The *dataforhistory.org* project : a proposal

Francesco Beretta (CNRS/Université de Lyon, Laboratoire de recherche historique Rhône-Alpes)

George Bruseker (FORTH, Center for Cultural Heritage - Institute of Computer Science, Heraklion)

Workshop on the creation of an international Data for History consortium

École normale supérieure de Lyon, 23-24 November 2017

symogih.org

Ontologie symogih.org
Types d'informations
et de contenus

symogih.org

Interface
d'alimentation

Notices
d'autorité

Instance 3

Notices
d'autorité

Instance 2

Notices
d'autorité

Base
d'hébergement
de projets (BHP)

Stockage collaboratif
de textes XML-TEI
(eXist-db)

Patrons
de France

Siprojuris

geo-larhra.org

...

Édition numérique des Actes
des églises réformées
de Bourgogne au XVIIe siècle

Society Religion Science.
Digital intellectual history

Édition critique numérique
des Mémoires de Léonard Michon
notable lyonnais (1675-1746)

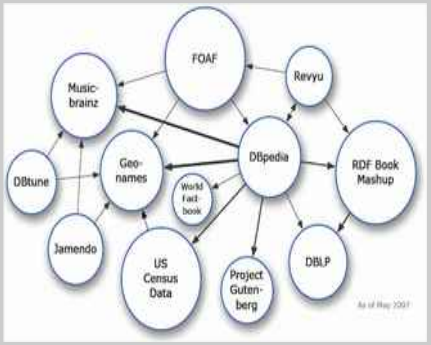
VIAF

ISNI

Notices
d'autorité

IdREF

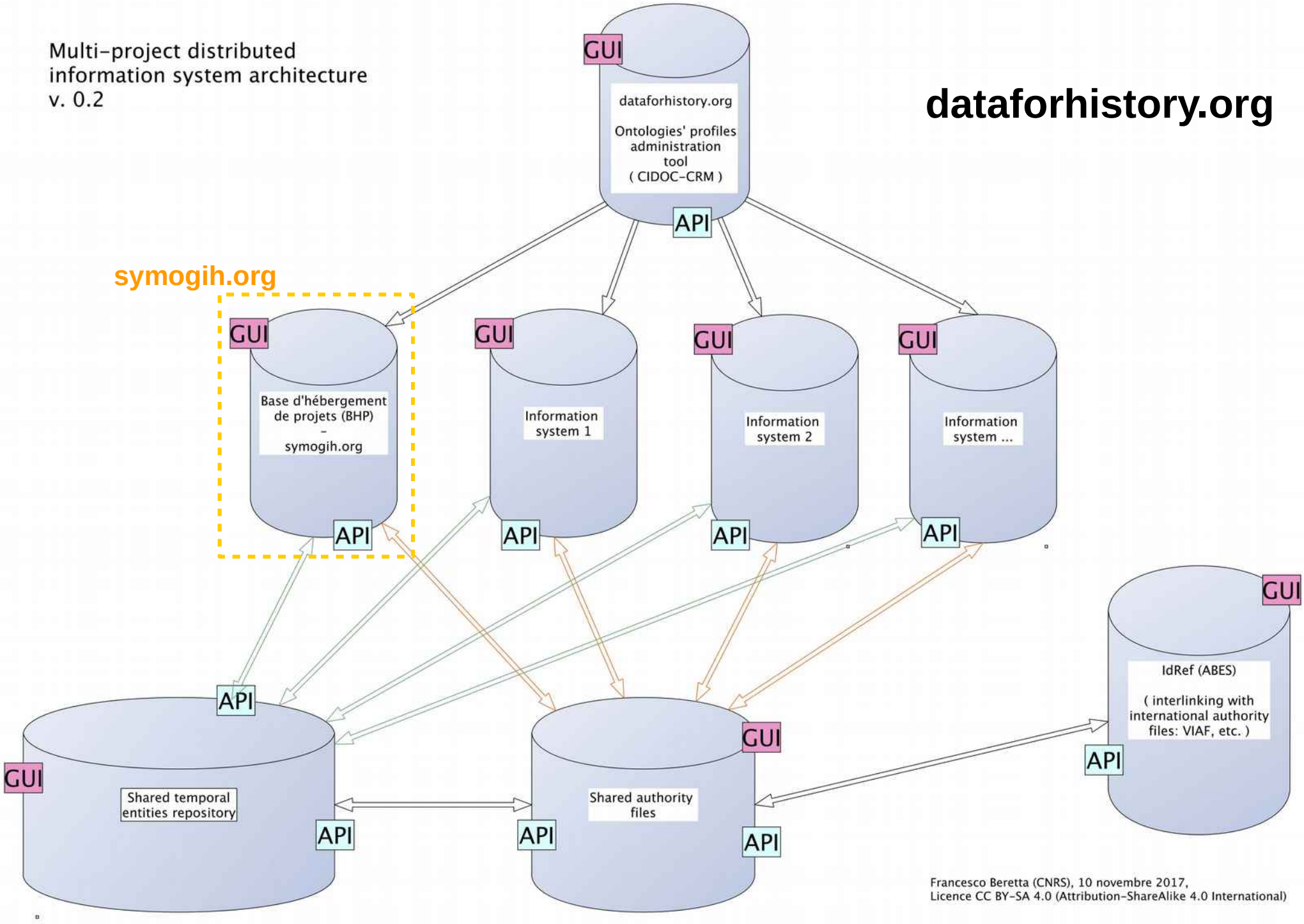
BNF



SPARQL
Endpoint

Digital research environment and data curation platform since 2008

- 50 users
- 20 hosted research projects


Multi-project distributed
information system architecture
v. 0.2
dataforhistory.org
GUI
dataforhistory.org
Ontologies' profiles
administration
tool
(CIDOC-CRM)
API
symogih.org
GUI
Base d'hébergement
de projets (BHP)
-
symogih.org
API
GUI
Information
system 1
API
GUI
Information
system 2
API
GUI
Information
system ...
API
GUI
IdRef (ABES)
(interlinking with
international authority
files: VIAF, etc.)
API
API
GUI
Shared temporal
entities repository
API
GUI
Shared authority
files
API
API
Francesco Beretta (CNRS), 10 novembre 2017,
Licence CC BY-SA 4.0 (Attribution-ShareAlike 4.0 International)

Interlinking URIs of resources

Kepler, Johannes

http://symogih.org/resource/Actr195

Actr195

Année de naissance: 1571 - Année de mort: 1630

Biographie – documentation

Biographie | Informations | Contenus | Carte | Documentation | Liens

Date	Ressource
2005	Depondt, Philippe / Véricourt, Guillemette de, Kepler. L'ort… Editions du Rouergue, 2005)
2003	Bucciantini, Massimo, Galileo e Keplero. Filosofia, cosmol… Einaudi, 2003)
1979	Simon, Gérard, Kepler: astronome, astrologue (Paris, Galli…

Affichage de 1 à 3 sur 3

Kepler, Johannes

Actr195

Année de naissance: 1571 - Année de mort: 1630

Biographie – documentation

Biographie | Informations | Contenus | Carte | Documentation | Liens

Idref – URL identifiant un objet : 026947676

Autorités BnF – identifiant pérenne : cb11909597m

DBPedia Live – URL de ressource : Johannes_Kepler

owl:sameAs

Kepler, Johannes (1571-1630)

