

# THE FUNCTION AND USE OF CIDOC CRM AND ITS EXTENSIONS

---

George Bruseker  
Zurich,  
Anais Guillem

(Digital Society Initiative/Univ. of  
ETH Zurich, Takin.solutions)  
(UC Merced, ETH Zurich)

DONIPAT

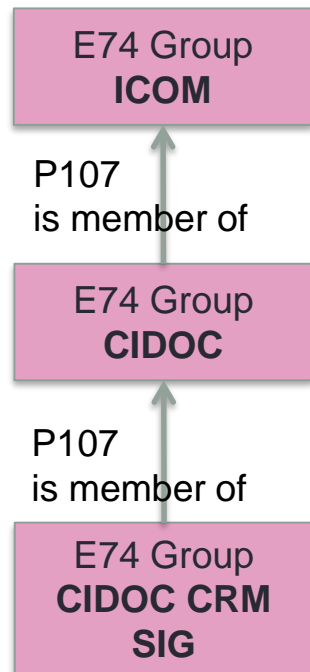
# Table of Contents

1. Background and Intended Function
2. Method of Modelling
3. Event-Based Modelling
4. Family Models
5. Tools, Implementation and More

# 1. BACKGROUND AND INTENDED FUNCTION

---

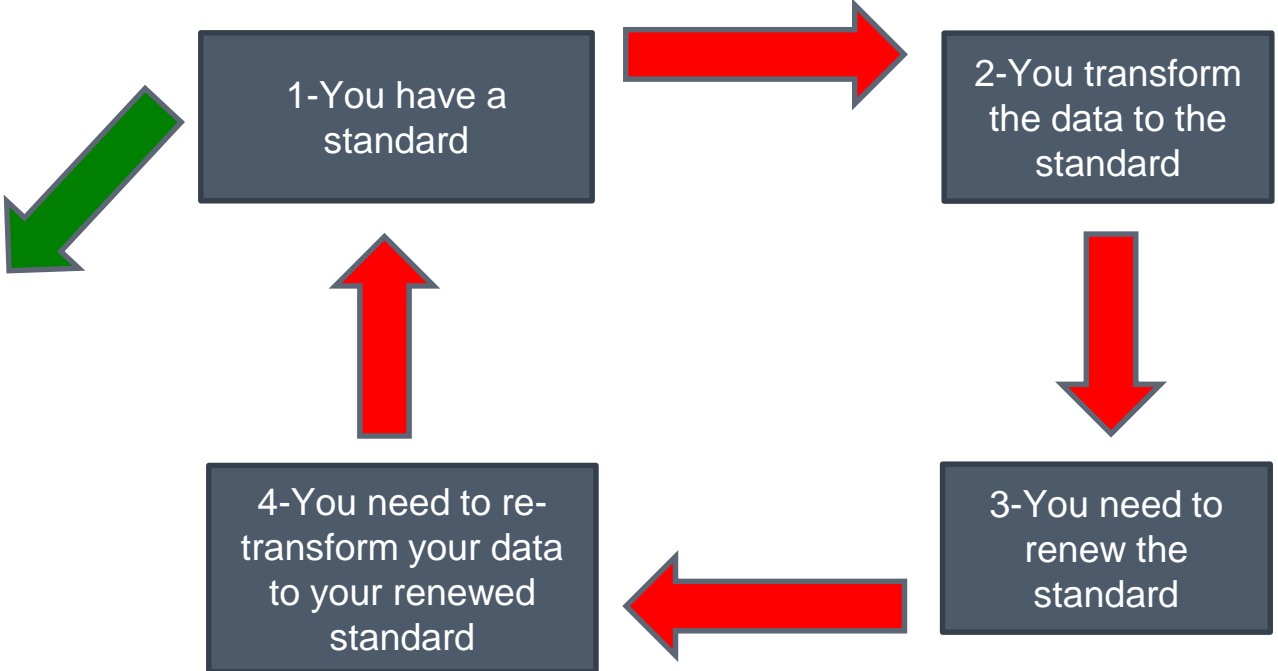
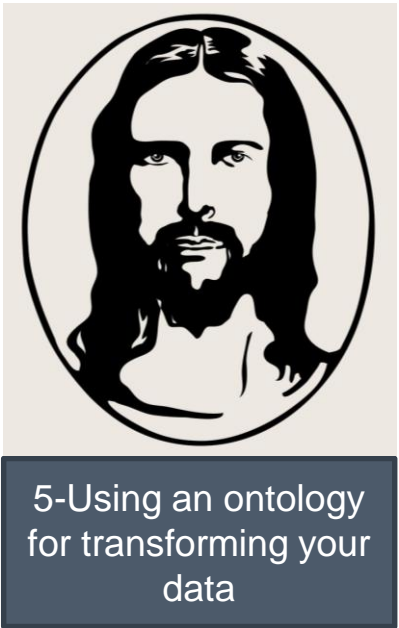
# The CIDOC Conceptual Reference Model



- a **core ontology** describing the underlying semantics of over a hundred database schemata and structures from all museum disciplines, archives and libraries.
- Recognized **ISO Standard** since 2006 (ISO21127:2014)
- the **result of 20 years** of interdisciplinary work and (dis)agreement
- a **generic model of recording of “what has happened”** in human scale
- **can generates huge, meaningful networks of knowledge by a simple abstraction: history as meetings of people, things and information.**

# Standards, Mapping and Data Transformation

## Making Standards

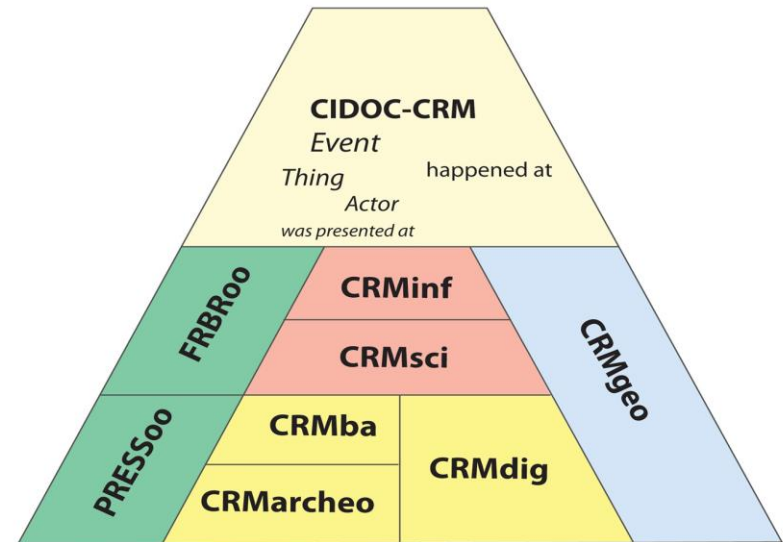


# CIDOC CRM: Description

Type	Top Level Ontology
Scope	Cultural Heritage and E-Sciences
Classes	90+-
Relations	150+-
Version	6
Maintained by	CIDOC CRM SIG
Official Extensions	8
Access	<a href="http://www.cidoc-crm.org/">http://www.cidoc-crm.org/</a>



*CIDOC-CRM family of Models*



# And standardized schemas?

## A Standard Schema

*Standardized schema proposed to solve problems of schema irregularity*

- Serve to give form to bodies of data/metadata
- Prescribes a practice
- Different standards for different objects
- Different standards for the same object
- Choice is pragmatic, focussed on goals and always local



# And standardized thesauri/vocabularies?



## A standard vocabulary

*Standardized thesauri and vocabs proposed to solve problem of data value and reference irregularity*

- Provides authoritative hierarchical terms and reference
- Offers alternative, preferred and alternative forms of term
- Prescribes a practice
- Domain focused
- Irreducible

# How ontologies, schemas and thesauri fit

The right solution  
for the right problem

- 1) A common agreed method of Recording some specific thing = **standard schema**
- 2) A common agreed set of terms/references for naming/classifying some specific thing = **thesauri**
- 3) A common way to integrate data expressed in heterogeneous schemas = **ontology**

