Dixit – ICS-FORTH Knowledge Exchange

George Bruseker (ICS-FORTH)

March 28, 2017

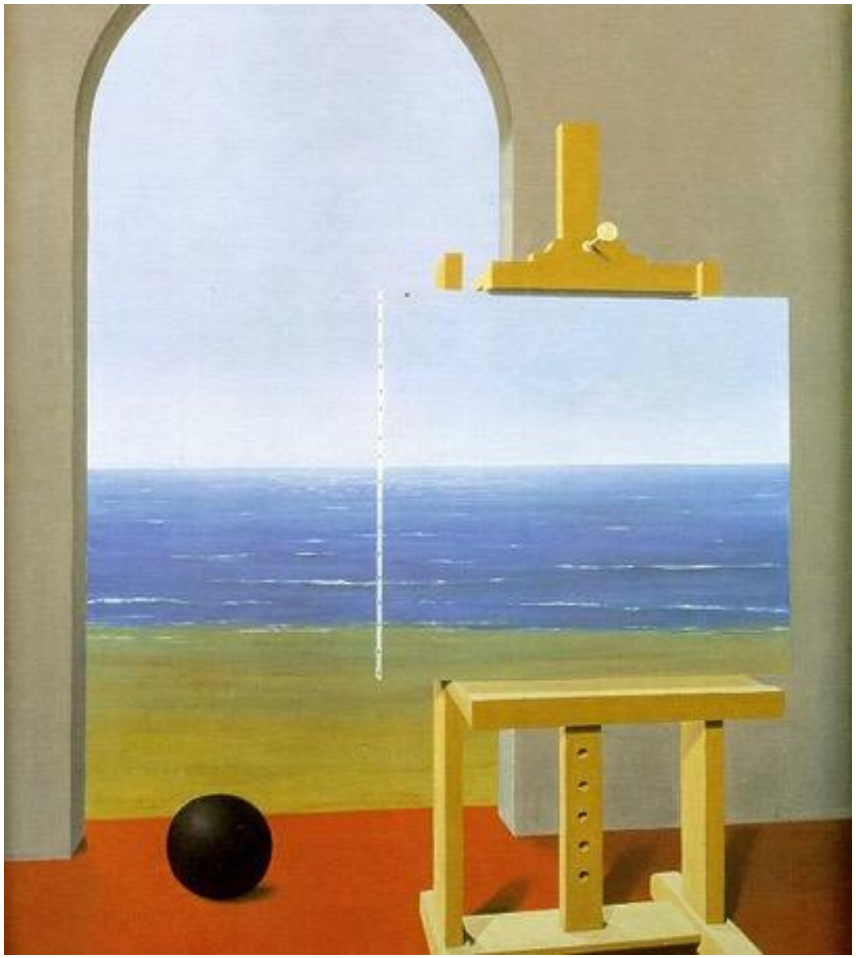
Crete, Greece

Agenda

1. What is a formal ontology and what is it good for?
2. High Level Overview to CIDOC CRM
3. Discussion of CIDOC CRM Extensions
   a. FRBRoo: for documenting creative processes
   b. CRMinf: for documenting argumentation
4. How to use CIDOC CRM

1. WHAT IS A FORMAL ONTOLOGY?

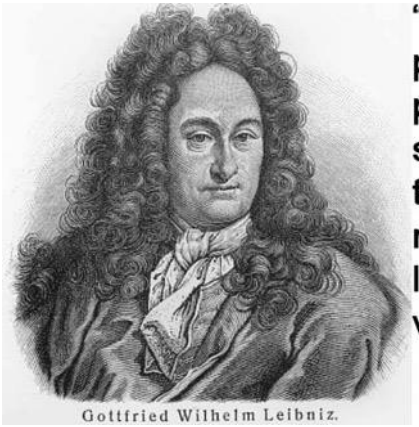
What is an ontology? (What is it not?)



- A window / frame which allows expression of the world according to the kinds of statements used by a domain(s)
- Not a description of the world as such (not Ontology)
- Plural not singular
- Heidegger's 'domain ontology'

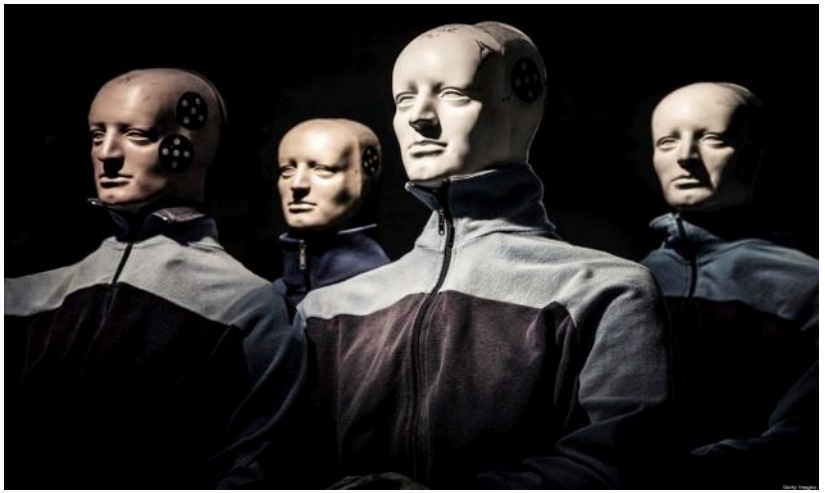
What is an ontology? (What is it not?)

- Mix of computer science and philosophy
- Bridge between computer science world and data producers/researchers
- Formalization of knowledge domain
- Bound to a reality
- Creates a *pidgin* or *lingua franca*
- Machine Processable, Human Readable



"A Universal Symbolism, very popular, very agreeable to the people...might be introduced if small figures were employed in the place of words, which would represent visible things by their lines, and the invisible, by the visible, which accompany them."

— Leibniz, 1679



Increased knowledge?

Why?
Specialists
(stuck?)
in their worlds

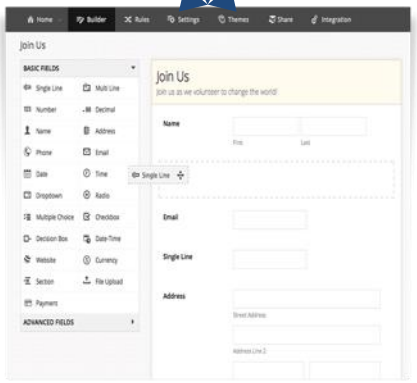
observe

publish

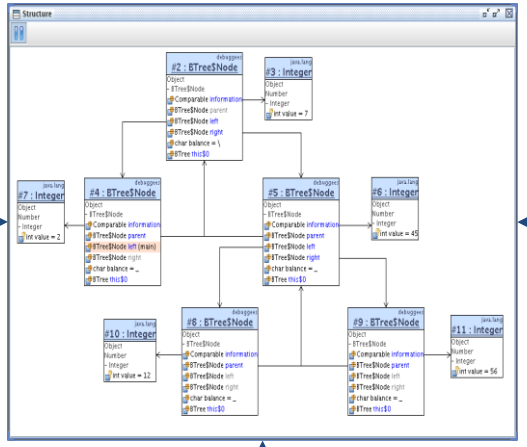
ground

input

Reach /
understand
/ critique



store



The real data/research situation on the ground

One world

Heterogeneity of Actors

- Many researchers
  - Many research questions
  - Many research methods
- Many institutions
  - Many institutional policies

about

Heterogeneity of Technology

- Many data standards
- Many data formats
- Many data entry/storage/retrieval tools

leads to

Intractable Data Heterogeneity

requires

Heterogeneity of Means

- Different financial constraints
- Different levels of ICT support

Methodological Solution

The Crimea Conference Example

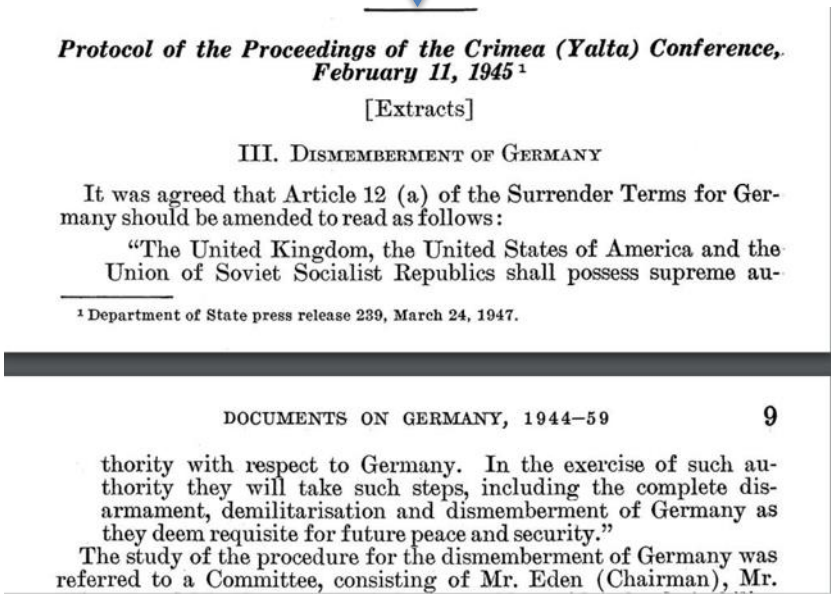
- Researcher wants to find all evidence and relevant material to Crimea conference
- Researcher wants to use textual sources, photographic archive and relevant geographic data to understand context and elements of this historic moment
- One standard won't do, and two or more standards leaves the data unrelated

Crimea Conference
Historical Archives….

Standard: Dublin Core OR EAD?

metadata for

Field	Value
Type	Text
Title	Protocol of Proceedings of Crimea Conference
Subtitle	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
Publisher	State Department
Subject	Postwar division of Europe and Japan



Protocol of the Proceedings of the Crimea (Yalta) Conference, February 11, 1945 [1]

[Extracts]

III. DISMEMBERMENT OF GERMANY

It was agreed that Article 12 (a) of the Surrender Terms for Germany should be amended to read as follows:

"The United Kingdom, the United States of America and the Union of Soviet Socialist Republics shall possess supreme au-

[1] Department of State press release 239, March 24, 1947.

DOCUMENTS ON GERMANY, 1944–59 9

thority with respect to Germany. In the exercise of such authority they will take such steps, including the complete disarmament, demilitarisation and dismemberment of Germany as they deem requisite for future peace and security."

The study of the procedure for the dismemberment of Germany was referred to a Committee, consisting of Mr. Eden (Chairman), Mr.

Documents

Crimea Conference

Images, non-textual objects…

Standard: Dublin Core OR IPTC?

about

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

Photographs

Crimea Conference
Places and Objects

about

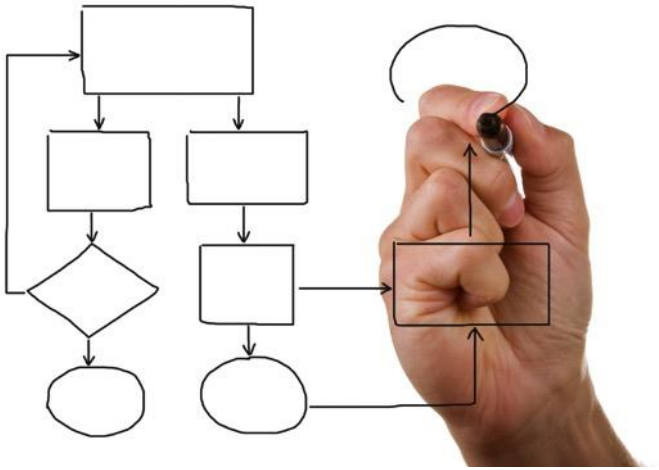
Standard: TGN OR Geonames?

Field	Value
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

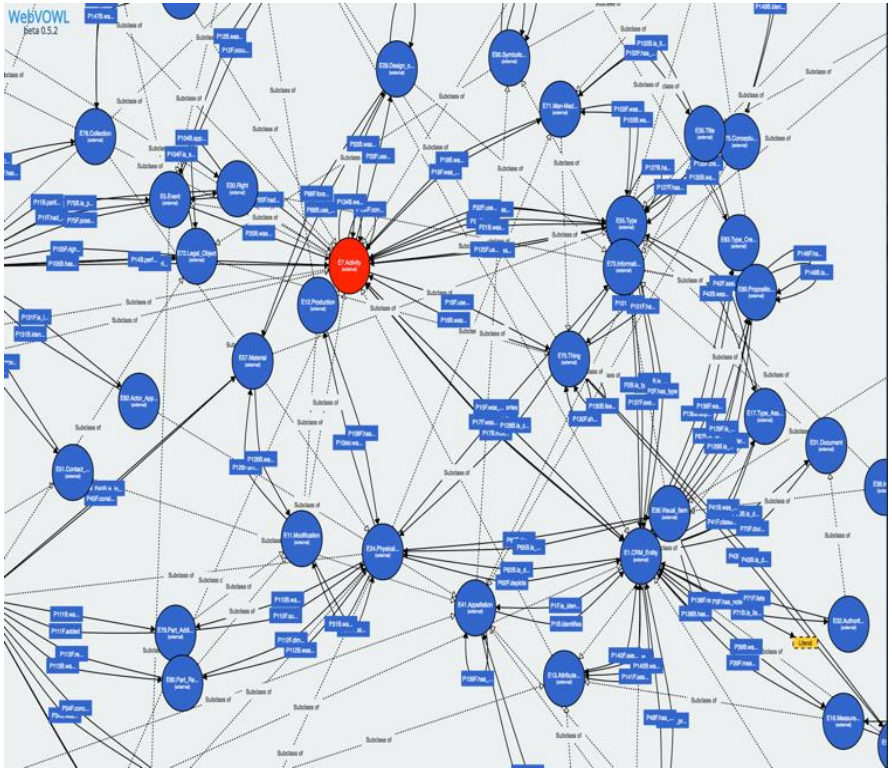
We need: a different approach

How is an ontology a solution? What can we hope for from it?



- Creates a general data form to which multiple heterogeneous data formats can be mapped
- Helps control polysemy problems (at schema level)
- In its production builds general concepts which can be used for high level information recall
- Extensible so that new knowledge can be integrated without requiring rethinking the ontology's basic divisions
- Creates compatibility
- Allows for automatic reasoning

What does an ontology look like?



Scope: a definition of the intended field of discourse/reality that the formal ontology should cover.

Classes: universals meant to represent some set of entities in the world of discourse, that have a distinct, identifiable behaviour and identity.

Properties: the relations that exist between classes in the ontology. These formally define the possible relations between classes and their meaning.

