DIP Interface Specification

Document information

Workproducts contained	APP 030: Interface Specifications
Author:	H. Verhoeven
Version:	2.6
Date:	10-02-2006
Classification:	Unclassified

Revision history

Revision Number	Revision date	Author	Reviewer	Summary of changes
0.1	10-01-2002	Han van Straaten	Hans Verhoeven	Initial version
0.2.	01-02-2002	Han van Straaten	Hans Verhoeven	Added comment
0.3	10-07-2002	Han van Straaten		Update from build cycle
1.1	07-10-2003	F Taylor Parkins	Hans Verhoeven	Converted DIP Interface to specification
1.2	21-11-2003	F Taylor Parkins	Hans Verhoeven	Processed review comment
1.3	10-09-2004	Hans Verhoeven	Frank Taylor Parkins	Additions and clarifications as a result of review comments from Kopal (DDB, SUB) project members.
1.4	21-04-2005	Maarten de Wit	Hans Verhoeven & Tanja van Tuijn	Internal review comments processed.
2.0	01-06-2005	Maarten de Wit		Included support for DIAS-METS Universal Object Format and multi-organization
2.1	14-06-2005	Maarten de Wit	H. Verhoeven, F. Taylor Parkins & H. van Straaten	Internal review comments processed
2.2	23-06-2005	Maarten de Wit		Added DIP examples and DTD
2.3	28-06-2005	Maarten de Wit	H. Verhoeven	Internal review comments processed
2.4	11-07-2005	H. Verhoeven		Removed gzip as archive format as it is only a compression format (REQUPD-02).
2.5	19-08-2005	H. Verhoeven		Reworked comments from KOPAL review, resolved PTM0042A (Retrieve->Retriever) and updated DTDs and XML-examples.
2.6	10-02-2006	Han van Straaten		Added note for the combination of FullMetadata, HTML without DipFmt. Added note for the file size limit to direct access for the unpacked DIP

Table of contents

1 Introduction

1.1 Purpose

This document describes the Dissemination Information Package (DIP) interface of DIAS-Core.

This document is intended for customers of DIAS-core who want to setup an interface to retrieve assets or metadata from DIAS-Core as well as the development group of DIAS-core who are responsible for the interface development.

See the DIAS-Core Capabilities Specification document [ref 1] for more information about DIAS-Core.

1.2 Scope and Approach

The Digital Information Archiving System (DIAS) is based on the Open Archival Information System ([OAIS]) ISO Reference Model that has been defined by the Consultative Committee of Space Data Systems (CCSDS), see ref [2]. Interfacing with DIAS means exchanging digital assets and / or metadata by data packages. The Dissemination Information Package (DIP) is used to retrieve metadata and the digital asset out of the DIAS System. The DIP data packages are transferred over the DIP interfaces.

The requirements of the DIP Interface are labelled with unique numbered items called Interface Items to facilitate tracing and tracking of these items. The Interface Items conform to the following format IFI_DIPnnnX for the DIAS-KB specific requirements or UOF.xx.xxx for the DIAS-METS specific requirements.

1.3 Document Overview and References

Chapter 2 introduces the DIP data package.
Chapter 3 describes the DIP interface.
Chapter 4 details the DIP data package.

The following documents are referenced:

DIAS	**[1]** DIAS-Core Capabilities Specification. **[5]** DIAS SIP Interface definition.
General	**[2]** Reference Model for an Open Archival Information System (OAIS), Jan 2002, (http://www.ccsds.org/documents/650x0b1.pdf). **[3]** METS: The technical metadata in DIAS-METS format is based on version 1.4 of the Metadata Encoding and Transmission Standard of the Library of Congress (http://www.loc.gov/standards/mets). **[4]** LMER: The lmer version 1.2 standard is described in http://www.ddb.de/standards/lmer/index.htm

2 DIP data package

The DIP data package is a self describing package of data. Meaning the DIP data package is not just a file with some data. Instead it is a bundle of technical metadata and / or a digital asset.

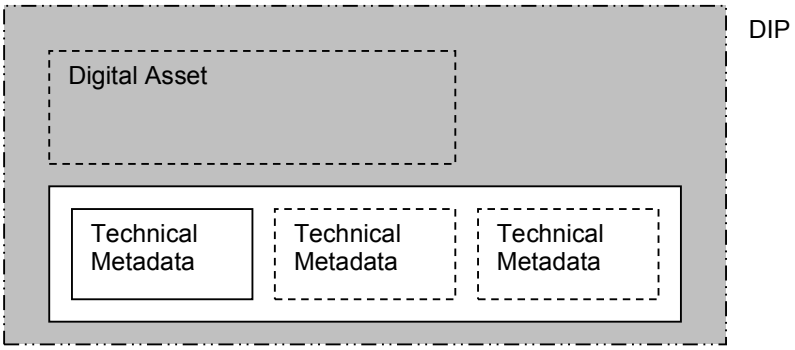

Figure 1: DIP data package

The format of the DIP data package provides a flexible mechanism for constructing a single file that is a container for a digital asset and/or technical metadata of one or even more assets.

2.1 The digital asset

The digital asset, or shortly called the asset, can be as simple as a single file or as complex as a collection of directory trees and files. This is the main content.

2.2 The technical metadata

The technical metadata consists of standard and custom specific elements describing the asset and the associated directories and files in the asset. See chapter 4, for more information.

The DIP data package can contain technical metadata of multiple, related assets or for one asset only.

2.3 Appearances

The DIP data package can appear unpacked in a File structure made available through HTTP in a location of the DIAS-Core system or as a single file in a popular archive format. More on this topic in detail in chapter 4, ‘DIP data package in detail’.

