Francesco Beretta
(CNRS UMR5190 LARHRA / Université de Lyon)

Making Research Data FAIR:

Building a CIDOC CRM Extension for Humanities and Social Sciences (SDHSS)

–

ontome.net

OAI12
The Geneva Workshop
on Innovations in Scholarly Communication

7 September 2021

LARHRA UMR 5190 LABORATOIRE DE RECHERCHE HISTORIQUE RHÔNE-ALPES

Accueil | Membres » | Recherche » | Manifestations » | Publications » | Ressources » | Valorisation »

LE LABORATOIRE

Le Laboratoire de Recherche Historique Rhône-Alpes est spécialisé en histoire moderne et contemporaine. Lire la suite

L'AGENDA DU MOIS

Toutes les manifestations à venir : colloques, séminaires, ateliers, journées d'étude, soutenances de thèses, etc... Voir l'agenda

SEMINAIRES 2020-2021

Les équipes du LARHRA organisent régulièrement des séminaires sur des sujets d'études très variés. Lire la suite

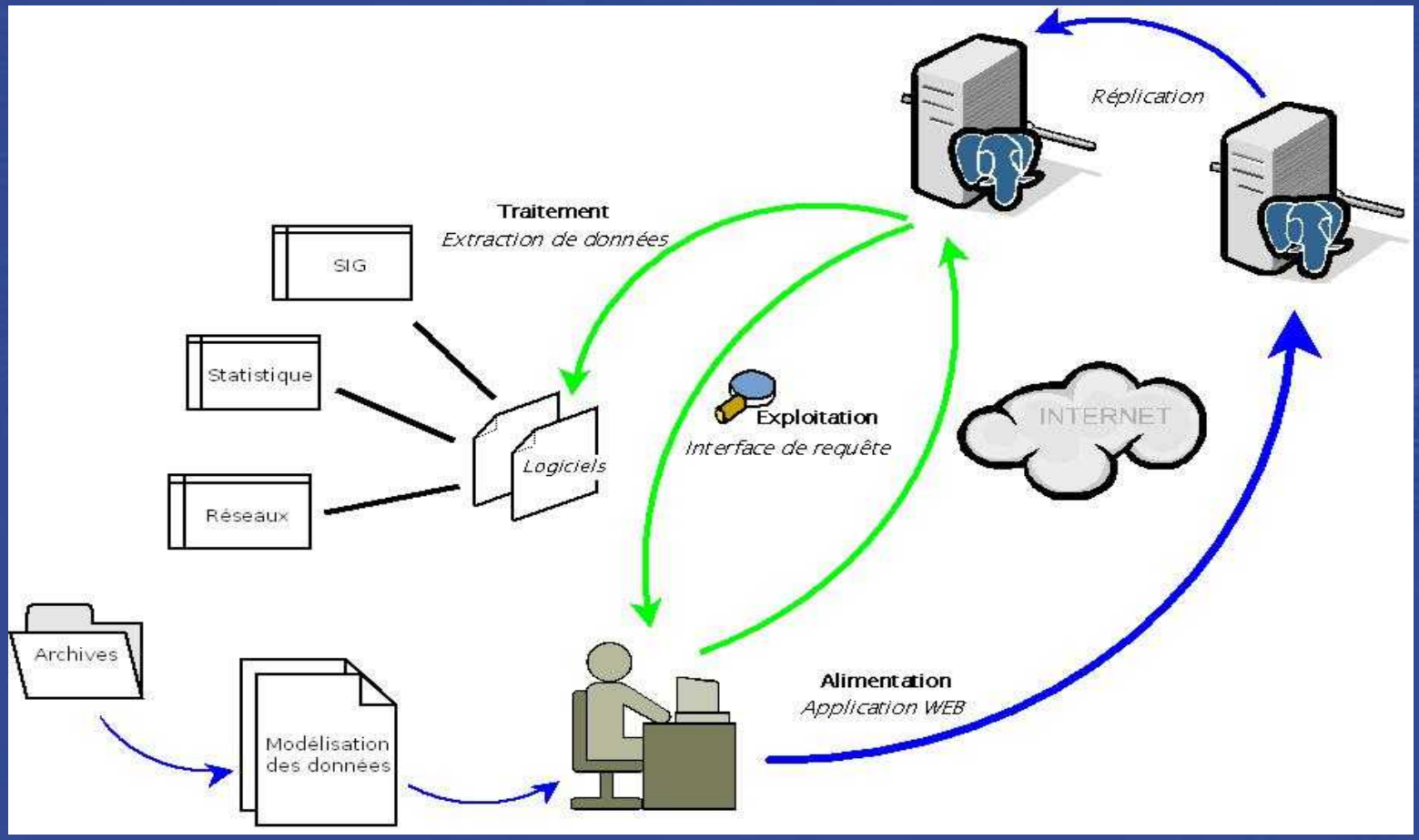
Francesco Beretta (CNRS/Université de Lyon), 7 September 2021 CC BY-NC-SA 4.0

1.

The *symogih.org* project :

a collaborative **Virtual Research Environment**
in the context of the Semantic Web
and the FAIR principles

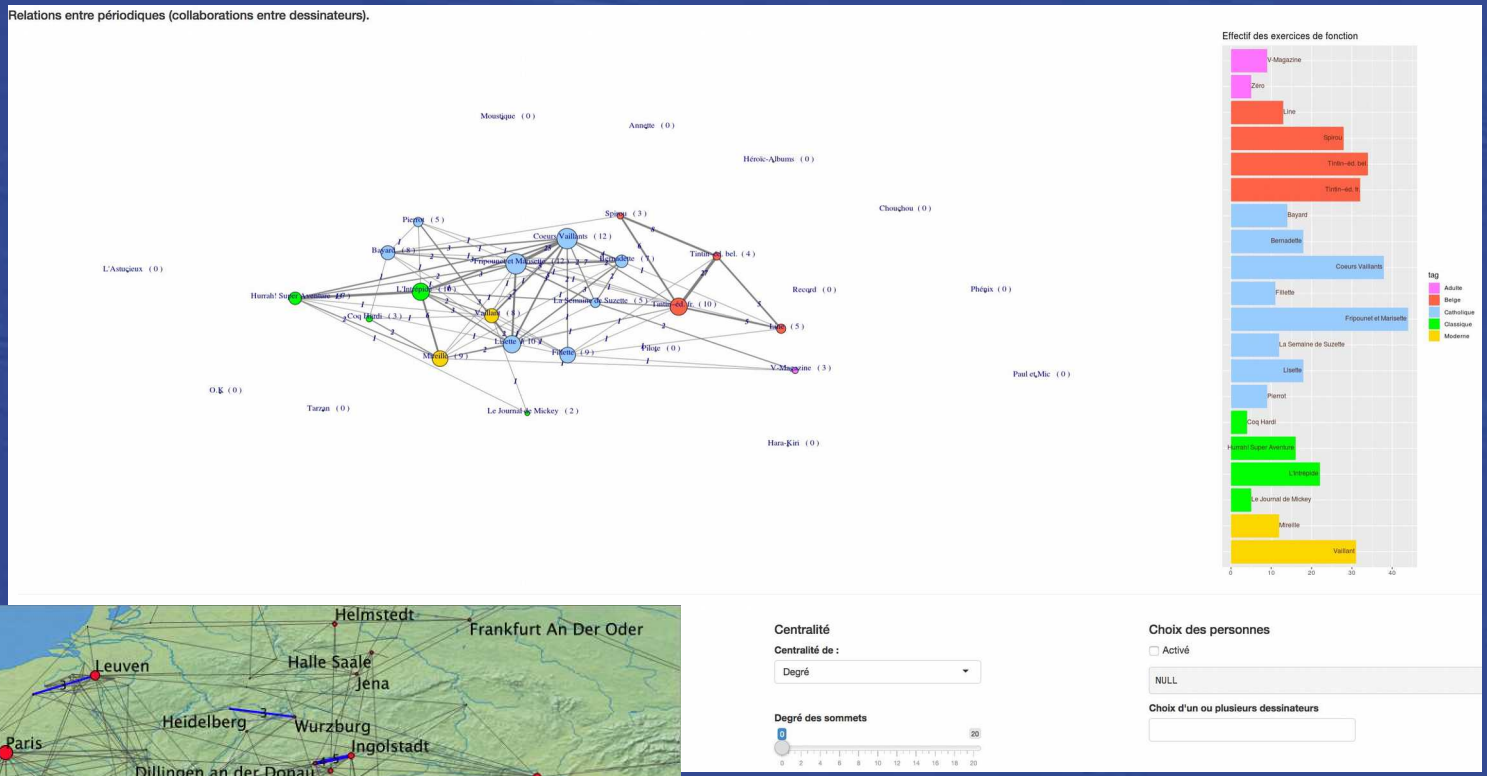
The *symogih.org* virtual research environment

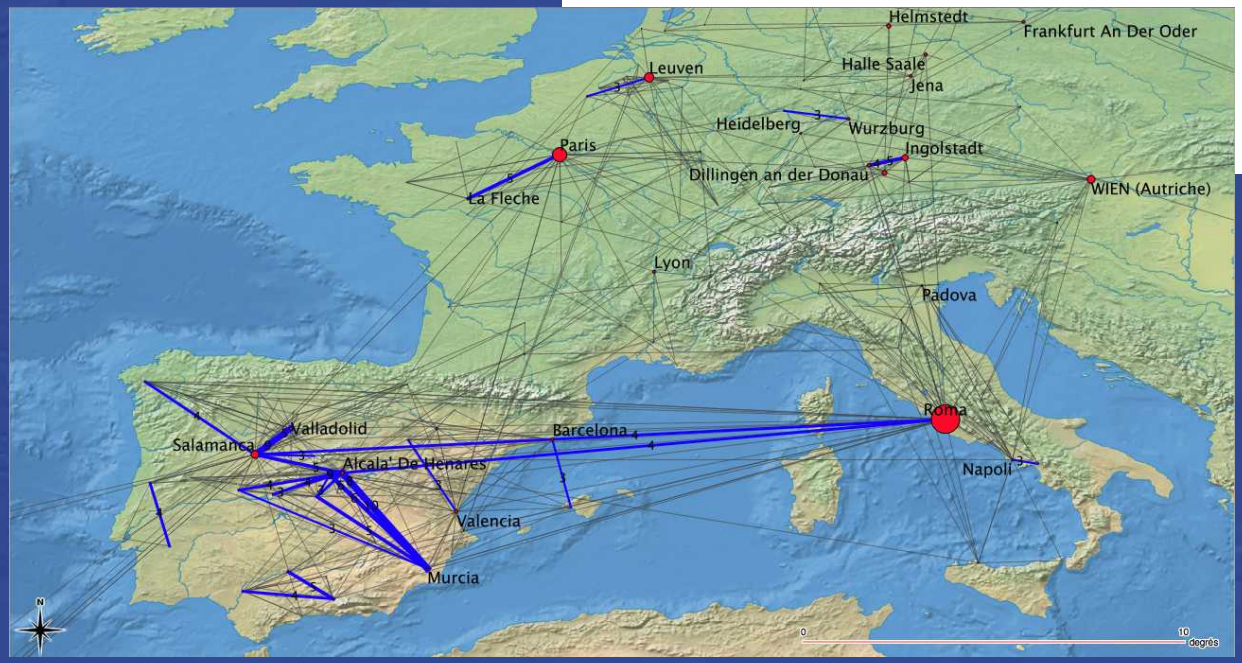


The « Système modulaire de gestion de l'information historique », symogih.org project was started in 2008.

About 50 scholars and students, and 15 national and international research projects used or are currently using the collaborative database to store and share historical information

Interactive and collaborative data analysis
Huma-Num RStudio-Shiny :https://tinyurl.com/phn-shiny





Local, individual analysis
QGIS and network analysis

symogih.org project website : documentation of the common model and published (open) data

http://symogih.org

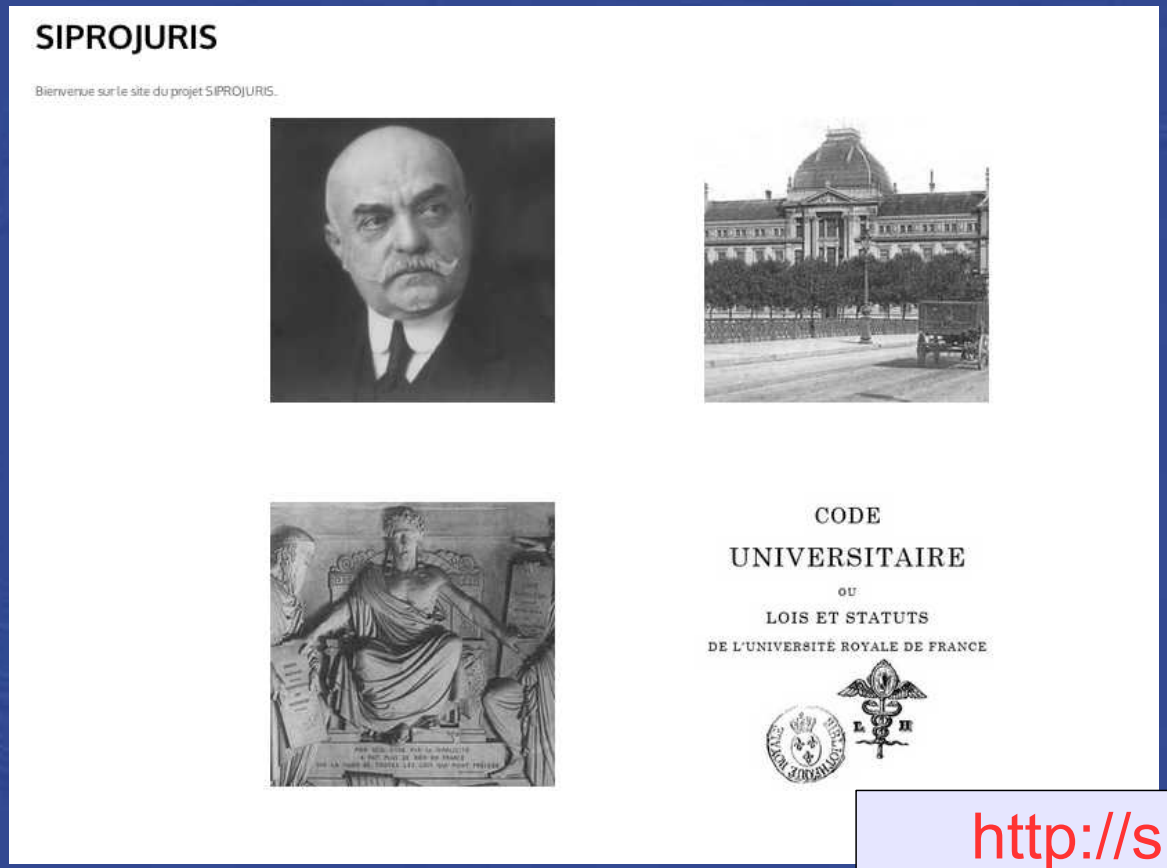
Project specific websites : Patrons de France

http://patronsdefrance.fr/

SPARQL Endpoint

Défi données MaDICS-ADOC 2018 : tinyurl.com/data-challenge-2018

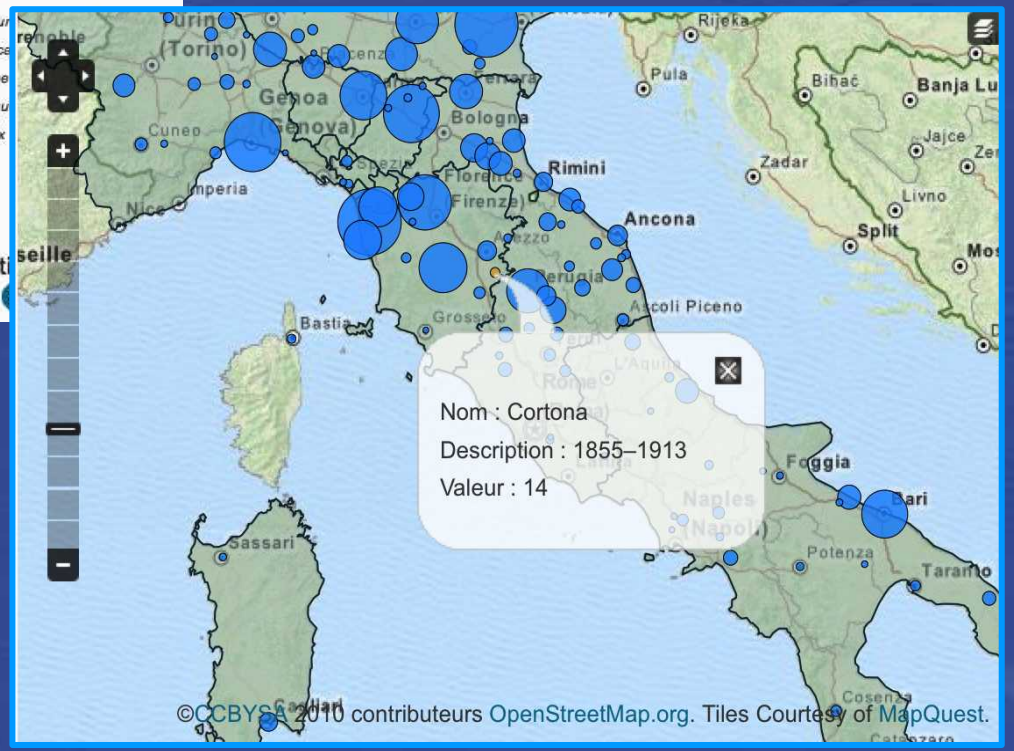
Enrichir et exploiter un corpus de données historiques
publiées sous forme de LOD.
Le projet *SIPROJURIS*.
Système d'information des professeurs de droit (1804-1950)



http://siprojuris.symogih.org

Linked Open Research Data :

use the semantic web for historical research



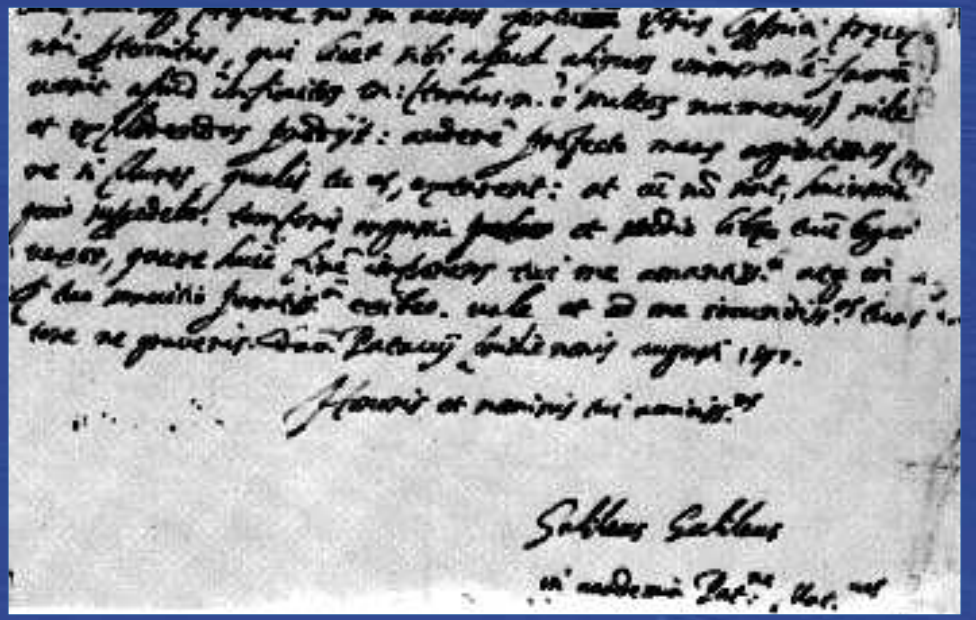
SPARQL – endpoint B3Kat

Bayerische Staatsbibliothek, Bibliotheksverbund Bayern, Kooperative Bibliotheksverbund Berlin-Brandenburg

https://lod.b3kat.de/doc/sparql-endpoint

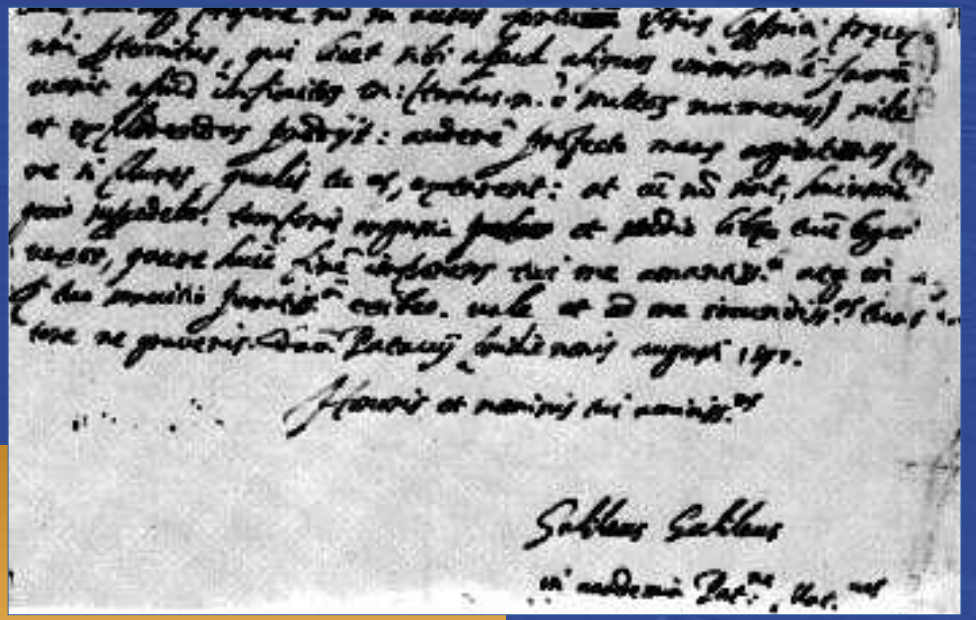
Galileo Galilei, Letter to Johannes Kepler, Padua, 4. August 1597

(Image source: Wikimedia Commons, Public Domain)



Galileo Galilei, Letter to Johannes Kepler, Padua, 4. August 1597

(Image source: Wikimedia Commons, Public Domain)



SYMOGIH.ORG

Références

Accueil | Actualités | Documentation | Membres

[info_label]

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 - Lettre adressée à Kepler, Johannes

Info3366
Type d'information: Lettre - TyIn1
Date: 1597-08-04

Composantes de l'information

Rôles | Textes | Sources

Libellé de l'objet	Type de rôle	Clé du rôle
Galilei, Galileo	auteur (être l')	InRo8646
Lettre à Giovanni Kepler, 04/08/1597	création (être une)	InRo8647
Kepler, Johannes	destination (être la)	InRo9279
Padova	localiser	InRo55897
Graz	destination (être la)	InRo55898

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

Factoïds : mentions of states of affairs in the sources

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.

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

Factual information: states of affairs as such

2.

