

# **The Agios Pavlos Extensions -**

## **Add-ons for the Completion of the CIDOC CRM**

Editor: Martin Doerr  
Date : 25 July 2000

Participants of the Agios Pavlos Meeting:

Nick Crofts  
Ifigenia Dionyssiadou  
Martin Doerr  
Tony Gill  
Siegfried Krause  
Angela Spinazze  
Steven Stead  
Matthew Stiff

### ***Contextual Note:***

This Document should be read together with the report from the Agios Pavlos Meeting and the Formal Definition of the CIDOC Conceptual Reference Model, version 2.2. or 2.3.

Following presentations of experience from the mapping of the AMICO datamodel, EAD, the Dublin Core Elements, including all so-far proposed Qualifiers, and mapping of the SPECTRUM data fields, presentation of the CIDOC Archeological data model, and on the background of the participants' data model needs, the Agios Pavlos Group proposes the following set of extensions to the CIDOC CRM version, that has been submitted for ballot to ISO TC46, SC 4.

This set of extensions has the power to **cover completely** the Dublin Core Elements and the so far proposed Qualifiers. It covers widely the scope of the CRM set in the Agios Pavlos Meeting report. It is proposed as a well-defined stage for the further discussion and development of the CRM in the ISO framework into its final form. A complete list of the further open issues about the CRM contents identified in the Meeting will be presented on the CIDOC CRM Test Project pages.

### ***Introduction:***

On widening the scope from that of the CIDOC Information Categories, set as a starting point, the CRM needs several additional properties and entities in order to complete temporal/causal reasoning and acquire sensible symmetry with respect to the development of the fundamental categories. In a few cases properties have been risen to a higher abstraction level either in the domain or in the range. These changes are all but two conservative, as they do not invalidate instances of the previous release. Thereby the stability of the CRM and the value of the applied object-oriented methodology has been confirmed.

In more detail, the changes refer to:

- temporal/causal reasoning with events of unknown type
- symmetries and commonalities between Physical and Conceptual Object.
- extension of the decomposition and appellation logic to Actor
- reasoning about Type
- completion of basic description of Actor
- improvement of temporal reasoning

Non-conservative changes are:

Property “**has title (is title of): Title**” was taken over from E19 Physical Object onto E71 Man Made Stuff. This change disallows title relations for non-man-made items. Those may have Appellations in general. The idea behind is, that titles are assigned to characterize creations. This may be debatable.

The re-encoding of Time-Span does not change the information contents, but needs rearrangement of data between instances of the current CRM and this proposal.

During the discussion, three methodological principles have been identified as appropriate, which motivate the style of the following extensions:

1. “Abstract classes” or “abstract entities” are not seen strictly in the sense, that they cannot be directly instantiated. Instantiating such an entity means, that the information fragment mapped by the respective instantiation was so incomplete, that no concrete subclass could be identified for it. In this sense, may be the notion “abstract” should be banned from the model.
2. Specialization of any entity into a set of subclasses does **not** mean, that these subclasses **cover** the scope of their superclass. Therefore definition of **complementary** entities should be carefully **avoided**, like “Legal Body” – “Non Legal Body”. The philosophical reason behind is the open world assumption necessary for cultural models. Even in cases like splitting Person into male and female, we cannot clearly argue that the complement is empty.
3. Entities should only be declared, if they are combined with the declaration of one or more properties specific to them. Else the Type-hierarchy should be used. This is another reason not to declare complementary entities.
4. Properties should be declared between the most specific entities, that cover the applicability of the perceived meaning according to the current understanding. If later wider cases appear, it is a conservative change to relax the model. This practice ensures the specificity of the model and the reasoning it supports. It comes together with the need, to define a set of highest level (abstract??) properties as containers for underspecified information.

In a total, 14 new entities and 23 new properties have been declared, domain or range of 8 properties was widened, 3 properties have been redefined, 2 properties deleted, 2 entities and 1 property renamed. A respective revised formal definition with complete cross-references will be produced in a while.

