fRamework for safE, opEn, collaboratiVe And incLUsive digitisAtion and managemenT of cultural heritagE

Grant Agreement No 101132389

DELIVERABLE 1.4:

Cultural Artefacts' Contextual Ontology (CACAO)

Work Package: 1

LEAD BENEFICIARY:

KU Leuven (KUL)

Delivery Date: 18.12.2024

Document Sheet

Project acronym	REEVALUATE
Project full title	Framework for safe, open, collaborative and inclusive digitization and management of cultural heritage
Programme	Horizon Europe
Topic	HORIZON-CL2-2023-HERITAGE-01-03
Type of Action	HORIZON-Research and Innovation Actions
Grant Agreement	101132389
Start day	1 January 2024
Duration	36 months

Document Information

Deliverable number	D1.4
Deliverable name	Cultural Artefacts’ Contextual Ontology
Lead beneficiary	KU Leuven
WP	1
Related task(s)	T1.4
Type	R
Reviewers (Organisation)	LINKS, FOKUS
Delivery date	18.12.2024
Main author(s)	Ruben Peeters (KUL), Anastasia Dimou (KUL)
Contributor(s)	Xuemin Duan (KUL), Giuseppe Rizzo (LINKS), Giacomo Blanco (LINKS), Georgios Botsoglou (CERTH), Theodoros Semertzidis (CERTH), Christian Fuhrhop (FOKUS)

Dissemination level

PU	Public	X
SEN	Sensitive, limited under the conditions of the Grant Agreement	
Classified R-UE/EU-R	EU RESTRICTED under the Commission Decision No2015/444	
Classified C-UE/EU-C	EU CONFIDENTIAL under the Commission Decision No2015/444	
Classified S-UE/EU-S	EU SECRET under the Commission Decision No2015/444	

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

Document history

Version	Date	Changes	Reviewer/Contributor
0.1	29.11.2024	First version	**Ruben Peeters (KUL), Anastasia Dimou (KUL), Xuemin Duan (KUL)**
0.2	06.12.2024	CERTH review	**Georgios Botsoglou (CERTH), Theodoros Semertzidis (CERTH)**
0.3	06.12.2024	LINKS review	**Giacomo Blanco (LINKS), Giuseppe Rizzo (LINKS)**
0.4	14.12.2024	Changes intra-task review	**Ruben Peeters (KUL), Anastasia Dimou (KUL)**
0.5	17.12.2024	FOKUS review	**Christian Fuhrhop (Fraunhofer FOKUS)**
1.0	**18.12.2024**	**Final draft**	**Ruben Peeters (KUL)**

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

Table of Contents

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

List of Figures

List of Tables

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

Abbreviations

Abbreviations	Full name
CACAO	Cultural Artefacts' Contextual Ontology
ccREL	Creative Commons Rights Expression Language
CH	Cultural Heritage
CH	Cultural Heritage
CIDOC	International Council for Documentation
DAML	DARPA Agent Markup Language
DOLCE	Descriptive Ontology for Linguistic and Cognitive Engineering
ECCCH	European Collaborative Cloud for Cultural Heritage
EKAW	European Knowledge Acquisition Workship
ESWC	Extended Semantic Web Conference
FOAF	Friend Of A Friend Ontology
FRBR	Functional Requirements for Bibliographic Records
FRBRoo	FRBR-object oriented
GLAM	Galleries Libraries Archives and Museums
ICOM	International Council of Museums
ISO	International Organization for Standardization
ISWC	International Semantic Web Conference
JWS	Journal of Web Semantics
K-CAP	International Conference on Knowledge Capture
LIDO	Lightweight Information Describing Objects
LOT	Linked Open Terms
LOV	Linked Open Vocabularies
MoMU	AG Culturele Instellingen Antwerpen/Erfgoed *(Modemuseum)*
OBO	Open Biological and Biomedical Ontologies

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

ODK	Ontology Development Toolkit
OLS	Ontology Lookup Service
OWL	Ontology Web Language
PROV-O	Provenance Ontology
RDF	Resource Description Framework
RO	Relations Ontology
SIG	Special Interest Group
SPK	Stiftung Preussischer Kulturbesitz
SUMO	Suggested Upper Merged Ontology
SWJ	Semantic Web Journal
XML	Extensible Markup Language

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

Publishable summary

This document provides an overview of the methodology followed and the on-going work on the Cultural Artefacts’ Contextual Ontology (CACAO). The work is grounded in its relevance to both the project's objectives and the broader field of cultural heritage.

Firstly, the deliverable describes related work both in the context of ontologies as well as in the context of ontology development methodologies. Relevant upper-level and domain-specific ontologies are described, and a high-level overview of representative ontology development methodologies is given.

Secondly, it is discussed how the chosen methodology is applied for the development of the CACAO ontology. The first steps of this methodology, requirements specification and ontology implementation, are described in detail, touching on the different inputs that served as inspiration for the requirements, as well as an extensive description of how those inputs were translated to functional and non-functional requirements. As this document describes on-going work, the last steps of the methodology, ontology publication and ontology maintenance, are not elaborated upon.

Lastly, the ontology itself is explained, elaborating on the five most relevant classes: the digital artefact, the physical artefact, context, intellectual property rights, and the user. Each of the classes is described with example properties and given a place in the larger ontology.

By reading this deliverable, the readers will gain insights into the requirements process, the design decisions made, and the ontology’s structure and use in representing cultural heritage data.

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

1 Introduction

The cultural heritage (CH) domain faces several challenges with regards to digitization, one of the most prominent being data heterogeneity, which leads to interoperability issues [1]. In an ideal scenario, a perfect, common, standardized data model enriched with semantics would be used by all CH institutions, ensuring that data is interoperable and understandable by all users, both human and artificial. With this ideal in mind, it is important to gather inputs from as many domain experts as possible to create consensus, satisfaction, and a growing community willing to use the ontology. To make such collaboration feasible, it is crucial to build in the open and make contributing to the ontology as easy as possible for domain experts.

In recent years, the CH sector has emphasized digitizing its collections. Digitization allows Galleries, Libraries, Archives, and Museums (GLAMs) to improve the accessibility of their collections and support digital preservation, ensuring artefacts remain available even after physical deterioration. The CH domain increasingly leverages ontologies to achieve interoperability, enable rich semantic annotation, and facilitate metadata creation. However, to the best of our knowledge, current ontologies in the CH sector, e.g. CIDOC-CRM [2], EDM [3], do not fully capture the rich context of CH artefacts and the intellectual property rights (IPR) associated with their digital counterparts.

The Cultural Artefacts' Contextual Ontology (CACAO) is the ontology introduced by the REEVALUATE project, providing the backbone for the technical enablers that will be developed during the project. CACAO aims to address the gaps in historical contextualization and IPR by extending the widely recognized CH domain ontology CIDOC-CRM with other ontologies such as FOAF [4], ODRL [5], and schema.org [6]. This document aims to provide an overview of on-going work on CACAO. It outlines related research, the development methodology, a summary of the current ontology, and conclusions with directions for future work.

2 Related work

In this section, existing works are analyzed regarding the development of an ontology for context and IPR in the CH sector. It is imperative to reuse relevant, well-established ontologies, should there be such ontologies available. Additionally, when extending existing ontologies or defining a new one, we aim to follow an established ontology development methodology to ensure the outcome of the process is well-defined, structured and consistent. For these reasons, in section 2.1 existing ontologies in the CH field are analyzed, in section 2.2 we analyze ontology development methods and in section 2.9 tooling is briefly discussed.

2.1 Ontologies

An ontology is an explicit specification of a conceptualization [7]. It provides a shared vocabulary and way of representing information, ensuring that both humans and machines can understand and process it. An ontology serves multiple purposes. In the context of the REEVALUATE project, the following three are most important. Firstly, it enables interoperability of data between different systems, minimizing loss of meaning or accuracy when sharing data between parties. Secondly, it enables rich semantic annotation by linking metadata with precise descriptions, as well as relations between the different concepts. Lastly, ontologies allow advanced algorithms such as machine learning and natural language processing, to reason about the data.

Four types of ontologies exist, namely top-level, domain, task, and application ontologies [8].

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

- *Top-level ontologies*, or *upper ontologies*, describe very general concepts like space, time, matter, object, event, action, etc. which are agnostic of a domain [8], such as BFO [9] or DOLCE [10].
- *Domain ontologies* focus on a specific domain or area of knowledge [8], such as the CH domain, e.g., CIDOC-CRM, the clinical medicine domain, e.g., SNOMED CT [11] or the machine learning domain, e.g. MLSea [12].
- *Task ontologies* are similar to domain ontologies in the sense that they are focused on describing the vocabulary of a task [8], e.g. the Medical Action Ontology [13] or the Generic Task Ontology [14].
- *Application ontologies* are specialized ontologies that contain definitions specific to a particular application. They combine elements from both domain and task ontologies [8]. An example is given in [15].

Generally, top-level ontologies are intended to be reused and extended for a particular domain to create a *domain ontology* [16]. For that reason, an overview of top-level or upper ontologies is provided in section 2.3. In addition to the general, upper-level ontologies, several domains of interest were identified during the requirement gathering process described in section 3.2. The REEVALUATE framework proposes several enablers for the digitization lifecycle of an artefact such as the contextualisation and collaboration enabler. In the former, artefacts from the CH domain will be enriched with context by a human-driven and AI-driven process. In the latter, the enriched artefacts will be made available for reuse for users of the proposed marketplace. Therefore, ontologies related to the CH domain, context, intellectual property rights and user profiling, as well as general purpose ontologies, are required by the REEVALUATE platform.

