Candidate approaches for describing ORE Aggregations in METS

PRELIMINARY DISCUSSION DRAFT

7 January 2009
Thomas Habing
Timothy Cole

Linked from this document are illustrations of two possible ways to express descriptive information contained in an ORE Resource Map in METS:

1. An ORE Resource Map represented by a specific structMap in a METS document (allows multiple ReMs in single METS document)
2. An ORE Resource Map represented by a single METS document (one-to-one correspondence)

While the OAI-ORE and METS initiatives each have different and distinctive goals and objectives, it has been suggested by members of both communities that it might be useful to describe in METS the essential structure, scope and relationships/properties of an ORE aggregation as described by an ORE Resource Map (ReM). Similarly, it has been suggested that it might be useful to represent a complex or compound digital object described by a METS instance as one or more ORE Aggregations -- i.e., it might be useful to include information in a METS instance (e.g., one or more structMap elements) that could directly be translated into one or more ORE ReMs. We suggest here two possible approaches to accomplish these objectives. Undoubtedly there are additional viable approaches, but to facilitate application development, we feel the communities should focus on, develop, and promulgate at most only a few. From the starting point provided here we hope that interested members of both communities can come together in an ad hoc working group to do just that.

Each proposed approach shown here has its pros and cons. The first approach posits a mapping in METS such that a specific structMap within a METS document could be thought of as corresponding in essence to an ORE ReM. Information in fileSec and dmdSec elements linked from the structMap would also map into the ORE Rem. This approach allows for situations where ORE Aggregations need to be surfaced from already existing METS documents, or where you need to wrap multiple ORE ReMs within a single METS document. It is assumed that the METS document may separately contain information not meant to map to anything in associated ORE ReM(s), allowing implementers to be explicit about which object components being described by a METS instance belong to any associated ORE Aggregation they wish to assert.

The second approach posits a mapping where the METS document itself is alternative serialization of the ORE ReM. In asserting selected equivalences (at least for narrow purposes) between ORE and METS semantics, this approach attempts to make optimal use of METS elements and attributes for which there seems a natural mapping and correspondence to ORE properties and types. This approach is potentially more useful than the first when the only objective is to make ORE Aggregations more accessible to METS-based applications and tools; however, asserting METS as another way to serialize an ORE ReM may be controversial for some.

In developing these two illustrations, we wish to highlight a few key points:

- There are issues to do with assigning values for ORE URI-R (URI of the ReM) and URI-A (URI of the Aggregation). In the case of approach 1 (potentially multiple ReM-like structMaps in a single METS instance), it is difficult to define an appropriate URI-R. We recommend identifying the "ReM" structMap using the METS file URI and a fragment identifier based on the ID attribute of the structMap element, but such URIs are possibly problematic as ORE URI-Rs and would at the very least preclude any reliance on a METS-embedded structMap as an "Authoritative Resource Map" (see: ORE Abstract Data Model). In contrast, approach 2 does allow for construction of URI-R and URI-A values supporting the use of the METS-serialized ReM as an Authoritative Resource Map.

- By design, METS can be used to express more than just RDF graph triples. Also a given triple often can be expressed in more than one section of a METS document (e.g., a given triple could be expressed either in the structMap or in the dmdSec or even in the amdSec). In contrast, ORE ReMs deal only with description and assertions which can be represented as part of an RDF graph. This requires the implementer to use care when constructing METS instances or parts of METS instances that are meant to be transformed into an ORE ReM. We relied on RDF syntax and semantics exclusively throughout dmdSec elements used in our illustrations. As a general rule, we chose to express as much of ORE ReM equivalent description as feasible in METS structMap elements, limiting our use of metsHdr, dmdSec and fileSec. Our examples make no use at all of amdSec, structLink or behaviorSec.

- In constructing our ReM-corresponding structMap elements, we made use of striped syntax to represent relationships between resources, including between ReM, Aggregation, and Aggregated Resource. This means that the root div element of our structMap in approach one corresponds to an ORE ReM. A child div is used to express a "Describes" relationship to the next-layer-in div representing the Aggregation, which in turn has child div elements expressing an "Aggregates" relationship to subsequently nested div elements which represent Aggregated Resources. In approach two, since the entire METS document is posited as equivalent to the ReM, the root div of the structMap represents the Aggregation (i.e., the "Describes" relationship is implicit), and it continues from there as in approach one.

By way of illustrating practical potential of the approaches linked from this page, an XSLT is provided for each approach implementation that transforms an ORE Atom Entry Resource Map into the corresponding METS Resource Map. While this should work well for most ATOM ReM serializations, we have not extensively nor systematically tested across full range of possible ATOM ReMs. (This is left as a potential "to-do" item.)

This preliminary discussion draft (we believe significant additional work, correction and/or refinement may be needed) is provided as a starting point for further collaborative work. We would suggest that it be shared with members of both communities (e.g., members of the ORE Technical Committee and Liaison Group and members of the METS Editorial Board), and that volunteers from both communities be sought to continue (albeit virtually) ad hoc work to refine and advance these ideas to a natural conclusion, including possibly, endorsed suggestions or best practices useful to would-be implementers working across both METS and ORE.