

# Serializing and Exposing Resource Maps

## Abstract

*Serializing and exposing Resource Maps is where the “rubber meets the road” in ORE. As described in chapter 1, to serialize a Resource Map essentially means to write it down in a format that can be transmitted and read by a machine. The ORE specification documents three different formats for Resource Map serialization: RDF/XML, RDFa, and Atom XML, although there are others.<sup>1</sup> In this chapter of “Object Reuse and Exchange,” we will continue the example of the National Digital Newspaper Program and demonstrate the serialization of Resource Maps in RDF/XML for the RDF Graphs of an Aggregation and its Resource Map, metadata, Aggregated Resources, and a nested Aggregation. In addition to RDF/XML, RDFa and Atom XML serialization will be introduced. Strategies for exposing Resource Maps using the World Wide Web architecture and HTTP will be explored, including 303 redirection (with and without content negotiation), hash URIs, and RDFa. Lastly, we will discuss some mechanisms that are suggested by the ORE specification for enabling the discovery of Resource Maps, including batch discovery using protocols such as the OAI-PMH and resource embedding in webpages.*

## Serializing a Resource Map as RDF/XML

A Resource Map describes an Aggregation and its Aggregated Resources and relationships, which we explored in the previous chapter as RDF Graphs. RDF Graphs can be expressed as RDF/XML by translating the triples represented by the circles and arrows into XML. The subjects, predicates, and objects of the RDF triples are represented in RDF/XML as XML elements and

attributes with their associated names and values. The syntax of RDF/XML allows triples to be expressed in different ways, so the same Resource Map can be expressed differently in RDF/XML but contain the same semantics.<sup>2</sup>

More detailed recommendations and syntax can be found in the “ORE User Guide: Resource Map Implementation in RDF/XML.” To demonstrate some of the more common elements, we will iterate the triples and then serialize the RDF Graphs from chapter 2 for page and issue Aggregations from the National Digital Newspaper Program’s implementation, which uses RDF/XML. It may be helpful to flip back to the graphs and reference them with the Resource Maps as we construct them.

## Aggregations and Resource Maps

We begin with an Aggregation for a newspaper page and its Resource Map. The same requirements for the graphed representations apply to Resource Maps. For instance, a Resource Map is required to express its relationship to the Aggregation it describes using the `ore:describes` predicate, and the subject of the triple must be the URI of the Aggregation. In this example, the Aggregation and Resource Map also declare their types. You can compare these triples that express the relationship between the Resource Map and its Aggregation in figure 10, which displays the entire Resource Map for a page Aggregation, encoded in RDF/XML.

