ORE Specification - Resource Map Implementation in JSON-LD

2014-08-14

This version: http://www.openarchives.org/ore/0.9/jsonld
Latest version: http://www.openarchives.org/ore/jsonld

Editors
Simeon Warner, Cornell University
Oscar Corcho, Universidad Politécnica de Madrid

Contributors

Abstract

Open Archives Initiative Object Reuse and Exchange (OAI-ORE) defines standards for the description and exchange of aggregations of web resources. OAI-ORE introduces the notion of a Resource Map, an RDF Graph [RDF Concepts] which describes the Aggregation, the aggregated Resources of which it is composed, and the relationships between them (and/or the relationships between these and other resources).

Since a Resource Map is an RDF Graph, it can be serialized using any RDF syntax. This document outlines the use of one such syntax for the serialization of Resource Maps: [JSON-LD] (ORE JSON-LD Frame). Two companion documents [ORE RDFa] and [ORE RDF/XML Documentation] outline the serialization of Resource Maps in RDFa Syntax and RDF/XML Syntax respectively.

This document is intended for implementers who have an understanding of ORE concepts and are responsible for the development of applications which generate or process Resource Maps using [JSON-LD]. Readers who want a high-level understanding of the motivation for ORE, and of the solution it provides, should read the Primer [ORE Primer].

This is one of several documents comprising the OAI-ORE specifications and user guides.

This is a beta specification released for public comment. Please send feedback to the OAI-ORE Google Group.

Table of Contents

1. Introduction
2. Overview of ORE in JSON-LD syntax
3. ORE JSON-LD context
4. Complete example
5. ORE JSON-LD frame
6. References
B. Change Log

1. Introduction

Open Archives Initiative Object Reuse and Exchange (OAI-ORE) defines standards for the description and exchange of aggregations of web resources. The OAI-ORE Abstract Data Model [ORE Model] introduces the notion of a Resource Map, an RDF Graph [RDF Concepts] which describes the Aggregation, the aggregated Resources of which it is composed, and the relationships between them (and/or the relationships between these and other resources). A Resource Map is an RDF Graph and can thus be serialized using any RDF syntax. This document outlines the use of the [JSON-LD] syntax for serialization of Resource Maps.

The [JSON-LD] specification describes a format for serialising RDF graphs as [JSON [RFC4627]]. In short, the [JSON-LD] format allows the use of [JSON Objects] (maps/dictionaries) to represent an RDF resource. The [JSON-LD] reserved key jsonld defines the IRI for the resource, and [jsonld] defines the rdf:type (class of the resource). Other RDF properties are defined using additional keys on the resource, and can point to either other [JSON Objects], [JSON Lists] (multiple property values) or to plain JSON literals. The [JSON-LD Context] can define prefixes and short-hand names, thus avoiding the need to use full IRIs for identifying properties and types. This specification includes a [ORE JSON-LD context] for this purpose.

JSON-LD allows for several equivalent representations of the same graph triples, and a compliant ORE [JSON-LD] document might therefore not follow the JSON structure exactly as listed in the examples of this document. See the Framing section for measures to shape an ORE [JSON-LD] document into a predictable JSON structure.

1.1 Notational Conventions

The keywords "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119]. This specification uses the common term URI to mean both IRI [RFC3987] and URI [RFC3986]. JSON-LD natively supports IRIs without any special measures.

1.2 Namespaces

This specification uses the following namespaces and prefixes:

