

Francesco Beretta
(CNRS UMR5190 LARHRA – Université de Lyon)

**Historical Information,
Epistemology and
Foundational Ontologies**

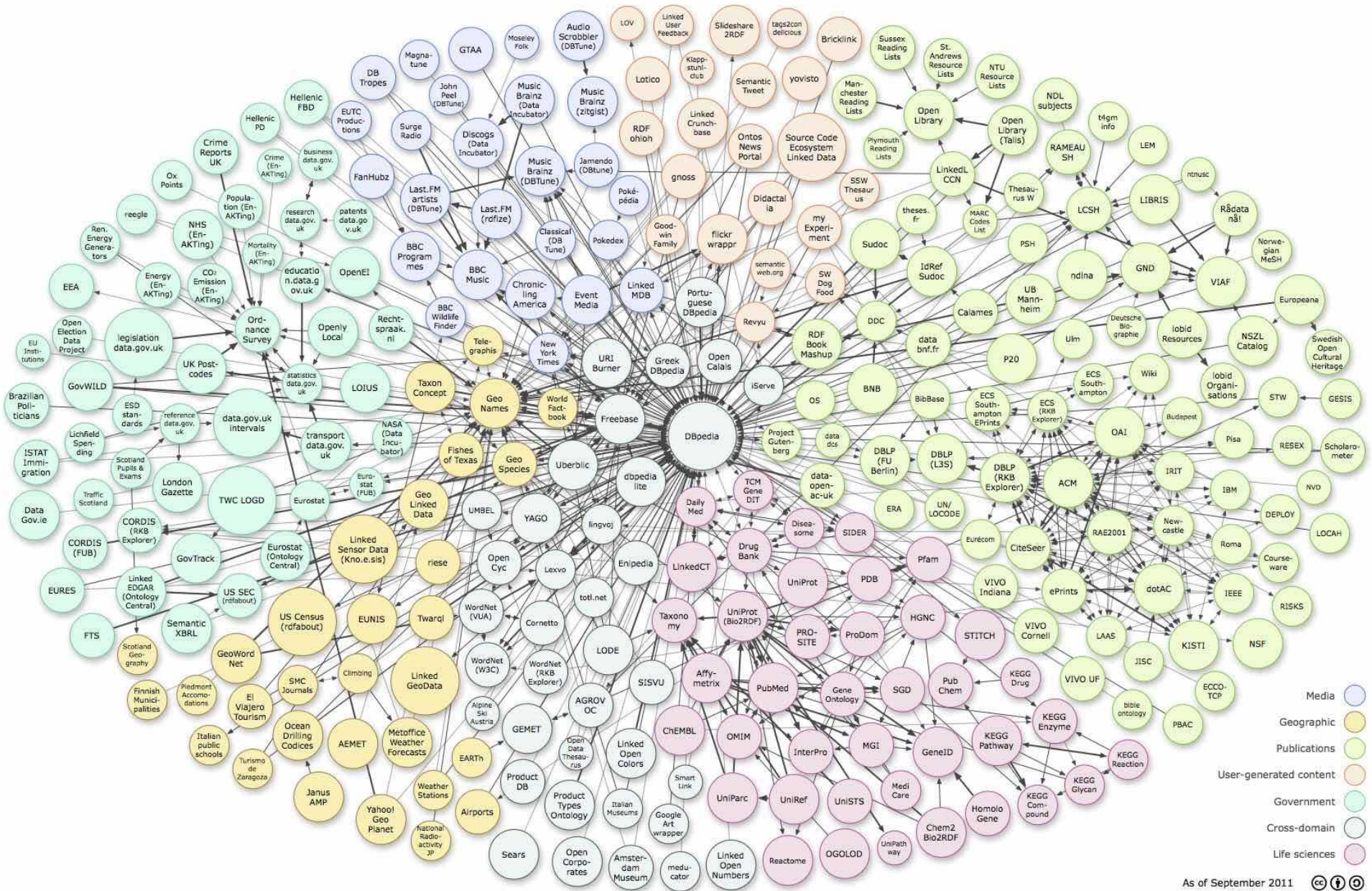
**Semantic Heritage Research Data – SHERD
Virtual Online Seminar**

20 Novembre 2020

1.

The *symogih.org* project :

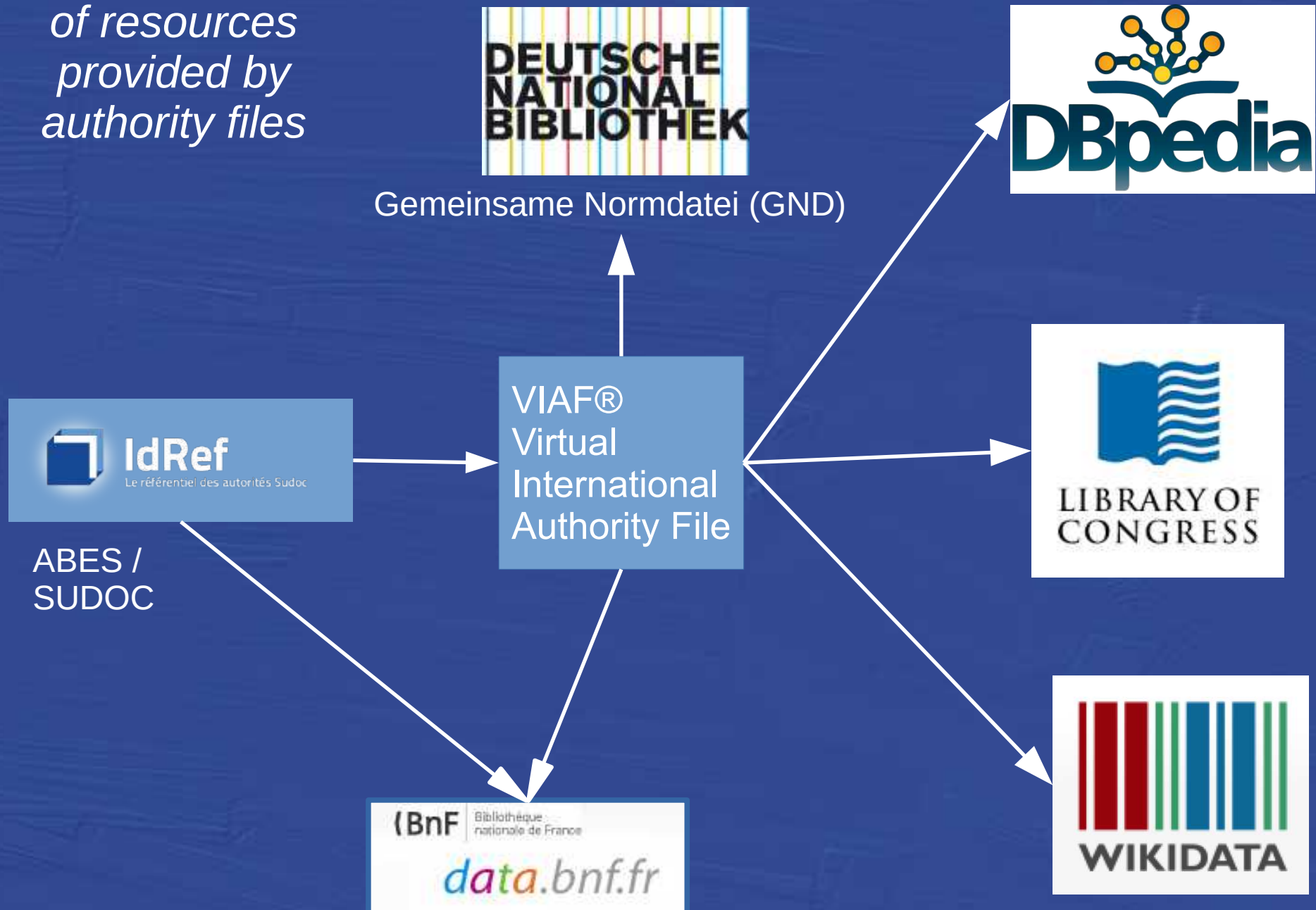
a collaborative VRE
in the context of the Semantic Web
and the FAIR principles

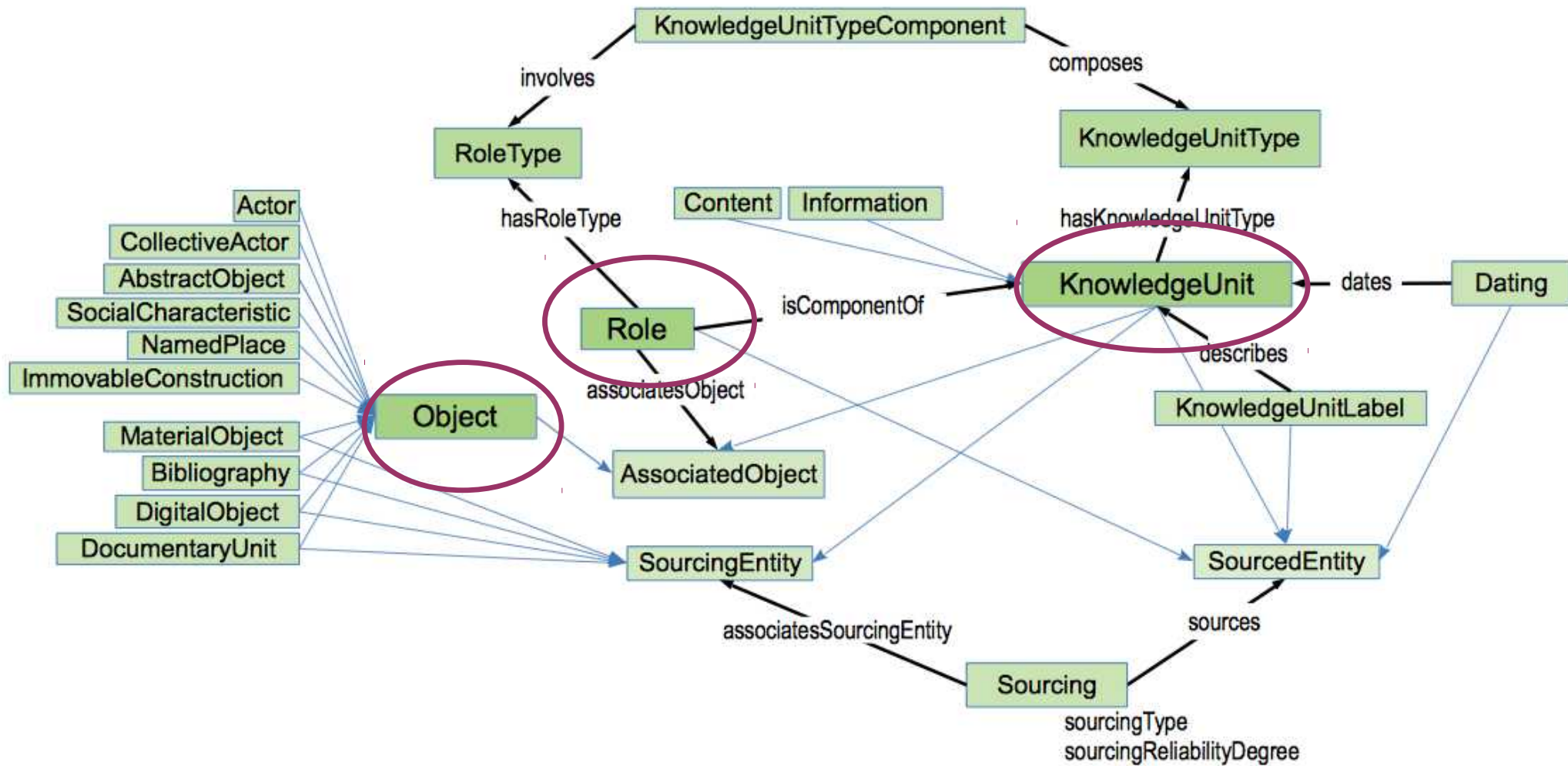


LOD diagram (2011)

by Anja Jentsch - Own work, CC BY-SA 3.0 <https://commons.wikimedia.org/w/index.php?curid=16574061>

*Interlinking the URIs
of resources
provided by
authority files*





The symogih.org ontology

SPARQL Endpoint

The symogih.org project's *modelling patterns* (more then 150 in 10 years)

SYMOGIH

Références

Accueil Documentation Membres

Références

- Arborescence des classes de types d'unités de connaissances
- Types d'informations**
- Types de contenus**

Objets

- Acteurs**
- Acteurs collectifs**
- Objets abstraits**
- Caractères sociaux**

Classes de

Chercher une

• Biographi

- Ense
- Exer
- Fin
- Lien
- Loca
- Rite
- Vie
- Vie

Enseignement

TyIn97

Exercer la fonction d'enseigner, avec indication de l'institution auprès de laquelle s'exerce l'enseignement et des matières enseignées.

Il s'agit d'un cas particulier du TyIn 'Exercice d'une fonction' : cf. Classe TyIn 'Exercice d'une fonction'.

Attention : ne pas renseigner le lieu si on peut localiser l'institution elle-même.

Liste des types de rôles associés

Libellé du type de rôle ▲	Clé du TyRo	Description
concerner	TyRo21	Institution auprès de laquelle s'exerce l'enseignement. Ce rôle a été gardé pour être l'équivalent du TyIn7 : Exercice d'une fonction.
enseigné (être)	TyRo131	La matière enseignée (un objet abstrait). On peut en associer plusieurs si on enseigne en même temps plusieurs matières. En revanche, il faut créer plusieurs informations si les enseignements des différentes matières se succèdent ou si le contexte institutionnel est différent
exercé (être)	TyRo47	Qualification de l'enseignement : professeur, chargé de cours, etc. Ce rôle a été gardé pour être l'équivalent du TyIn7 : Exercice d'une fonction.
exercer	TyRo12	Ce rôle a été gardé pour être l'équivalent du TyIn7 : Exercice d'une fonction.
localiser	TyRo8	Ne pas renseigner si l'institution auprès de laquelle s'effectue l'enseignement est déjà localisée.
occasionner la fin	TyRo176	Associe l'information ou le AbOb qui explique la fin de l'enseignement
origine (être l')	TyRo16	Associe l'information (nomination, élection, ...) ou l'objet abstrait qui indiquent la cause de l'enseignement
typer	TyRo98	A utiliser dans le contexte de ce TyIn pour spécifier la nature de l'enseignement grâce à un AbOb (cours magistral, séminaire, etc.).

MCD disponible(s)

Télécharger ce MCD

Patterns tree

Findable

Accessible

Interoperable

Re-usable

«There is an urgent need to improve the infrastructure supporting the **reuse** of scholarly data »

Wilkinson, Mark D., Michel Dumontier, Ijsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. “*The FAIR Guiding Principles for Scientific Data Management and Stewardship.*” *Scientific Data* 3 (March 15, 2016): 160018.

The FAIR Data Principles

To be Interoperable:

- I1. (meta)data use a *formal, accessible, shared, and broadly applicable language for **knowledge representation***.
- I2. (meta)data use *vocabularies that follow FAIR principles*.
- I3. (meta)data include qualified references to other (meta)data.

To be Re-usable:

- R1. meta(data) have a plurality of accurate and relevant attributes.
 - R1.1. (meta)data are released with a *clear and accessible data usage license*.
 - R1.2. (meta)data are associated with their *provenance*.
 - R1.3. (meta)data meet **domain-relevant community standards**.

2.

Modelling historical factual information

A relevant distinction : information is not knowledge

Historical factual information

“In 1592, Galileo Galilei was hired by the University of Padua, where he taught mathematics until 1610.”

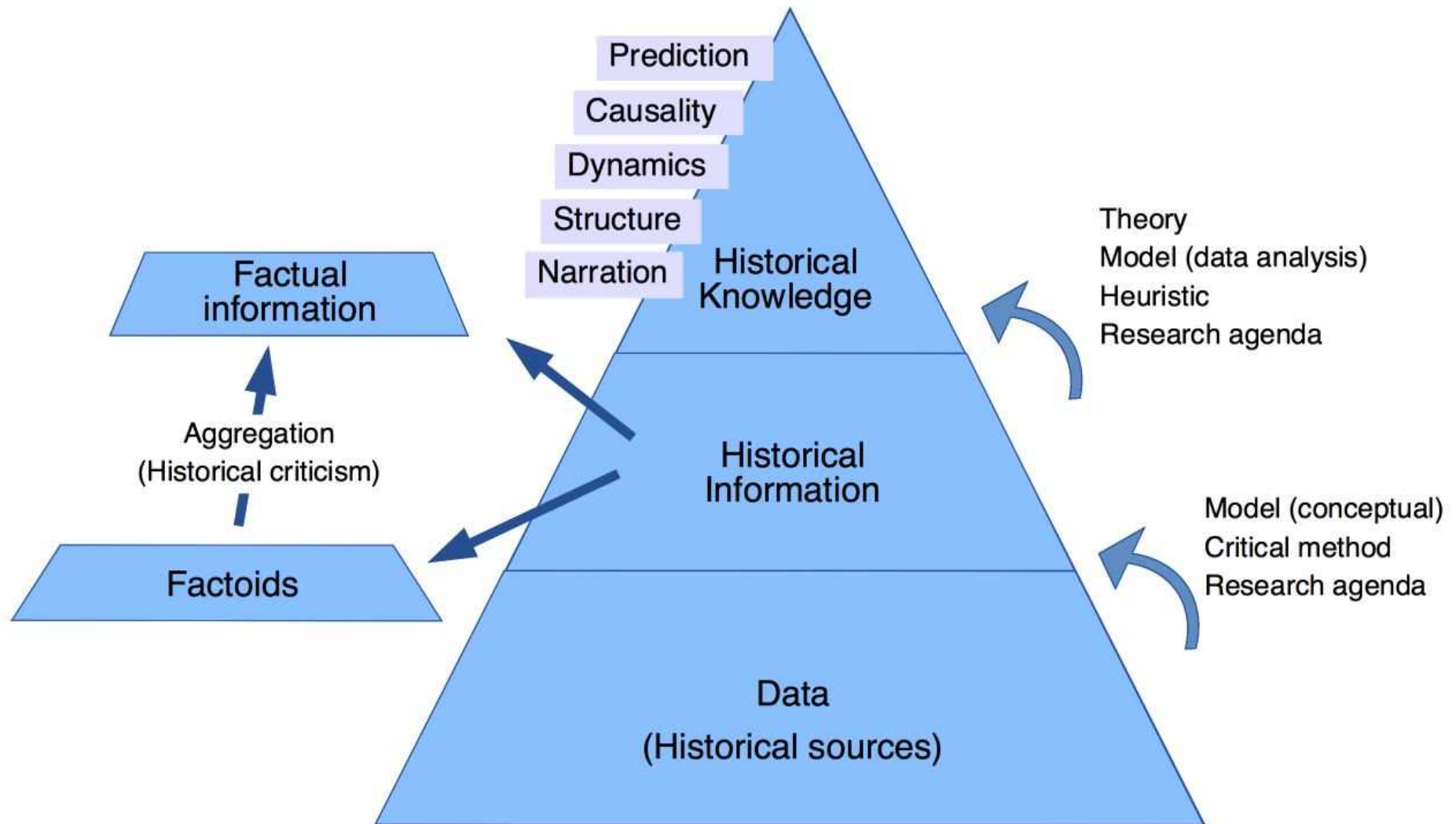
Letter by A to B, 11 March 1616 / Dictionary of astronomers, 2020

Historical knowledge

“Galileo Galilei was convicted in 1633 not because of heresy but for political reasons, in the context of the Thirty Years' War.”