1. The Aggregation is a newspaper page.  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1#page>  
`rdf:type`  
`rdf:resource=http://chroniclingamerica.loc.gov/terms#Page`

```

<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF
  xml:base="http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:exif="http://www.w3.org/2003/12/exif/ns#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:ndnp="http://chroniclingamerica.loc.gov/terms#"
  xmlns:ore="http://www.openarchives.org/ore/terms/"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
>
<rdf:Description rdf:about="/lccn/sn90061457/1918-12-14/ed-1/seq-1/ocr.xml">
  <dc:format>text/xml</dc:format>
</rdf:Description>
<rdf:Description rdf:about="/lccn/sn90061457/1918-12-14/ed-1/seq-1/thumbnail.jpg">
  <dc:format>image/jpeg</dc:format>
</rdf:Description>
<rdf:Description rdf:about="/lccn/sn90061457/1918-12-14/ed-1/seq-1#page">
  <dcterms:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#date">1918-12-14</
  dcterms:issued>
  <ndnp:sequence rdf:datatype="http://www.w3.org/2001/XMLSchema#long">1</ndnp:sequence>
  <dcterms:title>The St. Joseph observer. - 1918-12-14 - 1</dcterms:title>
  <ore:isDescribedBy rdf:resource="/lccn/sn90061457/1918-12-14/ed-1/seq-1.rdf"/>
  <ore:aggregates rdf:resource="/lccn/sn90061457/1918-12-14/ed-1/seq-1/ocr.xml"/>
  <ore:aggregates rdf:resource="/lccn/sn90061457/1918-12-14/ed-1/seq-1/ocr.txt"/>
  <ore:aggregates rdf:resource="/lccn/sn90061457/1918-12-14/ed-1/seq-1.jp2"/>
  <ore:aggregates rdf:resource="/lccn/sn90061457/1918-12-14/ed-1/seq-1/thumbnail.jpg"/>
  <ore:aggregates rdf:resource="/lccn/sn90061457/1918-12-14/ed-1/seq-1.pdf"/>
  <foaf:depiction rdf:resource="/lccn/sn90061457/1918-12-14/ed-1/seq-1/thumbnail.jpg"/>
  <ore:isAggregatedBy rdf:resource="/lccn/sn90061457/1918-12-14/ed-1#issue"/>
  <rdf:type rdf:resource="http://chroniclingamerica.loc.gov/terms#Page"/>
</rdf:Description>
<rdf:Description rdf:about="/lccn/sn90061457/1918-12-14/ed-1/seq-1.rdf">
  <ore:describes rdf:resource="/lccn/sn90061457/1918-12-14/ed-1/seq-1#page"/>
  <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">
    2010-02-13T13:06:31-04:00</dcterms:modified>
  <rdf:type rdf:resource="http://www.openarchives.org/ore/terms/ResourceMap"/>
  <dcterms:created rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">
    2010-02-13T13:06:31-04:00</dcterms:created>
  <dcterms:creator rdf:resource="http://chroniclingamerica.loc.gov/awardees/dlc#awardee"/>
</rdf:Description>
<rdf:Description rdf:about="/lccn/sn90061457/1918-12-14/ed-1/seq-1.pdf">
  <dc:format>application/pdf</dc:format>
</rdf:Description>
<rdf:Description rdf:about="/lccn/sn90061457/1918-12-14/ed-1/seq-1/ocr.txt">
  <dc:format>text/plain</dc:format>
</rdf:Description>
<rdf:Description rdf:about="/lccn/sn90061457/1918-12-14/ed-1/seq-1.jp2">
  <dc:format>image/jp2</dc:format>
  <exif:height rdf:datatype="http://www.w3.org/2001/XMLSchema#long">6175</exif:height>
  <exif:width rdf:datatype="http://www.w3.org/2001/XMLSchema#long">5036</exif:width>
</rdf:Description>
</rdf:RDF>

```

**Figure 9**  
Aggregation Graph for a newspaper page Aggregation serialized as a Resource Map in RDF/XML

2. *Seq-1.rdf is a Resource Map.*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1.rdf>  
`rdf:type`  
`rdf:resource="http://www.openarchives.org/ore/terms/ResourceMap"`
3. *The page Aggregation is described by a Resource Map, seq-1.rdf.*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1#page>  
`ore:isDescribedBy`  
`rdf:resource=http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1.rdf`
4. *The Resource Map, seq-1.rdf, describes the page Aggregation.*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1.rdf>  
`ore:describes`  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1#page>

### Adding Metadata

A Resource Map is required to express some basic metadata about itself, such as its creator and the last time and date it was modified. Compare these triples that express metadata in figure 10, which displays the entire RDF/XML of the Resource Map for a page Aggregation.

1. *The Resource Map, seq-1.rdf, has a creator of dlc (the OCLC symbol of the Library of Congress).*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1.rdf>  
`dcterms:creator`  
`rdf:resource=http://chroniclingamerica.loc.gov/awardees/dlc#awardee`
2. *The Resource Map, seq-1.rdf, was last modified on February 13, 2010 at 1:06 p.m.*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1.rdf>  
`dcterms:modified`  
`rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime" 2010-02-13T13:06:31-04:00`
3. *The Resource Map, seq-1.rdf, was created on February 13, 2010 at 1:06 p.m.*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1.rdf>  
`dcterms:created`  
`rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime" 2010-02-13T13:06:31-04:00`

A Resource Map can also express metadata about the

Aggregation it describes.

