

Member Nodes - Task #3567

MNDeployment # 3559 (Operational): ONEShare

Unresolvable content on CN from ONEShare

2013-02-11 18:36 - Skye Roseboom

Status: New	Start date: 2013-02-11
Priority: Normal	Due date:
Assignee:	% Done: 0%
Category:	Estimated time: 0.00 hour
Target version: Operational	
Story Points:	

Description

Content discovered in ONEMercury in the CN's search index and having meta data on the CN but do not exist on the MN any longer:

ark:/99999/fk4h42dbf/1/SecondFile-algae.csv
ark:/99999/fk4h42dbf/1/SecondFile-pop_size.csv
ark:/99999/fk4h42dbf/1/SecondFile-Oct_7to15.csv
ark:/99999/fk4h42dbf/1/SecondFile-scope_data.csv
ark:/99999/fk4h42dbf/1/mrt-eml.xml
ark:/99999/fk4h42dbf/1/mrt-dataone-map.rdf
ark:/99999/fk40c5hpb/1/17.01.2013_-_18.01.2013-DailyReport.csv
ark:/99999/fk40c5hpb/1/mrt-eml.xml
ark:/99999/fk40c5hpb/1/mrt-dataone-map.rdf

Content should be archived so it no longer appears in search results or provide /object service for metadata. Also so /resolve service behaves properly.

History

#1 - 2013-02-11 18:38 - Skye Roseboom

- Parent task set to #3568

#2 - 2013-02-18 16:42 - Skye Roseboom

- Assignee deleted (Skye Roseboom)

#3 - 2013-02-22 17:39 - Skye Roseboom

- Category set to 296

- Assignee set to John Kunze

First step is to determine if this content should be available from the MN. If it should no longer be available on the CN, we need to 'archive' the pids to remove these docs from search index and other services.

#4 - 2013-02-22 18:00 - Skye Roseboom

- Target version set to Operational

#5 - 2013-02-22 18:00 - Skye Roseboom

- Parent task changed from #3568 to #3559

- Project changed from Infrastructure to Member Nodes

- Category deleted (296)

- Assignee deleted (John Kunze)

#6 - 2013-02-22 18:26 - Skye Roseboom

An example of a document that has been harvested by CN from ONEShare:

<https://cn-unm-1.dataone.org/cn/v1/meta/ark:/99999/fk4h42dbf/1/mrt-eml.xml>

however resolving that document will not work:

<https://cn-unm-1.dataone.org/cn/v1/resolve/ark:/99999/fk4h42dbf/1/mrt-eml.xml>

<https://oneshare.unm.edu/knb/d1/mn/v1/object/ark:%2F99999%2Ffk4h42dbf%2F1%2Fmrt-eml.xml>

<https://oneshare.unm.edu/knb/d1/mn/v1/meta/ark:%2F99999%2Ffk4h42dbf%2F1%2Fmrt-eml.xml>

#7 - 2013-02-22 20:43 - John Kunze

Here is a list of the PIDs that are most likely still in the CN, but have been removed from our ONEShare MN. I think it's the same as the original list.

They were submitted in error by our tech writer and should all no longer be available on the CN, so please 'archive' those PIDS.

ark:/99999/fk40c5hpb/1/17.01.2013_-_18.01.2013-DailyReport.csv
ark:/99999/fk40c5hpb/1/mrt-dataone-map.rdf
ark:/99999/fk40c5hpb/1/mrt-eml.xml
ark:/99999/fk4h42dbf/1/mrt-dataone-map.rdf
ark:/99999/fk4h42dbf/1/mrt-eml.xml
ark:/99999/fk4h42dbf/1/SecondFile-algae.csv
ark:/99999/fk4h42dbf/1/SecondFile-Oct_7to15.csv
ark:/99999/fk4h42dbf/1/SecondFile-pop_size.csv
ark:/99999/fk4h42dbf/1/SecondFile-scope_data.csv

Is there a way we can do this ourselves in the future?

#8 - 2016-06-03 13:43 - Laura Moyers

This is what I am seeing as of 6/2/16, which is not exactly the same as what we were seeing before (bolded ones are the same):

ark:/99999/fk40c5hpb/1/17.01.2013_-_18.01.2013-DailyReport.csv
ark:/99999/fk4d813mb/1/mrt-dataone-map.rdf
ark:/99999/fk4d813mb/1/mrt-eml.xml
ark:/99999/fk4h42dbf/1/SecondFile-algae.csv
ark:/99999/fk4h42dbf/1/SecondFile-Oct_7to15.csv
ark:/99999/fk4h42dbf/1/SecondFile-pop_size.csv
ark:/99999/fk4h42dbf/1/SecondFile-scope_data.csv
ark:/99999/fk4d813mb/1/CTD_data3.xlsx

It is likely that I have missed some in my cursory review. We need someone from ONEShare to investigate the content currently on the production ONEShare MN (i.e. discoverable via DataONE) and determine which objects should be archived. DataONE can help by showing/telling the ONEShare person how to accomplish this.