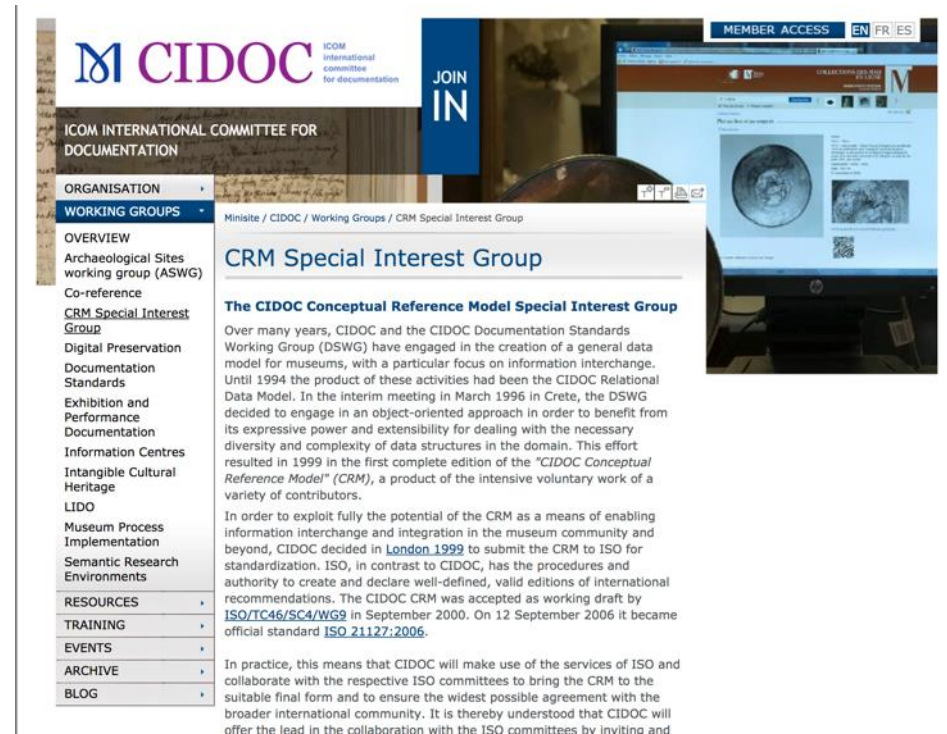


## 2. Method and Goal of Ontology Production

---

# The CIDOC CRM SIG

- Meets 3 - 4 times a year
- Undertakes to support the maintenance, development and extension of the CIDOC CRM Standard
- Made up of experts from different fields
- Maintains continuity through institutional membership
- Works through issue reporting on an email list / official website
- Decisions taken according to democratic vote after presentation / consultation



The image shows a screenshot of the CIDOC CRM SIG website. The header features the CIDOC logo (a stylized 'M' followed by 'CIDOC' and 'ICOM International Committee for Documentation') and a 'JOIN IN' button. Below the header is a navigation menu with categories: ORGANISATION, WORKING GROUPS, OVERVIEW, ARCHIVE, and BLOG. The 'WORKING GROUPS' category is expanded, showing a list of groups including 'CRM Special Interest Group'. The main content area is titled 'CRM Special Interest Group' and contains the following text:

**The CIDOC Conceptual Reference Model Special Interest Group**

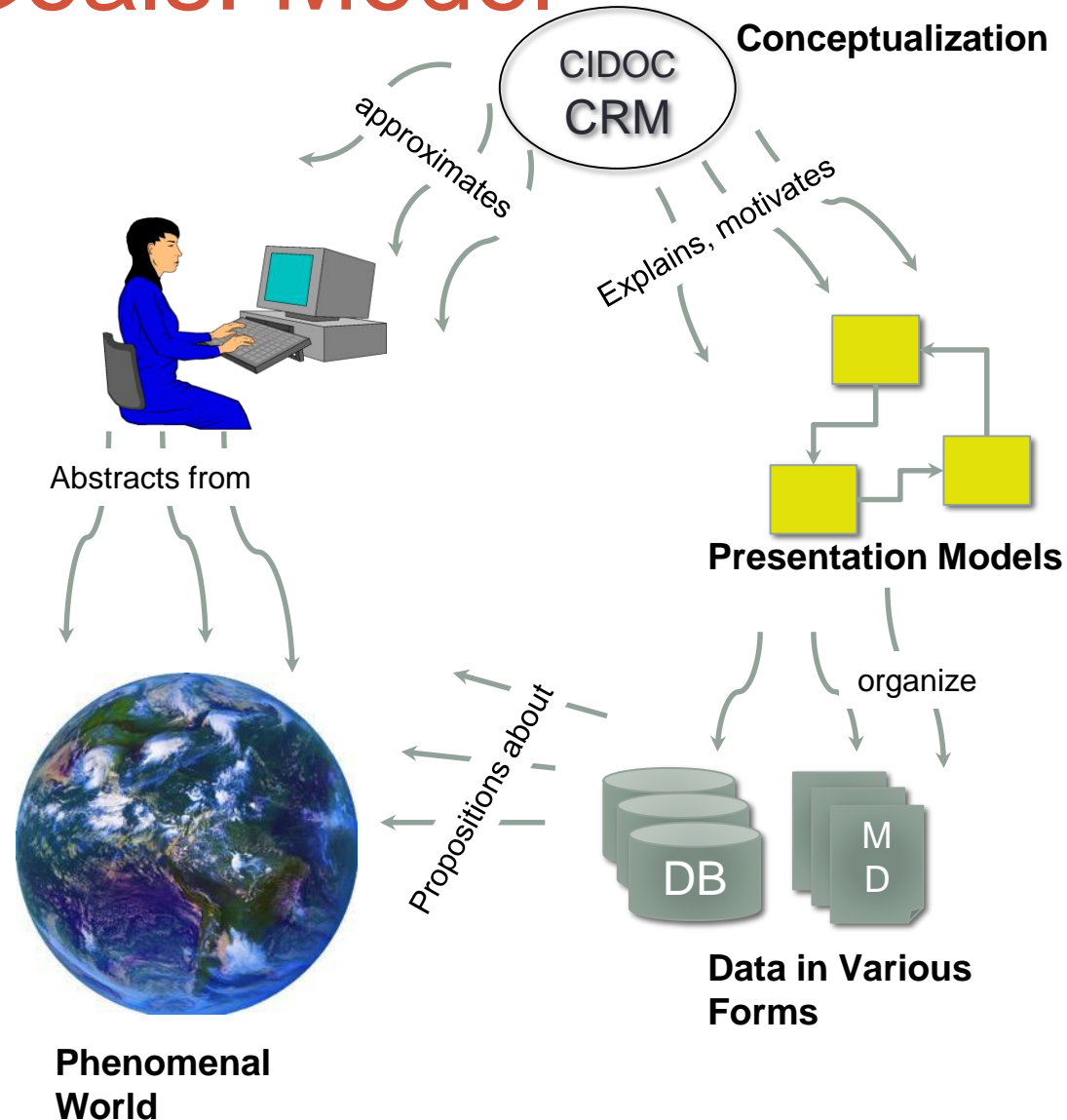
Over many years, CIDOC and the CIDOC Documentation Standards Working Group (DSWG) have engaged in the creation of a general data model for museums, with a particular focus on information interchange. Until 1994 the product of these activities had been the CIDOC Relational Data Model. In the interim meeting in March 1996 in Crete, the DSWG decided to engage in an object-oriented approach in order to benefit from its expressive power and extensibility for dealing with the necessary diversity and complexity of data structures in the domain. This effort resulted in 1999 in the first complete edition of the "CIDOC Conceptual Reference Model" (CRM), a product of the intensive voluntary work of a variety of contributors.

In order to exploit fully the potential of the CRM as a means of enabling information interchange and integration in the museum community and beyond, CIDOC decided in London 1999 to submit the CRM to ISO for standardization. ISO, in contrast to CIDOC, has the procedures and authority to create and declare well-defined, valid editions of international recommendations. The CIDOC CRM was accepted as working draft by [ISO/TC46/SC4/WG9](#) in September 2000. On 12 September 2006 it became official standard [ISO 21127:2006](#).

