## Issue 646

**Introduction**

This document presents CRMinf, an extension of the CIDOC Conceptual Reference Model (CRM, ISO21127) created to support the documentation of scholarly and scientific arguments for documented propositions about the past. The making of documented propositions and their arguments are seen as historical facts regardless their relevance. The purpose of documenting the argumentation is safeguarding and understanding the provenance of knowledge, for future assessments of authenticity and for providing sufficient information for reassessing the validity of an argument and its conclusions based on given or new evidence of whatever kind. CRMinf *does not* aim at promoting the application of formal logical reasoning about historical facts or replacing scholarly arguments by automation. Whereas the results of formal logical reasoning can be documented in CRMinf, it commits rather to an epistemology of “inference to the best explanation” (J.Ladyman XXXX).

**Scope**

CRMinf regards as “knowledge” to be anything someone says and can justify as “I know that X”, regardless whether X is regarded to be true, false, probable, etc., whereas X itself is regarded as information or “data”. In this sense, knowledge resides in humans, the ones which can relate the symbols in information to states of affairs in current or past reality. “Knowledge representation” is regarded as a particular form of encoded information, for instance, a CRM compatible form. Consequently, CRMinf aims at connecting the people who know something to the information representing their knowledge, and its justification. The model supposes scientific ethics and is not concerned with beliefs of people using CRMinf different from what they state, but it can quite well be used to reason about deliberately false statements in historical sources.

CRMinf is inspired by the IAM model in Doerr, Kritsotaki and Boutsika (2011), which in turn draws on a background of other argumentation models under the aspect of application *to knowledge about the past*, among them being the “logicist” approach (XXXX) in use in France for archeological data. As the IAM, CRMinf deals with the sources of knowledge for facts stated in explicit propositions. It simplifies IAM by making the general theories used for inferences (such as a mathematical proof, universal properties etc) and the belief in their correct application an implicit part of an argumentation event (possibly represented in a text). CRMinf is also less formal than IAM with respect to inference chains (i.e., using conclusions as premises for the next inference) of different granularity. As in IAM, a documented chain of inferences represents a state of knowledge at a point in time, and *not* the historical order of finding its elements. The latter is given explicitly by the time of argument making, which is taken to be a historical fact.

CRMinf makes a basic distinction between three kinds of sources of knowledge because of the way it can be acquired, justified or falsified. These are (1) observation, (2) belief adoption and (3) inference making.

*Observation* results in knowledge acquired by human senses or by technical devices at a particular place and time. Verification or falsification may reexamine the same environment or things, if sufficiently unaltered, examine observation protocols and the functionality of employed devices and compare with independent observations. Observation is the ultimate primary soure of all knowledge. The complexity of observation processes, in particular with calibrated means, it out the scope of CRMinf, which is primarily concerned with the origin and further history of the observation results, thus providing a common generalization for other extensions, notably CRMsci.

*Belief adoption* is used in CRMinf as a term for the use of information someone has heard, read or seen presented in symbolic form and accepts it as own knowledge. It is the major source of all our communicated knowledge, including reports from observations. It is supported or questioned by assessing the provenance of the source and trust in its credibility. In case of inconsistencies between reported facts, trust argument may be used to decide for the one or the other. Therefore, CRMinf has developed the concept of Belief Adoption into much more detail than IAM, and created an “articulation” (ontological connection) to the decyphering and reading of original texts addressed by the CRM externsion *CRMtex*, in order to be able to represent critical methods in historical research. Subsequent activities of belief adoption form endless networks of information transfer, which are of great importance for historical research,

*Inference making*, the third kind of acquiring knowledge, means that one concludes from the belief in the truth or likelihood of one or more propositions , the premises, that other propositions are true or likely, using background theories, such as common logic, laws of nature or assumption about general human behavior. The peculiarity of this knowledge is that it is relative to the truth of the premise. Therefore, it may be verified or falsified by revising the truth of the premises and the validity of the background assumptions for the given context and the correct application of the background theory, such as the common errors in applying logic. Note, that an inference may conclude that at least one of the premises must be wrong. In IAM it is described as “recursive inference”, but for reasons of simplicity not distinguished in CRMinf.

The knowledge itself is represented by an instance of I2 Belief, which relates an E39 Actor to a set of propositions (I4 Proposition Set) believed forming one context and holding the same truth value (I6 Belief Value) as explicitly stated by the Actor. It comes into existence as conclusion of  an instance of I1 Argumentation (through one of its sub-classes, S4 Observation, I5 Inference Making, or I7 Belief Adoption), and ends with any modification of its truth value and propositions. Only one E39 Actor may hold a particular instance of I2 Belief, though the E38 Actor may, of course, be an instance of E74 Group. Such an instance of E74 Group may lose or gain members (via one or more instances of E85 Joining or E86 Leaving) without affecting the belief the group representatively maintains. The members supporting the common belief may not necessarily be individually convinced of it. This does not invalidate the (explicitly stated) belief of the Group, for instance, in a publication.

The disciplines addressed by CRMinf are what (Turner XXXX) calls “historical sciences”,i.e., cultural heritage studies, human and natural history, archaeology, but also descriptive empirical sciences, such as biodiversity, ethnology, geology, cultural heritage conservation, even clinical studies, etc., in their focus on documenting particular states of affairs now and in the past.

If scientists and scholars, and in particular curators, would start documenting for each information source the provenance of its immediate sources in publicly accessible systems, this partial knowledge of provenance could be “stitched together” to more and more complete networks of provenance, similar to the way these days citations in scientific publications are processed. This is a major motivation for CRMinf, the other is to make the way transparent how knowledge was acquired for enabling justified future revisions, and who is supporting contested propositions.

**Status**

CRMinf uses and extends the CIDOC CRM (ISO21127) as a general ontology of human activity, things and events happening in space-time. It uses the same encoding-neutral formalism of knowledge representation (“data model” in the sense of computer science) as the CIDOC CRM, which can be implemented in RDFS, OWL, on RDBMS and other forms of encoding. Since the model reuses, whenever appropriate, parts of CIDOC CRM, we provide in this document also a comprehensive list of all constructs used from ISO201127 following the version 7.1.2 maintained by CIDOC.

CRMinf has so far been validated in the British Museum, and by the European-funded project RICOTRANS. This document describes the first consolidated version from this experience and reviewed by CRM SIG.