

# Baglt AIP Format

## General Baglt AIP Structure

- Site AIP (Sample - [SITE@123456789-0.zip](#))
  - Human readable information about the Site describing the DSpace instance
  - Contains a listing of all Groups and EPeople translated into the DSpaceRoles schema
  - Contains a listing of all handles of the DSpace objects for the site
- Community AIP (Sample - [COMMUNITY@123456789-1.zip](#))
  - Persistent ID information about the Community and a Logo if it exists
  - All Community permissions/policies translated to a policy.xml schema
  - Permissions/policies translated to the policy.xml schema
  - Groups and EPeople objects translate into the DSpaceRoles schema
- Collection AIP (Sample - [COLLECTION@123456789-2.zip](#))
  - Persistent ID information about the Collection and a Logo if it exists
  - Collection specific metadata translated to the metadata.xml schema
  - Permissions/policies translated to the policy.xml schema
  - Groups and EPeople objects translate into the DSpaceRoles schema
- Item AIP (Sample - [ITEM@123456789-3.zip](#))
  - All Bitstream files
  - Persistent ID information about the Item
  - Metadata for the Item and all Bitstreams translated to a metadata.xml schema
  - All Item and Bitstream permissions/policies translated to a policy.xml schema
- Deletion Catalog (Sample - [DELETION-RECORD@123456789-6.zip](#))
  - Catalog of a deleted DSpace object
  - Contains a listing of the handle of the DSpace object and all handles of child objects

### Notes

- The Deletion Catalog is only created if the Automation Options are set for the Replication Task Suite

## Baglt Payload Detailed Structure

- `members` file
  - For Sites and Deletion Catalogs, a listing of all child DSpace object Handles
- `object.properties` file
  - `bagType` The type of bag written, either AIP or man
  - `objectType` The DSpace object type, one of site, community, item, or deletion
  - `objectId` The handle of the DSpace object associated for the AIP
  - `ownerId` The handle for the parent DSpace object, if available
  - `created` The time at which the Catalog was created
- `dspace.properties` file
  - `Site-Handle` The handle of the DSpace Site
  - `DSpace-Version` The version of DSpace running
- `roles.xml`
  - See DSPACE-ROLES Schema in the [DSpace AIP Format](#)
- `policy.xml`
  - See [Baglt AIP Format#Policy Schema](#)
- `metadata.xml`
  - See [Baglt AIP Format#Metadata Schema](#)

## Policy Schema

The policy schema is a way to capture permission and policy information belonging to DSpace objects into a `policy.xml` file. It provides a one-to-one mapping to the `ResourcePolicy` object on a DSpace object. The hierarchy is as follows:

- `policies` element
  - `policy` - a `ResourcePolicy` serialized as an XML tag

## Policy Attribute Mapping

XML Attribute	ResourcePolicy field
name	name
type	rpType
group	group
action	action (represented as a String)
eperson	ePerson#getEmail

end-date	endDate
start-date	startDate
description	rpDescription

### Example Schema

The following is an example `policy.xml` which belongs to a single DSpace object. It contains 4 policies: one for a custom group, one for the Administrator group, one for the Anonymous group, and one for an EPerson.

