Francesco Beretta
(CNRS UMR5190 LARHRA – Université de Lyon)

Interoperability of historical data and FAIR principles: an ontology management environment (OntoME) for sharing and aligning data models

ontome.dataforhistory.org

EADH 2018 Conference
National University of Ireland,
Galway, 9 December 2018

CC BY-NC-SA 4.0

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.

CC BY-NC-SA 4.0

The FAIR data principles

To be **Findable**:

F1. (meta)data are assigned a globally unique and eternally *persistent identifier*.
F2. data are described with rich *metadata*.
F3. (meta)data are registered or indexed in a *searchable resource*.
F4. metadata specify the data identifier.

To be **Accessible**:

A1 (meta)data are retrievable by their identifier using a *standardized communications protocol*.
A1.1 the protocol is open, free, and universally implementable.
A1.2 the protocol allows for an authentication and authorization procedure, where necessary.
A2 *metadata are accessible*, even when the data are no longer available.

CC BY-NC-SA 4.0

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*.

CC BY-NC-SA 4.0

Requirements for structured data interoperability :

CC BY-NC-SA 4.0

Requirements for structured data interoperability :

* Share and interlink dictionaries of identified entities (authority files or gazetteers) : people, places, organisations, books, …

CC BY-NC-SA 4.0

« In 1592, he [Galileo Galilei] moved to the University of Padua where he taught [mathematics] until 1610»

[https://en.wikipedia.org/wiki/Galileo_Galilei#Reference-Sharratt-1994]

CC BY-NC-SA 4.0

« In 1592, he [Galileo Galilei] moved to the University of Padua where he taught [mathematics] until 1610»

[https://en.wikipedia.org/wiki/Galileo_Galilei#Reference-Sharratt-1994]

Record linkage : is my Galilei your Galilei ?

CC BY-NC-SA 4.0

« In 1592, he [Galileo Galilei] moved to the University of Padua where he taught [mathematics] until 1610»

[https://en.wikipedia.org/wiki/Galileo_Galilei#Reference-Sharratt-1994]

Record linkage : is my Galilei your Galilei ?

Galileo Galilei :

- *1564, mathematician and astronomer

CC BY-NC-SA 4.0

« In 1592, he [Galileo Galilei] moved to the University of Padua where he taught [mathematics] until 1610»

[https://en.wikipedia.org/wiki/Galileo_Galilei#Reference-Sharratt-1994]

Record linkage : is my Galilei your Galilei ?

Galileo Galilei :

- *1564, mathematician and astronomer
- *1370, medicine lecturer at Florence *studio*

CC BY-NC-SA 4.0

« In 1592, he [Galileo Galilei] moved to the University of Padua where he taught [mathematics] until 1610»

[https://en.wikipedia.org/wiki/Galileo_Galilei#Reference-Sharratt-1994]

Record linkage : is my Galilei your Galilei ?

Galileo Galilei : **http://symogih.org/resource/Actr161**

- *1564, mathematician and astronomer
- *1370, medicine lecturer at Florence *studio*

http://symogih.org/resource/Actr646

CC BY-NC-SA 4.0

« In 1592, he [Galileo Galilei] moved to the University of Padua where he taught [mathematics] until 1610»

[https://en.wikipedia.org/wiki/Galileo_Galilei#Reference-Sharratt-1994]

Record linkage : is my Galilei your Galilei ?

Galileo Galilei :

- *1564, mathematician and astronomer
- *1370, medicine lecturer at Florence *studio*

Who was Johannes Teutonicus ?

CC BY-NC-SA 4.0

« In 1592, he [Galileo Galilei] moved to the University of Padua where he taught [mathematics] until 1610»

[https://en.wikipedia.org/wiki/Galileo_Galilei#Reference-Sharratt-1994]

Record linkage : is my Galilei your Galilei ?

Galileo Galilei :

- *1564, mathematician and astronomer
- *1370, medicine lecturer at Florence *studio*

Who was Johannes Teutonicus ?

What about places, organizations, books ?

CC BY-NC-SA 4.0

Requirements for structured data interoperability :

* Share and interlink dictionaries of identified entities (authority files or gazetteers) : people, places, organisations, books, …

* Share controlled vocabulaires (thesauri and taxonomies of ‘types’) that are used to classify entities (and provide their identity)

CC BY-NC-SA 4.0

Galilei → Mathematician and astronomer

CC BY-NC-SA 4.0

Galilei → Mathematician and astronomer

Mathematicians :

- Louis de Broglie (1892-1987), French mathematician and physicist, Nobel Prize for Physics in 1929

CC BY-NC-SA 4.0

Galilei → Mathematician and astronomer

Mathematicians :

- Louis de Broglie (1892-1987), French mathematician and physicist, Nobel Prize for Physics in 1929
- Stephen Smale (*1930), American mathematician. His research concerns topology, dynamical systems and mathematical economics.

CC BY-NC-SA 4.0

Galilei → Mathematician and astronomer

Mathematicians :

- Louis de Broglie (1892-1987), French mathematician and physicist, Nobel Prize for Physics in 1929
- Stephen Smale (*1930), American mathematician. His research concerns topology, dynamical systems and mathematical economics.

Mathematics (17th century)

CC BY-NC-SA 4.0

Galilei → Mathematician and astronomer

Mathematicians :

- Louis de Broglie (1892-1987), French mathematician and physicist, Nobel Prize for Physics in 1929
- Stephen Smale (*1930), American mathematician. His research concerns topology, dynamical systems and mathematical economics.

Mathematics (17th century) Mathematics (20th century)

CC BY-NC-SA 4.0

Galilei → Mathematician and astronomer

Mathematicians :

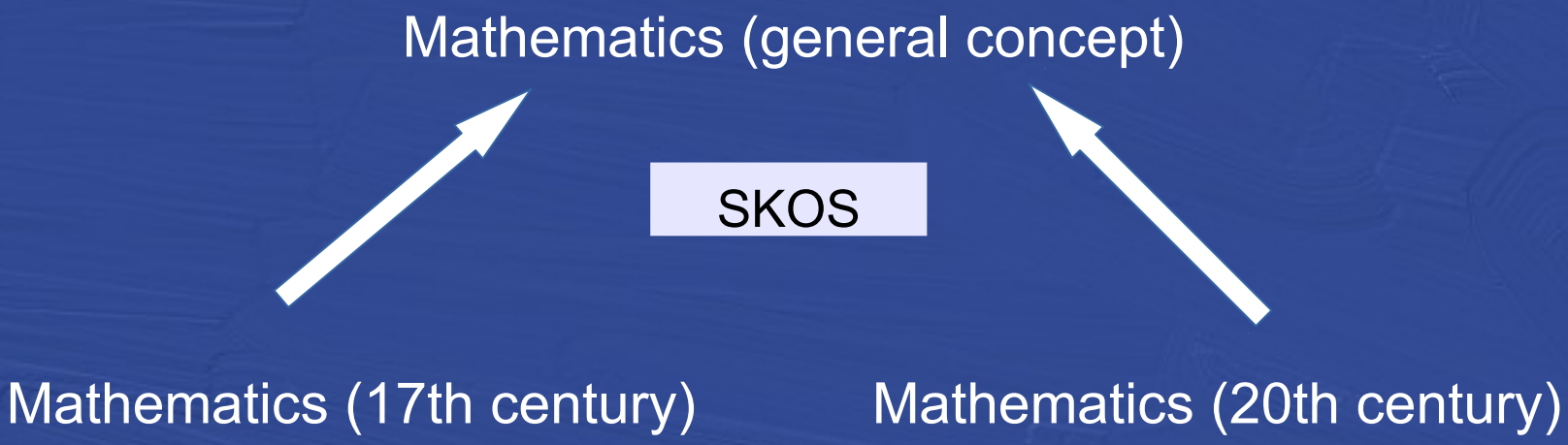
- Louis de Broglie (1892-1987), French mathematician and physicist, Nobel Prize for Physics in 1929
- Stephen Smale (*1930), American mathematician. His research concerns topology, dynamical systems and mathematical economics.

CC BY-NC-SA 4.0

Galilei → Mathematician and astronomer

Mathematicians :

- Louis de Broglie (1892-1987), French mathematician and physicist, Nobel Prize for Physics in 1929
- Stephen Smale (*1930), American mathematician. His research concerns topology, dynamical systems and mathematical economics.



CC BY-NC-SA 4.0

Requirements for structured data interoperability :

* Share and interlink dictionaries of identified entities (authority files or gazetteers) : people, places, organisations, books, …

* Share controlled vocabulaires (thesauri and taxonomies of ‘types’ that are used to classify entities (and provide their identity)

* Share information about entities :
  a common, multi-tiered and extensible data model, i.e.
  an ontology as a shared conceptualization of the ‘world’ and its geo-historical development (cultural heritage domain)

CC BY-NC-SA 4.0

« In 1592, he [Galileo Galilei] moved to the University of Padua where he taught [mathematics] until 1610»

[https://en.wikipedia.org/wiki/Galileo_Galilei#Reference-Sharratt-1994]

CC BY-NC-SA 4.0

« In 1592, he [Galileo Galilei] moved to the University of Padua where he taught [mathematics] until 1610»

[https://en.wikipedia.org/wiki/Galileo_Galilei#Reference-Sharratt-1994]

Galilei moves to Padua

CC BY-NC-SA 4.0

« In 1592, he [Galileo Galilei] moved to the University of Padua where he taught [mathematics] until 1610»

[https://en.wikipedia.org/wiki/Galileo_Galilei#Reference-Sharratt-1994]

Galilei moves to Padua

Galilei is hired by the University of Padua

CC BY-NC-SA 4.0

« In 1592, he [Galileo Galilei] moved to the University of Padua where he taught [mathematics] until 1610»

[https://en.wikipedia.org/wiki/Galileo_Galilei#Reference-Sharratt-1994]

Galilei moves to Padua

Galilei is hired by the University of Padua

Galilei teaches mathematics...

CC BY-NC-SA 4.0

« In 1592, he [Galileo Galilei] moved to the University of Padua where he taught [mathematics] until 1610»

[https://en.wikipedia.org/wiki/Galileo_Galilei#Reference-Sharratt-1994]

Galilei moves to Padua

Galilei is hired by the University of Padua

Galilei teaches mathematics...

- Galilei has the status of professor of mathematics

CC BY-NC-SA 4.0

« In 1592, he [Galileo Galilei] moved to the University of Padua where he taught [mathematics] until 1610»

[https://en.wikipedia.org/wiki/Galileo_Galilei#Reference-Sharratt-1994]

Galilei moves to Padua

Galilei is hired by the University of Padua

Galilei teaches mathematics...

- Galilei has the status of professor of mathematics
- Galilei teaches mathematics courses but he has a sabbatical year

CC BY-NC-SA 4.0

« In 1592, he [Galileo Galilei] moved to the University of Padua where he taught [mathematics] until 1610»

[https://en.wikipedia.org/wiki/Galileo_Galilei#Reference-Sharratt-1994]

Galilei moves to Padua

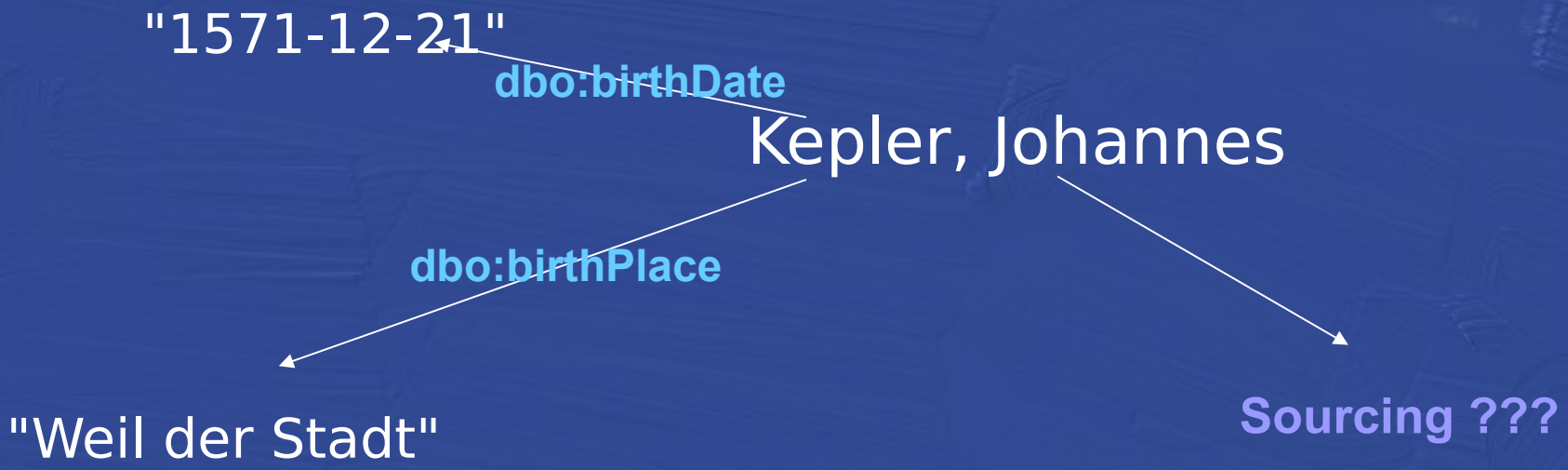
Galilei is hired by the University of Padua

Galilei teaches mathematics...

- Galilei has the status of professor
- Galilei teaches mathematics courses but he has a sabbatical year
- During this semester Galilei regularly teaches mathematics

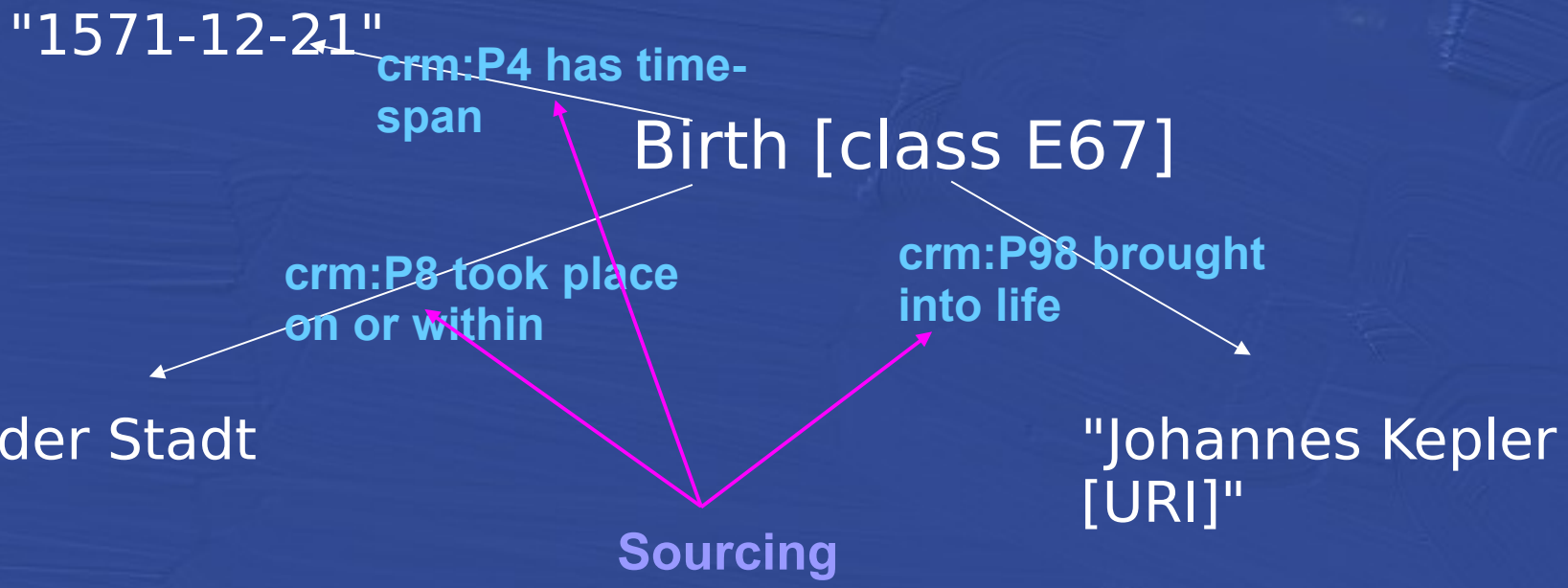
CC BY-NC-SA 4.0

Adopt a resource-centered model...