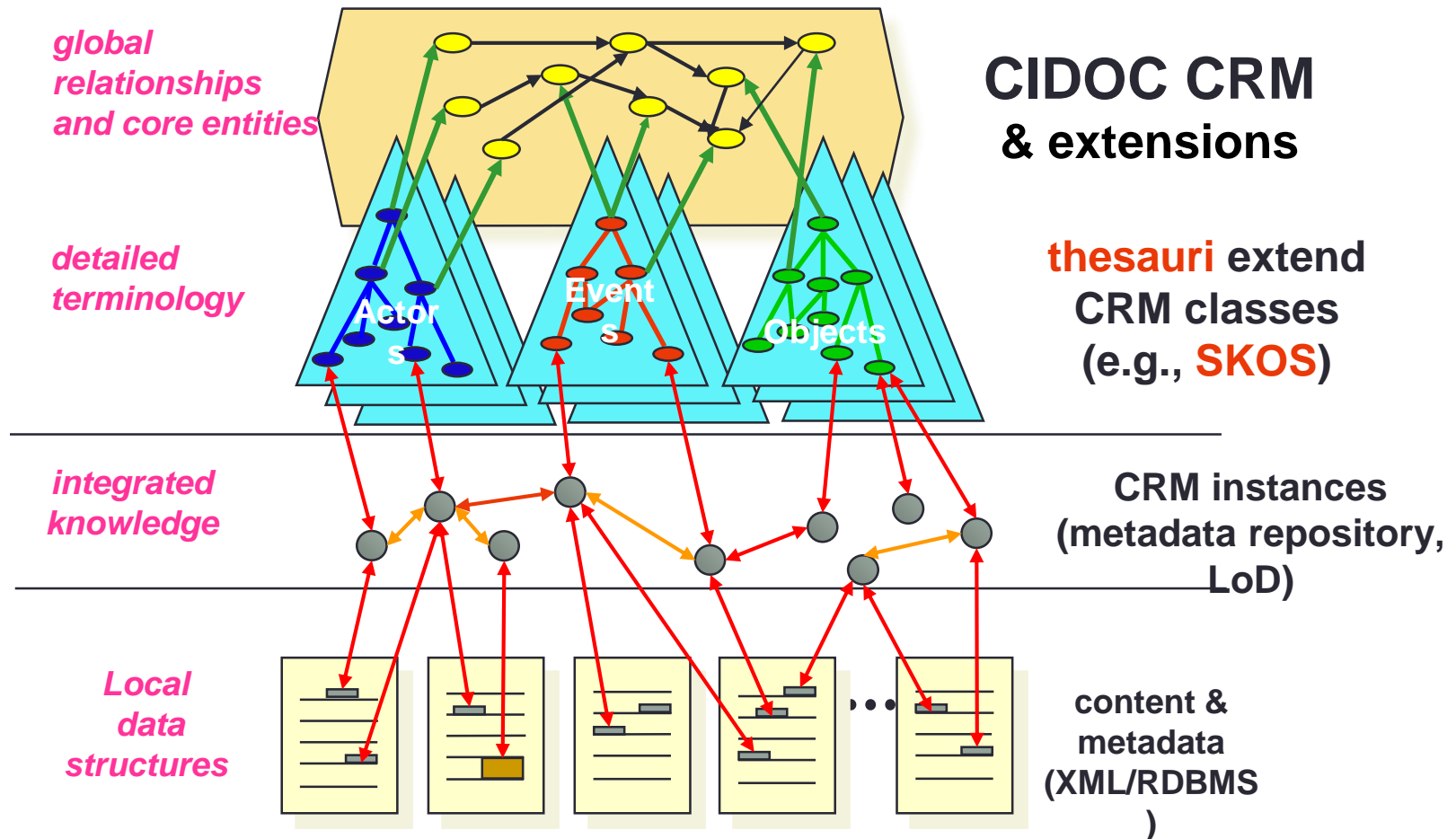
In practice, this means that CIDOC will make use of the services of ISO and collaborate with the respective ISO committees to bring the CRM to the suitable final form and to ensure the widest possible agreement with the broader international community. It is thereby understood that CIDOC will offer the lead in the collaboration with the ISO committees by inviting and

# CIDOC CRM Goals: Model

- Create a conceptual model by which multiple heterogeneous data formats can be translated /mapped to each other
- Support high level information recall
- Monotonic Extensibility
- Evidence based, not by fiat



# CIDOC CRM Goals: Data



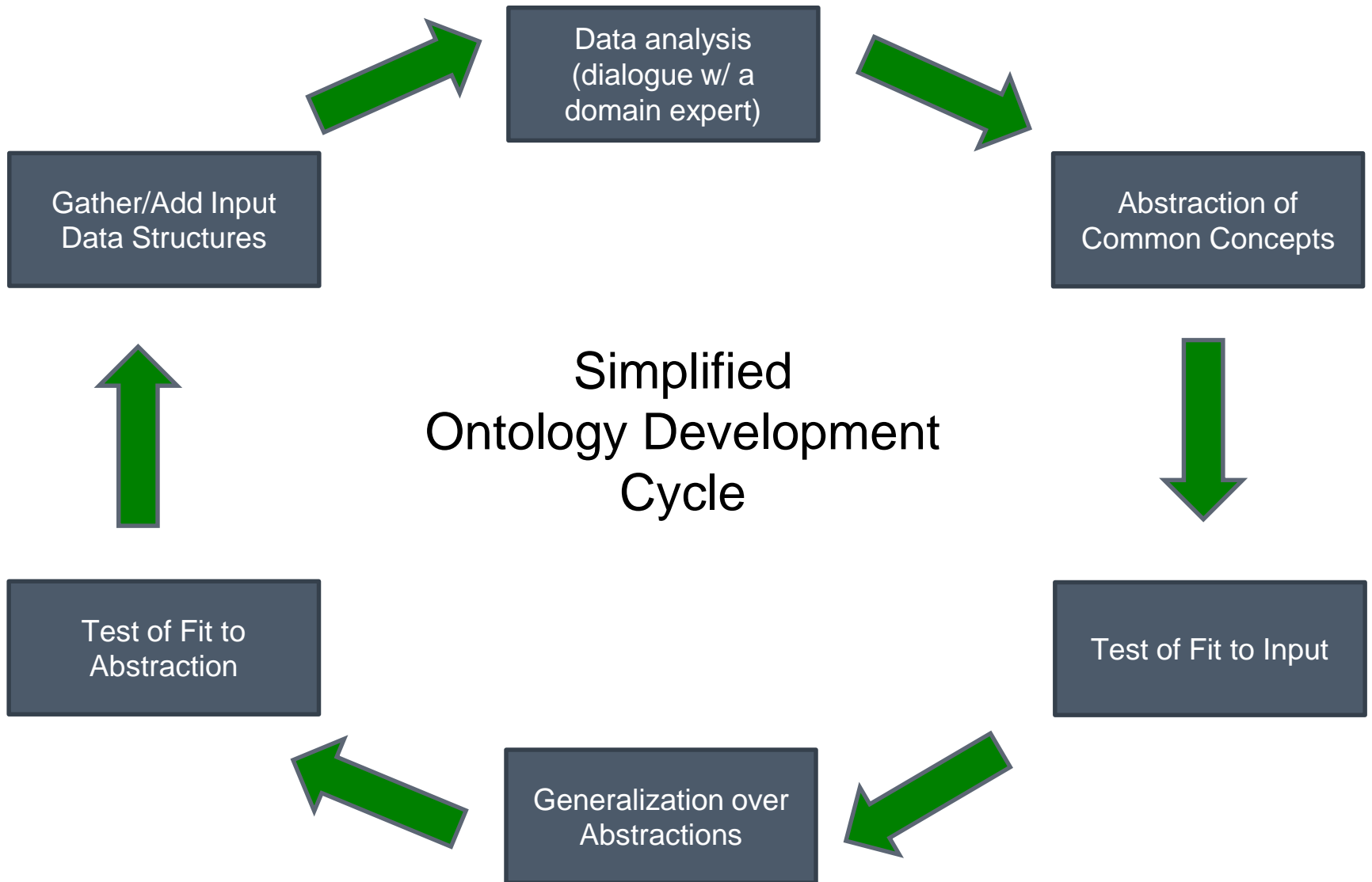
- Support interoperability of mutually relevant data sources
- enable sourced and verifiable facts from datasets
- foster referenceability and reusability of data
- foster structured argumentation on top of facts

# CIDOC CRM: How does it come about?

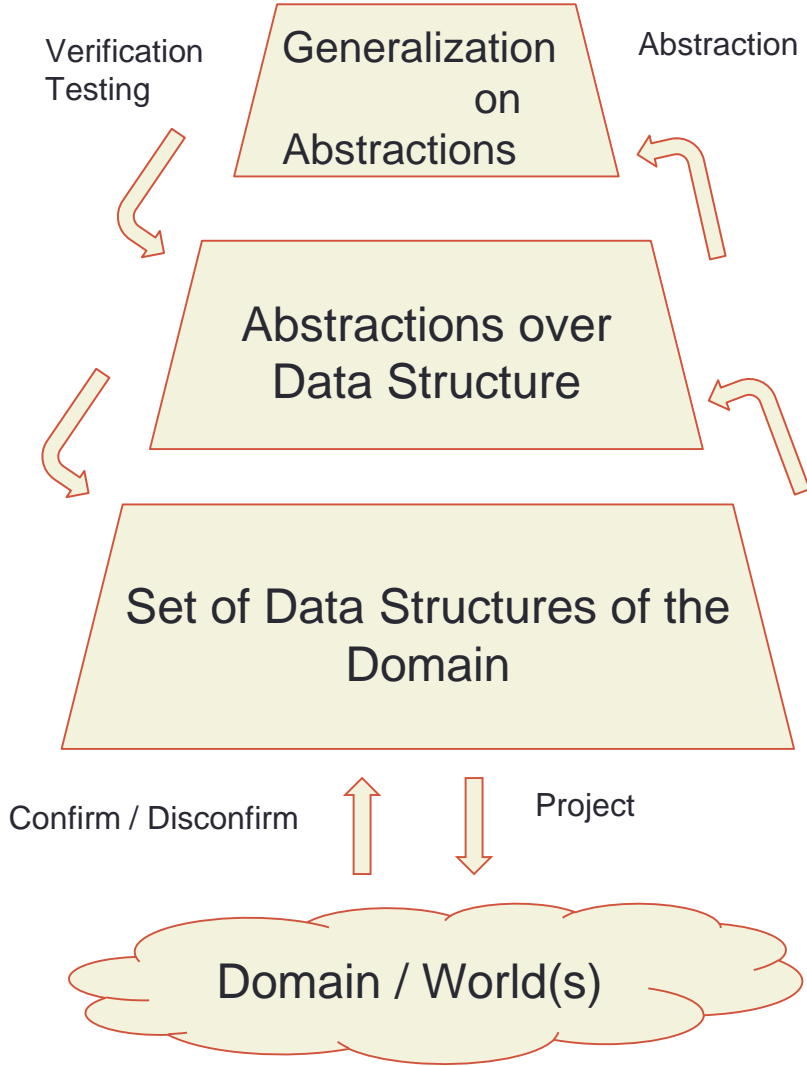
- Empirical study of data structures, to understand their semantic content
- Dialogue with domain specialists, to test conceptualizations, understand argumentation
- Elicitation of competence questions, to have a metric against which to measure the success of the effort



