D2.5

Upper-level Cultural Heritage Ontology

2018-03-29

Andreas Vlachidis (UCL); Antonis Bikakis (UCL); Melissa Terras (UCL); Yannick Naudet (LIST); Louis Deladiennee (LIST); Daphne Manessi (TEI-A); Evgenia Vasilakaki (TEI-A); Ioannis Triantafyllou (TEI-A); Joseph Padfield (NG); Kalliopi Kontiza (NG)

This deliverable presents the Upper-level Ontology and the other ontological schemas and vocabularies that we used to model the semantics of the "world" of CrossCult and its four pilots. It consists of two documents: a report describing the rationale and structure of the ontology and a PDF file containing the definitions of the classes and properties of the CrossCult ontologies in the syntax of Description Logics.

www.crosscult.eu

Project acronym	**CrossCult**
Full title	CrossCult: Empowering reuse of digital cultural heritage in context-aware crosscuts of European history
Grant agreement number	693150
Funding scheme	Innovation Action (IA)
Work programme topic	H2020-REFLECTIVE-2014-2015/REFLECTIVE-6-2015: Innovation ecosystems of digital cultural assets
Project start date	2016-03-01
Project duration	36 months
Workpackage	WP2
Deliverable lead organisation	University College London
Authors	Andreas Vlachidis (UCL); Antonis Bikakis (UCL); Melissa Terras (UCL); Yannick Naudet (LIST); Louis Deladiennee (LIST); Daphne Manessi (TEI-A); Evgenia Vasilakaki (TEI-A); Ioannis Triantafyllou (TEI-A); Joseph Padfield (NG); Kalliopi Kontiza (NG)
Reviewers	José Juan Pazos Arias (UVIGO); Alberto Gil Solla (UVIGO)
Version	1.0
Status	for EU review
Dissemination level	PU: Public
Due date	M25 (2018/03/31)
Delivery date	2018/03/29

Table of Contents

List of Tables

List of Figures

1. Introduction

The report presents a collaborative effort from the project partners, which took place in months M2-M25 of the project and focused on the delivery of an Upper-level conceptual structure that captures common concepts and relationships across the four project pilots. The deliverable is focused on i) reviewing the ontology requirements, ii) investigating ontology standards relevant to the project, iii) presenting the Upper-level Ontology and the other ontological schemas and vocabularies that we used to model the semantics of the "world" of CrossCult.

Compared to the first version of this deliverable, which was submitted in M8, this report additionally presents:

- The overall architecture of the CrossCult Knowledge Base (Section 2)
- The refinements of the Upper-level ontology (since M8), especially with respect to the representation of the concept of reflection and its relation to other project-related concepts (Section 3).
- The ontological representation of venues and users (the Venue and User ontologies, Sections 4 and 5)
- The CrossCult Classification Scheme, a faceted vocabulary that we developed for addressing the vocabulary needs of the project (Section 6).

The deliverable begins with an overview of the conceptual modelling requirements of CrossCult and highlighting relevant ontology standards (rest of Section 1). It then provides an overview of the architecture and ontological components of the CrossCult Knowledge Base (Section 2). This is followed by a more detailed presentation of its different components: the Upper-level ontology (Section 3); the Venue Ontology (Section 4); the User Ontology (Section 5); and the CrossCult Classification Scheme (Section 6). It concludes with a summary of the deliverable and a brief description of its relation with the other deliverables of the project (Section 7).

The accompanying PDF file contains the definitions of the classes and properties of the CrossCult ontologies in the syntax of Description Logics. The OWL file containing the code of the CrossCult ontologies in the XML syntax is available on the CrossCult website: http://www.crosscult.eu/en/resources/datasets/

1.1. Ontology Requirements and Standards

The CrossCult ontology is defined as a generic Upper-level conceptual structure that captures common concepts and relationships across the four pilots of the project. As such, the ontology delivers formalisms that describe the "world" of CrossCult, which accommodates common conceptual arrangements and enables augmentation, linking, semantic-based reasoning and retrieval across disparate data resources. The CrossCult ontology requirements are summarised as follows:

- a single and generic Upper-level structure that acts as a semantic layer of common concepts and relationships across pilots,
- a robust ontological definition that enables efficient semantic-based reasoning and

retrieval,

- a scalable structure that can be formally extended to support specialised conceptual needs when required,
- a knowledge representation model that builds on standard Semantic Web technologies facilitating interoperability and linking with Linked Data resources,
- a knowledge representation model that makes maximum reuse of established semantic web resources and standards.

With respect to the above requirements, the project has concluded to use OWL2 (https://www.w3.org/TR/owl2-syntax/) as the underlying language for the CrossCult ontology for different reasons:

- OWL2 is the standard ontology language for the Semantic Web, and therefore, adopting its semantics enables the reuse of other Semantic Web resources (ontologies, vocabularies, datasets, etc.) and better interoperability with other Semantic Web-based systems and applications.
- OWL2 meets the knowledge representation needs of the project. Specifically, it enables the representation of generic concepts of the cultural heritage domain (as OWL classes), individual objects of the project pilots (as OWL individuals), instance relationships among the individual objects and the ontology classes, relationships among individual objects (object properties), attributes of individual objects (as datatype properties), hierarchical relationships among the concepts (*subClassOf*) and the properties (*subPropertyOf*), equivalence (*sameAs*) relationships between the pilots' metadata and external resources, symmetric and transitive relationships, which are useful especially for spatiotemporal representation, etc.
- OWL2 enables the creation of data models at different levels of abstraction, and provides tools for their further extension, refinement or specialisation.
- OWL2 has efficient reasoning support. There are several efficient reasoning tools that we can use to automatically perform several useful tasks such as checking the consistency of the ontology, detecting unintended classification of instances or unintended relations between classes, inferring further ontological relationships based on the existing ontological definitions, etc.

The project has also decided to adopt elements from the following standard Semantic Web ontologies that support the aims of the CrossCult ontology.

1.2. CIDOC-CRM

The Conceptual Reference Model (CRM) of the International Council of Museums (ICOM) – International Committee for Documentation (CIDOC) [1] is a well-established ISO standard (ISO 21127:2006) in the modelling of cultural heritage information. CIDOC-CRM provides an extensible semantic framework that any cultural heritage information can be mapped to. It provides a framework for matching instances of people, places, things, events and periods using the information and context around these entities. In addition, it provides a "semantic glue" needed to mediate between different sources of cultural heritage information. The applicability of the CIDOC-CRM in information systems of the broader cultural heritage domain is evident in the literature on numerous large-scale projects[1]. For the needs of our project we will adopt the OWL2 version of CIDOC-CRM as defined by the Erlangen implementation of the model (http://erlangen-crm.org/).

Adopting CIDOC-CRM as the core conceptual layer of the CrossCult Upper-level ontology guarantees the use of well-defined and interoperable semantics, which support the generic aims of the Upper-level structure whilst providing specialisations that can benefit the individual needs of pilots. On the other hand, CIDOC-CRM as a formal and generic structure of concepts and relationships is not tied to any particular vocabulary of types, terms and individuals. This level of abstraction, albeit useful for the semantics of the broader cultural heritage domain, does not cover the need for a finer definition of types, terms and appellations. The need for an additional level of vocabulary semantics is addressed by connecting thesauri and glossary concepts with CIDOC-CRM.

1.3. SKOS

Simple Knowledge Organization System (SKOS, https://www.w3.org/2004/02/skos/) is a W3C recommendation designed for representation of thesauri, classification schemes, taxonomies, or any other type of structured controlled vocabulary. It builds upon RDF and RDFS, and its main objective is to enable easy publication and use of such vocabularies as linked data. SKOS structures can be linked to CIDOC-CRM concepts to provide a specialised vocabulary to instances of the ontology. It is an established approach that has been followed by numerous projects in the domain of cultural heritage. In its most common form, individuals of CIDOC-CRM concepts are connected via the *E55.Type* entity to SKOS thesauri concepts that provide further semantic relationships (e.g. broader / narrower term).

The Upper-level ontology incorporates the SKOS semantics, specifically the SKOS *Concept* and *Concept Scheme* classes and their associated properties, to provide access to specialised vocabularies.

1.4. Reuse of standard Semantic Web vocabularies

The Upper-level ontology also mediates and enables connections to additional semantics from specialised ontologies. For example, the FOAF (Friend-Of-A-Friend) ontology (http://xmlns.com/foaf/spec/) is a machine-readable ontology describing persons and their activities, and is suitable for linking to a specialised user ontology. The connection is enabled through a cherry-picked selection of useful ontology constructs (classes and properties) that participate in the Upper-level ontology, such as the FOAF *Person* class and the *interest_topic* property. The Upper-level ontology adopts also elements from the Dublin Core Schema (http://dublincore.org) to describe periods of time and specialised datatypes.

A range of specialised vocabularies has also been identified as relevant to the needs of the ontology. Such thesauri are primarily relevant to the broader cultural heritage domain but also span to more generic vocabularies and include:

- The Art & Architecture Thesaurus: a structured vocabulary of approximately 44,000 concepts of art, architecture and culture items.
  http://www.getty.edu/research/tools/vocabularies/aat/

[1] http://www.cidoc-crm.org/useCasesPage

- EuroVoc: a multilingual, multidisciplinary thesaurus, aiming to support the information management and dissemination services of the EU and its members. http://eurovoc.europa.eu
- Library of Congress Subject Authority Records http://id.loc.gov/authorities/subjects.html
- DBpedia: a crowd-sourced generic dataset containing information created in various Wikimedia projects structured in RDF. http://wiki.dbpedia.org

2. The CrossCult Knowledge-Base

The CrossCult Knowledge Base [2] (hereafter CCKB) is a comprehensive structure of semantic definitions and formalisms, developed for facilitating interoperable connections between cultural heritage data. Based on maximum reuse of well-established technologies, the CCKB incorporates a set of standard Semantic Web technologies and formats to support the data modelling requirements and objectives of CrossCult. The CCKB stack (Figure 1) illustrates the architecture of the knowledge base, where each section carries different semantics: a) the bottom section carries the semantics of different standard ontological schemas adopted in the CCKB; b) the middle section accommodates the project-specific cultural heritage semantics; c) the side section refers to the complementary CrossCult Classification Scheme (CCCS) vocabulary; and d) the top section to the representation of venues and users.