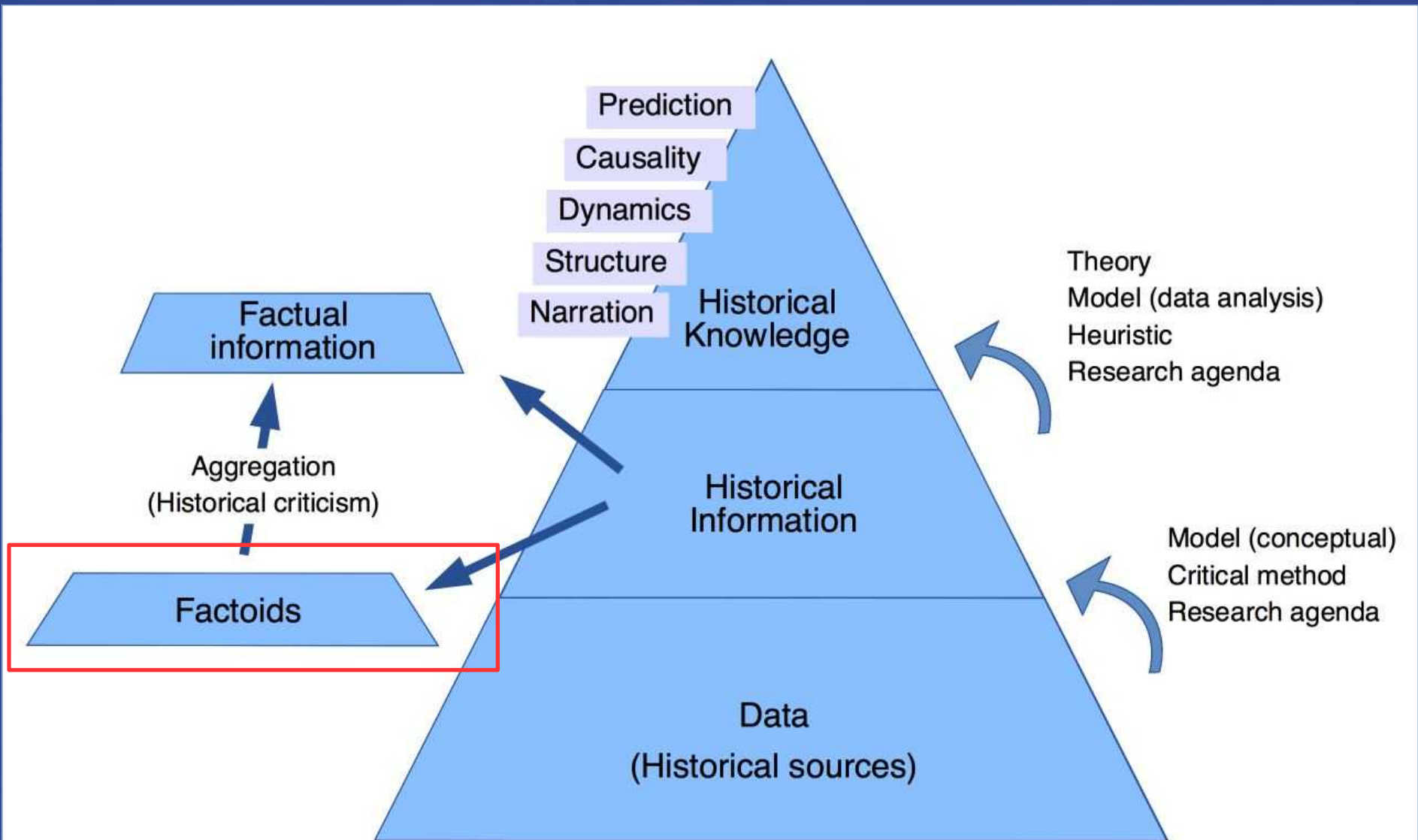
Dictionary of astronomers, 2020

Information is not knowledge : DIK(W) pyramid



Francesco Beretta (CNRS/Université de Lyon), 7 July 2020 CC BY-NC-SA 4.0

Factoids : the content and perspective of the sources



Francesco Beretta (CNRS/Université de Lyon), 7 July 2020 CC BY-NC-SA 4.0

Mentions of states of affairs in the text:

Galileo Galilei taught mathematics at the University of Padua from 1592 and 1610

```
<corresp="Info94542">In 1592, <rs ref="Actr161">he</rs> [Galileo Galilei] moved to the <name ref="CoAc54">University of Padua</name> where he taught geometry, mechanics, and astronomy until 1610.</p>
```

Galilei, Galileo - Enseigne : Mathématiques, auprès de : Université de Padoue

Info94542

Type d'information: Enseignement - TyIn97

Date: 1592

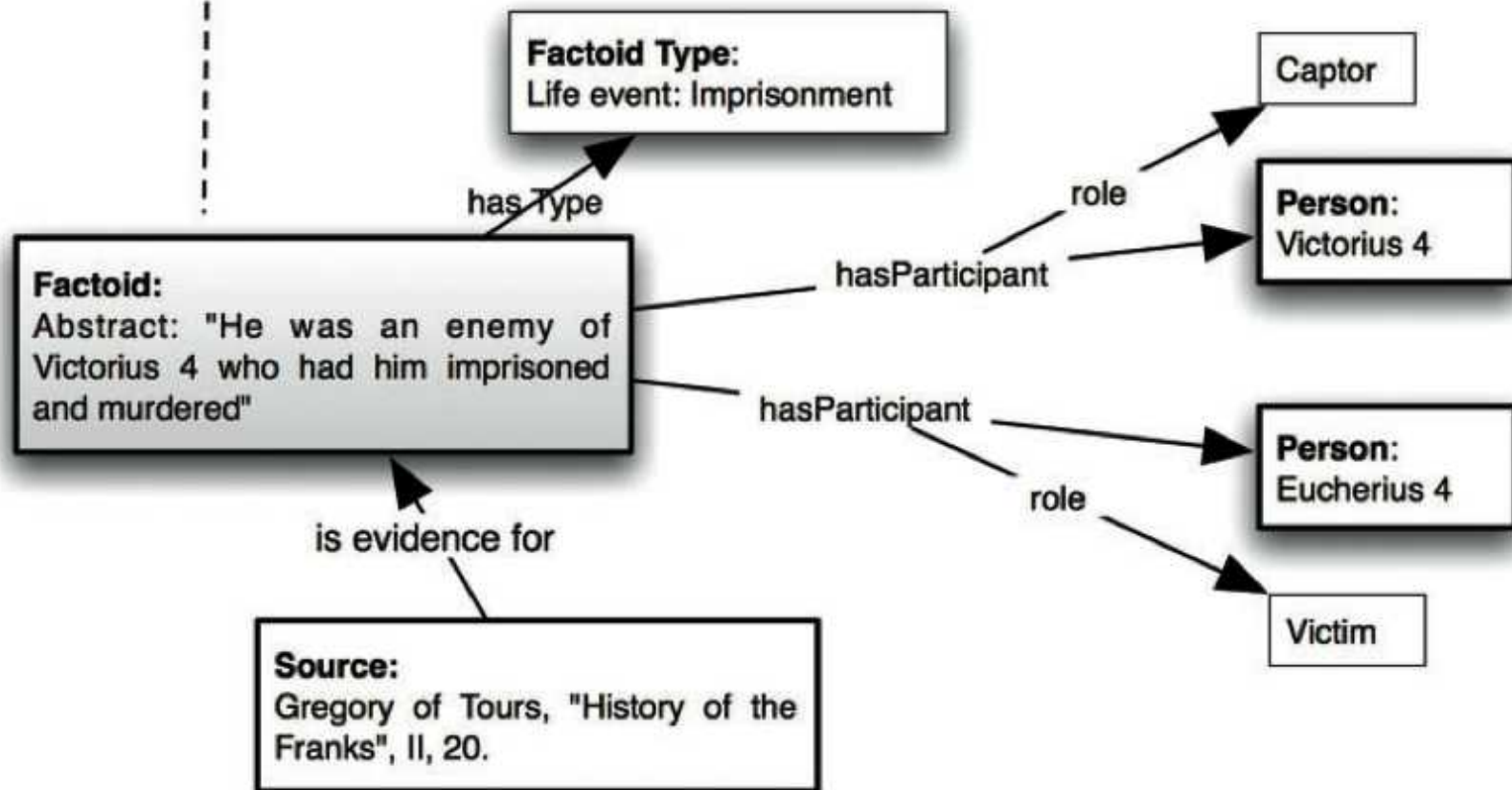
Composantes de l'information

Rôles

Textes

Sources

Libellé de l'objet	Type de rôle	Clé du rôle
Galilei, Galileo	exercer	InRo261100
Université de Padoue	concerner	InRo261101
<u>Mathématiques</u>	enseigné (être)	InRo261102

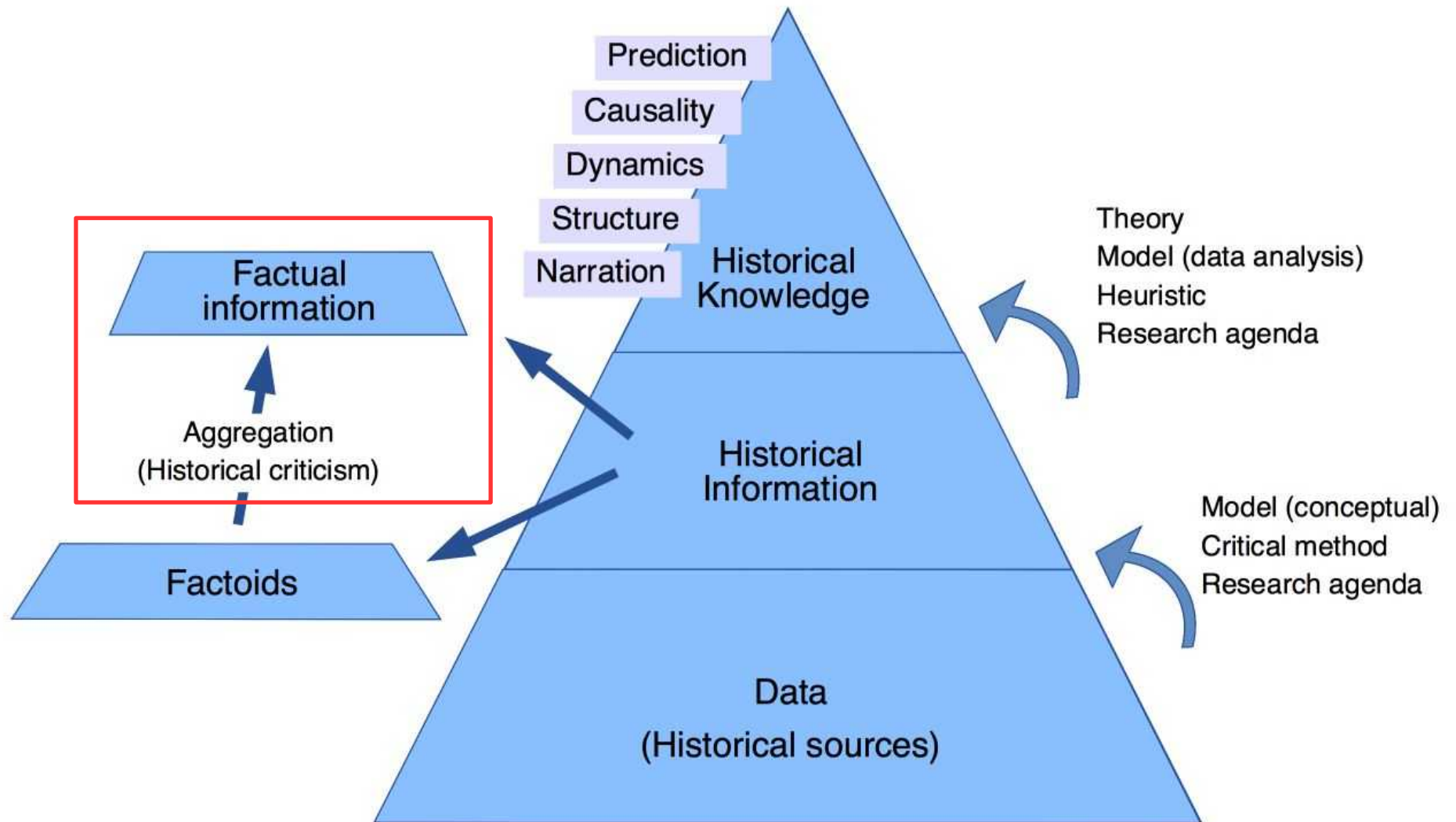


Factoid ontology

KCL's Department of Digital Humanities (DDH : Prosopographies of the Byzantine World (PBE and PBW), Anglo-Saxon England (PASE), Medieval Scotland (PoMS), Anglo-Scottish cross-border society ('Breaking of Britain': BoB)

Michele Pasin and John Bradley, Factoid-based prosopography and computer ontologies: Towards an integrated approach, Literary and Linguistic Computing Advance Access published June 29, 2013

Factual information as a representation of states of affairs in the world



Francesco Beretta (CNRS/Université de Lyon), 7 July 2020 CC BY-NC-SA 4.0

SYMOGIH

Références

Accueil

Actualités

Documentation

Membres

Références

- Arborescence des classes de types d'unités de connaissances
- Types d'informations
- Types de contenus

Objets

- Acteurs
- Acteurs collectifs
- Lieux
- Objets abstraits
- Caractères sociaux
- Formes concrètes

Galilei, Galileo - Enseigne : Mathématiques, auprès de : Université de Padoue

Info94542

Type d'information: Enseignement - TyIn97

Date: 1592

Composantes de l'information

Rôles

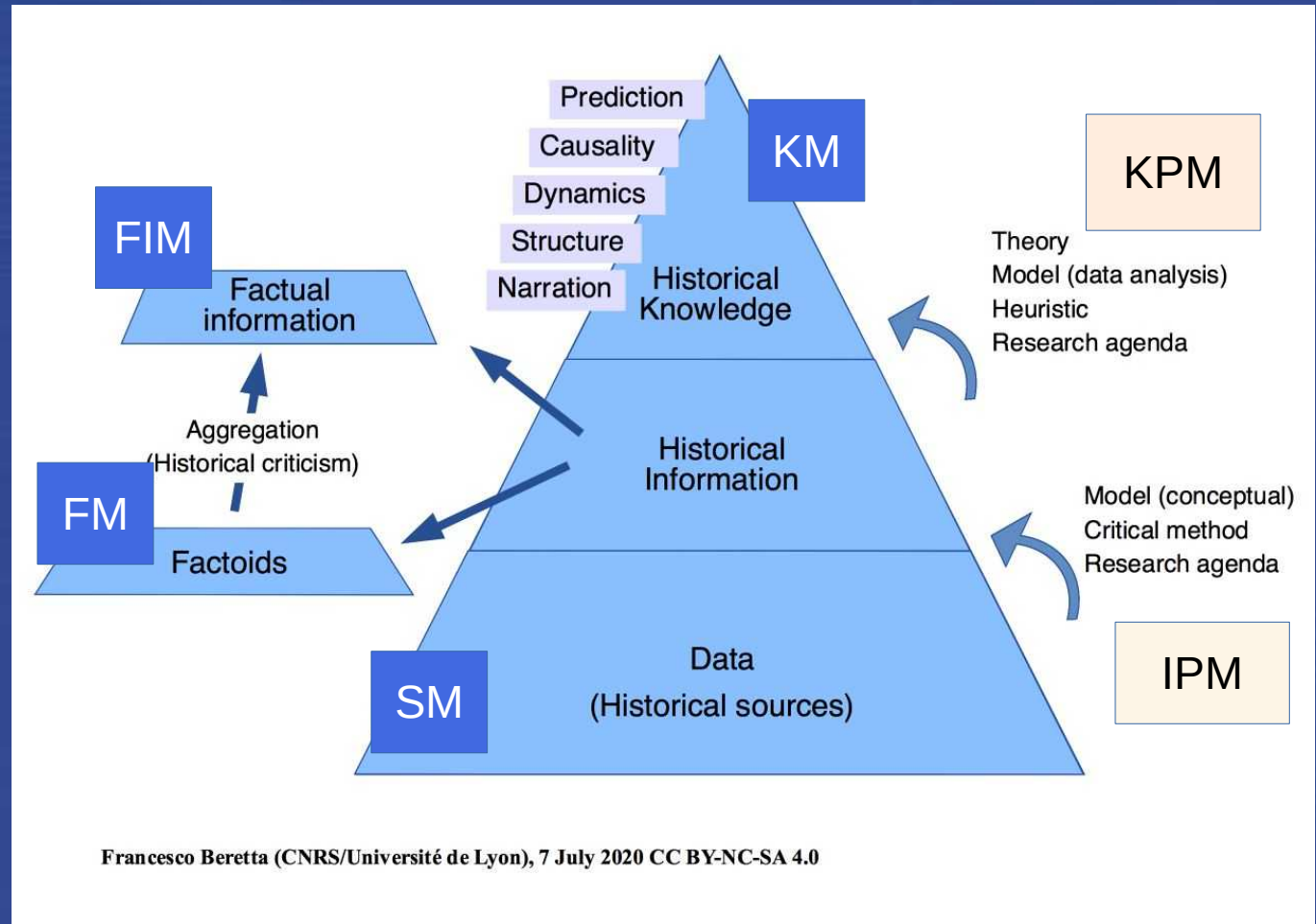
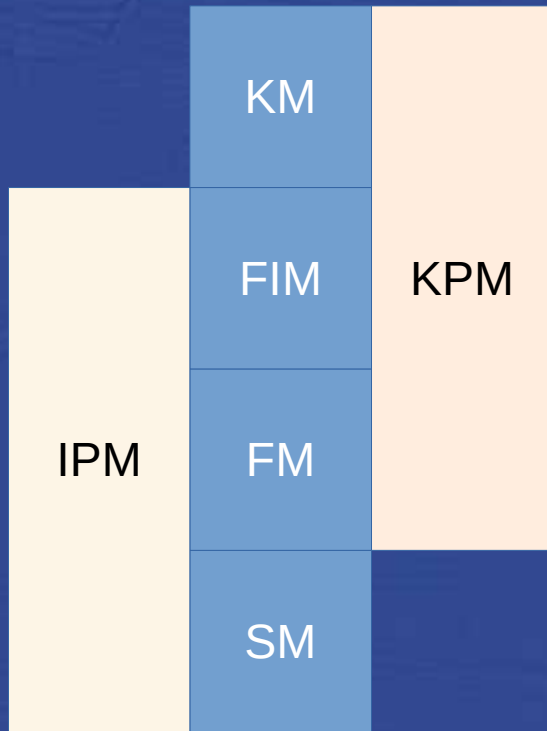
Textes

Sources

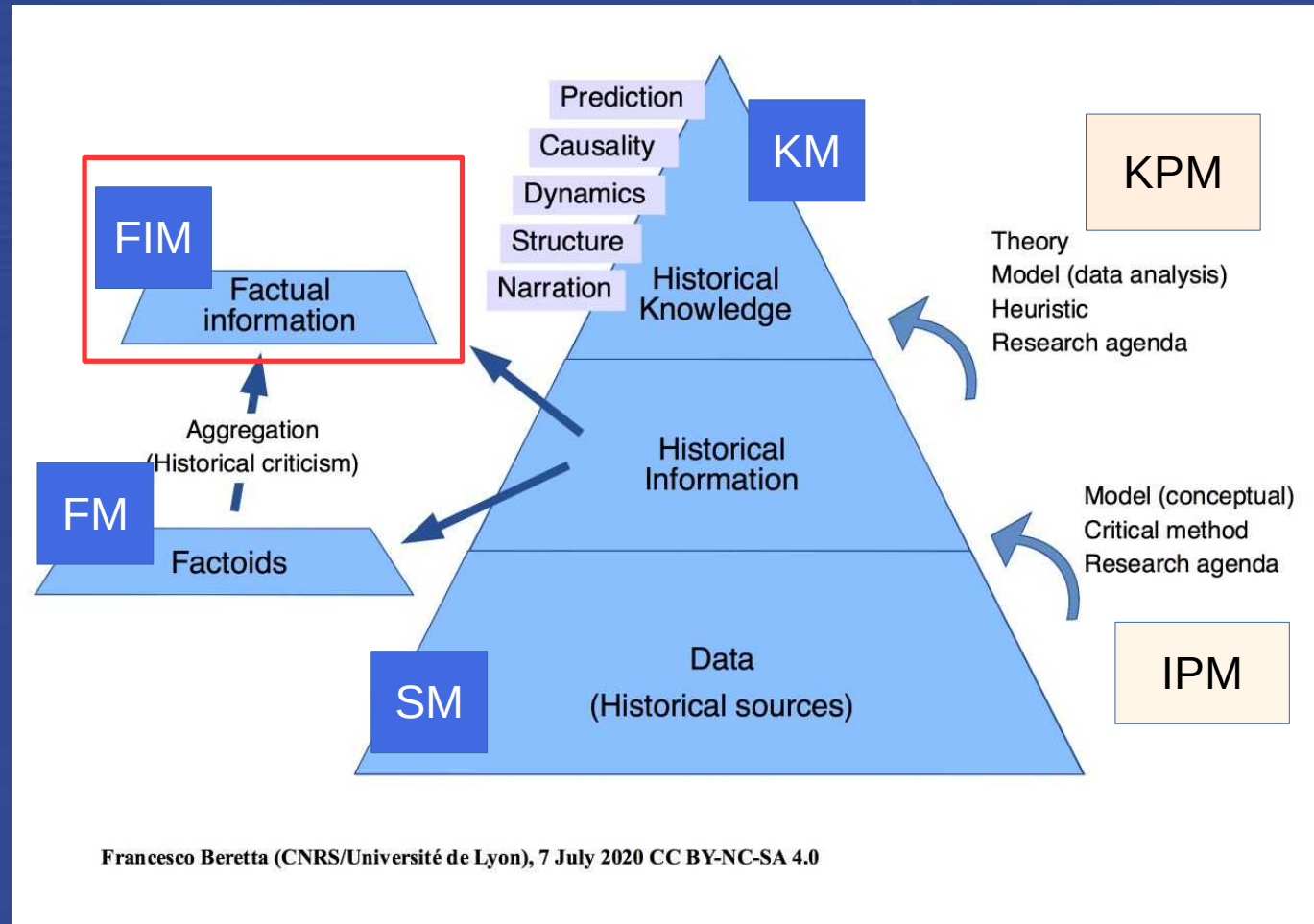
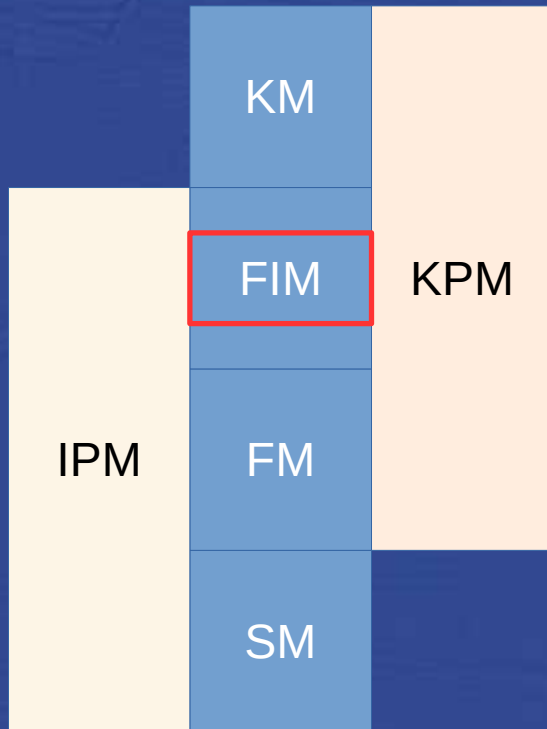
Libellé de l'objet	Type de rôle	Clé du rôle
Galilei, Galileo	exercer	InRo261100
Université de Padoue	concerner	InRo261101
Mathématiques	enseigné (être)	InRo261102

