

Common Specification (FGS-PUBL) for deposit of single electronic publications to the National Library of Sweden - Kungl. biblioteket (KB)

Reference to this document:
http://www.kb.se/namespace/digark/deliveryspecification/deposit/fgs-publ/v1/FGS-PUBL_1_1_eng.pdf

1. Introduction

This document is a specification for deposit of single electronic publications (resources) to the National Library of Sweden - Kungliga biblioteket (KB). The specification is based upon *Förvaltningsgemensam specifikation (FGS) för paketstruktur för e-arkiv*, produced as a result of the project eARD¹ by the Swedish National Archive – Riksarkivet.

This specification (FGS-PUBL) is particularly aimed to be used for delivery of "finalized" electronic documents published by an authority or government agency, that is for document types such as e-books, reports, e-journal issues, digital images or audio files, in accordance with the Swedish law of legal deposit of electronic material. The specification (FGS-PUBL) may also be used by any publishing organization and for publications not covered by this law, but delivered to the National Library (KB) by special agreement.

Change history

1.1. Version 1.1

This version has added the optional element `<transformFile TRANSFORMTYPE="decryption" TRANSFORMKEY="[decryption key]" TRANSFORM-ALGORITHM="[decryption algorithm]">` with its decryption key and/or decryption algorithms as attributes.

¹ FGS in the project eARD, see: http://riksarkivet.se/Media/pdf-filer/Projekt/eARD_informationstext_eng.pdf

2. Package structure and delivery

- A delivery can hold one or several delivery *packages*. Each package can contain one or several *objects* (data files or data streams) pertaining to a publication, i.e. to a *resource*, e.g. a book, a journal issue, an article. Each resource in a delivery must, consequently, have its own package; i.e. each package must contain one and only one resource.
- According to this specification *metadata* about the resource must be included in each package. Metadata must be formatted in accordance with METS² in a separate xml-file with the file name '*sip.xml*'.
- Packages are transferred to the National Library (KB) via FTP (IP-address, username and password is given to the document provider by KB at sign-up) or, in some events, via other file transfer methods agreed upon. A delivery is transferred as a *tar-file*, the file name of which represents the external delivery-ID (i.e. the provider's reference).
- Receipts from all packages and deliveries to the National Library (KB) are published via a REST-API, where they can be searched and queried by means of an external delivery-ID.

3. Limitations

For web-articles and other continuously updated documents published at the web-site of the provider RSS-feeds may in some cases be a better method and format for delivery. More information about RSS as a method of delivery is found at

http://www.kb.se/namespace/digark/deliveryspecification/deposit/rssfeeds/rssfeeds_eng.pdf.

² METS (Metadata Encoding and Transmission Standard),
<http://www.loc.gov/standards/mets/>

4. Metadata

This specification contains a selection of elements from the Common Specification (FGS) for digital archives. The standard used is METS³. For the bibliographic description in the section *mets:dmdSec* some other metadata standard must be used as a complement. The National Library (KB) has already a specification for MODS⁴ to be used together with FGS-PUBL/METS and we are working on another specification for Dublin Core⁵, possibly in a combination with RDF.

4.1. Namespaces and schemas

References to valid namespaces and xml schemas for each metadata standard being used in the metadata file *sip.xml*, must be included in the root element *mets:mets*. The namespaces and schemas to be referenced are declared by the National Library (KB).

4.2. Metadata for Submission Information Packages (SIPs)

Metadata for a delivery package, i.e. a submission information package (SIP), are recorded as specified data elements in the metadata file named "*sip.xml*". A metadata file "*sip.xml*" is to be included in each delivery package. Metadata in this file are common to all types of resources. Eleven metadata elements are mandatory:

1. Identifier
2. Package type
3. Date-Time stamp
4. Delivery type
5. Delivery specification
6. Submission agreement

³ METS (Metadata Encoding and Transmission Standard),
<http://www.loc.gov/standards/mets/>

⁴ MODS (Metadata Object Description Schema), <http://www.loc.gov/standards/mods/>

⁵ DC (Dublin Core, DCMI Metadata Terms), <http://dublincore.org/documents/dcmi-terms/>

7. Name of archivist (publisher)
8. Identity code for archivist (publisher)
9. Name of system software
10. Name of delivering agent (supplier)
11. Identity code for delivering agent (supplier)

The tables that follow show data elements that **must** (Card. **1..1** or **1..N**) or should (Card. 0..1) be used to describe a SIP.

Cardinalities "**1..1**" and "**1..N**" signify that a data element is **mandatory**. Cardinality 0..1 signifies that a data element is optional, but recommended.



