Mapping of the AMICO data dictionary to the CIDOC CRM

Martin Doerr

***Institute of Computer Science, Foundation for Research and Technology – Hellas
Science and Technology Park of Crete
P.O. Box 1385, GR 711 10, Heraklion, Crete, Greece***
martin@ics.forth.gr

Abstract: The CIDOC CRM is the first ontology designed to mediate contents in the area of material cultural heritage and beyond, which has been submitted to ISO as proposal for an international standard. The AMICO data dictionary version 1.2 can be regarded as a characteristic, mature example of encoding art museum information. This report describes the semantic mapping of AMICO data elements to the CIDOC CRM and its latest extension. This work represents a proof of concept for the functionality the CIDOC CRM is designed for.

Keywords: Ontology, metadata, data structure mapping, CIDOC CRM

1. Introduction

The CIDOC Conceptual Reference Model (see [CRM1999], [CRM2001], [Doerr99]), in the following only referred as "CRM", is outcome of an effort of the Documentation Standards Group of the CIDOC Committee (see "http://www.cidoc.icom.org, http://cidoc.ics.forth.gr") of ICOM, the International Council of Museums beginning in 1996. It is a domain ontology formulated in the form of an object-oriented semantic model aiming to solve the problem of semantic interoperability between museum data of various kinds and their relations to archive and library material. It has been accepted by ISO TC46, SC4 as work item in September 2000 [CRM1999].

After the model has found a stable form in 1998, and was accepted by CIDOC on the CIDOC Conference in Melbourne in this year, the Group decided that the model has to undergo a series of practice test in order to verify its utility and completeness in a so-called "Correlation Test Project". Particular importance was seen in the demonstration that and how the CRM is able the mediate between current relevant standards of the domain, that are regarded relevant to the scope set for the CRM. This is done via sets of mapping rules, that define how to transform descriptions given in the structures of one standard into an equivalent description in the CRM with the same intended meaning.

The Art Museum Image Consortium (AMICO, http://amico.org) is a not-for-profit organization of institutions with collections of art, collaborating to enable educational use of museum multimedia. Together, AMICO Members are building The AMICO Library, a licensed digital educational resource available under subscription universities and colleges, public libraries, elementary and secondary schools, and museums. The AMICO Library is the compilation of digital multimedia documentation of works of art contributed by the more than 30 AMICO members. The AMICO data dictionary version 1.2 defines the data structure of the AMICO library.

The first mapping of the AMICO fields in 1999 and other mapping experiments has motivated a proposal for extensions, which has become part of the proposed CIDOC CRM version 3.0 [CRM2001]. The mapping below is based on the version 3.0. It is recommended to read the mapping of the DC Element Set [Doerr00] before this document. This document can only be read together with the Definition of the CRM version 3.0 and the AMICO data dictionary version 1.2.

We want to express our warmest thanks to David Bearman and Jennifer Trant for their explanations of the AMICO data fields.

2. Formalism:

The AMICO data dictionary version 1.2 describes a structured record. The CRM is a semantic network. In a structured record,
- no distinction is made between the nature of an attribute value and kind of relation expressed by the attribute between the record as a whole and the attribute value.
- Referred discrete entities are not explicitly identified, but rather referred to by some representative values.

A semantic net,
- declares each attribute value as an a priori independent entity in the universe of discourse and connects such values symmetrically by directed links/arcs .
- The entities acquire identity independent from the composition of descriptional elements in some data structure.

In the following, CRM entities are referred by the unique id and name , eventually followed by a comment in parenthesis, like:

E7 Activity (of type "exhibition")

CRM links are referred by the unique id of the a*ppl*i*cabl*e entity in the given context, followed by the unique id of the entity from which the link is in*h*erite*d*, the name of the link and the name of the referred entity, like:

E22 (E7) was used for: Activity

The name of the link is only the part associated with the direction applicable for the given context. E.g. we present the link "E11 Modification.has produced (was produced by)" as

E22 (E11) was produced by: Modification

if we want to read it from the product rather than from the event.