Galileo Galilei taught mathematics at the University of Padua from 1592 to 1610

Different models for different aspects of historical knowledge production



Different models for different aspects of historical knowledge production



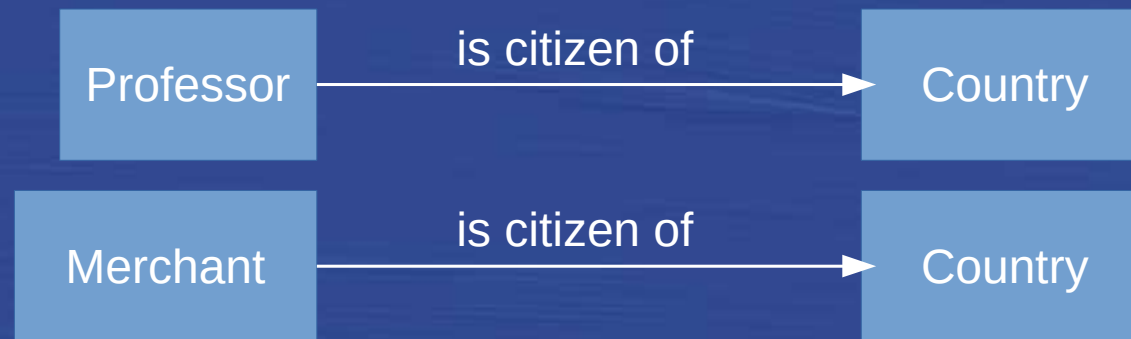
The FAIR data challenge :

modelling *factual historical information*,
then reusing it for a different research agenda

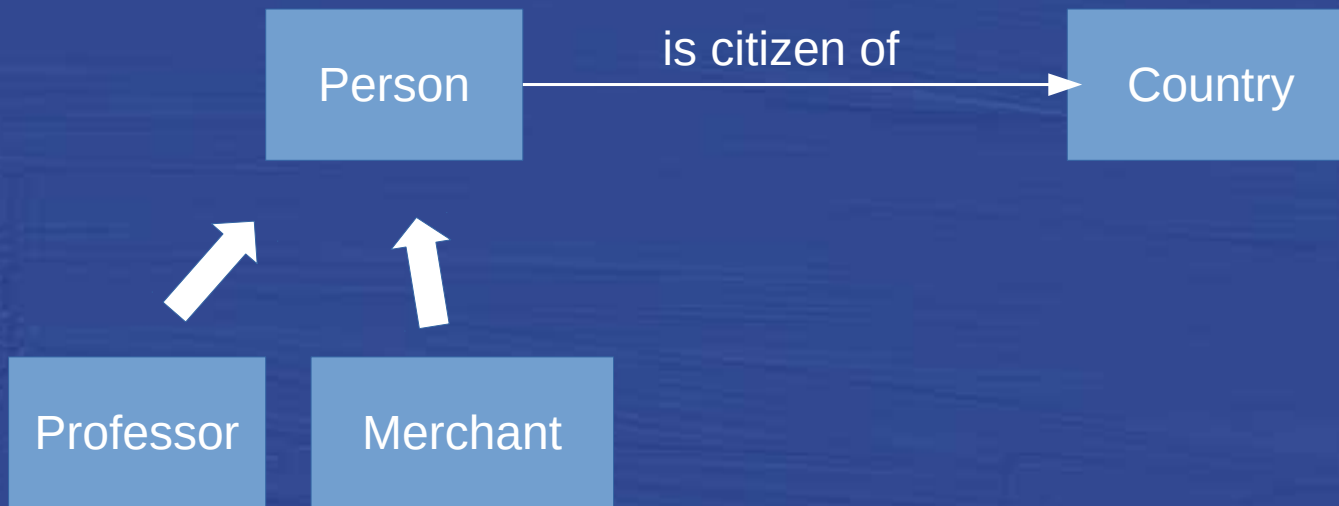
“In 1592, Galileo Galilei was hired by the University of Padua, where he taught mathematics until 1610.”

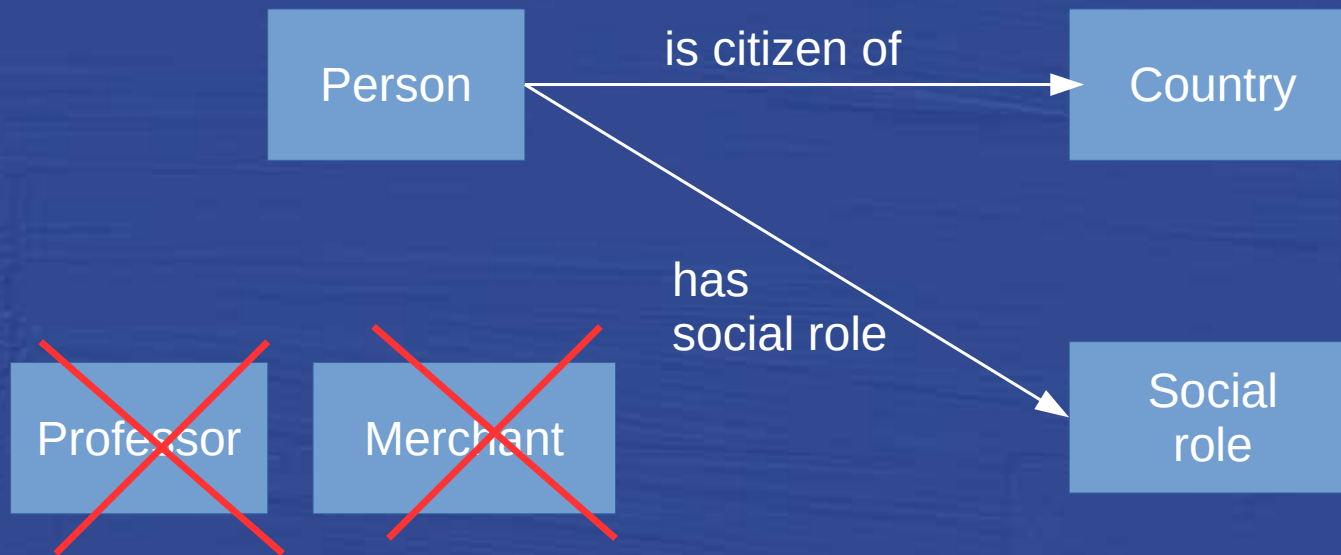
Different factual information pieces → different modelling patterns:

- Galileo Galilei taught mathematics at the University of Padua from 1592 to 1610
- that Galileo from 1592 resided in the city of Padua
- that he was hired by and was member of the University
- that he held the title of professor regardless of whether or not he was effectively teaching

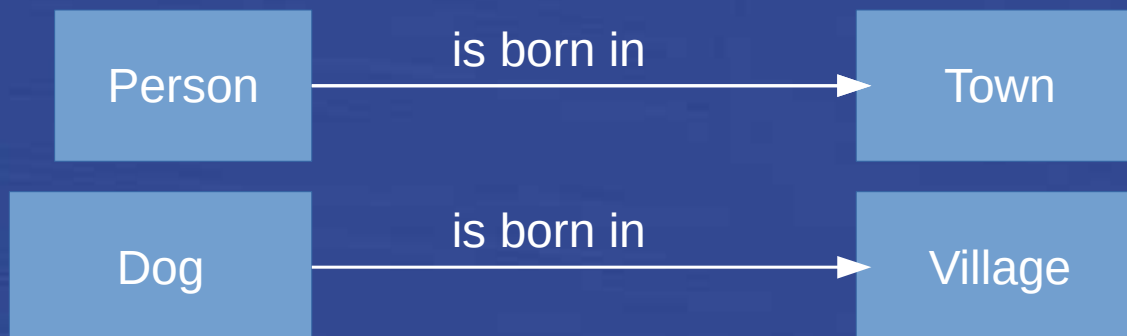
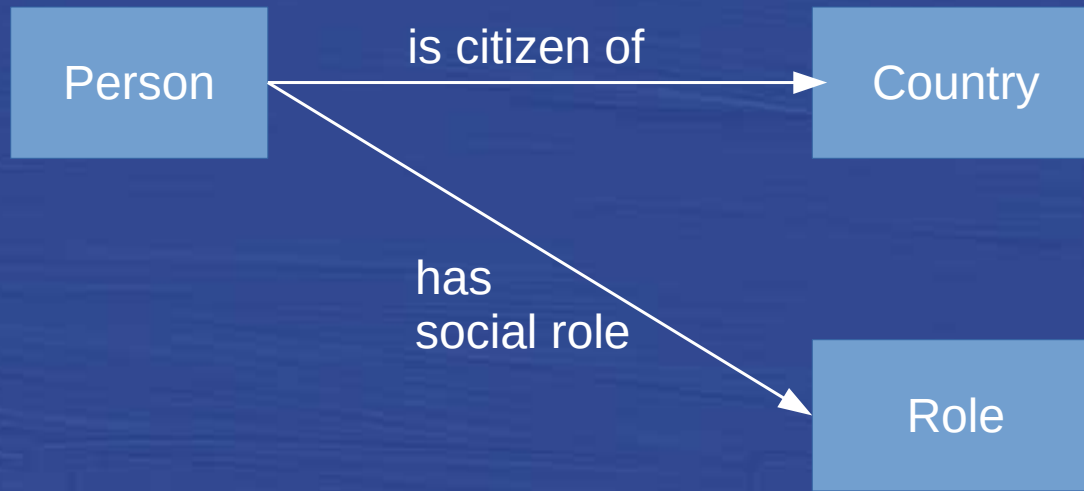


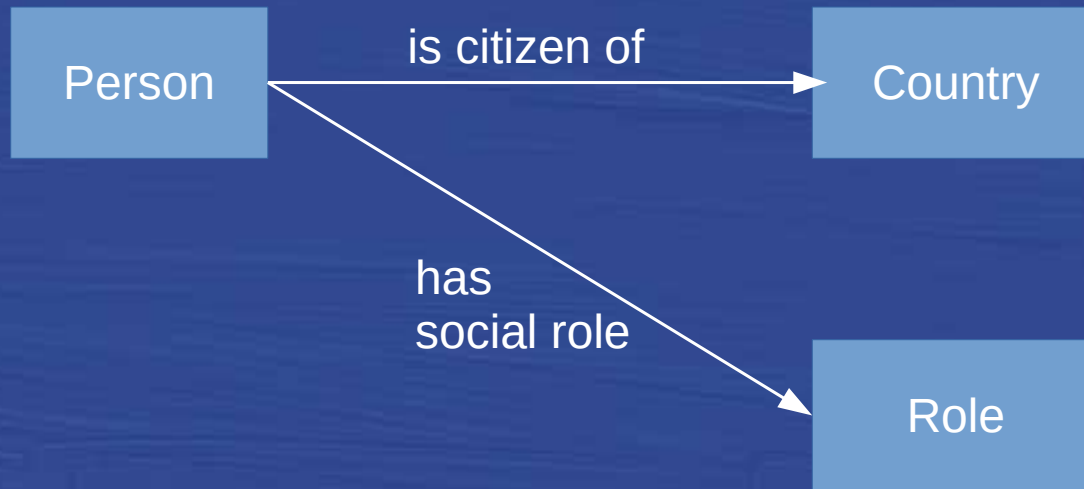
The issue of the suitable *conceptualization*



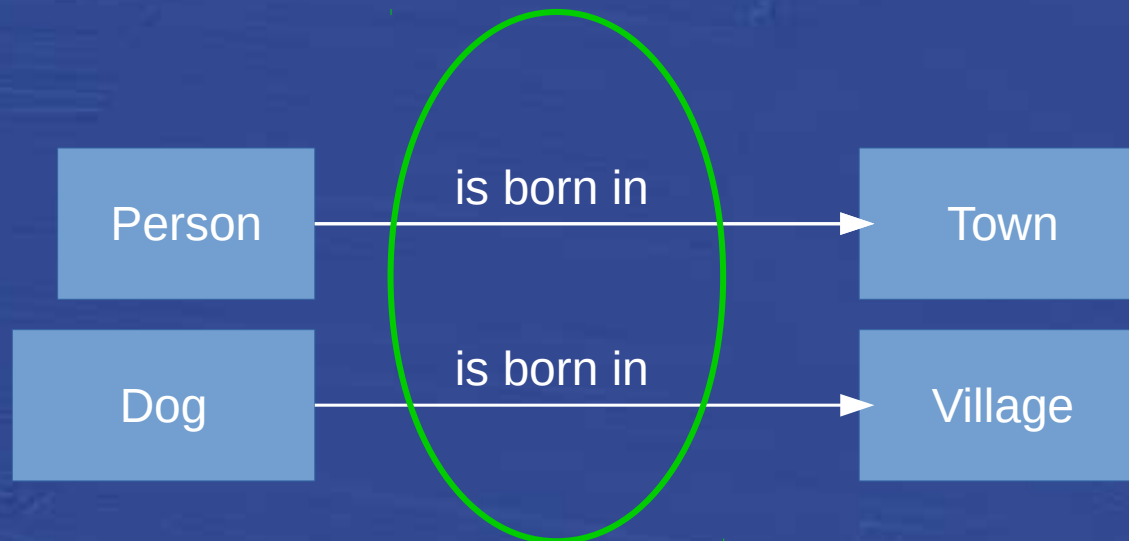


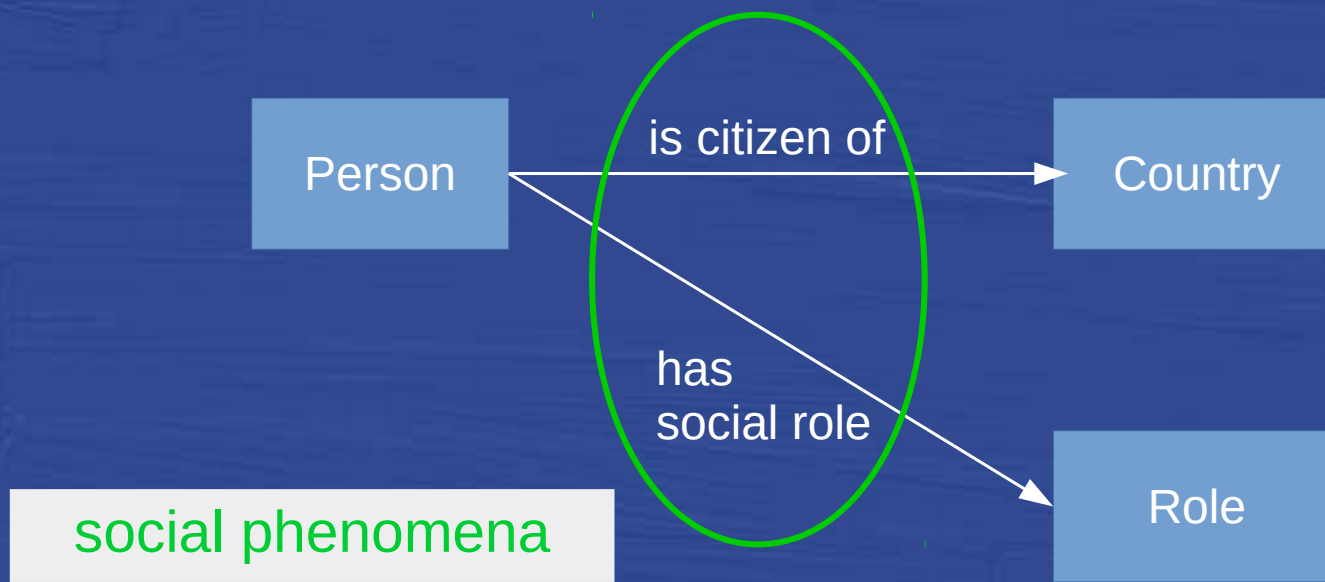
The issue of the suitable *conceptualization*



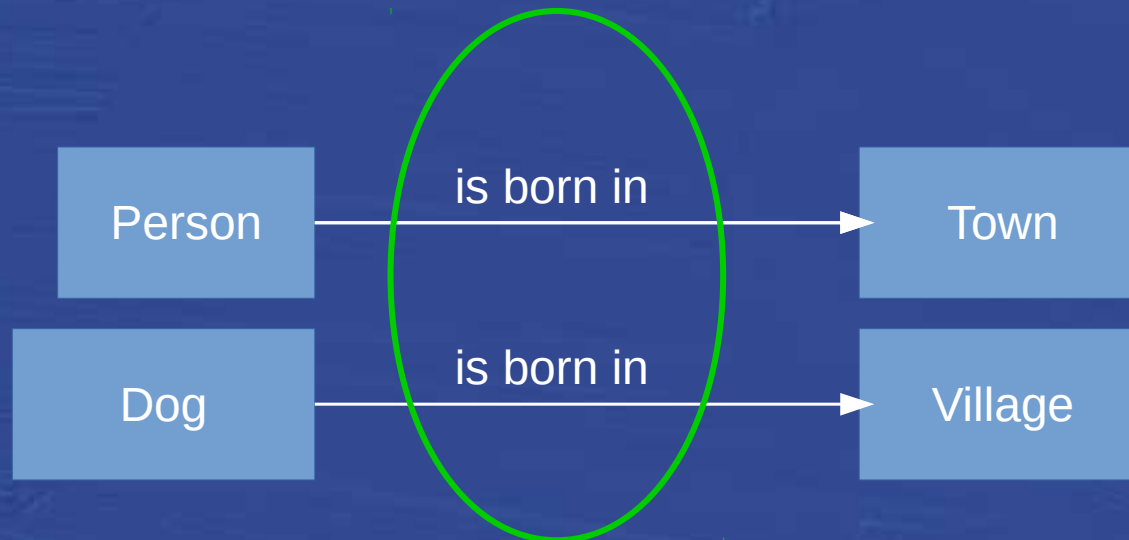


physical or biological phenomena





physical or biological phenomena



The issue of the suitable *conceptualization*

3.

Foundational ontologies

and

historical research data

“An ontology is
a formal explicit specification
of a shared conceptualization
of a domain of interest”

- « Formality – ... a knowledge representation language that is based on the grounds of **formal semantics**. »
- « Consensus – ... an agreement on a domain conceptualization among people in a community. »
- « Conceptuality – ... in terms of conceptual symbols that can be intuitively grasped by humans, as they correspond to the elements in their **mental models**. »
- « Domain Specificity – ... limited to knowledge about a particular **domain of interest**. »

[Domingue et al. 2011, p. 510-511]

The relevant domain : historical knowledge production

Historical factual information

“In 1592, Galileo Galilei was hired by the University of Padua, where he taught mathematics until 1610.”