http://www.idref.fr/autorites/autorites.html

AMELIORER la notice | SIGNALER une erreur

Précédent | Suivant

026947676

Lien permanent

http://www.idref.fr/026947676

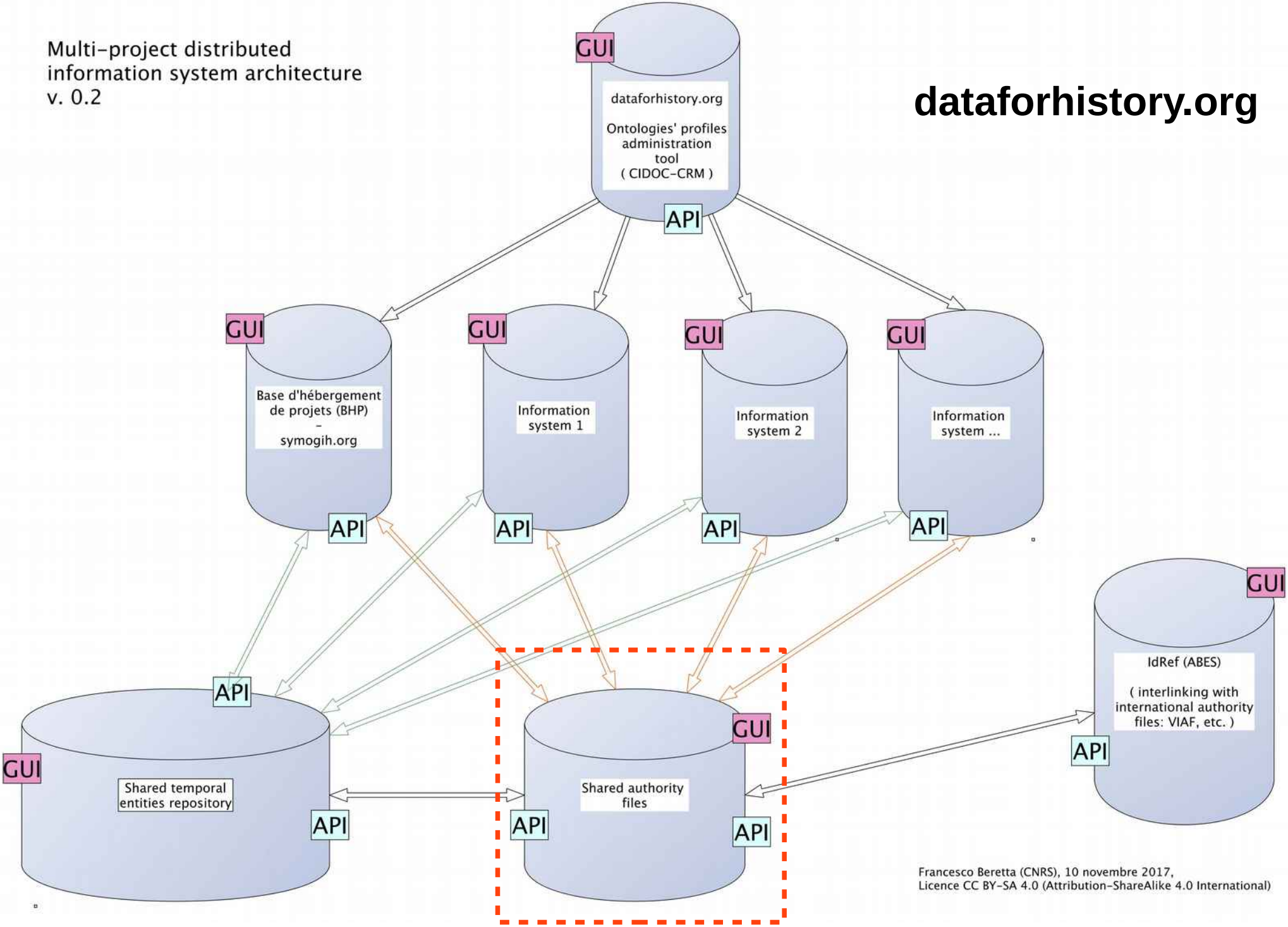
Notice de type
Personne

Forme retenue

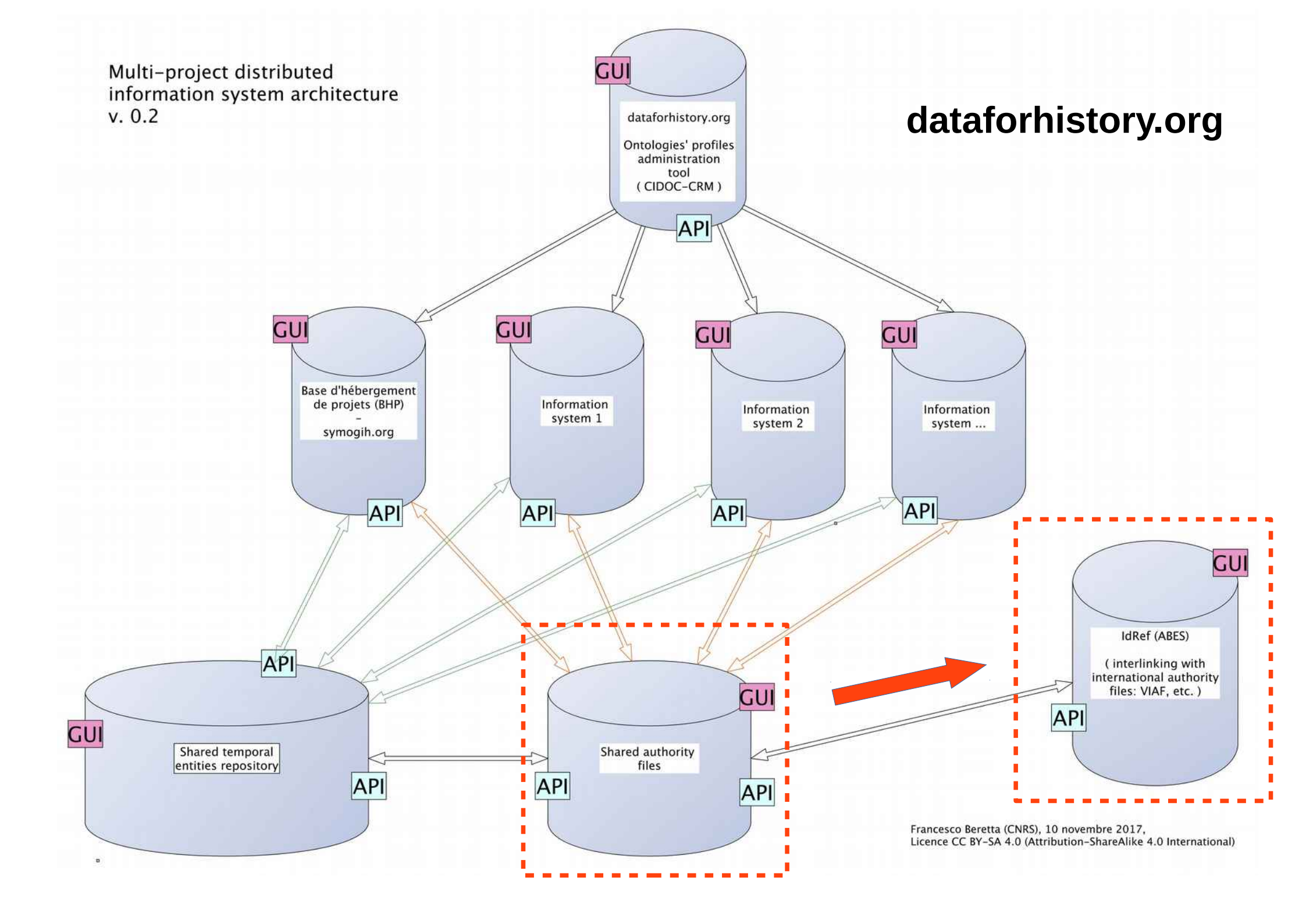
Kepler, Johannes (1571-1630)

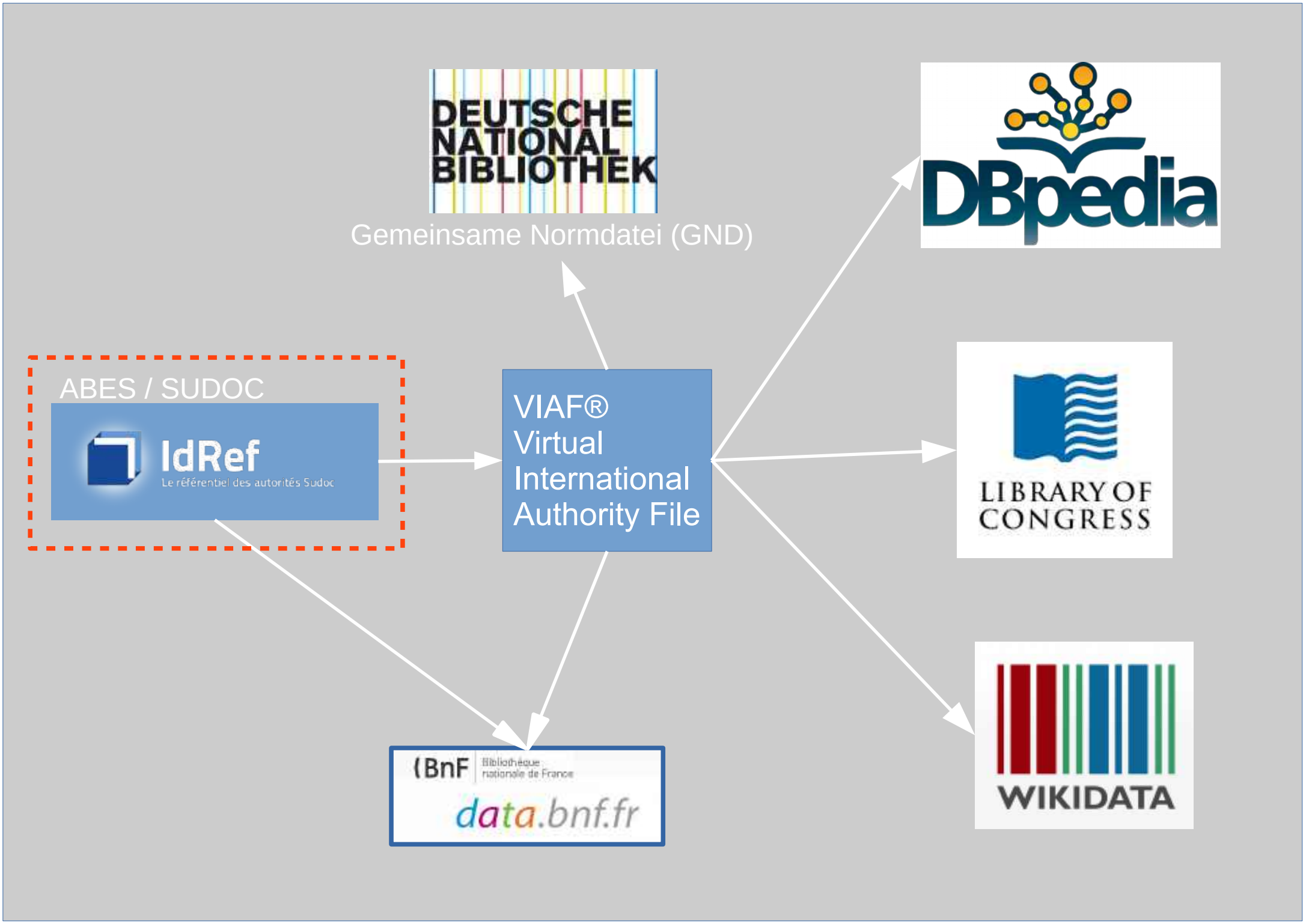
Multi-project distributed
information system architecture
v. 0.2

