The CIDOC CRM, a Standard for the Integration of Cultural Information

Stephen Stead

CIDOC Conceptual Reference Model Special Interest Group

ICS-FORTH, Crete, Greece
November, 2008

The CIDOC CRM
Outline

- **Problem statement – information diversity**

- **Motivation example – the Yalta Conference**

- **The goal and form of the CIDOC CRM**

- **Presentation of contents**

- **About using the CIDOC CRM**

- **State of development**

- **Conclusion**

The CIDOC CRM
Cultural Diversity and Data Standards

❏ **Cultural information is more than a domain:**

◆ Collection description (art, archeology, natural history….)

◆ Archives and literature (records, treaties, letters, artful works..)

◆ Administration, preservation, conservation of material heritage

◆ Science and scholarship – investigation, interpretation

◆ Presentation – exhibition making, teaching, publication

❏ **But how to make a documentation standard?**

◆ Each aspect needs its methods, forms, communication means

◆ Data overlap, but do not fit in one schema

◆ Understanding lives from relationships, but how to express them?

The CIDOC CRM
Historical Archives….

Type:	**Text**
Title:	**Protocol of Proceedings of Crimea Conference**
Title.Subtitle:	**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
Publisher:	**State Department**
Subject:	**Postwar division of Europe and Japan**

Metadata

Documents

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…… "**

The CIDOC CRM
Images, non-verbose…

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

The CIDOC CRM
Places and Objects

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

About…

Title: Yalta, Crimean Peninsula
Publisher: Kurgan-Lisnet
Source: Liaison Agency

The CIDOC CRM
The Integration Problem (1)

❑ **Problem 1- Identity:**

◆ **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

The CIDOC CRM
The Integration Problem (2)

❑ **Problem 2- hidden entities (typically "title"):**

◆ **Actors**
— Allied leader, Allied power
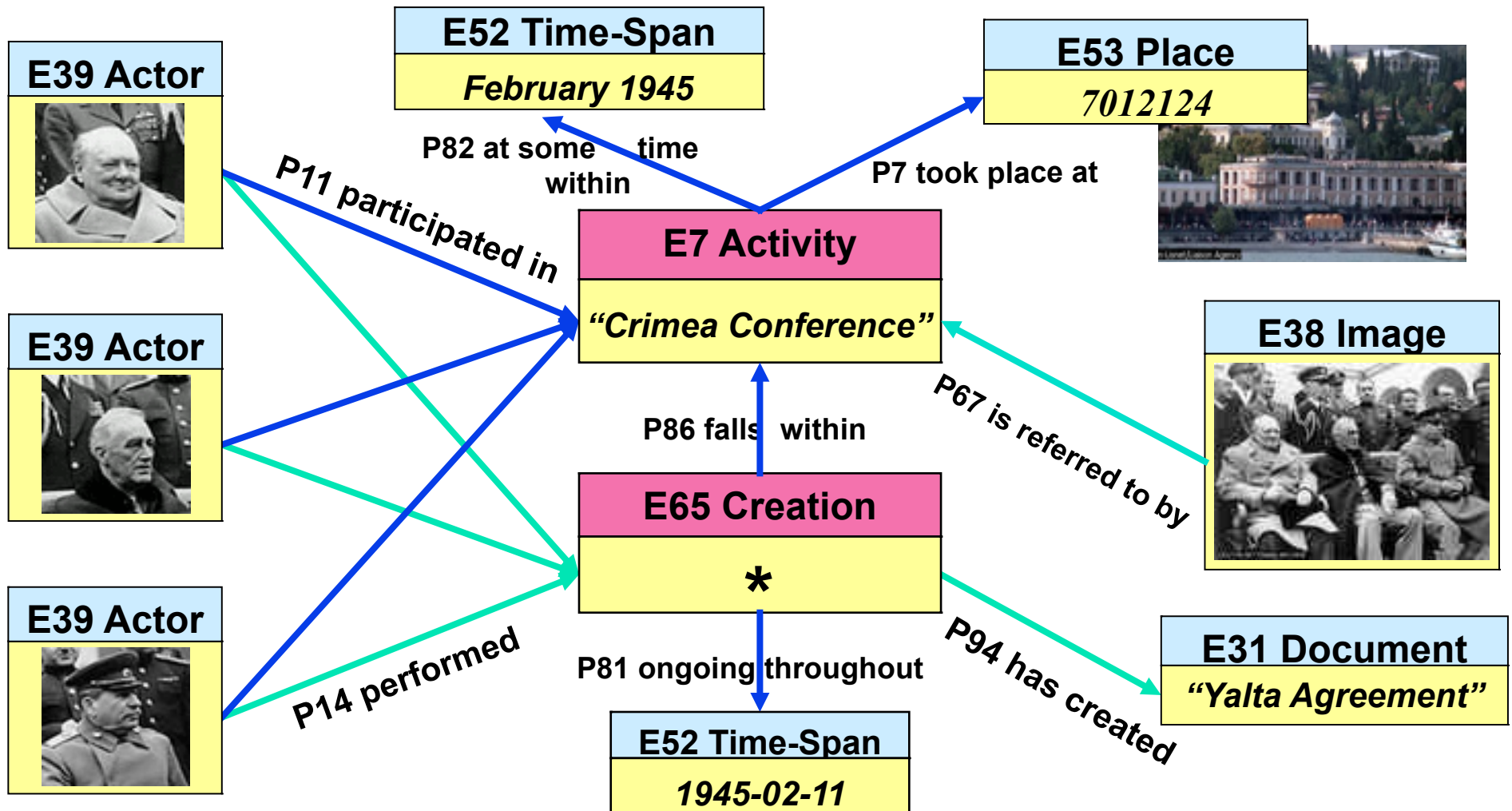
◆ **Places**
— Yalta, Crimea

◆ **Events**
— Crimea Conference, "Allied Leaders at Yalta","… conference between Allied powers" "Postwar division"

❑ **Solution:**

◆ **Change metadata structures: but what are the relevant elements?**

The CIDOC CRM
Explicit Events, Object Identity, Symmetry

E52 Time-Span
February 1945

E53 Place
7012124

E39 Actor

E39 Actor

E39 Actor

E7 Activity
"Crimea Conference"

E38 Image

E65 Creation
*

E31 Document
"Yalta Agreement"

E52 Time-Span
1945-02-11

P11 participated in

P82 at some time within

P7 took place at

P67 is referred to by

P86 falls within

P14 performed

P81 ongoing throughout

P94 has created

The CIDOC CRM...

❑ ...captures the underlying semantics of relevant documentation structures in a **formal ontology**

❑ Ontologies are **formalized knowledge**: clearly defined concepts and relationships about **possible states of affairs** in a domain

❑ They can be understood by people and processed by machines to enable data exchange, data integration, query mediation etc.

❑ Semantic interoperability in cultural heritage can be achieved with an **"extensible ontology of relationships"** and explicit **event** modeling

❑This provides **shared explanation** rather than the prescription of a common data structure

❑ The ontology is the **language** that S/W developers and museum experts can **share**. Therefore it needed interdisciplinary work. That is what CIDOC has provided

The CIDOC CRM
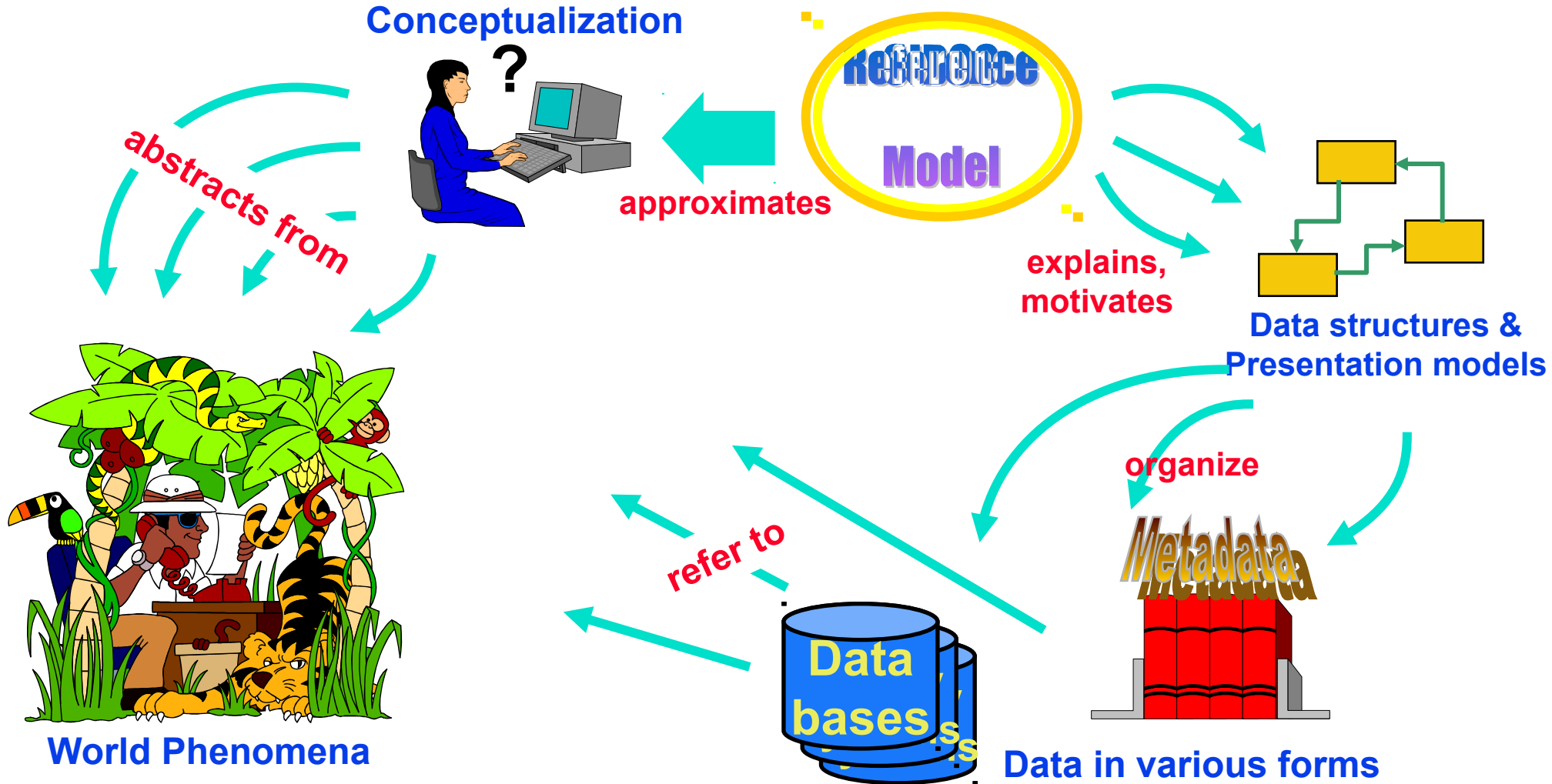## Outcomes

❑ **The CIDOC Conceptual Reference Model**

◆ A **collaboration** with the International Council of Museums

◆ An ontology of 86 classes and 137 properties for **culture** and **more**

◆ With the capacity to **explain** hundreds of (meta)data formats

◆ Accepted by ISO TC46 in September 2000

◆ International standard since 2006 - ISO 21127:2006

❑ **Serving as:**

◆ **intellectual guide** to create schemata, formats, profiles

◆ A language for analysis of existing sources for integration/mediation

"Identify elements with **common meaning**"

◆ **Transportation format** for data integration / migration / Internet

The CIDOC CRM
The Intellectual Role of the CRM

Conceptualization

Reference Model

abstracts from

approximates

explains, motivates

Data structures & Presentation models

organize

refer to

World Phenomena

Metadata

Data bases

Data in various forms
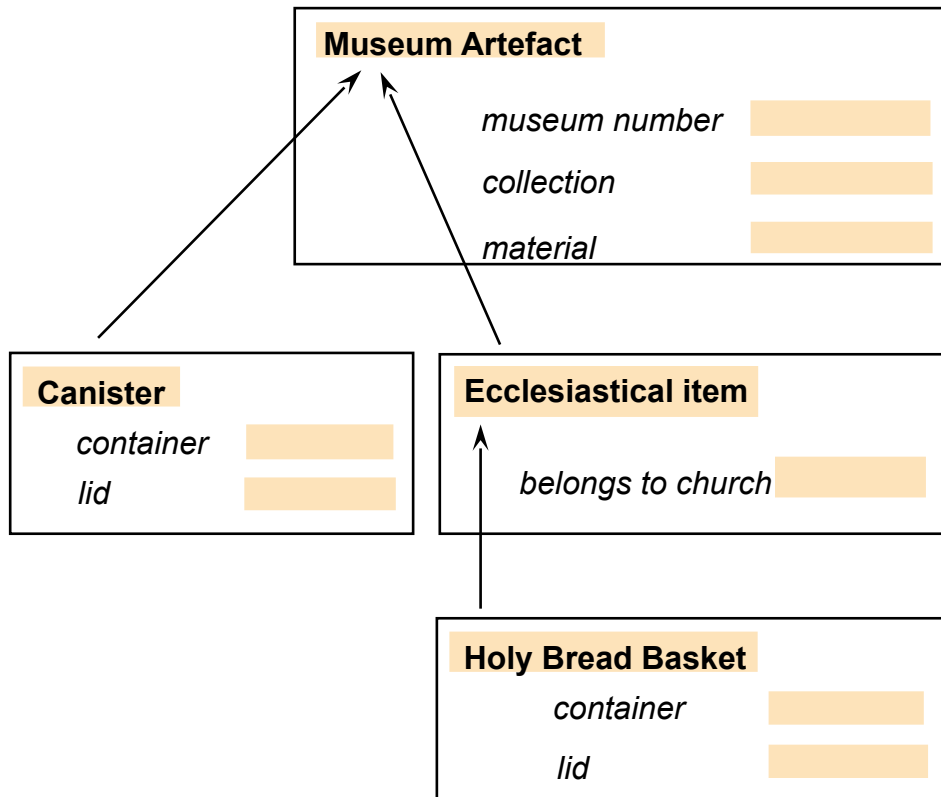
The CIDOC CRM
Encoding of the CIDOC CRM

❑ **The CIDOC CRM is a formal ontology (defined in TELOS)**

   ◆ **But CRM instances can be encoded in many forms: RDBMS, ooDBMS, XML, RDF(S)**

   ◆ **Uses Multiple isA – to achieve uniqueness of properties in the schema**

   ◆ **Uses multiple instantiation – to be able to combine not always valid combinations (e.g. destruction – activity)**

   ◆ **Uses Multiple isA for properties to capture different abstraction of relationships**

❑ **Methodological aspects:**

   ◆ **Entities are introduced as anchors of properties (and if structurally relevant)**

   ◆ **Frequent joins (short-cuts) of complex data paths for data found in different degrees of detail are modeled explicitly**
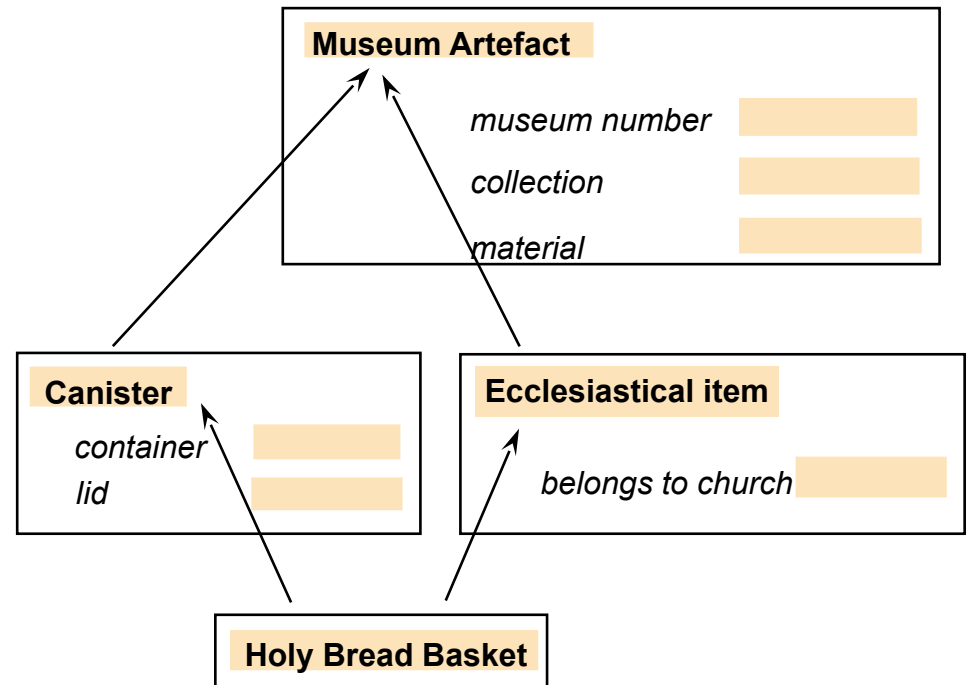
The CIDOC CRM
Justifying Multiple Inheritance