The rest of this document is loosely grouped into units, that represent some sort of reasoning and/or scope.

### **First unit**

This unit has to do with the ways of coming into existence and ending an existence.

#### **1. E63 Beginning of Existence**

subclass of: **Event**

superclass of: **Conceptual Creation**

**Production**

**Formation**

**Birth**

Scope note: Entity for temporal reasoning about things beginning to exist –

Intellectual products, physical items, groups of people, living beings , periods ?–

A hook for termini postquem and antequem. It may turn out useful to define more specializations of this entity for Natural History.

Properties:

**brought into existence (was brought into existence by) : Entity**

property note: this is the superproperty of “has produced”, “has formed”, “brought to life”, “has created”.

## 2. E64 End of Existence

subclass of: **Event**

superclass of: **Destruction**

**Dissolution**

**Death**

Scope note: Entity for temporal reasoning about things stopping to exist – physical items, groups of people, living beings –  
A hook for termini postquem and antequem.

Properties:

**took out of existence (was taken out of existence by) : Entity**

property note: this is the superproperty of “destroyed”, “dissolved” “was death of”.

## 3. E65 Conceptual Creation

Subclass of: **Beginning of Existence**

**Activity**

Scope note : The creation of an immaterial product – Text, music, image, movie, Law etc.

Properties:

**has created (was created by) : Conceptual Object**

property note: this is a subproperty of “brought into existence”

## 4. E66 Formation

Subclass of: **Beginning of Existence**

**Activity**

Scope note: The formation or foundation of a formal or informal group of people

Properties:

**has formed (was formed by) : Group**

property note: this is a subproperty of “brought into existence”

## 5. E67 Birth

Subclass of: **Beginning of Existence**

Scope note: The birth of a human being.

Properties:

**brought into life (was born) : Person**

property note: this is a subproperty of “brought into existence”

**from father (was father for) : Person**

**by mother (gave birth) : Person**

## 6. E68 Dissolution

Subclass of: **End of Existence**

Scope note: The formal or informal end of a group of people. If it was a deliberate act, the instance should also be instantiated as an activity

Properties:

**dissolved (was dissolved by) : Group**

property note: this is a subproperty of “took out of existence”

## 7. E69 Death

Subclass of: **End of Existence**

Scope note: The death of a human being. If the person was killed, the instance should also be instantiated as an activity

Properties:

*was death of (died in)* : **Person**

property note: this is a subproperty of “took out of existence”

### **Second unit:**

Comparing Physical and Conceptual Objects, several problems appear. Some properties are common, some are only analogous. In particular, “has title”, “is subject to : Right”, “right held by”, “is composed of”, “is identified by”, and the use/motivation/reference properties need a careful generalization.

This led to the introduction of three high-level concepts, the “Stuff”, “Man-Made Stuff”, and “Legal Object”. The lower-level entity “Information Object” serves to make a distinction between ideas, like types etc., and well-structured immaterial objects. Even though the revised isA hierarchy has a higher complexity, the number of properties per entity (in particular Physical Object) is reduced, which creates confidence, that the domains/ranges of these properties are captured by far more precise than before.

1. E24 Man-Made Entity renamed to:  
**E24 Physical Man-Made Stuff**

2. E18 Physical Entity renamed to:  
**E18 Physical Stuff**

3. **E70 Stuff**

Subclass of: **CIDOC Entity**

Superclass of: **Man-Made Stuff**  
**Physical Stuff**

Scope note: An identifiable, discrete, persistent item which constitutes a unit for documentation, be it an intellectual product or a physical thing. Such items are characterized by a relative stability, i.e. either a solid physical form, an electronic encoding, a reproducible pattern, a logical concept or structure. It introduces the concept of **being used**.

Excluded are liquids flowing, gases moving around etc. A packaged liquid however fits this entity.