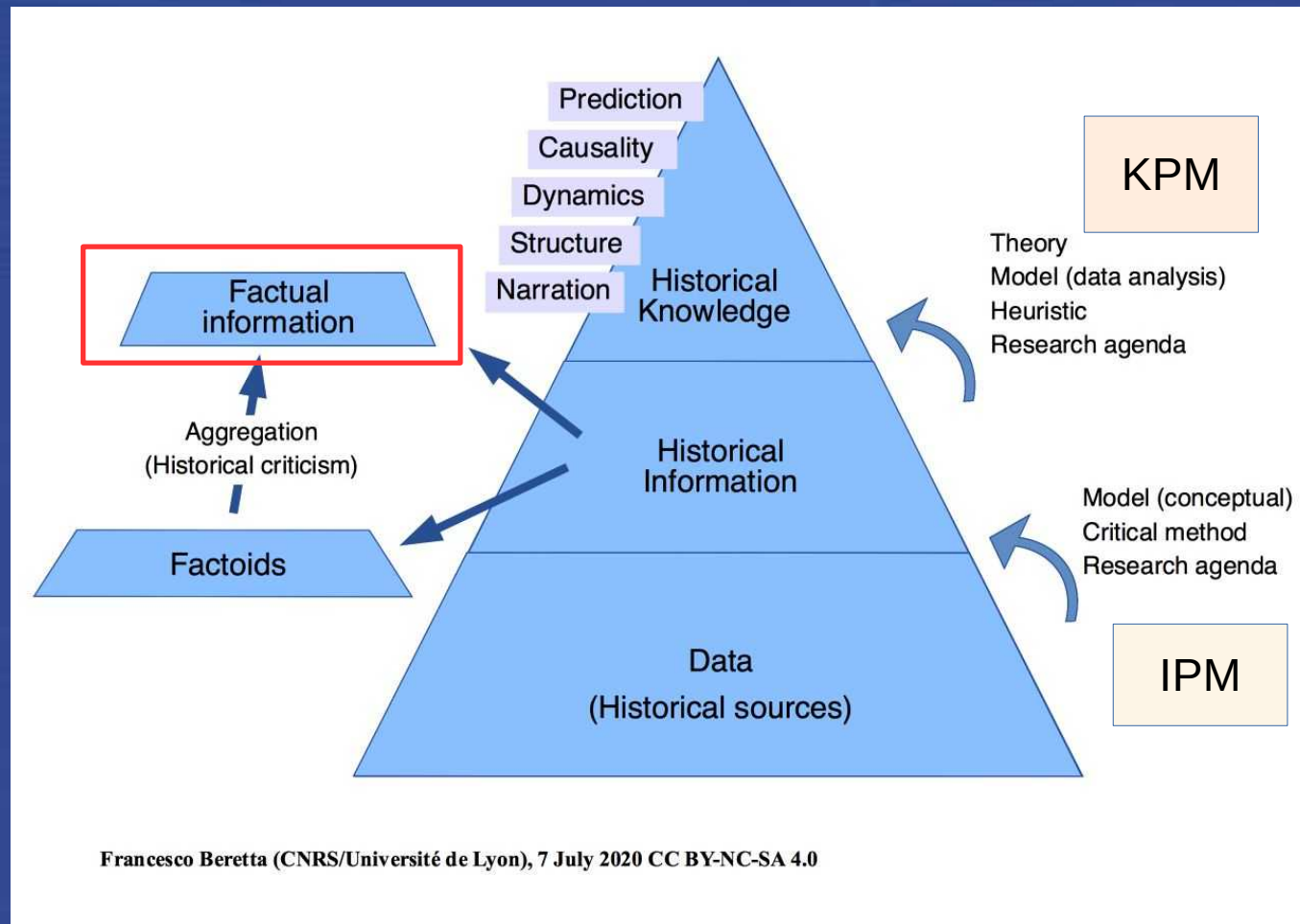
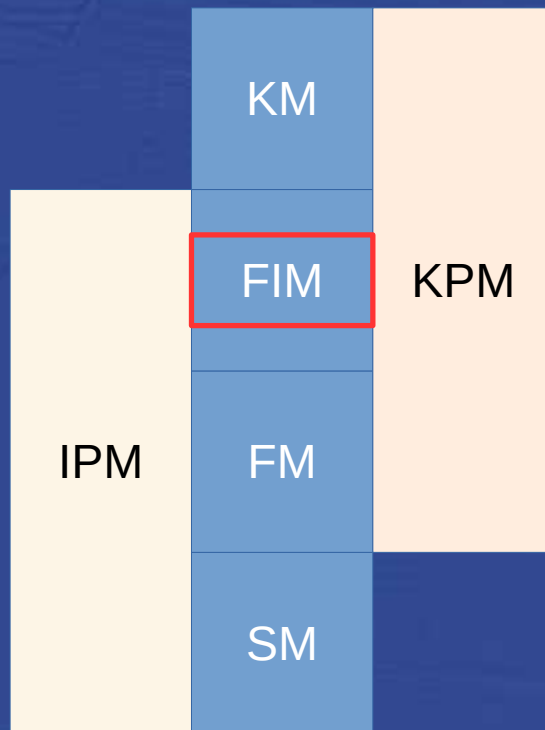
Letter by A to B, 11 March 1616 / Dictionary of astronomers, 2020

Historical knowledge

“Galileo Galilei was convicted in 1633 not because of heresy but for political reasons, in the context of the Thirty Years' War.”

Dictionary of astronomers, 2020

The focus : modelling factual information

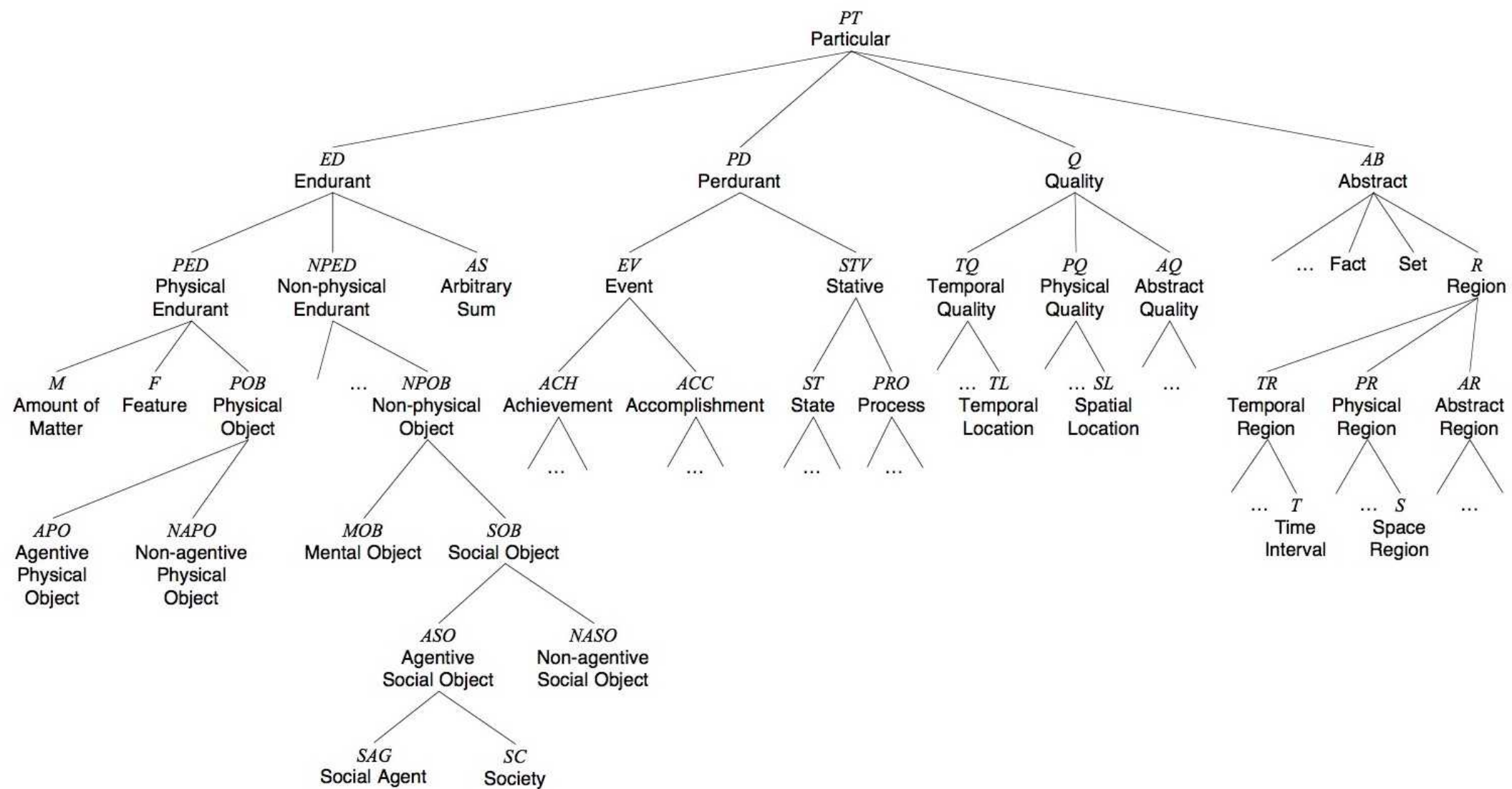


The issue of a *suitable conceptualization* :

Foundational ontologies

were developed to support the
verification and improvement

of the **conceptualization** of a **domain of discourse**.



Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) – a foundational ontology designed in 2002 in the context of the WonderWeb EU project, developed by Nicola Guarino and his associates at the Laboratory for Applied Ontology (LOA) – WonderWeb Deliverable D18, p.14

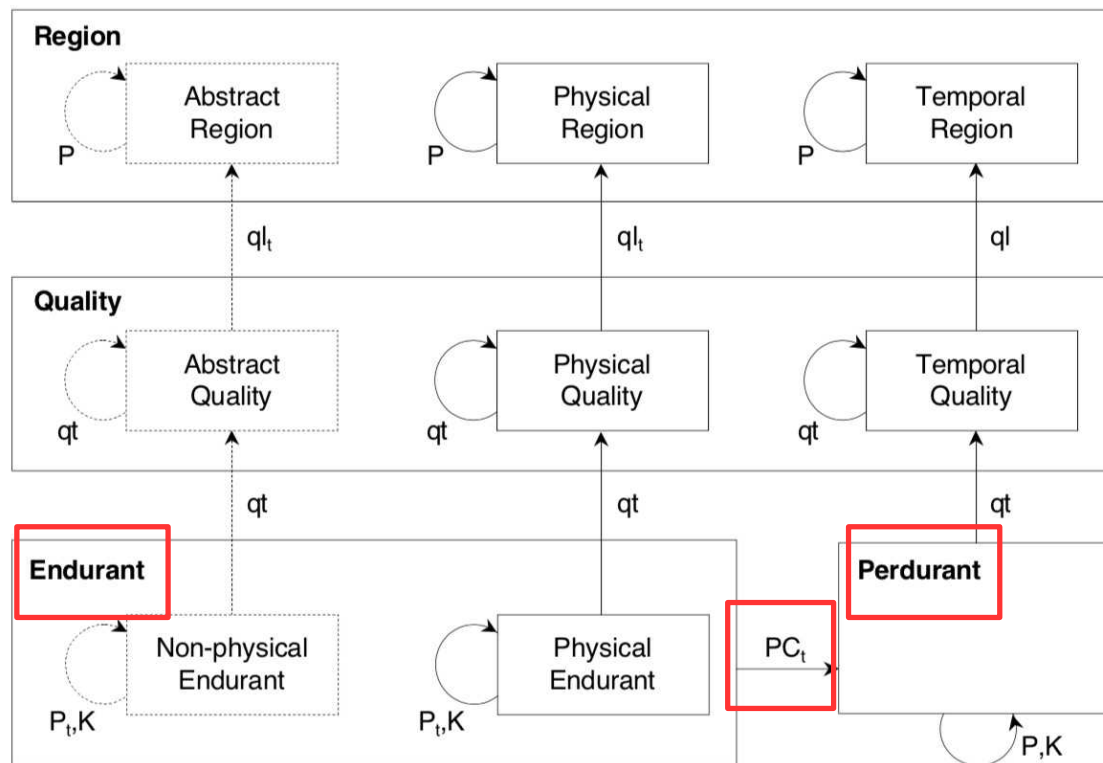


Figure 4: Primitive relations between basic categories (the dotted lines to the left indicate that we are less confident with what concerns non-physical endurants).

4.2.7 Participation

$$(Dd63) \quad PC_C(x, y) \triangleq \exists t (PRE(y, t)) \wedge \forall t (PRE(y, t) \rightarrow PC(x, y, t)) \quad (\text{Const. Participation})$$

$$(Dd64) \quad PC_T(x, y, t) \triangleq PD(y) \wedge \forall z ((P(z, y) \wedge PRE(z, t)) \rightarrow PC(x, z, t)) \\ (\text{Temporary Total Participation})$$

$$(Dd65) \quad PC_T(x, y) \triangleq \exists t (ql_T(t, y) \wedge PC_T(x, y, t)) \quad (\text{Total Participation})$$

$$(Dd66) \quad mpc(x, y) \triangleq x = \sigma_z (PC_T(z, y)) \quad (\text{Maximal Participant})$$

$$(Dd67) \quad mppc(x, y) \triangleq x = \sigma_z (PC_T(z, y) \wedge PED(z)) \quad (\text{Maximal Physical Participant})$$

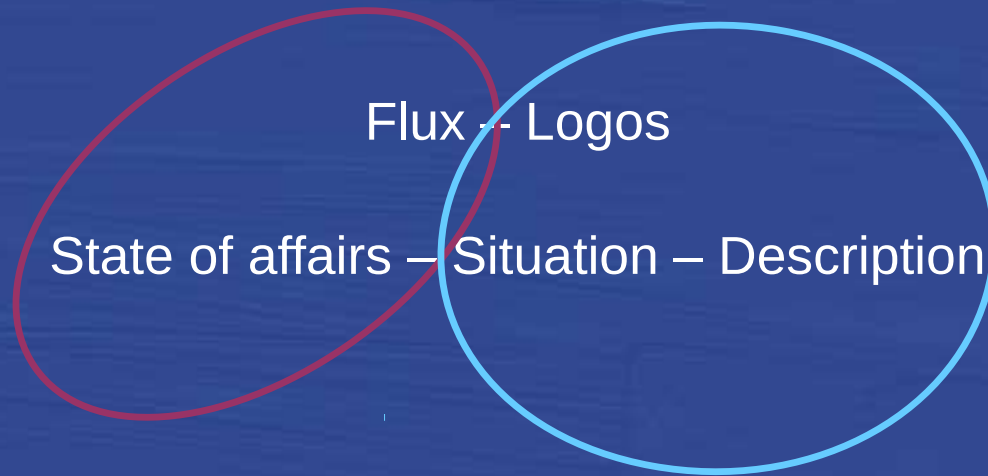
$$(Dd68) \quad lf(x, y) \triangleq x = \sigma_z (PC_T(y, z)) \quad (\text{Life})$$

Descriptions and Situations (DnS)

Flux – Logos

State of affairs – Situation – Description

Descriptions and Situations (DnS)



Descriptions and Situations (DnS)

Flux – Logos

State of affairs – Situation – Description



Descriptions and Situations (DnS)

Flux – Logos

State of affairs – Situation – Description



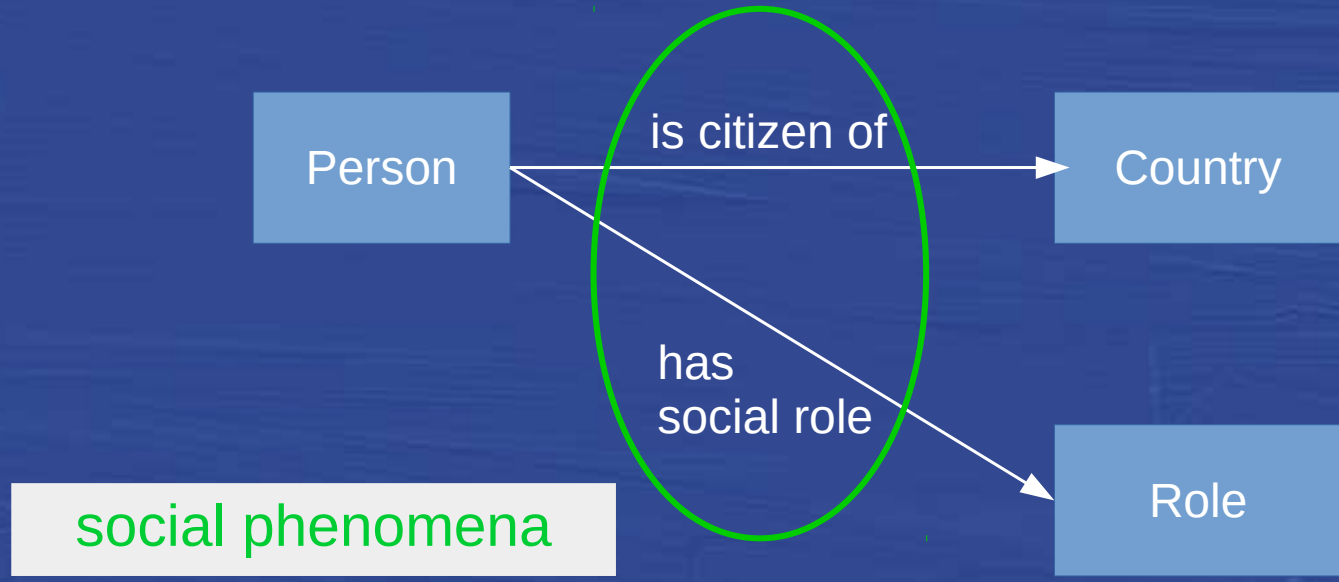
Social entities, social roles, collectives, membership, property, etc.

« If you want to express the legal constraints imposed by norms and regulations on the domain of your ground ontology, you have to extend the latter and add to it a DnS description of social reality under a legal perspective. Such DnS description makes it possible to describe the ideal (legal) view on the behaviour of your **social entities** (a situation), according to a given legal system. [...]