Single Inheritance form:

Museum Artefact

museum number

collection

material

Canister

container

lid

Ecclesiastical item

belongs to church

Holy Bread Basket

container

lid

Repetition of properties

Multiple Inheritance form:

Museum Artefact

museum number

collection

material

Canister

container
lid

Ecclesiastical item

belongs to church

Holy Bread Basket

Unique identity of properties

The CIDOC CRM
Data example (e.g. from extraction)

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   *Multiple Instantiation*

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)

P82 at some time within

1923 - 1928   (entity E61 Time Primitive)

P7 took place at

Greece   (entity E53 Place)

P2 has type

nation (entity E55 Type)

republic (entity E55 Type)
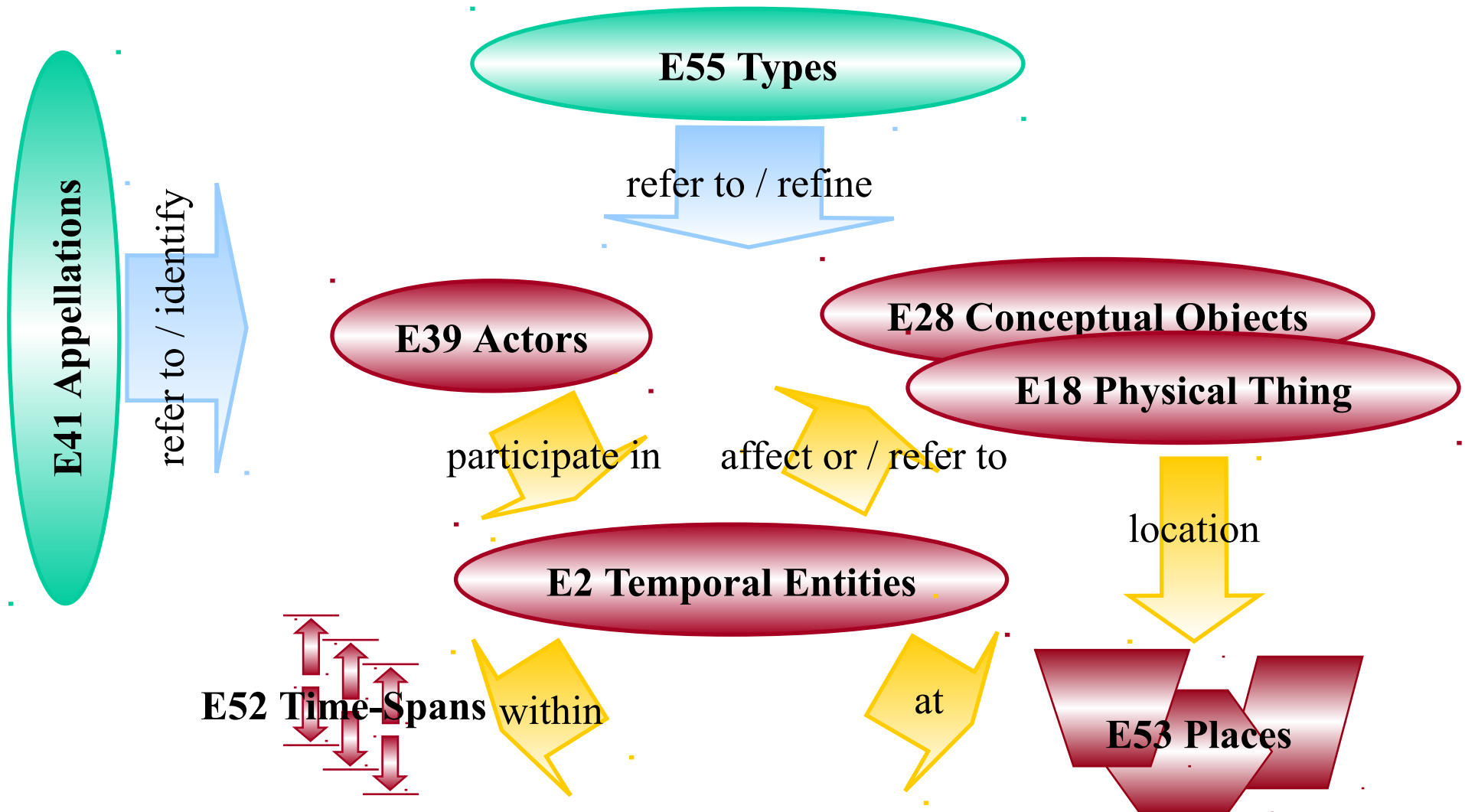
P89 falls within

Europe (entity E53 Place)

P2 has type

continent (entity E55 Type)

TGN data

15

The CIDOC CRM
Top-level classes useful for integration

E55 Types

refer to / refine

E41 Appellations

refer to / identify

E39 Actors

E28 Conceptual Objects

E18 Physical Thing

participate in

affect or / refer to

location

E2 Temporal Entities

E52 Time-Spans within

at

E53 Places

The CIDOC CRM
The types of relationships

◆ **Identification** of real world items by real world names

◆ **Observation** and **Classification** of real world items

◆ **Part-decomposition** and structural properties of Conceptual & Physical Objects, Periods, Actors, Places and Times

◆ **Participation** of persistent items in temporal entities

— creates a notion of history: "world-lines" meeting in space-time

◆ **Location** of periods in space-time and physical objects in space

◆ **Influence** of objects on activities and products and vice-versa

◆ **Reference** of information objects to any real-world item

The CIDOC CRM
The E2 Temporal Entity Hierarchy
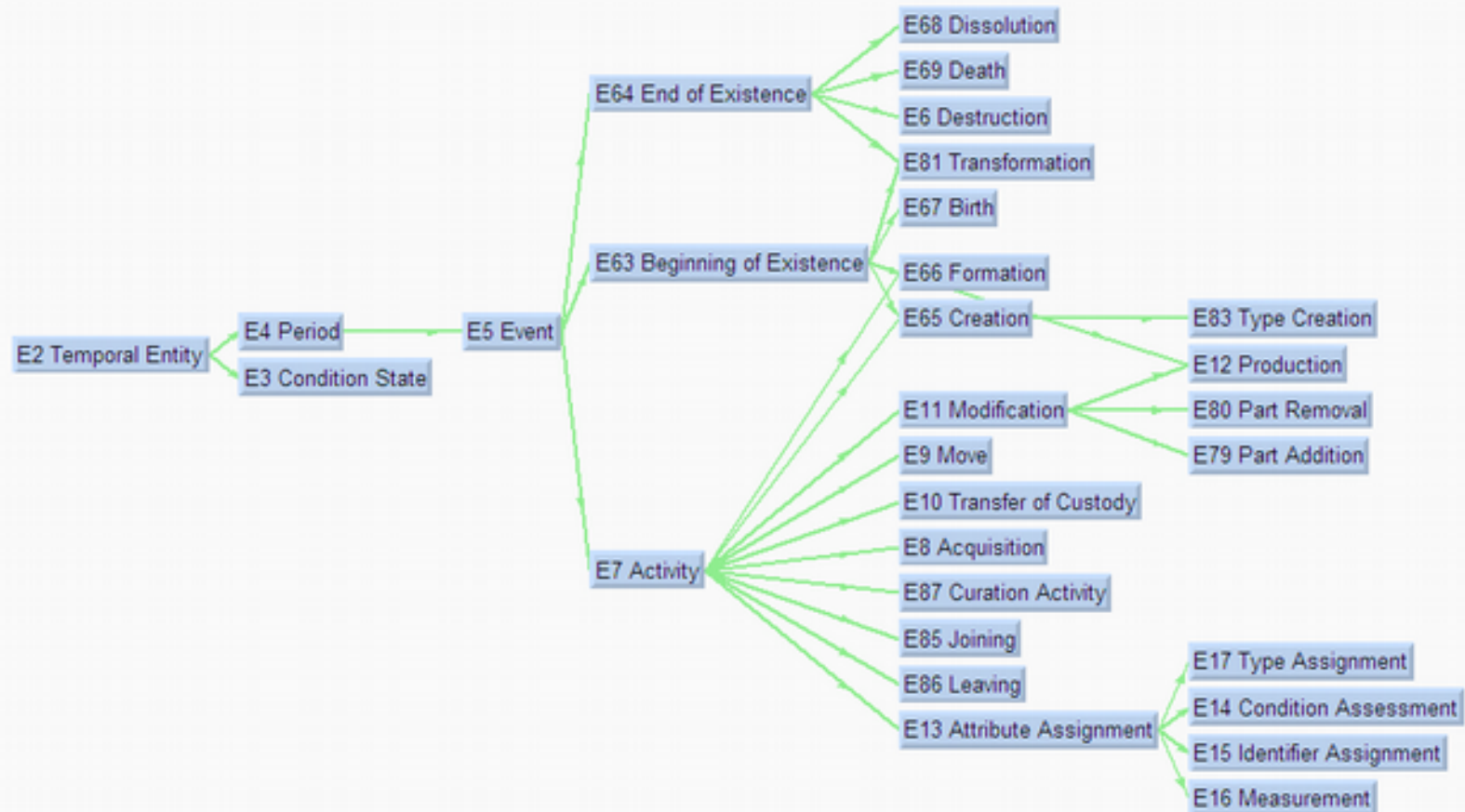
The CIDOC CRM
Scope note example: E2 Temporal Entity

❑**E2 Temporal Entity**

◆**Scope Note:**

This class comprises all phenomena, such as the instances of E4 Periods, E5 Events and states, which happen over a limited extent in time.

In some contexts, these are also called perdurants. This class is disjoint from E77 Persistent Item. This is an abstract class and has no direct instances. E2 Temporal Entity is specialized into E4 Period, which applies to a particular geographic area (defined with a greater or lesser degree of precision), and E3 Condition State, which applies to instances of E18 Physical Thing.

— Is limited in time, is the only link to time, but is not time itself

— spreads out over a place or object

— the core of a model of physical history, open for unlimited specialisation

The CIDOC CRM
Temporal Entity- Subclasses

❑**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 process**

❑ **E7 Activity**

◆ **adds intention, influence and purpose**
◆ **adds tools**

The CIDOC CRM
Temporal Entity- Main Properties

☐ **E2 Temporal Entity**
- ◆ **Properties: P4 has time-span (is time-span of):**   **E52 Time-Span**

☐ **E4 Period**
- ◆ **Properties: P7 took place at (witnessed):**   **E53 Place**
  P9 consists of (forms part of):   **E4 Period**
  P10 falls within (contains):   **E4 Period**

☐ **E5 Event**
- ◆ **Properties: P11 had participant (participated in):**   **E39 Actor**
  P12 occurred in the presence of (was present at): E77 Persistent Item

☐ **E7 Activity**
- ◆ **Properties: P14 carried out by (performed):**   **E39 Actor**
  P20 had specific purpose (was purpose of): E5 Event
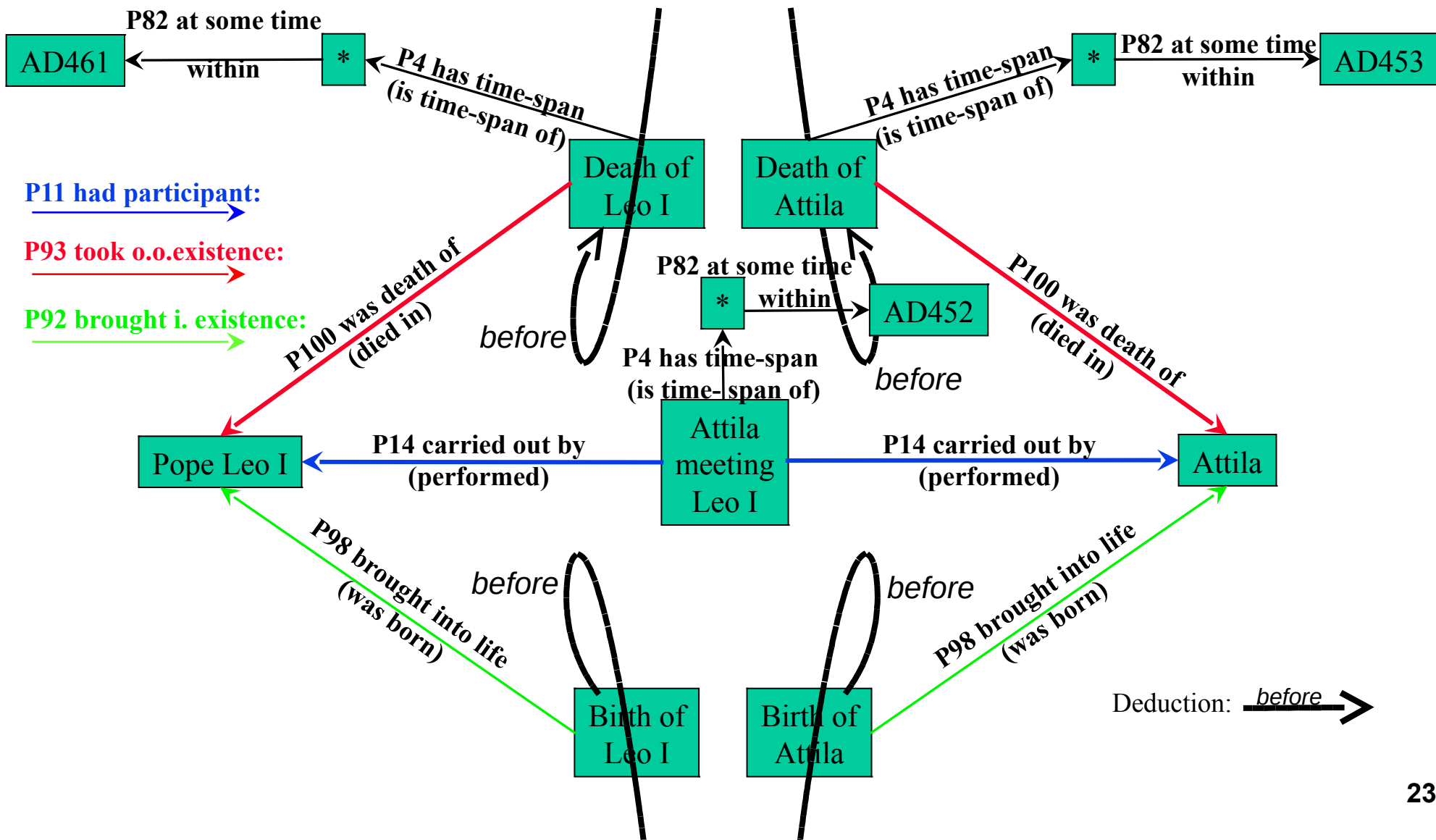 **P21 had general purpose (was purpose of): E55 Type**

The CIDOC CRM
The Participation Properties

The CIDOC CRM
Termini postquem / antequem

P82 at some time within — AD461 ← * ← **P4 has time-span (is time-span of)** — Death of Leo I

Death of Attila — **P4 has time-span (is time-span of)** → * — **P82 at some time within** → AD453

P11 had participant: (blue arrow)

P93 took o.o.existence: (red arrow)

P92 brought i. existence: (green arrow)

Death of Leo I — **P100 was death of (died in)** → Pope Leo I

Death of Attila — **P100 was death of (died in)** → Attila

before

P82 at some time within — * — **within** → AD452

P4 has time-span (is time-span of) — Attila meeting Leo I

before

Pope Leo I ← **P14 carried out by (performed)** — Attila meeting Leo I — **P14 carried out by (performed)** → Attila