Element	Definition	Explanation and rules	Card.	METS
Identifier	Identifier of a package	A code that uniquely identifies the SIP. A UUID or GUID can be used as globally unique identifiers. Example: "UUID:550e8400-e29b-41d4-a716-446655440004"	1..1	<mets OBJID>
Description	Description of a package	Short description of the contents of the package. The attribute is optional, but we recommend using the same content here as for <title> in the bibliographic description in <dmdSec> below.	0..1	<mets LABEL>
Package type	AIP/SIP/DIP	Indicates the place in the OAIS model where the package belongs. Delivery packages always have the value "SIP".	1..1	<mets TYPE>
Date-Time stamp	Date and time for the creation of the package	Date-Time stamp for the package according to XML-standard.. Specifies when the SIP and the metadata file "sip.xml" were created. Written in W3CDTF time format: YYYY-MM-DDThh:mm:ss.sTZD. Example: "2012-04-26T12:45:00+01:00"	1..1	<metsHdr CREATEDATE>
Status	Package status	Optional status of a SIP. Value options for this attribute are defined through a vocabulary in the METS profile." REPLACEMENT" or "SUPPLEMENT" indicate packages that are replacing or supplementing earlier delivered packages. Other values used for this attribute are "NEW", "VERSION" and "TEST".	0..1	<metsHdr RECORDSTATUS>
Delivery type	Delivery type of a package	Indicates delivery type of a SIP. A SIP can have one and only one delivery type. Value options for this data element are defined through a vocabulary in the METS profile. Valid values for this FGS-PUBL are: "DEPOSIT" (for resources delivered according to the Swedish law on legal deposit of electronic document); "AGREEMENT" (for other electronic resources delivered to the National Library by special agreement).	1..1	<altrecordID TYPE="DELIVERYTYPE">



Element	Definition	Explanation and rules	Card.	METS
Delivery specification	Delivery specification for the delivery type.	A URI reference to the specification document for the designated delivery type. (Most likely the URI for this very document). The National Library (KB) supplies value URIs to be used.	1..1	<altrecordID TYPE=" DELIVERY- SPECIFICATION">
Submission agreement	Submission agreement according to which a SIP is delivered.	A URI reference to the submission agreement. The National Library (KB) supplies value URI to be used.	1..1	<altrecordID TYPE="SUBMISSION- AGREEMENT">
Archivist Name	Name of archivist.	In this specification (FGS-PUBL) the archivist is the institution responsible for making the publication (resource) accessible on the internet, i.e. most often the publisher. Example: <i>Förslagsmyndigheten</i>	1..1	<agent ROLE="ARCHIVIST" TYPE= "ORGANIZATION"> <name>
Archivist Identifier	Unique identification code for the archivist	Identification with URI + unique corporate identity no (Swedish "organisationsnummer"). Prefix "URI:" Example: "URI:http://id.kb.se/organisations/SE2021001710"	1..1	<agent ROLE=" ARCHIVIST" TYPE=" ORGANIZATION"> <note>
System Name	Name of system from which files of a delivery were exported	States which system the files in a delivery were exported from. If there is no given name of the system, a short description is provided instead.	1..1	<agent ROLE=" ARCHIVIST" TYPE="OTHER" OTHERTYPE= "SOFTWARE"> <name>
System Version	Version of system from which files of a delivery were exported.	States version of system that the files in a delivery were exported from. Example: "Version 2.76"	0..1	<agent ROLE=" ARCHIVIST" TYPE="OTHER" OTHERTYPE= "SOFTWARE"> <note>



Element	Definition	Explanation and rules	Card.	METS
Delivering agent Name	Name of organization delivering the SIP to the National Library (KB).	Name of organization delivering the SIP to the National Library, often identical to name of archivist.	1..1	<agent ROLE="CREATOR" TYPE=" ORGANIZATION"> <name>
Delivering agent Identifier	Unique identification code for the delivering organization	Identification with URI + unique corporate identity no (Swedish "organisationsnummer"). Prefix "URI:". Example: "URI:http://id.kb.se/organisations/SE2021001710"	1..1	<agent ROLE=" CREATOR" TYPE=" ORGANIZATION"> <note>

4.3. Bibliographical metadata for the resource (Descriptive Metadata Section)

Element	Definition	Explanation and rules	Card.	METS
Bibliographical metadata	Descriptive metadata about the resource.	At least one embedded bibliographic description of the delivered resource is mandatory. Examples of recommended metadata standards to be used for this section are DC (Dublin Core) or MODS.	1..N	<dmdSec> <mdWrap MDTYPE="[name of metadata standard]" <xmlData>

4.4. Reference to metadata outside of *sip.xml*

Element	Definition	Explanation and rules	Card.	METS
Referenced metadata	-	It is possible to refer to metadata in other formats outside METS. For example, a SIP may hold a separate metadata file, referenced in the <i>sip.xml</i> . This is permissible according to <i>Paket-FGS</i> ("Package-FGS") and might be used in certain cases after agreement with the National Library (KB).	0..1	<mdRef: MDTYPE="[name of metadata standard]" Xlink:href>

4.5. Metadata for files referenced in the file sip.xml

Each physical object (datafile) included in the delivery package must have one and only one unique reference in the metadata file sip.xml. Each file thus referenced is further described by metadata elements and attributes in the METS format.

