

# Methodological tips for mappings to CIDOC CRM

Maria Theodoridou, George Bruseker, Maria Daskalaki, Martin Doerr

FORTH - Institute of Computer Science {bruseker, martin, maria}@ics.forth.gr daskalakim@hotmail.com



The CIDOC Conceptual Reference Model (ISO 21127:2006) has been chosen as the core model for use in several Cultural Heritage projects



http://www.parthenos-project.eu/



http://www.researchspace.org/



http://wiss-ki.eu/



http://www.ariadne-infrastructure.eu/



http://americanartcollaborative.org/



http://www.itn-dch.eu/



# CIDOC CRM extension suite





# Mapping one schema to another

- Foundational activity
  - Convert data stored in existing (source) schemata to an expression in a form (rdf) compatible with CIDOC CRM and its extensions (target schema)
  - It's not a matching exercise, it's not matching of terms, but interpreting relations
  - It is a consideration of the meaning of the source and a re-expression of that in the target
  - Common point of reference in interpreting source and target as being descriptions of the same reality

#### 🕨 Goal

- Preserve as much as possible the initial 'meaning' of the data
- Enable information exchange and integration between heterogeneous sources of cultural heritage information



OSLO CAA 2016, March 30, 2016



# Mapping & Training Activities

Several mapping activities were initiated trying to convert existing schemata of Cultural heritage data to CIDOC CRM

- ARIADNE: diverse archaeological databases
- Research Space: British Museum and Rijksmuseum data

#### Training activities

- Research Space Workshops in Yale and in Oxford Universities
- ITN-DCH CIDOC CRM board game
- ARIADNE Summer Schools and Transnational Access







CIDOC CRM MAPPING WORKSHOP AT OXFORD UNIVERSITY CIDOC CRM Mapping workshop for humanities scholars and cultural heritage

professionalsInaugural European workshop hosted at University of Oxford e-Research Centre Monday 9th & Tuesday 10th November 2015-Aimed at ... Posted 11 Oct 2015, 05:26 by Dominic Oldman



# General Mapping Principles & Approach



#### Define primary domain

Define primary domain - typically represents the focus of interest Identify the CIDOC CRM class that models it best

- Is there a need for a new Class?
- Do we need properties that are not available in CIDOC CRM?

The introduction of a new class should comply with the "Minimality" modelling principle of CIDOC CRM (CIDOC CRM, 2015):

"A class is not declared unless it is required as the domain or range of a property not appropriate to its superclass, or it is a key concept in the practical scope"



## Define primary domain

Example: Two approaches for defining Coin

- Introduce a specialization of E22 Man-Made Object:
   Exx Coin subclass of E22 Man-Made Object
- Define the Type of E22 Man-Made Object:
   E22 Man-Made Object. P2 has type: E55 Type = "Coin"

To choose we need to answer the question:



Does the new class Coin have new properties that are not available in E22?

Create a new Class, work with it and if it comes out that it has no new properties, revert to its CRM superclass



# Practical Mapping Tips



# Common database fields: Local identifiers

We map local identifiers in relational database tables explicitly only if these identifiers are visible in the user interface and used in other documents as well. Alternatively, we use the local database identifiers only for generating URIs for the record instance, if the thing identified belongs to our holdings or was created/invented by us. Otherwise we better use a global authority.

#### Golden rules for identifiers:

- they should be widely known for the item they identify
- they should be related to those who can confirm the relation between the identifier and the identified (museum object identifiers)



# Common database fields: Local identifiers





# Common database fields: Local identifiers

Whenever possible use a global authority El Greco (Dominico Theotokopoulos)

VIAF:

E39 Actor http://viaf.org/viaf/100215785/

or

GETTY ULAN:

E39 Actor

http://vocab.getty.edu/ulan/500010916



# Common database fields: Appellations

Label and Appellation are alternative implementations in rdf

For contemporary names (not historical) the use of rdfs:label + language attribute is sufficient

We use the E41 Appellation class only if there is need to assign some property to the Appellation



## Common database fields: Yes/No fields

Yes/No fields of relational databases are typically mapped to types.