P98 brought into life (was born) — Birth of Leo I → Pope Leo I

P98 brought into life (was born) — Birth of Attila → Attila

before

before

Deduction: *before* →

23

The CIDOC CRM
Historical events as meetings
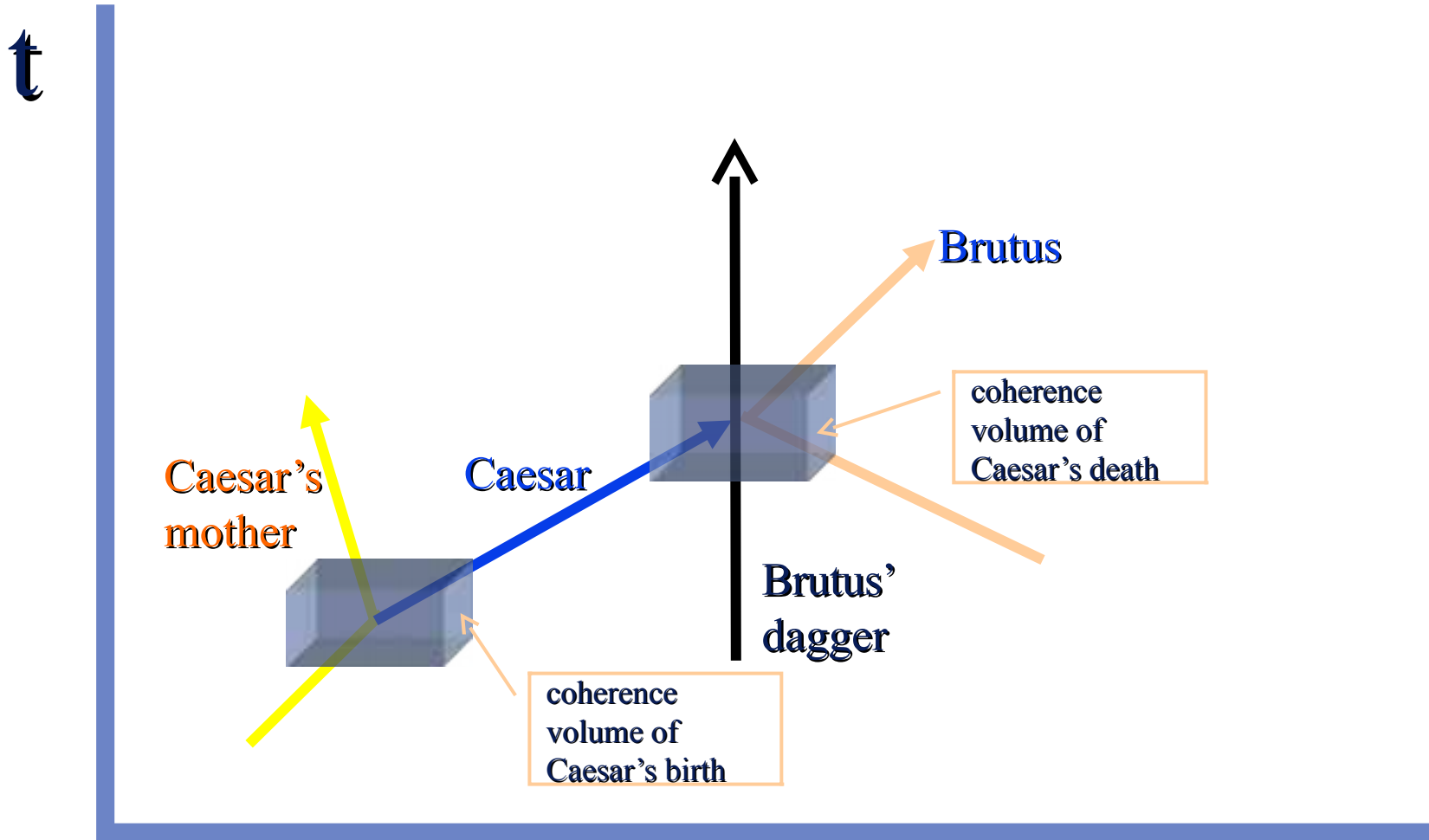
The CIDOC CRM
Depositional events as meetings

t

lava and
ruins

ancient
Santorinian

coherence volume of
volcano eruption

house

volcano

coherence
volume of house
building
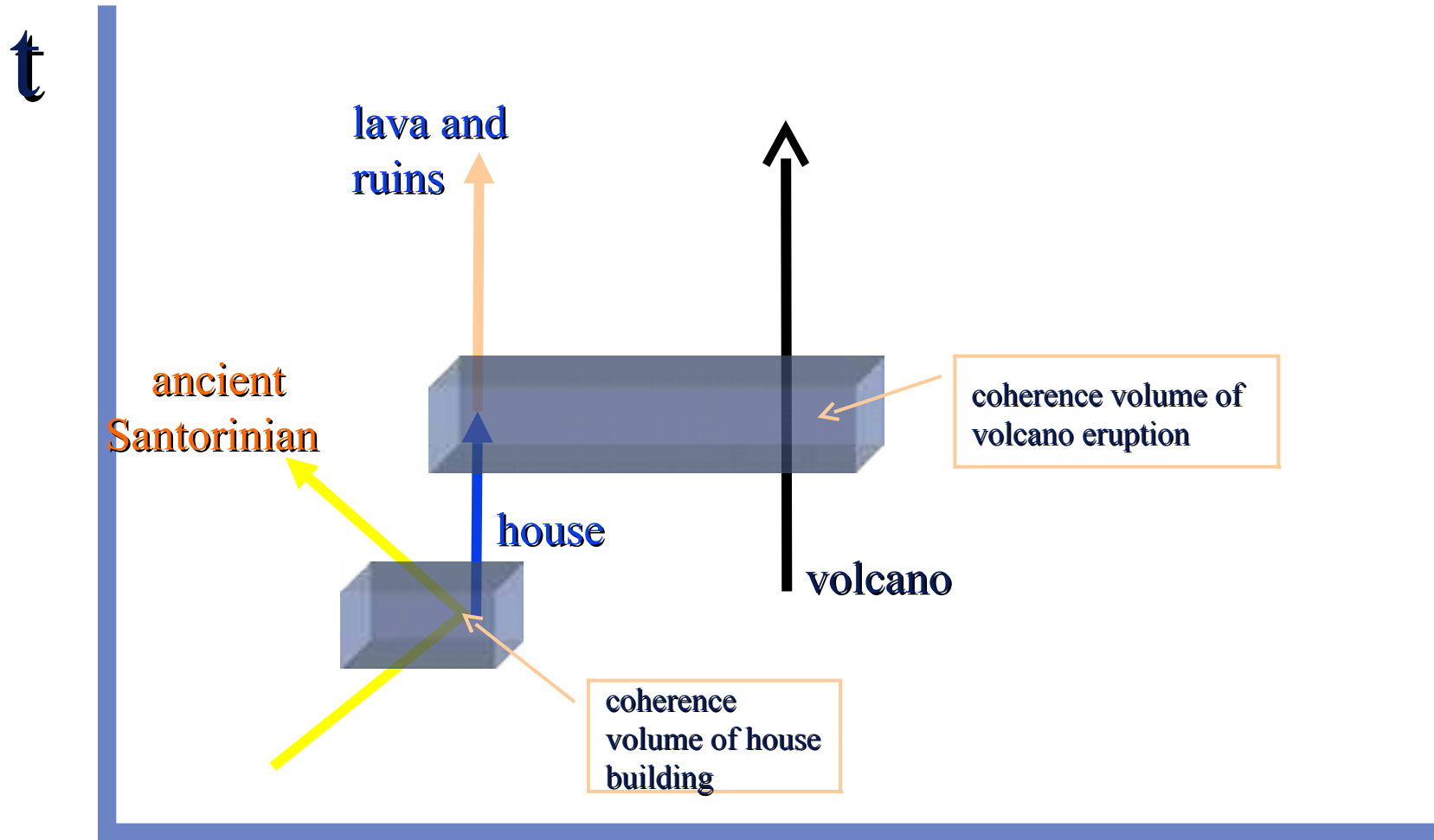
Santorini - Akrotiti

S

The CIDOC CRM
Exchanges of information as meetings

t

Victory!!!

coherence volume of
second announcement

coherence volume of
first announcement

2nd Athenian

Victory!!!

1st Athenian

other
Soldiers

runner

coherence volume of
the battle of
Marathon

Marathon   Athens

S
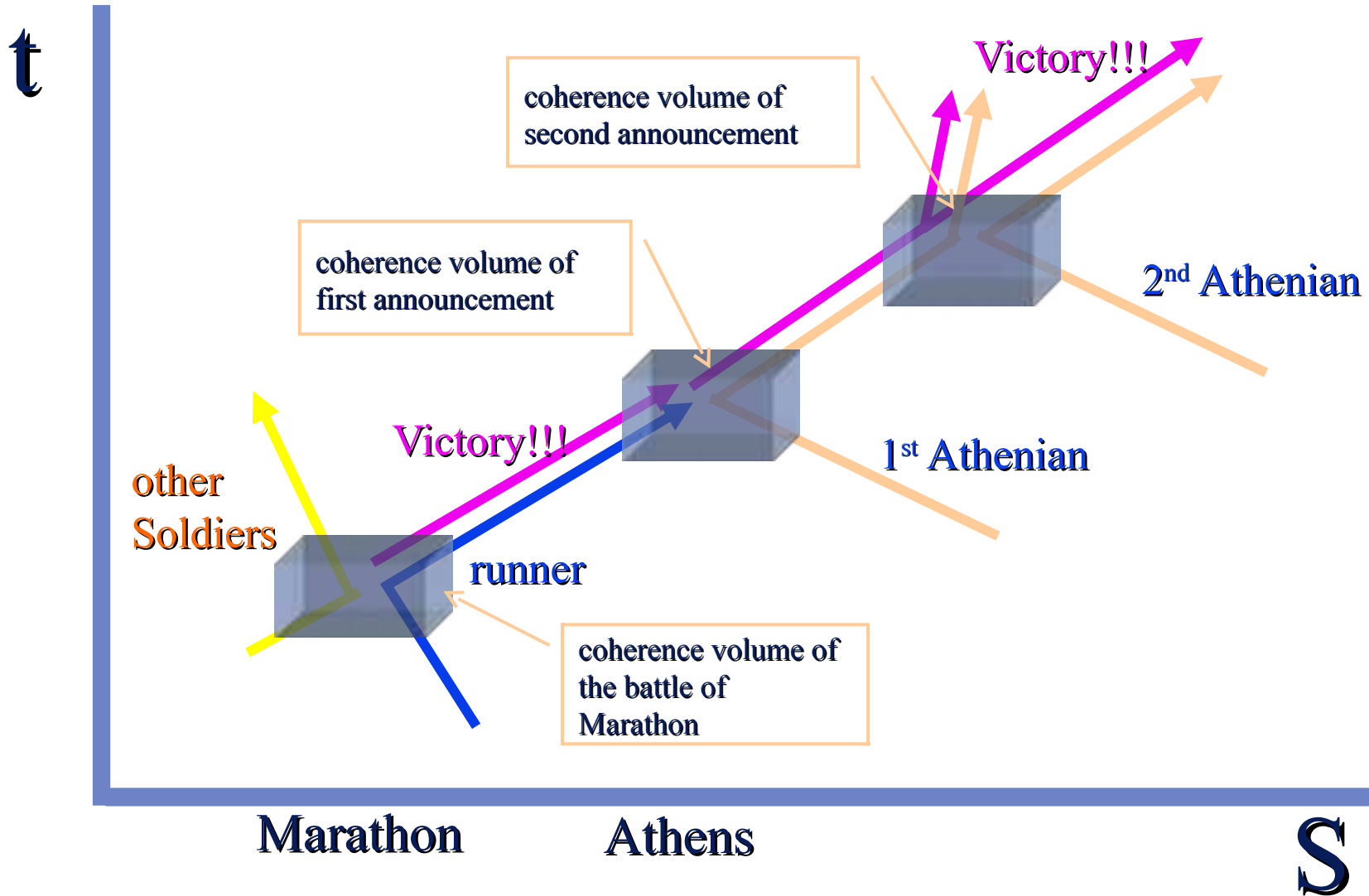
The CIDOC CRM
Time Uncertainty, Certainty and Duration
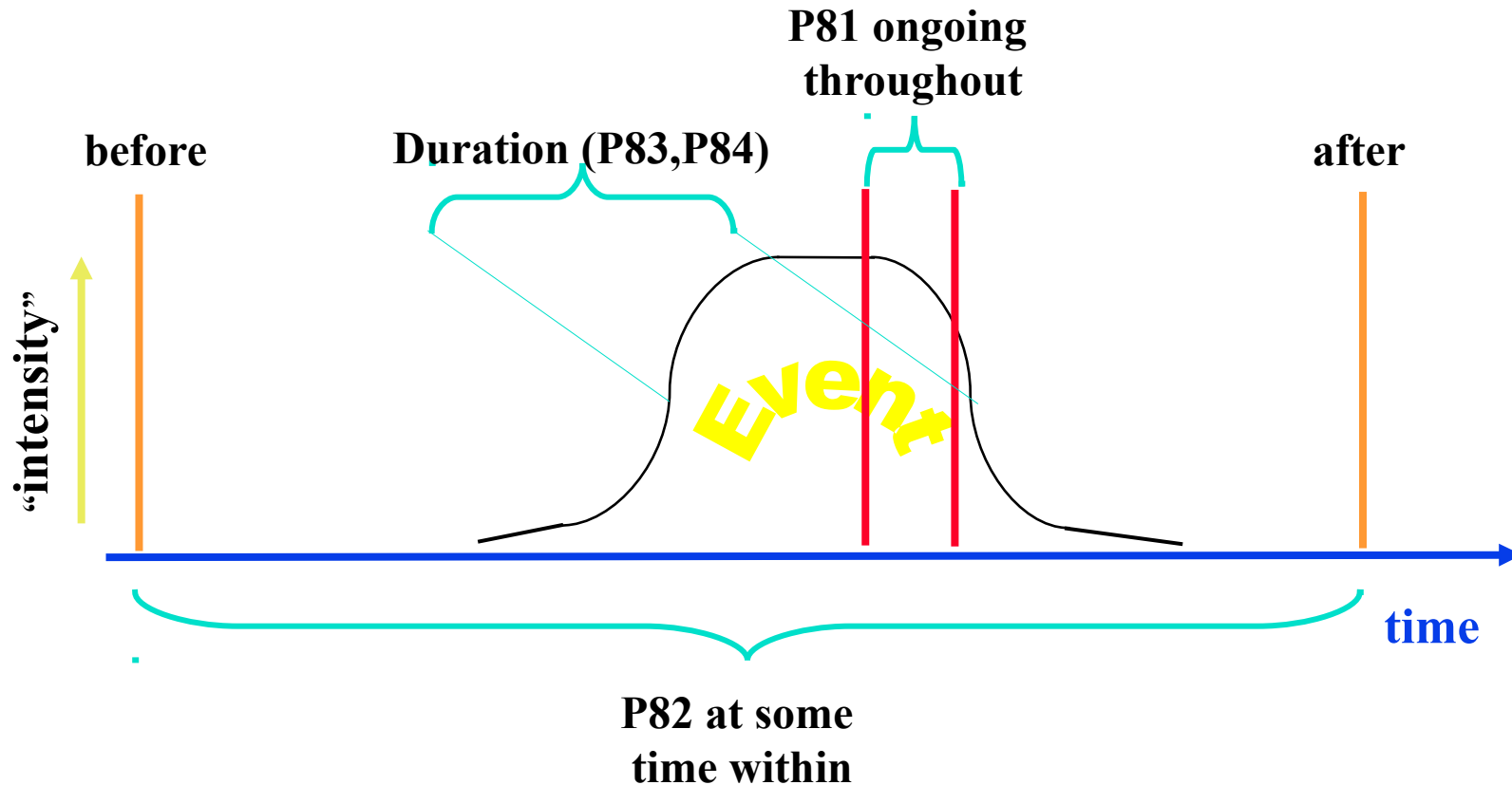
The CIDOC CRM
E52 Time-Span
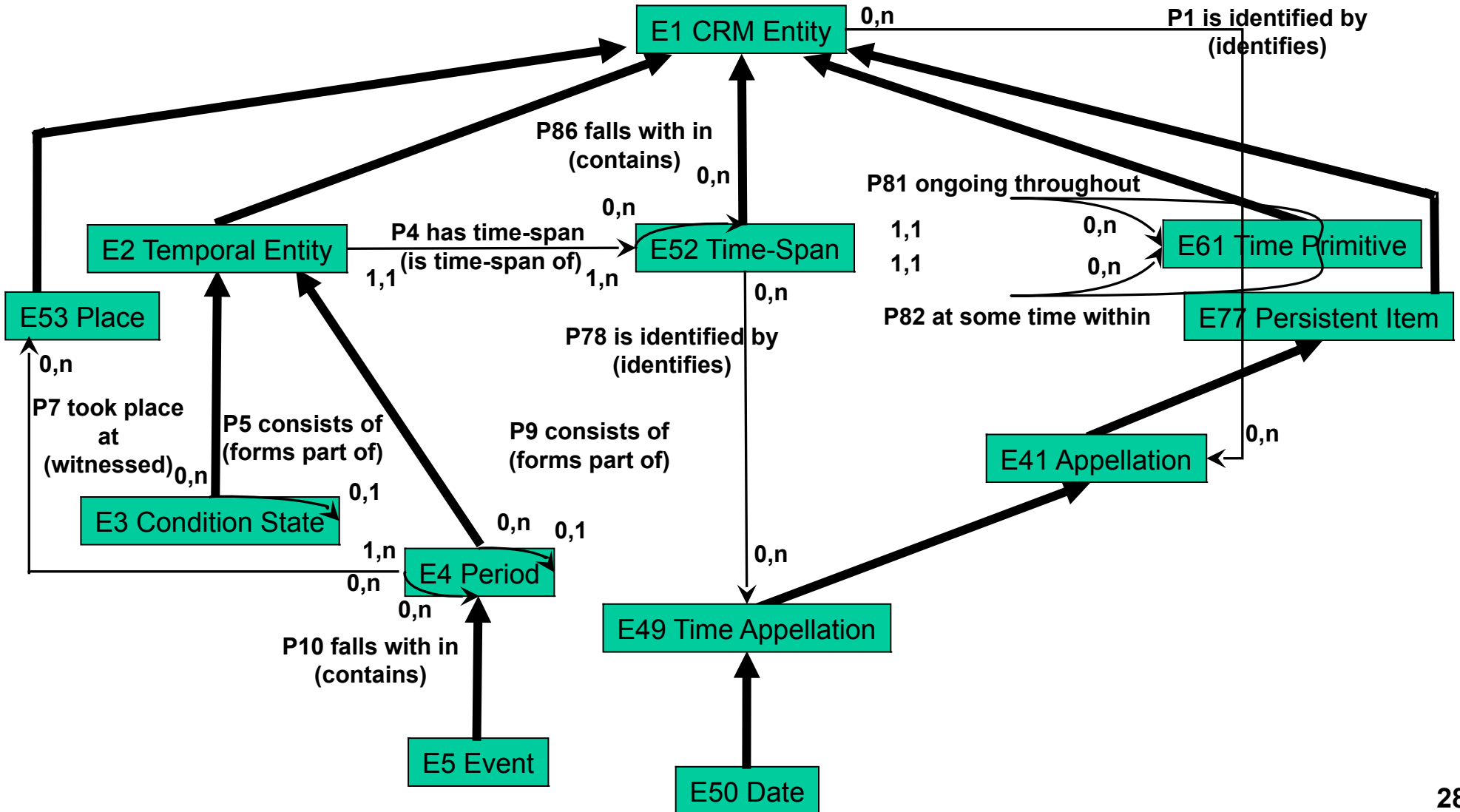
The CIDOC CRM
E7 Activity and inherited properties