IsA Hierarchy: identifies generalization/specialization relations

Anatomy of a Class

The Label: arbitrary but identifying

Subclass/Superclass: Place in IsA

The Scope Note: gives the meaning, the intension. First thing to check!

The Examples: helps to verify… do others think/do it like you do

The Properties: more verification of appropriateness.

How does it relate to other concepts? Is this how my concept behaves?

E5 Event

Subclass of: E4 Period
Superclass of: E7 Activity
E63 Beginning of Existence
E64 End of Existence

Scope Note: This class comprises changes of states in cultural, social or physical systems, regardless of scale, brought about by a series or group of coherent physical, cultural, technological or legal phenomena. Such changes of state will affect instances of E77 Persistent Item or its subclasses.

The distinction between an E5 Event and an E4 Period is partly a question of the scale of observation. Viewed at a coarse level of detail, an E5 Event is an 'instantaneous' change of state. At a fine level, the E5 Event can be analysed into its component phenomena within a space and time frame, and as such can be seen as an E4 Period. The reverse is not necessarily the case: not all instances of E4 Period give rise to a noteworthy change of state.

Examples:

* the birth of Cleopatra (E67)
* the destruction of Lisbon by earthquake in 1755 (E6)
* World War II (E7)
* the Battle of Stalingrad (E7)
* the Yalta Conference (E7)
* my birthday celebration 28-6-1995 (E7)
* the falling of a tile from my roof last Sunday
* the CIDOC Conference 2003 (E7)

Properties

P1 is identified by (identifies): E41 Appellation

P137 exemplifies (is exemplified by): E55 Type
(P137.1 in the taxonomic role: E55 Type)
P2 has type (is type of): E55 Type

P11 had participant (participated in): E39 Actor

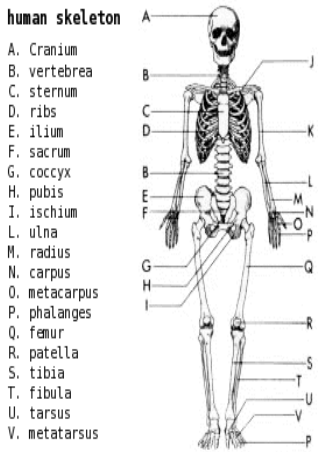
P12 occurred in the presence of (was present at): E77 Persistent Item

P8 took place on or within (witnessed): E19 Physical Object
(short cut of: E46 Section Definition)
P7 took place at (witnessed): E53 Place

P4 has time-span (is time-span of): E52 Time-Span

P120 occurs before (occurs after): E2 Temporal Entity
P119 meets in time with (is met in time by): E2 Temporal Entity
P118 overlaps in time with (is overlapped in time by): E2 Temporal Entity
P117 occurs during (includes): E2 Temporal Entity
P116 starts (is started by): E2 Temporal Entity
P115 finishes (is finished by): E2 Temporal Entity
P114 is equal in time to: E2 Temporal Entity
P10 falls within (contains): E4 Period
P9 consists of (forms part of): E4 Period
P132 overlaps with: E4 Period
P133 is separated from: E4 Period

P3 has note: E62 String
(P3.1 has type: E55 Type)



Anatomy of a Property

The Label: arbitrary but identifying

The Domain: The set of classes from which the property can originate

The Range: the set of classes to which the property can join the domain class

Superproperty/subproperty: Place in IsA Hierarchy

The Scope Note: gives the meaning, the intension. First thing to check!

The Examples: helps to verify…
do others think/do it like you do

P1 is identified by (identifies)

Domain: E1 CRM Entity
Range: E41 Appellation
Superproperty of: E1 CRM Entity. P48 has preferred identifier (is preferred identifier of): E42 Identifier
E52 Time-Span. P78 is identified by (identifies): E49 Time Appellation
E53 Place. P87 is identified by (identifies): E44 Place Appellation
E71 Man-Made Thing. P102 has title (is title of): E35 Title
E39 Actor. P131 is identified by (identifies): E82 Actor Appellation
E28 Conceptual Object.P149 is identified by (identifies): E75 Conceptual Object Appellation
Quantification: many to many (0,n:0,n)

Scope note: This property describes the naming or identification of any real world item by a name or any other identifier.

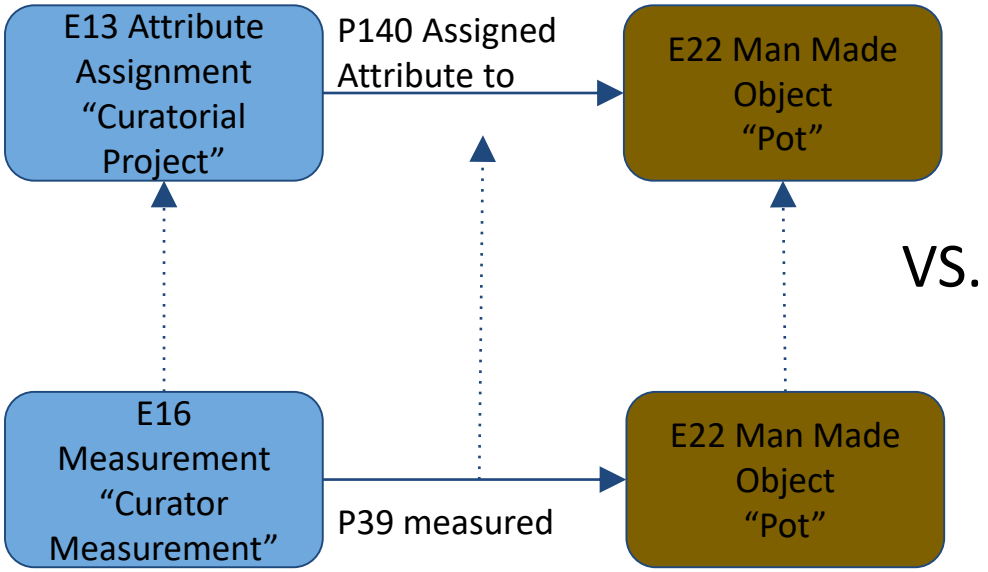
This property is intended for identifiers in general use, which form part of the world the model intends to describe, and not merely for internal database identifiers which are specific to a technical system, unless these latter also have a more general use outside the technical context. This property includes in particular identification by mathematical expressions such as coordinate systems used for the identification of instances of E53 Place. The property does not reveal anything about when, where and by whom this identifier was used. A more detailed representation can be made using the fully developed (i.e. indirect) path through E15 Identifier Assignment.

Examples:

- the capital of Italy (E53) *is identified by* "Rome" (E48)
- text 25014–32 (E33) *is identified by* "The Decline and Fall of the Roman Empire" (E35)

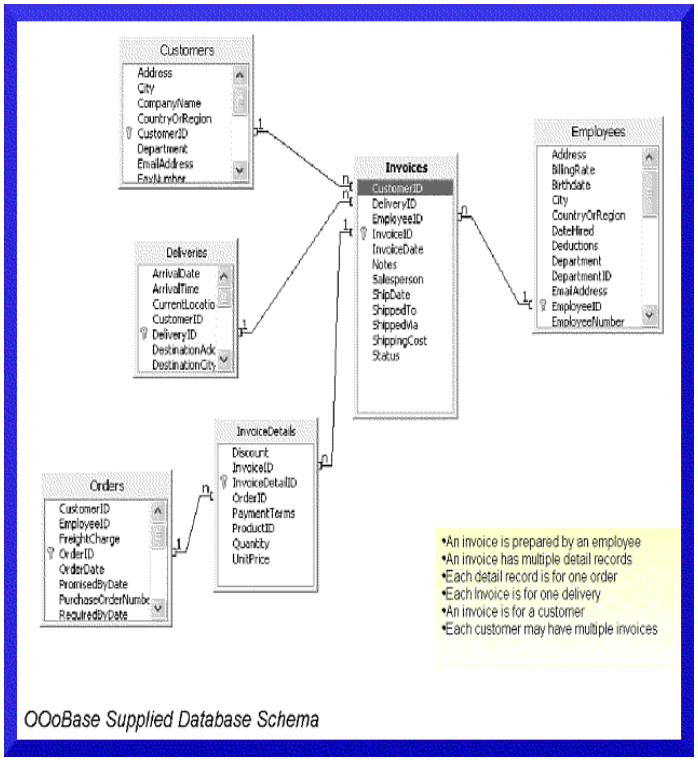
What kind of data does a formal ontology produce?

Triples (Subject - Verb - Object)



VS.

Traditional RDBMS

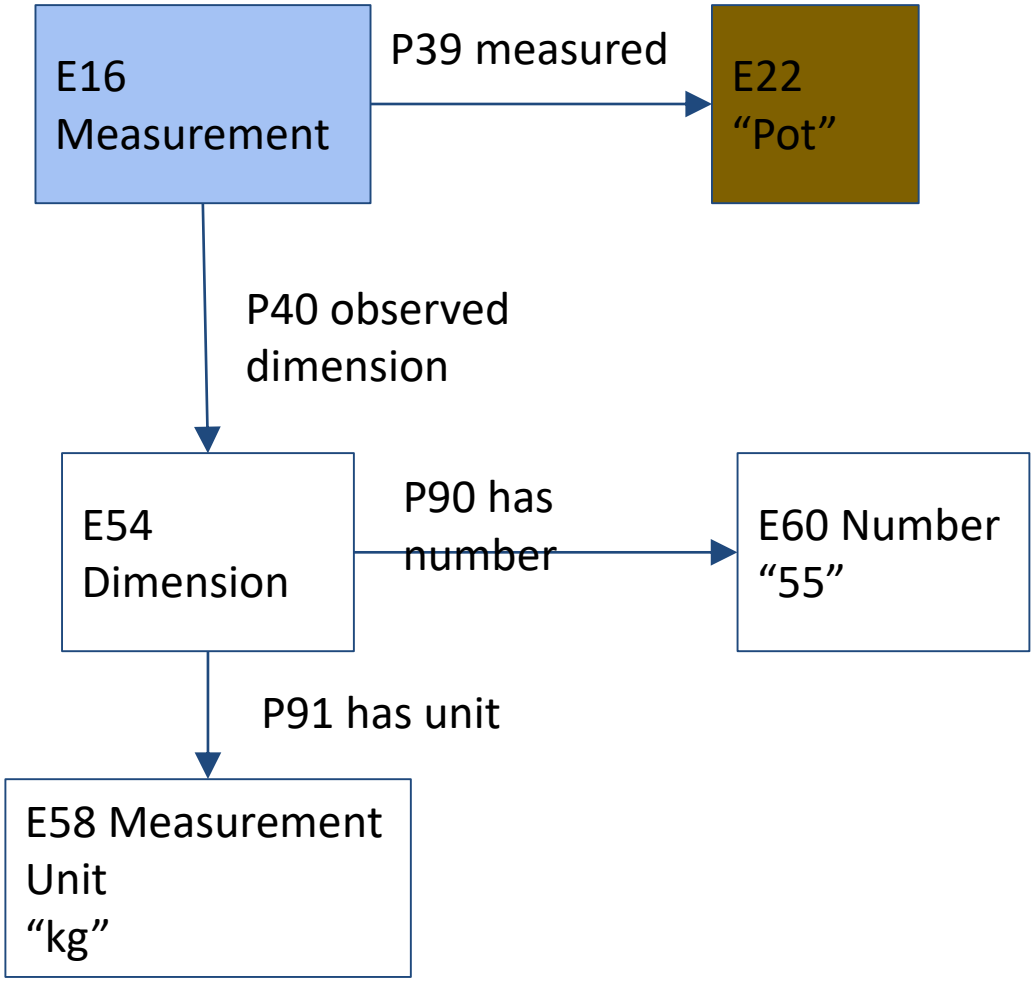


In a graph, supporting subsumption relation

What kind of data does a formal ontology produce?

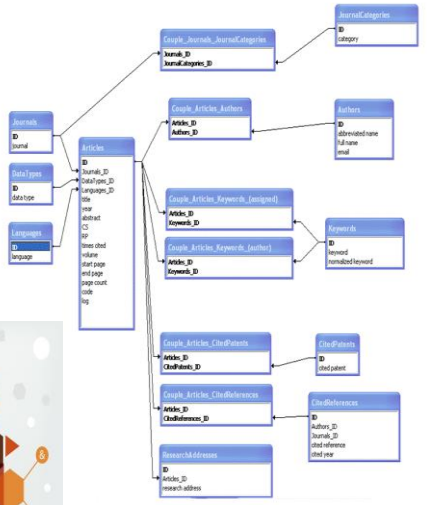
Object Table	
ID	101
Name	Pot
Weight (kg)	55

Rendering the implicit, explicit

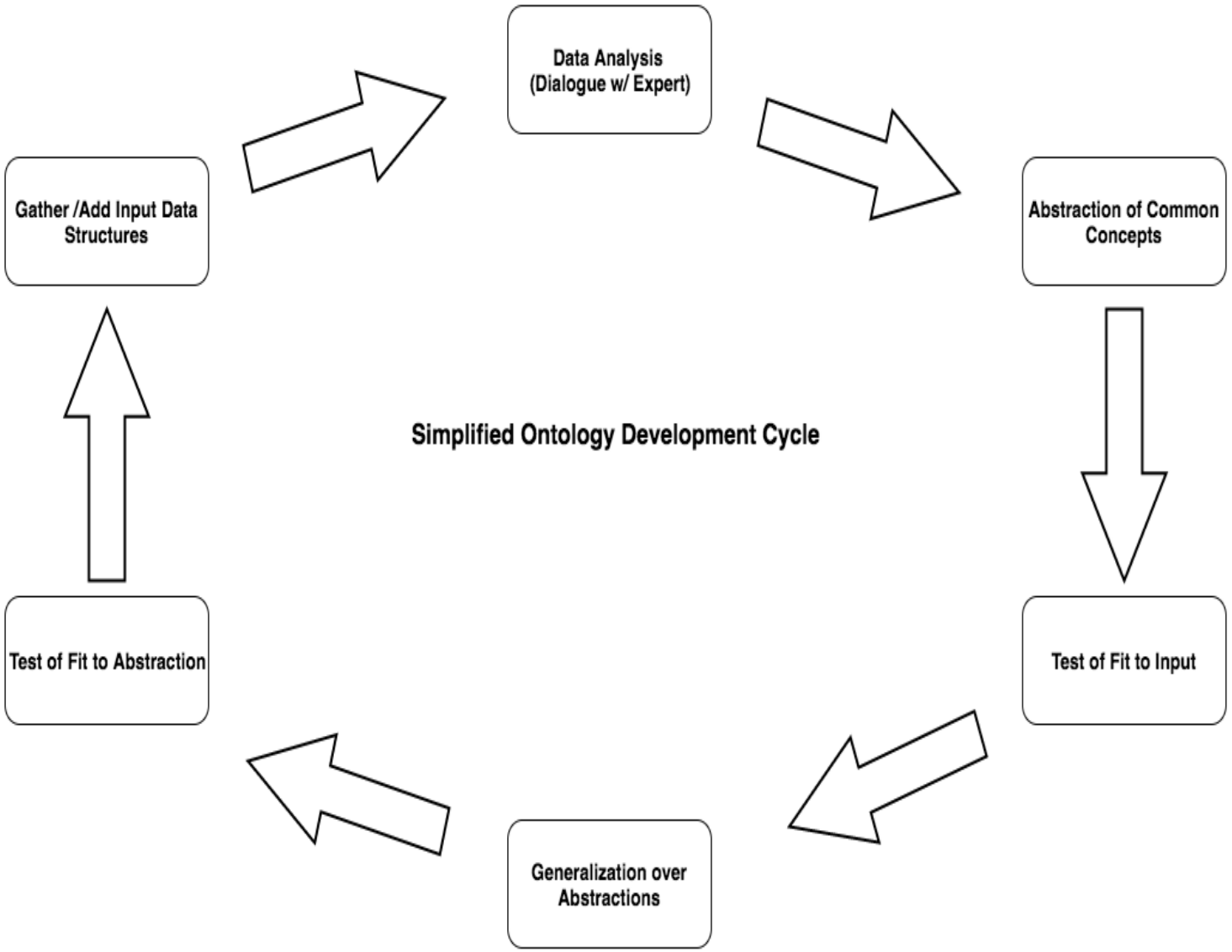


Formal Ontologies: How do they come about?

