREMIS-conformant metadata with another repository, values for mandatory semantic units must always be provided.

Values for optional semantic units are encouraged but not required.

If a container unit is optional, but a semantic component within that container is mandatory, the semantic component must be supplied if and only if the container unit exists. That is, if a value for any of the optional or mandatory semantic units in the container is supplied, a value for all of the mandatory semantic units in the container must be supplied.

- **Creation/Maintenance notes:** Notes about how the values for the semantic unit may be obtained and/or updated.
- **Usage notes:** Information about the intended use of the semantic unit, or clarification of the definition.

Limits to the scope of the Data Dictionary

Descriptive metadata: Typically, descriptive metadata is used to describe Intellectual Entities. Nearly all preservation repositories either include descriptive metadata or link to descriptive metadata located outside the repository itself. Such metadata may identify a resource by publication information such as creator and title, or may characterize its intellectual content through classification, subject terms, and so on. Descriptive metadata can be important both for discovery of archived resources and for helping decision makers during preservation planning. However, the Data Dictionary does not focus on descriptive elements for two reasons.

First, descriptive metadata is well served by existing standards. MARC¹⁶, MODS¹⁷, the Dublin Core Metadata Element Set¹¹, the Content Standard for Digital Geospatial Metadata¹⁸, the VRA Core¹⁹, the Encoded Archival Description (EAD)²⁰, and the Data Documentation Initiative²¹ schemas are only some of the standards that define descriptive metadata elements. The working

THE PREMIS DATA DICTIONARY

group did not want to add another set of descriptive elements to an already crowded field. Second, descriptive metadata is often domain specific. For the purposes of preservation it is less crucial that a common set of elements describe, for example, satellite telemetry and digital Picassos than that communities of interest be able to capture and exchange information in a form that reflects their materials and interests appropriately.

Agents: PREMIS does not define the characteristics of Agents in any detail. Metadata describing people, organizations, and other entities that can act as Agents has been defined in many existing formats and standards, such as MARC¹⁶, vCard²², MADS²³, and several other schemes currently under development. As long as a preservation repository can properly identify Agents that have acted upon Objects in its care, additional Agent characteristics will be determined by local requirements; many can be modeled on existing standard metadata element sets.

Rights: PREMIS primarily defines characteristics of rights and permissions concerned with preservation activities, not those associated with access and/or distribution. This revision broadens the semantic units used for rights information and allows for extensibility to use an external rights metadata scheme.

Technical metadata: Technical metadata describes the physical rather than intellectual characteristics of digital objects. Detailed, format-specific technical metadata is clearly necessary for implementing most preservation strategies, but the group had neither the time nor the expertise to tackle format-specific technical metadata for various types of digital files. Therefore, it restricted the technical metadata included in the Data Dictionary to the semantic units it believed apply to objects in all formats. Further development of technical metadata is left to format experts. An extensibility mechanism is provided by including the semantic unit *objectCharacteristicsExtension*, which may be used with an external technical metadata scheme.

Media or hardware details: The working group did not attempt to define metadata for detailed documentation of media or hardware. For example, PREMIS defines a semantic unit for identifying the medium on which an object is stored. A preservation repository will probably want to know more detailed information about the media employed. If the repository stores data on DVDs, for example, it may need to know the specific technical characteristics of the specific DVD units, such as manufacturer, dye material, and dye thickness. PREMIS leaves the definition of metadata for describing media and hardware characteristics to specialists in these areas.

Business rules: The working group made no attempt to describe the business rules of a repository, although certainly this metadata is essential for preservation within the repository. Business rules codify the application of preservation strategies and document repository policies, services, charges, and roles. Retention periods, disposition, risk assessment, permanence ratings, schedules for media refreshment, and so on are pertinent to objects but are not actual properties of Objects. A single exception was made for the level of preservation treatment to be accorded an object (*preservationLevel*) because this was felt to be critical information for any preservation repository. A more thorough treatment of business rules could be added to the data model by defining a Rules entity similar to Rights, although this is not included in the current revision.

Object Entity

The Object entity aggregates information about a digital object held by a preservation repository and describes those characteristics relevant to preservation management.

The only mandatory semantic unit that applies to all categories of object (representation, file, and bitstream) is *objectIdentifier*.

Entity types

- **Representation:** A digital object instantiating or embodying an Intellectual Entity. A representation is the set of stored digital files and structural metadata needed to provide a complete and reasonable rendition of the Intellectual Entity.
- **File:** A named and ordered sequence of bytes that is known to an operating system.
- **Bitstream:** Contiguous or non-contiguous data within a file that has meaningful properties for preservation purposes.

Entity properties

- Can be associated with one or more rights statements.
- Can participate in one or more events.
- Links between entities may be recorded from either direction and need not be bi-directional.

Entity semantic units

- 1.1 objectIdentifier (M, R)
 - 1.1.1 objectIdentifierType (M, NR)
 - 1.1.2 objectIdentifierValue (M, NR)
- 1.2 objectCategory (M, NR)
- 1.3 preservationLevel (O, R) [representation, file]
 - 1.3.1 preservationLevelValue (M, NR) [representation, file]
 - 1.3.2 preservationLevelRole (O, NR) [representation, file]
 - 1.3.3 preservationLevelRationale (O, R) [representation, file]
 - 1.3.4 preservationLevelDateAssigned (O, NR) [representation, file]
- 1.4 significantProperties (O, R)
 - 1.4.1 significantPropertiesType (O, NR)
 - 1.4.2 significantPropertiesValue (O, NR)
 - 1.4.3 significantPropertiesExtension (O, R)
- 1.5 objectCharacteristics (M, R) [file, bitstream]
 - 1.5.1 compositionLevel (M, NR) [file, bitstream]

THE PREMIS DATA DICTIONARY

- 1.5.2 fixity (O, R) [file, bitstream]
 - 1.5.2.1 messageDigestAlgorithm (M, NR) [file, bitstream]
 - 1.5.2.2 messageDigest (M, NR) [file, bitstream]
 - 1.5.2.3 messageDigestOriginator (O, NR) [file, bitstream]
- 1.5.3 size (O, NR) [file, bitstream]
- 1.5.4 format (M, R) [file, bitstream]
 - 1.5.4.1 formatDesignation (O, NR) [file, bitstream]
 - 1.5.4.1.1 formatName (M, NR) [file, bitstream]
 - 1.5.4.1.2 formatVersion (O, NR) [file, bitstream]
 - 1.5.4.2 formatRegistry (O, NR) [file, bitstream]
 - 1.5.4.2.1 formatRegistryName (M, NR) [file, bitstream]
 - 1.5.4.2.2 formatRegistryKey (M, NR) [file, bitstream]
 - 1.5.4.2.3 formatRegistryRole (O, NR) [file, bitstream]
 - 1.5.4.3 formatNote (O, R) [file, bitstream]
- 1.5.5 creatingApplication (O, R) [file, bitstream]
 - 1.5.5.1 creatingApplicationName (O, NR) [file, bitstream]
 - 1.5.5.2 creatingApplicationVersion (O, NR) [file, bitstream]
 - 1.5.5.3 dateCreatedByApplication (O, NR) [file, bitstream]
 - 1.5.5.4 creatingApplicationExtension (O, R) [file, bitstream]
- 1.5.6 inhibitors (O, R) [file, bitstream]
 - 1.5.6.1 inhibitorType (M, NR) [file, bitstream]
 - 1.5.6.2 inhibitorTarget (O, R) [file, bitstream]
 - 1.5.6.3 inhibitorKey (O, NR) [file, bitstream]
- 1.5.7 objectCharacteristicsExtension (O, R) [file, bitstream]
- 1.6 originalName (O, NR) [representation, file]
- 1.7 storage (M, R) [file, bitstream]
 - 1.7.1 contentLocation (O, NR) [file, bitstream]
 - 1.7.1.1 contentLocationType (M, NR) [file, bitstream]
 - 1.7.1.2 contentLocationValue (M, NR) [file, bitstream]
 - 1.7.2 storageMedium (O, NR) [file, bitstream]
- 1.8 environment (O, R)
 - 1.8.1 environmentCharacteristic (O, NR)
 - 1.8.2 environmentPurpose (O, R)
 - 1.8.3 environmentNote (O, R)
 - 1.8.4 dependency (O, R)
 - 1.8.4.1 dependencyName (O, R)
 - 1.8.4.2 dependencyIdentifier (O, R)
 - 1.8.4.2.1 dependencyIdentifierType (M, NR)
 - 1.8.4.2.2 dependencyIdentifierValue (M, NR)
 - 1.8.5 software (O, R)
 - 1.8.5.1 swName (M, NR)

- 1.8.5.2 swVersion (O, NR)
- 1.8.5.3 swType (M, NR)
- 1.8.5.4 swOtherInformation (O, R)
- 1.8.5.5 swDependency (O, R)
- 1.8.6 hardware (O, R)
 - 1.8.6.1 hwName (M, NR)
 - 1.8.6.2 hwType (M, NR)
 - 1.8.6.3 hwOtherInformation (O, R)
- 1.8.7 environmentExtension (O, R)
- 1.9 signatureInformation (O, R) [file, bitstream]
 - 1.9.1 signature (O, R)
 - 1.9.1.1 signatureEncoding (M, NR) [file, bitstream]
 - 1.9.1.2 signer (O, NR) [file, bitstream]
 - 1.9.1.3 signatureMethod (M, NR) [file, bitstream]
 - 1.9.1.4 signatureValue (M, NR) [file, bitstream]
 - 1.9.1.5 signatureValidationRules (M, NR) [file, bitstream]
 - 1.9.1.6 signatureProperties (O, R) [file, bitstream]
 - 1.9.1.7 keyInformation (O, NR) [file, bitstream]
 - 1.9.2 signatureInformationExtension (O, R) [file, bitstream]
- 1.10 relationship (O, R)
 - 1.10.1 relationshipType (M, NR)
 - 1.10.2 relationshipSubType (M, NR)
 - 1.10.3 relatedObjectIdentification (M, R)
 - 1.10.3.1 relatedObjectIdentifierType (M, NR)
 - 1.10.3.2 relatedObjectIdentifierValue (M, NR)
 - 1.10.3.3 relatedObjectSequence (O, NR)
 - 1.10.4 relatedEventIdentification (O, R)
 - 1.10.4.1 relatedEventIdentifierType (M, NR)
 - 1.10.4.2 relatedEventIdentifierValue (M, NR)
 - 1.10.4.3 relatedEventSequence (O, NR)
- 1.11 linkingEventIdentifier (O, R)
 - 1.11.1 linkingEventIdentifierType (M, NR)
 - 1.11.2 linkingEventIdentifierValue (M, NR)
- 1.12 linkingIntellectualEntityIdentifier (O, R)
 - 1.12.1 linkingIntellectualEntityIdentifierType (M, NR)
 - 1.12.2 linkingIntellectualEntityIdentifierValue (M, NR)
- 1.13 linkingRightsStatementIdentifier (O, R)
 - 1.13.1 linkingRightsStatementIdentifierType (M, NR)
 - 1.13.2 linkingRightsStatementIdentifierValue (M, NR)

THE PREMIS DATA DICTIONARY

Semantic unit	1.1 objectIdentifier		
Semantic components	1.1.1 objectIdentifierType 1.1.2 objectIdentifierValue		
Definition	A designation used to uniquely identify the object within the preservation repository system in which it is stored.		
Rationale	Each data object held in the preservation repository must have a unique identifier to relate it to descriptive, technical, and other metadata.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Mandatory	Mandatory	Mandatory
Creation / Maintenance notes	An identifier may be created by the repository system at the time of ingest, or it may be created or assigned outside of the repository and submitted with an object as metadata. Similarly, identifiers can be automatically or manually generated. Recommended practice is for repositories to use identifiers automatically created by the repository as the primary identifier in order to ensure that identifiers are unique and usable by the repository. Externally assigned identifiers can be used as secondary identifiers in order to link an object to information held outside the repository.		
Usage notes	<p>The <i>objectIdentifier</i> is mandatory if the preservation repository stores and manages objects at that level (i.e., representation, file, bitstream).</p> <p>The <i>objectIdentifier</i> is repeatable in order to allow both repository-assigned and externally-assigned identifiers to be recorded. See Creation/Maintenance note above.</p> <p>Identifiers must be unique within the repository. They may be preexisting, and in use in other digital object management systems.</p> <p>Identifiers used to identify a class of objects (e.g., the way an ISBN identifies all books in the same edition) are not acceptable as identifiers in the context of the preservation repository, which must identify the specific object in the repository.</p> <p>A preservation repository needs to know both the type of object identifier and the value. If the value itself contains the identifier type (e.g., “oai:lib.uchicago.edu:1”), the identifier type does not need to be explicitly recorded. Similarly, if the repository uses only one type of</p>		

	<p>identifier, the type can be assumed and does not need to be explicitly recorded.</p> <p>A persistent identifier should be used, but the particular identifier scheme is an implementation specific decision.</p>
--	---

THE PREMIS DATA DICTIONARY

Semantic unit	1.1.1 objectIdentifierType		
Semantic components	None		
Definition	A designation of the domain within which the object identifier is unique.		
Rationale	Identifier values cannot be assumed to be unique across domains; the combination of <i>objectIdentifierType</i> and <i>objectIdentifierValue</i> should ensure uniqueness.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples	DLC DRS hdl:4263537	DLC DRS hdl:4263537	DLC DRS hdl:4263537
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory
Usage notes	The type of the identifier may be implicit within the repository as long it is can be explicitly communicated when the digital object is disseminated outside of it.		

Semantic unit	1.1.2 objectIdentifierValue		
Semantic components	None		
Definition	The value of the <i>objectIdentifier</i> .		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples	0000000312	IU2440 WAC1943.56 AMNH CD269/CD269/70/10 596.PCD CDS-VDEP- 200211119- 24879.734 1001/dig/pres/2004- 024 http://nrs.harvard.edu /urn- 3:FHCL.Loeb:sa1	IU2440-1 IU2440-2
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	1.2 objectCategory		
Semantic components	None		
Definition	The category of object to which the metadata applies.		
Rationale	Preservation repositories are likely to treat different categories of objects (representations, files, and bitstreams) differently in terms of metadata and data management functions.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples	representation	file	bitstream
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory
Usage notes	Suggested values: representation, file, bitstream. A filestream should be considered a file.		

