LDP4ROs: Managing Research Objects with the W3C Linked Data Platform

Daniel Garijo
Ontology Engineering Group
Universidad Politécnica de Madrid
dgarijo@fi.upm.es

Nandana Mihindukulasooriya
Ontology Engineering Group
Universidad Politécnica de Madrid
nmihindu@fi.upm.es

Oscar Corcho
Ontology Engineering Group
Universidad Politécnica de Madrid
ocorcho@fi.upm.es

ABSTRACT

In this demo we present LDP4ROs, a prototype implementation that allows creating, browsing and updating Research Objects (ROs) and their contents using typical HTTP operations. This is achieved by aligning the RO model with the W3C Linked Data Platform (LDP).

Keywords

Linked Data Platform, Research Object, LDP, RO

1. INTRODUCTION

It is widely recognised that raw PDF files are usually insufficient to allow researchers to check, reuse and reproduce the methods exposed in scientific publications[1]. In this context, the Research Object model (RO) [1] has been proposed as a possible means to aggregate, link and describe the resources (e.g., code, scripts, datasets and references to them) associated to scientific publications, what may help (although not necessarily solve completely) these tasks.

Different applications have been developed to create, manage and access ROs[2] [4]. However, they rely on custom services and APIs, making it difficult to access, edit or create some of their resources and contents using standard HTTP operations, affecting interoperability among them.

In this demo we present LDP4ROs[3], a prototype implementation of our alternative approach to create, access and update ROs and their associated resources, which considers them as Web objects. For this purpose we align the RO model to the W3C Linked Data Platform (LDP) [5]. With our approach, one can create, edit or access resources as Linked Data using standard HTTP operations (e.g., GET plus content negotiation) instead of using SPARQL or custom REST APIs.

[1]https://www.force11.org/beyondthepdf2
[2]http://www.researchobject.org/specifications/
[3]https://github.com/oeg-upm/LDP4RO

SAVE-SD 2015 WWW, Florence, Italy

Section 2 describes the alignment between LDP and the RO model, Section 3 describes our working prototype and Section 4 discusses future work.

2. ALIGNING LDP WITH THE RO MODEL

This section describes the mapping between ROs to LDP concepts. First we describe the main features of these models (Sections 2.1 and 2.2) and then we illustrate the results of the mapping with a simple example (Section 2.3).

2.1 The Research Object Model

An RO can be defined as a bundle that aggregates a number of resources that are used or produced during a scientific investigation [1]. The RO ontology[4] extends the Object Reuse and Exchange model (ORE)[5] as shown in Figure 1: an ro:ResearchObject extends ore:Aggregation, and is described by an ro:Manifest (which extends the ore:ResourceMap concept). Finally, each aggregated ro:Resource extends the ore:AggregatedResource concept.

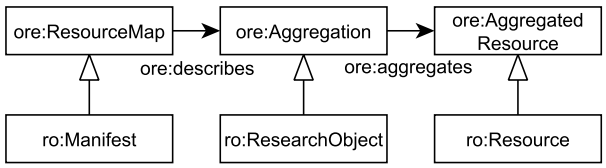

Figure 1: The RO model extending the ORE model

2.2 The Linked Data Platform

The Linked Data Platform (LDP) [5] aims to provide an architecture for read-write Linked Data on the web. LDP describes aggregations of linked documents as LDP containers (LDPC). There are three types of containers: Basic, Direct, and Indirect containers, which provide different levels of flexibility:

- **Basic Containers** maintain a containment triple. It follows a fixed pattern i.e., it will always have the LDPC itself as the subject, ldp:contains as the predicate, and the resource created using the container as the object.

[4]http://purl.org/wf4ever/ro
[5]http://www.openarchives.org/ore/1.0/toc.html

- **Direct Containers** specialize Basic Containers by introducing another set of triples managed by the container, membership triples. These membership triples provide more flexibility by allowing the subject and the predicate of the triple to be configured using the container definition.
- **Indirect Containers** maintain both containment triples and membership triples as the previous two types of containers but provide more flexibility by allowing the object of the membership triple to be configured too. By doing so, the resources to be created can provide any URI to be used as the object of the membership triple

2.3 From ROs to Web Objects

Both the RO model and LDP contain concepts for defining resource aggregations (ro:ResearchObject and ldp:Container). Out of the three types of LDP Containers, we decided to map ROs to the LDP Direct Container. We found it to be the most suitable type of container for the mapping, as the Basic container does not allow customizing the ldp:contains relationship, and the indirect container added too much complexity for our use case.