Reified constraints and theories are classified as social objects, which hold various properties like, for instance, a(n indirect) location in space and time. »

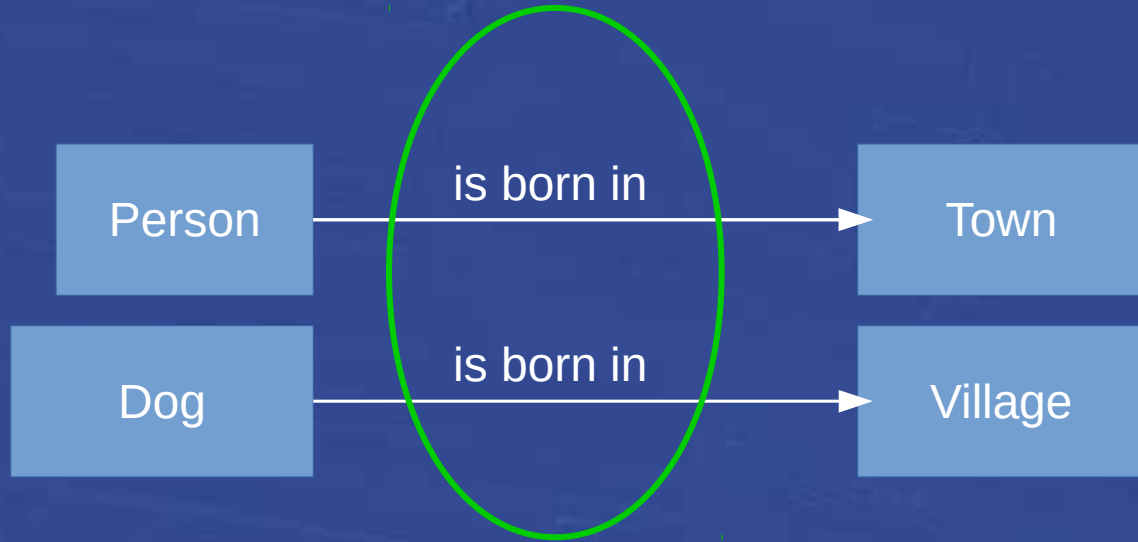
Bottazzi E., Catenacci C., Gangemi A., Lehmann J.(2006), p. 194-195

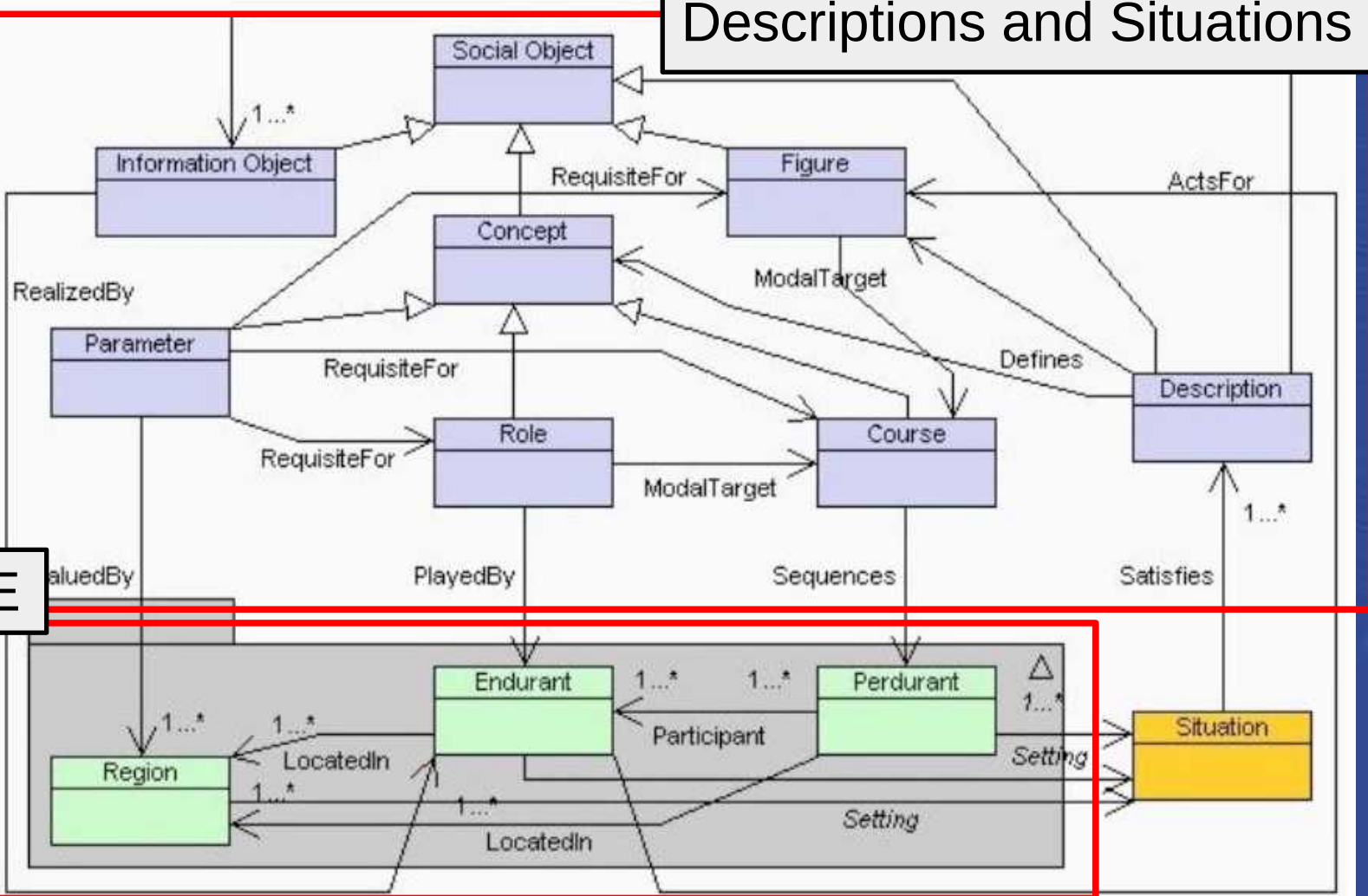
Descriptions
and
Situations –
DnS



physical or biological phenomena

DOLCE



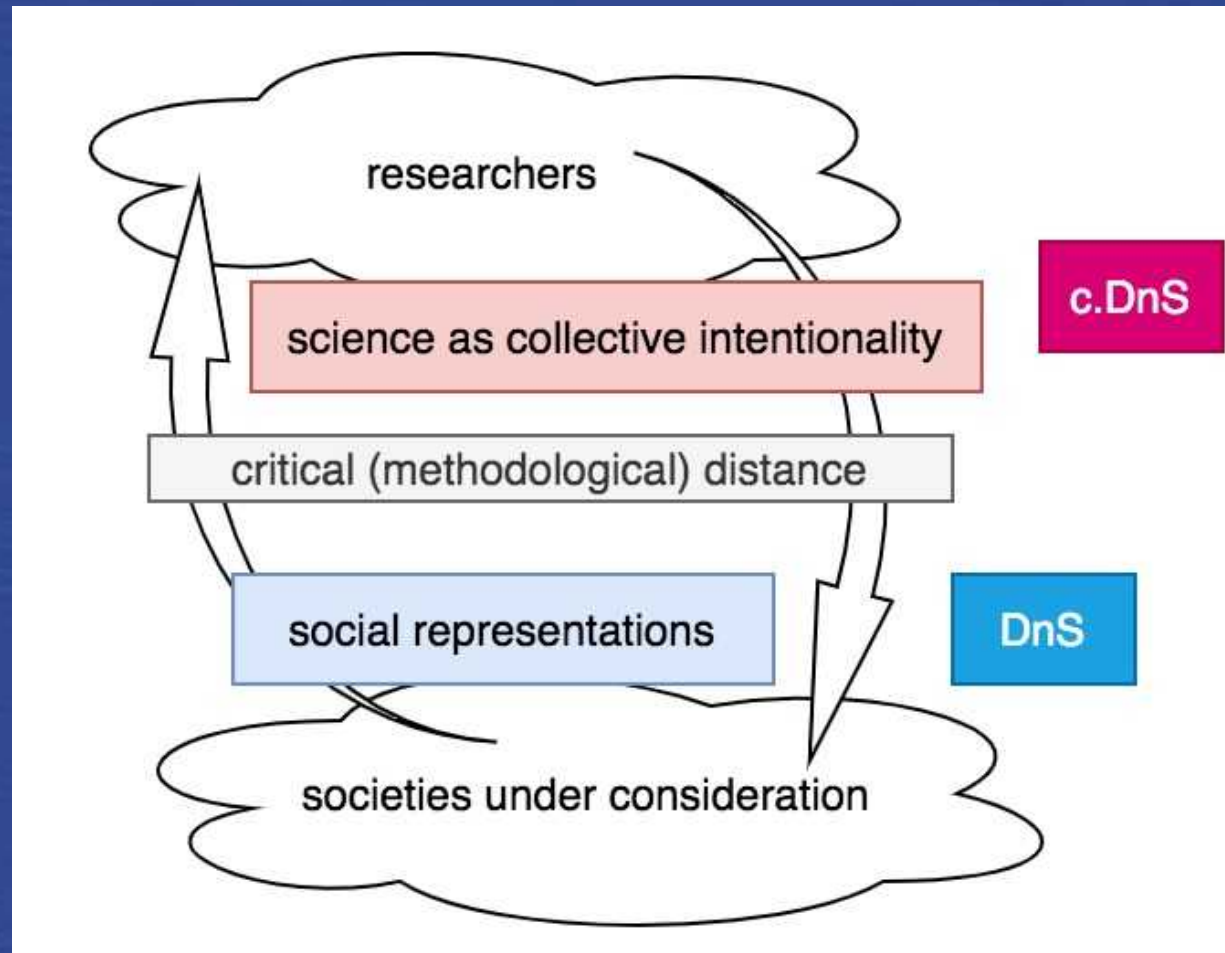


DOLCE

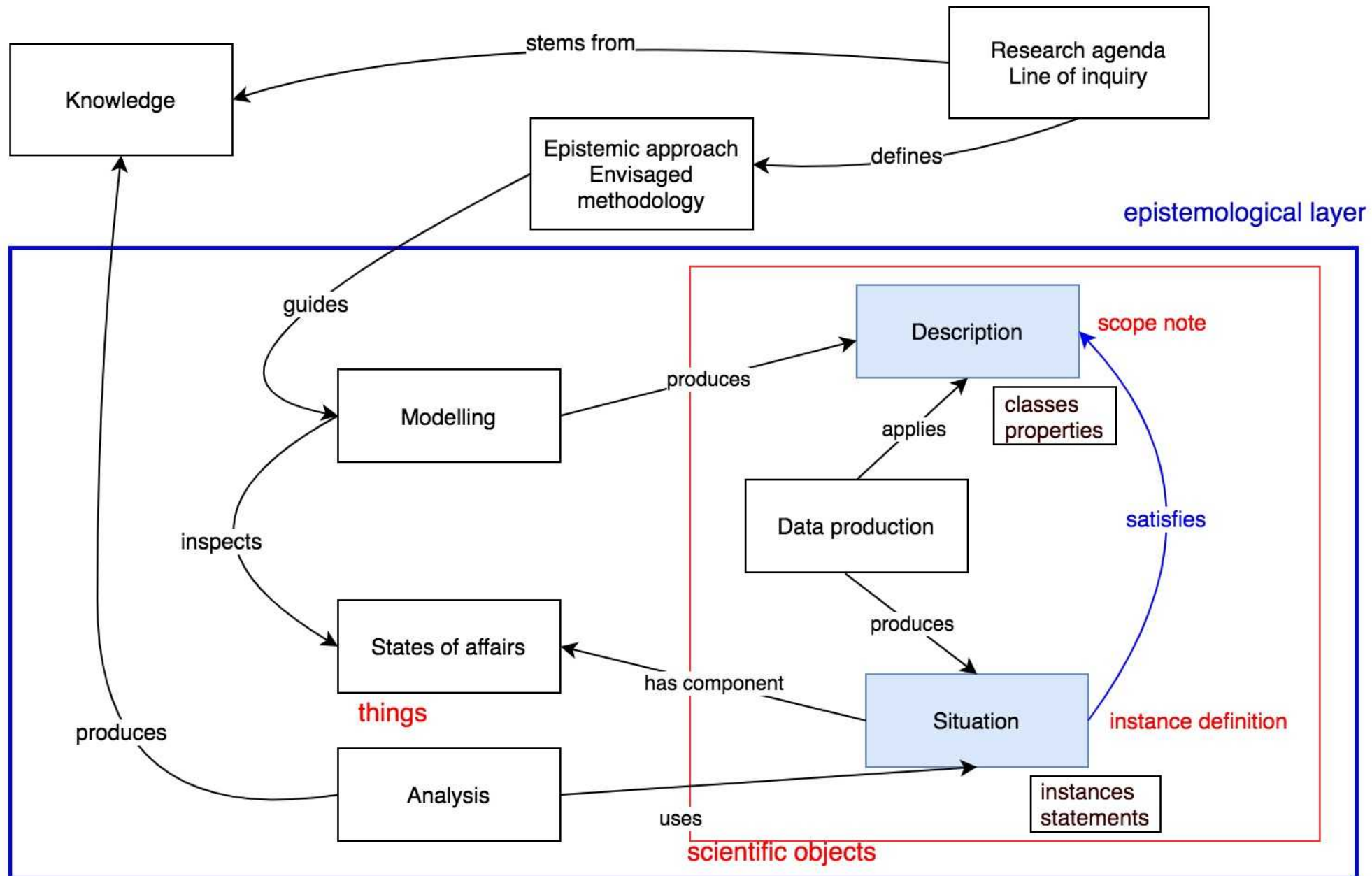
Fig. 2. A UML class diagram for D&S. The lower part of the pattern (within the grey package) is called the *ground ontology*, the higher is called the *descriptive ontology*; a situation satisfies a description if the two parts match according to the axioms specified for the concepts defined by the description.

Bottazzi E., Catenacci C., Gangemi A., Lehmann J.(2006) (from pre-print, not in published version)

Whose collective intentionality is to be modelled? Scientific knowledge and social representations



c.DnS : epistemological layering



4.

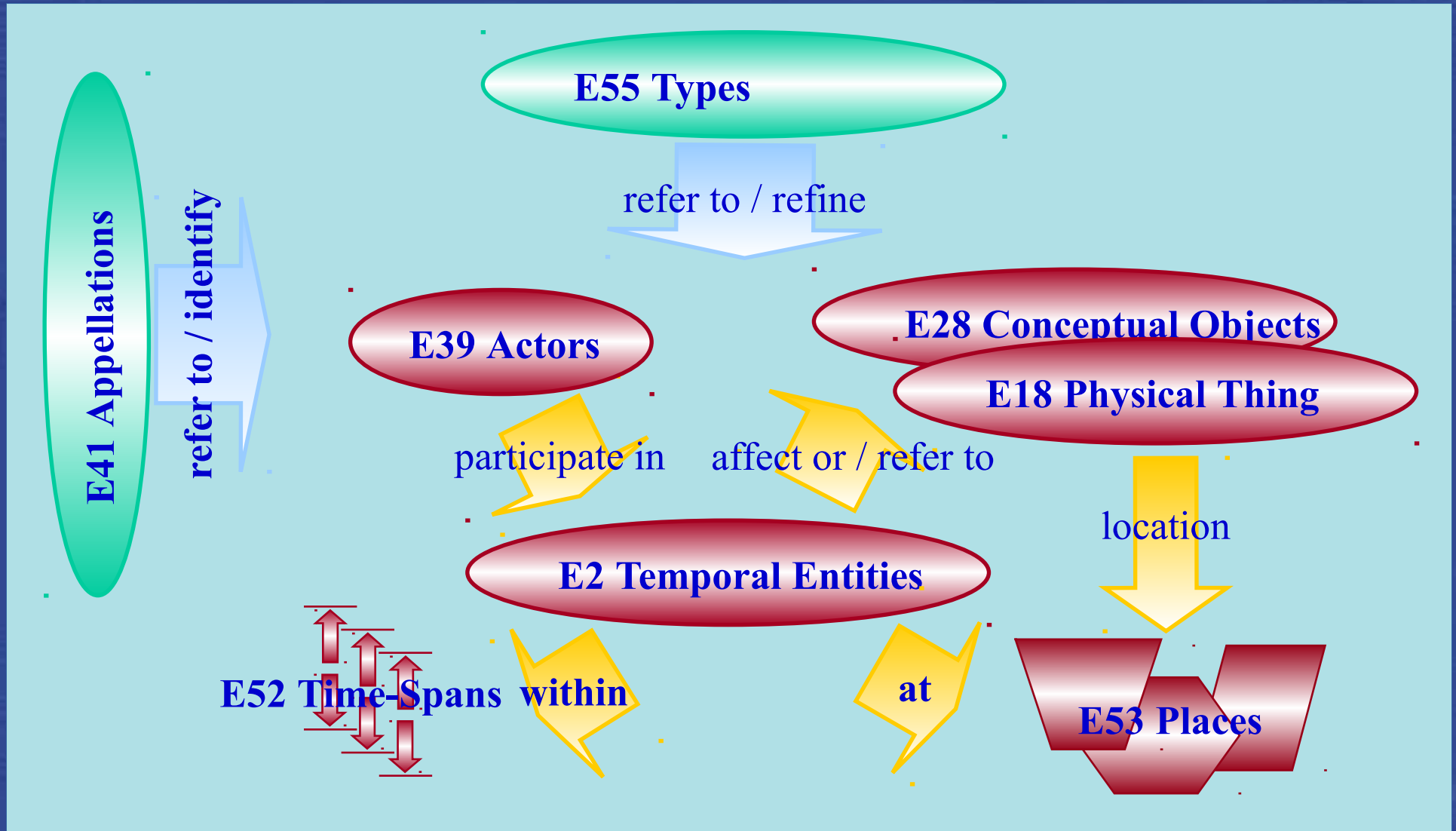
The CIDOC CRM

and

factual historical information conceptualization

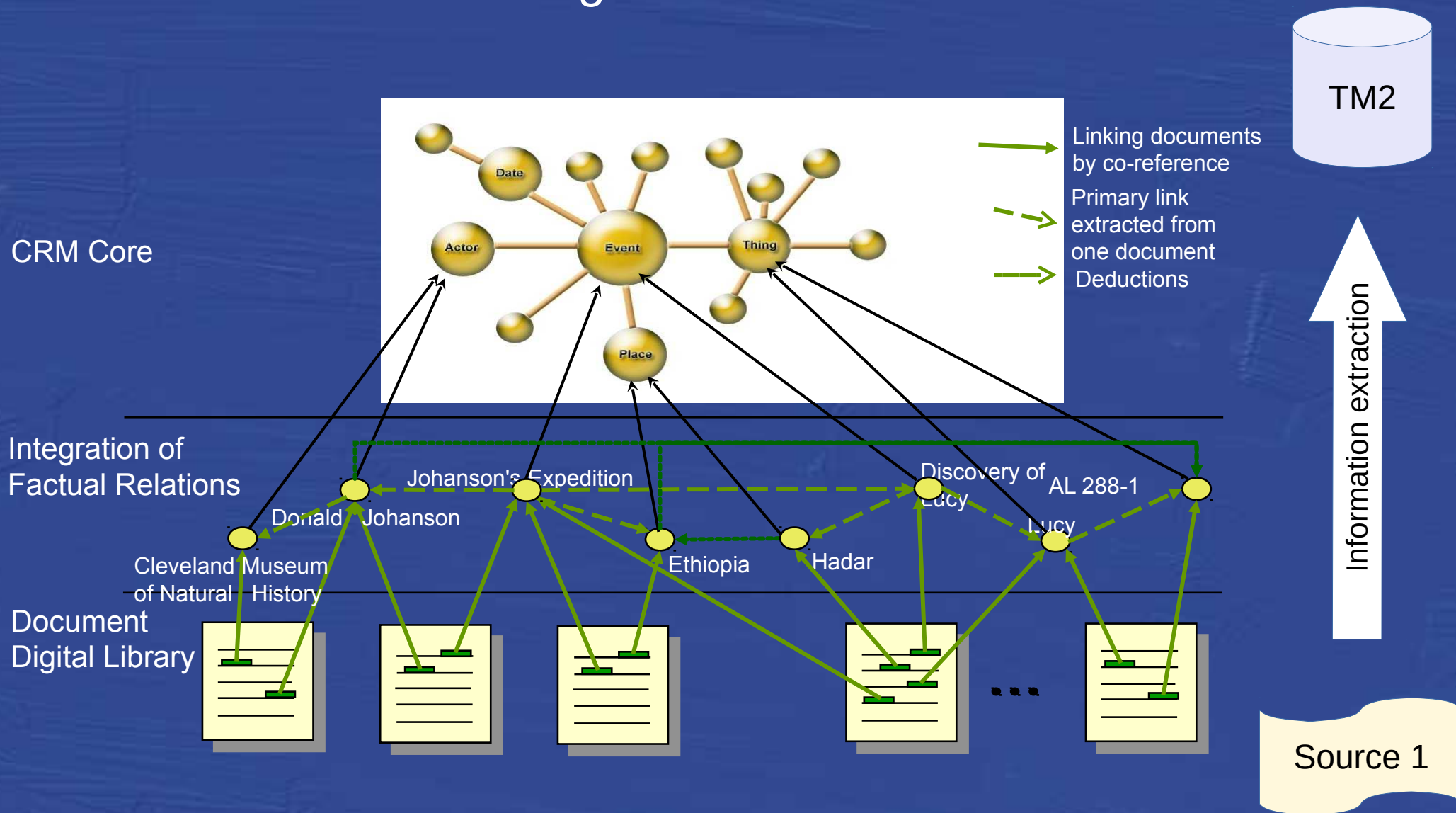
The CIDOC CRM (ISO21127:2006)

A semantic framework that provides *interoperability* between different sources of cultural heritage information



