heritage

MDPI

Review

Semantic Bridging of Cultural Heritage Disciplines and Tasks

Efthymia Moraitou *,†, John Aliprantis †, Yannis Christodoulou †, Alexandros Teneketzis † and George Caridakis †

School of Social Science, Cultural Technology and Communication, University of the Aegean, University Hill, 81132 Mytilene, Greece; jalip@aegean.gr (J.A.); ychris@aegean.gr (Y.C.); ateneketzis@aegean.gr (A.T.); gcari@aegean.gr (G.C.)

* Correspondence: e.moraitou@aegean.gr; Tel.: +30-698-448-1408

† These authors contributed equally to this work.

check for updates

Abstract: The Cultural Heritage (CH) domain encompasses a wide range of different disciplines, serving the study, interpretation, curation, and preservation of objects, collections, archives, sites, and the dissemination of related knowledge. In this context, stakeholders generate, retrieve, and share a vast amount of diverse information. Therefore, information interoperability has been considered a crucial task, especially in terms of semantics. In this way, the CIDOC CRM (International Committee for Documentation Conceptual Reference Model) has been widely used as an underlying model that offers interoperability between CH domain metadata standards and ontologies. To the best of our knowledge, an overall review of mapping, merging, and extending this core ontology, as well as an aggregate table which classifies and correlates those ontologies and standards, has not yet been presented. Our study conducts an aggregate review of relevant published efforts and outlines the various associations between them, encapsulating the CIDOC CRM and its specialized models, as well. This work aims to further clarify the field and scope of the different works, identify their methods, and highlight the semantic overlap, or differences, between them.

Keywords: CIDOC CRM; merging; mapping and extending; aggregation table; overview

1. Introduction

The study and preservation of Cultural Heritage (CH) requires the collection, storage, and processing of diverse cultural information related to different disciplines and activities within the domain. Combining this heterogeneous information and interconnecting the various data sources of CH institutions, research centers, and individuals is considered indispensable in this process, though not trivial to accomplish. Therefore, services such as the unified management, search, and retrieval of cultural content, as well as data integration, data mining, and knowledge extraction can all benefit from semantic web technologies [1], which can deliver interoperability. Particularly, top-level, or core, ontologies define basic concepts and their interrelations, which are common among different sub-domains, thereby facilitating the mapping of knowledge representation and metadata schemas between different disciplines and tasks, while the included entities could be extended by being described in more specific terms, thereby leading to more specialized conceptualizations.

In this way, the International Committee for Documentation Conceptual Reference Model (CIDOC CRM) constitutes a core ontology of the CH domain, developed by a multidisciplinary team of domain experts. The CIDOC CRM can be characterized as compact and efficient, considering the different sub-domains it aims to cover. Moreover, its structure is based on temporal events, whereas the structure of documentation reports or metadata schemas within the domain tends to be item-centric. Thus, the

CIDOC CRM provides semantic definitions for the optimal integration, mediation, and interchange of diverse cultural information through the Web, which is located in different local, and often isolated, information sources. Eventually, the CIDOC CRM may serve as a global schema, query mediation guideline for the conceptual modeling, or building, of information systems, or be utilized for tagging schemes development, thereby offering value in organizing and reusing cultural information [2].

Research projects, academics, and professionals in the CH community tend to support the usage of a common conceptual model in order to conduct knowledge representation and data integration. The CIDOC CRM has been widely used for the mapping and merging of metadata standards and ontologies related to the CH domain. Additionally, it has served as the basis for developing extensions in order to meet the needs of specialized fields and tasks. Therefore, our primary objective is the collection and organization of the work that has been done over the last twelve years in CIDOC CRM merging, mapping, and extension, without restricting our research to a specific sub-domain. This contribution is considered beneficial, since it aims to capture the whole universe lying under the semantic umbrella of the CIDOC CRM. Additionally, we aim to clarify the scope and context of each approach in regard to the respective discipline or task, as well as the methods and technologies used. Based on the aforementioned analysis, we intend to outline the potential conceptual relevance of the different approaches; this outline is considered imperative for achieving data interoperability within the domain. Finally, we intend to highlight the semantic overlap between linked ontologies and metadata schemas, as well as some emerging fields for further development and exploitation.

In the remainder of this paper, we first briefly present the resources and method of our review and specify the different disciplines and tasks of cultural information, based on the various existing attempts to employ the CIDOC CRM (Section 2). Thereafter, we present published works related to either extending the CIDOC CRM or employing it for mapping or merging ontologies within the domain (Section 3). Additionally, we describe a possible correlation between the different ontologies and metadata standards and highlight some fields for further development (Section 4). Finally, we conclude with a discussion of our observations as well as future work on the conceptual representation and data-semantic interoperability of the CH domain (Section 5).

2. Method

2.1. Data Sources Collection and Study

In the context of our review, we came across interesting projects and articles, which have been published over the past twelve years, about the mapping, merging, and extension of the CIDOC CRM. We chose this particular period due to the milestone year of 2006, when the CIDOC CRM became an official standard[1]. The data sources of our research include Semantic Scholar, Springer Link, Getty, Science Direct, as well as the resources from the official CIDOC CRM site.

Our study has been conducted over a six-month period, from June to November 2018. Aiming to cover the largest possible portion of the CH domain, we gathered, studied, and discussed the published works. Therefore, we tried to equally present the approaches of different years and sub-domains, situating the use of CIDOC CRM in merging, mapping, and extension approaches as the benchmark of our study. Additionally, it was our intention to present both older and later works in order to have a complete and updated view of our research field.

At this point, it is important to mention that it was outside of our scope of research to include works proposing ad-hoc combinations of CIDOC CRM entities and relations for the knowledge representation and information modeling of specific disciplines and tasks of the CH domain. Though

---

[1] ISO Standard "Information and Documentation: A Reference Ontology for the Interchange of Cultural Heritage Information" (ISO 21127:2006).

we recognize the importance of these projects, we intended to focus our study on the aspect of data interoperability rather than data interpretation.

2.2. Identifying Disciplines and Tasks

The CH domain embraces several disciplines, such as cultural management, curation and museology, library and archival science, preservation science, archaeology, architecture, history, art history, geography, as well as artistic practice itself. Considering the diversity of these aspects, it is easy to perceive the particularity of cultural information. Even if the aforementioned sciences are interconnected and essentially complementary, their methodologies and aims differ.

As an example, let us examine archaeology and conservation. Archaeology science refers to the scientific excavation, study, interpretation, and virtual reconstitution of material human remains and relevant human activity, while conservation science refers to studies and actions aiming to preserve tangible CH in the best possible condition for future generations [3,4]. While these two disciplines have different scopes and methods, it is common knowledge that they benefit from exchanging their scientific findings in order to reach more efficient conclusions. However, the documentation of the respective scientists' and professionals' observations, conclusions, and applied tasks may require different levels of detail and focus particularly on different fields of interest. This means that even if they share a common ground, the information produced may differ.

Several different activities and processes can be identified in the context of CH study, restoration, preservation, and presentation. These activities and processes may often be shared among CH sub-domains and are conducted by CH institutions and research centers. These may include scientific observation, analysis and diagnosis, dating, conservation treatment, information provenance, information retrieval, and scientific reasoning, and may be approached through different methodologies and means [5]. For example, the dating of a painting may be achieved through analysis of its iconography and production techniques, as well as chemical analysis of its pigments and binders. In this context, the different aspects of this task may be conducted by different CH specialists and lead to more complete and efficient conclusions.

In the documentation processes of CH, the activities of digitization, multimedia production, and visualization have held a prominent position, especially during the last decade. The media produced and used may include images, 2D designs, 3D models, diagrams, video, and audio [6]. Furthermore, these media are not always used exclusively for scientific documentation and information communication among the scientific community, but also for the promotion of CH to the public. 3D reconstructions, animations, virtual museums, and narratives with multimedia content are produced by presenting and reusing cultural information.

Given these diverse fields of interest, methods, and tasks, as well as the multimodality of cultural information, it is obvious that there is an imperative need for semantic interoperability. To this end, CIDOC CRM can be perceived as the "semantic glue", which connects this heterogeneous material [1]. Until now, there have been a significant number of research projects that have used CIDOC CRM for the integration of cultural data and sources, and for semantic querying/retrieval, such as ResearchSpace (museum collections data) [7], Sampo Model (cultural content, novels and historical data) [8], STAR (Semantic Technologies for Archaeological Resources) project (cultural and temporal data) [9], the DECHO (Digital Exploration of Cultural Heritage Objects) Project (archaeological data and visual media) [10], ARCHES (cultural and spatial data) [11], Wisski (Wissenschaftliche KommunikationsInfrastruktur) project (communication and documentation data) [12], DALI (Data Aggregation and Linking Interface) (preservation data) [13], DOC-CULTURE (diagnosis data) [5], PARTHENOS Project (history, linguistic studies, cultural heritage, archaeology, and related fields across the digital humanities) [14], ARIADNE (Advanced Research Infrastructure for Archaeological Dataset Networking in Europe) (archaeological data infrastructure) [15], 3D-COFORM (3D Collection Formation) (3D-digitisation data) [16], CASPAR (Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval) (digital preservation) [17], PARCOURS project

(conservation-restoration data) [18], GROPLAN Project (archaeology and 3D modelling) [19], the Papyrus Project (history of science and technology) [20], DOREMUS (musical works) [21], and the Invisibilia Project (contemporary art) [22].