Semantic unit	1.3 preservationLevel		
Semantic components	<p>1.3.1 preservationLevelValue</p> <p>1.3.2 preservationLevelRole</p> <p>1.3.3 preservationLevelRationale</p> <p>1.3.4 preservationLevelDateAssigned</p>		
Definition	Information indicating the decision or policy on the set of preservation functions to be applied to an object and the context in which the decision or policy was made.		
Rationale	Some preservation repositories will offer multiple preservation options depending on factors such as the value or uniqueness of the material, the “preservability” of the format, the amount the customer is willing to pay, etc. The context surrounding the choice of a particular preservation option for an object may also require further explanation.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Not applicable
Repeatability	Repeatable	Repeatable	
Obligation	Optional	Optional	
Creation / Maintenance notes	The preservation level may be assigned by the repository or requested by the depositor and submitted as metadata. The repository may also choose to record additional metadata indicating the context for the assignment of the preservation level.		
Usage notes	<p>If the repository offers only a single preservation level, this value does not need to be explicitly recorded within the repository.</p> <p>Application of a particular set of <i>preservationLevel</i> semantic units may only cover a single representation of an object: representations in other technical forms or serving other functions may have a different <i>preservationLevel</i> applied.</p> <p>The container may be repeated if a preservation level value needs to be recorded in additional contexts (see <i>preservationLevelRole</i>, page 35).</p>		

THE PREMIS DATA DICTIONARY

Semantic unit	1.3.1 preservationLevelValue		
Semantic components	None		
Definition	A value indicating the set of preservation functions expected to be applied to the object.		
Rationale	Some preservation repositories will offer multiple preservation options depending on factors such as the value or uniqueness of the material, the “preservability” of the format, the amount the customer is willing to pay, etc.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Not applicable
Examples	bit-level full 0 1 2	bit-level full 0 fully supported with future migrations	
Repeatability	Not repeatable	Not repeatable	
Obligation	Mandatory	Mandatory	
Creation / Maintenance notes	The preservation level may be assigned by the repository or requested by the depositor and submitted as metadata.		
Usage notes	Only one <i>preservationLevelValue</i> may be recorded per <i>preservationLevel</i> container. If a further <i>preservationLevelValue</i> applies to the object in a different context, a separate <i>preservationLevel</i> container should be repeated.		

Semantic unit	1.3.2 preservationLevelRole		
Semantic components	None		
Definition	A value indicating the context in which a set of preservation options is applicable.		
Rationale	Repositories may assign <i>preservationLevelValues</i> in different contexts which must be differentiated, and may need to record more than one context.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Not applicable
Examples	requirement intention capability	requirement intention capability	
Repeatability	Not repeatable	Not repeatable	
Obligation	Optional	Optional	
Usage notes	<p>This optional semantic unit qualifies the sense or context in which the <i>preservationLevelValue</i> in the current <i>preservationLevel</i> container is applied.</p> <p>For example, a repository may have a legislated obligation to “fully preserve” object X (which is of format F) but is presently only capable of preserving objects of format F at a “bit-level”. The repository may need to record both the required or intended level of preservation (e.g. <i>preservationLevelRole</i>=“requirement”) and the current capability (e.g. <i>preservationLevelRole</i>=“capability”).</p> <p>In transferring custody of material from one repository to another, it may also be important for the receiving repository to know the sense in which <i>preservationLevelValue</i> should be understood. A receiving repository may not need to know a “capability” preservation level of which the transferring repository was capable (as this will have little bearing on its own capabilities), but it needs to know any preservation level “requirements” for material for which it is now taking responsibility.</p> <p>It is good practice to specify <i>preservationLevelRole</i> for clarity even if the repository only assigns <i>preservationLevelValue</i> in one sense or context. If more than one <i>preservationLevel</i> is recorded, <i>preservationLevelRole</i> should always be supplied.</p>		

THE PREMIS DATA DICTIONARY

	If more than one sense or context needs to be expressed for the same object, (e.g. both the “requirement” and “capability” are recorded), separate <i>preservationLevel</i> containers should be used.
--	--

Semantic unit	1.3.3 preservationLevelRationale		
Semantic components	None		
Definition	The reason a particular <i>preservationLevelValue</i> was applied to the object.		
Rationale	Application of a particular <i>preservationLevelValue</i> may require justification, especially if it differs from that usually applied according to repository policy.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Not applicable
Examples	user pays legislation	defective file bit-level preservation only available for this format	
Repeatability	Repeatable	Repeatable	
Obligation	Optional	Optional	
Usage notes	<p>This optional semantic unit records the reason for applying the <i>preservationLevelValue</i>.</p> <p>This information can be particularly important when the assigned <i>preservationLevelValue</i> differs from usual repository policy.</p> <p>For example, a repository may normally assign a <i>preservationLevelValue</i> of “full preservation” for JPEG2000 files, but detects that a particular file is defective. This may mean that the repository’s preservation strategy for JPEG2000 may not be effective for this particular file, so the repository may assign a <i>preservationLevelValue</i> of “bit-level preservation” to this file, recording “defective file” as the rationale.</p> <p>Similarly, legislative requirements or contractual agreements may require a higher level of preservation to be assigned to a particular object than would be assigned to that class of object according to usual policy. In this case, the rationale for the assignment may be recorded as “legislation” or “user pays”, for example.</p> <p><i>preservationLevelRationale</i> may be repeated if more than one reason needs to be recorded.</p>		

THE PREMIS DATA DICTIONARY

Semantic unit	1.3.4 preservationLevelDateAssigned		
Semantic components	None		
Definition	The date, or date and time, when a particular <i>preservationLevelValue</i> was assigned to the object.		
Rationale	The <i>preservationLevel</i> applicable to an object is expected to be reviewed and changed over time, in response to changes in repository preservation requirements, policies, or capabilities relevant to the object. The date that the current <i>preservationLevelValue</i> was assigned aids review of decisions.		
Data constraint	To aid machine processing, value should use a structured form. To facilitate exchange of PREMIS-conformant metadata, use of standard conventions, for instance as used in the date elements in the PREMIS schema, is recommended.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Not applicable
Examples	2007-11-05 2007-11-05T08:15:30-05:00 20080315	2007-11-05 2007-11-05T08:15:30-05:00 20080315	
Repeatability	Not repeatable	Not repeatable	
Obligation	Optional	Optional	

Semantic unit	1.4 significantProperties		
Semantic components	1.4.1 significantPropertiesType 1.4.2 significantPropertiesValue 1.4.3 significantPropertiesExtension		
Definition	Characteristics of a particular object subjectively determined to be important to maintain through preservation actions.		
Rationale	Objects that have the same technical properties may still differ as to the properties that should be preserved for future presentation or use.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional
Creation / Maintenance notes	Significant properties may pertain to all objects of a certain class; for example, the repository can decide that for all PDF files, only the content need be preserved. In other cases, for example, for media art, the significant properties may be unique to each individual object. Where values are unique, they must be supplied by the submitter or provided by the curatorial staff of the repository.		
Usage notes	<p>All of this semantic unit’s subunits are optional. At least one of the <i>significantPropertiesValue</i> and <i>significantPropertiesExtension</i> subunits must be present if this container is included or both may be used.</p> <p>Significant properties may be objective technical characteristics subjectively considered important, or subjectively determined characteristics. For example, a PDF may contain links that are not considered important and JavaScript that is considered important. Or future migrations of a TIFF image may require optimization for line clarity or for color; the option chosen would depend upon a curatorial judgment of the significant properties of the image.</p> <p>Listing significant properties implies that the repository plans to preserve these properties across time and requires them to acceptably survive preservation action; for example, to be maintained during emulation or after format migration. It also implies that the repository would note when preservation action results in modification of significant properties.</p> <p>In practice, significant properties might be used as measures of</p>		

	<p>preservation success, as part of quality checking the results of a preservation action or evaluating the efficacy of a preservation method. For example, if the listed significant properties are not maintained after application of a particular preservation method, it may indicate a failure of the process or that the method is not well suited to the type of material.</p> <p>More experience with digital preservation is needed to determine the best ways of representing significant properties in general, and of representing modification of significant properties.</p> <p>The semantic units included in the <i>significantProperties</i> container aim to provide a flexible structure for describing significant properties, allowing general types of aspects, facets or attributes of an object to be declared and to be paired with specific significant details about the object pertaining to that aspect, facet or attribute.</p> <p>For example, some repositories may define significant properties for objects related to facets of content, appearance, structure, behavior, and context. Examples of facet:detail pairs in this case could include:</p> <p><i>significantPropertiesType</i> = “content” <i>significantPropertiesValue</i> = “all textual content and images” <i>significantPropertiesType</i> = “behavior” <i>significantPropertiesValue</i> = “editable”</p> <p>Other repositories may choose to describe significant properties at a more granular attribute level; for example:</p> <p><i>significantPropertiesType</i> = “page count” <i>significantPropertiesValue</i> = “7” <i>significantPropertiesType</i> = “page width” <i>significantPropertiesValue</i> = “210 mm”</p> <p>Each facet:detail pair should be contained in a separate, repeated <i>significantProperties</i> container.</p> <p>Further work on determining and describing significant properties may yield more detailed schemes to facilitate general description.</p> <p>Representing modification of significant properties as a result of preservation action also requires further work. One possible way involves the use of Object and Event information: Object A has significant properties volume and timing, which are recorded as <i>significantProperties</i> of A. In migrated version B, the timing is modified, which is noted in the <i>eventOutcome</i> of the migration event. Only volume is listed as a significant property of B.</p>
--	--

Semantic unit	1.4.1 significantPropertiesType		
Semantic components	None		
Definition	The aspect, facet, or attribute of an object about which significant properties are being described.		
Rationale	Repositories may choose to describe significant properties based on a particular aspect or attribute of an object.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples	content structure behavior page count page width typeface hyperlinks image count	content structure behavior page count page width typeface	[for an embedded image] color space
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Optional	Optional	Optional
Usage notes	This semantic unit is optional and may be used as part of a facet:detail pair with <i>significantPropertiesValue</i> .		

THE PREMIS DATA DICTIONARY

Semantic unit	1.4.2 significantPropertiesValue		
Semantic components	None		
Definition	Description of the characteristics of a particular object subjectively determined to be important to maintain through preservation actions.		
Rationale	Repositories may choose to describe significant properties based on a particular aspect or attribute of an object.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples	<p>[for a Web page containing animation that is not considered essential] Content only.</p> <p>[For detail associated with a <i>significantProperties</i> Type of “behavior”] “hyperlinks traversable”</p>	<p>[for a word processed document with embedded links that are not considered essential] Content only.</p> <p>[For detail associated with a <i>significantProperties</i> Type of “behavior”] “editable”</p> <p>[For detail associated with a <i>significantProperties</i> Type of “page width”] 210 mm</p>	<p>[for a PDF with an embedded graph, where the lines’ color determines the lines’ meaning] Color.</p> <p>[For detail associated with a <i>significantProperties</i> Type of “appearance”] Color.</p>
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Optional	Optional	Optional
Usage notes	<p>If facet:detail pairs are used, the content of <i>significantPropertiesValue</i> should describe the significant properties of object relevant to the aspect, facet, or attribute declared in the <i>significantPropertiesType</i> with which it is paired.</p> <p>If facet:detail pairs are not used, <i>significantPropertiesValue</i> may be used to freely describe any characteristic of an object.</p> <p><i>significantPropertiesValue</i> is not repeatable. Multiple significant properties should be described in separate, repeated <i>significantProperties</i> container units.</p>		

Semantic unit	1.4.3 significantPropertiesExtension		
Semantic components	Defined externally		
Definition	A container to include semantic units defined outside of PREMIS for significant properties.		
Rationale	There may be a need to replace or extend PREMIS defined units.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional
Usage notes	<p>All of this semantic unit's subunits are optional. At least one of the <i>significantPropertiesValue</i> and <i>significantPropertiesExtension</i> subunits must be present if this container is included.</p> <p>If the <i>significantPropertiesExtension</i> container needs to be associated explicitly with any PREMIS subunit under <i>significantProperties</i>, the container <i>significantProperties</i> is repeated. If extensions from different external schemas are needed, <i>significantProperties</i> should also be repeated.</p>		

THE PREMIS DATA DICTIONARY

Semantic unit	1.5 objectCharacteristics		
Semantic components	1.5.1 compositionLevel 1.5.2 fixity 1.5.3 size 1.5.4 format 1.5.5 creatingApplication 1.5.6 inhibitors 1.5.7 objectCharacteristicsExtension		
Definition	Technical properties of a file or bitstream that are applicable to all or most formats.		
Rationale	There are some important technical properties that apply to objects of any format. Detailed definition of format-specific properties is outside the scope of this Data Dictionary, although such properties may be included within objectCharacteristicsExtension .		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Repeatability		Repeatable	Repeatable
Obligation		Mandatory	Mandatory
Usage notes	<p>The semantic units included in <i>objectCharacteristics</i> should be treated as a set of information that pertains to a single object at a single <i>compositionLevel</i>. Object characteristics may be repeated when an object was created by applying two or more encodings, such as compression and encryption. In this case each repetition of <i>objectCharacteristics</i> would have an incrementally higher <i>compositionLevel</i>.</p> <p>When encryption is applied, the <i>objectCharacteristics</i> block must include an inhibitors semantic unit.</p> <p>A bitstream embedded within a file may have different object characteristics than the file. Where these characteristics are relevant for preservation, they should be recorded.</p> <p>When a single file is equivalent to a representation, <i>objectCharacteristics</i> may be applied and thus associated with the representation. In these cases, the relationship between the file comprising the representation and other associated files may be expressed using <i>relationshipSubType</i> (see page 112).</p>		

Semantic unit	1.5.1 compositionLevel		
Semantic components	None		
Definition	An indication of whether the object is subject to one or more processes of decoding or unbundling.		
Rationale	A file or bitstream can be encoded with compression, encryption, etc., or bundled with other files or bitstreams into larger packages. Knowing the order in which these actions are taken is important if the original object or objects must be recovered.		
Data constraint	Non-negative integers		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		0 1 2	0 1 2
Repeatability		Not repeatable	Not repeatable
Obligation		Mandatory	Mandatory
Creation / Maintenance notes	Composition level will generally be supplied by the repository, which should attempt to supply this value automatically. If the object was created by the repository, the creating routine knows the composition level and can supply this metadata. If the object is being ingested by the repository, repository programs will have to attempt to identify the composition level from the object itself or from externally supplied metadata.		
Usage notes	<p>A file or bitstream can be subject to multiple encodings that must be decoded in reverse order (highest to lowest). For example, file A may be compressed to create file B, which is encrypted to create file C. To recreate a copy of the base file A, one would have to unencrypt file C to create file B and then uncompress file B to create file A.</p> <p>A <i>compositionLevel</i> of zero indicates that the object is a base object and not subject to further decoding, while a level of 1 or higher indicates that one or more decodings must be applied.</p> <p>Numbering goes lowest to highest (first encoded = 0). 0 is base object; 1-n are subsequent encodings.</p> <p>Use 0 as the default if there is only one <i>compositionLevel</i>.</p> <p>When multiple file objects are bundled together as filestreams within</p>		

THE PREMIS DATA DICTIONARY

	<p>a package file object (e.g., a ZIP file), the individual filestream objects are <i>not</i> composition levels of the package file object. They should be considered separate objects, each with their own composition levels. For example, two encrypted files zipped together and stored in an archive as one file object would be described as three separate objects, each with its own associated metadata. The storage location of the two inner objects would point to the ZIP file, but the ZIP file itself would have only a single composition level (of zero) whose format would be “zip.” See “Object characteristics and composition level,” page199..</p>
--	---