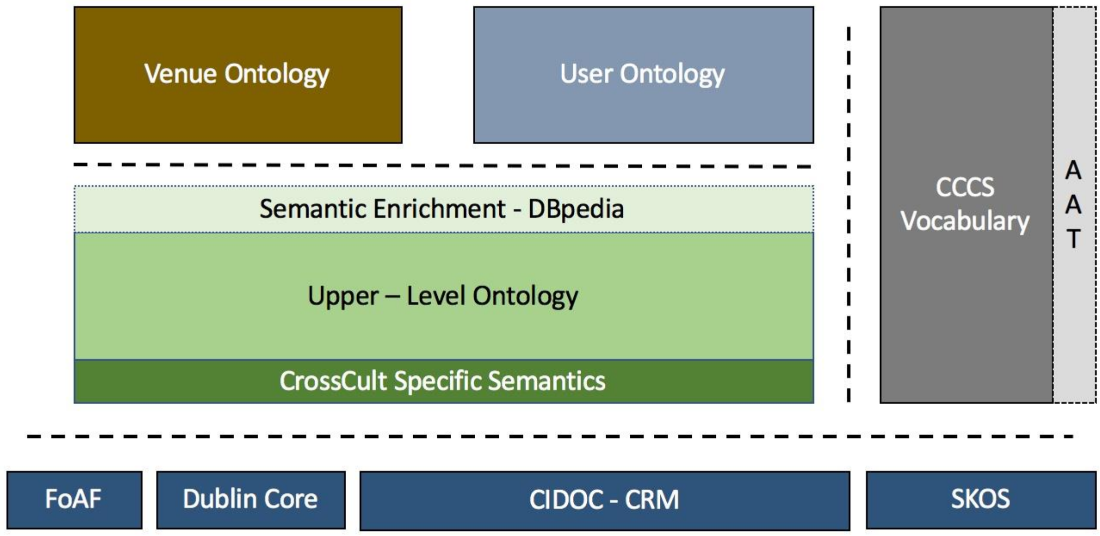

Figure 1: The architecture of the CrossCult Knowledge Base.

The four ontological schemas of the bottom section constitute the foundation of the architecture with CIDOC-CRM being the most prominent. The framework is complemented by the semantics of the Simple Knowledge Organization System (SKOS); the Dublin Core Schema, a standard vocabulary for describing web resources; and the FOAF (Friend-Of-A-Friend) ontology, which is used for mediating the semantics between the User Ontology layer and the Upper-Level Ontology layer in terms of describing user related entities and their interests.

The middle layer accommodates the semantics of the Upper-level ontology, which is defined as a generic conceptual structure that captures common concepts and relationships across a diverse range of cultural heritage data. Being written in OWL2 (the standard ontology language for the Semantic Web), it enables augmentation, semantic linking, semantic-based reasoning and retrieval across disparate data resources. It is driven by a core-subset of CIDOC-CRM semantics complemented with and extended by a set of project-specific definitions for handling the requirements of reflection, holistic understanding and reinterpretation of the European history. The instances of the Upper-level ontology are enriched with links to DBpedia concepts, which extend and enrich the middle layer with additional the interoperable properties.

The need for an additional level of vocabulary-based semantics, which is not covered by CIDOC-CRM, is addressed by the side section, which provides thesauri and glossary concepts to the middle and top layer of the architecture based on SKOS definitions. The side section accommodates the CrossCult Classification Scheme (CCCS), which is described as a faceted vocabulary structure. It aggregates terminology from standard thesauri resources such as, the Arts and Architecture Thesaurus of Getty (AAT), the EUROVOC, the UNESCO Thesaurus and the Library of Congress Subject Authorities (LC) vocabulary. The CCCS incorporates a limited number of CrossCult specific terminology entries, designed to accommodate particular needs of the project, such as types of multimedia elements, types of dimension and other.

The top section of the architecture contains the Venue and the User ontologies. The Venue ontology is a fully CIDOC-CRM compliant structure, which aims to model the spatial arrangements of the different venues that participate in the project. Key elements of the Venue ontology are the subclasses of the E18.Physical Thing and the E53.Place class, which is used for modelling the spatial elements of a venue. Place instances can be combined together to form larger spaces whereas spatial coordinates, dimensions, conditions and appellations are used for modelling the details of such spaces.

The User ontology is a CrossCult centric structure aimed at supporting the user modelling requirements of the project with respect to the user interests, visit experience, user background and other demographic information. The ontology combines elements from the Friend of a Friend (FOAF) and CIDOC-CRM models while it introduces project-specific classes and properties to address particular user modelling requirements, such as fatigue, prior knowledge, behaviour, etc.

3. Upper-level Ontology

3.1. Aim and Design Rationale

The aim of the Upper-level ontology is to accommodate the semantic description of the cultural heritage resources of the four CrossCult pilots. The project reviewed the pilots' datasets as they are discussed in deliverable D2.1 and engaged in a series of meetings before concluding to a set of common entities across the four pilots, which are discussed below.

3.1.1. Reflective Topic

This is a central concept to the CrossCult experience that carries the semantics of all those subjects and topics of interest that drive the reflection and reinterpretation qualities of the application. The notion of Reflective Topic acts as a hub that could connect under a common theme physical items, multimedia content and users. Specialisations of the class can be topics like Immigration, Women in Society, Healing, Painting Style, etc. The Reflective Topic could be related a range of ontology classes via properties which will realise relationships of interest. For example, the topic of Healing could be linked to a Place (spa), a Physical Item (fountain), a Date (1st Century BC), Multimedia Content (image) and a User Interest (healthcare).

3.1.2. Physical Item

A Physical item is any museum artefact, painting, venue item or landmark that has some relation to the reflection topics and can be associated with one or more MM Contents. Specialisations of the class can be things like a painting, a public square, a museum exhibit, etc. Several other non-top level classes will be related to the class such as material, title, dimension, etc.

3.1.3. Digital MM Content

Any MM (Multi-Media) content could relate to one or more Physical Items. Subclasses can be realised as Image, Text, Video, etc. Some existing metadata such as copyright, title, creator, etc., can be accommodated by specialised classes.

3.1.4. Place

Places in the ontology have a spatial focus, which could refer to the location of an object in space, a place of artefact production, a place of item display, a depicted place on a painting, etc. The properties of a Place and its relationships to the ontology classes dictate the type of a Place, for example, the location of an object, place of origin, place depicted on artwork, etc.

3.1.5. Actor

An Actor can be any person or organisation that relates to one or more ontology classes. The properties of an Actor and its relationships to the ontology classes describe its type; for example, creator of a painting, a historic person related with an artefact, a user interested in an exhibit and an institution owning an artwork.

3.1.6. Temporal Entity

Any form of temporal definition such as date and period that relates to one or more ontology classes. The properties of a Temporal entity and its relationships to the ontology classes describe its type; for example, the date of an artist's birth, the production date of an artefact, the historic period of an exhibit, the date of a visit, etc.

3.2. Ontology Structure

To address the data modelling requirements discussed in the previous section, we defined the Upper-level ontology as a subset of CIDOC-CRM enhanced with additional semantics from the SKOS and Dublin Core Schemas. The structure maintains full compatibility with CIDOC-CRM containing the least minimum set of CRM concepts as described in the latest specification document version 6.2.3[2]. In addition, the structure contains SKOS and Dublin-Core concepts and introduces some project-specific entities to capture the notion of Reflective Topic. In the rest of this section, we present the structure of the ontology (the class and property hierarchies). A diagram presenting the core elements (classes and properties) of the ontology is depicted in Figure 13 (see Appendix). Specific data modelling examples from the four pilots are presented in Deliverables D2.2 and D2.4.

3.2.1. Class Hierarchy

Table 1 depicts the hierarchy of a subset of the classes of the CrossCult Upper-level ontology, specifically those that we regard as more important for the needs of the project.

Table 1: CrossCult Ontology Entity Structure (project-specific classes are in bold)

owl:Thing							
-	**Reflective Topic**						
-	skos:Concept						
-	skos:Concept Scheme						
-	E1 CRM Entity						
-	-	E2 Temporal Entity					
-	-	-	E4 Period				
-	-	-	-	E5 Event			
-	-	-	-	-	E7 Activity		
-	-	-	-	-	-	E11 Modification	
-	-	-	-	-	-	-	E12 Production
-	-	-	-	-	-	E13 Attribute Assignment	
-	-	-	-	-	-	E65 Creation	
-	-	-	-	-	E63 Beginning of Existence		
-	-	-	-	-	-	E12 Production	
-	-	-	-	-	-	E65 Creation	
-	-	-	-	-	-	E67 Birth	
-	-	-	-	-	E64 End of Existence		
-	-	-	-	-	-	E6 Destruction	
-	-	-	-	-	-	E68 Dissolution	
-	-	-	-	-	-	E69 Death	
-	-	E77 Persistent Item					
-	-	-	E70 Thing				
-	-	-	-	E72 Legal Object			
-	-	-	-	-	E18 Physical Thing		

[2] http://www.cidoc-crm.org/sites/default/files/2017-12-30%23CIDOC%20CRM_v6.2.3_esIP.pdf

-	-	-	-	-	-	E19 Physical Object		
-	-	-	-	-	-	-	E20 Biological Object	
-	-	-	-	-	-	-	-	E21 Person
-	-	-	-	-	-	-	E22 Man Made Object	
-	-	-	-	-	-	-	-	E84 Information Carrier
-	-	-	-	-	-	E24 Physical Man-Made Thing		
-	-	-	-	-	-	-	E22 Man Made Object	
-	-	-	-	-	-	-	-	E84 Information Carrier
-	-	-	-	-	-	-	E25 Man Made Feature	
-	-	-	-	-	-	-	E76 Collection	
-	-	-	-	-	-	E26 Physical Feature		
-	-	-	-	-	-	-	E25 Man Made Feature	
-	-	-	-	-	-	-	E27 Site	
-	-	-	-	-	E90 Symbolic Object			
-	-	-	-	E71 Man-Made Thing				
-	-	-	-	-	E24 Physical Man-Made Thing			
-	-	-	-	-	-	E22 Man Made Object		
-	-	-	-	-	-	-	E84 Information Carrier	
-	-	-	-	-	-	E25 Man Made Feature		
-	-	-	-	-	-	E76 Collection		
-	-	-	-	-	E28 Conceptual Object			
-	-	-	-	-	-	E89 Propositional Object		
-	-	-	-	-	-	-	E30 Right	
-	-	-	-	-	-	-	E73 Information Object	
-	-	-	-	-	-	E90 Symbolic Object		
-	-	-	-	-	-	-	E41 Appellation	
-	-	-	-	-	-	-	E73 Information Object	
-	-	-	-	-	-	E55 Type		
-	-	-	E39 Actor					
-	-	-	-	Group				
-	-	E52 Time-Span						
-	-	E53 Place						
-	-	E54 Dimension						

3.2.2. Property Hierarchy

Table 2 depicts the hierarchy of a subset of the properties of the CrossCult ontology, specifically those that we regard as more important for the needs of the project.