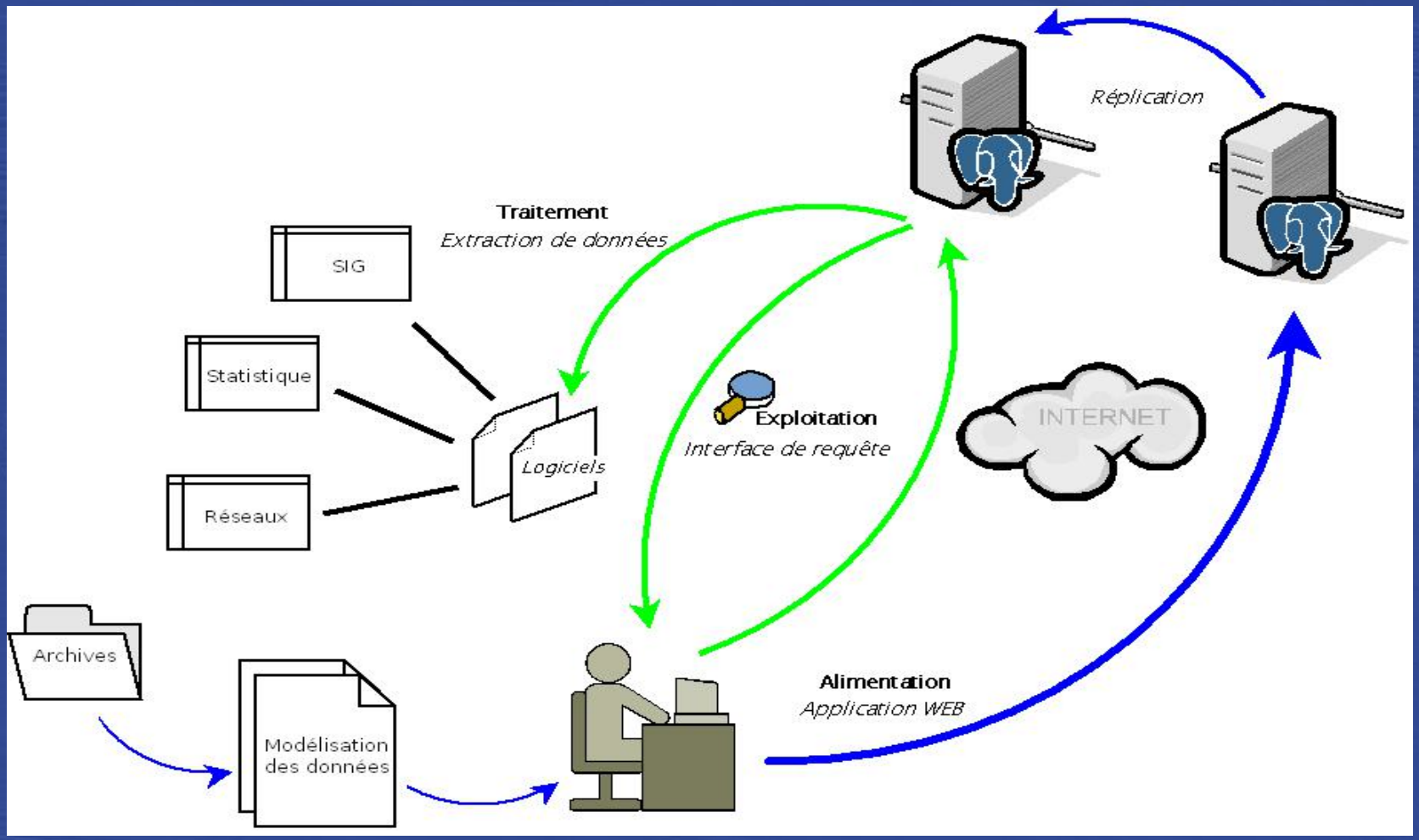
CC BY-NC-SA 4.0

… or prefer a more robust event-centered model !



CC BY-NC-SA 4.0

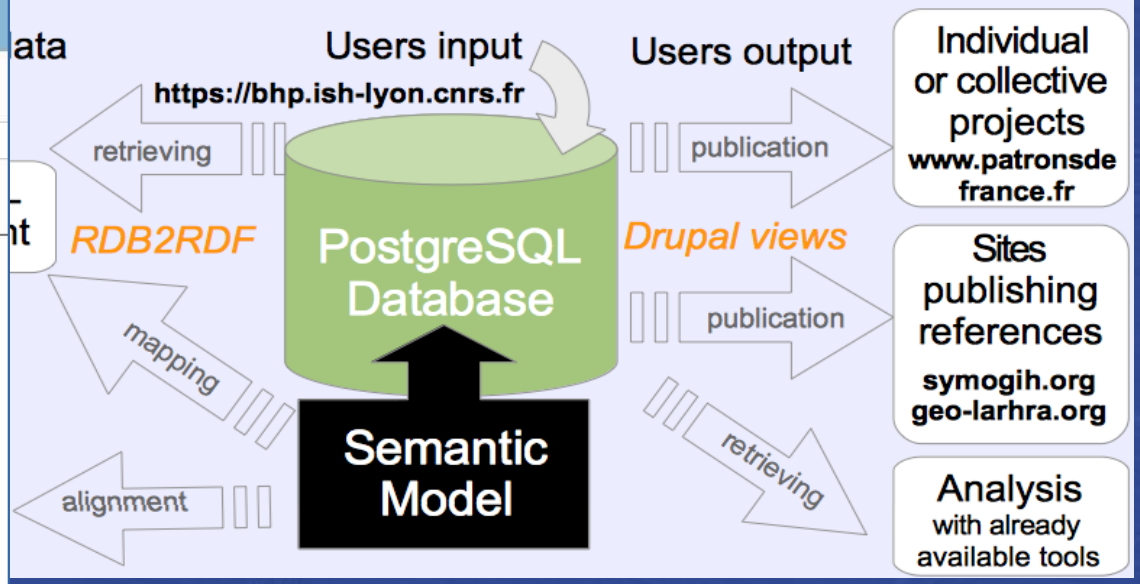
The *symogih.org* project : modular system for historical information management



A collaborative and cumulative information system for storing historical data: produce, share, visualise and analyse stuctured historical data

CC BY-NC-SA 4.0

Open, modular, collaborative platform for storing, analyzing and publishing historical data and texts



Virtuoso SPARQL Query Editor

Default Data Set Name (Graph IRI)

Query Text

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX viaf: <http://viaf.org/viaf/>
PREFIX geo: <http://www.w3.org/2003/01/geo/wgs84_pos#>
PREFIX sym: <http://symogih.org/ontology/>
PREFIX syr: <http://symogih.org/resource/>

SELECT DISTINCT ?i ?stLabel ?stDate ?KUTyLabel ?KUTy
WHERE
{ { ?r sym:associatesObject syr:AbOb213 .
?r sym:isComponentOf ?i .
?i sym:knowledgeUnitStandardLabel ?stLabel .
?i sym:knowledgeUnitStandardDate ?stDate .
?i sym:hasKnowledgeUnitType ?KUTy.
?KUTy rdfs:label ?KUTyLabel.