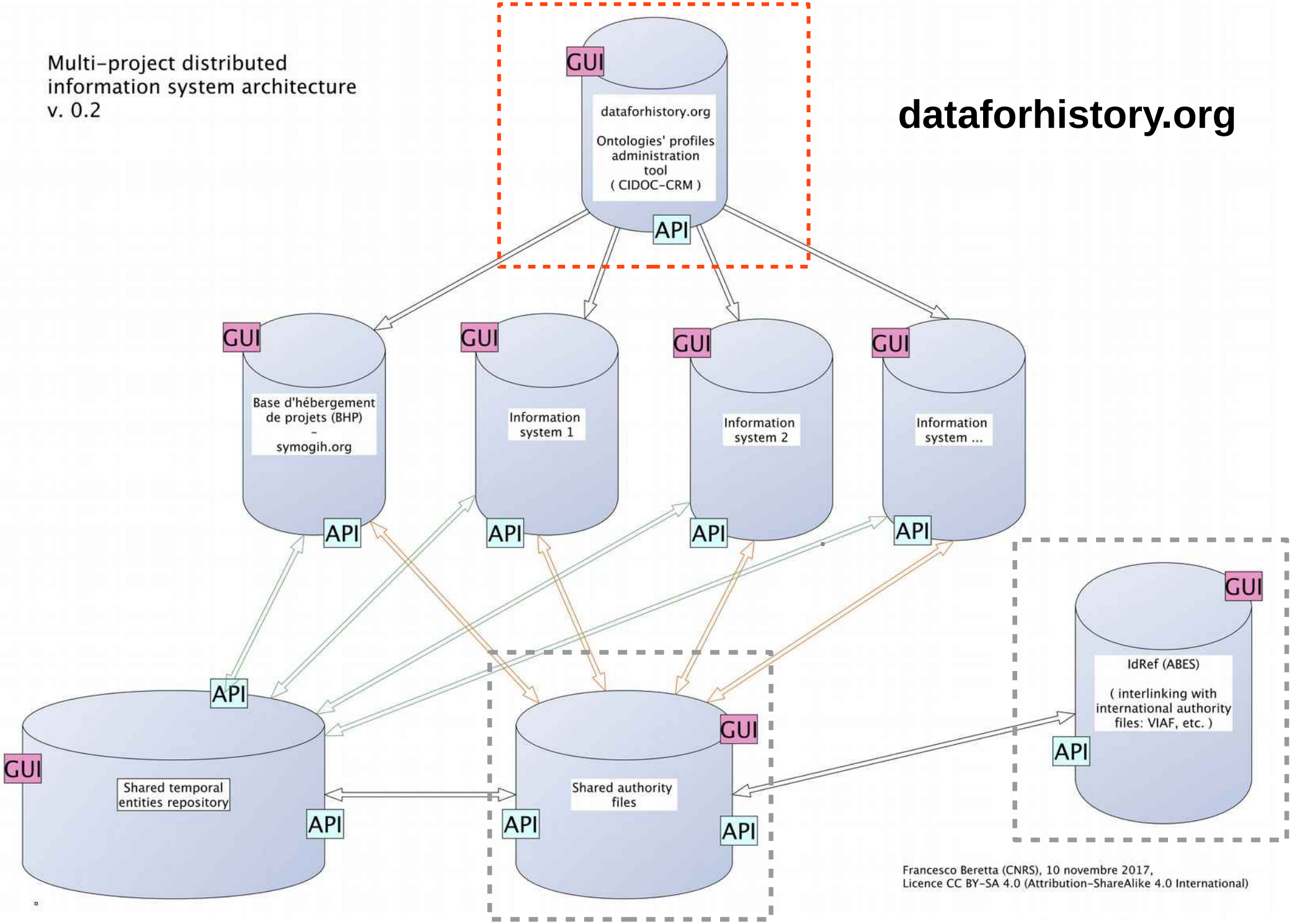
dataforhistory.org

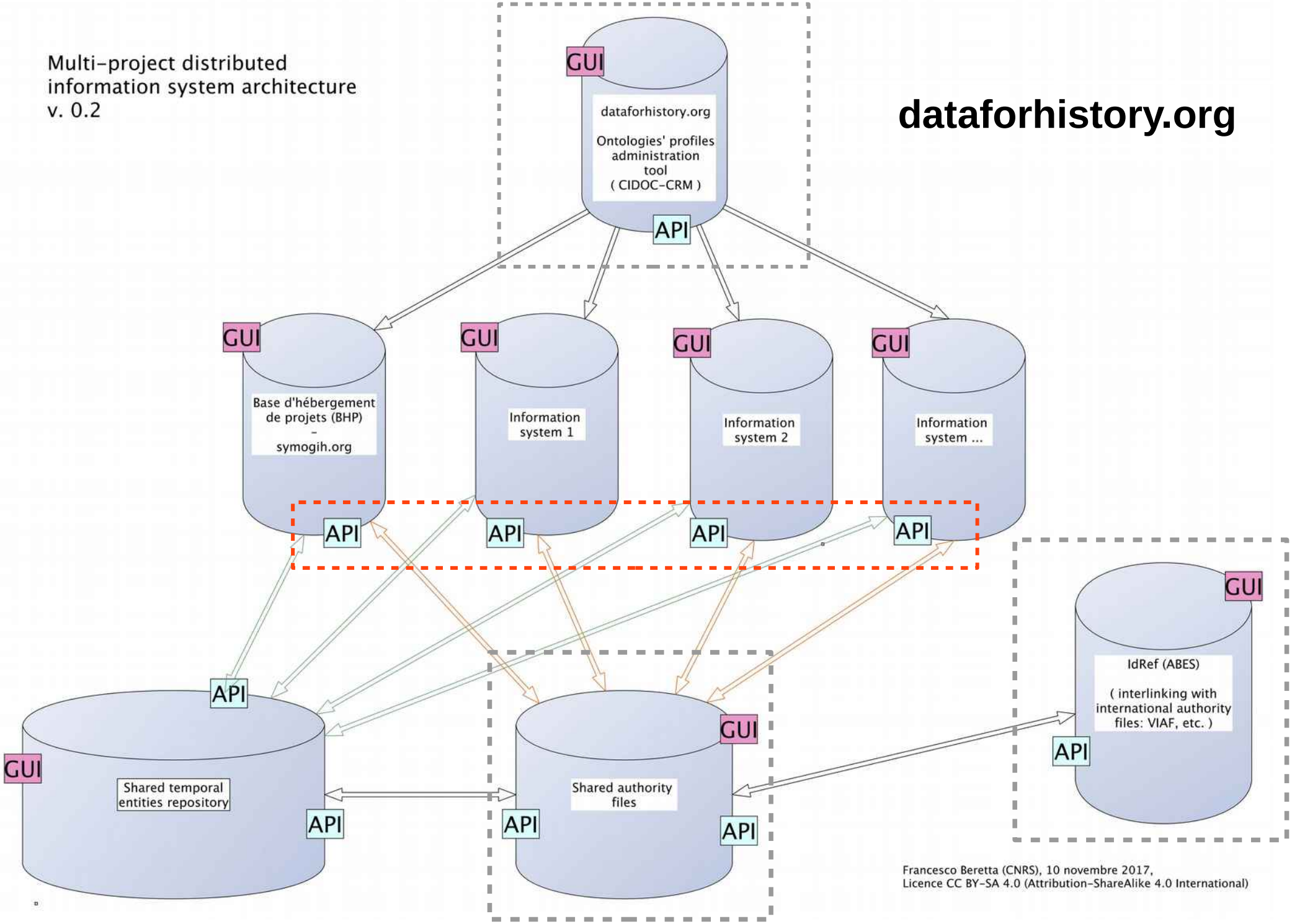


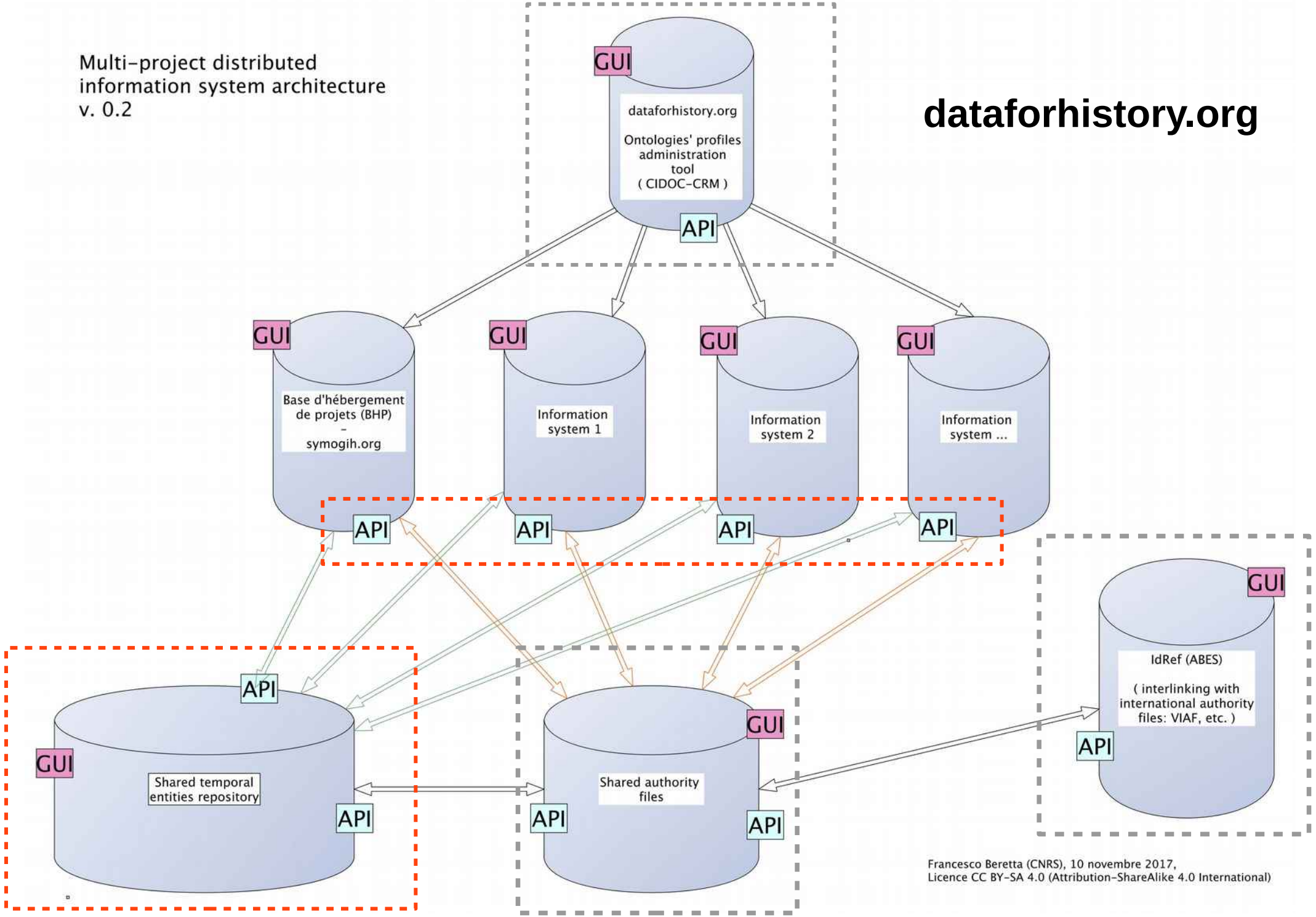
Francesco Beretta (CNRS), 10 novembre 2017,
Licence CC BY-SA 4.0 (Attribution-ShareAlike 4.0 International)


Multi-project distributed
information system architecture
v. 0.2
dataforhistory.org
GUI
dataforhistory.org
Ontologies' profiles
administration
tool
(CIDOC-CRM)
API
GUI
Base d'hébergement
de projets (BHP)
-
symogih.org
API
GUI
Information
system 1
API
GUI
Information
system 2
API
GUI
Information
system ...
API
API
GUI
Shared temporal
entities repository
API
GUI
Shared authority
files
API
API
GUI
IdRef (ABES)
(interlinking with
international authority
files: VIAF, etc.)
API
Francesco Beretta (CNRS), 10 novembre 2017,
Licence CC BY-SA 4.0 (Attribution-ShareAlike 4.0 International)