For each of these domains, an overview of available ontologies is provided in section 2.4, 2.5, 2.6, and 2.1.6 respectively. To identify the relevant ontologies, ontology lookup services were used, which are described in section 2.2, and scientific publications were gathered using platforms such as Google Scholar[1] and dblp[2]. Additionally, specific conferences, such as International Semantic Web Conference (ISWC), Extended Semantic Web Conference (ESWC), International Conference on Knowledge Engineering and Knowledge Management (EKAW), International Conference on Knowledge Capture (K-CAP), and journals, e.g., Semantic Web Journal (SWJ), Journal of Web Semantics (JWS), were consulted.

2.2 Ontology lookup services

To find the most suitable ontologies for reuse in the context of our use-cases, it is important to have a broad view of those that are available. To provide such an overview several ontology lookup services attempt to catalogue the vast number of ontologies, either generally available or specific to a domain. During the review process, several of these ontology lookup services were used. In the following subsection, a short description of the services and their focus is given. The domain with the strongest use of ontologies is currently the biomedical domain, hence there is a strong bias to this domain in lookup services.

Linked Open Vocabularies (LOV)

LOV, available at https://lov.linkeddata.es/dataset/lov, is a comprehensive ontology lookup service that catalogues and organizes a wide range of vocabularies and ontologies. It currently hosts vocabularies covering diverse domains such as metadata, industry, IoT, geography, environment, society, biology, government, and more.

[1] https://scholar.google.com/
[2] https://dblp.org/

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

OBO Foundry

The OBO (Open Biological and Biomedical Ontologies) Foundry, available at https://obofoundry.org/, is a collaborative initiative that aims to develop a family of interoperable ontologies intended primarily for the biological and biomedical domains. A key feature is the OBO ontology principles[3], which are intended as an evaluation metric for the ontologies submitted to the foundry. General-purpose ontologies such as PROV-O, the provenance ontology, can also be found in the OBO Foundry.

BioPortal

BioPortal is a comprehensive repository of biomedical ontologies. BioPortal is available at https://bioportal.bioontology.org/. It provides a wider range of ontologies than the OBO Foundry. Similar to the OBO Foundry, it also provides access to more general-purpose ontologies.

Ontology Lookup Service (OLS)

Like the OBO Foundry and BioPortal, OLS, available at https://www.ebi.ac.uk/ols4, is a repository primarily focused on biomedical ontologies, but additionally provides access to general-purpose ontologies such as the Relations Ontology (RO)

2.3 Upper ontologies

In addition to achieving interoperability between the partners in REEVALUATE and the CH domain, it is just as important to achieve interoperability with other domains. This is especially apparent in the context of the emerging Common European Data Spaces[4] , to create inter data-space operability. As such, a mapping to an upper ontology will provide a solid foundation for the development of CACAO. This section provides a description of the three most prominent upper ontologies in the CH domain.

Basic Formal Ontology (BFO)

The Basic Formal Ontology (BFO) [9] is an upper ontology recommended by the OBO Foundry and is therefore widely adopted in the biomedical domain. However, its use is limited in the CH domain. It is used as the foundation for the NFDIcore ontology [17] which is in turn extended by the NFDI4Culture [18] ontology.

BFO distinguishes between universals and particulars, as well as occurrents and continuants—the latter indicating how particulars relate to time. These distinctions provide a clear separation of static and dynamic aspects of reality. Compared to other upper-level ontologies, BFO is relatively small in scope, as it focuses on high-level, general concepts. This minimalist approach maximizes broad applicability and avoids unnecessary complexity. It is also recognized as an International Organization for Standardization (ISO) standard[5] and is available in OWL format [9].

Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE)

The Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) [10] is an upper-level ontology developed in 2002/2003 and has remained stable since its initial release. As its name suggests, DOLCE finds its inspiration in cognitive and linguistics considerations [19]. Therefore, additional alignment efforts might be necessary when integrating with ontologies that have different foundational perspectives. DOLCE has been adopted in fields such as industrial engineering [19], and has served as a means to improve the CIDOC-CRM ontology [20] [21]. To the best of our knowledge, DOLCE is not used as a foundation for any CH domain

[3] https://obofoundry.org/principles/fp-000-summary.html
[4] https://digital-strategy.ec.europa.eu/en/policies/data-spaces
[5] https://www.iso.org/standard/74572.html, last accessed 21/11/2024

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

ontologies. Like BFO, DOLCE is recognized as an ISO standard[6]. While it is not fully available in OWL format, some implementations exist. However, these implementations are re-engineered versions that do not fully represent the formalized ontology[7] [10]. DOLCE is an ontology of particulars where the distinction is made between endurants, perdurants, and abstracts [22].

Suggested Upper Merged Ontology (SUMO)

The Suggested Upper Merged Ontology (SUMO) [23] is a comprehensive upper-level ontology that aims to provide a foundation for semantic integration across various domains. Developed through the merging of different ontology modules and theories, SUMO is organized in a hierarchical structure with "Entity" as the topmost concept. The ontology covers a wide range of areas, from abstract concepts to specific domains, and is formalized using the Standard Upper Ontology Knowledge Interchange Format (SUO-KIF) [24]. SUMO's structure allows for the integration of domain-specific ontologies, making it a versatile tool for knowledge representation and reasoning. However, its lack of a strict ontological commitment and some inconsistencies in its axiomatization pose challenges for its application in certain contexts in and outside of the CH domain. Despite these limitations, SUMO's comprehensive nature and its mapping to lexical resources like WordNet [25] make it a valuable resource for various applications in artificial intelligence, natural language processing, and information retrieval [26], [27], [28], [29].

2.4 CH data models and domain ontologies

CIDOC Conceptual Reference Model (CIDOC-CRM)

The CIDOC Conceptual Reference Model (CIDOC-CRM) [20] was developed by the CIDOC CRM Special Interest Group (SIG), a committee of the International Council for Documentation (CIDOC), under the International Council of Museums (ICOM). It is a widely used ontology within the CH domain. While not officially classified as an upper ontology, CIDOC-CRM provides a high-level framework for organizing cultural heritage information. It is designed to be extended and specialized for specific use cases in the CH domain, such as documenting archaeological buildings [30]. At its highest level, CIDOC-CRM distinguishes between temporal entities, timespans, places, dimensions, persistent items, and spacetime volumes. The model structures information around events rather than static objects, making it a dynamic and process-oriented approach.

Furthermore, CIDOC-CRM is recognized as an ISO standard[8]. It is officially published through a specification document, which notes that the OWL file is a derivative. The version available in the OWL file [31] currently aligns with the latest iteration of the model (version 7.1.3).

Europeana Data Model (EDM)

The Europeana Data Model (EDM) [3] is the data model developed by Europeana[9] and serves as the backbone for their repository of over 50 million CH artefacts. Europeana will form the heart of the European Collaborative Cloud for Cultural Heritage (ECCCH). As such, its data model is a highly valuable resource for guaranteeing interoperability in the CH domain. It is published in the form of a specification document and has an OWL file available. EDM defines 11 top-level classes, six of which, Agent, Event, Information Resource, Physical Thing, Place, and Time Span, are equivalent to classes defined in CIDOC-CRM and are mapped accordingly. The other five classes are Europeana Aggregation, Europeana Object, Non-Information Resource, Provided Cultural Heritage Object, and Web Resource.

[6] https://www.iso.org/standard/78927.html, last accessed 21/11/2024
[7] https://www.loa.istc.cnr.it/dolce/overview.html, last accessed 21/11/2024
[8] https://www.iso.org/standard/85100.html, last accessed 21/11/2024
[9] https://www.europeana.eu/en, last accessed 21/11/2024

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

EDM version 5.2.8 is the latest release of the data model and its definition was published in 2017. However, at the time of writing, the ontology lags two minor versions behind the specification document (version 5.2.6). For these reasons, it is hard to justify the reuse of the data model as basis for CACAO, although it is widely used in the domain, as evidenced by the multitude of CH objects in the platform. In addition, it is not possible to reuse the latest version of the data model in the version of an ontology as, to the best of our knowledge, it does not exist.

Lightweight Information Describing Objects (LIDO)

Lightweight Information Describing Objects (LIDO) [32] is an XML schema designed for harvesting and delivering metadata about museum objects. Similar to CIDOC-CRM, LIDO is maintained by the International Committee for Documentation (CIDOC)[10], a committee of the International Council of Museums (ICOM). In fact, LIDO is a specific application of CIDOC-CRM. Consequently, LIDO relies heavily on CIDOC-CRM classes and adopts its event-centric approach.

Functional Requirements for Bibliographic Records (FRBR & FRBRoo)

Functional Requirements for Bibliographic Records (FRBR) is an entity-relationship model for metadata concerning information objects. As its name implies, it is primarily used for bibliographic information. FRBR was developed concurrently with the CIDOC Conceptual Reference Model (CIDOC-CRM). Efforts to align the two models resulted in FRBR Object-Oriented (FRBRoo) [33], an object-oriented extension of FRBR. FRBRoo incorporates the dynamic aspects of CIDOC-CRM into FRBR [34], effectively serving as an extension of CIDOC-CRM. The foundational concepts of the FRBR model are *Work, Expression, Manifestation,* and *Item*.

2.5 CH context domain ontologies

The Merriam-Webster dictionary defines *context* as *"the parts of a discourse that surround a word or passage and can throw light on its meaning"*[11]. Given such a broad definition, many ontologies could be viewed as contributing to the modelling of context, depending on their purpose and scope. Hence, in many existing ontologies, context is often treated as a catch-all field, represented through optional, unstructured free-text annotations (e.g. [35], [36]). This approach, while flexible, limits the potential for systematic and interoperable contextual modelling.