Properties (moved up from E19 Physical Object):

had as general use (was use of): Type

4. **E71 Man-Made Stuff**

Subclass of: **Stuff**

Superclass of: **Physical Man-Made Stuff**  
**Conceptual Object**

Scope note: An identifiable, discrete, persistent item deliberately produced, invented or created by humans which constitutes a unit for documentation, be an intellectual product, idea or a physical thing.

Properties (taken over from E19 Physical Object):

has title (is title of): Title

(has type : Type)

Properties (moved up from E22 Man-Made Object):

was intended for (was intention of): Type

5. **E72 Legal Object**

Subclass of: **CIDOC Entity**

Superclass of: **Information Object**  
**Physical Stuff**

Scope note: An identifiable item which can be owned or people can have a right on. It is not restricted to Stuff. May be Legal Bodies should be included.

Properties (taken over from Physical Object):

is subject to (applies to): Right

right held by (has right on): Actor

(has type: Type)

(has note: String)

6. **E73 Information Object**

Subclass of: **Conceptual Object**  
**Legal Object**

Superclass of: **Iconographic Object**  
**Linguistic Object**  
**Visual Item**  
**Document**  
**Design or Procedure**

Scope note: An identifiable immaterial item which constitutes a unit for documentation and has an objectively recognizable structure. Examples are data sets, images, texts, procedure prescriptions. It does not include types, ideas etc.

Properties: **is composed of (forms part of): Information Object**

7. Redirection and renaming of :

E31 Document : refers to (is referred to by): Physical Object  
into:

**E31 Document: *documents (is documented in)* : CIDOC Entity**

This property aims at identifying references that contain descriptive information about some entity. It is subproperty of E28 Conceptual Object: refers to .

8. The use and motivations reasoning is completed by the following properties, attached to activity due to the priority rule:

E7 Activity :

Additional Properties:

**took into account (was taken into account by) : Conceptual Object**

property note: The equivalent of using something physical. May be better "used: Stuff".

**motivated the creation of (was created for) : Conceptual Object**  
**was motivation for (motivated) : Conceptual Object**

The latter may be better at "Event", or event higher.

9. In order to cover at highest level the notions of participation, E5 Event should have two high-level properties, which are to be specialized in the subclasses:

E5 Event :

Additional Properties:

**occurred in the presence of (was present at): Stuff**  
**had participants (participated in): Actor**

These two properties complete the metamodel, and are understood as abstract properties, that comprise all specializations below like passive or active roles and are only instantiated in cases of insufficient knowledge. This can happen, if data are taken from poor "core models", or from unspecific sources. These links are also useful for basic causal/temporal reasoning. The explicit list of subproperties will be given in a separate document.

10. Finally the idea of references is generalized intellectually over the idea of a documentary reference to allusions etc. References in general can support temporal and spatial reasoning:

E28 Conceptual Object :

Additional Properties:

**refers to ( is referred to by) : CIDOC Entity**  
**(has type: Type)**  
**refers to concept (is referred to by) : Type**

### **Third unit:**

This unit provides a core model for human organization.

#### **1. E74 Group**

Subclass of: **Actor**

Superclass of: **Legal Body**

Scope note: A group is any gathering of people that acts collectively or in a similar way due to any kind of social bounds or contact.  
Nationality can be handled as a link to a group of appropriate type. Note the distinction between citizenship and ethnic group and other subtleties not easily expressible by nationality adjectives.  
Examples: A group of people painting together at some happening,  
A tribe of indigenous people, an artist workshop, a *museum*, "Those at the Bastille", a nation, *a government*, *a company*.