# The process: iterative abstraction/harmonization



# The process: testing against objective domain

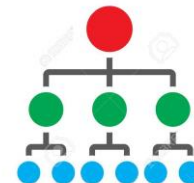


# What can I do with it?



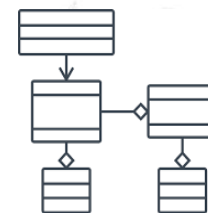
## Ontology Extension

- Extend the CRM



## Conceptual Model Creation

- Mind map model of existing or projected data
- Adopt CRM to express your model, harmonize [maybe extend CRM]



## Mapping

- Transformation
- Query Map

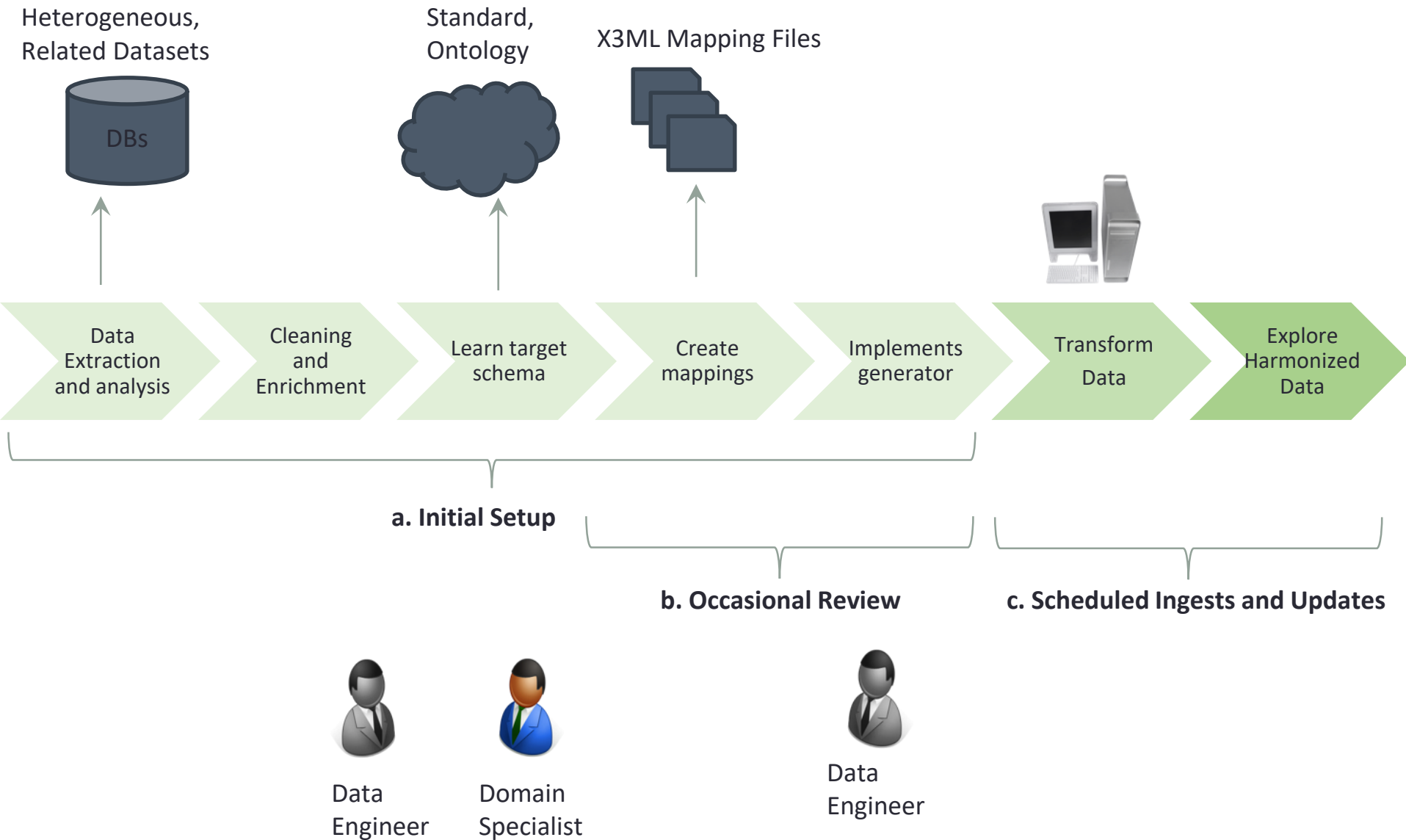


## Semantic Knowledge

- Create data models and forms
- Generate born semantic data



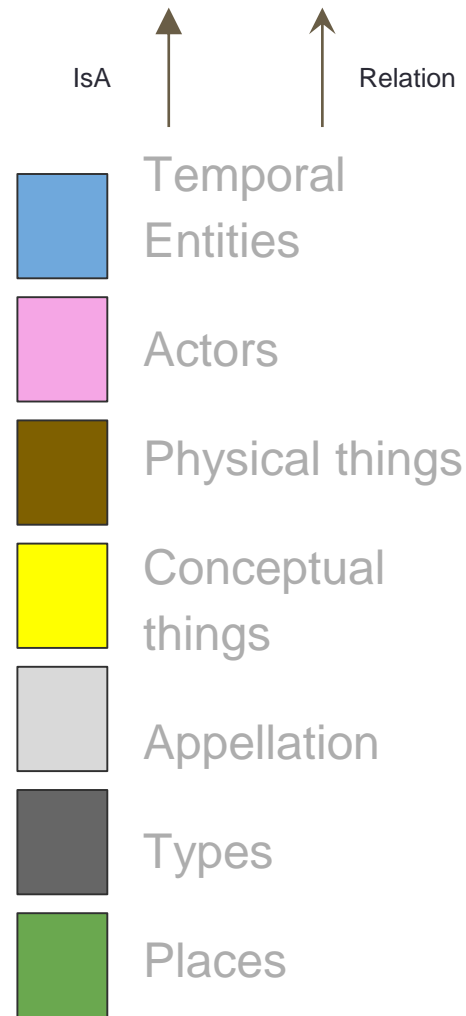
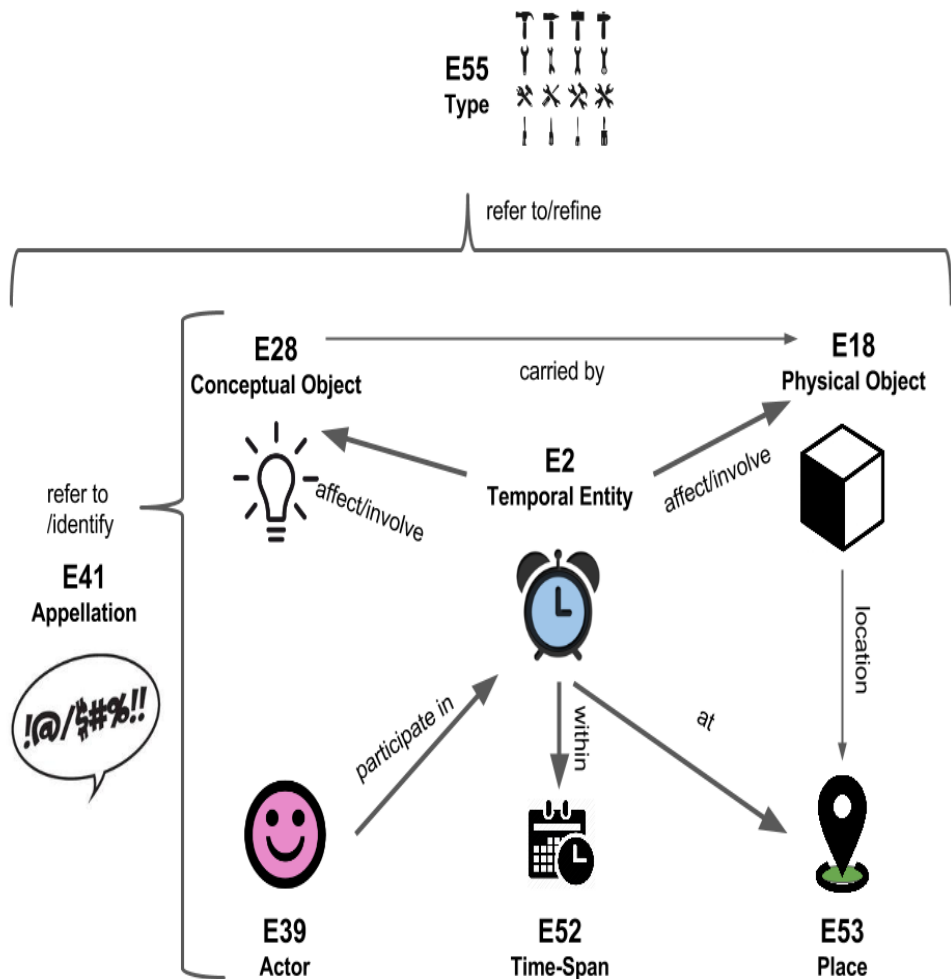
# Workflow