Previous works have attempted to model context explicitly. These efforts are often domain-specific, focusing on areas such as intelligent and mobile environments [37], [38], and are therefore not directly applicable to cultural heritage. By contrast, large, open knowledge bases such as Wikidata [39] and DBpedia [40] offer broad, domain-agnostic data models in RDF format. Their structured representation of diverse information, combined with their wide applicability, makes them valuable resources for contextual information. For instance, the rich semantic links in these knowledge bases can provide contextual relationships that are relevant for modelling cultural heritage data.

2.6 IPR domain ontologies

In the context of Intellectual Property Rights (IPR), Rights Expression Languages (REL) [41] offer a formal, machine-readable framework that expresses the terms and conditions under which intellectual property can be used, shared, or transferred [5], [42]. RELs are essential tools for managing and automating compliance with intellectual property rights, particularly in digital environments, where clear definitions of permissions and restrictions are crucial for content distribution.

[10] https://icom.museum/en/committee/international-committee-for-documentation/
[11] https://www.merriam-webster.com/dictionary/context, last accessed 21/11/2024

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

Two of the most prominent RELs are the Creative Commons Rights Expression Language (ccREL) [42] and the Open Digital Rights Language (ODRL) [5].

Creative Commons Rights Expression Language (ccREL)

ccREL is structured around two main property types: work properties and license properties. Work properties describe the resource itself and include fields such as title and creator. License properties define the conditions of the license, categorizing them into Permissions, Restrictions, and Obligations [42].

Open Digital Rights Language (ODRL)

The ODRL Core Model is built around the Policy class, which encompasses Permissions, Prohibitions, and Duties. These components are associated with specific Parties, Assets, and Actions, allowing for a nuanced representation of rights and obligations within digital content distribution frameworks [5].

2.7 User profile domain ontologies

A recent systematic review of user profile ontologies identified a range of models for representing user information, including the General User Modeling Ontology (GUMO) [43], User Profile Ontology (UPO) [44], OntobUMf [45], Holistic Persona Ontology (HPO) [46], Persona Ontology (PO) [47], Grapple ontology [48], User Modelling Meta-ontology [49] and the Friend Of A Friend Ontology (FOAF)[12]. These ontologies were evaluated as potential frameworks for user profile modelling, with the User Profile Meta-Ontology proposed as a unifying solution.

Despite the existence of OWL specification documents for these ontologies, the associated OWL files are no longer readily available, limiting their potential for reuse. An exception is the FOAF ontology, which thus serves as a valuable resource for representing user information.

2.8 Methodologies

A wide range of development methodologies is available for designing and implementing ontologies. Five methodologies deemed most representative by the study are analysed in [50]. This systematic literature review concluded that, while these methodologies adopt different approaches, their conceptual steps are fundamentally similar. Specifically, these steps include:

1. Determining the domain and scope of the ontology
2. Implementation (of the ontology)
3. Evaluation (of the ontology)
4. Documentation and maintenance (of the ontology)

The specific methodologies may further subdivide these steps into smaller, more detailed components.

Although the systematic literature review does not explicitly mention it, the Linked Open Terms (LOT) methodology [51] aligns closely with these steps. The LOT methodology comprises:

1. Ontology requirements specification
2. Ontology implementation
3. Ontology publication
4. Ontology maintenance

In the LOT methodology, ontology evaluation is integrated into the implementation step.

[12] http://xmlns.com/foaf/spec/, last accessed 22/11/2024

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

2.9 Tooling

The biomedical domain faces similar challenges as the CH domain regarding data management and data integration. Specifically, data is heterogeneous, and domain experts' input is necessary throughout the entire lifecycle of ontology development. Additionally, it is noted that domain experts are often not trained in software engineering, making it more difficult for them to contribute to these ontologies [52].

The Ontology Development Kit (ODK) [52] is a toolkit for building, maintaining, and standardizing biomedical ontologies. It provides several standardized, customizable workflows and packages, readily available to execute locally or using automated workflows on common repository hosting services. Although ODK was developed for the biomedical domain, it attempts to address the root problems apparent in both the biomedical and CH domains, making it applicable to the REEVALUATE use-case. By utilizing repository hosting services, one can develop in public. Additionally, domain experts can provide their input in plain language using integrated issue trackers, lowering the barrier to contribution. Moreover, ODK not only makes it easier for less technical users to contribute but also reduces the burden on ontology engineers and more technically inclined contributors due to its automation capabilities.

3 Methodology

Here we describe how the selected ontology development methodology is applied to the REEVALUATE project. The requirement gathering process is described in section 3.2. In section 3.3, the ontology implementation choice is described and underpinned by an overview of current and foreseen problems related to development of and contributing to an ontology in the CH domain.

3.1 Linked Open Terms methodology

Due to its industrial applicability, the authors' previous positive experience with LOT, and its satisfactory overlap with the methodologies considered most representative in [50], the LOT methodology was selected for developing CACAO [53]. Currently, (1) the requirements specification and (2) the ontology implementation are in a final stage.

3.2 Ontology requirements specifications

Review & interview findings (D1.1 & D1.2)

The literature review in D1.1, as well as the output of the interviews and surveys of CH professionals in D1.2 provide valuable insights into gaps within the CH domain, such as lack of technical expertise and reliance on existing partnerships. From the perspective of specifying requirements for the ontology, we can consider these findings primarily from a non-functional standpoint. Specifically, we have extracted the following non-functional requirements, where RX stands for Requirement X:

1. **R1 Ease of Contribution:** Many CH institutions perceive themselves as falling behind in adopting the latest technological advancements due to limited funding, lack of technical expertise, or resistance to change withing the organisation; therefore, it should be easy for non-technical users to contribute to the ontology.
2. **R2 Accessibility and Openness:** In the spirit of democratizing access to CH artefacts, it is only natural that the ontology should be as widely accessible as possible and developed openly.
3. **R3 Comprehensive Documentation:** The ontology should be well-documented to facilitate ease of use and adoption by users.

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

4. **R4 Community Feedback and Consensus:** Since CH institutions rely heavily on established partnerships, it is important to enable the broader community to provide feedback and reach consensus on the ontology's implementation.

User needs & usage scenarios (D1.3)

Deliverable D1.3 discusses Goals (G), Requirements (R), Obstacles (O), Needs (N), and Domain Hypotheses (D). In relation to the framework and technological enablers that will be developed in the REEVALUATE project. Goals are high-level objectives that are subdivided into smaller components: requirements and needs. Requirements are low-level tasks that need to be executed by the project' enablers, while needs pertain to user-facing (front-end) aspects. A domain hypothesis is an assumed general truth about the domain. An obstacle is a barrier to achieving a goal.

With these definitions in mind, we should focus on requirements and needs during the user requirements phase. Each of the user requirements were analysed and the relevant ones were translated into functional or non-functional ontology requirements. The needs are analysed in a similar manner.

This analysis results in the following required components for the ontology:

Table 1 Functional requirements generated from D1.3

Class	Properties
User profile	- Demographics - Affiliation - Interests
Digital artefact	- Existing metadata - Source - Related artefacts - Context - IPR - Reuse restrictions - Image(s)/media

Additionally, the following requirements were extracted:

- **R5** Solid descriptions
- **R6** Follow W3C guidelines
- **R7** Maximum reuse of existing ontologies
- **R8** Language tags
- **R9** Efficient querying (to reduce the ecological footprint)

Architectural schema & mock-ups (D1.5)

The architecture and mockups were created with consideration of the user needs and usage scenarios defined in D1.3. Nonetheless, we analyzed the D1.5 deliverable to identify any additional functional or non-

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

functional requirements that might arise from the architecture, independent of the user needs and usage scenarios. The requirements derived from D1.3 were confirmed by the proposed architecture (D1.5).

Gathered data sources (T2.1)

T2.1 is responsible for gathering the artefacts to be used in the REEVALUATE framework. REEVALUATE will run a number of pilots during the project's duration. As a result of these pilots, we have the privilege of working with four CH institutions, namely Stiftung Preussischer Kulturbesitz (SPK), Olympiako Mouseio, AG Culturele Instellingen Antwerpen/Erfgoed (MoMU) and Fondazione Aquileia, and thus have access to real-life CH data. Thus, an additional requirement is defined: **R10 fit all existing data shared by the institutions.**

The data provided generally consists of images and relevant metadata. In some cases, 3D models are also provided. We analysed the preliminary results of T2.1 and additional data gathered from the partners to compile a list of ontology terms and properties. In the first phase, candidate terms from existing ontologies were mapped onto the extracted terms; this analysis is available in Table 2. In the second phase, a single candidate ontology, CIDOC-CRM, was selected to map the existing data and metadata too. Except for two properties, all the terms and properties extracted from the data sources were mappable to CIDOC-CRM. The mappings are available in Table 3.

Table 2 Possible mappings of extracted terms to ontologies. The following prefixes are used: crm: http://www.cidoc-crm.org/cidoc-crm/, schema: https://schema.org/, wiki: https://www.wikidata.org/wiki/, dct: http://purl.org/dc/terms/, SKOS: http://www.w3.org/2004/02/skos/core# .