The issue of *research data interoperability*

in historical sciences and, more in general,

in the humanities and social sciences

Findable

Accessible

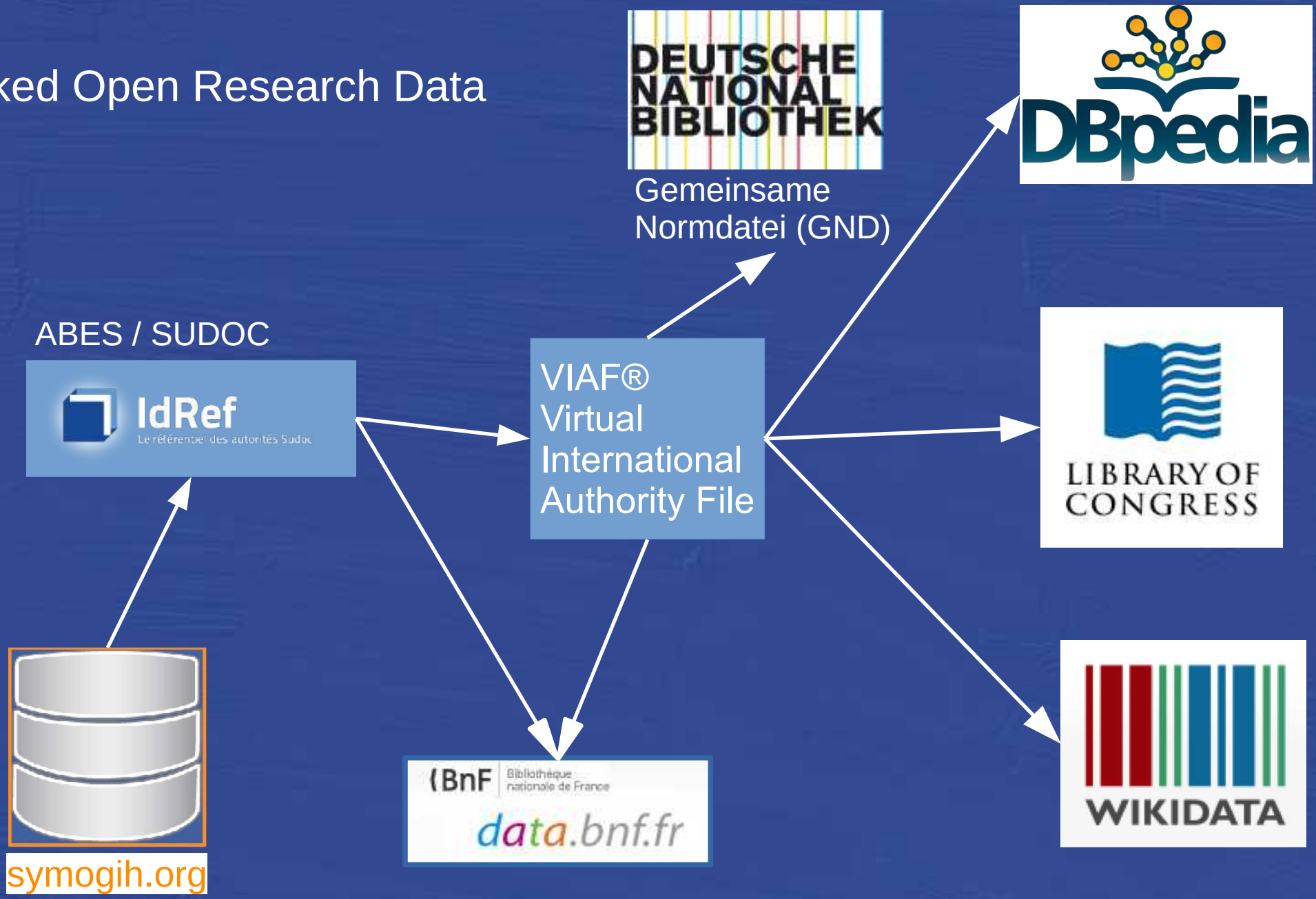
Interoperable

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.

Linked Open Research Data

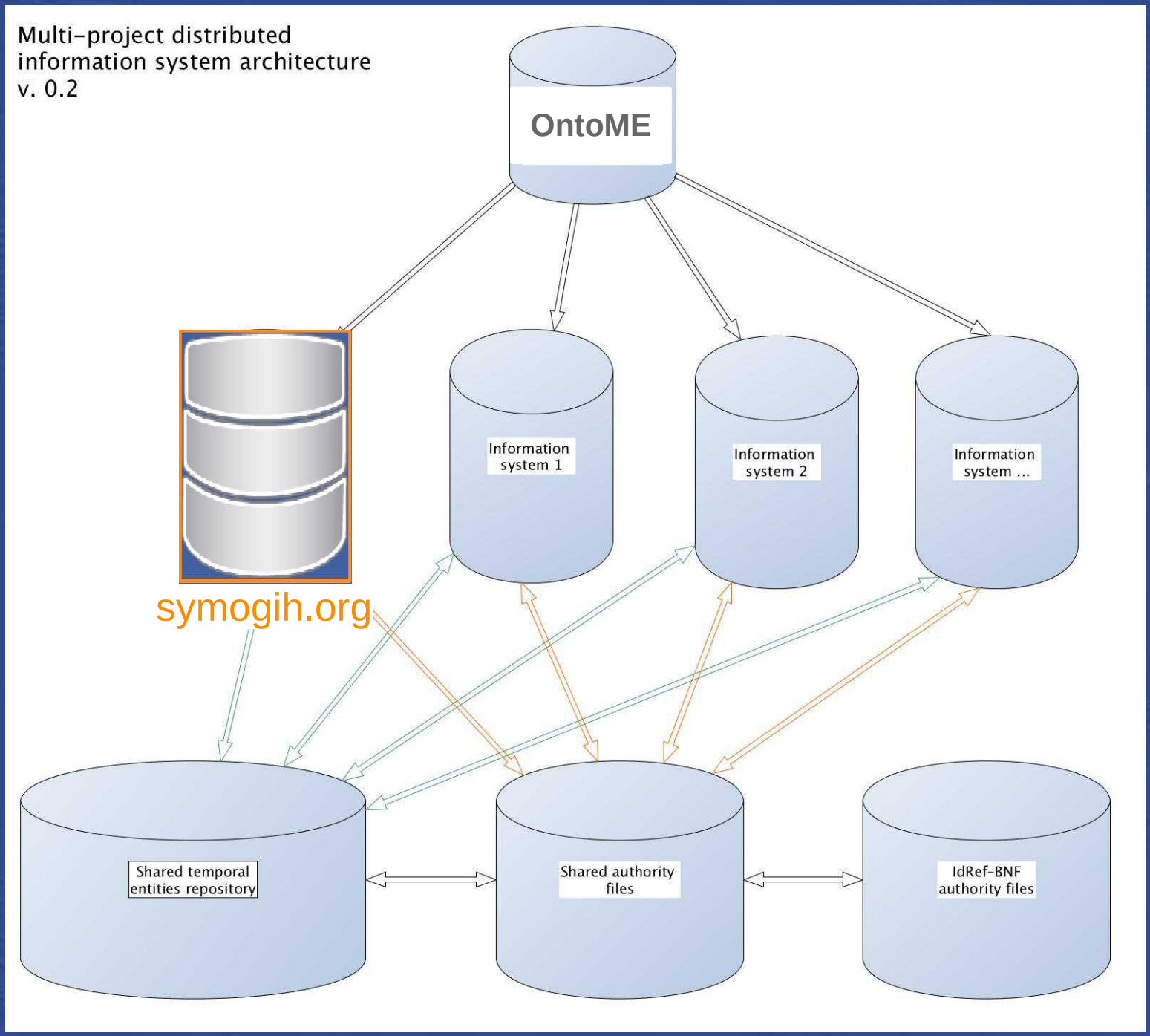


symogih.org

Research agenda

Research specific data model

Research data


Multi-project distributed
information system architecture
v. 0.2
OntoME
Information system 1
Information system 2
Information system ...
symogih.org
Shared temporal entities repository
Shared authority files
IdRef-BNF authority files

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.

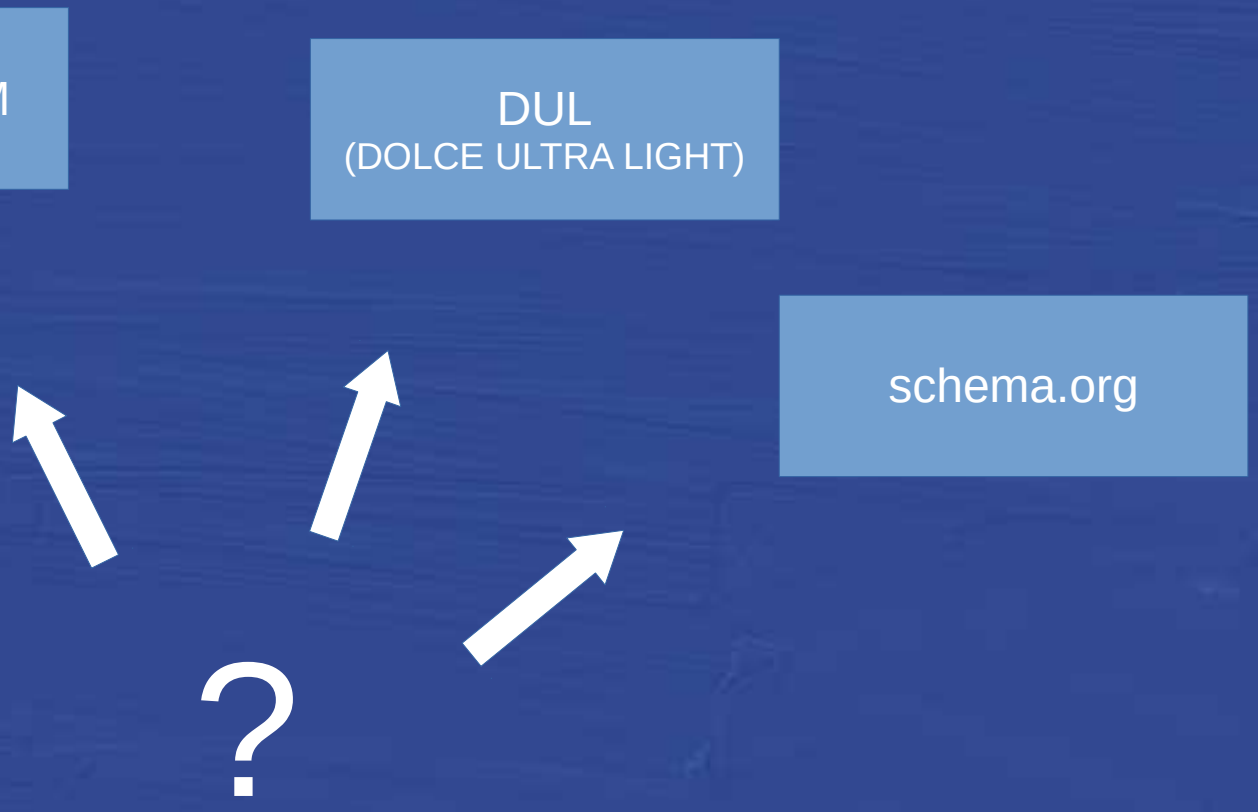
To be **Re-usable**:

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***.



Research agenda

Research specific data model

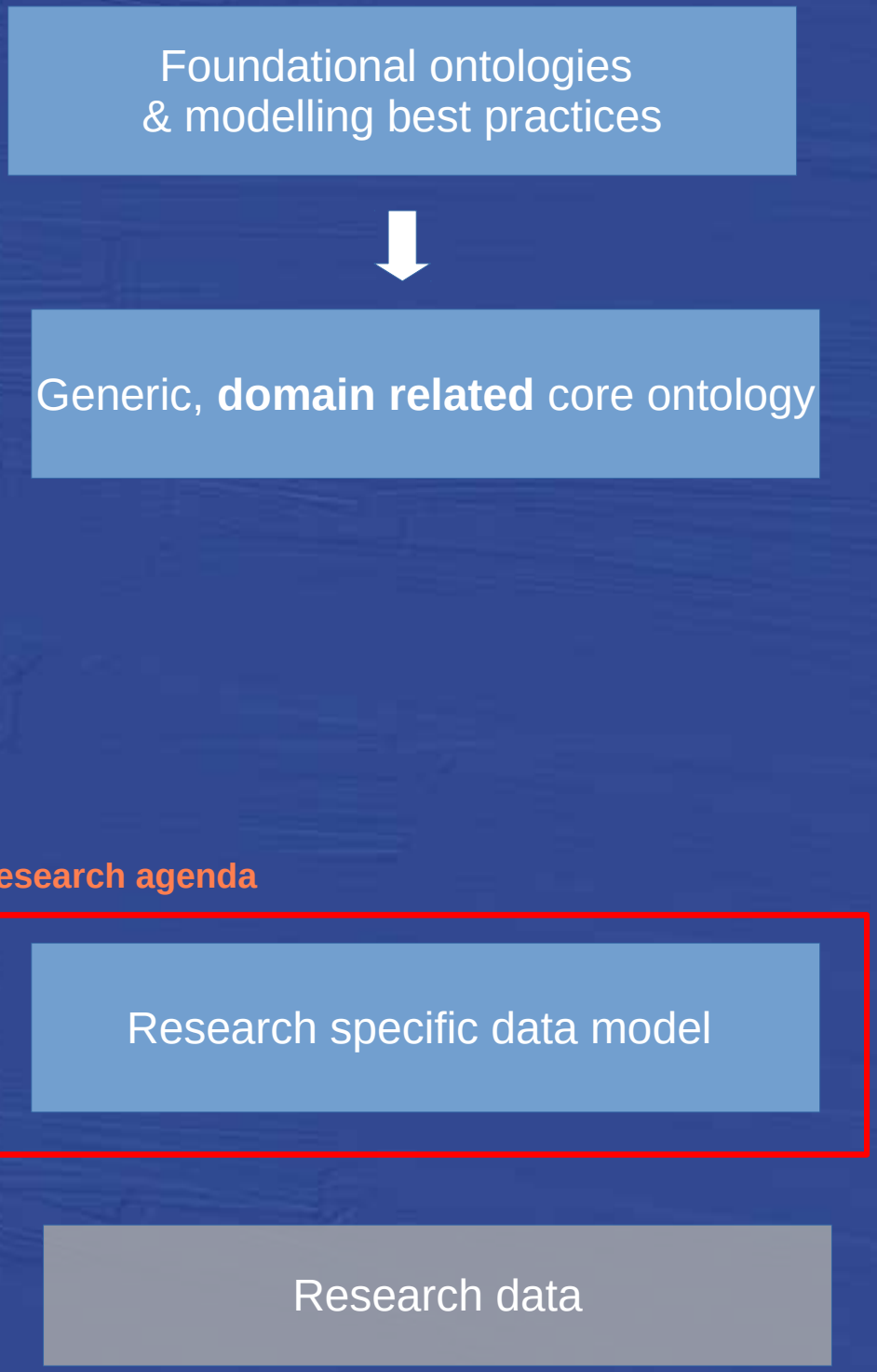
Research data

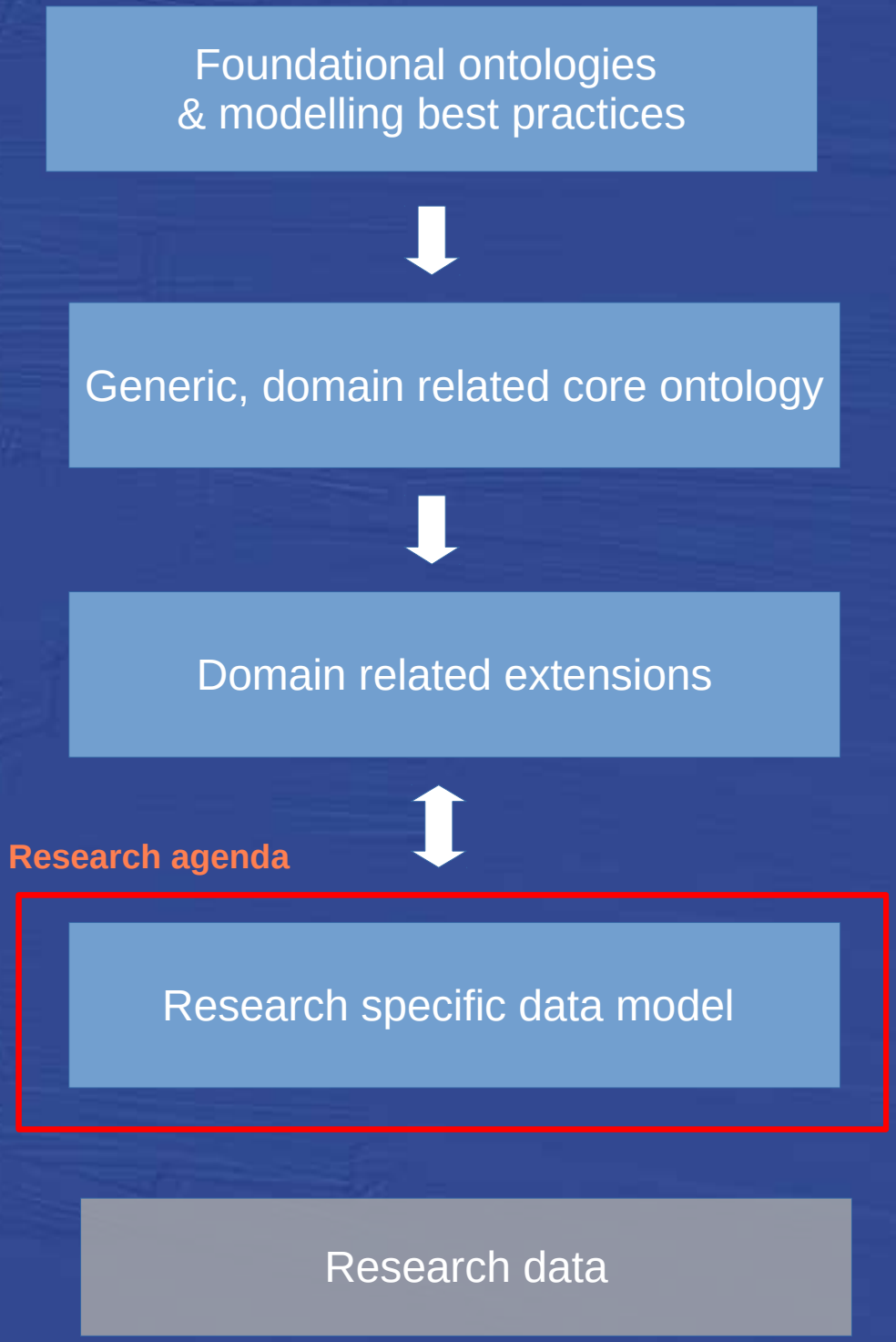
Foundational ontologies
& modelling best practices

Research agenda

Research specific data model

Research data


Foundational ontologies
& modelling best practices
Generic, **domain related** core ontology
Research agenda
Research specific data model
Research data


Foundational ontologies
& modelling best practices
Generic, domain related core ontology
Domain related extensions
Research agenda
Research specific data model
Research data

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."

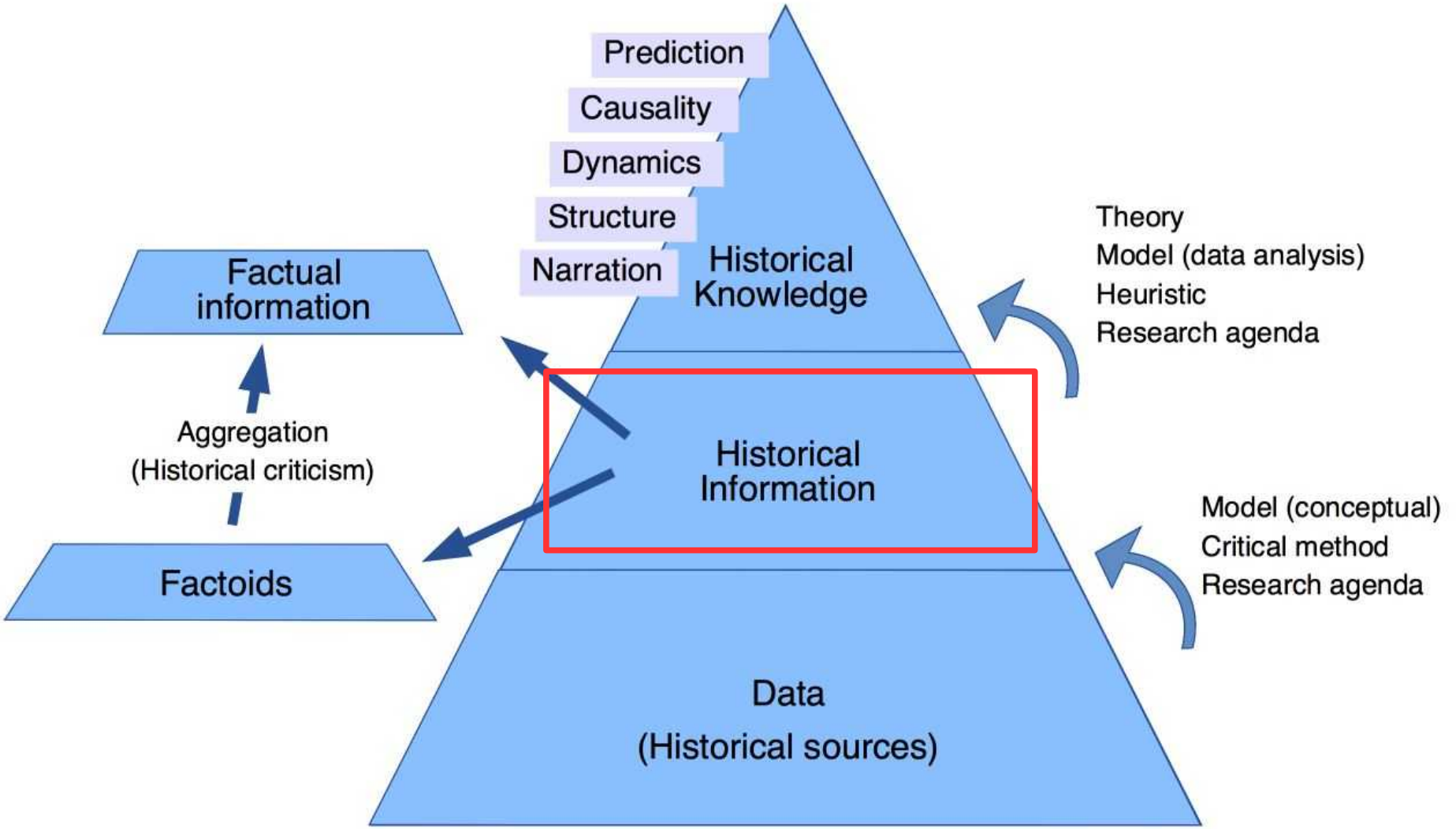
Letter by A to B, 11 March 1616 / Dictionary of astronomers, 2020

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

Factual information as the basis of the production of knowledge



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

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

Factoïds : mentions of states of affairs in the sources

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.