E59 Primitive Value

E62 String

0,1 **P3 has note** **0,n** E1 CRM Entity **0,n**

P3.1 has type

E55 Type

**P2 has type
(is type of)**

0,n

E5 Event

E55 Type   *E.g., "Field Collection"*

E55 Type   *E.g., "photographer"*

P14.1 in the role of

E7 Activity **0,n** **P14 carried out by
(performed)** **1,n** E39 Actor
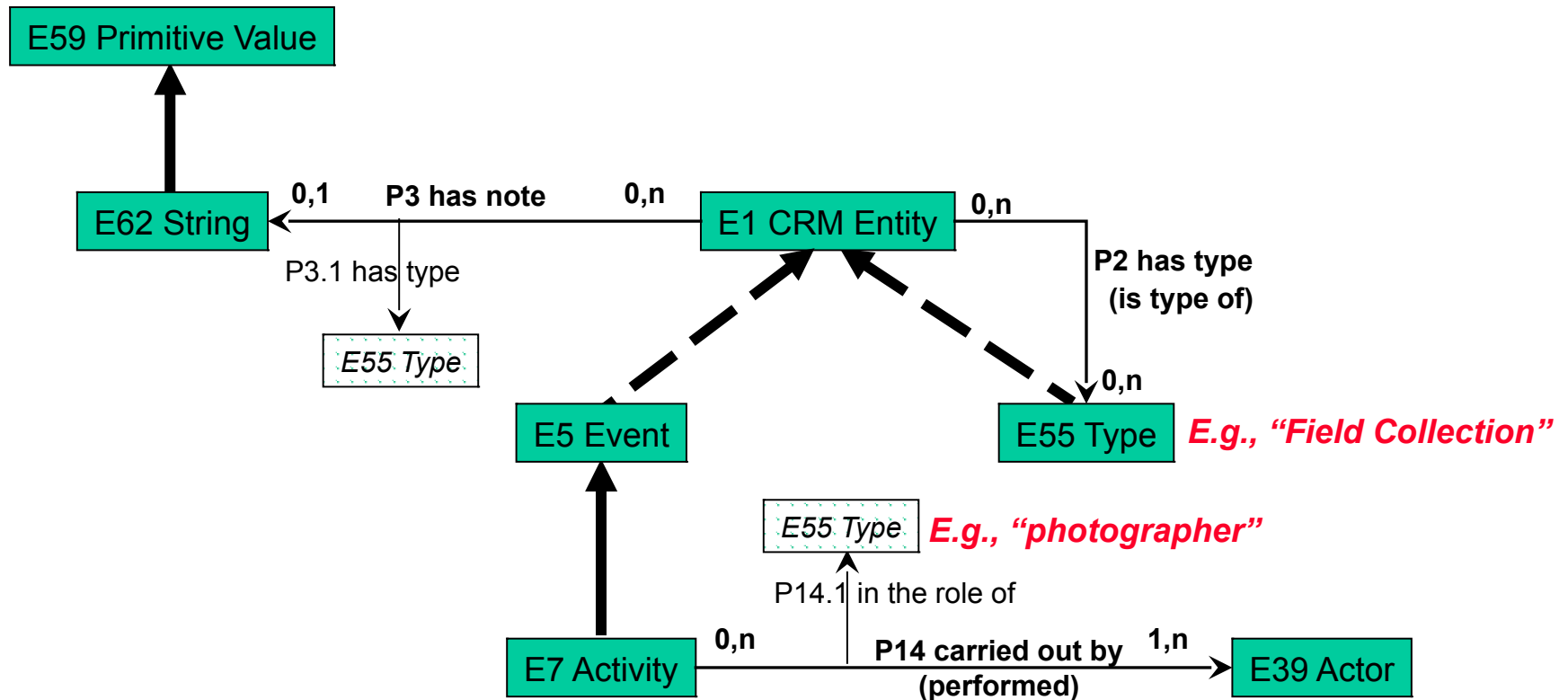
The CIDOC CRM
Activities: E16 Measurement
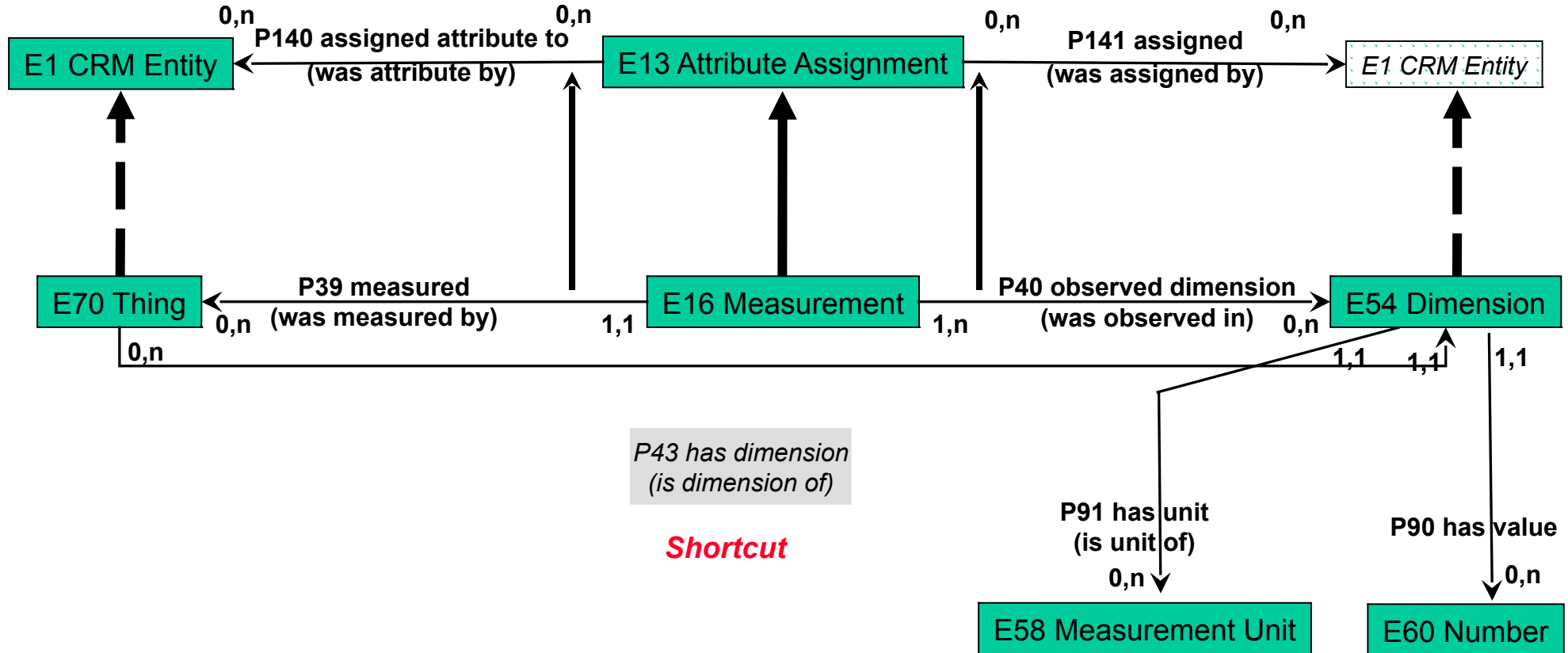
The CIDOC CRM

Activities: E14 Condition Assessment

P2 has type
(is type of)   0,n

E1 CRM Entity
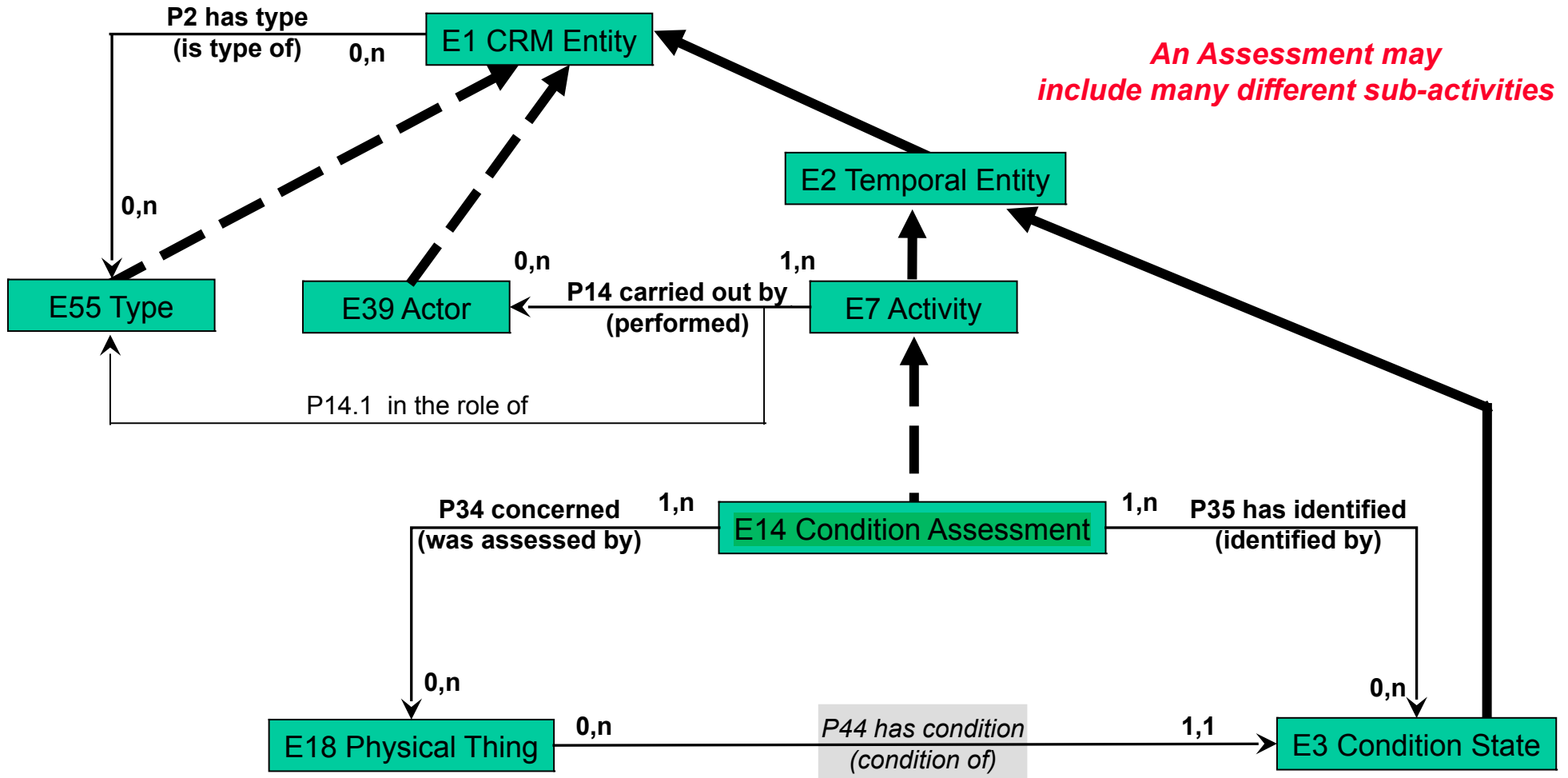
An Assessment may
include many different sub-activities

0,n

E55 Type

E2 Temporal Entity

0,n

E39 Actor   ←   **P14 carried out by**   1,n   E7 Activity
(performed)

P14.1 in the role of

P34 concerned   1,n   E14 Condition Assessment   1,n   **P35 has identified**
(was assessed by)   **(identified by)**

0,n   0,n

E18 Physical Thing   0,n   *P44 has condition*   1,1   E3 Condition State
(condition of)

Condition State is a Situation.
Its type is the "condition"

31

The CIDOC CRM
Activities: E8 Acquisition

E62 String ← **0,1** — **P3 has note** — **0,n** → E1 CRM Entity — **0,n** — **P2 has type (is type of)** — **0,n** → E55 Type

P3.1 has type → *E55 Type*

E5 Event

P14.1 in the role of — **1,n**

P14 carried out by (performed)

E7 Activity

No buying and selling, just one transfer

P22 transferred title to (acquired title through) — **0,n**

E8 Acquisition — **1,n** — **P24 transferred title of (changed ownership through)**

P23 transferred title from (surrendered title through) — **0,n**

0,n **0,n** **0,n**

E39 Actor

P51 has former or current owner (is former or current owner of) — **0,n**

P52 has current owner (is current owner of) — **0,n**

E18 Physical Thing — **0,n**

0,n

The CIDOC CRM
Activities: E9 Move

P7 took place at (witnessed) — 1,n — **E5 Period**

P20 had specific purpose (was purpose of) — 0,n

E7 Activity — 0,n — **P21 had general purpose** (was purpose of) — 0,n — **E55 Type**

0,n

P26 moved to (was destination of) — 1,n

E9 Move — 1,n — **P25 moved** (moved by)

P27 moved from (was origin of) — 1,n

P53 has former or current location (is former or current location of) — 1,n — **E18 Physical Thing**

0,n

P55 has current location (currently holds)

E53 Place — 0,n — **E19 Physical Object** — 0,1

0,n

P54 has current permanent location (is current permanent location of) — 0,1

0,n

The CIDOC CRM

Activities: E11 Modification/ E12 Production
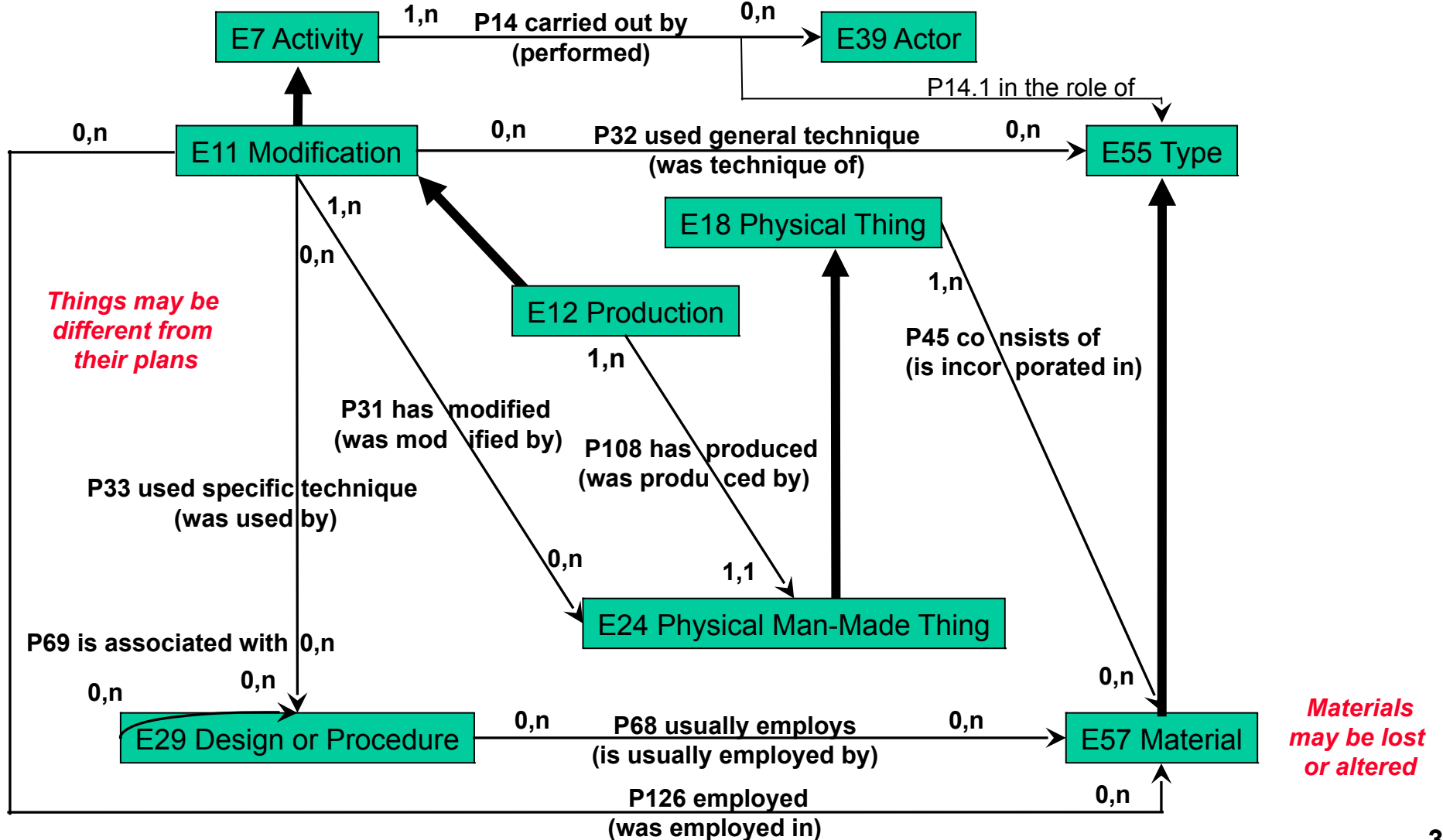
The CIDOC CRM
Inheriting Properties: E11 Modification