}
}
ORDER BY ?stDate
```

Sponging: Use only local data (including data retrieved before), but do

Results Format: HTML

Execution timeout: 0 milliseconds *(values less than 100*

Options: Strict checking of void variables

(The result can only be sent back to browser, not saved on the server, see details)

Run Query Reset

SPARQL-endpoint:

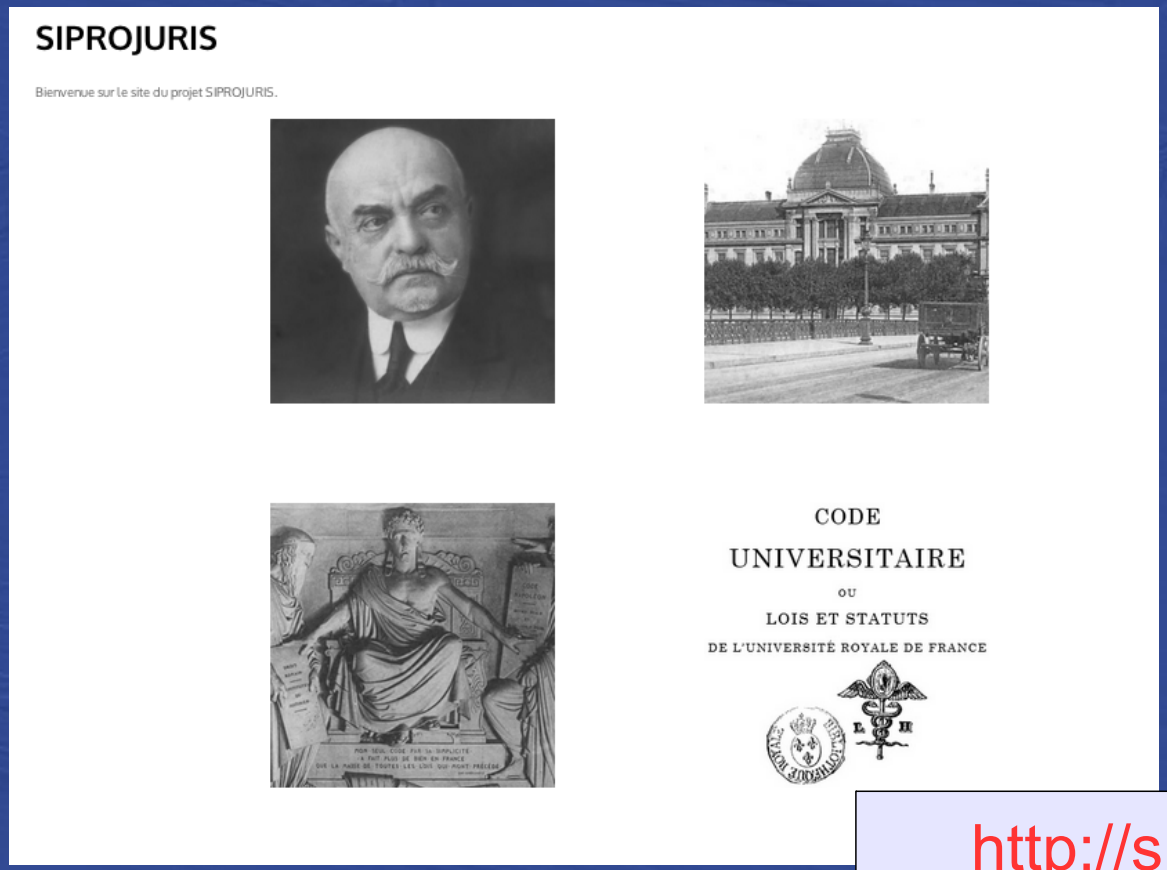
http://symogih.org/?q=rdf-publication

CC BY-NC-SA 4.0

Data challenge MaDICS-ADOC 2018

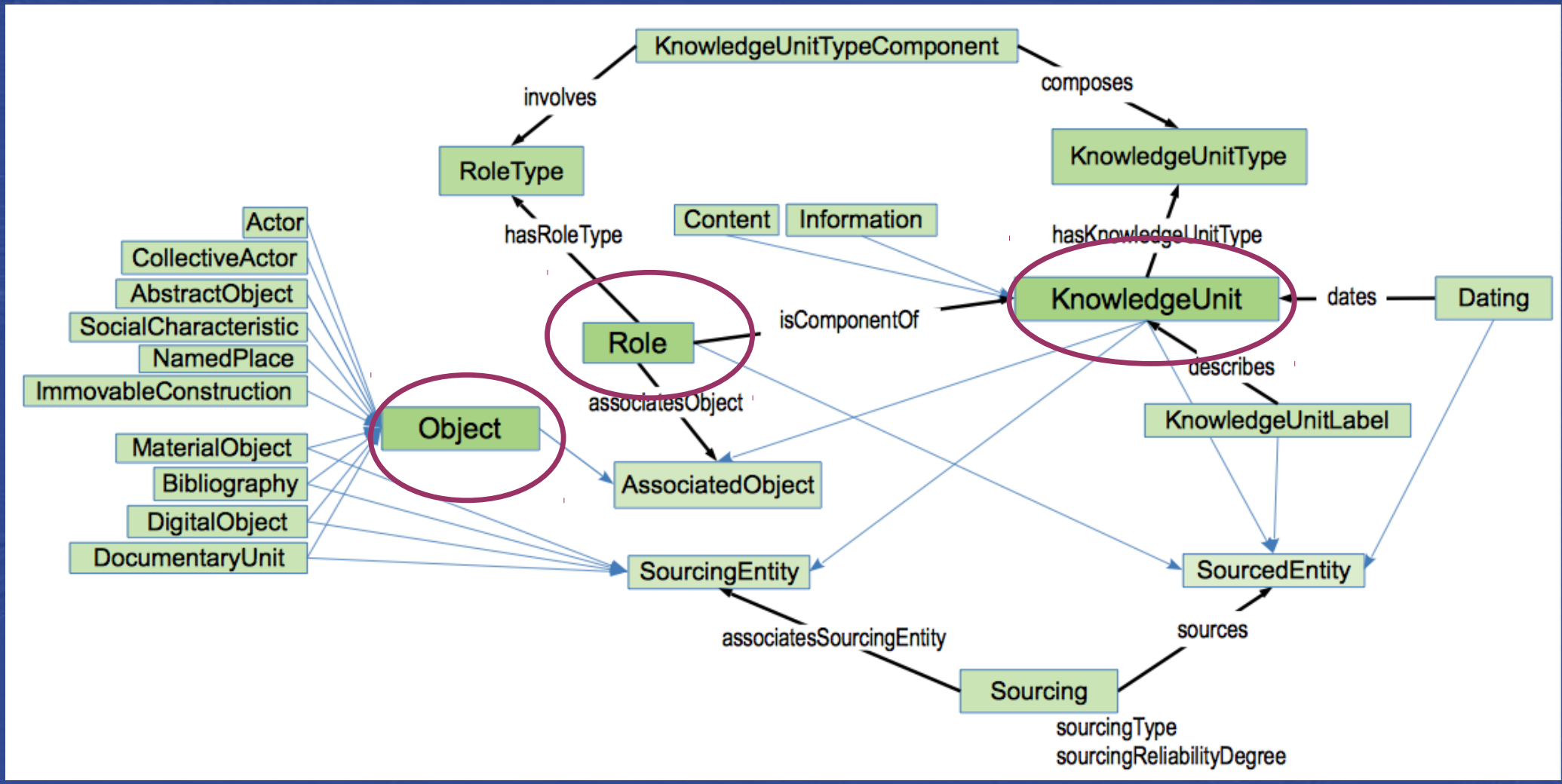
tinyurl.com/data-challenge-2018

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

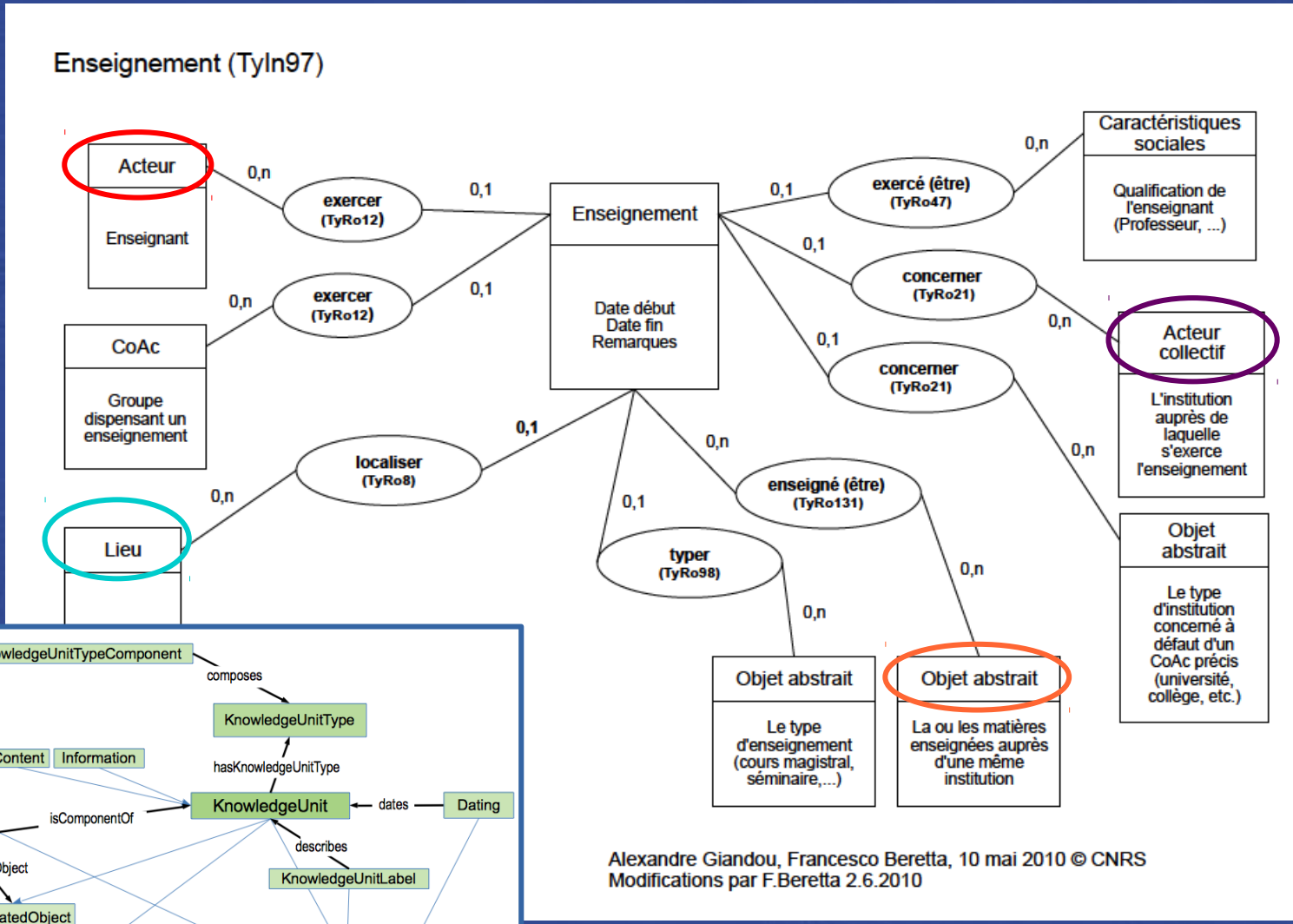
CC BY-NC-SA 4.0

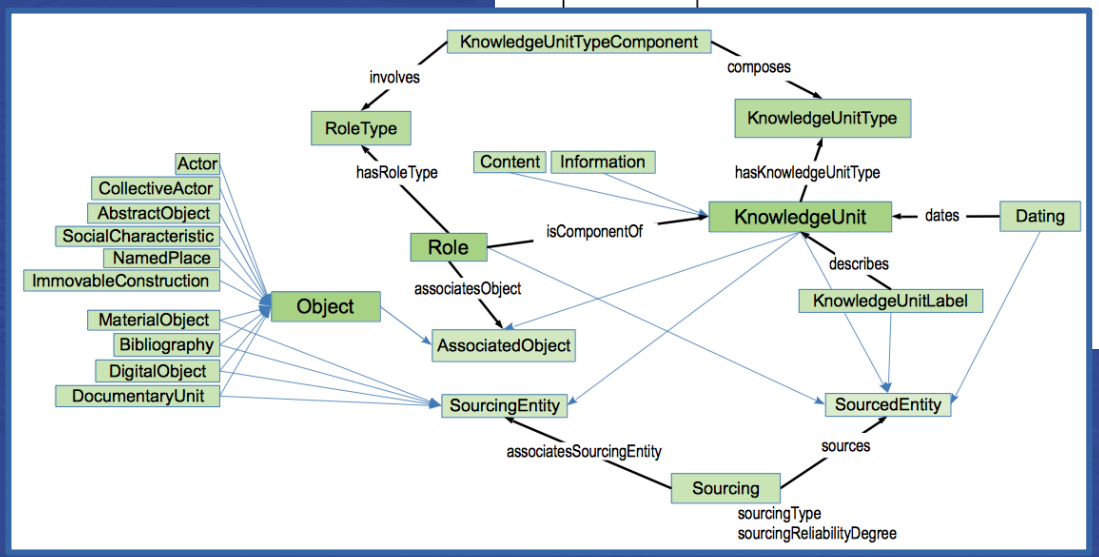


The symogih.org ontology

CC BY-NC-SA 4.0

« Galileo Galilei taught mathematics at the University of Padua from 1592 and 1610 » [Dizionario biografico degli italiani, vol. 51]





CC BY-NC-SA 4.0

The definition of each instance ot the data model is publicly available

SYMOGIH

Références

http://symogih.org

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 ty

Chercher une cl

- Biographie
  - Enseig
  - Exerci
  - Fin de
  - Liens a
  - Localis
  - Rites so
  - Vie mil
  - Vie pro

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

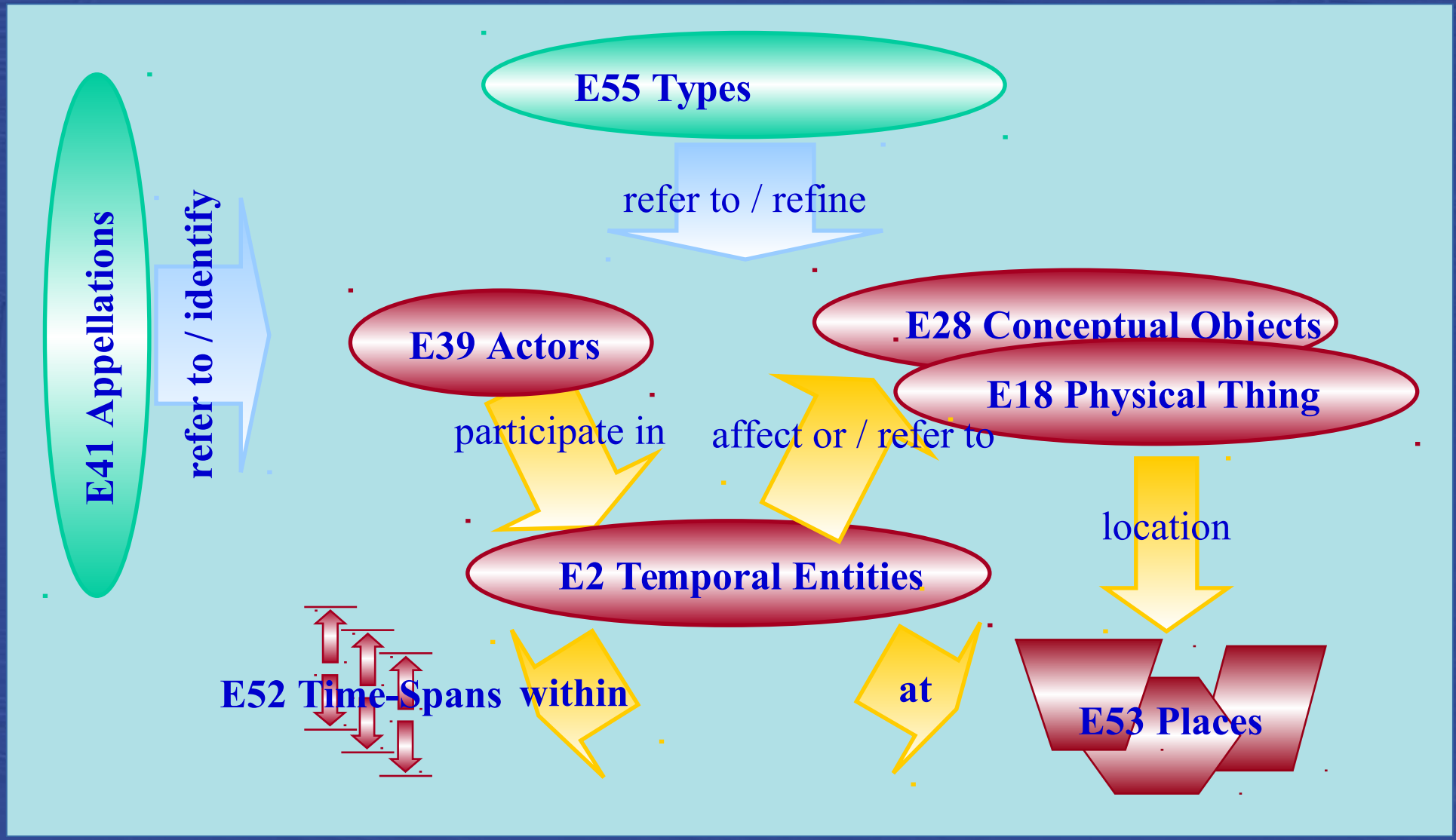
