

The structure of CIDOC CRM Model and Extensions

The structure of CIDOC CRM Model and Extensions	1
Background	2
Kinds of CIDOC CRM model	2
How to Propose a New CIDOC CRM Extension: CRM Compatible and Harmonized Extensions	4
How to Propose a new CIDOC CRM Compatible Model	4
Consistency	5
Open Source	5
External Responsibility	5
Full Documentation	6
Commitment to On-Going Integration Efforts	6
How to Propose a new CIDOC CRM Harmonized Extensions	7
Consistency	7
Open Source	7
Responsibility	8
Full Documentation	8

Background

The maintenance of all possible CRM compatible models by the CRM-SIG is not possible. CRM-SIG currently manages a set of harmonized model extensions, following strict intellectual and logical rules. On the other side, users expect clear rules of compatibility which teams with sufficient background can apply consistently.

Therefore this text proposes a management structure of three levels with respect to the certification of respective models. In more detail:

CIDOC CRM is what it is because of the principles of its development, two of the most foundational of which are that it favours the function of information integration with long-term validity and stresses the connection of this information with knowledge groups. It is because of this that it has developed slowly and steadily and inter alia become an ISO standard which people want to use because it is reliable. This rigour is what founds the CIDOC CRM's reputation. The CIDOC CRM base standard is therefore developed slowly and conservatively in order to ensure its reliability and long-term reusability. Concepts and relations are introduced into the standard by developers model only when they believe they have fully understood these and the implications of their inclusion against this high standard of long term reusability.

On the other hand, the needs of live data project are here and now, and they push the edge of what can be said, what has been fully thought through etc. They need to move fast(er) and CIDOC CRM needs to move, deliberately, slower.

For this reason we determine the need for different approaches and standards for CIDOC CRM extension maintenance depending on their relative function and origin.

Kinds of CIDOC CRM model

Specifically, we identify three kinds of model in the CIDOC CRM compatible information space:

- 1) The base model
- 2) Harmonized Extensions
- 3) Compatible Extensions

CRM base provides the foundation of the system, giving top level information and patterns, the settled understanding that comes from the general modelling experience. This is the core business of the CRM SIG, to maintain it, develop it and explain it. We change it sparingly, when we learn new things and it must be changed. We do so trying to make at all times monotonic(i.e. backwards compatible) changes

Under this and still in the direct purview of the CRM SIG are **CIDOC CRM harmonized extensions**. "Harmonization" denotes a process of ontology integration by modifying constituent concepts until they describe their combined domains consistently as one. Harmonized extensions are those which are made compatible with CIDOC CRM base but are furthermore made compatible amongst themselves and follow the general strict rules for ontology development set out for CRMbase. Thus while they address more specific domains of information than CRM base they are meant to be generally adoptable and stable, providing long term information integration guaranteed at the same level as CRMbase itself.

The third layer, consists of **CIDOC CRM compatible models**. A CIDOC CRM compatible extension is a designation that can only be given by CRM SIG. Compatibility consists in the model's technical adherence to CIDOC CRM, and its correct use of the logic of the model, correctly interpreting and reusing classes and properties it relates to. This does not mean then that the SIG endorses the model as such but that it acknowledges it as a coherent model that can be used as a compatible extension of CRM base and **harmonized extensions**.

How to Propose a New CIDOC CRM Extension: CRM Compatible and Harmonized Extensions

How to Propose a new CIDOC CRM Compatible Model

CIDOC CRM SIG welcomes the submission of not explicitly harmonized, CRM compatible extensions of the CIDOC CRM, in short “CRM compatible models”, for listing and reference on the CIDOC CRM site. This listing is intended to support a consolidated space wherein to share ideas and methods for the modeling of cultural heritage and historic data that goes beyond the present scope and aims of the CRM SIG, manifested in its present base ontology and harmonized extensions, which are guided by the goal of making standards with long-term validity.