Table 2: CrossCult Ontology Property Structure (project-specific properties are in bold)

Property Name	Domain	Range
reflects	Reflective Topic	owl:Thing
is reflected by	owl:Thing	Reflective Topic
has first	Reflective Topic	Reflective Topic
has last	Reflective Topic	Reflective Topic
has next	Reflective Topic	Reflective Topic
has previous	Reflective Topic	Reflective Topic
is first of	Reflective Topic	Reflective Topic

Property Name	Domain	Range
is last of	Reflective Topic	Reflective Topic
skos:has member	skos:Concept	skos:Concept
skos:has member list	skos:Concept	skos:Concept
skos:has top concept	skos:Concept	skos:Concept
skos:is in scheme	skos:Concept	skos:Concept Scheme
- skos:is top concept in scheme	skos:Concept	skos:Concept Scheme
skos:is in semantic relation	skos:Concept	skos:Concept
- skos:has broader transitive	skos:Concept	skos:Concept
- - skos:has broader	skos:Concept	skos:Concept
- skos:has narrower transitive	skos:Concept	skos:Concept
- - skos:has narrower	skos:Concept	skos:Concept
- skos:has related	skos:Concept	skos:Concept
- skos:is in mapping relation with	skos:Concept	skos:Concept
P1 is identified by (identifies)	E1 CRM Entity	E41 Appellation
P2 has type (is type of)	E1 CRM Entity	E55 Type
P4 has time-span (is time-span of)	E2 Temporal Entity	E52 Time-Span
P7 took place at (witnessed)	E4 Period	E53 Place
P12 occurred in the presence of (was present at)	E5 Event	E77 Persistent Item
- P11 had participant (participated in)	E5 Event	E39 Actor
- - P14 carried out (performed) by	E7 Activity	E39 Actor
- P16 used specific object (was used for)	E7 Activity	E70 Thing
- P31 has modified (was modified by)	E11 Modification	E24 Physical Man-Made Th.
- - P108 has produced (was produced by)	E12 Production	E24 Physical Man-Made Th.
- P92 brought into existence (was brought into existence by)	E63 Beginning of Existence	E77 Persistent Item
- - P108 has produced (was produced by)	E12 Production	E24 Physical Man-Made
- - P94 has created (was created by)	E65 Creation	E28 Conceptual Object
- P93 took out of existence (was taken out of existence by)	E64 End of Existence	E77 Persistent Item
P15 was influenced by (influenced)	E7 Activity	E1 CRM Entity
- P16 used specific object (was used for)	E7 Activity	E70 Thing
P20 had specific purpose (was purpose of)	E7 Activity	E5 Event
P43 has dimension (is dimension of)	E70 Thing	E54 Dimension
P46 is composed of (forms part of)	E18 Physical Thing	E18 Physical Thing
P59 has section (is located on or within)	E18 Physical Thing	E53 Place
P67 refers to (is referred to by)	E89 Prop. Object	E1 CRM Entity
P75 possesses (is possessed by)	E39 Actor	E30 Right
P89 falls within (contains)	E53 Place	E53 Place
P104 is subject to (applies to)	E72 Legal Object	E30 Right
P106 is composed of (forms part of)	E90 Symbolic Object	E90 Symbolic Object

Property Name	Domain	Range
P107 has current or former member (is current or former member of)	E74 Group	E39 Actor
P127 has broader term (has narrower term)	E55 Type	E55 Type
P128 carries (is carried by)	E18 Physical Thing	E90 Symbolic Object
P130 shows features of (features are also found on)	E70 Thing	E70 Thing
P140 assigned attribute to (was attributed by)	E13 Attribute	E1 CRM Entity
P141 assigned (was assigned by)	E13 Attribute	E1 CRM Entity
P148 has component (is component of)	E89 Prop. Object	E89 Propositional Object

3.2.3. Ontological Representation of Pilots' Cultural Heritage Objects

Figure 2 presents the modelling arrangements of the common semantics across the four project pilots for modelling cultural heritage objects. At the core of the model resides the CIDOC-CRM entity *E18 Physical Item*, which comprises all persistent physical items with a relatively stable form, man-made or natural. The entity enables the representation of a vast range of items of interest, such as museum exhibits, gallery paintings, artefacts, monuments and points of interest, whilst providing extensions to specialised entity definitions of targeted semantics for man-made objects, physical objects and physical features. The arrangement benefits from a range of relationships between *E18 Physical Item* and a set of entities that describe the static parameters of an item, such as dimension, unique identifier, title, and type. The model also allows the description of more complex objects through a composition of individual items (i.e. P46 is composed of). Moreover, the well-defined semantics enable rendering of rich relationships between the physical item and entities describing the item in terms of ownership, production, location, and other conceptual associations. The project-specific property *reflects* enables specific, direct connections between existing concepts and the CrossCult class *Reflective Topic*.

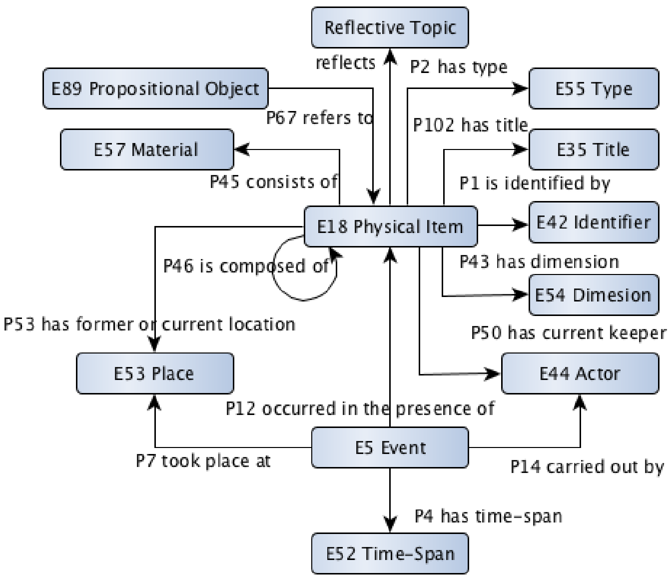

Figure 2: Elements of the Upper-level Ontology for modelling the pilots' cultural heritage objects

3.3. The Semantics of Reflective Topic

As shown in Tables 1 and 2, further to the elements of CIDOC-CRM, the Upper-level ontology contains the project-specific class (*Reflective Topic*) and a set of adjunct properties that support its conceptual arrangements. The class encompasses all those connections that can be made to create a network of points of view, aiding reflection and prospective interpretation over a topic that enable interconnection between physical or conceptual things of man-made or natural origin. Examples of reflective topics in CrossCult are "Daily Life", "Migration and Industrial Revolution in Europe", "Mortality and Immortality", "Religion and Pilgrimage" and others.

The definition of a reflective topic as an abstract theme (e.g. Daily Life) requires an additional form of contextualisation for accommodating a range of semantics relevant to a theme. As illustrated in Figure 3, the class connects to other Upper-level ontology classes via a set of well-defined semantics, some of which constitute project-specific extensions of standard CIDOC-CRM properties. In detail, the *Reflective Topic* class can be understood as extension of the *E89.Propositional Object*[3] class, extended by the project-specific property *reflects* (and its inverse property, *is reflected by*). The property sets a reflective topic instance as the primary subject of reflection of a physical or conceptual source. For example, the Eiffel tower can be used to drive a reflection about engineering and industrial revolution, hence, the physical object "Eiffel tower" -*reflects*- the Reflective Topic "Engineering Marvels of Europe".

A broader reflective topic can be composed by more specific (narrower) topics. The property *P148.has_component* allows for this kind of recursive composition, which can be experienced sequentially via the semantics of the *has_first*, *has_next* and *has_last* properties (and their inverse properties, *is_first_of*, *has_previous* and *is_last_of*). Multimedia elements, modelled as *E73.Information Object*, contextualise a topic by accommodating text and audio-visual materials. Such elements fell into three categories, which are distinguished via properties. The *P67_2_has_media* property is the most generic and is assigned to elements that simply complement the topic. The *P67_3_has_intro* property is assigned to those media that introduce the topic or act as a trigger for engaging with a topic, and the *P67_4_has_narrative* is assigned to the elements that drive reflection through a narrative which can have a textual or an audio form. A reflective topic is further contextualised by the contents of a title (*E35.Title)* and by linking to CCCS terms (*skosConcepts*), which they provide an additional layer of semantics, organised as a subject heading vocabulary scheme of broader and narrower concepts.

[3] This class comprises immaterial items, including but not limited to stories, plots, procedural prescriptions, algorithms, laws of physics or images that are, or represent in some sense, sets of propositions about real or imaginary things and that are documented as single units or serve as topic of discourse

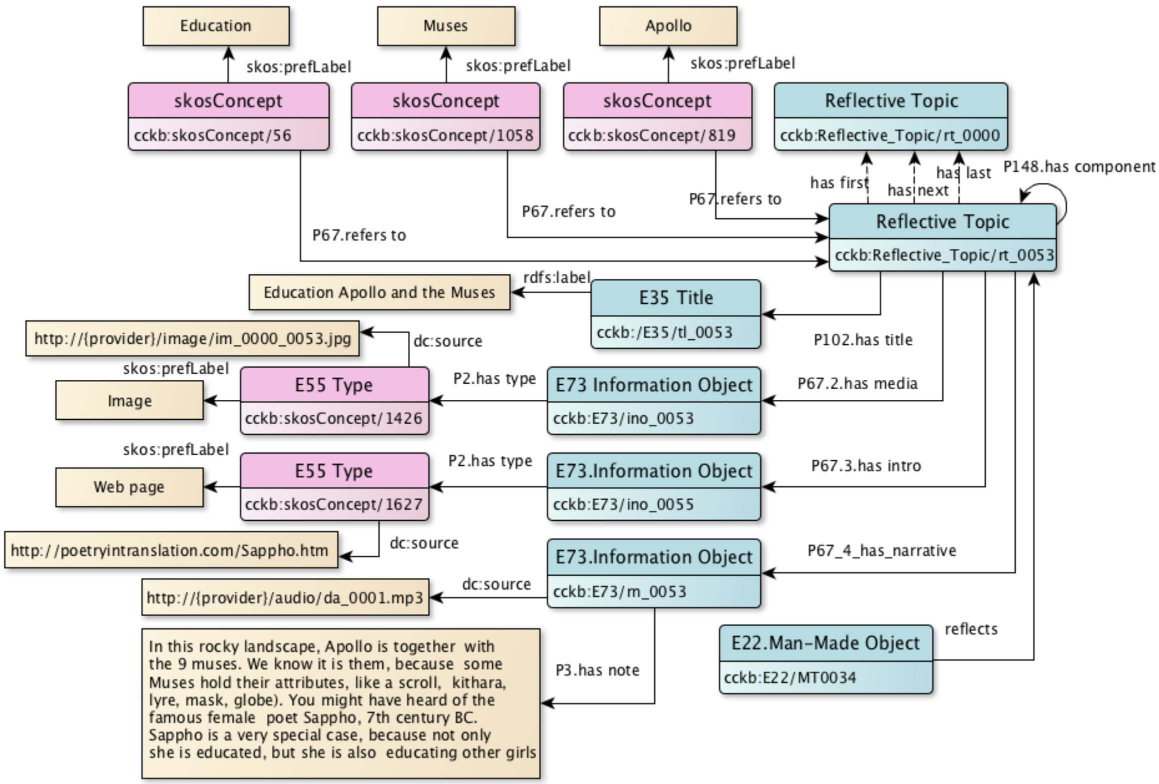

Figure 3: The semantics of Reflective Topic through an example.

