Infrastructure - Task #693

Design the R client component(s)

2010-06-22 02:56 - Dave Vieglais

Status:	Closed		**Start date:**	
Priority:	High		**Due date:**	
Assignee:	Matthew Jones		**% Done:**	100%
Category:	d1_client_r		**Estimated time:**	0.00 hour
Target version:	CCI-0.4			
Milestone:	None		**Story Points:**	
Product Version:	*			

Description
Design the basic integration of a DataONE client library (either Java or Python) with R. Ensure that the necessary methods can be called from within R, and identify operations that should be contained in a higher level API (e.g. get = resolve + download + insert into R environment). Also issues such as how to support viewing potentially large result sets from search.

History

#1 - 2010-06-24 03:15 - Matthew Jones

Initial exploration and prototyping was done using the java binding to R called 'rJava'. Initial difficulties stemmed from two issues: 1) the version of rJava I was using had a broken .javaclassPath() function, which was fixed in the source version; 2) running rJava on the Mac seems to insist on running under JDK 1.5, which causes a binary incompatibility with classes compiled under Java 1.6. Targeting 1.5 makes this problem go away, but we'll need to work on making sure that the 1.6 can be used on MacOS. In the meantime, developing under linux is probably recommended.

Now that the basic framework for calling Java from R has been established, writing a simple client should largely be a matter of creating the right pass-through calls to Java.

#2 - 2010-06-25 01:54 - Matthew Jones

After more troubleshooting, I determined that the reason the D1Client class was not being called was not due to classpath issues -- rather, it was because import dependencies were not being met, but were being shown as NoClassDefFound errors. By running the .jconstructors() method in rJava, one can get a much better error message about why a class is not loading. After adding required jar files, I am now able to call d1-java classes. Now need to finish design of the client and work on the implementation in R. The prototype d1r module in ITK shows some basic methods for using the d1-java client methods that can be expanded upon.

#3 - 2010-07-17 20:37 - Matthew Jones

Added basic package design to the architecture docs. See http://dev-testing.dataone.org:8080/hudson/job/API%20Documentation/javadoc/itk-d1r.html for details. The current documentation specifies a bit more complete of an R package than will likely be completed by July 31. For example, I will likely only complete the CSV support, and it will be rudimentary at that; other data types like NetCDF and raster types will have to wait till phase 2.

#4 - 2010-08-26 00:36 - Matthew Jones

Basic R client implemented. Will open new tickets with additional new features and redesigns.