# Introduction

The CIDOC CRM ontology is an international standard currently widely accepted and adopted by different research communities and digital infrastructures to manage heterogeneous documentation (ARIADNE, PARTHENOS, 3D COFORM, 3D ICONS, iMARINE, to cite a few)[1]. It fosters interoperability among different data structures by providing the semantic definitions needed to transform different and confined information sources into a coherent and global resource, and offering a flexible system that does not impose the use of a unique standard. The CRM is used to describe the documentation process and to express the implicit and explicit concepts and relationships typically assumed in the cultural heritage documentation. By providing a common and extensible semantic framework, to which any cultural heritage information can be mapped, it prevents semantic information loss, a phenomenon that usually occurs when integrating heterogeneous resources.

Although the CIDOC CRM ontology proposes high-level concepts, which was, together with the abstractness of its concepts, one of the criticisms addressed to the ontology until recent[2], it offers the possibility to create extensions at any degree of detail, necessary to capture the full richness of the cultural heritage datasets. On the other hand, the core CRM is the common framework on which domain-specific specializations rely enabling non-domain interoperability.

Recently various extensions have been released, focusing on geographical concepts, digital provenance preservation, scientific applications and reasoning, archaeology and built structures documentation. Moreover, the CIDOC CRM ontology has proved to be fundamental to Natural Language Processing (NLP) of heritage datasets and grey literature, gazetteers and thesauri, Linked Open Data and other methods for the use and re-use of datasets and collections of digital humanities, historical and archaeological resources.
Based on these considerations, the aim of the workshop was to start a constructive debate with participants and to collect insights, issues and suggestions. This activity provided a deeper understanding on users' requirements, which were taken into account and reported to the Special Interest Group of CIDOC CRM and the ARIADNE Metadata and Standards SIG, in which the workshop orga-

---

[1] ARIADNE: `http://ariadne-infrastructure.eu`
PARTHENOS: `http://www.parthenos-project.eu`
3D COFORM: `http://www.3d-coform.eu`
3D ICONS: `http://www.3dicons-project.eu`
iMARINE: `http://www.i-marine.eu`

[2] Nussbaumer, Philipp and Haslhofer, Bernhard, 2007. *CIDOC CRM in Action - Experiences and Challenges*. In Research and Advanced Technology for Digital Libraries: Proceedings of the 11th European Conference, ECDL 2007, Budapest, Hungary, September 16-21, 2007. Springer Berlin Heidelberg. URL `http://dx.doi.org/10.1007/978-3-540-74851-9_61`

nizers are fully involved, providing relevant inputs and contributing in making recommendations.

During the workshop participants shared their approaches describing and showcasing systems using the CRM, exchanging experience about its practical uses and describing difficulties in its application.

This proceeding includes six selected papers presented at the workshop "Extending, Mapping, and Focusing the CRM". Every paper received two reviews, provided by the program committee members. Most of the submitted works discussed about practical applications of the CIDOC CRM global ontology for the implementation of case studies directly connected with the ARIADNE project activities.

The reason stays in that the CIDOC CRM ontology has been chosen by the ARIADNE project as the global ontology to which the archaeological datasets and collections, made available by the partner institutions, were mapped. This activity allowed the identification of common concepts and relations, which were fundamental for the implementation of the archaeological extension of the ontology, the CRMarchaeo. These mapping activities were performed using the 3M tool developed by FORTH, which provides users with a powerful graphical interface that overcomes the complexity of the global model and allows using the CRM extensions. It also acts as guide to advice users in the mapping process. The X3ML data exchange framework tool is presented in the paper *"X3ML Framework: An effective suite for supporting data mappings"*, authored by Nikos Minadakis, Yannis Marketakis, Haridimos Kondylakis, Giorgos Flouris, Maria Theodoridou, Martin Doerr, and Gerald de Jong.

Furthermore, CIDOC CRM has been chosen as the backbone ontology for the integration of heterogeneous datasets at the level of single records.
A description of this activity is reported in the paper *"Integrating heterogeneous coin datasets in the context of archaeological research"* (by Achille Felicetti, Philipp Gerth, Carlo Meghini, and Maria Theodoridou), which demonstrates the item-level integration process of archaeological archives through the use of semantic technologies.
For the implementation of this case study, a sub set of ancient coin records, provided by several European archaeological institutions, was selected. The subset thus created, was analysed to identify similar concepts and common metadata elements to enable their integration. CIDOC CRM was chosen as the conceptual model for encoding the identified entities, while some important numismatic vocabularies have been employed to improve standardisation.

Another ARIADNE-related work, *"Integrating terminological tools and semantic archaeological information: the ICCD RA Schema and Thesaurus"* (by Achille Felicetti, Ilenia Galluccio, Cinzia Luddi, Maria Letizia Mancinelli, Tiziana Scarselli, and Antonio Davide Madonna), describes the process of mapping, translation and publication in SKOS format of the RA Thesaurus. The RA Thesaurus, developed by the Italian Ministry of Cultural Heritage (MiBACT),

provides a unified and meaningful terminology for the description of archaeological objects according to the MiBACT official cataloguing standards. A detailed description of the thesaurus, is provided within the paper, together with the technologies used for the publication of the thesaurus on the web.

The paper *"Dati.CulturaItalia: a use case of publishing Linked Open Data based on CIDOC-CRM"* (by Sara Di Giorgio, Achille Felicetti, Patrizia Martini and Emilia Masci) describes the pilot project dati.culturaitalia.it, aimed at building a Linked Open Data (LOD) Service that would make open datasets from the web-portal CulturaItalia, available. The CIDOC CRM ontology was used in this case study, to transform and represent cultural heritage data. The RDF triples mapped to the CRM Erlangen were enriched with links to URIs identifying instances of internationally established RDF resources for geographic names, and instances of authority files for personal and corporate names.

The work by Achille Felicetti, Francesca Murano, Paola Ronzino, and Franco Niccolucci, *"CIDOC CRM and Epigraphy: a hermeneutic challenge"*, proposes an extension of the CIDOC CRM to encode epigraphic concepts and to model the scientific process of investigation in this domain. After identifying the main concepts involved in the study of epigraphy, and analysing the existing CIDOC CRM entities, together with those provided by the CRMsci and CRMarchaeo extensions, the authors propose to introduce the CRMepi extension. With the new classes and properties developed ad hoc, CRMepi aims at contributing to the specific needs of epigraphic documentation.

A methodological contribution to temporal knowledge is provided by the paper *"Temporal Primitives, an Alternative to Allen Operators"* (by Manos Papadakis, and Martin Doerr). The paper discusses the limits of the Allen Interval Algebra set of operators, which fails in observation-driven fields like stratigraphy. In such cases, incomplete temporal information yields a disjunctive set of Allen operators, which affects RDF reasoning since it leads to expensive queries containing unions. To address this deficiency, the authors introduce a set of basic temporal primitives which are employed in an extension of CIDOC CRM. The flexible representation proposed by the authors can describe any Allen operator as well as scenarios with further temporal generalization using conjunctions of primitives. An extension to the basic set of primitives is also proposed, introducing fuzzy primitives that can model temporal topologies with imprecise boundaries that generalize over precise boundary models.

Paola Ronzino

PIN, Prato, Italy