Semantic unit	1.5.2 fixity		
Semantic components	1.5.2.1 messageDigestAlgorithm 1.5.2.2 messageDigest 1.5.2.3 messageDigestOriginator		
Definition	Information used to verify whether an object has been altered in an undocumented or unauthorized way.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Not applicable (see usage note)	Applicable	Applicable (see usage note)
Repeatability		Repeatable	Repeatable
Obligation		Optional	Optional
Creation / Maintenance notes	Automatically calculated and recorded by repository.		
Usage notes	<p>To perform a fixity check, a message digest calculated at some earlier time is compared with a message digest calculated at a later time. If the digests are the same, the object was not altered in the interim. Recommended practice is to use two or more message digests calculated by different algorithms. (Note that the terms “message digest” and “checksum” are commonly used interchangeably. However, the term “checksum” is more correctly used for the product of a cyclical redundancy check (CRC), whereas the term “message digest” refers to the result of a cryptographic hash function, which is what is referred to here.)</p> <p>The act of performing a fixity check and the date it occurred would be recorded as an Event. The result of the check would be recorded as the <i>eventOutcome</i>. Therefore, only the <i>messageDigestAlgorithm</i> and <i>messageDigest</i> need to be recorded as <i>objectCharacteristics</i> for future comparison.</p> <p>Representation level: It could be argued that if a representation consists of a single file or if all the files comprised by a representation are combined (e.g., zipped) into a single file, then a fixity check could be performed on the representation. However, in both cases the fixity check is actually being performed on a file, which in this case happens to be coincidental with a representation.</p> <p>Bitstream level: Message digests can be computed for bitstreams although they are not as common as with files. For example, the JPX format, which is a JPEG2000 format, supports the inclusion of MD5</p>		

THE PREMIS DATA DICTIONARY

	<p>or SHA-1 message digests in internal metadata that was calculated on any range of bytes of the file.</p> <p>See “Fixity, integrity, authenticity,” page 200..</p>
--	--

Semantic unit	1.5.2.1 messageDigestAlgorithm		
Semantic components	None		
Definition	The specific algorithm used to construct the message digest for the digital object.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		MD5 Adler-32 HAVAL SHA-1 SHA-256 SHA-384 SHA-512 TIGER WHIRLPOOL	
Repeatability		Not repeatable	Not repeatable
Obligation		Mandatory	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	1.5.2.2 messageDigest		
Semantic components	None		
Definition	The output of the message digest algorithm.		
Rationale	This must be stored so that it can be compared in future fixity checks.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		7c9b35da4f2ebd436f 1cf88e5a39b3a257ed f4a22be3c955ac49da 2e2107b67a1924419 563	
Repeatability		Not repeatable	Not repeatable
Obligation		Mandatory	Mandatory

Semantic unit	1.5.2.3 messageDigestOriginator		
Semantic components	None		
Definition	The agent that created the original message digest that is compared in a fixity check.		
Rationale	A preservation repository may ingest files that have had message digests calculated by the submitter; checking these ensures that the file as received is the same as the file as sent. The repository may also ingest files that do not have message digests, and so must calculate the initial value upon ingest. It can be useful to know who calculated the initial value of the message digest.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		DRS A0000978	
Repeatability		Not repeatable	Not repeatable
Obligation		Optional	Optional
Creation / Maintenance notes	If the calculation of the initial message digest is treated by the repository as an Event, this information could be obtained from an Event record.		
Usage notes	The originator of the message digest could be represented by a string representing the agent (e.g., "NRS" referring to the archive itself) or a pointer to an agent description (e.g., "A0000987" taken here to be an <i>agentIdentifierValue</i>).		

THE PREMIS DATA DICTIONARY

Semantic unit	1.5.3 size		
Semantic components	None		
Definition	The size in bytes of the file or bitstream stored in the repository.		
Rationale	Size is useful for ensuring the correct number of bytes from storage have been retrieved and that an application has enough room to move or process files. It might also be used when billing for storage.		
Data constraint	Integer		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		2038937	
Repeatability		Not repeatable	Not repeatable
Obligation		Optional	Optional
Creation / Maintenance notes	Automatically obtained by the repository.		
Usage notes	Defining this semantic unit as size in bytes makes it unnecessary to record a unit of measurement. However, for the purpose of data exchange the unit of measurement should be stated or understood by both partners.		

Semantic unit	1.5.4 format		
Semantic components	1.5.4.1 formatDesignation 1.5.4.2 formatRegistry 1.5.4.3 formatNote		
Definition	Identification of the format of a file or bitstream where format is the organization of digital information according to preset specifications.		
Rationale	Many preservation activities depend on detailed knowledge about the format of the digital object. An accurate identification of format is essential. The identification provided, whether by name or pointer into a format registry, should be sufficient to associate the object with more detailed format information.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Repeatability		Repeatable	Repeatable
Obligation		Mandatory	Mandatory
Creation / Maintenance notes	The format of a file or bitstream should be ascertained by the repository on ingest. Even if this information is provided by the submitter, directly in metadata or indirectly via the file name extension, recommended practice is to independently identify the format by parsing the file when possible. If the format cannot be identified at the time of ingest, it is valid to record that it is unknown, but the repository should subsequently make an effort to identify the format, even if manual intervention is required.		
Usage notes	<p>A bitstream embedded within a file may have different characteristics than the larger file. For example, a bitstream in LaTeX format could be embedded within an SGML file, or multiple images using different colorspace could be embedded within a TIFF file. <i>format</i> must be recorded for every object. When the bitstream format can be recognized by the repository and the repository might want to treat the bitstream differently from the embedding file for preservation purposes, <i>format</i> can be recorded for embedded bitstreams.</p> <p>Although this semantic unit is mandatory, both of its subunits are optional. At least one subunit (i.e. either <i>formatDesignation</i> or <i>formatRegistry</i>) must be present if this container is included or both may be used. If the subunit (<i>formatDesignation</i> or <i>formatRegistry</i>) needs to be repeated, the entire <i>format</i> container is repeated. This</p>		

THE PREMIS DATA DICTIONARY

	<p>allows for association of format designation with a particular set of format registry information. For example, if the precise format cannot be determined and two <i>format</i> designations are recorded, each is given within a separate <i>format</i> container. The <i>format</i> container may also be repeated for multiple format registry entries.</p> <p>See “Format information,” page195.</p>
--	--

Semantic unit	1.5.4.1 formatDesignation		
Semantic components	1.5.4.1.1 formatName 1.5.4.1.2 formatVersion		
Definition	An identification of the format of the object.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Repeatability		Not repeatable	Not repeatable
Obligation		Optional	Optional
Usage notes	<p>Either <i>formatDesignation</i> or at least one instance of <i>formatRegistry</i> is required. Both may be included.</p> <p>The most specific format (or format profile) should be recorded. A repository (or formats registry) may wish to use multipart format names (e.g., “TIFF_GeoTIFF” or “WAVE_MPEG_BWF”) to achieve this specificity.</p>		

THE PREMIS DATA DICTIONARY

Semantic unit	1.5.4.1.1 formatName		
Semantic components	None		
Definition	A designation of the format of the file or bitstream.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		Text/sgml image/tiff/geotiff Adobe PDF DES PGP base64 unknown	LaTeX
Repeatability		Not repeatable	Not repeatable
Obligation		Mandatory	Mandatory
Usage notes	For unidentified formats, <i>formatName</i> may be recorded as “unknown”.		

Semantic unit	1.5.4.1.2 formatVersion		
Semantic components	None		
Definition	The version of the format named in <i>formatName</i> .		
Rationale	Many authority lists of format names are not granular enough to indicate version, for example, MIME Media types.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		6.0 2003	
Repeatability		Not repeatable	Not repeatable
Obligation		Optional	Optional
Usage notes	If the format is versioned, <i>formatVersion</i> should be recorded. It can be either a numeric or chronological designation.		

THE PREMIS DATA DICTIONARY

Semantic unit	1.5.4.2 formatRegistry		
Semantic components	1.5.4.2.1 formatRegistryName 1.5.4.2.2 formatRegistryKey 1.5.4.2.3 formatRegistryRole		
Definition	Identifies and/or gives further information about the format by reference to an entry in a format registry.		
Rationale	If central format registries are available to the preservation repository, they may provide an excellent way of referencing detailed format information.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Repeatability		Not repeatable	Not repeatable
Obligation		Optional	Optional
Usage notes	<p>Either <i>formatDesignation</i> or at least one instance of <i>formatRegistry</i> is required. If more than one <i>formatRegistry</i> needs to be recorded the format container should be repeated to include each additional set of <i>formatRegistry</i> information.</p> <p>The PREMIS working group assumed that a number of format registries will be developed and maintained to support digital preservation efforts. The proposal for a Global Digital Format Registry (GDFR) (http://hul.harvard.edu/gdfr/documents.html#data), for example, would create a network-accessible registry designed to store detailed specifications on formats and profiles.</p>		

Semantic unit	1.5.4.2.1 formatRegistryName		
Semantic components	None		
Definition	A designation identifying the referenced format registry.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		PRONOM www.nationalarchives.gov.uk/pronom Representation Information Registry Repository	FRED: A format registry demonstration, release 0.07
Repeatability		Not repeatable	Not repeatable
Obligation		Mandatory	Mandatory
Usage notes	This can be a formal name, internally used name, or URI.		

THE PREMIS DATA DICTIONARY

Semantic unit	1.5.4.2.2 formatRegistryKey		
Semantic components	None		
Definition	The unique key used to reference an entry for this format in a format registry.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		info:gdfir/fred/f/tiff TIFF/6.0	
Repeatability		Not repeatable	Not repeatable
Obligation		Mandatory	Mandatory

Semantic unit	1.5.4.2.3 formatRegistryRole		
Semantic components	None		
Definition	The purpose or expected use of the registry.		
Rationale	The same <i>format</i> may be defined in different registries for different purposes. For example, one registry may give detailed format specifications while another has profile information. If multiple registries are recorded, this semantic unit can be used to distinguish among them.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		Specification Validation profile	
Repeatability		Not repeatable	Not repeatable
Obligation		Optional	Optional

THE PREMIS DATA DICTIONARY

Semantic unit	1.5.4.3 formatNote		
Semantic components	None		
Definition	Additional information about format.		
Rationale	Qualifying information may be needed to supplement format designation and registry information or record a status for identification		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		tentative identification disjunction multiple format identifications found	
Repeatability		Repeatable	Repeatable
Obligation		Optional	Optional
Usage notes	The <i>formatNote</i> may contain free text, a reference pointer, or a value from a controlled list.		

Semantic unit	1.5.5 creatingApplication		
Semantic components	1.5.5.1 creatingApplicationName 1.5.5.2 creatingApplicationVersion 1.5.5.3 dateCreatedByApplication 1.5.5.4 creatingApplicationExtension		
Definition	Information about the application that created the object.		
Rationale	<p>Information about the creating application, including the version of the application and the date the file was created, can be useful for problem solving purposes. For example, it is not uncommon for certain versions of software to be known for causing conversion errors or introducing artifacts. It is also useful to determine which rendering software is available for the digital object. For example, if you know that the Distiller program created the PDF file, you know it will be renderable with (among other programs) Adobe Reader.</p>		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Repeatability		Repeatable	Repeatable
Obligation		Optional	Optional
Creation / Maintenance notes	<p>If the object was created by the repository, assignment of creating application information should be straightforward.</p> <p>If the object was created outside the repository, it is possible this information could be supplied by the depositor. It might also be extracted from the file itself; the name of the creating application is often embedded within the file.</p>		
Usage notes	<p>This semantic unit applies to both objects created external to the repository and subsequently ingested, and to objects created by the repository, for example, through migration events.</p> <p>The <i>creatingApplication</i> container is repeatable if more than one application processed the object in turn. For example, a file could be created by Microsoft Word and later turned into a PDF using Adobe Acrobat. Details of both the Word and Acrobat applications may be recorded. However, if both files are stored in the repository, each file should be completely described as an Object entity and linked by using relationship information with a <i>relationshipType</i> “derivation.”</p> <p>It may also be repeated to record the creating application before the</p>		

THE PREMIS DATA DICTIONARY

	<p>object was ingested as well as the creating application used as part of the ingest process. For example, an HTML file was created pre-ingest using Dreamweaver, and the Web crawler Heritrix then captured a snapshot of the files as part of the ingest.</p> <p>The amount of information needed for <i>creatingApplication</i> given here is minimal. For more granularity, extensibility is provided.</p> <p>Rather than having each repository record this locally, it would be preferable to have a registry of this information similar to format or environment registries.</p>
--	---

Semantic unit	1.5.5.1 creatingApplicationName		
Semantic components	None		
Definition	A designation for the name of the software program that created the object.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		MSWord	
Repeatability		Not repeatable	Not repeatable
Obligation		Optional	Optional
Usage notes	The <i>creatingApplication</i> is the application that created the object in its current format, not the application that created the copy written to storage. For example, if a document is created by Microsoft Word and subsequently copied to archive storage by a repository's Ingest program, the <i>creatingApplication</i> is Word, not the Ingest program.		