DEUTSCHE NATIONAL BIBLIOTHEK
Gemeinsame Normdatei (GND)
DBpedia
ABES / SUDOC
IdRef
Le référentiel des autorités Sudoc
VIAF®
Virtual
International
Authority File
LIBRARY OF CONGRESS
BnF Bibliothèque nationale de France
data.bnf.fr
WIKIDATA

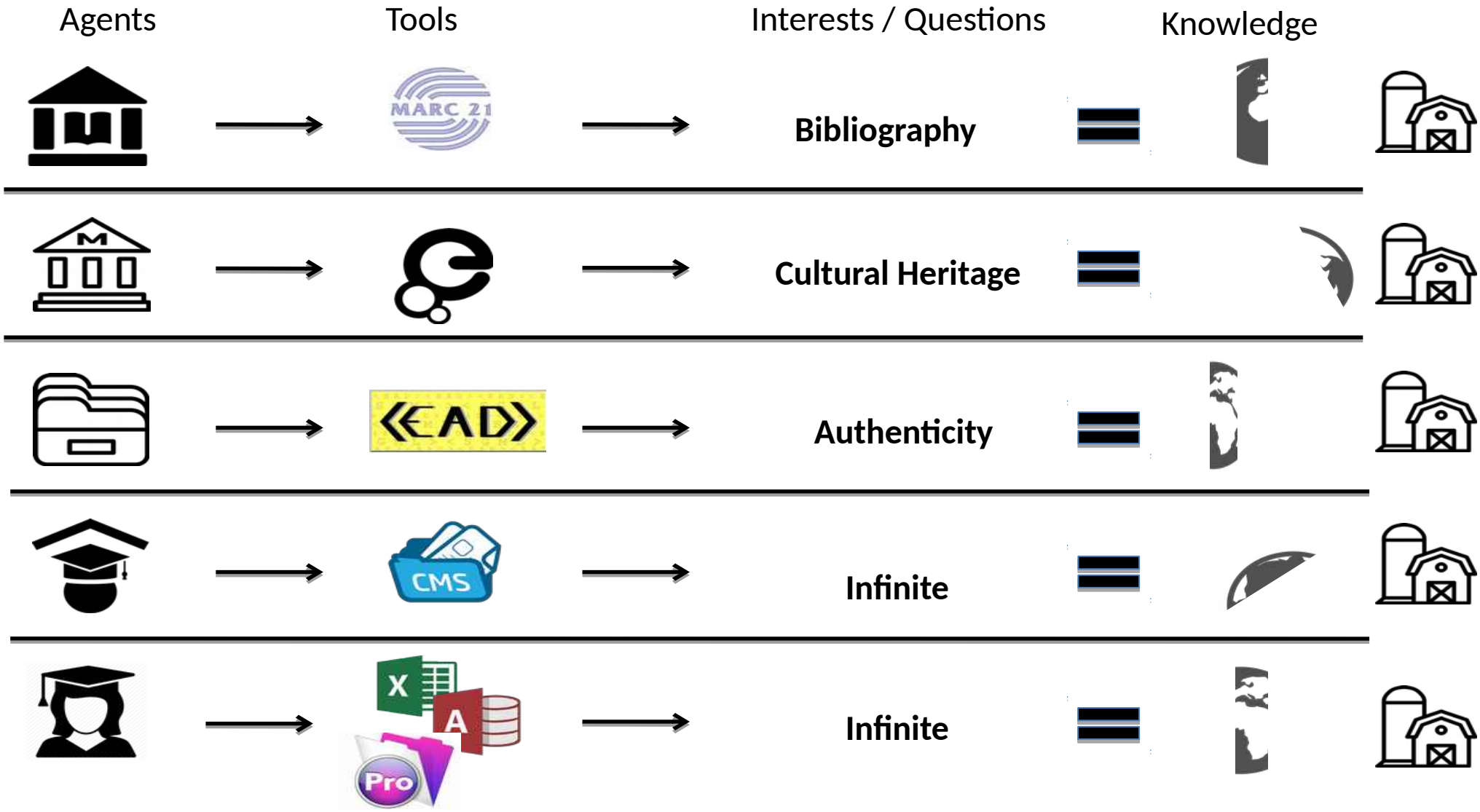

Multi-project distributed
information system architecture
v. 0.2
dataforhistory.org
GUI
dataforhistory.org
Ontologies' profiles
administration
tool
(CIDOC-CRM)
API
GUI
Base d'hébergement
de projets (BHP)
-
symogih.org
API
GUI
Information
system 1
API
GUI
Information
system 2
API
GUI
Information
system ...
API
API
GUI
Shared temporal
entities repository
API
GUI
Shared authority
files
API
API
GUI
IdRef (ABES)
(interlinking with
international authority
files: VIAF, etc.)
API
Francesco Beretta (CNRS), 10 novembre 2017,
Licence CC BY-SA 4.0 (Attribution-ShareAlike 4.0 International)


Multi-project distributed
information system architecture
v. 0.2
dataforhistory.org
GUI
dataforhistory.org
Ontologies' profiles
administration
tool
(CIDOC-CRM)
API
GUI
Base d'hébergement
de projets (BHP)
-
symogih.org
API
GUI
Information
system 1
API
GUI
Information
system 2
API
GUI
Information
system ...
API
API
GUI
Shared temporal
entities repository
API
GUI
Shared authority
files
API
API
GUI
IdRef (ABES)
(interlinking with
international authority
files: VIAF, etc.)
API
Francesco Beretta (CNRS), 10 novembre 2017,
Licence CC BY-SA 4.0 (Attribution-ShareAlike 4.0 International)


Multi-project distributed
information system architecture
v. 0.2
dataforhistory.org
GUI
dataforhistory.org
Ontologies' profiles
administration
tool
(CIDOC-CRM)
API
GUI
Base d'hébergement
de projets (BHP)
-
symogih.org
API
GUI
Information
system 1
API
GUI
Information
system 2
API
GUI
Information
system ...
API
GUI
Shared temporal
entities repository
API
API
GUI
Shared authority
files
API
API
GUI
IdRef (ABES)
(interlinking with
international authority
files: VIAF, etc.)
API
Francesco Beretta (CNRS), 10 novembre 2017,
Licence CC BY-SA 4.0 (Attribution-ShareAlike 4.0 International)

Historical Knowledge Production in Digital Environments

Actual



The CIDOC Conceptual Reference Model

ICOM international council of museums

CIDOC CRM CONCEPTUAL REFERENCE MODEL

1996
- Formation of CIDOC CRM Special Interest Group (SIG) under ICOM

2003
- International Working Group on FRBR/CIDOC CRM Harmonization formed

2006
- Acceptance as ISO Standard ISO21127:2006
- First Draft of FRBRoo

2014
- CRMdig Released
- ISO Standard reaffirmed

2015
- CRMarchaeo formulated under Ariadne Project
- CRMgeo released

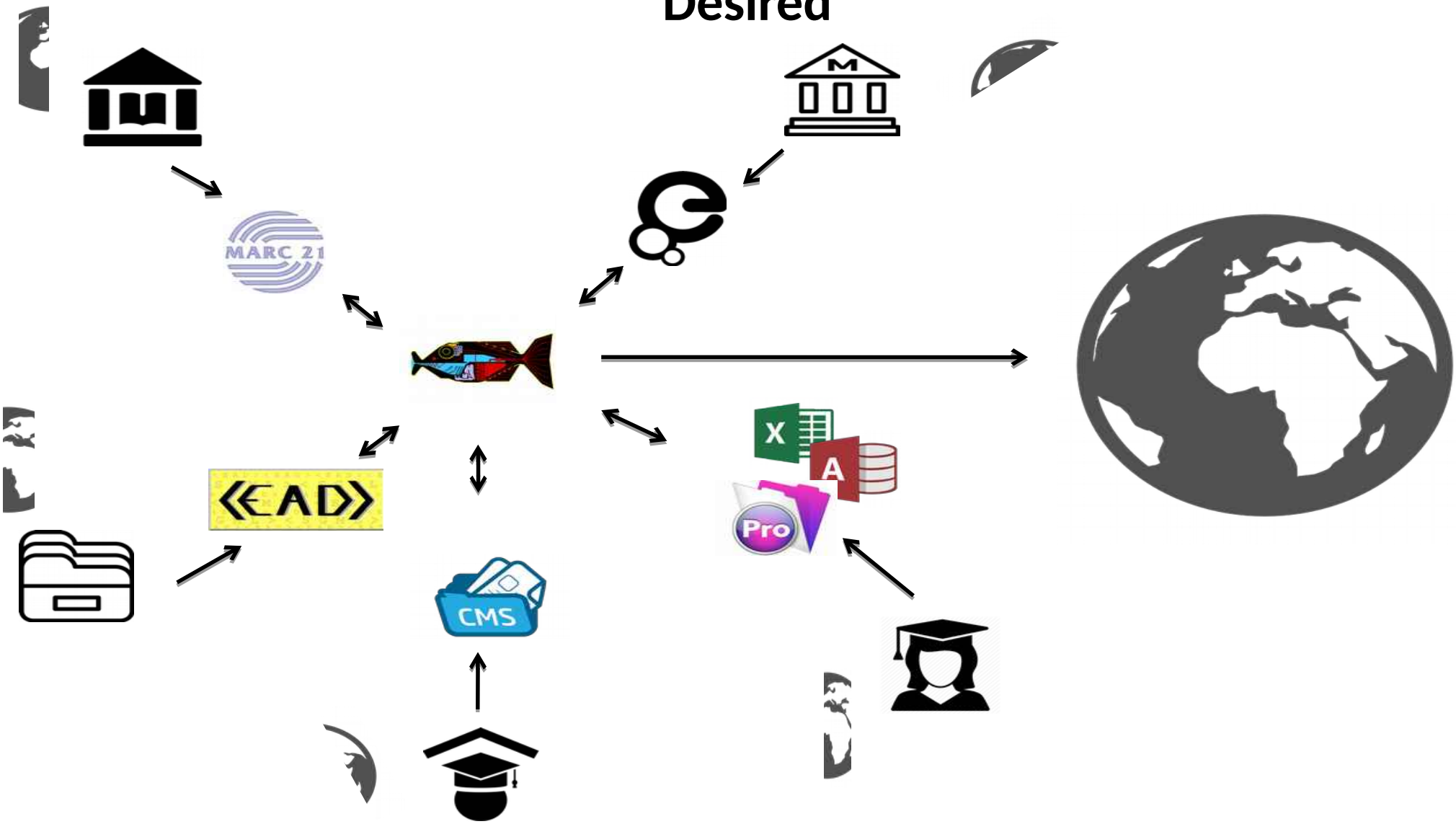
2017
- Data for History Launch

What it is and what it does

- *a **core ontology** describing the underlying semantics of database schemata and structures from historical knowledge producing organizations and individuals.*
- *a **generic model of recording of "what has happened"** in human scale*
- ***generates huge, meaningful networks of knowledge by a simple abstraction**: history as meetings of people, things and information*
- *the **result of 20 years** of interdisciplinary work and agreement*

