Rosetta AIP Data Model

ExLibris The bridge to knowledge

CONFIDENTIAL INFORMATION

The information herein is the property of Ex Libris Ltd. or its affiliates and any misuse or abuse will result in economic loss. DO NOT COPY UNLESS YOU HAVE BEEN GIVEN SPECIFIC WRITTEN AUTHORIZATION FROM EX LIBRIS LTD.

This document is provided for limited and restricted purposes in accordance with a binding contract with Ex Libris Ltd. or an affiliate. The information herein includes trade secrets and is confidential.

DISCLAIMER

The information in this document will be subject to periodic change and updating. Please confirm that you have the most current documentation. There are no warranties of any kind, express or implied, provided in this documentation, other than those expressly agreed upon in the applicable Ex Libris contract. This information is provided AS IS. Unless otherwise agreed, Ex Libris shall not be liable for any damages for use of this document, including, without limitation, consequential, punitive, indirect or direct damages.

Any references in this document to third-party material (including third-party Web sites) are provided for convenience only and do not in any manner serve as an endorsement of that third-party material or those Web sites. The third-party materials are not part of the materials for this Ex Libris product and Ex Libris has no liability for such materials.

TRADEMARKS

"Ex Libris," the Ex Libris bridge, Primo, Aleph, Alephino, Voyager, SFX, MetaLib, Verde, DigiTool, Preservation, URM, Voyager, ENCompass, Endeavor eZConnect, WebVoyage, Citation Server, LinkFinder and LinkFinder Plus, and other marks are trademarks or registered trademarks of Ex Libris Ltd. or its affiliates.

The absence of a name or logo in this list does not constitute a waiver of any and all intellectual property rights that Ex Libris Ltd. or its affiliates have established in any of its products, features, or service names or logos.

Trademarks of various third-party products, which may include the following, are referenced in this documentation. Ex Libris does not claim any rights in these trademarks. Use of these marks does not imply endorsement by Ex Libris of these third-party products, or endorsement by these third parties of Ex Libris products.

Oracle is a registered trademark of Oracle Corporation.

UNIX is a registered trademark in the United States and other countries, licensed exclusively through X/Open Company Ltd.

Microsoft, the Microsoft logo, MS, MS-DOS, Microsoft PowerPoint, Visual Basic, Visual C++, Win32,

Microsoft Windows, the Windows logo, Microsoft Notepad, Microsoft Windows Explorer, Microsoft Internet Explorer, and Windows NT are registered trademarks and ActiveX is a trademark of the Microsoft Corporation in the United States and/or other countries.

Unicode and the Unicode logo are registered trademarks of Unicode, Inc.

Google is a registered trademark of Google, Inc.

Copyright Ex Libris Limited, 2022. All rights reserved.

Document released: June 2022

Web address: http://www.exlibrisgroup.com

Table of Contents

Introduction **6**

Purpose 6

Scope 6

Data Model Structure **8**

Intellectual Entity *10*
Representations *10*
Files *10*
Bitstreams *11*

METS – Metadata Encoding and Transmission Standard **12**

Descriptive *12*
Administrative *13*
Structural Map *13*

METS Sections 13

Header (metsHdr) *13*
Descriptive Metadata (dmdSec) *13*
Administrative Metadata *14*
File Groups *16*
Structural Map *17*
Structural Links *19*
Behavior *19*
Events *19*
Access Rights *22*

METS XML Sections 22

Collections in METS 41

Rosetta DNX Profile **44**

DNX and PREMIS 45

Differences in the Data Model *45*

DNX Section Structure 45

Structure of a Repeatable Section 46

Events within DNX 46

Provenance Events 47
Access Rights Within DNX 48
Significant Properties of Files Within DNX 49
DNX Sections 50
General IE/Rep/File Characteristics 52
(Rosetta) Object Characteristics 53
CMS 53
Web Harvesting 54
Producer 54
Producer Agent 54
Access Rights Policy 55
Granted Rights Statement 55
Metadata (Deprecated) 56
Retention Policy 56
Internal Identifier 57
Object Identifier 57
Preservation Level 58
Significant Properties 58
File Fixity 59
File Format 59
File Virus Check 59
File Validation 60
File Technical Metadata Extraction 60
Validation Stack Outcome 61
Creating Application 61
Inhibitors 62
Object Characteristics Extension 62
Environment 62
Environment Dependencies 63
Environment Software 63
Environment Software Registry 64
Environment Hardware 64
Environment Hardware Registry 64
Environment Extension 65
Signature Information 65
Signature Information Extension 66
Relationship 66
IE Relationship 66

Linking IE Identifier 67
Event 67
Linking Rights Statement Identifier 68
Collection 68

Appendix A – METS XML **69**

Appendix B – DNX Data Dictionary **70**

DNX Controlled Lists 99

1

Introduction

Purpose

Ex Libris Rosetta is a digital-object preservation solution that conforms to the ISO-recognized Open Archival Information System (OAIS). The purpose of this document is to describe the uniqueness of the AIP data model in Rosetta and the way it implements the PREMIS reference model using METS.

Scope

The OAIS reference model defines three types of information packages that need to be managed by a preservation system:

- SIP – Submission Information Package
- AIP – Archival Information Package
- DIP – Dissemination Information Package

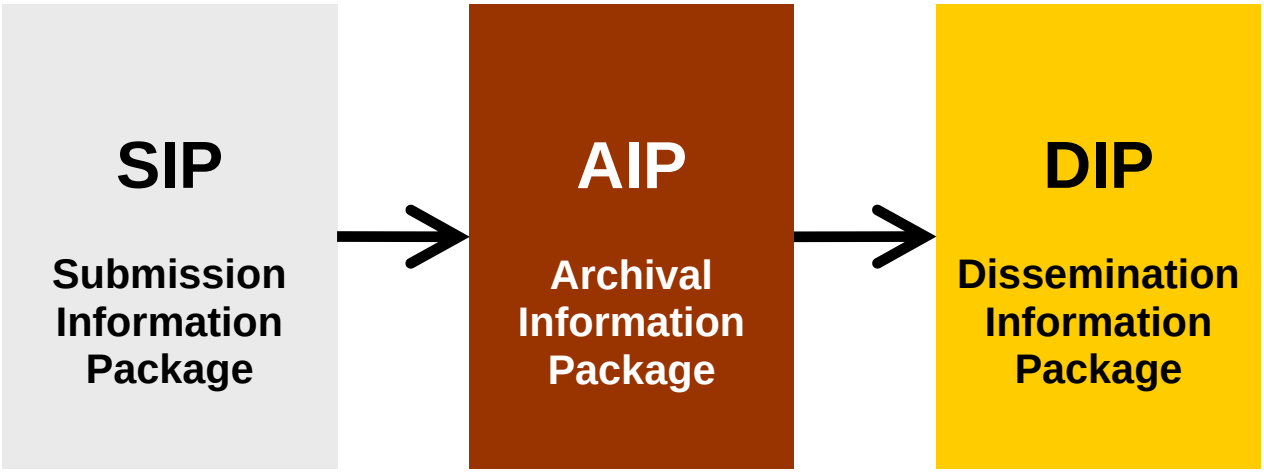
Rosetta allows the SIP and the DIP to have a variety of formats and structures and provides an SDK to support this (refer to the SDK documentation).

This document describes and refers to the AIP, which is stored in a METS XML file in Rosetta's permanent repository module. Each AIP describes one intellectual entity (IE).

The METS XML is generated in the Staging module during the SIP processing. During processing, the IE information is kept and managed in the database. By the time the SIP is moved to the permanent repository, the METS XML contains all the information regarding the IE, collected from the different database tables.

The information on the METS XML can be reloaded back into the database when the IE is brought from the permanent repository for maintenance (preservation actions, adding representations, and so forth).

The following diagram shows the flow between the three types of information packages:



The AIP is stored in a METS XML file and can be viewed in the permanent repository module of Rosetta. Each AIP relates to one IE.

Note: The METS includes only the representations that are not derivative copies of the IE. These representations are only available in the UI when reviewing the IE through the Web Editor.

For example, an IE that represents a scanned book has one Preservation Master representation that includes TIFF files and one derivative copy representation that includes low resolution JPEG files.

Whether the derivative copy representation was part of the IE when it was ingested by a submission application or whether it was generated by Rosetta as part of the SIP processing, the files of the representation remain in the Staging area. There is no mention of the derivative copy in the METS file since it is not part of the AIP.

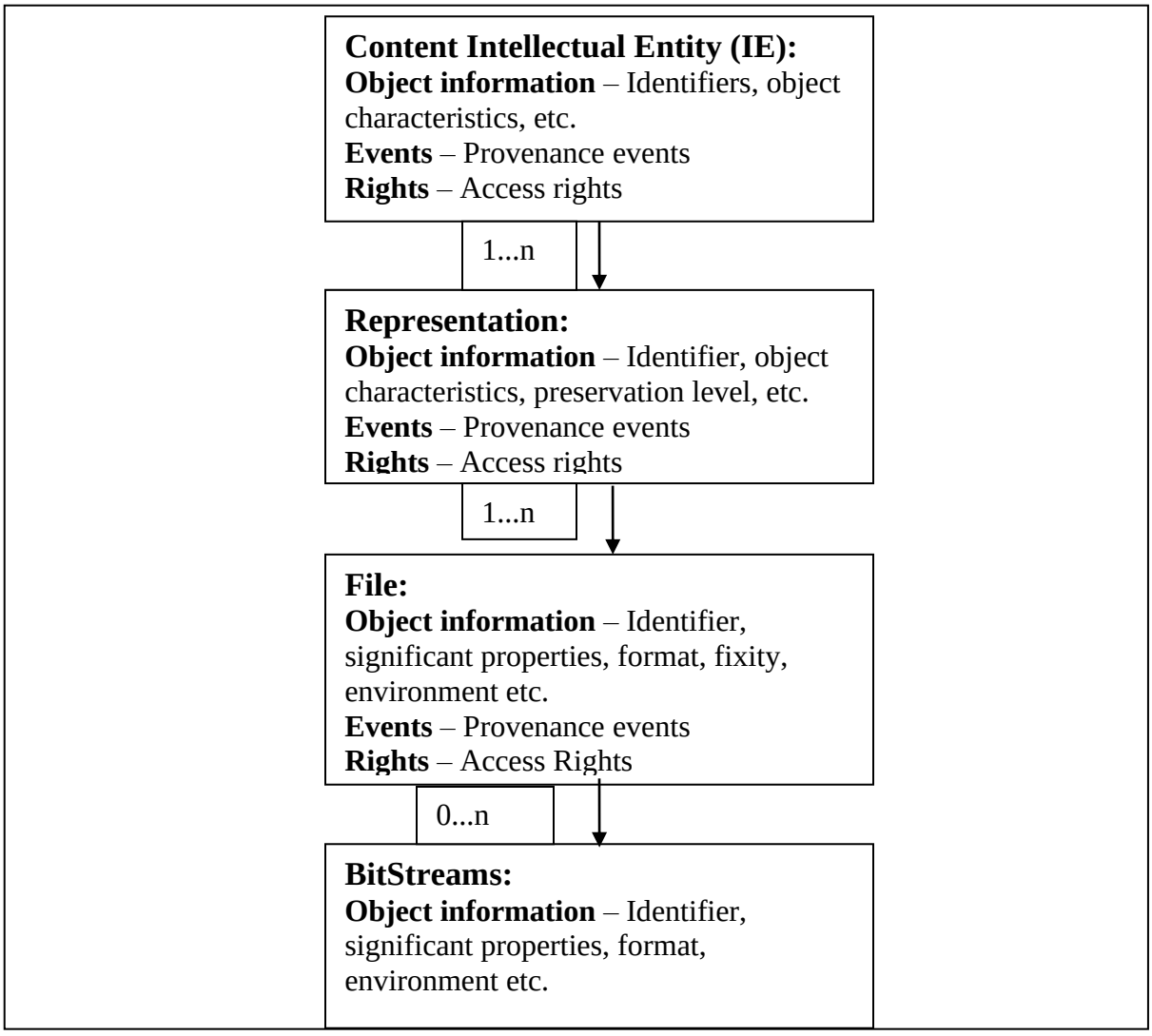
If the IE is exported from the UI, the derivative copy is exported as well, as part of the IE.

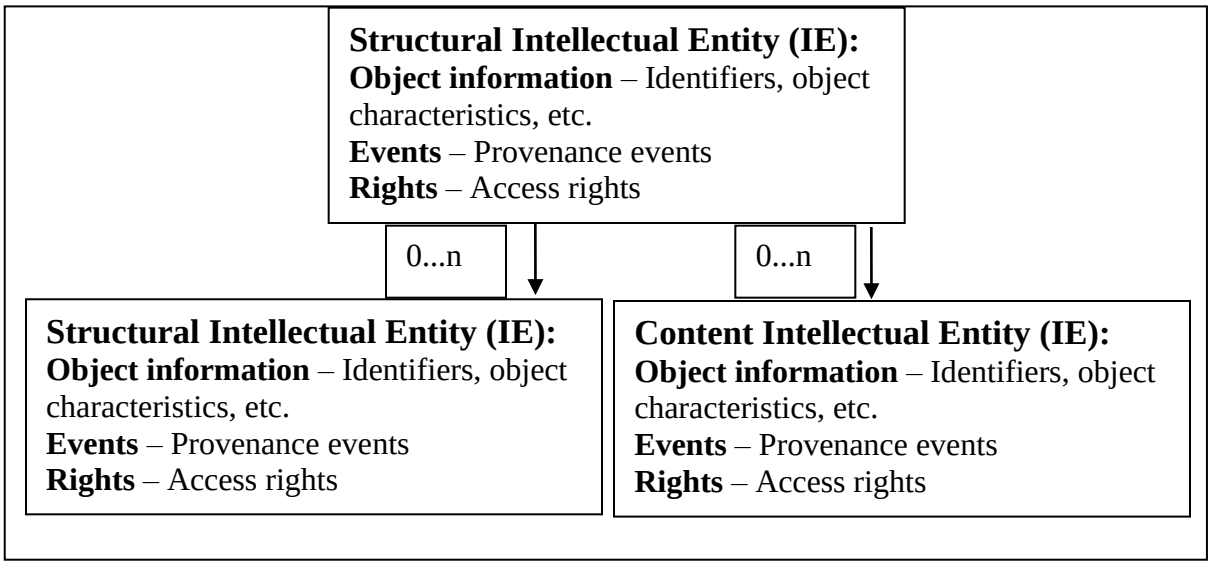
2

Data Model Structure

This chapter describes the hierarchical structure of the AIP data model that is based on the PREMIS reference model. It covers four levels of objects that are the basis for the Rosetta data model. Further information related to the PREMIS reference model can be found at: http://www.loc.gov/standards/premis/

The following two diagrams illustrate the entities of the AIP data model:


Content Intellectual Entity (IE):
Object information – Identifiers, object characteristics, etc.
Events – Provenance events
Rights – Access rights
1...n
Representation:
Object information – Identifier, object characteristics, preservation level, etc.
Events – Provenance events
Rights – Access rights
1...n
File:
Object information – Identifier, significant properties, format, fixity, environment etc.
Events – Provenance events
Rights – Access Rights
0...n
BitStreams:
Object information – Identifier, significant properties, format, environment etc.