4. The Venue Ontology

4.1. Aim and Design Rationale

The CrossCult Venue Ontology is a fully CIDOC-CRM compliant structure, which aims to model the spatial arrangements of the different venues that participate in the four project pilots: the National Gallery in London (pilot 1); the archaeological site of the Roman healing spa of Lugo in Spain, the archaeological sites of Chaves in Portugal and Montegrotto Terme in Italy; the Ancient theatre of Epidaurus in Greece (pilot 2); the Museum of Tripolis in Greece (pilot 3); and the Cities of Luxemburg in Luxemburg and Valetta in Malta (pilot 4).

4.1.1. The Venues Characteristics

The venues of the four pilots can be clustered broadly as indoor and outdoor "exhibitions" of POIs (Points of Interest), with similar characteristics:

i) Pilot 1, an indoor gallery with a large multi-thematic collection spread over 66 rooms and 2 floors.

ii) Pilot 2, four open air archaeological sites with location and POIs alterations over the various historical periods starting from the classical period and the Roman times.

iii) Pilot 3, a small museum with dense displays of archaeological exhibits confined in a small number of rooms.

iv) Pilot 4, two whole cities with disperse POIs located on façades of buildings, near bridges, in crossroads, near statues, on top of columns etc.

Although the purposes of the different venues are quite different, they are characterised by similarities that allow the construction of a common model that describes their spatial arrangements. The semantic representation of the city's structure conceptualised as an outdoor exhibition has similar characteristics to the indoor gallery and the small museum. It is composed of sections filled with other elements; for example, buildings composed of walls, floors, ceilings- that have dimensions and materiality- windows and doorways – spaces that are completely void.

In all venues the POIs, within a building or outdoors, are also characterised by events; POIs are moved from one location to another to serve for example the needs of exhibitions. They are also moved to receive treatment or for the needs of rehanging or changing the display of objects at a specific part of the building's structure.

Finally, the POIs move as the city's structure changes or as the result of constant alterations throughout time. Historic buildings and archaeological venues are, in most cases, the result of a series of matter addition and removal due to construction and destruction activities that modified their appearance over the various historical periods. The identification of these processes, together with the analysis of the different building techniques and the materials utilised over its existence, provides historians with an understanding of the continuity and discontinuity of matter and activities on a built structure.

All these strands of information can be used to produce a detailed understanding of the development of the historical provenance of any building, whether standing or in ruins, and to identify significant phases of the monument's appearance throughout the centuries.

4.1.2. The Ontology Development Process

Metadata standards for the documentation of the built heritage and the archaeological complexes attempt to record the semantics of the building's components but fail in describing the completeness of information about the building and the relationships among the parts and the whole. The aims of the conceptual model of venues in CrossCult are to i) describe and understand the correlation between the parts of a building and the whole; ii) to record and express the semantic relationships among the building components with the building as a whole; and iii) to be able to record the accurate location of objects within space and capture their provenance in terms of changes of location.

The process of building the Venue Ontology involved first developing the appropriate underlying conceptual model to support the requirements of the four venues and, second, populating the model with sufficient detail to realise its full potential. We kept the resulting model as generic as possible and we progressed with the task of populating the model with examples. The data for populating the ontology came from a variety of sources and differed in their underlying structures, accuracy and the level of detail in the representation of the places. Therefore, as more data was included in the process, the model was further specialised to meet the specific needs of each Venue.

The proposed CrossCult Venue Ontology attempts to address these emerging data modelling requirements and has been inspired from the CIDOC-CRMba, an extension of CIDOC CRM that has been proposed for approval by CIDOC CRM-SIG to support buildings archaeology documentation[4]. We decided on CIDOC CRM as the integrating framework, as a sensible first step on the road to interoperability. From the modelling process outlined above, we concluded that the resulting Venue Ontology does cover the basic needs and characteristics of the four pilot venues in terms of their spatial arrangements.

Finally, if we need to scope the needs of all our indoor and outdoor venues in more detail and cater for additional functionalities (for example model the spatial semantics related to the alterations of buildings that modified their appearance over the various historical periods), then the Venue Ontology should be enhanced with additional classes and properties from the CIDOC-CRMba (presented in Figure 4). The CIDOC-CRMba incorporates parts of the CRMgeo, a detailed model of generic spatio-temporal topology and geometric description [3]; parts of CRMsci, a model for scientific observation, measurements and processed data in descriptive and empirical sciences (such as biology, geology, geography, cultural heritage conservation, etc.); and CRMarcheo, a model developed for the documentation of archaeological excavations.

[4] http://icom.museum/resources/publications-database/publication/definition-of-the-crmba-an-extension-of-cidoc-crm-to-support-buildings-archaeology-documentation/print/1/

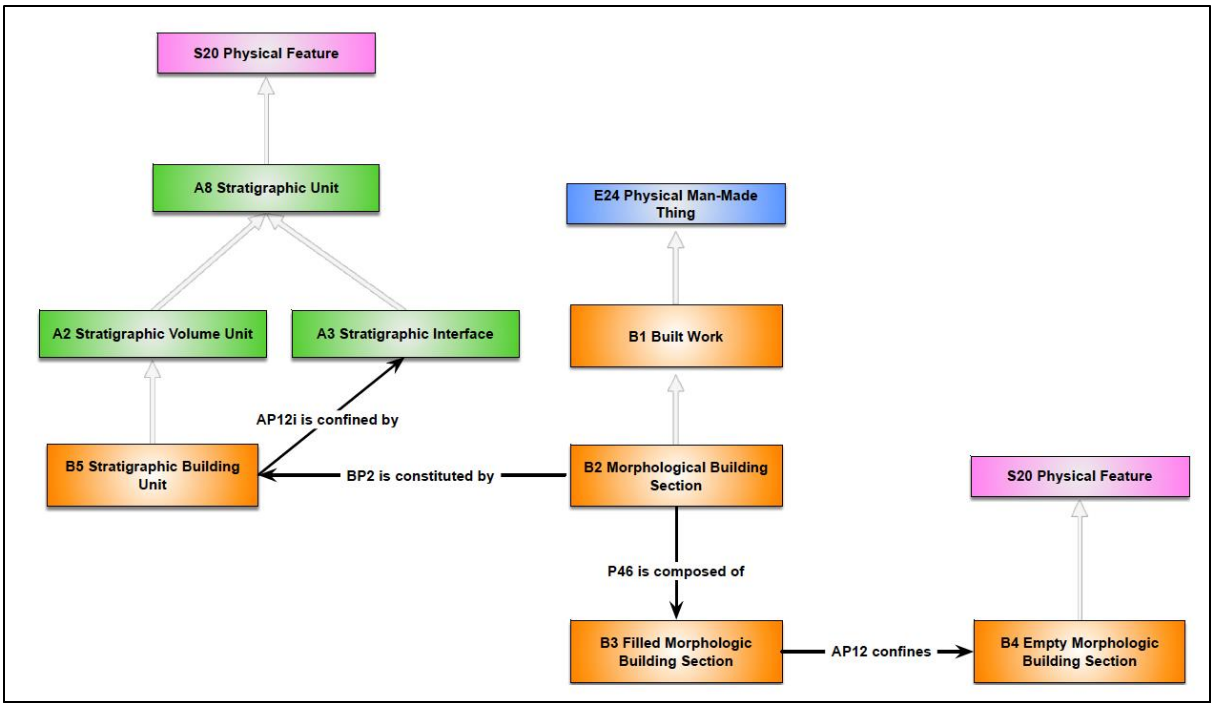

Figure 4: The CRMba conceptual model structure[5]

4.2. Ontology Structure

To address the data modelling requirements discussed in 4.1, we defined the Venue Ontology as a subset of CIDOC-CRM. Similar to the Upper- level Ontology, the structure maintains full compatibility with CIDOC-CRM containing the least minimum set of CRM concepts as described in the latest specification document version 6.2.3[6].

In the rest of this section, we present the structure of the venue ontology (the class and property hierarchies) and Figure 5 depicts its graphical representation. Specific data modelling examples from the four pilots are presented in Deliverables D2.2 and D2.4.

[5] Picture taken from http://www.cidoc-crm.org/crmba/sites/default/files/2016-12-3%23CRMba_v1.4.1_UR.pdf
[6] http://www.cidoc-crm.org/sites/default/files/2017-12-30%23CIDOC%20CRM_v6.2.3_esIP.pdf

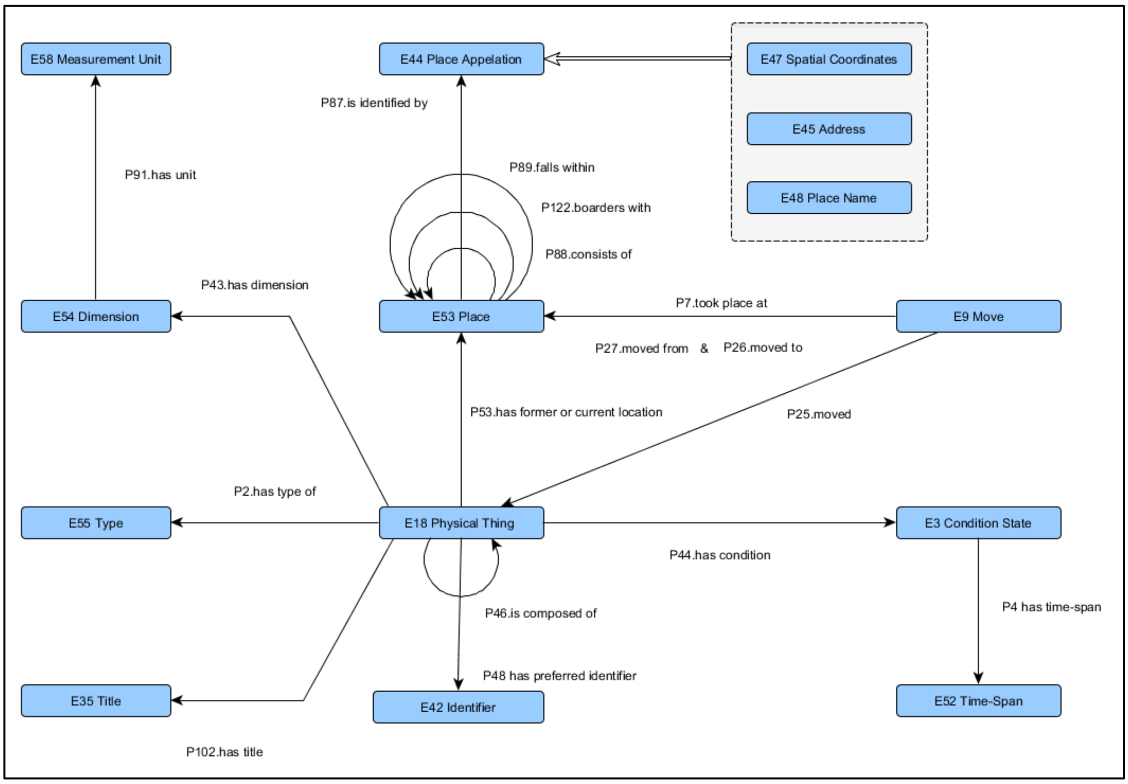

Figure 5: The top level view of the Venue Ontology

4.2.1. Class Hierarchy

Table 3 depicts in bold the complementary to the Upper level Ontology hierarchy subset of the classes of the CrossCult Venue Ontology (for reference the full table is available in section **Error! Reference source not found.**).