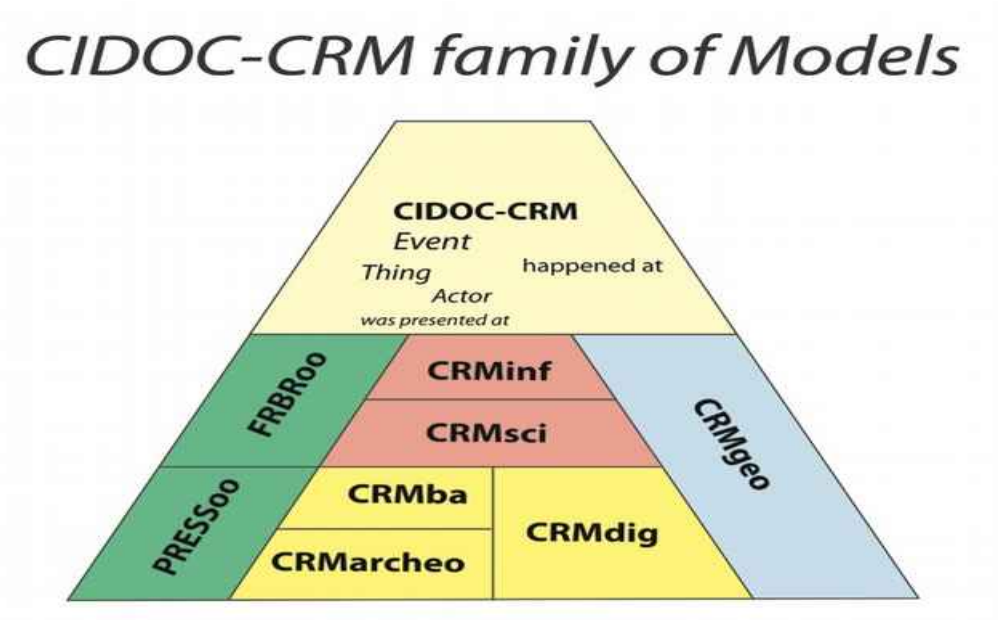
Historical Knowledge Production in Digital Environments:

Desired



Present State

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

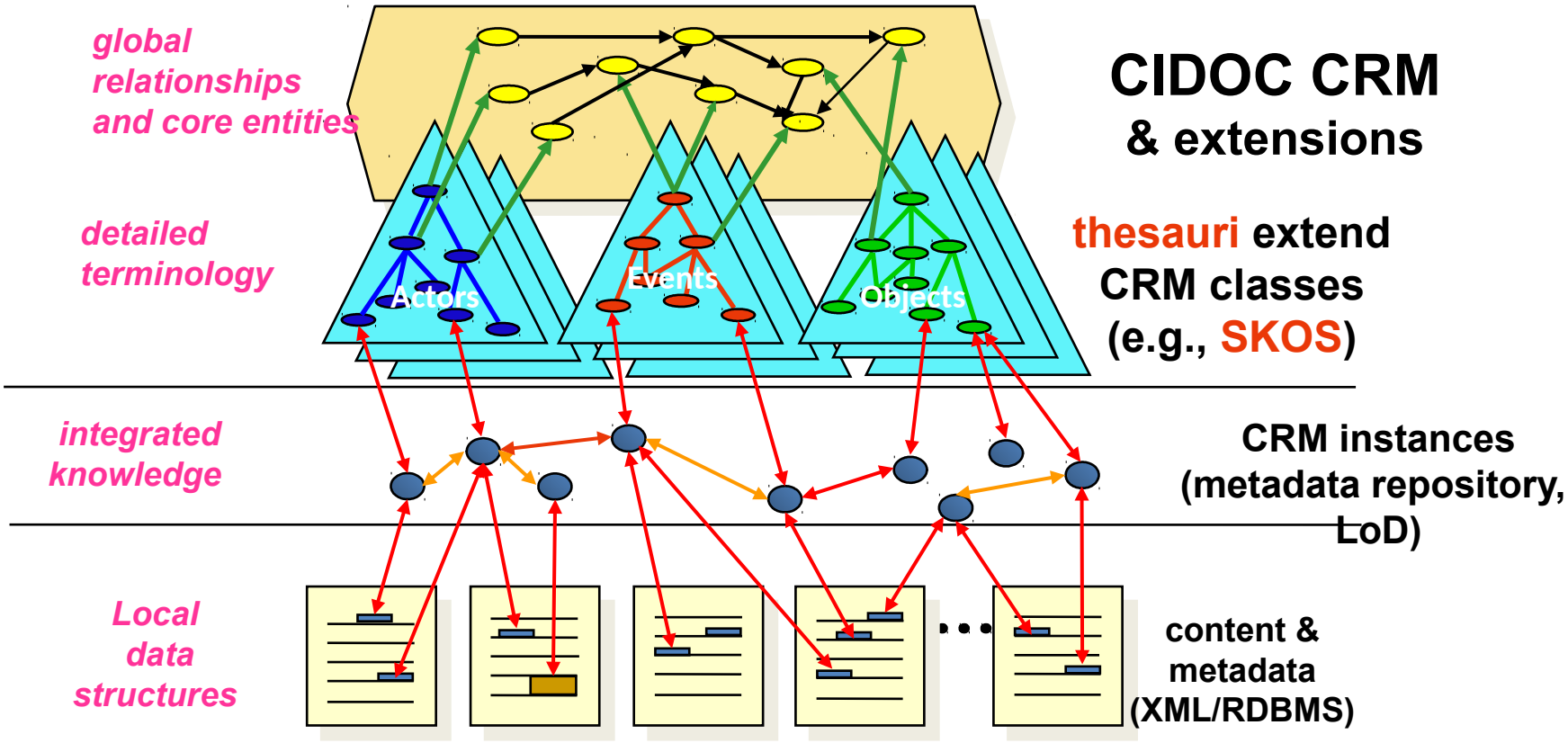


High Recall harmonized to detailed knowledge practice modelling

Extensions

Name	Scope	Community of Practice	Version
CRMBase	CH and E Science	CH Specialists	6.2.2
FRBRoo	Bibliographic Data and Creative Processes	Librarians	2.4
PreSSoo	Serials Data	Librarians	1.2
CRMinf	Argumentation		0.7
CRMsci	E-sciences	Analytic heritage science community	1.2.4
CRMdig	Digitization processes	3D modelling community	3.2.1
CRMarchaeo	Excavation practice	Archaeologists	1.4.3
CRMba	Building archaeology	Archaeologists	1.4
CRMgeo	Geophysics and geolocation	Archaeologists and Geophysicists	1.2

Function



- Support interoperability of mutually relevant data sources
- enable sourced and verifiable facts from datasets
- foster referenceability and reusability of data
- foster structured argumentation on top of facts

