Morpho - Bug #5843

Failed to open the data package - https://mn-demo-5.test.dataone.org/knb/d1/mn/v1/object/doi:10.6085/AA/YBHX00_XXXITBDXMMR01_ 20040720.50.5

02/01/2013 03:53 PM - Jing Tao

Status:	Resolved		**Start date:**	02/01/2013
Priority:	Normal		**Due date:**	
Assignee:	Jing Tao		**% Done:**	0%
Category:	morpho - general		**Estimated time:**	0.00 hour
Target version:	2.0.0		**Spent time:**	0.00 hour
Bugzilla-Id:	5843			

Description

When I search "PISCO" and get the data package:
https://mn-demo-5.test.dataone.org/knb/d1/mn/v1/object/doi:10.6085/AA/YBHX00_XXXITBDXMMR01_20040720.50.5

However, when i clicked it, i can't open it.

History

#1 - 02/01/2013 04:13 PM - Jing Tao

the error message is:
[java] Could not read package: No system metadata could be found for given PID: YBHX00_XXXITBDXMMR01_20040720.40.2
[java] Exception in thread "Thread-11" java.lang.NullPointerException
[java]   at edu.ucsb.nceas.morpho.datapackage.DataPackagePlugin.openDataPackage(DataPackagePlugin.java:987)
[java]   at edu.ucsb.nceas.morpho.query.OpenPackageCommand$1.construct(OpenPackageCommand.java:170)
[java]   at edu.ucsb.nceas.morpho.framework.SwingWorker$2.run(SwingWorker.java:140)
[java]   at java.lang.Thread.run(Thread.java:680)

#2 - 02/01/2013 04:15 PM - Jing Tao

It seems that the failure of one entity downloading causes the failure of the open. We should still open a package even though some or all entities fail to be downloaded.

#3 - 02/01/2013 05:08 PM - Jing Tao

I added code to handle the failure of one entity downloading gracefully and can save it locally.

However, the id was changed to YBHX00_XXXITBDXMMR01_20040720.40.2, the last part of the DOI.

I dug it a little bit. This is an eml 2.0 document, if we don't transform it, the package shows the whole DOI. If we transform it to eml 2.1, it shows the last part of DOI. By the way, the pacakageId in the eml is the last part of the doi.

#4 - 02/01/2013 05:22 PM - ben leinfelder

The ID was not changed -- that is the data entity ID originally used in PISCO and encoded in the ecogrid url:
<url function="download">ecogrid://knb/YBHX00_XXXITBDXMMR01_20040720.40.2</url>

I don't know how we can handle this case at the moment. In theory, the ORE package would tell us what the data file is, but we also expect that the data ID in the EML matches the ID used in the ORE package (and what is saved on the server). In this case the EML has the data ID without the DOI "shoulder" (as expected) but Morpho has no way to "know" that there is more to it than that ID.

#5 - 02/01/2013 05:42 PM - Jing Tao

Sorry, it is my fault. After transform to the new verion (eml 2.1) document, the main morpho frame shows: DataPackage: YBHX00_XXXITBDXMMR01_20040720.50.5
Accession Number: YBHX00_XXXITBDXMMR01_20040720.50.5
And the data package location is blank (this is correct).

(it is not the YBHX00_XXXITBDXMMR01_20040720.40.2, the data file id. It is the last part of DOI package). Note the eml has the packageId=YBHX00_XXXITBDXMMR01_20040720.40.2.

Now i saved the package into local, it changes to a UUID. This is great.

We only need to figure out why it shows the last part of DOI after the eml version transform.

#6 - 02/01/2013 06:12 PM - Jing Tao

A good news: i can use the open id menu item to open a package with a DOI id.

#7 - 02/01/2013 09:33 PM - Jing Tao

As I suspected: during the version transformation, the original identifier (doi:10.6085/AA/YBHX00_XXXITBDXMMR01_20040720.50.5) was replaced by the attribute value of the eml (YBHX00_XXXITBDXMMR01_20040720.50.5). After the transformation, the identifier changed to YBHX00_XXXITBDXMMR01_20040720.50.5.

I added code to preserve the id. After the transformation, the id still be doi:10.6085/AA/YBHX00_XXXITBDXMMR01_20040720.50.5 and the location is "". Then i saved it and it created a local copy of doi:10.6085/AA/YBHX00_XXXITBDXMMR01_20040720.50.5.

This cause an issue: both local and network has a copy of doi:10.6085/AA/YBHX00_XXXITBDXMMR01_20040720.50.5. But they are different! The newwork copy is eml 2.0.1 and the local copy is eml 2.1.0!

So I think so after eml version transformation, we should assign a new id for this data package. Because The transformation is a update.

#8 - 02/01/2013 11:08 PM - ben leinfelder

I agree that transforming the EML from one version to another is an "update" and therefore requires a new identifier be generated for it.

#9 - 02/04/2013 12:00 PM - Jing Tao

After update the eml to the newest version, morpho assigns a new uuid for the package. Now you can update and save the package correctly.

#10 - 03/27/2013 02:31 PM - Redmine Admin

Original Bugzilla ID was 5843