

Metacat - Bug #1984

add support for LSID identifiers

02/17/2005 05:50 PM - Matt Jones

Status:	Resolved	Start date:	02/17/2005
Priority:	Normal	Due date:	
Assignee:	Matt Jones	% Done:	0%
Category:	metacat	Estimated time:	0.00 hour
Target version:	2.1.0	Spent time:	0.00 hour
Bugzilla-Id:	1984		
Description			
<p>Metacat currently supports identifiers of the form 'scope.id.revision' with a 'system' attribute in the metadata. We need to modify metacat to support the Life Science Identifier specification (LSID). LSIDs have the form:</p> <p>urn:lsid:authority:namespace:localidentifier:revision</p> <p>This maps directly onto our existing scheme, and so there should be a one to one correspondence except that we fail to store the EML system attribute in our existing scheme. To support LSID fully, we need to:</p> <ol style="list-style-type: none">1) Accept EML documents that use an LSID in their packageId or other Id fields2) Accept LSID in the docid input parameters to the Metacat interfaces, including insert, update, delete, and read among others3) Implement an LSID resolver that is associated with Metacat that allows the standard LSID web services to respond with information about the document4) Allow LSIDs to be placed as identifiers within the "url" field of EML to reference the data objects that are associated with metadata documents <p>We probably also need to:</p> <ol style="list-style-type: none">5) Allow data objects to be inserted with LSID identifiers6) Allow LSID resolver calls for a dataset id to be able to locate the metadata associated with the data file in order to properly respond to the resolver request <p>The last point requires some new info to be tracked, because right now we maintain the linkage between metadata and data only in the EML documents -- we probably need to extract this when ingesting an eml document and store it for later retrieval.</p>			
Related issues:			
Blocked by Morpho - Bug #1985: add morpho support for LSID identifiers		New	02/17/2005

History

#1 - 02/17/2005 05:52 PM - Matt Jones

This probably also means that we need to be able to generate candidate LSIDs for clients to use for insertion, check LSIDs for uniqueness, and modify Morpho to use the LSID format for data file references (instead of the existing, obsolete 'ecogrid:' syntax).

#2 - 05/01/2007 10:50 AM - Matt Jones

Update on status: LSID support has been partially implemented. Use 'ant deploy-lsid' after building and installing metacat to install an LSID authority in the 'authority' context. This authority extracts a metacat docid from an LSID and uses the metacat client API to look up that document in the Metacat db. The metadata for the document is converted to a simple RDF format using an XSLT stylesheet and returned to the LSID client when requested using the LSID getMetadata call. The data is returned to the client when using the LSID getData call. For docids that represent actual data objects (type BIN in xml_documents), the actual data stream is returned. For docids that represent EML documents, the full EML document is returned. This is a bit confusing because for EML documents both getMetadata() and getData() return the same content (RDF for the former, and EML/XML for the latter).

A major outstanding problem is not being able to support more than one authority from a single metacat. Right now, metacat only tracks scope, id, and revision for documents. So two different authorities that issue the same scope.id.rev could in theory conflict with one another. To fully support LSIDs, we need to modify metacat to also track the authority, and then modify the metacat LSID resolver to make sure that authority and scope.id.rev match when responding to LSID client requests. This is a significant refactoring exercise because docids are used so pervasively in the metacat

model, often as primary and foreign keys.

#3 - 05/04/2007 12:17 PM - Matt Jones

Note: We need to improve the stylesheet that converts EML -> RDF. It is currently extremely minimalistic.

#4 - 05/04/2007 01:15 PM - Matt Jones

TDWG has proposed the use of a proxy service for LSID identifiers that allows LSIDs to be resolved using standard web browsers:

<http://wiki.tdwg.org/wiki/bin/view/GUID/LsidHttpProxyUsageRecommendation>

Given that Metacat has support for LSIDs built in, we should probably consider following their recommendations.

I tried out the proxy resolver that they put in place (<http://lsid.tdwg.org/>) -- my only complaint is that it is not particularly well formatted for the resources described, if only because its a generic service. It would be nice to consider how to include formatting/styling information in a way that allowed the proxy to do a better job presenting the information.

Here's a metacat LSID being resolved by their proxy:

Get the RDF: <http://lsid.tdwg.org/urn:lsid:esa.org:esa:8:7>

Get RDF in HTML: <http://lsid.tdwg.org/summary/urn:lsid:esa.org:esa:8:7>

Note the latter provides direct links to the GetMetadata and GetData endpoints that we provide from the esa.org LSID authority.

#5 - 01/24/2013 11:52 AM - ben leinfelder

Given our D1 API and DOI support, I believe we have gone a different direction with this. Please comment then close.

#6 - 01/24/2013 11:56 AM - Matt Jones

Agreed. LSIDs are dead. We need to maintain our ability to find existing content from the ESA skin based on its LSID, but we should stop creating any new ones, in favor of using the DOI system. Closing this bug as completed, in that we have done as much as we will. Another bug may be needed to determine a successional strategy away from LSIDs that we have already created.

#7 - 03/27/2013 02:18 PM - Redmine Admin

Original Bugzilla ID was 1984