Furthermore, it is important to mention some notable attempts at mapping cultural data collections with the CIDOC CRM, on behalf of several major art/cultural organizations and institutions around the globe. The British Museum has released a Semantic Web version[2] of its database, which complements the Collection Online search facility; this database complies with the CIDOC CRM conceptual framework. In this way, the potential of semantic harmonization with other organizations is greatly enhanced, allowing researchers to perform data searches much more precisely and efficiently. The Linked Data Finland Research Initiative is working on an open and linked web platform, the LDF.fi platform,[3] in order to provide the public with aggregated, harmonized, and enriched cultural/historical data in a cost-efficient way using Semantic Web technologies [23]. The Virtual Emigration Museum has developed a reduced CRM-compatible form ontology, based on the CIDOC CRM, in order to extract knowledge from its hosted data and conceptualize relevant emigration documents stored in a relational database [24]. The ontology is instantiated through a parser that automatically translates a plain text description of emigration data into RDF (Resource Description Framework). Furthermore, the RDF triples are saved in order to enable querying of the RDF data using SPARQL (Simple Protocol and RDF Query Language). Finally, Cultura-Italia, the Italian national cultural aggregator, has transformed its digital cultural resources into records structured according to the CIDOC CRM, by defining a set of mapping rules [25].

These approaches illustrate the capabilities and implementation range of the CIDOC CRM in the various CH sub-domains and activities, and address some of the aforementioned sciences, tasks, documentation media, and organizational requirements (e.g., in terms of time and space). Moreover, most of the aforementioned projects, as well as additional research projects and relevant publications, propose different cases of the CIDOC CRM mapping/merging with metadata schemas or ontologies, as well as its extension for accommodating specific informational needs (see Table 1 and Section 3). These works have been categorized, and a conceptual correlation between them, which is considered valuable for future work on knowledge representation and management of the domain, is proposed.

Table 1. Projects related to merging, mapping, and extending the CIDOC Conceptual Reference Model (CRM), grouped by discipline (for a detailed presentation see Section 3).

Disciplines	Projects
Archaeology	Star Project [9], DECHO Project [10], ARCHES Project [11], PARTHENOS Project [14], ARIADNE Project [15], GROPLAN Project [19]
Architecture	ARCHES Project [11]
CH (in general)	"Sampo" portals [8], PARTHENOS Project [14], 3D-COFORM Project [16]
History	"Sampo" portals [8], PARTHENOS Project [14], Papyrus Project [20]
Library and Archival Science	WissKI Project [12], DOREMUS Project [21]
Museology	ResearchSpace [7], Invisibilia Project [22]
Preservation Science	DALI Project [13], DOC-CULTURE Project [5], CASPAR Project [17], PARCOURS Project [18]

---

2   http://collection.britishmuseum.org
3   http://www.LDF.fi

3. Merging, Mapping and Extending the CIDOC CRM

Reviewing publications related to the CIDOC CRM's implementations within the CH domain, we discerned some interesting cases on the model's mapping and merging with metadata standards and ontologies, as well as its extension with thesauri, specialized concepts, and relations. In particular, the merging approaches presented in Section 3.2 define new merged ontologies (e.g., the ABC-CIDOC CRM merged ontology is a result of the alignment between an ontology of the CH domain and the CIDOC CRM [26]). On the other hand, mapping approaches (Section 3.2) may refer to (i) metadata paths or elements mapped to conceptually equivalent CIDOC CRM entities and relation paths (e.g., Encoded Archival Description (EAD) mapping to the CIDOC CRM or the Italian Central Institute for Catalogue and Documentation (ICCD) RA (Schema – Archaeological Finds) mapping to the CIDOC CRM), (ii) metadata concepts mapped to CIDOC CRM entities and metadata relations mapped to CIDOC CRM properties (e.g., MPEG7 (Motion Picture Experts Group) -CIDOC CRM mapping), (iii) ontology classes mapped to corresponding classes of the CIDOC CRM and CIDOC CRM family models (e.g., the Building Conservation Ontology mapping to the CIDOC CRM, CRMsci, CEMdig and CRMinf). Finally, extension approaches (Section 3.3) refer to the addition of new classes and properties in the structure of the CIDOC CRM and/or its family models, thereby expanding its capabilities regarding the knowledge representation of CH disciplines and specialized tasks.

To be thoroughly presented and organized, various approaches have been categorized based on applied methods and identified disciplines and tasks (Section 2). As such, the current section first includes a presentation of CIDOC CRM official family models, and then describes the published works about CIDOC CRM mapping and extension.

3.1. CIDOC CRM Family Models

Due to its structure and level of abstraction, the CIDOC CRM is able to expand and cover different CH sciences and processes related to the investigation of the human past and activity. Until now, a number of compatible models have been developed and proposed by the CIDOC CRM Special Interest Group, extending the model's main entities by defining more specialized concepts.

At the higher levels of the family models hierarchy[4] lies the CRMsci model, which serves as a global schema for data integration related to scientific observation, measurements, and processed data, as well as the CRMinf model, which facilitates data integration related to argumentation and reasoning in the context of descriptive and empirical sciences involved in CH [27,28]. Encoded in RDFS, CRMsci constitutes an ontology for integrating the results of empirical sciences. It distinguishes the observation process from its results and the observed object from a sample one. Similarly, CRMinf focuses on the different arguments and states of belief based on observations and results, and can, therefore, be perceived as an extension of CRMsci [29].

At a conceptually lower level lie the CRMba, CRMarchaeo and CRMdig family models. CRMba is an ontology, encoded in RDFS, for documenting information about archaeological buildings. For that purpose, CRMba supports the identification and understanding of a building's components and structures, as well as its stratigraphy, different building phases, and different uses. CRMarchaeo supports the documentation of the excavation process and related activities, providing entities and properties for describing the stratigraphy and individual layers, events, or activities relevant to their creation or modification, as well as discovered human remains or artifacts. Finally, CRMdig is another RDFS-encoded model, used for encoding data relevant to measurements and digitization processes, production devices and methods, digitization products provenance, and complex digital representations (2D models, 3D models, etc.) [30,31].

---

[4] https://www.ics.forth.gr/index_main.php?l=e&searchtype=st_p&c=229&i=

In order to support the correlation between cultural and spatial information, the CRMgeo family model has been developed [32]. CRMgeo also has an RDFS encoding and is the result of integration of the OGC/ISO standards for geographic information with the CIDOC CRM. CRMgeo focuses on the integration of spatiotemporal features of temporal entities with persistent items, facilitating the reasoning over space in the geophysical sense. Additionally, the compatible FRBRoo model is the result of a merging between the FRBR ("Functional Requirements for Bibliographic Records") and the CIDOC CRM. As a CIDOC CRM extension, FRBRoo represents semantics about bibliographic information and facilitates the integration, mediation, and interchange between bibliographic and museum information [33]. Accordingly, PRESSoo constitutes a further extension of FRBRoo, in order to cover the semantics of bibliographic information about periodicals [34]. CRMgeo, FRBRoo, and PRESSoo may be partly incorporated with other family models, in order to complement the documentation process of their specific domain or task.

Eventually, the ongoing Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies (PARTHENOS) project proposes a CRMpe compatible extension, and member, of the family models [14]. The PARTHENOS entities model extends the CIDOC CRM based on the analysis of data structures of the repositories used in the project. The primary scope of this approach is the implementation of a semantic structures for supporting the rich relationships that emerge from the relevant research. Based on an analysis of basic relevant concepts, a typical model for the research infrastructure sub-domain was developed. First, the identified concepts and their relations were represented as classes and properties, forming a domain-specific ontology. Then, the basic relations and class extensions were added, and finally, the model was harmonized according to the CIDOC CRM. Encoded in RDF(S), CRMpe aims to provide a common understanding among the various sources and to ensure that data generated in heterogeneous forms will have a semantic re-expression that facilitates their interoperability with other datasets.