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Namespace URI</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ore</td>
<td><a href="http://www.openarchives.org/ore/terms">http://www.openarchives.org/ore/terms</a></td>
<td>ORE vocabulary</td>
</tr>
<tr>
<td>rdfs</td>
<td><a href="http://www.w3.org/2000/01/rdf-schema#">http://www.w3.org/2000/01/rdf-schema#</a></td>
<td>RDF schema vocabulary</td>
</tr>
<tr>
<td>foaf</td>
<td><a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/</a></td>
<td>FOAF vocabulary</td>
</tr>
<tr>
<td>fabio</td>
<td><a href="http://purl.org/spar/fabio/">http://purl.org/spar/fabio/</a></td>
<td>Fabio, the FRBR-aligned Bibliographic Ontology</td>
</tr>
<tr>
<td>dcmitype</td>
<td><a href="http://purl.org/dc/dcmitype/">http://purl.org/dc/dcmitype/</a></td>
<td>Dublin Core DCMI types</td>
</tr>
<tr>
<td>dcterms</td>
<td><a href="http://purl.org/dc/terms/">http://purl.org/dc/terms/</a></td>
<td>Dublin Core Terms</td>
</tr>
<tr>
<td>owl</td>
<td><a href="http://www.w3.org/2002/07/owl#">http://www.w3.org/2002/07/owl#</a></td>
<td>OWL vocabulary</td>
</tr>
<tr>
<td>jsonld</td>
<td><a href="http://www.json-ld.org/0.9/syntax">http://www.json-ld.org/0.9/syntax</a></td>
<td>JSON-LD vocabulary</td>
</tr>
<tr>
<td>pydr</td>
<td><a href="http://pydr.org/api/">http://pydr.org/api/</a></td>
<td>PyDr Framework</td>
</tr>
<tr>
<td>frbr</td>
<td><a href="http://www.lib.rpi.edu/FRBR">http://www.lib.rpi.edu/FRBR</a></td>
<td>FRBR vocabulary</td>
</tr>
<tr>
<td>rdfs</td>
<td><a href="http://www.w3.org/2000/01/rdf-schema#">http://www.w3.org/2000/01/rdf-schema#</a></td>
<td>RDF schema vocabulary</td>
</tr>
<tr>
<td>jsonld</td>
<td><a href="http://www.json-ld.org/0.9/syntax">http://www.json-ld.org/0.9/syntax</a></td>
<td>JSON-LD vocabulary</td>
</tr>
</tbody>
</table>
1.3 Conformance

The Internet Media Type (MIME type) for JSON-LD is `application/ld+json`. Where JSON-LD is used to provide representations of ORE Resource Maps, that media type SHOULD be used. For backwards compatibility, the JSON-LD specification allows for the alternative content-type of `application/json`, in which case the JSON-LD context Link header MUST be present.

A consumer of ORE Resource Maps encoded in JSON-LD SHOULD support all of the syntactic conventions defined by JSON-LD. This allowance for different syntactic conventions is supported by conformant JSON-LD parsers (see JSON-LD Community).

2. Overview of ORE in JSON-LD syntax

An ORE Resource Map expressed as a JSON-LD resource MUST be valid JSON [RFC4627], and MUST be valid JSON-LD [JSON-LD].

An ORE JSON-LD resource MUST define a JSON-LD @context, which SHOULD be `https://w3id.org/ore/context` (see ORE JSON-LD context below).

The @type SHOULD be provided for ORE resources of type Aggregation, ResourceMap, Proxy, and MAY be provided for resources of type AggregatedResource and others.

ResourceMap, describes, Aggregation

The ResourceMap SHOULD be the top-level JSON object (indicating that the JSON-LD document is the resource map), in which case it MUST indicate which Aggregation it describes:

```json
{  "@context": "https://w3id.org/ore/context",
  "@type": "ResourceMap",
  "@id": 
  "describes": {
    "@id": "http://example.com/aggregation-1",
    "@type": "Aggregation"
  }
}
```

The Aggregation MUST have an @id to identify it; this URI SHOULD be absolute. The ResourceMap MAY have an identifier using @id, which MUST be different from the Aggregation URI.

isDescribedBy

The Aggregation MAY list other resource maps it isDescribedBy, which SHOULD be given as a JSON list of URIs:

```json
{  "@context": "https://w3id.org/ore/context",
  "@type": "ResourceMap",
  "describes": {
    "@id": "http://example.com/aggregation-1",
    "@type": "Aggregation",
    "isDescribedBy": [
      "http://example.com/aggregation-1.rdf",
      "http://example.com/aggregation-1.jsonld"
    ]
  }
}
```

similarTo

If an Aggregation can be considered similar to another resource, like a journal article identified with a DOI, then this MAY be indicated using similarTo:

```json
{  "@context": "https://w3id.org/ore/context",
  "@type": "ResourceMap",
  "describes": {
    "@id": "http://example.com/aggregation-1",
    "@type": "Aggregation",
    "similarTo": "http://dx.doi.org/10.1002/cpe.1594"
  }
}
```

aggregates

All resources aggregated by this Aggregation MUST be listed using aggregates, which SHOULD be given as a JSON list. The list MUST contain either URI strings or JSON Objects, which SHOULD specify their URI with @id:

```json
{  "@context": "https://w3id.org/ore/context",
  "@type": "ResourceMap",
  "describes": {
    "@id": "http://example.com/aggregation-1",
    "@type": "Aggregation",
    "aggregates": [  
      "http://example.com/document-1",
      {  "@id": "http://other.example.org/data-2",
        "@type": "AggregatedResource"
      }
    ]
  }
}
```

Aggregated resources MAY specify their @type: "AggregatedResource". If a more specific type of the resource is known (e.g. `dcmitype:MovingImage` or `fabio:MagazineNewsItem`), it is RECOMMENDED to specify this using @type and the full URI of the class. A JSON List MAY be used to indicate multiple types.