Libellé du type de rôle	Clé du TyRo	Description
concerner	TyRo21	Institution auprès de laquelle s'exerce l'enseignement. Ce rôle a été gardé pour être l'équivalent du TyIn7 : Exercice d'une fonction.
enseigné (être)	TyRo131	La matière enseignée (un objet abstrait). On peut en associer plusieurs si on 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

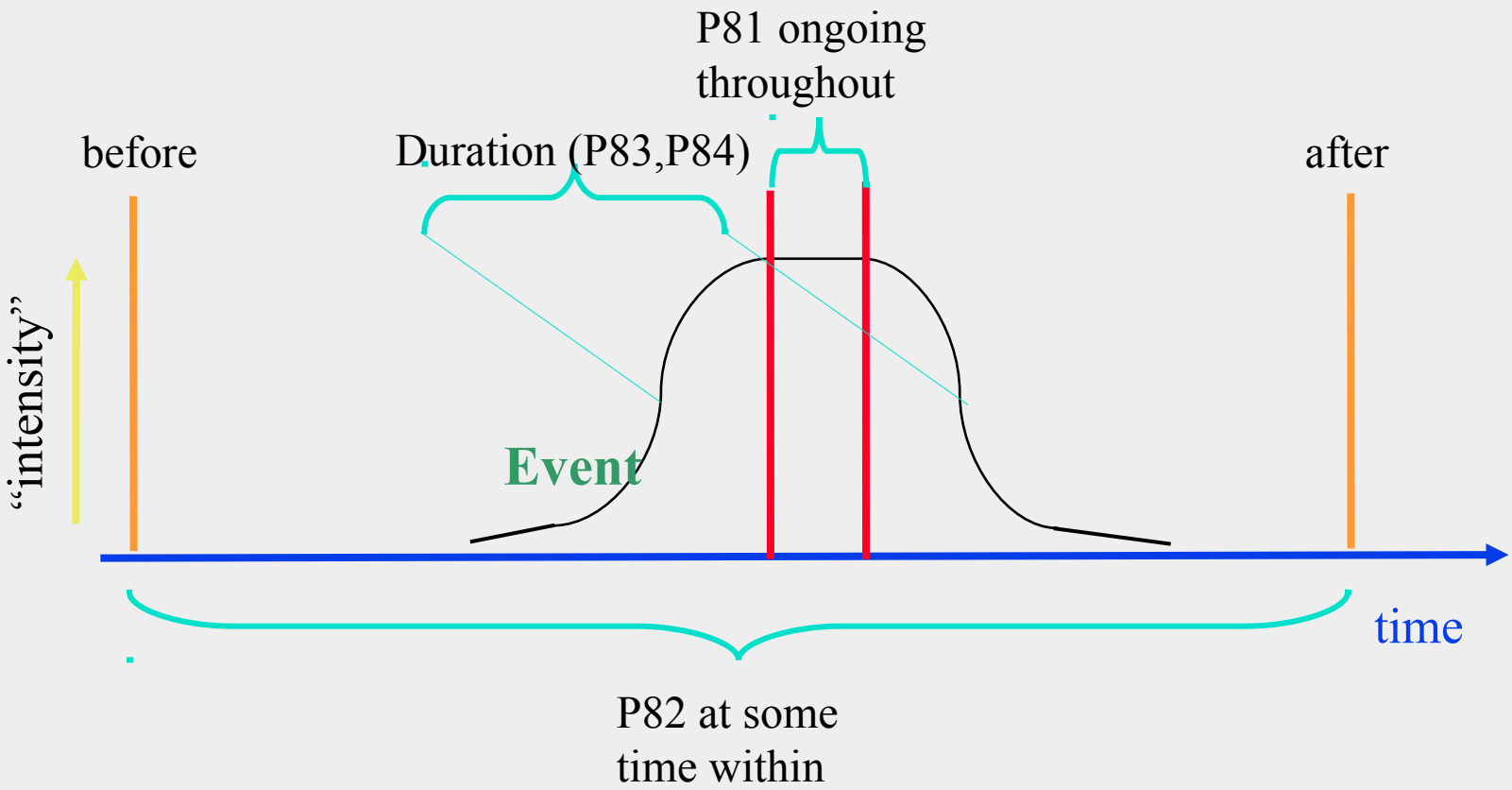
CC BY-NC-SA 4.0

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



CC BY-NC-SA 4.0

The CIDOC CRM
Time Uncertainty, Certainty and Duration



CC BY-NC-SA 4.0

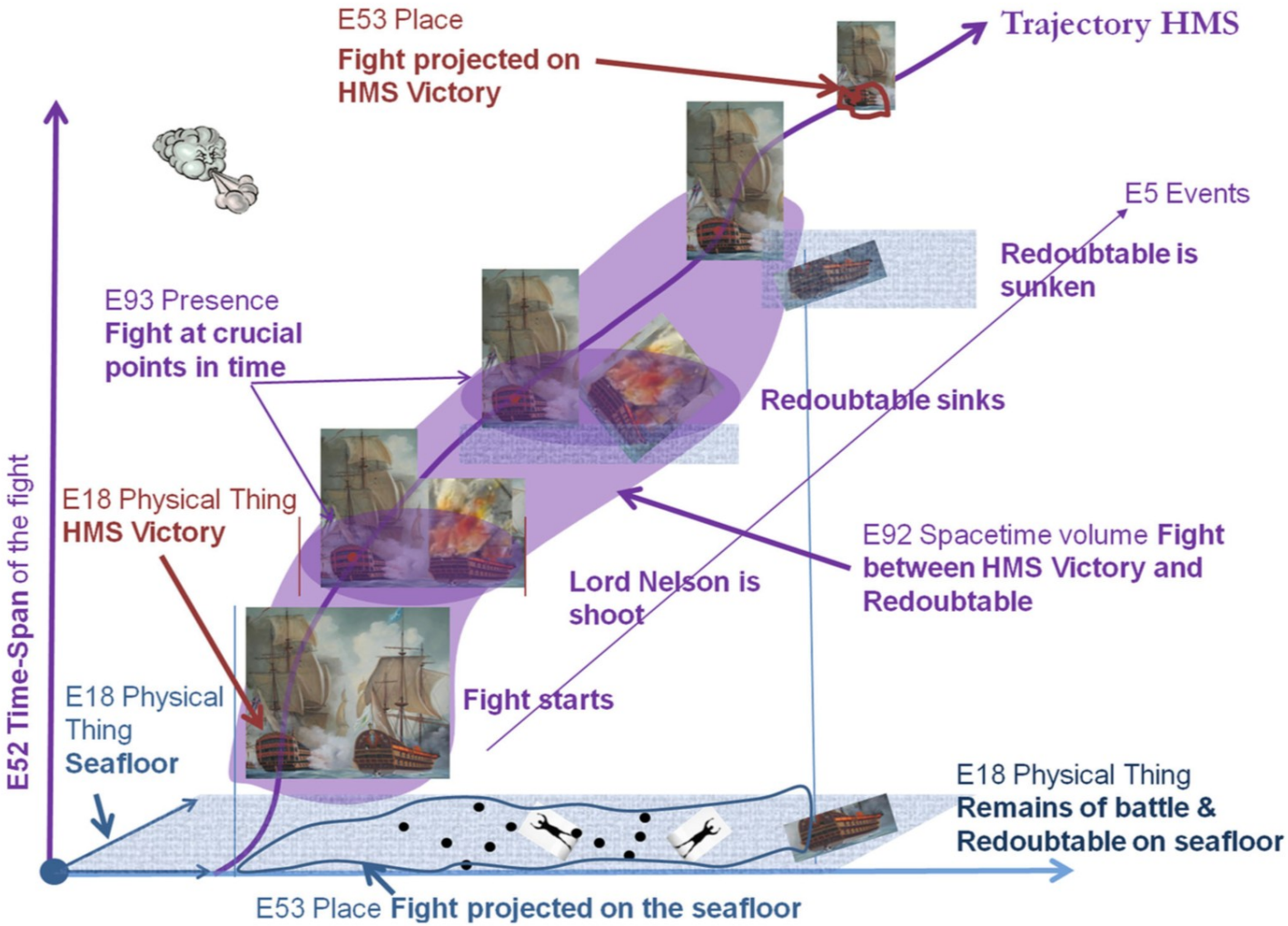
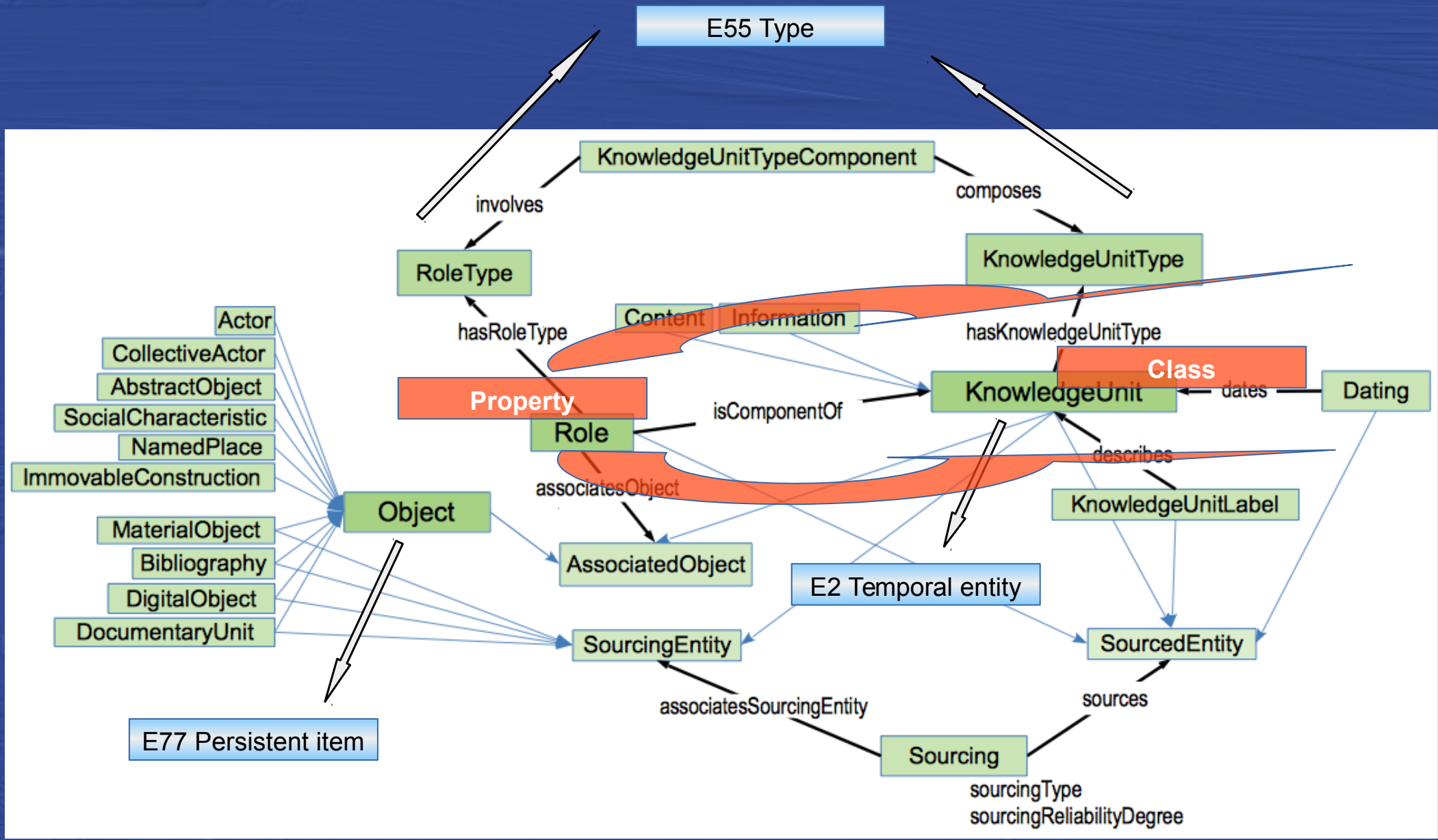


Fig. 2 The fight of the HMS Victory and the Redoubtable in the Battle of Trafalgar illustrating Spacetime volume (E92) and Presence (E93) and their projection to different Places (E53)

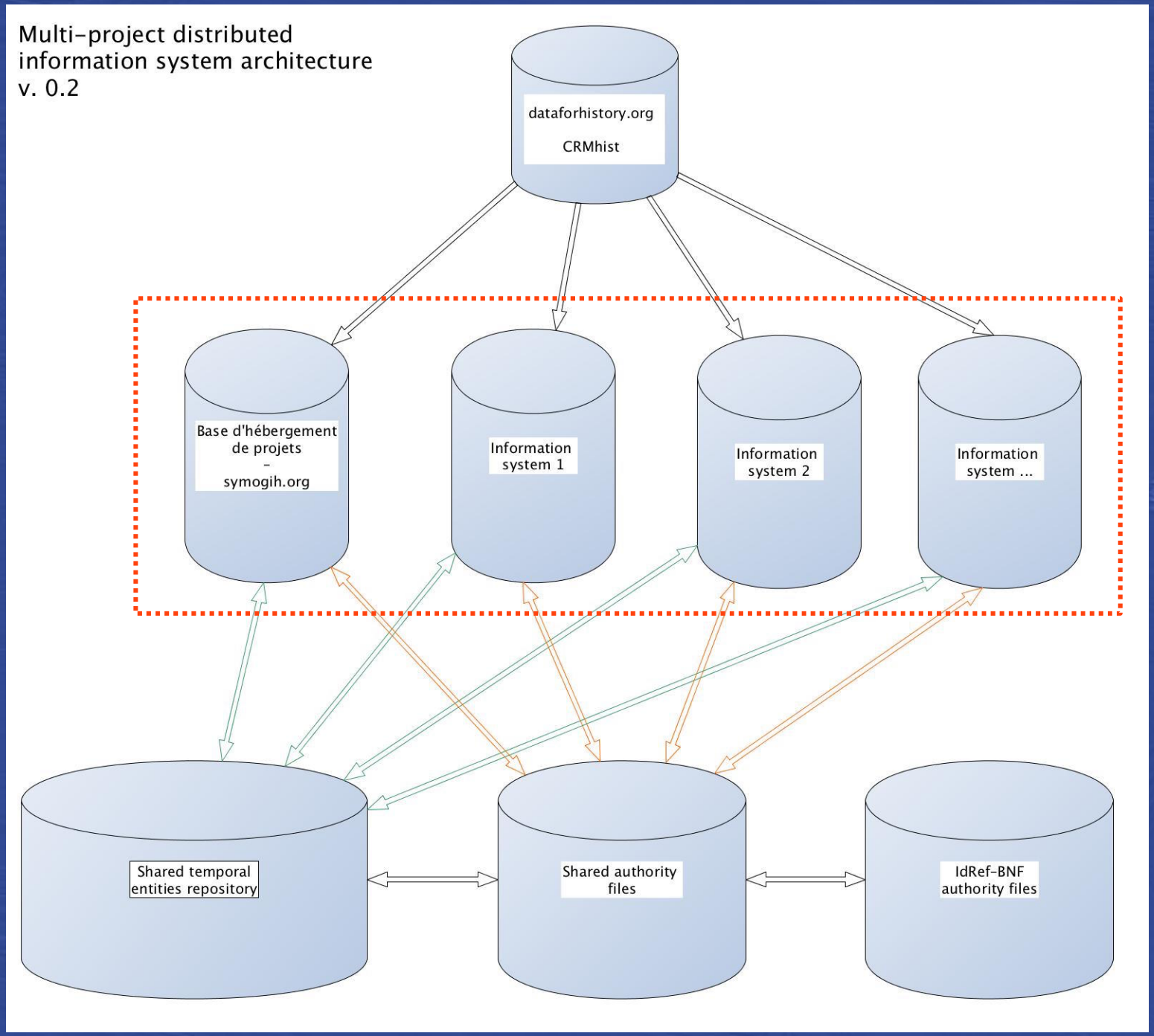
DOI 10.1007/s00799-016-0192-4

CRMgeo: A spatiotemporal extension of CIDOC-CRM Gerald Hiebel1 · Martin Doerr2 · Øyvind Eide3

CC BY-NC-SA 4.0

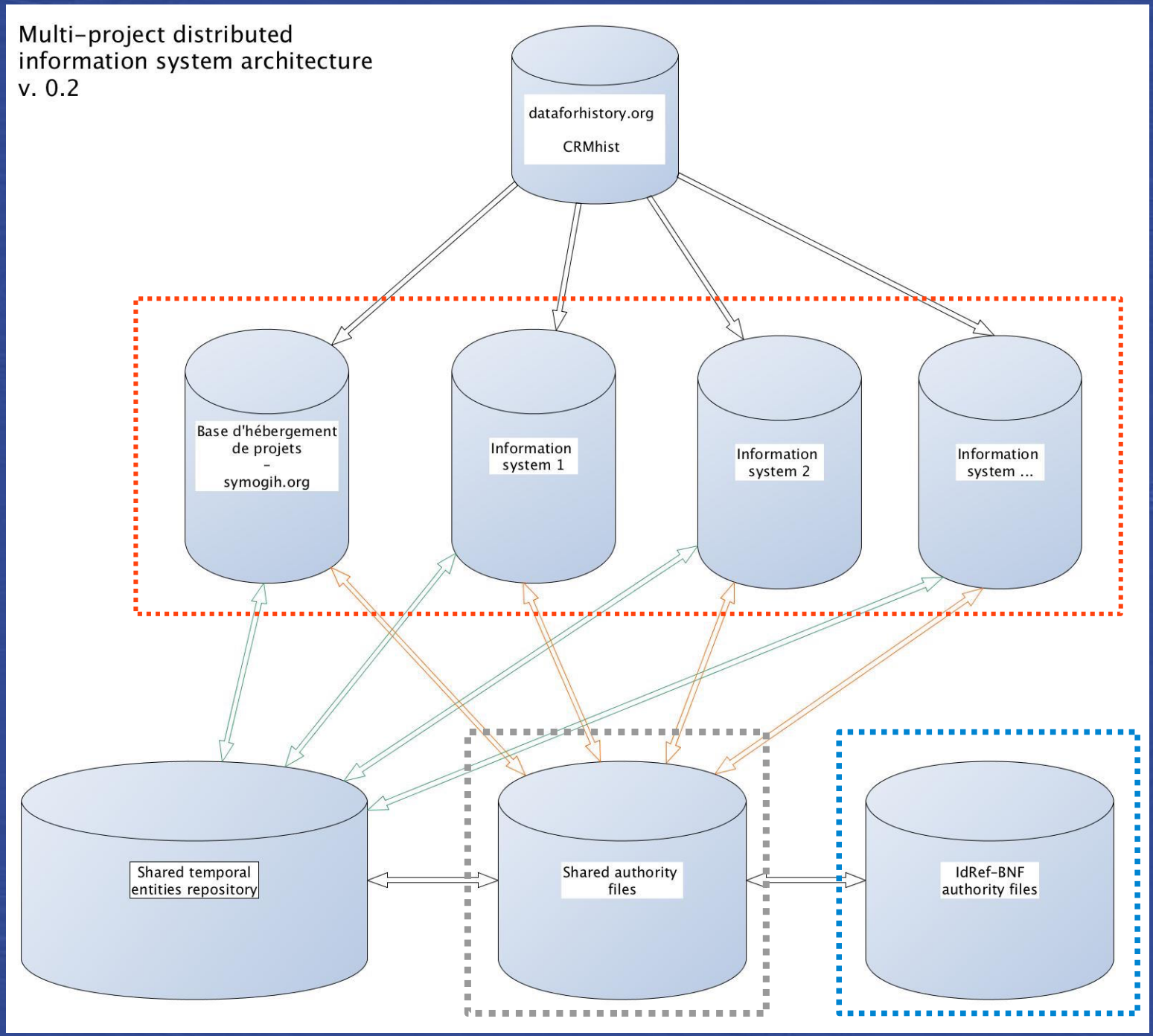

E55 Type
KnowledgeUnitTypeComponent
involves
composes
RoleType
KnowledgeUnitType
hasRoleType
Content
Information
hasKnowledgeUnitType
Class
Actor
CollectiveActor
AbstractObject
SocialCharacteristic
NamedPlace
ImmovableConstruction
Property
KnowledgeUnit
dates