3.2. Merging and Mapping Approaches

The cultural content of memory institutions is often described by metadata standards, which define a set of mandatory or optional elements, such as properties and data formats that can be applied (e.g., Dublin Core, VRA (Visual Resources Associations)). Although their consistent use is important for data management, access, combination, and interoperability in general, they are often not enough to achieve data integration [29,35]. As already mentioned, data integration has been a challenging and increasingly interesting task within the CH domain and is focused not only on the coding level but also on the semantic level, from which the particular conceptual characteristics of the domain are derived. To this end, semantic web technologies, especially ontologies, may facilitate data integration and contribute to the confrontation of semantic heterogeneity. In our case, CIDOC CRM has served as a global schema through which metadata of different sources can be mapped and ontologies can be merged, thereby harmonizing the different views of knowledge between diverse, yet related, communities.

In this context, a common ontology has been developed through the mapping of the CIDOC CRM to ABC concepts, which resulted in the ABC-CIDOC CRM merged ontology. ABC ontology facilitates bibliographic information exchange and integration, and therefore bridges and connects CH and library knowledge [2]. The basic idea of the ABC-CIDOC CRM project was the preservation of ontologies' prime use and their integration to a wider, more coherent, model. This wider model would be valuable for the integration of the two different domains, as well as for retrieving unified information from remote sources. An interesting remark regarding these ontologies is that although they both model the concept of change over time, the nature of change is different in each case. On the one hand, ABC describes how bibliographic objects change over time (e.g., the versioning of digital objects or the production of derivative works). On the other hand, the CIDOC CRM covers changes in context and interpretation of cultural objects over time due to the appearance of new related research findings. The comparison and convergence of the two ontologies was achieved through the OntoClean

method, and the final harmonization included assigning "equivalent to", "similar to", "subclass of", and "subproperty of" relations between the classes and relations of the two ontologies.

Furthermore, regarding library and archival science, certain metadata standards, such as Encoded Archival Description (EAD) and Dublin Core (DC), have been mapped to the CIDOC CRM. In particular, as presented in [35], the CIDOC CRM was considered a mediator for the semantic integration and elimination of semantic heterogeneities between EAD and DC. According to the path-oriented methodology presented in that work, metadata paths were created and mapped to conceptually equivalent CIDOC CRM entities and relations paths. In the context of that work, a metadata path is defined as a sequence of elements, sub-elements (or element refinements), encoding schemes, and vocabulary terms, while a CIDOC CRM path is defined as a chain of the form "entity-property-entity". In some cases, new classes and properties were added in order to cover particular metadata fields. An interesting issue highlighted by the authors is the event orientation of CIDOC CRM, a feature that led to the mandatory intermediation of an event or activity in order to interlink an actor with the described object. An approach similar to the aforementioned one, involving the mapping of DCMI Type Vocabulary to the CIDOC CRM, is described in [36]. DC is widely used in the domains of library, archival science and CH, and the whole process of specifying DC metadata paths and respective CIDOC CRM paths contributed to the organization of DC elements.

CIDOC CRM has also been exploited in archaeology for metadata mapping and data integration. In the context of the ARIADNE project [15], the Italian Central Institute for Catalogue and Documentation (ICCD) was successfully mapped to the CIDOC CRM. The primary scope of the ICCD is the establishment of a national catalogue of Italian cultural heritage. The ICCD cataloguing standards include control tools, such as vocabularies and lists of terms, as well as rules and guidelines, which illustrate best practices and methods for the acquisition and generation of CH documentation. According to the analysis of ICCD presented in [15], the RA Schema-Archaeological Finds was considered the most significant model of the ICCD archaeological cataloguing system, and is, therefore, used as a starting point for mapping ICCD to the CIDOC CRM. Furthermore, a conceptual mapping between the RA model and CIDOC CRM is described, as well as an example of mapping between ICCD RA elements and vocabularies, and CIDOC CRM entities and relations paths.

In the same context, the mapping of the International Core Data Standard for Archaeological and Architectural Heritage (CDS) to the CIDOC CRM was carried out during the development of the Arches system [11]. CDS is the result of collaboration between CIDOC and the archaeology documentation group of the Council of Europe, and provides a starting point for structuring data in Arches. Utilizing this model, Arches maps the identified user data and their respective data fields to CIDOC CRM entities. Structured as a graph, the Arches system consists of i) the Arches Server which provides system administration and information management services, and ii) an implementation of Data Management Packages. The initial package is the CDS Package, through which graphs identify the CRM classes (nodes) and CRM properties (edges), creating a framework for storing information. Creating graphs using the CDS Package necessitated representing CDS using classes and properties as defined by the CIDOC CRM. To create machine-readable graphs, a spreadsheet was created, identifying the CDS field name represented, the CRM class used to represent it, and its path (starting from the top node). The mapping of the CDS to the CRM was based on a mapping of MIDAS Heritage. Moreover, Gephi visualization and analysis software was used for visualizing the mappings when necessary.

In the history sub-domain of CH, Nasir and Noor [37] present a mapping and merging between the Traditional Malay Textile Knowledge Model (TMT) and the CIDOC CRM. The TMT Knowledge Model focuses on textile descriptions and related historical factors of the Malay Peninsula, aiming to capture the related semantics. First, the TMT concepts and properties underwent a refinement process in order to match the CIDOC CRM structure. Next, manual mapping was conducted in order to identify and align similar or overlapping concepts and achieve ontology similarity. Lastly, a foundation for a common understanding in automating the whole process of mapping between the TMT and the CIDOC CRM was provided. As a result of that work, the final harmonization between

the two ontologies included "equivalent to", "similar to", and "subclass of" relations between each other's classes.

Cultural data consists not only of textual information, but also of various types of multimedia. Although there exist metadata schemas for organizing and managing multimedia objects and their corresponding models, no model has yet been defined to describe museum multimedia content. As such, many efforts have been made to integrate and map the MPEG-7 multimedia content standard to the CIDOC CRM, such as the one mentioned in [38]. The model of mapping MPEG-7 to the CIDOC CRM, allowing the transformation of MPEG-7 descriptions to CIDOC CRM descriptions presented in the aforementioned work, facilitates the exploitation of annotations in multimedia content hosted in CH digital libraries. According to the MPEG-7 CIDOC mapping model, MPEG-7 types (MPEG-7 concepts) are mapped to semantically correspondent CIDOC CRM entities, while the MPEG-7 relations are mapped to CIDOC CRM properties. Based on this mapping model, the authors developed a toolkit for the automatic transformation of CIDOC CRM descriptions into MPEG-7 syntax.

In the same direction, mapping between the CIDOC CRM and VRA Core 4.0 is presented in [39]. VRA Core 4.0, originally created by the Visual Resources Association's Data Standards Committee, is a metadata schema used by the Heritage Community. VRA Core 4.0 provides guidance for describing works, such as paintings, statues or other artwork, visual representations of artwork, and collections of groups of works or images. The method proposed for mapping VRA to the CIDOC CRM is based on paths and uses the Mapping Description Language (MDL) to define semantic rules for mapping from the source schema to the target schema. Each schema element is represented as a VRA path based on XPath, which is then semantically translated to an equivalent path of CIDOC CRM classes and relations.

Combining conservation science with the digitization and visualization of relevant information, Messaoudi [40] proposes a domain ontology model mapped to the CIDOC CRM for the reality-based 3D semantic annotations of building conservation states. The domain ontology considers both qualitative and quantitative aspects of historical buildings by merging three dimensions (semantic, spatial, and morphological), efficiently bridging conservation science and information visualization. This domain ontology has been developed based on Lassila's method, while two glossary types were used for representing stone deterioration patterns and architecture. First, a thesaurus and a lightweight ontology were developed, and then a heavyweight ontology was implemented by adding more terms, inference rules, constraints, and axioms. Finally, some classes of the domain ontology were mapped to corresponding classes of the CIDOC CRM core, CRMsci, CRMdig, and CRMinf.

Finally, considering the information provenance and the development of research infrastructures that gather and integrate heterogeneous datasets, a mapping between the CLARIN's (European Research Infrastructure for Language Resources and Technology) underlying metadata model (CMDI—Component Metadata Infrastructure) [41] and the CRMpe extension of the CIDOC CRM, developed during the PARTHENOS Project [14], has been proposed in [42] and is currently under development. CMDI is not just one schema, but a framework for creating and reusing self-defined metadata schemas in order to meet the various needs of data providers. Therefore, the proposed mapping strategy relies on semantic interoperability established in the CLARIN infrastructure, alongside X3ML mapping files and a Java application. The purpose of mapping the CLARIN metadata to the PE (Parthenos Entities) model is to deliver information about CLARIN resources to the PARTHENOS Project, expressing these resources in a well-established, high-level conceptual model like the CIDOC CRM.

3.3. Extending Approaches

Being a core ontology, the CIDOC CRM includes concepts that are common in different sub-domains of a wider "universe of interest", in this case CH. As such, it can provide the basis for the specialization of concepts and vocabularies of various CH sub-domains and tasks [2]. According to

our research, several projects have extended the CIDOC CRM in order to capture particular concepts and cover the requirements of different fields of interest.