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

Info94542

Type d'information: Enseignement - TyIn97

Date: 1592

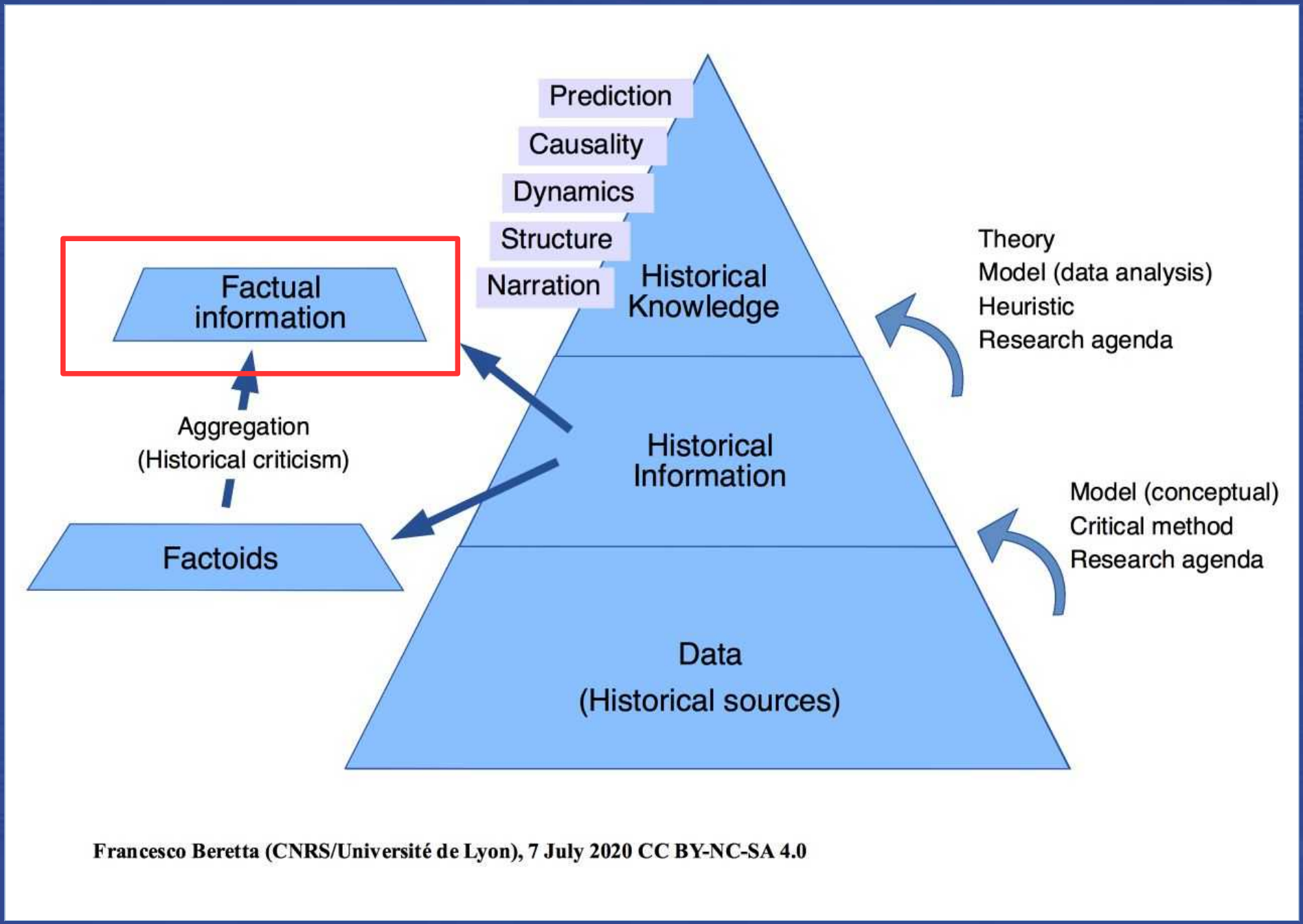
Composantes de l'information

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

Factual information: states of affairs as such

Our focus : modelling factual information



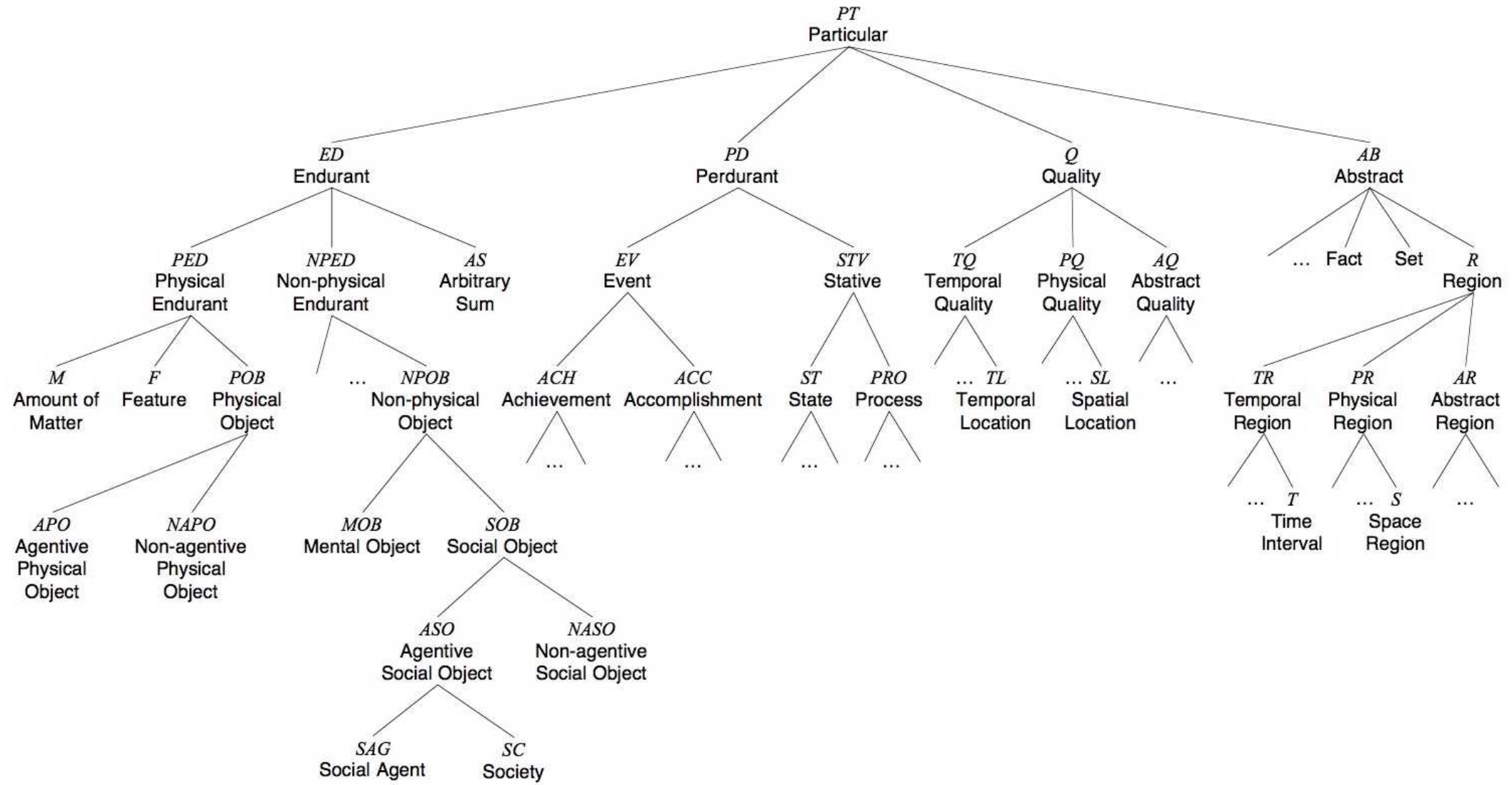
Francesco Beretta (CNRS/Université de Lyon), 7 July 2020 CC BY-NC-SA 4.0

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

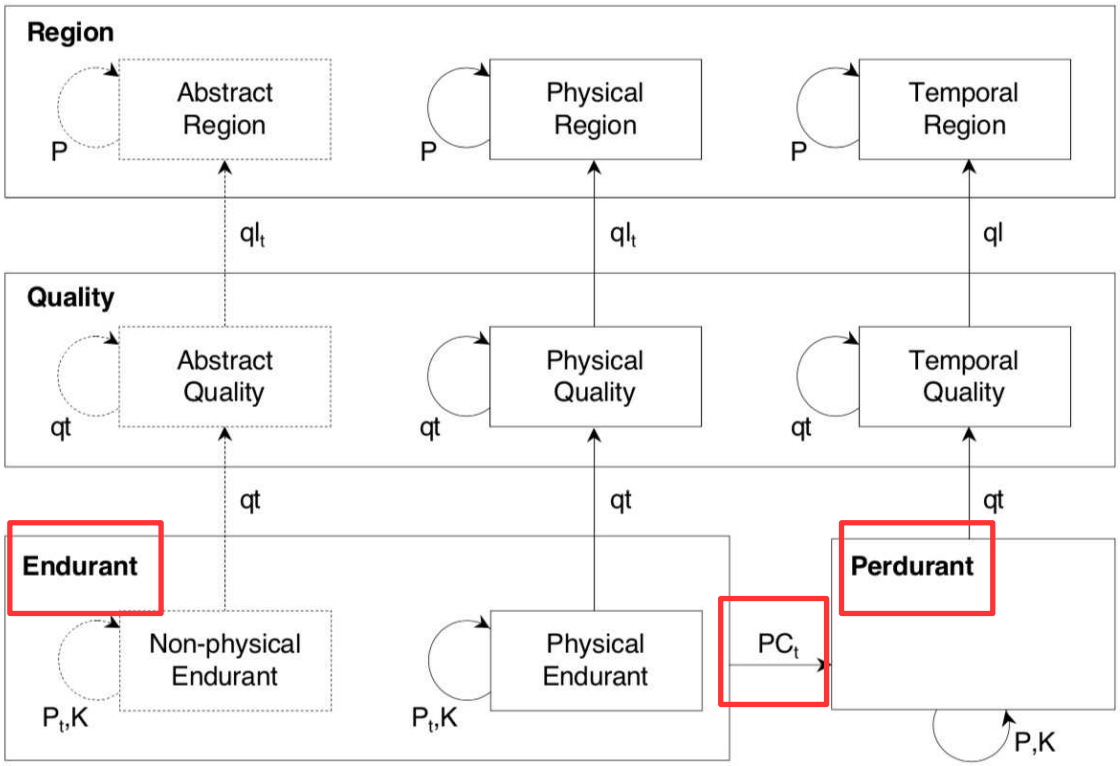


Figure 4: Primitive relations between basic categories (the dotted lines to the left indicate that we are less confident with what concerns non-physical endurants.

4.2.7 Participation

(Dd63) $\mathsf{PC}_{\mathsf{C}}(x,y) \triangleq \exists t(\mathsf{PRE}(y,t)) \wedge \forall t(\mathsf{PRE}(y,t) \rightarrow \mathsf{PC}(x,y,t))$ *(Const. Participation)*

(Dd64) $\mathsf{PC}_{\mathsf{T}}(x,y,t) \triangleq PD(y) \wedge \forall z((\mathsf{P}(z,y) \wedge \mathsf{PRE}(z,t)) \rightarrow \mathsf{PC}(x,z,t))$ *(Temporary Total Participation)*

(Dd65) $\mathsf{PC}_{\mathsf{T}}(x,y) \triangleq \exists t(\mathsf{ql}_T(t,y) \wedge \mathsf{PC}_{\mathsf{T}}(x,y,t))$ *(Total Participation)*

(Dd66) $\mathsf{mpc}(x,y) \triangleq x = \sigma_t z(\mathsf{PC}_{\mathsf{T}}(z,y))$ *(Maximal Participant)*

(Dd67) $\mathsf{mppc}(x,y) \triangleq x = \sigma_t z(\mathsf{PC}_{\mathsf{T}}(z,y) \wedge PED(z))$ *(Maximal Physical Participant)*

(Dd68) $\mathsf{lf}(x,y) \triangleq x = \sigma z(\mathsf{PC}_{\mathsf{T}}(y,z))$ *(Life)*

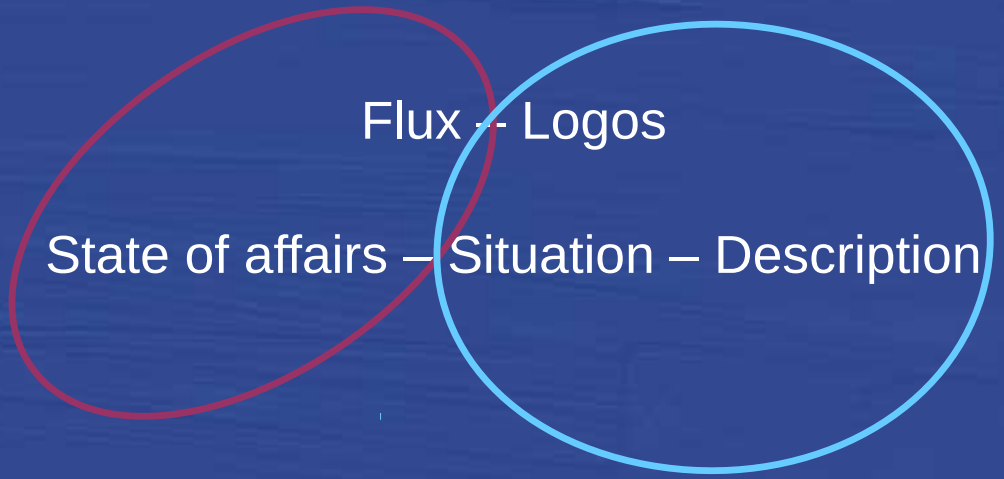
WonderWeb Deliverable D18, p.30

Descriptions and Situations (DnS)

Flux – Logos

State of affairs – Situation – Description

Descriptions and Situations (DnS)

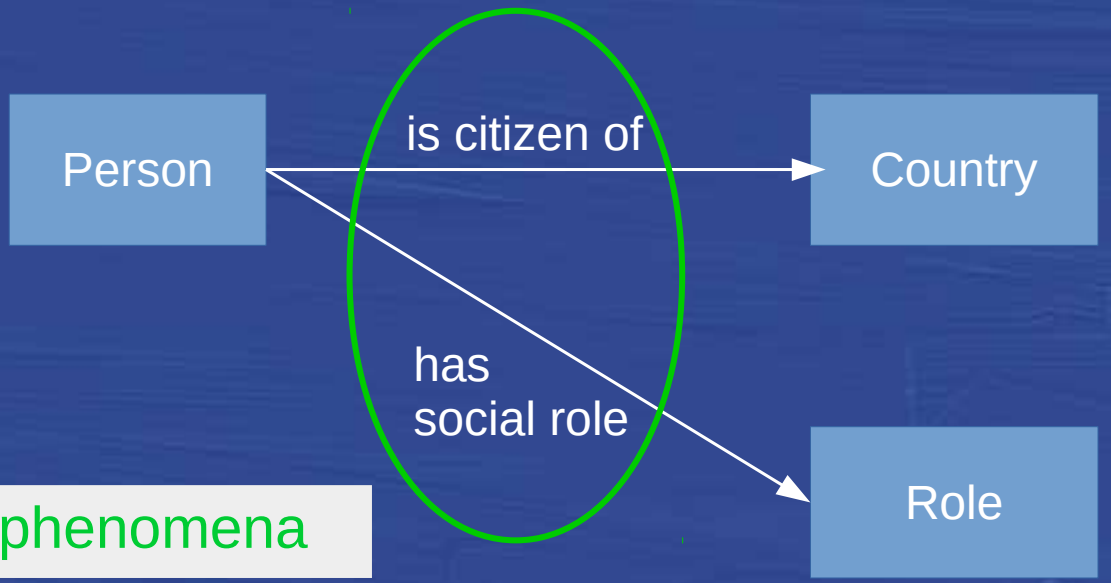


Descriptions and Situations (DnS)

Flux – Logos

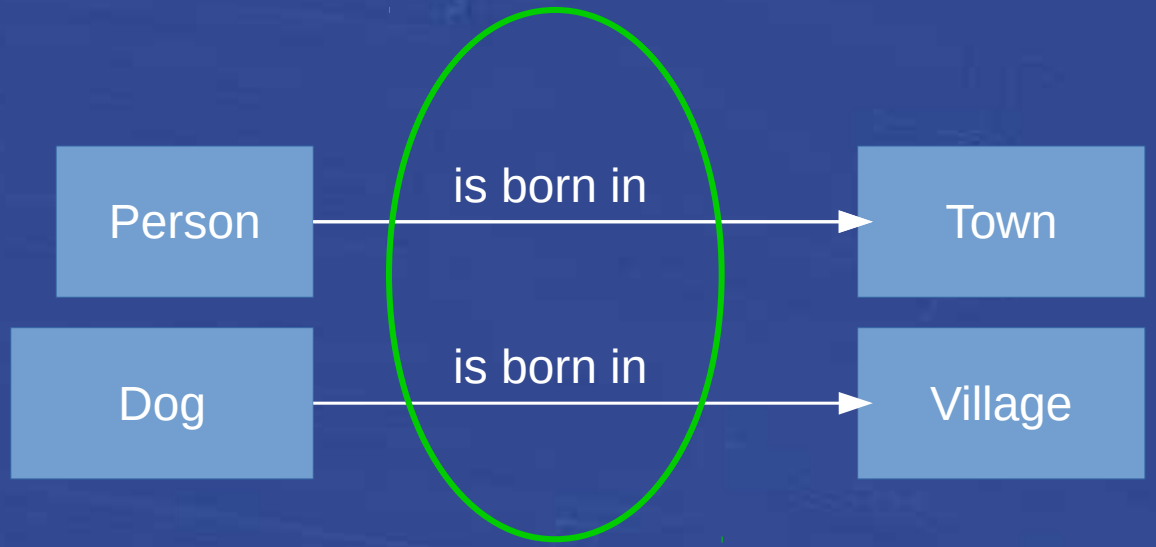
State of affairs – Situation – Description

Descriptions and Situations – DnS



physical or biological phenomena

DOLCE



Descriptions and Situations – DnS

DOLCE

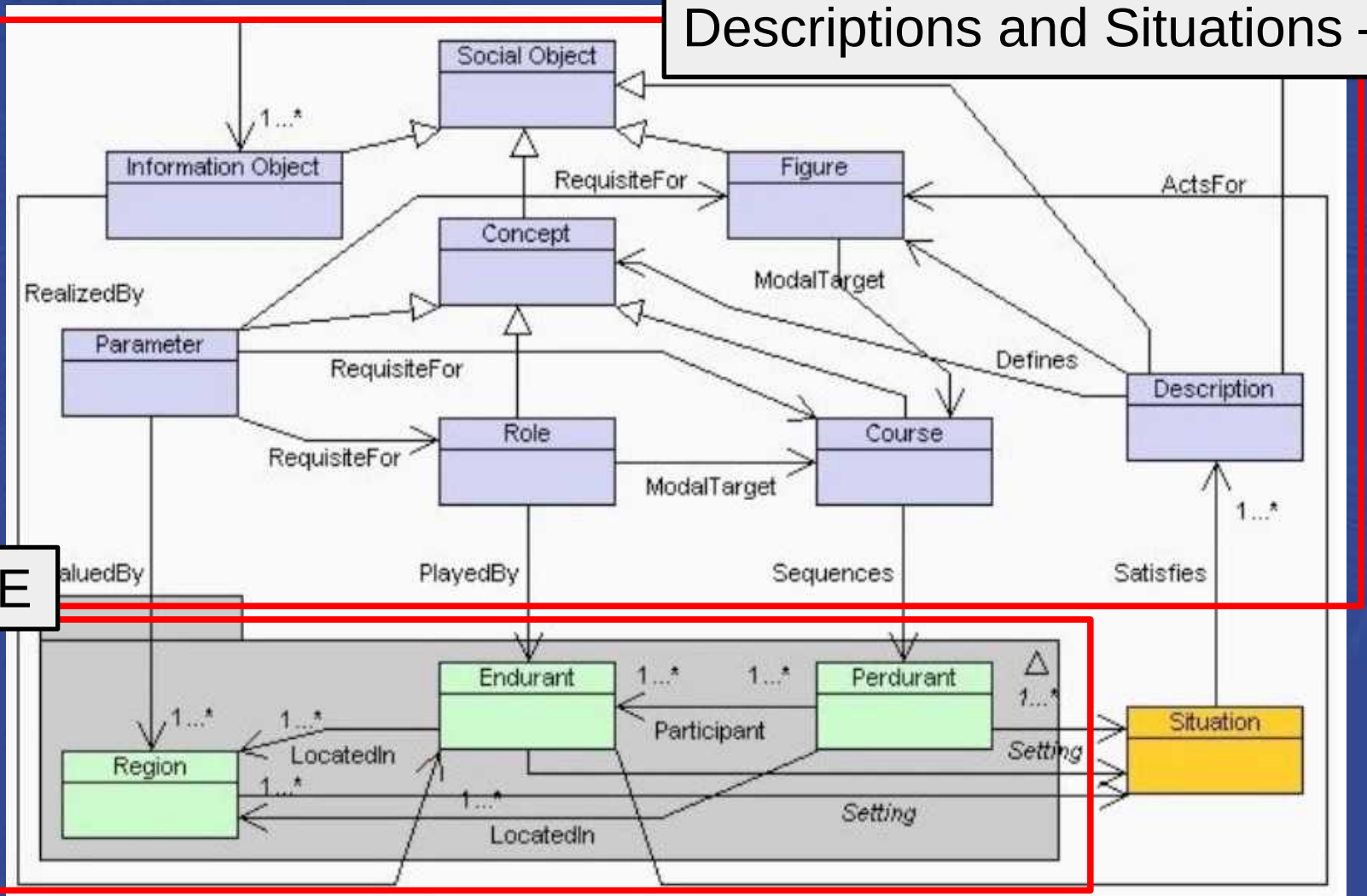
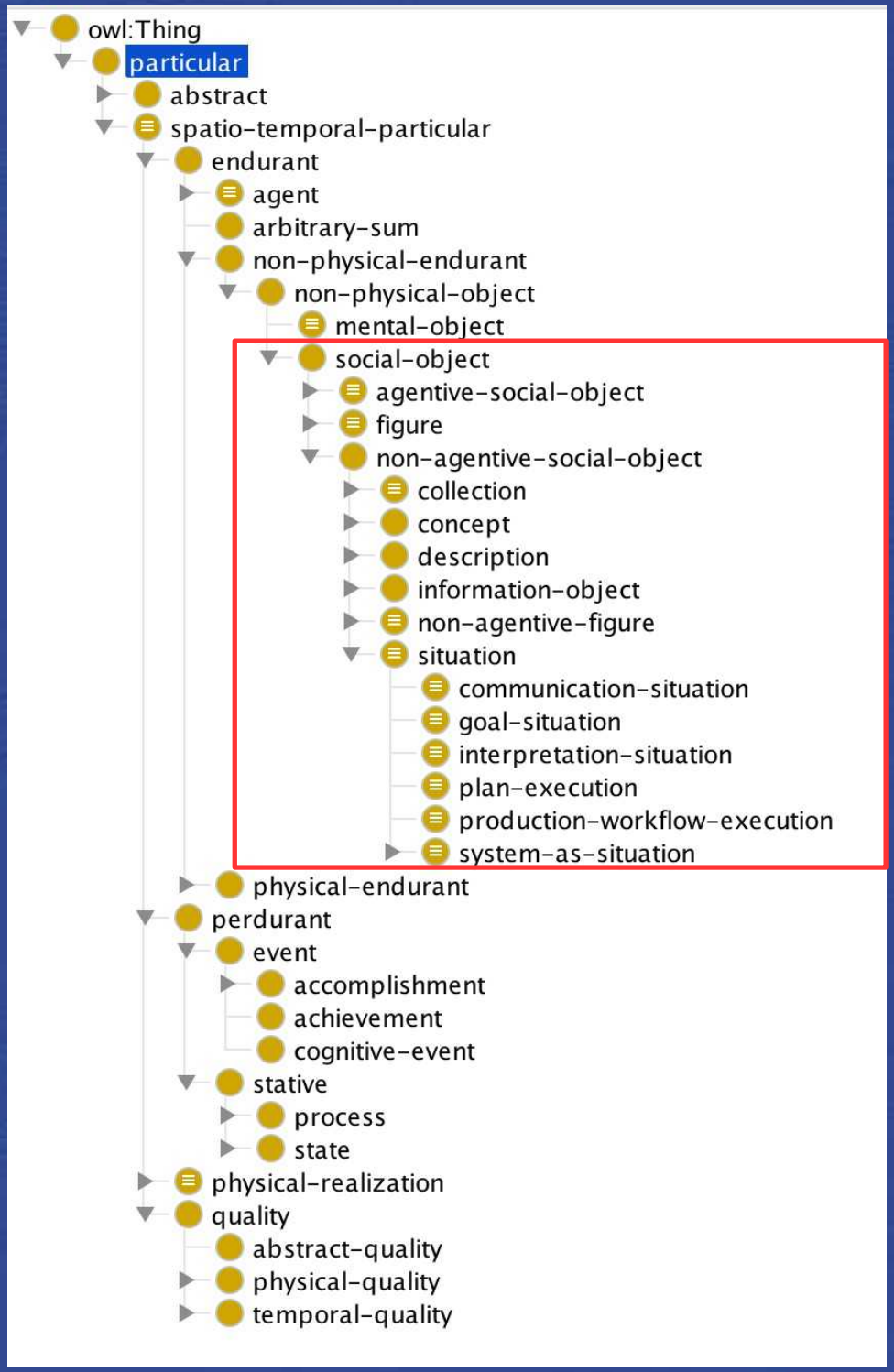


Fig. 2. A UML class diagram for D&S. The lower part of the pattern (within the grey package) is called the *ground ontology*, the higher is called the *descriptive ontology*; a situation satisfies a description if the two parts match according to the axioms specified for the concepts defined by the description.