Structural Intellectual Entity (IE):
Object information – Identifiers, object characteristics, etc.
Events – Provenance events
Rights – Access rights
0...n
0...n
Structural Intellectual Entity (IE):
Object information – Identifiers, object characteristics, etc.
Events – Provenance events
Rights – Access rights
Content Intellectual Entity (IE):
Object information – Identifiers, object characteristics, etc.
Events – Provenance events
Rights – Access rights

The entities in the data model are defined as follows.

Intellectual Entity

An intellectual entity is a distinct intellectual or artistic creation that is considered relevant to a designated community in the context of digital preservation. An intellectual entity can include other intellectual entities or alternatively, it may have one or more representations with files. There are two types of intellectual entities – structural and content:

- Structural IEs –Structural IEs represent a complex object (such as datasets with various content items or whole journals with multiple issues to preserve). They hold the metadata and structure of the complex object with relations to either other structural IEs, creating a nested hierarchy of objects for preservation, or content IEs (for example, the dataset metadata with the structure of its various content items or the journals and multiple issues metadata and structure). The actual digital content (for example, the various items under a dataset with their own metadata, the articles under the journal, and issues that are preserved) is stored and represented by content IEs. Therefore, structural IEs can contain relationships to other IEs, either structural or content, but not representations with files. The relationships are kept in a tree structure with each parent IE keeping the relationships to its immediate child IEs. A child IE does not hold a relationship to its parent IE and circular relationships are not allowed.
- Content IEs – Content IEs hold the actual digital content (for example, the various items under a dataset with their own metadata, the articles under the journal, and issues that are preserved, books, maps, photographs, databases, etc.). Content IEs can stand by themselves or be related to a structural IE. Content IEs hold the IE's metadata and representations with files but cannot hold relationships to other IEs.

Representations

A representation is the set of files, including structural metadata, needed for a complete and reasonable rendition of an intellectual entity. There can be more than one representation for the same intellectual entity. For example, a journal article may be complete in one PDF file and this single file will then constitute the representation. However, another journal article may consist of one SGML file and two image files. In this case, these three files will constitute the representation. A third article may be represented by one TIFF image for each of 12 pages plus an XML file of structural metadata showing the order of the pages. In this case, 13 files will constitute the representation. (PREMIS data dictionary, p. 14)

Files

A file is a named and ordered sequence of bytes that is known by an operating system. A file can be zero or more bytes and has a file format, access permissions, and file system characteristics such as size and last modification date.

Bitstreams

A bitstream is contiguous or non-contiguous data within a file that has meaningful common properties for preservation purposes. A bitstream cannot be transformed into a standalone file without the addition of file structure (headers, and so forth) and/or reformatting to comply with a particular file format.

A bitstream is defined in the PREMIS data model as a set of bits embedded within a file. This differs from common usage, where a bitstream could, in theory, span more than one file.

A good example of a file with embedded bitstreams is a TIFF file containing two images.

According to the TIFF file format specification, a TIFF file must contain a header that includes information about the file. It may then contain one or more images. In the data model, each of these images is a bitstream and can have properties such as identifiers, location, inhibitors, and detailed technical metadata (for example, color space).

Some bitstreams have the same properties as files and some do not. The image embedded within the TIFF file clearly has properties that are different from the file itself. However, three TIFF files can also be aggregated within a larger TAR file. In this case, the three TIFF files are filestreams, but they have all the properties of TIFF files. [1]

Rosetta bitstream functionality is limited to filestream only. Real bitstreams (embedded objects within a file) are functionally not supported. However, from a Data Model perspective, the Data Model serves both types of bitstreams.

[1] http://www.loc.gov/standards/premis/v2/premis-2-1.pdf

3

METS – Metadata Encoding and Transmission Standard

Rosetta uses METS as a container for the IE as an AIP. Further information related to the METS reference model can be found at: http://www.loc.gov/standards/mets/

Note: For the use of METS as a container for collections, see below.

The METS schema contains three types of metadata: Descriptive, Administrative, and Structural Map.

The following table illustrates which metadata type is applicable to each object type:

	Descriptive	Administrative	Structural Map
IE	√	√	
Representation		√	√
File	√	√	
Bitstream		√	

Descriptive

Information relating to the intellectual contents of the object, akin to much of the content of a standard catalogue record. This enables the user of a digital library to find the object and assess its relevance. Rosetta supports Dublin Core (DC) as the standard for descriptive metadata.

The descriptive metadata can be viewed any time the IE is edited or accessed. It can be published or exported to external systems and viewed by external users – as a whole or in parts only, based on the configuration. (Configuration is performed in the Rosetta Administration module.)

The descriptive metadata can be edited by a staff user who uses the Rosetta Web Editor. The edited metadata is written in a new version of the METS file.

Administrative

Information is necessary for the manager of the electronic collection to administer the object, including information on intellectual property rights and technical information on the object and the files or other objects (structural IE relationships) that comprise it.

The administrative metadata is mostly generated by Rosetta throughout the SIP processing, and some of it can be edited by the staff user. (See below in the Rosetta DNX Profile section for which metadata can be edited.)

Structural Map

Information on how the individual components that make up the object relate to each other, including the order in which they should be presented to the user – for example, how should still image files that comprise a digitized version of a print volume be ordered.

The structural map can be edited in the Web Editor by the staff user, or outside of Rosetta, if the IE is loaded through the submission application.

Note that once the IE is in the permanent repository, it can be edited only in the Web Editor.

METS Sections

A METS file consists of seven major sections, each describing a different facet of the digital object:

Header (metsHdr)

The header section is not in use in Rosetta and is not included in the Rosetta data model.

Descriptive Metadata (dmdSec)

The attributes of the descriptive metadata are:

- **IE and File level** – the descriptive metadata is stored only on the IE and File level
- **Dublin Core schema** – Rosetta uses Dublin Core (DC) and Qualified DC standards, or other schemas that are not hierarchical (such as MODS).

- **METS section <dmdSec>** – Descriptive metadata is held within sections of the METS file named <dmdSec>. Although METS allows this metadata either to be held in external files that are referenced from within the METS file, or to be embedded directly with it, Rosetta requires it to be embedded directly with it for preservation reasons.

The descriptive metadata in Rosetta is embedded directly in the METS file, using an <mdWrap> element to contain it, as illustrated below:

```
<mets:dmdSec ID="ie-dmd">
    <mets:mdWrap MDTYPE="DC">
        <mets:xmlData>
            <dc:record xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dcterms="http://purl.org/dc/terms/"
            xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
            xmlns:mods="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-0.xsd">
                <dc:title>My Photos</dc:title>
                <dc:description>my first photo album</dc:description>
                <dcterms:created>10/07/2013</dcterms:created>
            </dc:record>
        </mets:xmlData>
    </mets:mdWrap>
</mets:dmdSec>
<mets:dmdSec ID="FL1415-dmd">
    <mets:mdWrap MDTYPE="DC">
        <mets:xmlData>
            <dc:record xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dcterms="http://purl.org/dc/terms/"
            xmlns:mods="http://www.loc.gov/mods/v3 http://www.loc.gov/standards/mods/v3/mods-3-0.xsd"
            xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
                <dc:title>Flowers</dc:title>
                <dc:description>photo of my marigolds</dc:description>
                <dcterms:created>11/04/2013</dcterms:created>
            </dc:record>
        </mets:xmlData>
    </mets:mdWrap>
</mets:dmdSec>
```

Administrative Metadata

Administrative metadata includes all the information that is not descriptive for each object that is part of the intellectual entity. It includes technical attributes of the stream files (image resolution, file size, and so forth), access rights for delivery, important events that are relevant for preservation (provenance events), the relationships structure of a structural IE and metadata that arrived with the IE and should be kept in its original structure (not normalized, such as MIX technical metadata for image files).

The administrative elements, each identified by an ID, are used to record this metadata, which may be held in external files or embedded within the METS file using the <mdWrap> element. Rosetta uses the DNX format for holding all the technical information, events, and access rights.

Rosetta uses the following sections of METS within the <amdSec>:

- technical (mets:techMD) – Holds the technical information of the object, within the <mdWrap>, in DNX format.
- rights (mets:rightsMD) – Holds the access rights within the <mdWrap>, in DNX format.
- events (mets:digiprovMD) – Holds the provenance events, within the <mdWrap>, in DNX format.

- source (mets:sourceMD) – Within the `<mdWrap>`, there could be any type of source metadata as specified in the Rosetta METS Profile and defined in the `Other Source Metadata Subtype` code table. Currently, this section holds the details of the DB record of the `<amdSec>`, such as the record ID, date of creation, and modification date.

For more information regarding the DNX format, see the Rosetta DNX Profile section below.

Administrative Section for IE, Representation, File, and Bitstream

Each object level (IE, representation, file, and bitstream) has its own `<amdSec>` that includes the four sub-sections described above (technical, rights, events, and source), even if some of these sections are empty.

As illustrated in the following table, the content of the sections differs between each object level, according to the PREMIS data dictionary:

	techMd	rightsMd	digiProvMd	sourceMD
IE	Identifiers, Control information, Retention policy, Structural IE relationships	Access rights	Provenance events	Source descriptive – MODS, MARC
Representation	Preservation type, usage type, revision no.	Access rights	Provenance events – Add Representation	
File	Significant properties, Validation stack outcome	Access rights	Provenance events – Validation Stack	Source technical metadata – MIX, NISO; descriptive – MODS, MARC
Bitstream	Significant properties, Validation stack outcome			

The sections that are empty will look like the following:

```
- <mets:rightsMD ID="REP1122-amd-rights">
  - <mets:mdWrap MDTYPE="OTHER" OTHERMDTYPE="dnx">
    - <mets:xmlData>
        <dnx xmlns="http://www.exlibrisgroup.com/dps/dnx" />
      </mets:xmlData>
    </mets:mdWrap>
  </mets:rightsMD>
```

Implementation of PREMIS within METS

As mentioned above, Rosetta implements the PREMIS data model within METS.

In each object level of each METS section, there are DNX sections and fields that match the PREMIS semantic units.

The PREMIS entities are located in the METS sections in the following way:

- Objects – As explained above, each object has its own `<amdSec>` section in which its administrative metadata is specified.
- Events – In Rosetta, the events are related to objects. Each object has its relevant events specified in the `<amd-digiProv>` section, within its `<amdSec>`.
- Agents – In Rosetta, the agents' elements are represented as attributes of the entity of which they are agents. For example, each event has its agents that are linked to it – user, software, or hardware.
- Rights – The Rosetta AIP stored in the METS contains the access rights of the IE, representation, and file. The access rights are stored in the `<amd-rights>` section within the `<amdSec>` section on the IE/REP level and these rights are relevant for the delivery of any part of the IE. This is currently an Ex Libris proprietary format. For more information, see the Access Rights Within DNX section below.

For more information about PREMIS implementation in Rosetta DNX format, see below in *Rosetta DNX Profile* chapter.

Note: The classification type of an IE, content or structural, is not kept in the METS, but in the database tables. When submitting an IE with the submission application, it is inferred from the existence or nonexistence of representations and files.

File Groups

The `<mets:fileSec>` section and `<mets:fileGrp>` sections are applicable only for content IEs. The `<mets:fileSec>` section includes the `<mets:fileGrp>` sections in which each section holds the content of a representation (the list of files that are grouped in the representation).

This METS section holds the information about all the files, and some information about the representation.

- Representation information:

- **USE** – The usage of this representation. In Rosetta, it will be View even though METS allows more values, such as Thumbnail or ALTO.
- **ID** – The unique ID of the representation.
- **ADMID** – The ID of the administrative section that describes the representation.

- File information:
  - **File ID** – The unique ID of the file.
  - **ADMID** – The ID of the administrative section that describes the file.
  - **GROUPID** – Attribute to determine relationship between files within separate Representations.
  - **MIMETYPE** Currently not in use. MIME type is inferred from the file itself and kept in the fileFormat DNX section.
  - **<mets:FLocat>** - The file location element, `<FLocat>`, provides a pointer to the location of a content file. It uses the XLink reference syntax to provide linking information indicating the actual location of the content file, along with other attributes specifying additional linking information. Only local references are currently supported.

Note: `<FLocat>` is an empty element. The location of the resource pointed to must be stored in the `xlin:href` attribute.

The following is an example of a `<mets:fileGrp>` section within the `<mets:fileSec>` section:

```
<mets:fileSec>
    <mets:fileGrp ID="REP3970" ADMID="REP3970-amd">
        <mets:file ID="FL3974" GROUPID="1" ADMID="FL3974-amd">
            <mets:FLocat xmlns:xlin="http://www.w3.org/1999/xlink" LOCTYPE="URL" xlin:href=
            "/exlibris/dps/d4_1/profile/permanent//file/storage1/2014/07/03/file_5/V1-FL3974.pdf"/>
        </mets:file>
    </mets:fileGrp>
</mets:fileSec>
```

Structural Map

The Structural Map section `<mets:structMap>` is applicable only for content IEs. This part of a METS file is a description of the structure of each representation and contains information on how the files relate to each other hierarchically. For example, if the IE is a digitized book, one of its representations will be 150 scanned images of the book's pages. To show how the pages are structured in chapters, this section will show the divisions by chapters, and for each division there is a label that describes the chapter's name.