In Class	Possible mappings
Artefact	▪ crm: E18_Physical_thing ▪ schema: thing ▪ wiki: artificial object (Q16686448)
Identifier	▪ crm: E42_Identifier ▪ wiki: identifier (Q853614)
Type of object / classification	▪ crm: E55_Type ▪ wiki: type (Q21146257)
Title	▪ crm: E35_Title ▪ dct: title ▪ schema: title ▪ wiki: title (Q216353)
Repository	▪ wiki:repository (Q108296843)
Place	▪ crm: E53_Place ▪ schema: place ▪ wiki: geographic location (Q2221906) ▪ dct: location
Measurement	▪ crm: E16_Measurement ▪ wiki: measurement (Q12453)
Value	▪ crm:E60_Number ▪ schema:value ▪ wiki: numeric value (Q10388960)

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

Unit	▪ crm: E58_Measurement_unit ▪ schema: unitText ▪ wiki: unit of measurement (Q47574)
Event	▪ crm:E5_Event ▪ wiki: occurrence (Q1190554) ▪ schema: event ▪ dct: event
Date	▪ crm:E61_Time_Primitive ▪ wiki: point in time (Q186408) ▪ schema: Date ▪ dct: date
Time span	▪ crm:E52_Time-Span ▪ wiki: time interval (Q186081)
Material	▪ crm:E57_Material ▪ wiki: material (Q214609) ▪ schema: material
Actor	▪ crm:E39_Actor ▪ wiki: agent (Q24229398) ▪ schema: Person
Nationality (of the actor)	▪ schema:nationality ▪ wiki: nationality (Q231002)
Role (of the actor)	▪ schema:Role ▪ wiki: role (Q4897819)
Related work	▪ wiki: related match (Q39894604) ▪ SKOS: relatedMatch
Legal body	▪ crm:E74_Group ▪ wiki: legal person (Q3778211)
Record	▪ crm: E90_Symbolic_Object ▪ wiki: digital representation (Q42396623)
Rights	▪ crm:E30_Right ▪ wiki: license (Q79719) ▪ schema: License ▪ dct: License
Resource	▪ crm:E73_Information_Object ▪ wiki: information object (Q23698381)

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

Table 3 Mapping of data sources to CIDOC-CRM

	Original term	Subject	Predicate	Object
Digital Artefact	Digital Artefact	Digital Artefact	a	E73_Information_Object
	Name		P1_is_identified_by	E41_Appellation
	Type		P2_Has_Type	E55_Type
	Note		P3_has_note	rdfs:Literal
	Preferred identifier		P48_has_preferred_identifier	E42_Identifier
	Rights		P104_is_subject_to	E30_Right
	Holding organisation		P50_has_current_keeper	E39_Actor
	Digital representation of ..		P67_refers_to	E1_CRM_Entity
	Related artefacts		P130_shows_features_of	E70_Thing
	Legal		P104_is_subject_to	E30_Right
	Image(s)		P138i_has_representation	E36_Visual_Item
Physical Artefact	Physical artefact	Physical artefact	a	E22_Human-Made_Object
	Type		P2_Has_Type	E55_Type
	Description		schema:description	rdfs:Literal
	Preferred identifier		P48_has_preferred_identifier	E42_Identifier
	Title		P102_has_title	E35_Title
	Related artefact		P130_shows_features_of	E70_Thing
	Current (permanent) location		P54_has_current_permanent_location	E53_Place
	Part (of resource, resource = digital artefact)		- P106_is_composed_of - P46_is_composed_of	- E90_Symbolic_Object - E18_Physical_Thing

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

	Subject (object name)		P1_is_identified_by	E41_Appellation
	Has current location (most specific location)		P55_has_current_location	E53_Place
	Creator		P94i_was_created_by	E65_Creation
		E65_Creation	P14_carried_out_by	E39_Actor
	Date (of creation)	E65_Creation	P4_has_time-span	E52_Time-Span
	Medium		P2_Has_Type	E57_Material
	Measurements of the object		P39i_was_measured_by	E16_Measurement
		E16_Measurement	P40_observed_dimension	E54_Dimension
		E54_Dimension	P90_has_value	rdfs:Literal
		E54_Dimension	P91_has_unit	E58_Measurement_Unit
	Publisher/ Holding institution		- P49_has_former_or_current_keeper - P50_has_current_keeper	- E74_Group
	Note		P3_has_note	rdfs:literal
	Condition		P44_has_condition	E3_Condition_State
Place	Place	Place	a	E53_Place
	Name		P1_is_identified_by	E41_Appellation
	Preferred identifier		P48_has_preferred_identifier	E42_Identifier
	Falls within		P89_falls_within	E53_Place
Rights	Rights	Rights	a	E30_Right
	Name		P1_is_identified_by	E41_Appellation
	Preferred identifier		P48_has_preferred_identifier	E42_Identifier

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

Appellation	Appellation	Appellation	a	E41_Appellation
	Content		P190_has_symbolic_content	rdfs:Literal
Actor	Actor	Actor	a	E21_Person
	Preferred identifier		P48_has_preferred_identifier	E42_Identifier
	Name		P1_is_identified_by	E41_Appellation
	Nationality		schema:nationality	rdfs:literal
	Role		schema:Role	rdfs:Literal
	Birth		P98i_was_born	E67_Birth
	Death		P100i_died_in	E69_Death
Legal body	Legal body	Legal body	a	E74_Group
	Name		P1_is_identified_by	E41_Appellation
	Preferred identifier		P48_has_preferred_identifier	E42_Identifier
Event	Event	Event	a	E5_Event
	Name		P1_is_identified_by	E41_Appellation
	Type		P2_Has_Type	E55_Type
	Location		P7_took_place_at	E53_Place
	Carried out by		P11_had_participant	E39_Actor
Resources (image/video/text/audio/…)	Resource	Resource	a	E73_Information_Object
	Name		P1_is_identified_by	E41_Appellation
	Type		P2_Has_Type	E55_Type
	Preferred identifier		P48_has_preferred_identifier	E42_Identifier
	Date of creation		P94i_was_created_by	E65_Creation
		E65_Creation	P4_has_time-span	E52_Time-Span

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

		E52_Time-Span	P79_beginning_is_qualified_by	rdfs:Literal
		E52_Time-Span	P80_end_is_qualified_by	rdfs:Literal
	Holding organisation		P105_right_held_by	E74_Group
	Rights		P104_is_subject_to	E30_Righ
Repositories	Repository	Repository	a	E78_Curated_Holding
	Name		P1_is_identified_by	E41_Appellation
	Location		P54_has_current_permanent_location	E53_Place
	Preferred identifier		P48_has_preferred_identifier	E42_Identifier
	Has rights		P75_possesses	E30_Right

Data management plan (D6.3)

The data management plan (D6.3) provides insight into different types of software licenses. It distinguishes between copyleft, permissive, and public domain licenses. Various licenses such as the GNU General Public License [54], the MIT License [55], and Creative Commons licenses [56] are available to specify the rights associated with these license types. Additionally, a custom license can be defined if the licenser's needs are not covered by any pre-existing licenses.

This analysis leads to the following requirements:

- **R11** Ability to specify one of the open-source licenses.
- **R12** Ability to specify a Creative Commons (CC) license.
- **R13** Ability to indicate if something belongs to the public domain.
- **R14** Ability to specify a custom license.

Definition of context

It is important to understand what the broader CH community means when referring to context. To this end, discussions were organized with the pilot partners in the project, where each partner provided their input separately. These inputs were then consolidated into common requirements.

Based on the inputs from the pilot partners, we have identified several key requirements for the ontology:

- **R15 Relevant Domain:** The ontology should specify the domain, or category, to which each artefact belongs.
- **R16 Relation to Historical Events:** It should capture the artefact's connection to significant historical events, highlighting its historical importance.
- **R17 Social Significance:** The ontology should represent the artefact's relevance to social movements and its impact on society.

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

- **R18 Religious Significance:** It should include information about the artefact's association with religious movements and its spiritual importance.
- **R19 Cultural Significance:** The ontology should address the artefact's role in cultural movements, reflecting its influence on cultural development.
- **R20 Political Significance:** It should reflect the artefact's connection to political movements and its significance in political history.
- **R21 Economic Significance:** The ontology should encompass the artefact's impact on economic history, including its role in economic developments.
- **R22 Meaning of Material:** There should be provisions for open-text descriptions that explain the significance of the materials used in the artefact.
- **R23 Meaning of Shapes:** The ontology should capture the symbolic meanings associated with the shapes present in the artefact.
- **R24 Link to Other Artefacts:** It should enable linking the artefact to other related artefacts, facilitating a network of connected cultural heritage items.

These requirements aim to ensure that the ontology comprehensively represents the contextual dimensions of cultural heritage artefacts, thereby enhancing understanding and accessibility for users.

Competency questions

In this section, for each of the functional requirements, competency questions are defined. An overview can be found in Table 4.

Table 4 Competency Questions

Functional Requirement	Competency Questions
R10 fit all existing data shared by the institutions	- What is the location of the artefact? - What are the dimensions of the artefact? - Who holds the right of this artefact? - What resources are available for a distinct artefact? - Which digital artefact is a representation of a distinct physical artefact
R11 Ability to specify one of the open-source licenses	- Is there an open-source license specified for a distinct artefact? - What artefacts are available under a distinct open-source artefact?
R12 Ability to specify a Creative Commons (CC) license	- Is there a Creative Commons license specified for a distinct artefact? - What artefacts are available under a distinct Creative Commons artefact?
R13 Ability to indicate if something belongs to the public domain	- Does a distinct artefact belong to the public domain? - Which artefacts belong to the public domain?
R14 Ability to specify a custom license.	- Under which custom license is a distinct artefact available? - What are the duties related to a license? - What are the requirements related to a license? - What are the prohibitions related to a license?
R15 Relevant Domain	- What is the domain to which an artefact belongs? - What domains are defined within the ontology's structure? - What artefacts are classified in this domain?
R16 Relation to Historical Events	- What historical events are relevant to a given artefact? - What artefacts are relevant to a given historical event?
R17 Social Significance	- To what social movement is this artefact relevant?

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

	- Which social movements are relevant to this knowledge base?
R18 Religious Significance	- To what religious movement is this artefact relevant? - Which religious movements are relevant to this knowledge base?
R19 Cultural Significance	- To what cultural movement is this artefact relevant? - Which cultural movements are relevant to this knowledge base?
R20 Political Significance	- To what political movement is this artefact relevant? - Which political movements are relevant to this knowledge base?
R21 Economic Significance	- How is this artefact economically relevant?
R22 Meaning of Material	- What is the significance or symbolism of materials documented?
R23 Meaning of Shapes	- What is the meaning of shapes present on a distinct artefact?
R24 Link to Other Artefacts	- To what other artefacts is a distinct artefact related?

