

Member Nodes - Task #5942

MNDeployment # 3708 (Operational): Minnesota Population Center

Task # 5921 (Closed): MPC: Testing

Task # 5922 (Closed): MPC: Registration in environment

Task # 5933 (Closed): MPC: Content Review

Task # 5941 (Closed): MPC: Verify Resource Maps

MPC: Verify Resource Map content

2014-07-18 15:31 - Laura Moyers

Status:	Closed	Start date:	2014-07-18
Priority:	Normal	Due date:	
Assignee:	Chris Jones	% Done:	100%
Category:		Estimated time:	0.00 hour
Target version:	Operational		
Story Points:			
Description			
Verify that Resource Maps are complete, correctly formatted and represent the intended associations.			
Related issues:			
Blocks Member Nodes - Task #5944: MPC: Verify that Resource Maps are correctl...		Closed	2014-07-18
Blocks Member Nodes - Task #5932: MPC: Set up synchronization of the MN		Rejected	2014-07-18

History

#1 - 2014-08-04 18:47 - Laura Moyers

- Target version changed from Deploy by end of Y5Q4 to Deploy by end of Y1Q1

#2 - 2014-08-26 16:27 - Chris Jones

- Assignee set to Chris Jones

- Status changed from New to In Progress

I'm seeing a few issues with the MPC resource maps at

<https://dataone-test.pop.umn.edu/mn/v1/object?formatId=http://www.openarchives.org/ore/terms> :

- 1) Some system metadata for resource maps have a formatId of 'application/octet-stream'. This needs to be changed to '<http://www.openarchives.org/ore/terms>'. See https://dataone-test.pop.umn.edu/mn/v1/object/ipumsi_6-3_cr_1984.rm.xml as an example.
- 2) The serialized resource maps include an aggregation statement where the resource map itself is aggregated, which I don't think is correct. The aggregation should only include the science data and science metadata triple statements (or other types of metadata, like provenance, etc.).
- 3) The triple statements in the resource maps don't point to CN-resolvable URIs, which is a requirement for DataONE data packages. See <http://mule1.dataone.org/ArchitectureDocs-current/design/DataPackage.html#generating-resource-maps> . An example is in https://dataone-test.pop.umn.edu/mn/v1/object/ipumsi_6-3_cr_1984.rm.xml, where both the subjects and objects in the triple statements point to URIs like <http://international.ipums.org/> ...

#3 - 2014-08-28 20:02 - Bruce Wilson

I communicated the formatId issue to Fabio and Wendy today (2014-08-28 2:00 PM EDT).

I think that there's a fourth issue to address:

3b) Identifiers appear to be using mixed case and underscores, but the MPC identifiers are all lower case and dots. For example, in https://dataone-test.pop.umn.edu/mn/v1/object/ipumsi_6-3_cr_1984.rm.xml one of the items in the aggregation is @@. But the object that this likely refers to has an identifier ipumsi_6-3_cr_1984.dc.xml

#4 - 2014-09-06 00:36 - Chris Jones

- File ipumsi_6-3_at_2001-without-data.rdf.xml added
- File ipumsi_6-3_at_2001-without-data.rdf.xml added

#5 - 2014-09-06 00:46 - Chris Jones

- File ipumsi_6-3_at_2001-with-data.rdf.xml added
- File ipumsi_6-3_at_2001-with-data.rdf.xml added

I've attached two example resource maps to help clarify the content of the MPC resource maps.

The first describes an aggregation that has no science data in it, but rather three metadata files (Dublin Core file, DDIC XML file, and DDIC HTML file). Because only the Dublin Core metadata file is formatType METADATA in our object format list, its fields will get parsed into the search index. The DDIC XML file, for now, can stay with a formatId of application/octet-stream, and its fields won't be parsed. The DDIC HTML transform should have a formatId of text/html, and it too won't have its fields parsed. However, all three of these files will be available for download by scientists since they are part of the aggregation (Data Package).

The second example describes an aggregation that contains one science metadata file (Dublin Core), and two science data files (CSVs). This resource map shows how the one science metadata file 'cito:documents' the two science data files, and all three are members of the aggregation (Data Package).

#6 - 2014-09-19 13:46 - Chris Jones

- translation missing: en.field_remaining_hours set to 0.0
- Status changed from In Progress to Closed

Skye and I have worked with Wendy to correct resource map issues, and now the resource maps are being parsed correctly. Closing this ticket.

#7 - 2014-09-29 21:30 - Laura Moyers

- Target version changed from Deploy by end of Y1Q1 to Deploy by end of NCTE

#8 - 2015-01-02 16:42 - Laura Moyers

- Target version changed from Deploy by end of NCTE to Operational

#9 - 2015-01-28 15:38 - Laura Moyers

- % Done changed from 0 to 100
- Estimated time set to 0.00

Files

ipumsi_6-3_at_2001-without-data.rdf.xml	4.46 KB	2014-09-06	Chris Jones
ipumsi_6-3_at_2001-with-data.rdf.xml	4.95 KB	2014-09-06	Chris Jones