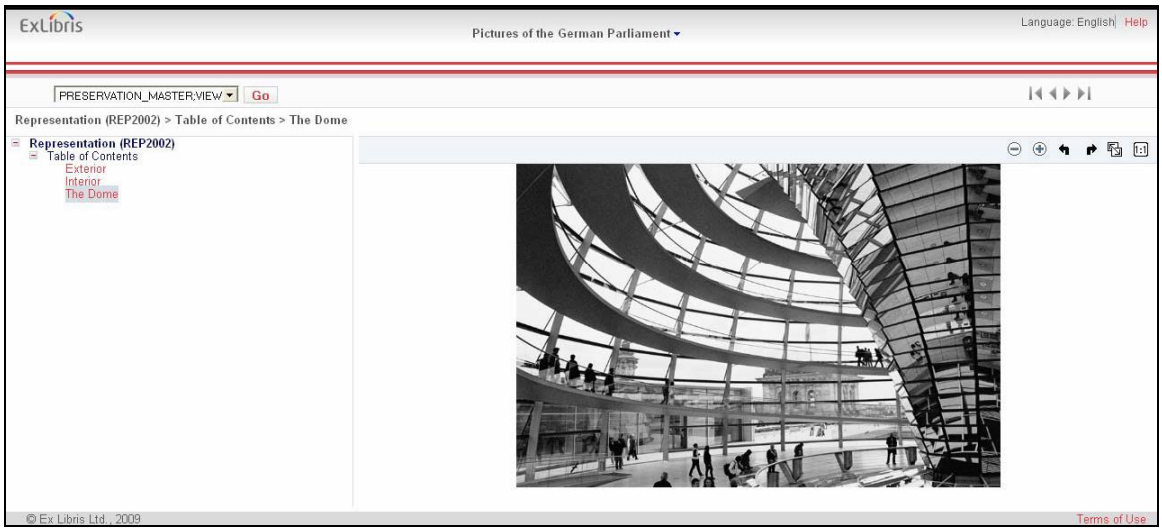
There can be multiple structural maps for each representation if the same files are structured differently (that is, one for chapters and one for content). The example below shows two structural maps for the same representation of three files: One is physical, and the other is logical (where the files are divided in chapters):

```
- <mets:structMap ID="rep1-1" TYPE="PHYSICAL">
  - <mets:div LABEL="PRESERVATION_MASTER;VIEW">
    - <mets:div LABEL="Table of Contents">
      - <mets:div LABEL="Serenity now" TYPE="FILE">
          <mets:fptr FILEID="fid1-1" />
        </mets:div>
      - <mets:div LABEL="Blue hills" TYPE="FILE">
          <mets:fptr FILEID="fid2-1" />
        </mets:div>
      - <mets:div LABEL="Sunset" TYPE="FILE">
          <mets:fptr FILEID="fid3-1" />
        </mets:div>
      </mets:div>
    </mets:div>
  </mets:structMap>
- <mets:structMap ID="rep1-2" TYPE="LOGICAL">
  - <mets:div LABEL="PRESERVATION_MASTER;VIEW">
    - <mets:div LABEL="Table of Contents">
      - <mets:div LABEL="Chapter 1">
        - <mets:div LABEL="Serenity now" TYPE="FILE">
            <mets:fptr FILEID="fid1-1" />
          </mets:div>
        - <mets:div LABEL="Blue hills" TYPE="FILE">
            <mets:fptr FILEID="fid2-1" />
          </mets:div>
        </mets:div>
      - <mets:div LABEL="Chapter 2">
        - <mets:div LABEL="Sunset" TYPE="FILE">
            <mets:fptr FILEID="fid3-1" />
          </mets:div>
        </mets:div>
      </mets:div>
    </mets:div>
  </mets:structMap>
```

This structural map shows that the representation is referred in the structMap ID (rep1), and each structMap of the same representation has its own identifier (rep1-1, rep1-2). Within these sections are pointers to the files that hold the images of the pages; these are referenced by the FILEID attribute within the `<fptr>` (file pointer) element.

Each file pointer is wrapped inside a section `<mets:div>` that contains a label. This label is created when the METS is generated (by Rosetta or by the Rosetta SDK), and it holds the file label that was entered by the Producer Agent when depositing the file. If there is no file label, this label holds the file name.

The structural map provides a logical layout for the structure of the whole object, and one that is easy to navigate using any XML-compatible software. In the Delivery module of Rosetta, the viewer displays the files' labels according to their order:



Structural Links

This section of the METS is not in use in the Rosetta AIP data model.

Behavior

This section of the METS is not in use in the Rosetta AIP data model.

Events

The events metadata holds the information about actions that affect the object. Each object level has different types of actions that should be captured. In Rosetta, the events that are recorded in the AIP are provenance events, while many other events are captured in the system but do not become part of the AIP metadata.

The following types of events are considered provenance events:

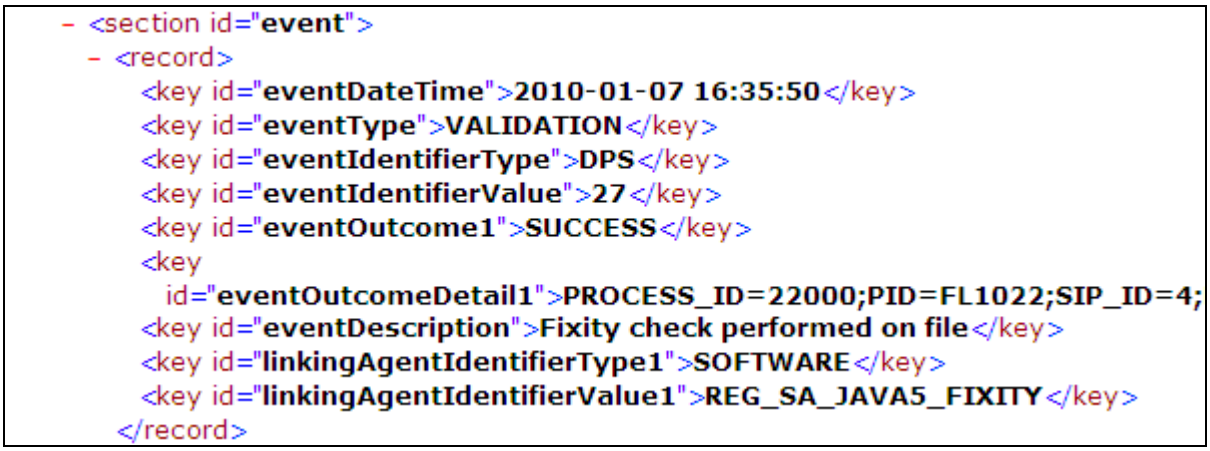
- New version of the IE – a result of adding a new representation or metadata (descriptive, access rights)
- Validation checks – validity and integrity checks on files

Each such event will be written in the events (`mets:digiprovMD`) section belonging to the relevant object level (IE, representation, or file).

Each event will be written in the DNX format and will include the following:

- Agent – The agent that triggered this event. An agent is not necessarily a person. An agent may also refer to a process, plug-in tool, and so forth.
- Event details – Such as the creation date, a description, the parameters, and so forth

The following is an example of an event that is stored in the `digiProvMD` section of a file. This section holds the events in DNX format:


- <section id="event">
- <record>
<key id="eventDateTime">2010-01-07 16:35:50</key>
<key id="eventType">VALIDATION</key>
<key id="eventIdentifierType">DPS</key>
<key id="eventIdentifierValue">27</key>
<key id="eventOutcome1">SUCCESS</key>
<key
id="eventOutcomeDetail1">PROCESS_ID=22000;PID=FL1022;SIP_ID=4;
<key id="eventDescription">Fixity check performed on file</key>
<key id="linkingAgentIdentifierType1">SOFTWARE</key>
<key id="linkingAgentIdentifierValue1">REG_SA_JAVA5_FIXITY</key>
</record>

In addition to events, the `digiprovMD` section on the IE level stores the details of the Producer and the Producer Agent who deposited the IE.

```
- <mets:digiprovMD ID="ie-amd-digiprov">
  - <mets:mdWrap MDTYPE="OTHER" OTHERMDTYPE="dnx">
    - <mets:xmlData>
      - <dnx xmlns="http://www.exlibrisgroup.com/dps/dnx">
        - <section id="producer">
          - <record>
              <key id="userName" />
              <key id="address1">6740</key>
              <key id="address2">Willow Lane</key>
              <key id="address3">Dallas</key>
              <key id="address4">Texas</key>
              <key id="address5">U.S.A</key>
              <key id="defaultLanguage">en</key>
              <key id="emailAddress">marek.melichar@nkp.com</key>
              <key id="firstName">University of Oklahoma</key>
              <key id="jobTitle" />
              <key id="lastName">Legal Department</key>
              <key id="middleName" />
              <key id="telephone1">972-456-6547</key>
              <key id="telephone2" />
              <key id="authorativeName">Library of Legal Department</key>
              <key id="producerId">34366</key>
              <key id="userIdAppId">34362</key>
              <key id="webSiteUrl" />
              <key id="zip" />
            </record>
          </section>
        - <section id="producerAgent">
          - <record>
              <key id="firstName">John</key>
              <key id="lastName">Smith</key>
              <key id="middleName" />
            </record>
          </section>
        </dnx>
      </mets:xmlData>
    </mets:mdWrap>
  </mets:digiprovMD>
```

Access Rights

The Rosetta METS XML can hold two types of access rights information: PREMIS and non-PREMIS:

- PREMIS rights – Information regarding an external system that manages the IE's rights:
  - `linkingRightsStatementIdentifierType` – A designation of the domain within which the `linkingRightsStatementIdentifier` is unique
  - `linkingRightsStatementIdentifierValue` – The value of the `linkingRightsStatementIdentifier`
- Non PREMIS rights – Information regarding the access rights policy managed by Rosetta:
- PolicyID – The unique ID of the different access rights managed by Rosetta. For example: AR_EMBARGOED_FOR_5_YEARS, AR_5_CONCURRENT_USERS
- Policy parameters – If the policy requires any parameters
- Policy description – Description of the policyID. For example: AR_EMBARGOED_FOR_5_YEARS – Embargoed for 5 years, AR_5_CONCURRENT_USERS – Limited access according to copyright law

METS XML Sections

Declaration	<?xml ... ?>	
Attributes	`version`	`"1.0"`
	`encoding`	`"utf-8"`
Obligation	Mandatory	
Repeatable	No	

Element	<mets>	
Attributes	xmlns	"http://www.loc.gov/METS/"
Content	<dmdSec ID="ie-dmd"> <amdSec ID="ie-amd"> <amdSec ID="REP1…n"> <amdSec ID="FL1…n"> <fileSec> <structMap ID=REP1…n TYPE="[PHYSICAL|LOGICAL]">	
Obligation	Mandatory	
Repeatable	No	

Element	<dmdSec>	
Description	Descriptive metadata (author, title, and so forth) describing the intellectual entity	
Attributes	ID	"ie-dmd"
Content	<mdWrap MDTYPE="DC">	
Obligation	Mandatory	
Repeatable	No	

Element	<mdWrap>	
Description	Container element for DC (Dublin Core) descriptive metadata	
Attributes	MDTYPE	"DC"
Content	<xmlData>	
Obligation	Mandatory	
Repeatable	No	

Element	<xmlData>
Content	< dc:record>
Obligation	Mandatory
Repeatable	No

Element	<dc:record>	
Attributes	`xmlns:dc`	`"http://purl.org/dc/elements/1.1/"`
Content	All valid DC elements	
Obligation	Mandatory	
Repeatable	No	

Element	<amdSec>	
Description	Information is necessary for the manager of the electronic collection to administer the object, including information on intellectual property rights and technical information on the object and the files that comprise it. The administrative metadata is mostly generated by Rosetta throughout the SIP processing, and some of it can be edited by the staff user.	
Attributes	`ID`	`"REP1…n-amd"`
Content	`<techMD>`	
Obligation	Mandatory	
Repeatable	No	

Element	<techMD>	
Description	Holds the technical information of the object, within the <mdWrap>, in DNX format	
Attributes	ID	"REP1…n-amd-tech"
Content	<mdWrap>	
Obligation	Mandatory	
Repeatable	No	

Element	<mdWrap>	
Description	Container element for the representation DNX	
Attributes	MDTYPE	"OTHER"
	OTHERMDTYPE	"dnx"
Content	<xmlData>	
Obligation	Mandatory	
Repeatable	No	

Element	<xmlData>
Content	<dnx>
Obligation	Mandatory
Repeatable	No

Element	<dnx>	
Attributes	xmlns	"http://www.exlibrisgroup.com/dps/dnx"
Content	DNX Sections (See Appendix B – DNX Profile)	
Obligation	Mandatory	
Repeatable	No	

Element	<rightsMD>	
Description	Holds the rights information of the object, within the <mdWrap>, in DNX format	
Attributes	ID	"REP1…n-amd-rights"
Content	<mdWrap>	
Obligation	Mandatory	
Repeatable	No	

Element	<mdWrap>	
Description	Container element for the representation DNX	
Attributes	MDTYPE	"OTHER"
	OTHERMDTYPE	"dnx"
Content	<xmlData>	
Obligation	Mandatory	
Repeatable	No	

Element	<xmlData>
Content	<dnx>
Obligation	Mandatory
Repeatable	No

Element	<dnx>	
Attributes	xmlns	"http://www.exlibrisgroup.com/dps/dnx"
Obligation	Mandatory	
Repeatable	No	
Note	If there are no access rights on the representation level, the DNX is empty.	