Table 3: CrossCult Venue Ontology Entity Structure

owl:Thing							
-	*E1 CRM Entity*						
-	-	*E2 Temporal Entity*					
-	-	-	**E3 Condition State**				
-	-	-	*E4 Period*				
-	-	-	-	*E5 Event*			
-	-	-	-	-	-	**E9 Move**	
-	-	-	-	-	*E18 Physical Thing*		
-	-	-	-	-	-	*E19 Physical Object*	
-	-	-	-	-	-	-	*E22 Man Made Object*
-	-	-	-	-	-	*E24 Physical Man-Made Thing*	
-	-	-	-	-	-	-	*E22 Man Made Object*
-	-	-	-	-	-	-	*E25 Man Made Feature*
-	-	-	-	-	-	-	*E76 Collection*
-	-	-	-	-	-	*E26 Physical Feature*	
-	-	-	-	-	-	-	*E25 Man Made Feature*
-	-	-	-	-	-	-	*E27 Site*
-	-	-	-	*E71 Man-Made Thing*			
-	-	-	-	-	-	*E24 Physical Man-Made Thing*	

-	-	-	-	-	-	-	*E25 Man Made Feature*		
-	-	-	-	-	-	-	*E76 Collection*		
-	-	-	-	-	-	-	*E22 Man Made Object*		
-	-	-	-	-	-	*E55 Type*			
-	-	-	-	-	-	-	*E41 Appellation*		
-	-	-	-	-	-	-	-	**E44 Place Appellation**	
-	-	-	-	-	-	-	-	-	**E48 Place Name**
-	-	-	-	-	-	-	-	-	**E47 Spatial Coordinates**
-	-	-	-	-	-	-	-	**E51 Contact Point**	
-	-	-	-	-	-	-	-	-	**E45 Address**
-	-	-	-	-	-	-	**E58 Measurement Unit**		
-	-	*E52 Time-Span*							
-	-	*E53 Place*							
-	-	*E54 Dimension*							

4.2.2. Property Hierarchy

Table 4 depicts in bold the complementary to the Upper level Ontology hierarchy subset of the of the CrossCult Venue Ontology properties (for reference the full table is available in **Error! Reference source not found.**).

Table 4: CrossCult Venue Ontology Property Structure

Property Name	**Domain**	**Range**
P1 is identified by (identifies)	*E1 CRM*	*E41 Appellation*
- **P87 is identified by (identifies)**	**E53 Place**	**E44 Place Appellation**
P2 has type (is type of)	*E1 CRM Entity*	*E55 Type*
P4 has time-span (is time-span of)	*E2 Temporal Entity*	*E52 Time-Span*
P7 took place at (witnessed)	*E4 Period*	*E53 Place*
P8 took place on or within	**E4 Period**	**E53 Place**
- **P26 moved to**	**E9 Move**	**E53 Place**
- **P27 moved from**	**E9 Move**	**E53 Place**
P12 occurred in the presence of (was present at)	*E5 Event*	*E77 Persistent Item*
- **P25 moved**	**E9 Move**	**E19 Physical Object**
P20 had specific purpose (was purpose of)	*E7 Activity*	*E5 Event*
P43 has dimension (is dimension of)	*E70 Thing*	*E54 Dimension*
P44 has condition (is condition of)	**E18 Physical Thing**	**E3 Condition State**
P46 is composed of (forms part of)	*E18 Physical Thing*	*E18 Physical Thing*
P53 has former or current location (is former or current location of)	**E18 Physical**	**E53 Place**
P59 has section (is located on or within)	*E18 Physical Thing*	*E53 Place*
P67 refers to (is referred to by)	*E89 Prop. Object*	*E1 CRM Entity*
P89 falls within (contains)	*E53 Place*	*E53 Place*
P91 has unit (is unit of)	**E54 Dimension**	**E58 Measurement Unit**
P122 borders with	**E53 Place**	**E53 Place**

Property Name	Domain	Range
P1 is identified by (identifies)	*E1 CRM*	*E41 Appellation*
P148 has component (is component of)	*E89 Prop. Object*	*E89 Propositional Object*

4.3. Discussion on the Venue Ontology

4.3.1. Representation of Venues as Physical Things

Major components of the Venue ontology arrangements are the subclasses of the *E18.Physical Thing*, *E19.Physical Object*, *E26.Physical Feature* and *E24.Physical Man Made Thing*, which are used to model physical objects and features as well as man-made structures. Physical thing and Physical man-made thing Instances such as a "Building", a "Room", a "Floor", a "Wall", etc. can also be combined together to form more complex structures. These classes are further related to other ontology classes to model the physical and man-made structures' dimensions, conditions or events. The class *E.55Type* has also been employed to differentiate between the functionalities of a room in a museum as a "Gallery", a "Cafe", a "Temporary exhibition" room, etc.

4.3.2. Representation of Venues as Places

Complementary to the notion of the *E19.Physical Object* and *E24.Physical Man Made Thing* classes is the *E53.Place* class, which is used to model the different types of the venue spaces. Place instances can be combined together to form complex spaces, whereas spatial coordinates and appellations are used to model the details of such spaces.

4.3.3. Movement of Things within the Venues

We use the *E9.Move* class to describe changes of the physical location of the instances of *E19.Physical Object*, for example the movement of a painting from one room to another. This class inherits the property *P7_took_place_at (witnessed),* which has range *E53.Place.* We use this property to describe the larger area within which a move takes place, whereas the properties *P26.moved_to (was_destination_of)* and *P27.moved_from* (*was_origin_of*) describe the start and end points only. For example, (E9) "Movement of the painting" moved the (E19) "Painting"; (E53) "East Wall location" is the origin of the (E9) "Movement of the painting" and (E53) "West wall location" is the destination of the movement; the (E9) "Movement of the painting" took place at (E53) "the location of Room 9". In some cases, we can also use the *P8.took_place_on* or *within (witnessed)* which has range *E19.Physical Object.* This property is in effect a special case of *P7.took_place_at* and we can use it to describe, for example, a movement that can be located with respect to the space defined by an *E19.Physical Object* such as a "Building", a "Room" or a "Wall".

5. The User Ontology

The User ontology (Figure 6) allows the modelling of users involved in activities with CrossCult applications. It is grounded on CRM-CIDOC and the Friend-of-a-Friend (FOAF) ontology, which is commonly used for user modelling in the Linked Data Cloud, while introducing new concepts. It is centred around the concept of *cc:Visitor*, which is a subclass of *cc:User*. Both are linked to *foaf:Person* in a specific way formalising their semantics, and *foaf:Person* is mapped to E21 Person in CRM. A visitor is a person who does a visit: <*foaf:Person*, *P14 performs*, *cc:Visit*>, while a user is a person that uses an application: <*foaf:Person*, *cc:uses*, *cc:Application*>. The ontology then models characteristics of a person, and characteristics that are specific to a person as a visitor. Since a visitor is also a user, the application(s) they use can be specified if needed.

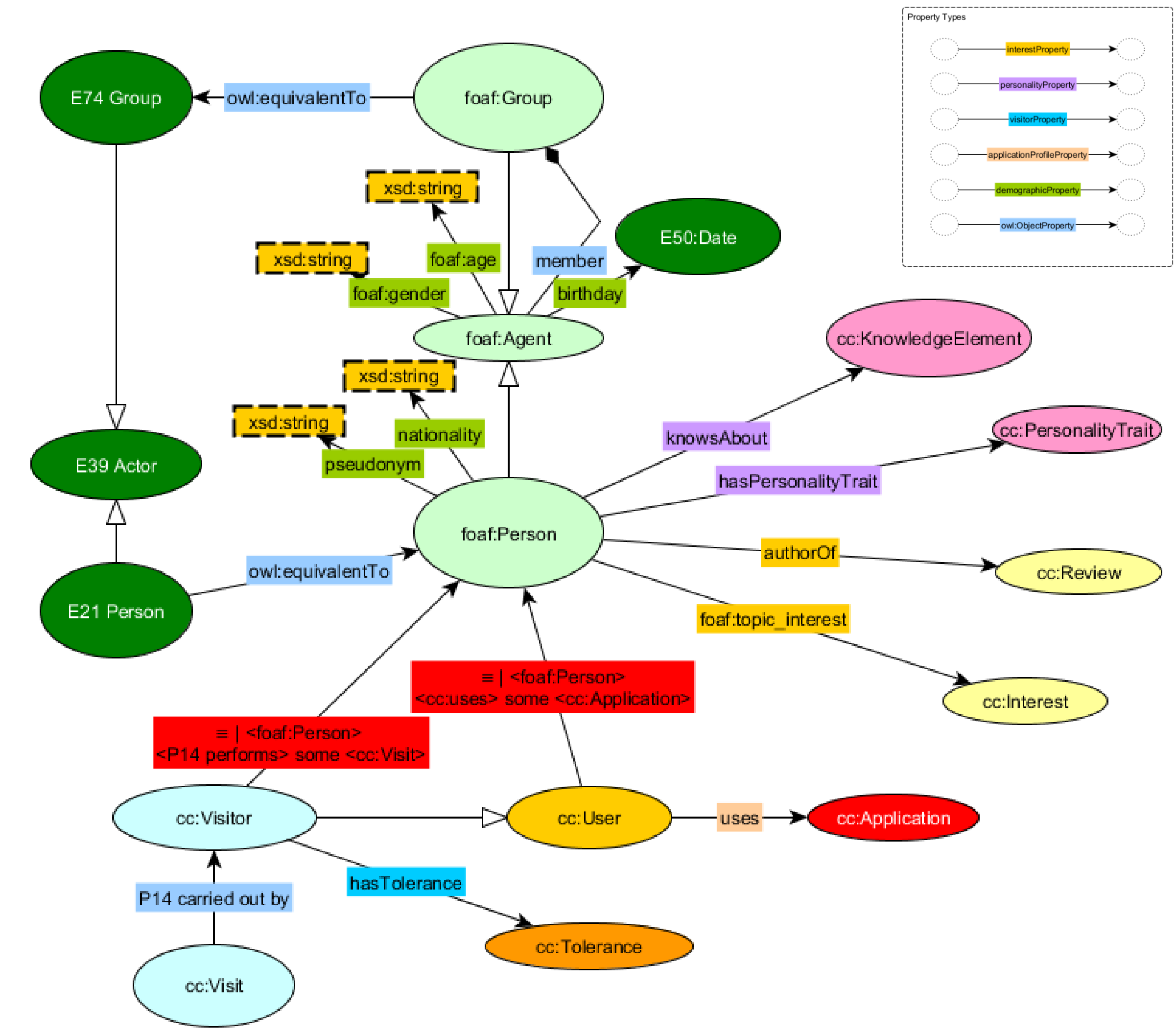

Figure 6: User ontology main concepts

Although any property of FOAF can be used, we reuse in particular those properties that capture demographic characteristics such as gender or age, which we completed with birthday, nationality and *cc:pseudonym*, which allows mapping a person to a pseudonym instead of a name. Although when asked to give a name, anyone is free to give the actual or a false one, using explicitly the term pseudonym conveys a clear semantics asking more for a false name. It can also be filled-in automatically by applications to serve as a pseudonymous identifier. Persons can be attached to groups using the member property of *foaf:Group* (equivalent to *E74 Group*).