3.3 Ontology implementation

To implement the ontology, the ODK toolkit was used in combination with GitHub[13] and Protégé[14]. Using ODK, terms from external ontologies are imported into the ontology-edit file, which can then be edited using Protégé. After the desired edits are saved, the update is pushed to GitHub. When a new release of the ontology is ready, an automatic workflow is run to create the desired output files, available at https://github.com/REEVALUATE/CACAO/tree/main, after which a GitHub release is created.

To request new terms or propose changes to the ontology, GitHub's integrated issue tracker can be used, by creating an issue or a pull request with implemented changes. This will maximize the outreach of the ontology while facilitating the acquisition of users' feedback that will help in evolving the ontology.

4 Results

In the results section, the current implementation of the CACAO ontology is described. Firstly, the ontology is explained from a high-level perspective, after which subsequent sections explain in more detail the five major parts of the ontology.

4.1 The Cultural Artefacts' Contextual Ontology (CACAO)

Because of its widespread use in the CH domain, as well as its availability as an OWL formalization, CIDOC-CRM was chosen as the foundation for the CACAO ontology. Using all the requirements gathered, classes and properties that were deemed relevant were imported from CIDOC-CRM. The definitions of these terms can be found in the official documentation[15]. When the label of a class starts with 'E' or a property label starts with 'P', this indicates that the class or property is imported from CIDOC-CRM.

The ontology is based around five central concepts:

- The physical artefact
- The digital artefact
- A user (of the marketplace, for example)
- Context

[13] https://github.com/
[14] https://protege.stanford.edu/
[15] https://cidoc-crm.org/Version/version-7.1.3

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

- Rights

In this ontology, a digital artefact is the digitized representation of a physical artefact, and context is attached to the digital artefact. Both the digital and physical artefacts are subject to rights, which are held by high-level actors, including users and institutions. A user is interested in certain digital artefacts, events, and movements; events can, in turn, be connected to specific artefacts.

A high-level overview is illustrated in Figure 1. Due to the hierarchy inherent to the CIDOC-CRM ontology, most of the classes and properties have been omitted, only showing the most representative examples needed to gain a general understanding of the ontology structure and modelling approach.

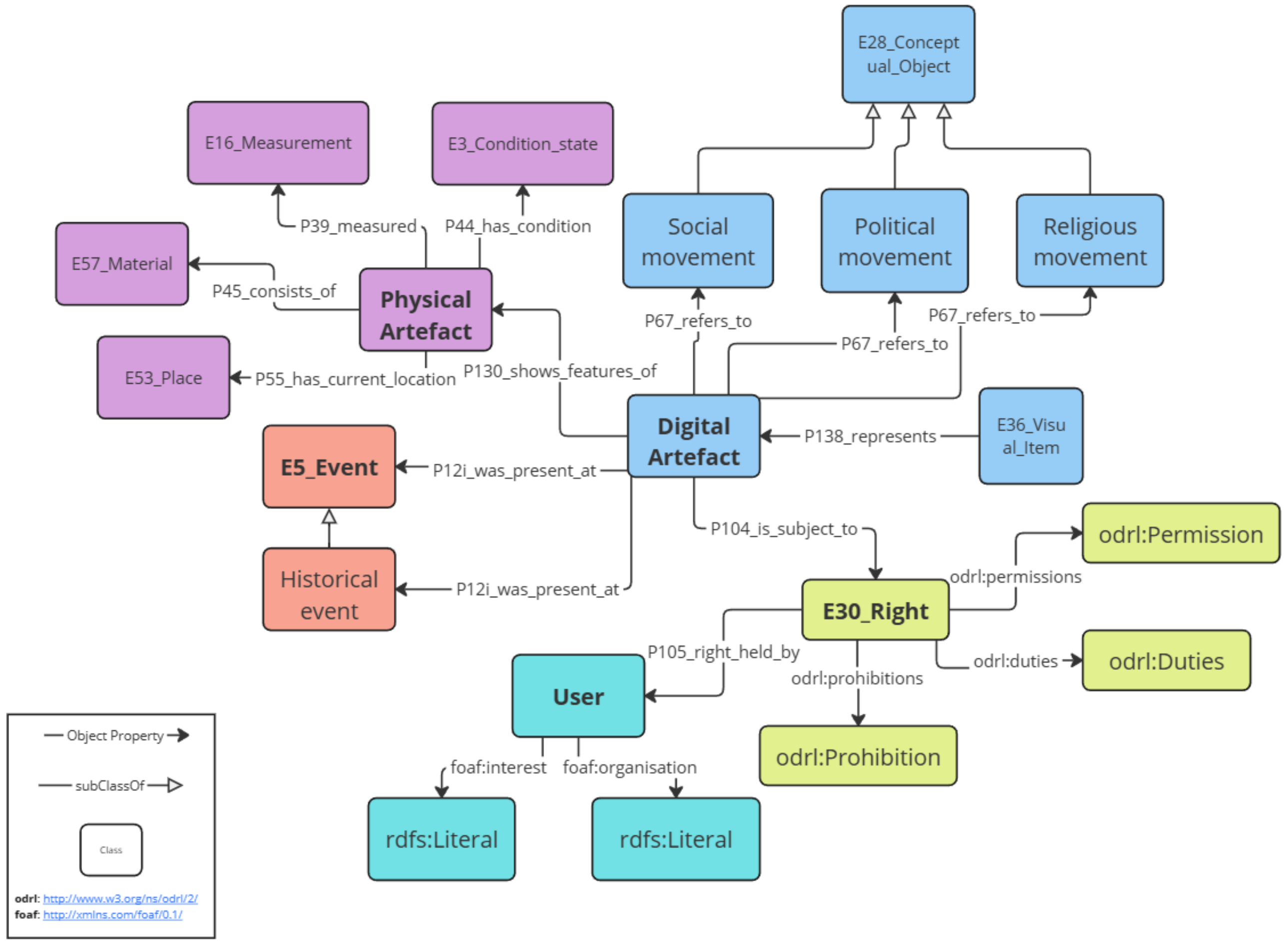

Figure 1 CACAO High-level overview

4.2 Physical artefact

The physical artefact, which is directly mapped to CIDOC CRM's E22_Human-Made_Object, is illustrated in Figure 2.

The concept of the physical artefact encompasses information about its physical attributes, such as location, dimensions, and material. A physical artefact is identified using classes like *E41_Appellation*, *E35_Title*, and *E42_Identifier*. The condition, category, and material of the object are described using classes such as *E3_Condition_State*, *E55_Type*, and *E57_Material*, respectively. Rights statements can be expressed using the *E30_Right* class. In addition, the current physical holder, likely a CH institute, is expressed as a *39_Actor*.

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

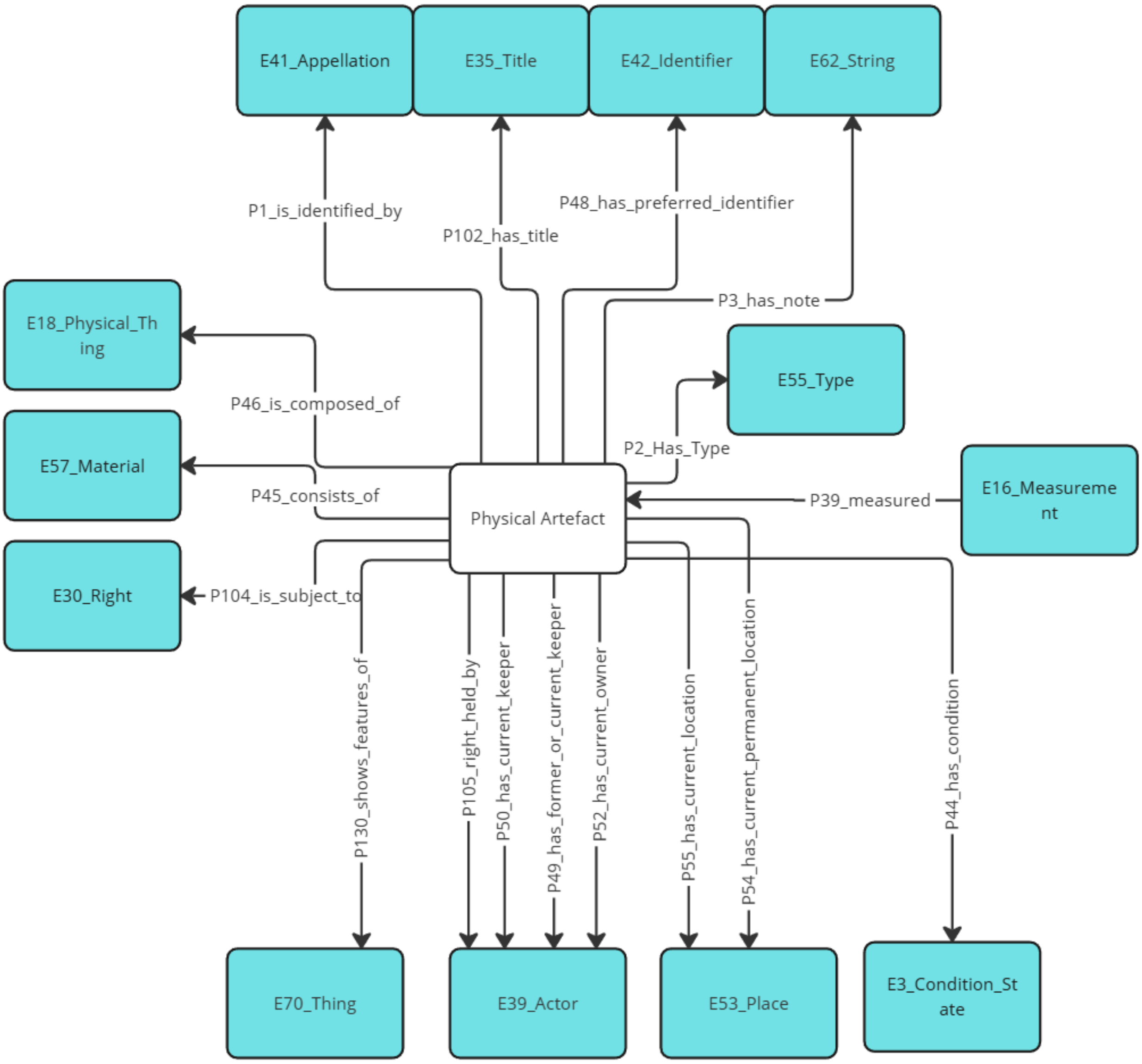

Figure 2 CACAO Physical Artefact

4.3 Digital artefact

The digital artefact, which is directly mapped to *E73_Information_Object*, is illustrated in Figure 3.