To begin with, an extension of the CIDOC CRM for the definition of archaeological periods and chronologies was presented during the Semantic Technologies of Archaeological Resources (STAR) project [9]. In the context of archaeology, textual or numerical data referring to time periods can be managed and organized using controlled vocabularies. Although controlled vocabularies can be semantically exploited with the use of the Simple Knowledge Organization System (SKOS), in order to facilitate reasoning, temporal relations need to be fully described. Consequently, CRM-EH environmental archaeology extension has been developed as an extension of the CIDOC CRM that covers the tasks of excavation and analysis [43]. An RDF implementation of the model, which complements the RDFS version 4.2 of the CIDOC CRM, is provided. Furthermore, the model has been combined with controlled vocabularies and thesauri, which have been transformed into a SKOS to define concepts and correlations between terms for describing periods and chronologies.

Regarding the sub-domain of archaeology, [19] presents an extension of the CIDOC CRM that was developed in the context of the GROPLAN project. This project aims to provide archaeologists with user-friendly measurement tools, combining 3D models of a site with integrated archaeological information. For that purpose, two ontologies have been aligned: the first ontology captures the photogrammetrical measurement and the geo-localization of measured items, while the second captures concepts about archaeological artifacts, extending the CIDOC CRM with new classes (in OWL 2 (Web Ontology Language) terms). Therefore, the extension of the CIDOC CRM, structured around the concept of the E22 Man-Made Object, ensured the coherence of measurement results used for 3D modeling based on photogrammetry [19].

Ontoceramic is one more extension of the CIDOC CRM that is relevant to the sub-domain of archaeology [44]. Encoded in OWL 2, Ontoceramic facilitates the cataloguing and classification of ancient ceramics, addressing problems related to knowledge management. Specifically, Ontoceramic introduces the concept of *location*, allowing correlations of locations with findings. Furthermore, it facilitates the description of object fragments, enabling the description of both the whole object as well as its individual fragments. As such, with Ontoceramic, one can describe how a unique ceramic is composed by specifying how its fragments are linked. Finally, the ontology represents the different shapes of findings through the "Shape" class and its subclasses.

Finally, regarding the sub-domain of archaeology and in particular the discipline of epigraphy, CRMepi is proposed as a discipline-specific extension of the CIDOC CRM [45]. Considering the main concepts of the epigraphy domain, as well as the compatible models of CRMsci and CRMarchaeo, CRMepi introduces some new concepts that are more appropriate for representing the epigraph itself, the activity of creating inscriptions, the locations of epigraphs, the transcription process, and systems used for writing and transliterating epigraphs. Furthermore, additional relations were defined to better correlate the relevant concepts.

Regarding the sub-domain of art history, Anh Tran and Isemann present an extension of the CIDOC CRM aimed at modeling creative influence between artists, and thus describing the associations between the historical artwork of the Dutch Renaissance [46]. The CIDOC CRM provides a common model for heterogeneous cultural data, which is significantly helpful in the aforementioned approach, since the creator-artwork associations are already modeled. Furthermore, the OWL implementation of CIDOC CRM provides the required expressiveness for defining axioms. Based on the Erlangen CRM, an OWL Description Language (DL) implementation of the CIDOC CRM defines cardinalities and constraints that are not explicitly defined in the CRM while also granting more axiomatic expressiveness [47]; the authors imported the Getty ULAN controlled vocabulary in order to capture information about artists, as well as an auxiliary ontology for likelihood and connection strength measurement. This ontology was populated through an RDF graph containing over 900 instances of individual connections between works of art, referencing external Linked Open Data from the Getty ULAN.

Regarding the sub-domain of history of science and technology, and the need for terminology awareness on behalf of history researchers about a past topic of interest, the Papyrus project presents an ontological approach for retrieving news content for historical research. Specifically, two ontologies were developed in order to model the disciplines of news and history. Specifically, the development of the Papyrus History ontology was based on the CIDOC CRM and captures a historical perspective on events and topics covered by the news. For that purpose, the CIDOC CRM was extended with abstract domains of periodization and classification, historiographical issues, time and evolution constructs for temporal information, and language constructs for a multilingualism model, using RDF and OWL [20].

As described in [48], biographical data can be harmonized and interlinked using Semantic Web technologies. Based on this idea, Bio CRM is presented as an extension of the CIDOC CRM for the sub-domain of biography, and is compatible with CH data hosted in museums, libraries, and archives. Bio CRM extends the CIDOC CRM by introducing classes and properties that facilitate the representation of events, people, their relationships, and their roles, while supporting role-centric modeling of information. The implementation of Bio CRM includes RDF and OWL encoding.

The CIDOC CRM has also been proven significantly useful in the CH sub-domain of conservation-restoration. Nevertheless, in some cases it is difficult to match the domain's specific concepts with the related processes through the CIDOC CRM. As witnessed during the ARIADNE project [15], sometimes it is difficult to classify the observed damages based on their degradation mechanisms. Therefore, capturing damage phenomenology using the CIDOC CRM is useful, though not always efficient, since it usually requires a more detailed level of documentation than the current conservation procedure has reached. In this context, and in order to fully cover the specific needs of the conservation-restoration sub-domain, several extensions have been proposed, introducing more specific terms and relations. A good example is the OPPRA ontology (Ontology of Paintings and Preservation of Art), which extends the CIDOC CRM, merging entities of chemistry ontologies, such as OreChem and OIA-ORE (Open Archives Initiative—Object Exchange and Reuse), in order to cover the needs of painting conservation and material analysis [49]. In particular, OPPRA aims to document descriptions of physical artifacts, events, damage mechanisms and related digital information within the sub-domain of painting conservation, in a reusable and machine-processable form. Its implementation is based on the OWL DL, offering maximum expressivity while facilitating reasoning for different aspects of art history and material science.

Likewise, to cover the requirements of non-destructive analysis and diagnosis methods of artwork during conservation, several terms that extend the CIDOC CRM have been proposed [5]. This extension, which is part of the DOC-Culture project, focuses on concepts of conservation and analysis procedures, and their sequence definitions. The three main CIDOC CRM concepts that have been extended are Physical Object Event, Document, and Measurements. The whole process is described, and conceptual maps of the new concepts and relations are presented, proposing an ontology for representing cultural artifacts of any type and format for conservation purposes.

Additionally, for the needs of the Polygnosis platform, an extension of CIDOC CRM with a thesaurus of terms relevant to the analysis of artwork materials has been developed and proposed [50]. For the purposes of the platform, a domain terminology was defined, and the individual terms were classified, forming a semantically linked thesaurus. In this case, CIDOC CRM was used for an ontology-driven faceted analysis method, defining the top-level concepts that can be extended in XML (eXtensible Markup Language). Therefore, the CIDOC CRM served as the backbone for thesaurus development and terms organization, extending four basic entities to more specialized concepts: Material Objects, Investigation Methods, Identifiable Features, and Data.

Nevertheless, the conservation-restoration domain encapsulates a wide range of processes and data that other approaches have tried to represent, using the CIDOC CRM as the core for the organization of new classes. The CORE (Conservation Reasoning) ontology builds upon, and extends, the CIDOC CRM with concepts and relations about materials and techniques, condition states, and conservation processes of artwork, particularly of byzantine icons [51]. The development of the

CORE ontology was based on empirical analysis, scientific knowledge, and existing vocabularies of the conservation domain, and has been implemented using OWL 2. It consists of a base of eleven classes, each of which branches into subclasses with semantic consistency. CIDOC CRM top-level classes capture the provenance information about an artwork while the CORE extensions capture domain-related knowledge.

Furthermore, the PARCOURS ontology [18] models information about cultural objects, phenomena and features of events, data related to scientific study and related instruments, and information about applied treatments. Similar to the aforementioned cases, the PARCOURS ontology extends the CIDOC CRM and CRMsci, defining new, more specialized entities and relations between them. Specifically, its modular framework includes i) top-level ontologies that capture the conservation-restoration objects and processes, and ii) a terminology layer for covering the conservation-restoration domain. Encoded in OWL 2 QL, PARCOURS integrates a set of thesauri that aim to manage potential mismatches, arising within the conservation-restoration terminology, at both the syntactic and semantic levels.

CIDOC CRM has been also utilized in visualizing CH domain knowledge. In the context of the Invisibilia project, the Intangible Component of Contemporary Art Ontology has been created by extending the CIDOC CRM to the sub-domain of the intangible component of contemporary art, in order to visualize stories involving artwork [22]. As noted in [22], apart from installations, contemporary art may include performances and interactive elements, which CIDOC CRM cannot fully capture at a conceptual level. As such, [22] proposes the introduction of new properties in order to describe the "invisible" components of contemporary artwork and their relations, using RDF/XML.