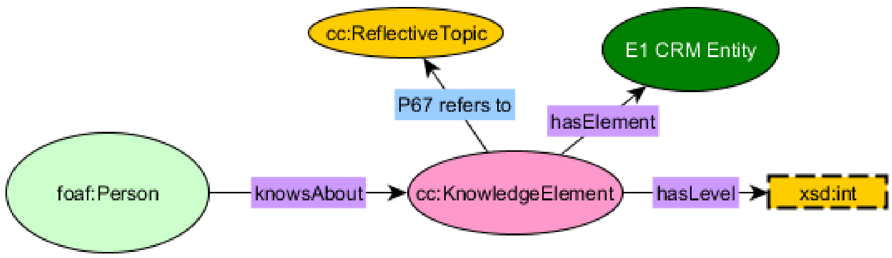

Figure 7: Concepts related to a person's knowledge

The knowledge possessed by a person can be formalised with <*foaf:Person*, *cc:knowsAbout*, *cc:KnowledgeElement*> (Figure 7). A knowledge element is a *E1* CRM *Entity*, referring to a *cc:ReflectiveTopic*, and to which a level can be attached, to quantify the expertise.

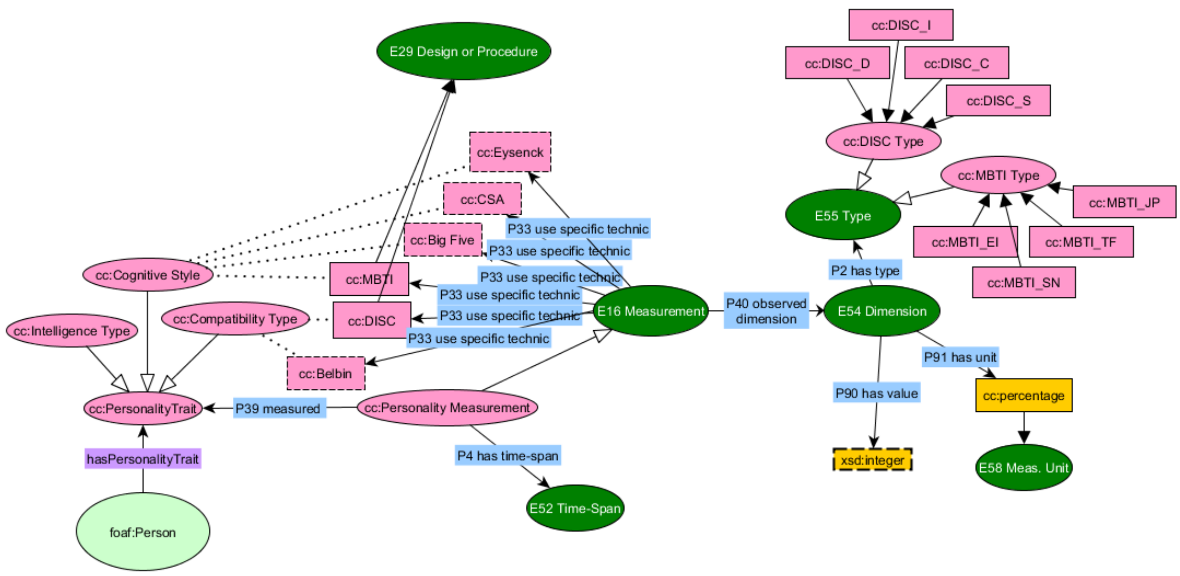

Figure 8: Concepts related to a person's personality

Another facet of a user as a person is their personality, which can be used for personalisation purposes. The CrossCult user ontology allows formalising different personality traits and their assessment with dedicated measurement methods, as depicted in Figure 8. Standard tools for evaluating the different kinds of personality traits are defined as instances of *E29 Design or Procedure*. It is expected that for the *cc:Personality Trait* class, only specific tools can be used (e.g. Cognitive Style can be measured using MBTI (Myers–Briggs Type Indicator) but not Compatibility Type). This is formalised as a constraint in the ontology. As an example, the cognitive style of a person as measured by a MBTI test can be modelled as follows:

<X, cc:hasPersonalityTrait, CS(X)>
<measCS(X), P39 measured, CS(X)>
<measCS(X), P33 use specific technique, cc:MBTI>

To formalise the values obtained for the different dimensions of a personality trait used in a particular method, we rely on new subclasses of *E55 Type*, which are specific to each method. Completing the preceding example, we would have the following for MBTI:

<measCS(X), P40 observed dimension, mbti_EI(X)>
<mbti_EI(X), P2 has type, cc:MBTI_EI>
<mbti_EI(X), P90 has value, 50>

The ontology models so far DISC (Dominance, Influence, Steadiness, Conscientiousness) and MBTI traits and their dimensions, as these are the ones used in our pilots so far. Others standard techniques are formalised as well, but not their dimensions, which are left for extension by who would need them.

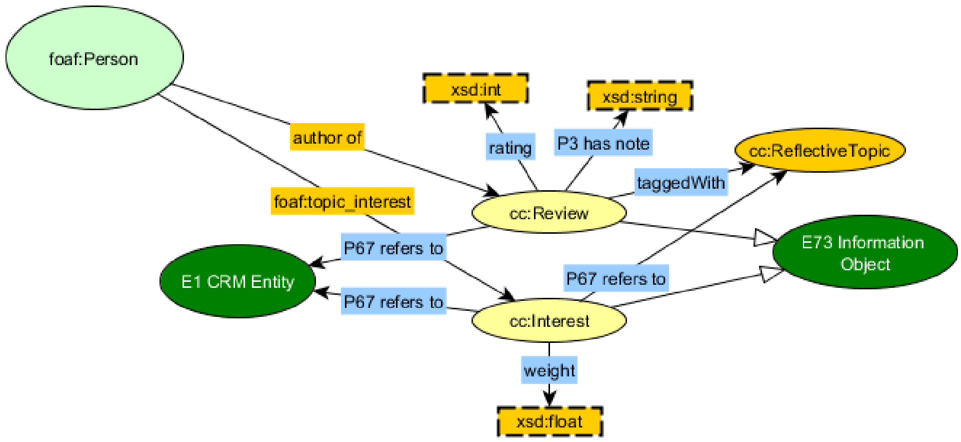

Figure 9: Concepts related to a user's interests and reviews given in an application

To allow for personalisation and evaluating reflection, the user profile captures user interests and reviews, respectively using the *cc:Interest* and *cc:Review* classes, which are both subclasses of *E73 Information Object* (Figure 9). They are formalised respectively with <*foaf:Person*, *foaf:topic_interest*, *cc:Interest*> and <*foaf:Person*, *author of*, *cc:Review*>. A *cc:Interest* has a weight attached and can refer to any cultural entity (*E1* CRM *Entity*) or a *cc: ReflectiveTopic*. A *cc:Review* can be either a rating or a detailed annotation (*P3 has note*). Reviews refer to cultural entities and can be tagged with instances of *cc:ReflectiveTopic*.

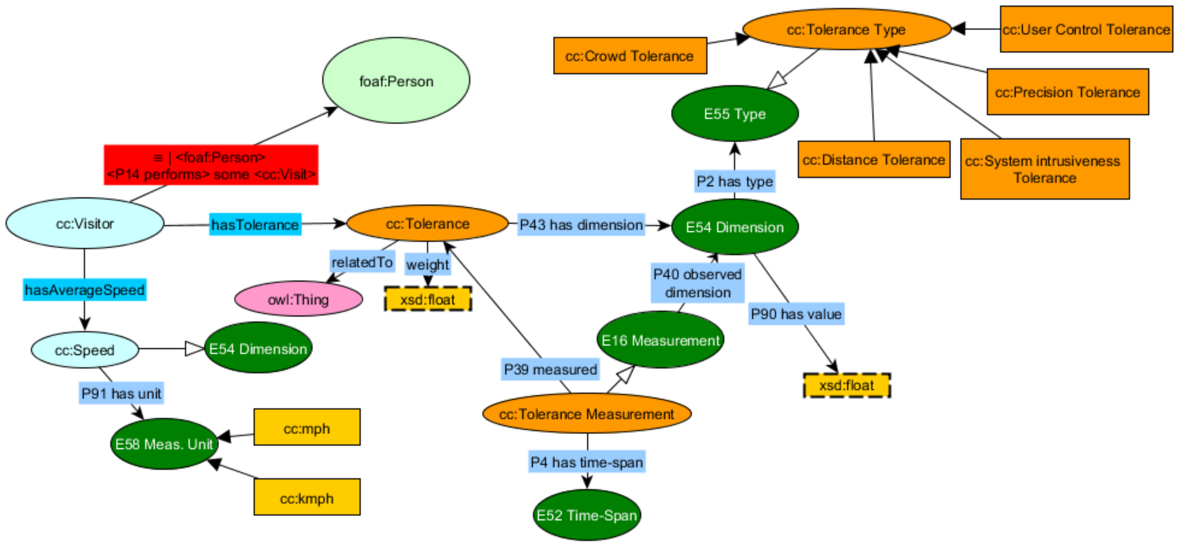

Figure 10: Concepts related to user's tolerances

We retained a set of features of a visitor that can be used for personalising the visit experience, and refer to the visitor and their visit (Figure 10). In addition to properties linked to *foaf:Person*, a *cc:Visitor* is characterised by an average speed (*cc:Speed*), which we assume true for all their visits. Of course, this is a modelling choice influenced by the needs of our applications. For a better granularity in the description, an extension of the user ontology could model an average speed for each visit or even an instantaneous speed between two elements of a route followed during a visit. The second kind of characteristics attached to a visitor are tolerances, for which we defined five types, using the same modelling structure used for personality traits: *cc:Distance Tolerance*, *cc:Crowd Tolerance*, *cc:System Intrusiveness Tolerance*, *cc: User Control Tolerance*, *cc: Precision Tolerance*. Details about their meaning can be found in [4]. A tolerance is instantiated as in the following example (only one dimension defined here, but multiple can be defined):