3 DIP Interface

This interface describes the procedure for the Consumer to retrieve metadata and / or assets from DIAS-Core within a DIP data package. The DIP Interface is showed in the following figure as IDIP (Interface-DIP):

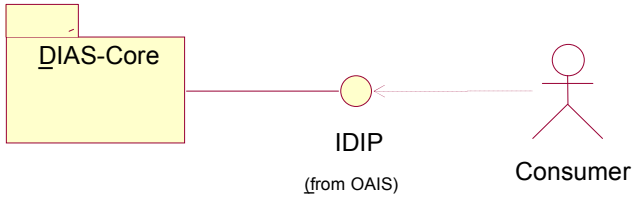

Figure 2: Component diagram DIP interface context

A typical request-response sequence looks as follows:

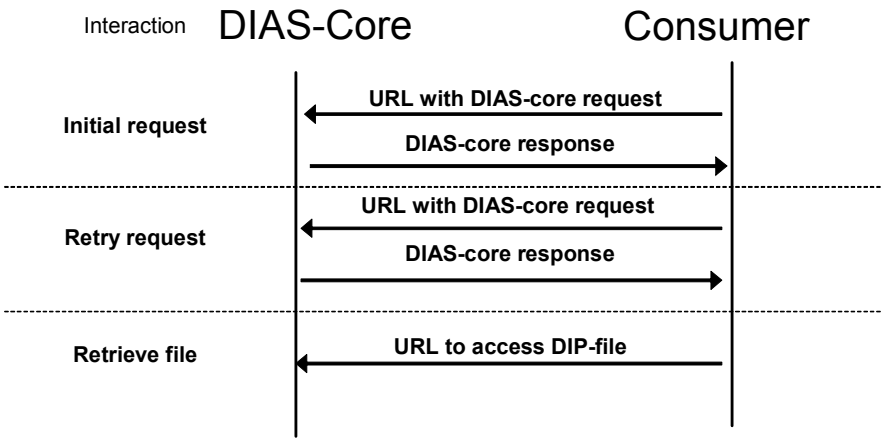

Figure 3: DIP Interface interaction

A URL is sent to DIAS-Core. The URL holds an identifier of the asset concerned, and the type of response wanted.
The response is given in the form of an output page in HTML or XML format.
In case of an html request and the DIAS-Core server is busy retrieving the requested asset from the archive a page will be returned which will automatically retry the request after some time. In case of an XML request a parameter indicating the expected wait time is returned.
If the asset has then been retrieved by DIAS-core a link to the DIP-file / DIP-archive on a HTTP accessible location in the DIAS-Core system will be returned otherwise another busy response may be given.
The consumer will have to use that link to retrieve the asset. The asset remains in that place for a configurable amount of time (typical 24h) but will be removed earlier in case the file system containing retrieved assets becomes too full.

The following interactions are supported via the DIP interface:

- DIAS-KB specific:
  - Request technical metadata of specific asset;
  - Request technical metadata related to a specific asset;
  - Request a specific asset.
- DIAS-METS specific:
  - Request technical metadata by external asset identifier;
  - Request technical metadata by internal asset identifier;
  - Request complete technical metadata by internal asset identifier;

- Request a specific asset.

UOF.MOS.F3 DIAS-Core will be able to handle multiple organizations in a single instance of the DIAS system. DIAS-Core provides a virtual repository for each organization so that assets from the various organizations are separated. Access controls are applied to ensure that users can only access assets and related information from their own organization.

3.1 DIAS-KB

3.1.1 Request technical metadata of a specific asset

3.1.1.1 Purpose

Request the technical metadata of one, and only one specific asset.

3.1.1.2 Request

Request is formed as URL.

IFI_DIP001A http://<diashost>/dias/Retriever?NBN=<*NBN*>&RespFmt=<*RespFmt*>&Request=Metadata

<diashost>	hostname of the dias system.
<NBN>	National Bibliographic Number identifying the asset.
<RespFmt>	Parameter indicating the format of the returned response (‘xml’ | ‘html’).

IFI_DIP001B The requests can be extended with the parameter ‘DipFmt’ to deliver the metadata in an archive of choice.
&DipFmt=<format>

<format>	archive format (‘zip’ or ‘tar’).

Note: The combination of FullMetadata, HTML and no DipFmt is not logical, hence this results in an error.

3.1.1.3 Response

Technical metadata of the asset in a DIP data package.

IFI_DIP002A Response on IFI_DIP001A is put into an html or xml page.

IFI_DIP002B The response on IFI_ DIP 001B is returned as a link to the DIP-archive into an html or xml page . The DIP-archive is placed on a HTTP accessible location in the DIAS-Core system.

3.1.2 Request technical metadata related to a specific asset

3.1.2.1 Purpose

Request metadata related to a specific asset (Converted assets from Long Term Preservation).

Remark: One OriginalEpublication can have multiple ConvertedEpublications when the digital object is transformed over time. There is a relation between an OriginalEpublication and its ConvertedEpublications. In addition, an OriginalEpublication can have a relation to the installed version that is archived as an InstalledEpublication. Via this request it is possible to retrieve the metadata of all derivates of the OriginalEpublication identified by it’s original NBN.

3.1.2.2 Request

Request is formed as URL.

IFI_DIP010A http://<diashost>/dias/Retriever?originalNBN=<*NBN*>&Respfmt=<*RespFmt*>&Request=Metadata

<diashost> hostname of the dias system.
<NBN> National Bibliographic Number identifying the original asset.
<RespFmt> Parameter indicating the format of the returned page ('xml' | 'html').

IFI_DIP010B The requests can be extended with the parameter 'DipFmt' to deliver the response in an archive of choice.
&DipFmt=<format>

<format> archive format ('zip' or 'tar').

Note: The combination of FullMetadata, HTML and no DipFmt is not logical, hence this results in an error.

3.1.2.3 Response

Technical metadata of the related assets bundled in a DIP data package.

IFI_DIP011A Response on IFI_DIP010A is put into an html or xml page.

IFI_DIP011B The response on IFI_ DIP 011B is returned as a link to the DIP-archive into an html or xml page. The DIP-archive is placed on a HTTP accessible location in the DIAS-Core system.