1. *The page Aggregation was issued on December 14, 1918.*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1#page>  
`dcterms:issued`  
`rdf:datatype="http://www.w3.org/2001/XMLSchema#date" 1918-12-14`
2. *The page Aggregation is titled "The St. Joseph observer. - 1918-12-14 - 1" (the title of the newspaper, its date, and the page number).*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1#page>  
`dcterms:title`  
`"The St. Joseph observer. - 1918-12-14 - 1"`
3. *The page Aggregation has a sequence number of 1 (i.e., it is the first page of the newspaper).*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1#page>  
`ndnp:sequence`  
`rdf:datatype="http://www.w3.org/2001/XMLSchema#long" 1`
4. *The page Aggregation is depicted by thumbnail.jpg (a thumbnail image can serve as metadata about an object).*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1#page>  
`foaf:depiction`  
`rdf:resource=http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1/thumbnail.jpg`

### Aggregated Resources

Now we change our focus to the relationship between the page Aggregation and its Aggregated Resources. Every page consists of a JPEG 2000 image, a PDF, a thumbnail JPEG, the raw text from the OCR, and the structured XML output of the OCR. The Resource Map describes that these files are aggregated into a page. You can see how these triples have been encoded in RDF/XML in figure 4, which presents a Resource Map for a page Aggregation.

1. *The page Aggregation aggregates a JPEG 2000 file, seq-1.jp2.*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1#page>  
`ore:aggregates`  
`rdf:resource=http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1.jp2`
2. *The page Aggregation aggregates a PDF file, seq-1.pdf.*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1#page>

```

<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF
  xml:base="http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:ore="http://www.openarchives.org/ore/terms/"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
>
  <rdf:Description rdf:about="/lccn/sn90061457/1918-12-14/ed-1.rdf">
    <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2010-02-
13T13:37:16-04:00</dcterms:modified>
    <dcterms:creator rdf:resource="http://chroniclingamerica.loc.gov/awardees/dlc#awardee"/>
    <dcterms:created rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2010-02-
13T13:37:16-04:00</dcterms:created>
    <rdf:type rdf:resource="http://www.openarchives.org/ore/terms/ResourceMap"/>
    <ore:describes rdf:resource="/lccn/sn90061457/1918-12-14/ed-1#issue"/>
  </rdf:Description>
  <rdf:Description rdf:about="/lccn/sn90061457/1918-12-14/ed-1#issue">
    <ore:isDescribedBy rdf:resource="/lccn/sn90061457/1918-12-14/ed-1.rdf"/>
    <dcterms:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#date">1918-12-14</
dcterms:issued>
    <ore:isAggregatedBy rdf:resource="/lccn/sn90061457#title"/>
    <ore:isAggregatedBy rdf:resource="/batches/batch_mohi_carver_ver01#batch"/>
    <dcterms:title>The St. Joseph observer. - 1918-12-14</dcterms:title>
    <rdf:type rdf:resource="http://purl.org/ontology/bibo/Issue"/>
    <ore:aggregates rdf:resource="/lccn/sn90061457/1918-12-14/ed-1/seq-5#page"/>
    <ore:aggregates rdf:resource="/lccn/sn90061457/1918-12-14/ed-1/seq-1#page"/>
    <ore:aggregates rdf:resource="/lccn/sn90061457/1918-12-14/ed-1/seq-2#page"/>
    <ore:aggregates rdf:resource="/lccn/sn90061457/1918-12-14/ed-1/seq-7#page"/>
    <ore:aggregates rdf:resource="/lccn/sn90061457/1918-12-14/ed-1/seq-3#page"/>
    <ore:aggregates rdf:resource="/lccn/sn90061457/1918-12-14/ed-1/seq-4#page"/>
    <ore:aggregates rdf:resource="/lccn/sn90061457/1918-12-14/ed-1/seq-8#page"/>
    <ore:aggregates rdf:resource="/lccn/sn90061457/1918-12-14/ed-1/seq-6#page"/>
  </rdf:Description>
</rdf:RDF>