CIDOC CRM CONCEPTUAL REFERENCE MODEL

Successful Implementations

Museums/Institutions

The British Museum

INSTITUTUM ARCHAEOLOGICUM GERMANICUM

GERMANISCHES NATIONAL MUSEUM

RIJKS MUSEUM

GETTY

European Networks

ARIADNE

DOREMUS
DOing REusable MUSical data

PARTHENOS
Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies

BlueBRIDGE

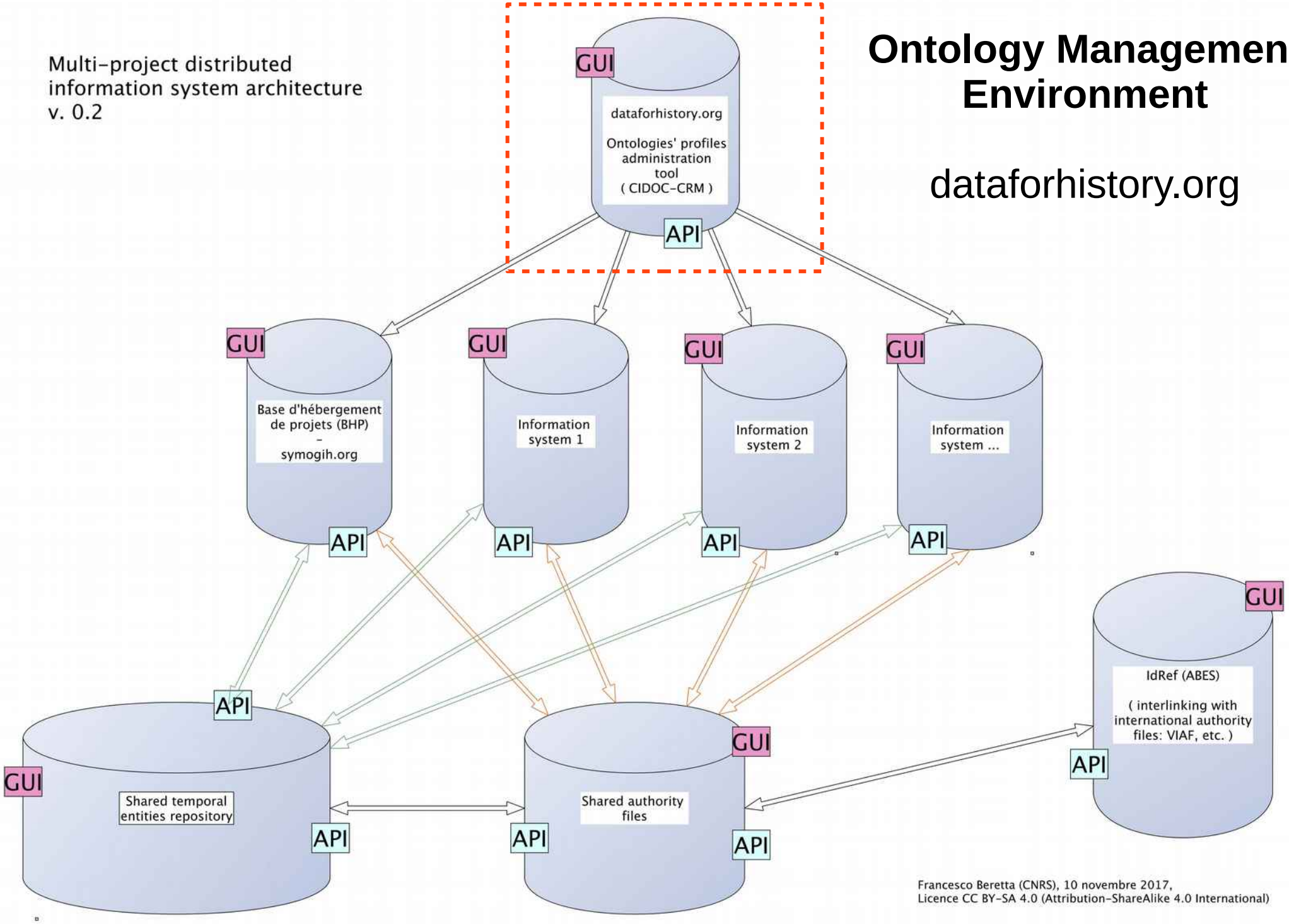
Global

american art COLLABORATIVE
Linked Open Data Initiative

Multi-project distributed
information system architecture
v. 0.2

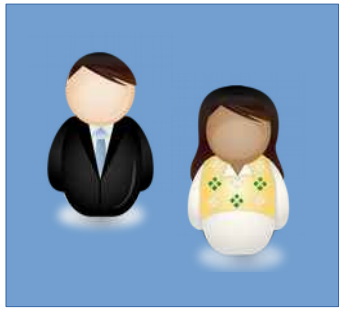
Ontology Management Environment

dataforhistory.org



Francesco Beretta (CNRS), 10 novembre 2017,
Licence CC BY-SA 4.0 (Attribution-ShareAlike 4.0 International)