Element	<sourceMD>	
Description	Holds the source metadata of the object, within the <mdWrap>, in DNX format	
Attributes	ID	"REP1…n-amd-source"
Content	<mdWrap>	
Obligation	Mandatory if applicable (source MD has been provided with the object)	
Repeatable	Yes	

Element	<mdWrap>	
Description	Container element for the representation DNX	
Attributes	MDTYPE	"OTHER"
	OTHERMDTYPE	"dnx"
Content	<xmlData>	
Obligation	Mandatory	
Repeatable	No	

Element	<xmlData>
Content	<dnx>
Obligation	Mandatory
Repeatable	No

Element	<dnx>	
Attributes	xmlns	"http://www.exlibrisgroup.com/dps/dnx"
Content	DNX Sections (See Appendix B - DNX Profile)	
Obligation	Mandatory	
Repeatable	No	

Element	<digiprovMD>	
Description	Holds the provenance events information of the object, within the <mdWrap>, in DNX format	
Attributes	ID	"REP1…n-amd-digiprov"
Content	<mdWrap>	
Obligation	Mandatory	
Repeatable	No	

Element	<mdWrap>	
Description	Container element for the representation DNX	
Attributes	MDTYPE	"OTHER"
	OTHERMDTYPE	"dnx"
Content	<xmlData>	
Obligation	Mandatory	
Repeatable	No	

Element	<xmlData>
Content	<dnx>
Obligation	Mandatory
Repeatable	No

Element	<dnx>	
Attributes	xmlns	"http://www.exlibrisgroup.com/dps/dnx"
Content	DNX Sections (See Appendix B - DNX Profile)	
Obligation	Mandatory	
Repeatable	No	

Element	<dmdSec>	
Description	Descriptive metadata (author, title, and so forth) describing the file	
Attributes	ID	"FL1…n-dmd"
Content	<mdWrap MDTYPE="DC">	
Obligation	Optional	
Repeatable	No	

Element	<amdSec>	
Description	Information is necessary for the manager of the electronic collection to administer the object, including information on intellectual property rights and technical information on the object and the files that comprise it. The administrative metadata is mostly generated by Rosetta throughout the SIP processing, and some of it can be edited by the staff user.	
Attributes	ID	"FL1…n-amd"
Content	<techMD>	
Obligation	Mandatory	
Repeatable	No	

Element	<techMD>	
Description	Holds the technical information of the object, within the <mdWrap>, in DNX format	
Attributes	ID	"FL1…n-amd-tech"
Content	<mdWrap>	
Obligation	Mandatory	
Repeatable	No	

Element	<mdWrap>	
Description	Container element for the representation DNX	
Attributes	MDTYPE	"OTHER"
	OTHERMDTYPE	"dnx"
Content	<xmlData>	
Obligation	Mandatory	
Repeatable	No	

Element	<xmlData>
Content	<dnx>
Obligation	Mandatory
Repeatable	No

Element	<dnx>	
Attributes	xmlns	"http://www.exlibrisgroup.com/dps/dnx"
Content	DNX Sections (See Appendix B - DNX Profile)	
Obligation	Mandatory	
Repeatable	No	

Element	<rightsMD>	
Description	Holds the rights information of the object, within the <mdWrap>, in DNX format	
Attributes	ID	"FL1…n-amd-rights"
Content	<mdWrap>	
Obligation	Mandatory	
Repeatable	No	

Element	<mdWrap>	
Description	Container element for the representation DNX	
Attributes	MDTYPE	"OTHER"
	OTHERMDTYPE	"dnx"
Content	<xmlData>	
Obligation	Mandatory	
Repeatable	No	

Element	<xmlData>
Content	<dnx>
Obligation	Mandatory
Repeatable	No

Element	<dnx>	
Attributes	xmlns	"http://www.exlibrisgroup.com/dps/dnx"
Obligation	Mandatory	
Repeatable	No	
Note	If there are no access rights on the file level, the DNX is empty	

Element	<sourceMD>	
Description	Holds the source metadata of the object, within the <mdWrap>, in DNX format	
Attributes	ID	"FL1…n-amd-source"
Content	<mdWrap>	
Obligation	Mandatory if applicable (source MD has been provided with the object)	
Repeatable	Yes	

Element	<mdWrap>	
Description	Container element for the representation DNX	
Attributes	MDTYPE	"OTHER"
	OTHERMDTYPE	"dnx"
Content	<xmlData>	
Obligation	Mandatory	
Repeatable	No	

Element	<xmlData>
Content	<dnx>
Obligation	Mandatory
Repeatable	No

Element	<dnx>	
Attributes	xmlns	"http://www.exlibrisgroup.com/dps/dnx"
Content	DNX Sections (See Appendix B - DNX Profile)	
Obligation	Mandatory	
Repeatable	No	

Element	<digiprovMD>	
Description	Holds the provenance events information of the object, within the <mdWrap>, in DNX format	
Attributes	ID	"FL1…n-amd-digiprov"
Content	<mdWrap>	
Obligation	Mandatory	
Repeatable	No	

Element	<mdWrap>	
Description	Container element for the representation DNX	
Attributes	MDTYPE	"OTHER"
	OTHERMDTYPE	"dnx"
Content	<xmlData>	
Obligation	Mandatory	
Repeatable	No	

Element	<xmlData>
Content	<dnx>
Obligation	Mandatory
Repeatable	No

Element	<dnx>	
Attributes	xmlns	"http://www.exlibrisgroup.com/dps/dnx"
Content	DNX Sections (See Appendix B - DNX Profile)	
Obligation	Mandatory	
Repeatable	No	

Element	<amdSec>	
Description	Information is necessary for the manager of the electronic collection to administer the object, including information on intellectual property rights and technical information on the object and the files and other objects that comprise it. The administrative metadata is mostly generated by Rosetta throughout the SIP processing, and some of it can be edited by the staff user.	
Attributes	ID	"ie-amd"
Content	<techMD>	
Obligation	Mandatory	
Repeatable	No	

Element	<techMD>	
Description	Holds the technical information of the object, within the <mdWrap>, in DNX format	
Attributes	ID	"ie-amd-tech"
Content	<mdWrap>	
Obligation	Mandatory	
Repeatable	No	

Element	<mdWrap>	
Description	Container element for the representation DNX	
Attributes	MDTYPE	"OTHER"
	OTHERMDTYPE	"dnx"
Content	<xmlData>	
Obligation	Mandatory	
Repeatable	No	

Element	<xmlData>
Content	<dnx>
Obligation	Mandatory
Repeatable	No

Element	<dnx>	
Attributes	xmlns	"http://www.exlibrisgroup.com/dps/dnx"
Content	DNX Sections (See Appendix B - DNX Profile)	
Obligation	Mandatory	
Repeatable	No	

Element	<rightsMD>	
Description	Holds the rights information of the object, within the <mdWrap>, in DNX format	
Attributes	ID	"ie-amd-rights"
Content	<mdWrap>	
Obligation	Mandatory	
Repeatable	No	

Element	<mdWrap>	
Description	Container element for the representation DNX	
Attributes	MDTYPE	"OTHER"
	OTHERMDTYPE	"dnx"
Content	<xmlData>	
Obligation	Mandatory	
Repeatable	No	

Element	<xmlData>
Content	<dnx>
Obligation	Mandatory
Repeatable	No

Element	<dnx>	
Attributes	xmlns	"http://www.exlibrisgroup.com/dps/dnx"
Content	DNX Sections (See Appendix B - DNX Profile)	
Obligation	Mandatory	
Repeatable	No	

Element	<sourceMD>	
Description	Holds the source metadata of the object, within the <mdWrap>, in DNX format	
Attributes	ID	"ie-amd-source[-source type][-1…n]"
Content	<mdWrap>	
Obligation	Mandatory if applicable (source MD has been provided with the object)	
Repeatable	Yes	
Note	ie-amd-source is reserved for DNX metadata. Other supported sourceMD types (specified in the Rosetta METS Profile) must be identified using the source type, for example, ie-amd-source-mods.	

Element	<mdWrap>	
Description	Container element for the representation DNX	
Attributes	MDTYPE	"OTHER"
	OTHERMDTYPE	"dnx"

Content	<xmlData>
Obligation	Mandatory
Repeatable	No

Element	<xmlData>
Content	<dnx>
Obligation	Mandatory
Repeatable	No

Element	<dnx>	
Attributes	xmlns	"http://www.exlibrisgroup.com/dps/dnx"
Content	DNX Sections (See Appendix B - DNX Profile)	
Obligation	Mandatory	
Repeatable	No	

Element	<digiprovMD>	
Description	Holds the provenance events information of the object, within the <mdWrap>, in DNX format	
Attributes	ID	"ie-amd-digiprov"
Content	<mdWrap>	
Obligation	Mandatory	
Repeatable	No	

Element	<mdWrap>	
Description	Container element for the representation DNX	
Attributes	MDTYPE	"OTHER"
	OTHERMDTYPE	"dnx"
Content	<xmlData>	
Obligation	Mandatory	
Repeatable	No	

Element	<xmlData>
Content	<dnx>
Obligation	Mandatory
Repeatable	No

Element	<dnx>	
Attributes	xmlns	"http://www.exlibrisgroup.com/dps/dnx"
Content	DNX Sections (See Appendix B - DNX Profile)	
Obligation	Mandatory	
Repeatable	No	

Element	<fileSec>
Description	Inventory of all of the files that make up the digital representation of work.
Content	<fileGrp>
Obligation	Mandatory
Repeatable	No

Element	<fileGrp>	
Description	Container for all of the files of a given representation.	
Content	<file>	
	USE	The usage of this representation. Currently not in use. Rosetta regards all file groups as used for viewing.
	ID	"REP1...n"
	ADMID	"REP1...n-amd"
Obligation	Mandatory	
Repeatable	Yes	
Note	If there is more than one <fileGrp> element, no specific order is required.	

Element	<file>	
Description	Properties of a file that provides a representation of a single page of work	
Attributes	ID	"FL1...n"
	MIMETYPE	Currently not in use. MIME type is inferred from the file itself and kept in the fileFormat DNX section:
	ADMID	"FL1...n-amd"
	GROUPID	String
Content	<FLocat>	
Obligation	Mandatory	
Repeatable	Yes	

Element	<FLocat>	
Attributes	LOCTYPE	"URL"
	xlin:href	"file://" + File Name
	xmlns:xlin	"http://www.w3.org/1999/xlink"

Content	
Obligation	
Repeatable	Yes

Element	<structMap>	
Description	Logical/physical organization of the files within the representation	
Attributes	ID	"REP1…n-N"
Attributes	TYPE	"PHYSICAL" or "LOGICAL"
Content	e.g.	
Obligation	Mandatory	
Repeatable	Yes	

Element		
Attributes	LABEL	PRESERVATION_MASTER;VIEW
Content		
Obligation	Mandatory	
Repeatable	No	

Element		
Description	A logical section of work	
Attributes	LABEL	Table Of Content
Content		
Obligation	Mandatory	
Repeatable	Yes	

Element		
Description	A single page of work	
Attributes	LABEL	Free text. Label may remain empty
	TYPE	"FILE"
Content	<fptr>	
Obligation	Mandatory	
Repeatable	Yes	

Element	<fptr>	
Description	Pointer to a certain file within the Structural Map	
Attributes	FILEID	PID of the file
Obligation	Mandatory	
Repeatable	No	

Collections in METS

Collections are managed in the operational repository in a dedicated table (called Collection). In the permanent repository each Collection record is contained in a METS file that is different than the IE METS.

The METS file includes the following sections:

mets:dmdSec ID="collection-dmd"	Descriptive metadata in DC
DC	The descriptive metadata in DC format
mets:amdSec ID="collection-amd"	**Administrative metadata section**
mets:techMD ID="collection-amd-tech"	**Technical metadata sub-section**
collection	DNX section that includes the collection's identifiers
objectCharacteristics	DNX section that includes the control information of the Collection
objectIdentifier	DNX section that can hold DOI

mets:sourceMD ID="collection-amd-source"	Source matadata of the Collection (repeatable section)
DC/MARC/MODS	Descriptive metadata in any format

An example for a Collection stored in METS is as follows:

```
<mets:mets xmlns:mets="http://www.loc.gov/METS/">
  <mets:dmdSec ID="collection-dmd">
    <mets:mdWrap MDTYPE="DC">
      <mets:xmlData>
        <dc:record xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:dcterms="http://purl.org/dc/terms/"
xmlns:mods="http://www.loc.gov/mods/v3
http://www.loc.gov/standards/mods/v3/mods-3-0.xsd"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
          <dc:title>Economist E-Journal </dc:title>
          <dc:creator>Jack Bauer</dc:creator>
          <dc:description>The electronic version of the
Economist</dc:description>
          <dc:type>E-Journal</dc:type>
          <dcterms:created>07/02/2011</dcterms:created>
          <dc:publisher/>
          <dc:description/>
        </dc:record>
      </mets:xmlData>
    </mets:mdWrap>
  </mets:dmdSec>
  <mets:amdSec ID="collection-amd">
    <mets:techMD ID="collection-amd-tech">
      <mets:mdWrap MDTYPE="OTHER" OTHERMDTYPE="dnx">
        <mets:xmlData>
          <dnx xmlns="http://www.exlibrisgroup.com/dps/dnx">
              <section id="collection">
            <record>
              <key id="externalSystem"></key>
              <key id="externalId"></key>
              <key id="collectionId">1820</key>
              <key id="Name">Economist E-Journal</key>

              </record>
            </section>
            <section id="objectCharacteristics">
              <record>
                <key id="objectType">COLLECTION</key>
                <key id="parentID"/>
                <key id="groupID"/>
                <key id="creationDate">2011-02-07 22:07:46</key>
                <key id="createdBy">admin1</key>
                <key id="modificationDate">2011-02-07 22:22:11</key>
                <key id="modifiedBy">admin1</key>
                <key id="owner">CRS00.INS00.DPR00</key>
              </record>
            </section>
           <section id="objectIdentifier">
              <record>
                <key id="objectIdentifierType">DOI</key>
```

```
                <key id=" objectIdentifierValue">DPS-123456-8</key>
              </record>
            </section>
          </dnx>
        </mets:xmlData>
      </mets:mdWrap>
    </mets:techMD>
    <mets:sourceMD ID="collection-amd-source">
      <mets:mdWrap MDTYPE="OTHER" OTHERMDTYPE="MODS">
        <mets:xmlData>
        </mets:xmlData>
      </mets:mdWrap>
    </mets:sourceMD>
  </mets:amdSec>
</mets:mets>
```

The METS file is written to the permanent repository every time it is changed in the UI. The UI shows the last version of the METS.

4

Rosetta DNX Profile

The DNX schema is a simple and unified XML schema that holds the administrative metadata of the IE in the permanent repository. It contains all the important data elements in a simple flat structure, divided between the different object levels (IE, representation, file and bitstreams), and includes the important technical metadata that is relevant for preservation.

The administrative metadata that needs to be stored arrives from various sources:

- Technical metadata that is being generated by the metadata extraction tools (JHOVE, NLNZ tools)
- Access rights associated with the material flow
- CMS information (system and record ID)
- Provenance information – Producer, Producer Agent information, events information
- Structural IE relationships – provided by the depositing or editing users
- Miscellaneous information – such as links to external events, or other intellectual entities

Since all this information comes from different sources with different standards, some of it is duplicated or organized in a way that is not useful. The DNX profile, therefore, is designed to hold all this information in a clear and organized way, with a clear mapping to the original source that enables converting it back and forth.