```

**Figure 10**  
Aggregation Graph for a newspaper issue Aggregation serialized as a Resource Map in RDF/XML

- ore:aggregates  
rdf:resource=http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1.pdf
3. *The page Aggregation aggregates a thumbnail image file, thumbnail.jpg.*  
http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1#page  
ore:aggregates  
rdf:resource=http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/thumbnail.jpg
  4. *The page Aggregation aggregates an OCR text file, ocr.txt.*  
http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1#page  
ore:aggregates  
rdf:resource=http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/ocr.txt
  5. *The page Aggregation aggregates a OCR XML file, ocr.xml.*

http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1#page  
ore:aggregates  
rdf:resource=http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/ocr.xml

### Nested Aggregations

The data model for the NDNP defines Aggregations for pages, issues, titles, and batches. Issues aggregate pages. Issues are aggregated by titles and batches. Let's take a look at a set of triples that describe an issue Aggregation. You can see that a Resource Map for it would need to express that the issue is being aggregated by a title Aggregation and a batch Aggregation. The full Resource Map serialized in RDF/XML can be found in figure 11.

1. *Ed-1.rdf describes an issue Aggregation.*  
http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1.rdf

- ore:describes  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1#issue>
2. *Ed-1.rdf is a Resource Map.*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1.rdf>  
 rdf:type  
 rdf:resource=<http://www.openarchives.org/ore/terms/ResourceMap>
  3. *The Resource Map, ed-1.rdf, was modified on February 13, 2010 at 1:27 p.m.*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1.rdf>  
 dcterms:modified  
 rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime" 2010-02-13T13:37:16-04:00
  4. *The Resource Map, ed-1.rdf, was created by dlc (the Library of Congress).*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1.rdf>  
 dcterms:creator  
 rdf:resource=<http://chroniclingamerica.loc.gov/awardees/dlc#awardee>

And now, some triples that describe the issue Aggregation:

1. *The Aggregation is described by the Resource Map, ed-1.rdf.*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1#issue>  
 ore:isDescribedBy  
 rdf:resource=<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1.rdf>
2. *The Aggregation is a newspaper issue.*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1#issue>  
 rdf:type  
 rdf:resource=<http://purl.org/ontology/bibo/Issue>
3. *The issue Aggregation was issued on December 14, 1918.*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1#issue>  
 dcterms:issued  
 rdf:datatype="http://www.w3.org/2001/XMLSchema#date" 1918-12-14
4. *The issue Aggregation is titled "The St. Joseph observer. - 1918-12-14".*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1#issue>  
 dcterms:title  
 "The St. Joseph observer. - 1918-12-14"

5. *The issue Aggregation is aggregated by a title Aggregation.*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1#issue>  
 ore:isAggregatedBy  
<http://chroniclingamerica.loc.gov/lccn/sn90061457#title>
6. *The issue Aggregation is also aggregated by a batch Aggregation.*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1#issue>  
 ore:isAggregatedBy  
[http://chroniclingamerica.loc.gov/lccn/batches/batch\\_mohi\\_carver\\_ver01#batch](http://chroniclingamerica.loc.gov/lccn/batches/batch_mohi_carver_ver01#batch)
7. *The issue Aggregation aggregates a page Aggregation, seq-1 (first page).*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1#issue>  
 ore:aggregates  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-1#page>
8. *The issue Aggregation aggregates a page Aggregation, seq-2 (second page, and so on...)*  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1#issue>  
 ore:aggregates  
<http://chroniclingamerica.loc.gov/lccn/sn90061457/1918-12-14/ed-1/seq-2#page>

## Other Serializations

RDF/XML was selected to demonstrate Resource Map serialization in this issue of *Library Technology Reports* because it was implemented in our example of the NDNP; however, this should not imply any endorsement of RDF/XML or give the impression that it is better or more widely adopted than other serializations. In fact, the ORE specification suggests that Atom may currently be the most widely used serialization.<sup>3</sup> In addition to RDF/XML, the ORE specification documents two other formats, RDFa and Atom XML, although others may be used.