3.1.3 Request a specific asset

3.1.3.1 Purpose

Request a specific asset.

3.1.3.2 Request

Request is formed as URL.

IFI_DIP020A http://<diashost>/dias/Retriever?originalNBN=<*NBN*>&RespFmt=<*RespFmt*>&Request=AIP

<diashost> hostname of the Dias system.
<NBN> National Bibliographic Number identifying the original asset.
<RespFmt> Parameter indicating the format of the returned page ('xml' | 'html').

IFI_DIP020B The requests can be extended with the parameter 'DipFmt' to deliver the response in an archive of choice.
&DipFmt=<format>

<format> archive format ('zip' or 'tar').

Note: The combination of FullMetadata, HTML and no DipFmt is not logical, hence this results in an error.

3.1.3.3 Response

Technical metadata of the asset and the asset itself is returned in a DIP data package.

IFI_DIP021A In response of IFI_ DIP020A the technical metadata is put into a html or xml page and the asset is placed in an unpacked form (DIP-file) on a HTTP accessible location in the DIAS-Core system.

IFI_DIP021B The response on IFI_ DIP 020B is returned as a link to the DIP-archive into an html or xml page. The DIP-archive is placed on a HTTP accessible location in the DIAS-Core system.

3.2 DIAS-METS

3.2.1 Request technical metadata by external asset identifier

3.2.1.1 Purpose

Request the technical metadata by the external asset identifier (ImerObject.persistentidentifier in SIP interface).

Remark: One External Asset identifier can refer to multiple Internal Assets when the digital object is transformed over time. The metadata of all related Internal Assets can be returned as part of this request.

3.2.1.2 Request

Request is formed as URL.

UOF.ac.MQ2 http://<diashost>/dias/Retriever?ExternalAssetID=<*ExternalAssetID*>&RespFmt=<*RespFmt*> &Request=Metadata

<diashost>	Hostname of the dias system.
<ExternalAssetID>	External id identifying the asset collection.
<RespFmt>	Parameter indicating the format of the returned page (‘xml’ | ‘html’).

UOF.ac.MQ2a The request can be extended with the parameter ‘List’ which can have the values ‘true’ or ‘false’. If List is set to ‘true’ then metadata is returned for all Assets in the collection of Assets with the specified ExternalAssetID. If the value is set to ‘false’ (which is the default value) only the metadata of the latest Asset with the specified externalAssetID will be returned.
&List=<’true’ | ’false’>

3.2.1.3 Response

Return technical metadata of the asset.

UOF.ac.MQ3a The result set returned for successful metadata queries must contain a collection of items with the following data elements:

- *internalAssetID;*
- *externalAssetID;*
- *assetVersion;*
- *assetDescription;*
- *AIPIsDeleted;*
- *AIPDeletionTimestamp;*
- *assetCreationTimestamp;*
- *assetIngestTimestamp;*
- *oldObjectIdentifier (if available);*
- *oldVersion (if available).*

Response on UOF.ac.MQ2 is put into an html or xml page.

3.2.2 Request technical metadata by internal asset identifier

3.2.2.1 Purpose

Request metadata related by the internal asset identifier.

3.2.2.2 Request

Request is formed as URL.

UOF.ac.MQ1 http://<diashost>/dias/Retriever?InternalAssetID=<*InternalAssetID*>&*RespFmt*=<*RespFmt*> &Request=Metadata

<diashost>	Hostname of the dias system.
<InternalAssetID>	Internal id identifying the asset.
<RespFmt>	Parameter indicating the format of the returned page ('xml' | 'html').

3.2.2.3 Response

Technical metadata of the requested asset.

UOF.ac.MQ3b The result set returned for successful metadata queries must contain a collection of items with the following data elements:

- *internalAssetID;*
- *externalAssetID;*
- *assetVersion;*
- *assetDescription;*
- *AIPIsDeleted;*
- *AIPDeletionTimestamp;*
- *assetCreationTimestamp;*
- *assetIngestTimestamp;*
- *oldObjectIdentifier (if available);*
- *oldVersion (if available).*

Response on UOF.ac.MQ1 is put into a html or xml output page.

3.2.3 Request complete technical metadata by internal asset identifier

3.2.3.1 Purpose

Request complete metadata related by the internal asset identifier.

3.2.3.2 Request

Request is formed as URL.

UOF.ac.MQ4 http://<diashost>/dias/Retriever?InternalAssetID=<*InternalAssetID*> &Respfmt=<*RespFmt*>&Request=FullMetadata

<diashost>	Hostname of the dias system.
<InternalAssetID>	Internal id identifying the asset.
<RespFmt>	Parameter indicating the format of the returned page ('xml' | 'html').

UOF.ac.MQ6 The request can be extended with the parameter 'DipFmt' to deliver the metadata in an archive of choice.
&DipFmt=<format>

<format> archive format ('zip', 'tar' or 'tar.gz').

Note: The combination of FullMetadata, HTML and no DipFmt is not logical, hence this results in an error.

3.2.3.3 Response

Technical metadata of the asset returned in a DIP data package.

UOF.ac.MQ7 The response on UOF.ac.MQ4 is returned as a link to the file containing the full technical metadata. The technical metadata is placed in an unpacked form on a HTTP accessible location in the DIAS-Core system.

UOF.ac.MQ8 The response on UOF.ac.MQ6 is returned as a link to the DIP-archive. The DIP-archive containing the full technical metadata is placed on a HTTP accessible location in the DIAS-Core system.

UOF.ac.MQ9 The response will contain DIAS-METS formatted technical metadata with all information that was present in the input technical metadata with the exception of data elements that will be actualised (see UOF.ac.AR2, AR3 and AR4).

3.2.4 Request complete technical metadata by external asset identifier

3.2.4.1 Purpose

Request complete metadata related by the external asset identifier. The complete metadata of the latest Asset with the specified externalAssetID will be returned.