Table name: Gefaesse (vessels) (E22 Man-Made Object)Field name: GefFragmen: yes/no field

If value of field is yes

then Gefaesse.GefFragment maps to E55 Type fragment





# Common database fields: Implicit and contextual information

All the hidden constants, the common context information of a db, can be added as constant node information at a suitable object type, typically in the domain node.

Quite often semantics are built into

- the User Interface
- > the queries
- values of the identifiers



OSLO CAA 2016, March 30, 2016



## Common database fields: Categorical vs factual info

- Need to separate categorical and factual data Inconsistent information:
  - Find spot -> for a specific coin
  - Historical facts -> for a category of coins

COINS – an archaeological source





## Common database fields: Categorical vs factual info





# People: Groups, Nationality, Origins

Nationality can be modeled as:



membership in an *E74 Group* = *citizenship, social participation* an actor may join and/or leave the group (former, current member)



E55 Type = behavior

suitable for ambiguous, fuzzy characterizations



## People: Groups, Nationality, Origins

Oil on canvas

El Grec

this 1577.



OSLO CAA 2016, March 30, 2016

Art Institute Chicago, Photo M. Doerr, 2010

Fl Greco (Domenico Theotokópoulos)

The Assumption of the Virgin, 1577/79

is Assumption of the Virgin was the central element

in his first major Spanish commission, a complex of paint-

story of a huge retable that filled the end of the sanctuary, the Assumption was flanked by paintings of single saints and crowned by a depiction of the Trinity. Side

altars with pictures of the Resurrection and the Adoration of the Shepherds completed the complex. The space before the high altar contained the tombs of the church's donor,

Doña Maria de Silva, and of Don Diego de Castilla, dean

of the cathedral of Toledo, who commissioned the work.

The triumphant program of the retable reflected their hope of salvation. Greco, who had previously worked on a much smaller scale, proudly attached his signature in Green at the

lower right: Domenikos Theotokopoulo, Cretan

gs for the church of the Cistercian convent of Santo Domingo el Antiguo in Toledo. Set at the center of the lower

Spanish, porn Greece, 1541-1614



#### People: Groups, Nationality, Origins

El Greco's Assumption of the Virgin was the central element in his first major Spanish commission, a complex of paintings for the church of the Cistercian convent of Santo Domingo el Antiguo in Toledo. Set at the center of the lower story of a huge retable that filled the end of the sanctuary, the Assumption was flanked by paintings of single saints and crowned by a depiction of the Trinity. Side altars with pictures of the Resurrection and the Adoration of the Shepherds completed the complex. The space before the high altar

JUNECTIONS < ADOUT THIS ARWORK

#### About This Artwork

#### Domenikos Theotokopoulos, called El Greco Greek, active in Spain, 1541–1614

The Assumption of the Virgin, 1577/79

Oil on canvas

403.2 x 211.8 cm (158 3/4 x 83 3/4 in.); original image, approximate: 396.4 x 202.5 cm (156 1/16 x 79 3/4 in.)

Inscribed on paper at lower right in Greek: (Domenikos Theotokopoulos, Cretan, displayed this in 1577)

Gift of Nancy Atwood Sprague in memory of Albert Arnold Sprague, 1906.99

European Painting and Sculpture Gallery 211



🖂 🔒 🖪 🔽

View enlargement Add to my collection License this image

contained the tombs of the church's donor, Doña Maria de Silva, and of Don Diego de Castilla, dean of the cathedral of Toledo, who commissioned the work. The triumphant program of the retable reflected their hope of salvation. Greco, who had previously worked on a much smaller scale, proudly attached his signature in Greek at the lower right. *Domenikos Theotokopoulos, Cretan, displayed this 1577.*

# Art Institute Chicago

http://www.artic.edu/aic/collections/artwork/87479



Accidental roles are roles that do not characterize an actor independently from a particular context of activity. These roles are properties of the P14 carried out property but this cannot be represented in rdf.



