

## Java Client - Feature #6369

### ResourceMapFactory.parseResourceMap does not return aggregated resources

2014-09-09 15:38 - Skye Roseboom

<b>Status:</b>	In Progress	<b>Start date:</b>	2014-09-25
<b>Priority:</b>	Normal	<b>Due date:</b>	2014-10-06
<b>Assignee:</b>	Rob Nahf	<b>% Done:</b>	30%
<b>Category:</b>	d1_libclient_java	<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>			
<b>Story Points:</b>			

#### Description

During testing an ORE that only defines aggregation relationships:

It appears that resources that are only 'aggregated' by the resource map - and not 'documentedBy' or 'document' another resource - do not get returned in the data structure created by the 'parseResourceMap' method.

Since this method is an encapsulation of org.dspace.forsite.ResourceMap - to allow an interface to hide the parsing method (XPath vs forsite) - it seems it should return all valid identifiers/relationships found in the ORE document.

#### History

##### #1 - 2014-09-25 17:24 - Dave Vieglais

- Start date set to 2014-09-25
- Target version set to Release Backlog
- Due date set to 2014-09-25

##### #2 - 2014-09-25 17:28 - Dave Vieglais

- Target version changed from Release Backlog to CCI-1.5.0

##### #3 - 2014-09-26 23:24 - Rob Nahf

- Status changed from New to In Progress
- Tracker changed from Bug to Feature
- Due date changed from 2014-09-25 to 2014-09-26
- Target version changed from CCI-1.5.0 to 329
- % Done changed from 0 to 20
- Product Version changed from \* to 1.4.0

this goes beyond the implementation capabilities, so I'm changing this to a feature request.

parseResourceMap was written to extract the metadata and data objects from a ResourceMap (serialized) in order to build a DataPackage object. Keys of the returned map are (generally) METADATA objects, defined in the ResourceMap as Subject of the cito:documents triple (and object of the cito:documentedBy triple), while values of the map are (generally) lists of DATA objects, defined by the inverse relationship triples.

The situation that raised this issue is a ResourceMap that aggregates 3 objects that represent 3 different formats of the same METADATA (probably with minor variations). Without the cito:documents/isDocumentedBy relationships, there isn't a way to know whether an object is a METADATA or DATA object, and thus whether it should be a key or value of the returned map.

A couple possible solutions are possible, both requiring incrementing the minor version:

1. define a separate method for retrieving resource map aggregated resources - it would not know how to differentiate between METADATA and DATA. The advantage of this is clarity, the downside is need to manage ResourceMap objects in memory (which can be much larger than their serialized form)
  1. Return objects not participating in the cito:documents/documentedBy relationships as keys in the existing map structure of parseResourceMap. The advantage of this is that the application layer doesn't have to manage ResourceMap instances - the serialized form can be completely parsed into a relatively small Map of Identifiers.

For retrieving all relationships, it might make more sense to work with the `foresite.ResourceMap` object directly:  
`ResourceMap resMap = ResourceMapFactory.deserialize(InputStream is);`  
`List trips = resMap.listAllTriples();`  
`trips.get(0).getSubject();`  
`trips.get(0).getPredicate();`  
`trips.get(0).getObject();`  
etc.

**#4 - 2014-09-26 23:25 - Rob Nahf**

- Target version changed from 329 to CCI-1.5.0

**#5 - 2014-10-02 19:36 - Skye Roseboom**

- Target version changed from CCI-1.5.0 to Release Backlog

- Due date changed from 2014-09-26 to 2014-10-02

**#6 - 2014-10-06 22:58 - Rob Nahf**

- Due date changed from 2014-10-02 to 2014-10-06

- Product Version deleted (1.4.0)

**#7 - 2015-02-11 20:29 - Rob Nahf**

- % Done changed from 20 to 30

- Target version changed from Release Backlog to CLJ

- Category changed from `d1_libclient_java` to `d1_libclient_java`

- Project changed from Infrastructure to Java Client

**#8 - 2015-02-11 20:56 - Rob Nahf**

- Target version deleted (CLJ)