6 elements are mandatory:

1. File identifier
2. File name
3. Date-Time stamp
4. MIME-type
5. File format and version
6. File size

Element	Definition	Explanation and rules	Card.	METS
File identifier	Identifier of file	A code that uniquely identifies the file within the METS-file (<i>sip.xml</i>). This code does not need to have any other function. A valid identifier code consists of a prefix "ID" immediately followed by a UUID or GUID. Example: "ID550e8400-e29b-41d4-a716-4466554400bg"	1..1	<file ID>
File name	Name of file	File name and search path and file suffix. File name must be unique within the SIP. File name must always be preceded by a prefix "file." Example: "file.nameoffile.pdf"	1..1	<file <flocat LOCTYPE="URL" xlink:type="simple" xlink:href="">
Date-Time stamp	Date-Time stamp of file	Date and time of file. Date-Time format according to XML standard. Reference is to time for latest update of file before upload to the delivered SIP. Example: "2012-04-20T13:30:00+01:00"	1..1	<file CREATED>



Element	Definition	Explanation and rules	Card.	METS
MIME-type	Simple designation of file format	A way of describing file type. For example "text/plain" for a simple text file. Example: <i>"text/xml"</i>	1..1	<file MIMETYPE>
File format and version	Format and format version of file	In addition to MIME-type there is a need for file format name, version of format and, when applicable, format key in a format registry, such as PRONOM. Format name is mandatory. Format registry values for this data element are defined in a value list in the METS-profile. Example: <i>"Extensible Markup Language;1.0;PRONOM:fmt/101"</i>	1..1	<file USE="[Format name]"; "[Format version]";"[Format registry]";"[Format key]">
File size	Size of a file in bytes	File size in bytes is mandatory for each file contained in the SIP. Example: <i>"8765324"</i>	1..1	<file SIZE>
Checksum	Computed checksum value	A checksum value that is unique to each file. Example: <i>"574b69cf71ceb5534c8a2547f5547dcc"</i>	0..1	<file CHECKSUM>
Checksum type	Algorithm used for computation of checksum	Valid checksum algorithms are taken from a value list in the METS-profile. Example: <i>'MD5'</i>	0..1	<file CHECKSUMTYPE>
Decryption key	Decryption key for an encrypted file	A key to be used with the transform algorithm for accessing the file's contents. Example: <i>"574b69cf71ceb5534c8a2547f5547d"</i>	0..1	<file> <transformFile TRANSFORMTYPE= "decryption" TRANSFORMKEY="[Decryption key]">
Decryption algorithm	Decryption algorithm for an encrypted file	Specifies the decompression or decryption routine used to access the contents of the file. Example: <i>"DES"</i>	0..1	<file> <transformFile TRANSFORM- TYPE="decryption" TRANSFORM- ALGORITHM="[Decryption algorithm]">

4.6. Use of a structMap

A METS-file must always include the mandatory element <structMap>. It outlines a hierarchical structure of the relationships between the files that are part of the SIP, the information package.

Element	Definition	Explanation and rules	Card.	METS
Physical structural map	An inventory that defines the physical relationships between the datafiles that constitute a SIP.		1..1	<structMap TYPE="physical"> <div TYPE="files">
Superior division	Superior level in a hierarchical inventory of the datafiles that constitute a SIP.	At least one <div>-element is mandatory. In a structMap the superior level <div> always has the attribute value TYPE="files".	1..1	<div TYPE="files">
Subordinate division	Subordinate level in a hierarchical inventory of the datafiles that constitute a SIP.	The inventory of files can have several levels of hierarchy, with different attribute values for TYPE. Valid TYPE-values at present are "publication", "coverpicture". The approved value list may be extended as needs arise.	0..N	<div TYPE="[value from value list]">
File reference	Reference to ID of file in the SIP.		1..N	<fptr FILEID="ID[value in <file ID=" ">]">

Two examples of the simplest form of a structMap for this FGS.

Example 1 (reference to 1 delivered file):

```
<mets:structMap TYPE="physical">  
  <mets:div TYPE="files">  
    <mets:fptr FILEID="ID1" />  
  </mets:div>  
</mets:structMap>
```

Example 2: (reference to 2 delivered files):

```
<mets:structMap TYPE="physical">  
  <mets:div TYPE="files">  
    <mets:fptr FILEID="ID1" />  
    <mets:fptr FILEID="ID2" />  
  </mets:div>  
</mets:structMap>
```

These two examples (above) would not provide more information than what is already inherent in the package structure. However, if a package (SIP) contains several files, with different content types or functions, and this is reflected in the structMap, the information provided by the structMap becomes much more valuable. By using the *div* element it is possible to create hierarchy levels named with the attribute *TYPE*, which may subsequently serve for purposes of searching and display in an end user interface (e.g. such as LIBRIS).

Example 3 (reference to a pdf-file, that constitutes the publication itself, and another reference to an accompanying cover picture):

```
<mets:structMap TYPE="physical">  
  <mets:div TYPE="files">  
    <mets:div TYPE="publication">  
      <mets:fptr FILEID="ID1" />  
    </mets:div>  
    <mets:div TYPE="coverpicture">  
      <mets:fptr FILEID="ID2" />  
    </mets:div>  
  </mets:div>  
</mets:structMap>
```