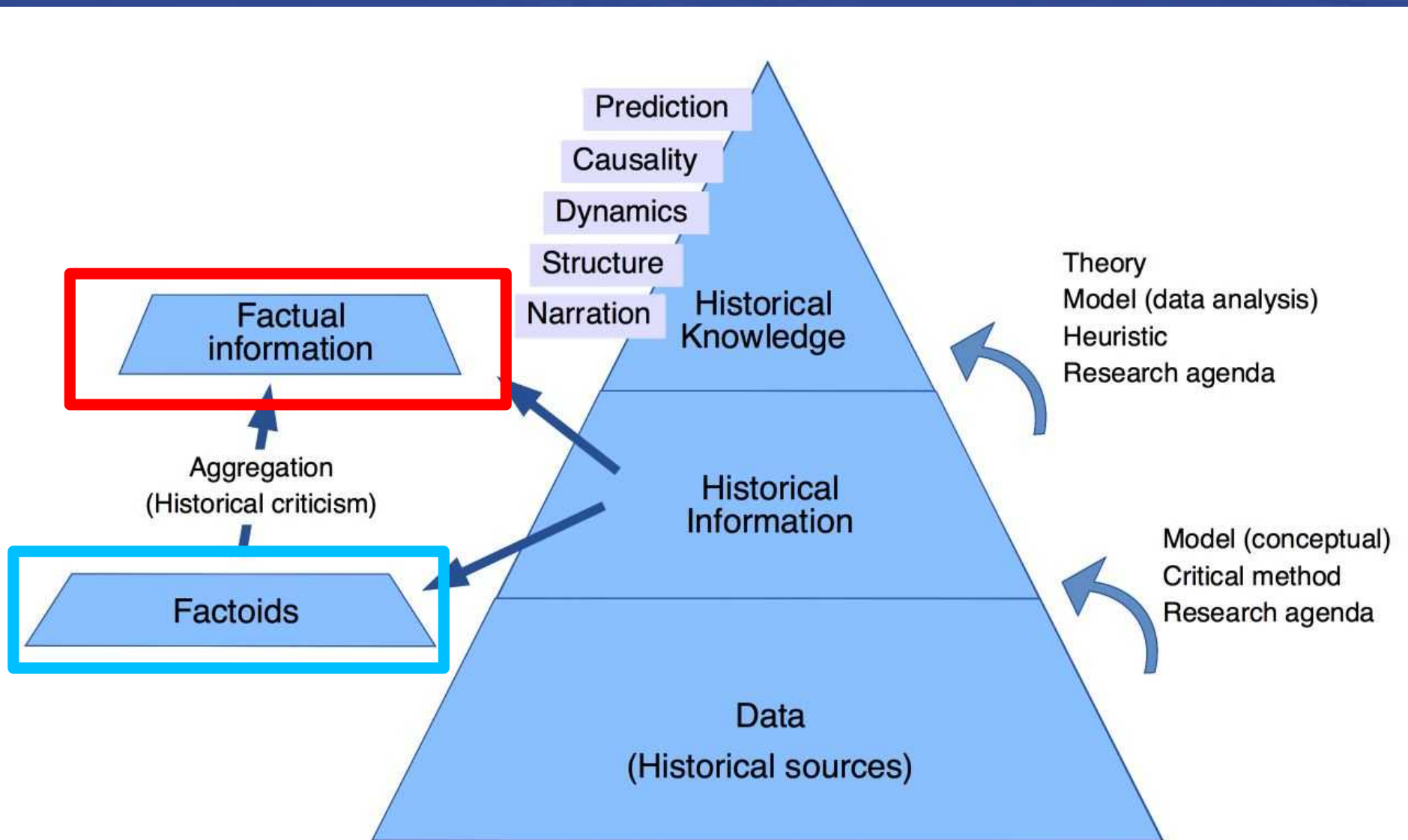
Stephen Stead (2008)

Integration of information extracted from documents using the CIDOC CRM



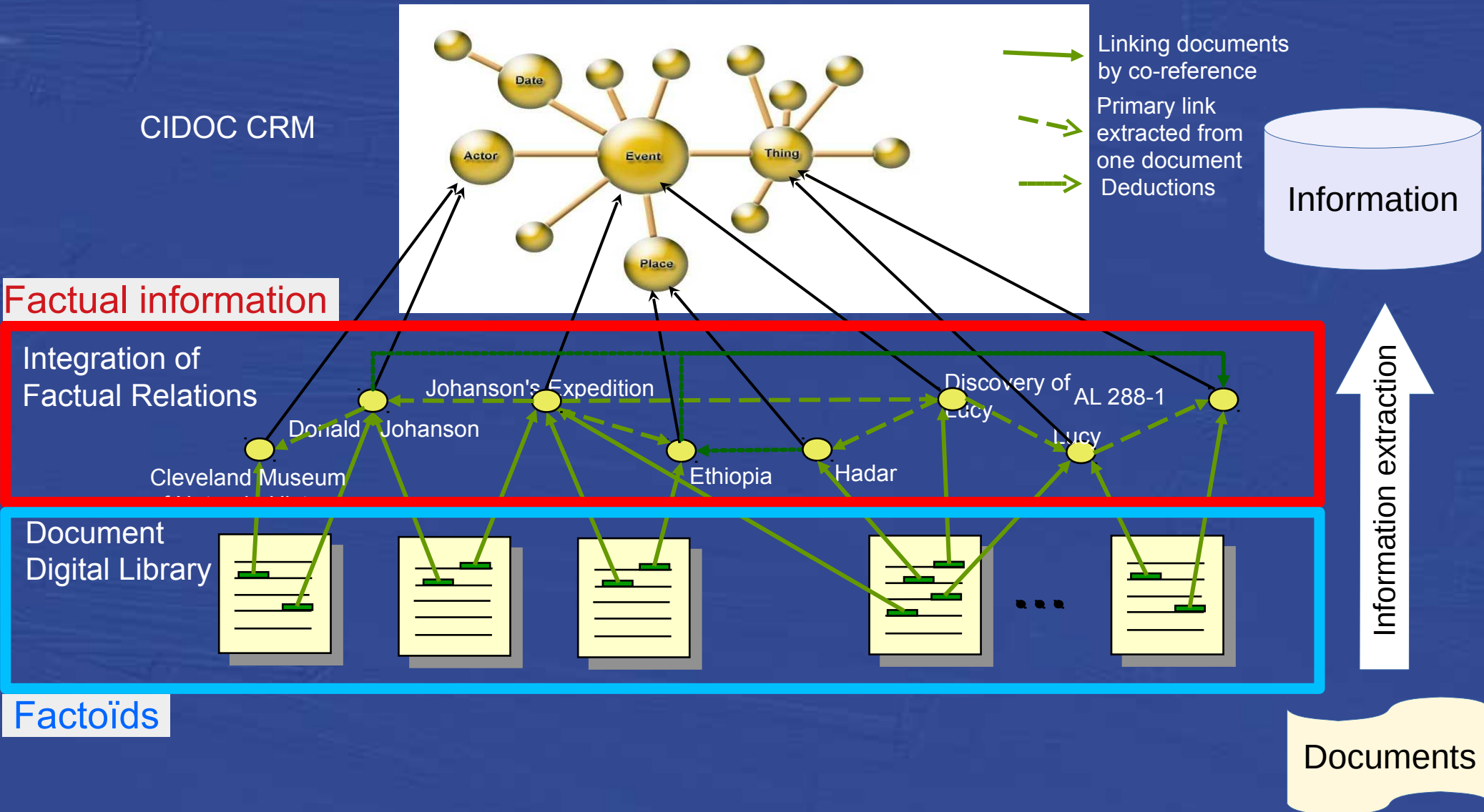
Stephen Stead (2008). Cf. Doerr M. (2003), fig. 4, p. 81

Integration of factual information



Francesco Beretta (CNRS/Université de Lyon), 7 July 2020 CC BY-NC-SA 4.0

Integration of information extracted from documents using the CIDOC CRM



« The CIDOC CRM on the other hand **does not include a general construct to model time periods** during which certain properties of an object are **static**.

The CRM encourages the description of *transitions* rather than the states or situations in between them.

This approach was inspired by **considerations of modern physics**, which considers Events as complex interactions with non-negligible temporal extent, leading into phases of stability with fuzzy (if any) temporal boundaries.

The observations necessary to acquire knowledge can themselves be regarded as Events that are part of the observed system.

It seems that an **analogous argument can be applied to historical research**. »

Doerr M., Hunter J., Lagoze C. (2003)

Person

is citizen of

Country

has social role

Role

Different domain of discourse ?

social phenomena



physical or biological phenomena

CIDOC CRM

Person

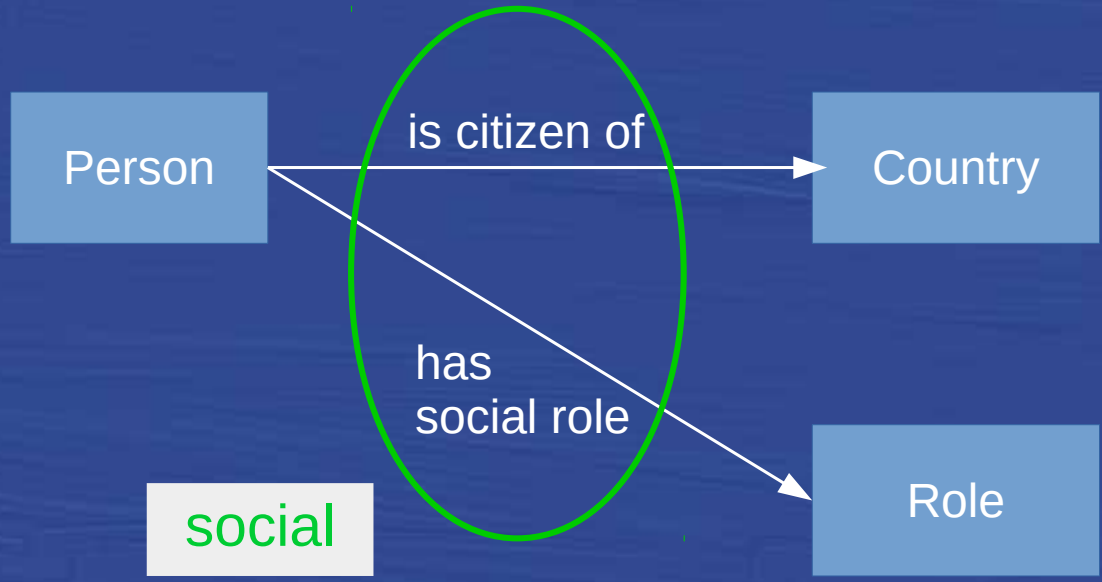
is born in

Town

Dog

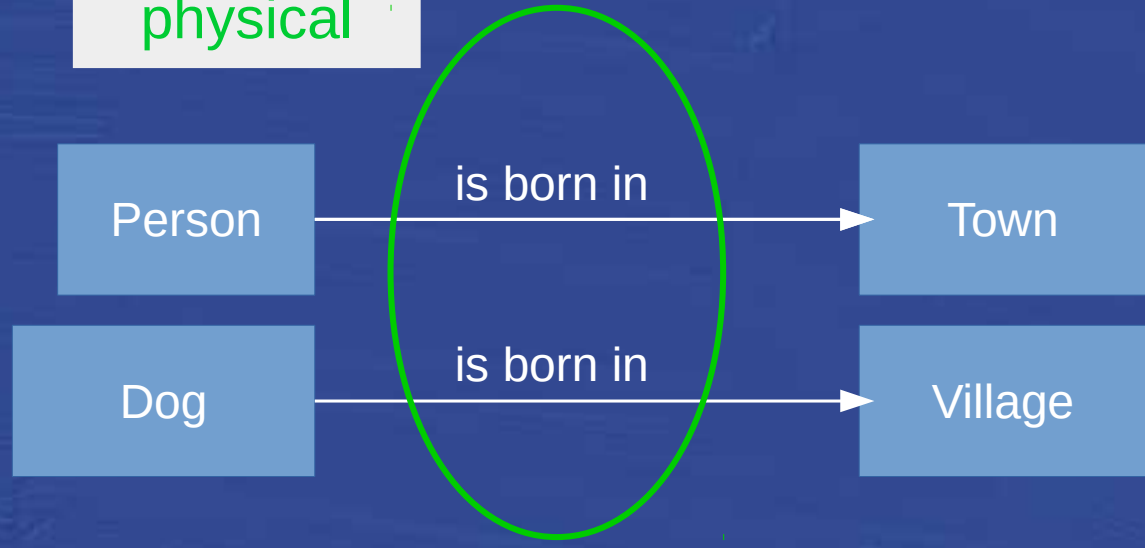
is born in

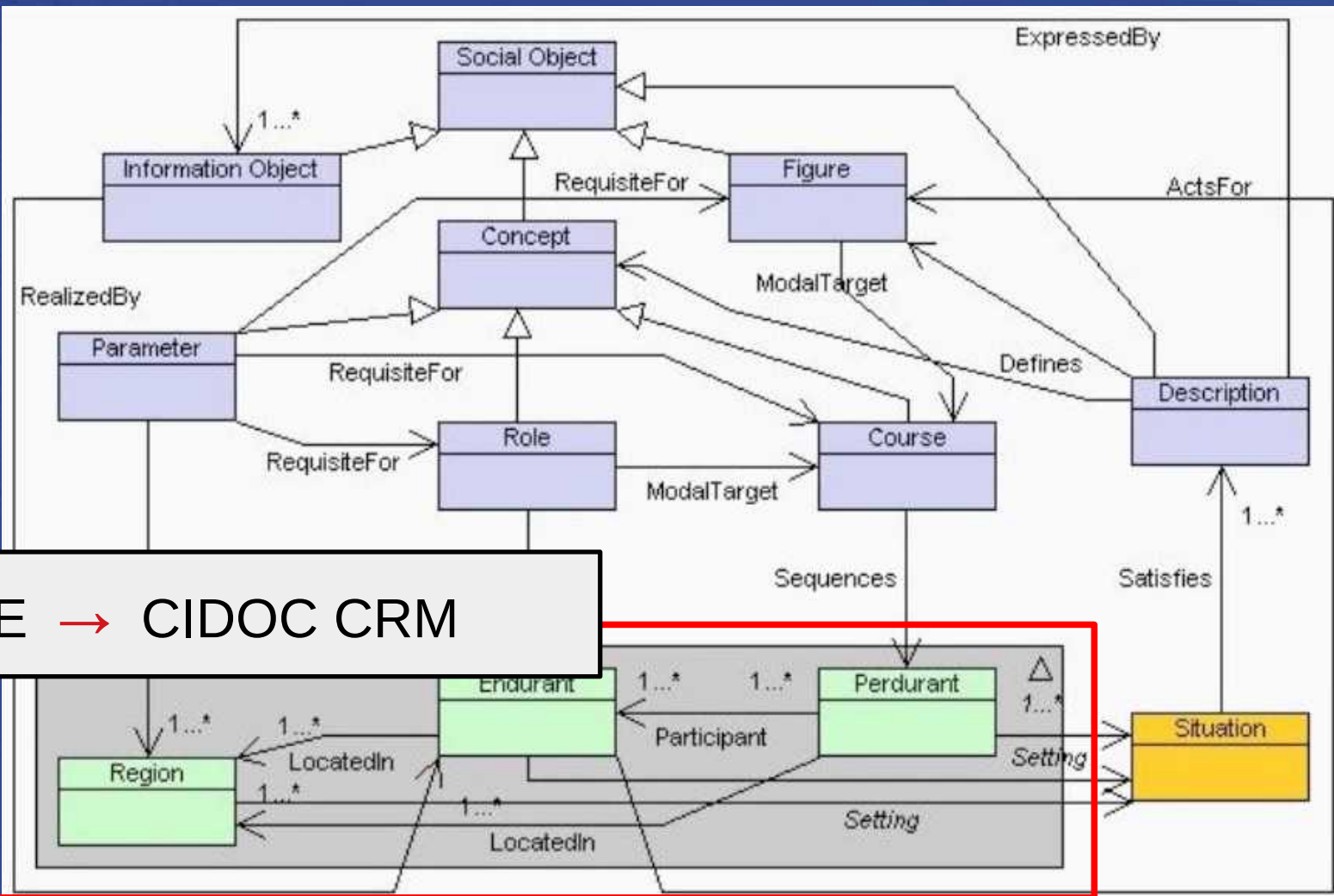
Village



DOLCE + DnS = DOLCE Lite PLUS

physical



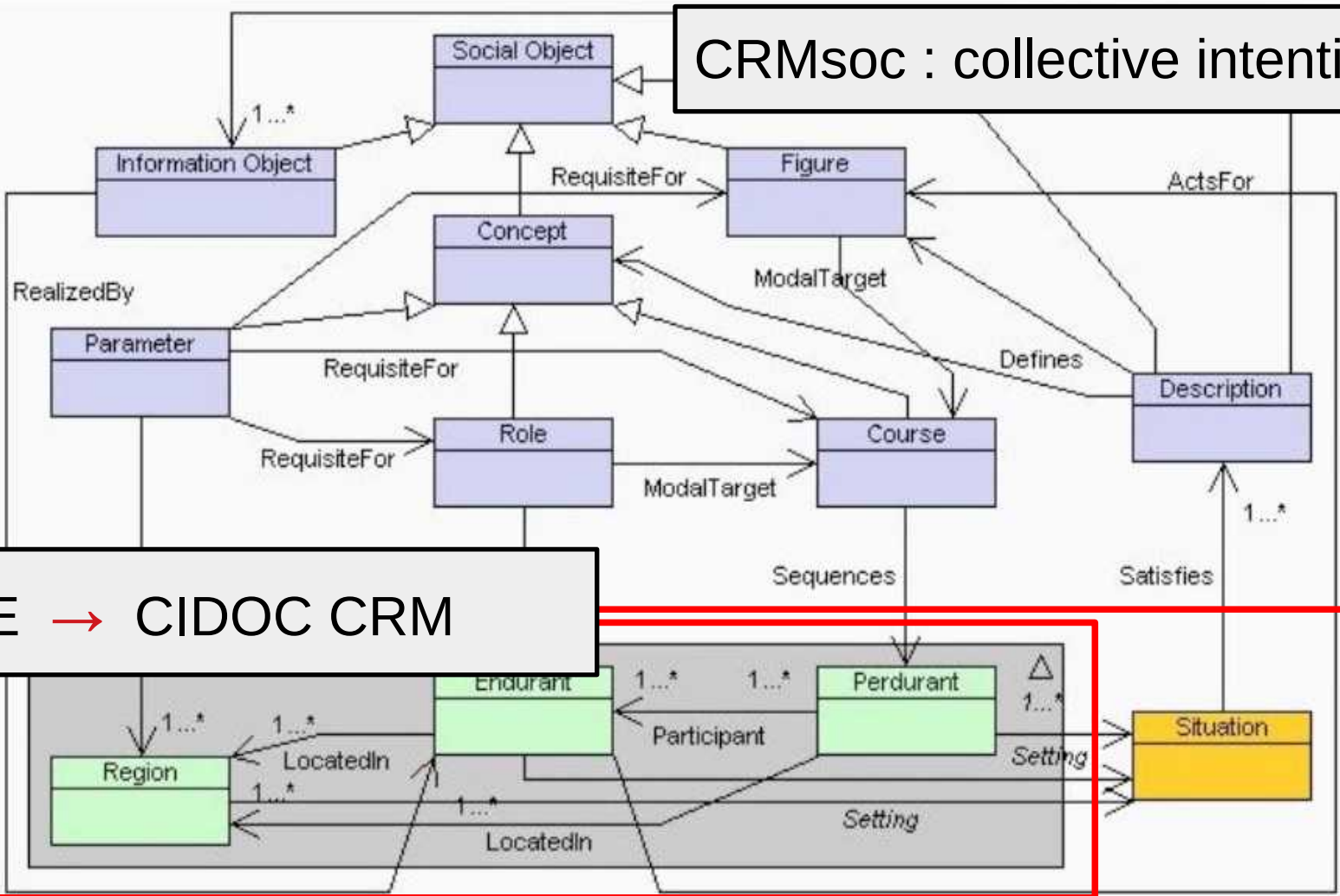


DOLCE → CIDOC CRM

Fig. 2. A UML class diagram for D&S. The lower part of the pattern (within the grey package) is called the *ground ontology*, the higher is called the *descriptive ontology*; a situation satisfies a description if the two parts match according to the axioms specified for the concepts defined by the description.

Bottazzi E., Catenacci C., Gangemi A., Lehmann J.(2006) (from pre-print, not in published version)

CRMsoc : collective intentionality



DOLCE → CIDOC CRM

Fig. 2. A UML class diagram for D&S. The lower part of the pattern (within the grey package) is called the *ground ontology*, the higher is called the *descriptive ontology*; a situation satisfies a description if the two parts match according to the axioms specified for the concepts defined by the description.

Bottazzi E., Catenacci C., Gangemi A., Lehmann J.(2006) (from pre-print, not in published version)

CRMsoc : Model for Social Phenomena

Compatible models & Collaborations

In this section you can find all compatible models and collaborations to the CIDOC CRM. Each one of them will navigate you to its corresponding web page.