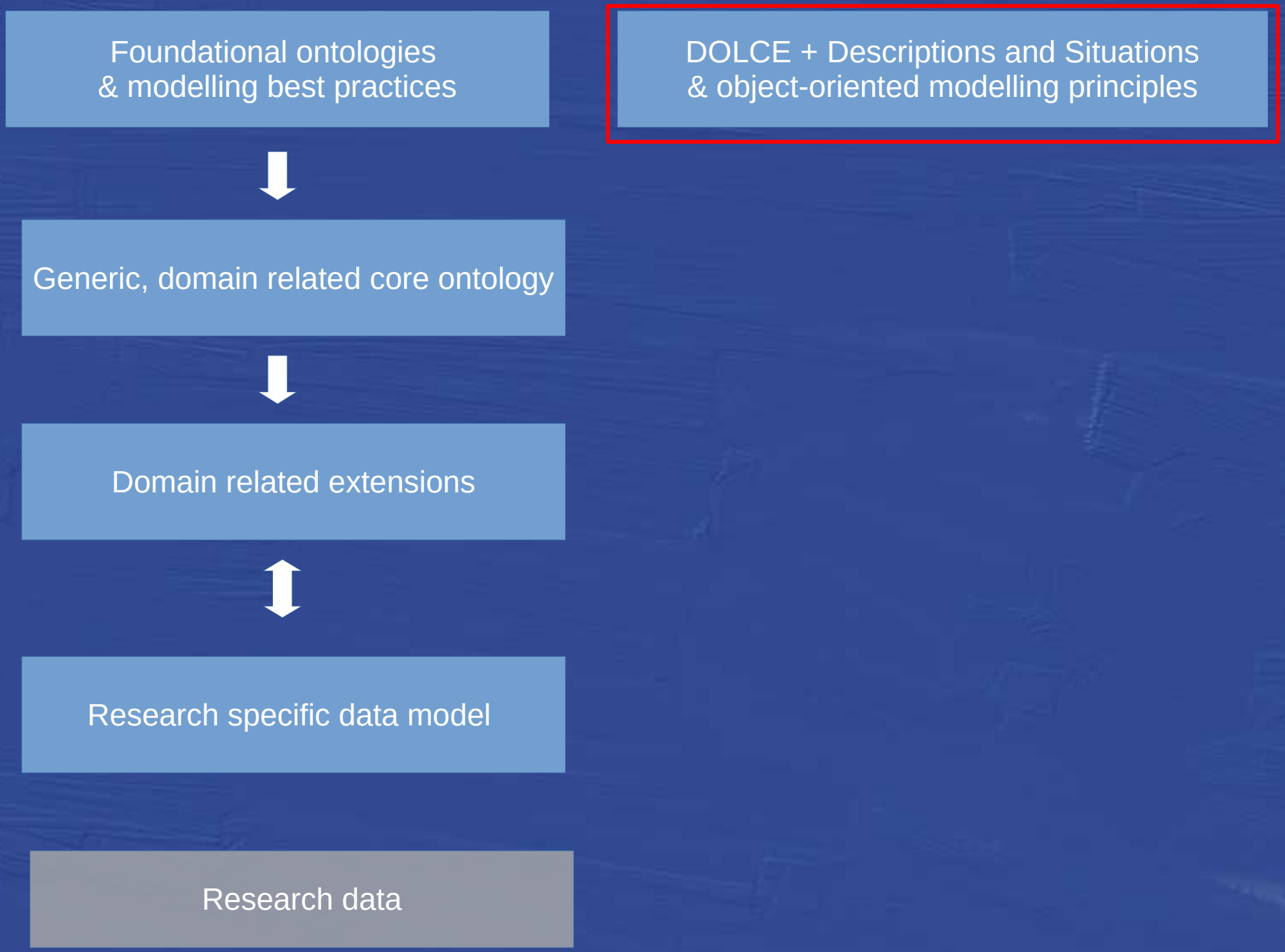
Bottazzi E., Catenacci C., Gangemi A., Lehmann J.(2006) (from pre-print, not in published version)



DOLCE + DnS

=

DOLCE Lite PLUS


Foundational ontologies
& modelling best practices
DOLCE + Descriptions and Situations
& object-oriented modelling principles
Generic, domain related core ontology
Domain related extensions
Research specific data model
Research data

Foundational ontologies
& modelling best practices

DOLCE + Descriptions and Situations
& object-oriented modelling principles

Generic, **domain related** core ontology

Research specific data model

Research data

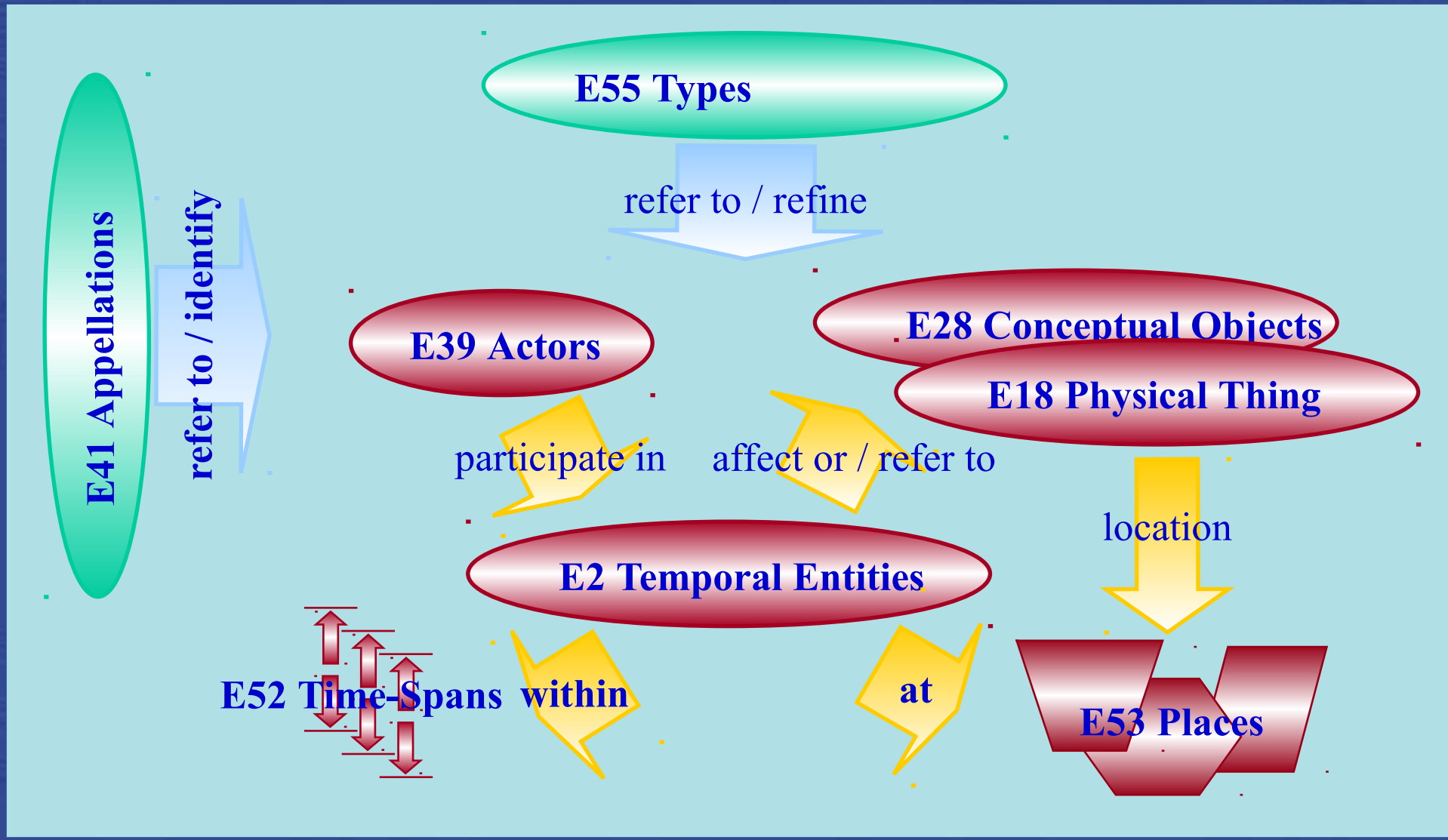
3.

The CIDOC CRM

and

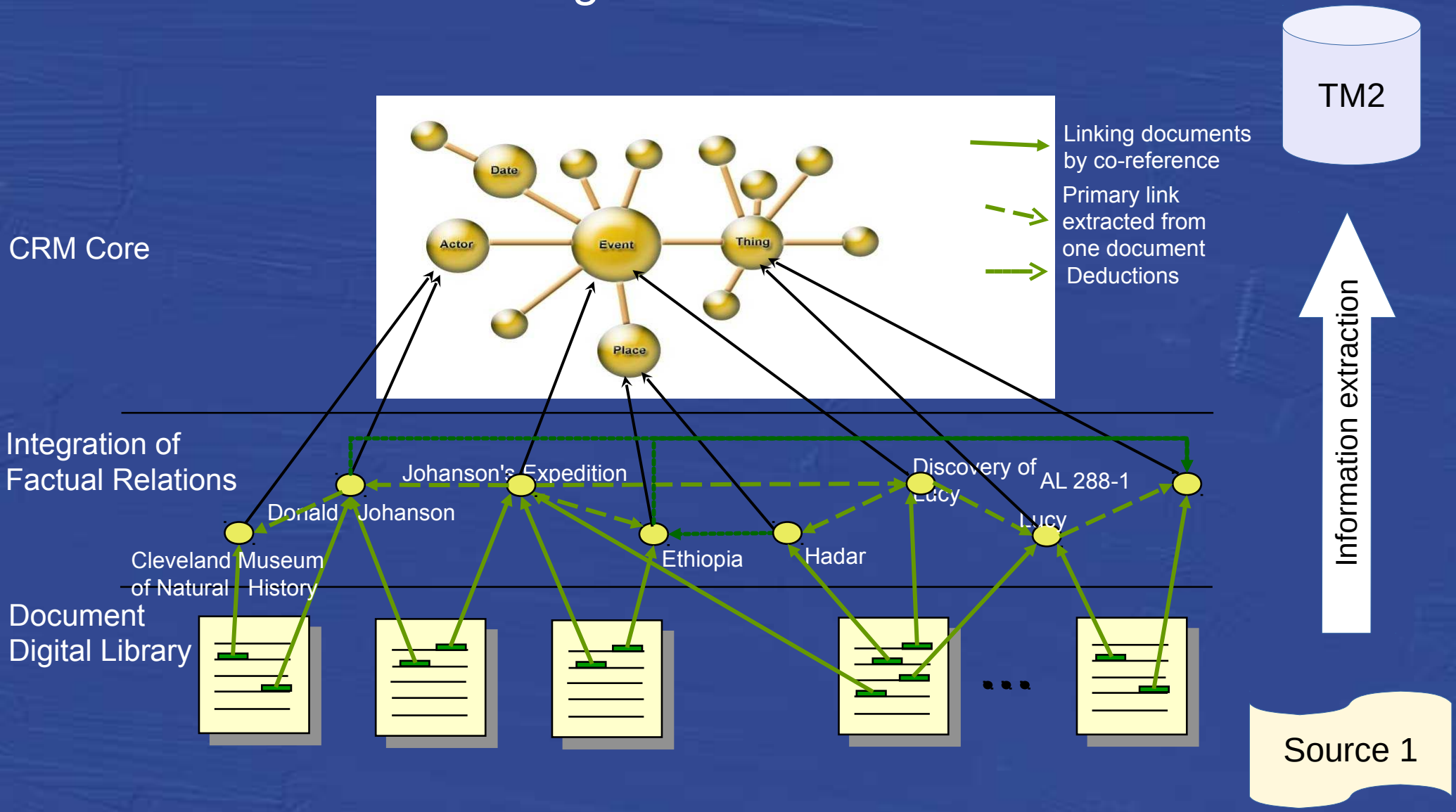
a high-level extension for humanities and social sciences

The CIDOC CRM (ISO21127:2006)
A semantic framework that provides *interoperability* between different sources of **cultural heritage information**



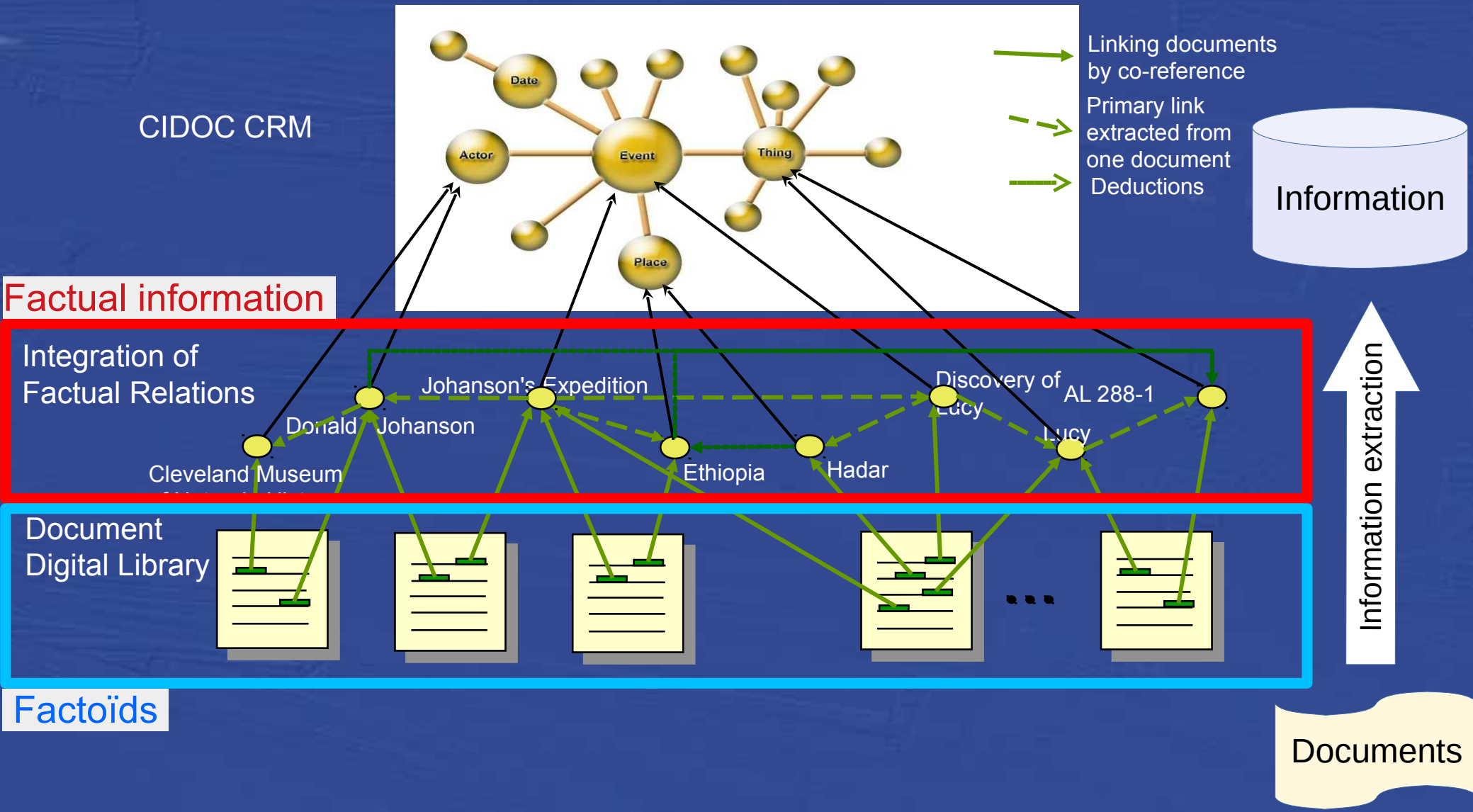
Stephen Stead (2008)