The digital artefact represents any digital format of the physical artefact and is related to it through the property *P62i_is_depicted_by*. It can comprise a multitude of *E90_Symbolic_Objects*, each representing a different digital resource such as an image, video, 3D object scan, audio recording or even another Digital Artefact. Additionally, for provenance purposes, the creation of a digital artefact can be described using *E65_Creation*. A digital artefact relates to events using the *P12i_was_present_at* property, and can have reference to other entities using the *P67_refers_to* and *P129_is_about* properties. The use of which will become apparent in section 4.5 on context.

Funded by
the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

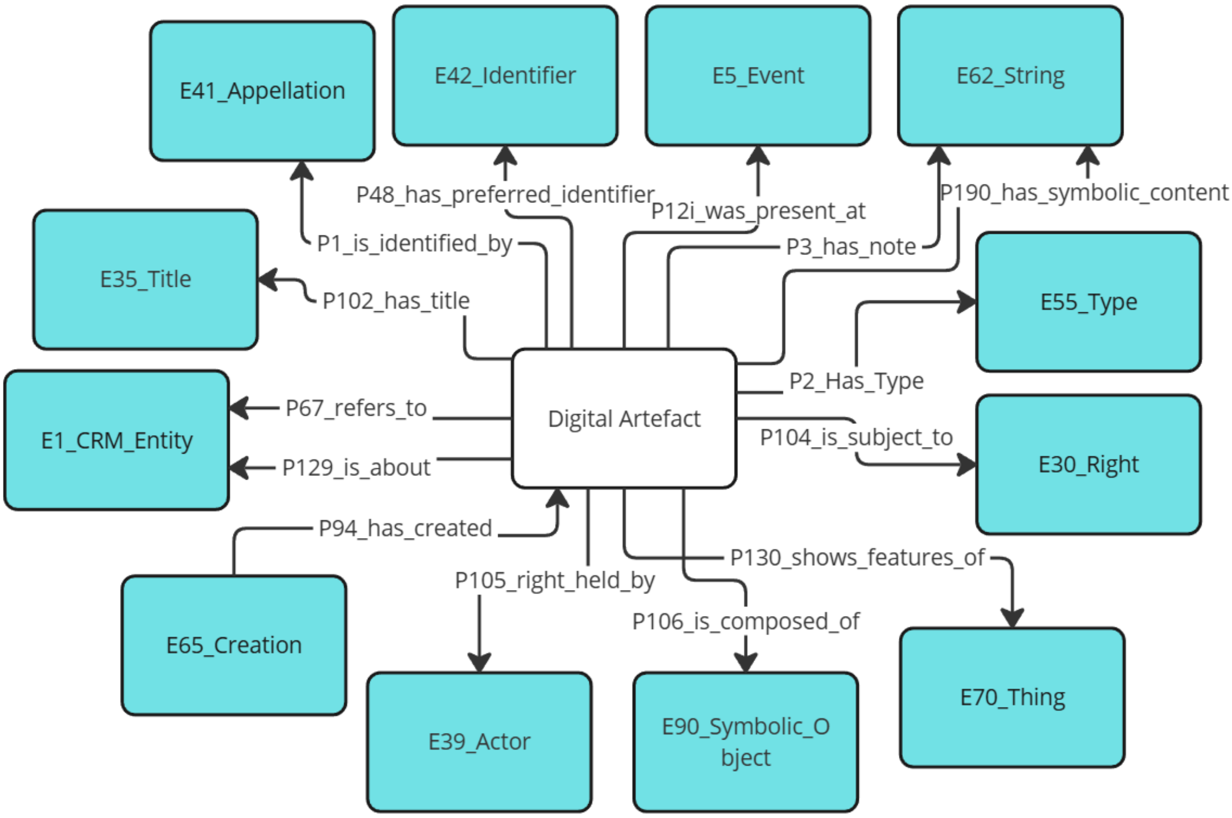

Figure 3 CACAO Digital Artefact

4.4 User

The user is a specialization of *E21_Person* and, consequently, a further specialization of *E39_Actor*. This implies that a user can hold rights regarding any legal object. In this context, legal objects are the digital and physical artefacts, although our focus is on the digital. To align with the requirements of the marketplace as well as the recommendation system, the user's name, age (via birth and death dates), and identifier (e.g., ORCID) can be expressed using instances of reused CIDOC-CRM classes such as *E67_Birth*, *E69_Death*, and *E41_Appellation*. Additionally, the user's nationality, organisation, and interests can be expressed using instances and properties of the FOAF ontology. The organisation of the user can be used to assess their credibility.

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

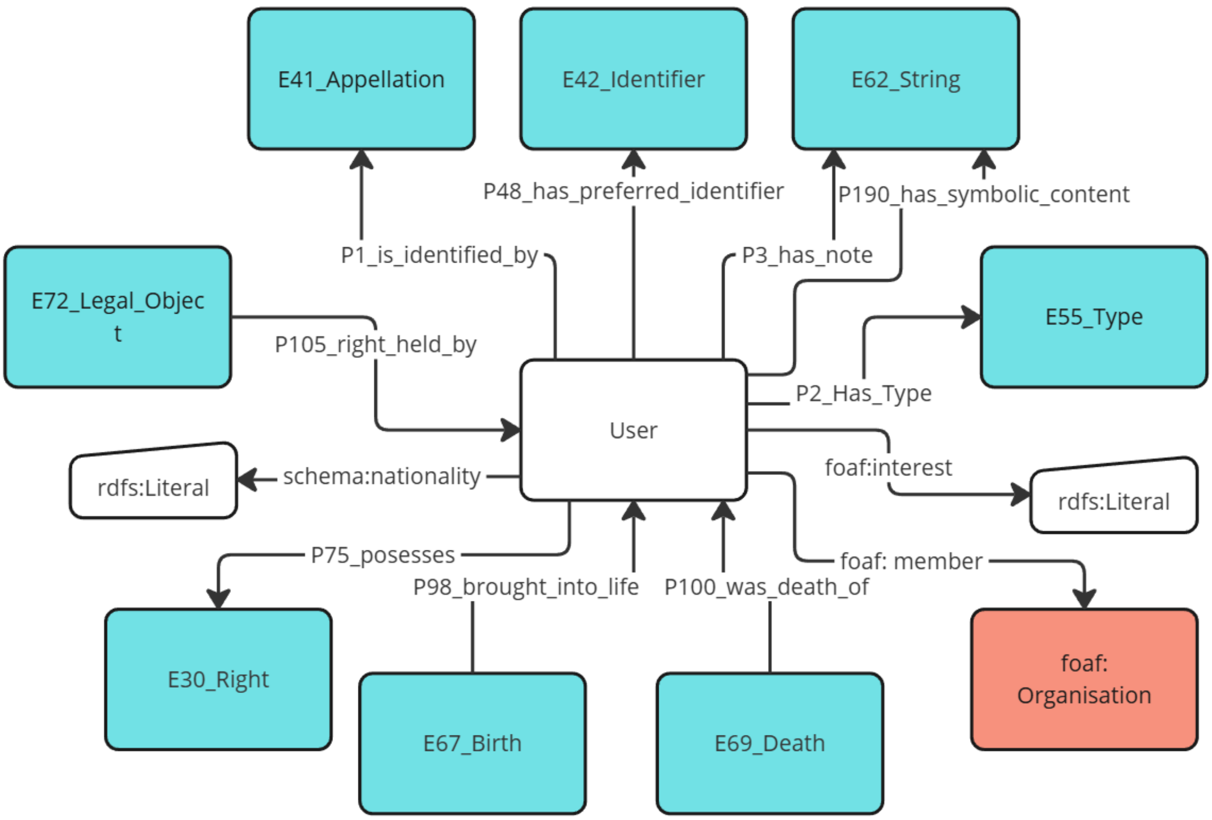

Figure 4 CACAO User

4.5 Context

Based on the gathered requirements, several classes are defined. To model the historical context, *Historical Event* is defined as a subclass of *E5_Event*, while *Social Movement*, *Cultural Movement*, *Political Movement, Religious Movement*, and other movements discussed earlier are defined as subclasses of *E28_Conceptual_Object*. Using these classes, context can be added to digital artefacts by employing, for example, the property *P12_occurred_in_the_presence_of*, where a Historical Event is related to a digital or physical artefact, or *P67_refers_to* when referring to one of the movements.