This space will serve as a marketplace of ideas wherein high-quality extensions to CIDOC CRM, made for particular purposes extending beyond the present scope of CIDOC CRM, can be shared between users of the CIDOC CRM. These extensions can serve as potential immediate solutions to modelling problems of active users of CIDOC CRM. They can also serve as a place of exchange, whereby best practices and modelling ideas can be shared in a decentralized manner across the community. Such conversations can then, in principle, feed back to the CIDOC CRM SIG as part of a process of the continuous maintenance and improving of the standard according to new areas of research and approaches within cultural heritage and historical disciplines.

Non-harmonized extensions are expected to meet the quality controls laid out below but are not restricted by the full set of modelling methodology and practices demanded for CIDOC CRM base and harmonized extensions. In this way, CIDOC CRM compatible models can represent more fluid and dynamic modelling approaches and may serve research purposes, while the CIDOC CRM and its extensions will continue to play their role as a stable, monotonically curated set of ontological models that allow the general expression and integration of CH and historic data.

CIDOC CRM compatible models may be contributed for listing on the CIDOC CRM website under the following principles.

- 1) Ontological Consistency
- 2) Open Source and Non Profit
- 3) External Responsibility
- 4) Full Documentation
- 5) Commitment to on-going integration efforts

Consistency

The submitted model should be aligned to the CIDOC CRM through sub and super class and property statements, joining the ontological model proposed to the overall logic and structure of the CIDOC CRM base ontology and harmonized extensions, as defined in the section “Extensions of CIDOC CRM” of the Definition of the CIDOC CRM. In this regard, the sub and super class and property declarations must be ontologically compatible with the semantics of the relative CIDOC CRM classes and properties referenced. It should further respect the basic disjunctions in the ontology which are temporal entities vs persistent items, conceptual objects versus physical objects, place, time and spacetime.

This should be checked automatically using ontology reasoner tools and the basic consistency of top level classes and properties double checked by the CIDOC CRM SIG community. The formal semantics should be compatible with that of the CIDOC CRM as defined in: Meghini, C. and Doerr, M. (2018) ‘A first-order logic expression of the CIDOC conceptual reference model’, *International Journal of Metadata, Semantics and Ontologies*, 13(2), pp. 131–149. doi: [10.1504/IJMSO.2018.098393](https://doi.org/10.1504/IJMSO.2018.098393).

Open Source

CIDOC CRM SIG is an organization responsible for the development and maintenance of ontology standards for cultural heritage data and as such it committed to the open source sharing and reuse of its ontological models and their potential extensions. As such, any extension proposed for listing as a compatible extension on the official website by default should be an open source document reusable by the public towards the generation of semantic data.

Therefore, every proposed extension must be accompanied by a valid open source copyright claim, such as CC BY 4.0 or a practically equivalent license.

External Responsibility

The CIDOC CRM SIG claims no responsibility for the models submitted with regards to their maintenance or longevity. Submitters of extensions take full responsibility for the maintenance and hosting of their extensions and may offer user support. Potential users of these models should consult the documentation of the model to understand the commitment of the maintainer of that model, with regards to the long term reusability and compatibility of that model.

This externality of responsibility will be noted on the CIDOC CRM site where Compatible Models are listed. Moreover, submitters of models are expected to have a description of their maintenance policy on the site where they host their ontology.

Full Documentation

The functionality of CIDOC CRM as a formal ontology for cultural heritage relies on the CRM SIG's development of a rigorous documentation methodology for creating and maintaining a well defined and transparent ontology that is first and foremost a specification document which is also elaborated as various serializations. Submitted extensions should follow the CIDOC CRM style specification document for the definition of the extensions scope, its classes and properties. The specification should be populated with examples for classes and properties. Ideally, extension will provide links to publicly available data to which their models apply or have been applied.

Commitment to On-Going Integration Efforts

Submitters of CRM Compatible Models are assumed to wish to have reference to their work as part of an on-going community process to develop integrated, harmonized standards for semantic data construction. As such, it is assumed that submitters of models will be members of the SIG and participate as possible in the life of the community but particularly in discussions regarding the development of harmonized extensions of CIDOC CRM that may provide integrated, long term CIDOC CRM approved extensions for constructs that may fall within the scope of contributed extensions.