Wehave therefore chosen a mapping format, where for each AMICO field we declare the correspondence of the **va***lu***e** to a CRM entity, like:

OEH　　　　= E7 Activity (of type "exhibition")

and in the sequence we declare the correspondence of the **re***l***ation** expressed by the field to a CRM link, like:

CR->OEH　　= E22 (E7) was used for: Activity

Where CR stands for the catalog record as a whole (like a root element in SGML). In cases of data groups, we have typically identified the abstract group identifier with a link:

CLG   = E22 (E1) has type: Type

If the data group corresponds to a CRM entity, we declare the correspondence of the **re/ation** expressed by the field with respect to the group identifier, rather than the catalog record, like:

MEG -> MDU = E54 value: Number

If the relation expressed by an AMICO field corresponds to a path with intermediate entities in the CRM, we have marked it as "(JOIN)", like:

CR -> OMS = E22 (E19) is composed of – E 19 (of type support) consists of:
                                                        Material (JOIN)

We have preceded each mapping group with a textual comment for the non-expert reader. The latter may be better to present the mapping to a wider public, whereas the degree of analysis presented here is suitable to implement data transfer routines.

The "has note" link appears very often. On one side, the AMICO format foresees many free texts that express a certain aspect. The CRM has deliberately not analyzed in the current stage the possible specializations of "has note". It is left to the user to specify the type ("has type" link) of the note. We have made certain proposals for note types below. The type "Definition" stands for a text explaining the nature of the entity it is attached to itself. Those could be subject to standardization at some future time. For the moment, they serve to preserve the semantics of distinct text fields if the data are transported as CRM instances. On the other side, an AMICO text field quite often refers implicitly to an independent entity that exists in the CRM. In that case we have mapped the text explicitly to that entity, like:

OCH corresponds to E14 Condition Assessment, represented as the text associated to one ore more E14 entities.
OCH   = E62 String
CR-> OCH   = E22 (E19) assessed by – E14 has note: String (JOIN)
                                           (has type: "Definition")

3. Mappings:

We regard two semantic levels. In the first, an AMICO dataset about an object is seen in its completeness as an instance of E31 Document, which can be identified by AID plus appropriate scope identifier.

On the second level, the contents of an AMICO dataset about an object is analyzed in terms of the real-world entities it refers and relates. Under this consideration, we do the following mappings:

Unique identification
1. The catalog record as a total is regarded as a 1:1 correspondance to the described object, in the following referred to as "CR". (We assume, that one object is documented by not more than one record)
    CR   = E22 Man-Made Object

2. AID represents the catalog record as a document, that refers to the object:
    AID -> CR = E31 documents: CRM Entity

What is it?
3. OTY (Object-Type) corresponds to the subclass of E55 Type for Man-Made Objects

```
        OTY      =  E55 Type (Man-Made Object Type)
   CR -> OTY  =  E22 (E1) has type: Type
```

4. OPP (Object-Parts/Pieces) can be decomposed into objects of a certain type and their cardinality:
```
   OPP        = E55 Type (Man-Made Object Type)
                  AND/OR
                  E60 Number
   CR -> OPP  = E22 (E19) is composed of - E22 (E1) has type: Type (JOIN)
                  AND/OR
                     E22 (E19) is composed of - E22 (E19) has number of parts: Number
                                                                      (JOIN)
```

4.b Alternatively it is just a text:
```
   OPP         = E62 String
   CR -> OPP   = E22 (E1) has note: String (Text attached to the object itself)
                                (has type: "Object-Parts/Pieces").
```
5. CLG (Classification Group) corresponds to the CRM link to the type of the object, and
    CLT(Classification-Term) to the type itself.
```
   CLG         = E22 (E1) has type: Type
   CLT         = E55 Type (Man-Made Object Type)
   CLS         = E32 Authority Document
   CLS -> CLT = E32 is part of: Authority Document
```

6. OTG (Object-Title/Name Group) corresponds to the CRM link to the title of the object, and OTN to
    the title itself.
```
   OTG         = E22 (E19) has title: Title
   OTN         = E35 Title
   OTT         =  E55 Type (Title Type)
   OTN -> OTT = E35 (E1) has type: Type
```

7. OST (State) corresponds to data around the single common process that created the multiple.
```
   OST         = E12 Production  (The event that created the series)
   CR-> OST    = E22 (E11) was produced by: Modification
```
    In this mapping, the number of objects created is in the E12 Production event note (or a reasonable
    extension of the CRM).