3.2.4.2 Request

Request is formed as URL.

UOF.ac.MQ4A http://<diashost>/dias/Retriever?ExternalAssetID=<*ExternalAssetID*>
&Respfmt=<*RespFmt*>&Request=FullMetadata

<diashost> Hostname of the dias system.
<ExternalAssetID> External id identifying the asset.
<RespFmt> Parameter indicating the format of the returned page ('xml' | 'html').

UOF.ac.MQ6A The request can be extended with the parameter 'DipFmt' to deliver the metadata in an archive of choice.
&DipFmt=<format>

<format> archive format ('zip', 'tar' or 'tar.gz').

Note: The combination of FullMetadata, HTML and no DipFmt is not logical, hence this results in an error.

3.2.4.3 Response

Technical metadata of the asset returned in a DIP data package.

UOF.ac.MQ7A The response on UOF.ac.MQ4A is returned as a link to the file containing the full technical metadata. The technical metadata is placed in an unpacked form on a HTTP accessible location in the DIAS-Core system.

UOF.ac.MQ8A The response on UOF.ac.MQ6A is returned as a link to the DIP-archive. The DIP-archive containing the full technical metadata is placed on a HTTP accessible location in the DIAS-Core system.

UOF.ac.MQ9A The response will contain DIAS-METS formatted technical metadata with all information that was present in the input technical metadata with the exception of data elements that will be actualised (see UOF.ac.AR2, AR3 and AR4).

3.2.5 Request a specific asset

3.2.5.1 Purpose

Request a specific asset.

3.2.5.2 Request

Request is formed as URL.

UOF.ac.AR1 http://<diashost>/dias/Retriever?InternalAssetID=<*InternalAssetID*>&*RespFmt*=<*RespFmt*> &Request=Asset

<diashost>	Hostname of the dias system.
<InternalAssetID>	Internal id identifying the asset.
<RespFmt>	Parameter indicating the format of the returned page ('xml' | 'html').

UOF.ac.AR6 The request can be extended with the parameter 'DipFmt' to deliver the metadata in an archive of choice.
&DipFmt=<format>

<format> archive format ('zip', 'tar' or 'tar.gz').

Note: The combination of FullMetadata, HTML and no DipFmt is not logical, hence this results in an error.

3.2.5.3 Response

Technical metadata of the asset and the asset itself is returned in a DIP data package.

UOF.ac.AR7 The response on UOF.ac.AR1 is returned as a link to the DIP-file containing the asset. The technical metadata and the asset is placed in an unpacked form on a HTTP accessible location in the DIAS-Core system.

UOF.ac.AR8 The response on UOF.ac.AR6 is returned as a link to a DIP-archive. The DIP-archive containing the asset is placed on a HTTP accessible location in the DIAS-Core system.

UOF.ac.AR1 The DIP data package for an Asset with *assetMDTypeID* of DIAS-METS that is returned for an Asset Retrieve Request will contain DIAS-METS formatted technical metadata with all information that was present in the input technical metadata with the exception of data elements that have to be actualised (see UOF.ac.AR2, AR3 and AR4).

UOF.ac.AR2 The following data elements will be actualised with current values from Data Management in the DIAS-METS formatted technical metadata:

- mets.OBJID must be set with the *internalAssetID*;
- mets.amdSec.techMD.mdWrap.xmlData.lmerObject.objectIdentifier must be set with the *internalAssetID*;
- mets.metsHdr.CREATEDATE must be set with the current *system-date/time*;
- mets.metsHdr.agent.ROLE must be set to "ARCHIVIST";
- mets.metsHdr.agent.TYPE must be set to "ORGANISATION";

- mets.metsHdr.agent.name must be set to the requesting organisation.

UOF.ac.AR3 DIAS-Core will update the ImerObject.status-element to “deleted” in the technical metadata in DIAS-METS format that is included in the DIP data package that is returned for a successful Asset Retrieve Request in case the AIP of the requested Asset has been deleted.

UOF.ac.AR4 DIAS-Core will remove all mets.fileSec.fileGrp.file-elements in the mets.fileSec.fileGrp-element identified by “ASSET” in the technical metadata in DIAS-METS format that is included in the DIP data package that is returned for a successful Asset Retrieve Request in case the AIP of the requested Asset has been deleted.

UOF.ac.AR5 The physical files in the DIP data package will be present in the same relative location from the Asset root as they were in the input SIP data package.

4 DIP data package in detail

As stated before, a DIP data package can reside in two forms; unpacked DIP-File or DIP-archive.
The following requirements are applicable for either format:

UOF.dip.F5 The DIP Data Packages must contain at root level an XML-document that contains metadata. This metadata is delivered in the same format as it had during Ingest:

- DIAS-KB (SIP_toc.xml).
  This is the DIAS V1.x supported technical metadata;
- DIAS-METS (mets.xml).
  This is the DIAS V2.0 supported technical metadata in METS V1.4 format.

UOF.dip.F6 Each DIP Data Package must contain a maximum of one Asset.

4.1 DIAS-KB

4.1.1 Unpacked DIP-file

The unpacked DIP-file is the default format for the response on a request for a specific asset (IFI_DIP020A). The Archival Information Package (AIP) with the asset is retrieved from the archive and the asset is extracted and stored in unpacked format and made available on a HTTP accessible location in the DIAS-Core system.

The unpacked DIP-file is placed in a unique directory for the request. In this directory there is an ‘unpacked’ subdirectory where the content of the asset is placed. The content of the asset is the same as it was ingested. This means the technical metadata file and additional files are included as well.

Note: Direct access of a file >2GB size results in an error page with Apache HTTP Server 1.3.26.