Research Project



creates

Application Profile

Ontology Management Environment

dataforhistory.org

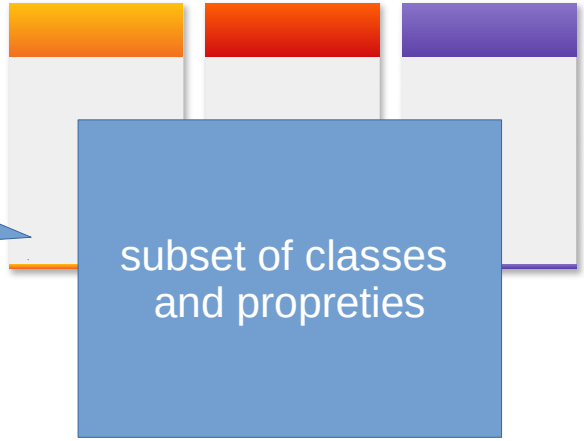
Research Project

creates

Application Profile

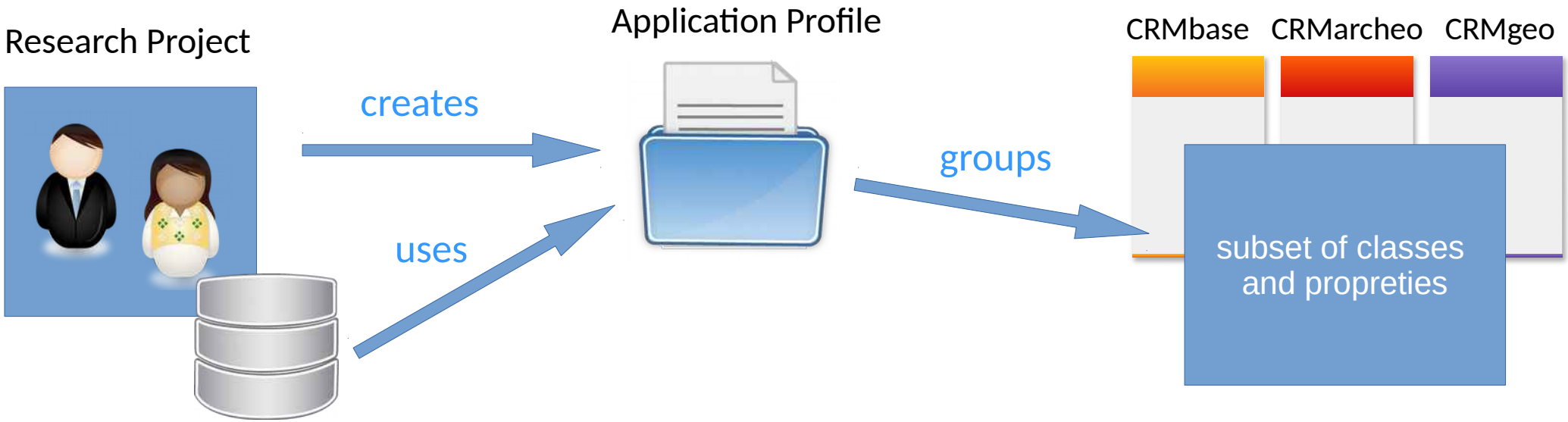
groups

CRMbase CRMarcheo CRMgeo



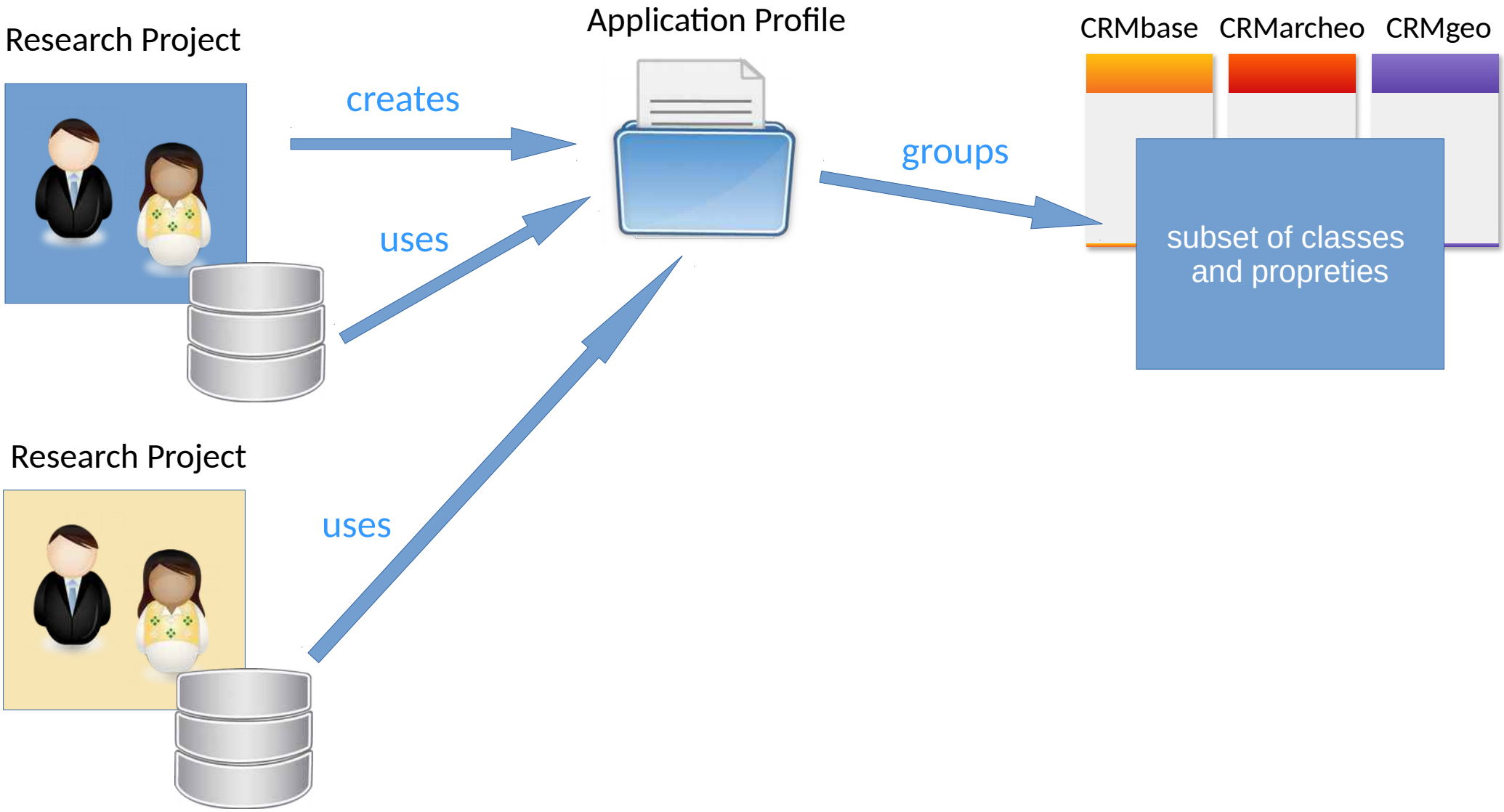
Ontology Management Environment

dataforhistory.org



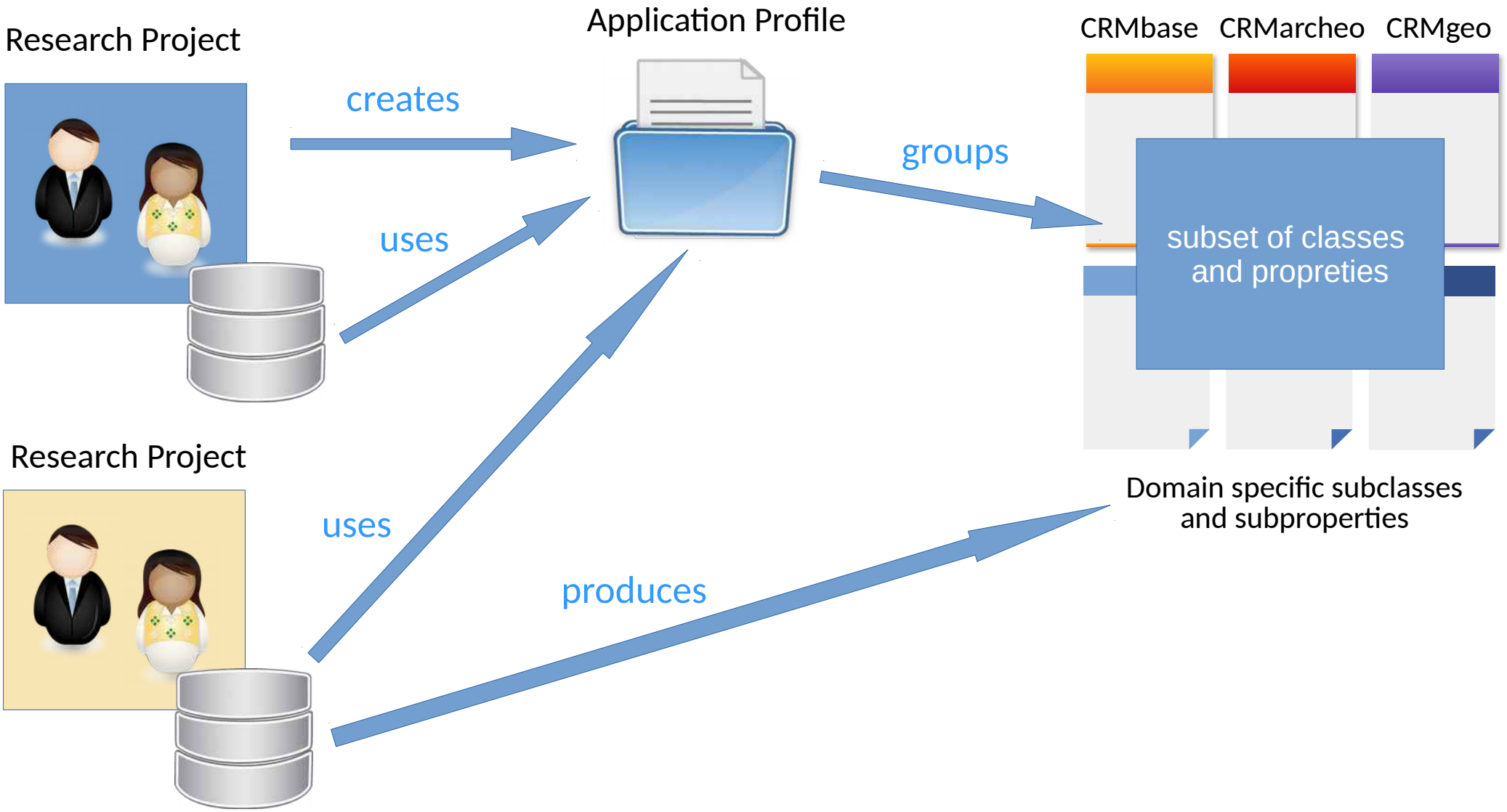
Ontology Management Environment

dataforhistory.org



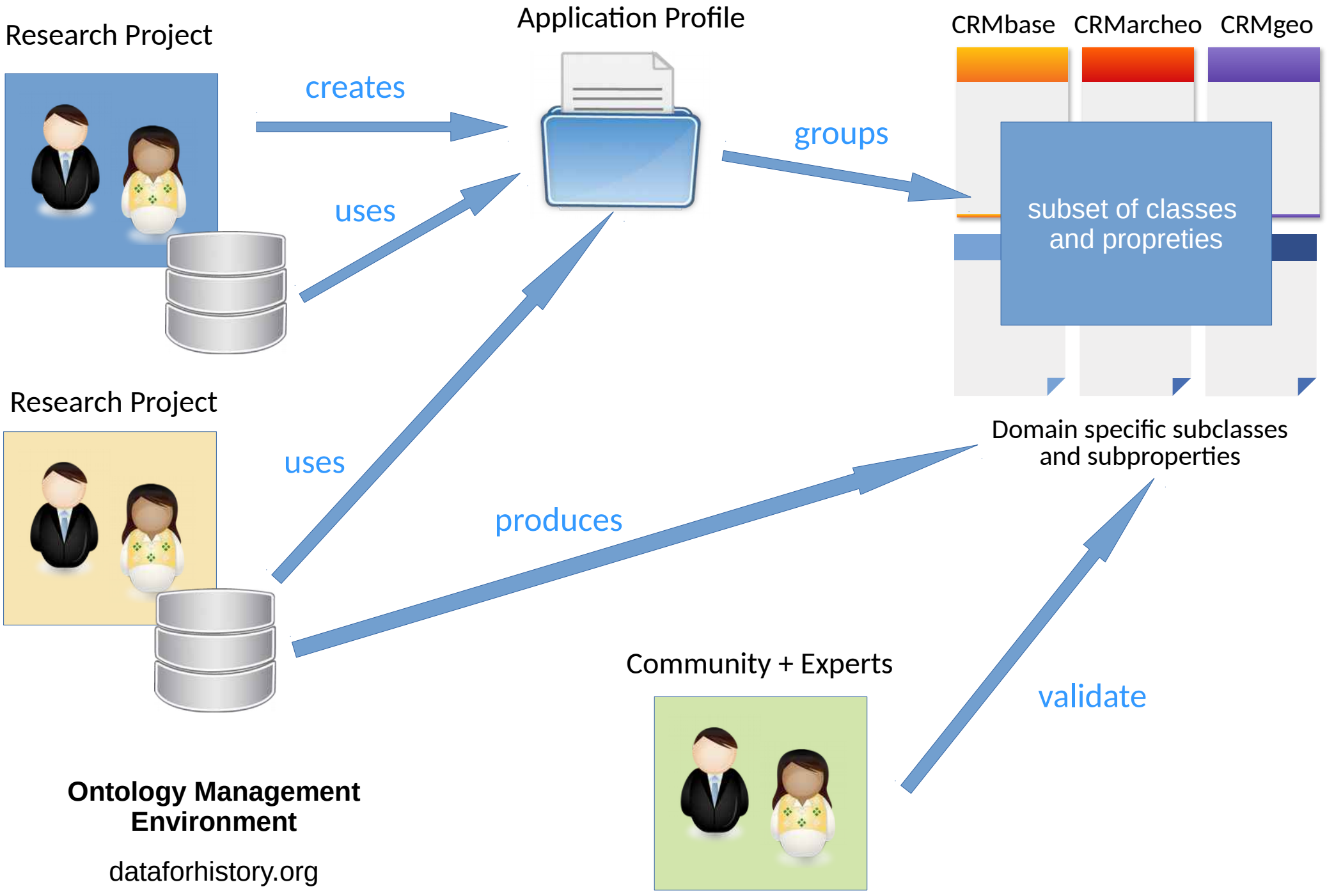
Ontology Management Environment

dataforhistory.org

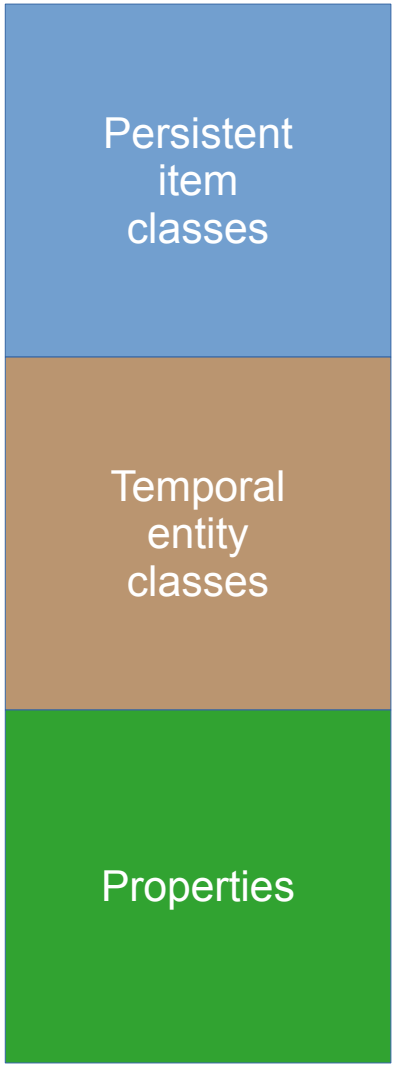


Ontology Management Environment

dataforhistory.org

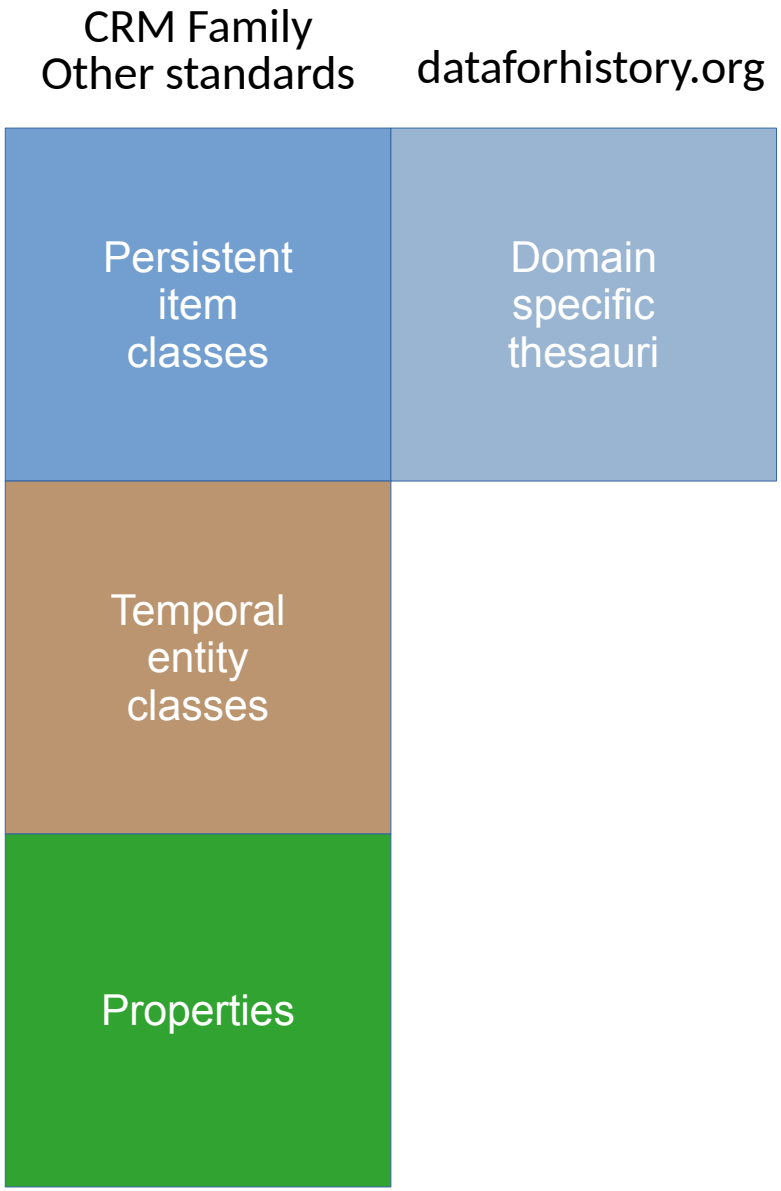

Research Project
creates
Application Profile
CRMbase
CRMarcheo
CRMgeo
groups
uses
subset of classes
and propreties
Research Project
uses
Domain specific subclasses
and subproperties
produces
Community + Experts
validate
Ontology Management
Environment
dataforhistory.org