THE PREMIS DATA DICTIONARY

Semantic unit	1.5.5.2 creatingApplicationVersion		
Semantic components	None		
Definition	The version of the software program that created the object.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		2000	1.4
Repeatability		Not repeatable	Not repeatable
Obligation		Optional	Optional

Semantic unit	1.5.5.3 dateCreatedByApplication		
Semantic components	None		
Definition	The actual or approximate date and time the object was created.		
Data constraint	To aid machine processing, value should use a structured form. To facilitate exchange of PREMIS-conformant metadata, use of standard conventions, for instance as used in the date elements in the PREMIS schema, is recommended.		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		2000-12-01 20030223T151047	
Repeatability		Not repeatable	Not repeatable
Obligation		Optional	Optional
Usage notes	<p>Use the most precise date available.</p> <p>This is the date the object was created by the creating application, not the date any copy was made externally or by the repository. For example, if a file is created by Microsoft Word in 2001 and two copies are made in 2003, the <i>dateCreatedByApplication</i> of all three files is 2001. The date a file is written to storage can be recorded as an Event.</p> <p>If the object itself contains internal creation and modification dates, the modification date should be used as <i>dateCreatedByApplication</i>.</p> <p>If the application is a Web harvester capturing an object at a point of time, use for date captured.</p>		

THE PREMIS DATA DICTIONARY

Semantic unit	1.5.5.4 creatingApplicationExtension		
Semantic components	Defined externally		
Definition	Creating application information using semantic units defined external to PREMIS.		
Rationale	There may be a need to supplement or replace PREMIS defined units.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Repeatability		Repeatable	Repeatable
Obligation		Optional	Optional
Usage notes	<p>For more granularity or use of externally defined semantic units, extensibility is provided. Either local semantic units or metadata using another specified metadata scheme may be included instead of or in addition to PREMIS defined semantic units. When using an externally defined schema, a reference to that schema must be provided. See further guidance “Extensibility,” page19.</p> <p>If <i>creatingApplicationExtension</i> container needs to be associated explicitly with any PREMIS subunit under <i>creatingApplication</i>, the container <i>creatingApplication</i> is repeated. If extensions from different external schemas are needed, <i>creatingApplication</i> should also be repeated.</p>		

Semantic unit	1.5.6 inhibitors		
Semantic components	1.5.6.1 inhibitorType 1.5.6.2 inhibitorTarget 1.5.6.3 inhibitorKey		
Definition	Features of the object intended to inhibit access, use, or migration.		
Rationale	Format information may indicate whether a file is encrypted, but the nature of the encryption also must be recorded, as well as the access key.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Repeatability		Repeatable	Repeatable
Obligation		Optional	Optional
Creation / Maintenance notes	Inhibitors are more likely to be present on an object ingested by the repository than applied by the repository itself. It is often not possible to tell that a file has been encrypted by parsing it; the file may appear to be ASCII text. Therefore, information about inhibitors should be supplied as metadata with submitted objects when possible.		
Usage notes	<p>Some file formats allow encryption for embedded bitstreams.</p> <p>Some file formats such as PDF use passwords to control access to content or specific functions. Although this is actually implemented at the bitstream level, for preservation purposes it is effectively managed at the file level; that is, passwords would not be recorded for individually addressable bitstreams.</p> <p>For certain types of inhibitor keys, more granularity may be required. If the inhibitor key information is identical to key information in digital signatures, use those semantic units.</p>		

THE PREMIS DATA DICTIONARY

Semantic unit	1.5.6.1 inhibitorType		
Semantic components	None		
Definition	The inhibitor method employed.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		DES PGP Blowfish Password protection	
Repeatability		Not repeatable	Not repeatable
Obligation		Mandatory	Mandatory
Usage notes	Common inhibitors are encryption and password protection. When encryption is used the type of encryption should be specifically indicated, that is, record “DES”, not “encryption”.		

Semantic unit	1.5.6.2 inhibitorTarget		
Semantic components	None		
Definition	The content or function protected by the inhibitor.		
Data constraint	Values should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		All content Function: Play Function: Print	
Repeatability		Repeatable	Repeatable
Obligation		Optional	Optional
Usage notes	If not supplied, assume that the target is the content of the object.		

THE PREMIS DATA DICTIONARY

Semantic unit	1.5.6.3 inhibitorKey		
Semantic components	None		
Definition	The decryption key or password.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		[DES decryption key]	
Repeatability		Not repeatable	Not repeatable
Obligation		Optional	Optional
Usage notes	The key should be provided if known. However, it is not advisable to actually store the <i>inhibitorKey</i> in plain text in an unsecure database.		

Semantic unit	1.5.7 objectCharacteristicsExtension		
Semantic components	Defined externally		
Definition	A container to include semantic units defined outside of PREMIS.		
Rationale	There may be a need to replace or extend PREMIS defined units.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Repeatability		Repeatable	Repeatable
Obligation		Optional	Optional
Usage notes	<p>For more granularity or use of externally defined semantic units, extensibility is provided. Either local semantic units or metadata using another specified metadata scheme may be included in addition to PREMIS defined semantic units. When using an extension schema, a reference to that schema must be provided. See further guidance in “Extensibility,” page 19.</p> <p><i>objectCharacteristicsExtension</i> is used for additional object characteristics not covered by PREMIS, for instance format specific metadata that is defined externally. It is not a replacement for units specified in PREMIS.</p> <p>If <i>objectCharacteristicsExtension</i> container needs to be associated explicitly with any PREMIS subunit under <i>objectCharacteristics</i>, the container <i>objectCharacteristics</i> is repeated. If extensions from different external schemas are needed, <i>objectCharacteristics</i> should also be repeated.</p>		

THE PREMIS DATA DICTIONARY

Semantic unit	1.6 originalName		
Semantic components	None		
Definition	The name of the object as submitted to or harvested by the repository, before any renaming by the repository.		
Rationale	The name used within the preservation repository may not be known outside of the repository. A depositor might need to request a file by its original name. Also, the repository may need to reconstruct internal links for dissemination.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Not applicable
Examples		N419.pdf	
Repeatability	Not repeatable	Not repeatable	
Obligation	Optional	Optional	
Creation / Maintenance notes	This value would always be supplied to the repository by the submitter or harvesting application. How much of the file path to preserve would be up to the repository.		
Usage notes	This is the name of the object as designated in the Submission Information Package (SIP). The object may have other names in different contexts. When two repositories are exchanging content, it would be important for the receiving repository to know and record the name of the representation at the originating repository. In the case of representations, this may be a directory name.		

Semantic unit	1.7 storage		
Semantic components	1.7.1 <i>contentLocation</i> 1.7.2 <i>storageMedium</i>		
Definition	Information about how and where a file is stored in the storage system.		
Rationale	It is necessary for a repository to associate the <i>contentLocation</i> with the <i>storageMedium</i> .		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Repeatability		Repeatable	Repeatable
Obligation		Mandatory	Mandatory
Usage notes	<p>Normally there would be a single storage location and medium for an object, because an object in another location would be considered a different object. The storage composite should be repeated if there are two or more copies that are identical bit-wise and managed as a unit except for the medium on which they are stored. They must have a single <i>objectIdentifier</i> and be managed as a single object by the repository.</p> <p>Although this semantic unit is mandatory, both of its subunits are optional. At least one subunit (i.e. either <i>contentLocation</i> or <i>storageMedium</i>) must be present or both may be used.</p>		

THE PREMIS DATA DICTIONARY

Semantic unit	1.7.1 contentLocation		
Semantic components	1.7.1.1 contentLocationType 1.7.1.2 contentLocationValue		
Definition	Information needed to retrieve a file from the storage system, or to access a bitstream within a file.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Repeatability		Not repeatable	Not repeatable
Obligation		Optional	Optional
Creation / Maintenance notes	A preservation repository should never refer to content that it does not control. Therefore, the PREMIS working group assumed that the repository will always assign the <i>contentLocation</i> , probably by program.		
Usage notes	If the preservation repository uses the <i>objectIdentifier</i> as a handle for retrieving data, <i>contentLocation</i> is implicit and does not need to be recorded.		

THE PREMIS DATA DICTIONARY

Semantic unit	1.7.1.1 contentLocationType		
Semantic components	None		
Definition	The means of referencing the location of the content.		
Rationale	To understand the meaning of the value it is necessary to know what location scheme is used.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		URI hdl NTFS EXT3	byte offset
Repeatability		Not repeatable	Not repeatable
Obligation		Mandatory	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	1.7.1.2 contentLocationValue		
Semantic components	None		
Definition	The reference to the location of the content used by the storage system.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		<p>http://wwasearch.loc.gov/107th/200212107035/http://house.gov/langevin/</p> <p>hdl:loc.pnp/cph.3b34188</p> <p>c:\apache2\htdocs\index.html</p> <p>/home/web/public_html/index.html</p>	64 [offset from start of file c:\apache2\htdocs\image\logo.gif]
Repeatability		Not repeatable	Not repeatable
Obligation		Mandatory	Mandatory
Usage notes	This could be a fully qualified path and filename, or the information used by a resolution system (e.g., a handle) or the native information used by a storage management system. For a bitstream or filestream, this would probably be the reference point and offset of the starting position of the bitstream. It is up to the repository to determine the level of granularity that should be recorded.		

Semantic unit	1.7.2 storageMedium		
Semantic components	None		
Definition	The physical medium on which the object is stored (e.g., magnetic tape, hard disk, CD-ROM, DVD).		
Rationale	The repository needs to know the medium on which an object is stored in order to know how and when to do media refreshment and media migration.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		Magnetic tape Hard disk TSM	Magnetic tape Hard disk TSM
Repeatability		Not Repeatable	Not Repeatable
Obligation		Optional	Optional
Usage notes	<p>In some cases this can be masked from direct repository management by storage management systems but the underlying assumption is that the repository ultimately is in control and needs to manage for technological obsolescence.</p> <p>In some cases the value may not be the specific medium, but the system that knows the medium, e.g., Tivoli Storage Manager (TSM).</p> <p>Knowing the storage medium is an internal requirement in order to trigger preservation actions. However, since this is not information that is used for exchange purposes, it is optional.</p>		

THE PREMIS DATA DICTIONARY

Semantic unit	1.8 environment		
Semantic components	1.8.1 environmentCharacteristic 1.8.2 environmentPurpose 1.8.3 environmentNote 1.8.4 dependency 1.8.5 software 1.8.6 hardware 1.8.7 environmentExtension		
Definition	Hardware/software combinations supporting use of the object.		
Rationale	Environment is the means by which the user renders and interacts with content. Separation of digital content from its environmental context can result in the content becoming unusable.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional
Creation / Maintenance notes	<p>This information may be omitted when the repository is doing only bit-level preservation on the object.</p> <p>Rather than having each repository record this locally, it would be preferable to have a registry of environment information similar to proposed registries of format information.</p> <p>Repositories may choose to design mechanisms for inheritance, so that if the environment required for each file within a representation is identical to the environment recorded for the representation as a whole, it is not necessary to store this information in each file.</p> <p>See “Environment,” page197.</p>		
Usage notes	All of this semantic units’ subunits are optional. At least one subunit (i.e. <i>environmentNote</i> , <i>dependency</i> , <i>software</i> , <i>hardware</i> , and/or <i>environmentExtension</i>) must be present if this container is included.		

Semantic unit	1.8.1 environmentCharacteristic		
Semantic components	None		
Definition	An assessment of the extent to which the described environment supports its purpose.		
Rationale	If multiple environments are described, this element can help to distinguish among them.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples	unspecified minimum	recommended minimum	
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Optional	Optional	Optional
Creation / Maintenance notes	This value could be supplied by the submitter or by the repository. If environment software and hardware information is obtained from an environments registry, <i>environmentCharacteristic</i> might also be obtained from the registry. Note however that the criteria for “recommended” may be different for different repositories.		
Usage notes	<p>Suggested values:</p> <p>unspecified = no attempt made to provide this value</p> <p>known to work = the object can be rendered in this environment</p> <p>minimum = the least demanding (in terms of components or resources needed) environment known to work by the repository</p> <p>recommended = an environment preferred for optional rendering</p> <p>If an environment is both “minimum” and “recommended,” use “recommended.”</p> <p>“Known to work” implies the object is supported by the described environment but the repository doesn’t know if this environment is minimum or recommended.</p>		

THE PREMIS DATA DICTIONARY

Semantic unit	1.8.2 environmentPurpose		
Semantic components	None		
Definition	The use(s) supported by the specified environment.		
Rationale	Different environments can support different uses of objects. For example, the environment needed to edit and modify a file can be quite different than the environment needed to render it.		
Data constraint	Values should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional
Creation / Maintenance notes	This value would have to be supplied by the agent that provided the hardware and software environment information, which might be the submitter, the repository, or an environments registry.		
Usage notes	Suggested values: render, edit. This list may need to be expanded. Other values might indicate the ability to transform, print, and manipulate by program.		

Semantic unit	1.8.3 environmentNote		
Semantic components	None		
Definition	Additional information about the environment.		
Rationale	There may be a need to give a textual description of the environment for additional explanation.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples		This environment assumes that the PDF will be stored locally and used with a standalone PDF reader.	
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional
Usage notes	<p>This note could be used to record the context of the environment information. For example, if a file can be rendered through a PC client application or through a browser with a plug-in, this note could be used to identify which situation applies.</p> <p>The note should not be used for a textual description of environment information recorded more rigorously elsewhere.</p>		

THE PREMIS DATA DICTIONARY

Semantic unit	1.8.4 dependency		
Semantic components	1.8.4.1 dependencyName 1.8.4.2 dependencyIdentifier		
Definition	Information about a non-software component or associated file needed in order to use or render the representation or file, for example, a schema, a DTD, or an entity file declaration.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional
Creation / Maintenance notes	Recommended practice is for a repository to archive objects on which other objects depend. These may be sent by the submitter of the primary object, or they may in some cases be automatically obtained by the repository. For example, a markup file will often contain links to other objects it requires such as DTDs or XML Schema. If it does, these objects can often be identified by the link and downloaded by the repository.		
Usage notes	<p>This semantic unit is for additional objects that are necessary to render a file or representation, not for required software or hardware. It may also be used for a non-executable component of the object, such as a font or style sheet. For things that the software requires, see <i>swDependency</i>, page 94.</p> <p>This semantic unit does not include objects required by structural relationships, such as child content objects (e.g., figures that are part of an article), which are recorded under relationship with a <i>relationshipType</i> of “structural”.</p> <p>It is up to the repository to determine what constitutes a dependency in the context of the designated community.</p> <p>The objects noted may be internal or external to the preservation repository.</p>		

Semantic unit	1.8.4.1 dependencyName		
Semantic components	None		
Definition	A designation for a component or associated file needed by the representation or file.		
Rationale	It may not be self-evident from the <i>dependencyIdentifier</i> what the name of the object actually is.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples		Additional Element Set for Language Corpora	
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional

THE PREMIS DATA DICTIONARY

Semantic unit	1.8.4.2 dependencyIdentifier		
Semantic components	1.8.4.2.1 dependencyIdentifierType 1.8.4.2.2 dependencyIdentifierValue		
Definition	A unique designation used to identify a dependent resource.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional
Usage notes	The <i>dependencyIdentifier</i> must be unique within the preservation repository, although it might not be globally unique.		

Semantic unit	1.8.4.2.1 dependencyIdentifierType		
Semantic components	None		
Definition	A designation of the domain in which the identifier of the dependent resource is unique.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples		URI	
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory
Usage notes	A preservation repository needs to know both the type of object identifier and the value. When the value itself contains the identifier type (e.g., “oai:lib.uchicago.edu:1”), the identifier type does not need to be recorded explicitly. Similarly, if the repository uses only one type of identifier, the type can be assumed and does not need to be recorded explicitly.		