# 3. EVENT BASED MODELLING

---

# CIDOC CRM: General Modelling Pattern



# Event-Centric Modelling Motivation

- Search for a means to **represent** the ontology of things we **experience** and know that allow us to reconstruct a possible past
- Empirical sciences begin from **primary data**, objects (physical or textual) **are traces of past events**
- **We seek to know what happened** (the event) through the traces of the past in the present
- The object points us towards the events that brought it into existence, modified it and/or led to its destruction
- **Events** are the most **powerful tool for inference**, knowing that something came to be and its position relative to another event allows powerful understanding of the past
- Event modelling **does not require complete knowledge** and explicitly makes room for revision



# Terracotta Warriors Object Evidence

Field	Value
Type	Object
Title	Terracotta Warrior
Date	210 to 209 BCE
Creator	Qin Shi Huang (秦始皇)
Place	Xianyang
Publisher	Xi'an Terracotta Army Museum
References	Ssu-ma Chien

metadata  
for

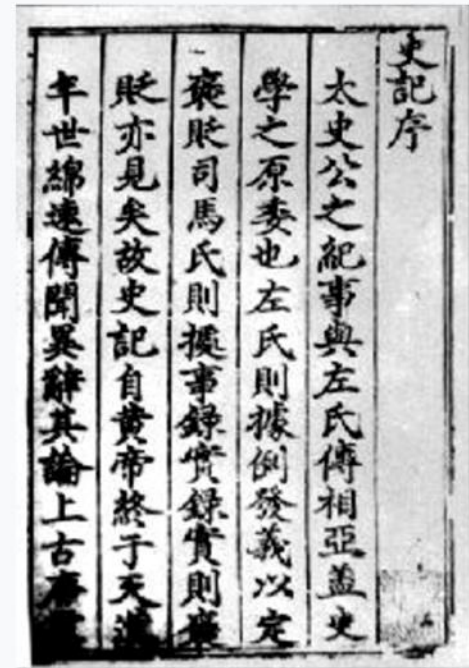


Object

# Terracotta Warriors Textual Evidence

Field	Value
Type	Document
Title	Shǐjì 史記
Alternative Title	Records of the Grand Historian
Author	Sima Qian
Translator	Watson, Burton
Date	c. 94 BC
Publisher	Columbia University Press
References	Tomb of First Emperor

Metadata  
for



Document

# Terracotta Warriors Place Documentation

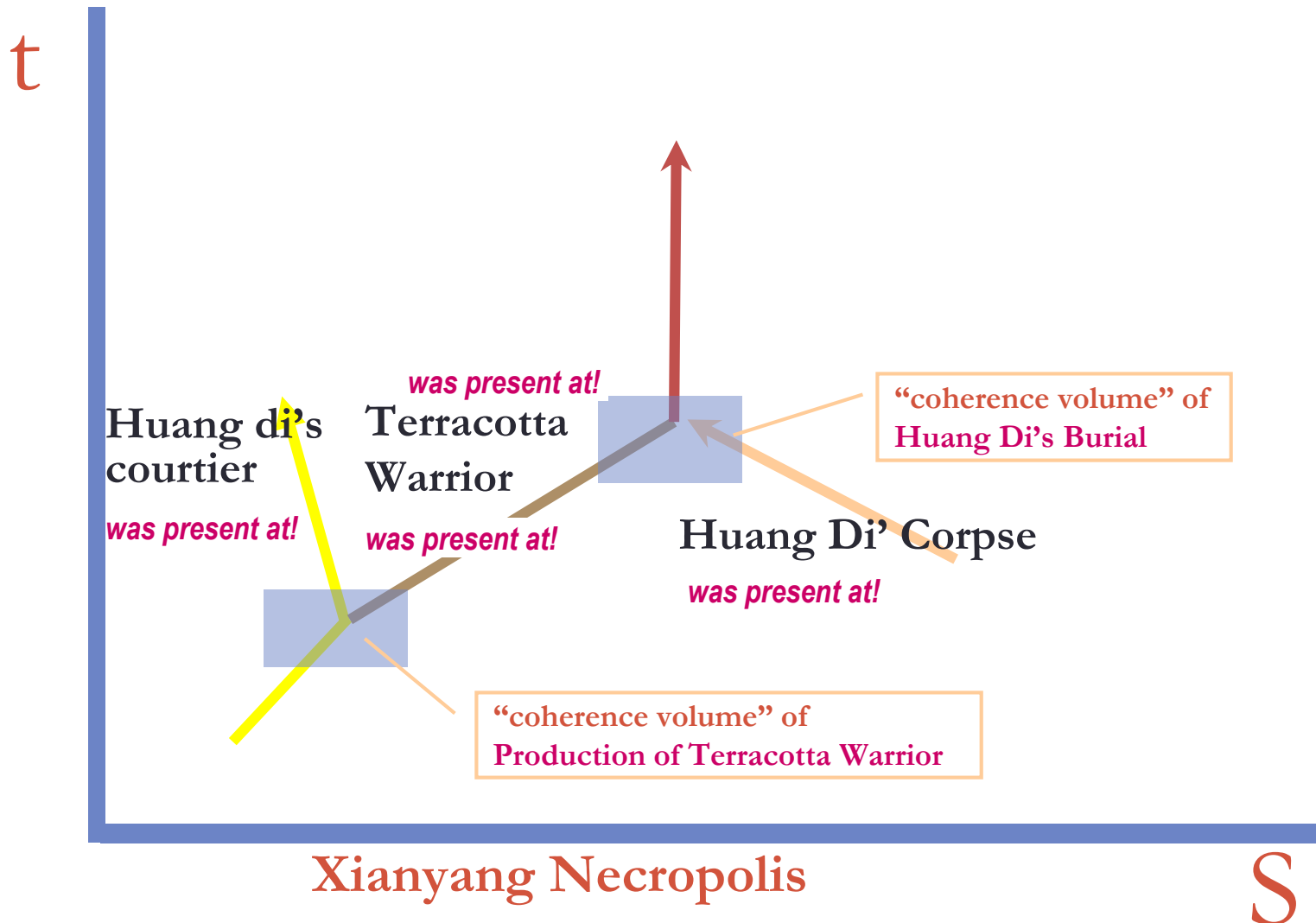
Field	Value
TGN ID	7001810
Names	Sian (C,V), Xi'an (C,V, Preferred)
Types	inhabited place(C), city (C)
Position	Lat: 44 30 N,Long: 034 10 E
Hierarchy	World (facet) -> Asia (continent) -> China (nation) (P) -> Shaanxi (province) (P)
Note	The empire capital under many dynasties, famed for the tomb of Shihuangdi, buried with terracotta soldiers...
Source	TGN, Thesaurus of Geographic Names