RDFa allows a Resource Map to be encoded in an XHTML document. One advantage to this approach is that the XHTML document can be presented as a webpage by a browser and can also be parsed by an ORE client to retrieve a machine-readable Resource Map. For example, the "splash page" for an Aggregation could be implemented as RDFa in XHTML and serve both a human user and a machine user-agent at the same time.

The Atom Syndication Format is a part of the Atom standard that is an XML format that describes lists of related items as web feeds. Items in the feed are known as entries, and extensible metadata can be attached to

an entry.<sup>4</sup> Atom was originally created as an alternative to Really Simple Syndication (commonly known by its indefinite acronym, RSS). In ORE, Resource Maps may be exposed as Atom entries or Resource Maps may simply be linked from entries. In this way, Resource Maps can be discoverable in batches through Atom feeds.

For more detailed information on RDFa and Atom XML serialization of Resource Maps, reference the resources in chapter 5, namely the ORE User Guides for Resource Map Implementation in RDFa and Atom.

## Exposing Resource Maps

Because of the pervasiveness of the World Wide Web, it is no surprise that the ORE specification recommends protocol-based URIs such as HTTP for Aggregations and Resource Maps.<sup>5</sup> When most people think about HTTP, the most common scenario that comes to mind is a web browser requesting an HTML page from a webserver. HTTP has also been put to use in developing and deploying Web Services that use it to request and return XML and other kinds of data besides HTML webpages. The same architecture of the World Wide Web is employed in an HTTP implementation of ORE. In this section we'll explore how Resource Maps can be made available by webserver to ORE clients after they have been serialized.

### HTTP 303

Most Web users are familiar with the “404 Page Not Found” response code that you get when you experience a broken link or try to access a webpage that is no longer available from a webserver. In fact, there are a number of different codes that can be exchanged between an HTTP client (e.g., a web browser) and a webserver. One such response code is “303 See Other,” in which the webserver can provide the client with the location of another resource, perhaps in a different format or a different language than the resource that was requested. In this way, a web client could also request the URI of an Aggregation that the webserver handles with a “303 See Other” and then provides the URI of a Resource Map. Information can be communicated between the web client and server in the HTTP headers; for example, the client may indicate it will accept particular formats by stating the acceptable MIME types in the header of its request.<sup>6</sup> For Aggregations that have more than one serialized Resource Map, this *content negotiation* enables the webserver to provide the format that is preferred by the client. For example, a web client could specify that it will accept “application/atom+xml,” in which case the server may try to respond with the Atom XML serialization of the Aggregation’s Resource Map.

In the event that more than one acceptable format is

available or no acceptable formats are available, it is left up to the webserver to determine which format to send in its response.<sup>7</sup> The ORE specification recommends that the webserver be configured to attempt to honor the format that is requested by the web client for an Aggregation, but if it can’t, to respond with a Resource Map (as opposed a non-machine-readable format or HTML). Furthermore, it recommends the webserver respond with a Resource Map in cases where an Aggregation is requested and no format preference is specified.<sup>8</sup>

A 303 redirection can also allow a human-readable “splash page” to be accessed as well as Resource Maps when one is available. The “splash page” should include HTML `<link>` elements to enable the discovery of Resource Maps that are associated with the Aggregation. Keep in mind that the client should be redirected to a Resource Map and not the splash page when the client does not specify an acceptable MIME type in its HTTP request.<sup>9</sup>