Dating
isComponentOf
Role
describes
associatesObject
KnowledgeUnitLabel
Object
MaterialObject
Bibliography
DigitalObject
DocumentaryUnit
AssociatedObject
E2 Temporal entity
SourcingEntity
SourcedEntity
associatesSourcingEntity
sources
E77 Persistent item
Sourcing
sourcingType
sourcingReliabilityDegree

CC BY-NC-SA 4.0



dataforhistory.org

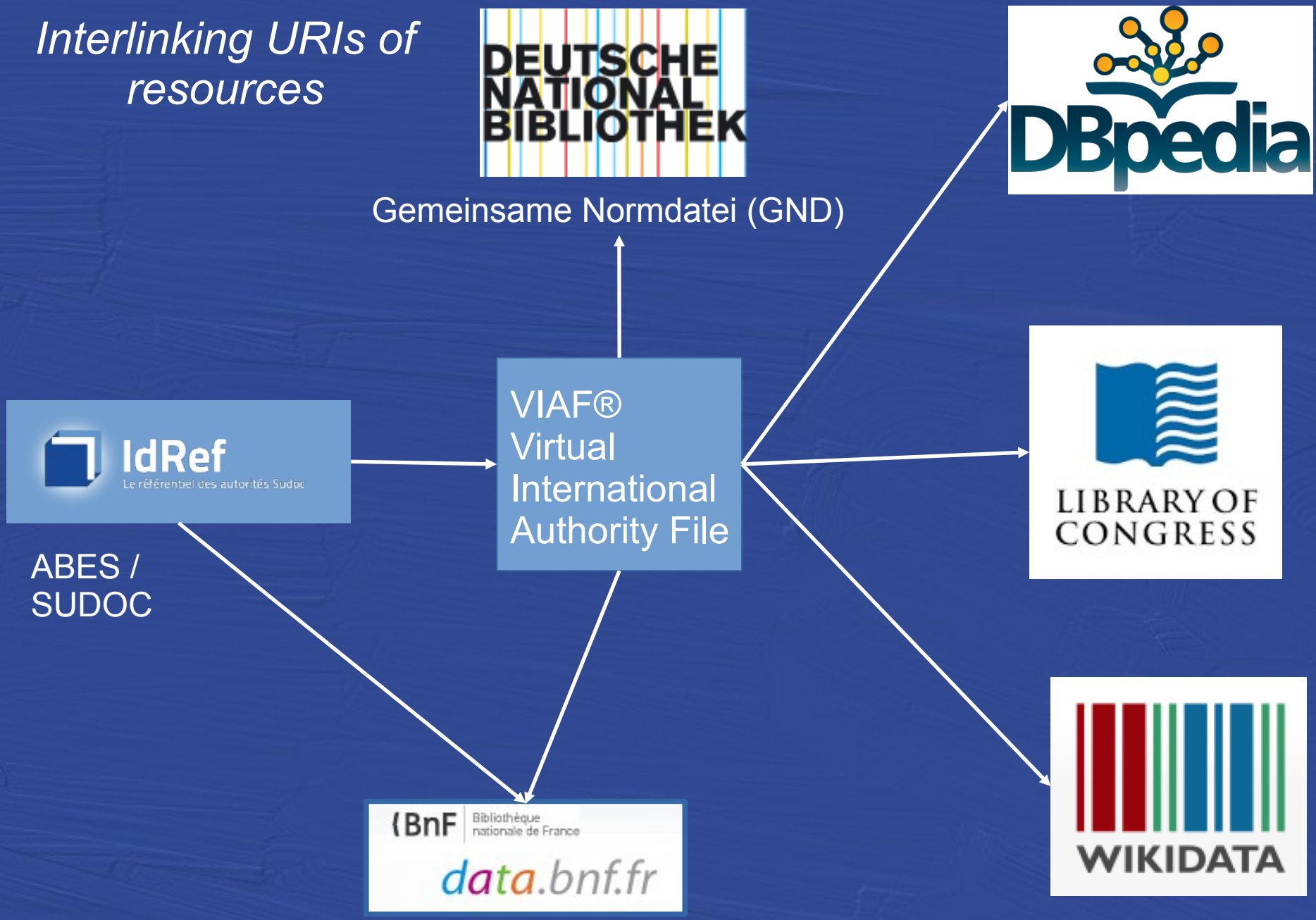
CC BY-NC-SA 4.0



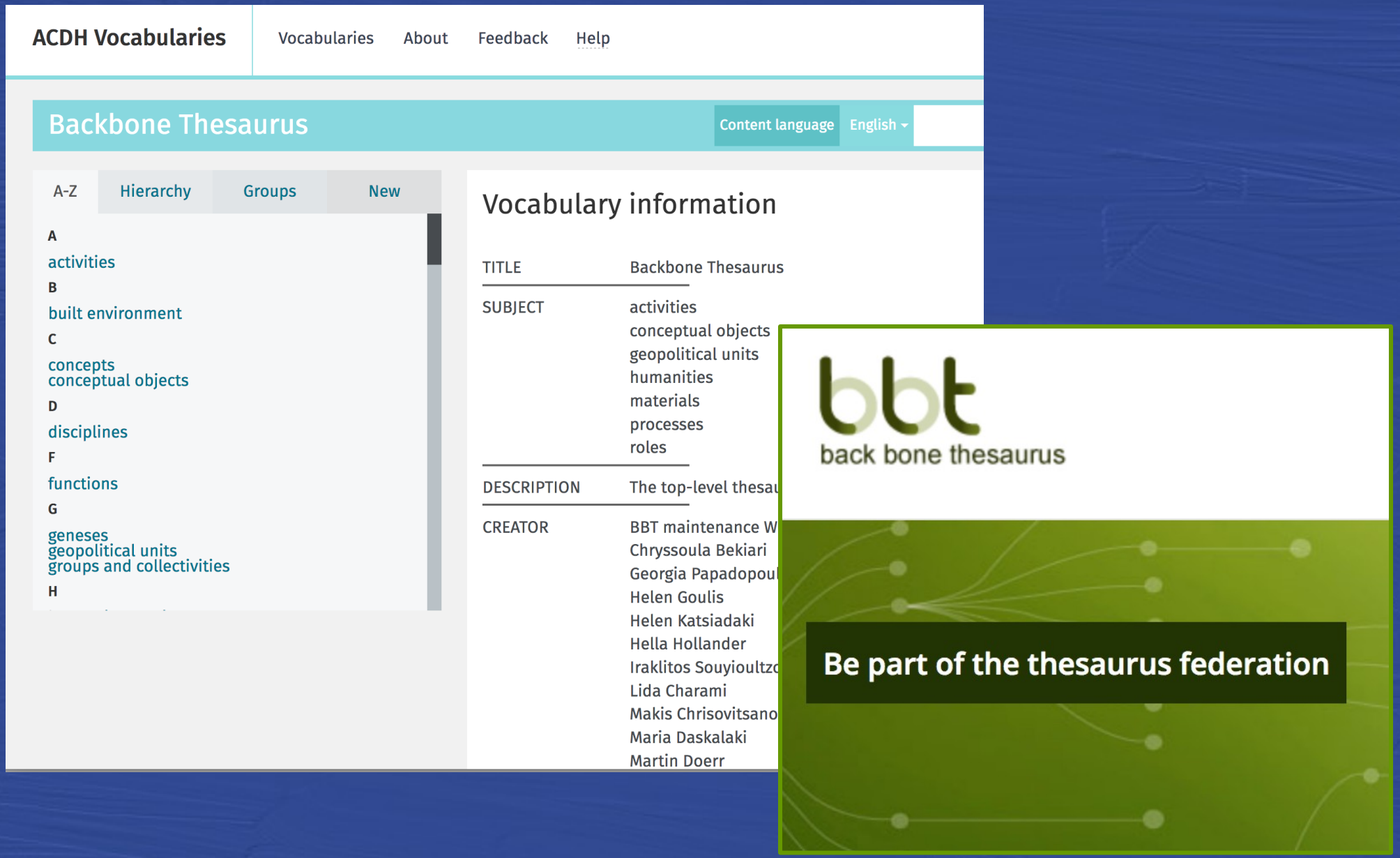
dataforhistory.org

CC BY-NC-SA 4.0

Interlinking URIs of resources

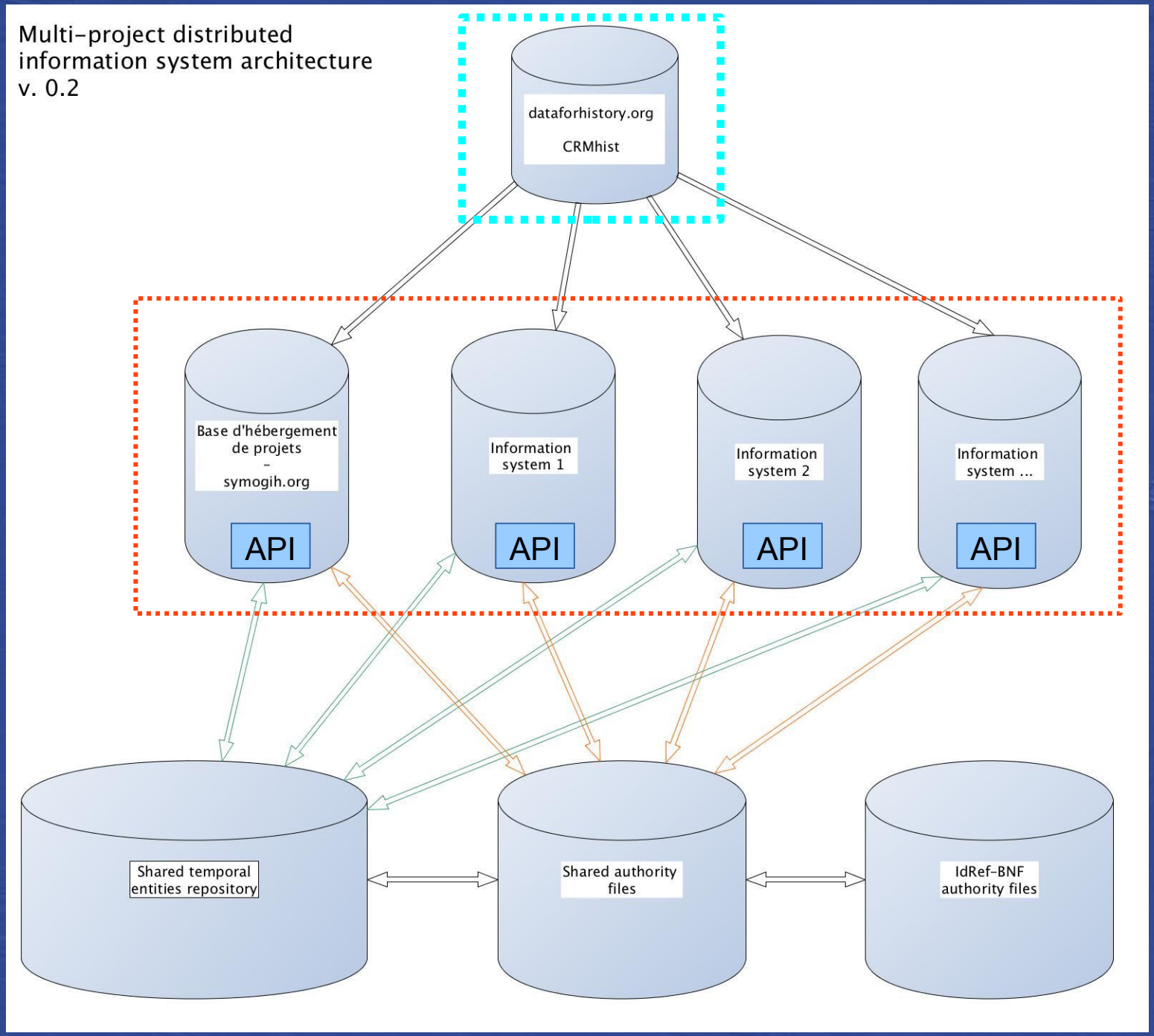


CC BY-NC-SA 4.0



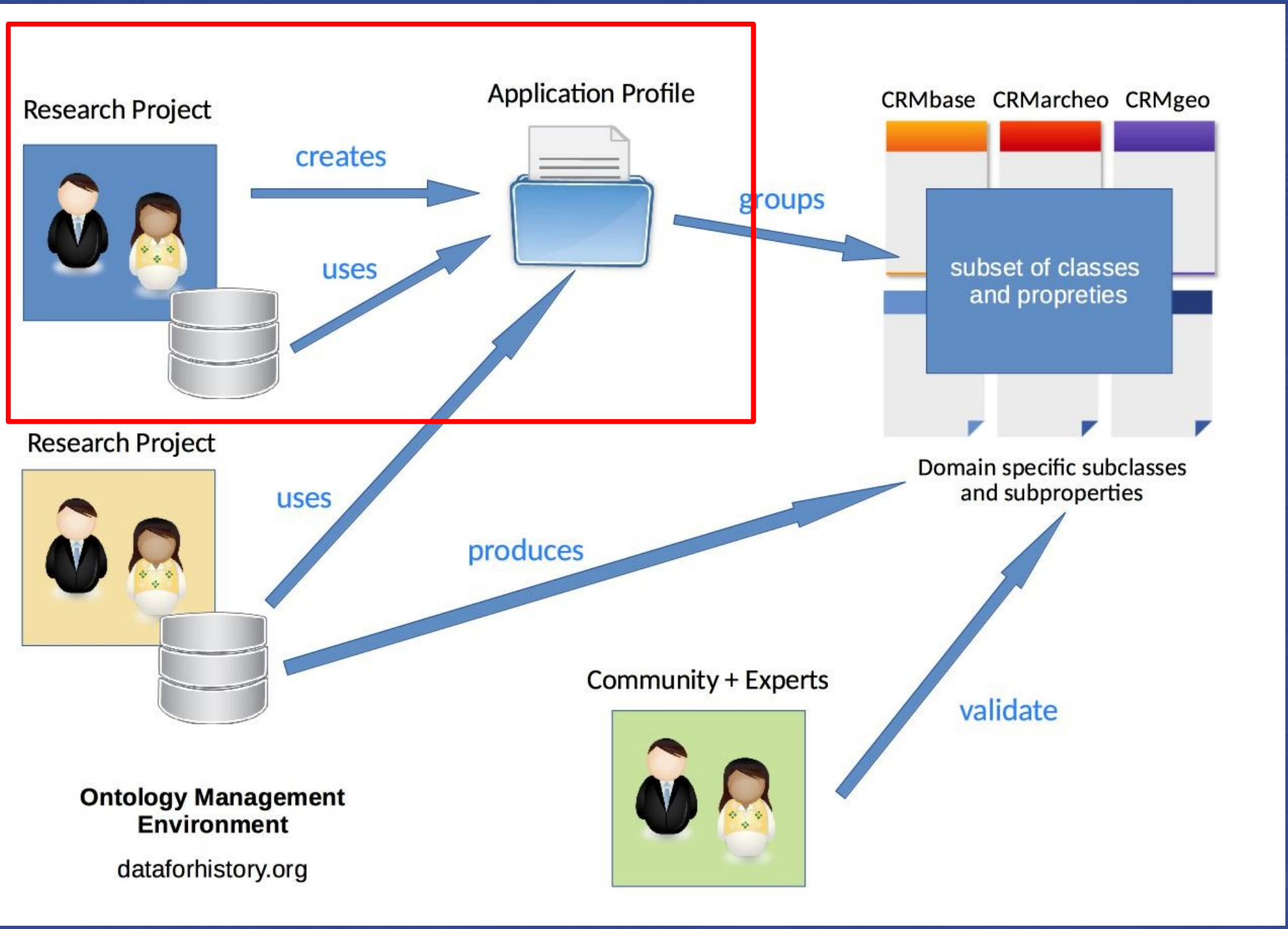
backbonethesaurus.eu

CC BY-NC-SA 4.0

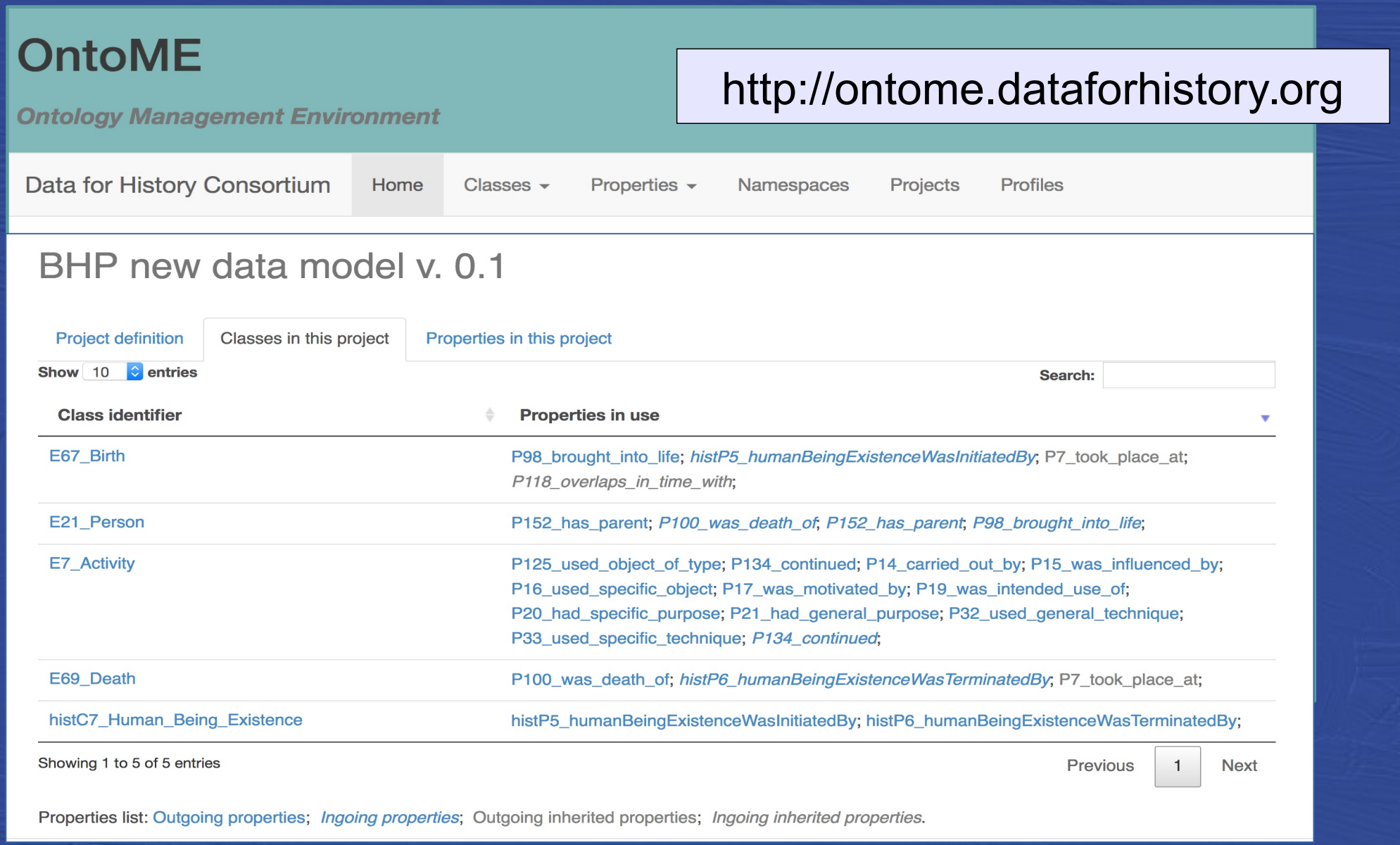


dataforhistory.org

CC BY-NC-SA 4.0


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

CC BY-NC-SA 4.0

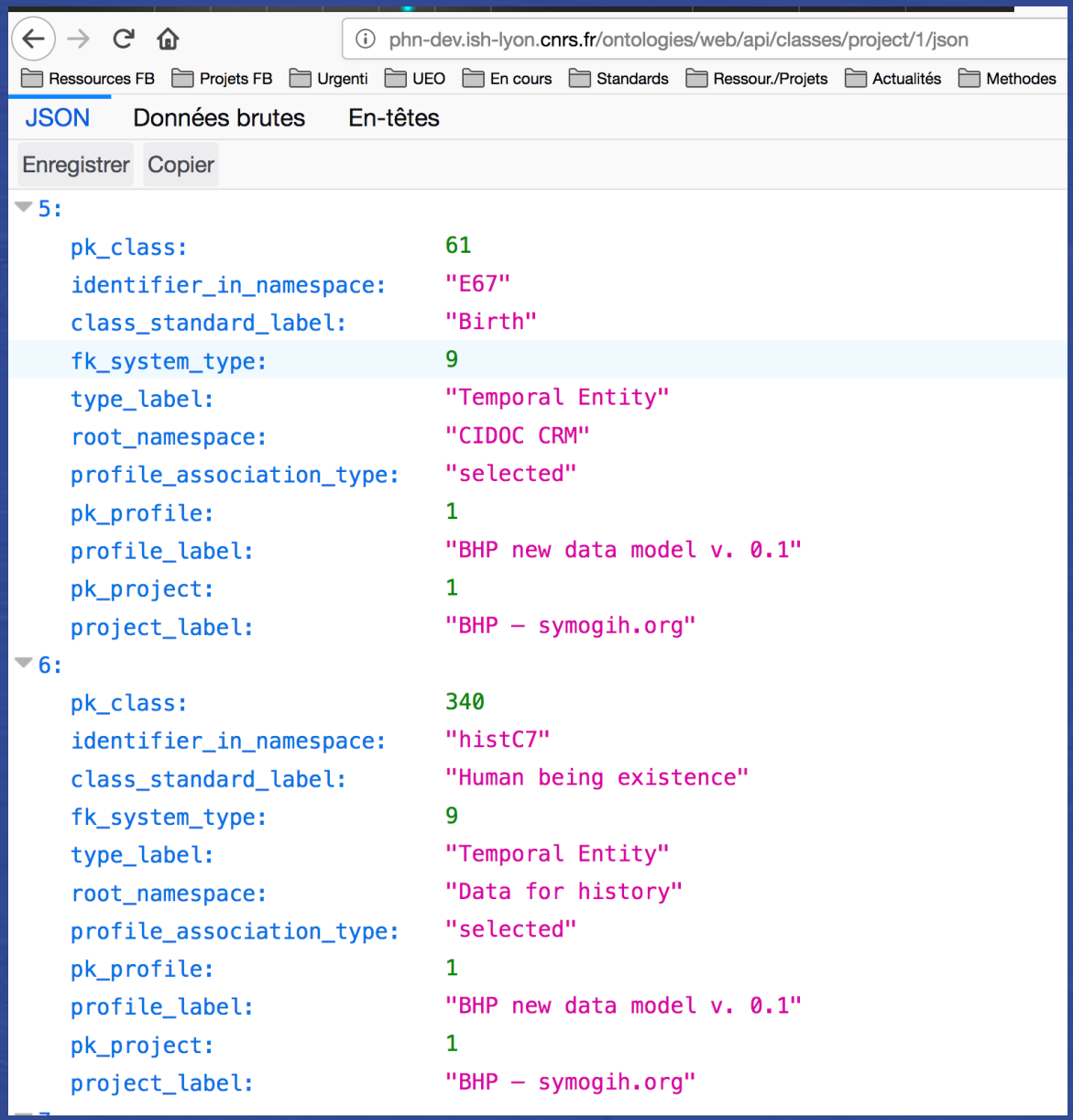


Define and customize the application profile of your own project

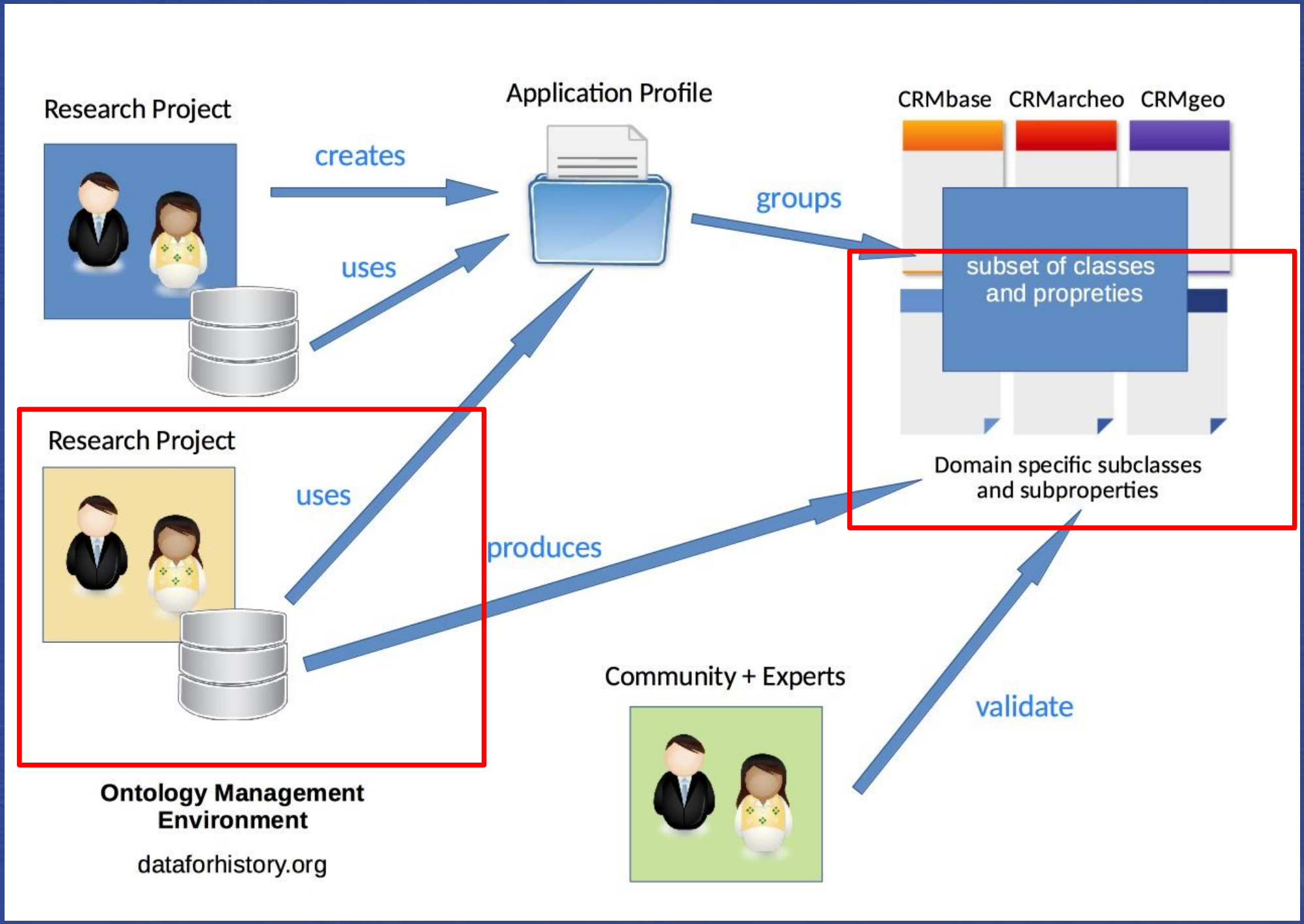
CC BY-NC-SA 4.0

Retrieve your project's application profiles from an API

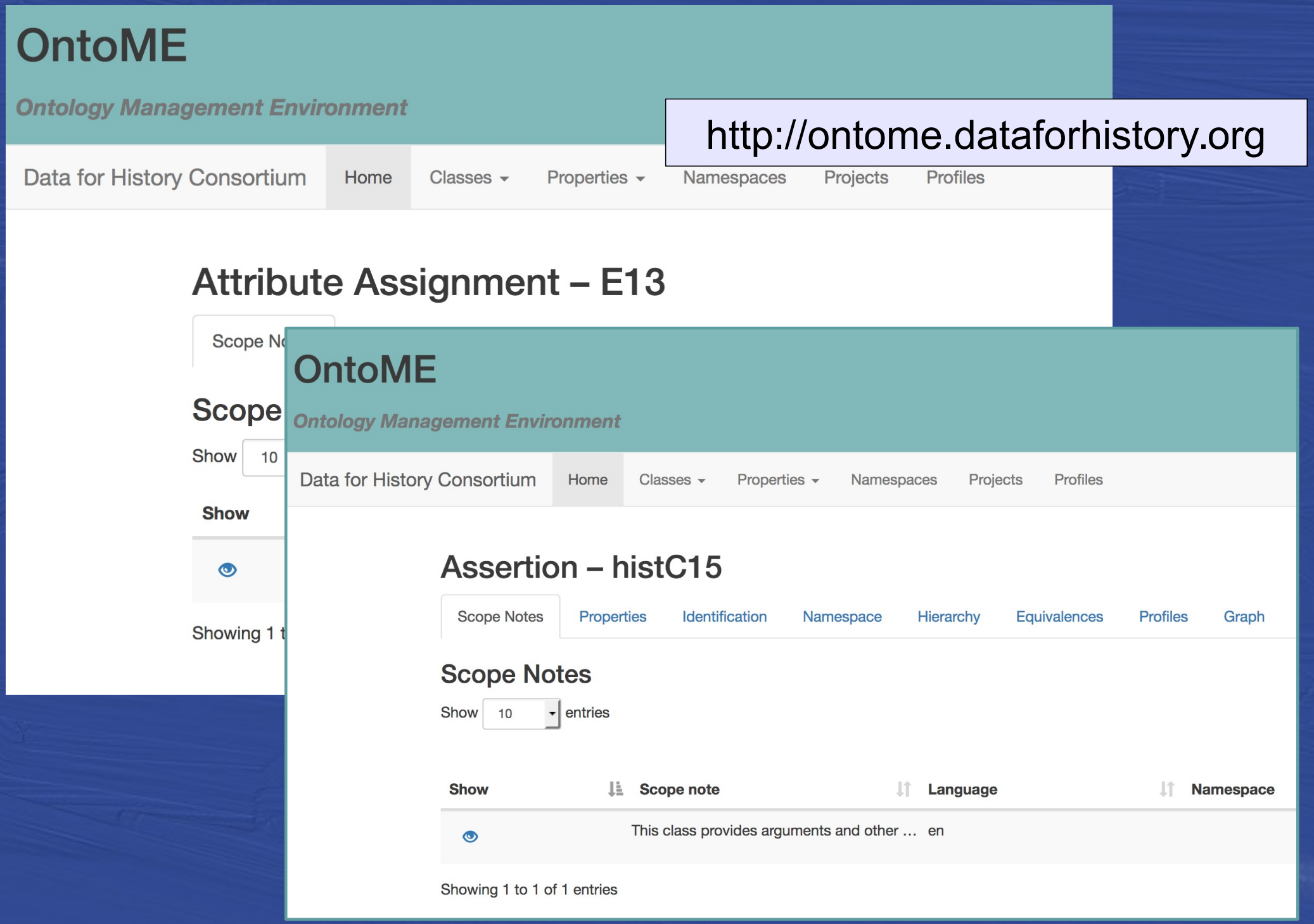
http://ontologies.dataforhistory.org/api/classes/project/1/json



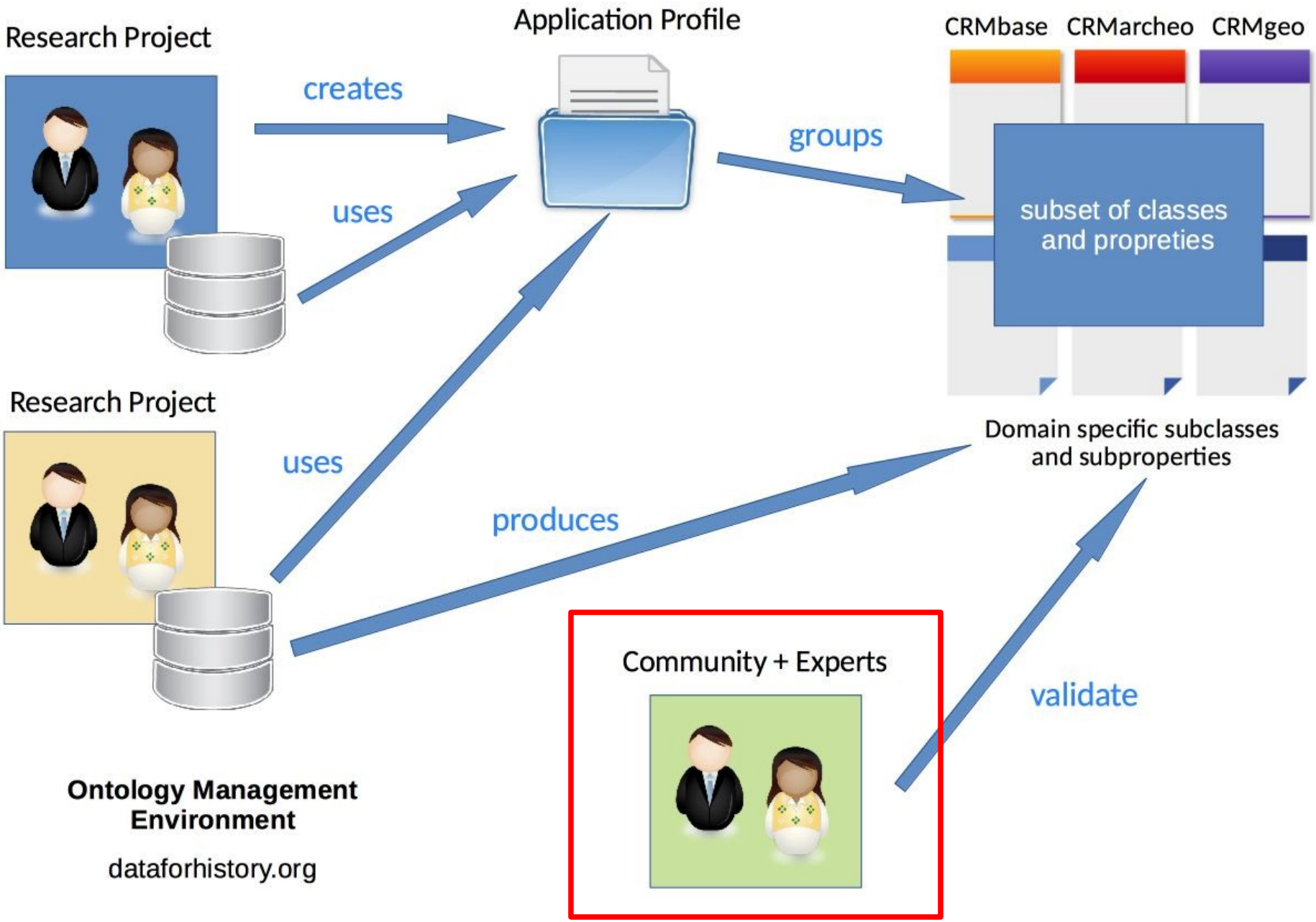
CC BY-NC-SA 4.0


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

CC BY-NC-SA 4.0


OntoME
Ontology Management Environment