4.6 Rights

In the context of the European data space initiatives, the Open Digital Rights Language (ODRL) is used to describe usage conditions. Additionally, the Creative Commons (CC) ontology can be used to express CC licenses and usage conditions. The essence of ODRL is the expression of policies through combinations of *Duty, Permission,* and *Prohibition*. CC licenses are expressed as combinations of *Requirement, Permission,* and *Prohibition*. Clearly, ODRL and CC adopt a similar approach in modelling usage policies. Therefore, we have chosen ODRL for IPR management in CACAO due to its backing from the European Union. ODRL uses the *Policy* class to define anything related to rights. *Policy* is related to *Permissions, Duties*, and *Prohibitions* through properties with the names *permission, obligation,* and *prohibition* respectively. These policies are then related to an *Asset* which is mapped onto a *E72_Legal_Object* in CIDOC-CRM. Both *Physical Artefact* and *Digital Artefact* are subclasses of *E72_Legal_Object.*

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

5 Conclusion & Future Work

This document has presented an overview of the CACAO ontology, focusing on physical and digital artefacts, rights, context, and users. We analyzed related work, specified requirements through the LOT methodology, and implemented and published the ontology.

The main benefits of the CACAO ontology include:

1. **A Common Data Model:** Providing interoperability among the REEVALUATE project's partners and CH institutions outside the project that are willing to use the ontology.
2. **Preservation of Original Context:** Enhancing the expression and preservation of the original context of digital and physical artefacts.
3. **Flexible Usage Policies:** Enabling the attachment of flexible usage policies to digital artefacts, allowing the project to express reuse context rules and licenses.

Moving forward, the CACAO ontology will be extended to reflect the project's further needs. Specifically:

1. **Mapping to a Foundational Ontology:** The ontology will be mapped to an upper ontology, such as BFO [9] to enhance its structural coherence.
2. **User Alignment:** The user component will be extended to align with marketplace user profiles.
3. **Context Elaboration:** Context will be further elaborated by providing vocabularies for different classes (e.g., social movements, cultural movements).
4. **Ontology Refinement:** The ontology will be refined based on feedback from the project partners.
5. **Ontology Alignment:** The ontology will be further aligned with other ontologies in the CH domain.

With this work, we have enabled the expression of Intellectual Property Rights (IPR) management in the context of cultural heritage and made significant progress toward developing a contextual ontology for the domain.

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

References

[1] E. Tóth-Czifra, '10. The Risk of Losing the Thick Description: Data Management Challenges Faced by the Arts and Humanities in the Evolving FAIR Data Ecosystem', in *Digital Technology and the Practices of Humanities Research*, 1st ed., J. Edmond, Ed., Cambridge, UK: Open Book Publishers, 2020, pp. 235–266. doi: 10.11647/OBP.0192.10.

[2] G. Bruseker, N. Carboni, and A. Guillem, 'Cultural Heritage Data Management: The Role of Formal Ontology and CIDOC CRM', in *Heritage and Archaeology in the Digital Age: Acquisition, Curation, and Dissemination of Spatial Cultural Heritage Data*, M. L. Vincent, V. M. López-Menchero Bendicho, M. Ioannides, and T. E. Levy, Eds., Cham: Springer International Publishing, 2017, pp. 93–131. doi: 10.1007/978-3-319-65370-9_6.

[3] M. Doerr, S. Gradmann, S. Hennicke, A. Isaac, C. Meghini, and H. Van de Sompel, 'The europeana data model (edm)', in *World Library and Information Congress: 76th IFLA general conference and assembly*, 2010, p. 15. Accessed: Dec. 04, 2024. [Online]. Available: https://iris.cnr.it/retrieve/2cb6de91-2e29-4b3d-865e-82ea8caa6caf/prod_295309-doc_84849.pdf

[4] D. Brickley and L. Miller, 'The Friend of a Friend (FOAF) project'. 2000.

[5] R. Ianella, 'Open digital rights language (ODRL)', *Open Content Licens. Cultiv. Creat. Commons*, 2007, Accessed: Dec. 02, 2024. [Online]. Available: https://www.academia.edu/download/67069103/C9_Ianella.pdf

[6] R. V. Guha, D. Brickley, and S. Macbeth, 'Schema.org: evolution of structured data on the web', *Commun ACM*, vol. 59, no. 2, pp. 44–51, Jan. 2016, doi: 10.1145/2844544.

[7] T. R. Gruber, 'A translation approach to portable ontology specifications', *Knowl. Acquis.*, vol. 5, no. 2, pp. 199–220, Jun. 1993, doi: 10.1006/knac.1993.1008.

[8] N. Guarino, *Formal Ontology in Information Systems: Proceedings of the 1st International Conference June 6-8, 1998, Trento, Italy*, 1st ed. NLD: IOS Press, 1998.

[9] J. N. Otte, J. Beverley, and A. Ruttenberg, 'BFO: Basic Formal Ontology', *Appl. Ontol.*, vol. 17, no. 1, pp. 17–43, Jan. 2022, doi: 10.3233/AO-220262.

[10] S. Borgo *et al.*, 'DOLCE: A descriptive ontology for linguistic and cognitive engineering', *Appl. Ontol.*, vol. 17, no. 1, pp. 45–69, Jan. 2022, doi: 10.3233/AO-210259.

[11] K. Donnelly, 'SNOMED-CT: The advanced terminology and coding system for eHealth', *Stud. Health Technol. Inform.*, vol. 121, p. 279, 2006.

[12] I. Dasoulas, D. Yang, and A. Dimou, 'MLSea: A Semantic Layer for Discoverable Machine Learning', in *The Semantic Web*, A. Meroño Peñuela, A. Dimou, R. Troncy, O. Hartig, M. Acosta, M. Alam, H. Paulheim, and P. Lisena, Eds., Cham: Springer Nature Switzerland, 2024, pp. 178–198. doi: 10.1007/978-3-031-60635-9_11.

[13] L. C. Carmody *et al.*, 'The Medical Action Ontology: A tool for annotating and analyzing treatments and clinical management of human disease', *Med*, vol. 4, no. 12, pp. 913-927.e3, Dec. 2023, doi: 10.1016/j.medj.2023.10.003.

[14] D. Rajpathak, E. Motta, and R. Roy, 'A generic task ontology for scheduling applications', 2001, Accessed: Dec. 04, 2024. [Online]. Available: https://oro.open.ac.uk/23737/4/generictaskontology.pdf

[15] H. Tan, A. Adlemo, V. Tarasov, and M. E. Johansson, 'Evaluation of an application ontology', in *Proceedings of the Joint Ontology Workshops 2017 Episode 3: The Tyrolean Autumn of Ontology Bozen-Bolzano, Italy, September 21–23, 2017*, CEUR-WS, 2017. Accessed: Dec. 04, 2024. [Online]. Available: https://www.diva-portal.org/smash/record.jsf?pid=diva2:1185058

[16] M. Sevcenko, 'Online Presentation of an Upper Ontology', 2003. Accessed: Dec. 02, 2024. [Online]. Available: https://www.semanticscholar.org/paper/Online-Presentation-of-an-Upper-Ontology-Sevcenko/2ba6de65fc254e80b9ce587cdc1ccf8f874b9bd5

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

[17] O. Bruns, T. Tietz, J. Waitelonis, E. Posthumus, and H. Sack, 'NFDIcore 2.0: A BFO-Compliant Ontology for Multi-Domain Research Infrastructures', Sep. 16, 2024, *arXiv*: arXiv:2410.01821. doi: 10.48550/arXiv.2410.01821.
[18] T. Tietz, O. Bruns, H. Fliegl, E. Posthumus, T. Schrade, and H. Sack, 'Knowledge Graph-basierte Forschungsdatenintegration in NFDI4Culture.', in *DHd*, 2023, p. 181.
[19] W. Terkaj, S. Borgo, and E. M. Sanfilippo, 'Ontology for Industrial Engineering: A DOLCE Compliant Approach.', in *FOMI*, 2022. Accessed: Dec. 05, 2024. [Online]. Available: https://ceur-ws.org/Vol-3240/paper2.pdf
[20] 'Version 7.1.3 | CIDOC CRM'. Accessed: Nov. 26, 2024. [Online]. Available: https://cidoc-crm.org/Version/version-7.1.3
[21] E. Sanfilippo, 'Ontology-based data modeling for cultural heritage', *Stud. Multidiscip. J.*, vol. 4, pp. 64–78, 2020.
[22] A. Gangemi, N. Guarino, C. Masolo, A. Oltramari, and L. Schneider, 'Sweetening Ontologies with DOLCE', in *Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web*, A. Gómez-Pérez and V. R. Benjamins, Eds., Berlin, Heidelberg: Springer, 2002, pp. 166–181. doi: 10.1007/3-540-45810-7_18.
[23] A. Pease, I. Niles, and J. Li, 'The suggested upper merged ontology: A large ontology for the semantic web and its applications', in *Working notes of the AAAI-2002 workshop on ontologies and the semantic web*, 2002, pp. 7–10. Accessed: Dec. 05, 2024. [Online]. Available: https://cdn.aaai.org/Workshops/2002/WS-02-11/WS02-11-011.pdf
[24] A. Pease, 'Standard upper ontology knowledge interchange format', *Unpubl. Lang. Man. Available Httpsigmakee Sourceforge Net*, 2004, Accessed: Dec. 05, 2024. [Online]. Available: http://ontolog.cim3.net/file/resource/reference/SIGMA-kee/suo-kif.pdf
[25] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. J. Miller, 'Introduction to WordNet: An on-line lexical database', *Int. J. Lexicogr.*, vol. 3, no. 4, pp. 235–244, 1990.
[26] I. Niles and A. Pease, 'Linking Lixicons and Ontologies: Mapping WordNet to the Suggested Upper Merged Ontology.', in *Ike*, 2003, pp. 412–416. Accessed: Dec. 02, 2024. [Online]. Available: https://www.researchgate.net/profile/Adam-Pease/publication/2880063_Linking_Lexicons_and_Ontologies_Mapping_WordNet_to_the_Suggested_Upper_Merged_Ontology/links/0fcfd511f0fd45b5b1000000/Linking-Lexicons-and-Ontologies-Mapping-WordNet-to-the-Suggested-Upper-Merged-Ontology.pdf
[27] R. Enache, 'Reasoning and language generation in the sumo ontology', Master's Thesis, Chalmers University of Technology, 2010.
[28] D. Oberle *et al.*, 'DOLCE ergo SUMO: On Foundational and Domain Models in SWIntO (Smart Web Integrated Ontology)'.
[29] I. Niles and A. Pease, 'Towards a standard upper ontology', in *Proceedings of the international conference on Formal Ontology in Information Systems - Volume 2001*, Ogunquit Maine USA: ACM, Oct. 2001, pp. 2–9. doi: 10.1145/505168.505170.
[30] P. Ronzino, 'Harmonizing the CRMba and CRMarchaeo models', *Int. J. Digit. Libr.*, vol. 18, no. 4, pp. 253–261, 2017.
[31] '7.1.3/owl/CIDOC_CRM_v7.1.3.owl · master · CIDOC-CRM / CIDOC CRM Base', GitLab. Accessed: Dec. 05, 2024. [Online]. Available: https://gitlab.isl.ics.forth.gr/cidoc-crm/cidoc_crm_rdf/-/blob/master/7.1.3/owl/CIDOC_CRM_v7.1.3.owl
[32] E. Coburn, R. Light, G. McKenna, R. Stein, and A. Vitzthum, 'LIDO - Lightweight Information Describing Objects Version 1.0'.
[33] C. Bekiari, M. Doerr, P. Le Boeuf, and P. Riva, 'Definition of FRBRoo: A conceptual model for bibliographic information in object-oriented formalism', 2017, Accessed: Dec. 05, 2024. [Online]. Available: http://repository.ifla.org/handle/123456789/659
[34] P. L. Boeuf, 'A Strange Model Named FRBROO', in *The FRBR Family of Conceptual Models*, Routledge, 2013.