The DNX is written to the AIP (METS XML file) based on the metadata that is stored in different tables in the Rosetta staging database. Most of the DNX data is generated by Rosetta, while some of the data in the DNX section is populated by the submission application, before the IE is deposited.

The provenance information is written in the DNX when the data is moved to the permanent stage, since the information is still gathered during the SIP processing stage.

The purpose of this document is to describe the DNX profile. This document includes all the information about the sections and elements of the DNX schema, such as the description of each field, the data source of the field, the matching PREMIS semantic unit, and the phase at which the IE lifecycle is created.

DNX and PREMIS

Most of the DNX sections and fields come from the PREMIS data dictionary. Rosetta implements PREMIS (PREMIS compliant), and most of the PREMIS semantic units are represented in the DNX profile. In case semantic units will be added to PREMIS, they will considerably be added to the DNX profile.

Note: Not all the PREMIS fields in the DNX are managed automatically by Rosetta. Some fields can only be filled in and monitored manually – for example, the fields that hold the relationships between different IEs (**relationship** DNX section).

Differences in the Data Model

The differences between the PREMIS data model and Rosetta's data model are as follows:

- Agent entity – In PREMIS, the **Agents** entity holds the details of an agent, which is a person, organization, or software program/system associated with events in the life of an object, or with rights attached to an object. In Rosetta, the agent is only an attribute of an external provenance event, since in the other areas, Rosetta is the agent associated with events in the life of the objects and the access rights attached to the IE.

DNX Section Structure

The DNX format is built from logical groups of metadata fields called **Sections**.

Each DNX section contains a group of fields that are related to each other. For example, the section **generalRepCharacteristics** (General Representation Characteristics) includes the fields that describe the parameters of the representation – Preservation Type, Usage Type, Revision Number, and so forth.

Most of the sections come from the PREMIS data dictionary, but some of them are unique to Rosetta. The structure of a DNX section is as follows:

```
<section id=" Section Name ">
<record>
 <key id="Field Name">Field Value</key>
…
</record>
</section>
```

Each record holds the fields of the section in the form of:

```
<key id=Name>Value</key>
```

The following example illustrates this:

```
- <dnx xmlns="http://www.exlibrisgroup.com/dps/dnx">
  - <section id="generalRepCharacteristics">
    - <record>
        <key id="preservationType">PRESERVATION_MASTER</key>
        <key id="usageType">VIEW</key>
        <key id="RevisionNumber">1</key>
        <key id="DigitalOriginal">false</key>
      </record>
    </section>
```

Structure of a Repeatable Section

If a DNX section is repeatable, there will be multiple records of the same structure, as shown in the following example:

```
  - <section id="internalIdentifier">
    - <record>
        <key id="internalIdentifierType">SIPID</key>
        <key id="internalIdentifierValue">20</key>
      </record>
    - <record>
        <key id="internalIdentifierType">PID</key>
        <key id="internalIdentifierValue">REP1101</key>
      </record>
    - <record>
        <key id="internalIdentifierType">DepositSetID</key>
        <key id="internalIdentifierValue">41</key>
      </record>
    </section>
```

Events within DNX

The event metadata holds the information about actions that affect the object. Each object level has different types of actions that should be captured. In Rosetta, the events that are recorded in the AIP are provenance events, while many other events are captured in the system but do not become part of the AIP metadata.

All events that are generated by the system are written to a database table. Events that are indicated as provenance (in the code, non-configurable) are copied from the table of events to the METS file, while the non-provenance events remain in the table.

The storage of events in a table allows the creation of reports that show the statistics regarding various activities.

Provenance Events

The following types of events are considered provenance events:

- Changes to the IE metadata – adding metadata to any of the IE levels (descriptive DC, source MD, access rights policy, structural map, DNX)
- Addition of a new Representation – new Representation that was added through the Web Editor or as a result of a Preservation Action
- Validation checks – validity and integrity checks on files (Note – Fixity check will not generate a provenance event unless calculated fixity is different than the previous one)
- Enrichment – generation of a persistent identifier

Each such event will be written in the events (`mets:digiprovMD`) section belonging to the relevant object level (IE, representation, or file).

Each event will be written in the DNX format and will include the following:

- Agent – The agent that triggered this event. An agent is not necessarily a person. An agent may also refer to a process, plug-in tool, and so forth.
- Event details – Such as the creation date, a description, the parameters, and so forth

Following is an example of an event that is stored in the `digiProvMD` section of a file. This section holds the events in DNX format:

```
- <section id="event">
  - <record>
      <key id="eventDateTime">2010-01-07 16:35:50</key>
      <key id="eventType">VALIDATION</key>
      <key id="eventIdentifierType">DPS</key>
      <key id="eventIdentifierValue">27</key>
      <key id="eventOutcome1">SUCCESS</key>
      <key
        id="eventOutcomeDetail1">PROCESS_ID=22000;PID=FL1022;SIP_ID=4;
      <key id="eventDescription">Fixity check performed on file</key>
      <key id="linkingAgentIdentifierType1">SOFTWARE</key>
      <key id="linkingAgentIdentifierValue1">REG_SA_JAVA5_FIXITY</key>
    </record>
```

In addition to events, the `digiprovMD` section on the IE level stores the details of the Producer and the Producer Agent who deposited the IE. This section is populated automatically for each

IE in Rosetta and includes all the information of the Producer as it exists in Rosetta at the time of the deposit:

```
- <mets:digiprovMD ID="ie-amd-digiprov">
  - <mets:mdWrap MDTYPE="OTHER" OTHERMDTYPE="dnx">
    - <mets:xmlData>
      - <dnx xmlns="http://www.exlibrisgroup.com/dps/dnx">
        - <section id="producer">
          - <record>
              <key id="userName" />
              <key id="address1">6740</key>
              <key id="address2">Willow Lane</key>
              <key id="address3">Dallas</key>
              <key id="address4">Texas</key>
              <key id="address5">U.S.A</key>
              <key id="defaultLanguage">en</key>
              <key id="emailAddress">marek.melichar@nkp.com</key>
              <key id="firstName">University of Oklahoma</key>
              <key id="jobTitle" />
              <key id="lastName">Legal Department</key>
              <key id="middleName" />
              <key id="telephone1">972-456-6547</key>
              <key id="telephone2" />
              <key id="authorativeName">Library of Legal Department</key>
              <key id="producerId">34366</key>
              <key id="userIdAppId">34362</key>
              <key id="webSiteUrl" />
              <key id="zip" />
            </record>
          </section>
        - <section id="producerAgent">
          - <record>
              <key id="firstName">John</key>
              <key id="lastName">Smith</key>
              <key id="middleName" />
            </record>
          </section>
        </dnx>
      </mets:xmlData>
    </mets:mdWrap>
  </mets:digiprovMD>
```

Access Rights Within DNX

Two types of rights are stored in the DNX sections: PREMIS and non-PREMIS.

- PREMIS rights (IE only) – Information regarding an external system that manages the IE's rights. Note that these rights are not mandatory, and they are not managed or enforced by Rosetta. There is one DNX section for holding the details of these rights:

- `linkingRightsStatementIdentifier` – Holds the type and the value of the statement identifier, if it is generated and stored in a repository other than Rosetta.
    - `linkingRightsStatementIdentifierType` – A designation of the domain within which the `linkingRightsStatementIdentifier` is unique
    - `linkingRightsStatementIdentifierValue` – The value of the `linkingRightsStatementIdentifier`
- Non PREMIS (IE, Representation, and File) – Information regarding the access rights policy managed by Rosetta. Note that it is mandatory for each IE to have an associated access rights policy, while for representation and file access rights are optional. The DNX section for holding this information is `accessRightsPolicy`. The following fields are part of this section:
- PolicyID – The unique ID of the different access rights managed by Rosetta. For example: AR_EMBARGOED_FOR_5_YEARS, AR_5_CONCURRENT_USERS
- Policy description – Description of the policyID. For example: AR_EMBARGOED_FOR_5_YEARS – Embargoed for 5 years, AR_5_CONCURRENT_USERS – Limited access according to copyright law
- Policy parameters – If the policy requires any parameters

Significant Properties of Files Within DNX

To have a scalable structure that supports additions of technical metadata over the years, the DNX section that holds the extracted technical metadata for each file has the following structure:

```
- <section id="significantProperties">
  - <record>
      <key id="significantPropertiesType">image.planarConfiguration</key>
      <key id="significantPropertiesValue">1</key>
      <key id="significantPropertiesExtension" />
    </record>
  - <record>
      <key id="significantPropertiesType">image.maxSampleValue</key>
      <key id="significantPropertiesValue">[1]</key>
      <key id="significantPropertiesExtension" />
    </record>
  - <record>
      <key id="significantPropertiesType">image.minSampleValue</key>
      <key id="significantPropertiesValue">[0]</key>
      <key id="significantPropertiesExtension" />
    </record>
  - <record>
      <key id="significantPropertiesType">image.newSubfileType</key>
      <key id="significantPropertiesValue">0</key>
      <key id="significantPropertiesExtension" />
    </record>
```

This structure allows defining the technical attributes as the values of the `significantPropertiesType` fields, and their values as the values of the `significantPropertiesValue` fields.

DNX Sections

IE level
generalIECharacteristics
objectCharacteristics
CMS
webHarvesting
internalIdentifier
objectIdentifier
significantProperties
linkingIEIdentifier
producer
producerAgent
event
linkingRightsStatementIdentifier
accessRightsPolicy
grantedRightsStatement
metadata
collection
retentionPeriodPolicy
IERelationship

Representation level
generalRepCharacteristics
objectCharacteristics
internalIdentifier
significantProperties
linkingIEIdentifier
event
linkingRightsStatementIdentifier
metadata
preservationLevel
environment
environmentDependencies
environmentSoftware
envSoftwareRegistry
environmentHardware
envHardwareRegistry
environmentExtension
relationship

File level
generalFileCharacteristics
objectCharacteristics
internalIdentifier
significantProperties
linkingIEIdentifier
event
linkingRightsStatementIdentifier
metadata
preservationLevel
environment
environmentDependencies
environmentSoftware
envSoftwareRegistry
environmentHardware
envHardwareRegistry
environmentExtension
relationship
fileFixity
fileFormat
fileVirusCheck
fileValidation
fileTechnicalMetadataExtraction
vsOutcome
creatingApplication
Inhibitors
objectCharacteristicsExtension

signatureInformation
signatureInformationExtension

BitStream level
generalFileCharacteristics
significantProperties

Below is the description for each of the DNX sections.

Note – Defining a section as 'Mandatory' means that the information stored in these sections is required by Rosetta for its functioning. For example, without the internal identifier, objects cannot be searched and found and without populated Format ID, Rosetta will not be able to perform any preservation activities.

This is not the meaning of 'Mandatory' according to PREMIS, and there is no contradiction between the two definitions – Rosetta allows its users to define which fields must be populated as part of the SIP processing. For more details regarding metadata validation see the *Rosetta Configuration Guide* document.

General IE/Rep/File Characteristics

Definition	The **generalIECharacteristics**, **generalRepCharacteristics**, **generalFileCharacteristics** sections contain administrative as well as control attributes that determine how objects are delivered, published, and searched.
Rosetta Mandatory	Yes – Not every field
Source	User
Repeatable	No
Level	IE, Representation, File and BitStream
METS section	techMD

(Rosetta) Object Characteristics

Definition	**objectCharacteristics** – This section can be on each level (IE, representation, and file) and it contains control attributes that are relevant on all levels, such as dates and user information.
Rosetta Mandatory	Yes
Source	System/User
Repeatable	No
Level	IE, Representation and File
METS section	techMD

CMS

Definition	This section holds the Collection Management System details. Each IE in Rosetta can have a "handle" to descriptive metadata that is managed in the ILS, such as Aleph or Voyager. Since this information might be relevant for many IEs and in order to allow a single point of update, the IE holds only the reference to this information, without the need to duplicate it in Rosetta.
Rosetta Mandatory	No
Source	User/System
Repeatable	No
Level	IE
METS section	techMD

Web Harvesting

Definition	**webHarvesting** – This section contains the information regarding Web harvesting. It describes the tool that was used for building the Web archive file and some other parameters of this action. (This section was added because there is no existing set of fields that can hold this metadata according to PREMIS)
Rosetta Mandatory	No
Source	User
Repeatable	No
Level	IE
METS section	techMD

Producer

Definition	This section holds the information of the Producer as it is stored in the staging DB.
Rosetta Mandatory	Yes
Source	System
Repeatable	No
Level	IE
METS section	digiprovMD

Producer Agent

Definition	**producerAgent** - This section holds the information of the Producer Agent who deposited the IE. (It contains only the name, not the entire user record)
Rosetta Mandatory	Yes
Source	System

Repeatable	No
Level	IE
METS section	digiprovMD

Access Rights Policy

Definition	**accessRightsPolicy** - This section holds the access rights policy details that are checked before delivery. The system analyzes whether the calling user is authorized to view the object.
Rosetta Mandatory	Yes
Source	System/User
Repeatable	No
Level	IE, Representation, File
METS section	rightsMD

Granted Rights Statement

Definition	**grantedRightsStatement** – This section holds the copyrights statement that was presented to the Producer Agent upon depositing the IE (boilerplates as part of the material flow). It is currently not in use.
Rosetta Mandatory	No (Currently not in use)
Source	System/User
Repeatable	Yes (no limits)
Level	IE
METS section	rightsMD

Metadata (Deprecated)

Definition	This table is deprecated and not in use. This record holds the details of the HDEMETADATA record that is kept in the sourceMD METS section. The details are used by the system to allow accurate matching between the data in the METS to the data in the DB, when the IE is loaded back to the staging DB from the permanent repository. The details include the ID and the type (DC, DNX_REP, and so forth) as well as the control dates (creation, modification).
Rosetta Mandatory	No
Source	System
Repeatable	Yes (no limits)
Level	IE, Representation and File
METS section	sourceMD

Retention Policy

Definition	Hold the details of Retention Policy ID which determines the duration required to preserve content, after which content will be deleted.
Rosetta Mandatory	No
Source	User
Repeatable	No
Level	IE
METS section	techMD

