

Implementing ResourceSync for DSpace

Richard Jones, Cottage Labs, richard@cottagelabs.com
Stuart Lewis, University of Edinburgh, stuart.lewis@ed.ac.uk

Abstract: Through UKOLN, members of the UK repository community have been working on the ResourceSync technical advisory board since 2012 to assist the core team in developing the standard. In addition, we were funded to develop code which demonstrated the effectiveness of the standard, and we chose to implement the OAI-PMH use case for DSpace using ResourceSync. This work is presented here, with a demonstration of the resulting DSpace module.

The ResourceSync specification describes a synchronization framework for the web consisting of various capabilities that allow third party systems to remain synchronized with a server's evolving resources. The capabilities can be combined in a modular manner to meet local or community requirements. The specification also describes how a server can advertise the synchronization capabilities it supports and how third party systems can discover this information. [1]

Through UKOLN [2], members of the UK repository community have been working on the ResourceSync technical advisory board since 2012 to assist the core team in developing the standard. Jisc[3] funded a number of participants to attend meetings and spend time considering the work, and then allocated additional budget for software development to prove the value of the resulting standard.

For the software development part of the project it was decided to carry out the following tasks, as far as was possible:

1. *Implement the OAI-PMH [4] use case with the ResourceSync standard.* This would be valuable because ResourceSync is also an Open Archives Initiative project, and one of the stated goals was to subsume and extend on the PMH domain. OAI-PMH is critical infrastructure for the repository community, as it fulfills a key use case. By doing this we would show fitness for purpose of the standard to meet that use case, and to extend beyond this into further useful areas such as full-text exchange.
2. *Develop a client environment which can be used to demonstrate that the ResourceSync development can be properly tested.* In order to make this more than just test code, we decided to develop a Pull-to-Push converter which consumes data from a ResourceSync pull-based endpoint, and re-packages the content using SWORDv2 [5] to push into another repository. This was to demonstrate the functionality of the ResourceSync service on DSpace, and show how ResourceSync and SWORDv2 are complementary technologies.

3. *Implement a proof-of-concept ResourceSync server endpoint against DSpace, supporting all of the features.* ResourceSync is a fairly substantial standard, so we proposed not to attempt to develop all the features in detail within the timeline of our project. Nonetheless, since ResourceSync is a highly modular standard, even a partial/incomplete implementation is valuable in testing the specification.

From the project we will produce a number of useful artifacts:

1. A documented understanding of the process of implementing ResourceSync in a real environment
2. A pattern for use of ResourceSync to replace OAI-PMH; how are Sets represented and used, how is metadata-only transferred, etc.
3. General Java-based code libraries for supporting servers in implementing ResourceSync
4. General PHP-based client libraries for supporting clients in implementing ResourceSync (and an integration of these libraries with the PHP-based SWORDv2 libraries [6])
5. A DSpace-specific ResourceSync web-app which can be used as an add-on module, and which may ultimately be suitable for distribution with the main DSpace codebase

At time of writing (March 2013) ResourceSync is approaching completion and the software development part of the project has just begun. By Open Repositories 2013 this will have been completed and we will come to the event with the final code outputs from the project, and a working demonstration.

The presentation will include a brief introduction to ResourceSync and discussion of how the terminology and features map onto DSpace. We will look at the requirements that we needed to implement in order to effectively meet the same use cases as OAI-PMH. There will be a demonstration of the code in action, both exposing data from DSpace using ResourceSync, and transferring it to another repository via SWORDv2.

We will conclude by discussing future development needed on the DSpace module, and also the opportunities for developing a client embedded into DSpace that can consume from a ResourceSync enabled service.

[1] ResourceSync specification: <http://www.openarchives.org/rs/0.5/resourcesync>

[2] UKOLN: <http://www.ukoln.ac.uk/>

[3] Jisc: <http://www.jisc.ac.uk>

[4] OAI-PMH: <http://www.openarchives.org/pmh/>

[5] SWORDv2: <http://swordapp.org>

[6] SWORDv2 PHP client library: <https://github.com/swordapp/swordappv2-php-library/>