The CRM-SIG is interested in socially widening the competence about the CRM, its application and extending it. Therefore, per default, the CRM SIG will trust and review only the declaration of the submitters about the intellectual and formal principles applied, as requested for the submission, and rely on the submitters to maintain sufficient coordination with the CRM SIG. If any user or CRM SIG member has questions to which degree the submitted model complies with the requested terms, suitable action should be discussed as an issue at the CRM SIG.

How to Propose a new CIDOC CRM Harmonized Extensions

When it has identified an area of sufficient agreement and expertise, the CRM SIG may choose as a group to begin the development of a harmonized extension. A harmonized extension is an ontological model which is consistent with CRMbase, the other harmonized models and which is maintained the CRM SIG under its own resources. The CRM SIG has maintained a list of harmonized models for over a decade and will continue to do so. These models are backed by the CRM SIG and their development falls within its remit.

Adding a new harmonized extension must be agreed by the whole SIG and sufficiently support by the appropriate expertise and data in order to work on the issue. Moreover, sufficient resources in terms of time and expertise must be ascertained to support the harmonized extension at the SIG level over the long run.

CIDOC CRM harmonized models may be added to the work of the CIDOC CRM SIG under the following principles.

1. Ontological Consistency
2. Open Source and Non Profit
3. SIG Responsibility
4. Full Documentation
5. Commitment to on-going integration efforts

Consistency

CRMbase and **CIDOC CRM harmonized extensions** are methodologically based on the **ontological commitment** to empirical reality, following the section “**Reality, Knowledge Bases and CIDOC CRM**”, and on unambiguous, diachronic, distinct identity of instances it aims at describing, verified by the cross-disciplinary dialogue which CRM-SIG maintains with domain experts across applications, schools of thought and scientific disciplines, validating real data, as warrant for the long-lasting ability to support cross-disciplinary information integration.

“Harmonization” denotes a process of ontology integration by modifying constituent concepts until they describe their combined domains consistently as one. A CIDOC CRM harmonized extension has to be harmonized with a particular version of CRM base, preferably the latest official one, and harmonized with all other extensions it has a conceptual overlap with, such that CRM base and its harmonized extensions can be used as a stable ontological framework guaranteed to be consistent with the rest of the CIDOC CRM world.

Open Source

CIDOC CRM SIG is an organization responsible for the development and maintenance of ontology standards for cultural heritage data and as such it committed to the open source

sharing and reuse of its ontological models and their potential extensions. As such, any extension proposed for listing as a compatible extension on the official website by default should be an open source document reusable by the public towards the generation of semantic data.

Therefore, every proposed extension must be accompanied by a valid open source copyright claim, such as CC BY 4.0 or a practically equivalent license.

Responsibility

CIDOC CRM harmonized extensions are named with the prefix “CRMxxx”. CIDOC CRM harmonized extensions are systematically maintained by the SIG to the degree its resources allow for. All requests for modifications constitute CRM SIG issues.

A CIDOC CRM harmonized extension should be declared 'official' by the SIG when it is mature, in the sense of verification of its ontological commitment in a cross-disciplinary dialogue, and is harmonized with an explicit version of CRMbase, plus the necessary extensions it has a conceptual overlap with.

CRM-SIG maintains their compatibility with the **latest versions** of CRMbase, and **implementations** in important KR language encodings, in cooperation with other teams.

Given the limited resources of CRM-SIG and the ambition to provide a set of models that is robust in the long term and highly backwards compatible after revisions, the following principles guide the development process:

“Do not model what you have not fully understood”. Robustness and ontological commitment overrule the wish for completeness of thematic coverage. The latter can be achieved by CRM compatible models as described above. This imposes a slow rhythm of development, and bottom-up development to a kind of ontological middle layer avoiding over-generalizations with blurred semantics.

“Don't open more of these harmonized models unless there is sufficient experience available, interest and support”.

Use the “emerging semantics” of critically reviewed, multiple CRM compatible models covering the same subject area for backing up the evolution of harmonized extensions.

Full Documentation

Documentation of CIDOC CRM harmonized extensions relies on the same rigorous documentation methodology applied by the SIG for creating and maintaining CRMbase and elaborating as various serializations.