In the sub-domains of visualization and virtual museums, the OntoMP ontology, an ontology for the Museum of the Person (a virtual museum that exhibits life stories of common people), is proposed for building exhibition rooms within a virtual space of the museum. The museum components are encoded using XML, forming a repository. OntoMP's objective is the extraction of information from this heterogeneous dataset, including people's interviews/narratives, for implementing the virtual environment [52]. Firstly, OntoMP was designed by specifying related concepts and relations. Next, OntoMP concepts and relations were mapped to the CIDOC CRM. The approach also included the extension of the CIDOC CRM with FOAF (Friend of a Friend) and DBpedia concepts and properties, thereby providing a vocabulary for describing individuals, their activities, and their relations with other people and objects.

Finally, also regarding visualizing knowledge within the CH domain, CulTO (Cultural heritage Tool based on Ontology) has been designed to characterize religious historical buildings [53]. In order to support the modelling of CH buildings alongside visual data annotation, and develop high-level applications for data curation, retrieval, and classification, CulTO extends the CIDOC CRM by adding subclasses relative to the characteristic structures of churches (e.g., columns, shafts, and altars) to the Physical Object and Physical Property classes. CulTO is encoded in OWL 2 and supports image annotation by linking user annotations through the Annotation class in order to associate sample images.

Regarding the music domain, the DOREMUS project implements novel tools to link and explore data hosted in three French institutions, for the purpose of better representing the domain. In this context, an extension of the FRBRoo model (an extension of the CIDOC CRM itself) is proposed, allowing the description of musical works and their publications, concerts, festivals, and recordings that are part of the activities of Radio France and the Philharmonie de Paris [21]. The DOREMUS ontology facilitates the description of future events and changes between original musical work and its performed versions. A working version of the ontology in RDF-Turtle format is available on Github.

4. Overview and Association of Different Approaches

Having completed the presentation of various approaches of merging, mapping and extending the CIDOC CRM in the context of the interdisciplinary domain of CH and its sub-domains, an aggregate

table overviewing and juxtaposing the different approaches was considered a useful tool for future reference. In Table 2, the CIDOC CRM-compatible models and the 27 published works discussed in the current review are organized in chronological order, from 2006 to 2018. In cases of works that have been published in the same year, their order in the table is based on their appearance in this paper. To present some important features of each approach in a compact fashion, we have included in the table six main information categories (columns): a short description or title of the approach (Description), the type of process (merging, mapping or extending) (Type), the relevant discipline (Discipline), the relevant task (Task), the representation technology (Encoding), and the publication year (Year).

Considering all the aforementioned approaches, as well as the main role of the CIDOC CRM model, Figure 1 outlines the reviewed ontologies and metadata schemas in a way that foregrounds the various associations between them. Setting the CIDOC CRM as the top-level ontology, and following the Semantic Web stack, we have defined two additional levels of abstraction. The first one includes ontologies that have been mapped to, or merged with, the CIDOC CRM (ABC and Building Conservation Ontology), and the CIDOC CRM extensions, which cover specialized disciplines and tasks. The second one covers metadata schemas mapped to CIDOC CRM core entities. The nodes in the transparent background denote extensions of the CIDOC CRM that constitute official family models. The solid lines connect ontologies that have been combined in the context of the works included in the current review, while dashed lines depict a potential conceptual relevance between ontologies. The different colors indicate the different sub-domains of interest (see illustration legend).
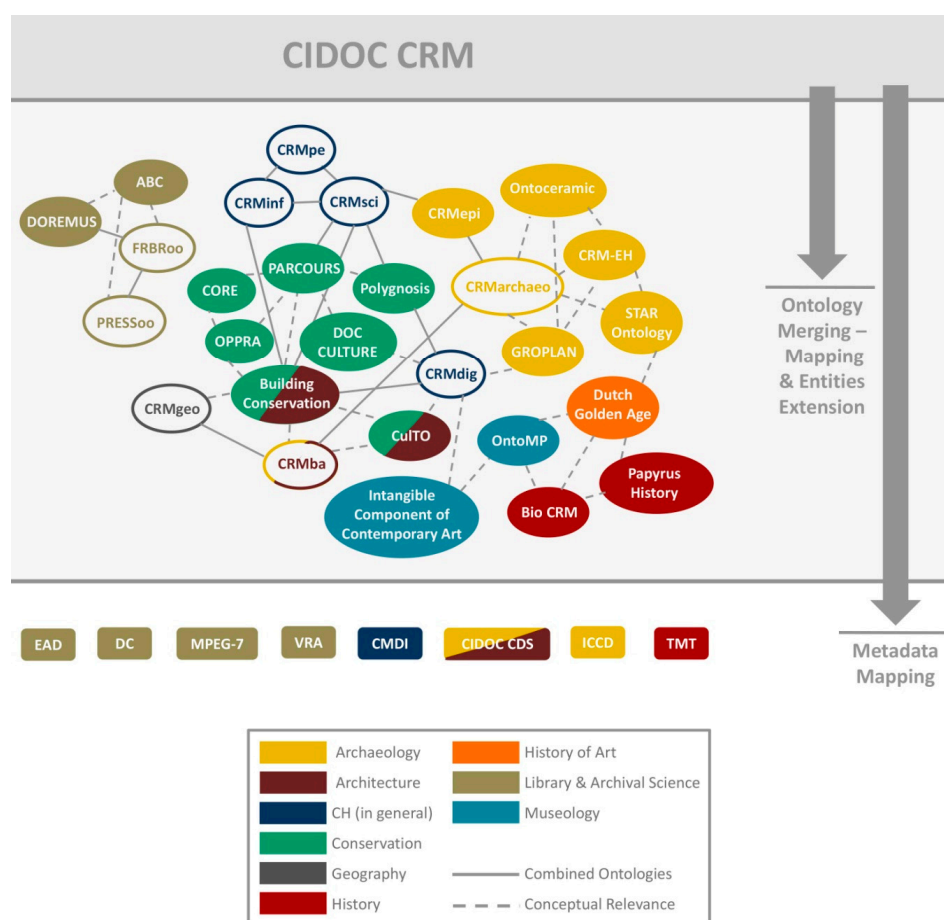

Figure 1. Organization of ontologies and metadata standards, in three abstraction levels: (1) The CIDOC CRM, as a top-level ontology, (2) associations among ontologies that have been mapped to or merged with the CIDOC CRM (ABC and Building Conservation Ontology) as well as the CIDOC CRM extensions, (3) metadata schemas mapped to CIDOC CRM core entities.

Table 2. Aggregation table about merging, mapping, and extending approaches of the CIDOC CRM.

Description	Type	Discipline	Task	Encoding	Year
ABC and CIDOC/CRM ontologies merging	Merging	Library and Archival Science	Information provenance; Multimedia	RDFS	2006
Mapping from the Encoded Archival Description (EAD) and Dublin Core (DC) metadata to CIDOC/CRM	Mapping	Library and Archival Science	-	(encoding neutral)	2007
FRBRoo is a conceptual model for bibliographic information	Extending	Library and Archival Science	-	RDF	2008
CRM-EH covers the archaeological excavation and analysis workflow	Extending	Archaeology	Scientific observation, analysis and diagnosis	RDF, SKOS	2008
CRMgeo provides an "articulation" between the standards of the geospatial and the cultural heritage community	Extending	Geography	-	RDFS	2009
STAR Project Ontology implements archaeological time periods using CIDOC CRM and SKOS	Extending	Archaeology	Dating	RDF	2010
Papyrus History ontology captures the history perspective on events and topics covered by the news	Extending	History	Information provenance, Information retrieval	RDF/OWL	2010
CRMdig encodes metadata about the steps and methods of production ("provenance") of digitization products and synthetic digital representations	Extending	CH (in general)	Scientific observation, analysis and diagnosis; Digitization	RDFS	2011
Mapping the Traditional Malay Textile Knowledge Model (TMT) for textile description to CIDOC CRM	Mapping/Merging	History	-	RDFS	2011
Combining the CIDOC CRM and MPEG-7 to describe multimedia in museums	Mapping	Library and Archival Science	Multimedia	(encoding neutral)	2011
Mapping the VRA Core, which provides guidance to describe artwork, to CIDOC CRM	Mapping	Library and Archival Science	Visualization	Path-oriented MDL	2011
Mapping the Italian archaeological documentation system (ICCD) to CIDOC CRM	Mapping	Archaeology	Scientific observation, analysis and diagnosis	(encoding neutral)	2013
PRESoo represents the underlying semantics of bibliographic information about periodicals	Extending	Library and Archival Science	-	RDF	2014
Arches project—Mapping CIDOC—CDS to CIDOC CRM	Mapping	Architecture; Archaeology	Scientific observation, analysis and diagnosis; Multimedia	(encoding neutral)	2014
Ontology of Paintings and Preservation of Art (OPPRA) covers the needs of paintings conservation and material analysis	Extending	Preservation Science	Conservation treatment; Scientific observation, analysis and diagnosis	OWL DL, OWL 2 RL	2014
Invisibilia project—Intangible Component of Contemporary Art Ontology	Extending	Museology	Visualization	RDF/XML	2014
CRMsci is a global schema about scientific observation, measurements and processed data	Extending	CH (in general)	Scientific observation, analysis and diagnosis	RDFS	2015
CRMinf is a global schema about argumentation and inference making	Extending	CH (in general)	Scientific reasoning	RDF	2015
CRMarchaeo supports the archaeological excavation process and the various entities and activities related to it	Extending	Archaeology	Scientific observation, analysis and diagnosis; Dating	RDFS	2015