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<policies>
  <policy action="ADMIN" group="COMMUNITY_hdl:123456789/1_ADMIN" />
  <policy start-date="2010-01-01" action="READ" group="Anonymous" />
  <policy start-date="2010-01-01" action="ADMIN" group="Administrator" />
  <policy type="TYPE_CUSTOM" eperson="dspace-user@localhost.localdomain" action="Add" />
</policies>
```

### Metadata Schema

The `metadata.xml` contains information about the DSpace object, typically user generated. The hierarchy is as follows:

- `metadata` element
  - `value` element - corresponds to the `MetadataValue` for the DSpace object or information captured from the object itself; described below for each applicable DSpace object

#### Value Attribute Mapping for Communities

For Communities, the following fields are translated to the metadata schema:

XML Attribute	Database Field
name	name
short_description	short_description
introductory_text	introductory_text
copyright_text	copyright_text
side_bar_text	side_bar_text

#### Value Attribute Mapping for Collections

For Collections, the following fields are translated to the metadata schema:

XML Attribute	Database Field
name	name
short_description	short_description
introductory_text	introductory_text
provenance_description	provenance_description
license	license
copyright_text	copyright_text
side_bar_text	side_bar_text

#### Value Attribute Mapping for Items

For Items, the following fields are translated to the metadata schema:

XML Attribute	MetadataValue Field
language	MetadataValue#getLanguage

qualifier	MetadataValue#getMetadataFeild#getQualifier
element	MetadataValue#getMetadataField#getElement
schema	MetadataValue#getMetadataField#getMetadataSchema#getName

### Value Attribute Mapping for Bitstreams

For Bitstreams, the following fields are translated to the metadata schema:

Bitstream Field
name
source
description
sequenceID
true if the Bitstream is equal to Bundle#getPrimaryBitstream

### Example metadata.xml (Item)

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?
>
<metadata>
  <value element="contributor" qualifier="author" schema="dc">User, DSpace<
/ value>
  <value element="date" qualifier="accessioned" schema="dc">2020-04-20T19:14:01Z<
/ value>
  <value element="date" qualifier="available" schema="dc">2020-04-20T19:14:01Z<
/ value>
  <value element="date" qualifier="issued" schema="dc">2019-02-10<
/ value>
  <value element="identifier" qualifier="uri" schema="dc">http://localhost:8080/xmlui/handle/123456789/3<
/ value>
  <value element="description" language="en" qualifier="provenance" schema="dc">Submitted by DSpace User
(dspace-test@localhost) on 2020-04-20T19:14:01Z
No. of bitstreams:
1
cover.jpg: 65692 bytes, checksum: ce271030c23b61c97f609e6cb8b55542 (MD5)<
/ value>
  <value element="description" language="en" qualifier="provenance" schema="dc">Made available in DSpace on
2020-04-20T19:14:01Z (GMT). No. of bitstreams: 1
cover.jpg: 65692 bytes, checksum: ce271030c23b61c97f609e6cb8b55542
(MD5)
  Previous issue date: 1989-02-10<
/ value>
  <value element="description" language="en" qualifier="provenance" schema="dc">Restored into DSpace on 2020-
04-20T21:31:39Z (GMT).</ value>
  <value element="language" language="en_US" qualifier="iso" schema="dc">en<
/ value>
  <value element="subject" language="en_US" schema="dc">Keyword<
/ value>
  <value element="title" language="en_US" schema="dc">A Title<
/ value>
  <value element="title" language="en_US" qualifier="alternative" schema="dc">Alternate Title<
/ value>
  <value element="type" language="en_US" schema="dc">Image<
/ value>
<
/ metadata>
```

## Binary Payloads

### Bitstreams

For Items or other objects which contain Bundles and Bitstreams, a directory is created for each Bundle under which Bitstreams are collected. Each bitstream is stored using a prefix of 'bitstream\_', the uuid of the Bitstream, and a file extension if it can be derived. e.g. for a Bundle ORIGINAL and Bitstream with a uuid 46f383da-15d7-423e-bd76-741d9c503d0f will be stored under data/ORIGINAL/46f383da-15d7-423e-bd76-741d9c503d0f.

### Logos

If a DSpace Object contains a logo, it is written directly under the data directory using the same filename format as Bitstreams.

## Examples

### Example Layout of a BagIt Site AIP

```
SITE@123456789-0
bag-info.txt
bagit.txt
data
  dspace.properties
  members
  object.properties
  roles.xml
manifest-md5.txt
tagmanifest-md5.txt
```

### Example Layout of a BagIt Community AIP

```
COMMUNITY@123456789-1
bag-info.txt
bagit.txt
data
  bitstream_883620c8-4486-4a72-a86b-242a1120f69c
  metadata.xml
  object.properties
  policy.xml
  roles.xml
manifest-md5.txt
tagmanifest-md5.txt
```

### Example Layout of a BagIt Collection AIP

```
COLLECTION@123456789-2
bag-info.txt
bagit.txt
data
  bitstream_4598cca5-9687-466c-9278-8a34f877abec
  metadata.xml
  object.properties
  policy.xml
  roles.xml
manifest-md5.txt
tagmanifest-md5.txt
```

### Example Layout of a BagIt Item AIP

ITEM@123456789-3

bagit.txt

bag-info.txt

data

LICENSE

bitstream\_e8609dc1-eecf-49a4-ab66-e82d6fe8d689

bitstream\_e8609dc1-eecf-49a4-ab66-e82d6fe8d689-metadata.xml

bitstream\_e8609dc1-eecf-49a4-ab66-e82d6fe8d689-policy.xml

ORIGINAL

bitstream\_46f383da-15d7-423e-bd76-741d9c503d0f.jpg

bitstream\_46f383da-15d7-423e-bd76-741d9c503d0f-metadata.xml

bitstream\_46f383da-15d7-423e-bd76-741d9c503d0f-policy.xml

bitstream\_a2140708-8398-4931-b6d7-26d2a217ec96.jpg

bitstream\_a2140708-8398-4931-b6d7-26d2a217ec96-metadata.xml

bitstream\_a2140708-8398-4931-b6d7-26d2a217ec96-policy.xml

policy.xml

metadata.xml

object.properties

manifest-md5.txt

tagmanifest-md5.txt