

Fedora METS

Table of Contents

1. [Introduction](#)
2. [Object Encoding Rules](#)

Introduction

As usual an example is worth a thousand words. So, please refer to the sample Fedora object that is encoded for ingest in METS 1.1: [mets-ingest-example.xml](#).

Fedora supports ingest of objects in a Fedora-specific extension of Metadata Encoding and Transmission Standard (METS). More information on METS can be found at <http://www.loc.gov/standards/mets/>. For specific information about the Fedora extension to METS, please see the [METS 1.1 schema](#).

Since METS was designed to be very generic and support a variety of uses, the rules of the METS Schema are very general-purpose. Fedora objects must conform to other rules that are beyond the scope of what is expressed in the METS schema. Therefore, the Fedora Object XML submissions will also be validated against a set of Fedora-specific rules that are expressed using the Schematron language. Internally, the repository will use Schematron to enforce these rules on incoming XML submission packages. The Schematron rules are expressed in XML and can be found in the Fedora server distribution at: %FEDORA_HOME%\server\schematron\metsExtRules1-1.xml.

For convenience and ease of understanding we have enumerated the Fedora rules in plain English below.

Object Encoding Rules

Encoding by hand requires a pretty good understanding of METS, although it can be done by following the patterns in the demo objects that come with the Fedora distribution. Demo objects are located at: %FEDORA_HOME%\client\demo.

General attributes

- On METS root element, the OBJID attribute will represent the Fedora object PID. Normally, this should be left empty so that the Fedora repository can generate a new PID. However, you can assign test/demo PIDs by inserting a value in OBJID that begins with "demo:" or "test:" for example, "demo:100"
- On METS root element, the value of the EXT_VERSION attribute must be "1.1".
- On METS root element, the value of LABEL serves as the official description of the object. If there is no Dublin Core record present in the object, the Fedora repository will use this label to populate the title element of a baseline Dublin Core record for the object.
- On METS root element, the PROFILE element can be used by institutions to classify different types of Fedora data objects.
- On the METS:metsHdr element the CREATEDATE attribute should be omitted since the Fedora repository will assign this at ingest time. Fedora dates are in the ISO 8601 format in milliseconds and with UTC time as follows: yyyy-MM-ddTHH:mm:ss.SSSZ. The same thing goes for LASTMODDATE.
- On the METS:metsHdr element the RECORDSTATUS should be set to "A", "I", or "D" to indicate that the object is in the "Active", "Inactive", or "Deleted" state. The usual state is "A" (Active). These states may be used by policy enforcement, for example, to prevent access to items in non-Active states. They may also be used by external tools, for example, to indicate whether an object's data should be indexed or not.

Datastreams

- To create a proper section for Datastreams in the METS file, the METS:fileSec must have a single child METS:fileGrp element whose ID attribute has the value "DATASTREAMS"
- Datastreams that are encoded in the METS:fileGrp must follow the following pattern to establish proper version groups and datastream IDs. Each datastream has its own METS:file element whose ID attribute is the official datastream ID. The recommended convention is ID="DSn" where n is a number (for example ID="DS1" or ID="DS2").
- Within a METS:fileGrp, there can be one or more METS:file elements to represent different versions of a datastream. As of Fedora 1.2, versioning of data objects is supported. The METS:file element for the datastream must have an ID attribute that represent the version number relative to the datastream ID set in the METS:fileGrp. The recommended convention is ID="DSn.v" where n is the number of the datastream and v is the version number (for example ID=DS1.0 or ID=DS1.1).
- The METS:file element for a datastream must have a MIMETYPE.
- The METS:file element for a datastream must have an OWNERID attribute. In Fedora, the OWNERID attribute is used to encode the Datastream Control Group. The following are valid values:
 - "M" - Managed Content. This tells the repository to store the datastream's content byte stream inside the repository. When the METS:file contains "M" on the OWNERID, the repository will resolve the URL associated with the METS:file element and pull the content into the repository for permanent storage. Fedora will establish its own local identifier for retrieving the content, and disregard the original URL that came in on the METS submission package.
 - "E" - External Referenced Content. This tells the repository to store the URL for the datastream content, not the content byte stream itself. For this type of datastream, Fedora does not actually manage or have custodianship of the content, but it manages the link to the content and some basic metadata about it.
 - "R" - Redirected Content. Like "E" this tells the repository to store the URL for the datastream content, not the content byte stream itself. More importantly, it tells the repository not to mediate or proxy this content at runtime. Instead, the repository will redirect to the URL at run time. This is desirable when a datastream points to a streaming media source, or to a complex web page where some components are lost during proxying.

Inline XML Datastreams

- Datastreams can also be encoded in the METS:dmdSecFedora and METS:amdSec. These are considered "inline XML datastreams" in Fedora. The METS:dmdSecFedora and METS:amdSec elements act as datastream version group containers just like the METS:fileGrp acts for regular datastreams. Within these elements, the METS "metadata section" elements (i.e., METS:techMD, METS:rightsMD, etc.) are used for the specific version instances of the inline metadata datastreams, just like the METS:file acts for regular datastreams. The datastream IDs work the same way, where the ID attribute on the container element acts as the datastream ID, and the ID on the metadata section element acts as the datastream version ID.
- Do not use the schemaLocation attribute in the root element of inline XML datastreams (within METS:mdWrap element).

Dublin Core Record Datastream

- A Dublin Core (DC) record is optional in the Fedora object submission package. If one is not provided the repository will automatically create a minimal DC record in the object by using the LABEL (on METS root) as the DC title element. It will also use the object PID as the DC identifier element.
- If a DC record is provided in the METS submission package it should be encoded within a METS:dmdSecFedora. The dmdSecFedora element will act as the datastream version group container. It MUST have an ID attribute whose value is "DC" to be recognized by Fedora!
- Within the METS:dmdSecFedora, there must be one METS:descMD element. This element is part of the Fedora extension of METS 1.1 and is used to encode a specific version of the DC datastream within the version group container. The ID attribute on the METS:descMD element MUST have the value "DC1.0" to be recognized by Fedora.
- The actual DC metadata should be encoded using the Open Archives Initiative (OAI) Dublin Core schema.