Table 2. *Cont.*

Description	Type	Discipline	Task	Encoding	Year
GROPLAN provides the management of the coherence of measurement results about 3D models that have been created based on photogrammetric processes	Extending	Archaeology	Scientific observation, analysis and diagnosis; Multimedia; Digitization	OWL 2	2015
DOC-Culture Ontology defines the conservation interventions and analysis methods performed on artefacts	Extending	Preservation Science	Conservation treatment; Scientific observation, analysis and diagnosis	(encoding neutral)	2015
Ontoceramic is an OWL ontology for ceramics cataloguing and classification	Extending	Archaeology	Information provenance; Scientific observation, analysis and diagnosis	OWL 2	2015
CRMepi encodes epigraphic concepts and models the scientific process of investigation related to the study of epigraphy	Extending	Archaeology	Scientific observation, analysis and diagnosis	XML	2015
CRMba encodes metadata about the documentation of archaeological buildings	Extending	Architecture; Archaeology	Scientific observation, analysis and diagnosis; Dating	RDF	2016
Ontology for Dutch Golden Age artwork importing Getty ULAN controlled vocabulary	Extending	History of Art	Scientific reasoning	RDF/OWL	2017
Bio CRM is an extension of CIDOC CRM into the sub-domain of biography	Extending	History	-	RDF/OWL	2017
Polygnosis is a thesaurus about conservation and laser technologies that uses CIDOC CRM as a backbone	Extending	Preservation Science	Conservation treatment; Scientific observation, analysis and diagnosis	XML/XSD	2017
PARCOURS conservation/restoration ontology focuses on conservation-restoration processes	Extending	Preservation Science	Conservation treatment; Scientific observation, analysis and diagnosis	OWL 2 QL	2017
Mapping CMDI metadata to Parthenos Entities model (CRMpe)	Mapping	CH (in general)	Language Resources	X3ML	2017
CRMpe (PARTHENOS Entities) is a super—model to support research into the human past	Extending	CH (in general)	Information provenance	RDFS	2017
CulTO ontology characterizes religious historical building and supports its modeling	Extending	Architecture; Preservation Science	Digitization; Multimedia; Scientific observation, analysis and diagnosis	OWL 2	2017
Building Conservation ontology facilitates the reality-based 3D semantic annotations of the building conservation state	Mapping	Preservation Science; Architecture	Conservation treatment; Scientific observation, analysis and diagnosis, Digitization; Multimedia	Protégé-OWL, SQL/SPARQL	2018
CORE focuses on materials and techniques, condition state and conservation processes of artwork, and particularly byzantine icons	Extending	Preservation Science	Conservation treatment; Scientific observation, analysis and diagnosis	OWL 2	2018
OntoMP uses CIDOC CRM, FOAF and DBPedia to create virtual rooms and enable visitors to lookup individual life stories and also inter-cross information among them	Extending	Museology	Visualization	RDF	2018
DOREMUS adds new elements to FRBRoo and CIDOC CRM whenever needed to precisely express any music-related concept or relationship	Extending	Library and Archival Science	Multimedia; Information provenance	RDF, LYAM++	2018

Our study indicates a significant thematic connection between the various approaches, since they share common field and/or research goals. As illustrated in Figure 1 and presented in Table 2, preservation and conservation science has been the center of interest for a considerable number of projects over the last four years. The specific requirements of artwork conservation treatment, production materials and techniques, deterioration, analysis, and diagnosis can be covered by several extensions, namely PARCOURS, CORE, OPPRA, Polygnosis and DOC-CULTURE, which focus on slightly different aspects, exploiting other ontologies and thesauri. Moreover, a conceptual relevance can be witnessed between the aforementioned extensions and the Building Conservation and CulTO ontologies. Since the two latter ontologies focus on historical and religious buildings and their deterioration, deformation, and state conservation, they outline a more specialized field that combines the sub-domains of conservation and architecture. In the same context, it is important to mention that the ontologies of the conservation domain can also be combined with compatible models such as CRMsci, CRMinf and CRMdig. The development of PARCOURS and Polygnosis, as well as the mapping of the Building Conservation ontology to the CIDOC CRM and its compatible models, justifies the aforementioned observation.

Regarding further development related to the sub-domain of conservation and the tasks of analysis and diagnosis, ontologies of other domains could be included in order to capture the required information. A good example is the OPPRA conservation ontology, which includes entities from the OreChem ontology that addresses the chemistry domain. While CRMsci has been proven efficient for representing scientific observations, measurements, and processed data in the conservation sub-domain, it is very probable that an officially compatible model of the CIDOC CRM, particularly for the chemistry domain, could better define and clarify some basic concepts and relations, facilitating their further specification using only vocabularies and thesauri. Likewise, since preventive conservation activities often exploit data produced during automatic environmental monitoring through sensors or sensor networks installed on indoor (museums) or outdoor (cultural spaces, monuments) sites, it would be interesting to combine ontologies related to conservation with sensor data ontologies [54].

One more sub-domain that has proven significantly fruitful due to the number of the different ontologies developed, is the archaeology domain. Firstly, there is CRMarchaeo, an officially compatible model which efficiently covers the domain. Additionally, CRMarchaeo is extended by CRMba, which presents an explicit correlation between the sub-domains of archaeology and architecture. Alongside the extensions of the CIDOC CRM community, there have been a number of efforts to extend the CIDOC CRM by focusing on specialized fields of archaeology. These efforts have produced extensions such as CRMepi for epigraphy, GROPLAN and Ontoceramic for archaeological artifacts and ceramics, CRM EH for tasks of excavation and analysis, and STAR for descriptions of periods and chronologies. These extensions can also be correlated with other CIDOC CRM extensions (e.g., CRMsci, CRMdig) or be further extended with concepts related to scientific measurement and observation, resulting—among other benefits—in a better representation of sub-domains, such as archaeometry and photogrammetry (e.g., the GROPLAN project).

Going back to the architecture sub-domain, a compatible model (CRMba) that mainly focuses on archaeological buildings has been developed. Likewise, a mapping that also covers the disciplines of archaeology and architecture has been proposed between the CIDOC CRM and CIDOC CDS standards, interrelating the two sub-domains. Since the CRMba, Building Conservation and CulTO ontologies all address a common discipline, we suggest a potential conceptual relevance between the first and latter two. Besides the aforementioned interrelation, there is no other explicit combination or further extension of CRMba. As such, architecture may prove to be an interesting sub-domain for further work, in order to represent its underlying knowledge in a more complete way.

Regarding the sub-domain of history of art, Figure 1 depicts only one relevant extension (the Dutch Golden Age), which focuses on the influence between the artists and the overall correlation between creators and artwork. This extension could be correlated with other extensions related to

the concept of time, such as STAR (archaeology discipline), as well as with extensions related to the representation of historic events, people, and their relationships, such as Bio CRM and Papyrus History (history discipline). Furthermore, the OntoMP ontology, which is related to the museology domain, exhibits a conceptual relevance with the domains of both art history and history, since it focuses on people's interviews and narratives.

In the context of the sub-domain of history of art, we consider of particular interest the attempt to correlate various techniques of different artists, by combining the discipline of history of art with CH tasks related to artwork analysis. Regarding the domain of history of art, it would be beneficial to represent concepts related to the iconographic analysis of artwork, focusing on the visual representation rather than the material aspect of the artwork.

According to our study, the library and archival science sub-domain was one of the first fields of interest for the CIDOC CRM community, as indicated by the merging of ABC and the CIDOC CRM. Additionally, this particular sub-domain presents a significant number of mapping efforts between the CIDOC CRM and relevant metadata standards of the CH domain, including EAD, DC, VRA, and MPEG-7, while two compatible models for bibliographic information and periodicals efficiently cover the sub-domain. It is important to mention that the aforementioned efforts were presented within the first eight years following the establishment of the CIDOC CRM as an official standard (from 2006 to 2014). The most recent work (2018), which presents an explicit correlation between the sub-domains of library and archival science and music, is the DOREMUS ontology. Since much work has already been done on the library and archival sub-domain, it would be interesting to correlate music with other sub-domains, such as the history sub-domain. For example, the Bio CRM approach indicates a combination of history with library and archival science.