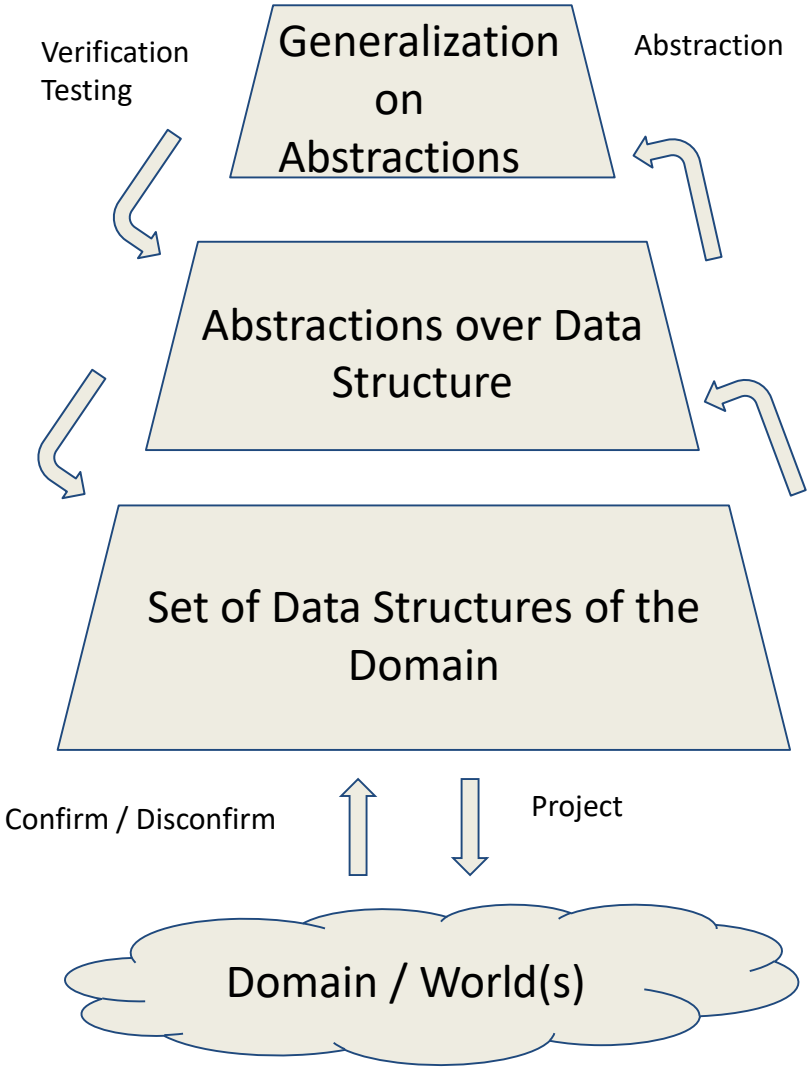
- Empirical study of data structures, to understand their semantic content
- Dialogue with domain specialists, to test conceptualizations, understand argumentation
- Elicitation of competence questions, to have a metric against which to measure the success of the effort



The process: iterative



The process: testing against objective domain



What, practically can be done with it?

Integration and Interoperability

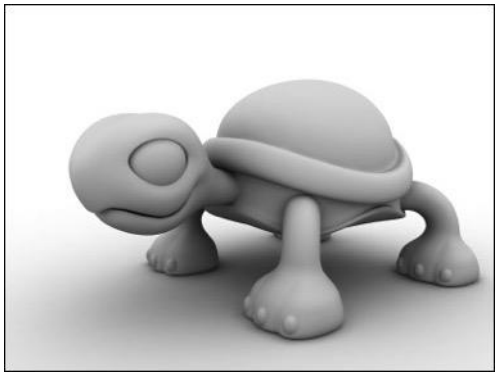
Modelling/Creating

1) Model from Scratch

No current ontology model exists, can start a modelling process for providing a lingua franca for your domain

2) **Model an Extension**

An ontology exists that is generically useful for your domain but cannot capture all relations to sufficient accuracy



Implementing/Using

1) Mapping Integration / Compatibility

An ontology exists, is adequate and you wish to map your existing data into it and take advantage of its capacities

2) **Native Expression / Compatibility**

An ontology exists, is adequate and you wish to express your data natively in that ontology and take advantage of its capacities

2. CIDOC CRM – HIGH LEVEL OVERVIEW

The CIDOC Conceptual Reference Model

ICOM international council of museums

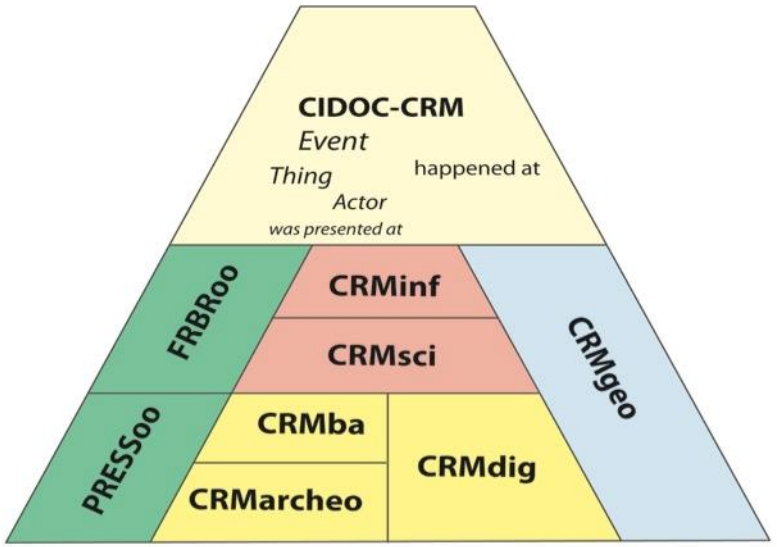
- *Developed by the **CRM Special Interest Group of the International Committee for Documentation** (CIDOC) of the International Council of Museums (ICOM), following an initiative of ICS-FORTH, Heraklion, Crete.*
- *a **core ontology** describing the underlying semantics of over a hundred database schemata and structures from all museum disciplines, archives and libraries.*
- *Recognized **ISO Standard** since 2006* (ISO21127:2006)
- *the **result of 20 years** of interdisciplinary work and agreement*
- *a **generic model of recording of "what has happened"** in human scale*
- ***generates huge, meaningful networks of knowledge by a simple abstraction**: history as meetings of people, things and information.*

CIDOC CRM: Description

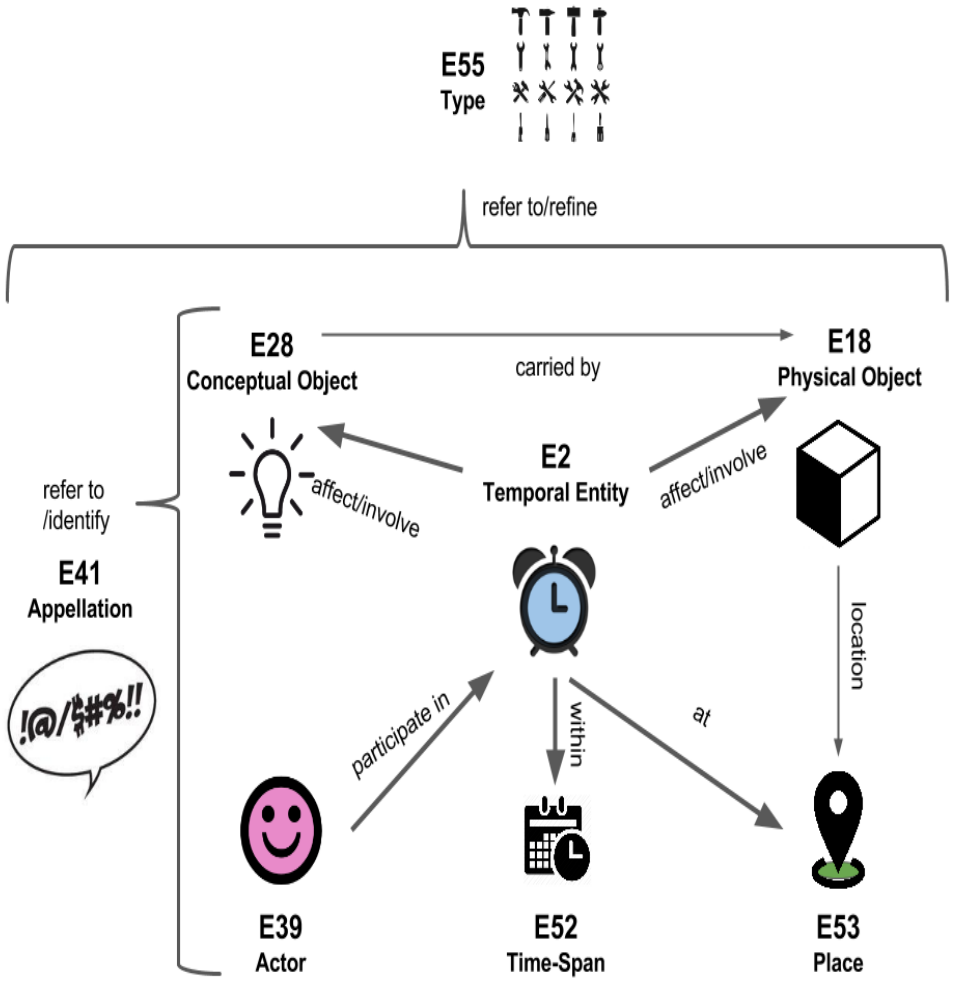
Type	Top Level Ontology
Scope	Cultural Heritage and E-Sciences
Classes	90+-
Relations	150+-
Version	6
Maintained by	CIDOC CRM SIG
Official Extensions	8
Access	http://www.cidoc-crm.org/