<X, cc:hasTolerance, Tol(X)>
<Tol(X), P43 has dimension, dimCrowdTol(X)>
<dimCrowdTol(X), P2 hasType, cc:Crowd Tolerance>
<dimCrowdTol(X), P90 has value, 0.6>

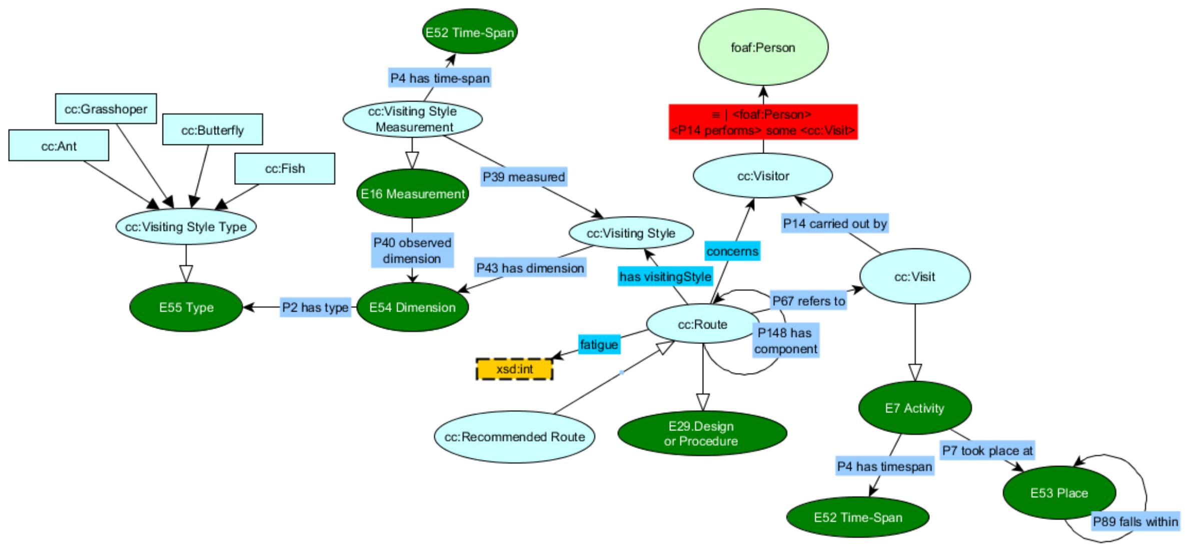

Figure 11: Concepts related to a person as a visitor and his/her visit

In the user ontology, a *cc:Visit* is modelled as an activity (subclass of *E7 Activity*) carried out by a visitor, which has a given timespan and happens at a specific place (Figure 11). A route that is followed by a visitor or recommended to them can also be formalised, with the *cc:Route* concept, a subclass of *E29 Design or Procedure*. A *cc:Route* is composed of other sub-routes (this is formalised with the *P148 component* property), with any possible granularity: a single visitor step can be considered as such a sub-route. To each component of a route (i.e. a sub-route), a fatigue value and a visiting style can be attached. Although they could have been attached to *cc:Visitor*, these are parameters that can vary from one visit to another and for each step of a visit. If it is evident that fatigue evolves during a visit, the theory of visiting styles [5] does not consider that the way visitors behave during a visit could be subject to changes. This is something that requires further study and the user ontology allows this possibility by attaching visiting style to a visit (more precisely a route followed by a visitor during a visit) instead of attaching it to a visitor. A visiting style can be instantiated as in the following example:

```
<Route(X), concerns, X>
<Route(X), has visitingStyle, VisitStyle(X)>
<VisitStyle(X), P43 has dimension, DimVisitStyle(X)>
< DimVisitStyle(X), P2 has type, cc:Butterfly>
```

6. The Crosscult Classification Scheme (CCCS)

6.1. Aim and Rationale

The CrossCult Classification Scheme (CCCS) is a terminological structure, supplementary to the ontology, which accommodates the keyword requirements of the project through a controlled vocabulary of concepts. The role of CCCS is not to classify objects according to their characteristics, which is handled by the ontology, but to provide a supplementary layer of terminology (as subjects, types etc.) that can be useful during retrieval. In this respect CCCS provides an additional access point for cultural heritage resources incorporated in the ontology, namely subject access focusing on enhancing and supporting reflection processes and social aspects of history instances. Another aim of the CCCS is to complement the flat list of Reflective Topics with vocabulary terms. Keywords from the CCCS can be linked with instances of Reflective Topic in order to provide to the topics additional vocabulary and conceptual meaning. Some relevant examples are provided in Table 5. Finally, while modelling the pilots' metadata, it became evident that apart from the terms representing subjects, other terms referring to time periods and types were of essence and were incorporated as separate facets.

Table 5: Examples of Associative Relationships between Reflective Topics and CCCS Terms

Reflective Topic	CCCS Terms
Human Senses and optical effects	Senses Sight
Objects as symbols in Altarpieces of the collection	Religious Work Iconography Altar Pieces
Costume_and_fashion_in_Paintings	Clothing
Thermal_Springs_and_Health	Healing Medical Treatment Healing Practices Mineral Water
Social Status	Rank People in power Status of Women Under authority
Thermal Bath Buildings	Hyperthermal Roman baths Roman bath spaces Mineral water Thermal bath buildings Pool buildings
Religion and Rituals	Births Votive offerings Deities Rites of passage People in religion
Religion and Thermal Springs	Shrines Ex voto shops Ablutions Ritual Events Sacred springs

	Deities Nymphae Apollo Asclepius Antiquity
The Role of women in society	Daily life of women Social Status
Medicine and healing in art	Illness Medical treatment Health
Human senses and audible performances	Senses Hearing
Colours and Pigments through the ages	Trade routes Pigment Colorant material Materials
Life in the Netherlands in 17th century	Daily life 17th century

6.2. Development methodology

The CCCS has the form of a faceted classification of terms. The development of CCCS relied to a large extent on the reuse of terms from standard and widely known controlled vocabularies. The reuse of standardised resources ensures the validity of the CCCS structure and the consistency in the use of its terms. CCCS incorporates terms from the following vocabularies.

- Terms referring to arts and cultural heritage elements (e.g. art collections, painters, iconography, etc.) have been drawn from AAT (http://www.getty.edu/research/tools/vocabularies/aat/). Due to the nature of the project, most of the keywords have been mapped to terms of this vocabulary.
- Terms related to policies, social issues (e.g. migration), politics and the environment have been drawn from EUROVOC (http://eurovoc.europa.eu), an EU Thesaurus that covers extensively the particular subjects.
- Terms related to social issues and social structures have been verified against the UNESCO Thesaurus (http://vocabularies.unesco.org/browser/thesaurus/)
- More specific terms that do not fall within any of the above controlled vocabularies have been verified against the LC Subject Authorities (http://id.loc.gov/authorities/subjects.html).

To ensure the comprehensiveness of CCCS and to maintain the project-specific focus of the terminology, the contributing terms were derived from the specifications of the four pilots (scenarios, games, narratives, etc.) and the descriptions of the pilots' cultural heritage objects, including their meaning, symbolism, materials, cultural context and construction techniques.

The general methodology for constructing CCCS consisted of the following steps:

S1. We identified the terms used in the four pilots (as explained above) and discard duplicate terms.

S2. We verified the terms against the authority vocabularies. For example, the term <*collectors*> used in pilot 1 is available from AAT (http://vocab.getty.edu/aat/300025234), the term <*emigration*> of pilot 4 available from EUROVOC (http://eurovoc.europa.eu/724), and the term <*women in society*> used in pilots 2 and 3 can be found in LC Authorities (https://lccn.loc.gov/n84736267).

S3. For each term that didn't directly match any of the terms in the authority vocabularies, we examined the best possible match. For example, the term <*Medical treatment*> is used in pilots 1 and 2 and is a reference term in both EUROVOC and LC with the indication to use the valid term <*therapeutics*>. In this case, the term <*therapeutics*> was incorporated in the structure as the preferred term.

S4. For terms that did not match any of the terms in the authority vocabularies nor had a possible close match such as above, we incorporated them as "project specific terms". For example, the term <*Gender specific education*> used in pilots 3 and 4 could not be matched to any of the controlled vocabulary sources, and was added as such and was placed in the facets of *Education* and *Gender*.

S5. We placed the terms in the CCCS hierarchy using appropriate SKOS properties (*has broader*, *has narrower, has close match*, etc.), taking into account the structures of the external vocabularies. This process is described in more detail below.

S6. We added further inter-term relationships, following the guidelines of the authority vocabularies. For example, the term <*Healing*> verified in LC is related to the term <*traditional medicine*> defined in AAT.

The hierarchy of the CCCS (i.e. the "broader term" and "narrower term" relationships among the CCCS terms) is consistent with the structures of AAT and EUROVOC, which comply with the ISO 25964 standards (http://www.niso.org/schemas/iso25964/) for thesaurus construction. We used the structures of the two vocabularies to construct CCCS, and specifically to:

- establish broader and narrower term relationships between the CCCS terms, following the AAT and EUROVOC hierarchies.
  For example, for the term <*pigment*> used in pilot 1, we established the following relationships, which are consistent with the AAT hierarchy.

  <*pigment*> (http://vocab.getty.edu/aat/300013109)
  has BT <*colorant (material)*> (http://vocab.getty.edu/aat/300013026)
  has NT <*blue pigment*> (http://vocab.getty.edu/aat/300013182)

  For a project-specific term, we first selected the relevant facet from AAT/EUROVOC and then selected the appropriate category, under which this term could be added. For example, the term <*women in society*> is not part of the authority records of AAT or EUROVOC but exists in LC Authorities. We could add the term under the broader EUROVOC category <*social status*>, therefore creating an original CrossCult structure within the Classification Scheme, i.e.

  <*social status*> (http://eurovoc.europa.eu/4277)
  has NT: <*women in society*> (https://lccn.loc.gov/n84736267)

- place at similar hierarchical level concepts of the same specificity. This was necessary as AAT and EUROVOC are fully developed vocabularies accommodating thousands of terms and are much broader in scope than CCCS. For example, the term <*costume and fashion*> has as narrower terms at the same level <*costume accessories*> and <*hair styles*>

Costume (mode of fashion) (http://vocab.getty.edu/aat/300178802)
has NT Costume accessories (http://vocab.getty.edu/aat/300209273)
has NT Hair styles (http://vocab.getty.edu/aat/300262903)

This is instead of having the full-fledged hierarchy of AAT, as it appears in http://vocab.getty.edu/aat/300178802.