http://ontome.dataforhistory.org
Data for History Consortium
Home
Classes
Properties
Namespaces
Projects
Profiles
Attribute Assignment – E13
Scope No
Scope
Show 10
Show
Showing 1 t
OntoME
Ontology Management Environment
Data for History Consortium
Home
Classes
Properties
Namespaces
Projects
Profiles
Assertion – histC15
Scope Notes
Properties
Identification
Namespace
Hierarchy
Equivalences
Profiles
Graph
Scope Notes
Show 10 entries
Show
Scope note
Language
Namespace
This class provides arguments and other …
en
Showing 1 to 1 of 1 entries

CC BY-NC-SA 4.0


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

CC BY-NC-SA 4.0

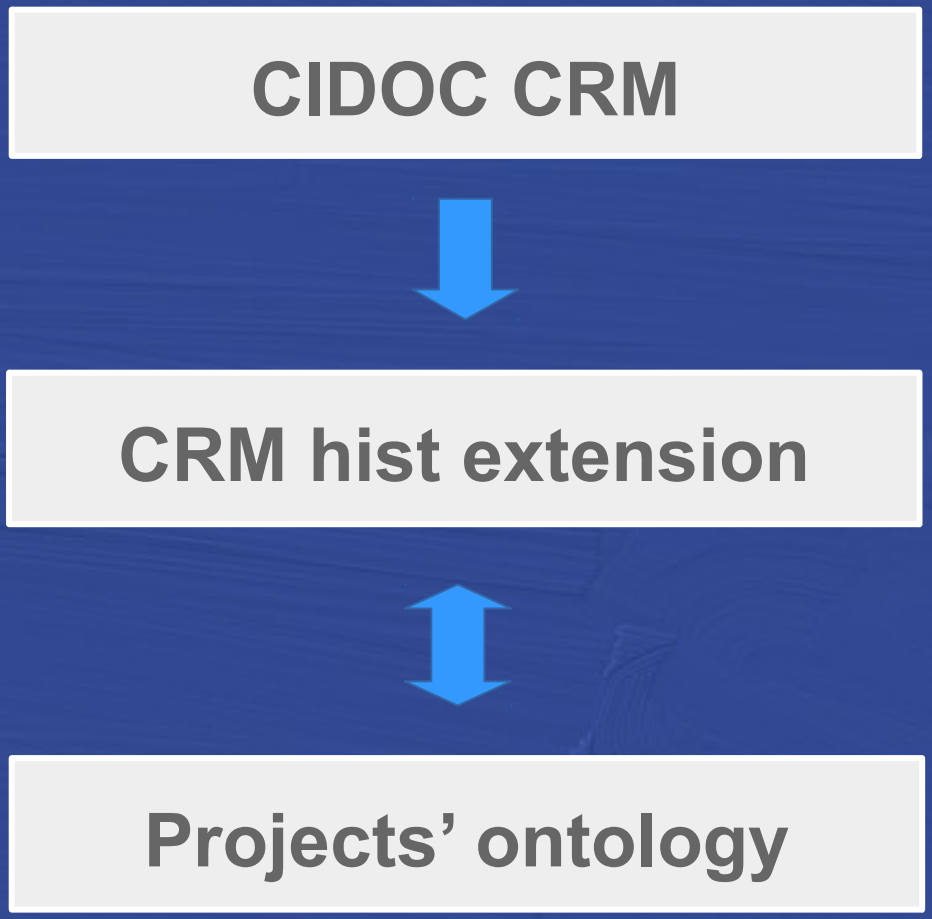
CIDOC CRM

CC BY-NC-SA 4.0

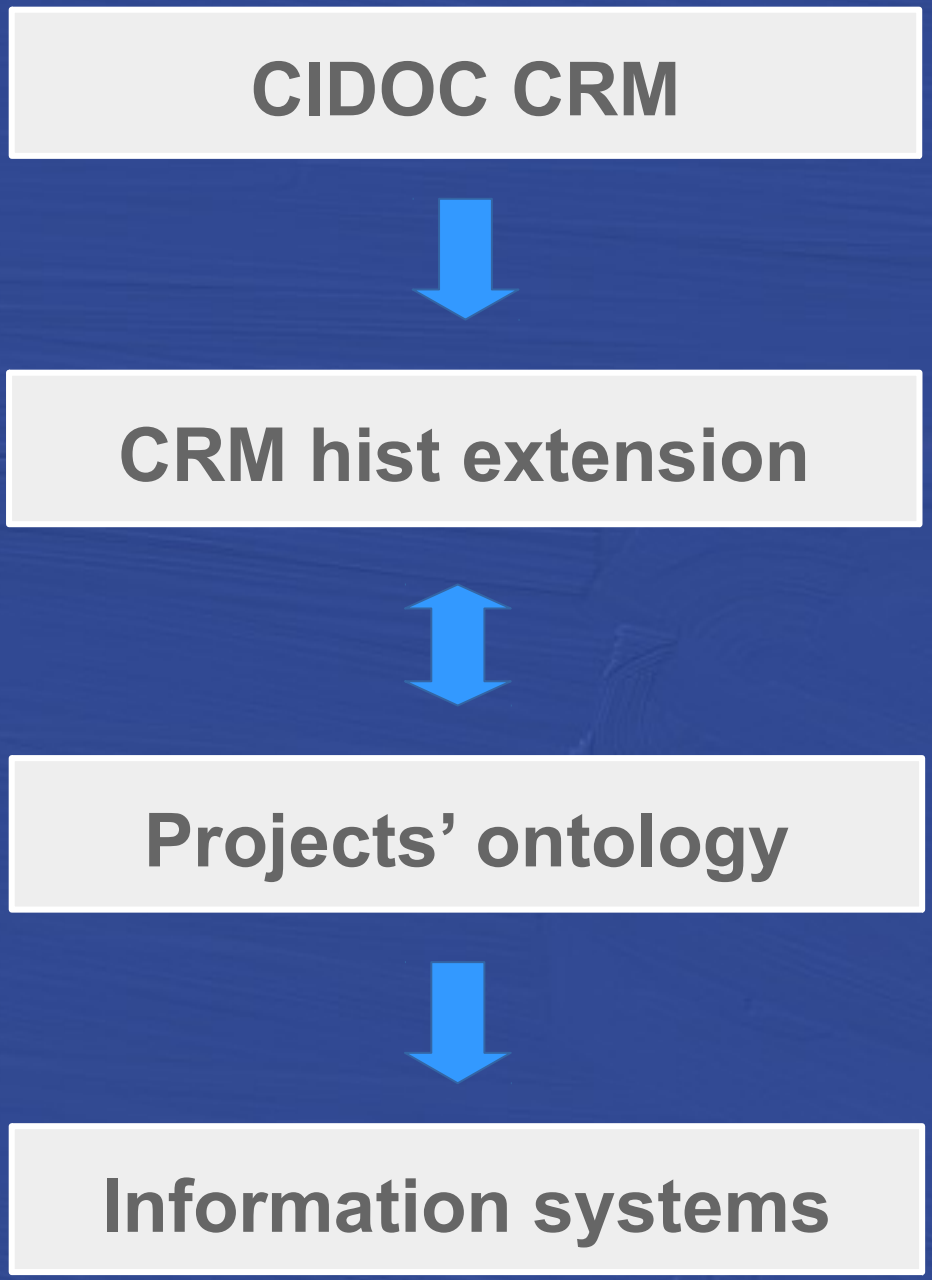
CIDOC CRM

CRM hist extension

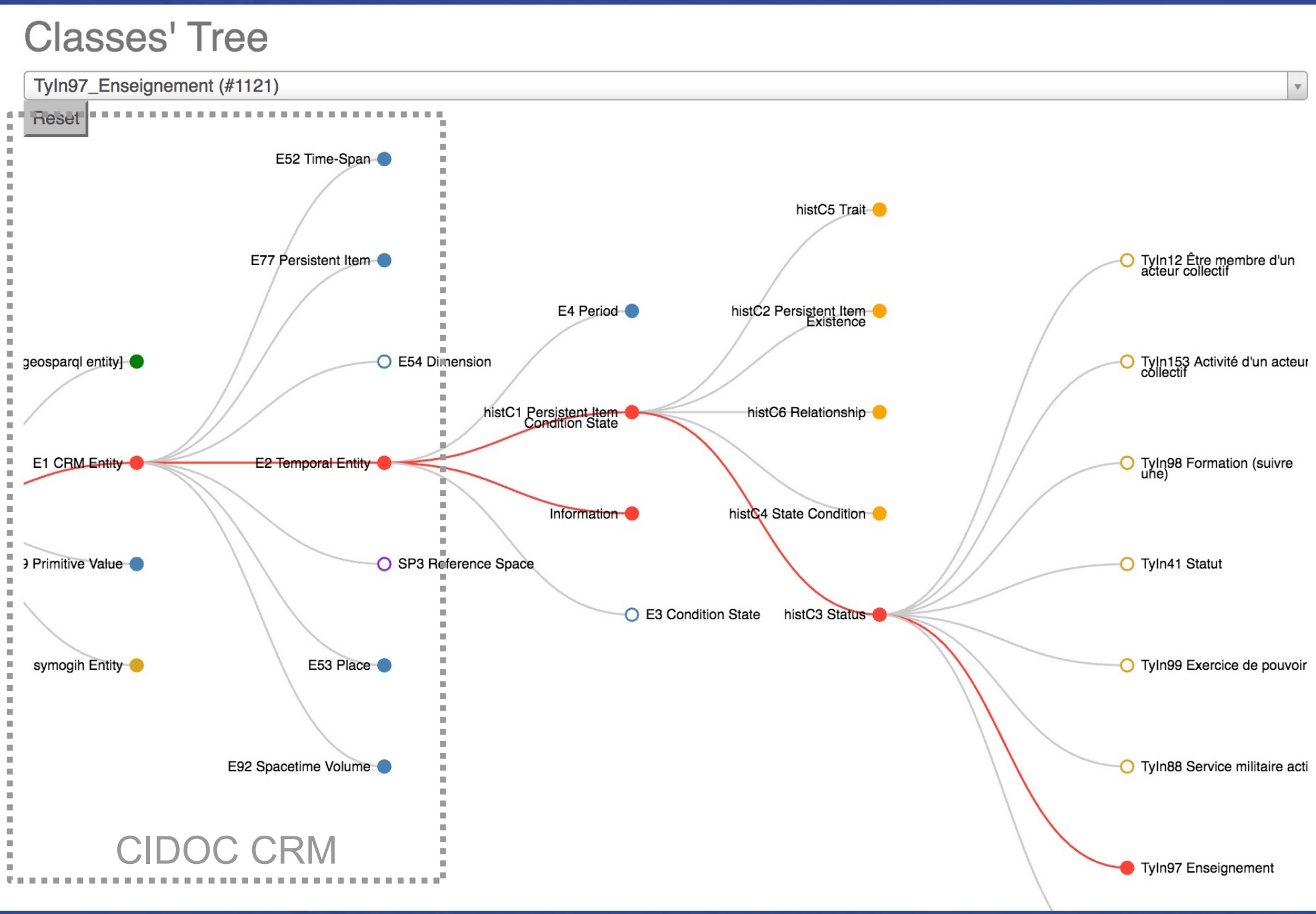
CC BY-NC-SA 4.0



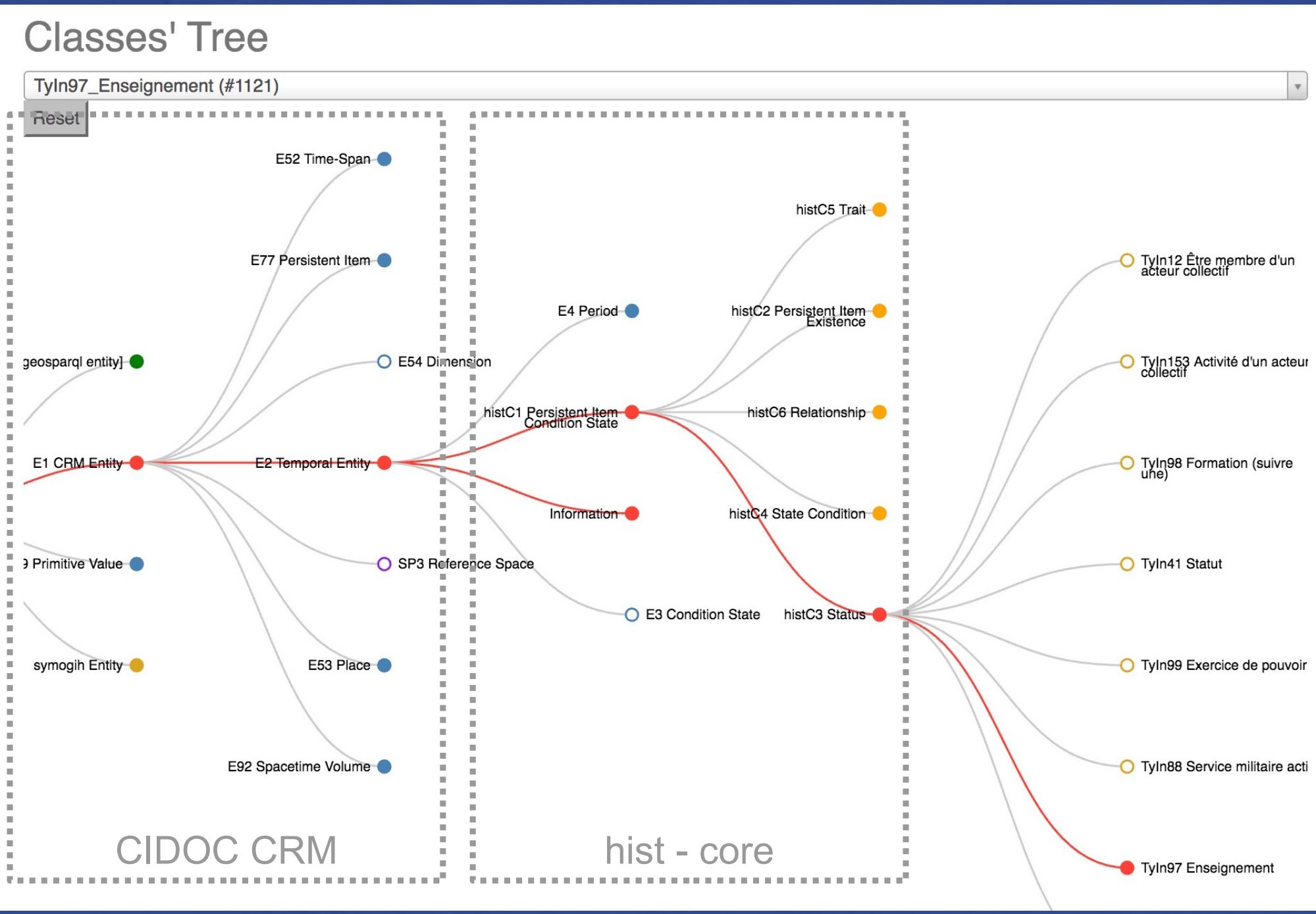
CC BY-NC-SA 4.0


CIDOC CRM
CRM hist extension
Projects' ontology
Information systems

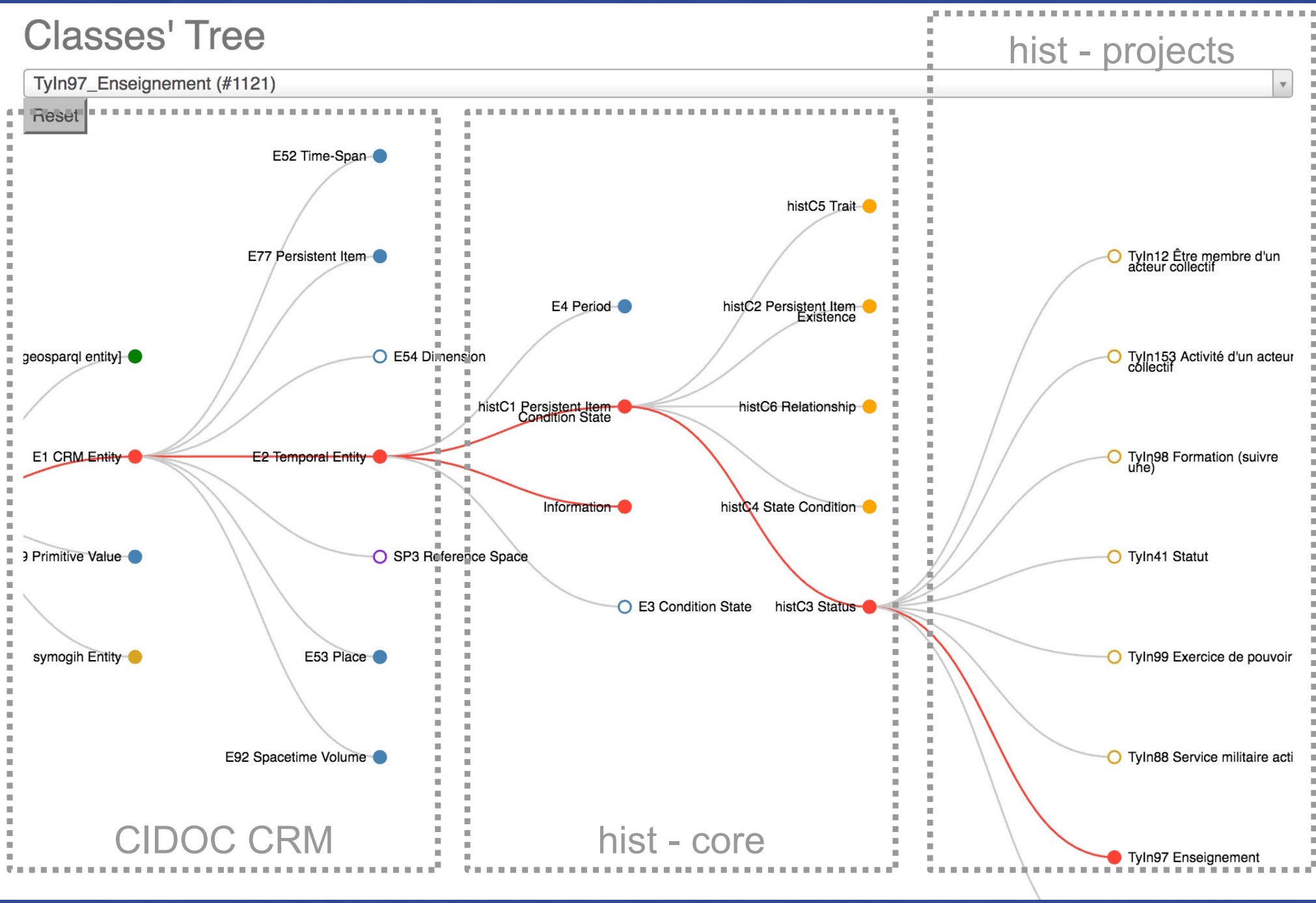
CC BY-NC-SA 4.0


Classes' Tree
TyIn97_Enseignement (#1121)
Reset
E52 Time-Span
E77 Persistent Item
geosparql entity]
E54 Dimension
E1 CRM Entity
E2 Temporal Entity
Primitive Value
SP3 Reference Space
symogih Entity
E53 Place
E92 Spacetime Volume
CIDOC CRM
E4 Period
histC1 Persistent Item Condition State
