One ring to rule them all
CIDOC CRM

Alan Poulter
CIG Conference 2010
Overview

• The Google ‘problem’ – how to create an alternative to a large collection of unstructured information
• An explanation of CIDOC CRM
• How CIDOC CRM has been applied to developing FRBR
• Future prospects
Intro

• There is increasing provision of high quality digital cultural content from trusted information providers such as museums, libraries and archives.

• Unlike much of the information available on the Web, digital cultural content from so-called "memory institutions" is typically authoritative, high quality, useful for education and research and has broad appeal to a wide range of audiences.
Intro

• But the information managed by museums, libraries and archives can vary according to factors such as:
  • Type of institution
  • Professional approaches
  • Academic discipline
  • Granularity of description
  • Level of detail of description
  • Descriptive data structures
  • Descriptive data content values
• Differences in descriptive schema across museums, libraries and archives can seriously hinder cross-domain discovery and interoperability of cultural information resources.

• Fortunately, differences are primarily at the level of data structure and syntax. Significant common conceptual elemental concepts exist: objects, people, places, events, and the interrelationships between them.

• Solution is to map everything to a simple schema with broad and universal semantics for the purposes of initial resource discovery (e.g. Dublin Core) – but lose richness/structure.
CIDOC Conceptual Reference Model

- CIDOC Conceptual Reference Model (hereafter CRM) is an object-oriented domain ontology for exchanging rich cultural heritage data.
- ISO standard 21127: 2006
- Class hierarchy of 81 named classes, interlinked by 132 named properties.
- Actors participate in Temporal Entities (e.g. events), which are affected by Physical Entities (i.e. things) and Conceptual Objects (i.e. ideas), and occur at Places within certain Time-Spans. Appellations (i.e. names) can be used to identify entities, and Types can be used to classify them.
CIDOC Conceptual Reference Model
CIDOC CRM Example

Type: Text
Title: Protocol of Proceedings of Crimea Conference II. Declaration of Liberated Europe
Date: February 11, 1945
Creator: The Premier of the Union of Soviet Socialist Republics
         The Prime Minister of the United Kingdom
         The President of the United States of America
         State Department
Subject: Postwar division of Europe and Japan

About...

"The following declaration has been approved: The Premier of the Union of Soviet Socialist Republics, the Prime Minister of the United Kingdom and the President of the United States of America have consulted with each other in the common interests of the people of their countries and those of liberated Europe. They jointly declare their mutual agreement to concert... and to ensure that Germany will never again be able to disturb the peace of the world...... "

Documents
CIDOC CRM Example

Type: Image
Title: Allied Leaders at Yalta
Date: 1945
Publisher: United Press International (UPI)
Source: The Bettmann Archive
Copyright: Corbis
References: Churchill, Roosevelt, Stalin

Metadata

About...

Photos, Persons
CIDOC CRM Example

**TGN Id:** 7012124  
**Names:** Yalta (C,V), Jalta (C,V)  
**Types:** inhabited place(C), city (C)  
**Position:** Lat: 44 30 N, Long: 034 10 E  
**Hierarchy:** Europe (continent) ← Ukrayina (nation) ← Krym (autonomous republic)  
**Note:** ...Site of conference between Allied powers in WW II in 1945; ....  
**Source:** TGN, Thesaurus of Geographic Names

**Places, Objects**

**Title:** Yalta, Crimean Peninsula  
**Publisher:** Kurgan-Lisnet  
**Source:** Liaison Agency
CIDOC CRM Example

- **Actors, Roles, proper names:**
  - The Premier of the Union of Soviet Socialist Republics
  - Allied leader, Allied power
  - Joseph Stalin….

- **Places**
  - Jalta, Yalta
  - Krym, Crimea

- **Events**
  - Crimea Conference, “Allied Leaders at Yalta”, “… conference between Allied powers” “Postwar division”

- **Objects and Documents:**
  - The photo, the agreement text
CIDOC CRM Example

- E39 Actor
- E52 Time-Span
  - February 1945
- E53 Place
  - 7012124
- E58 Image
- E38 Image
- E31 Document
  - “Yalta Agreement”
- E7 Activity
  - “Crimea Conference”
- E65 Creation
- P67 is referred to by
- P86 falls within
- P81 ongoing throughout
- P11 participated in
- P7 took place at
- P82 at some time within
- P94 has created
- P14 performed
- 1945-02-11
CIDOC Conceptual Reference Model
CIDOC CRM Example

Epitaphios GE34604 (entity E22 Man-Made Object)
P30 custody transferred through, P24 changed ownership through
Transfer of Epitaphios GE34604 (entity E10 Transfer of Custody, E8 Acquisition Event)
P28 custody surrendered by
   Metropolitan Church of the Greek Community of Ankara (entity E39 Actor)
P23 transferred title from
   Metropolitan Church of the Greek Community of Ankara (entity E39 Actor)
P29 custody received by
   Museum Benaki (entity E39 Actor)
P22 transferred title to
   Exchangeable Fund of Refugees (entity P40 Legal Body)
P2 has type
   national foundation (entity E55 Type)
P14 carried out by
   Exchangeable Fund of Refugees (entity E39 Actor)
P4 has time-span
   GE34604_transfer_time (entity E52 Time-Span)
P7 took place at
   Greece (entity E53 Place)
P2 has type
   nation (entity E55 Type)
P2 has type
   republic (entity E55 Type)
P89 falls within
   Europe (entity E53 Place)
P2 has type
   continent (entity E55 Type)
FRBR: the conceptual model for libraries

• “FRBR” is for “Functional Requirements for Bibliographic Records”
• Developed 1991-1997 & published 1998 by IFLA (International Federation of Library Associations and Institutions)
• Maintained by the IFLA FRBR Review Group
Key concepts of FRBR