    Alternatively:

7.b OST   = E62 String
```
   CR -> OST = E22 (E1) has note: String (text attached to the object itself)
                                (has type: "State").
```
8. OEN (Edition) corresponds to E29 Design or Procedure, i.e. the non-material process which devised
    the edition.
```
   OEN         = E29 Design or Procedure
   CR -> OEN   = E22 (E11) was produced by – E12 (E11) used specific technique:
                     Design or Procedure (JOIN)
```
    Actually this JOIN reuses the same Production event from 7.a

9. OPD, (Physical Description), OPA (Physical Orientation/Arrangement) are chapters of the text
    associated directly to the object.
```
   OPD         = E62 String
   CR -> OPD  = E22 (E1) has note: String (Text associated to the object itself)
                                (has type: "Physical Description").
   OPA         = E62 String
   CR -> OPA  = E22 (E1) has note: String (Text associated to the object itself)
```

(has type: "Physical Orientation/Arrangement").

10. MET (Measurements-Text) can be interpreted as the text associated with an act of measurement
    MET   = E62 String
    CR -> MET = E22 (E16) was measured – E16 has note:String (JOIN)
                             (has type: "Measurements-Text").

11. MEG (Measurements Group) corresponds to the entity E54 Dimension and its attributes
    MEG   = E22 (E18) has dimension: Dimension
    For the following, one E54 Dimension instance is created all other fields connect to.
    MCM   = E55 Type (Dimension Type)
    MEG -> MCM = E54 has type: Type
    MED   = E55 Type (Dimension Type)
    MEG -> MED = E54 has type: Type
    MDV   = E60 Number
    MEG -> MDU = E54 value: Number
    MDU   = E58 Measurement Unit
    MEG -> MDU = E54 unit: Measurement Unit

    MEQ (Measurement Qualifier): from a mathematical point of view, this should be encoded in the value by a standardardized interval notation. From a methodical point of view, it should be added as text to "MET" above.

12. OMG (Materials and Techniques Group) is generally associated with the production process (E12):
    OMG   = E12 Production

    If there is a specifically elaborated process associated, e.g. an existing plan, it corresponds to E29 Design or Procedure, and there is an aspect of "E18 consists of: Material" (OMM):
    OMG   = E22 (E11) was produced by – E11 used specific technique:
                                Design or Procedure (JOIN)
    OMD   = E62 String
    OMG-> OMD = E12 (E1) has note: String
                        (has type: "Materials and techniques").
     (the note is attached to the production event, not the object!)

    OMT   = E55 Type (Design or Procedure Type)
    OMG-> OMT = E12 (E11) used general technique: Type (JOIN)

    OMM   = E57 Material
    If the material is a consequence of the plan:
    OMG-> OMM = E12 (E11) used specific technique - E29 usually employs: Material (JOIN)
    CR ->   OMM = E22 (E18) consists of: Material

    OMS   = E55 Type (Design or Procedure Type)
              combined with
              E57 Material
    Both:
    CR -> OMS =E22 (E19) is composed of – E19(of type "support") was produced by
                    – E11 used general technique:Type (double JOIN)
    CR -> OMS = E22 (E19) is composed of – E19 (of type support) consists of:
                               Material (JOIN)
    Or simply summarized at the "whole" object:
    CR-> OMS = CR-> OMM OR OMG-> OMT

    OIN (Inscriptions and/or Marks) corresponds to E37 Mark, a specialization of E36 Visual Item, represented as the associated text to one ore more E37 entities.

OIN   = E62 String
CR-> OIN   = E22 (E24) shows visual item – E37 (E1) has note: String (JOIN)
                                              (has type: "Definition")

13. OCH (Condition/Examination History) corresponds to a set of E14 Condition Assessment events, and represented as the text associated with one ore more E14 entities.
OCH   = E62 String
CR-> OCH   = E22 (E19) assessed by – E14 has note: String (JOIN)
                                              (has type: "Definition")

14. OTH (Treatment/Conservation History) corresponds to a set of E11 Modification events (every treatment is a modification in the sense of the CRM!), and is represented as the associated text to one ore more E11 entities.
OTH   = E62 String
CR-> OHT   = E22 (E11) was produced by – E11 has note: String (JOIN)
                                              (has type: "Definition")

Who made it?

