Member Nodes - Task #7780

MNDeployment # 7051 (Operational): GRIIDC - Gulf of Mexico Research Initiative

GRIIDC: Register any new object formats in all environments

2016-05-10 01:24 - Laura Moyers

Status:	Closed		**Start date:**	2016-05-10
Priority:	Normal		**Due date:**	
Assignee:	Jing Tao		**% Done:**	100%
Category:			**Estimated time:**	0.00 hour
Target version:	Deploy by end of Y3Q1			
Story Points:				

Description
Add new object format(s) to dev, sandbox, stage, production once the new object format objects have been determined.

History

#1 - 2016-05-10 01:57 - Laura Moyers

- Assignee set to Chris Jones

- File griidcDataOneObjectFormats.xml added

- File griidcDataOneObjectFormats.txt added

Joe has sent a list of object formats in use at GRIIDC. Other formats may need to be added to the formats list in future, as identified by data submitters.

#2 - 2017-01-13 16:33 - Laura Moyers

- Status changed from New to Closed

- % Done changed from 0 to 100

No new object formats required.

#3 - 2017-01-13 17:14 - Chris Jones

- Assignee changed from Chris Jones to Jing Tao

Jing,

will you go through this list that Joe provided and add missing formats to the list? I see that the format names need to be more descriptive and compared with what's in PRONOM. Also, the @@ elements should be added, and the @@s should be vetted against the IANA list before we assign them.

Thanks!

#4 - 2017-01-13 17:15 - Laura Moyers

- % Done changed from 100 to 30

- Status changed from Closed to In Progress

Whoops, still verifying that some of the requested formats have been added.

#5 - 2017-01-13 23:32 - Jing Tao

We already have those format ids which exactly match the format ids in their list:

application/x-gzip
video/avi
application/x-bzip2
text/csv
application/msword
application/vnd.openxmlformats-officedocument.wordprocessingml.document
image/gif
application/x-gzip
image/jpeg
video/mp4
application/pdf
application/vnd.openxmlformats-officedocument.presentationml.presentation
application/x-rar-compressed
application/x-tar
text/plain
video/x-ms-wmv
application/vnd.ms-excel
application/vnd.openxmlformats-officedocument.spreadsheetml.sheet
application/xml
application/zip

#6 - 2017-01-13 23:39 - Jing Tao

They have the tab-separated-value in this way:

text/tab-separated-values
Tab Separated Values
DATA
.tsv

We have it in another way(different format id):

text/tsv
Tab Separated Values Text
DATA

tsv

#7 - 2017-01-13 23:43 - Jing Tao

They have this format id:

application/vnd.ms-excel.sheet.macroenabled.12
Microsoft Excel - Macro-Enabled Workbook
DATA
.xlsm

We don't have the exactly identical one, but a very close one. Can they use our one?

application/vnd.ms-excel.sheet.binary.macroEnabled.12

Microsoft Office Excel 2007 binary workbooks
DATA

xls

#8 - 2017-01-13 23:53 - Jing Tao

Their list has a general format id for netcdf:

application/x-netcdf
Network Common Data Form (NetCDF)
DATA
.nc

However, we has more specific netcdf version form ids:
from 1.0 to 1.4, 3 and 4.
Can they just use ours?

#9 - 2017-01-14 00:02 - Jing Tao

They have three items with the same format id:

application/octet-stream
NeuroExplorer
DATA
.nex

application/octet-stream
Origin Project
DATA
.opj

application/octet-stream
Sequence Read Archive
DATA
.sra

We have the general one:

application/octet-stream

Octet Stream
DATA

data

Can they just use our general one to replace the three ones?

#10 - 2017-01-14 00:10 - Jing Tao

Those are the format ids we don't have (totally 5 format ids):

application/x-7z-compressed
application/msaccess
magnus-internal/cgi
chemical/seq-aa-fasta (FASTA nucleic acid or protein sequences as character strings)
application/x-compressed (gzipped tar file)

Their current definitions are:

application/x-7z-compressed
7-Zip
DATA
.7z

application/msaccess
Microsoft Office Access 2007 Database
DATA
.accdb

magnus-internal/cgi
Common Gateway Interface
DATA
.cgi

chemical/seq-aa-fasta

FASTA nucleic acid or protein sequences as character strings
DATA
.fna

application/x-compressed
gzipped tar file
DATA
.tgz

#11 - 2018-03-06 15:26 - Mark Servilla

- % Done changed from 30 to 100

- Status changed from In Progress to Closed

No new formats have been identified since GRIIDC went into production.

Files

griidcDataOneObjectFormats.txt	11.3 KB	2016-05-10		Laura Moyers
griidcDataOneObjectFormats.xml	15.2 KB	2016-05-10		Laura Moyers