4.1.2 DIP-archive

IFI_DIP100 The DIAS-KB DIP data Packages must be supported in the following formats:

DIP Data Package Format	Description	File Extension
ZIP	ZIP format	.zip
GNU-TAR	GNU TAR format	.tar

Table 1: DIAS-KB DIP Data Package Formats

UOF.dip.F2a Each ZIP format used for DIP Data Packages must be compliant with PKZIP version 2.50 or higher but lower than 5.0.

UOF.dip.F3a The ZIP format used for DIP Data Packages must support the use of the standard ZIP-compression features.

UOF.dip.F4a Each DIP Data Package must have a file name that is based upon the NBN and must have a file extension that conforms to the package format.

UOF.dip.F7a The TAR format used for DIP Data Package must be compliant with GNU-TAR version 1.15 and use the ‘gnu’ format.

The DIP-archive is the response on a request with the DipFmt parameter. A DIP-archive is made available on a HTTP accessible location in the DIAS-Core system. The DIP-archive is placed in a unique directory for the request. In this directory there are two subdirectories: the ‘packed’ and the ‘unpacked’ directory. In the packed directory the DIP-archive is named like;

- Dip-<*ResponseType*>-<*NBN*>.archive extension

Where ResponseType is the type of response requested, (asset or metadata) and NBN is the National Bibliographic Number.

The Archival Information Package (AIP) with the asset is retrieved from the archive and the asset is extracted and stored in unpacked format in the unpacked directory. Then the asset is packet into the requested archive format and stored in the packed directory. The DIP-archive is made available on a HTTP accessible location in the DIAS-Core system. Dependent on request the DIP-archive contains metadata and/or the asset. When the asset is included the content of the asset is the same as it was ingested. This means the additional files and a descriptive file are included as well.

IFI_DIP101 The special files
There is a method for including categories of special files in which directory trees can be stored that can contain additional metadata files (like scanned images of CD-booklet) to describe the digital asset. The special files must be of supported DIAS file types in case these files have to be subject of long term preservation. Unknown or unsupported files can be stored. However, these files may not open correctly and are not subject to preservation processing. The special files method consists of applying an additional directory structure for the special files.

IFI_DIP102 Some attributes of the returned metadata will have been added by the DIAS Loader during Ingest. These are collName (Collection Name), which is defined when the libraryFunctionName rule is processed and indicates the actual storage medium, size which is the total size of the AIP-object, checksum which is the checksum of the AIP-object based on the used checkSumType and fileCount which indicates the number of files (directories are not counted) in the AIP. The metadata attribute EstimatedDownloadTime is calculated by the DIAS Retriever based on the size of the AIP, the number of files in the AIP and the actual storage medium of the AIP in combination with deterministic rules about retrieve processing.

4.1.3 Response dtd

IFI_DIP102 The request response in case the format is xml must conform to the following dtd:

```
<!ELEMENT DIASKBRESPONSE (DESCRIPTION?, REQUEST, DIP) >
<!ATTLIST DIASKBRESPONSE
 type (error | dip | busy)       #IMPLIED
 version (1.0)                   #IMPLIED
 >

<!ELEMENT REQUEST                (NBN?)>

<!ELEMENT DIP                    (METADATABLOCK?, ASSETBLOCK?) >
<!ELEMENT METADATABLOCK          (METADATA+) >

 <!ELEMENT METADATA              (NBN, TYPE?, SOURCEDESCRIPTION?, SUPPLIER?,
                                  SOURCETYPE?, ORIGINALNBN?, LIBRARYFUNCTION?,
                                  CREATIONDATE?, COLLNAME?, CHECKSUM?, CHECKSUMTYPE?,
                                  FILECOUNT?, REFPLATFORMNBN?, AIPSTATUS?,
                                  STARTERFILENAME?, LOCATION?, SIZE?,
                                  DIPRETRIEVEWAITTIME?)>
<!ELEMENT ASSETBLOCK             (ASSET?) >
<!ELEMENT ASSET                  (NBN, LOCATION?, SPECIALFILE*)>
<!ATTLIST ASSET type            CDATA #REQUIRED >
<!ELEMENT SPECIALFILE            (CATEGORY, LOCATION)>
<!ELEMENT LOCATION               (AUTHORITY, PATH)>

<!ELEMENT DESCRIPTION            (#PCDATA)>

<!ELEMENT NBN                    (#PCDATA)>
<!ELEMENT TYPE                   (#PCDATA)>
<!ELEMENT SOURCEDESCRIPTION      (#PCDATA)>
<!ELEMENT SUPPLIER               (#PCDATA)>
<!ELEMENT SOURCETYPE             (#PCDATA)>
<!ELEMENT ORIGINALNBN            (#PCDATA)>
<!ELEMENT LIBRARYFUNCTION        (#PCDATA)>
<!ELEMENT CREATIONDATE           (#PCDATA)>
<!ELEMENT COLLNAME               (#PCDATA)>
<!ELEMENT CHECKSUM               (#PCDATA)>
<!ELEMENT CHECKSUMTYPE           (#PCDATA)>
<!ELEMENT FILECOUNT              (#PCDATA)>
<!ELEMENT REFPLATFORMNBN         (#PCDATA)>
<!ELEMENT AIPSTATUS              (#PCDATA)>
<!ELEMENT STARTERFILENAME        (#PCDATA)>
<!ELEMENT SIZE                   (#PCDATA)>
<!ELEMENT DIPRETRIEVEWAITTIME    (#PCDATA)>

<!ELEMENT CATEGORY               (#PCDATA)>
<!ELEMENT AUTHORITY              (#PCDATA)>
<!ELEMENT PATH                   (#PCDATA)>
```