At this point, it is important to mention that CH sub-domains, such as geography, and tasks, such as scientific reasoning, measurements, scientific observations, and information provenance are largely covered by the official CIDOC CRM family models. This fact is rather beneficial for the wider CH domain, since these models must be common among the different disciplines and tasks in order to facilitate data integration and complement the documentation process. Nevertheless, since these models were published and approved by the community gradually, they were not always included in different CIDOC CRM extension, merging, and mapping projects through the years, and their use has recently tended to increase (e.g., the recent mapping of the CMDI to the CRMpe family model).

Eventually, as depicted in Figure 1, CRMdig was also used as a basic model to further describe multimedia that enriches the documentation of cultural artifacts. As mentioned in the previous section, there are a number of ontologies that explicitly or implicitly combine the cultural visualization field with various sub-domains, such as archaeology (GROPLAN), architecture (Building Conservation, CulTO), conservation (Polygnosis, Building Conservation, DOC-CULTURE, CulTO), museology (Intagible Component of Contemporary Art), and respective tasks, such as scientific documentation, 3D modeling of cultural artifacts and buildings, and virtual exhibitions. However, many of these ontologies were developed to accommodate specific applications, whereas more generic models that could have a wider use are still to be defined. A typical example is the lack of ontologies defining basic concepts and relations in mobile guides and augmented reality museum applications, extending the CIDOC CRM entities accordingly.

5. Conclusions

In this work, a review of CIDOC CRM mapping, merging, and extending efforts has been conducted, including older, as well as more recent, publications. Based on the CIDOC CRM, ample extensions have been developed to facilitate the various disciplines and tasks of the CH domain by defining new, specialized entities and relations. Additionally, the most popular metadata schemas and standards of the CH domain have been matched to CIDOC CRM entities. Overall, all these efforts validate the role of the CIDOC CRM as a top-level ontology for the CH community, enrich its expressivity, and ensure data interoperability.

In conclusion, knowledge of existing extensions, as well as deployments of CIDOC CRM for mapping/merging ontologies, is necessary for efficiently modeling information related to the various CH sub-disciplines and tasks. In this way, suggestions on using appropriate entities and relations (e.g., [55] makes suggestions for modeling arguments and relations for virtual reconstruction), which express more efficiently knowledge/information for different use-cases, are considered valuable. It is common knowledge that more specialized semantic representations of CH domain information will further facilitate its reuse and ensure its provenance, capturing a whole universe of perpetually produced information.

As we move from the public space of the museum to the private space of a home or a teaching room, and then back to a global space through the Internet, the reuse of the CIDOC CRM and its derivatives will pave the way for developing novel, semantically-enabled digital services to accommodate the different dimensions of cultural experience. It will also, for the first time, give CH institutes the opportunity to share knowledge, approach CH in an interdisciplinary fashion, and establish new fields of study and work. The CIDOC CRM can give cultural organizations the chance to enrich their range of research and activity, by opening new modes of communication and familiarization with the best international practices.

In addition to its contribution to the primary objectives of a cultural organization, i.e., preserving and presenting CH, the CIDOC CRM can also help to modernize the profile of a CH organization and redefine its practices concerning:

- Education and entertainment (edutainment)
- Approach and Information of the Community (national and international) and Development of new Communities
- Information Service
- Enhanced public access over the Internet
- Control of Content Management
- and Measurement of Results.

To put it in the words of André Malraux, the CIDOC CRM can help in the creation of "an imaginary museum that is to push to the end the imperfect mapping imposed by the true museums" [56].

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Häyrinen, A. Towards semantic modelling of cultural historical data. In Proceedings of the 20th European-Japanese Conference on Information Modelling and Knowledge Bases EJC2010, Jyväskylä, Finland, 31 May–4 June 2010; pp. 322–331.
2. Doerr, M.; Hunter, J.; Lagoze, C. Towards a Core Ontology for Information Integration. *J. Dig. Inf.* **2003**, *4*, 1.
3. International Council of Museums—Committee of Conservation (ICOM-CC). Terminology to Characterize the Conservation of Tangible Cultural Heritage. Available online: http://www.icom-cc.org/242/about/terminology-for-conservation/#Wim1LHlx3IU (accessed on 11 February 2019).
4. Renfrew, C.; Paul, B. Archaeology: the key concepts. In *Archaeology: The Key Concepts*; Taylor & Francis Group: London, UK, 2013; pp. 15–23.
5. Vassilakaki, E.; Zervos, S.; Giannakopoulos, G. CIDOC-CRM extensions for conservation processes: A methodological approach. *AIP Conf. Proc.* **2015**, *1644*, 185–192.
6. Hunter, J. Combining the CIDOC CRM and MPEG-7 to describe multimedia in museums. In Proceedings of the Museums and the Web international conference, Boston, MA, USA, 18–20 April 2002.

7. Alexiev, V.; Manov, D.; Parvanova, J.; Petrov, S. *Large-scale Reasoning with a Complex Cultural Heritage Ontology (CIDOC CRM), a Mapping of CIDOC CRM Events to German Wordnet for Event Detection in Texts*; CIDOC: Shanghai, China, 2013.

8. Hyvönen, E. Cultural heritage linked data on the semantic web: Three case studies using the sampo model. In Proceedings of the VIII Encounter of Documentation Centres of Contemporary Art: Open Linked Data and Integral Management of Information in Cultural Centres, Artium, Vitoria-Gasteiz, Spain, 19–20 October 2016.

9. Binding, C. *Implementing Archaeological Time Periods Using CIDOC CRM and SKOS*; Springer: Berlin/Heidelberg, Germany, 2010.

10. Aliaga, D.G.; Bertino, E.; Valtolina, S. DECHO-A framework for the digital exploration of cultural heritage object. *J. Comput. Cult. Herit.* **2011**, *3*, 12. [CrossRef]

11. Carlisle, P.K.; Avramides, I.; Dalgity, A.; Myers, D. The arches heritage inventory and management system: A standards-based approach to the management of cultural heritage information. In Proceedings of the Access and Understanding-Networking in the Digital Era, Dresden, Germany, 6–11 September 2014.

12. Hohmann, G.; Schiemann, B. An ontology-based communication system for cultural heritage: Approach and progress of the WissKI project. In *Scientific Computing and Cultural Heritage*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 127–135.

13. Hunter, J.; Odat, S. Building a Semantic Knowledge-Base for Painting Conservators. In Proceedings of the IEEE Seventh International Conference on e-Science Workshops (eScience2011), Stockholm, Sweden, 5–8 December 2011.

14. Bruseker, G.; Doerr, M.; Theodoridou, M. *Report on the common semantic framework (D5.1), PARTHENOS—Pooling activities, resources and tools for heritage e-searching networking, optimization and synergies*, PUBLIC with CC-BY distribution. 2017.

15. Felicetti, A.; Scarselli, T.; Mancinelli, M.; Niccolucci, F. Mapping ICCD archaeological data to CIDOC-CRM: the RA schema. In Proceedings of the 17th International Conference on Theory and Practice of Digital Libraries (TPDL 2013), Valetta, Malta, 26 September 2013.

16. CORDIS | European Commission. Available online: https://cordis.europa.eu/project/rcn/89256/factsheet/en (accessed on 26 December 2018).

17. Dcc.ac.uk. CASPAR | Digital Curation Centre. Available online: http://www.dcc.ac.uk/resources/briefing-papers/technology-watch-papers/caspar (accessed on 26 December 2018).

18. Niang, C.; Marinica, C.; Markhoff, B.; Leboucher, E.; Malavergne, O.; Bouiller, L.; Darrieumerlou, C.; Laissus, F. Supporting Semantic Interoperability in Conservation-Restoration Domain: The PARCOURS Project. *J. Comput. Cult. Herit.* **2017**, *10*, 1–20. [CrossRef]

19. Drap, P.; Merad, D.; Hijazi, B.; Gaoua, L.; Nawaf, M.M.; Saccone, M.; Castro, F. Underwater photogrammetry and object modeling: a case study of Xlendi Wreck in Malta. *Sensors* **2015**, *15*, 30351–30384. [CrossRef] [PubMed]

20. Kiyavitskaya, N.; Katifori, A.; Velegrakis, Y.; Tsinaraki, C.; Bykau, S.; Savaidou, E.; Koubarakis, M. *Modeling and Mapping Multilingual and Historically Diverse Content*; CIDOC: Shanghai, China, 2010.

21. Lisena, P.; Achichi, M.; Choffé, P.; Cecconi, C.; Todorov, K.; Jacquemin, B.; Troncy, R. Improving (Re-) Usability of Musical Datasets: An Overview of the DOREMUS Project. *Bibl. Forsch. Prax.* **2018**, *42*, 194–205. [CrossRef]

22. Damiano, R.; Lieto, A.; Lombardo, V. Ontology-based visualisation of cultural heritage. In Proceedings of the 2014 Eighth International Conference on Complex, Intelligent and Software Intensive Systems (CISIS), Birmingham, UK, 2–4 July 2014; pp. 558–563.