The following three units, "who made it", "when was it made", "where was it made" originate in the CRM in one or more E12 Production events. The attachment of these three units directly to the object does not allow to describe creations with phases of changing people, places and dates. Therefore we would recommend to AMICO to introduce a "Production" or "Creation" data group. Therefore in the following we necessarily assume ONE Production event per object, and all links below from an E12 entity are thought to originate at the same instance per object.

15. CRG (Creator Group) corresponds to E21Person, a specialization of E39 Actor. It seems, as if AMICO did not explicitly foresee groups of people creating objects without knowledge about the participant individuals. Alternatively, we could map CRG to Actor in general, even though those do not necessarily have birth and death.

CRG   = E22 (E11) was produced by – E12 (E7) carried out by: Actor (JOIN)

CRQ (Creator-Qualifier, attribution of the work to…) is seen as a property of the production event, and not of a person/actor:
CRQ   = E62 String
CRG-> CRQ = E12 (E1) has note: String (Text attached to the production event)
                              (has type: "Creator attribution").

CRN   = E21 Person (identifier of)

CRT   = E62 String ### better Appellation
CRG-> CRT = E21 (E1) has note: String (Text attached to the Actor/Person)
                              (has type: "Appellation").

CDT   = E62 String
CRG-> CDT = E21 (E1) has note: String (Text attached to the Actor/Person)
                              (has type: "Short Biography").
CRB   = E62 String
CRG-> CRB = E21 (E1) has note: String (Text attached to the Actor/Person)
                              (has type: "Extended Biography").
CNO   = E62 String
CRG-> CNO = E21 (E1) has note: String (Text attached to the Actor/Person)
                              (has type: Creator and Work").

CRC (Creator-Culture/Nationality) has three interpretations: Either the culture the creator was born into, or the group he belonged to during the creation, or the cultural context of the creation. The first can either be represented as a "Type" of Actor/Person or as the Period his Foundation(E66)/ Birth(E67) falls into. In the second the object can be seen as product of a group. The third is the Period the Production falls into: Note, that Period in the CRM is not defined by the date and place, but by the kinds of manifestation. So, Symbolism and Impressionism could be perceived as distinct periods. Note, that the CRM allows to perceive a whole nation as a group (Actor).

CRC   = E55 Type (Person Type)
CRG-> CRC = E21 (E1) has type: Type
AND/OR* if related to the creator's birth:
CRC   = E4 Period
CRG-> CRC = E21(E67) was born – E67 (E4) falls within: Period (JOIN)
AND/OR* if related to local influence during the production:
CRC   = E4 Period
CR-> CRC   = E22 (E11) was produced by – E11 (E4) falls within: Period (JOIN)
AND/OR* if only the group (E74, subclass of Actor) of the creator is known:
CRC   = E74 Group
CR-> CRC   = E22 (E11) was produced by – E11 (E7) carried out by: Actor (JOIN)

CGN (Creator-Gender) is a Type of person:
CGN   = E55 Type (Gender Type)
CRG-> CGN = E21 (E1) has type: Type

The following fields:
CBD (Creator-Birth-Date), CBP (Creator-Birth-Place) are links of the Birth (E67) event. CBQ (Creator-Birth-Qualifier) should be encoded in the date expression itself, but can also be given in free text:

CBD   = E52 Time-Span
CRG-> CBD = E21(E67) was born – E67 (E2) has time-span: Time-Span (JOIN)
CBQ   = E62 String
CRG-> CBQ = E21(E67) was born – E67 (E2) has time-span -E52 (E1) begins at qualify: String
                                              AND/OR*   E52 (E1) ends at qualify: String
                               (double JOIN)
CBP   = E53 Place
CRG-> CBD = E21(E67) was born – E67 (E4) took place at: Place (JOIN)

The following fields:
CDD (Creator-Death-Date), CDP (Creator-Death -Place) are links of the Death (E69) event. CDQ (Creator- Death-Qualifier) should be encoded in the date expression itself, but can also be given in free text:

CDD   = E52 Time-Span
CRG-> CDD = E21(E67) died in – E69 (E2) has time-span: Time-Span (JOIN)
CDQ   = E62 String
CRG-> CDQ = E21(E67) died in – E69 (E2) has time-span -E52 (E1) begins at qualify: String
                                              AND/OR*   E52 (E1) ends at qualify: String
                               (double JOIN)
CDP   = E53 Place
CRG-> CDP = E21(E67) died in – E69 (E4) took place at: Place (JOIN)

The fields CAD (Creator-Active-Date), CAP (Creator-Active-Place) correspond to one or multiple E7 Activity periods :

CAD   = E52 Time Span
CRG-> CAD = E21 (E7) performed – E7 (E2) has time-span: Time-Span (JOIN)
CAP   = E53 Place
CRG-> CAP = E21 (E7) performed – E7 (E4) took place at: Place (JOIN)

When was it made?