4.1.4 Examples

4.1.4.1 Example of XML-document as response on metadata request:

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE DIASRESPONSE SYSTEM "diasresponse.dtd" >
<DIASRESPONSE type="DIP" version="1.0">
  <DESCRIPTION>metadata response</DESCRIPTION>
  <REQUEST>
    <NBN>duiken4Org</NBN>
  </REQUEST>
  <DIP>
    </METADATABLOCK>
      <METADATA>
        <NBN>duiken4Org</NBN>
        <ORIGINALNBN>duiken4Inst</ORIGINALNBN>
        <SUPPLIER>bim</SUPPLIER>
        <SOURCEDESCRIPTION>duiken4 </SOURCEDESCRIPTION>
        <CREATIONDATE>2005-04-11-10-00-00-000000</CREATIONDATE>
        <CHECKSUM>304975828</CHECKSUM>
        <CHECKSUMTYPE>CRC32</CHECKSUMTYPE>
        <FILECOUNT>34</FILECOUNT>
        <TYPE>I</TYPE>
        <COLLNAME>DISKCOLL</COLLNAME>
        <LIBRARYFUNCTIONAME>Hosting</LIBRARYFUNCTIONAME>
        <DIPRETRIEVEWAITTIME>600</DIPRETRIEVEWAITTIME>
        <REFPLATFORMNBN>REFPLATFORM-10</REFPLATFORMNBN>
        <AIPSTATUS>A</AIPSTATUS>
        <STARTERFILENAME>duiken.exe</STARTERFILENAME>
        <SIZE>454031360</SIZE>
      </METADATA>
    </METADATABLOCK>
    <ASSETBLOCK>
    </ASSETBLOCK>
  </DIP>
</DIASRESPONSE>
```

4.2 DIAS-METS

4.2.1 Unpacked DIP-file

The unpacked DIP-file is the default format for the response on a request for a specific asset (UOF.ac.AR1). The Archival Information Package (AIP) with the asset is retrieved from the archive and the asset is extracted and stored in unpacked format and made available on a HTTP accessible location in the DIAS-Core system.

The unpacked DIP-file is placed in a unique directory for the request. In this directory there is an 'unpacked' subdirectory where the content of the asset is placed. The content of the asset is the same as it was ingested. This means the technical metadata file and additional files are included as well.

Note: Direct access of a file >2GB size results in an error page with Apache HTTP Server 1.3.26.

4.2.2 DIP-archive

UOF.dip.F1 The DIAS-METS DIP Data Packages must be supported in the following formats:

DIP Data Package Format	Description	File Extension
ZIP	ZIP format	.zip
GNU-TAR-ZIPPED	GNU-TAR-ZIPPED format	.tar.gz
GNU-TAR	GNU TAR format	.tar

Table 2: DIAS-METS DIP Data Package Formats

UOF.dip.F2 Each ZIP format used for DIP Data Packages must be compliant with PKZIP version 2.50 or higher but lower than 5.0.

UOF.dip.F3 The ZIP format used for DIP Data Packages must support the use of the standard ZIP-compression features.

UOF.dip.F4 Each DIP Data Package must have a file name that is based upon the internal asset identifier and must have a file extension that conforms to the package format.

UOF.dip.F7 The TAR format used for DIP Data Package must be compliant with GNU-TAR version 1.15 and use the 'gnu' format.

The DIP-archive is the response on a request with the DipFmt parameter. A DIP-archive is made available on a HTTP accessible location in the DIAS-Core system. The DIP-archive is placed in a unique directory for the request. In this directory there are two subdirectories: the 'packed' and the 'unpacked' directory. In the packed directory the DIP-archive is named like;

- Dip-<*ResponseType*>-<*unique number*>.archive extension

Where ResponseType is the type of response requested, (asset or metadata) and unique number is the Internal Asset Identifier.

The Archival Information Package (AIP) with the asset is retrieved from the archive and the asset is extracted and stored in unpacked format in the unpacked directory. Then the asset is packet into the requested archive format and stored in the packed directory. The DIP-archive is made available on a HTTP accessible location in the DIAS-Core system. Dependent on request the DIP-archive contains metadata and/or the asset. When the asset is included the content of the asset is the same as it was ingested. This means the additional files and a descriptive file are included as well.

4.2.3 Response dtd

UOF.dip.F8 The request response in case the format is xml must conform to the following dtd:

```
<!ELEMENT DIASMETSRESPONSE (DESCRIPTION, REQUEST, DIP) >
<!ATTLIST DIASMETSRESPONSE
 type (error | dip | busy)       #IMPLIED
 version (1.0)                   #IMPLIED
 >

<!ELEMENT REQUEST ((INTERNALASSETID | EXTERNALASSETID), LIST?)>

<!ELEMENT DIP                    (METADATABLOCK?, ASSETBLOCK?) >
<!ELEMENT METADATABLOCK          (METADATA+) >
<!ELEMENT METADATA               (INTERNALASSETID, EXTERNALASSETID, ASSETVERSION,
                                 ASSETDESCRIPTION, AIPISDELETED, AIPDELETIONTIMESTAMP,
                                 ASSETCREATIONTIMESTAMP, ASSETINGESTTIMESTAMP,
                                 OLDOBJECTIDENTIFIER?, OLDVERSION?, SIZE?,
                                 DIPRETRIEVEWAITTIME?)>
<!ELEMENT ASSETBLOCK             (ASSET?) >
<!ELEMENT ASSET                  (INTERNALASSETID, URI?)>
<!ELEMENT URI                    (#PCDATA)>

