



Executive Summary

D2.4 – Digital cultural resources: Selection, Processing & Linking

This report presents the results of a collaborative effort of the project partners, which took place in months M2-M25 of the project. It focuses on the production of semantically modelled digital cultural resources and their ingestion into the CrossCult Knowledge Base (CCKB). It presents the most recent version of the integrated data model, which builds on the refined version of the CCKB as discussed in deliverable D2.5. This final version expands to the full scale of the available data of cultural heritage items content, narratives and reflective topics, including refinements and semantic enrichment of the data with links to definitions of standard semantic web resources.

Four distinct pilots contribute data to the CrossCult project covering a unique range of cultural heritage venues across Europe (see D2.1). The project ingests a wide range of diverse data associated to cultural heritage objects, events and subjects that span from Antiquity to Modern times. Such disparate data means there is a wide array of formats, technologies, management and classification approaches relevant to each data provider or source. The adopted data modelling method caters for solutions that address issues relating to the diversity of content types, data formats, and levels of data detail. Data modelling in the context of this deliverable refers to the specific process of applying the conceptual arrangements and definitions of the CrossCult Upper-level ontology to a range of disparate data resources. The ontology provides a flexible and robust schema of standardised conceptual abstractions, capable of supporting the data modelling needs of cultural heritage data. The deliverable presents examples and use cases demonstrating the capability of the CCKB to support a range of complex and cross-searching scenarios.

The Data modelling process evolved in two separate iterations. The deliverable discusses both, putting emphasis on the second iteration since the outcome of the first iteration is described in detail in deliverable D2.2. The first iteration addressed a manageable and well-balanced set of representative data, which did not vary significantly across the four pilots. Overall, it delivered 80 uniquely identified items, which were composed of 102 Physical Man Made Objects and 17 Physical Man Made Things, delivering overall 3440 ontology (OWL) statements of named individual declaration and property assertion. The second iteration supported a homogenised automatic processing approach of data ingestion based on standard data schema definitions. It was focused on data ingestion of reflective topics, narratives, digital resources and semantic enrichment definitions. The narratives received 2140 semantic enrichments to unique DBpedia concepts whilst 1250 vocabulary terms of the CrossCult Classification Scheme were mapped to concepts of Getty AAT, Eurovoc, LOC and the Unesco Vocabulary. Overall, the final version of the unified CCKB contains 1,201,533 axioms and 282,010 ontology individuals including cultural heritage items, vocabulary entries and DBpedia enrichments.

As a final result, the digital cultural resources of the project pilots such as digital stories, images, hypertext resources and audio-visual elements are interweaved into rich compositions, fostering reflection within the semantic environment of the CCKB. Applications that use the CCKB can provide a synthesised view on a topic that can hardly be conveyed by the usual item description that accompanies a museum exhibit. The CrossCult knowledge base and ingested data have been deployed on the project's triple store.

www.crosscult.eu