They should **NOT** be assigned to the actor



#### Solution 1: Specializing P14 carried out by for each role

suitable if we know a priory all the possible values





Solution 2: Introducing class PC0 Typed CRM\_Property and its subclass PC14 carried out by and properties P01 has domain and P02 has range





Solution 2: equivalence with CIDOC CRM definition





#### Solution 3: Add a note to each Activity with the Person's role does not require any extension of the model





#### Solution 4: Assign a type to each Activity

does not require any extension of the model





People: Accidental roles - Issuer





#### Political office is modeled as an *E74 Group* with one member

an actor may join and/or leave the group (former, current member)

#### https://en.wikipedia.org/wiki/Alexis\_Tsipras

**Alexis Tsipras** is the 185th and current Prime Minister of Greece, having been sworn in on 21 September 2015. He previously served as the 183rd Prime Minister of Greece from 26 January 2015 to 27 August 2015.

#### E74 Group The Prime Minister of Greece

P144i gained member by E85 Joining P143 joined E39 Actor Alexis Tsipras P4 has time-span E52 Time-Span 26 January 2015
P146i lost member by E86 Leaving P145 separated E39 Actor Alexis Tsipras P4 has time-span E52 Time-Span 27 August 2015
P144i gained member by E85 Joining P143 joined E39 Actor Alexis Tsipras P4 has time-span E52 Time-Span 21 September 2015



Places: Countries

Geopolitical units map to a Period with type "State", etc.

Map them as Place but the id has the name of the country plus a date, when this name of country was assigned.

It is a snapshot that allows to be always in the right time space context.

always refer to a global authority e.g. TGN (falls within, overlaps, is identified by)



"Why hell is not a good place"

**E53 Place** represents always a **real** extent in space independent of time

An imaginary place can be represented as an **E89 Propositional Object** 

The ambiguity between real and imaginary is an ambiguity of the phrase and not an ambiguity in the mind of the author of the phrase (as long as he is sane)

Alternatives should be made explicit: principle of recall over precision



#### How many interpretations of hell?





Byzantine-Orthodox Hell Crete Catholic Hell, Pisa 13<sup>th</sup> century





William Blake's depiction of "The Vestibule of Hell and the Souls Mustering to Cross the Acheron" in his Illustrations to Dante's "Divine Comedy" object 5 c. 1824-27. The original for the work is held by the National Gallery of Victoria

Example taken from https://en.wikipedia.org/wiki/Acheron





Now Italian archaeologists working at the Greco-Roman site of ancient Hierapolis (modern-day Pamukkale) in Turkey have uncovered that city's gate to the underworld. Example taken from http://news.nationalgeographic.com/news/2013/04/130414-hell-underworld-archaeology-mount-olympus--greece/





# Objects: Pieces of a broken object

Pieces of a broken object are mapped as products of an event that destroyed the object and not as part\_of.

- For provenance reasons
- For semantic clarity

As long as the object existed the respective pieces of matter were not distinct, had no identity





# *Objects: Pieces of a set*

#### Pieces of a set are mapped as part of

Chess set has parts: Chess board and chess pieces

Collective objects in the general sense, like a tomb full of gifts, a folder with stamps or a set of chessmen, should be documented as instances of E19 Physical Object, and not as instances of E78 Collection. This is because they form wholes either because they are physically bound together or because they are kept together for their functionality.







#### Always consult the reference document Definition of the CIDOC Conceptual Reference Model http://www.cidoc-crm.org/comprehensive\_intro.html

Read carefully the scope note of the class and property that you intend to use. Learning to use a formal ontology requires understanding the basic principles expressed in its scope, classes and relations

#### http://www.cidoc-crm.org/index.html

New site (available for testing) http://new.cidoc-crm.org/



#### Conclusions

#### IF



# **ASK!**



# Thank you!

Maria Theodoridou maria@ics.forth.gr