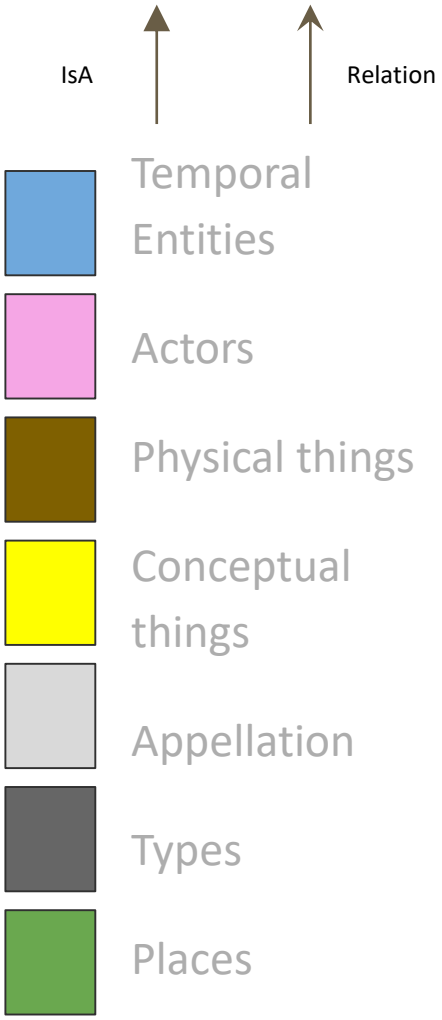
CIDOC CRM
CONCEPTUAL REFERENCE MODEL

CIDOC-CRM family of Models



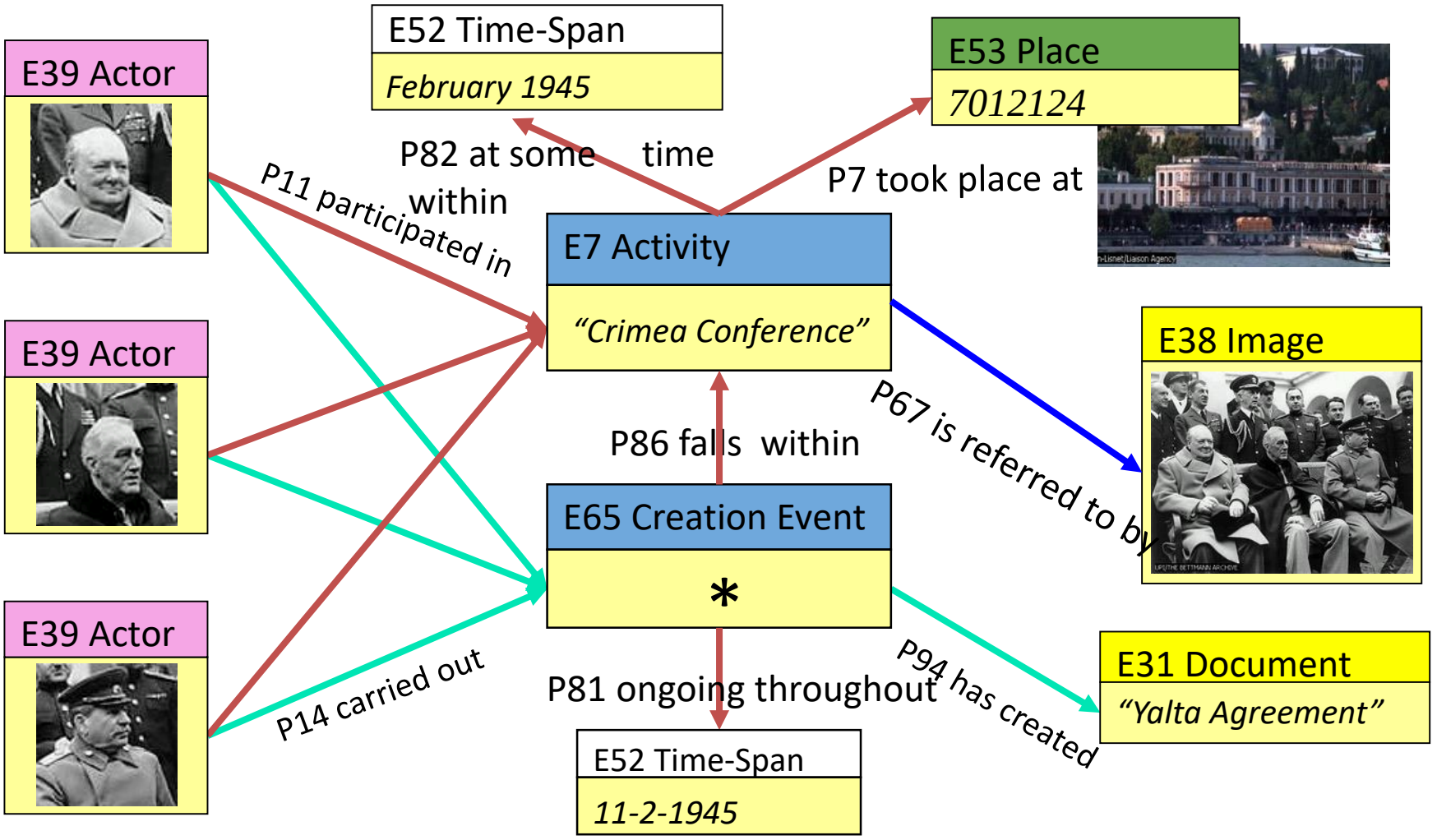
CIDOC CRM: General Modelling Pattern



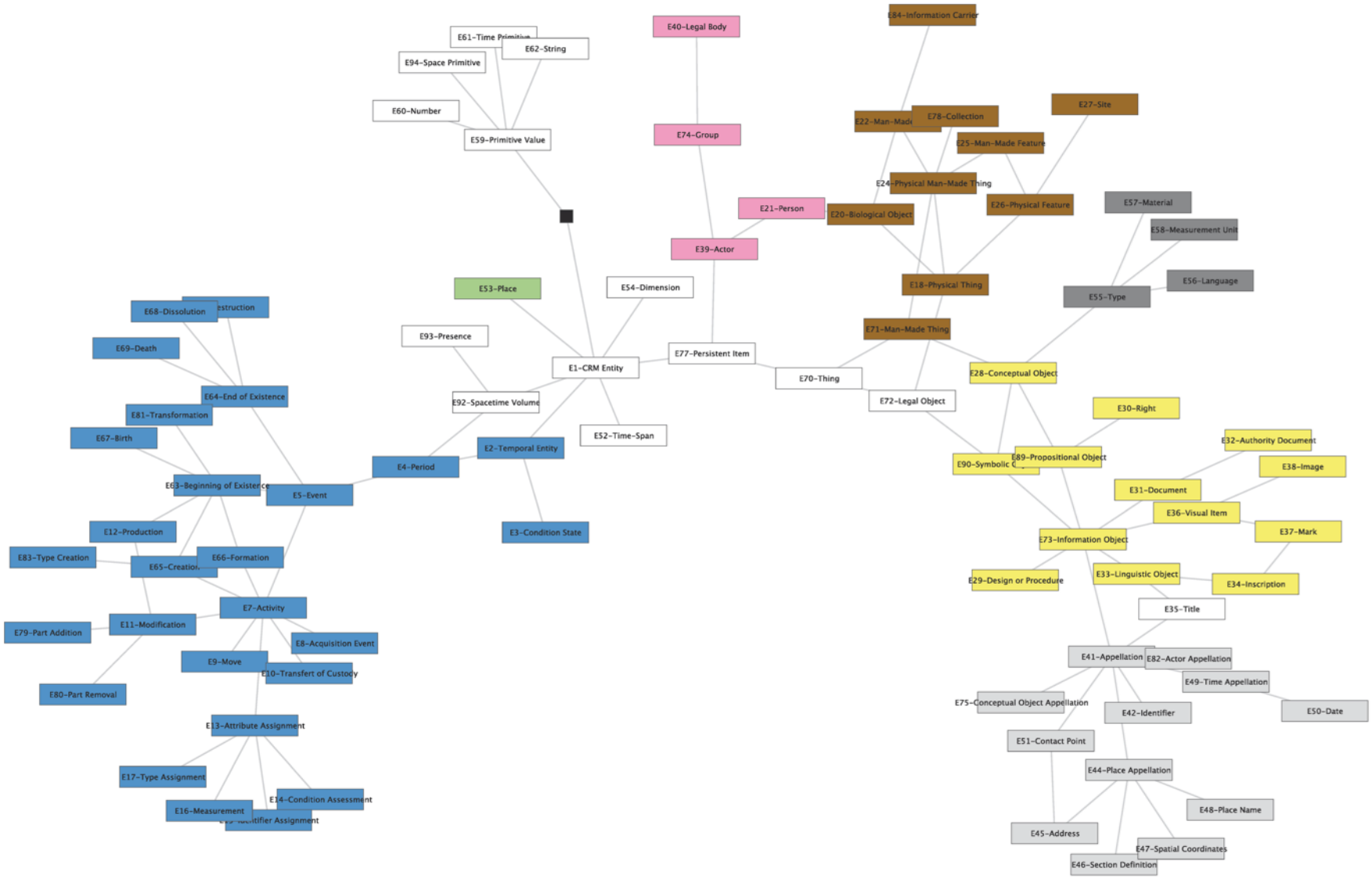


The Crimea Conference in CIDOC CRM

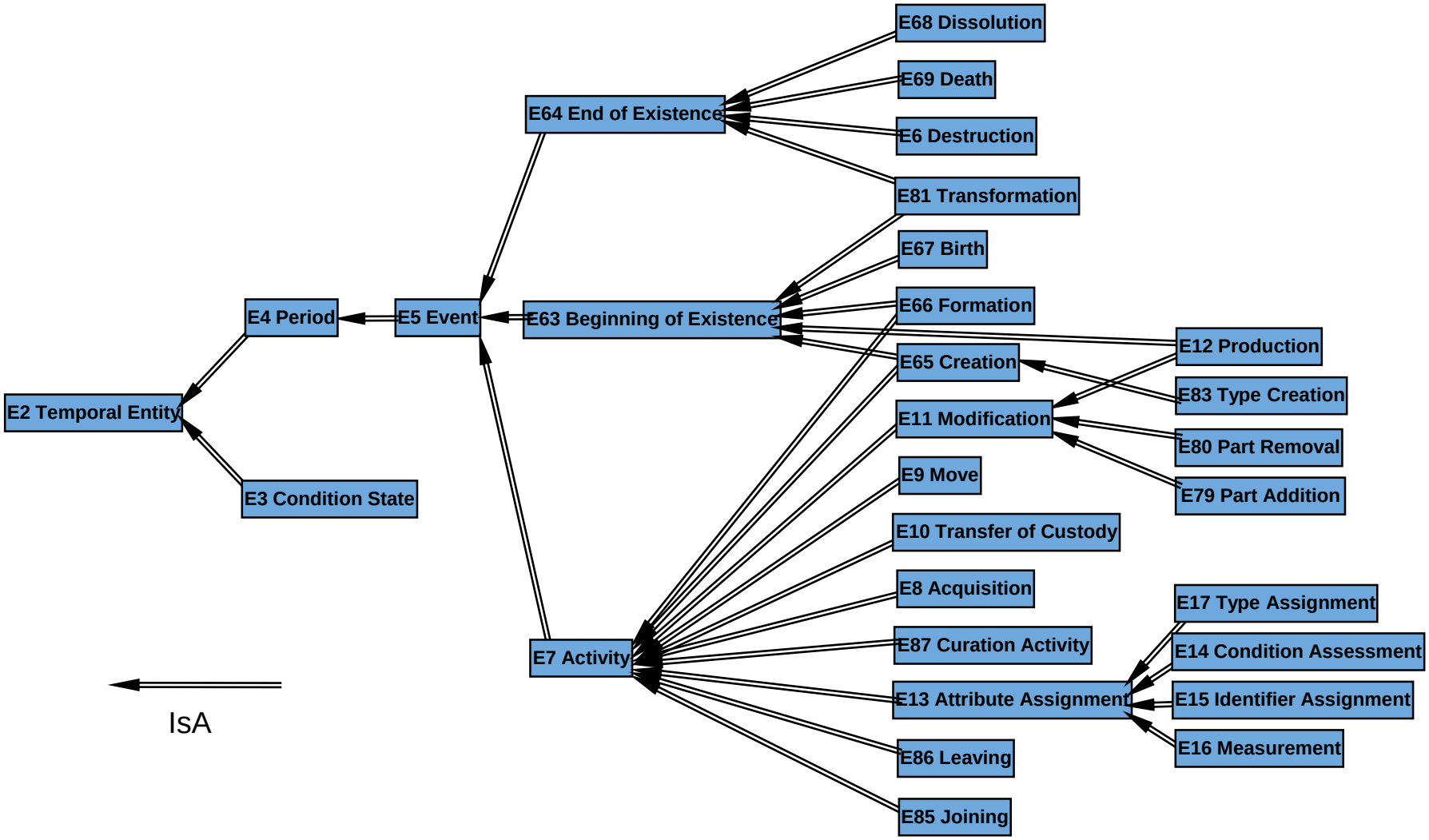
Explicit Events, Object Identity, Symmetry



CRM Overall Class Hierarchy



Temporal Classes: Reasoning over Time



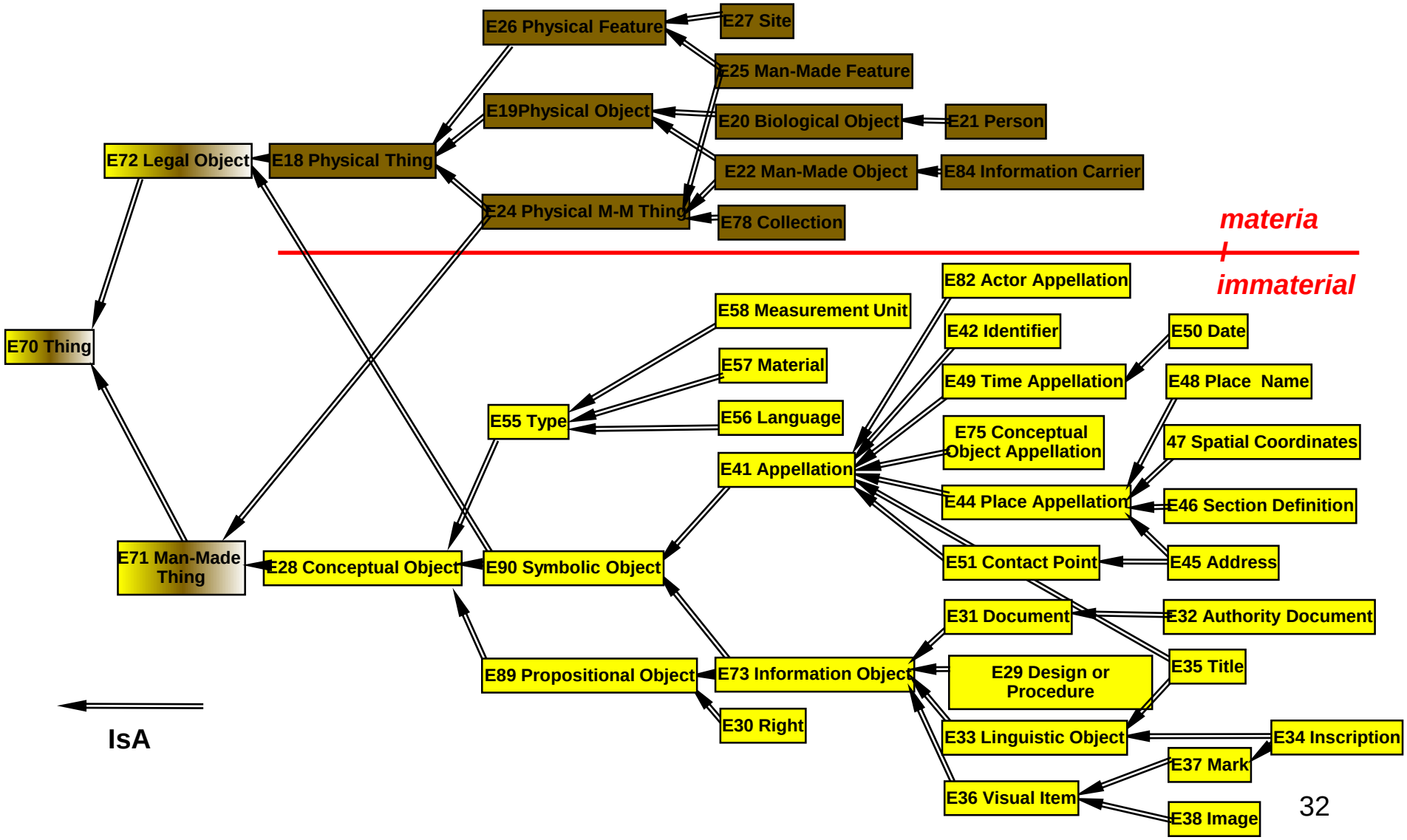
Temporal Classes: Reasoning over Time

- E4 Period
  - binds together related phenomena
  - introduces inclusion topologies - parts etc.
  - Is confined in space and time
  - the basic unit for temporal-spatial reasoning
- E5 Event
  - looks at the input and the outcome
  - introduces participation of people and presence of things
  - the basic unit for weak causal reasoning
  - each event is a period if we study the details of the process
- E7 Activity
  - adds intention, influence and purpose
  - adds tools