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

[35] H. J. Pandit, B. Esteves, G. P. Krog, P. Ryan, D. Golpayegani, and J. Flake, 'Data Privacy Vocabulary (DPV) -- Version 2', Aug. 27, 2024, *arXiv*: arXiv:2404.13426. doi: 10.48550/arXiv.2404.13426.

[36] J. P. McCrae, J. Bosque-Gil, J. Gracia, and P. Buitelaar, 'The OntoLex-Lemon Model: development and applications'.

[37] T. Gu, X. H. Wang, H. K. Pung, and D. Q. Zhang, 'An Ontology-based Context Model in Intelligent Environments', Mar. 06, 2020, *arXiv*: arXiv:2003.05055. doi: 10.48550/arXiv.2003.05055.

[38] M. Poveda-Villalón, M. Suárez-Figueroa, R. García-Castro, and A. Gómez-Pérez, 'A Context Ontology for Mobile Environments', presented at the CIAO@EKAW, Oct. 2010. Accessed: Nov. 21, 2024. [Online]. Available: https://www.semanticscholar.org/paper/A-Context-Ontology-for-Mobile-Environments-Poveda-Villal%C3%B3n-Su%C3%A1rez-Figueroa/855b19ba68e2f4cb81e1aa3c1af89c588e94543b

[39] D. Vrandečić, 'Wikidata: a new platform for collaborative data collection', in *Proceedings of the 21st International Conference on World Wide Web*, in WWW '12 Companion. New York, NY, USA: Association for Computing Machinery, Apr. 2012, pp. 1063–1064. doi: 10.1145/2187980.2188242.

[40] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives, 'DBpedia: A Nucleus for a Web of Open Data', in *The Semantic Web*, K. Aberer, K.-S. Choi, N. Noy, D. Allemang, K.-I. Lee, L. Nixon, J. Golbeck, P. Mika, D. Maynard, R. Mizoguchi, G. Schreiber, and P. Cudré-Mauroux, Eds., Berlin, Heidelberg: Springer, 2007, pp. 722–735. doi: 10.1007/978-3-540-76298-0_52.

[41] S. Guth, 'Rights Expression Languages', in *Digital Rights Management*, vol. 2770, E. Becker, W. Buhse, D. Günnewig, and N. Rump, Eds., in Lecture Notes in Computer Science, vol. 2770. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 101–112. doi: 10.1007/10941270_8.

[42] H. Abelson, B. Adida, M. Linksvayer, and N. Yergler, 'ccREL: The creative commons rights expression language'. Technical report, Creative Commons, 2008. Accessed: Dec. 02, 2024. [Online]. Available: https://maisonbisson.com/post/hnews-might-not-be-so-bad/wiki-creativecommons-org-Ccrel-1.0.pdf

[43] D. Heckmann, T. Schwartz, B. Brandherm, M. Schmitz, and M. Wilamowitz-Moellendorff, 'Gumo – The General User Model Ontology', Jul. 2005, pp. 428–432. doi: 10.1007/11527886_58.

[44] M. Golemati, A. Katifori, C. Vassilakis, G. Lepouras, and C. Halatsis, 'Creating an Ontology for the User Profile: Method and Applications'.

[45] L. Razmerita, 'An Ontology-Based Framework for Modeling User Behavior—A Case Study in Knowledge Management', *IEEE Trans. Syst. Man Cybern. - Part Syst. Hum.*, vol. 41, no. 4, pp. 772–783, Jul. 2011, doi: 10.1109/TSMCA.2011.2132712.

[46] F. Anvari and H. M. T. Tran, 'Persona ontology for user centred design professionals', in *Proceedings of the ICIME 4th International Conference on Information Management and Evaluation*, 2013, pp. 35–44. Accessed: Nov. 22, 2024. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=-ZkTBAAAQBAJ&oi=fnd&pg=PA35&dq=Anvari,+F.,+Tran,+H.M.T.:+Persona+ontology+for+user+centred+design+professionals.+&ots=raJTXun9fK&sig=ZrpD4oQY9BRm_qbJ9yjO-ylXQsI

[47] W. W. Sim and P. Brouse, 'Towards an Ontology-based Persona-driven Requirements and Knowledge Engineering', *Procedia Comput. Sci.*, vol. 36, pp. 314–321, Jan. 2014, doi: 10.1016/j.procs.2014.09.099.

[48] F. Abel *et al.*, 'A framework for flexible user profile mashups', *Proc. Int. Workshop Adapt. Pers. Web 20 APWEB 20 2009 Trento Italy June 22 2009 Conjunction UMAP09*, pp. 1–10, 2009.

[49] M. Yudelson, T. Gavrilova, and P. Brusilovsky, 'Towards User Modeling Meta-ontology', in *User Modeling 2005*, L. Ardissono, P. Brna, and A. Mitrovic, Eds., Berlin, Heidelberg: Springer, 2005, pp. 448–452. doi: 10.1007/11527886_62.

[50] Y. Alfaifi, 'Ontology Development Methodology: A Systematic Review and Case Study', in *2022 2nd International Conference on Computing and Information Technology (ICCIT)*, Jan. 2022, pp. 446–450. doi: 10.1109/ICCIT52419.2022.9711664.

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

[51] M. Poveda-Villalón, A. Fernández-Izquierdo, M. Fernández-López, and R. García-Castro, 'LOT: An industrial oriented ontology engineering framework', *Eng. Appl. Artif. Intell.*, vol. 111, p. 104755, May 2022, doi: 10.1016/j.engappai.2022.104755.
[52] N. Matentzoglu *et al.*, 'Ontology Development Kit: a toolkit for building, maintaining and standardizing biomedical ontologies', *Database*, vol. 2022, p. baac087, Jan. 2022, doi: 10.1093/database/baac087.
[53] 'REEVALUATE/CACAO'. Accessed: Dec. 05, 2024. [Online]. Available: https://github.com/REEVALUATE/CACAO
[54] G. G. P. License, 'Gnu general public license', *Retrieved Dec.*, vol. 25, p. 2014, 1989.
[55] J. H. Saltzer, 'The origin of the "MIT license"', *IEEE Ann. Hist. Comput.*, vol. 42, no. 4, pp. 94–98, 2020.
[56] M. Kim, 'The Creative Commons and copyright protection in the digital era: Uses of Creative Commons licenses', *J. Comput.-Mediat. Commun.*, vol. 13, no. 1, pp. 187–209, 2007.

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

PARTNER		SHORT NAME
	CENTRE FOR RESEARCH & TECHNOLOGY HELLAS	CERTH
	FRAUNHOFER INSTITUTE FOR OPEN COMMUNICATION SYSTEMS FOKUS	FOKUS
	FROMSCRATCH DESIGN STUDIO BV	FS
	FONDAZIONE LINKS - LEADING INNOVATION & KNOWLEDGE FOR SOCIETY	LINKS
	FFP FILM- & FERNSEHPRODUKTION GMBH	FFP
	NUROGAMES GMBH	NURO
	GVAM GUIAS INTERACTIVAS SL	GVAM
	STIFTUNG PREUSSISCHER KULTURBESITZ	SPK
	HYPERTECH	HYP
	OLYMPIAKO MOUSEIO	OLYMPIC
	ARTHUR'S LEGAL BV	ARTHUR
	KATHOLIEKE UNIVERSITEIT LEUVEN	KU Leuven

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389

	AG CULTURELE INSTELLINGEN ANTWERPEN/ ERFGOED	MoMU
	EUROPEAN FASHION HERITAGE ASSOCIATION	EFHA
	3D RESEARCH SRL	3DR
	FONDAZIONE AQUILEIA	AQUILEIA

Funded by the European Union

This project has received funding from the European Union Horizon Research and Innovation programme under grant agreement No 101132389