Internal Identifier

Definition	**internalIdentifier** – This section holds a record for each of the identifiers that are created by Rosetta, such as PID, SIP ID, and Deposit Set ID. Each object level has its own section of identifiers (there is a PID for each IE, representation, and file), while on the IE level there are other identifiers (such as SIP ID).
Rosetta Mandatory	Yes – All types of internal identifiers are Rosetta Mandatory since they are created and used by the system
Source	System
Repeatable	Yes (no limits)
Level	IE, Representation, and File
METS section	techMD

Object Identifier

Definition	**objectIdentifier** – This section holds the identifiers of the IE that are stored in an external system – for example, Handle and URN: NBN. These identifiers are not internal in the sense that in Rosetta they are used only as metadata, and not as identifiers. These identifiers can be generated in Rosetta by a plug-in or they can be populated pre-ingest by the submission application.
Rosetta Mandatory	No
Source	User/System
Repeatable	Yes (no limits)
Level	IE, Representation, and File
METS section	techMD

Preservation Level

Definition	**preservationLevel** – This section holds information indicating the decision or policy on the set of preservation functions to be applied to an IE and the context in which the decision or policy was made.
Rosetta Mandatory	No
Source	User
Repeatable	No
Level	Representation and File
METS section	techMD

Significant Properties

Definition	**significantProperties** – This section holds the extracted technical metadata for each file. However, it can be used in any of the other levels and it can hold other properties that were not extracted by the MD Extraction tool(s).
Rosetta Mandatory	No (Depends on the MD Extraction tool that is associated with the Format)
Source	System/User
Repeatable	Yes (no limits)
Level	IE, Representation File and BitStream
METS section	techMD

File Fixity

Definition	**fileFixity** – For each file, this section holds a record for each checksum algorithm that is used by the validation stack (SHA-1, CRC32, and MD5).
Rosetta Mandatory	No
Source	System
Repeatable	Yes – For every checksum algorithm in use by the Fixity task
Level	File
METS section	techMD

File Format

Definition	**fileFormat** – For each file, this section holds the format details as they were identified by the format identification task in the validation stack.
Rosetta Mandatory	Yes
Source	System/User
Repeatable	No
Level	File
METS section	techMD

File Virus Check

Definition	**fileVirusCheck** – For each file, this section holds the results of the virus check that was performed as part of the validation stack.
Rosetta Mandatory	No
Source	System
Repeatable	No

Level	File
METS section	techMD

File Validation

Definition	**fileValidation** – For each file, this section holds the details and the results (valid/invalid, well-formed/not well formed) of the format validation tool that was used by the Format Validation task (or the soon to be deprecated MD Extraction with Validation task) as part of the validation stack. Note that this section does not hold the actual output of the extraction tool (for example, JHOVE). The output is stored in the significant properties section and holds the information about the extraction tool.
Rosetta Mandatory	No
Source	System
Repeatable	No
Level	File
METS section	techMD

File Technical Metadata Extraction

Definition	**fileTechnicalMetadataExtraction** – For each file, this section holds the extraction tool information (agent name, plug-in name, errors when relevant) of the technical MD extraction tool that was used by the MD Extraction task as part of the validation stack. Note that this section does not hold the actual output of the extraction tool (for example, JHOVE). The output is stored in the significant properties section, while this section holds the information about the extraction tool.
Rosetta Mandatory	No
Source	System
Repeatable	No
Level	File

METS section	techMD

Validation Stack Outcome

Definition	**vsOutcome** – This section holds the information about the validation routines that were used to validate the files. The validation includes the following: a virus check, fixity check, format identification, technical metadata extraction and risk extraction. Different plug-ins can be used and their details are captured in this section.
Rosetta Mandatory	Yes
Source	System
Repeatable	Yes – Repeated for every task in the VS task chain
Level	File
METS section	techMD

Creating Application

Definition	**creatingApplication** – For each file, this section holds the information about the application that was used for creating the file, which was created before it was deposited or in Rosetta as part of a preservation action.
Rosetta Mandatory	No
Source	System/User
Repeatable	No
Level	File
METS section	techMD

Inhibitors

Definition	On a file level, this section holds the features intended to inhibit access, use, or migration.
Rosetta Mandatory	No
Source	User
Repeatable	Yes (no limits)
Level	File
METS section	techMD

Object Characteristics Extension

Definition	**objectCharacteristicsExtension** – On a file level, this is a container for including semantic units that are not DNX.
Rosetta Mandatory	No
Source	User
Repeatable	Yes (no limits)
Level	File
METS section	techMD

Environment

Definition	On a file or representation level, this section holds the details of hardware/software combination that supports the usage (rendering, viewing) of the representation/file.
Rosetta Mandatory	No
Source	User
Repeatable	Yes (no limits)

Level	Representation, File
METS section	techMD

Environment Dependencies

Definition	**environmentDependencies** - On a file or representation level, this section holds information about a non-software component or associated file required in order to use or render the representation or file - for example, a schema, DTD, or an entity file declaration.
Rosetta Mandatory	No
Source	User
Repeatable	Yes (no limits)
Level	Representation, File
METS section	techMD

Environment Software

Definition	**environmentSoftware** – This section holds the details of the software that is needed for rendering the object (file, representation). The details include name, version, type, and dependencies.
Rosetta Mandatory	No
Source	User
Repeatable	Yes (no limits)
Level	Representation, File
METS section	techMD

Environment Software Registry

Definition	**envSoftwareRegistry** – This section holds the details of the registry in which the environment software is registered.
Rosetta Mandatory	No
Source	User
Repeatable	Yes (no limits)
Level	Representation, File
METS section	techMD

Environment Hardware

Definition	**environmentHardware** – This section holds the details of the hardware that is required for rendering the object (file, representation). The details include name and type.
Rosetta Mandatory	No
Source	User
Repeatable	Yes (no limits)
Level	Representation, File
METS section	techMD

Environment Hardware Registry

Definition	**envHardwareRegistry** – This section holds the details of the registry in which the environment hardware is registered.
Rosetta Mandatory	No
Source	User
Repeatable	Yes (no limits)

Level	Representation, File
METS section	techMD

Environment Extension

Definition	**environmentExtension** – This section is a container for including semantic units that are not DNX.
Rosetta Mandatory	No
Source	User
Repeatable	Yes (no limits)
Level	Representation, File
METS section	techMD

Signature Information

Definition	**signatureInformation** – On a file level, this section can hold the information that is required for using a digital signature to authenticate the signer of an object and/or the information contained in the object.
Rosetta Mandatory	No
Source	User
Repeatable	No
Level	File
METS section	techMD

Signature Information Extension

Definition	**signatureInformationExtension** – This section holds digital signature information using semantic units that are not DNX.
Rosetta Mandatory	No
Source	User
Repeatable	Yes (no limits)
Level	File
METS section	techMD

Relationship

Definition	This section holds the relations between files or between representations, if there are any.
Rosetta Mandatory	No
Source	User/System (During Add Representation or Preservation Action)
Repeatable	Yes (no limits)
Level	File, Representation
METS section	techMD

IE Relationship

Definition	This section holds the structural IE relationships between a parent structural IE and its child IEs.
Rosetta Mandatory	No
Source	User
Repeatable	Yes (no limits)

Level	IE
METS section	techMD

Linking IE Identifier

Definition	**linkingIEIdentifier** – This section holds the identifier of a different IE that is related to the object (IE, representation, or file)
Rosetta Mandatory	No
Source	User
Repeatable	Yes (no limits)
Level	IE, Representation or File
METS section	techMD

Event

Definition	This section holds the provenance events on each level (IE, representation, and file)
Rosetta Mandatory	Yes – The provenance events are Rosetta Mandatory.
Source	User/System
Repeatable	Yes (no limits)
Level	IE, Representation or File
METS section	digiprovMD

Linking Rights Statement Identifier

Definition	**linkingRightsStatementIdentifier** – This section holds the identifier of a copyrights statement that may be stored outside of Rosetta.
Rosetta Mandatory	No
Source	User
Repeatable	Yes (no limits)
Level	IE, Representation, or File
METS section	rightsMD

Collection

Definition	collection – This section holds the information of the collection(s) that the IE is associated with. There could be multiple records pointing to multiple collections/sub-collections. The collection METS will have one record that holds the identifiers of the collection and the parent collection (if exists).
Rosetta Mandatory	No
Source	User
Repeatable	Yes for IE (no limits), no in case of collection METS
Level	IE, Collection
METS section	techMD

The full list of fields in each section is specified in Appendix B below.

A

Appendix A – METS XML

- The link to the METS example (in the Library of Congress - Rosetta METS Profile) is:

  http://www.loc.gov/standards/mets/profiles/00000042.xml

- The link to mets_rosetta.xsd is

  ```
  https:// github.com/ExLibrisGroup/Rosetta.dps-
  sdkprojects/blob/master/{version}/dps-sdk-projects/dps-sdk-
  deposit/src/xsd/mets_rosetta.xsd
  ```

  For example:

  https://github.com/ExLibrisGroup/Rosetta.dps-sdk-projects/blob/master/6.0/dps-sdk-projects/dps-sdk-deposit/src/xsd/mets_rosetta.xsd

B

Appendix B – DNX Data Dictionary

The source field indicates how the field is populated. The possible values are:

- User – The field can be updated either by the staff users via the Web editor, or by the submission application, before the SIP is submitted.
- System – The field is updated automatically by one of Rosetta modules. This table will include as much as possible the exact stage and task that updates the field.

Note: In sections that are mandatory for Rosetta, not all fields are mandatory. The fields that are will be indicated as 'Rosetta Mandatory' in the table below.

Name	Description	Source	PREMIS path
generalIECharacteristics	IE attributes		Non-PREMIS
submissionReason	The reason for submitting the IE (for example, donation of this collection to the National Library)	User – Pre ingest or through the UI (MD form, web editor)	
status	The status of the IE - Active or Suppressed (Only active IEs are published). A suppressed IE is not available for searches from outside of Rosetta (when the search is performed via the SRU protocol). It can be used as an alternative for deleting the IE altogether from Rosetta.	User – Pre-ingest or through the UI (web editor)	
generalIECharacteristics.Version	The version of the IE	System – when the IE is committed to permanent	
statusDate	The date of the status change.	System – Upon committing the new version of the IE.	
IEEntityType	Logical categories such as digitized books, e-journals, videos, and so forth.	User – Pre-ingest or through the web editor	
UserDefinedA	These fields are available for users for storing metadata that has no designated DNX field.	User – Pre-ingest or through the web editor	
UserDefinedB			
UserDefinedC			

Name	Description	Source	PREMIS path
generalRepCharacteristics	Representation attributes	User	Non-PREMIS
label	Label of the representation for display in Delivery.	Users can populate this section pre-ingest or through the Web editor	
preservationType	Possible values are: Preservation Master – The master representation for preservation – for example, TIFF files of the scanned book. Modified Master – The representation to be created out from the master representation – for example, text files for each page of the scanned book. Derivative Copy – Access copies that are not preserved; used for better and faster access to Delivery. **Note:** IEs must contain one preservation master and may contain up to one modified master. Other types are unlimited. *** Rosetta Mandatory**		
usageType	Currently, Rosetta supports only the **VIEW** value as a type of usage. A representation that is defined otherwise will not be available for Delivery. *** Rosetta Mandatory**		
representationEntityType	This field is used to distinguish between different structures of representations.		
contentType	This field is used to distinguish between different types of content.		
contextType	This field is used to distinguish between different types of context.		
hardwareUsed	The hardware that was used to create the content of this Representation.		

Name	Description	Source	PREMIS path
physicalCarrierMedia	The physical media that the content was originally on.		
derivedFromId	Used in migration and 'adding Representation' tasks to identify the source representation that was used to derive this representation.		
deliveryPriority	The priority of listing the representations in Delivery.		
orderingSequence	The ordering sequence of representations of the same priority.		
DigitalOriginal	Used to mark a Preservation Master as being the digital original of the IE.		
RevisionNumber	Used to sequence the set of Master Representations. The highest revision number is by definition the latest master.		
RepresentationCode	A code that categorizes a representation in terms of how it can be used in Delivery and external systems.		
TaskID	An identifier of a certain type of derivative copy, used to manage the derivative copy when a task operates on an IE.		
RepresentationOriginalName	The name of the object as submitted to or harvested by the repository, before any renaming by the repository.		OBJECT:OriginalName
UserDefinedA	These fields available for the users for storing metadata that has no designated DNX field.		
UserDefinedB			
UserDefinedC			
generalFileCharacteristics	File attributes	User/System	PREMIS
label	When there is no struct map as part of the ingest, a physical struct map is generated automatically using this label	User – Pre-ingest or as part of	

Name	Description	Source	PREMIS path
	for file delivery. Once generated, this label is no longer in effect.	deposit	
note	An internal note pertaining to the file.	User – Pre-ingest or as part of deposit	
FileEntityType	Sub-format (GeoTiff)	User – Pre-ingest	
compositionLevel	An indication of whether the object is subject to one or more processes of decoding or unbundling.	User – Pre-ingest	OBJECT:objectCharacteristics.compositionLevel
fileLocationType	The means of referencing the location of the content. During the Loading stage the value is set to 'FILE' as a default.	System - Loading stage	OBJECT:storage.contentLocation.contentLocationType
fileLocation	Could be used for reference to the location of the content. (Currently not in use)	User – Pre-ingest	OBJECT:storage.contentLocation.contentLocationValue
fileOriginalName	The name of the object as submitted to or harvested by the repository, before any renaming by the repository	User – Pre-ingest or during web deposit (bulk load)	OBJECT:OriginalName
fileOriginalPath	The original path from which the file was taken (user PC, server). Exported files are structured according to this path.	User – Pre-ingest or during web deposit (bulk load)	
fileOriginalID	Used to store the location of the file on the Deposit server.	System – Loading stage *** Rosetta Mandatory**	
fileExtension	The normalized and definitive file extension for the file.	System – loading stage, based on the file name	