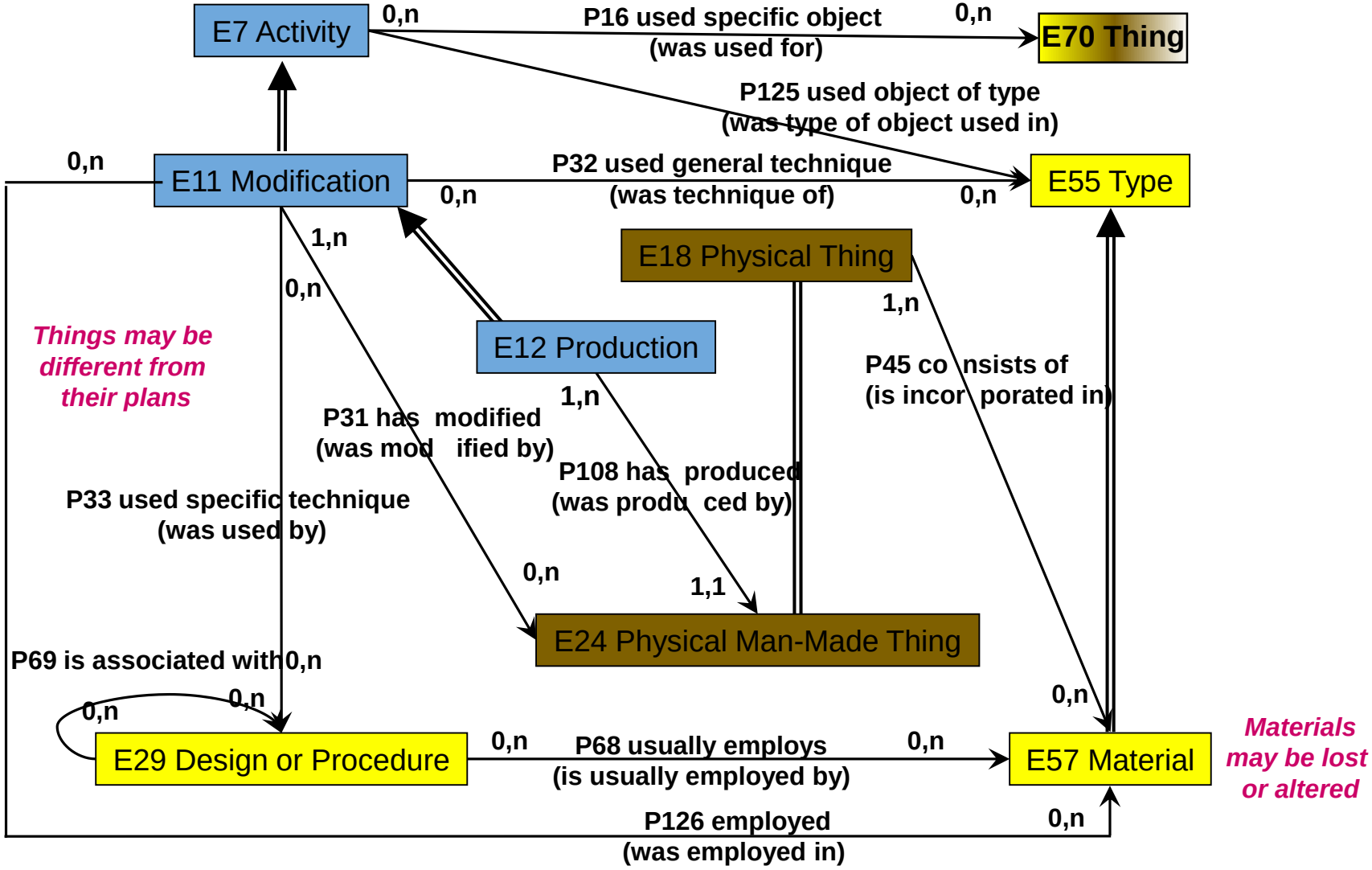
Temporal Classes: Reasoning over Time

Class	Property	Range Class
E2 Temporal Entity	P4 has time-span (is time-span of)	E52 Time-Span
E4 Period	P7 took place at (witnessed)	E53 Place
	P9 consists of (forms part of)	E4 Period
	P10 falls within (contains)	E4 Period
E5 Event	P12 occurred in the presence of (was present at)	E77 Persistent Item
	P11 had participant (participated in)	E39 Actor
	P14 carried out by (performed)	E39 Actor
	P20 had specific purpose (was purpose of)	E5 Event
	P21 had general purpose (was purpose of)	E55 Type
	P16 used specific object (was used for)	E70 Thing

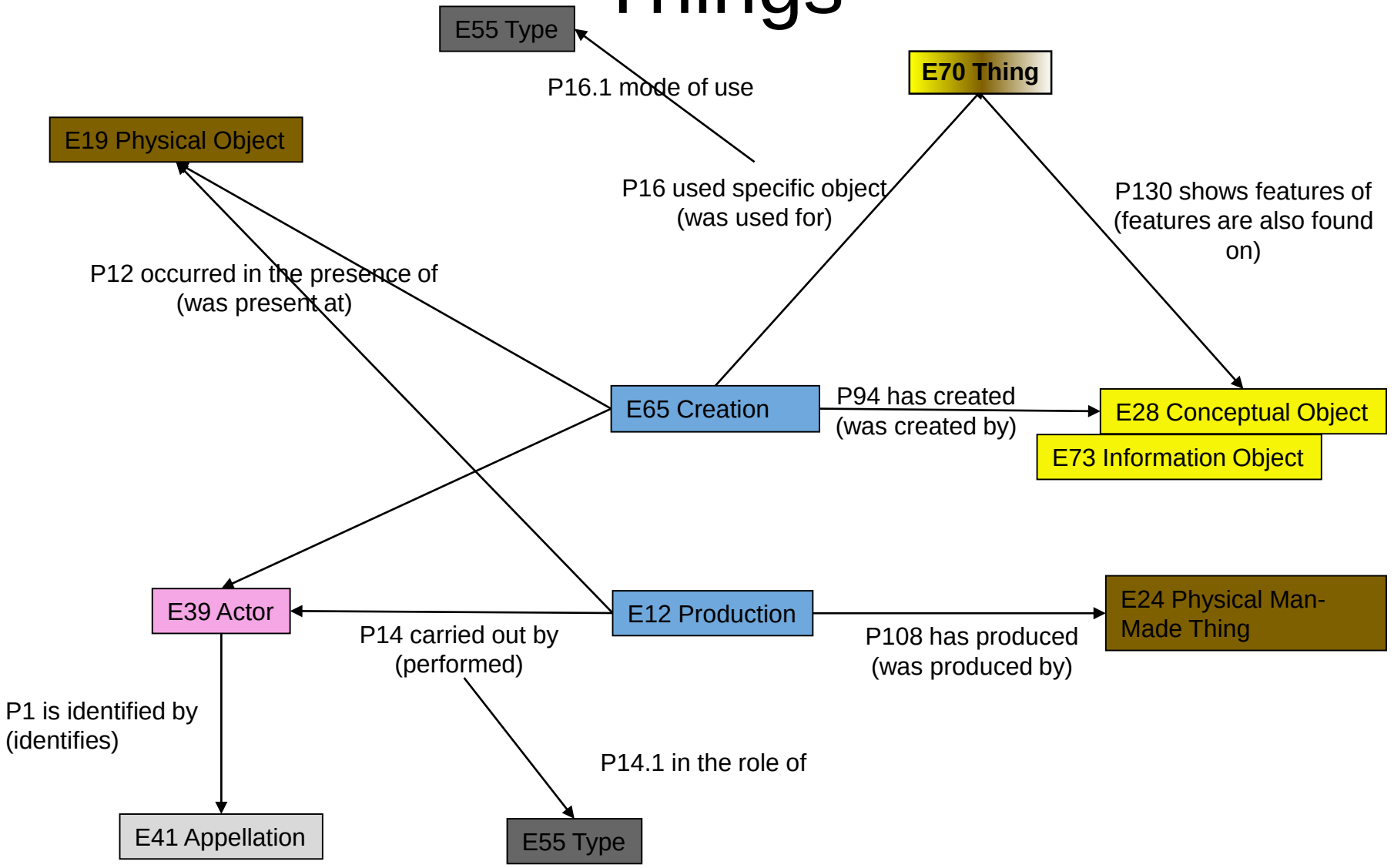
Endurant Classes: Reasoning over Things



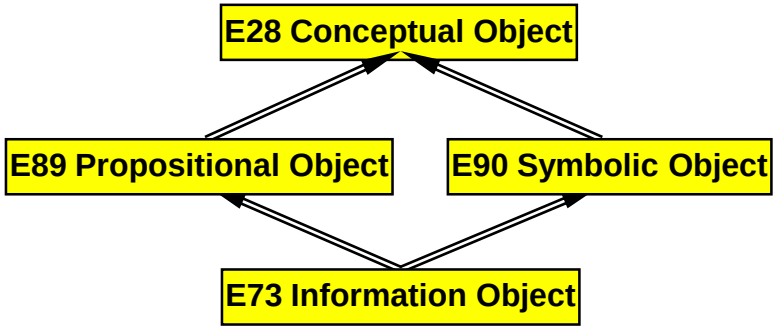
Things Coming to be in Time



Genesis of Conceptual vs. Physical Things



High Level Distinctions/Connections in Conceptual Objects



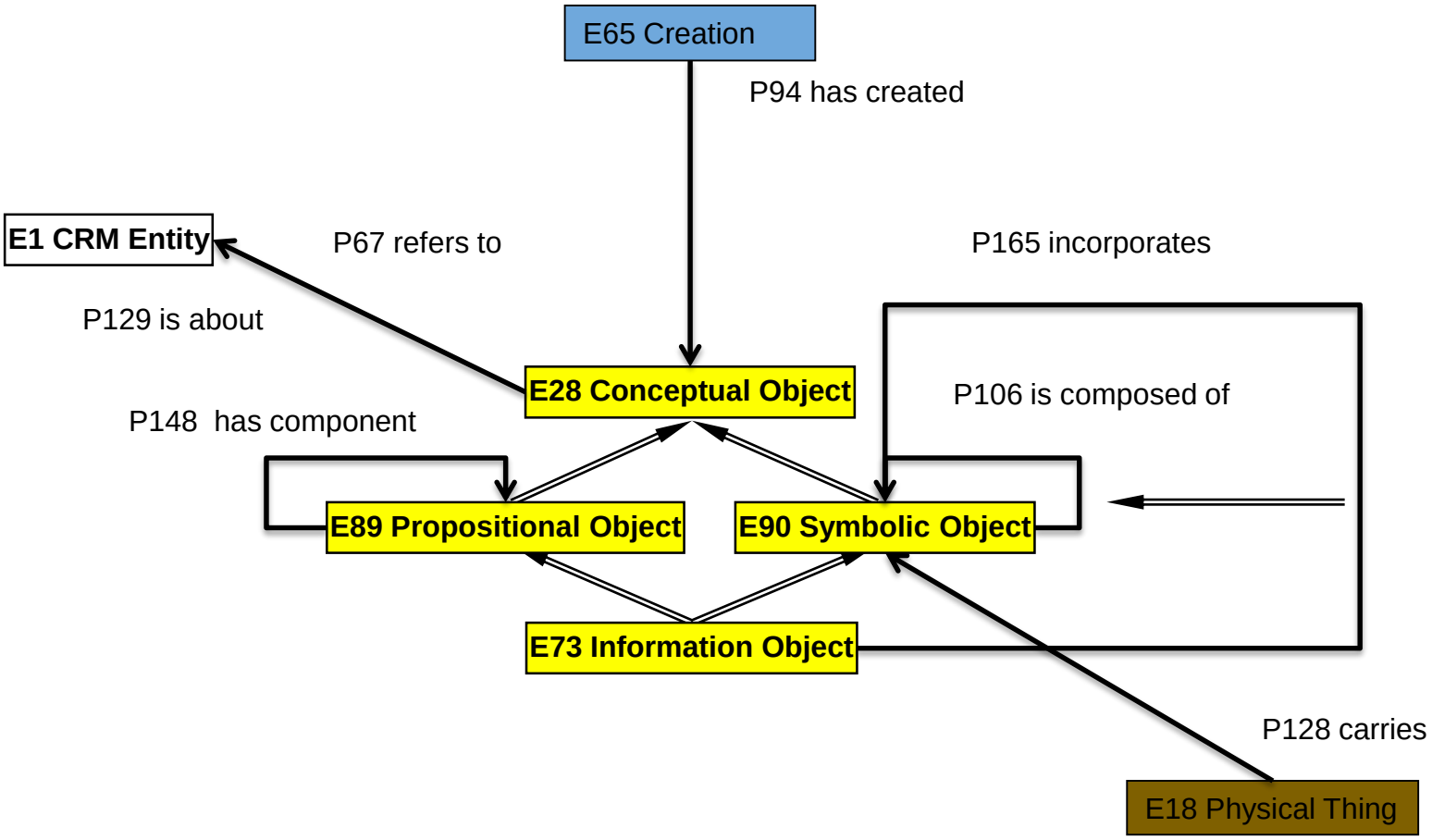
E28 Conceptual Object
'This class comprises non-material products of our minds and other human produced data that have become objects of a discourse about their identity, circumstances of creation or historical implication.'

E90 Symbolic Object
'This class comprises identifiable symbols and any aggregation of symbols, such as characters, identifiers, traffic signs, emblems, texts, data sets, images, musical scores, multimedia objects, computer program code or mathematical formulae that have an objectively recognizable structure and that are documented as single units.