about

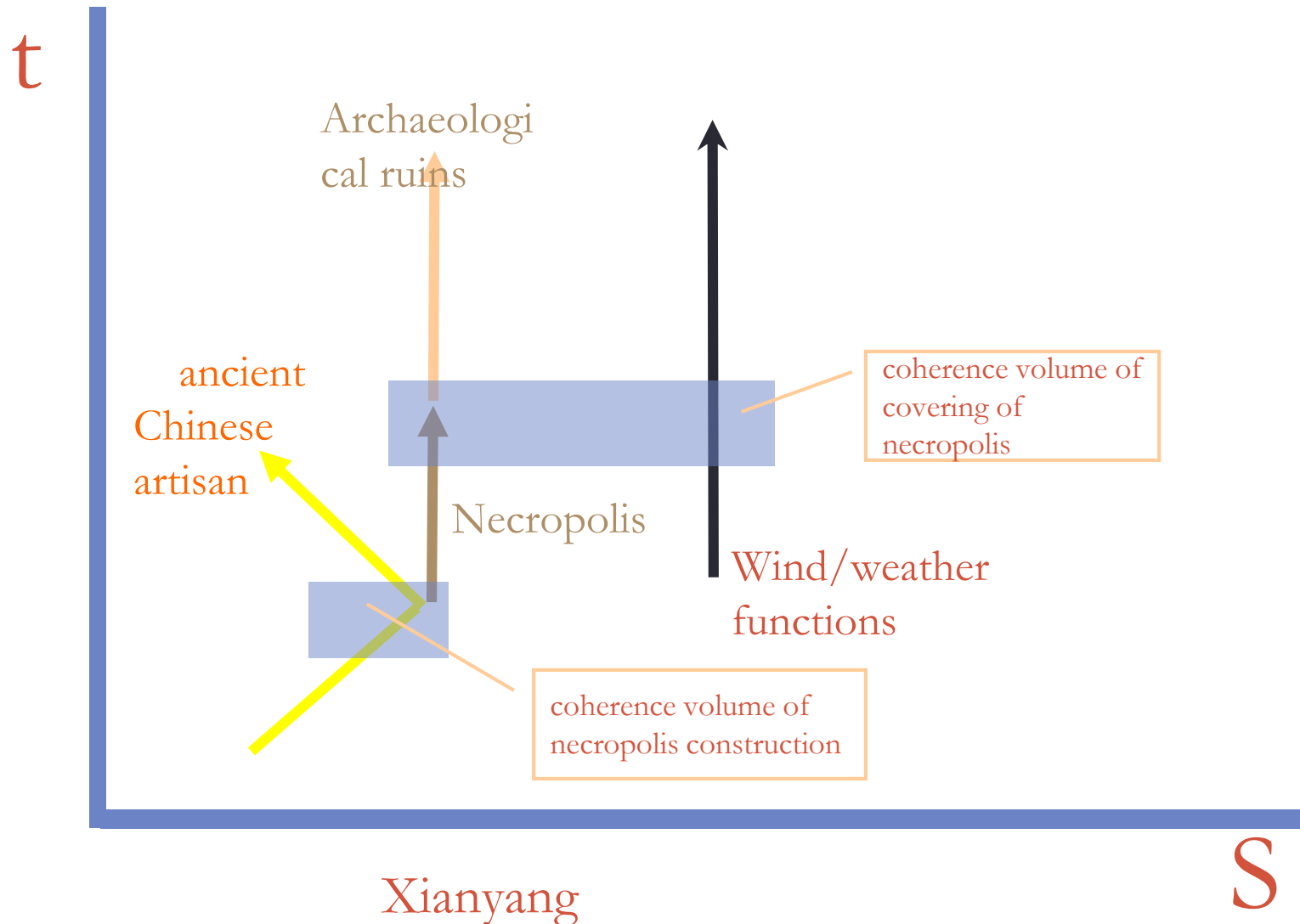


Place

# Historical Events as Meetings




# Depositional events as meetings



# Context of Production of Xi'an Warrior

**E39 Actor**



**E22 Man Made Object**



**E52 Time-Span**  
**221 to 206 BCE**

P11 participated in

**E5 Event**  
**Qin Dynasty**

P4 had time-span

P70i is documented in

**E31 Document**  
**Records of the Grand Historian**

P86 falls within

**E12 Production**  
\*

P7 took place at


**E53 Place**



P108 has produced

# Context of Finding Xi'an Warrior

E39 Actor



E5 Event

People's Republic of China

E31 Document

People's Daily

E7 Activity

Discovery

E53 Place



E52 Time-Span

1974

P11 participated in

P86 falls within

P70i is documented in

P12 witnessed

P7 took place at

P82 at some time within

E22 Man Made Object



# 4. FAMILY MODELS

---

# Extending the CRM

- The CRM standardizes only **stable concepts for information sharing**.
  - local extensions are encouraged for subjective concepts and local practices
- Forms 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”**

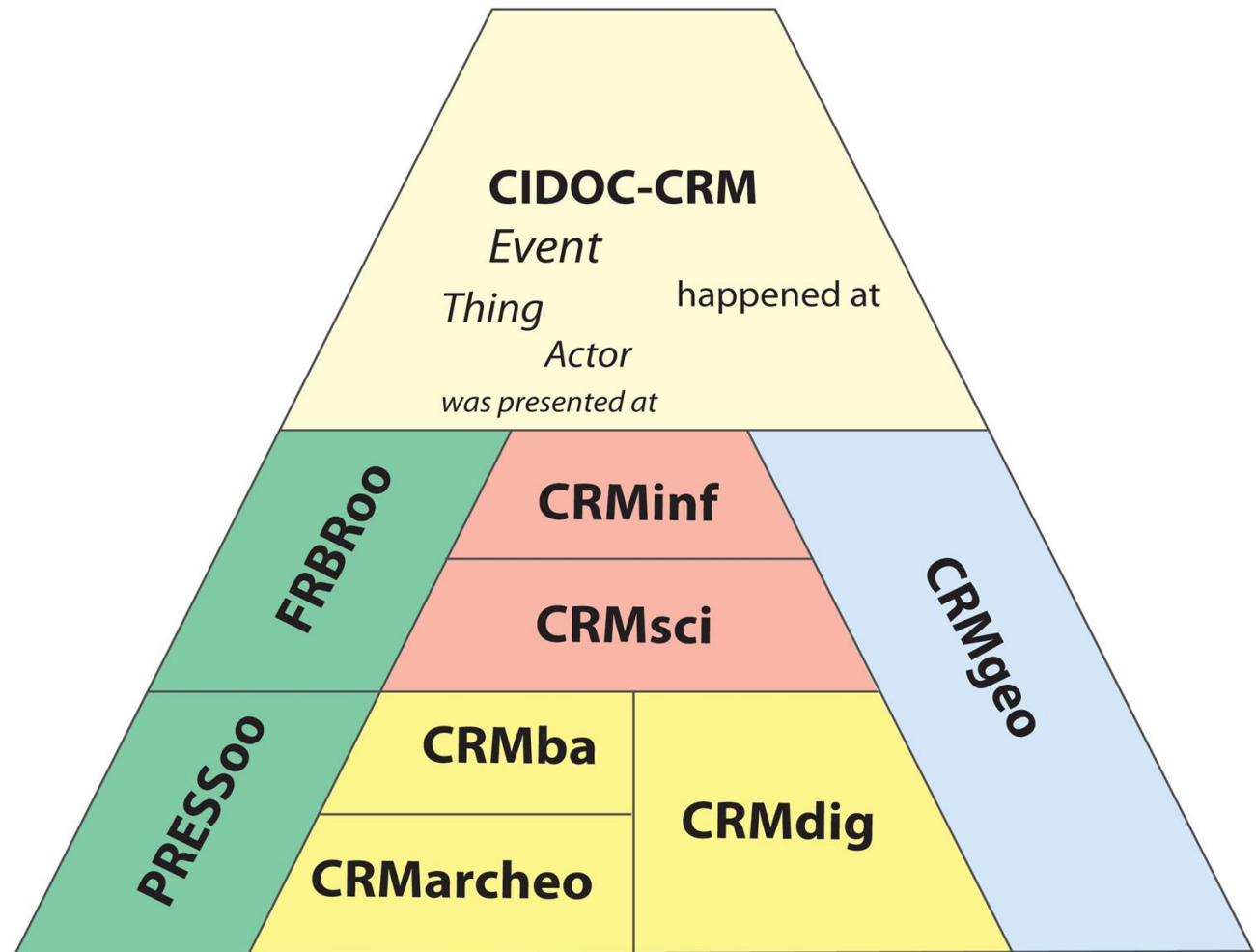
# CIDOC CRM extension suite

*CIDOC Conceptual Reference Model (CRM)*

by virtue of  
superproperties:

*Few concepts,  
high recall*

*Special concepts,  
high precision*



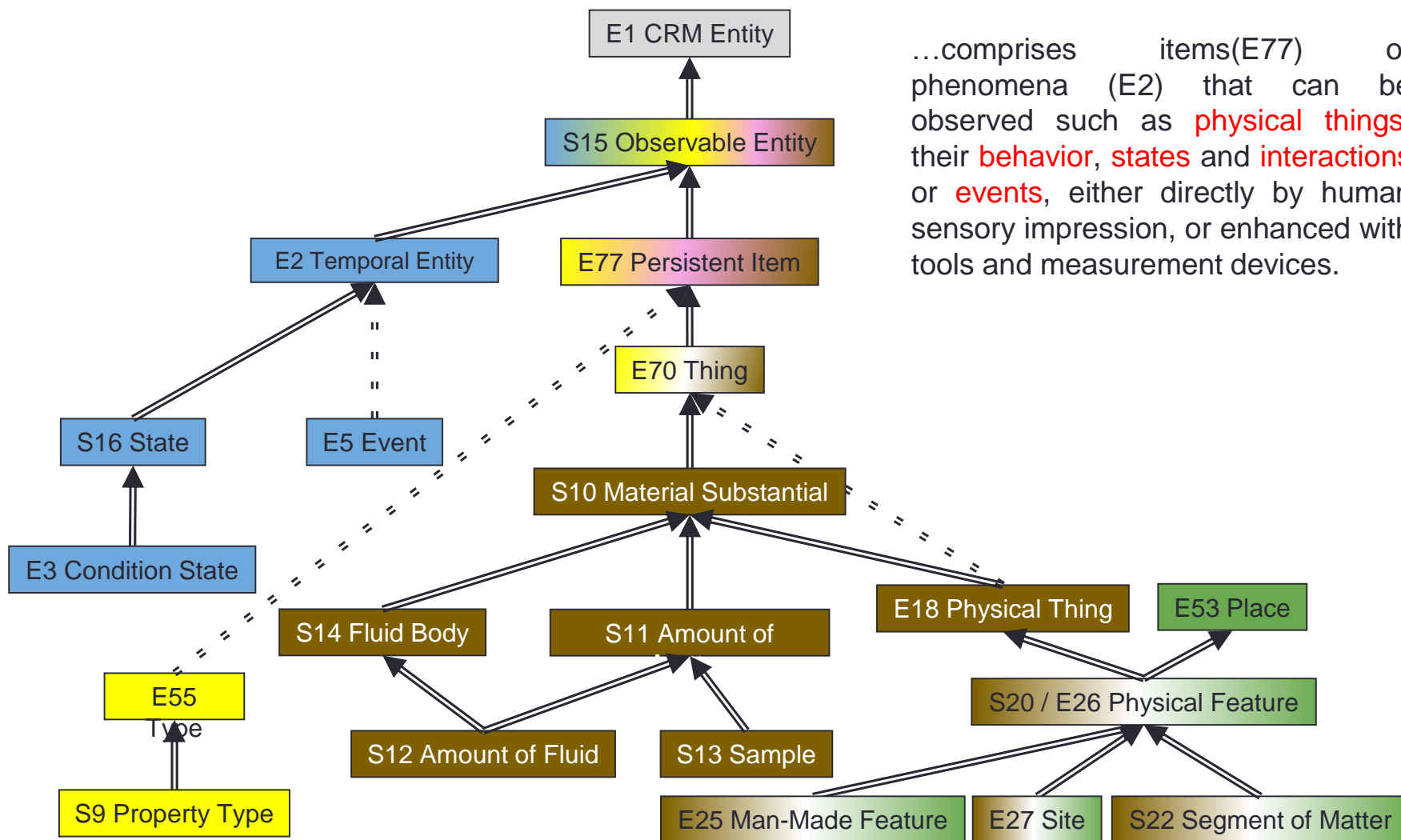
# CIDOC CRM Extensions

Name	Scope	Community of Practice	Related Project / Institution
FRBRoo	Bibliographic Data and Creative Processes	Librarians	IFLA
PreSSoo	Serials Data	Librarians	IFLA
CRMinf	Argumentation		
CRMsci	E-sciences	Analytic heritage science community	Kripis
CRMdig	Digitization processes	3D modelling community	3DCoform
CRMarchaeo	Excavation practice	Archaeologists	Ariadne
CRMba	Building archaeology	Archaeologists	Ariadne
CRMgeo	Geophysics and geolocation	Archaeologists and Geophysicists	Marie Curie Individual Project