<!ELEMENT DESCRIPTION            (#PCDATA)>

<!ELEMENT INTERNALASSETID        (#PCDATA)>
<!ELEMENT EXTERNALASSETID        (#PCDATA)>
<!ELEMENT LIST                   (#PCDATA)>
<!ELEMENT ASSETVERSION           (#PCDATA)>
<!ELEMENT ASSETDESCRIPTION       (#PCDATA)>
<!ELEMENT AIPISDELETED           (#PCDATA)>
<!ELEMENT AIPDELETIONTIMESTAMP   (#PCDATA)>
<!ELEMENT ASSETCREATIONTIMESTAMP (#PCDATA)>
<!ELEMENT ASSETINGESTTIMESTAMP   (#PCDATA)>
<!ELEMENT OLDOBJECTIDENTIFIER    (#PCDATA)>
<!ELEMENT OLDVERSION             (#PCDATA)>
<!ELEMENT SIZE                   (#PCDATA)>
<!ELEMENT DIPRETRIEVEWAITTIME    (#PCDATA)>
```

4.2.4 Examples

4.2.4.1 Example of XML-document as response on successful metadata request

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE DIASMETSRESPONSE …inline dtd removed… >
<DIASMETSRESPONSE type="dip" version="1.0">
    <DESCRIPTION>metadata response</DESCRIPTION>
    <REQUEST>

        <EXTERNALASSETID>urn:nbn:nl:ST_IA21-004</EXTERNALASSETID>

    </REQUEST>
    <DIP>
        <METADATABLOCK>
```

```
            <METADATA>
                <INTERNALASSETID>urn:diasid:ast:devos1:020050816113723795000O</INTERNALASSETID>
                <EXTERNALASSETID>urn:nbn:nl:ST_IA21-004</EXTERNALASSETID>
                <ASSETVERSION>1</ASSETVERSION>

                <ASSETDESCRIPTION>Domhof, Sebastian, Nichtparametrische relative
Effekte</ASSETDESCRIPTION>
                <AIPISDELETED>N</AIPISDELETED>
                <AIPDELETIONTIMESTAMP></AIPDELETIONTIMESTAMP>
                <ASSETCREATIONTIMESTAMP>2005-08-16T11:35:39.0Z</ASSETCREATIONTIMESTAMP>
                <ASSETINGESTTIMESTAMP>2005-08-16T11:37:23.76Z</ASSETINGESTTIMESTAMP>
                <OLDOBJECTIDENTIFIER></OLDOBJECTIDENTIFIER>
                <OLDVERSION></OLDVERSION>

                <SIZE>83845120</SIZE>
            </METADATA>
        </METADATABLOCK>
        <ASSETBLOCK>
        </ASSETBLOCK>
    </DIP>
</DIASMETSRESPONSE>
```

4.2.4.2 Example of XML-document as response on successful metadata as DIP Archive request

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE DIASMETSRESPONSE …inline dtd removed… >
<DIASMETSRESPONSE type="dip" version="1.0">
    <DESCRIPTION>metadata response</DESCRIPTION>
    <REQUEST>

        <EXTERNALASSETID>urn:nbn:nl:ST_IA21-004</EXTERNALASSETID>

    </REQUEST>
    <DIP>
        <METADATABLOCK>
            <METADATA>
                <INTERNALASSETID>urn:diasid:ast:devos1:020050816113723795000O</INTERNALASSETID>
                <EXTERNALASSETID>urn:nbn:nl:ST_IA21-004</EXTERNALASSETID>
                <ASSETVERSION>1</ASSETVERSION>

                <ASSETDESCRIPTION>Domhof, Sebastian, Nichtparametrische relative
Effekte</ASSETDESCRIPTION>
                <AIPISDELETED>N</AIPISDELETED>
                <AIPDELETIONTIMESTAMP></AIPDELETIONTIMESTAMP>
                <ASSETCREATIONTIMESTAMP>2005-08-16T11:35:39.0Z</ASSETCREATIONTIMESTAMP>
                <ASSETINGESTTIMESTAMP>2005-08-16T11:37:23.76Z</ASSETINGESTTIMESTAMP>
                <OLDOBJECTIDENTIFIER></OLDOBJECTIDENTIFIER>
                <OLDVERSION></OLDVERSION>

                <SIZE>83845120</SIZE>
            </METADATA>
        </METADATABLOCK>
```

```
        <ASSETBLOCK>
            <ASSET>
                <INTERNALASSETID>urn:diasid:ast:devos1:0200508161137237950000</INTERNALASSETID>
                <URI>https://9.142.108.32/asset//cmos_1124447570322/DipMetadata24754.zip</URI>

            </ASSET>
        </ASSETBLOCK>
    </DIP>
</DIASMETSRESPONSE>
```

4.2.4.3 Example of XML-document as response on successful full metadata request

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE DIASMETSRESPONSE …inline dtd removed… >
<DIASMETSRESPONSE type="dip" version="1.0">
    <DESCRIPTION>dip response</DESCRIPTION>
    <REQUEST>

        <EXTERNALASSETID>urn:nbn:nl:ST_IA21-004</EXTERNALASSETID>

    </REQUEST>
    <DIP>
        <METADATABLOCK>
            <METADATA>
                <INTERNALASSETID>urn:diasid:ast:devos1:0200508161137237950000</INTERNALASSETID>
                <EXTERNALASSETID>urn:nbn:nl:ST_IA21-004</EXTERNALASSETID>
                <ASSETVERSION>1</ASSETVERSION>

                <ASSETDESCRIPTION>Domhof, Sebastian, Nichtparametrische relative
Effekte</ASSETDESCRIPTION>
                <AIPISDELETED>N</AIPISDELETED>
                <AIPDELETIONTIMESTAMP></AIPDELETIONTIMESTAMP>
                <ASSETCREATIONTIMESTAMP>2005-08-16T11:35:39.0Z</ASSETCREATIONTIMESTAMP>
                <ASSETINGESTTIMESTAMP>2005-08-16T11:37:23.76Z</ASSETINGESTTIMESTAMP>
                <OLDOBJECTIDENTIFIER></OLDOBJECTIDENTIFIER>
                <OLDVERSION></OLDVERSION>


            </METADATA>
        </METADATABLOCK>
        <ASSETBLOCK>
          <ASSET>
                <INTERNALASSETID>urn:diasid:ast:devos1:0200508161137237950000</INTERNALASSETID>
                <URI>https://9.142.108.32/asset//cmos_1124447163885/unpacked/mets.xml</URI>
            </ASSET>
        </ASSETBLOCK>

    </DIP>
</DIASMETSRESPONSE>
```