Name	Description	Source	PREMIS path
		*** Rosetta Mandatory**	
fileMIMEType	The definitive MIME type for the file. Note – This value should be the same as in the fileFormat section. If the value was populate by the user, the system will not override it, and then it might mismatch the value in the fileFormat section.	User – can populate in Submission application System – During format validation task, based on format library *** Rosetta Mandatory**	
fileSizeBytes	File size in bytes.	System – Calculated by fixity task. (It can be populated pre-ingest as well)	OBJECT:objectCharacteristics.size
storageID	Currently not in use.		OBJECT:storage.storageMedium
streamRefId	Currently not in use.		
formatLibraryId	Reference to the file format library. The format library ID is used by the system to search and create sets, based on a given format. (For example, preservation sets).	System – Validation Stack – Format Identification task *** Rosetta Mandatory**	
riskLibraryIdentifiers	List of 'Tool Driven' risks associated with the file format. All the 'Tool Driven' risks that were found as relevant for the file are stored in this field (concatenated to a string). 'Property Driven' risks are not stored in the DNX.	System – Validation Stack – Risk Extraction task	

Name	Description	Source	PREMIS path
fileCreationDate	The date the file was created	User – during ingest System - during ingest (if not provided by user)	
fileModificationtionDate	The date the file was modified.	User – during ingest System - during ingest (if not provided by user)	
objectCharacteristics	Attributes applicable to all types	System	Non-PREMIS
objectType	INTELLECTUAL_ENTITY, REPRESENTATION, FILE	Pre-ingest or during loading stage *** Rosetta Mandatory**	
parentID	This field is currently not in use		
groupID	Link between objects on the same level that cross over representation boundaries	Pre-ingest or during loading stage	
creationDate	Creation date of the IE.	Loading stage *** Rosetta Mandatory**	
createdBy	User who created the IE.	Loading stage *** Rosetta Mandatory**	
modificationDate	Modification date.	Upon committing new version *** Rosetta Mandatory**	

Name	Description	Source	PREMIS path
modifiedBy	User who was the last to modify the IE.	Upon committing new version *** Rosetta Mandatory**	
owner	Consortia ownership path.	Loading stage – Based on the output of the SIP routing rule *** Rosetta Mandatory**	
CMS	Collection Management System	User / System	Non-PREMIS
system	The ILS system in which this IE is described - Aleph, Voyager, Taphui	User – Pre-ingest System – In the CMS Update task	
recordId	The record ID in the ILS.		
mId	The ID of the record in the HDEMETADATA table. (System generated)		
webHarvesting	Metadata of Web harvesting (NLNZ)	User	Non-PREMIS
primarySeedURL	The URL of the harvested domain.	Pre-ingest or through the web editor.	
WCTIdentifier	The Web curator tool that was used for the Web harvest.		
targetName	The name of the harvested site		
group	If a group of sites is being harvested, this is the name of the group.		
harvestDate	Date of the Web harvest.		
harvestTime	Time of the Web harvest		
internalIdentifier	Rosetta internal identifiers	System	Non-PREMIS
internalIdentifierType	For example, PID, UUID, SIPID, DepositSetID	This section is populated during	

Name	Description	Source	PREMIS path
internalIdentifierValue	Deposit ID: Generated ORACLE sequence SIP ID: Generated ORACLE sequence IE PID: IE + Generated ORACLE sequence REP PID: REP + Generated ORACLE sequence FILE PID: FL + Generated ORACLE sequence BITSTREAM PID: BS + Generated ORACLE sequence	the loading phase. *** Rosetta Mandatory**	
objectIdentifier	A designation used to uniquely identify the object within the preservation repository system in which it is stored.	System, User	PREMIS
objectIdentifierType	A designation of the domain within which the object identifier is unique. (for example, Handle, URN)	User – Pre-ingest System – as part of the PiGeneratorGenericTask task	OBJECT:objectIdentifier.objectIdentifierType
objectIdentifierValue	The value of the objectIdentifier.		OBJECT:objectIdentifier.objectIdentifierValue
preservationLevel	Preservation level	User	PREMIS
preservationLevelValue	A value indicating the set of preservation functions expected to be applied to the object (for example, bit-level, full)	Pre-ingest or through the web-editor	OBJECT:preservationLevel.preservationLevelValue
preservationLevelRole	A value indicating the context in which a set of preservation options is applicable (for example, requirement, intention, capability)		OBJECT:preservationLevel.preservationLevelRole

Name	Description	Source	PREMIS path
preservationLevelType	A value indicating the type of preservation functions expected to be applied to the object for this preservation level.		OBJECT:preservationLevel.preservationLevelType
preservationLevelRationale	The reason a particular preservationLevelValue was applied to the object.		OBJECT:preservationLevel.preservationLevelRationale
preservationLevelDateAssigned	The date, or date and time, when a particular preservationLevelValue was assigned to the object.		OBJECT:preservationLevel.preservationLevelDateAssigned
significantProperties	Significant properties	System/User	PREMIS
significantPropertiesType	The aspect, facet, or attribute of an object about which significant properties are being described (for example, content, structure, behavior, page count, page width)	This section is populated for files, by the MD Extraction task during the Validation Stack stage. However, users can use this section in the IE and Representation level and populate it pre-ingest	OBJECT:significantProperties.significantPropertiesType
significantPropertiesValue	Description of the characteristics of a particular object. These properties are important for characterizing objects, finding objects, identifying risk in objects, and evaluating the success of preservation actions.		OBJECT:significantProperties.significantPropertiesValue
significantPropertiesExtension	A container to include semantic units defined outside of PREMIS for significant properties.		OBJECT:significantProperties.significantPropertiesExtension
fileFixity	The file fixity	System	PREMIS
agent	The agent ID of the file fixity utility (such as JAVA SE v5) if performed internally, or externally provided agent information.	Populated by the Fixity task (part of the validation stack)	OBJECT:objectCharacteristics.fixity.messageDigestOriginato

Name	Description	Source	PREMIS path
		*** Rosetta Mandatory**	r
fixityType	The specific algorithm used to construct the message digest for the digital object (such as MD5, CRC32, and SHA-256).		OBJECT:objectCharacteristics.fixity.messageDigestAlgorithm
fixityValue	The output of the message digest algorithm.		OBJECT:objectCharacteristics.fixity.messageDigest
fileFormat	The file format	System	PREMIS
Agent	Agent ID of the file format utility (such as DROID v1 definition file v17). *** Rosetta Mandatory**	These fields are populated during the format validation task (runs as part of the validation task)	
formatRegistry	Type of format library record (such as PRONOM).		OBJECT:ObjectCharacteristics.format.formatRegistry.formatRegistryName
formatRegistryId	The registry ID.		OBJECT:ObjectCharacteristics.format.formatRegistry.formatRegistryKey
formatRegistryRole	The purpose or expected use of the registry: format identification.		OBJECT:ObjectCharacteristics.format.formatRegistry.formatRegistryRole
formatName	A designation of the format of the file or bitstream. (The ID of the Format in the Format Library e.g. fmt/7 for TIFF)		OBJECT:ObjectCharacteristics.format

Name	Description	Source	PREMIS path
	*** Rosetta Mandatory**		.formatDesignation.formatName
formatVersion	The version of the format listed in formatName. *** Rosetta Mandatory**		OBJECT:ObjectCharacteristics.format.formatDesignation.formatVersion
formatDescription	The format description from the format library entry. (e.g. 'Tagged Image File Format' for format fmt/7 (TIFF))		
formatNote	Contains the note as entered by the Technical Analyst upon the manual format assignment.		OBJECT:ObjectCharacteristics.format.formatNote
exactFormatIdentification	True – in case the format was identified by the format identification tool (e.g. DROID) False – in case the Format was not identified by the tool (multiple or tentative results) and it was populated based on auto-correction rule or manually by a staff user (Technical Analyst).		
mimeType	Based on the mime type in the format library.		
agentVersion	The version of the tool used for identifying the format (such as DROID). *** Rosetta Mandatory**		
agentSignatureVersion	The version of the format's signatures file that is used by the format identification tool. *** Rosetta Mandatory**		
IdentificationMethod	The method that the file format is identified. Possible values are:		

Name	Description	Source	PREMIS path
	signature, extension, container, manual, and rule		
fileVirusCheck	The file virus check	System	Non-PREMIS
Agent	Agent ID of the file virus check utility (such as MacAfee v32, def file 98).	This section is populated by the virus check task (part of the validation stack)	
Status	Passed/fail		
content	The descriptive output of the virus check utility.		
fileValidation	The file validation utility (such as jhove, NLNZ MD extractor)	System	Non-PREMIS
Agent	Name and version of the file validation utility *** Rosetta Mandatory**	This section is as part of the validation stack	
pluginName	Name of the plugin used for validation		
Status	This field is currently not in use.		
errorMessage	The error message received during the MD extraction (if received)		
errorId	The error ID received during the MD extraction (if received)		
Format	A designation of the format of the file or bitstream. (for example, TIFF) *** Rosetta Mandatory**		
version	The version of the format listed in format. (such as 1.0) *** Rosetta Mandatory**		
mimeType	Specific output of JHOVE/NLNZ. In case no mimeType was identified by the utility, this field will remain empty.		

Name	Description	Source	PREMIS path
Profile	This field is currently not in use.		
isValid	True/False indicator		
isWellFormed	True/False indicator		
agentVersion	This field is currently not in use.		
validationDetails	Further details regarding the validation (if applicable)		
fileTechnicalMetadataExtraction	The technical MD extraction utility	System	Non-PREMIS
agent	The agent that did the MD extraction	This section is populated during the MD extraction task	
pluginName	The plug-in that did the MD extraction		
errorMessage	The error message received during the MD extraction (if received)		
errorId	The error ID received during the MD extraction (if received)		
vsOutcome	The validation stack results - one record per check	System	Non-PREMIS
checkDate	The date of the validation stack *** Rosetta Mandatory**	This section is populated by the validation stack task chain.	
vsAgent	Same agent as in the output of specific sections *** Rosetta Mandatory**		
Type	Fixity, virus, validation, and so forth *** Rosetta Mandatory**		
Result	Passed, fail *** Rosetta Mandatory**		
resultDetails	The automatic output from the validation stack agent (such as: `/exlibris/dps/d4_1/profile/repository/storage1/2010/01/03/file_1/FL1123 is Virus Free`)		
vsEvaluation	Passed, fail		

Name	Description	Source	PREMIS path
	*** Rosetta Mandatory**		
vsEvaluationDetails	If the result was changed by the Technical Analyst, this field will capture the details.		
creatingApplication	Information about the application that created the object	User/System	PREMIS
creatingApplicationName	A designation for the name of the software program that created the object (such as MSWord)	This section can be populated by: • the user in the submission application, external preservation plan (descriptor file) or via the web editor • the system during preservation action (internal plan)	OBJECT:ObjectCharacteristics.creatingApplication.creatingApplicationName
creatingApplicationVersion	The version of the software program that created the object.		OBJECT:ObjectCharacteristics.creatingApplication.creatingApplicationVersion
dateCreatedByApplication	The actual or approximate date and time the object was created.		OBJECT:ObjectCharacteristics.creatingApplication.dateCreatedByApplication
creatingApplicationExtension	Application creation information using semantic units defined externally to PREMIS.		OBJECT:ObjectCharacteristics.creatingApplication.creatingApplicationExtension
inhibitors	Features of the object intended to inhibit access, use, or migration.	User	PREMIS
inhibitorType	The inhibitor method employed (such as DES, PGP, Blowfish Password	User can populate this	OBJECT:ObjectCharacteristics.inhibit

Name	Description	Source	PREMIS path
	protection).	section pre-ingest (in submission application) or via web editor	ors.inhibitor Type
inhibitorTarget	The content or function protected by the inhibitor (All content, Function: Play, Function: Print).		OBJECT:ObjectCharacteristics.inhibitors.inhibitorTarget
inhibitorKey	The decryption key or password.		OBJECT:ObjectCharacteristics.inhibitors.inhibitorKey
objectCharacteristicsExtension	A container to include semantic units defined outside of PREMIS.	User	PREMIS
objectCharacteristicsExtension		User can populate this section pre-ingest	OBJECT:ObjectCharacteristics.objectCharacteristicsExtension
environment	Hardware/software combinations supporting the use of the object.	User	PREMIS
environmentCharacteristic	An assessment of the extent to which the described environment supports its purpose (such as recommended, minimum).	User can populate this section pre-ingest (in submission application) or via web editor	OBJECT:environment.environmentCharacteristic
environmentPurpose	The use(s) supported by the specified environment (such as render, edit).		OBJECT:environment.environmentPurpose
environmentNote	Additional information about the environment.		OBJECT:environment.environmentNote
environmentDependencies	Information about a non-software component or associated file needed in order to use or render the representation or file - for example, a schema, DTD, or an entity file declaration.	User	PREMIS
dependencyName	A designation for a component or		OBJECT:env

Name	Description	Source	PREMIS path
	associated file needed by the representation or file.	User can populate this section pre-ingest (in submission application) or via web editor	ironment.dependency.dependencyName
dependencyIdentifierType1	A designation of the domain in which the identifier of the dependent resource is unique (for example, URI).		OBJECT:environment.dependency.dependencyIdentifier.dependencyIdentifierType
dependencyIdentifierValue1	The value of the dependencyIdentifier (for example, http://www.teic.org/P4X/DTD/teicorp2.dtd)		OBJECT:environment.dependency.dependencyIdentifier.dependencyIdentifierValue
dependencyIdentifierType2	These fields allow holding another 2 sets of environment dependencies.		
dependencyIdentifierValue2			
dependencyIdentifierType3			
dependencyIdentifierValue3			
environmentSoftware	Software required for rendering or using the object.	User	PREMIS
softwareName	Manufacturer and title of the software application (for example, Adobe Photoshop).	User can populate this section pre-ingest (in submission application) or via web editor	OBJECT:environment.software.swName
softwareVersion	The version or versions of the software referenced in softwareName (for example, >=2.2.0).		OBJECT:environment.software.swVersion

Name	Description	Source	PREMIS path
softwareType	Class or category of software (for example, renderer, operating system).		OBJECT:environment.software.swType
softwareOtherInformation	Additional requirements or instructions related to the software referenced in softwareName.		OBJECT:environment.software.swOtherInformation
softwareDependancy	The name and, if applicable, version of any software component needed by the software referenced in softwareName in the context of using this object.		OBJECT:environment.software.swDependancy
envSoftwareRegistry	Software required for rendering or using the object.	User	Non-PREMIS
registryId		User can populate this section pre-ingest (in submission application) or via web editor	