Information
E3 Condition State
histC5 Trait
histC2 Persistent Item Existence
histC6 Relationship
histC4 State Condition
histC3 Status
TyIn12 Être membre d'un acteur collectif
TyIn153 Activité d'un acteur collectif
TyIn98 Formation (suivre une)
TyIn41 Statut
TyIn99 Exercice de pouvoir
TyIn88 Service militaire acti
TyIn97 Enseignement

CC BY-NC-SA 4.0


Classes' Tree
TyIn97_Enseignement (#1121)
Reset
E52 Time-Span
E77 Persistent Item
geosparql entity]
E54 Dimension
E1 CRM Entity
E2 Temporal Entity
Primitive Value
SP3 Reference Space
symogih Entity
E53 Place
E92 Spacetime Volume
CIDOC CRM
E4 Period
histC1 Persistent Item Condition State
Information
E3 Condition State
histC5 Trait
histC2 Persistent Item Existence
histC6 Relationship
histC4 State Condition
histC3 Status
hist - core
TyIn12 Être membre d'un acteur collectif
TyIn153 Activité d'un acteur collectif
TyIn98 Formation (suivre une)
TyIn41 Statut
TyIn99 Exercice de pouvoir
TyIn88 Service militaire acti
TyIn97 Enseignement

CC BY-NC-SA 4.0


Classes' Tree
TyIn97_Enseignement (#1121)
Reset
hist - projects
E52 Time-Span
E77 Persistent Item
E54 Dimension
geosparql entity]
E1 CRM Entity
E2 Temporal Entity
Primitive Value
SP3 Reference Space
symogih Entity
E53 Place
E92 Spacetime Volume
CIDOC CRM
E4 Period
histC1 Persistent Item Condition State
