Kepler - Bug #5474

Make the eml actor on kepler handle multiple physical representations in an entity

08/19/2011 09:17 AM - Jing Tao

Status:	New		**Start date:**	08/19/2011
Priority:	Normal		**Due date:**	
Assignee:	David Welker		**% Done:**	0%
Category:	actors		**Estimated time:**	0.00 hour
Target version:	Unspecified		**Spent time:**	0.00 hour
Bugzilla-Id:	5474			

Description

Wade Sheldon from GCE-LTER brought an question if the eml actor can handle mutiple physical representations in an entity.

I took a look at code and found kepler can't handle this correctly. In org.ecoinformatics.seek.datasource.eml.eml2.Eml200Parser class, we always use

physical/dataFormat/textFormat/simpleDelimited/collapseDelimiters, physical/distribution/online/url/@function and et al directly. And also we only get the value from the node which index is 0.

This will handle entity with one physical representation correctly. But it will failed when an entity has more than one physical representation.

For example, a dataTable has two physicals. The first one isn't textFormat and has a downloadable url. The second one is textFormat and has a downloadable url.

Kepler will download the first non-text data object and ignore the second text data object. And the worst case is that kepler will mismatch the non-text data object with the text format information from the second text data object.

History

#1 - 08/19/2011 09:21 AM - Jing Tao

Matt suggests the possible solution:

probably collect all of the objects for an entity, then pick the best one for the app (e.g., text/csv over excel)

#3 - 03/27/2013 02:30 PM - Redmine Admin

Original Bugzilla ID was 5474

Files

knb-lter-gce.327.13.xml	40.7 KB	08/19/2011	Jing Tao