16. OCG corresponds to E52 Time-Span attached to the E12 Production event. OCS (Creation-Date-Start) and OCE (Creation-Date-End) should be encoded into a time-interval primitive enclosing the possible creation date:

OCG   = E22 (E11) was produced by – E12 (E2) has time-span:
                                        Time-Span (JOIN)
OCT   = E62 String
OCG->OCT = E52 (E1) has note: String (Text attached to the time-span)
                           (has type: "Definition")
OCS/OCE   = E61 Time Primitive
OCG->OCS/OCE   = E52 (E1) at most within: Time Primitive
OCQ   = E62 String
OCG->OCQ = E52 (E1) begins at qualify: String
                AND/_OR_
           E52 (E1) ends at qualify: String

Where was it made?

17. OCP corresponds to E53 Place attached to the E12 Production event.

OCP   = E53 Place
CR -> OCP = E22 (E11) was produced by – E12 (E4) took place at: Place

What is it about

18. STG (Style/Period Group) corresponds to text and types of the object itself, the STG group itself does not map. From the point of view of the CRM, style and period are distinct. The first is a morphological notion about the object, the second a physical notion about the production event. The morphological notion is seen as a type of objects, which **ten**_d_ to appear in a certain period, but are not necessarily bound to it. Periods often imply characteristic styles named the same. Here the mapping of Style:

STD   = E62 String
CR-> STD = E22 (E1) has note: String (Text attached to the object itself)
                         (has type: "Style")
STT   = E55 Type (Man-Made Object Type)
CR-> STT = E22 (E1) has type: Type

19. SUG (Subject Matter Group) corresponds to the "depicts" links of the object, and textual notes to the object:

SUP   = E62 String

```
CR-> SUP     = E22 (E1) has note String (Text attached to the object itself)
                           (has type: "Subject")
SUI          = E62 String
CR-> SUI     = E22 (E1) has note: String (Text attached to the object itself)
                           (has type: "Iconography")
```

SUT maps to any combination of three different links, depending on the kind of subject. The CRM devotes special links to identifiable items in the subject in order to cross-reference depicted objects to where they might be described independently from this depiction.

For non-real subjects (e.g.: "a" cow, Crucification, Holy Virgin with Child, joy, ICONCLASS codes):
```
SUT          = E55 Type (subject types)
CR-> SUT     = E22 (E24) depicts concept: Type
```
For objects, "snapshots" of places on earth and any living or dead individual:
```
SUT          = E18 Physical Entity
CR-> SUT     = E22 (E24) depicts object: Physical Entity
```
For events, i.e. snapshots understood as a scene from some process of interest:
```
SUT          = E5 Event
CR-> SUT     = E22 (E24) depicts event: Event
```

20. CXG (Context Group) corresponds to a series of E7 Activities, if not to a simple text attached to the object. If CXD is not a text attached directly to the object itself, it can be "distributed" to the individual activities. The AMICO format may loose the per event connection between person, date and place, if one CXG contains more than one activity. This may hinder relevant reasoning on these data in the future. The connection between the object and the activity can be either through use or purpose, and activities can relate by purpose (e.g. orders, gifts etc.):

```
CXG          = E7 Activity
CR->CXG      = E22 (E7) was used for: Activity
                     AND/OR
               E22 (E7) was made for - E7 had specific purpose: Activity (JOIN)

CXD          = E62 String
CR->CXG      = E22 (E1) has note: String (Text attached to object)
                           (has type: "History")
AND/OR
CXG-> CXD    = E7 (E1) has note: String (Text attached to each activity)
                           (has type: "Definition")
CXP          = E39 Actor
CXG->CXP     = E7 carried out by: Actor
CXS          = E53 Place
CXG->CXS     = E7 (E4) took place at: Place
CXT          = E52 Time-Span
CXG->CXT     = E7 (E2) has time-span: Time-Span
```

What does it mean?