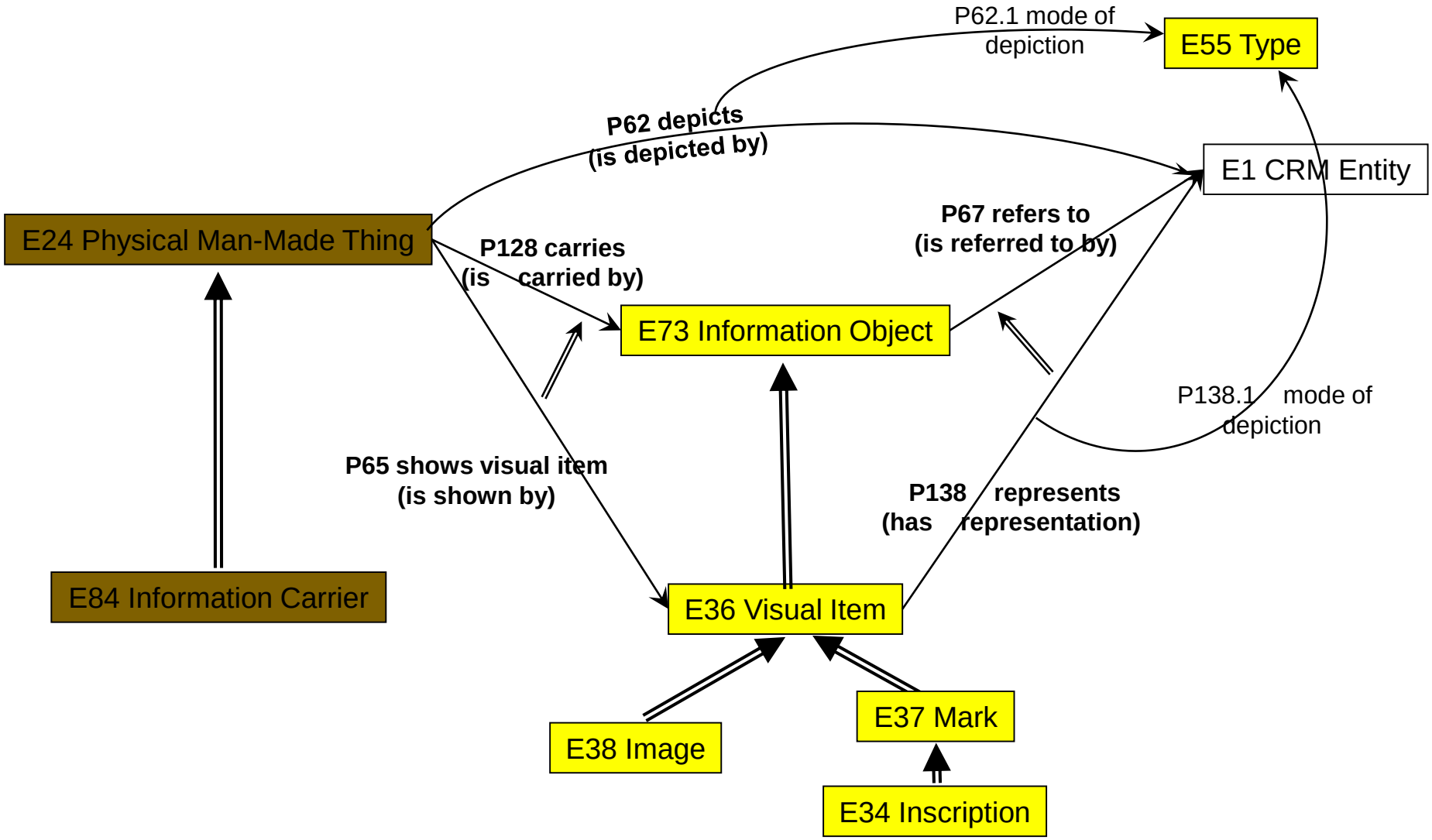
E89 Propositional Object
This class comprises immaterial items, including but not limited to stories, plots, procedural prescriptions, algorithms, laws of physics or images that are, or represent in some sense, sets of propositions about real or imaginary things and that are documented as single units or serve as topics of discourse.

E73 Information Object
This class comprises identifiable immaterial items, such as a poems, jokes, data sets, images, texts, multimedia objects, procedural prescriptions, computer program code, algorithm or mathematical formulae, that have an objectively recognizable structure and are documented as single units.

High Level Relations between Conceptual Objects



Conceptual, Physical Relations and Reasoning



4. CIDOC CRM – HOW TO USE

Tools for Learning the CRM

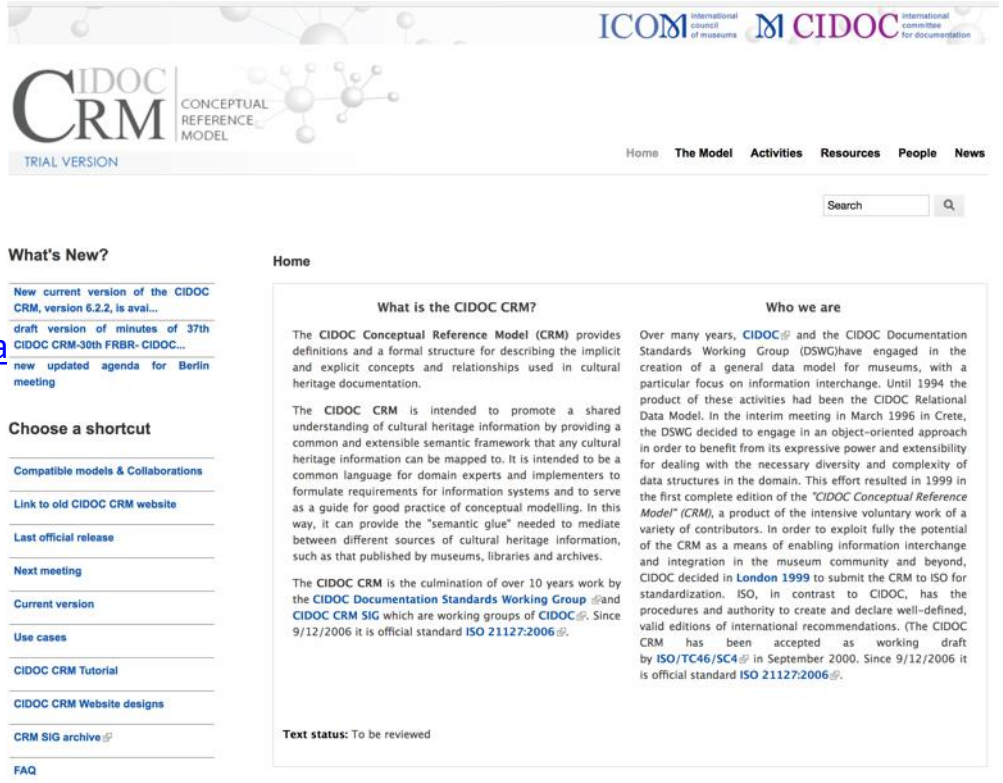
Reference Materials

- CIDOC CRM Specification

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



Tutorials

- One video
- Many powerpoints

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

- Mailing list

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

Tools for Using the CRM

Mapping Tools

- 3M (http://139.91.183.3/3M/)

Data Entry / Navigation Tools

- Wiss-ki Project (http://wiss-ki.eu/)
- Research Space Project (http://www.researchspace.org/)

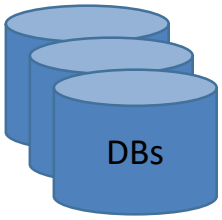
N.B.: CIDOC CRM is neutral with respect to software. The above represent projects that actively deploy technologies to run CIDOC CRM and thus have an experience base.

Workflow

Heterogeneous, Related Datasets

Standard, Ontology

X3ML Mapping Files



Domain Specialist

IT Specialist

Identify Common Need

Agree target schema

Learn target schema

Create mappings

Implements generator

Transform Data

Explore Harmonized Data

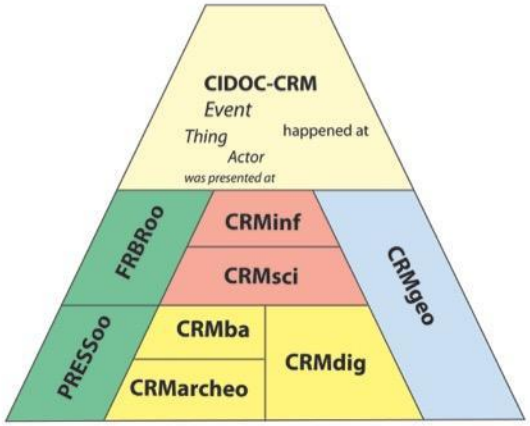
a. Initial Setup

b. Occasional Review

c. Scheduled Ingests and Updates

Mapping Path

Ontology



Encoding

KIF, RDFS, OWL

query

Source Data

dlib

KE EMu ELECTRONIC MUSEUM

PAST PERFECT MUSEUM SOFTWARE

omeka

LibLime

ExLibris The bridge to knowledge

map

Mapping / Transformation Tool

3M Mapping Memory Manager

transform

Triple Store / Graph DB

blazegraph by SYSTAP

neo4j

metaphacts

ontotext

What do I need to a mapping?

The Tools

- Planning
  - Pen and Paper
  - Time
  - My Schema / Examples
  - Ontology Schema / Examples

- Executing
  - A Mapping software
  - Time
  - 3M @ http://139.91.183.3/3M/Login

Mapping
Memory
Manager

George Bruseker

Main Menu

Mappings

Help

Quick Start Guide
Manual

More

Search

Mappings

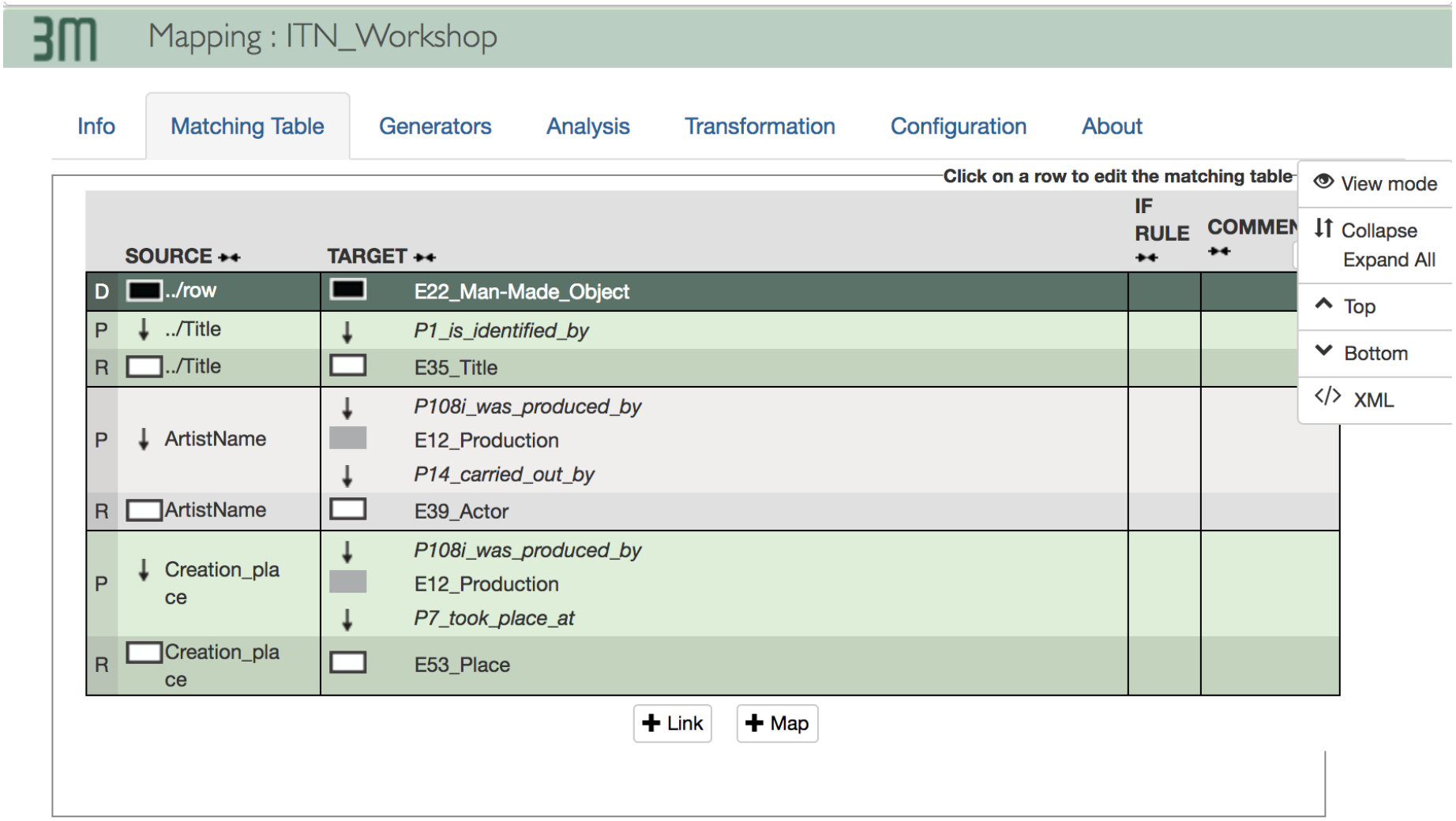
Showing: All

Filter Table

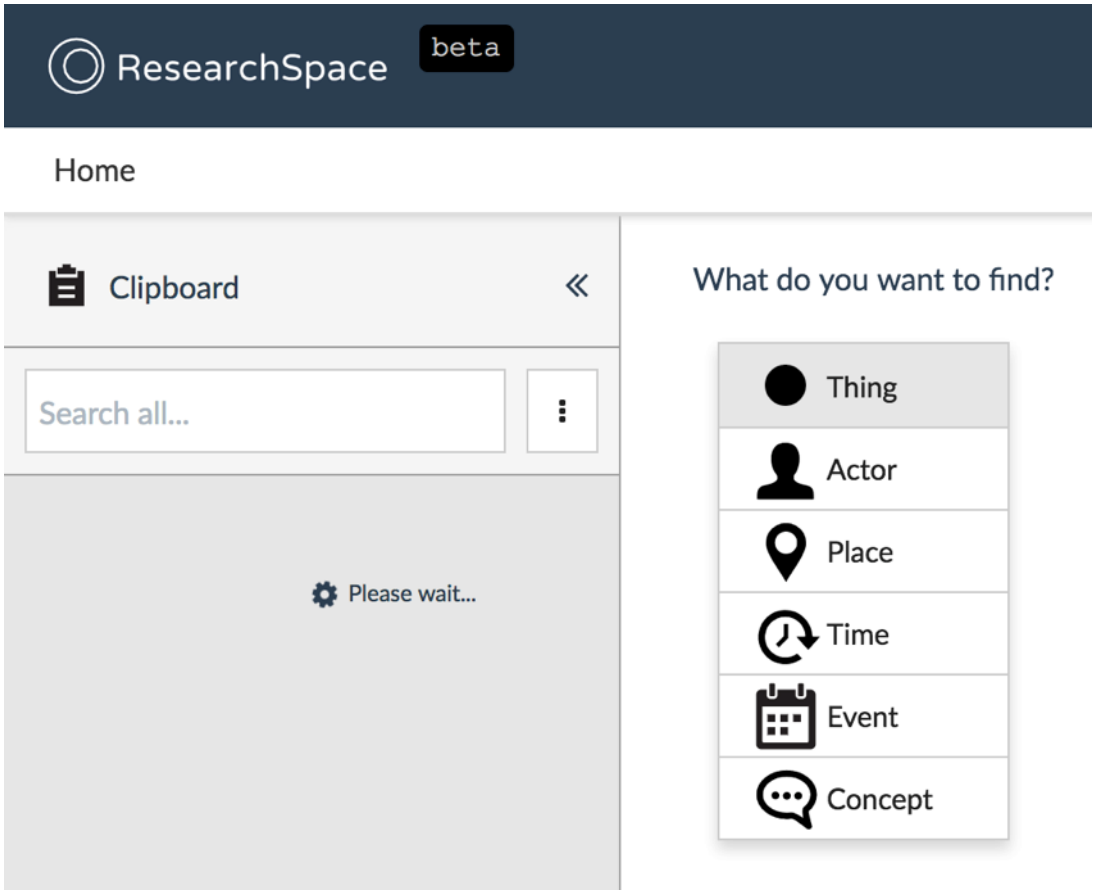
Showing 10 **entries**

Title	Source Schema	Target Schema	General Description	Creator	Card Status	Id	
LIDO to CIDOC (version 2.0)	LIDO	CIDOC-CRM	The mapping was implemented within the ATHENA project (www.athenaeurope.org). The ATHENA project is bringing together re...	admin	Unpublished	Mapping/1	
Mapping of iDAIfield_Ausgrabung to CIDOC-CRM		CIDOC-CRM crmarchaeo crmsci	Mapping of the excavation table of DAI's field research infrastructure iDAI.field.	math	Unpublished	Mapping/103	
A mapping of LIDO V1.0 to CIDOC 5.0.4		CIDOC-CRM	Descriptive Metadata Only	lida	Unpublished	Mapping/106	
Thunau		CIDOC-CRM		anja	Unpublished	Mapping/11	
YCBA LIDO v.2 to CRM		CIDOC-CRM	This is for our YCBA's version 2 of LIDO that includes provenance (blurb/dsiplay label only), exhibition history, biblio...	edgartdata	Unpublished	Mapping/110	
Training1		CIDOC-CRM		smengler	Unpublished	Mapping/116	
OpenDig	OpenDig	CIDOC-CRM CRMarchaeo	OpenDig Archaeological Recording Application	mlvincent	Unpublished	Mapping/119	
DC to CIDOC-CRM	Dublin Core	CIDOC-CRM CRMroles	Mapping of Dublin Core to CIDOC-CRM	math	Unpublished	Mapping/12	
Test Bruseker Instance		CIDOC-CRM		bruseker	Unpublished	Mapping/121	