- Concept
- Object
- Event
- Place
- Person
- Corporate Body
- Work
- Expression
- Manifestation
- Item
FRBR/CRM Harmonisation Group

• Formed 2003
• Representatives from:
  – the IFLA FRBR Review Group
  – the CRM Special Interest Group
• Aims:
  – To reach a common view of cultural heritage information
  – To check FRBR’s internal consistency
• Output: FRBRoo
FRBRoo Works

- E28 Conceptual Object
- F1Work
  - F46 Individual Work
  - F21 Complex Work
  - F48 Container Work
    - F43 Publication Work
  - F22 Serial Work
Individual Work

• subclass of Work:
  – “This class comprises works that are realised by one and only one self-contained expression, i.e., works representing the concept as expressed by precisely this expression, and that do not have other works as parts.”
  – Example: Abstract content of Giovanni Battista Piranesi’s ‘Carcere XVI: the pier with chains: 1st state’
Complex Work

• subclass of Work:
  – “can be taken up by any creator who acquires the spirit of its concept, it is never finished in an absolute sense.”
  – Example: Work entitled ‘Hamlet’ by William Shakespeare
Container Work

• subclass of Work:
  – “comprises works whose essence is to enhance or add value to expressions from one or more other works without altering them, by the selection, arrangement and/or addition of features of different form.”
  – Example: The concept for the layout created by printer Guido Morris for the text of Michael Hamburger’s English translation of 12 poems by Georg Trakl for publication in 1952 [an Publication Work]
Publication Work

• subclass of Container Work:
  – “This class comprises works that have been planned to result in a manifestation product type and that pertain to the rendering of expressions from other works.”
  – Example: The concept of publishing Stephen Crane’s complete poems (as edited by Joseph Katz), which includes the idea that every time a stanza jumps over a page change, the statement ‘[NO STANZA BREAK]’ should be printed as a warning for readers that the new page continues the same stanza
Serial Work

• subclass of Complex Work, Publication Work:
  – “This class comprises works that are, or have been, planned to result in sequences of manifestations with common features.”
  – Example: The periodical entitled ‘The UNESCO Courier’, ISSN ‘0041-5278’
FRBRoo Works

- E28 Conceptual Object
- F1Work
  - F46 Individual Work
  - F21 Complex Work
  - F48 Container Work
    - F43 Publication Work
  - F22 Serial Work
FRBRoo Expressions

- E73 Information Object
  - F2 Expression
    - F20 Self-contained Expression
    - F41 Publication Expression
    - F23 Expression Fragment
Self-contained Expression

• subclass of Expression:
  – “comprises the immaterial realisations of individual works at a particular time that are regarded as a complete whole. The quality of wholeness reflects the intention of its creator that this expression should convey the concept of the work.”
  – Example: The Italian text of Dante’s ‘Inferno’ as found in the authoritative critical edition *La Commedia secondo l’antica vulgata a cura di Giorgio Petrocchi*, Milano: Mondadori, 1966-67
Publication Expression

• subclass of Self-Contained Expression:
  – “comprises complete sets of signs present in publications, reflecting publishers’ final decisions as to both content and layout of the publications”
  – Example: The overall content of the book identified by ISBN ‘0-8014-9130-4’:
Expression Fragment

• subclass of Expression:
  – “This class comprises parts of Expressions and these parts are not Self-contained Expressions themselves”
  – Example: The only remnants of Sappho’s poems
FRBRoo Expressions

1. E73 Information Object
   - F2 Expression
     - F20 Self-contained Expression
     - F41 Publication Expression
     - F23 Expression Fragment
FRBRoo Manifestations

E55 Type

F3 Manifestation Product Type

E84 Information Carrier
Item

E24 Physical Man-Made Stuff
Manifestation Singleton

immaterial

material
Manifestation Product Type

- “This class comprises the definitions of publication products.”

- Example: The publication product containing the cartographic resource titled ‘Ordnance Survey Explorer Map 213, Aberystwyth & Cwm Rheidol’, issued in May 2005 by the publisher named ‘Ordnance Survey’ and identified by ISBN ‘0-319-23640-4’ (folded), 1:25,000 scale
Manifestation Singleton

– “This class comprises physical objects that each carry an instance of Expression, and that were produced as unique objects, with no siblings intended in the course of their production”

– Example: The manuscript known as ‘The Book of Kells’
FRBRoo Manifestations
FRBRoo Publishing example

Author's Level

- F20 Self-Contained Expression
  - "Oliver Twist" Text

Publisher's Level

- F43 Publication Work
  - "Oliver Twist" Publication Concept
  - R56 is realized in
    - R11B forms part of
      - P128B is carried by
        - F46 Individual Work
          - "Oliver Twist" Completed

- F45 Publication Event
  - "Oliver Twist" published!

- F41 Publication Expression
  - "Oliver Twist" Publication content
  - CLR5 should carry
    - F3 Manifestation-Product Type
      - "Oliver Twist" Publication
      - P128B is carried by
        - F39 Production Plan
          - "Oliver Twist" Plan
          - R38 produced things of type
            - R40 Carrier Production Event
              - "Oliver Twist" Printing

- F40 Carrier Production Event
  - "Oliver Twist" Printing
  - R41 produced (was produced by)
  - R39B was followed by
    - R55 created production plan
      - P128B is carried by
        - F4 Individual Work
          - "Oliver Twist" Completed

- F5 Item
  - "Oliver Twist" 20th Book
Conclusions

• FRBR has been improved using CRM
• CRM has great potential as a universal model for all types of ‘information’
• Can a ‘structured’ search engine be the end result of better/unified models/descriptions?
Acknowledgements


Questions?