Properties:

P1 is identified by (identifies): E41 Appellation
P2 has type (is type of): E55 Type
P11 had participant (participated in): E39 Actor
P14 carried out by (performed): E39 Actor
       (P14.1 in the role of : E55 Type)

P31 has modified (was modified by): E24 Physical Man-Made Thing

P12 occurred in the presence of (was present at): E77 Persistent Item
P16 used specific object (was used for): E70 Thing
   (P16.1 mode of use: E55 Type)

P32 used general technique (was technique of): E55 Type
P33 used specific technique (was used by): E29 Design or Procedure

P17 was motivated by (motivated): E1 CRM Entity
P19 was intended use of (was made for): E71 Man-Made Thing
   (P19.1 mode of use: E55 Type)
P20 had specific purpose (was purpose of): E5 Event
P21 had general purpose (was purpose of): E55 Type

P126 employed (was employed in): E57 Material

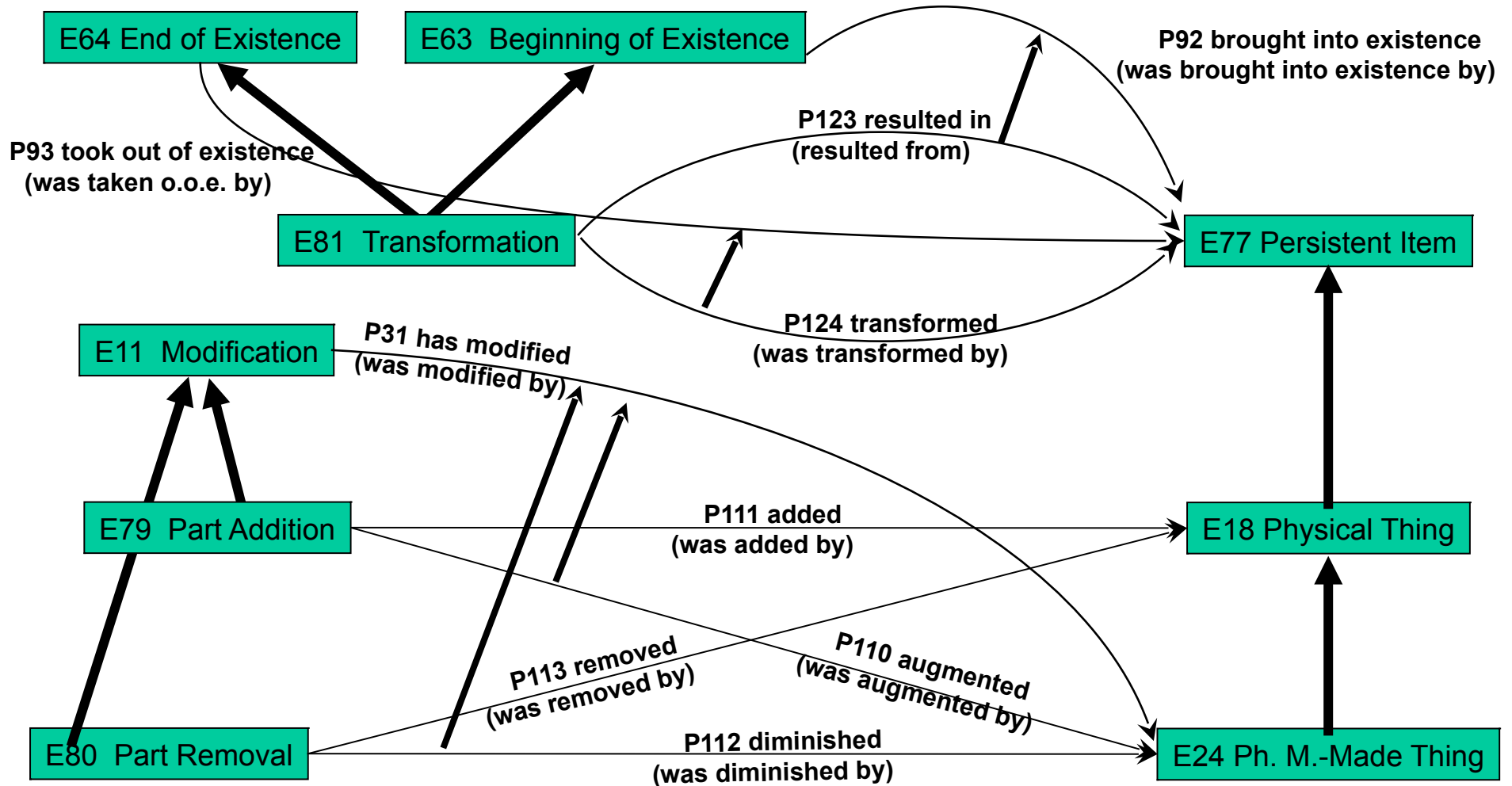*inherited properties*

declared property

inherited properties
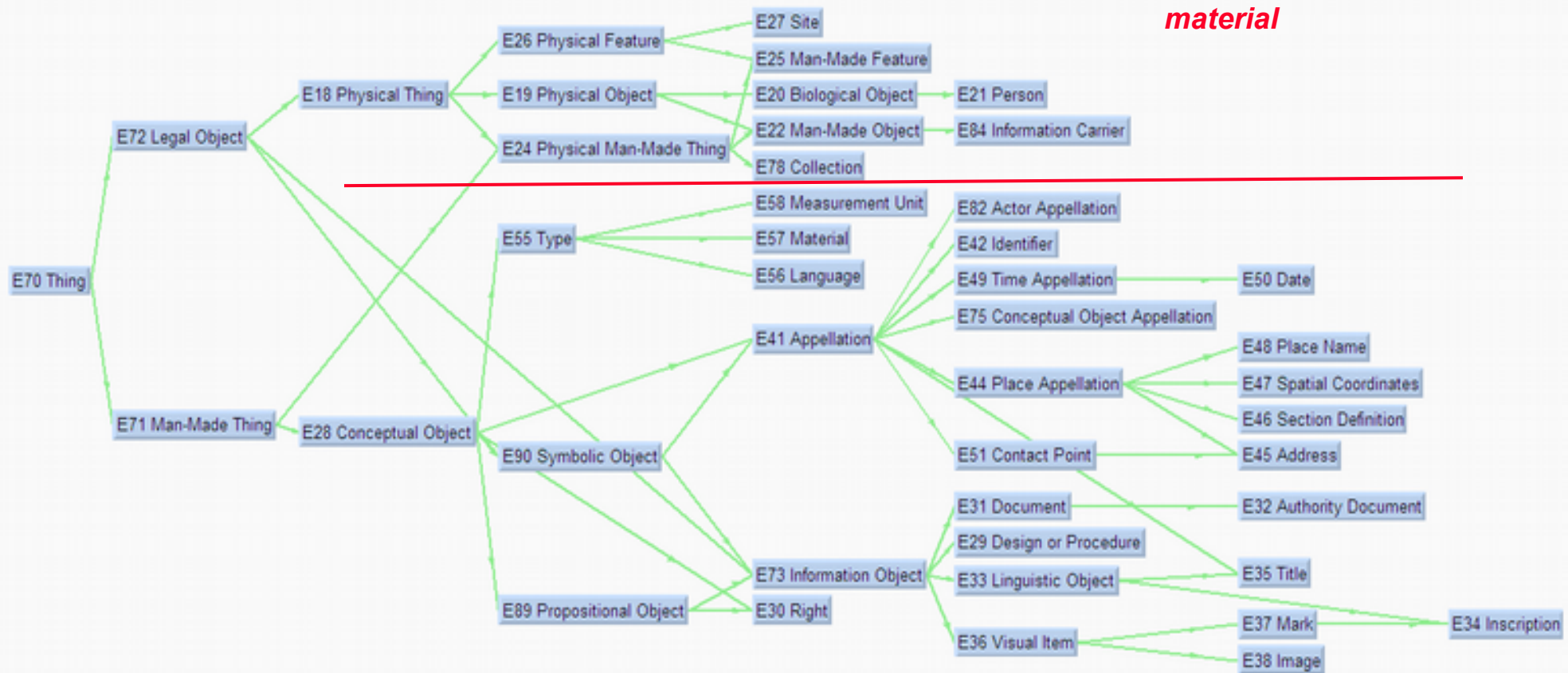
declared properties

inherited properties

declared property
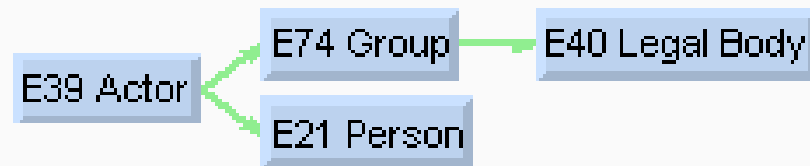
The CIDOC CRM
Ways of Changing Things

E64 End of Existence

E63 Beginning of Existence

P92 brought into existence
(was brought into existence by)

P93 took out of existence
(was taken o.o.e. by)

P123 resulted in
(resulted from)

E81 Transformation

E77 Persistent Item

P124 transformed
(was transformed by)

E11 Modification

P31 has modified
(was modified by)

E79 Part Addition

P111 added
(was added by)

E18 Physical Thing

P113 removed
(was removed by)

P110 augmented
(was augmented by)

E80 Part Removal

P112 diminished
(was diminished by)

E24 Ph. M.-Made Thing

The CIDOC CRM
Taxonomic discourse

E7 Activity

E65 Creation Event

E83 Type Creation

E17 Type Assignment

E1 CRM Entity

E28 Conceptual Object

E55 Type

P94 has created
(was created by)

P136 was based on
(supported type creation)

P135 created type
(was created by)

P41 classified
(was classified by)

P42 assigned
(was assigned by)

P136.1 in the
taxonomic role

P137 is exemplified
by (exemplifies)

P137.1 in the
taxonomic role

The CIDOC CRM
E70 Thing
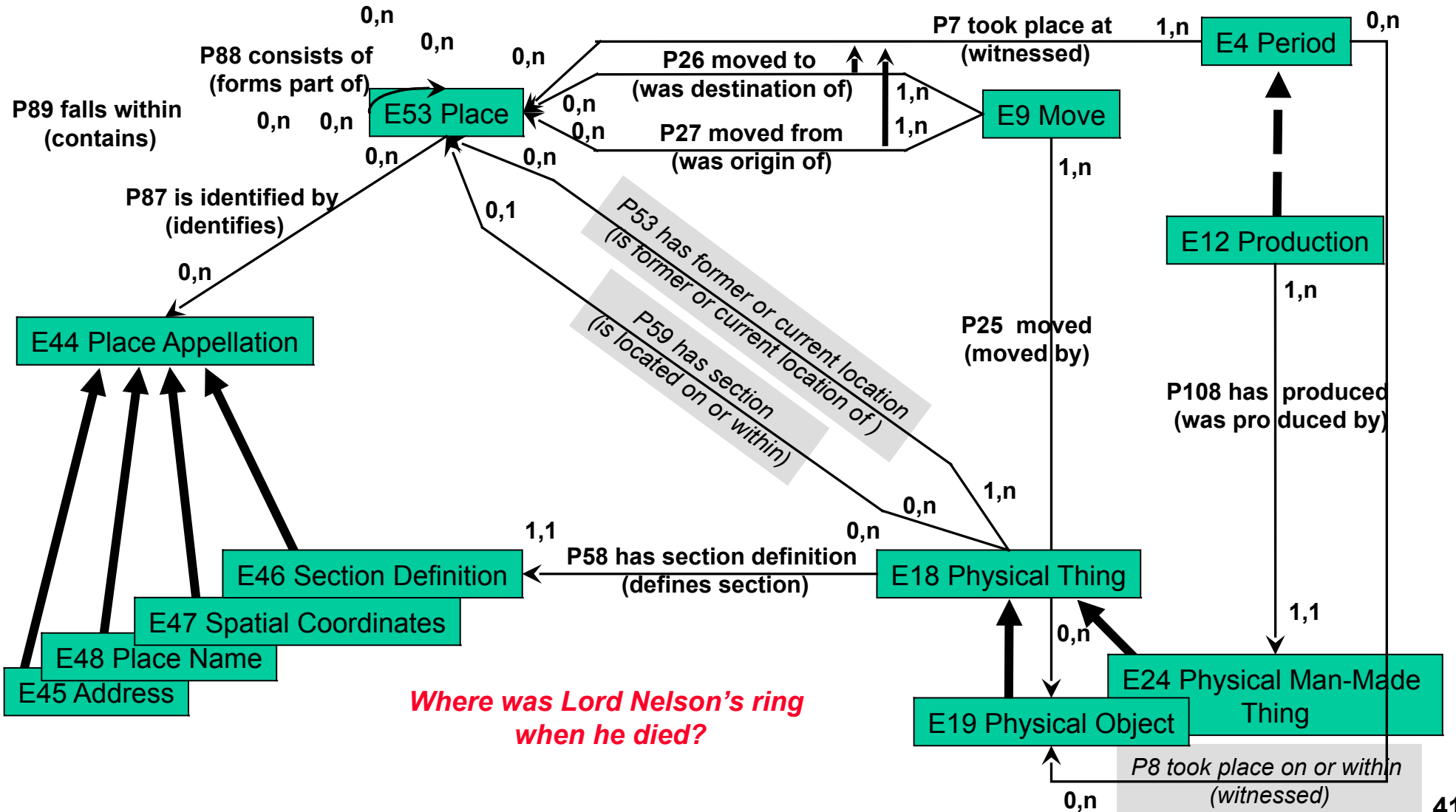
The CIDOC CRM
E39 Actor

The CIDOC CRM
E53 Place

☐ E53 Place

◆ **A place is an extent in space, determined diachronically with regard to a larger, persistent constellation of matter, often continents -**

by coordinates, geophysical features, artefacts, communities, political systems, objects - but not identical to

◆ **A "CRM Place" is not a landscape, not a seat - it is an abstraction from temporal changes - "the place where…"**