23. Hyvönen, E.; Tuominen, J.; Alonen, M.; Mäkelä, E. Linked Data Finland: A 7-star model and platform for publishing and re-using linked datasets. In *European Semantic Web Conference*; Springer: Cham, Germany, 2014; pp. 226–230.

24. Martini, R.G.; Araújo, C.; Librelotto, G.R.; Henriques, P.R. A reduced CRM-compatible form ontology for the virtual Emigration Museum. In *New Advances in Information Systems and Technologies*; Springer: Cham, Germany, 2016; pp. 401–410.

25. Cultura Italia: The Italian National Cultural Aggregator. Available online: http://www.culturaitalia.it/opencms/documentazione_tecnica_en.jsp?language=en&tematica=static (accessed on 21 January 2019).

26. Kalfoglou, Y.; Schorlemmer, M. Ontology mapping: The state of the art. *Knowl. Eng. Rev.* **2003**, *18*, 1–31. [CrossRef]

27. Doerr, M.; Kritsotaki, A.; Rousakis, Y.; Hiebel, G.; Theodoridou, M. *Definition of the CRMsci an Extension of CIDOC-CRM to Support Scientific Observation*; Technical Report 1.2.3; ICS-FORTH: Heraklion, Greece, 2015.

28. Stead, S.; Doerr, M. CRMinf: the Argumentation Model. An Extension of CIDOC-CRM to Support Argumentation. Available online: http://www.ics.forth.gr/isl/CRMext/CRMinf/docs/CRMinf-0.7.pdf (accessed on 26 December 2018).

29. Bruseker, G.; Carboni, N.; Guillem, A. Cultural Heritage Data Management: The Role of Formal Ontology and CIDOC CRM. In *Heritage and Archaeology in the DigitalAge*; Springer: Cham, Germany, 2017; pp. 93–131.

30. Doerr, M.; Theodoridou, M. CRMdig: A Generic Digital Provenance Model for Scientific Observation. In Proceedings of the 3rd USENIX Workshop on the Theory and Practice of Provenance, Heraklion, Greece, 20–21 June 2011.

31. Ronzino, P.; Niccolucci, F.; Felicetti, A.; Doerr, M. CRMba a CRM extension for the documentation of standing buildings. *Int. J. Digit. Libr.* **2016**, *17*, 71–78. [CrossRef]

32. Hiebel, G.; Doerr, M.; Eide, O. CRMgeo: A spatiotemporal extension of CIDOC-CRM. *Int. J. Digit. Libr.* **2017**, *18*, 271. [CrossRef]

33. Doerr, M.; Bekiari, C.; LeBoeuf, P.; nationale de France, B. FRBRoo, a conceptual model for performing arts. In Proceedings of the 2008 Annual Conference of CIDOC, Athens, Greece, 15–18 September 2008.

34. CIDOC CRM Family Models and Collaborations. PRESSoo—Modeling of Bibliographic Information Pertaining to Continuing Resources. Available online: http://www.cidoc-crm.org/pressoo/ (accessed on 26 December 2018).

35. Stasinopoulou, T.; Bountouri, L.; Kakali, C.; Lourdi, I.; Papatheodorou, C.; Doerr, M.; Gergatsoulis, M. Ontology-based metadata integration in the cultural heritage domain. In *International Conference on Asian Digital Libraries*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 165–175.

36. Kakali, C.; Lourdi, I.; Stasinopoulou, T.; Bountouri, L.; Papatheodorou, C.; Doerr, M.; Gergatsoulis, M. Integrating Dublin Core metadata for cultural heritage collections using ontologies. In Proceedings of the International Conference on Dublin Core and Metadata Applications, Singapore, 27–31 August 2007.

37. Nasir, S.A.M.; Noor, N.L.M. Automating the mapping process of traditional Malay textile knowledge model with the core ontology. *Am. J. Econ. Bus. Admin.* **2011**, *3*, 191–196. [CrossRef]

38. Angelopoulou, A.; Tsinaraki, C.; Christodoulakis, S. Mapping MPEG-7 to CIDOC/CRM. In Proceedings of the International Conference on Theory and Practice of Digital Libraries, Berlin, Germany, 25–29 September 2011.

39. Gaitanou, P.; Gergatsoulis, M. Mapping VRA Core 4.0 to the CIDOC CRM Ontology. Available online: http://eprints.rclis.org/15843/ (accessed on 26 December 2018).

40. Messaoudi, T.; Véron, P.; Halin, G.; Luca, L.D. An ontological model for the reality-based 3D annotation of heritage building conservation state. *J. Cult. Herit.* **2018**, *29*, 100–112. [CrossRef]

41. Broeder, D.; Schonefeld, O.; Trippel, T.; Van Uytvanck, D.; Witt, A. A pragmatic approach to XML interoperability—the Component Metadata Infrastructure (CMDI). In Proceedings of the Balisage: The Markup Conference, Balisage Series on Markup Technologies, Montréal, QC, Canada, 2–5 August 2011.

42. Durco, M.; Lorenzini, M.; Sugimoto, G. Something will be connected-Semantic mapping from CMDI to Parthenos Entities. In Proceedings of the CLARIN Annual Conference 2017, Budapest, Hungary, 18–20 September 2017.

43. Binding, C.; May, K.; Tudhope, D. Semantic interoperability in archaeological datasets: Data mapping and extraction via the CIDOC CRM. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2008.

44. Cantone, D.; Asmundo, M.N.; Santamaria, D.F.; Trapani, F. *Ontoceramic: An OWL Ontology for Ceramics Classification*; CILC: Genova, Italy, 2015; pp. 122–127.

45. Felicetti, A.; Murano, F.; Ronzino, P.; Niccolucci, F. *CIDOC CRM and Epigraphy: A Hermeneutic Challenge*; EMF-CRM@ TPDL: Poznan, Poland, 2015; pp. 55–68.

46. Tran, T.A.; Isemann, D. Modeling creative influence between artists and their works with OWL and CIDOC CRM. In *INFORMATIK*; Gesellschaft für Informatik: Bonn, Germany, 2017.

47. Hohmann, G.; Schiemann, B. *Scientific Computing and Cultural Heritage*; Springer: Berlin, Germany, 2013.

48. Tuominen, J.; Hyvönen, E.; Leskinen, P. *Bio CRM: A Data Model for Representing Biographical Data for Prosopographical Research*; Biographical Data in a Digital World (BD2017); CEUR: Linz, Austria, 2017.

49. Odat, S.A. A Semantic e-Science Platform for 20th Century Paint Conservation. Ph.D. Thesis, University of Queensland, Brisbane, Australia, 2014. Available online: http://espace.library.uq.edu.au/view/UQ:341015 (accessed on 26 December 2018).

50. Platia, N.; Chatzidakis, M.; Doerr, C.; Charami, L.; Bekiari, C.; Melessanaki, K.; Pouli, P. "POLYGNOSIS": The development of a thesaurus in an Educational Web Platform on optical and laser-based investigation methods for cultural heritage analysis and diagnosis. *Herit. Sci.* **2017**, *5*, 50. [CrossRef]

51. Moraitou, E.; Kavakli, E. Knowledge Management Using Ontology on the Domain of Artworks Conservation. In *Digital Cultural Heritage*; Springer: Cham, Germany, 2018; pp. 50–62.

52. Araújo, C.; Martini, R.G.; Henriques, P.R.; Almeida, J.J. Annotated Documents and Expanded CIDOC-CRM Ontology in the Automatic Construction of a Virtual Museum. In *Developments and Advances in Intelligent Systems and Applications*; Rocha, A., Reis, L.P., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 91–110.

53. Garozzo, R.; Murabito, F.; Santagati, C.; Pino, C.; Spampinato, C. Culto: An Ontology-Based Annotation Tool for Data Curation in Cultural Heritage. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* **2017**, *42*, W5. [CrossRef]

54. Peralta, L.R.; de Brito, L.P.L.; Gouveia, B.T.; de Sousa, D.G.; Alves, C.D.S. Automatic monitoring and control of museums environment based on wireless sensor networks. *Electron. J. Struct. Eng.* **2010**, *EJSE Special Issue: Wireless Sensor Networks and Practical Applications*, 12–34.

55. Bruseker, G.; Guillem, A.; Carboni, N. Semantically Documenting Virtual Reconstruction: Building a Path to Knowledge Provenance. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* **2015**, *II-5/W3*, 33–40. [CrossRef]

56. Malraux, A. *Le Musée Imaginaire*, 1st ed.; Gallimard: Paris, France, 1965.