Integration of information extracted from documents using the CIDOC CRM

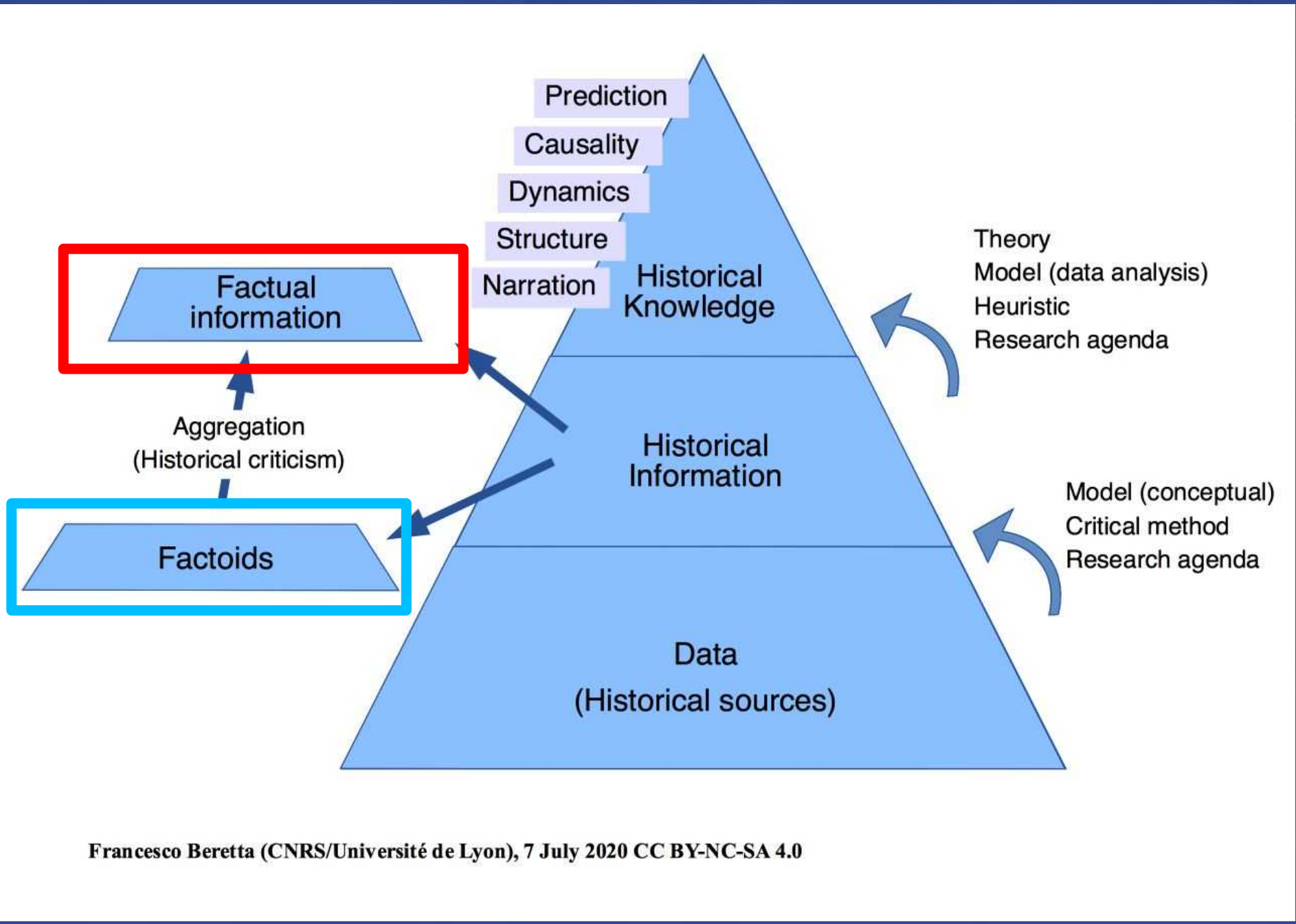


Stephen Stead (2008). Cf. Doerr M. (2003), fig. 4, p. 81

Integration of information extracted from documents using the CIDOC CRM



Integration of factual information



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

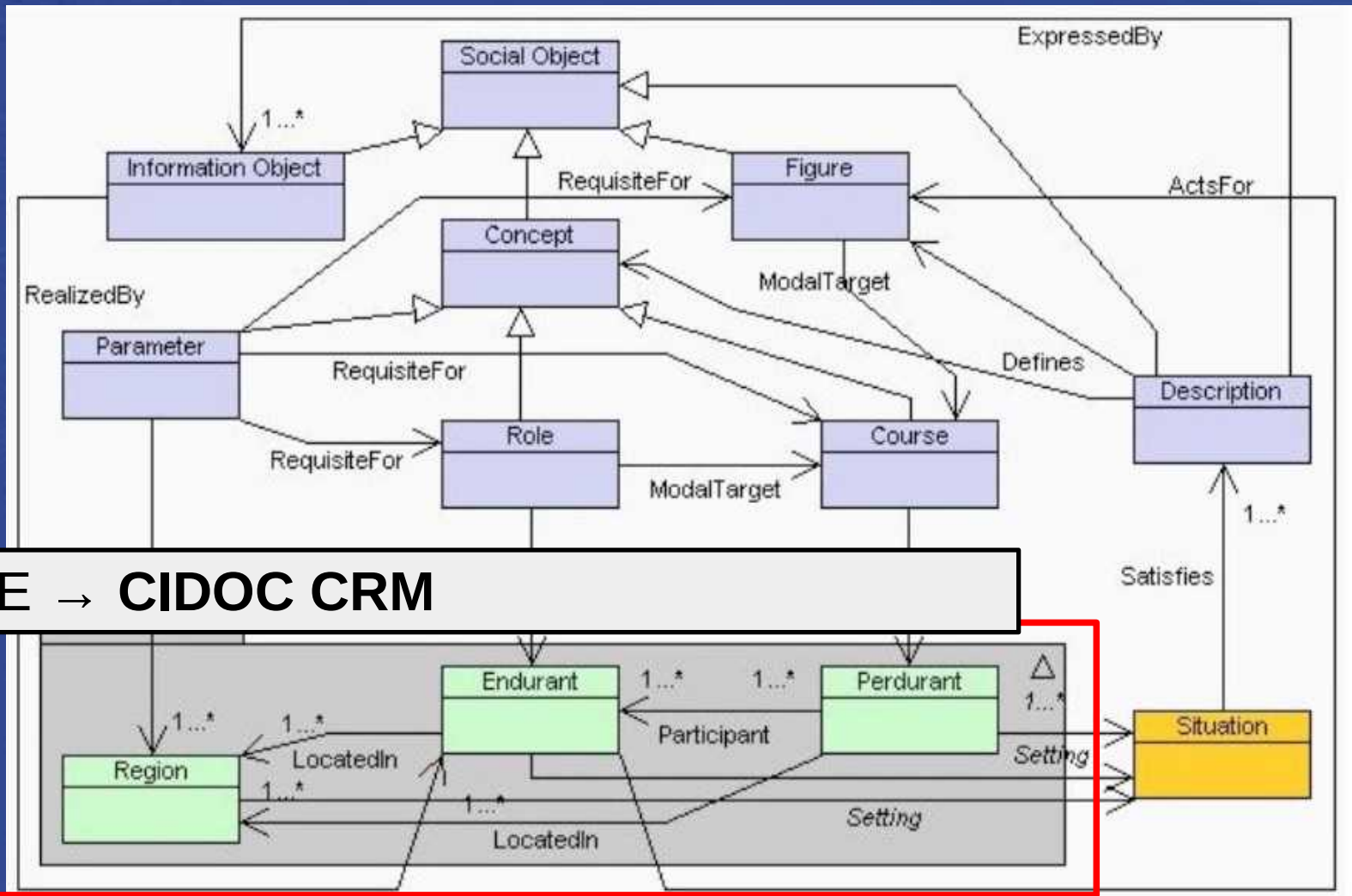
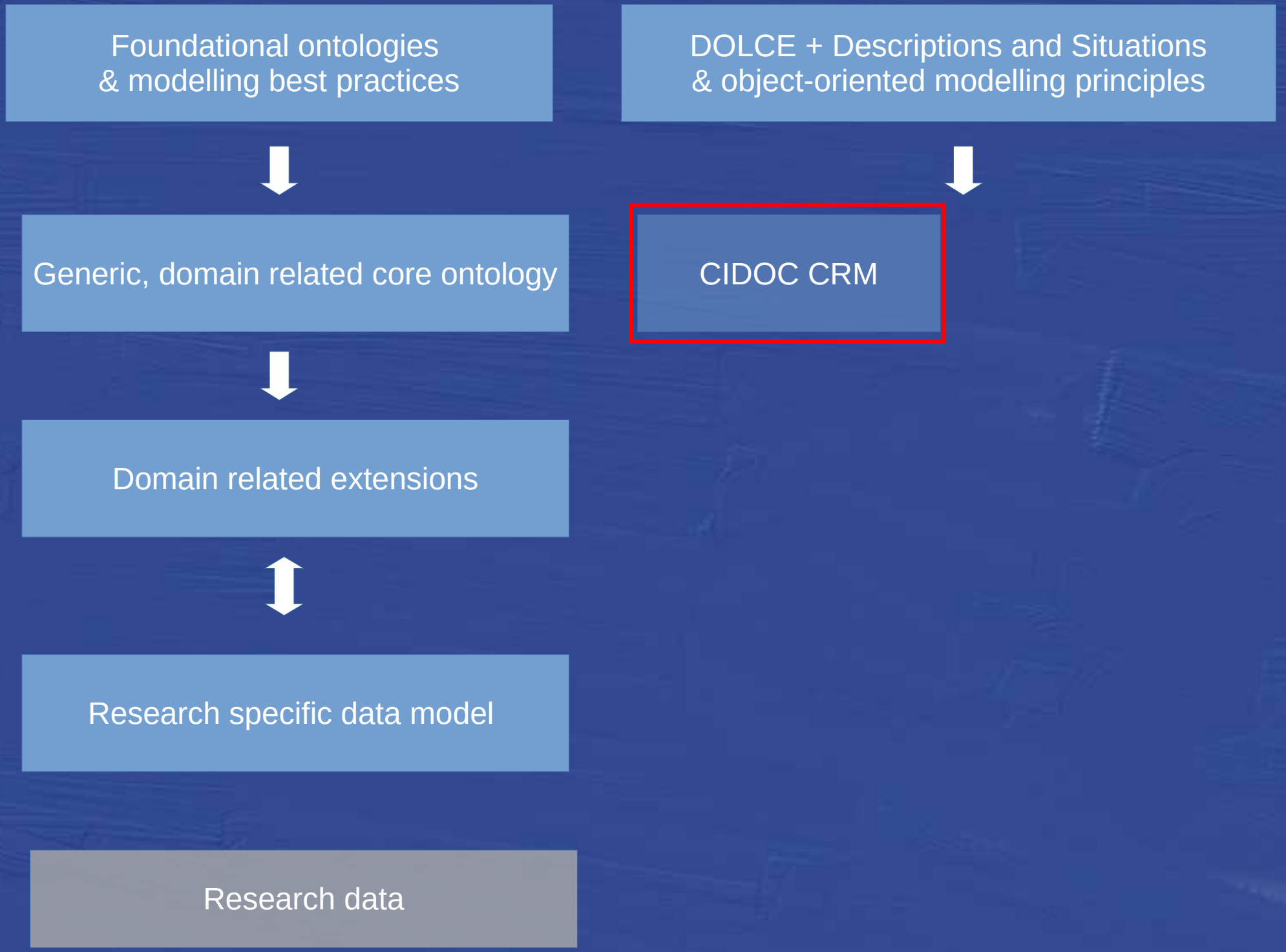
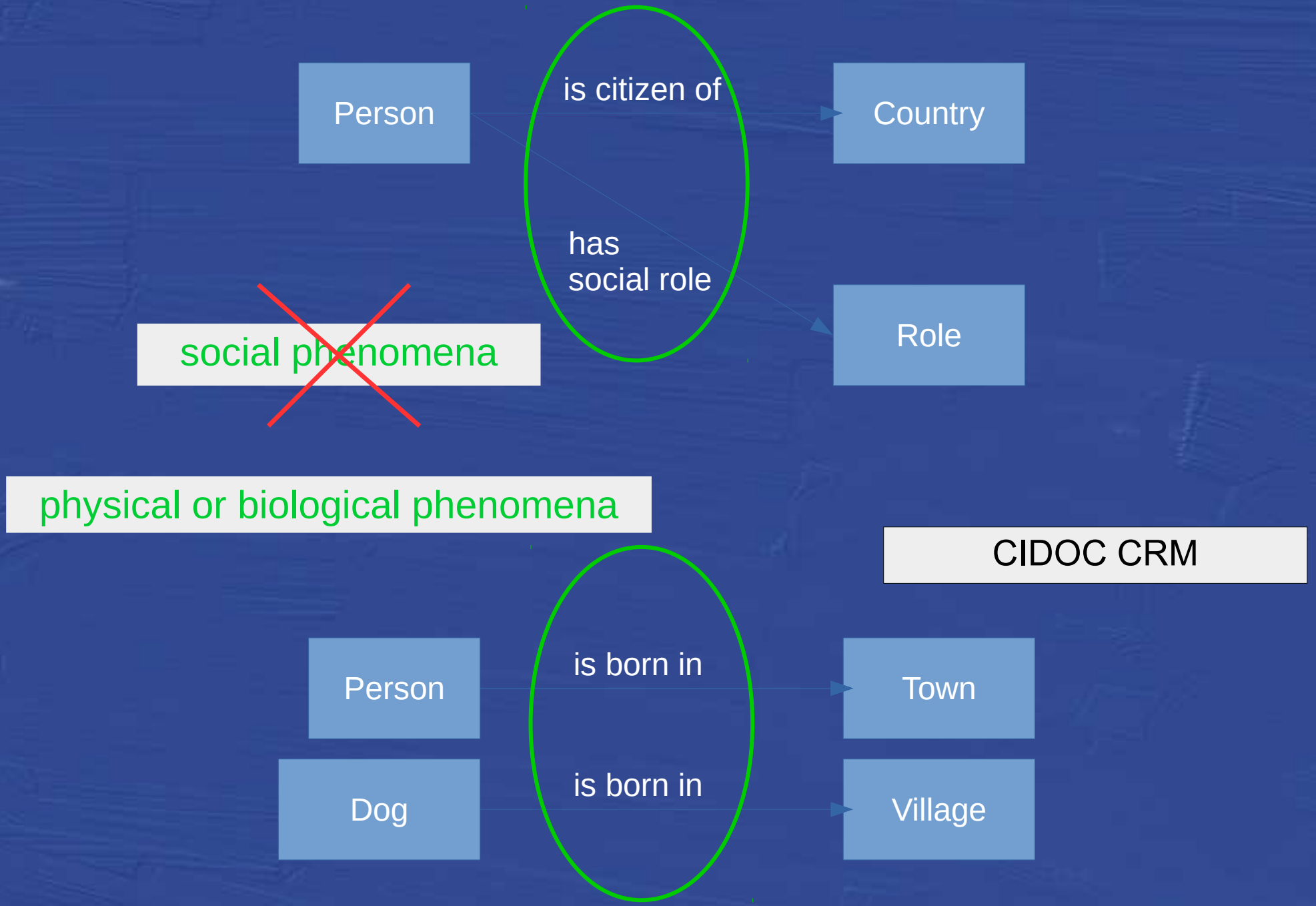


Fig. 2. A UML class diagram for D&S. The lower part of the pattern (within the grey package) is called the *ground ontology*, the higher is called the *descriptive ontology*; a situation satisfies a description if the two parts match according to the axioms specified for the concepts defined by the description.

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


Foundational ontologies
& modelling best practices
DOLCE + Descriptions and Situations
& object-oriented modelling principles
Generic, domain related core ontology
CIDOC CRM
Domain related extensions
Research specific data model
Research data


Person
is citizen of
Country
has
social role
Role
social phenomena
physical or biological phenomena
CIDOC CRM
Person
is born in
Town
Dog
is born in
Village

Semantic Data for Humanities and Social Sciences (SDHSS) CIDOC CRM Top-Level Extension

Semantic Data for Humanities and Social Sciences (SDHSS) CIDOC CRM Top-Level Extension

Description:

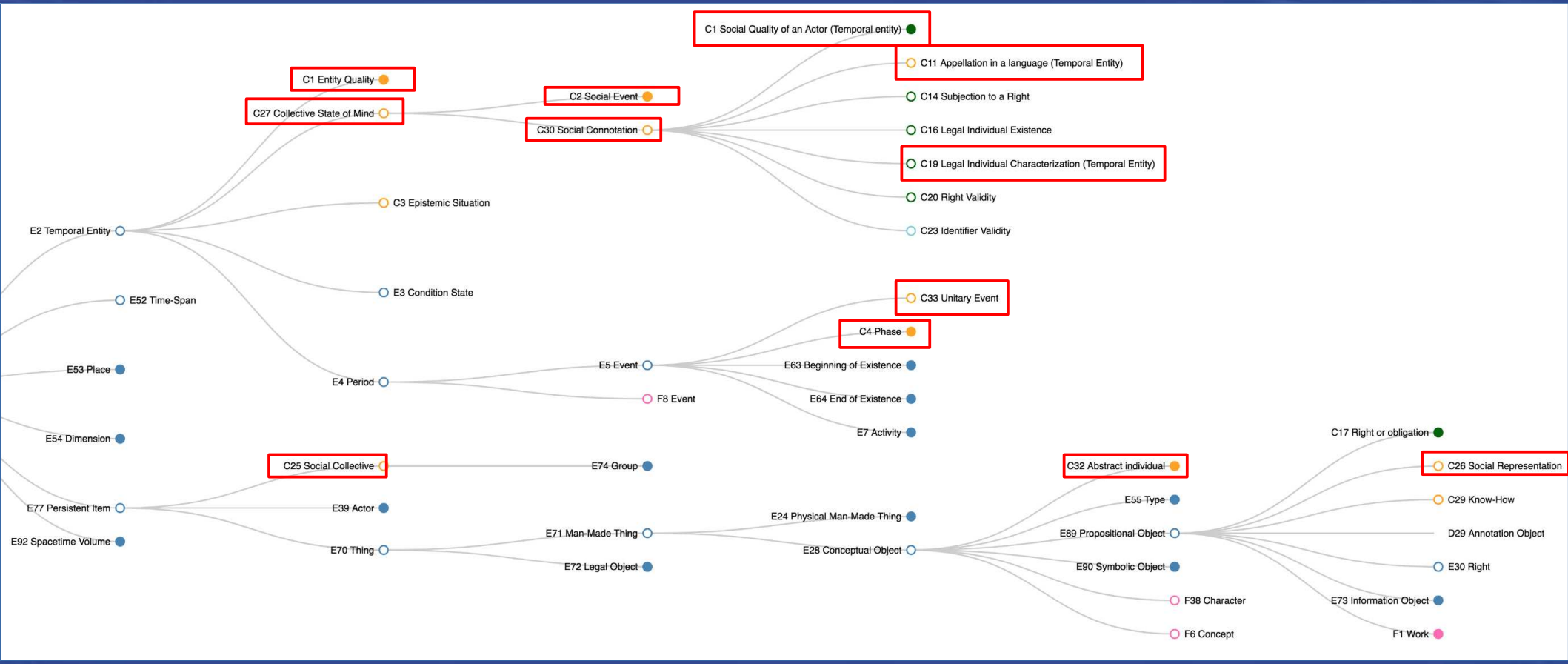
Published by Francesco Beretta (CNRS/Université de Lyon), 7 December 2020. Last revised on March 30 2021. (CC BY-SA 4.0)

The extension of CIDOC CRM for semantic data for humanities and social sciences (SDHSS) stems from the need to conceptualise the reality in the world, and more specifically factual information, from the point of view of historical research. The ontological commitment is therefore related to the domain of discourse of history but insofar as history, as a discipline that studies the life of humans and societies in the past, is interested in all the different aspects of social, economic, political, religious, literary and cultural life, the scope of this extension could be defined as the whole of social and human life, apprehended from the descriptive point of view, and global approach to reality, that characterises historical research.

This definition of the scope or domain modelled is based on the conviction that in a constructivist approach of scientific knowledge, a conceptualisation and data model can only be developed from the point of view of a specific discipline because *scientific objects* do not exist in the absolute but depend on the method and research agenda. They depend on the perspective or epistemic context researchers adopt in considering states of affairs: *scientific objects*, and semantic models modelling them, are not declared to be the only appropriate and exclusive representation of *things* in the pre-Kantian sense but defined as *intentional objects* constructed from the point of view of a discipline and methodological approach in relation to things in the world. Scientific objects are not the things in the world themselves, even if they must necessarily refer to them by way of observation or experimentation, if a scientific and therefore realistic approach is to be maintained. This corresponds to the notion of inter-objectivity in social sciences relying on the distinction between things in themselves and things as perceived, experienced and discussed by human subjects, in their shared intentionality and in relation to their social practices and context.