◆ **A means to reason about the "where" in multiple reference systems.**

◆ **Examples:**

—figures from the bow of a ship

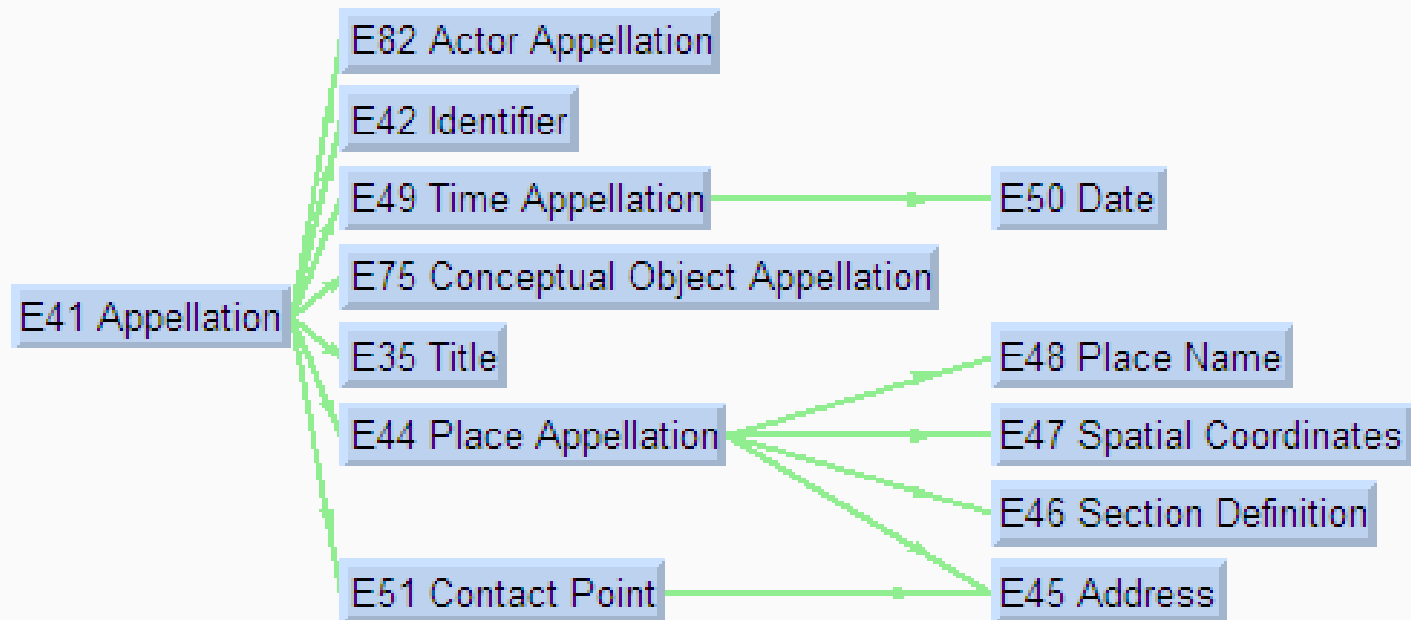—African dinosaur foot-prints in Portugal

—where Nelson died

The CIDOC CRM
Properties of E53 Place

P88 consists of (forms part of) 0,n 0,n 0,n

P7 took place at (witnessed) 1,n — E4 Period 0,n

P26 moved to (was destination of) 1,n

P27 moved from (was origin of) 1,n

E53 Place 0,n 0,n — E9 Move 1,n

P89 falls within (contains) 0,n 0,n 0,n 0,n

P87 is identified by (identifies) 0,n

E44 Place Appellation

P53 has former or current location (is former or current location of)

P59 has section (is located on or within) 0,1

P25 moved (moved by) 1,n

E4 Period

E12 Production 1,n

P108 has produced (was produced by) 1,1

0,n 1,n

E46 Section Definition

P58 has section definition (defines section) 1,1

E18 Physical Thing 0,n

E47 Spatial Coordinates

E48 Place Name

E45 Address

Where was Lord Nelson's ring when he died?

E19 Physical Object 0,n

E24 Physical Man-Made Thing

P8 took place on or within (witnessed) 0,n

The CIDOC CRM
E41 Appellation

The CIDOC CRM
Extension Example: Getty's TGN

P89 falls within

E53 Place

P87 is identified by (identifies) → E44 Place Appellation

E39 Actor

E13 Attribute Assignment

E74 Group

carries out

to place

P7 took place at

assigns name

identified by

Place Naming

P4 has time-span

to community

E4 Period

E52 Time-Span

P4 has time-span ←

Community
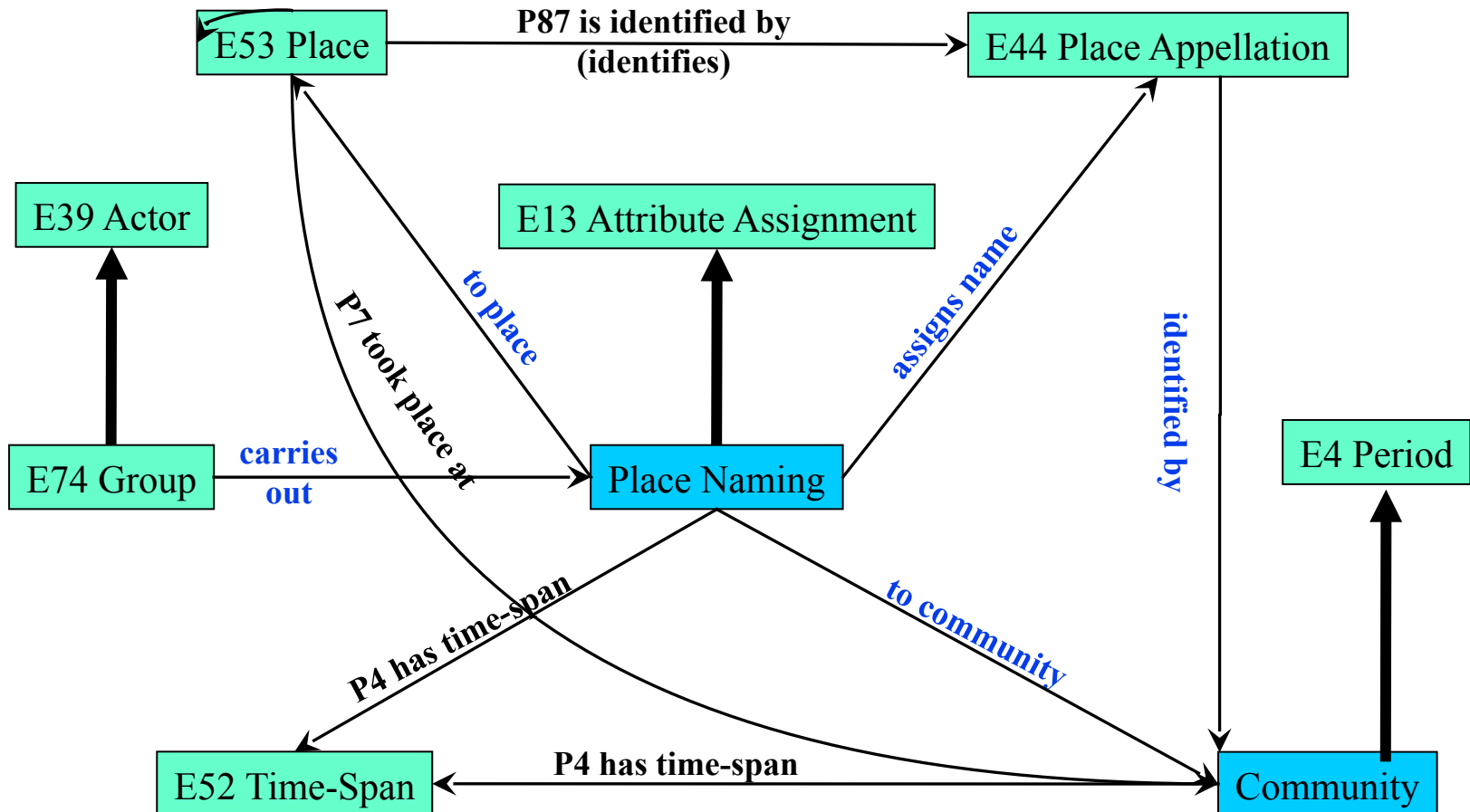
The CIDOC CRM
Sample of the TGN extension
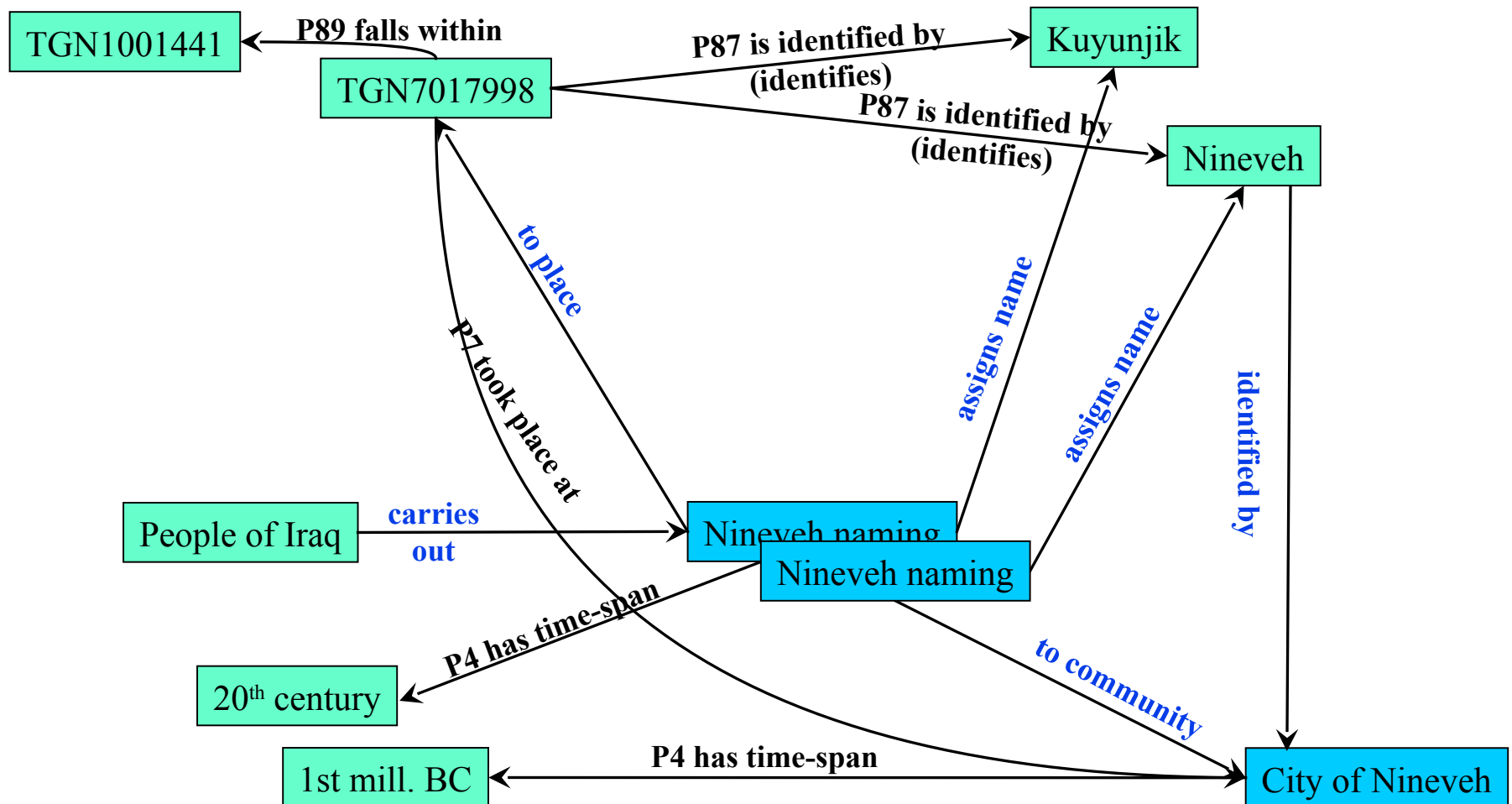
The CIDOC CRM
Visual Content and Subject
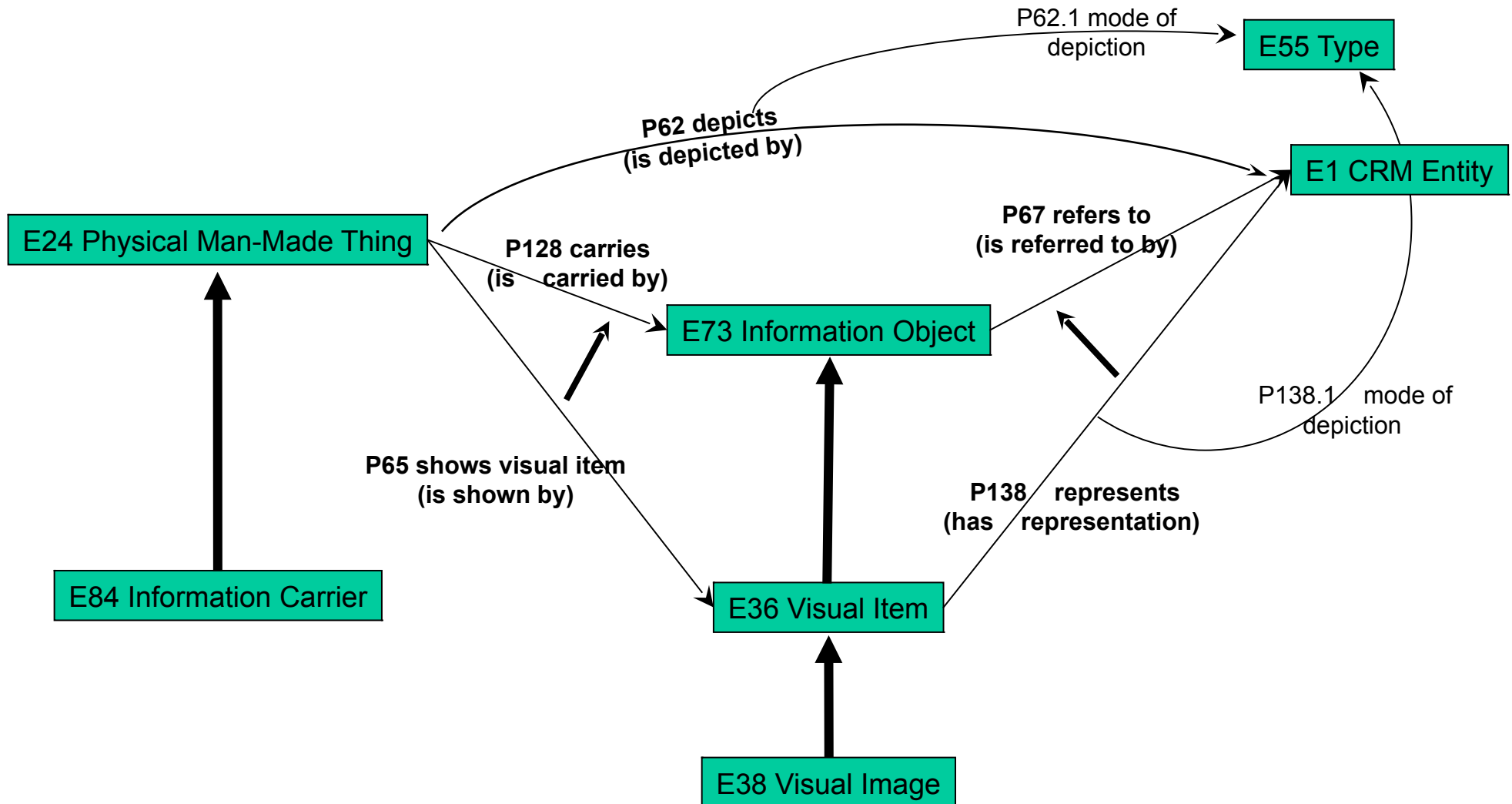
The CIDOC CRM
Application: Mapping DC to the CRM (1)

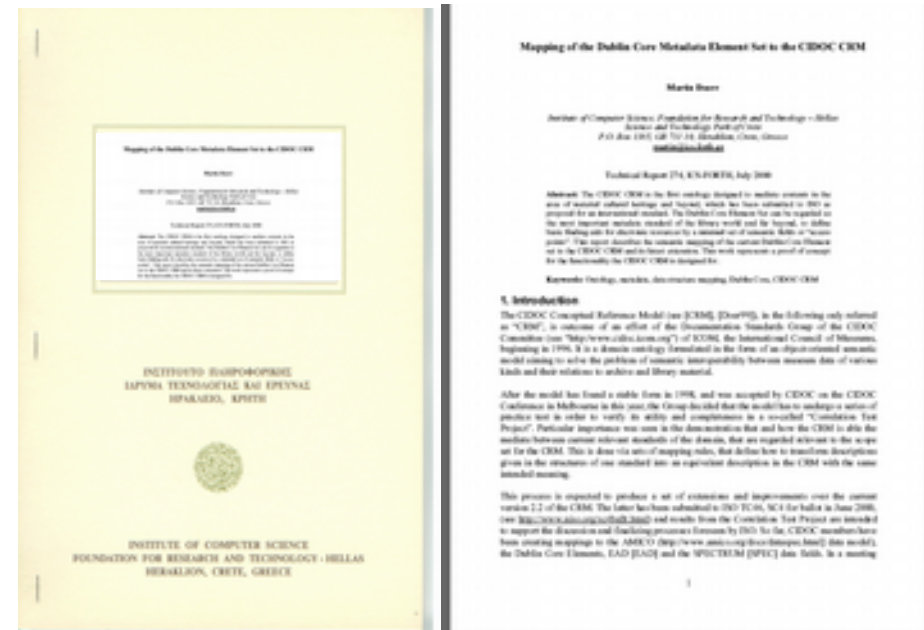**Example: DC Record about a Technical Report**