Example:

```json
{  "@context": "https://w3id.org/ore/context",
  "@type": "ResourceMap",
  "describes": {
    "@id": "http://example.com/aggregation-1",
    "@type": "Aggregation",
    "aggregates": [  
      "http://example.com/document-1",
      {  "@id": "http://other.example.org/data-2",
        "@type": ["AggregatedResource",
                    "http://purl.org/dc/dcmitype/Dataset",
                    "http://www.w3.org/ns/dcat#Dataset"],
        "http://www.openarchives.org/ore/0.9/jsonld"}
Additional descriptions of aggregated resources MAY be provided using other RDF vocabularies following the JSON-LD syntax. The example below shows how to indicate the provenance of an aggregated resource using PROV-O (PROV) and PAV (PAV). When using third-party vocabularies it may be desirable to add prefixes using additional contexts as a list.

```json
{
  "@context": [
    "https://w3id.org/ore/context",
    {
      "foaf": "http://xmlns.com/foaf/0.1",
      "schema": "http://schema.org",
      "prov": "http://www.w3.org/ns/prov",
      "pav": "http://purl.org/pav"
    }
  ],
  "@type": "ResourceMap",
  "describes": {
    "@id": "http://example.com/aggregation-1",
    "@type": "Aggregation",
    "aggregates": [
      { "@id": "http://example.com/document-1",
        "@type": "schema:BlogPosting",
        "prov:wasDerivedFrom": { "@id": "http://example.com/document-1" },
        "pav:authoredBy": { "@id": "http://orcid.org/0000-0001-9842-9718", "foaf:name": "Stian Soiland-Reyes" }
      }
    ]
  }
}
```

isAggregatedBy, isDescribedBy

If a resource is also part of a different Aggregation, then this can be indicated with isAggregatedBy:

```json
{
  "@context": "https://w3id.org/ore/context",
  "@type": "ResourceMap",
  "describes": {
    "@id": "http://example.com/aggregation-1",
    "@type": "Aggregation",
    "aggregates": { 
      { 
        "@id": "http://other.example.org/data-2",
        "isAggregatedBy": { 
          "@id": "http://other.example.org/aggregation-2",
          "@type": "Aggregation"
        }
      }
    }
  }
}
```

It is NOT RECOMMENDED to detail the aggregates of the other Aggregation, but one MAY use isDescribedBy to indicate its ResourceMap representations.

In the special case of aggregating another Aggregation, that SHOULD be typed as an Aggregation, and there SHOULD be at least one ResourceMap representation indicated with isDescribedBy:

```json
{
  "@context": "https://w3id.org/ore/context",
  "@type": "ResourceMap",
  "describes": {
    "@id": "http://example.com/aggregation-1",
    "@type": "Aggregation",
    "aggregates": { 
      { 
        "@id": "http://other.example.org/aggregation-2",
        "isDescribedBy": [ "http://other.example.org/aggregation-2.rdf", "http://other.example.org/aggregation-2.jsonld" ]
      }
    }
  }
}
```

proxies, proxyFor, proxyIn

A Proxy represents an AggregatedResource as it exists in a specific Aggregation, which can be used as a subject in other annotations and aggregations. Proxies SHOULD be provided for each aggregated resource:

```json
{
  "@context": "https://w3id.org/ore/context",
  "@type": "ResourceMap",
  "describes": {
    "@id": "http://example.com/aggregation-1",
    "@type": "Aggregation",
    "aggregates": { 
      { 
        "@id": "http://example.com/document-1"
      }
    }
  }
}
```

The property proxies SHOULD be a JSON List of resources, each of which SHOULD have "@type": "Proxy" and MUST identify the aggregated resource using proxyFor. The resource identified by proxyFor MUST be listed under aggregates.

The proxy MUST be identified with an URI using @id, which MAY be a randomly generated UUID (UUID) using the urn:uuid: scheme.

Note that, unlike in other serializations of ORE resource maps, the property proxies SHOULD NOT be given for proxies listed under proxies (it is already given as a reverse property for proxyIn); however in the special case of aggregating a proxy, that MUST specify proxyIn, which MUST be a different Aggregation:

```json
{
  "@context": "https://w3id.org/ore/context",
  "@type": "ResourceMap",
  "describes": {
    "@id": "http://example.com/aggregation-1",
    "@type": "Aggregation",
    "aggregates": { 
      { 
        "@id": "urn:uuid:06e63599-d28d-4966-8006-db9985a65f12",
        "@type": "Proxy",
        "proxyFor": "http://example.com/document-1"
      }
    }
  }
}
```
Such "external" proxies MUST NOT be listed under proxies as a proxy can only be proxyFor for a single Aggregation, and an aggregation SHOULD NOT aggregate its own proxies. The proxyIn MAY be expanded to a JSON Object with @id and @type: "Aggregation" to provide lineage and provenance of proxy

If an aggregated resource was discovered in a different aggregation, this lineage can be described between the corresponding proxies:

Other kinds of provenance of a proxy (e.g. who aggregated the given resource) MAY be stated using vocabularies like PROV-O and PAV. The example below says that Matthew aggregated document-1 (e.g. created its proxy in this aggregation) on the 17th of May 2014 (without implying that he created document-1):
3.2 Content of ORE JSON-LD context

The content of the ORE JSON-LD context at time of writing is:

```json
{
  "@context": {
    "Proxy": "http://www.openarchives.org/ore/terms/Proxy",
    "proxyFor": {
      "@id": "http://www.openarchives.org/ore/terms/proxyFor",
      "@type": "@id"
    },
    "lineage": {
      "@id": "http://www.openarchives.org/ore/terms/lineage",
      "@type": "@id"
    },
    "describes": {
      "@id": "http://www.openarchives.org/ore/terms/describes",
      "@type": "@id"
    },
    "AggregatedResource": {
      "@id": "http://www.openarchives.org/ore/terms/AggregatedResource"
    },
    "ResourceMap": {
      "@id": "http://www.openarchives.org/ore/terms/ResourceMap"
    },
    "similarTo": {
      "@id": "http://www.openarchives.org/ore/terms/similarTo",
      "@type": "@id"
    },
    "Aggregation": {
      "@id": "http://www.openarchives.org/ore/terms/Aggregation"
    },
    "isAggregatedBy": {
      "@id": "http://www.openarchives.org/ore/terms/isAggregatedBy",
      "@type": "@id"
    },
    "proxies": {"@reverse": "http://www.openarchives.org/ore/terms/proxyIn"}
  }
}
```

4. Complete example

The following complete example shows a ResourceMap describing aggregation-1 containing 4 resources (document-1, data-2, aggregation-3 and a proxy for document-1 within aggregation-3). data-2 is also known to be aggregated by aggregation-2.

Finally our aggregation has its own set of identified proxies, one for each aggregated resource. Here, the lineage of data-2's aggregation is shown to be originating from aggregation-2.

4.1 Example JSON-LD

```json
{
  "@context": "https://w3id.org/ore/context",
  "@id": ",
  "@type": "ResourceMap",
  "describes": {
    "@id": "http://example.com/aggregation-1",
    "@type": "Aggregation",
    "isDescribedBy": ["http://example.com/aggregation-1.rdf",
                     "http://example.com/aggregation-1.jsonld"],
    "aggregates": {
      "@id": "http://example.com/document-1",
      "@type": "Proxy",
      "proxyFor": "http://example.com/document-1",
      "proxyIn": "http://other.example.org/aggregation-3"
    }
  },
  "proxies": {
    "@id": "urn:uuid:09561248-bf55-4c85-930a-9a7a60e81602",
    "@type": "Proxy",
    "proxyFor": "http://example.com/document-1",
    "proxyIn": "http://other.example.org/aggregation-3"
  }
}
```
For example, say we have a

5.1 Example framing

By facilitating structures than the one defined in this document. This could particularly be the case if converting an ORE Resource Map to JSON-LD using regular RDF libraries.

JSON-LD, as all RDF serialization, allow many degrees of freedom for representing the same RDF graph. Thus, an ORE Resource Map could appear in other JSON

If processed with the ORE JSON-LD context and using a JSON-LD library to generate RDF, the following triples are expressed:

\[
\text{\textbf{ORE JSON-LD frame}}
\]

5. ORE JSON-LD frame

JSON-LD, as all RDF serialization, allow many degrees of freedom for representing the same RDF graph. Thus, an ORE Resource Map could appear in other JSON structures than the one defined in this document. This could particularly be the case if converting an ORE Resource Map to JSON-LD using regular RDF libraries. By facilitating JSON-LD Framing [JSON-LD Framing] with the ORE JSON-LD Frame, followed by JSON-LD Compaction, a predictable JSON structure is returned that can be used by consumers without requiring any further JSON-LD or RDF knowledge.

Note: At time of writing, JSON-LD Framing a draft specification which is still subject to change. This section is therefore to be considered as experimental. The ORE JSON-LD Frame at [http://www.openarchives.org/ore/0.9/jsonld] might subsequently be updated independently of this document in order to be compliant with later updates of the JSON-LD Framing specification.

The ORE JSON-LD frame is available at https://w3id.org/ore/frame, which at time of writing is:

\[
\text{\textbf{ORE JSON-LD Frame}}
\]

5.1 Example framing

For example, say we have a "unstructured" JSON-LD ORE manifest after serializing a Resource Map as triples in JSON-LD:

\[
\text{\textbf{ORE Specification - Resource Map Implementat... http://www.openarchives.org/ore/0.9/jsonld}
\]
If we expand this JSON-LD document using https://w3id.org/ore/context, then frame it using https://w3id.org/ore/frame we get the framed JSON:

```json
{
  "@context": "https://w3id.org/ore/context",
  "@graph": [
    {
      "@id": "index.html",
      "@type": "ResourceMap",
      "http://www.openarchives.org/ore/terms/describes": [
        {
          "@id": "http://example.com/aggregation-1"
        }
      ],
      "http://www.openarchives.org/ore/terms/isAggregatedBy": [
        {
          "@id": "urn:uuid:360ea070-efee-11e3-ac10-0800200c9a66"
        },
        {
          "@id": "urn:uuid:318745e0-5ea4-4a68-8a0a-ab27ab60e82a"
        },
        {
          "@id": "urn:uuid:360ea070-efee-11e3-ac10-0800200c9a66"
        }
      ],
      "http://www.openarchives.org/ore/terms/proxyFor": [
        {
          "@id": "http://example.com/aggregation-1"
        },
        {
          "@id": "http://example.com/aggregation-1"
        },
        {
          "@id": "http://example.com/aggregation-1"
        },
        {
          "@id": "http://example.com/aggregation-1"
        },
        {
          "@id": "http://example.com/aggregation-1"
        }
      ],
      "http://www.openarchives.org/ore/terms/proxyIn": [
        {
          "@id": "http://example.com/aggregation-1"
        }
      ]
    }
  ]
}
```
Note that JSON-LD Framing always gives a \@graph JSON List in case there are several Resource Maps that match the pattern. The properties \texttt{aggregates} and \texttt{isDescribedBy} if present, will after processing always be a JSON list, even if there is only a single resource.

5.2 Framing proxies

Note: \texttt{@reverse} properties like \texttt{proxies} is not at the time of writing commonly supported by JSON-LD Framing processors, and therefore have to be extracted separately using this frame for proxies:

```json
{
  "@context": "https://w3id.org/ore/context",
  "@type": "Proxy",
  "proxyFor": {
    "@embed": false
  },
  "proxyIn": {
    "@embed": false
  },
  "lineage": {
    "@embed": false,
    "@omitDefault": true
  }
}
```

When the above frame is processed against the earlier example it will return a list of all the proxies (including those in other aggregations):

```json
{
  "@context": "https://w3id.org/ore/context",
  "@graph": {
    "@id": "urn:uuid:05bd5e0c-94c7-4856-a53f-7f6cf8756751",
    "@type": "Proxy",
    "proxyFor": "http://other.example.org/data-2",
    "proxyIn": "http://example.com/aggregation-1"
  },
  "@id": "urn:uuid:09561248-bf55-4c85-93ba-9a7ad6b81682",
  "@type": "Proxy",
  "proxyFor": "http://example.com/document-1",
  "proxyIn": "http://other.example.org/aggregation-3",
  "lineage": {
    "@embed": false,
    "@omitDefault": true
  }
}
```

6. References

[JSON-LD]


Latest version at http://www.w3.org/TR/json-ld/
A. Acknowledgements
The Open Archives Initiative Object Reuse and Exchange specifications are the work of the Open Archives Initiative. Work on this document by Stanisław-Reyes was funded by the EU FP7 project BioVeL (FP7 RI-283359). Funding for Open Archives Initiative Object Reuse and Exchange was provided by the Andrew W. Mellon Foundation, Microsoft, the National Science Foundation, and the Coalition for Networked Information.

B. Change Log

<table>
<thead>
<tr>
<th>Date</th>
<th>Editor</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014-08-14</td>
<td>simeon</td>
<td>Integrate with other OAI-ORE specs. v0.9</td>
</tr>
<tr>
<td>2014-08-05</td>
<td>simeon</td>
<td>Fixing syntax issues</td>
</tr>
<tr>
<td>2014-08-05</td>
<td>simeon</td>
<td>First draft</td>
</tr>
</tbody>
</table>

See also the current editors' draft. For detailed changelog, see the Github commits.

This work is licensed under a Creative Commons Attribution-Share Alike 3.0 Unported License.