ontome.net/namespace/11

SDHSS and extension for social life



ontome.net/namespace/11

SDHSS – ontome.net/namespace/11

CIDOC CRM

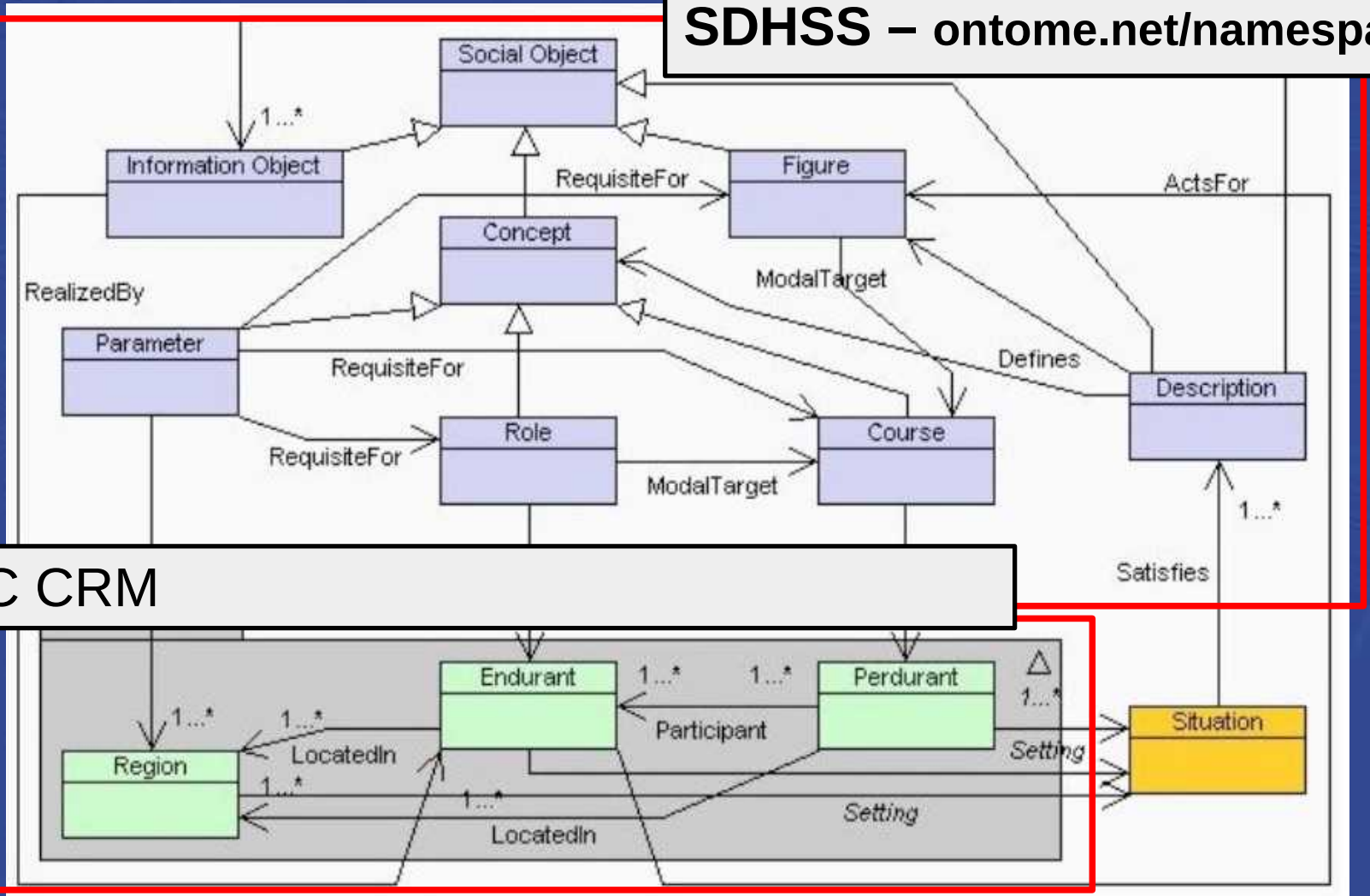
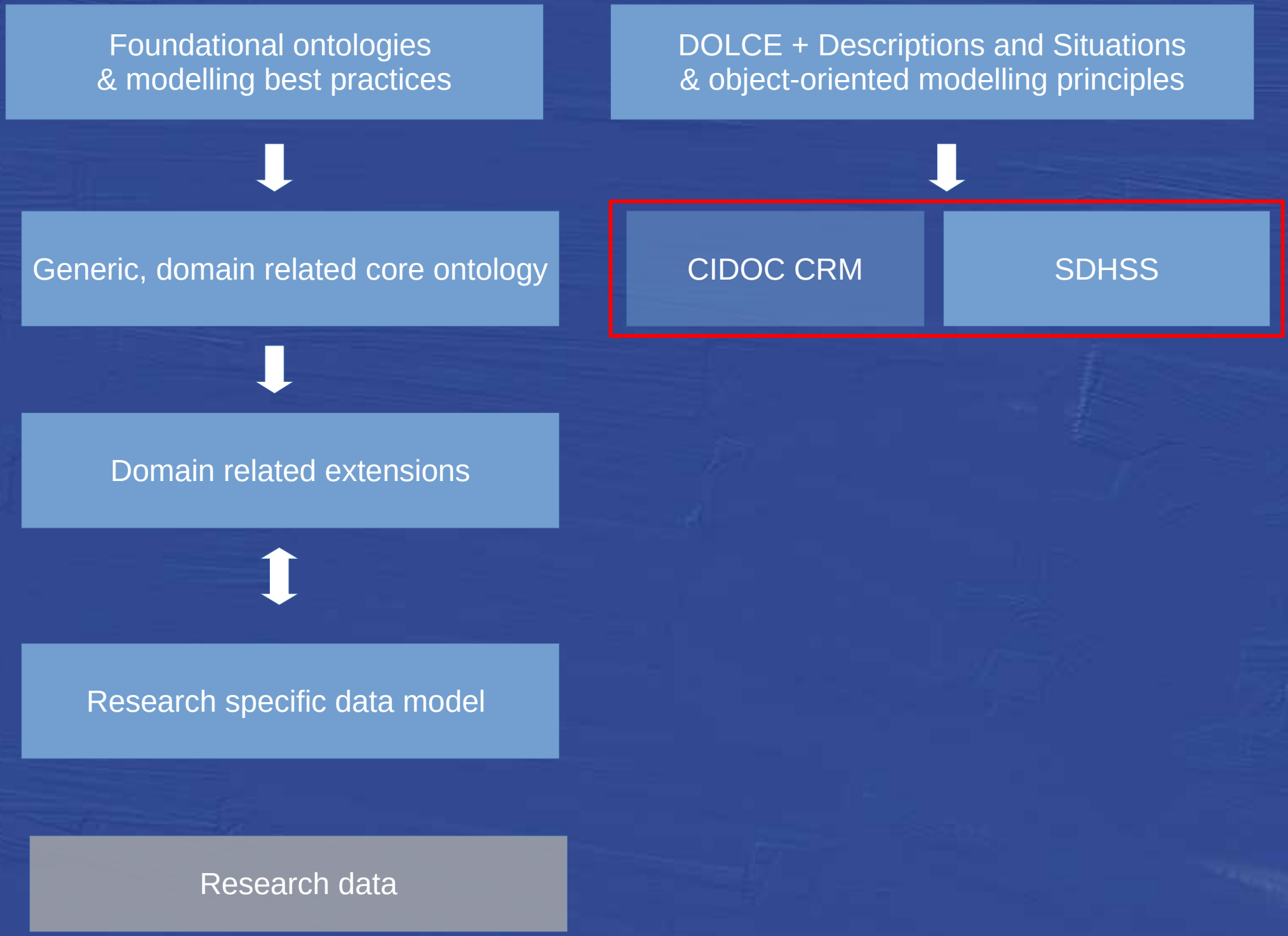
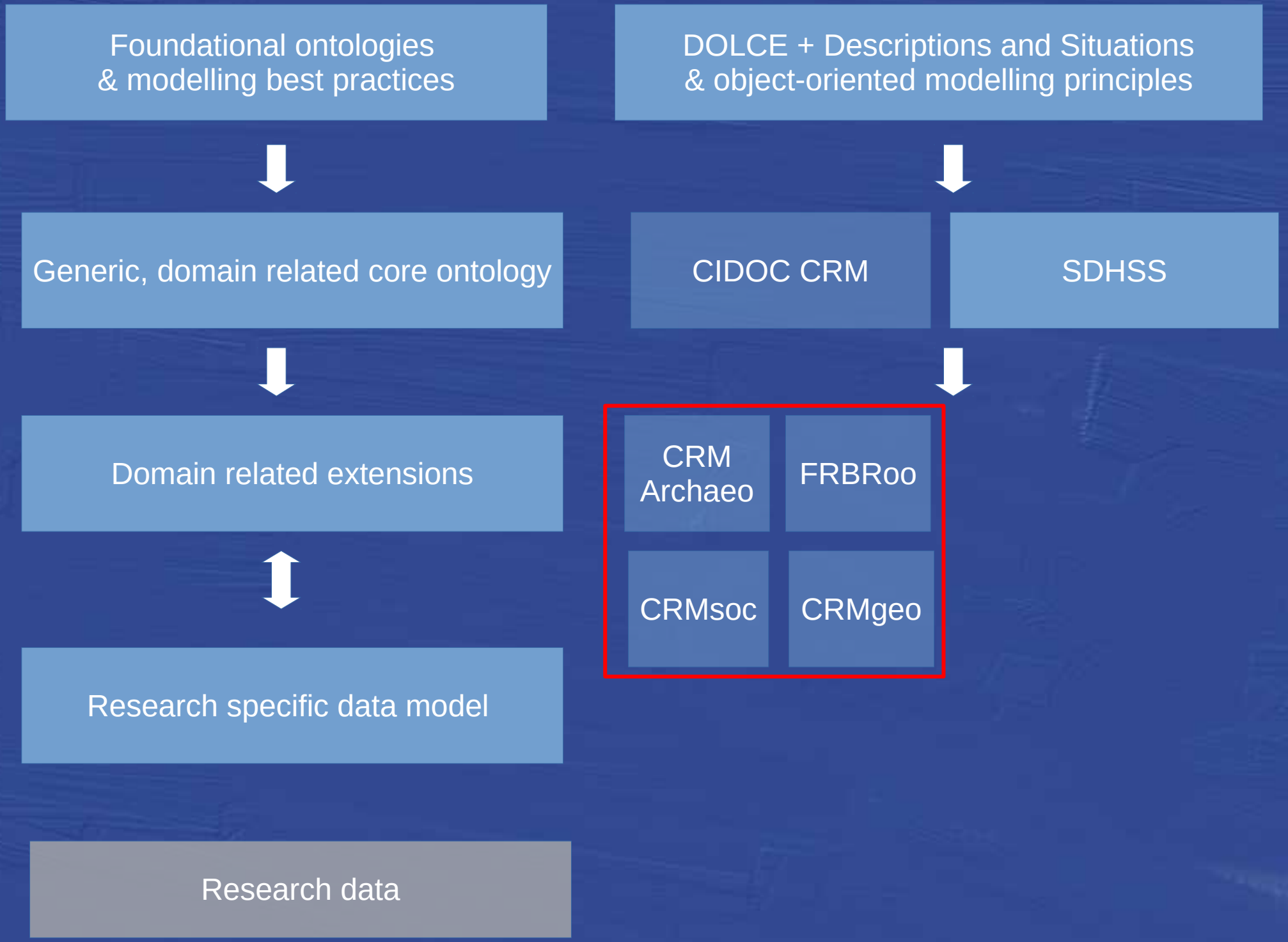
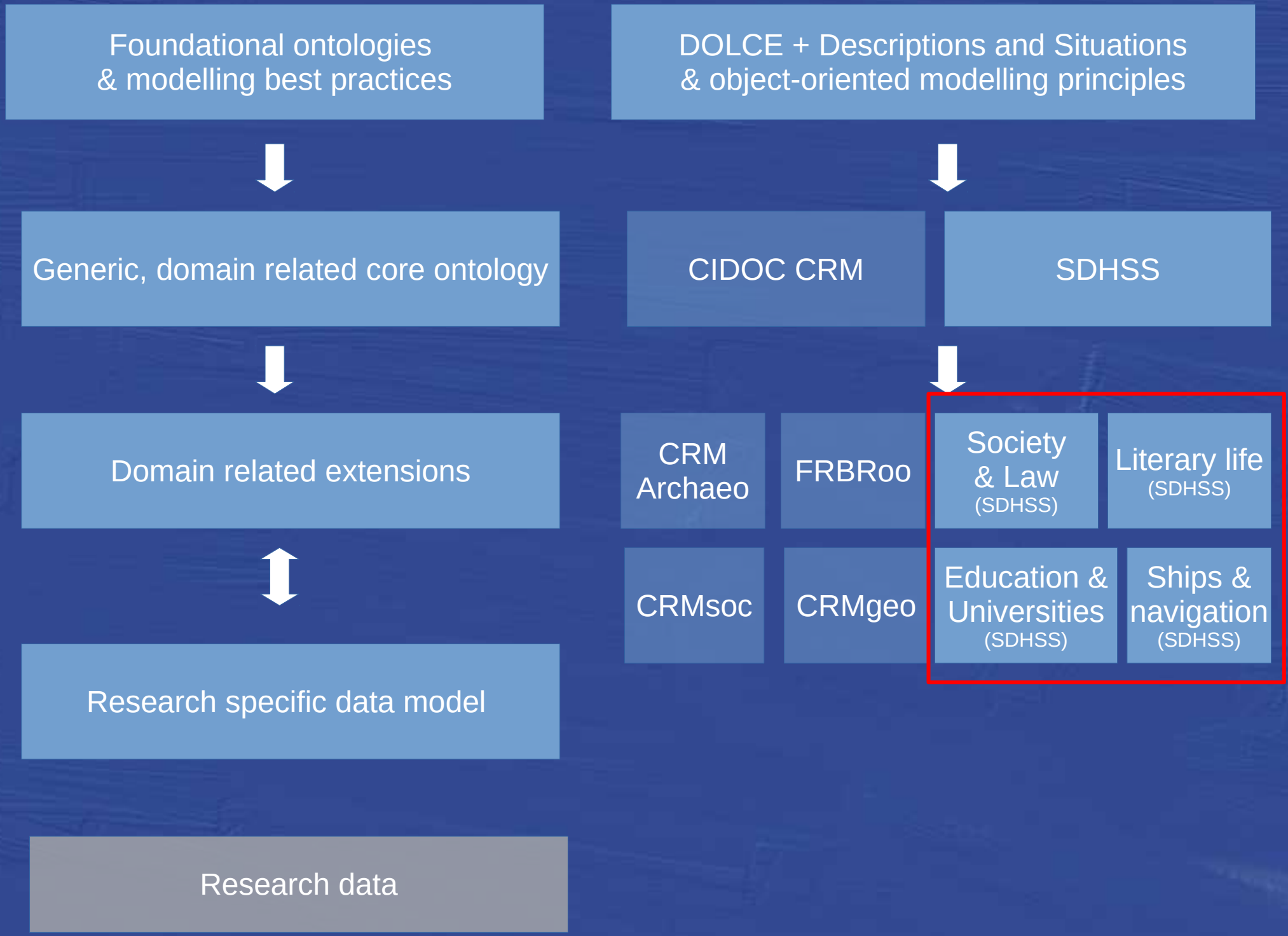


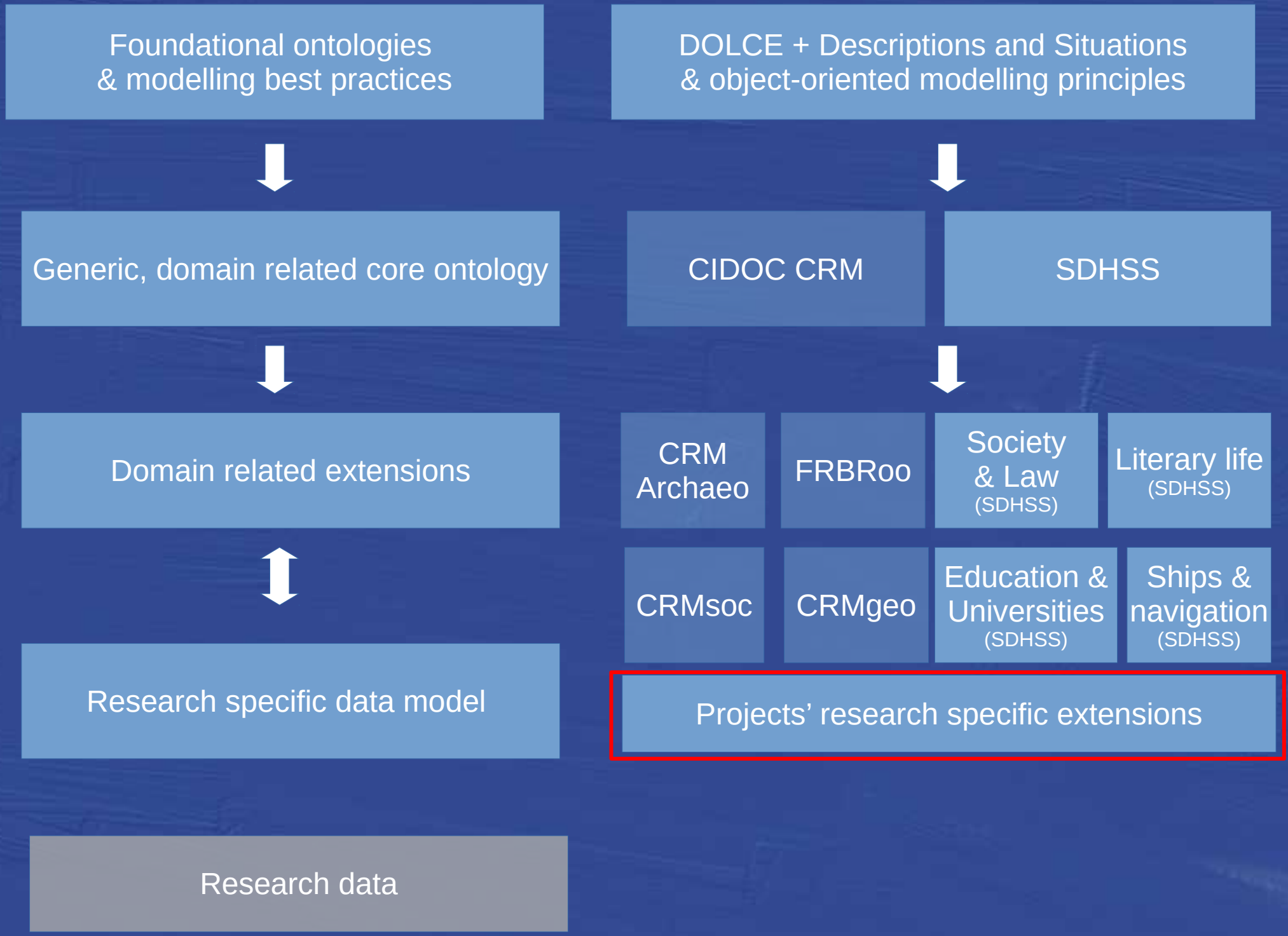
Fig. 2. A UML class diagram for D&S. The lower part of the pattern (within the grey package) is called the *ground ontology*, the higher is called the *descriptive ontology*; a situation satisfies a description if the two parts match according to the axioms specified for the concepts defined by the description.

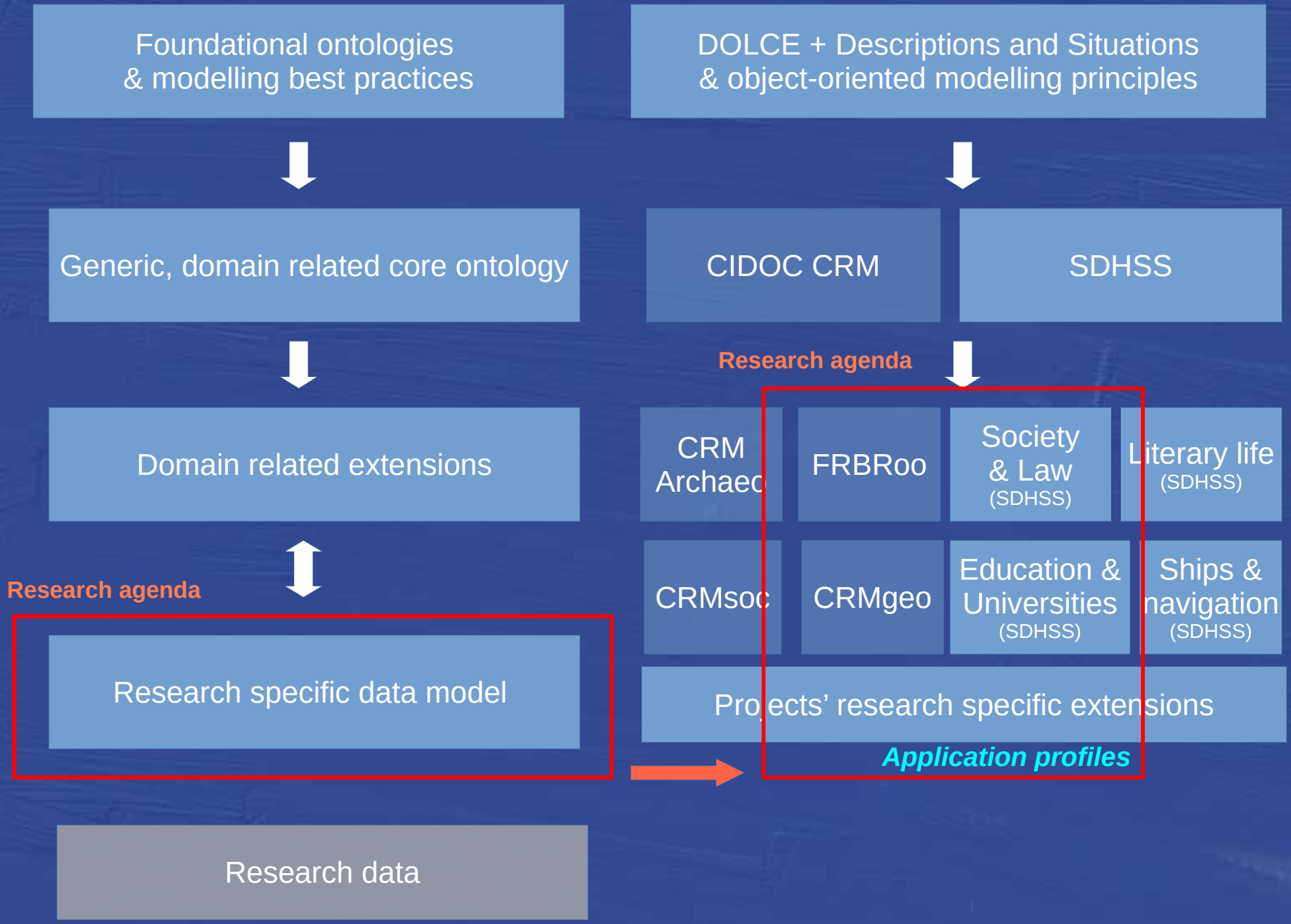
Bottazzi E., Catenacci C., Gangemi A., Lehmann J.(2006) (from pre-print, not in published version)

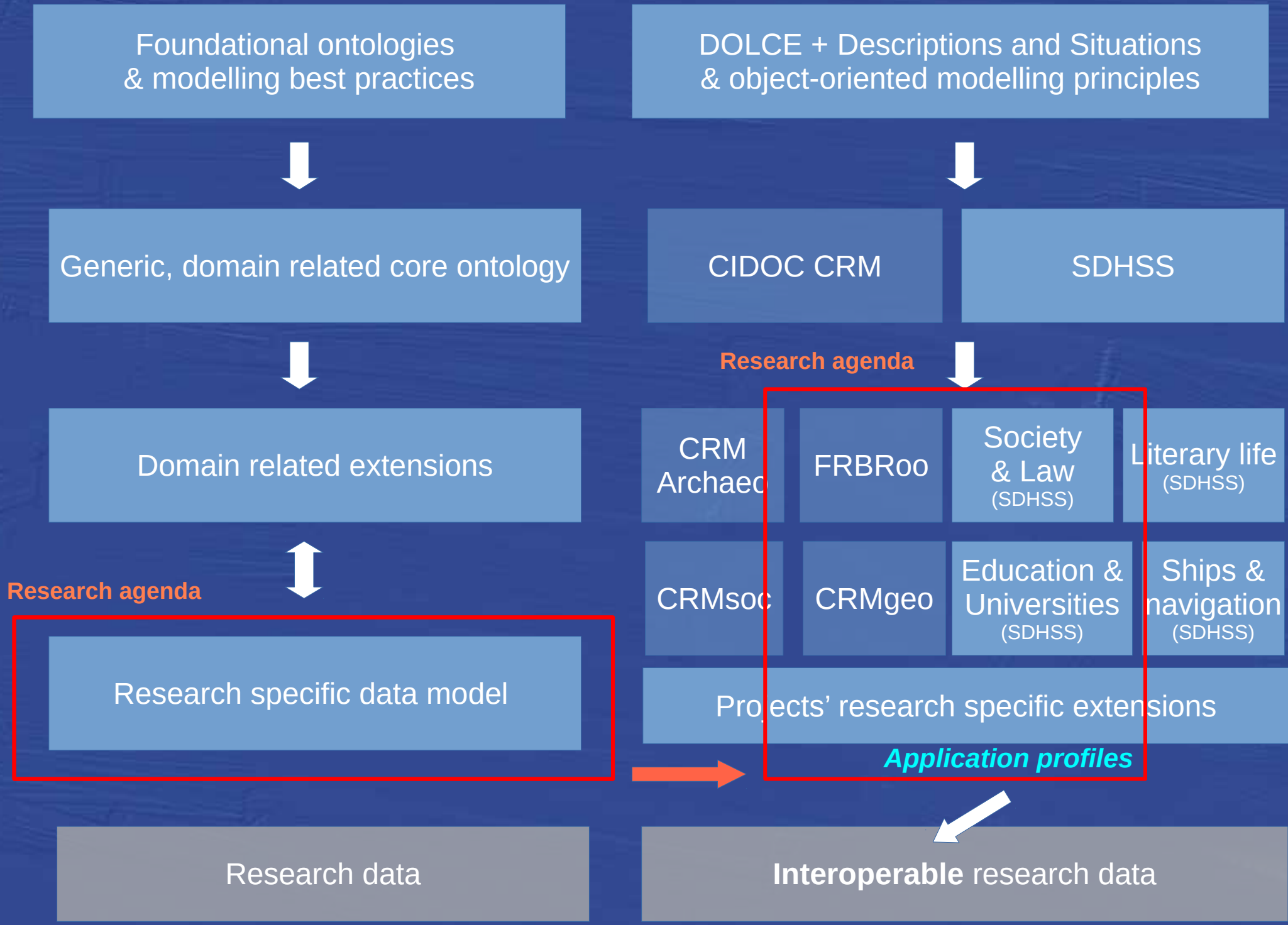

Foundational ontologies
& modelling best practices
DOLCE + Descriptions and Situations
& object-oriented modelling principles
Generic, domain related core ontology
CIDOC CRM
SDHSS
Domain related extensions
Research specific data model
Research data


Foundational ontologies
& modelling best practices
DOLCE + Descriptions and Situations
& object-oriented modelling principles
Generic, domain related core ontology
CIDOC CRM
SDHSS
Domain related extensions
CRM
Archaeo
FRBRoo
CRMsoc
CRMgeo
Research specific data model
Research data


Foundational ontologies
& modelling best practices
DOLCE + Descriptions and Situations
& object-oriented modelling principles
Generic, domain related core ontology
CIDOC CRM
SDHSS
Domain related extensions
CRM Archaeo
FRBRoo
Society & Law (SDHSS)
Literary life (SDHSS)
CRMsoc
CRMgeo
Education & Universities (SDHSS)
Ships & navigation (SDHSS)
Research specific data model
Research data


Foundational ontologies
& modelling best practices
DOLCE + Descriptions and Situations
& object-oriented modelling principles
Generic, domain related core ontology
CIDOC CRM
SDHSS
Domain related extensions
CRM Archaeo
FRBRoo
Society & Law (SDHSS)
Literary life (SDHSS)
CRMsoc
CRMgeo
Education & Universities (SDHSS)
Ships & navigation (SDHSS)
Research specific data model
Projects' research specific extensions
Research data


Foundational ontologies
& modelling best practices
DOLCE + Descriptions and Situations
& object-oriented modelling principles
Generic, domain related core ontology
CIDOC CRM
SDHSS
Research agenda
Domain related extensions
CRM Archaeo
FRBRoo
Society & Law (SDHSS)
Literary life (SDHSS)
Research agenda
CRMsoc
CRMgeo
Education & Universities (SDHSS)
Ships & navigation (SDHSS)
Research specific data model
Projects' research specific extensions
Application profiles
Research data