Type: text
Title: Mapping of the Dublin Core Metadata Element Set to the CIDOC CRM
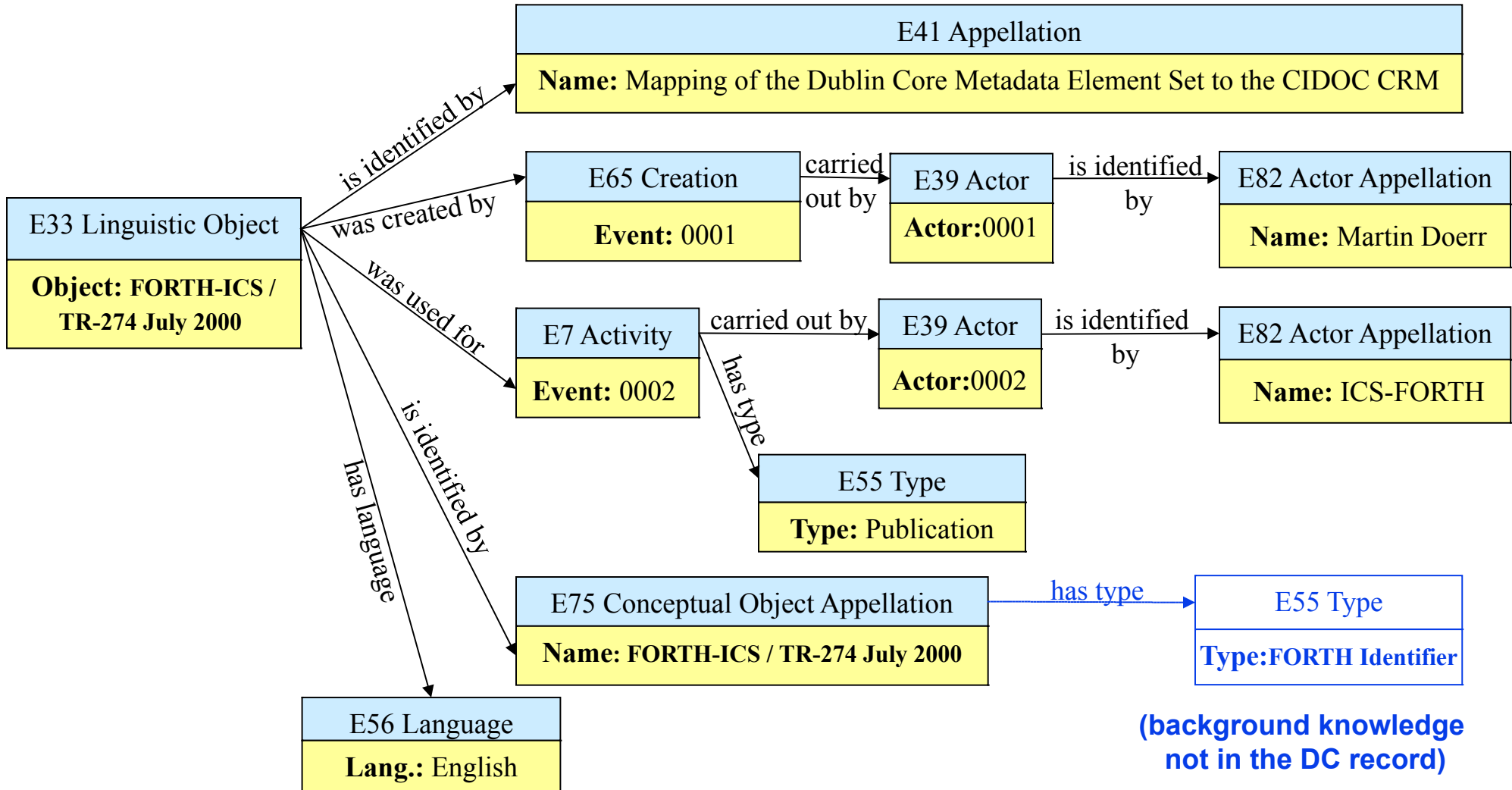**Creator:** Martin Doerr
Publisher: ICS-FORTH
Identifier: FORTH-ICS / TR 274 July 2000
Language: English

The CIDOC CRM
Application: Mapping DC to the CRM (2)

The CIDOC CRM
Application: Mapping DC to the CRM (3)
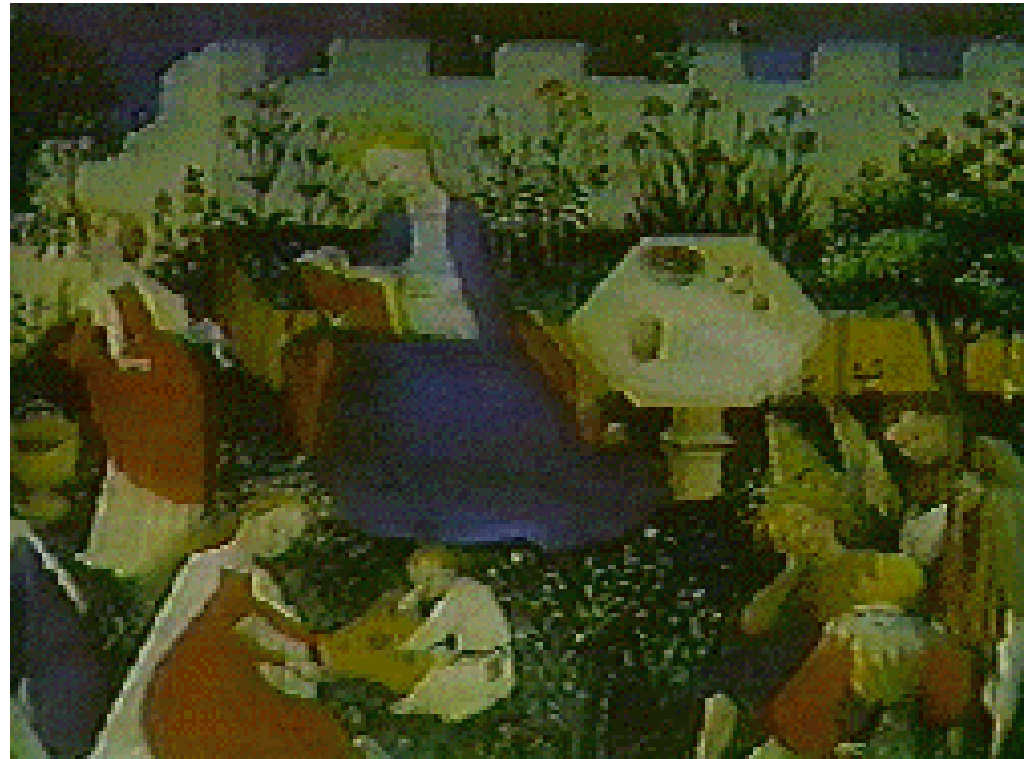
Example: DC Record about a painting

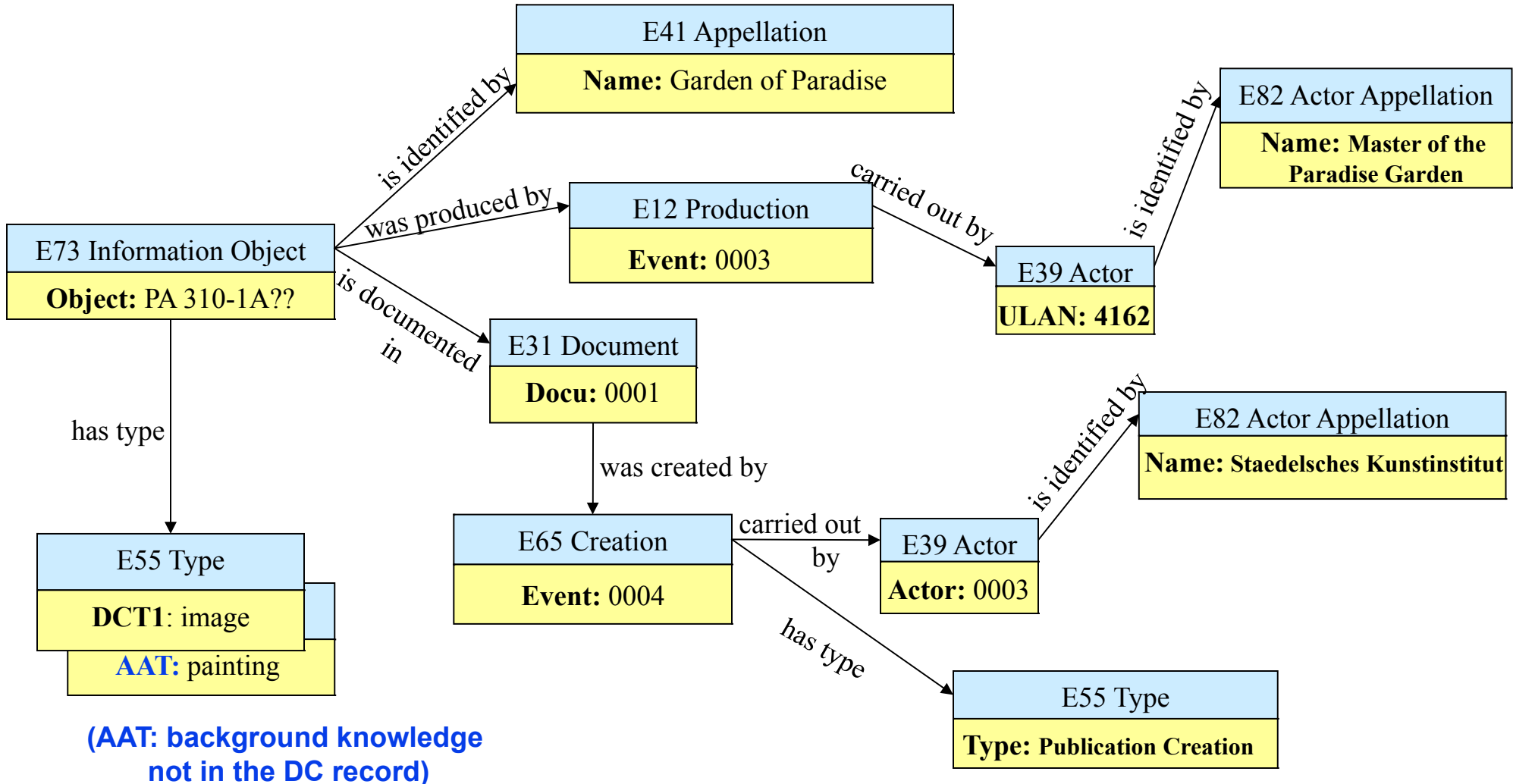Type.DCT1: **image**
Type: **painting**
Title: **Garden of Paradise**
Creator: **Master of the Paradise Garden**
Publisher: **Staedelsches Kunstinstitut**

The CIDOC CRM
Application: Mapping DC to the CRM (4)

E41 Appellation
Name: Garden of Paradise

E73 Information Object
Object: PA 310-1A??

— is identified by →

— was produced by →

E12 Production
Event: 0003

— carried out by →

E39 Actor
ULAN: 4162

— is identified by →

E82 Actor Appellation
Name: Master of the Paradise Garden

— is documented in →

E31 Document
Docu: 0001

— was created by →

E65 Creation
Event: 0004

— carried out by →

E39 Actor
Actor: 0003

— is identified by →

E82 Actor Appellation
Name: Staedelsches Kunstinstitut

— has type →

E55 Type
Type: Publication Creation

has type →

E55 Type
DCT1: image
AAT: painting

(AAT: background knowledge not in the DC record)

49

The CIDOC CRM
Lessons from mapping experiences

❑ **Semantic Interoperability can defined by the capability of mapping**

❑ **Mapping for epistemic networks is relatively simple:**

◆ *Specialist / primary information databases frequently employ a flat schema, reducing complex relationships into simple fields*

◆ *Source fields frequently map to composite paths under the CRM, making semantics explicit using a small set of primitives*

◆ *Intermediate nodes are postulated or deduced (e.g., "birth" from "person"). They are the hooks for integration with complementary sources*

◆ *Cardinality constraints must not be enforced= Alternative or incomplete knowledge*

❑ **Domain experts easily learn schema mapping**

◆ *IT experts may not understand meaning, underestimate it or are bored by it!*

❑ **Intuitive tools for domain experts needed:**

◆ **Separate identifier matching from schema mapping**
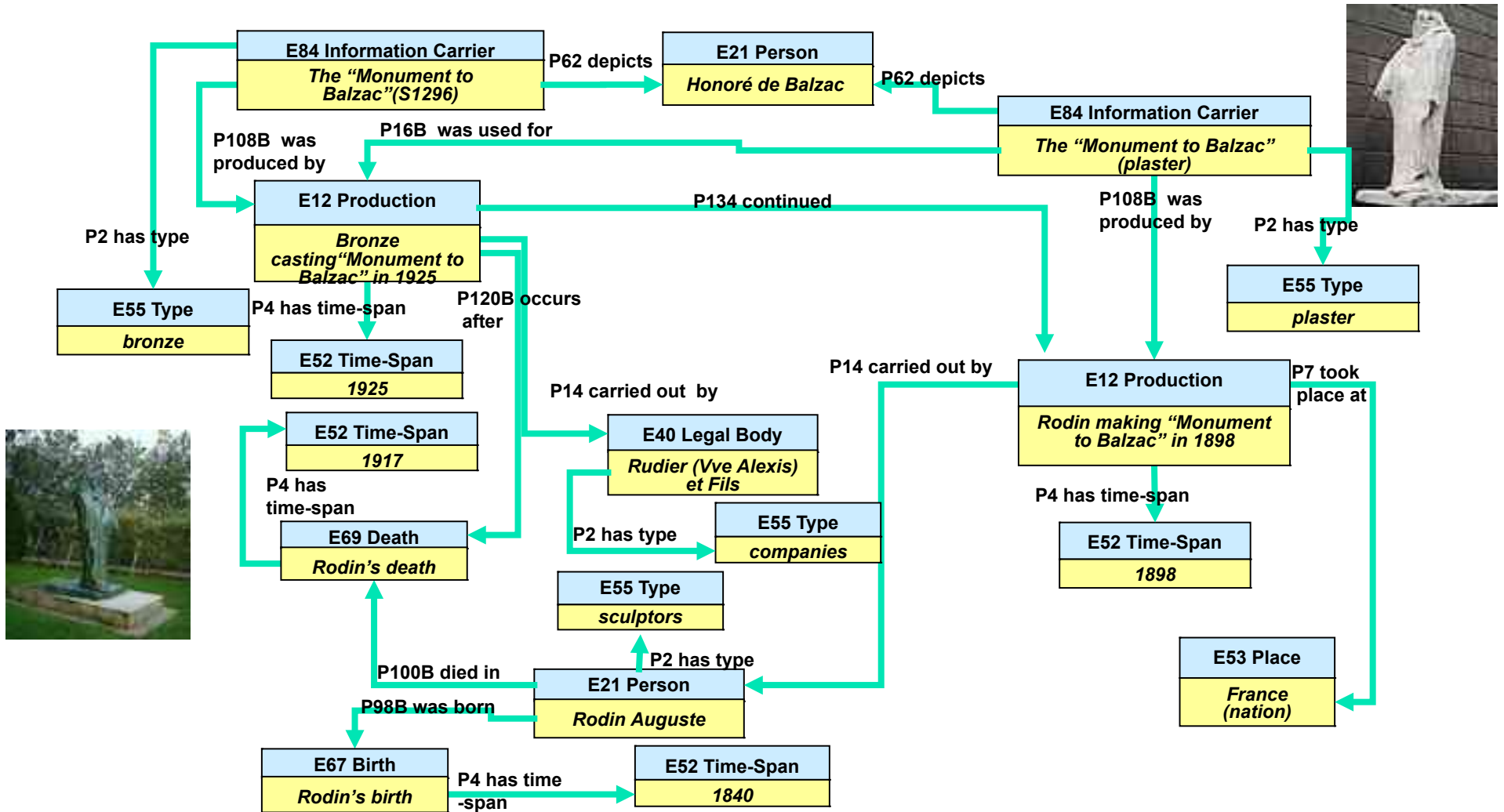
◆ **Separate terminology mediation from schema mapping**