THE PREMIS DATA DICTIONARY

Semantic unit	1.8.4.2.2 dependencyIdentifierValue		
Semantic components	None		
Definition	The value of the <i>dependencyIdentifier</i> .		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples		http://www.tei-c.org/P4X/DTD/teicorp2.dtd	
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory

Semantic unit	1.8.5 software		
Semantic components	1.8.5.1 swName 1.8.5.2 swVersion 1.8.5.3 swType 1.8.5.4 swOtherInformation 1.8.5.5 swDependency		
Definition	Software required to render or use the object.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional
Creation / Maintenance notes	<p>If recording this explicitly, many different software environments may apply; for example, a particular object such as a PDF file may be viewable by several versions of several applications running under several operating systems and operating system versions. Although at least one software environment should be recorded, it is not necessary to record them all and each repository will have to make its own decisions about which software environments to record.</p> <p>Also, what appears to the user as a single rendering program can have many dependencies, including system utilities, runtime libraries, and so on, which each might have their own dependencies in turn.</p> <p>As with environment, metadata may be more efficiently managed in conjunction with a format registry either internal or external to a repository. In the absence of a global mechanism, repositories may be forced to develop their own local “registries” relating format to software environment.</p>		

THE PREMIS DATA DICTIONARY

Semantic unit	1.8.5.1 swName		
Semantic components	None		
Definition	Manufacturer and title of the software application.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples	Sybase	Adobe Photoshop Adobe Acrobat Reader	
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory
Usage notes	Include manufacturer when this helps to identify or disambiguate the product, for example, use “Adobe Photoshop” rather than “Photoshop.”		

THE PREMIS DATA DICTIONARY

Semantic unit	1.8.5.2 swVersion		
Semantic components	None		
Definition	The version or versions of the software referenced in <i>swName</i> .		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples		>=2.2.0 6.0 2000	
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Optional	Optional	Optional
Usage notes	If there is no formal version, the date of issuance can be used.		

THE PREMIS DATA DICTIONARY

Semantic unit	1.8.5.3 swType		
Semantic components	None		
Definition	Class or category of <i>software</i> .		
Rationale	Several different layers of software can be required to support an object.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory
Usage notes	<p>Suggested values:</p> <p>renderer = application that can display/play/execute the format instance, e.g., image viewer, video player, Java virtual machine (when the format instance is a Java class file)</p> <p>ancillary = required ancillary software, e.g., run time libraries, browser plug-ins, compression/decompression routines, utilities, operating system emulators, etc.</p> <p>operatingSystem = software that supports application execution, process scheduling, memory management, file systems, etc.</p> <p>driver = software with the primary function of communicating between hardware and the operating system or other software</p>		

Semantic unit	1.8.5.4 swOtherInformation		
Semantic components	None		
Definition	Additional requirements or instructions related to the software referenced in <i>swName</i> .		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples		Install Acroread (Adobe Acrobat) first; copy nppdf.so (the plug-in) to your Mozilla plug-ins directory, and make sure a copy of (or symlink to) Acroread is in your PATH.	
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional
Usage notes	This could be a reliable persistent identifier or URI pointing to software documentation within or outside the repository.		

THE PREMIS DATA DICTIONARY

Semantic unit	1.8.5.5 swDependency		
Semantic components	None		
Definition	The name and, if applicable, version of any software component needed by the software referenced in <i>swName</i> in the context of using this object.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples		GNU gcc >= 2.7.2	
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional
Usage notes	The value should be constructed in a way that is consistent with the construction of <i>swName</i> and <i>swVersion</i> . This semantic unit identifies the software that is needed by what is recorded in <i>swName</i> , for example, a Perl script that depends on a Perl module. In this case the Perl script is listed in <i>swName</i> , with the module in <i>swDependency</i> within a software container.		

Semantic unit	1.8.6 hardware		
Semantic components	1.8.6.1 hwName 1.8.6.2 hwType 1.8.6.3 hwOtherInformation		
Definition	Hardware components needed by the software referenced in <i>swName</i> or the human user of the referenced software.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional
Creation / Maintenance notes	<p>Hardware environment information can be very difficult to provide. Many different hardware environments may apply; there are a huge number of combinations of maker and type of CPU, memory, video drivers, and so on. Although at least one hardware environment should be recorded, it is not necessary to record them all and each repository will have to make its own decisions about which hardware environments to record.</p> <p>Because of the difficulty recording this information comprehensively, it would be optimal if central registries of environment information existed. In many cases the environment of a file object is directly associated with the format, making registry lookup by format feasible. In the absence of a global mechanism, repositories may be forced to develop their own local “registries” relating format to <i>hwEnvironment</i>.</p>		

THE PREMIS DATA DICTIONARY

Semantic unit	1.8.6.1 hwName		
Semantic components	None		
Definition	Manufacturer, model, and version (if applicable) of the hardware.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples		Intel Pentium III 1 GB DRAM Windows XP-compatible joystick	
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory
Usage notes	<p>Include manufacturer when this helps to identify or disambiguate the product.</p> <p>Include version for firmware or other components where that information is pertinent.</p>		

Semantic unit	1.8.6.2 hwType		
Semantic components	None		
Definition	Class or category of the hardware.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory
Usage notes	Suggested values: processor, memory, input/output device, storage device.		

THE PREMIS DATA DICTIONARY

Semantic unit	1.8.6.3 hwOtherInformation		
Semantic components	None		
Definition	Additional requirements or instructions related to the hardware referenced in <i>hwName</i> .		
Rationale	For hardware, the amount of computing resource needed (such as memory, storage, processor speed, etc.) may need to be documented. In addition, more detailed instructions may be needed to install and/or operate the hardware.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples	32MB minimum	32MB minimum Required RAM for Apache is unknown	
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional
Usage notes	This could be an identifier or URI used to point to hardware documentation.		

Semantic unit	1.8.7 environmentExtension		
Semantic components	Defined externally		
Definition	A container to include semantic units defined outside of PREMIS.		
Rationale	There may be a need to replace or extend PREMIS defined units.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples			
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional
Usage notes	<p>For more granularity or use of externally defined semantic units, extensibility is provided. Either local semantic units or metadata using another specified metadata scheme may be included instead of or in addition to PREMIS defined semantic units. When using an extension schema, a reference to that schema must be provided. See further guidance in “Extensibility,” page19.</p> <p>If <i>environmentExtension</i> container needs to be associated explicitly with any PREMIS subunit under <i>environment</i>, the container <i>environment</i> is repeated. If extensions from different external schemas are needed, <i>environment</i> should also be repeated.</p>		

THE PREMIS DATA DICTIONARY

Semantic unit	1.9 signatureInformation		
Semantic components	1.9.1 signature 1.9.2 signatureInformationExtension		
Definition	A container for PREMIS defined and externally defined digital signature information, used to authenticate the signer of an object and/or the information contained in the object.		
Rationale	A repository may have a policy of generating digital signatures for files on ingest, or may have a need to store and later validate incoming digital signatures.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Repeatability		Repeatable	Repeatable
Obligation		Optional	Optional
Usage notes	Either <i>signature</i> or <i>signatureInformationExtension</i> may be used. Use of <i>signatureInformationExtension</i> with the schema defined in W3C's <i>XML-Signature Syntax and Processing</i> (www.w3.org/TR/2002/REC-xmlsig-core-20020212/) is encouraged when applicable. See the discussion of digital signatures on page 201 for more information on use of both PREMIS-defined and externally-defined semantic units.		

Semantic unit	1.9.1 signature		
Semantic components	1.9.1.1 signatureEncoding 1.9.1.2 signer 1.9.1.3 signatureMethod 1.9.1.4 signatureValue 1.9.1.5 signatureValidationRules 1.9.1.6 signatureProperties 1.9.1.7 keyInformation		
Definition	Information needed to use a digital signature to authenticate the signer of an object and/or the information contained in the object.		
Rationale	A repository may have a policy of generating digital signatures for files on ingest, or may have a need to store and later validate incoming digital signatures.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Repeatability		Repeatable	Repeatable
Obligation		Optional	Optional
Usage notes	Several of the semantic components of <i>signatureInformation</i> are taken from the W3C's <i>XML-Signature Syntax and Processing</i> ; see www.w3.org/TR/2002/REC-xmlsig-core-20020212/ for more information on the structure and application of these semantic units. (See also the discussion of digital signatures, page201.)		

THE PREMIS DATA DICTIONARY

Semantic unit	1.9.1.1 signatureEncoding		
Semantic components	None		
Definition	The encoding used for the values of <i>signatureValue</i> , <i>keyInformation</i> .		
Rationale	These values cannot be interpreted correctly if the encoding is unknown.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		Base64 Ds:CryptoBinary	
Repeatability		Not repeatable	Not repeatable
Obligation		Mandatory	Mandatory

Semantic unit	1.9.1.2 signer		
Semantic components	None		
Definition	The individual, institution, or authority responsible for generating the signature.		
Rationale	The signer might also be carried in the <i>keyInformation</i> , but it can be accessed more conveniently if recorded here.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Repeatability		Not repeatable	Not repeatable
Obligation		Optional	Optional
Usage notes	If the signer is an <i>Agent</i> known to the repository, an <i>agentIdentifier</i> can be used here.		

THE PREMIS DATA DICTIONARY

Semantic unit	1.9.1.3 signatureMethod		
Semantic components	None		
Definition	A designation for the encryption and hash algorithms used for signature generation.		
Rationale	The same algorithms must be used for signature validation.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		DSA-SHA1 RSA-SHA1	
Repeatability		Not repeatable	Not repeatable
Obligation		Mandatory	Mandatory
Usage notes	Recommended practice is to encode the encryption algorithm first, followed by a hyphen, followed by the hash (message digest) algorithm.		

THE PREMIS DATA DICTIONARY

Semantic unit	1.9.1.4 signatureValue		
Semantic components	None		
Definition	The digital signature; a value generated from the application of a private key to a message digest.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Examples		juS5RhJ884qoFR 8flVXd/rbrSDVGn 40CapgB7qeQiT +rr0NekEQ6BHh UA8dT3+BCTBU QI0dBjlm19lwzEN XvS83zRECjzXb MRTUtVZiPZG2p qKPnL2YU3A964 5UCjTXU+jgFum v7k78hieAGDzNc i+PQ9KRmm//icT 7JaYztgt4=	
Repeatability		Not repeatable	Not repeatable
Obligation		Mandatory	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	1.9.1.5 signatureValidationRules		
Semantic components	None		
Definition	The operations to be performed in order to validate the digital signature.		
Rationale	The repository should not assume that the procedure for validating any particular signature will be known many years in the future without documentation.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability		Applicable	Applicable
Repeatability		Not repeatable	Not repeatable
Obligation		Mandatory	Mandatory
Usage notes	<p>This may include the canonicalization method used before calculating the message digest, if the object was normalized before signing.</p> <p>This value could also be a pointer to archive documentation.</p>		

Semantic unit	1.9.1.6 signatureProperties		
Semantic components	None		
Definition	Additional information about the generation of the signature.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Repeatability		Repeatable	Repeatable
Obligation		Optional	Optional
Usage notes	This may include the date/time of signature generation, the serial number of the cryptographic hardware used, or other information related to the generation of the signature. Repositories will likely want to define a suitably granular structure to <i>signatureProperties</i> .		

THE PREMIS DATA DICTIONARY

Semantic unit	1.9.1.7 keyInformation		
Semantic components	Extensible container		
Definition	Information about the signer's public key needed to validate the digital signature.		
Rationale	To validate a digital signature for an object, one first recalculates the message digest for the object, and then uses the public key of the signer to verify that the value of the signature (<i>signatureValue</i>) is correct. The repository must therefore have the public key value and some assurance that it truly belongs to the signer.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Not applicable	Applicable	Applicable
Repeatability		Not repeatable	Not repeatable
Obligation		Optional	Optional
Usage notes	Different types of keys will have different structures and parameters. PREMIS does not define structure for this container. Recommended practice is to represent key values as defined for "KeyInfo" in the W3C's <i>XML-Signature Syntax and Processing</i> (www.w3.org/TR/2002/REC-xmlsig-core-20020212/).		

Semantic unit	1.9.2 signatureInformationExtension		
Semantic components	Defined externally		
Definition	Digital signature information using semantic units defined outside of PREMIS.		
Rationale	There may be a need to replace or extend PREMIS defined units.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability		Applicable	Applicable
Repeatability		Repeatable	Repeatable
Obligation		Optional	Optional
Usage notes	<p>For more granularity or use of externally defined semantic units, extensibility is provided. Either local semantic units or metadata using another specified metadata scheme may be included instead of or in addition to PREMIS defined semantic units. When using an extension schema, a reference to that schema must be provided. See further guidance in “Extensibility,” page19.</p> <p>If <i>signatureInformationExtension</i> container needs to be associated explicitly with any PREMIS subunit under <i>signatureInformation</i>, the container <i>signatureInformation</i> is repeated. If extensions from different external schemas are needed, <i>signatureInformation</i> should also be repeated.</p> <p>Use of the W3C’s <i>XML-Signature Syntax and Processing</i> (www.w3.org/TR/2002/REC-xmlsig-core-20020212/) is encouraged when applicable.</p>		

THE PREMIS DATA DICTIONARY

Semantic unit	1.10 relationship		
Semantic components	1.10.1 relationshipType 1.10.2 relationshipSubType 1.10.3 relatedObjectIdentification 1.10.4 relatedEventIdentification		
Definition	Information about a relationship between this object and one or more other objects.		
Rationale	A preservation repository must know how to assemble complex objects from component parts (structural relationships) and rigorously track digital provenance (derivation relationships). Documentation about relationships between different objects is crucial to these purposes.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional
Usage notes	<p>Most preservation repositories will want to record all relevant relationships.</p> <p>In complex scenarios, PREMIS might not be able to express rich enough structural relationships to be the only source of structural metadata.</p> <p>Many formats for representing structural information may be used instead of the semantic units specified here. This information must be known, and some implementations may know it by using other structures.</p> <p>Structural relationships at the file level are necessary to reconstruct a representation in order to ascertain that the representation is renderable.</p> <p>A record of structural relationships at the representation level may be necessary to render the representation.</p> <p>Structural relationships at the bitstream level can relate bitstreams within a file.</p> <p>Derivative relationships at the file and representation level are important for documenting digital provenance.</p>		

Semantic unit	1.10.1 relationshipType		
Semantic components	None		
Definition	A high-level categorization of the nature of the relationship.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory
Usage notes	<p>Suggested values:</p> <p>structural = a relationship between parts of an object</p> <p>derivation = a relationship where one object is the result of a transformation performed on the related object</p> <p>A repository may find it necessary to define additional relationship types.</p>		

THE PREMIS DATA DICTIONARY

Semantic unit	1.10.2 relationshipSubType		
Semantic components	None		
Definition	A specific characterization of the nature of the relationship documented in <i>relationshipType</i> .		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory
Usage notes	<p>Suggested values:</p> <p>has sibling = the object shares a common parent with the related object</p> <p>is part of = the object is contained by the related object (when these are the same entity types)</p> <p>has part = the object contains the related object (when these are the same entity types)</p> <p>is source of = the related object is a version of this object created by a transformation</p> <p>has source = the object is derived from the related object as a result of a transformation</p> <p>has root = for a representation only, the related object is the file that must be processed first in order to render the representation</p> <p>includes = for the relationship of a representation to a file, , or a file to a bitstream, the described object includes the referenced object</p> <p>is included in = for the relationship of a file to a representation, or a bitstream to a file, the described object is included in the referenced object</p> <p>A repository may find it necessary to define more or less granular relationships. For derivation relationships, note that the precise relationship may be indicated by the type of the related event.</p> <p>The relationship “has root” is applicable only to the representation, because it implies that a compound object (i.e., one made up of multiple files) requires that one file be picked up first as its root to render it. In the metadata for the representation, “has root” identifies that particular file.</p>		

Semantic unit	1.10.3 relatedObjectIdentification		
Semantic components	1.10.3.1 relatedObjectIdentifierType 1.10.3.2 relatedObjectIdentifierValue 1.10.3.3 relatedObjectSequence		
Definition	The identifier and sequential context of the related resource.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Mandatory	Mandatory	Mandatory
Usage notes	The related object may or may not be held within the preservation repository. Recommended practice is that objects reside within the repository unless there is a good reason to reference an object outside. Internal and external references should be clear.		