Name	Description	Source	PREMIS path
environmentHardware	Hardware required for rendering or using the object.	User	PREMIS
hardwareName	Manufacturer, model, and version (if applicable) of the hardware (for example, Intel Pentium III).	User can populate this section pre-ingest (in submission application) or via web editor	OBJECT:environment.software.hwName
hardwareType	Class or category of the hardware (for example, processor, memory, input/output device).		OBJECT:environment.software.hwType
hardwareOtherInformation	Additional requirements or instructions related to the hardware referenced in hardwareName (for example, 32MB minimum).		OBJECT:environment.software.hwOtherInformation
envHardwareRegistry	Hardware required for rendering or using the object.	User	Non-PREMIS
registryId		User can populate this section pre-ingest (in submission application) or via web editor	
environmentExtension	A container to include semantic units defined outside of PREMIS.	User	Non-PREMIS
environmentExtension		User can populate this section pre-ingest (in submission application) or via web editor	OBJECT:environment.environmentExtension
signatureInformation	Information needed to use a digital signature to authenticate the signer of an object and/or the information contained in the object.	User	PREMIS
signatureInformationEncoding	The encoding used for the values of signatureValue, keyInformation (for	User can populate this	OBJECT:signatureInfor

Name	Description	Source	PREMIS path
	example, Base64, Ds:CrytoBinary).	section pre-ingest (in submission application) or via web editor	mation.signature.signatureInformationEncoding
signer	The individual, institution, or authority responsible for generating the signature.		OBJECT:signatureInformation.signature.signer
signatureMethod	A designation for the encryption and hash algorithms used for signature generation (for example, DSA-SHA-256,DSA-SHA-256).		OBJECT:signatureInformation.signature.signatureMethod
signatureValue	The digital signature; a value generated from the application of a private key to a message digest.		OBJECT:signatureInformation.signature.signatureValue
signatureValidationRules	The operations to be performed in order to validate the digital signature.		OBJECT:signatureInformation.signature.signatureValidationRules
signatureProperties	Additional information about the generation of the signature.		OBJECT:signatureInformation.signature.signatureProperties
keyInformation	Information about the signer's public key needed to validate the digital signature.		OBJECT:signatureInformation.signature.keyInformation

Name	Description	Source	PREMIS path
signatureInformationExtension	Digital signature information using semantic units defined outside of PREMIS.	User	PREMIS
signatureInformationExtension		User can populate this section pre-ingest (in submission application)	OBJECT:signatureInformation.signatureInformationExtension
relationship		User	PREMIS
relationshipType	A high-level categorization of the nature of the relationship.	User can populate this section pre-ingest (in submission application) or via web editor	OBJECT:relationship.relationshipType
relationshipSubType	A specific characterization of the nature of the relationship documented in relationshipType.		OBJECT:relationship.relationshipSubType
relatedObjectIdentifierType1	A designation of the domain within which the identifier is unique.		OBJECT:relationship.relatedObjectIdentification.relatedObjectIdentifierType
relatedObjectIdentifierValue1	The value of the related object identifier.		OBJECT:relationship.relatedObjectIdentification.relatedObjectIdentifierType
relatedObjectSequence1	The order of the related object relative to other objects with the same type of relationship.		OBJECT:relationship.relatedObjectIdentification.relatedObjectSequence
relatedObjectIdentifierTyp			

Name	Description	Source	PREMIS path
e2	These fields allow holding another 2 sets of relations.		
relatedObjectIdentifierValue2			
relatedObjectSequence2			
relatedObjectIdentifierType3			
relatedObjectIdentifierValue3			
relatedObjectSequence3			
IERelationship		User	PREMIS
relationshipType	A high-level categorization of the nature of the relationship. Accepted value "STRUCTURAL".	User can populate this section pre-ingest (in submission application) or via web editor	OBJECT:relationship.relationshipType
relationshipSubType	A specific characterization of the nature of the relationship. Accepted value "HAS_PART".		OBJECT:relationship.relationshipSubType
relatedObjectIdentifierType	The type of the identifier for which the related child IE identifier is unique. Accepted value "PID".		OBJECT:relationship.relatedObjectIdentifier.relatedObjectIdentifierType
relatedObjectIdentifierValue	The value of the related child IE identifier.		OBJECT:relationship.relatedObjectIdentifier.relatedObjectIdentifierValue
relatedObjectSequence	The order of the related child IE relative to other child IEs under the parent IE. Accepted value is a positive integer.		OBJECT:relationship.relatedObjectIdentifier.relatedObjectSequ

Name	Description	Source	PREMIS path
			ence
linkingIEIdentifier	An identifier for an intellectual entity associated with the object.	User	PREMIS
linkingIEIdentifierType	The type of the linked object.	User can populate this section pre-ingest (in submission application) or via web editor	OBJECT:linkingIntellectualEntityIdentifier.linkingIntellectualEntityIdentifierType
linkingIEIdentifierValue	The ID of the linked object.		OBJECT:linkingIntellectualEntityIdentifier.linkingIntellectualEntityIdentifierValue
producer	Producer information	System	Non-PREMIS
producerId	The Producer ID in the Rosetta DB. *** Rosetta Mandatory**	This section is populated by the system during the loading stage, as part of the SIP processing.	
userIdAppId	The user ID in the Rosetta DB. *** Rosetta Mandatory**		
defaultLanguage	Default language of the Producer.		
authorativeName	Authoritative name of the Producer. *** Rosetta Mandatory**		
userName	The user name (if the Producer is not an organization).		
firstName	First name of the Producer. *** Rosetta Mandatory**		
lastName	Last name of the Producer. *** Rosetta Mandatory**		
middleName	Middle name of the Producer.		

Name	Description	Source	PREMIS path
jobTitle	Job title (if the Producer has a user record of a staff member).		
address1	The Producer's address (street, neighborhood, city, state). *** Rosetta Mandatory**		
address2			
address3			
address4			
address5			
zip	Zip code.		
emailAddress	E-mail address. *** Rosetta Mandatory**		
webSiteUrl	Producer's Web site.		
telephone1	Phone number. *** Rosetta Mandatory**		
telephone2			
producerAgent	Producer Agent information	System	Non-PREMIS
firstName	First name of the Producer Agent. *** Rosetta Mandatory**	This section is populated by the system during the loading stage, as part of the SIP processing.	
lastName	Last name of the Producer Agent. *** Rosetta Mandatory**		
middleName	Middle name of the Producer Agent.		
event	Events that were created during the SIP processing	System/User	PREMIS
eventIdentifierType	Type of event (for example, Rosetta) *** Rosetta Mandatory**	Events are	EVENT:eventIdentifier.eventIdentif

Name	Description	Source	PREMIS path
		created by the system. All the provenance events are written into the DNX upon creation and copied to the METS file upon moving to permanent. Pre-ingest events can be created by the users and written in the METS as part of the submission application.	ierType
eventIdentifierValue	ID of the event in the generating system *** Rosetta Mandatory**		EVENT:eventIdentifier.eventIdentifierValue
eventType	Event category in the generating system (processing, deposit) *** Rosetta Mandatory**		EVENT:eventType
eventDescription	Description of the event in the generating system. *** Rosetta Mandatory**		
eventDateTime	Event date. *** Rosetta Mandatory**		EVENT:eventDateTime
eventOutcome1	Event outcome can be Success or Failure.		EVENT:eventOutcomeInformation.eventOutcome
eventOutcomeDetail1	A detailed description of the result or product of the event in textual form.		EVENT:eventOutcomeInformation.eventOutcomeDetail.eventOutcomeDetailNote
eventOutcomeDetailExtension1	If the event contains an XML file that should be stored as is. Note: XML should be wrapped as CDATA		EVENT:eventOutcomeInformation.eventOutcomeDetail.eventOutcomeDetailExtension
eventOutcome2	The section structure allows to have 3 sets of eventOutcome , eventOutcomeDetail and		
eventOutcomeDetail2			

Name	Description	Source	PREMIS path
eventOutcomeDetailExtension2	eventOutcomeDetailExtension as part of the same event.		
eventOutcome3			
eventOutcomeDetail3			
eventOutcomeDetailExtension3			
linkingAgentIdentifierXMLID1	If there is an XML that is generated by the agent.		
linkingAgentIdentifierType1	If the event is being generated by different agents, such as specific software		EVENT:linkingAgentIdentifier.linkingAgentIdentifierType
linkingAgentIdentifierValue1	The ID of the agent - software tool name.		EVENT:linkingAgentIdentifier.linkingAgentIdentifierValue
linkingAgentRole1	The role of the agent - migration, rendering, and so forth.		EVENT:linkingAgentIdentifier.linkingAgentRole
linkingAgentIdentifierXMLID2	The section structure allows to have 3 sets of linkingAgentIdentifierXMLID, linkingAgentIdentifierType, linkingAgentIdentifierValue and linkingAgentRole as part of the same event.		
linkingAgentIdentifierType2			
linkingAgentIdentifierValue2			
linkingAgentRole2			
linkingAgentIdentifierXMLID3			
linkingAgentIdentifierTyp			

Name	Description	Source	PREMIS path
e3			
linkingAgentIdentifierValue3			
linkingAgentRole3			

Name	Description	Source	PREMIS path
linkingRightsStatementIdentifier	An identifier for a rights statement associated with the object.	User	PREMIS
linkingRightsStatementIdentifierType	A designation of the domain within which the linkingRightsStatementIdentifier is unique.	User can populate this section pre-ingest (in submission application)	OBJECT: linkingRightsStatementIdentifier. linkingRightsStatementIdentifierType
linkingRightsStatementIdentifierValue	The value of the linkingRightsStatementIdentifier.		OBJECT: linkingRightsStatementIdentifier. linkingRightsStatementIdentifierValue
accessRightsPolicy	Access rights Policy	System/User	Non-PREMIS
policyId	The ID of the policy in its shared MD format (in the hdemetadata table in the staging DB and in the permanent repository). Note that this ID can be a string (for example, OPEN_ACCESS).	The system assigns the IE with the AR policy of the material flow. The AR policy can be re-assigned by the user in the web editor (3A, editor) *** Rosetta Mandatory**	
policyParameters	Modifying parameters of the access rights policy record (future use)		
policyDescription	Textual brief description of the access rights policy record		
retentionPeriodPolicy	Retention Period	System/User	Non-PREMIS
policyId	The ID of the policy in its shared MD format (in the hdemetadata table in the staging DB and in the permanent		

Name	Description	Source	PREMIS path
	repository). Note that this ID can be a string (for example, FIVE YEARS).		
policyDescription	Textual brief description of the retention rights policy record		
grantedRightsStatement	The deposit boilerplate statement	User	Non-PREMIS
grantedRightsStatementIdentifier	The ID of the boilerplate that is related to the material flow through which this IE was inserted. This ID is generated in Rosetta when creating a new copyrights statement.	User can populate this section pre-ingest (in submission application)	
grantedRightsStatementValue	Actual content of the statement		
metadata (Deprecated)	The Metadata record	System	Non-Premis
This table is deprecated and not in use.			
MID	The ID of the metadata record as it stored in the PID/MID table. *** Rosetta Mandatory**	This section is populated in the METS file upon the move to permanent.	
UUID	The unique ORACLE sequence ID of the metadata record in the HDEMETADATA table. *** Rosetta Mandatory**		
creationDate	Creation date of the metadata record in Rosetta. *** Rosetta Mandatory**		
createdBy	User who created the metadata record. *** Rosetta Mandatory**		
modificationDate	Modification date. *** Rosetta Mandatory**		
modifiedBy	User who modified the record. *** Rosetta Mandatory**		
metadataType	Metadata type (DNX, descriptive DC,		

Name	Description	Source	PREMIS path
	access rights). *** Rosetta Mandatory**		
description	This field is currently not in use.		
externalSystem	The external system from which this metadata record was brought.		
externalRecordId	The ID of the MD in an external system.		
collection	Collection information	System	Non-Premis
collectionID	The unique ID of the collection. *** Rosetta Mandatory**		
collectionName	The name of the Collection *** Rosetta Mandatory**		
externalSystem	Future use - The external system from which this Collection record was brought.		
externalRecordId	Future use - The ID of the Collection in an external system.		
parentCollectionId	The Collection ID of the parent collection, in case it exists.		

DNX Controlled Lists

The following table lists the DNX fields which can hold only values from a predefined list. The list of available values is defined using Rosetta Code Table mechanism.

Field Name	Code Table Name	Editable[2]
status	IEStatus	Yes
IEEntityType	IEEntityType	Yes

[2] The code table can be edited by a staff user

Field Name	Code Table Name	Editable[2]
preservationType	PreservationType	Some values are read=only but new ones can be added
PreservationLevelType	PreservationLevelType	Yes
usageType	UsageType	No
representationEntityType	representationEntityType	Yes
RepresentationCode	RepresentationCode	Yes
FileEntityType	FileEntityType	Yes
fileLocationType	fileLocationType	No
fileExtension	com.exlibris.preservation.format.FormatsExtensionsPopulator	No
fileMIMEType	com.exlibris.preservation.format.FormatsMimeTypesPopulator	No
objectType	objectType	No
system	ExternalSystems	Yes
internalIdentifierType	internalIdentifierType	No
objectIdentifierType	objectIdentifierType	Yes
fixityType	fixityType	Some values are read only but new ones can be added
type	vsOutcomeType	No
result	vsResult	No
vsEvaluation	vsEvaluation	No
creatingApplicationExtension	SubmissionFormatFileExtensions	No
inhibitorType	inhibitorType	No

Field Name	Code Table Name	Editable[2]
dependencyIdentifierType1	dependencyIdentifierType	No
dependencyIdentifierType2	dependencyIdentifierType	No
dependencyIdentifierType3	dependencyIdentifierType	No
relationshipType	relationshipType	No
relationshipSubType	relationshipSubType	No
relatedObjectIdentifierType1	relatedObjectIdentifierType	No
relatedObjectIdentifierType2	relatedObjectIdentifierType	No
relatedObjectIdentifierType3	relatedObjectIdentifierType	No
linkingIEIdentifierType	linkingIEIdentifierType	No
defaultLanguage	UserLanguage	Yes
eventType	eventTypeCategory	No
linkingAgentIdentifierType1	linkingAgentIdentifierType	No
linkingAgentIdentifierType2	linkingAgentIdentifierType	No
linkingAgentIdentifierType3	linkingAgentIdentifierType	No
linkingRightsStatementIdentifierType	linkingRightsStatementIdentifierType	No
linkingAgentRole	linkingAgentRole	No