21. OCR (Critical Responses) corresponds to textual notes to the object

```
OCR          = E62 String
CR-> OCR     = E22 (E1) has note: String (Text attached to the object itself)
                           (has type: "Critical Responses")
```

Who showed it?

22. OEH (Exhibition or Loan History) corresponds in CRM to context history:

OEH   = E7 Activity (of type "exhibition")
CR->OEH   = E22 (E7) was used for: Activity

Who owns it?

23. OOG (Owner Group) corresponds in CRM to the distinct notions of custody and legal ownership, as many museums keep objects for third parties.

OON   = E39 Actor
OOG   = E22 (E19) has current owner: Actor
                AND/*OR*
            E22 (E19) has current keeper: Actor

OOP (Owner-Place): The CRM foresees a complete "Address" object rather than a place name only. This link has the meaning of contacting the Actor rather than locating him physically (An address could be a P.O.Box etc.). Hence either:

OOP   = E45 Address
OOG->OOP = E39 has contact points: Contact Point

Or OOP can be mapped to the link to where an actor resides:
OOP   = E53 Place
OOG-> OOP = E39 has current or former residence: Place

OOA   = E42 Object Identifier
CR->OOA   = E22 (E19) is identified by: Object Identifier
                AND/*OR*
            E22 (E19) preferred identifier is: Object Identifier

OOC   = E62 String
CR-> OOC   = E22 (E1) has note: String (Text attached to the object itself)
                        (has type: "Iconography")

24. OPO (Provenance/Prior Owners-Text) corresponds to E39 Actor, the previous owners. The CRM links summarize all owners. More detailed information can be mapped to specific activities.

OPO   = E39 Actor
CR-> OPO   = E22 (E19) has former/current owner: Actor
                AND/*OR*
            E22 (E19) has former/current keeper: Actor

25. ORG (Rights/Copyright Group) corresponds to E30 Right

ORG   = E30 Right
CR-> ORG   = E22 (E72) is subject to: Right

```
CRS            = E62 String
CR-> CRS   = E30 (E1) has note: String (Text attached to the Right)
                            (has type: "Definition")
CRL            = E62 String
CR-> CRL   = E30 (E1) has note: String (Text attached to the Right)
                            (has type: "Reference")
```

What is it related to?

26. RWG (Related Works of Art Group) corresponds to E22 Man-Made Object, however such a generic relation is not present in the current CRM. All causal relations should be mapped to the E12 Production entity, like derivatives and parallel product. Others, like part-of, exist explicitly in the CRM. Respectively:

```
RWD            = E62 String
CR-> RWD   = E22 (E1) has note: String (Text attached to the object itself)
                            (has type: "Relations")
RWL corresponds to a link to an E31 Document about that work
RWL            = E22 (E31) is documented in: Document

RWR            does not map directly, but must be interpreted accordingly. As AMICO recommends DC
```
Relation Type Values, details can be found in the Mapping of the DC Elements [Doerr2000].

27. RIG (Related Images Group) corresponds to E38 Image, subclass of E28 Conceptual Object.

```
RIL            = E38 Image
RIG            = E22 (E28) is referred to by: Conceptual Object
RID            = E62 String
RIG-> RID     = E38 (E1) has note: String (Text attached to the documentary image)
                            (has type: "View Values")
RIP            = E62 String
RIG-> RIP     = E38 (E1) has note: String (Text attached to the documentary image)
RIR            = E55 Type
RIG-> RIR      = E38 (E1) has type: Type (Type of depiction mode)
```

28. RMG, RDG correspond to E31 Document (the differentiation is made in the type of document)

```
RML            = E31 Document
RMG            = E22 (E31) is documented in: Document
RMD            = E62 String
RMG-> RMD  = E31 (E1) has note: String (Text attached to the multimedia)
                            (has type: "Definition")
RDL            = E31 Document
RDG            = E22 (E31) is documented in: Document
RDD            = E62 String
RDG-> RDD  = E31 (E1) has note: String (Text attached to the document)
                            (has type: "Definition")
```
RMR, RDR does not map directly, but must be interpreted accordingly. As AMICO recommends DC Relation Type Values, details can be found in the Mapping of the DC Elements [Doerr2000].