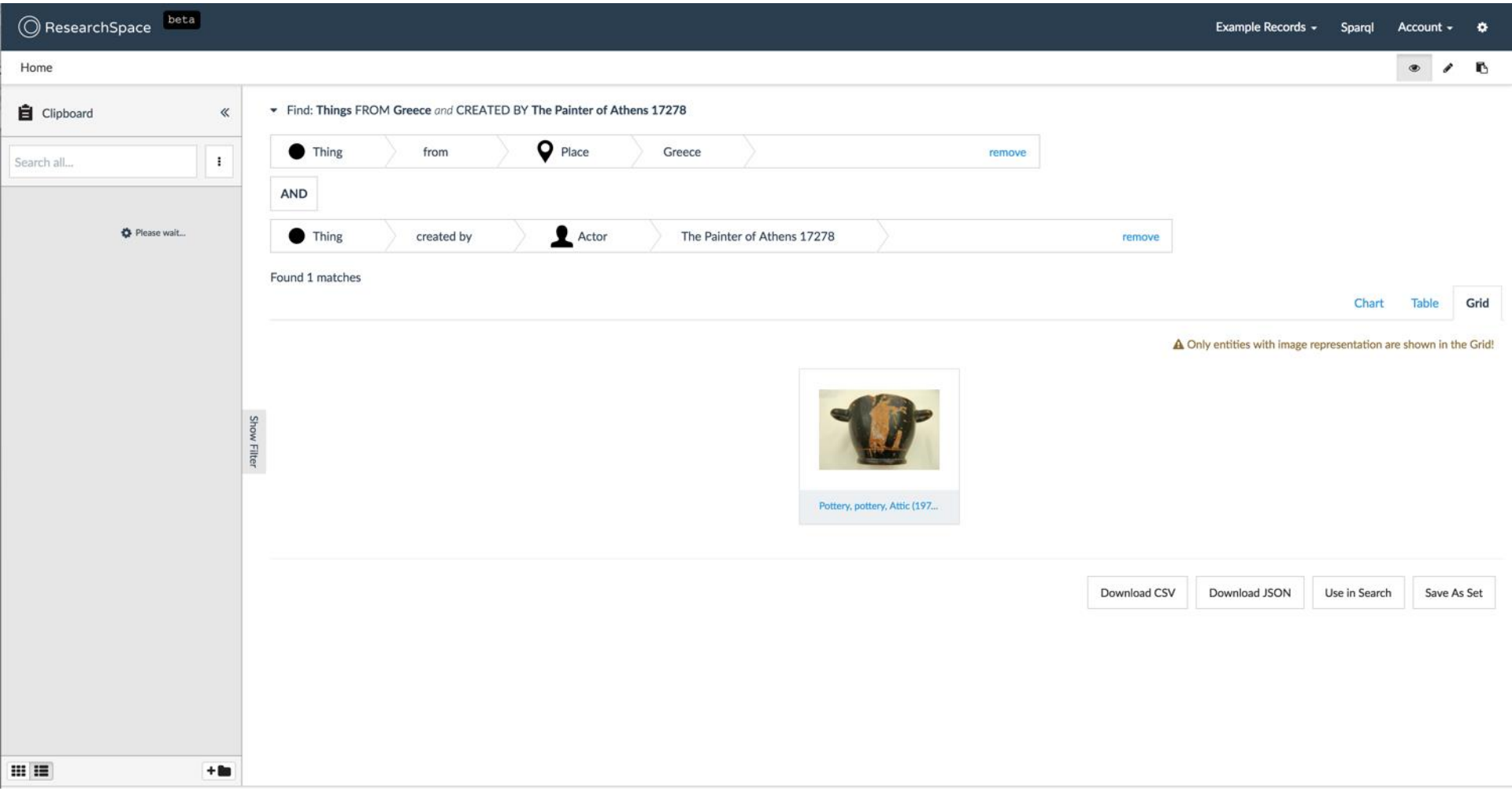
Mapping Interface



Metaphacts: Query Tool



Metaphacts: Complex Query



Metaphacts: Item Level View

ResearchSpace beta

Example Records | Sparql | Account

Home / Pottery, pottery, Attic (1978,0117.5)

Clipboard

Search all...

Please wait...

Pottery, pottery, Attic (1978,0117.5)

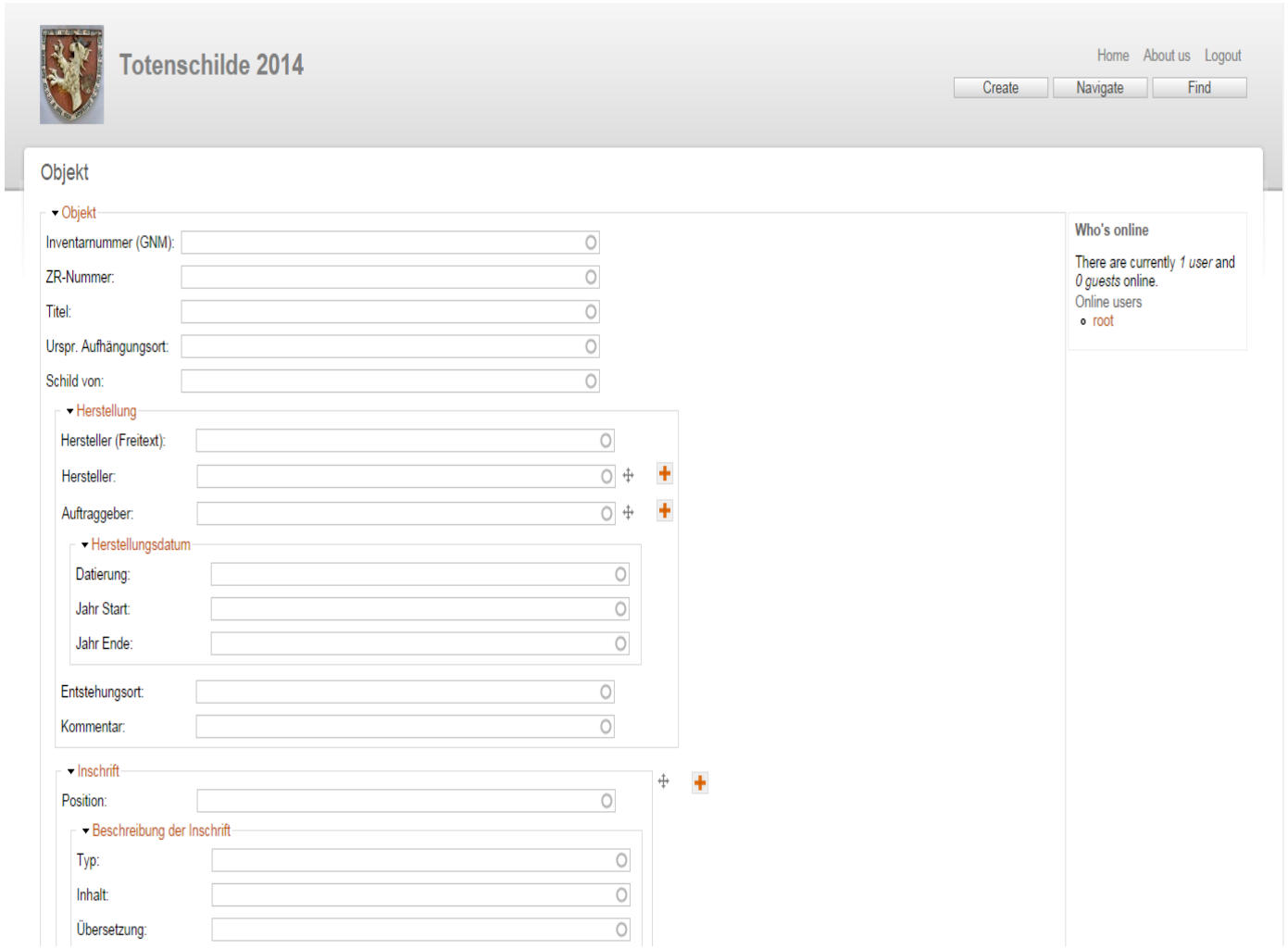
Object (Man-Made Object) record, British Museum

Summary | Annotations 0 | Explore Related

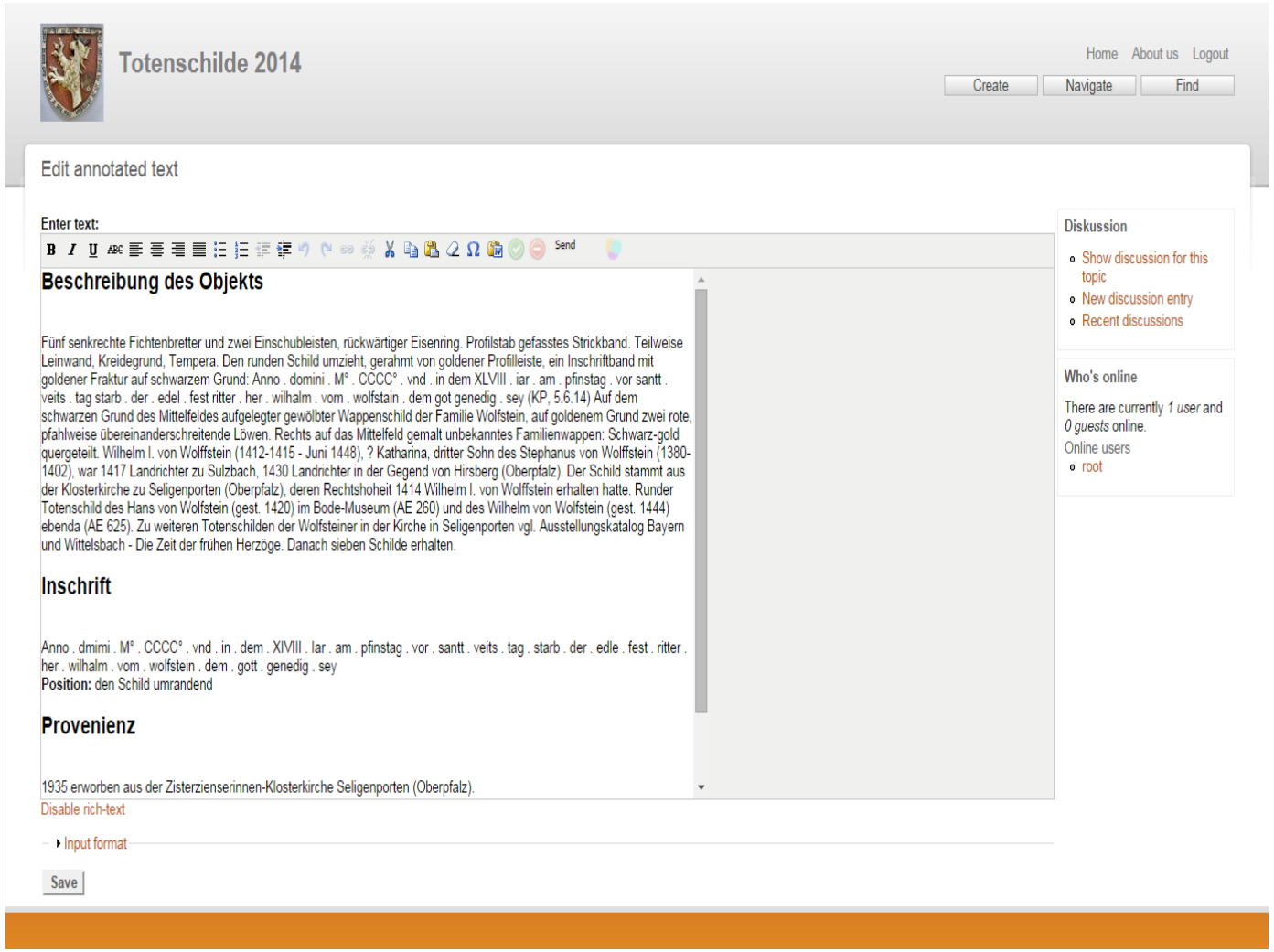
Field	Value	Annotations	Assertions
Title	Pottery, pottery, Attic (1978,0117.5)	0	0
Description	Pottery: red-figured skyphos: youth and pillar.	0	0
Object Type	skyphos	0	0
Asset ID	GAA7154	0	0
Material	pottery (en)	0	0
Dimensions	Diameter: 16.7 Centimeter Height: 12.4 Centimeter	0	0
Place of Production	Attica Greece	0	0
Date of Production	420BC :: circa	0	0
Production Period	Attic	0	0
Production Technique	painted	0	0
Production School	The Painter of Athens 17278	0	0
Current Owner	The British Museum	0	0