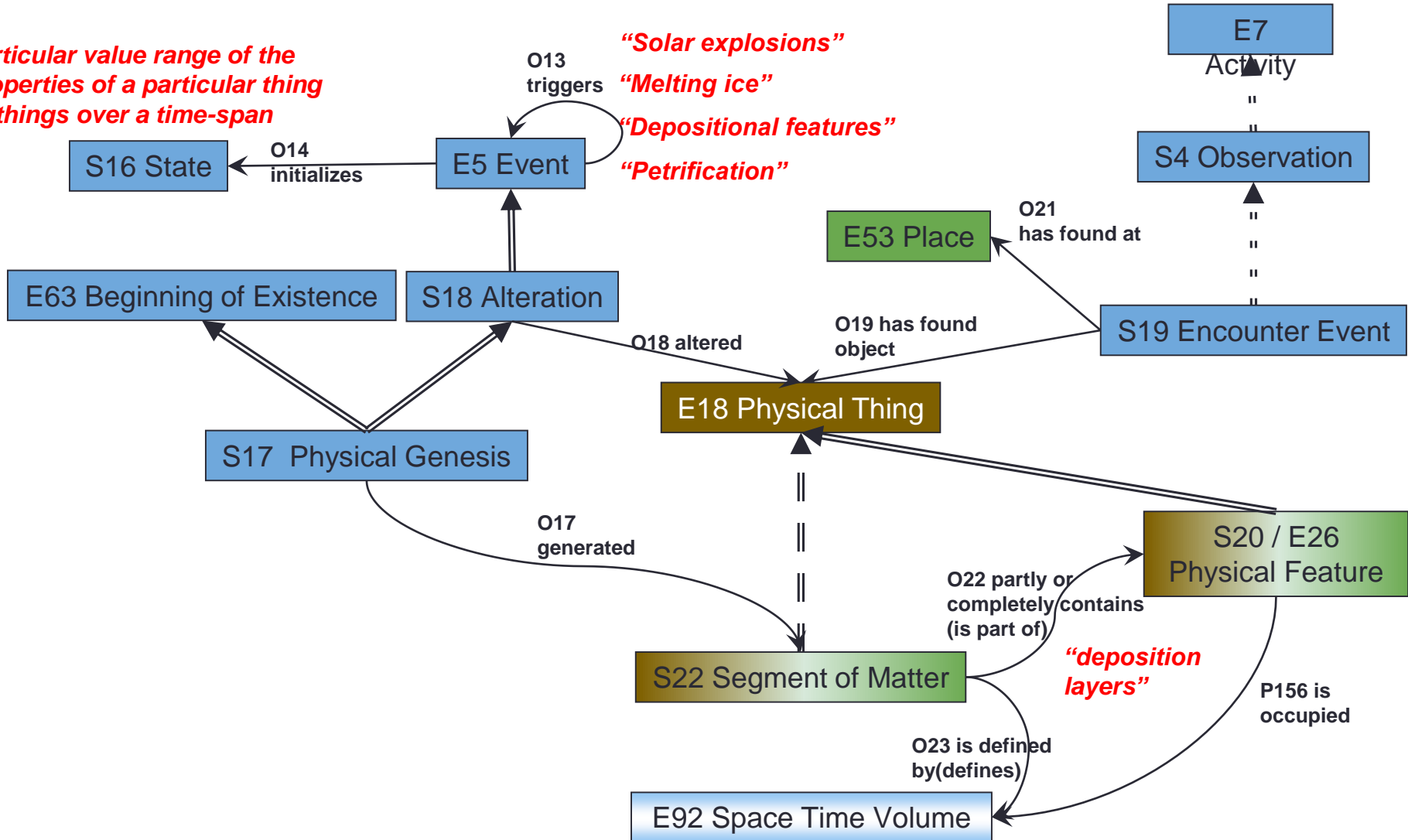
# CRMsci: what's observable



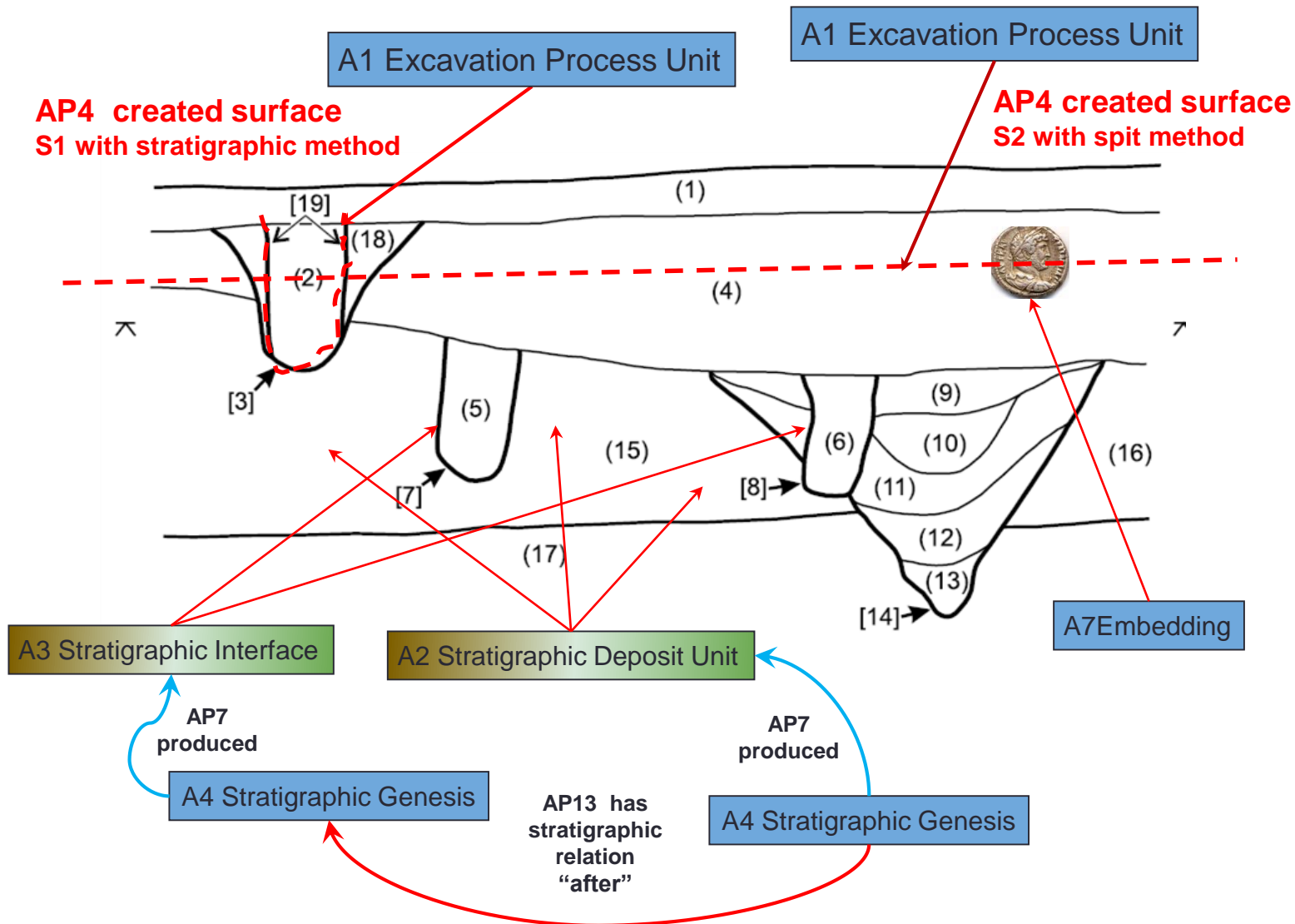
...comprises items(E77) or phenomena (E2) that can be observed such as **physical things**, their **behavior**, **states** and **interactions** or **events**, either directly by human sensory impression, or enhanced with tools and measurement devices.

# CRMsci: physical genesis and discovery

*particular value range of the properties of a particular thing or things over a time-span*



# CRMarchaeo: excavation as observation and destruction



# 5. TOOLS, IMPLEMENTATION & MORE

---

# A Selection of Semantic DM Tools

## Ontological Development



OntoME

## Mapping Tools



OnTop

**Karma**

## Triple Stores / Graph DBs



## Semantic Data Management Platforms



# Implementation Examples

## Museums/Institutions

The British  
Museum

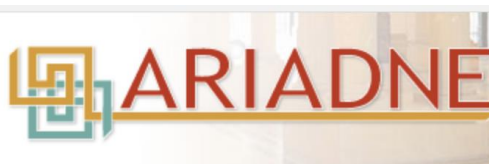
GERMANISCHES  
NATIONAL  
MUSEUM



RIJKS MUSEUM



## European Networks



PARTHENOS

Pooling Activities, Resources and Tools  
for Heritage E-research Networking,  
Optimization and Synergies

BlueBRIDGE

## Global



american art  
**COLLABORATIVE**  
Linked Open Data Initiative



# More information?

## Reference Materials

- CIDOC CRM Specification:

[http://www.cidoc-crm.org/releases\\_table](http://www.cidoc-crm.org/releases_table)

- Visual Charts:

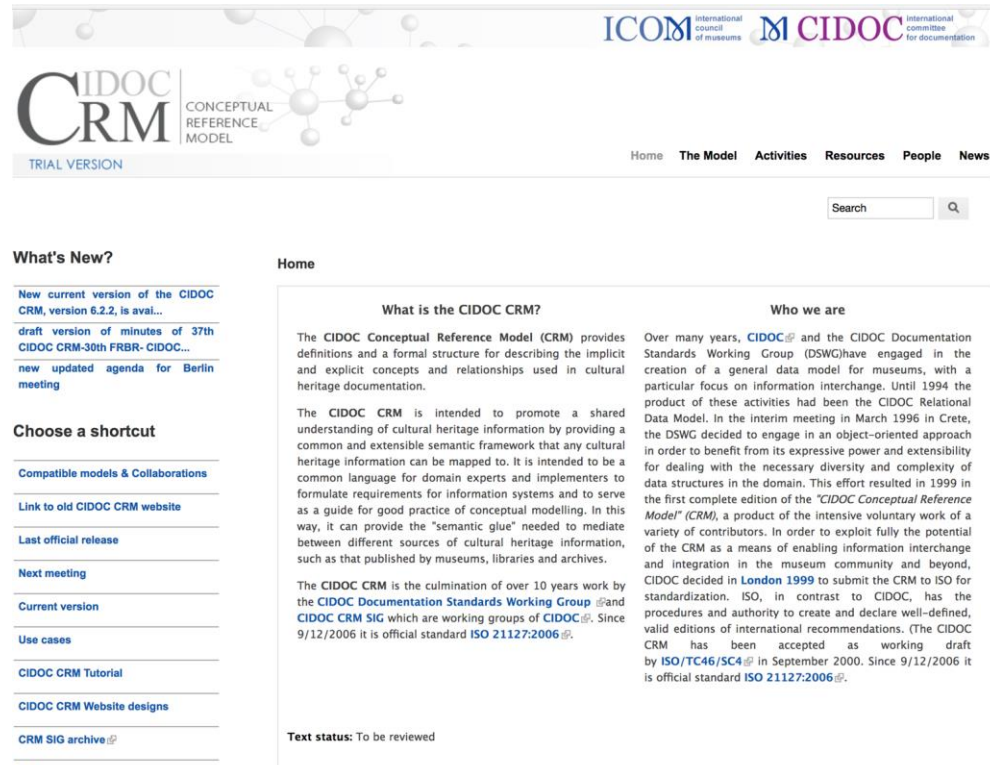
[http://old.cidoc-crm.org/cidoc\\_graphical\\_representation\\_v\\_5\\_1/graphical\\_representation\\_5\\_0\\_1.html](http://old.cidoc-crm.org/cidoc_graphical_representation_v_5_1/graphical_representation_5_0_1.html)

## Tutorials

- One video
- Many powerpoints:

<http://www.cidoc-crm.org/tutorialPage>

- Mailing list



The screenshot shows the homepage of the CIDOC CRM website. At the top, there are logos for ICOM (International Council of Museums) and CIDOC (International Committee for Documentation). The main header features the 'CIDOC CRM' logo with the tagline 'CONCEPTUAL REFERENCE MODEL' and a 'TRIAL VERSION' label. A navigation menu includes 'Home', 'The Model', 'Activities', 'Resources', 'People', and 'News'. A search bar is located on the right side. The main content area is divided into two columns. The left column, titled 'What's New?', lists several updates: 'New current version of the CIDOC CRM, version 6.2.2, is available...', 'draft version of minutes of 37th CIDOC CRM-30th FRBR- CIDOC...', and 'new updated agenda for Berlin meeting'. Below this is a 'Choose a shortcut' section with links for 'Compatible models & Collaborations', 'Link to old CIDOC CRM website', 'Last official release', 'Next meeting', 'Current version', 'Use cases', 'CIDOC CRM Tutorial', 'CIDOC CRM Website designs', 'CRM SIG archive', and 'FAQ'. The right column, titled 'Home', contains two sub-sections: 'What is the CIDOC CRM?' and 'Who we are'. The 'What is the CIDOC CRM?' section explains that the CIDOC Conceptual Reference Model (CRM) provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation. It also states that the CRM is intended to promote a shared understanding of cultural heritage information by providing a common and extensible semantic framework. The 'Who we are' section describes the CIDOC CRM as the culmination of over 10 years of work by the CIDOC Documentation Standards Working Group (DSWG) and the CIDOC CRM SIG. It mentions that the CRM was submitted to ISO for standardization in London 1999 and became an official standard ISO 21127:2006 in September 2006. A 'Text status: To be reviewed' message is displayed at the bottom of the right column.

<http://www.cidoc-crm.org/>

# THANK YOU, QUESTIONS?

---

George Bruseker ([george.bruseker@gmail.com](mailto:george.bruseker@gmail.com))

Anais Guillem ([anais.guillem@gmail.com](mailto:anais.guillem@gmail.com))