HTTP 303 with content negotiation is the recommended implementation when

- Your webserver supports 303 and content negotiation
- You require support for multiple Resource Maps
- You wish to include “splash pages” along with Resource Maps
- You wish to allow easy extensibility to future, additional Resource Maps or “splash pages”<sup>10</sup>

Note that 303s with content renegotiation is the only recommended HTTP implementation that supports serving multiple Resource Maps for an Aggregation. If your webserver does not support content negotiation and you only have one Resource Map, the ORE specification recommends that 303 redirection without content negotiation be implemented. In this case, Resource Maps can become aware of each other by linking to each other using `ore:isDescribedBy` predicates and HTML `<link>` elements in HTML “splash pages.”<sup>11</sup>

### Hash URIs

In some situations, your webserver may not support 303s or your Resource Maps may be hosted on a webserver that is not under your control. One simple work-around is to construct URIs for Aggregations using hash notation, where the URI to the Resource Map is followed by the pound sign (“#”) and a fragment that qualifies the URI as being for the Aggregation. For example:

- URI for Aggregation: `http://example.com/resource-map.rdf#aggregation`
- URI for Resource Map: `http://www.example.com/resourcemap.rdf`

This same notation is commonly used for “jump links” in HTML, to direct a browser to scroll down to a particular named anchor in a webpage. Because this is a client-side behavior, the web client does not send the fragment to the server in its request. And so, a client requesting or following a link to an Aggregation (<http://example.com/resourcemap.rdf#aggregation>) is only effectively sending the URI to the Resource Map (<http://example.com/resourcemap.rdf>) because it chops off the fragment (`#aggregation`). After the client receives a response that is the Resource Map, it may attempt to resolve the fragment but will fail, harmlessly. If the web client discards the fragment, you may be asking yourself, why bother with it in the first place? The reason is that the hash URI allows you to satisfy the requirement that an Aggregation have its own unique URI, while allowing an easy way to provide a Resource Map for it.<sup>12</sup>

### RDFa

RDFa enables a Resource Map to be contained inside an XHTML document, which can be rendered by a web browser as a normal webpage and also be used by an ORE client as a Resource Map. The previous HTTP implementations (303s with or without content negotiation and hash URIs) work the same way with RDFa as other serializations of Resource Maps. The only difference is when there are multiple serializations available; the Resource Map in RDFa should be used as the default because it fulfills the requirement of being machine-readable with the additional value of being presentable as a webpage.<sup>13</sup>

## Discovering Resource Maps

There are a variety of ways in which Resource Maps (and thus, Aggregations) can be discovered on the Web after they have been serialized and exposed for ORE client applications, such as harvesters and crawlers, to find. The ORE specification suggests two categories of discovery, resource embedding and batch discovery, although new mechanisms or categories may emerge and evolve over time.<sup>14</sup>

### Resource Embedding

One strategy for exposing Resource Maps for discovery is to link to them from within the `<head>` tags of an



**Figure 11**  
Enabling discovery of Resource Maps using the `<link>` element in HTML

HTML webpage. Web crawlers that find the page can parse `<link rel="resourcemap">` elements, identify them as MIME-typed Resource Maps (e.g., `application/rdf+xml`) and follow their URIs to Resource Maps. Multiple Resource Maps and serializations of the same Resource Map can be provided with multiple `<link>` elements.<sup>15</sup> This is illustrated in Figure 11 for a webpage that displays an issue of a newspaper in the NDNP.