THE PREMIS DATA DICTIONARY

Semantic unit	1.10.3.1 relatedObjectIdentifierType		
Semantic components	None		
Definition	A designation of the domain within which the identifier is unique.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples	[see examples for <i>objectIdentifierType</i>]	[see examples for <i>objectIdentifierType</i>]	[see examples for <i>objectIdentifierType</i>]
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory
Usage notes	If the related object is held within the preservation repository, this should be the value of that object's <i>objectIdentifierType</i> .		

Semantic unit	1.10.3.2 relatedObjectIdentifierValue		
Semantic components	None		
Definition	The value of the related object identifier.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples	[see examples for <i>objectIdentifierValue</i>]	[see examples for <i>objectIdentifierValue</i>]	[see examples for <i>objectIdentifierValue</i>]
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory
Usage notes	If the related object is held within the preservation repository, this should be the value of that object's <i>objectIdentifierValue</i> .		

THE PREMIS DATA DICTIONARY

Semantic unit	1.10.3.3 relatedObjectSequence		
Semantic components	None		
Definition	The order of the related object relative to other objects with the same type of relationship.		
Rationale	This semantic unit is particularly useful for structural relationships. In order to reconstruct a representation, it may be necessary to know the order of components with sibling or part-whole relationships. For example, to render a page-image book, it is necessary to know the order of files representing pages.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples		1 2 3	
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Optional	Optional	Optional
Usage notes	<p>This semantic unit could be implemented in several ways. It might be recorded explicitly in metadata as a sequence number or as a pointer. It might be implicit in some other ordering of objects, for example, incrementing identifier values. The value of <i>relationshipSubType</i> might imply the sequence (e.g., “is preceding sibling,” “is following sibling”).</p> <p>There is no requirement that sequence numbers must be unique or sequential.</p> <p>Some related objects have no inherent sequence, for example, unordered Web pages making up a Web site. In this case all related objects can be given the “dummy” sequence number zero.</p> <p>This semantic unit is applicable only for structural relationships and is thus optional.</p>		

Semantic unit	1.10.4 relatedEventIdentification		
Semantic components	1.10.4.1 relatedEventIdentifierType 1.10.4.2 relatedEventIdentifierValue 1.10.4.3 relatedEventSequence		
Definition	The identifier and contextual sequence of an event associated with the relationship.		
Rationale	An object may be related to another object because of an event, for example, migration.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional
Usage notes	For derivative relationships between objects <i>relatedEventIdentification</i> must be recorded.		

THE PREMIS DATA DICTIONARY

Semantic unit	1.10.4.1 relatedEventIdentifierType		
Semantic components	None		
Definition	The <i>eventIdentifierType</i> of the related event.		
Data constraint	Must be an existing <i>eventIdentifierType</i> value.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples	[see examples for <i>eventIdentifierType</i>]	[see examples for <i>eventIdentifierType</i>]	[see examples for <i>eventIdentifierType</i>]
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory
Usage notes	For most preservation repositories, the <i>eventIdentifierType</i> will simply be its own internal numbering system. It can be implicit within the system and provided explicitly only if the data is exported.		

Semantic unit	1.10.4.2 relatedEventIdentifierValue		
Semantic components	None		
Definition	The <i>eventIdentifierValue</i> of the related event.		
Data constraint	Must be an existing <i>eventIdentifierValue</i> value.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples	[see examples for <i>eventIdentifierValue</i>]	[see examples for <i>eventIdentifierValue</i>]	[see examples for <i>eventIdentifierValue</i>]
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	1.10.4.3 relatedEventSequence		
Semantic components	None		
Definition	The order of the related event.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples		1 2 3	
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Optional	Optional	Optional
Usage notes	The sequence of a related event can be inferred from the <i>eventDateTime</i> associated with the related event.		

Semantic unit	1.11 linkingEventIdentifier		
Semantic components	1.11.1 linkingEventIdentifierType 1.11.2 linkingEventIdentifierValue		
Definition	The <i>eventIdentifier</i> of an event associated with the object.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional
Usage notes	Use to link to events that are not associated with relationships between objects, such as format validation, virus checking, etc.		

THE PREMIS DATA DICTIONARY

Semantic unit	1.11.1 linkingEventIdentifierType		
Semantic components	None		
Definition	The <i>eventIdentifierType</i> value of the related event.		
Data constraint	Must be an existing <i>eventIdentifierType</i> value.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples	[see examples for <i>eventIdentifierType</i>]	[see examples for <i>eventIdentifierType</i>]	[see examples for <i>eventIdentifierType</i>]
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory
Usage notes	For most preservation repositories, the <i>eventIdentifierType</i> will simply be their own internal numbering system. It can be implicit within the system and provided explicitly only if the data is exported.		

Semantic unit	1.11.2 linkingEventIdentifierValue		
Semantic components	None		
Definition	The <i>eventIdentifierValue</i> value of the related event.		
Data constraint	Must be an existing <i>eventIdentifierValue</i> value.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples	[see examples for <i>eventIdentifierValue</i>]	[see examples for <i>eventIdentifierValue</i>]	[see examples for <i>eventIdentifierValue</i>]
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	1.12 linkingIntellectualEntityIdentifier		
Semantic components	1.12.1 linkingIntellectualEntityIdentifierType 1.12.2 linkingIntellectualEntityIdentifierValue		
Definition	An identifier for an intellectual entity associated with the object.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional
Usage notes	Use to link to an intellectual entity that is related to the object. This may be a link to descriptive metadata that describes the intellectual entity or some other surrogate for it that can be referenced. This link will likely be to an identifier of an object that is at a higher conceptual level than the object for which the metadata is provided, for example, to a collection or parent object.		

Semantic unit	1.12.1 linkingIntellectualEntityIdentifierType		
Semantic components	None		
Definition	A designation of the domain within which the <i>linkingIntellectualEntityIdentifier</i> is unique.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples		URI LCCN	
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	1.12.2 linkingIntellectualEntityIdentifierValue		
Semantic components	None		
Definition	The value of the <i>linkingIntellectualEntityIdentifier</i> .		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples	hdl:loc.natlib/mrva00 02.0495 info:lccn/19018302		
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory

Semantic unit	1.13 linkingRightsStatementIdentifier		
Semantic components	1.13.1 linkingRightsStatementIdentifierType 1.13.2 linkingRightsStatementIdentifierValue		
Definition	An identifier for a rights statement associated with the object.		
Rationale	A repository may choose to link from a rights statement to an object or from an object to a rights statement or both.		
Data constraint	Container		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Repeatable	Repeatable	Repeatable
Obligation	Optional	Optional	Optional

THE PREMIS DATA DICTIONARY

Semantic unit	1.13.1 linkingRightsStatementIdentifierType		
Semantic components	None		
Definition	A designation of the domain within which the <i>linkingRightsStatementIdentifier</i> is unique.		
Data constraint	Value should be taken from a controlled vocabulary.		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Examples		URI LCCN	
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	1.13.2 linkingRightsStatementIdentifierValue		
Semantic components	None		
Definition	The value of the <i>linkingRightsStatementIdentifier</i> .		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Applicable
Repeatability	Not repeatable	Not repeatable	Not repeatable
Obligation	Mandatory	Mandatory	Mandatory

THE PREMIS DATA DICTIONARY

Event Entity

The Event entity aggregates information about an action that involves one or more Object entities. Metadata about an Event would normally be recorded and stored separately from the digital object.

Whether or not a preservation repository records an Event depends upon the importance of the event. Actions that modify objects should always be recorded. Other actions such as copying an object for backup purposes may be recorded in system logs or an audit trail but not necessarily in an Event entity.

Mandatory semantic units are: *eventIdentifier*, *eventType*, and *eventDateTime*.

Entity properties

- Must be related to one or more objects.
- Can be related to one or more agents.
- Links between entities may be recorded from either direction and need not be bi-directional.

Entity semantic units

- 2.1 eventIdentifier (M, NR)
 - 2.1.1 eventIdentifierType (M, NR)
 - 2.1.2 eventIdentifierValue (M, NR)
- 2.2 eventType (M, NR)
- 2.3 eventDateTime (M, NR)
- 2.4 eventDetail (O, NR)
- 2.5 eventOutcomeInformation (O, R)
 - 2.5.1 eventOutcome (O, NR)
 - 2.5.2 eventOutcomeDetail (O, R)
 - 2.5.2.1 eventOutcomeDetailNote (O, NR)
 - 2.5.2.2 eventOutcomeDetailExtension (O, R)
- 2.6 linkingAgentIdentifier (O, R)
 - 2.6.1 linkingAgentIdentifierType (M, NR)
 - 2.6.2 linkingAgentIdentifierValue (M, NR)
 - 2.6.3 linkingAgentRole (O, R)
- 2.7 linkingObjectIdentifier (O, R)
 - 2.7.1 linkingObjectIdentifierType (M, NR)
 - 2.7.2 linkingObjectIdentifierValue (M, NR)
 - 2.7.3 linkingObjectRole (O, R)

Semantic unit	2.1 eventIdentifier
Semantic components	2.1.1 eventIdentifierType 2.1.2 eventIdentifierValue
Definition	A designation used to uniquely identify the event within the preservation repository system.
Rationale	Each event recorded by the preservation archive must have a unique identifier to allow it to be related to objects, agents, and other events.
Data constraint	Container
Repeatability	Not repeatable
Obligation	Mandatory
Creation / Maintenance notes	The <i>eventIdentifier</i> is likely to be system generated. There is no global scheme or standard for these identifiers. The identifier is therefore not repeatable.

THE PREMIS DATA DICTIONARY

Semantic unit	2.1.1 eventIdentifierType
Semantic components	None
Definition	A designation of the domain within which the event identifier is unique.
Data constraint	None
Examples	FDA Stanford Repository Event ID UUID
Repeatability	Not repeatable
Obligation	Mandatory
Creation / Maintenance notes	For most preservation repositories, the <i>eventIdentifierType</i> will be its own internal numbering system. It can be implicit within the system and provided explicitly only if the data is exported.

Semantic unit	2.1.2 eventIdentifierValue
Semantic components	None
Definition	The value of the <i>eventIdentifier</i> .
Data constraint	None
Examples	[a binary integer] E-2004-11-13-000119 58f202ac-22cf-11d1-b12d-002035b29092
Repeatability	Not repeatable
Obligation	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	2.2 eventType
Semantic components	None
Definition	A categorization of the nature of the event.
Rationale	Categorizing events will aid the preservation repository in machine processing of event information, particularly in reporting.
Data constraint	Value should be taken from a controlled vocabulary.
Examples	E77 [a code used within a repository for a particular event type] Ingest
Repeatability	Not repeatable
Obligation	Mandatory
Usage notes	<p>Each repository should define its own controlled vocabulary of <i>eventType</i> values. A suggested starter list for consideration (see also the Glossary for more detailed definitions):</p> <p>capture = the process whereby a repository actively obtains an object</p> <p>compression = the process of coding data to save storage space or transmission time</p> <p>creation = the act of creating a new object</p> <p>deaccession = the process of removing an object from the inventory of a repository</p> <p>decompression = the process of reversing the effects of compression</p> <p>decryption = the process of converting encrypted data to plaintext</p> <p>deletion = the process of removing an object from repository storage</p> <p>digital signature validation = the process of determining that a decrypted digital signature matches an expected value</p> <p>dissemination = the process of retrieving an object from repository storage and making it available to users</p> <p>fixity check = the process of verifying that an object has not been changed in a given period</p> <p>ingestion = the process of adding objects to a preservation repository</p> <p>message digest calculation = the process by which a message digest (“hash”) is created</p> <p>migration = a transformation of an object creating a version in a more contemporary format</p>

	<p>normalization = a transformation of an object creating a version more conducive to preservation</p> <p>replication = the process of creating a copy of an object that is, bit-wise, identical to the original</p> <p>validation = the process of comparing an object with a standard and noting compliance or exceptions</p> <p>virus check = the process of scanning a file for malicious programs</p> <p>Note that migration, normalization, and replication are more precise subtypes of the creation event. “Creation” can be used when more precise terms do not apply, for example, when a digital object was first created by scanning from paper.</p> <p>In general, the level of specificity in recording the type of event (e.g., whether the <i>eventType</i> indicates a transformation, a migration or a particular method of migration) is implementation specific and will depend upon how reporting and processing is done.</p> <p>Recommended practice is to record detailed information about the event itself in <i>eventDetail</i> rather than using a very granular value for <i>eventType</i>.</p>
--	---

THE PREMIS DATA DICTIONARY

Semantic unit	2.3 eventDateTime
Semantic components	None
Definition	The single date and time, or date and time range, at or during which the event occurred.
Data constraint	To aid machine processing, value should use a structured form. To facilitate exchange of PREMIS-conformant metadata, use of standard conventions, for instance as used in the date elements in the PREMIS schema, is recommended.
Examples	20050704T071530-0500 [July 4, 2005 at 7:15:30 a.m. EST] 2006-07-16T19:20:30+01:00 20050705T0715-0500/20050705T0720-0500 [from 7:15 a.m. EST to 7:20 a.m. EST on July 4, 2005] 2004-03-17 [March 17, 2004, only the date is known]
Repeatability	Not repeatable
Obligation	Mandatory
Usage notes	Recommended practice is to record the most specific time possible and to designate the time zone.

Semantic unit	2.4 eventDetail
Semantic components	None
Definition	Additional information about the event.
Data constraint	None
Examples	Object permanently withdrawn by request of Caroline Hunt. Program="MIGJP2JP2K"; version="2.2"
Repeatability	Not repeatable
Obligation	Optional
Usage notes	<i>eventDetail</i> is not intended to be processed by machine. It may record any information about an event and/or point to information stored elsewhere.