FRBR_{oo}

Functional
Requirements
for Bibliographic
Records

PRESS_{oo}

Modeling of
bibliographic
information

CRM_{inf}

Argumentation
model

CRM_{archaeo}

Excavation
model

CRM_{sci}

Scientific
observation
model

CRM_{geo}

Spatiotemporal
model

CRM_{dig}

Model for
provenance
metadata

CRM_{ba}

Model for
Archaeological
Buildings

CRM_{tex}

Model
for the study
of ancient texts

CRM_{soc}

Model for
Social
Phenomena

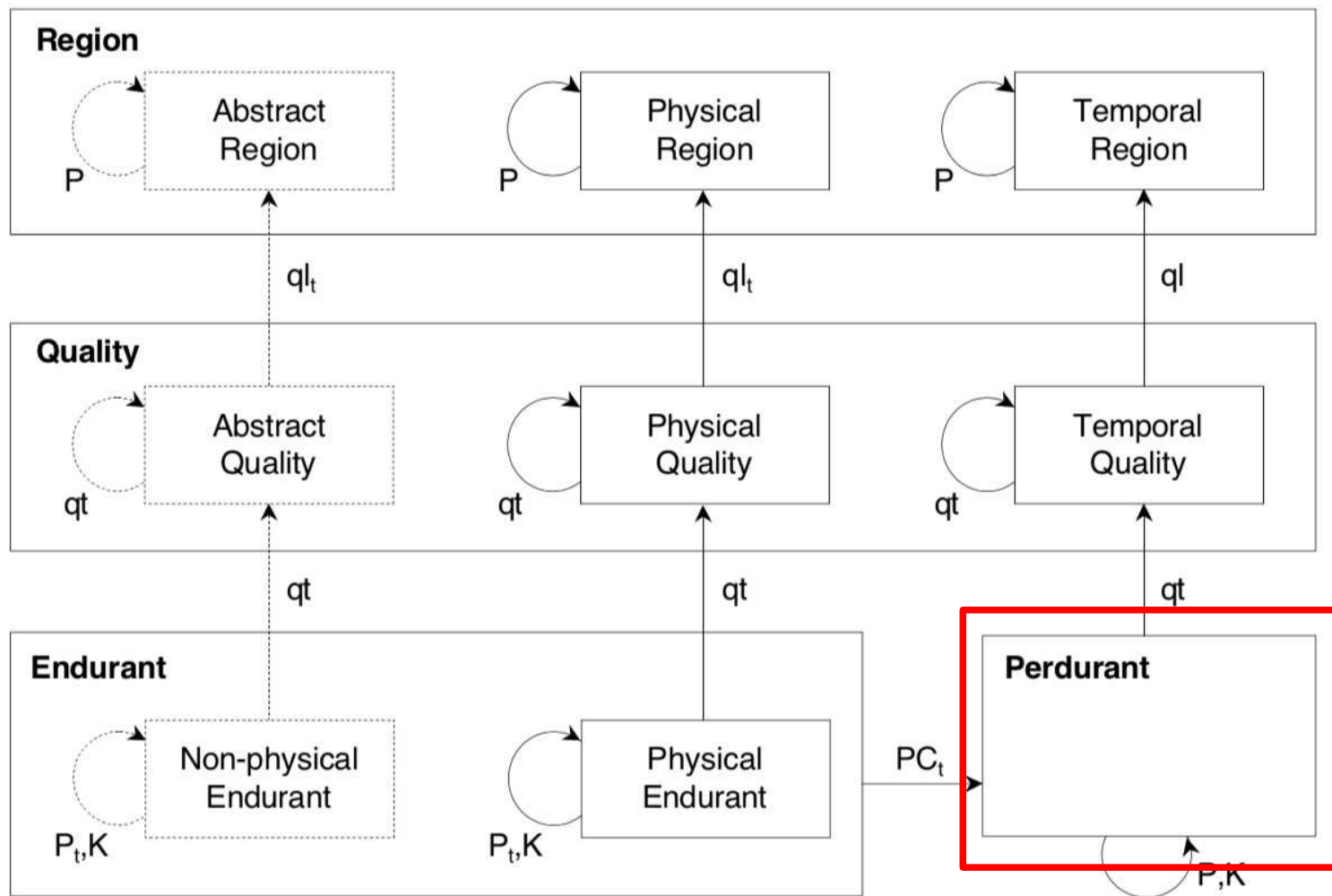
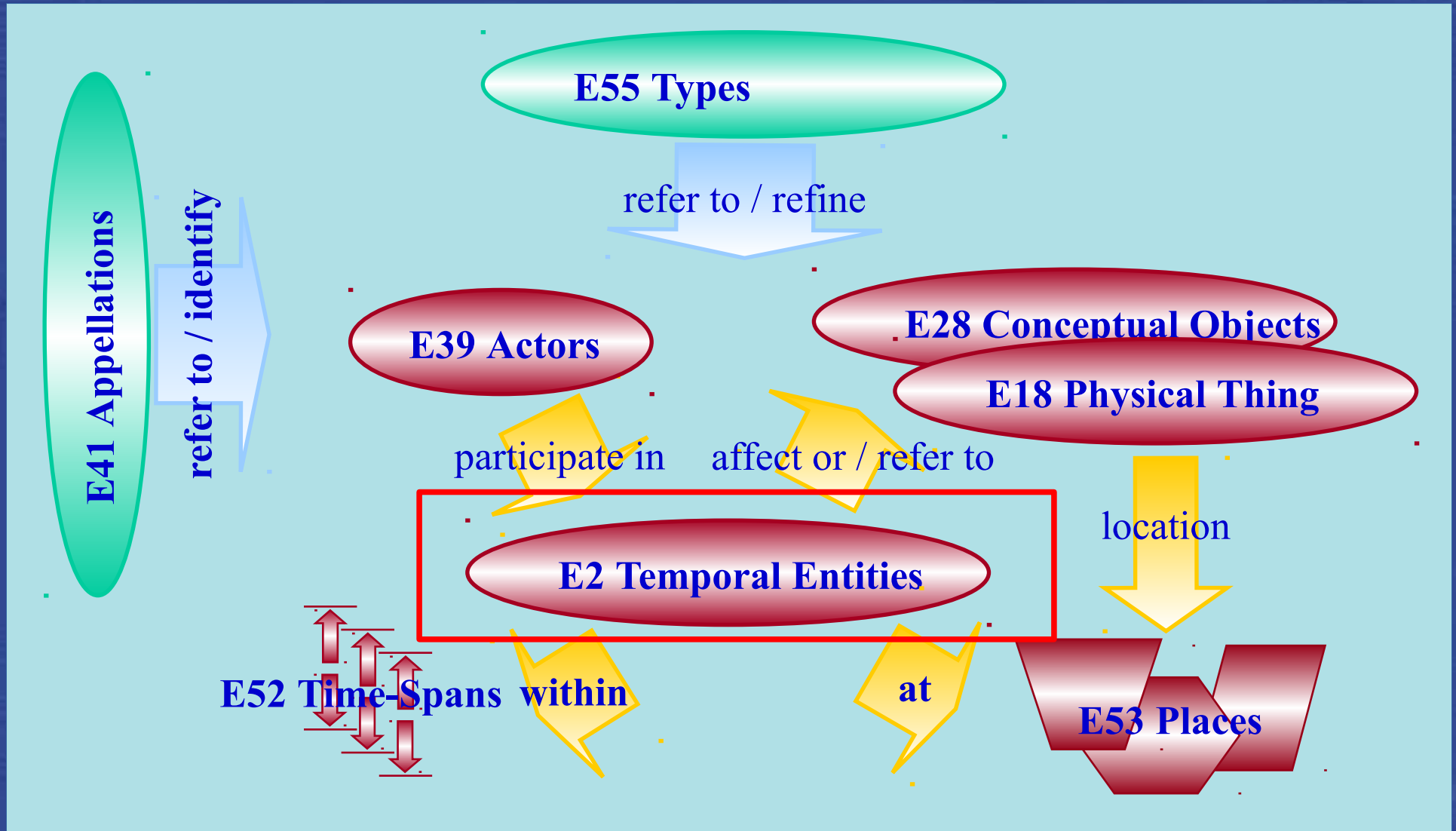


Figure 4: Primitive relations between basic categories (the dotted lines to the left indicate that we are less confident with what concerns non-physical endurants).

The CIDOC CRM (ISO21127:2006)

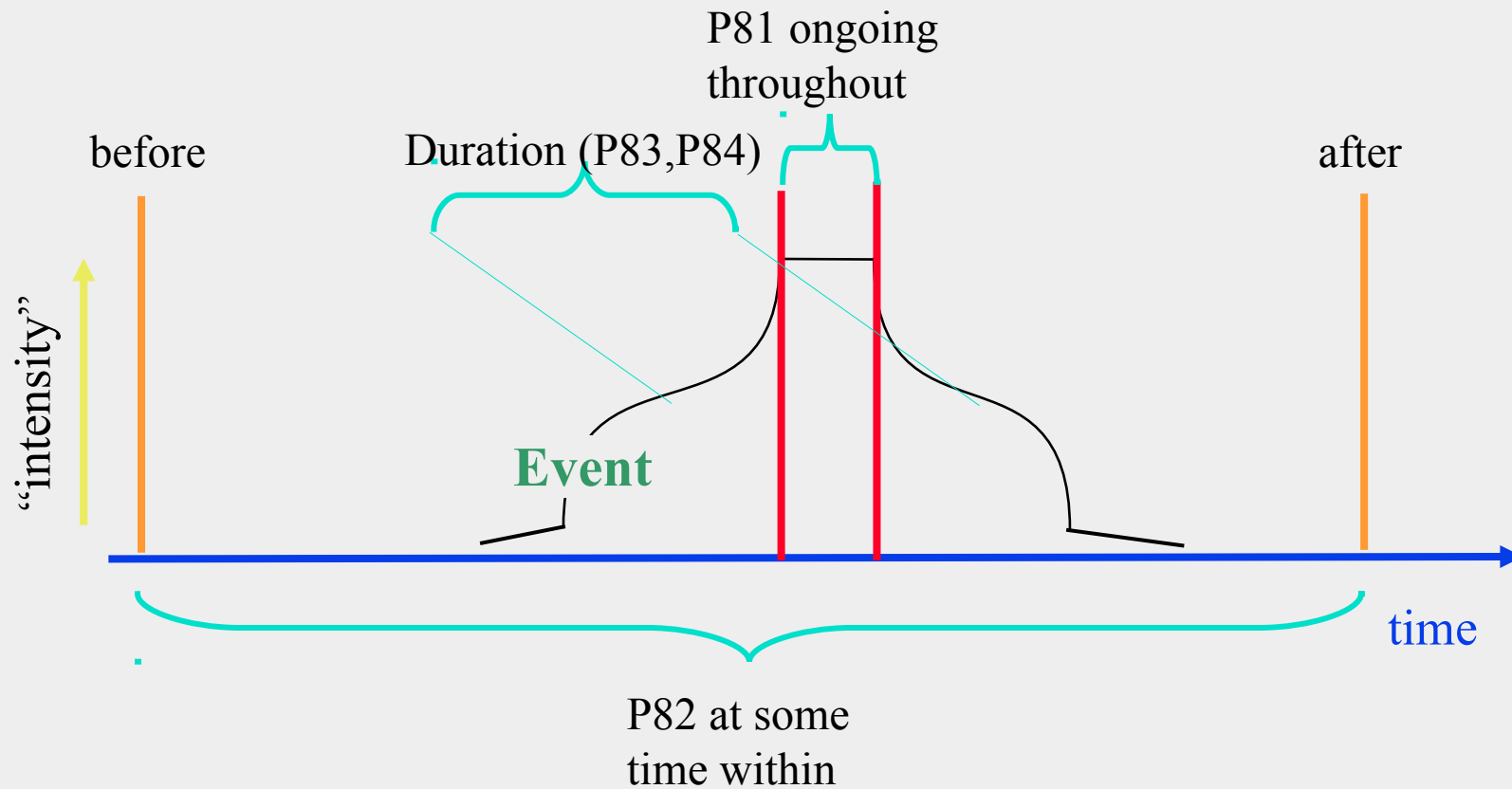
A semantic framework that provides *interoperability* between different sources of cultural heritage information



Stephen Stead (2008)

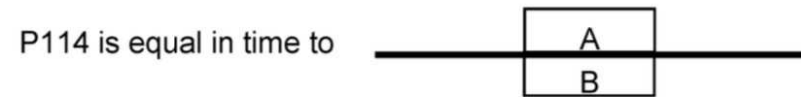
The CIDOC CRM

Time Uncertainty, Certainty and Duration

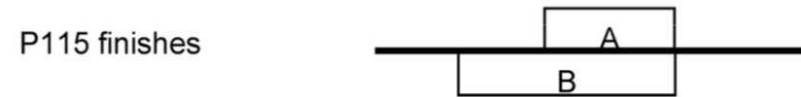


Stephen Stead (2008)

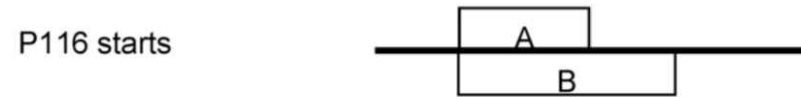
The CIDOC CRM P114-120 'Allen' properties (James F. Allen)



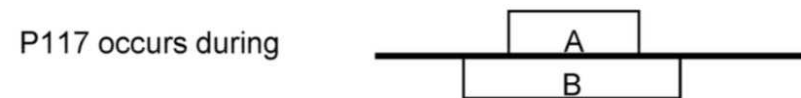
$$A_s = B_s \ \& \ A_e = B_e \Leftrightarrow (A_{ss} = B_{ss} \ \& \ A_{se} = B_{se}) \ \& \ (A_{es} = B_{es} \ \& \ A_{ee} = B_{ee})$$



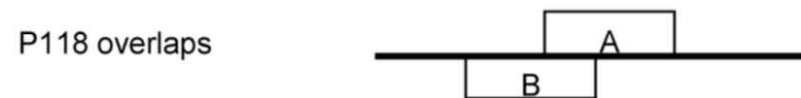
$$A_s > B_s \ \& \ A_e = B_e \Leftrightarrow (B_{ss} < A_{ss} \ \& \ B_{se} < A_{se}) \ \& \ (B_{es} = A_{es} \ \& \ B_{ee} = A_{ee})$$



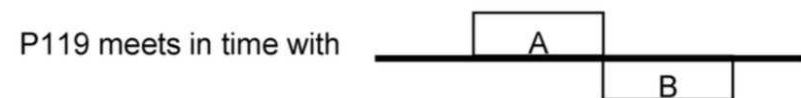
$$A_e < B_e \ \& \ A_s = B_s \Leftrightarrow (A_{ss} = B_{ss} \ \& \ A_{se} = B_{se}) \ \& \ (A_{ee} < B_{ee} \ \& \ A_{es} < B_{es})$$



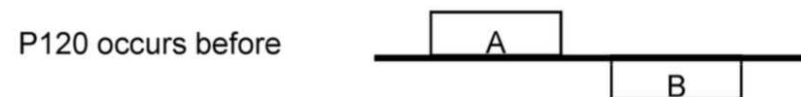
$$B_s < A_s \ \& \ A_e < B_e \Leftrightarrow (B_{ss} < A_{ss} \ \& \ B_{se} < A_{se}) \ \& \ (A_{es} < B_{es} \ \& \ A_{ee} < B_{ee})$$



$$B_s < A_s \ \& \ A_s < B_e \ \& \ B_e < A_e \Leftrightarrow (B_{ss} < A_{ss} \ \& \ B_{se} < A_{se}) \ \& \ (A_{ss} < B_{es} \ \& \ A_{se} < B_{ee}) \ \& \ (B_{es} < A_{es} \ \& \ B_{ee} < A_{ee})$$



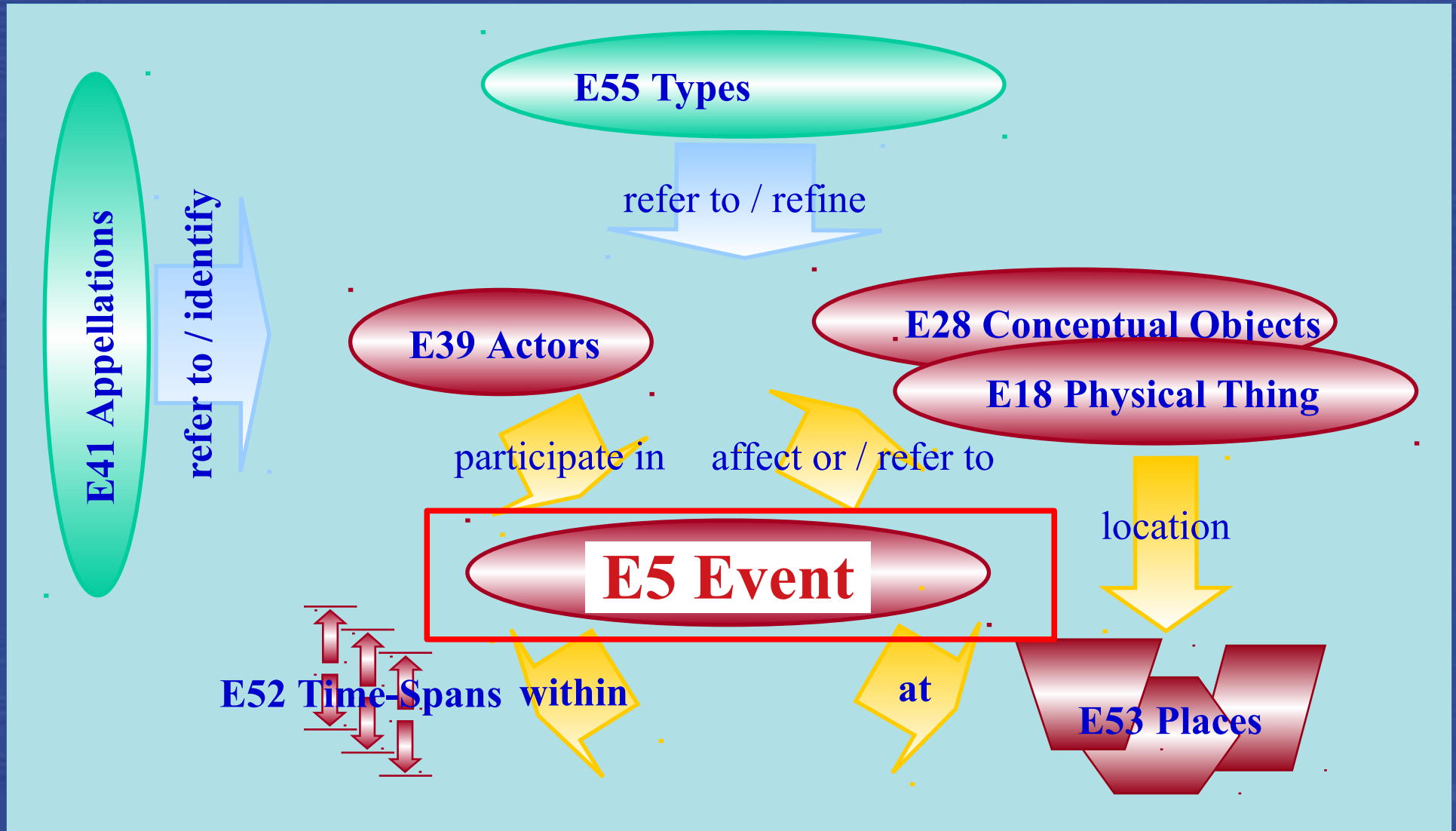
$$A_e = B_s \Leftrightarrow A_{es} = B_{ss} \ \& \ A_{ee} = B_{se}$$

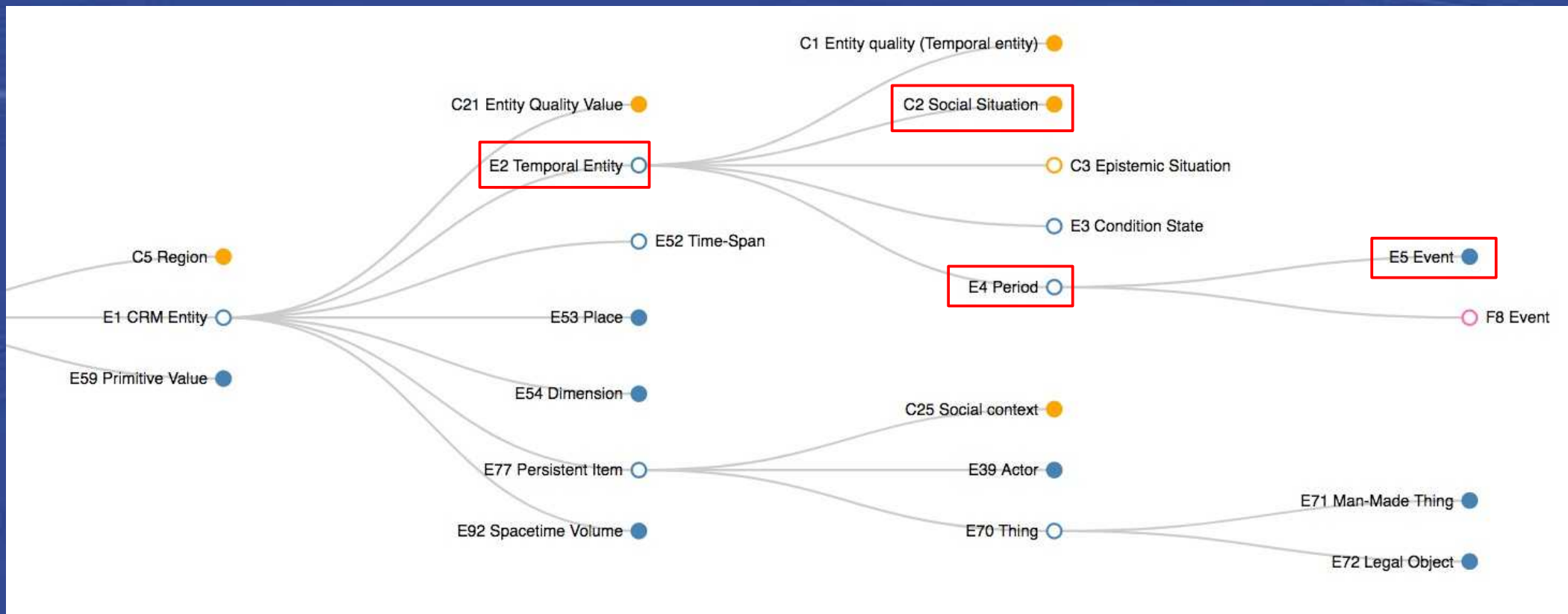


$$A_e < B_s \Leftrightarrow A_{es} < B_{ss} \ \& \ A_{ee} < B_{se}$$

J. Holmen et Ch.-E. Ore (2010)

Time, space and participants : crm:E5 Event





<http://ontome.dataforhistory.org>

HistDMI – Generic Classes and Properties – ongoing



[Summary](#)
[Identification](#)
[Definition](#)
[Hierarchy](#)
[Classes](#)
[Properties](#)
[Graph](#)
[Comments](#)

Classes

Show entries



Search:

Class	Last updated
Appellation in a language (Temporal Entity) – C11	2020-08-10
Appellation in a language type – C12	2020-08-10
Built work type – C18	2020-08-10
Built work – C17	2020-08-10
Composition – C22	2020-08-24
Description – C26	2020-10-28
Entity quality (Temporal entity) – C1	2020-08-21

<http://ontome.dataforhistory.org>

5.