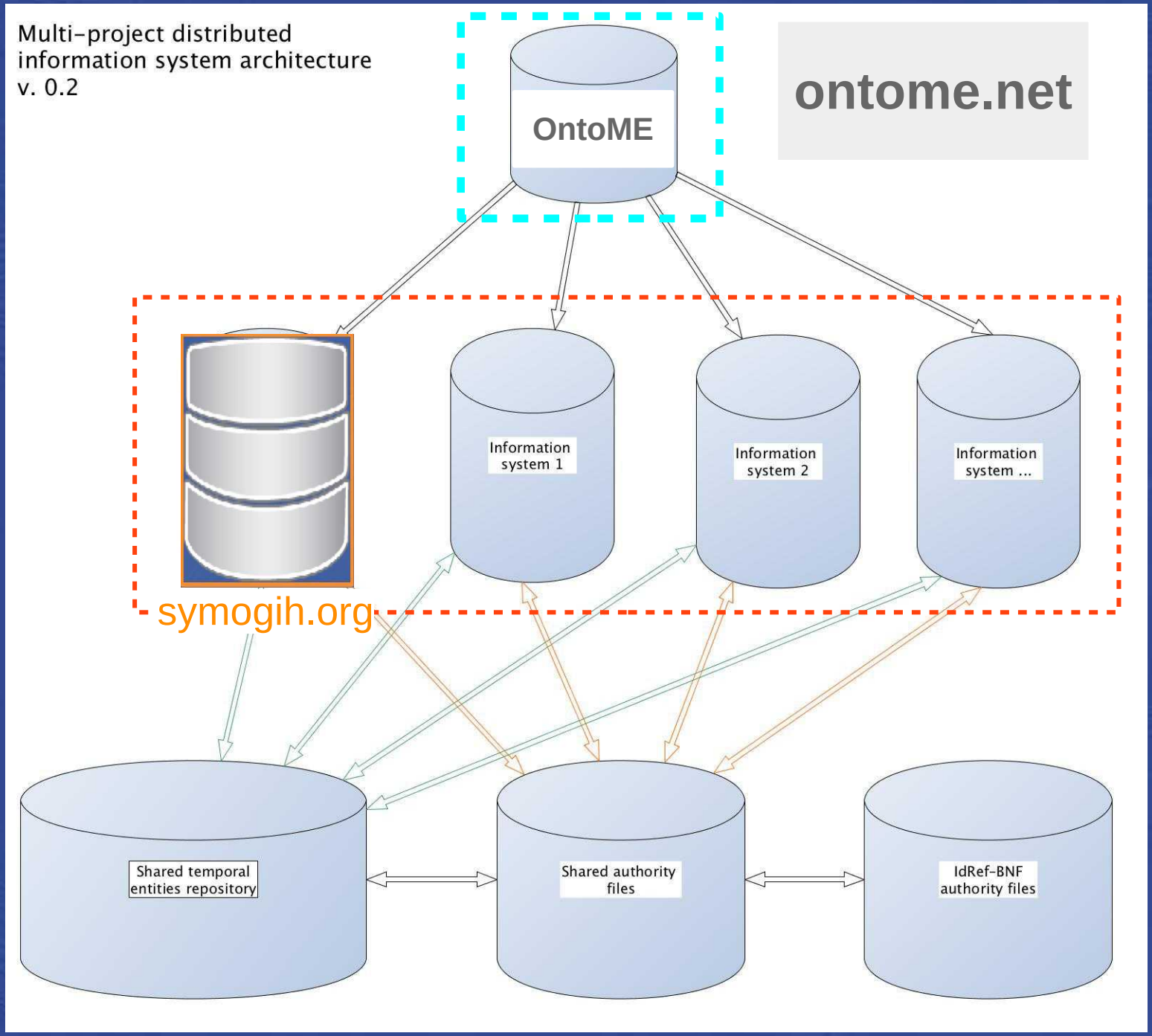

Foundational ontologies
& modelling best practices
DOLCE + Descriptions and Situations
& object-oriented modelling principles
Generic, domain related core ontology
CIDOC CRM
SDHSS
Research agenda
Domain related extensions
CRM Archaeo
FRBRoo
Society & Law (SDHSS)
Literary life (SDHSS)
Research agenda
CRMsoc
CRMgeo
Education & Universities (SDHSS)
Ships & navigation (SDHSS)
Research specific data model
Projects' research specific extensions
Application profiles
Research data
Interoperable research data

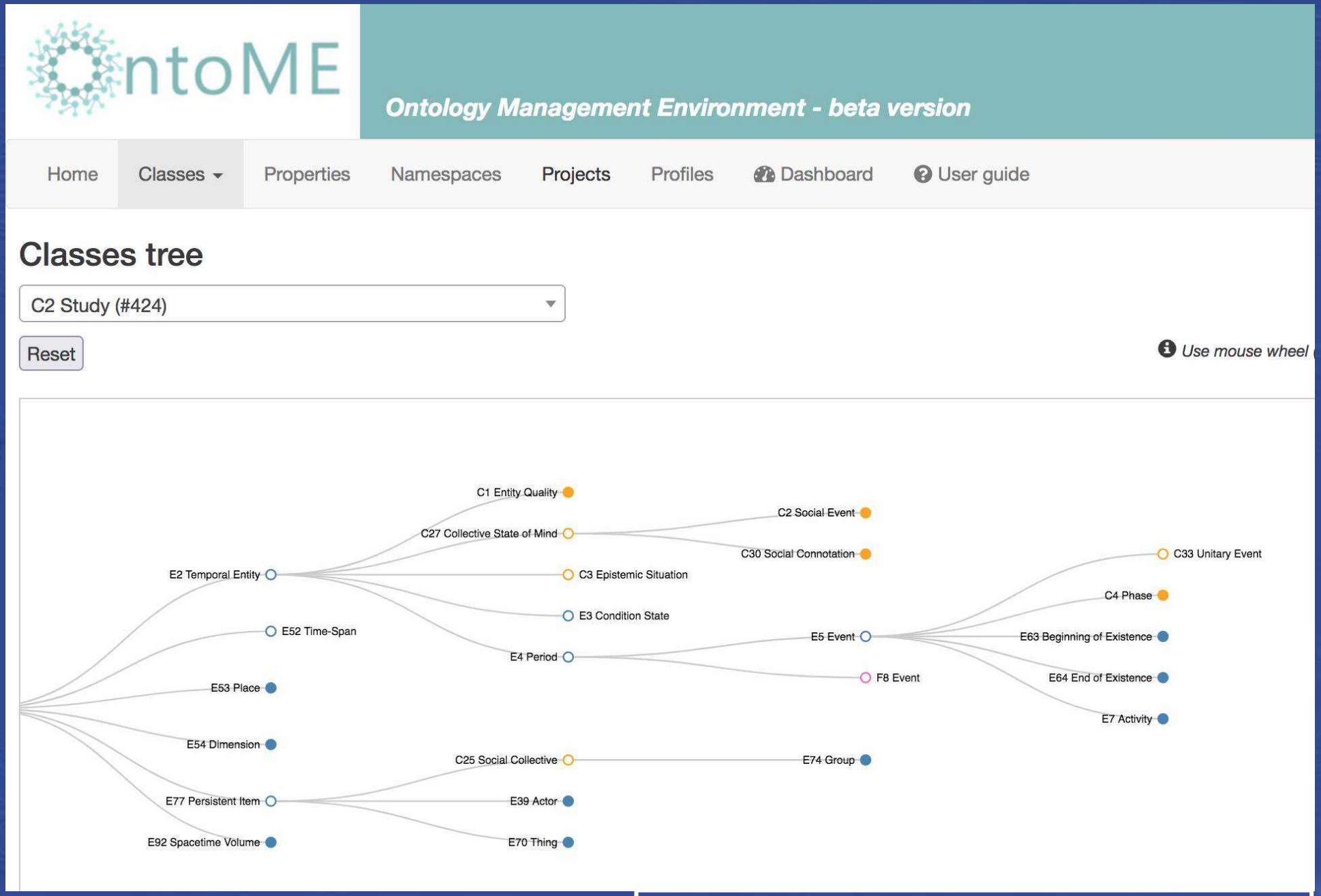
3.

Developing extensions devoted to specific sub-domains

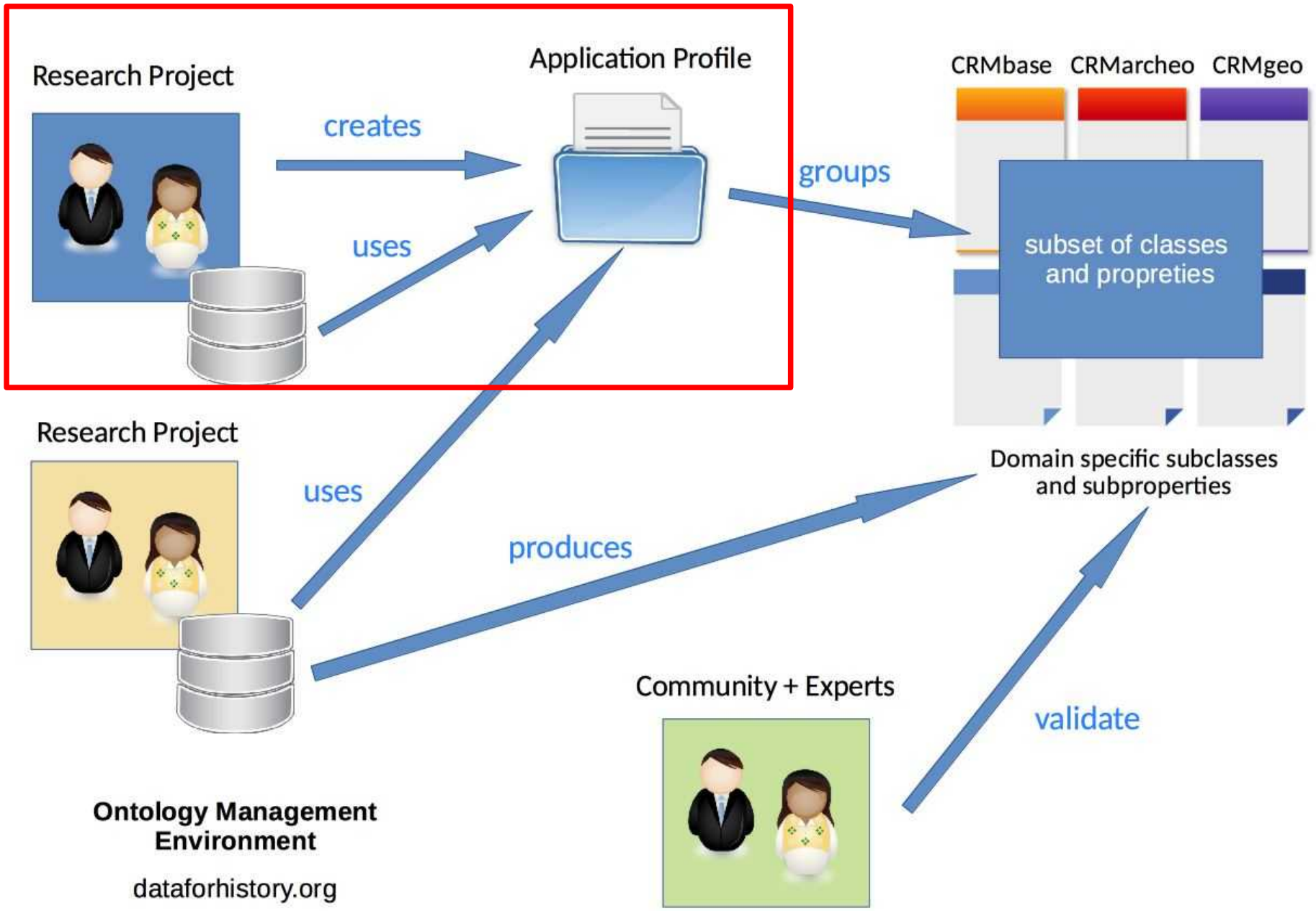
and managing a communal ontology development

ontome.net


Multi-project distributed
information system architecture
v. 0.2
OntoME
ontome.net
Information
system 1
Information
system 2
Information
system ...
symogih.org
Shared temporal
entities repository
Shared authority
files
IdRef-BNF
authority files


OntoME
Ontology Management Environment - beta version
Home
Classes
Properties
Namespaces
Projects
Profiles
Dashboard
User guide
Classes tree
C2 Study (#424)
Reset
Use mouse wheel
C1 Entity Quality
C27 Collective State of Mind
C2 Social Event
C30 Social Connotation
C33 Unitary Event
E2 Temporal Entity
C3 Epistemic Situation
C4 Phase
E3 Condition State
E52 Time-Span
E5 Event
E63 Beginning of Existence
E4 Period
F8 Event
E64 End of Existence
E53 Place
E7 Activity
E54 Dimension
C25 Social Collective
E74 Group
E77 Persistent Item
E39 Actor
E92 Spacetime Volume
E70 Thing

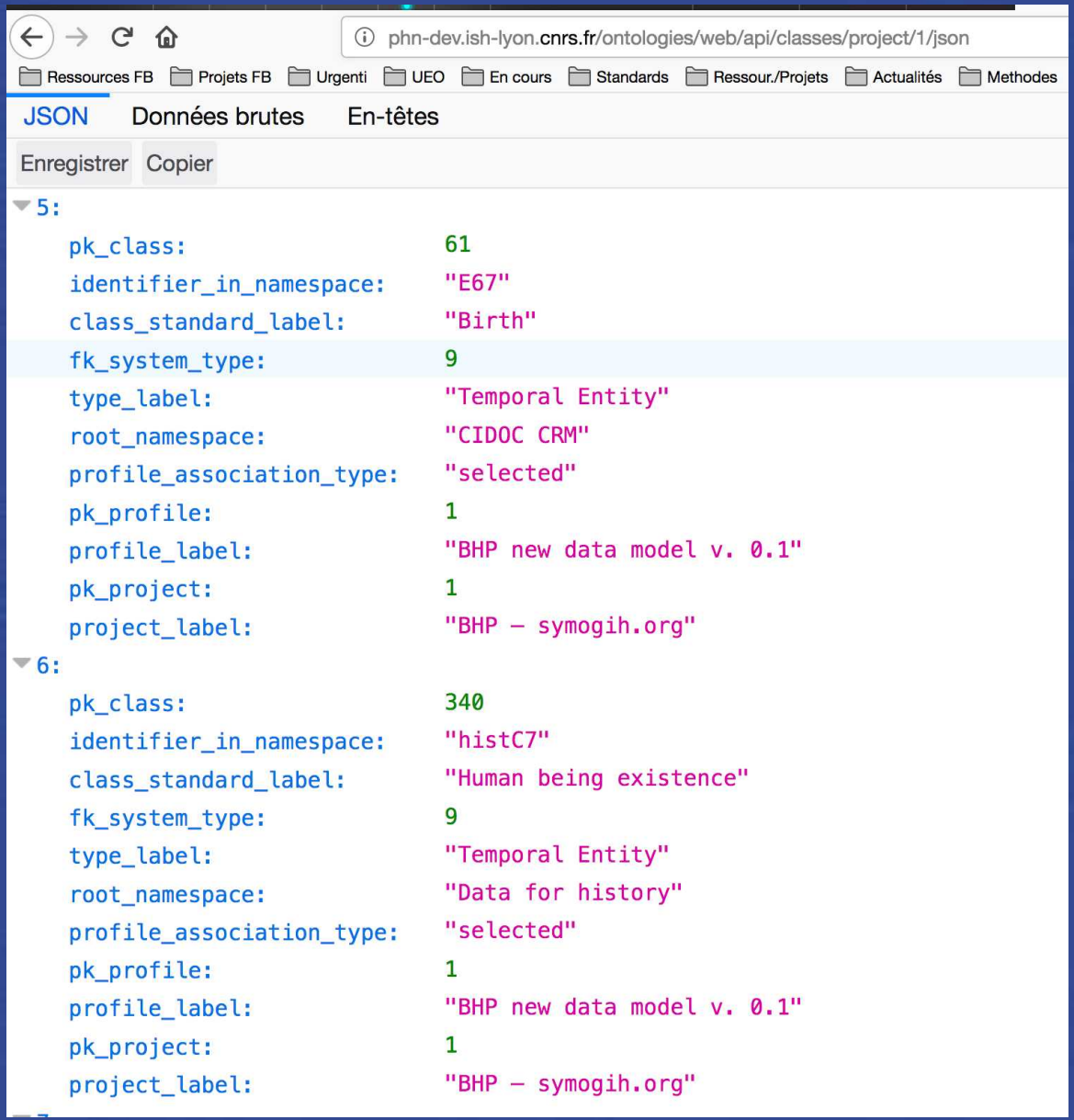
ontome.net

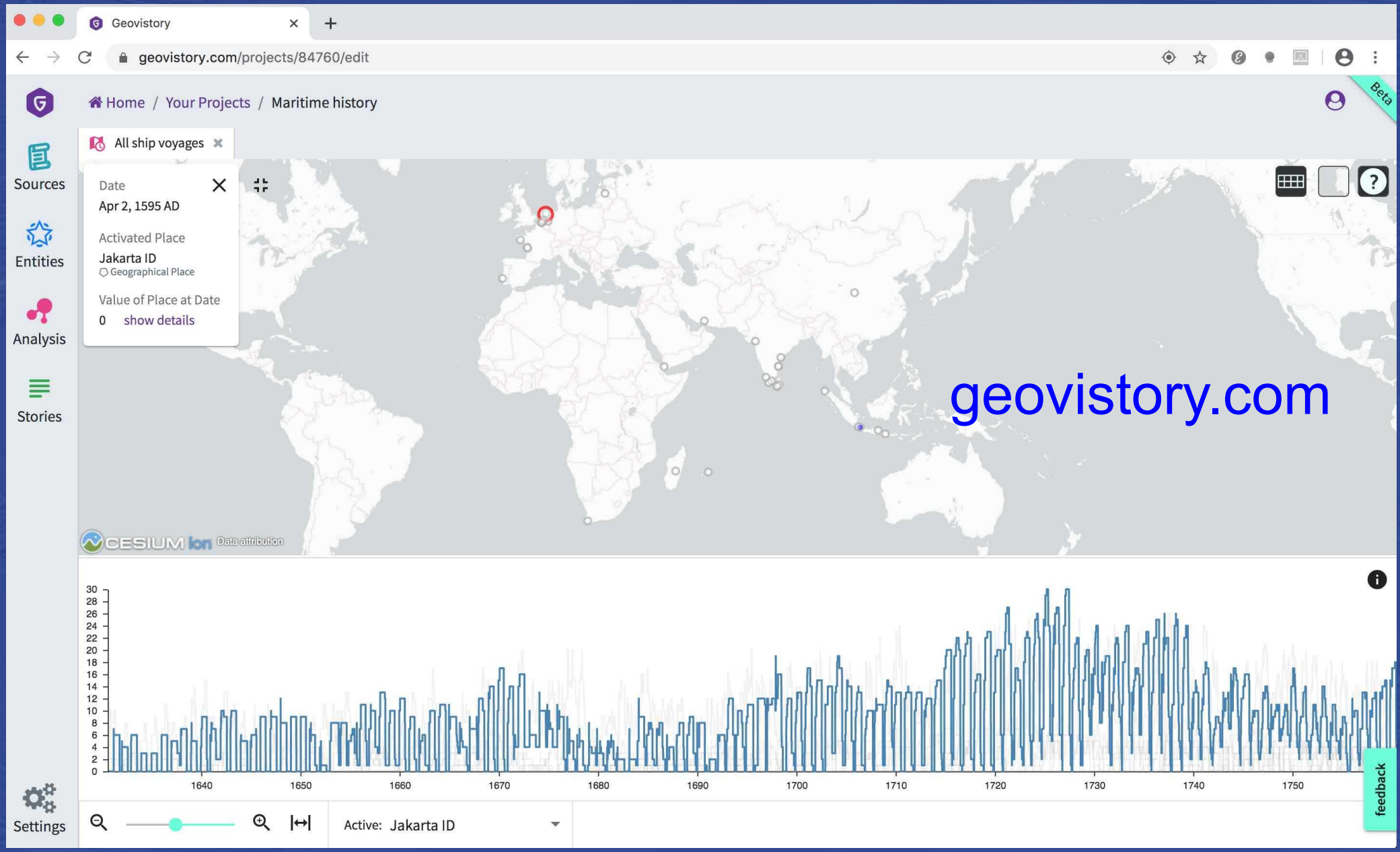

Research Project
Application Profile
creates
uses
groups
CRMbase
CRMarcheo
CRMgeo
subset of classes
and propreties
Research Project
uses
produces
Domain specific subclasses
and subproperties
Community + Experts
validate
Ontology Management
Environment
dataforhistory.org

ontome.net

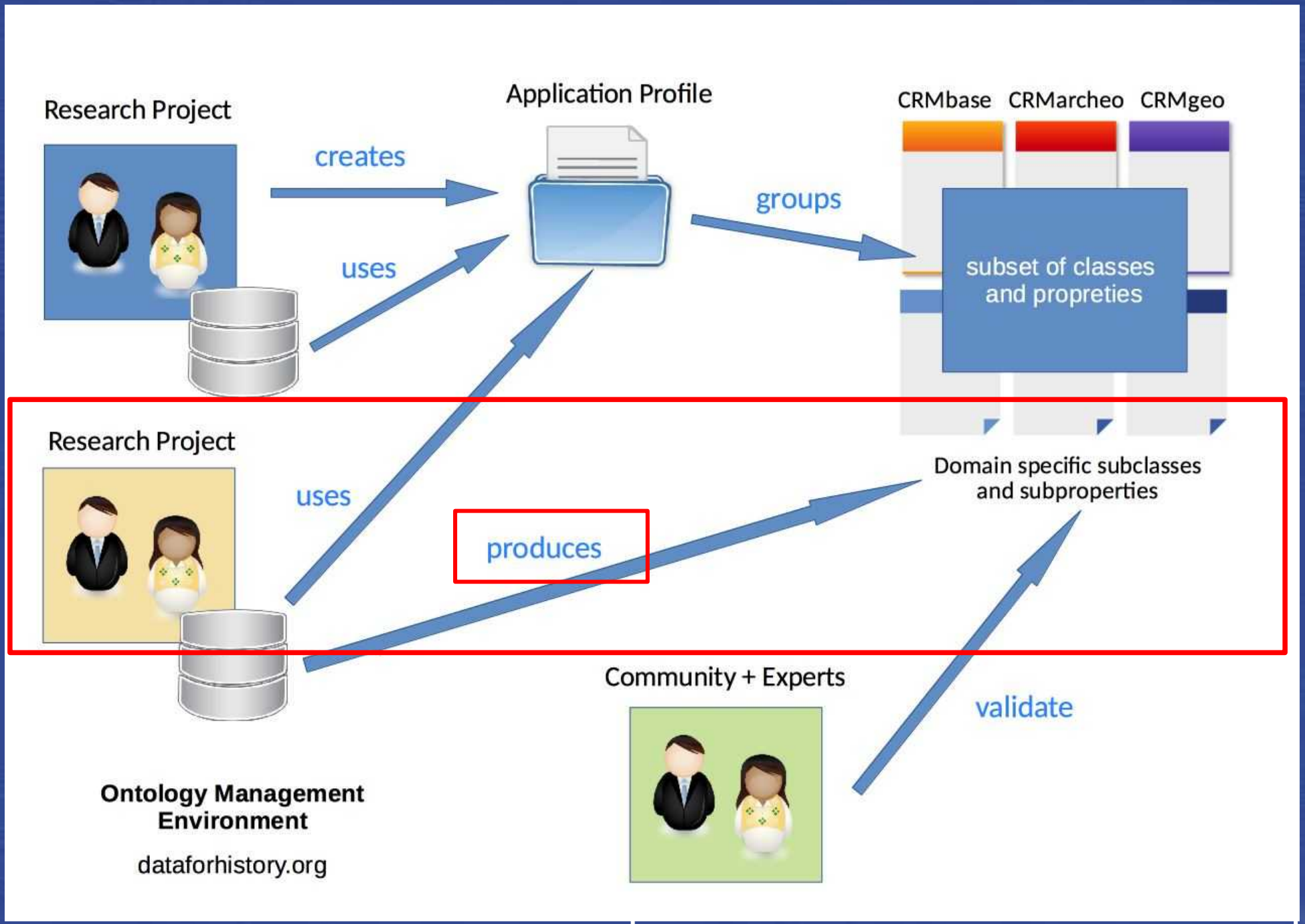
Retrieve your project's application profiles from an API