The CIDOC CRM
Differences to other ontologies

- ❑ Generally: Many ontologies:-
  - ◆ lack an empirical base
  - ◆ have a functionally insufficient system of relationships (terminology driven)
  - ◆ Have a lack of functional specifications
- ❑ The CRM misses concepts not in the empirical base (e.g., contracts), but detects concepts that are not lexicalized (e.g."Persistent Item"), because they are functionally required
- ❑ **DOLCE**: Lexical base, intuition. Very good theoretically motivated logical description. Foundational relationships. Over specified relationships (e.g. modes of participation). Bad model of space-time. Strong overlap with CRM
- ❑ **BFO**: Philosophically motivated. Poor model of relationships. Notion of a precise, deterministic underlying reality. Empirical verification difficult. Strong overlap with CRM
- ❑ **IndeCs, ABC Harmony**: Small ontologies, event centric, strong overlap with CRM (harmonized!)
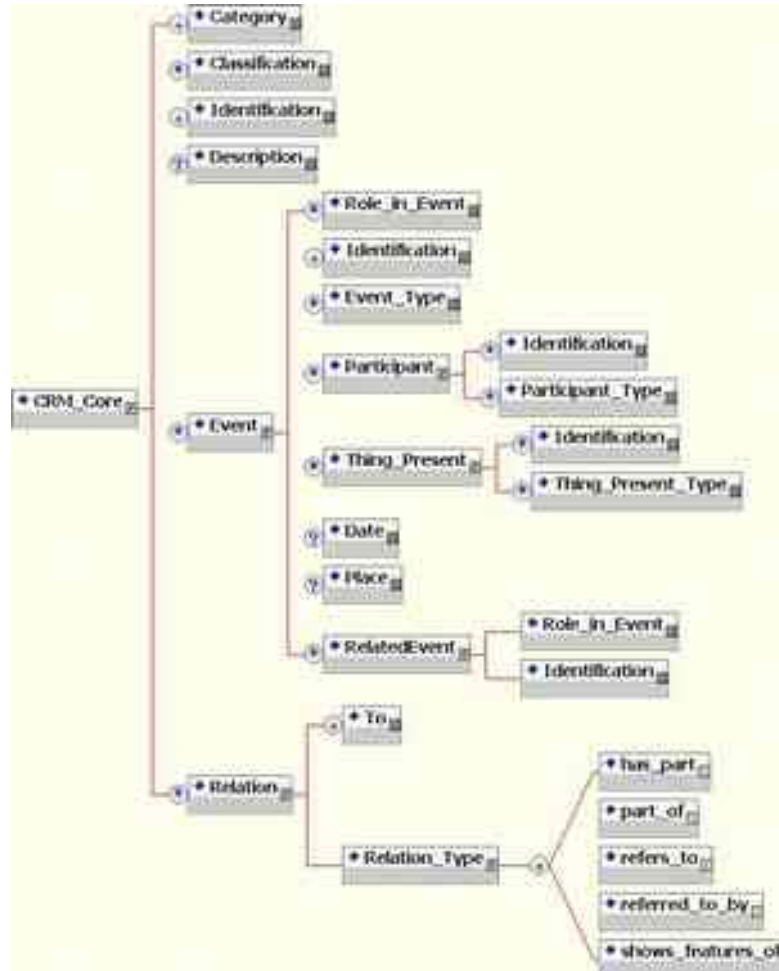- ❑ **SUMO**: Large aggregation of concepts without functional specifications

The CIDOC CRM
Applications: Integration with CRM Core (1)

E84 Information Carrier
The "Monument to Balzac"(S1296)

P62 depicts →

E21 Person
Honoré de Balzac

← **P62 depicts**

E84 Information Carrier
The "Monument to Balzac" (plaster)

P16B was used for

P108B was produced by

E12 Production
Bronze casting "Monument to Balzac" in 1925

P134 continued

P108B was produced by

P2 has type

E55 Type
bronze

P4 has time-span

P120B occurs after

E52 Time-Span
1925

P2 has type

E55 Type
plaster

P14 carried out by

P14 carried out by

E12 Production
Rodin making "Monument to Balzac" in 1898

P7 took place at

E52 Time-Span
1917

P14 carried out by

E40 Legal Body
Rudier (Vve Alexis) et Fils

P4 has time-span

P4 has time-span

E69 Death
Rodin's death

P2 has type

E55 Type
companies

E52 Time-Span
1898

E55 Type
sculptors

P100B died in

P98B was born

E21 Person
Rodin Auguste

P2 has type

E53 Place
France (nation)

E67 Birth
Rodin's birth

P4 has time -span

E52 Time-Span
1840

The CIDOC CRM
Applications: Integration with CRM Core (2)

CRM Core
A minimal metadata element set

Artist (CRM Core).

<u>Category</u> = E21 Person

Classification = artists

Classification = sculptors

Identification =Rodin, Auguste

Identification =ID: 500016619

Event

  Role in Event =P98B was born

  Identification= Rodin's birth

  Event Type = E67 Birth

  Date = 1840

Event

  Role in Event =P100B died in

  Identification= Rodin's death

  Event Type = E69_Death

  Date = 1917

  Related event

    Role in Event =P120 occurs before

    Identification= Bronze casting Monument to Balzac in 1925

Work (CRM Core).

<u>Category</u> = E84 Information Carrier

Classification =sculpture (visual work)

Classification =plaster

Identification =The Monument to Balzac (plaster)

Description =Commissioned to honor one of France's greatest novelists, Rodin spent seven years preparing for Monument to Balzac. When the plaster original was exhibited in Paris in 1898, it was widely attacked. Rodin retired the plaster model to his home in the Paris suburbs. It was not cast in bronze until years after his death.

Event

  Role in Event =P108B was produced by

  Identification= Rodin making Monument to Balzac in 1898

  Event Type = E12 Production

  Participant

    Identification =Rodin, Auguste

    Identification =ID: 500016619

    Participant Type = artists

    Participant Type = sculptors

  Date = 1898

  Place = France (nation)

  Related event

    Role in Event =P134B was continued by

    Identification= Bronze casting Monument to Balzac in 1925

Event

  Role in Event =P16B was used for

  Identification= Bronze casting Monument to Balzac in 1925

  Event Type = E12 Production

  Participant

    Identification =Rudier (Vve Alexis) et Fils

    Participant Type = companies

  Thing Present

    Identification =The Monument to Balzac (S.1296)

    Thing Present Type =bronze

    Thing Present Type =sculpture (visual work)

  Date = 1925

  Related event

    Role in Event =P120B occurs after

    Identification= Rodin's death

Relation

  To = Honore de Balzac

  Relation type

    refers to

The CIDOC CRM
Methodological aspects

❑ **The CRM aims at semantic integration beyond context.**

❑ **It aims at pulling together all relevant sources and data to evaluate a scientific or scholarly question not answered by an individual document**

❑ **Based on the CRM, effective integration schemata can be defined, such as "CRM Core", the full CRM or extensions of the CRM**

❑ **The CRM can fit rich and poor models under one common logical frame-work . For instance Dublin Core (DC) maps to the CRM**

❑ **Idea: Not being prescriptive creates lots of flexibility**

◆ **It does not propose what to describe. It allows interpretation of what museums and archives actually describe**

The CIDOC CRM
Documents and Knowledge

❑ **Scientific and scholarly work produces knowledge by argumentation**

◆ This comes in **closed** units, "**documents**"
◆ They have a history of evolution, "**versions**"
◆ The knowledge is "directed"
◆ It can only be evaluated in **context**
— document about Mona Lisa
— theory about the origin of the Minoan people

❑ **It should be possible to map primary document structures to the CRM. This is easy:**

◆ E.g. **good is**: "creator - creation place - creation date"
    bad is : "provenance", "place associations - life-cycle dates" etc.

◆ **Good** document structures map **easily**

◆ **No completeness** requirements

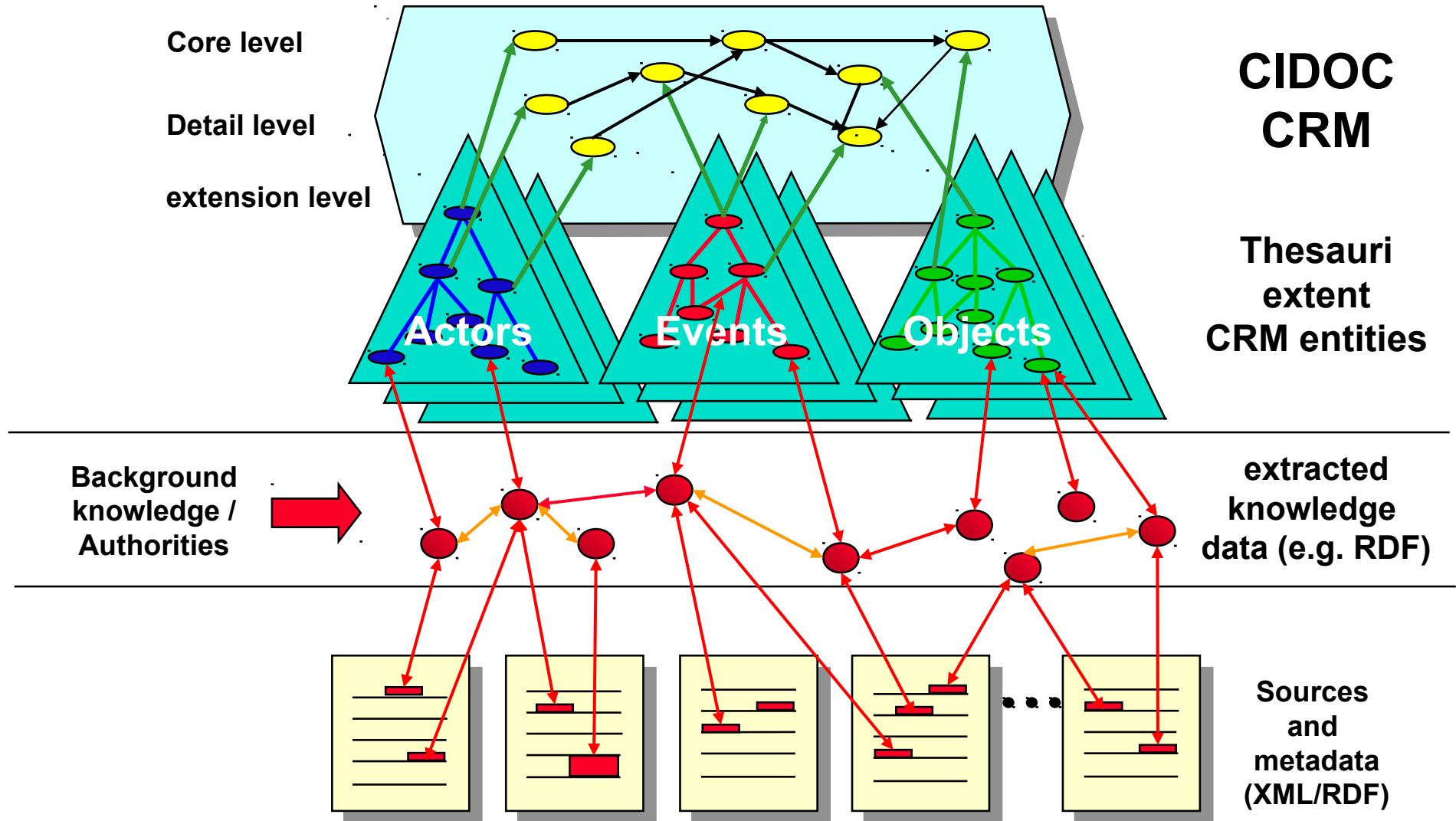# The CIDOC CRM
Knowledge management

❑ **Three-level knowledge management:**

◆ **Acquisition (can be motivated by the CRM):**
  — **sequence and order**, completeness, constraints to guide and control data entry.
  — ergonomic, case-specific language, optimized to specialist needs
  — often working on series of analogous items
  — Low interoperability needs (**capability to be mapped!**)

◆ **Integration / comprehension (target of the CRM):**
  — break up document boundaries, relate facts to wider context
  — match **shared identifiers** of items, **aggregate alternatives**
  — no preference direction of search, no cardinality constraints
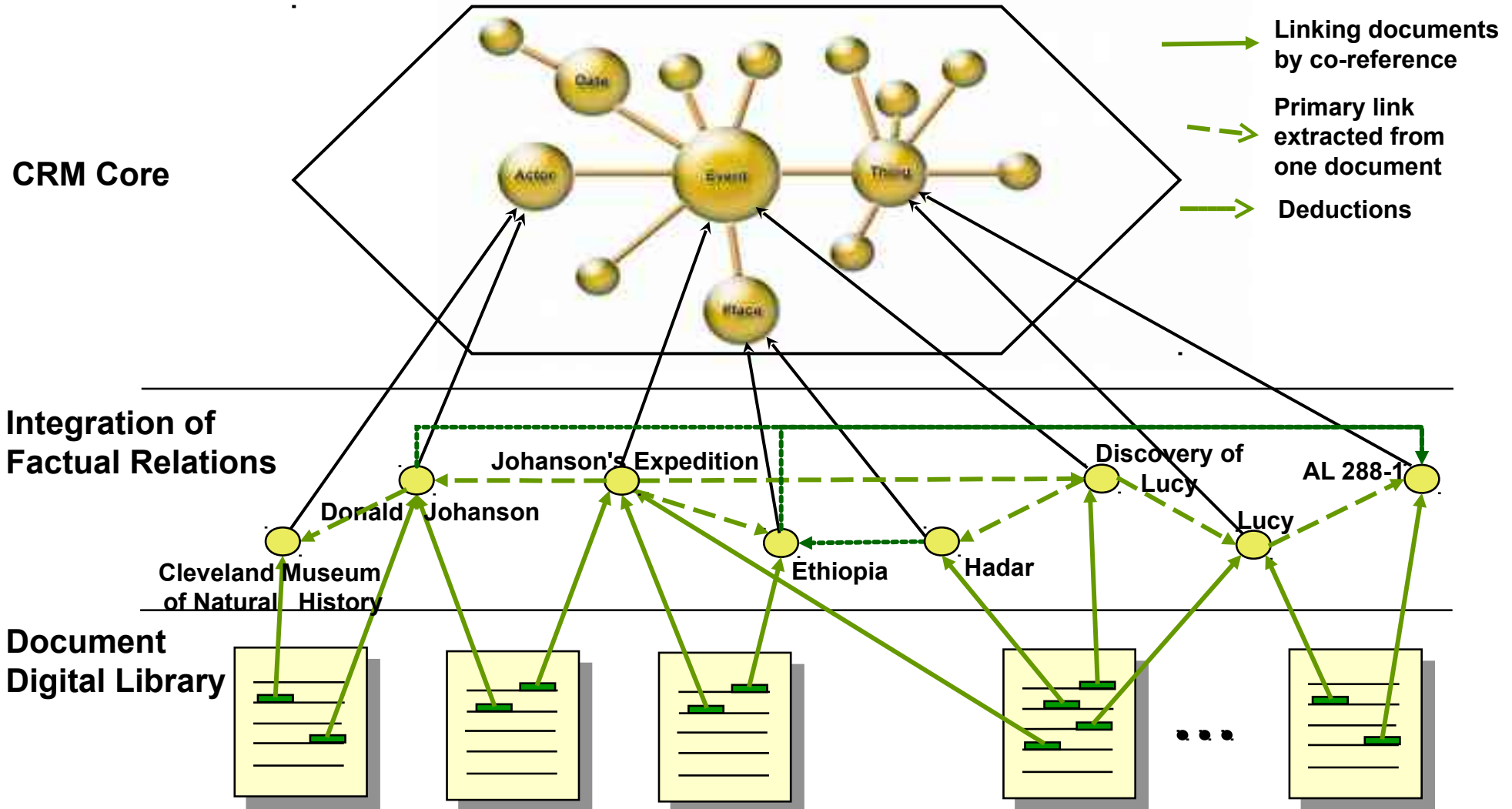 — High interoperability needs (**mapping to a global schema**)

◆ **Presentation, story-telling (can be based on CRM)**
  — explore context, paths, analogies orthogonal to data acquisition
  — **present in order**, allow for digestion
  — deduction and induction

The CIDOC CRM -Application
Repository Indexing

Core level

Detail level

extension level

CIDOC CRM

Actors   Events   Objects

Thesauri extent CRM entities

Background knowledge / Authorities

extracted knowledge data (e.g. RDF)

Sources and metadata (XML/RDF)

The CIDOC CRM
Documents and Factual Knowledge

The CIDOC CRM
Benefits of the CRM (From Tony Gill)

❑ Elegant and simple compared to comparable Entity-Relationship models

❑ Coherently integrates information at varying degrees of detail

❑ Readily extensible through O-O class typing and specializations

❑ Richer semantic content; allows inferences to be made from underspecified data elements

❑ Designed for mediation of cultural heritage information

The CIDOC CRM
State of Development

❑ **Publication as ISO 21127:2006 in October 2006**

❑ **Work on extension covering FRBR, FRAD and CRM resulted in "FRBR$_{OO}$", accepted by IFLA and CIDOC**

❑ **Ongoing work on TEI – CRM harmonization**

❑ **Application models (CRM Core, good and rich data exchange formats, extensions)**

❑ **OWL version being finalized**

The CIDOC CRM
Conclusions

❑ **Doing all that, we encounter a surprise compared with common preconceptions:**

◆ **Nearly no domain specificity (e.g."current permanent location"), generic concepts appear in medicine, biodiversity etc.**

◆ **Rather a notion of scientific method emerges, such as "retrospective analysis", "taxonomic discourse" etc.**

◆ **Extraordinary small set of concepts**

◆ **Extraordinary convergence: adding dozens of new formats hardly introduces any new concept**

❑ **This approach is economic, investment pays off**

➢**The CRM should become our language for semantic interoperability,**