The Data for History vision

and a collaborative

Ontology Management Environment (OntoME)

<http://ontome.dataforhistory.org>

CIDOC CRM

CIDOC CRM



CRM hist extension

CIDOC CRM



CRM hist extension



Projects' ontology

CIDOC CRM



CRM hist extension



Projects' ontology



Information systems

<http://ontome.dataforhistory.org>

Man-Made Object – E22

Ship – C2

C2 Ship

Subclass of: [E22 Man-Made Object](#)

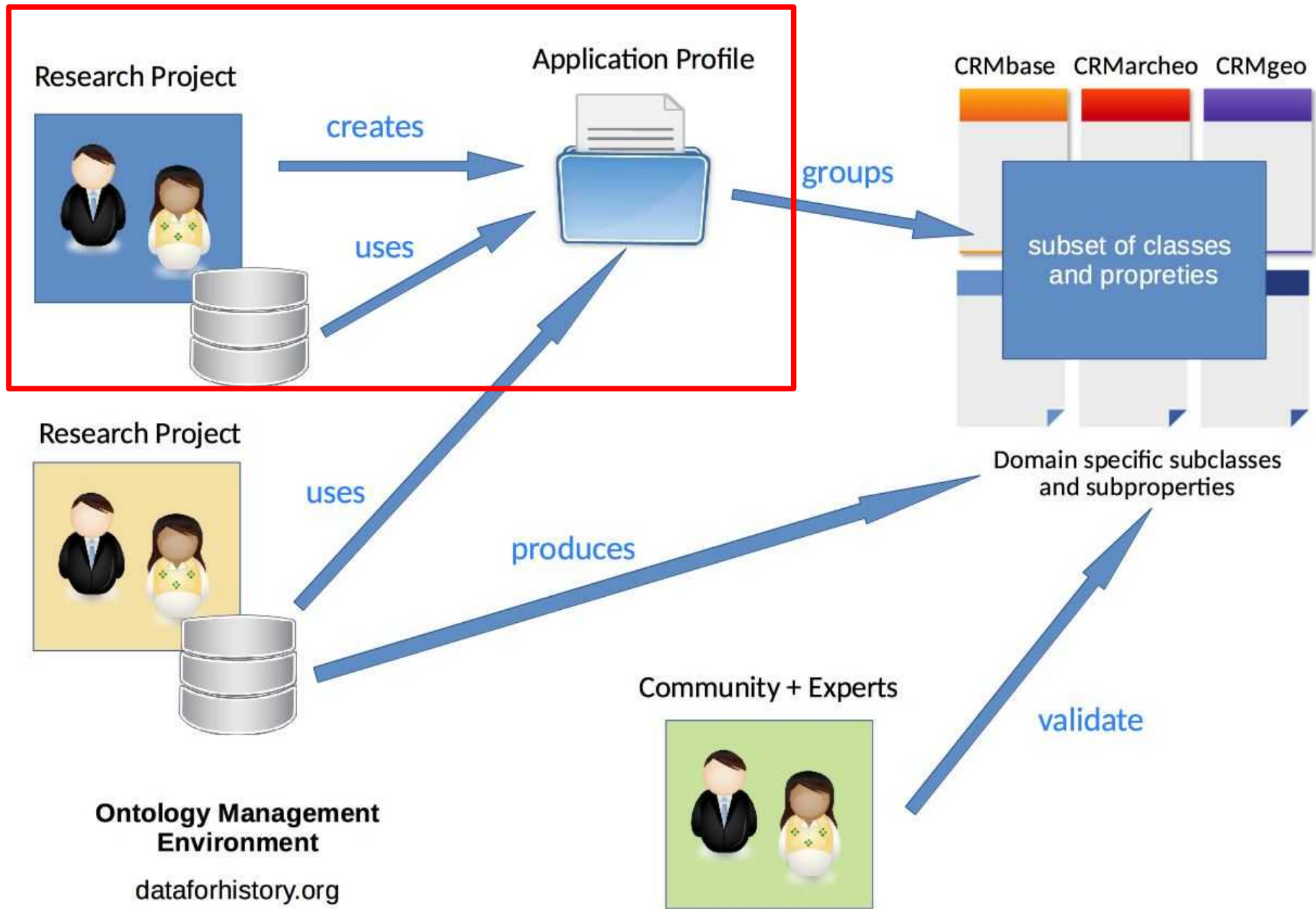
Scope note: Used to denote a watercraft that travels the world's oceans and other sufficiently deep waterways, carrying passengers or goods, or in support of specialized missions, such as defense, research and fishing.

Examples: tba

In First Order Logic: $C2(x) \supset E22(x)$

Outgoing properties: [P6 has ship type](#) → [C3 Ship type](#)

Incoming properties: [C1 Ship voyage](#) → [P3 carried out by](#)
[C12 Shipbuilding](#) → [P7 has built](#)



ontome.dataforhistory.org

Retrieve your project's application profiles from an API

<https://ontome.dataforhistory.org/api/classes-profile.json?lang=en&available-in-profile=8>

```
5:
  pk_class: 61
  identifieur_in_namespace: "E67"
  class_standard_label: "Birth"
  fk_system_type: 9
  type_label: "Temporal Entity"
  root_namespace: "CIDOC CRM"
  profile_association_type: "selected"
  pk_profile: 1
  profile_label: "BHP new data model v. 0.1"
  pk_project: 1
  project_label: "BHP - symogih.org"

6:
  pk_class: 340
  identifieur_in_namespace: "histC7"
  class_standard_label: "Human being existence"
  fk_system_type: 9
  type_label: "Temporal Entity"
  root_namespace: "Data for history"
  profile_association_type: "selected"
  pk_profile: 1
  profile_label: "BHP new data model v. 0.1"
  pk_project: 1
  project_label: "BHP - symogih.org"
```


Geovistory

geovistory.com/projects/84760/edit

Home / Your Projects / Maritime history

All ship voyages

Date: Apr 2, 1595 AD

Activated Place: Jakarta ID (Geographical Place)

Value of Place at Date: 0 (show details)

geovistory.com

CESIUM Ion Data attribution

30, 28, 26, 24, 22, 20, 18, 16, 14, 12, 10, 8, 6, 4, 2, 0

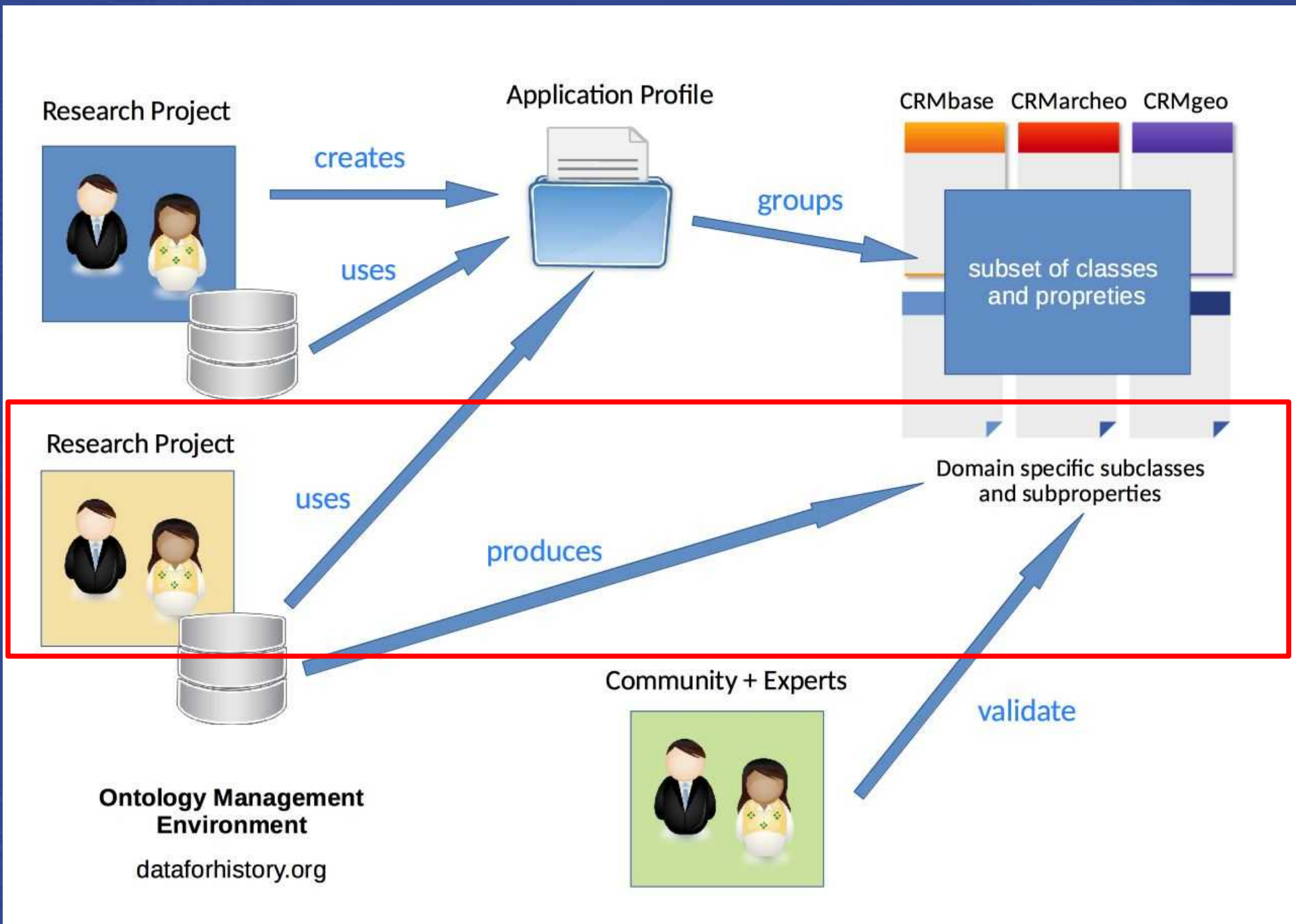
1640, 1650, 1660, 1670, 1680, 1690, 1700, 1710, 1720, 1730, 1740, 1750

Settings

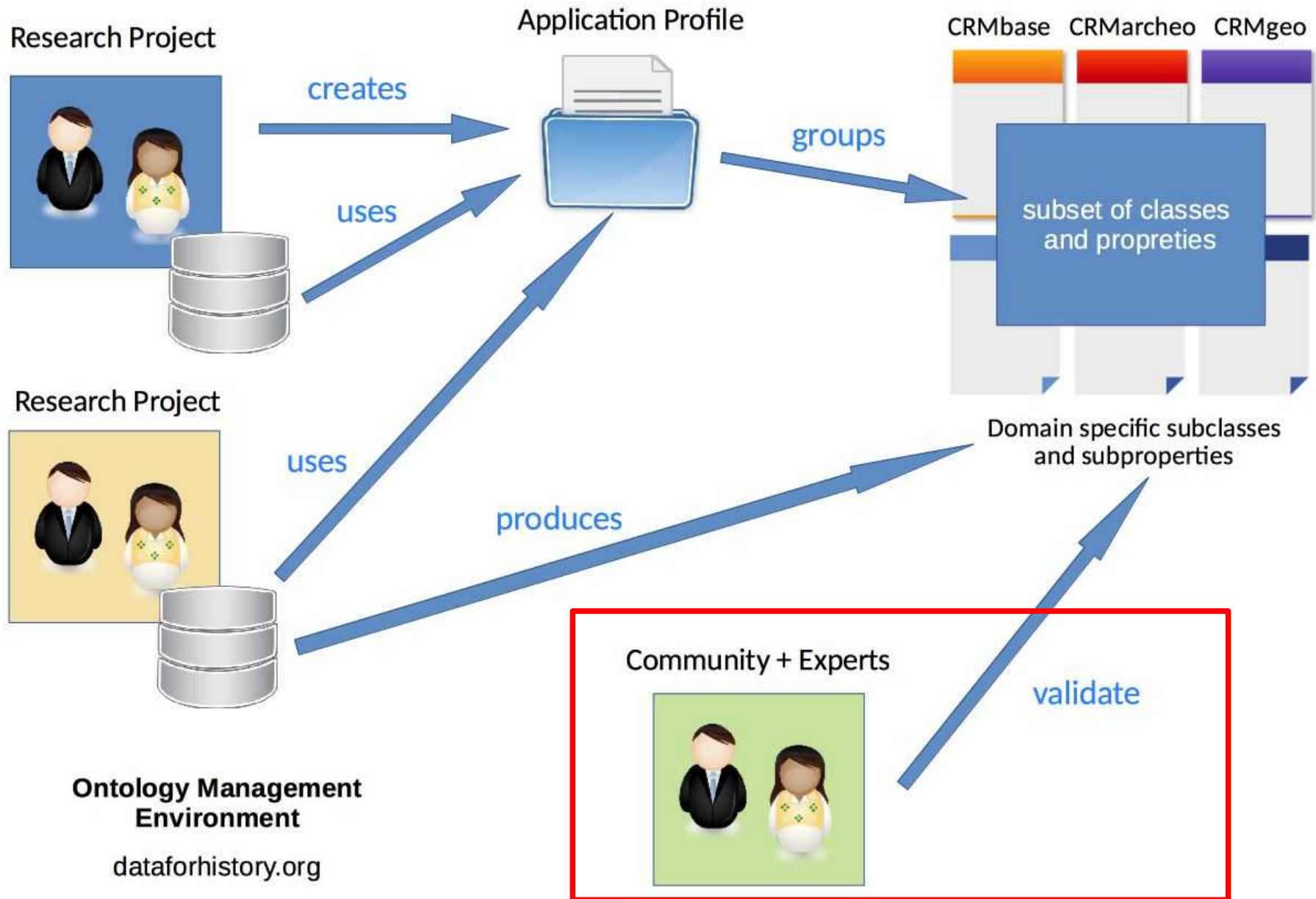
Active: Jakarta ID

feedback

Geovistory : a new VRE for the symogih.org project
developed by *kleiolab.ch* (Basel)

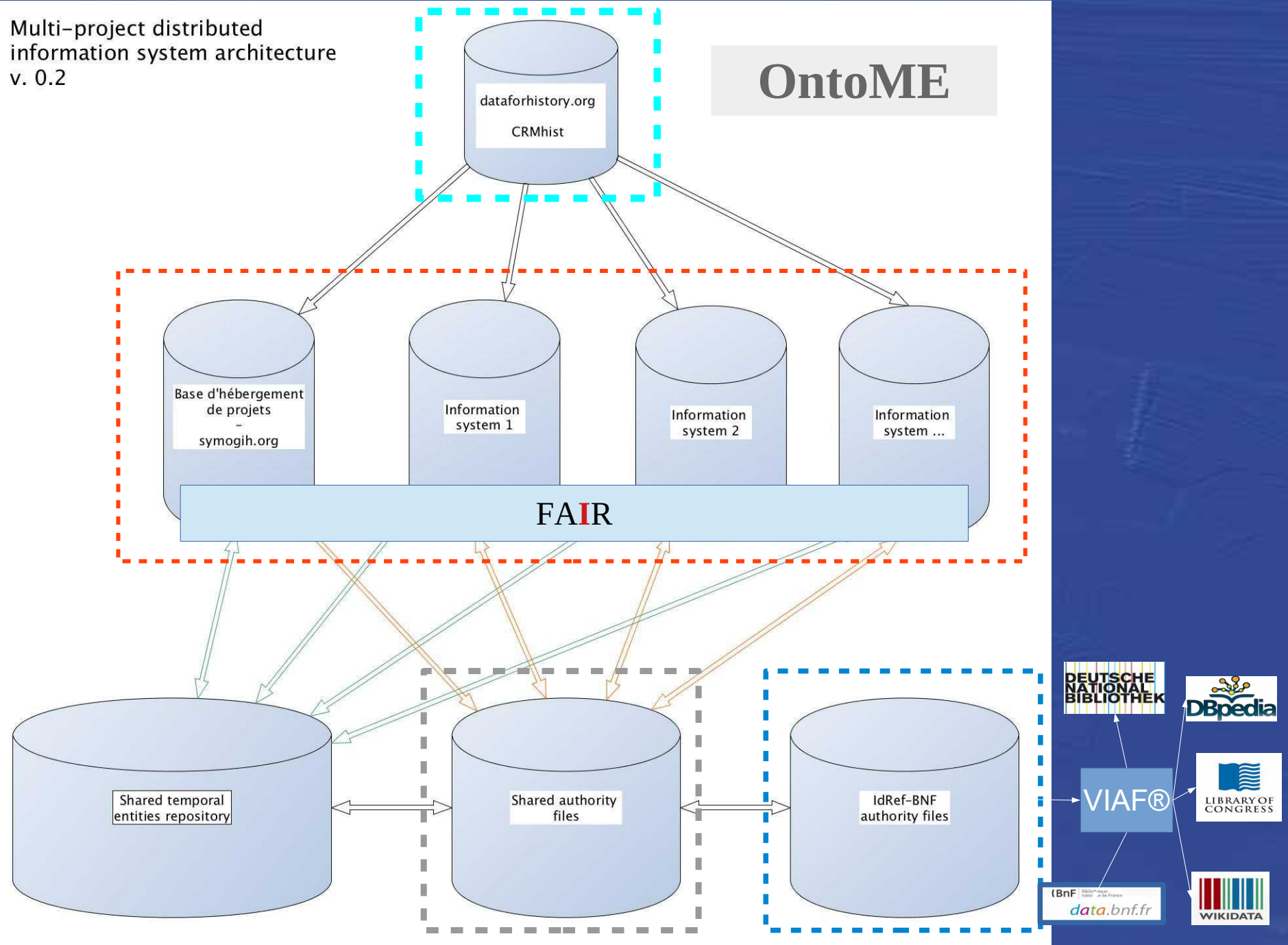


ontome.dataforhistory.org



ontome.dataforhistory.org

Multi-project distributed
information system architecture
v. 0.2



dataforhistory.org