Another detail to take into consideration is the ro:Manifest (the entity describing the ore:Aggregation), which does not have a mapping to any LDP concept. LDP merges the description of the container with the container itself, which is not allowed in the ORE model. However, this issue can be avoided by following the ORE user guide, which proposes using hash URIs for defining the aggregations[6]. An example is shown below:

```
@prefix ore:<http://www.openarchives.org/ore/terms/> .
@prefix ro: <http://purl.org/wf4ever/ro#> .
@prefix ldp: <http://www.w3.org/ns/ldp#> .
@base <http://example.org/ro/> .

<#agr> a ro:ResearchObject, ore:Aggregation,
        ldp:Container, ldp:DirectContainer;
    ldp:membershipResource <#agr>;
    ldp:hasMemberRelation ore:aggregates;
    ldp:contains <res1>, <res2>, <res3>;
    ore:aggregates <res1>, <res2>, <res3>,
            <http://example.org/data.xlsx>;
    ore:isDescribedBy <> .

<> a ore:ResourceMap .
```

In the example, a Research Object (<#agr>) aggregates four resources. Three of those resources (<res1>, <res2> and <res3>) have been created directly by POSTing the resources to the LDP container, while the fourth one (data.xlsx) is an external resource. All the descriptions about the RO are contained in the resource map (<>). Therefore, both resource map and Research Object are resolved against the same URL, but they have different identifiers (<> and <#agr> respectively). The full specification with the mapping, examples and use cases can be accessed online[7].

3. LDP4ROS

We have developed a simple Web application[8] for creating and browsing ROs. After creating an RO with the online application, a user is able to access it or its resources directly with their URL, being able to obtain different representations through content negotiation. Figure 2 shows a picture of the architecture used in the prototype, based on GET and POST requests to handle the data. In some cases, intermediate servlets have been added to handle file uploads and feed the proper URIs to LDP.

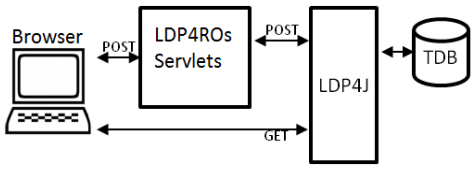

Figure 2: Architecture of the prototype

Regarding the implementation details, the frontend uses the Bootstrap[9] framework and JSPs to handle data and transform it for LDP. As a backend, we use the LDP4J application [2], so as to manage ROs as LDP Containters, and the Jena TDB triplestore[10] to store RDF data.

4. CONCLUSIONS AND FUTURE WORK

In this demo paper we have introduced a working prototype (LDP4ROs) of our alignment between the Research Object model and the Linked Data Platform. The prototype allows creating Research Objects and exploring existing ones by using typical HTTP operations.

As future work, we plan on expanding the prototype to handle folders and other serializations (e.g., the RO bundle[11]), to allow editing ROs on the fly and storing RO resources in popular repositories like FigShare [3].

5. ACKNOWLEDGMENTS

This work is partially supported by an FPU grant (Formación de Profesorado Universitario) from the Spanish Ministry of Science and Innovation (MICINN).

6. REFERENCES

[1] K. Belhajjame, O. Corcho, D. Garijo, J. Zhao, et al. Workflow-Centric Research Objects: First Class Citizens in Scholarly Discourse. In *SePublica2012 workshop at ESWC2012*, 2012.

[2] M. Esteban-Gutiérrez, N. Mihindukulasooriya, and R. García-Castro. LDP4j: A framework for the development of interoperable read-write Linked Data applications. In *Proceedings of the 1st ISWC Developers Workshop*, Riva del Garda, Italy, Oct 2014.

[3] J. Singh. Figshare. *Journal of pharmacology & pharmacotherapeutics*, 2(2):138, 2011.

[4] S. Soiland-Reyes, D. Cruickshank, F. Bacall, J. Zhao, K. Belhajjame, D. D. Roure, and C. A. Goble. myExperiment Research Objects: Beyond Workflows and Packs. Technical report, 2013.

[5] S. Speicher, J. Arwe, and A. Malhotra. Linked Data Platform 1.0 W3C Recommendation 26 February 2015. Technical report, W3C, 2014.

[6] http://www.openarchives.org/ore/1.0/http#Simple
[7] http://purl.org/net/ldp4ro/spec
[8] http://purl.org/net/ldp4ro
[9] http://getbootstrap.com/
[10] http://jena.apache.org/documentation/tdb/
[11] https://w3id.org/bundle/