THE PREMIS DATA DICTIONARY

Semantic unit	2.5 eventOutcomeInformation
Semantic components	2.5.1 eventOutcome 2.5.2 eventOutcomeDetail
Definition	Information about the outcome of an event.
Data constraint	Container
Repeatability	Repeatable
Obligation	Optional
Usage notes	<p>A repository may wish to supplement a coded <i>eventOutcome</i> value with additional information in <i>eventOutcomeDetail</i>. Since events may have more than one outcome, the container is repeatable.</p> <p>All subunits of this semantic unit are optional. At least one subunit (i.e. <i>eventOutcome</i> or <i>eventOutcomeDetail</i>) must be present if this container is included.</p>

Semantic unit	2.5.1 eventOutcome
Semantic components	None
Definition	A categorization of the overall result of the event in terms of success, partial success, or failure.
Rationale	A coded way of representing the outcome of an event may be useful for machine processing and reporting. If, for example, a fixity check fails, the event record provides both an actionable and a permanent record.
Data constraint	Value should be taken from a controlled vocabulary.
Examples	00 [a code meaning “action successfully completed”] CV-01 [a code meaning “checksum validated”]
Repeatability	Not repeatable
Obligation	Optional
Usage notes	Recommended practice is to use a controlled vocabulary that a system can act upon automatically. More detail about the outcome may be recorded in <i>eventOutcomeDetail</i> . Recommended practice is to define events with sufficient granularity that each event has a single outcome.

THE PREMIS DATA DICTIONARY

Semantic unit	2.5.2 eventOutcomeDetail
Semantic components	2.5.2.1 eventOutcomeDetailNote 2.5.2.2 eventOutcomeDetailExtension
Definition	A detailed description of the result or product of the event.
Rationale	An event outcome may be sufficiently complex that a coded description is not adequate to document it.
Data constraint	Container
Repeatability	Repeatable
Obligation	Optional
Usage notes	<p>This may be used to record all error and warning messages issued by a program involved in the event or to record a pointer to an error log.</p> <p>If the event was a validity check (e.g., profile conformance) any anomalies or quirks discovered would be recorded here.</p> <p>All subunits of this semantic unit are optional. At least one subunit (i.e. <i>eventOutcomeDetailNote</i> and/or <i>eventOutcomeDetailExtension</i>) must be present if this container is included.</p>

Semantic unit	2.5.2.1 eventOutcomeDetailNote
Semantic components	None
Definition	A detailed description of the result or product of the event in textual form.
Rationale	Additional information in textual form may be needed about the outcome of the event.
Data constraint	None
Examples	LZW compressed file Non-standard tags found in header
Repeatability	Not repeatable
Obligation	Optional

THE PREMIS DATA DICTIONARY

Semantic unit	2.5.2.2 eventOutcomeDetailExtension
Semantic components	Defined externally
Definition	A container to include semantic units defined outside of PREMIS.
Rationale	There may be a need to replace or extend PREMIS defined units.
Data constraint	Container
Repeatability	Repeatable
Obligation	Optional
Usage notes	<p>For more granularity or use of externally defined semantic units, extensibility is provided. Either local semantic units or metadata using another specified metadata scheme may be included instead of or in addition to PREMIS defined semantic units. When using an extension schema, a reference to that schema must be provided. See further guidance in “Extensibility,” page19.</p> <p>If <i>eventOutcomeDetailExtension</i> container needs to be associated explicitly with any PREMIS subunit under <i>eventOutcomeDetail</i>, the container <i>eventOutcomeDetail</i> is repeated. If extensions from different external schemas are needed, <i>eventOutcomeDetail</i> should also be repeated.</p>

Semantic unit	2.6 linkingAgentIdentifier
Semantic components	2.6.1 linkingAgentIdentifierType 2.6.2 linkingAgentIdentifierValue 2.6.3 linkingAgentRole
Definition	Information about an agent associated with an event.
Rationale	Digital provenance requires often that relationships between agents and events are documented.
Data constraint	Container
Repeatability	Repeatable
Obligation	Optional
Usage notes	Recommended practice is to record the agent if possible.

THE PREMIS DATA DICTIONARY

Semantic unit	2.6.1 linkingAgentIdentifierType
Semantic components	None
Definition	A designation of the domain in which the linking agent identifier is unique.
Data constraint	Value should be taken from a controlled vocabulary.
Examples	[see examples for agentIdentifierType]
Repeatability	Not repeatable
Obligation	Mandatory

Semantic unit	2.6.2 linkingAgentIdentifierValue
Semantic components	None
Definition	The value of the linking agent identifier.
Data constraint	None
Examples	[see examples for <i>agentIdentifierValue</i>]
Repeatability	Not repeatable
Obligation	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	2.6.3 linkingAgentRole
Semantic components	None
Definition	The role of the agent in relation to this event.
Rationale	Events can have more than one agent associated with them. The role of each agent may need to be documented.
Data constraint	Values should be taken from a controlled vocabulary.
Examples	Authorizer Implementer Validator Executing program
Repeatability	Repeatable
Obligation	Optional

Semantic unit	2.7 linkingObjectIdentifier
Semantic components	2.7.1 linkingObjectIdentifierType 2.7.2 linkingObjectIdentifierValue 2.7.3 linkingObjectRole
Definition	Information about an object associated with an event.
Rationale	Digital provenance often requires that relationships between objects and events are documented.
Data constraint	Container
Repeatability	Repeatable
Obligation	Optional

THE PREMIS DATA DICTIONARY

Semantic unit	2.7.1 linkingObjectIdentifierType
Semantic components	None
Definition	A designation of the domain in which the linking object identifier is unique.
Data constraint	Value should be taken from a controlled vocabulary.
Examples	[see examples for <i>objectIdentifierType</i>]
Repeatability	Not repeatable
Obligation	Mandatory

Semantic unit	2.7.2 linkingObjectIdentifierValue
Semantic components	None
Definition	The value of the linking object identifier.
Data constraint	None
Examples	[see examples for <i>objectIdentifierValue</i>]
Repeatability	Not repeatable
Obligation	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	2.7.3 linkingObjectRole
Semantic components	None
Definition	The role of the object associated with an event.
Rationale	Distinguishes the role of the object in relation to an event. If this is not explicit it is necessary to analyze the relationship between objects in the object metadata.
Data constraint	None
Examples	source outcome
Repeatability	Repeatable
Obligation	Optional

Agent Entity

The Agent entity aggregates information about attributes or characteristics of agents (persons, organizations, or software) associated with rights management and preservation events in the life of a data object. Agent information serves to identify an agent unambiguously from all other Agent entities.

The only mandatory semantic unit is *agentIdentifier*.

Entity properties

- May hold or grant one or more rights.
- May carry out, authorize, or compel one or more events.
- May create or act upon one or more objects through an event or with respect to a rights statement.

Entity semantic units

- 3.1 agentIdentifier (R, M)
 - 3.1.1 agentIdentifierType (M, NR)
 - 3.1.2 agentIdentifierValue (M, NR)
- 3.2 agentName (O, R)
- 3.3 agentType (O, NR)

THE PREMIS DATA DICTIONARY

Semantic unit	3.1 agentIdentifier
Semantic components	3.1.1 agentIdentifierType 3.1.2 agentIdentifierValue
Definition	The designation used to uniquely identify the agent within a preservation repository system.
Rationale	Each agent associated with the preservation repository must have a unique identifier to allow it to be related to events and rights statements.
Data constraint	Container
Repeatability	Repeatable
Obligation	Mandatory
Creation / Maintenance notes	An identifier may be created by the repository system, or it may be created or assigned outside of the repository. Similarly, identifiers can be automatically or manually generated. Recommended practice is for repositories to use an identifier automatically created by the repository as the primary identifier in order to ensure that identifiers are unique and usable by the repository. Externally assigned identifiers can be used as secondary identifiers in order to link an agent to information held outside the repository.
Usage notes	Identifiers must be unique within the repository. The <i>agentIdentifier</i> is repeatable in order to allow both repository-assigned and externally-assigned identifiers to be recorded. See Creation/Maintenance note above.

Semantic unit	3.1.1 agentIdentifierType
Semantic components	None
Definition	A designation of the domain in which the agent identifier is unique.
Data constraint	Value should be taken from a controlled vocabulary.
Examples	LCNAF SAN MARC Organization Codes URI
Repeatability	Not repeatable
Obligation	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	3.1.2 agentIdentifierValue
Semantic components	None
Definition	The value of the <i>agentIdentifier</i> .
Data constraint	None
Examples	92-79971 Owens, Erik C. 234-5676 MH-CS info:lccn/n78890351
Repeatability	Not repeatable
Obligation	Mandatory
Usage notes	May be a unique key or a controlled textual form of name.

Semantic unit	3.2 agentName
Semantic components	None
Definition	A text string which could be used in addition to <i>agentIdentifier</i> to identify an agent.
Rationale	This semantic unit provides a more reader-friendly version of the agent identified by the agentIdentifier.
Data constraint	None
Examples	Erik Owens Woodyard Pc
Repeatability	Repeatable
Obligation	Optional
Usage notes	The value is not necessarily unique.

THE PREMIS DATA DICTIONARY

Semantic unit	3.3 agentType
Semantic components	None
Definition	A high-level characterization of the type of agent.
Data constraint	Value should be taken from a controlled vocabulary.
Repeatability	Not repeatable
Obligation	Optional
Usage notes	Suggested values: person organization software

Rights Entity

For the purpose of the PREMIS Data Dictionary, statements of rights and permissions are taken to be constructs that can be described as the Rights entity. Rights are entitlements allowed to agents by copyright or other intellectual property law. Permissions are powers or privileges granted by agreement between a rightsholder and another party or parties.

A repository might wish to record a variety of rights information including abstract rights statements and statements of permissions that apply to external agents and to objects not held within the repository. The minimum core rights information that a preservation repository must know, however, is what rights or permissions a repository has to carry out actions related to objects within the repository. These may be granted by copyright law, by statute, or by a license agreement with the rightsholder.

If the repository records rights information, either *rightsStatement* or *rightsExtension* must be present.

Entity properties

- May be related to one or more objects.
- May be related to one or more agents.
- Links between entities may be recorded from either direction and need not be bi-directional.

Entity semantic units

- 4.1 *rightsStatement* (O, R)
 - 4.1.1 *rightsStatementIdentifier* (M, NR)
 - 4.1.1.1 *rightsStatementIdentifierType* (M, NR)
 - 4.1.1.2 *rightsStatementIdentifierValue* (M, NR)
 - 4.1.2 *rightsBasis* (M, NR)
 - 4.1.3 *copyrightInformation* (O, NR)
 - 4.1.3.1 *copyrightStatus* (M, NR)
 - 4.1.3.2 *copyrightJurisdiction* (M, NR)
 - 4.1.3.3 *copyrightStatusDeterminationDate* (O, NR)
 - 4.1.3.4 *copyrightNote* (O, R)
 - 4.1.4 *licenseInformation* (O, NR)
 - 4.1.4.1 *licenseIdentifier* (O, NR)
 - 4.1.4.1.1 *licenseIdentifierType* (M, NR)
 - 4.1.4.1.2 *licenseIdentifierValue* (M, NR)
 - 4.1.4.2 *licenseTerms* (O, NR)
 - 4.1.4.3 *licenseNote* (O, R)
 - 4.1.5 *statuteInformation* (O, R)
 - 4.1.5.1 *statuteJurisdiction* (M, NR)

THE PREMIS DATA DICTIONARY

- 4.1.5.2 statuteCitation (M, NR)
- 4.1.5.3 statuteInformationDeterminationDate (O, NR)
- 4.1.5.4 statuteNote (O, R)
- 4.1.6 rightsGranted (O, R)
 - 4.1.6.1 act (M, NR)
 - 4.1.6.2 restriction (O, R)
 - 4.1.6.3 termOfGrant (M, NR)
 - 4.1.6.3.1 startDate (M, NR)
 - 4.1.6.3.2 endDate (O, NR)
 - 4.1.6.4 rightsGrantedNote (O, R)
- 4.1.7 linkingObjectIdentifier (O, R)
 - 4.1.7.1 linkingObjectIdentifierType (M, NR)
 - 4.1.7.2 linkingObjectIdentifierValue (M, NR)
- 4.1.8 linkingAgentIdentifier (O, R)
 - 4.1.8.1 linkingAgentIdentifierType (M, NR)
 - 4.1.8.2 linkingAgentIdentifierValue (M, NR)
 - 4.1.8.3 linkingAgentRole (M, NR)
- 4.2 rightsExtension (O, R)

Semantic unit	4.1 rightsStatement
Semantic components	<p>4.1.1 rightsStatementIdentifier</p> <p>4.1.2 rightsBasis</p> <p>4.1.3 copyrightInformation</p> <p>4.1.4 licenseInformation</p> <p>4.1.5 statuteInformation</p> <p>4.1.6 rightsGranted</p> <p>4.1.7 linkingObjectIdentifier</p> <p>4.1.8 linkingAgentIdentifier</p>
Definition	Documentation of the repository's right to perform one or more acts.
Data constraint	Container
Repeatability	Repeatable
Obligation	Optional
Usage notes	<p>This semantic unit is optional because in some cases rights may be unknown. Institutions are encouraged to record rights information when possible.</p> <p>Either <i>rightsStatement</i> or <i>rightsExtension</i> must be present if the Rights entity is included.</p> <p>The <i>rightsStatement</i> should be repeated when the act(s) described has more than one basis, or when different acts have different bases.</p>

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.1 rightsStatementIdentifier
Semantic components	4.1.1.1 rightsStatementIdentifierType 4.1.1.2 rightsStatementIdentifierValue
Definition	The designation used to uniquely identify the rights statement within a preservation repository system.
Rationale	Each statement of rights associated with the preservation repository must have a unique identifier to allow it to be related to events and agents.
Data constraint	Container
Repeatability	Not repeatable
Obligation	Mandatory
Creation/ maintenance notes	The <i>rightsStatementIdentifier</i> is likely to be system generated. There is no global scheme or standard for these identifiers. The identifier is therefore not repeatable.
Usage notes	Identifiers must be unique within the repository.

Semantic unit	4.1.1.1 rightsStatementIdentifierType
Semantic components	None
Definition	A designation of the domain within which the rights statement identifier is unique.
Data constraint	Value should be taken from a controlled vocabulary.
Repeatability	Not repeatable
Obligation	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.1.2 rightsStatementIdentifierValue
Semantic components	None
Definition	The value of the <i>rightsStatementIdentifier</i> .
Data constraint	None
Repeatability	Not repeatable
Obligation	Mandatory

Semantic unit	4.1.2 rightsBasis
Semantic components	None
Definition	Designation of the basis for the right or permission described in the <i>rightsStatementIdentifier</i> .
Data constraint	Values should be taken from a controlled vocabulary.
Repeatability	Not repeatable
Obligation	Mandatory
Usage notes	<p>Suggested values: copyright, license, statute.</p> <p>When <i>rightsBasis</i> is “copyright”, <i>copyrightInformation</i> should be provided.</p> <p>When <i>rightsBasis</i> is “license”, <i>licenseInformation</i> should be provided.</p> <p>When <i>rightsBasis</i> is “statute”, <i>statuteInformation</i> should be provided.</p> <p>If the basis for the rights is the item is public domain, use “copyright”. If the basis is Fair Use, use “statute”.</p> <p>If more than one basis applies, the entire rights entity should be repeated.</p>

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.3 copyrightInformation
Semantic components	4.1.3.1 copyrightStatus 4.1.3.2 copyrightJurisdiction 4.1.3.3 copyrightStatusDeterminationDate 4.1.3.4 copyrightNote
Definition	Information about the copyright status of the object(s).
Data constraint	Container
Repeatability	Not repeatable
Obligation	Optional
Usage notes	<p>When <i>rightsBasis</i> is “copyright”, <i>copyrightInformation</i> should be provided.</p> <p>Repositories may need to extend this with more detailed information. See the California Digital Library's copyrightMD schema (www.cdlib.org/inside/projects/rights/schema/) for an example of a more detailed scheme.</p>

Semantic unit	4.1.3.1 copyrightStatus
Semantic components	None
Definition	A coded designation for the copyright status of the object at the time the rights statement is recorded.
Data constraint	Values should be taken from a controlled vocabulary.
Repeatability	Not repeatable
Obligation	Mandatory
Usage notes	Suggested values: copyrighted = Under copyright. publicdomain = In the public domain. unknown = Copyright status of the resource is unknown.