Who documented it?

29. As we map AID to the catalog record as a document, DCG (Documentation/Cataloguing-History) maps to the E65Conceptual Creation event associated with this document.

DCB   = E39 Actor
AID->DCB = E31 (E65) was created by – E65 (E7) carried out by: Actor (JOIN)
DCD   = E52 Time-Span
AID->DCD = E31 (E65) was created by – E65 (E2) has time-span: Time-Span (JOIN)

Version Control?

30. The following fields can be regarded to be out of the scope of the CRM, but could be mapped as activities:

AVD, AVV, ADP, DEL, ALY.

4. Conclusions

The above mappings can be used to create CRM instances from AMICO Records, which preserve the semantics of the original records. In many cases, the AMICO information contains textual fields, which can be analyzed in more detail in the CRM. It may not always be possible to make analysis from free text automatically. It is not the concern of the CRM, if a better mapping must be done manually, or which technology (like natural language processing) is the best to do it. Here we propose alternatives for both, simple automatic mapping or an optimal, may-be manual process. The most analytic mapping will allow for the richest reasoning on these data, and the optimal comparison with data from other sources. It may be worthwhile to standardize the note-types for interoperability of the various junks of associated texts.

Let us summarize the most important structural differences between AMICO and the CRM:

1. In Number 4.a, "OPP" has no 1:1 equivalent to CRM. It is the summary proposed in the CIDOC Information Groups and Categories, 1995. The CRM abandoned this form in favor of dealing with parts as full objects. On the other hand, the CRM foresees all more general statements about parthood to be described in a note. Both methods can be applied for AMICO.

2. MEQ: We propose that all values (times, coordinates, measures etc.) should be given as value ranges expressing the uncertainty, may be using middle values in addition.

3. OMG: The CRM distinguishes materials specific only to the object instance, and those specific to the technique applied. This results in a slightly different linking, but is nevertheless compatible in contents. The CRM tries to accommodate different documentation practices by that. The rather domain specific field about the "support" of a work of art can either be mapped introducing a derived "support object", or just be ignored, i.e. summarized as text at the object. The first is richer, the second poorer than AMICO.

4. The CRM distinguishes identifiable items in the subject in order to cross-reference depicted objects to where they might be described independently from this depiction.

5. The groups Who made it? When/Where was it made may be better grouped into production events/phases. In general, the CRM suggests that data become easier convertible to other formats if events are made explicit wherever possible. This holds also for "context" etc. The cases of multiple creators, and multiple phases of creation can only be documented correctly, if the events are made explicit. E.g. the difference between the forming of the mold and the casting of a bronze item. The case of groups or institutions as creators seems not to have been foreseen in AMICO.

The mapping of AMICO has motivated besides others, several extensions to the CRM. Those were in particular biographical entities and the creation of non-material objects. This mapping exercise is a successful proof of concept for the CRM. On the other side it proves the compatibility of the AMICO data with a future CRM-based ISO standard. Future work could go into practical data transfer experiments, and study to which degree textual fields can be analyzed efficiently into more detailed explicit concepts. Another application problem on the way to complete data interchange and merging is the matching of all associated events, places, persons with data from other records.

5. References

[CRM2001] Nick Crofts, Ifigenia Dionissiadou, Martin Doerr, Matthew Stiff (editors), "Definition of the CIDOC object-oriented Conceptual Reference Model ", March 2001 (version 3.0). See locations: http://cidoc.ics.forth.gr/docs/crm_version_3.0.rtf

[CRM1999] Nick Crofts, Ifigenia Dionissiadou, Martin Doerr, Matthew Stiff (editors), "Definition of the CIDOC object-oriented Conceptual Reference Model" , September 1999 (version 2.1) http://cidoc.ics.forth.gr/docs/CRM_version_2_1.rtf

[Doerr99] Martin Doerr and Nicholas Crofts "Electronic Esperanto: The Role of the Object Oriented CIDOC Reference Model", Proc. of the ICHIM'99, Washington, DC, September 22-26, 1999.

[Doerr00] Martin Doerr, "Mapping of the Dublin Core Metadata Element Set to the CIDOC CRM" , Technical Report, 274, ICS-FORTH, Heraklion, Crete, July 2000. http://cidoc.ics.forth.gr/docs/dc_to_crm_mapping.rtf