Properties:

**had member (was member of): Actor**  
property note: Any kind of belonging to a group. May be seen as shortcut of a respective temporal entity

2. The following link completes spatial reasoning, may be more dynamic notions should be added.

E39 Actor:

**has current or former residence (is current or former residence of): Place**

3. The following link seems to be socially relevant:

E21 Person:

**has gender (is gender of): Gender**

#### **Fourth unit:**

This unit deals with identification. The understanding of the group was, that the Appellations of instances of the fundamental categories (Stuff, Actor, Temporal Entity etc.) are not intrinsically different from each other. On the opposite, names are often used for multiple categories by association, e.g. names of saints typically denote churches, icons, communities, and the saint himself. However, there are specific constructs used to name objects of a specific category, like ISBN numbers, geographical coordinates, date expressions, museum numbers etc. This constraint, i.e. that appellation can be used for any entity instance, but a specific subclass of an appellation is only applicable to a specific subentity instance, is not easy to express in an RDF-like semantic model. Therefore we refer to it only verbally.

In the sequence, all properties with the name “is identified by (identifies)” are merged into one. Thereby all entities where the current CRM hasn’t yet foreseen a name can be specifically identified. It must be noted, that any entity instance in the CRM is by itself uniquely identified within the scope of an application, following the object-oriented methodology. The Appellation instances model the often non-unique or ambiguous naming practice in the real world as a **historical fact**, and **NOT** the implementation of the model.

#### **1. E1 CIDOC Entity : is identified by (identifies) : Appellation**

All subclasses of Appellation are meant to gather specific cases, for instance : Place Appellation is an entity gathering naming constructs applicable only to Place instances, however Place instances may quite well have other names as well. Following this thought, "Period Appellation" is deleted as no specific form/format exists for Periods.

2. A specific new entity is introduced for identifiers typical for immaterial items:

#### **E75 Conceptual Object Appellation**

Subclass of: **Appellation**

Scope note: Specific Identifiers of an intellectual product or standardized pattern

Examples: ISBN 3-7913-1418-1, ISO2788-1986 (E)

#### **Fifth unit:**

This unit improves temporal reasoning. The entity Time-Span is numerically determined by two Time Primitive values, which in turn implement intervals. Whereas the properties “begins at”, “ends at” denote the edges of a duration, the Time Primitive expresses an uncertainty. This makes calculus with uncertain dates complex, and needs again two property instances to define a Time-Span instance without explicit duration, i.e. begins at = ends at. On the other hand, a real duration cannot be specified, as the uncertainty of the date can be by far greater than the duration, which may be quite well known. Take e.g. a battle of have a day somewhere in 1204. Therefore we propose to define an **outer bound** of a Time-Span as uncertainty interval, and an **inner bound** as a least certainty interval, if applicable. In addition, two duration bounds are added:

#### E55 Time-Span

New properties:

**at most within: Time Primitive**

property note: the outer bound (ensuring recall)

**at least covering: Time Primitive**

property note: the inner bound, can be omitted

**had at least duration: Dimension**

**had at most duration: Dimension**

Deleted properties:

begins at: Time Primitive

ends at: Time Primitive

This deletion is formally not conservative, as it needs reencoding of instances of the current CRM, the process to do so can however be automated, as the information content is not altered.

### **6<sup>th</sup> Unit:**

This unit deals with Types and other auxiliary concepts. Types actually have a double nature, like designs and procedures. On one side, they act in a mathematical sense to organize our knowledge about the universe around us in “handy” categories, on the other side, they are product of our creativity.

1. As such, it is proposed:

E55 Type

Subclass of: **CIDOC Entity**

This creates the recursion of Types having Types, and adds notes. Alternatively, we could put it even lower under Conceptual Object. Here we may cause some curious inheritance, that “Types” have title etc.

2. The definition of Gender above is a Type not corresponding to an CIDOC Entity, like Material and Language. Following the rule, it must be declared explicitly:

**E76 Gender**

Subclass of: Type

3. The “has note” property of E1 Entity covers numerous facets of descriptions, about parts, history, function etc. If semi-structured textual data are transferred into a CRM instance, such semantics may get lost. Therefore this property should be typed (“subtyped”):

E1 has note : String

**(has type : Type)**