In some cases, the webpage may itself be an Aggregated Resource that has one or more `<link>` elements that point to Resource Maps for Aggregations that include it. If the Resource Maps are available in an Atom feed, it may be exposed for autodiscovery using `<link rel="alternate">`. It is also possible to link to the URI of an Aggregation by specifying `<link rel="aggregation">` without a type. Note that it is up to the crawler or other ORE client application to follow the links to the Resource Maps and process them in order to determine the relationships of the Aggregations and Aggregated Resources.<sup>16</sup> This cannot be inferred by the placement of the `<link>` elements or any other information in the ORE documentation that encourages displaying the URIs of Aggregations in the content of web pages, for example, to allow users to copy and paste them into blog posts, e-mails, and other environments.<sup>17</sup>

### Batch Discovery

Beyond linking or embedding Resource Maps into webpages, Resource Maps may be made discoverable in batches more explicitly. If Resource Maps have been serialized using Atom XML, they may be exposed directly in Atom

## Exercise

Resource Maps for newspaper title and batch Aggregations have been serialized from the National Digital Newspaper Program as additional examples in Appendix 1 and 2. Can you understand and graph them with pencil and paper?

feeds, which supply their own mechanism for autodiscovery. Resource Maps can be linked from Atom entries or may be supplied as entries themselves.<sup>18</sup> The URIs for Resource Maps and Aggregations may also be included in a Sitemap, which is an XML document that lists all of the URLs in a particular website along with some metadata describing them to a web crawler or other web client.<sup>19</sup> The last batch discovery mechanism suggested by the ORE specification is the OAI-PMH (explained in chapter 1), which is a protocol for metadata harvesting. The serialization format of the Resource Map can be specified as a metadata Prefix, enabling the harvest of Resource Maps instead of or in addition to Dublin Core metadata records. Even if the full Resource Maps aren't exposed to be harvested, another type of metadata record may be harvested that includes links to Resource Maps or Aggregations and serve the purpose of facilitating their discovery.<sup>20</sup>

## Notes

1. Carl Lagoze et al., "ORE User Guide: Primer," Oct. 17, 2008, Open Archives Initiative Object Reuse and Exchange website, <http://www.openarchives.org/ore/1.0/primer> (accessed March 6, 2010).
2. Carl Lagoze et al., "ORE User Guide: Resource Map Implementation in RDF/XML," Oct. 17, 2008, Open

- Archives Initiative Object Reuse and Exchange website, <http://www.openarchives.org/ore/1.0/rdfxml> (accessed March 6, 2010).
3. Carl Lagoze et al., "ORE User Guide: HTTP Implementation," Oct. 17, 2008, Open Archives Initiative Object Reuse and Exchange website, <http://www.openarchives.org/ore/1.0/http.html> (accessed March 6, 2010).
4. M. Nottingham and R. Sayre, eds., "RFC 4287: The Atom Syndication Format," Dec. 2005, The Internet Engineering Task Force, <http://www.ietf.org/rfc/rfc4287.txt> (accessed March 6, 2010).
5. Lagoze et al., "ORE User Guide: HTTP Implementation."
6. Leo Sauermann and Richard Cyganiak, eds., "Cool URIs for the Semantic Web," World Wide Web Consortium website, Dec. 3, 2008, <http://www.w3.org/TR/cooluris> (accessed March 6, 2010).
7. K. Holtman and A. Mutz, "RFC 2295: Transparent Content Negotiation in HTTP," March 1998, The Internet Engineering Task Force, <http://www.ietf.org/rfc/rfc2295.txt> (accessed March 6, 2010).
8. Lagoze et al., "ORE User Guide: HTTP Implementation."
9. Ibid.
10. Ibid.
11. Ibid.
12. Ibid.
13. Ibid.
14. Carl Lagoze et al., "ORE User Guide: Resource Map Discovery," Oct. 17, 2008, Open Archives Initiative Object Reuse and Exchange website, <http://www.openarchives.org/ore/1.0/discovery> (accessed March 6, 2010).
15. Ibid.
16. Ibid.
17. Ibid.
18. Ibid.
19. "What Are Sitemaps?" sitemaps.org website, <http://www.sitemaps.org> (accessed March 6, 2010).
20. Lagoze et al., "ORE User Guide: Resource Map Discovery."