CRM extensions

$\text{FRBR}_{\text{OO}} \text{ and CRMI}_{\text{nf}}$

DiXiT Farewell Party

Martin Doerr

Center for Cultural Informatics
Institute of Computer Science
Foundation for Research and Technology - Hellas

Heraklion, Crete, Greece
March 28, 2017
Extending the CRM

- The CRM standardizes only **stable** concepts for information **sharing**.
  - local extensions are encouraged for subjective concepts and local practices
  - using the CRM starts with 1 property and does **not restrict** data to CRM

- We have now created a **modular** structure
  - Maintaining a core so that all extensions are (property) specializations
  - All more detailed facts can be reached by querying core concepts
  - For being interoperable, **no more restriction** of data to a “core vocabulary”!

- **What is “core” is not** historical, not community domination, but the **dynamic** result of applying **functional** principles.

- CRM is an **open invitation to extend** it by sharing, respecting and evolving common concepts:
  - **The CRM becomes an open “family of models”**
**Outcome: CRM compatible Extensions**

- **FRBRoo**: modelling the new library practice of IFLA (approved)
  - a causal model of intellectual creation and derivation
  - how to identify intellectual content
  - the thing and the word: integrating museum and library perspectives

- **PRESSoo**: modelling journals and serials (approved)

- **CRMInf**: who said that? – from data to knowledge (under review)
  - integrating data with their scholarly justification
  - being validated with scholarly annotations
Outcome: CRM compatible Extensions

- **CRMsci**: a Scientific Observation model (under review)
  - generalizes over INSPIRE, OBOE, SEEK, Darwin Core
  - generalizes concepts of units of matter and their “(physical) genesis”
  - introduces concept of observation and data evaluation
  - validated in archeology, biodiversity and geology

- **CRMarachaeo/CRM_{BA}**: an Excavation model (under review)
  - introduces concepts of stratigraphy and excavation
  - being validated by archaeological records
CRMgeo: a Spatiotemporal model (to be reviewed)
- integrates CRM with OGC standards
- a complete model of phenomena occupying spacetime (consistent with modern physics)
- integrates geometry- and semantics-derived topological relations
- core concepts being integrated into CRM

CRMdig: a model of Digitization processes (to be reviewed)
- validated in European & US projects, to be adapted to CRMsci

...give us your extension for review and approval!
CIDOC CRM extension suite

CIDOC
Conceptual Reference Model (CRM)

Few concepts, high recall

Special concepts, high precision

CIDOC June 10, 2012
FRBR_{oo}:

“who’s idea was that?”

“do you have a translation of…?”
The Functional Requirements for Bibliographic Records (FRBR)

- developed 1992-1997 by IFLA (now being complemented by the Functional Requirements for (Subject) Authority Data (FRAD/FRSAD))
- A core ER model to integrate library objects by content relation: 4+6 Entities, 97 attributes, 34 relationships
- Based on new requirement to make derivation/intellectual relations/equivalences accessible to bibliographic access/search?
- Might result in a new library practice

Innovations:

- Definition of stages/abstraction levels of intellectual products: Work, Expression, Manifestation, Item (precise distinctions debated for many years).
- Clusters publications and items around the notion of derivation and common conceptual origin across stages/abstraction levels.

Lacks: any explicit notion of the processes behind.
“a distinct intellectual or artistic creation… there is no single material object one can point to as the work...”

“the intellectual or artistic realization of a work in the form of alpha-numeric, musical, or choreographic notation, sound, image, object, movement, etc”

“the physical embodiment of an expression of a work...all the physical objects that bear the same characteristics... ...may be only a single physical exemplar...”

“a single exemplar of a manifestation...”
Conception of FRBR_{OO}:

- **Scope of FRBR_{OO}**:  
  - processes creating and mediating content,
  - evidence of content,
  - inferred abstraction of content (expression => work!)

- **Innovations**:  
  - Theory of identifying immaterial objects
  - Connecting material and intellectual production (“externalization”)
  - Explicating the publication work (publishing houses, graphics designers, illustrators)
  - Coherent theory of performances, content and recordings
  - A new intellectual content relationship: “incorporation”!
  - Adaptation of CRM:
    - (Information Object => Propositional /Symbolic divide)
    - Identifier definition & construction (librarians have the best models of identifiers!)
The Process View: “Externalization”

**F1 Work**
- **F27 Work Conception**
  - **F28 Expression Creation**
    - **F4 Manifestation Singleton**

**E39 Actor**
- **E52 Time**
- **E53 Place**

**Work elaboration**

**Work conception**
- **produces an idea**

**Expression creation**
- **produces a work**
  - **produces (simultaneously) an Expression and a Manifestation-Singleton**

CIDOC June 10, 2012
Publication Work - An “Identity Crisis”

Publisher’s Level

F19 Publication Work
“Oliver Twist” Publication Concept
R3 is realized in (realises)

F24 Publication Expression
“Oliver Twist” Publication content
CLR6 should carry (should be carried by)

F3 Manifestation-Product Type
“Oliver Twist” Publication
R7 is example of (has example)

F5 Item
“Oliver Twist” 20th Book

F30 Publication Event
“Oliver Twist” published!
R23 created a realization of (was realized through)

F32 Carrier Production Event
“Oliver Twist” Printing
R26 produced things of type (was produced by)
R28 produced (was produced by)
R27 used as source material (was used by)

Author’s Level

F15 Complex Work
“Oliver Twist” Ideas
R10 has member (is member of)

F14 Individual Work
“Oliver Twist” Version
R3 is realized in (realises)

F22 Self-Contained Expression
“Oliver Twist” Text
R14 incorporates (is incorporated in)

F4 Manifestation Singleton
“Oliver Twist” Manuscript

P128 carries (is carried by)

CIDOC June 10, 2012
Performing Arts: An “Added Value” Chain

CIDOC June 10, 2012
Example according to FRBR\textsubscript{OO} by Patrick LeBoeuf
CRMinf:

“why is it true that?”

A model for Annotation, Argumentation and Provenance Documentation
The Three Sources of Scientific Knowledge

**Here is CRMInf:**
- Observation
- Data Evaluation
- Simulation

**Here is CRMSsci:**

**CRMSsci:**

Knowledge from observation, data evaluation and (computer)simulation

(engineered from OBOE, SEEK, INSPIRE Darwin Core etc)
Simplified Inference Making

I4 Proposition Set

I2 Belief

J2 concluded that

I5 Inference Making

E39 Actor P14 carried out by

I3 Inference Logic

I6 Belief Value

J4 that

J5 holds to be

J1 used as premise

J3 applies

can be “common sense”, “feeling”…
Modelling beliefs about old inventory numbers

Hans Sloane collection inventory entry
Proposed Epistemological Core Model of Sciences

Categorical Models
(terminologies)

Possible Worlds
and mechanisms
of processes

Collective world view

Collective material behavior
Collective psychological behavior

Factual world

Material world (CRM)
Socio-Psychological world

Observable world
Non-discrete unlimited and unclassified material phenomena

Reality

Reasoning are also real world activities (CRM)!!