4.2.4.4 Example of XML-document as response on successful asset request

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE DIASMETSRESPONSE …inline dtd removed… >
<DIASMETSRESPONSE type="dip" version="1.0">
    <DESCRIPTION>dip response</DESCRIPTION>
    <REQUEST>

        <EXTERNALASSETID>urn:nbn:nl:ST_IA21-004</EXTERNALASSETID>

    </REQUEST>
    <DIP>
        <METADATABLOCK>
            <METADATA>
                <INTERNALASSETID>urn:diasid:ast:devos1:02005081611372379500000</INTERNALASSETID>
                <EXTERNALASSETID>urn:nbn:nl:ST_IA21-004</EXTERNALASSETID>
                <ASSETVERSION>1</ASSETVERSION>

                <ASSETDESCRIPTION>Domhof, Sebastian, Nichtparametrische relative
Effekte</ASSETDESCRIPTION>
                <AIPISDELETED>N</AIPISDELETED>
                <AIPDELETIONTIMESTAMP></AIPDELETIONTIMESTAMP>
                <ASSETCREATIONTIMESTAMP>2005-08-16T11:35:39.0Z</ASSETCREATIONTIMESTAMP>
                <ASSETINGESTTIMESTAMP>2005-08-16T11:37:23.76Z</ASSETINGESTTIMESTAMP>
                <OLDOBJECTIDENTIFIER></OLDOBJECTIDENTIFIER>
                <OLDVERSION></OLDVERSION>


                <SIZE>83845120</SIZE>

            </METADATA>
        </METADATABLOCK>
        <ASSETBLOCK>
          <ASSET>
                <INTERNALASSETID>urn:diasid:ast:devos1:02005081611372379500000</INTERNALASSETID>
                <URI>https://9.142.108.32/asset//cmos_1124446827334/unpacked/</URI>

            </ASSET>
        </ASSETBLOCK>
    </DIP>
</DIASMETSRESPONSE>
```

4.2.4.5 Example of XML-document as response on server busy with full metadata request

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE DIASMETSRESPONSE …inline dtd removed… >
<DIASMETSRESPONSE type="busy" version="1.0">
    <DESCRIPTION>in progress response</DESCRIPTION>
    <REQUEST>

        <EXTERNALASSETID>urn:nbn:nl:ST_IA21-004</EXTERNALASSETID>

    </REQUEST>
    <DIP>
```

```
        <METADATABLOCK>
            <METADATA>
                <INTERNALASSETID>urn:diasid:ast:devos1:0200508161137237950000</INTERNALASSETID>
                <EXTERNALASSETID>urn:nbn:nl:ST_IA21-004</EXTERNALASSETID>
                <ASSETVERSION>1</ASSETVERSION>

                <ASSETDESCRIPTION>Domhof, Sebastian, Nichtparametrische relative
Effekte</ASSETDESCRIPTION>
                <AIPISDELETED>N</AIPISDELETED>
                <AIPDELETIONTIMESTAMP></AIPDELETIONTIMESTAMP>
                <ASSETCREATIONTIMESTAMP>2005-08-16T11:35:39.0Z</ASSETCREATIONTIMESTAMP>
                <ASSETINGESTTIMESTAMP>2005-08-16T11:37:23.76Z</ASSETINGESTTIMESTAMP>
                <OLDOBJECTIDENTIFIER></OLDOBJECTIDENTIFIER>
                <OLDVERSION></OLDVERSION>


            </METADATA>
        </METADATABLOCK>
        <ASSETBLOCK>
        </ASSETBLOCK>
    </DIP>
</DIASMETSRESPONSE>
```

4.2.4.6 Example of XML-document as response on server busy with asset request

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE DIASMETSRESPONSE …inline dtd removed… >
<DIASMETSRESPONSE type="busy" version="1.0">
    <DESCRIPTION>in progress response</DESCRIPTION>
    <REQUEST>

        <EXTERNALASSETID>urn:nbn:nl:ST_IA21-004</EXTERNALASSETID>

    </REQUEST>
    <DIP>
        <METADATABLOCK>
            <METADATA>
                <INTERNALASSETID>urn:diasid:ast:devos1:0200508161137237950000</INTERNALASSETID>
                <EXTERNALASSETID>urn:nbn:nl:ST_IA21-004</EXTERNALASSETID>
                <ASSETVERSION>1</ASSETVERSION>

                <ASSETDESCRIPTION>Domhof, Sebastian, Nichtparametrische relative
Effekte</ASSETDESCRIPTION>
                <AIPISDELETED>N</AIPISDELETED>
                <AIPDELETIONTIMESTAMP></AIPDELETIONTIMESTAMP>
                <ASSETCREATIONTIMESTAMP>2005-08-16T11:35:39.0Z</ASSETCREATIONTIMESTAMP>
                <ASSETINGESTTIMESTAMP>2005-08-16T11:37:23.76Z</ASSETINGESTTIMESTAMP>
                <OLDOBJECTIDENTIFIER></OLDOBJECTIDENTIFIER>
                <OLDVERSION></OLDVERSION>


                <SIZE>83845120</SIZE>
                <DIPRETRIEVEWAITTIME>20</DIPRETRIEVEWAITTIME>
```

```
            </METADATA>
        </METADATABLOCK>
        <ASSETBLOCK>
        </ASSETBLOCK>
    </DIP>
</DIASMETSRESPONSE>
```

4.2.4.7 Example of XML-document as response on unsuccessful request

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE DIASMETSRESPONSE …inline dtd removed… >
<DIASMETSRESPONSE type="error" version="1.0">
    <DESCRIPTION>error. details in retriever logfile</DESCRIPTION>
    <REQUEST>

        <EXTERNALASSETID>urn:nbn:nl:abc:233-449</EXTERNALASSETID>

    </REQUEST>
    <DIP>
    </DIP>
</DIASMETSRESPONSE>
```