



→ EDM Case Study: Mapping EAD to EDM



EuropeanaConnect¹ (May 2009 to October 2011) has carried out multiple experiments to test the feasibility of using EDM with various kinds of data. This included the Amsterdam Museum Linked Data case² as well as an experiment investigating how metadata can be represented with EDM following the important principles of the archive domain.



The latter experiment, reported upon in a technical paper,³ focused on the Encoded Archival Description (EAD) standard⁴ as it was used in the APENet project⁵ (January 2009 – December 2011) that contributed data to Europeana. It concluded that EDM offers a feasible approach to represent fairly complex metadata structures such as those in EAD.

EAD indeed represents archives' *finding aids*, which describe collections that consist of different levels of elements (classes, series, subseries, files and items). They may be organised in a deep hierarchy that provides them with a context of provenance and interpretation. Each level can provide descriptive metadata and be associated with digital representations (for example, digitised book pages).

¹ <http://europeanaconnect.eu>

² <http://pro.europeana.eu/thoughtlab/linked-open-data>

³ Steffen Hennicke, Victor de Boer, Antoine Isaac, Marlies Olensky, Jan Wielemaker. *Conversion of EAD into EDM Linked Data*. First International Workshop on Semantic Digital Archives, TPD conference. Berlin, 2011. Available at <http://ceur-ws.org/Vol-801/>

⁴ <http://www.loc.gov/ead/>

⁵ <http://www.apenet.eu/>



EDM copes with the representation of hierarchical and sequential order in EAD. Representing elements in the hierarchy as fully-fledged, distinct resources allows the creation of explicit links between them. Together with the properties (types of links) that EDM re-uses or mints in order to reflect ‘part-of’ or ‘next-in-sequence’ relationships, this feature enables the precise representation needed.

The experiment used more advanced RDF modelling functions that were not planned in the first Europeana implementation of EDM data ingestion. In particular, it was possible to make different levels of information granularity co-exist in an interoperable fashion. Namely, the most general level of semantic detail, as represented by the Dublin Core properties EDM re-uses by default, e.g. ‘identifier’, can be specialised to capture the finer distinctions made in the original EAD data, e.g. ‘call number’.

The report also discusses other conversion decisions Europeana EAD data providers could make when mapping their data to EDM. One important example was the possibility of representing the EAD elements without digital representations as mere ‘contextual’ entities, not fully-fledged items to be considered as retrievable objects by Europeana.

We finally observed that an EDM data provider must aim at providing data that is not ‘polluted’ by information needed only for materialising the data. Such information can be useful for some purposes (for instance, data display) but it should only be made explicit at later stages. This relates well to a previous report from the National Archives of the Netherlands.⁶ This report showed how difficult it was to enforce such a best practice with very simple models such as ESE, which does not fit well with the rich structure of EAD and forces a provider to express the same data in several places. EDM will allow this to be changed.

⁶Go Sugimoto, Wim van Dongen. *Technical Report: Archival Digital Object Ingestion into Europeana (ESE-EAD harmonisation) v1.0*. Available at <http://pro.europeana.eu/technical-requirements>