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.3.2 copyrightJurisdiction
Semantic components	None
Definition	The country whose copyright laws apply.
Rationale	Copyright law can vary from country to country.
Data constraint	Values should be taken from ISO 3166.
Example	us de
Repeatability	Not repeatable
Obligation	Mandatory

Semantic unit	4.1.3.3 copyrightStatusDeterminationDate
Semantic components	None
Definition	The date that the copyright status recorded in <i>copyrightStatus</i> was determined.
Data constraint	To aid machine processing, value should use a structured form. To facilitate exchange of PREMIS-conformant metadata, use of standard conventions, for instance as used in the date elements in the PREMIS schema, is recommended.
Examples	20070608
Repeatability	Not repeatable
Obligation	Optional

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.3.4 copyrightNote
Semantic components	None
Definition	Additional information about the copyright status of the object.
Data constraint	None
Examples	Copyright expiration expected in 2010 unless renewed. Copyright statement is embedded in file header.
Repeatability	Repeatable
Obligation	Optional

Semantic unit	4.1.4 licenseInformation
Semantic components	4.1.4.1 licenseIdentifier 4.1.4.2 licenseTerms 4.1.4.3 licenseNote
Definition	Information about a license or other agreement granting permissions related to an object.
Data constraint	Container
Repeatability	Not repeatable
Obligation	Optional
Usage notes	When <i>rightsBasis</i> is “license”, <i>licenseInformation</i> should be provided.

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.4.1 licenseIdentifier
Semantic components	4.1.4.1.1 licenseIdentifierType 4.1.4.1.2 licenseIdentifierValue
Definition	A designation used to identify the granting agreement uniquely within the repository system.
Data constraint	Container
Repeatability	Not repeatable
Obligation	Optional
Usage notes	<p>This semantic unit is intended to refer to a document recording the granting of permission. For some repositories this may be a formal signed contract with a customer. If the granting agreement is verbal, this could point to a memo by the repository documenting the verbal agreement.</p> <p>The identifier is optional because the agreement may not be stored in a repository with an identifier. In the case of a verbal agreement, for example, the entire agreement may be included or described in the <i>licenseTerms</i>.</p>

Semantic unit	4.1.4.1.1 licenseIdentifierType
Semantic components	None
Definition	A designation of the domain within which the license identifier is unique.
Data constraint	Value should be taken from a controlled vocabulary.
Repeatability	Not repeatable
Obligation	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.4.1.2 licenseIdentifierValue
Semantic components	None
Definition	The value of the <i>licenseIdentifier</i> .
Data constraint	None
Repeatability	Not repeatable
Obligation	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.4.2 licenseTerms
Semantic components	None
Definition	Text describing the license or agreement by which permission was granted.
Data constraint	None
Repeatability	Not repeatable
Obligation	Optional
Usage notes	This could contain the actual text of the license or agreement or a paraphrase or summary.

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.4.3 licenseNote
Semantic components	None
Definition	Additional information about the license.
Data constraint	None
Example	License is embedded in XMP block in file header.
Repeatability	Repeatable
Obligation	Optional
Usage notes	Information about the terms of the license should go in <i>licenseTerms</i> . <i>licenseNotes</i> is intended for other types of information related to the license, such as contact persons, action dates, or interpretations. The note may also indicate the location of the license, for example, if it is available online or embedded in the object itself.

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.5 statuteInformation
Semantic components	4.1.5.1 statuteJurisdiction 4.1.5.2 statuteCitation 4.1.5.3 statuteInformationDeterminationDate 4.1.5.4 statuteNote
Definition	Information about the statute allowing use of the object.
Data constraint	Container
Repeatability	Repeatable
Obligation	Optional
Usage notes	When <i>rightsBasis</i> is “statute”, <i>statuteInformation</i> should be provided.

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.5.1 statuteJurisdiction
Semantic components	None
Definition	The country or other political body enacting the statute.
Rationale	The connection between the object and the rights granted is based on jurisdiction.
Data constraint	Values should be taken from a controlled vocabulary.
Example	us de
Repeatability	Not repeatable
Obligation	Mandatory

Semantic unit	4.1.5.2 statuteCitation
Semantic components	None
Definition	An identifying designation for the statute.
Data constraint	None
Example	Legal Deposit (Jersey) Law 200- National Library of New Zealand (Te Puna Mātauranga o Aotearoa) Act 2003 no 19 part 4 s 34
Repeatability	Not repeatable
Obligation	Mandatory
Usage notes	Use standard citation form when applicable.

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.5.3 statuteInformationDeterminationDate
Semantic components	None
Definition	The date that the determination was made that the statute authorized the permission(s) noted.
Rationale	The permission in question may be the subject of some interpretation. These assessments are made within a specific context and at a specific time. At another time the context, and therefore the assessment, could change. For this reason it can be important to record the date of the decision.
Data constraint	To aid machine processing, value should use a structured form. To facilitate exchange of PREMIS-conformant metadata, use of standard conventions, for instance as used in the date elements in the PREMIS schema, is recommended.
Examples	2007-12-01 20040223151047.0
Repeatability	Not repeatable
Obligation	Optional

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.5.4 statuteNote
Semantic components	None
Definition	Additional information about the statute.
Data constraint	None
Example	Applicability to web-published content sent for review by general counsel 9/19/2008.
Repeatability	Repeatable
Obligation	Optional

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.6 rightsGranted
Semantic components	4.6.1 act 4.6.2 restriction 4.6.3 termOfGrant 4.6.4 rightsGrantedNote
Definition	The action(s) that the granting agency has allowed the repository.
Data constraint	Container
Repeatability	Repeatable
Obligation	Optional

Semantic unit	4.1.6.1 act
Semantic components	None
Definition	The action the preservation repository is allowed to take.
Data constraint	Value should be taken from a controlled vocabulary.
Repeatability	Not repeatable
Obligation	Mandatory
Usage notes	<p>Suggested values:</p> <p>replicate = make an exact copy</p> <p>migrate = make a copy identical in content in a different file format</p> <p>modify = make a version different in content</p> <p>use = read without copying or modifying (e.g., to validate a file or run a program)</p> <p>disseminate = create a copy or version for use outside of the preservation repository</p> <p>delete = remove from the repository</p> <p>It is up to the preservation repository to decide how granular the controlled vocabulary should be. It may be useful to employ the same controlled values that the repository uses for <i>eventType</i>.</p>

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.6.2 restriction
Semantic components	None
Definition	A condition or limitation on the act.
Data constraint	None
Examples	No more than three Allowed only after one year of archival retention has elapsed Rightsholder must be notified after completion of act
Repeatability	Repeatable
Obligation	Optional

Semantic unit	4.1.6.3 termOfGrant
Semantic components	4.1.6.3.1 startDate 4.1.6.3.2 endDate
Definition	The time period for the permissions granted.
Rationale	The permission to preserve may be time bounded.
Data constraint	Container
Repeatability	Not repeatable
Obligation	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.6.3.1 startDate
Semantic components	None
Definition	The beginning date of the permission granted.
Data constraint	To aid machine processing, value should use a structured form. To facilitate exchange of PREMIS-conformant metadata, use of standard conventions, for instance as used in the date elements in the PREMIS schema, is recommended.
Examples	2006-01-02 20050723
Repeatability	Not repeatable
Obligation	Mandatory

Semantic unit	4.1.6.3.2 endDate
Semantic components	None
Definition	The ending date of the permission granted.
Data constraint	To aid machine processing, value should use a structured form. To facilitate exchange of PREMIS-conformant metadata, use of standard conventions, for instance as used in the date elements in the PREMIS schema, is recommended.
Examples	2010-01-02 20120723
Repeatability	Not repeatable
Obligation	Optional
Usage notes	Use “OPEN” for an open ended term of grant. Omit <i>endDate</i> if the ending date is unknown or the permission statement applies to many objects with different end dates.

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.6.4 rightsGrantedNote
Semantic components	None
Definition	Additional information about the rights granted.
Rationale	A textual description of the rights granted may be needed for additional explanation.
Data constraint	None
Repeatability	Repeatable
Obligation	Optional
Usage notes	This semantic unit may include a statement about risk assessment, for example, when a repository is not certain about what permissions have been granted.

Semantic unit	4.1.7 linkingObjectIdentifier
Semantic components	4.1.7.1 linkingObjectIdentifierType 4.1.7.2 linkingObjectIdentifierValue
Definition	The identifier on an object associated with the rights statement.
Rationale	Rights statements must be associated with the objects to which they pertain, either by linking from the rights statement to the object(s) or by linking from the object(s) to the rights statement. This provides the mechanism for the link from the rights statement to an object.
Data constraint	Container
Repeatability	Repeatable
Obligation	Optional
Usage notes	<i>linkingObjectIdentifier</i> is optional because in some cases it will be more practical to link from the object(s) to the rights statement; for example, a repository may have a single rights statement covering thousands of public domain objects.

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.7.1 linkingObjectIdentifierType
Semantic components	None
Definition	A designation of the domain in which the linking object identifier is unique.
Data constraint	Value should be taken from a controlled vocabulary.
Examples	[see examples for <i>objectIdentifierType</i>]
Repeatability	Not repeatable
Obligation	Mandatory

Semantic unit	4.1.7.2 linkingObjectIdentifierValue
Semantic components	None
Definition	The value of the <i>linkingObjectIdentifier</i> .
Data constraint	None
Examples	[see examples for <i>objectIdentifierValue</i>]
Repeatability	Not repeatable
Obligation	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.8 linkingAgentIdentifier
Semantic components	4.1.8.1 linkingAgentIdentifierType 4.1.8.2 linkingAgentIdentifierValue 4.1.8.3 linkingAgentRole
Definition	Identification of one or more agents associated with the rights statement.
Rationale	Rights statements may be associated with related agents, either by linking from the rights statement to the agent(s) or by linking from the agents(s) to the rights statement. This provides the mechanism for the link from the rights statement to the agent.
Data constraint	Container
Repeatability	Repeatable
Obligation	Optional
Usage notes	<i>linkingAgentIdentifier</i> is optional because a relevant agent may be unknown, or in no agent may be relevant. The latter is likely when the rights basis is statute.

Semantic unit	4.1.8.1 linkingAgentIdentifierType
Semantic components	None
Definition	A designation of the domain in which the linking agent identifier is unique.
Data constraint	Value should be taken from a controlled vocabulary.
Examples	[see examples for agentIdentifierType]
Repeatability	Not repeatable
Obligation	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	4.1.8.2 linkingAgentIdentifierValue
Semantic components	None
Definition	The value of the <i>linkingAgentIdentifier</i> .
Data constraint	None
Examples	[see examples for agentIdentifierValue]
Repeatability	Not repeatable
Obligation	Mandatory

Semantic unit	4.1.8.3 linkingAgentRole
Semantic components	None
Definition	The role of the agent in relation to the rights statement.
Data constraint	Values should be taken from a controlled vocabulary.
Examples	contact creator publisher rightsholder grantor
Repeatability	Not repeatable
Obligation	Mandatory

THE PREMIS DATA DICTIONARY

Semantic unit	4.2 rightsExtension
Semantic components	Defined externally
Definition	A container to include semantic units defined outside of PREMIS.
Rationale	There may be a need to replace or extend PREMIS defined units.
Data constraint	Container
Repeatability	Repeatable
Obligation	Optional
Usage notes	<p>For more granularity or use of externally defined semantic units, extensibility is provided. Either local semantic units or metadata using another specified metadata scheme may be included instead of or in addition to PREMIS defined semantic units. When using an extension schema, a reference to that schema must be provided. See further guidance in “Extensibility,” page19.</p> <p>Either <i>rightsStatement</i> or <i>rightsExtension</i> must be present if the Rights entity is included.</p> <p>If <i>rightsExtension</i> container needs to be associated explicitly with any PREMIS subunit under <i>rights</i>, the container <i>rights</i> is repeated. If extensions from different external schemas are needed, <i>rights</i> should also be repeated.</p>

SPECIAL TOPICS

The PREMIS Special Topics section (pages 195-207) has been removed from this excerpt; it is available in a separate excerpt and in the full document. All the PREMIS documents are available online at: <http://www.loc.gov/standards/premis/>

METHODOLOGY

METHODOLOGY

The PREMIS Methodology section (page 208) has been removed from this excerpt; it is available in a separate excerpt and in the full document. All the PREMIS documents are available online at: <http://www.loc.gov/standards/premis/>

GLOSSARY

The PREMIS Glossary section (pages 209-215) has been removed from this excerpt; it is available in a separate excerpt and in the full document. All the PREMIS documents are available online at: <http://www.loc.gov/standards/premis/>

NOTES

NOTES

The notes contain only those references used in this excerpt but the reference numbers from the full document were retained. As a result, there are some missing numbers, which are for notes that are not relevant to this excerpt.

¹⁰ Metadata Encoding & Transmission Standard (METS), <http://www.loc.gov/standards/mets/>.

¹¹ The Dublin Core Metadata Element Set, <http://www.dublincore.org/documents/dces/>.

¹⁶ MARC 21, <http://www.loc.gov/marc/>.

¹⁷ Metadata Object Description Schema (MODS), <http://www.loc.gov/standards/mods/>.

¹⁸ *Content Standard for Digital Geospatial Metadata*, FGDC-STD-001-1998, <http://www.fgdc.gov/metadata/csdlgm/>.

¹⁹ VRA Core 4.0, <http://www.vraweb.org/projects/vracore4/>.

²⁰ Encoded Archival Description (EAD), <http://www.loc.gov/ead/>.

²¹ Data Documentation Initiative (DDI), <http://www.ddialliance.org/>.

²² vCard, <http://www.imc.org/pdi/>.

²³ Metadata Authority Description Schema (MADS), <http://www.loc.gov/standards/mads/>.