https://ontome.net/api/classes-profile.json?lang=en&available-in-profile=8

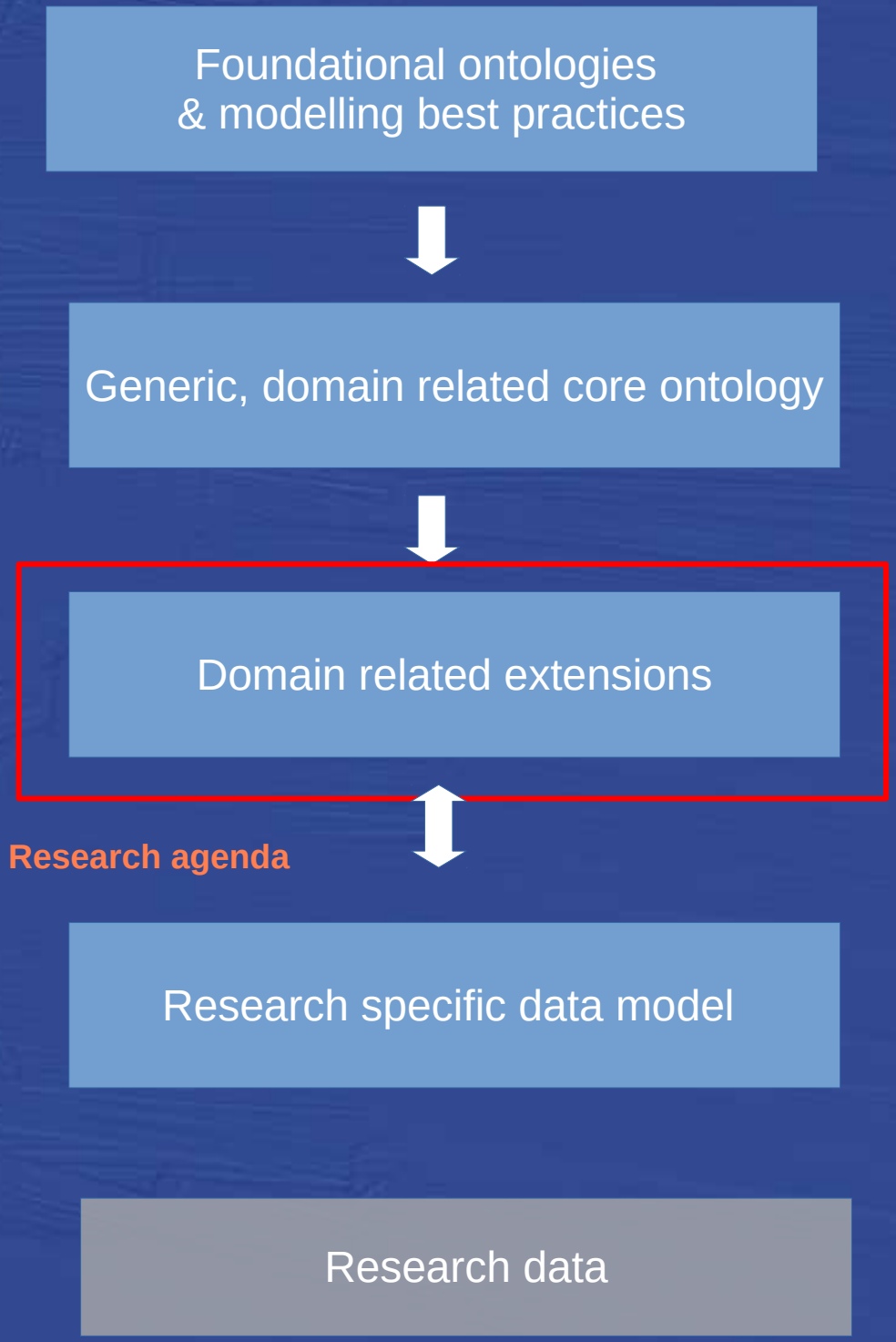


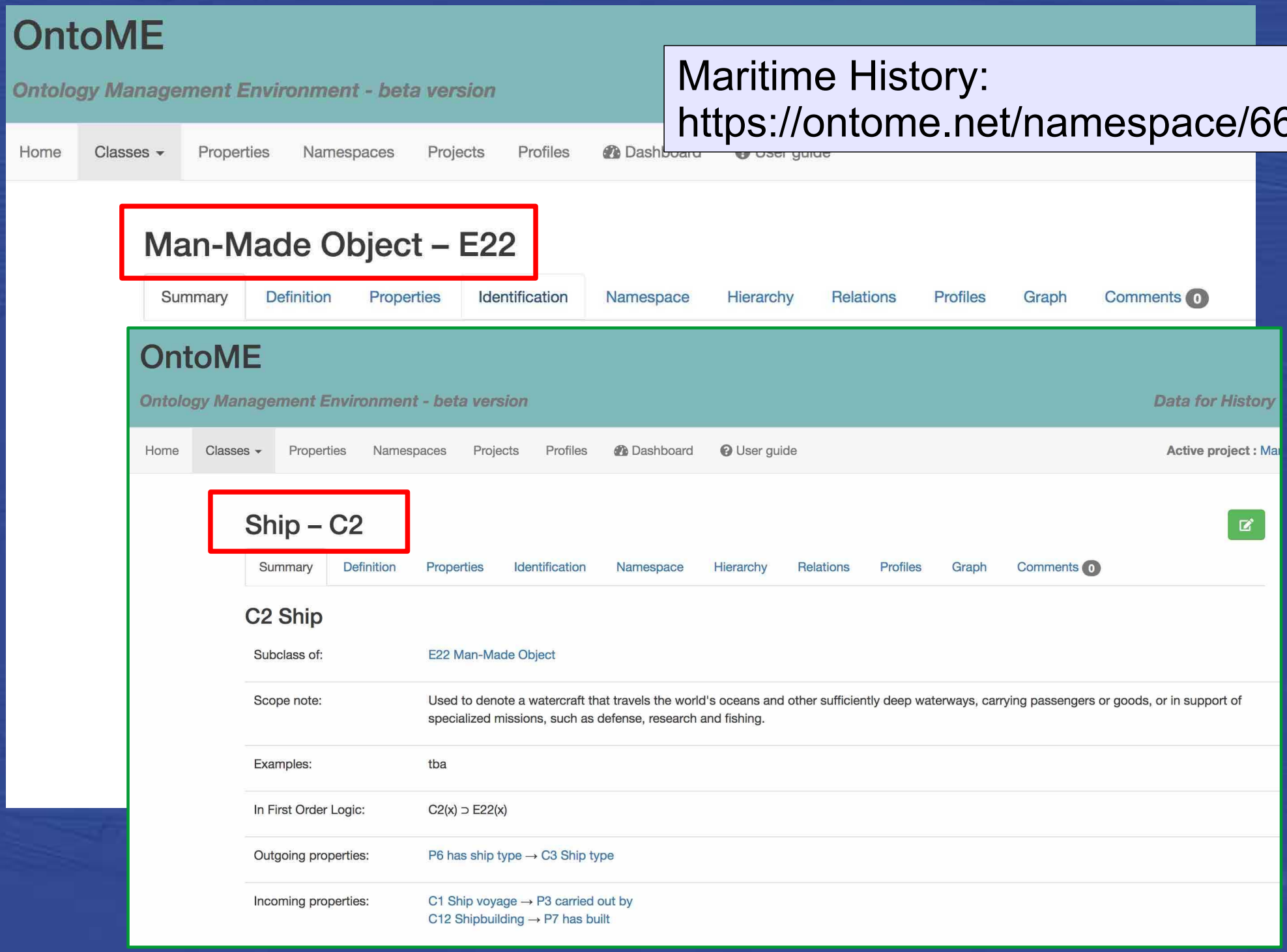


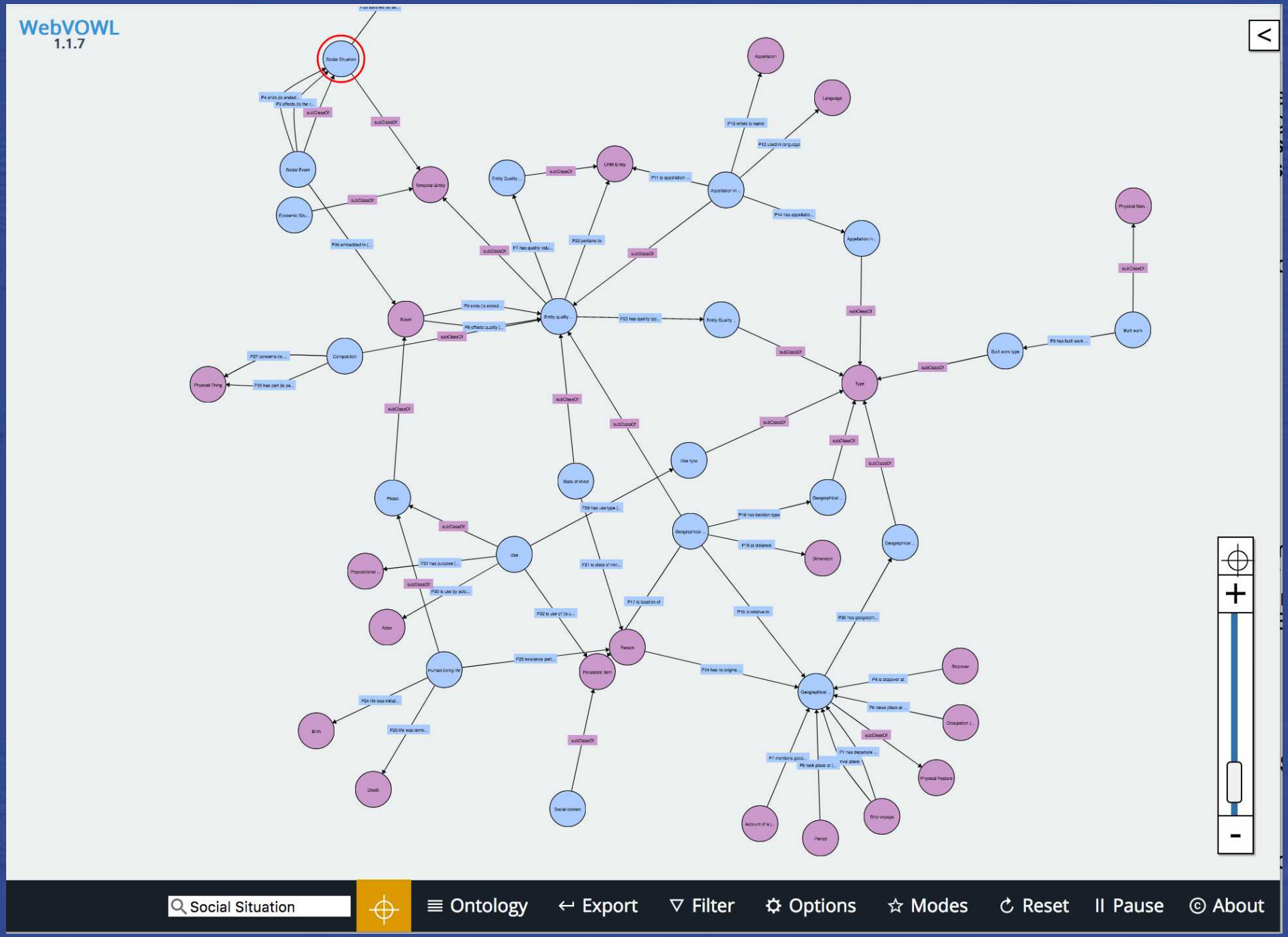
Geovistory : a new VRE for the symogih.org project developed by ***kleiolab.ch*** (Basel)

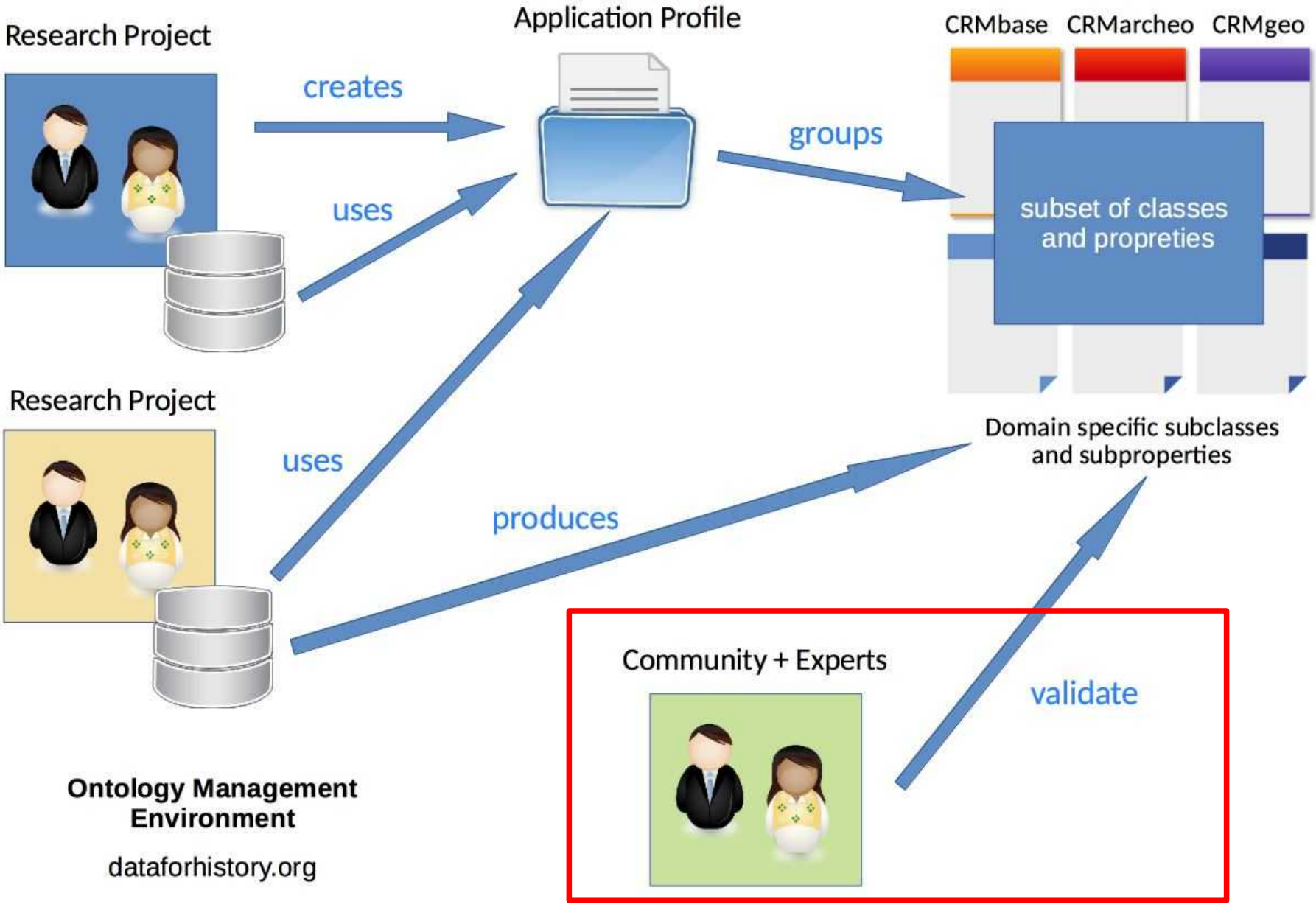

Research Project
creates
Application Profile
groups
CRMbase
CRMarcheo
CRMgeo
uses
subset of classes
and propreties
Research Project
uses
produces
Domain specific subclasses
and subproperties
Community + Experts
validate
Ontology Management
Environment
dataforhistory.org

ontome.net

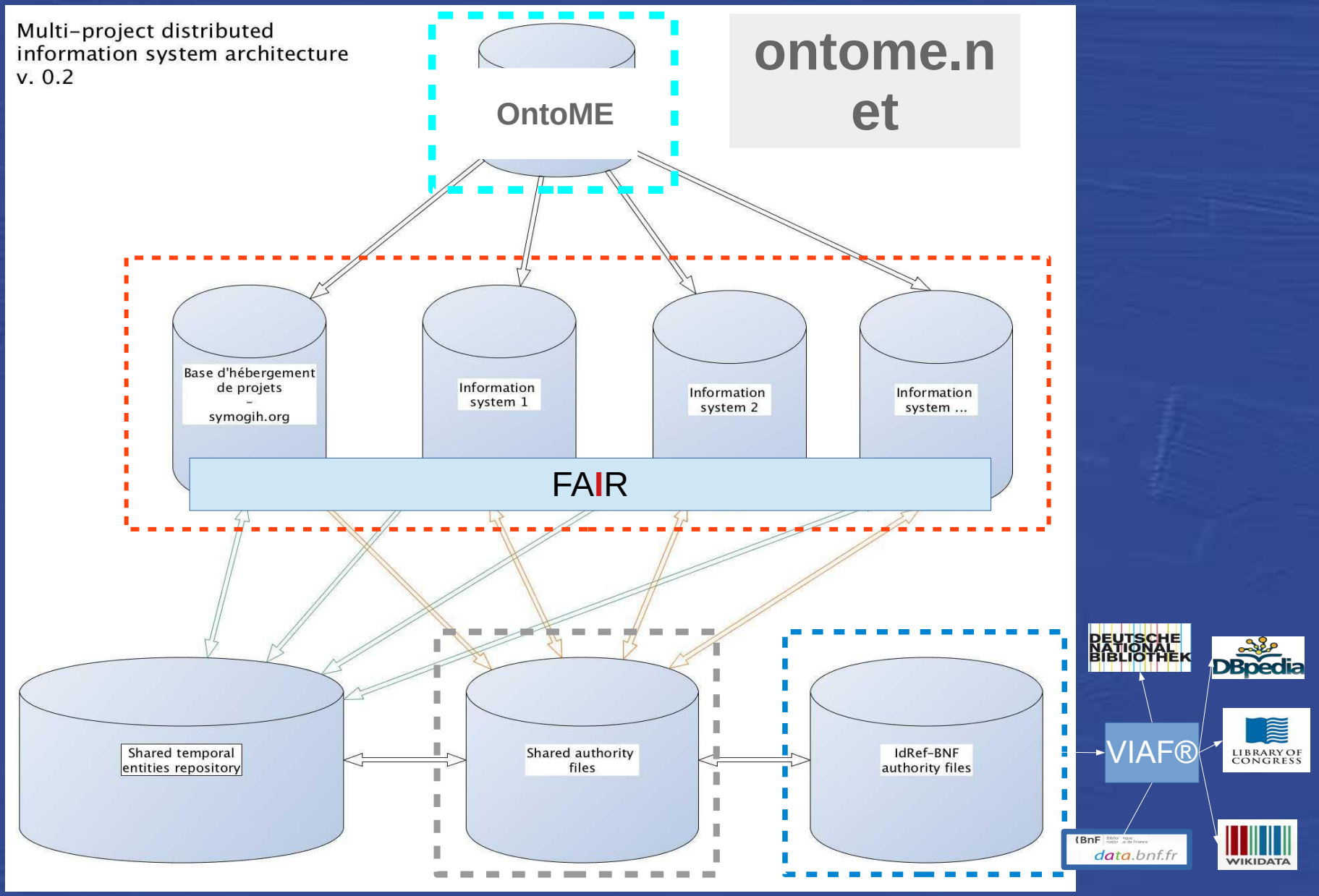

Foundational ontologies
& modelling best practices
Generic, domain related core ontology
Domain related extensions
Research agenda
Research specific data model
Research data


OntoME
Ontology Management Environment - beta version
Maritime History:
https://ontome.net/namespace/66
Home
Classes
Properties
Namespaces
Projects
Profiles
Man-Made Object – E22
Summary
Definition
Properties
Identification
Namespace
Hierarchy
Relations
Profiles
Graph
Comments 0
OntoME
Ontology Management Environment - beta version
Data for History
Home
Classes
Properties
Namespaces
Projects
Profiles
Dashboard
User guide
Active project :
Ship – C2
Summary
Definition
Properties
Identification
Namespace
Hierarchy
Relations
Profiles
Graph
Comments 0
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) ⊃ 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


WebVOWL
1.1.7
Social Situation
Appellation
Language
Social Event
Temporal Entity
Entity Quality
CRM Entity
Appellation in ...
Physical Mark...
Event
Entity quality
Built work
Composition
Physical Thing
Type
Built work type
Phase
Use
State of mind
Use type
Geographical
Dimension
Propositional ...
Actor
Person
Human being life
Persistent Item
Stopover
Occupation (...
Physical Feature
Birth
Death
Social context
Ship voyage
Period
Social Situation
Ontology
Export
Filter
Options
Modes
Reset
Pause
About


Research Project
creates
Application Profile
CRMbase
CRMarcheo
CRMgeo
groups
uses
subset of classes
and propreties
Research Project
uses
Domain specific subclasses
and subproperties
produces
Community + Experts
validate
Ontology Management
Environment
dataforhistory.org

ontome.net



dataforhistory.org – dataforhumanities.org

ontome.net – what we offer

- Contact us on :
  - https://colibris.link/ontome
  - contact@ontome.net
- OntoME trainings
- Semantic engineering workshops :
  align your research data models with the community driven CIDOC CRM – SDHSS extensions
- Join the OntoME project and developers community

dataforhistory.org – dataforhumanities.org