Information
E3 Condition State
histC5 Trait
histC2 Persistent Item Existence
histC6 Relationship
histC4 State Condition
histC3 Status
hist - core
TyIn12 Être membre d'un acteur collectif
TyIn153 Activité d'un acteur collectif
TyIn98 Formation (suivre une)
TyIn41 Statut
TyIn99 Exercice de pouvoir
TyIn88 Service militaire acti
TyIn97 Enseignement

CC BY-NC-SA 4.0

Data for History Consortium

Meetings

- Lyon, November 2017 (founding meeting)
- Lyon, May 2018
- Panel presenting the Data for History vision and tools
  at the EADH 2018 conference in Galway (December 2018)
- Leipzig, April 2019 (planned)

The Data for History consortium

is in the process of being formally established

and is **open to all interested institutions and researchers**

with the aim of sharing FAIR data, best practices, tools and platforms

CC BY-NC-SA 4.0

The FAIR data principles

To be **Interoperable**:

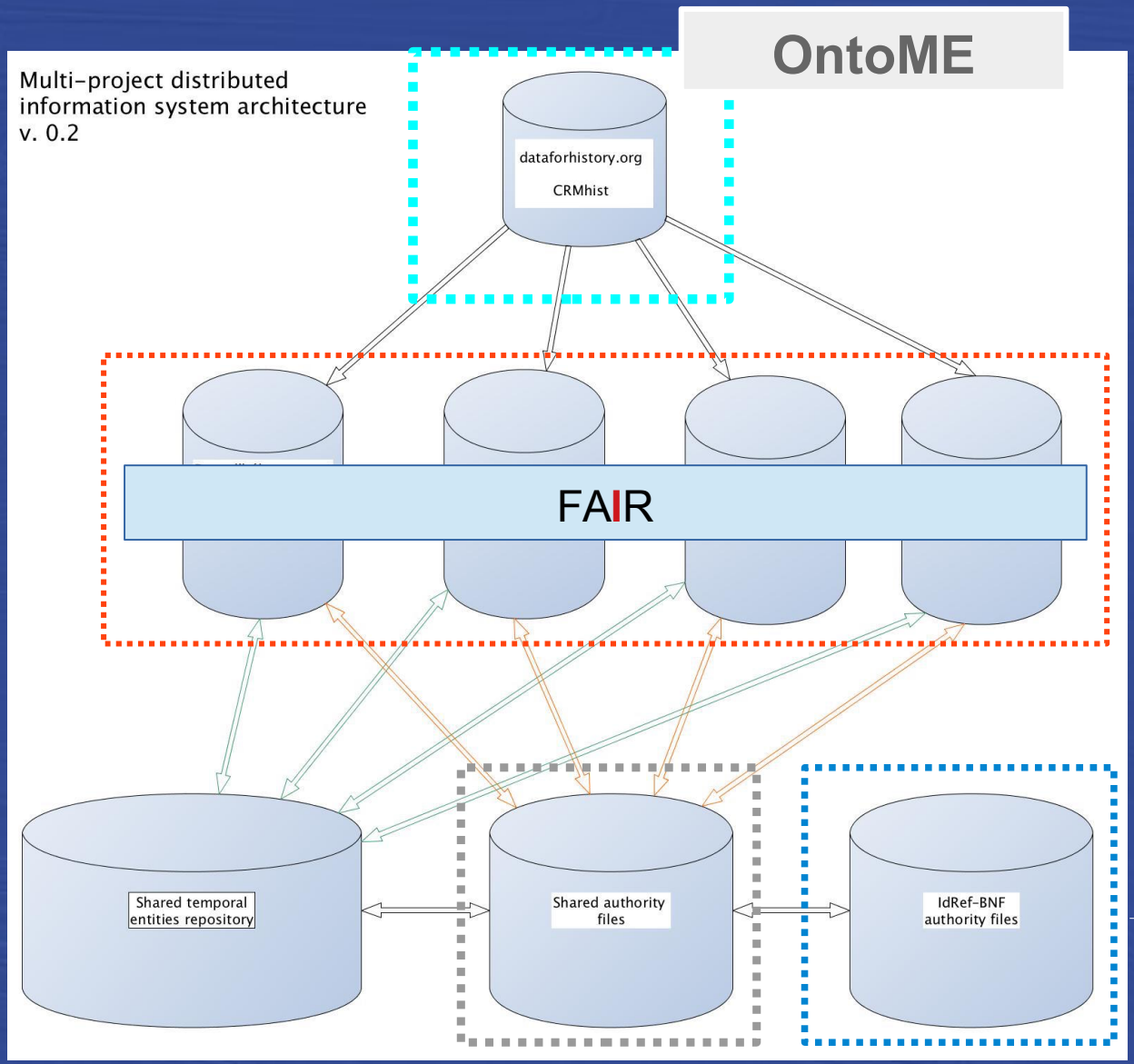
I1. (meta)data use a *formal, accessible, shared, and broadly applicable* ***language for knowledge representation***.

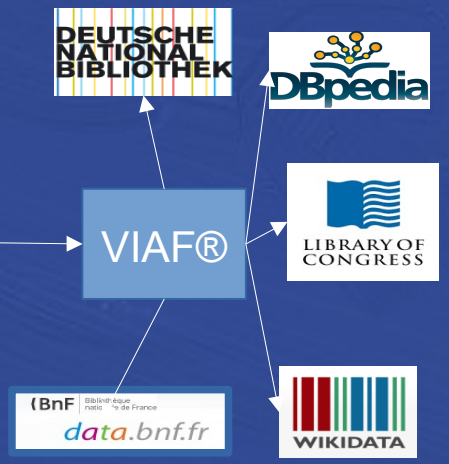
I2. (meta)data use ***vocabularies*** *that follow FAIR principles*.

I3. (meta)data include **qualified references** to other (meta)data.

CC BY-NC-SA 4.0

dataforhistory.org





CC BY-NC-SA 4.0