Object summary | All properties

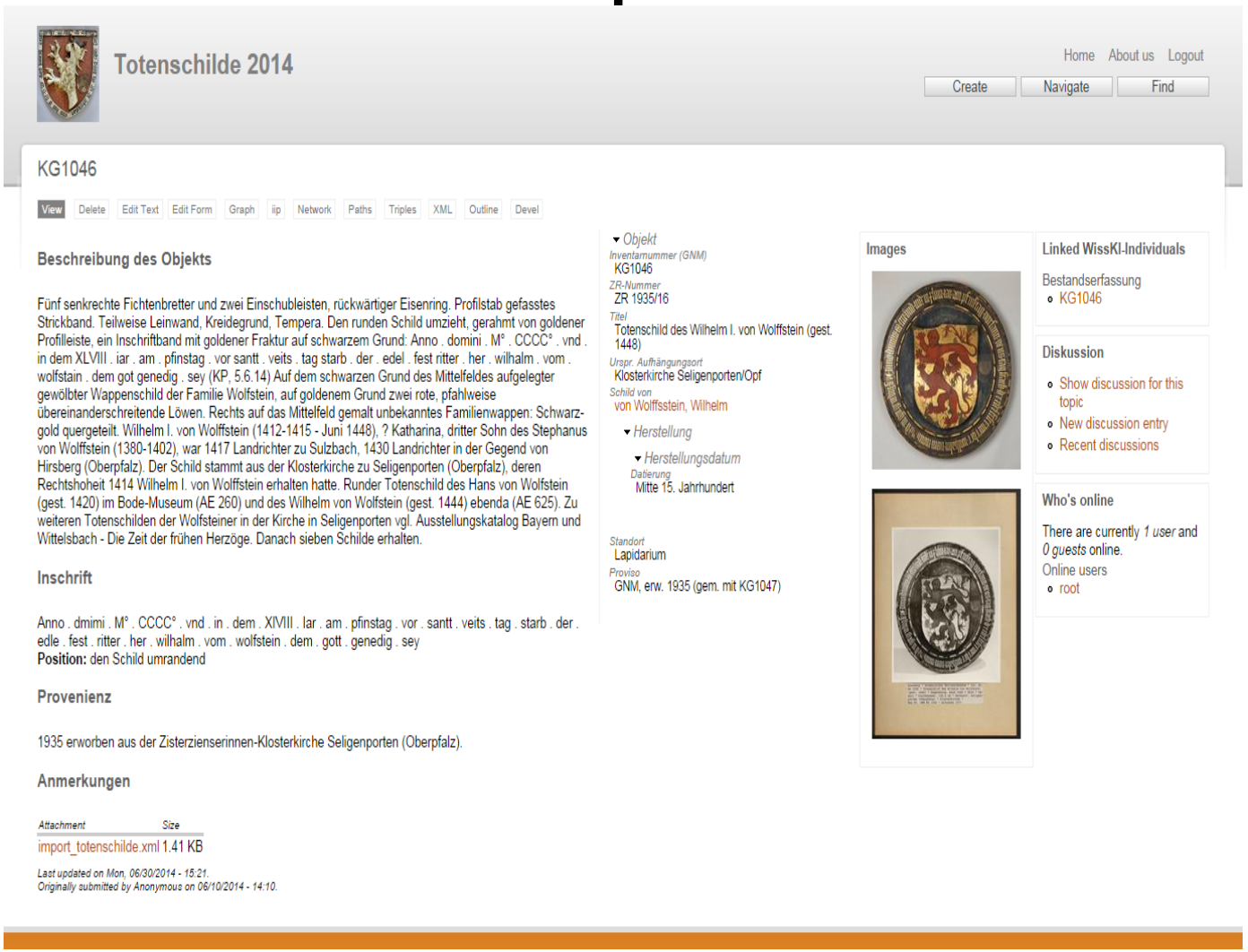
Form based data acquisition



Text based data acquisition



Data presentation



Totenschilde 2014

Home About us Logout

Create Navigate Find

KG1046

View Delete Edit Text Edit Form Graph iip Network Paths Triples XML Outline Devel

Beschreibung des Objekts

Fünf senkrechte Fichtenbretter und zwei Einschubleisten, rückwärtiger Eisenring. Profilstab gefasstes Strickband. Teilweise Leinwand, Kreidegrund, Tempera. Den runden Schild umzieht, gerahmt von goldener Profilleiste, ein Inschriftband mit goldener Fraktur auf schwarzem Grund: Anno . domini . M° . CCCC° . vnd . in dem XLVIII . iar . am . pfinstag . vor santt . veits . tag starb . der . edel . fest ritter . her . wilhalm . vom . wolfstain . dem got genedig . sey (KP, 5.6.14) Auf dem schwarzen Grund des Mittelfeldes aufgelegter gewölbter Wappenschild der Familie Wolfstein, auf goldenem Grund zwei rote, pfahlweise übereinanderschreitende Löwen. Rechts auf das Mittelfeld gemalt unbekanntes Familienwappen: Schwarz-gold quergeteilt. Wilhelm I. von Wolffstein (1412-1415 - Juni 1448), ? Katharina, dritter Sohn des Stephanus von Wolffstein (1380-1402), war 1417 Landrichter zu Sulzbach, 1430 Landrichter in der Gegend von Hirsberg (Oberpfalz). Der Schild stammt aus der Klosterkirche zu Seligenporten (Oberpfalz), deren Rechtshoheit 1414 Wilhelm I. von Wolffstein erhalten hatte. Runder Totenschild des Hans von Wolfstein (gest. 1420) im Bode-Museum (AE 260) und des Wilhelm von Wolfstein (gest. 1444) ebenda (AE 625). Zu weiteren Totenschilden der Wolfsteiner in der Kirche in Seligenporten vgl. Ausstellungskatalog Bayern und Wittelsbach - Die Zeit der frühen Herzöge. Danach sieben Schilde erhalten.

Inschrift

Anno . dmimi . M° . CCCC° . vnd . in . dem . XIVIII . iar . am . pfinstag . vor . santt . veits . tag . starb . der . edle . fest . ritter . her . wilhalm . vom . wolfstein . dem . gott . genedig . sey
Position: den Schild umrandend

Provenienz

1935 erworben aus der Zisterzienserinnen-Klosterkirche Seligenporten (Oberpfalz).

Anmerkungen

Attachment	Size
import_totenschilde.xml	1.41 KB

Last updated on Mon, 06/30/2014 - 15:21.
Originally submitted by Anonymous on 06/10/2014 - 14:10.

▾ *Objekt*
Inventarnummer (GNM)
KG1046
ZR-Nummer
ZR 1935/16
Titel
Totenschild des Wilhelm I. von Wolffstein (gest. 1448)
Urspr. Aufhängungsort
Klosterkirche Seligenporten/Opf
Schild von
von Wolffsstein, Wilhelm

▾ *Herstellung*

▾ *Herstellungsdatum*
Datierung
Mitte 15. Jahrhundert

Standort
Lapidarium
Proviso
GNM, erw. 1935 (gem. mit KG1047)

Images

Linked WissKI-Individuals

Bestandserfassung
- KG1046

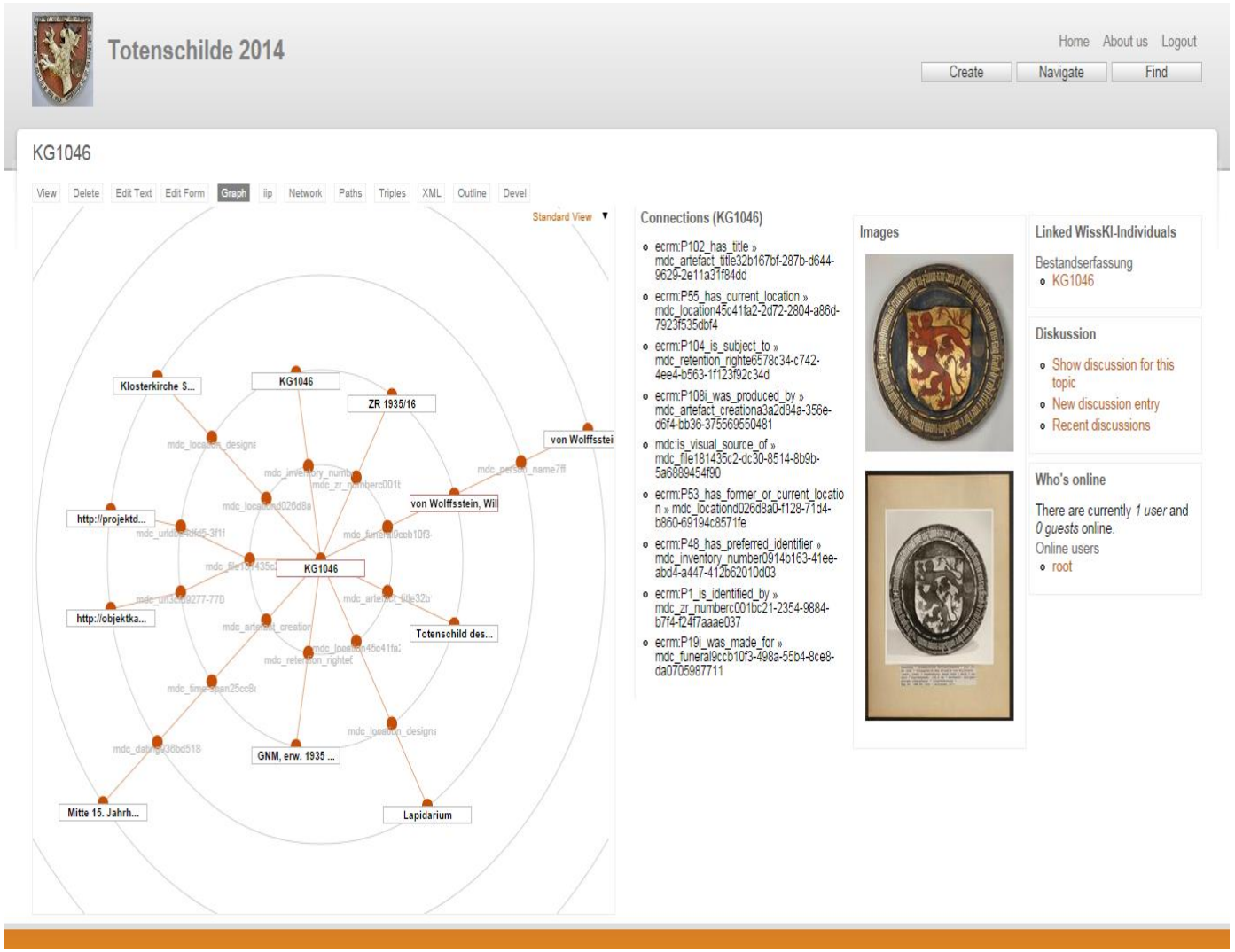
Diskussion

- Show discussion for this topic
- New discussion entry
- Recent discussions

Who's online

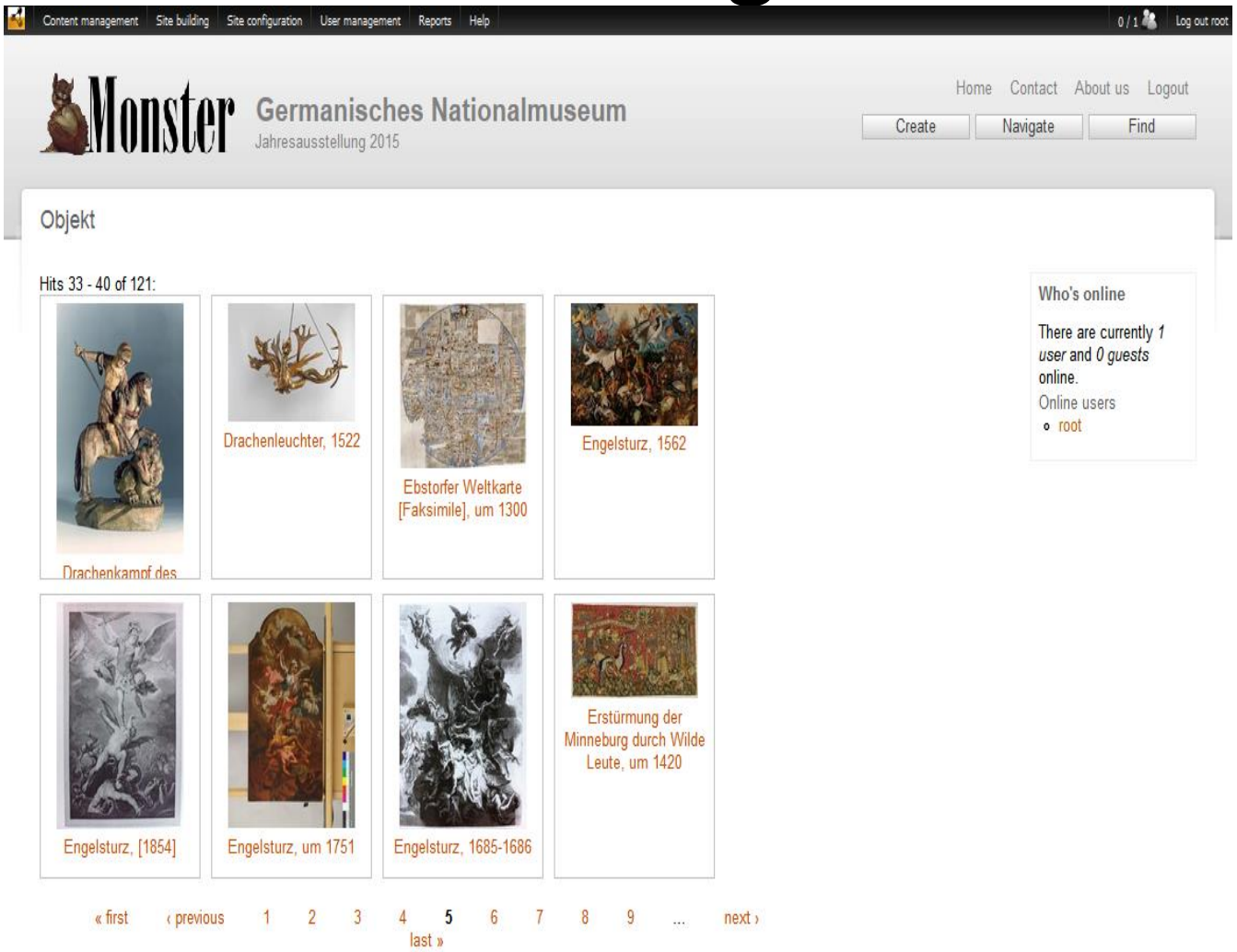
There are currently *1 user* and *0 guests* online.
Online users
- root

OWL/XML in the triplestore



WissKI

Navigation



Full-text search and detailed search