- create hierarchical relationships between project-specific terms. For example, the term <*appearance*> (http://www.crosscult.eu/KB#appearance) used in pilots 1, 2 and 3 does not appear in any of the controlled vocabularies. We added the term in the CCCS hierarchy under <*culture related concepts*> (http://vocab.getty.edu/aat/300073689) as it refers to a person's appearance that is perceived in various cultures.
- to complete CCCS with concepts that were not originally identified in the four pilots. For example, the term <*jewelry and accessory components*> (http://vocab.getty.edu/aat/300387426) accommodates a variety of object descriptions and it was considered as a useful addition for all pilots, albeit it was not originally identified in any of the pilots.

S7. A separate facet was created to accommodate time periods and provide an access point using time in history.

S8. A separate facet was created to incorporate specific types such as dimensions, functionality and media.

The CCCS was developed using an open-source software, the TemaTres controlled vocabulary server, and is publically available at http://ccdev.uop.gr:5180/vocab/index.php. TemaTres is a web application for the management of formal representations of knowledge, thesauri, taxonomies and multilingual vocabularies (https://sourceforge.net/projects/tematres/). It was installed in the UoP server and has been maintained and used by the pilots and the WP2 team who is responsible for extracting, verifying, normalizing, maintaining and updating the vocabulary.

TemaTres enabled us to create a unique project-specific URI (http://crosscult.eu/vocab/) for each term. It accommodates notes (cataloguers', scope notes and historical notes), preferred and non-preferred terms, the use of equivalent terms (multilingual) and allows both hierarchical and relational associations between terms. The current version of the CCCS incorporates 1451 terms in 36 facets. Hierarchical relations go as far as deep level 8 but most of the terms are in deep levels 2, 3, 4 and 5 thus, accommodating hierarchies without "hiding" the very specific terms within their broader ones. TemaTres was continuously updated as it grew in parallel to the pilot's data models. Special effort was made to "normalise" vocabulary rather than adding to the CCCS excess terms, thus maintaining a manageable pool of terms.

6.3. Integration of CCCS terms into the Upper-level Ontology

The CCCS was integrated into the Upper-Level ontology delivering a unified Knowledge Base resource, as depicted in Figure 12. Terms referring to types were classified under *E55 Type* and were associated to the individuals they describe via the *P2 has type* property. Terms that represent subjects used to enrich the semantic description of cultural heritage objects or places, were classified under *E89 Propositional Object*. They were then associated to the relevant ontology individuals via the *P67 refers to* property. Finally, vocabulary terms referring to Reflective Topics were classified under the project-specific ontology class *Reflective Topic*, and were associated to the entities that drive the corresponding reflection via the *reflects* property.

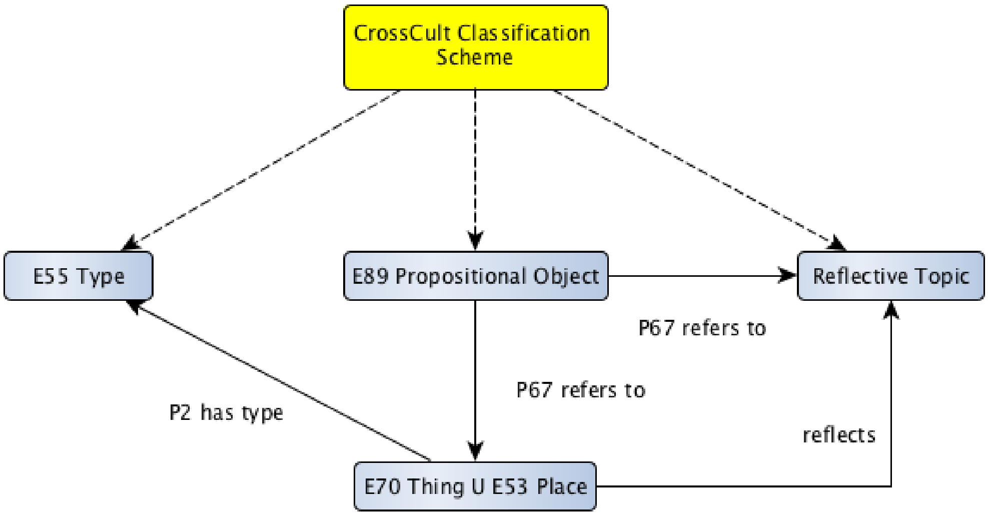

Figure 12: Relationships between the CCCS and the Upper-level Ontology.

7. Conclusion

This deliverable describes the final status of the CrossCult Upper-level ontology and the other ontological schemas (the Venue and User ontologies) and vocabularies (the CrossCult Classification Scheme) that we used to model the semantics of the "world" of CrossCult.

The accompanying PDF file contains the definitions of the classes and properties of the CrossCult ontologies in the syntax of Description Logics. The OWL file containing the code of the CrossCult ontologies in the XML syntax is available The OWL file containing the code of the CrossCult ontologies in the XML syntax is available on the CrossCult website: http://www.crosscult.eu/en/resources/datasets/

The CrossCult ontologies were defined based on the requirements of the four pilots described in D2.1, and were further refined based on the needs of the end-user applications, described in D5.1. Their use for semantically modelling the cultural heritage resources of the four pilots is described in detail in D2.2 and D2.4, and their use for supporting the different technological modules of the CrossCult platform is described in D3.4.

The CrossCult knowledge base has been deployed on the project's triple store. Details on the installation and deployment of the CCKB are provided in deliverable D4.1, while different ways of accessing the CCKB are described in D2.4, D4.2 and D4.3.

References

[1]. Doerr, M.: The CIDOC Conceptual Reference Module: An ontological approach to semantic interoperability of metadata. AI Magazine 24(3), 75–92 (2003).

[2]. Vlachidis, A., Bikakis, A., Kyriaki-Manessi, D., Triantafyllou, I., Antoniou, A.: The CrossCult Knowledge Base: A co-inhabitant of cultural heritage ontology and vocabulary classification. In: Proceedings of 2nd International Workshop on Semantic Web for Cultural Heritage (SW4CH), in conjunction with 21st European Conference on Advances in Databases and Information Systems (ADBIS). Nicosia, Cyprus (2017)

[3]. Doerr, M., & Hiebel, G.H (2013). Where did the Varus battle take place? - A spatial refinement for the CIDOC CRM ontology. Seventh World Archaeological Congress, The Dead Sea, Jordan, January 13th - 18th 2013.

[4]. P.-E. Osche, S. Castagnos, A. Napoli, and Y. Naudet, "Walk the line: Toward an efficient user model for recommendations in museums," in 11th International Workshop on Semantic and Social Media Adaptation and Personalization, SMAP 2016, Thessaloniki, Greece, October 20-21, 2016, 2016, pp. 83–88.

[5]. E.V.a.M. Levasseur, Ethnographie de l'exposition: l'espace, le corps et le sens, in, Centre Georges Pompidou, Paris: Bibliothèque Publique d'Information, 1989.

Appendix

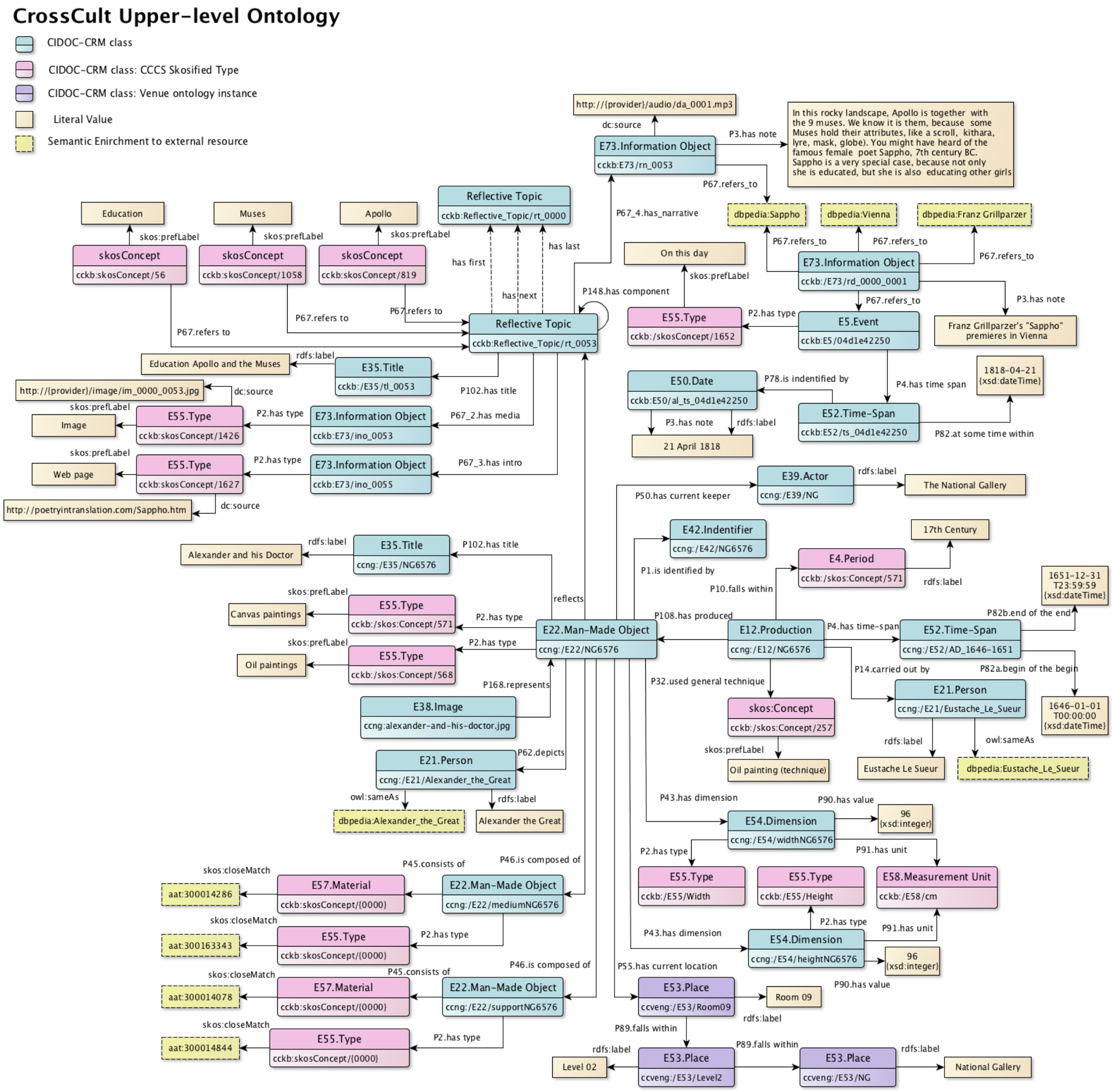

Figure 13: The core elements of the CrossCult Upper-level Ontology