CRM Family
Other standards



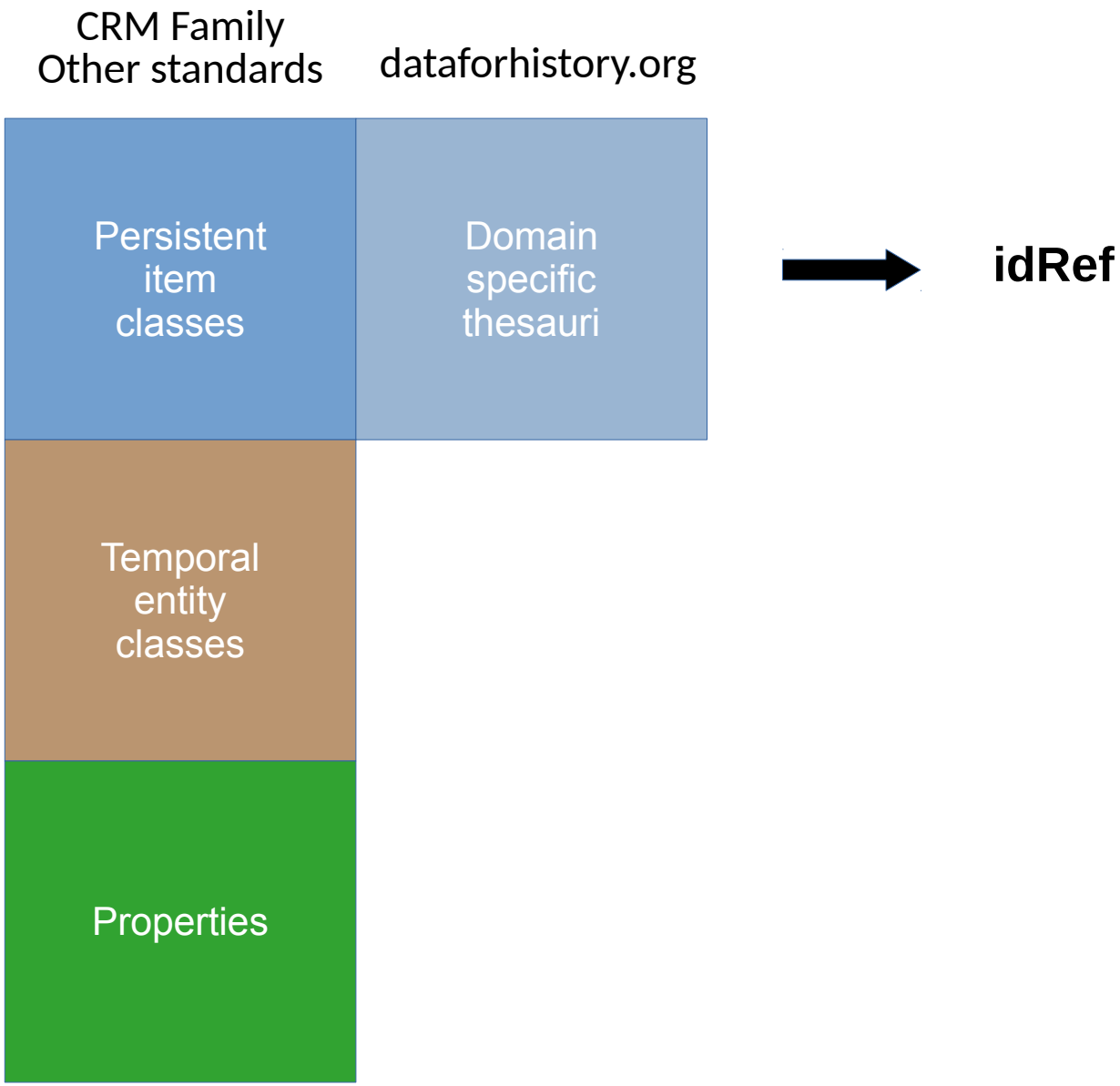
Ontology Management Environment

dataforhistory.org



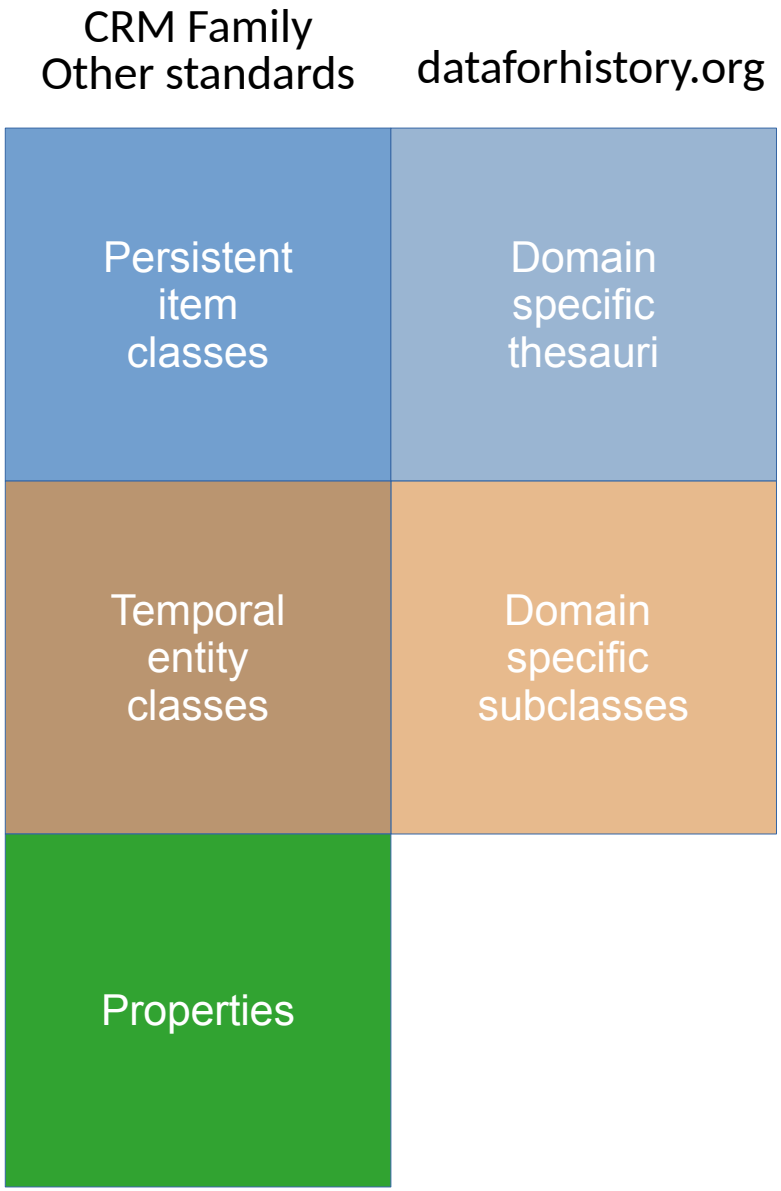
Ontology Management Environment

dataforhistory.org



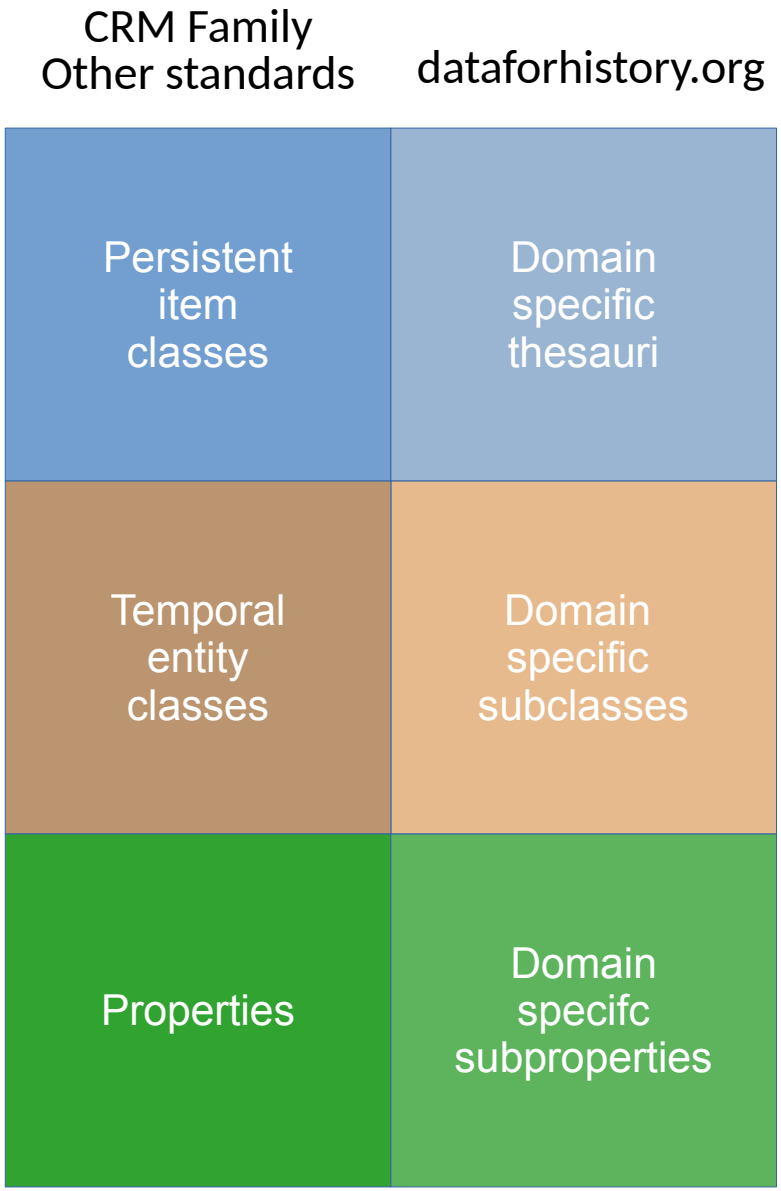
Ontology Management Environment

dataforhistory.org



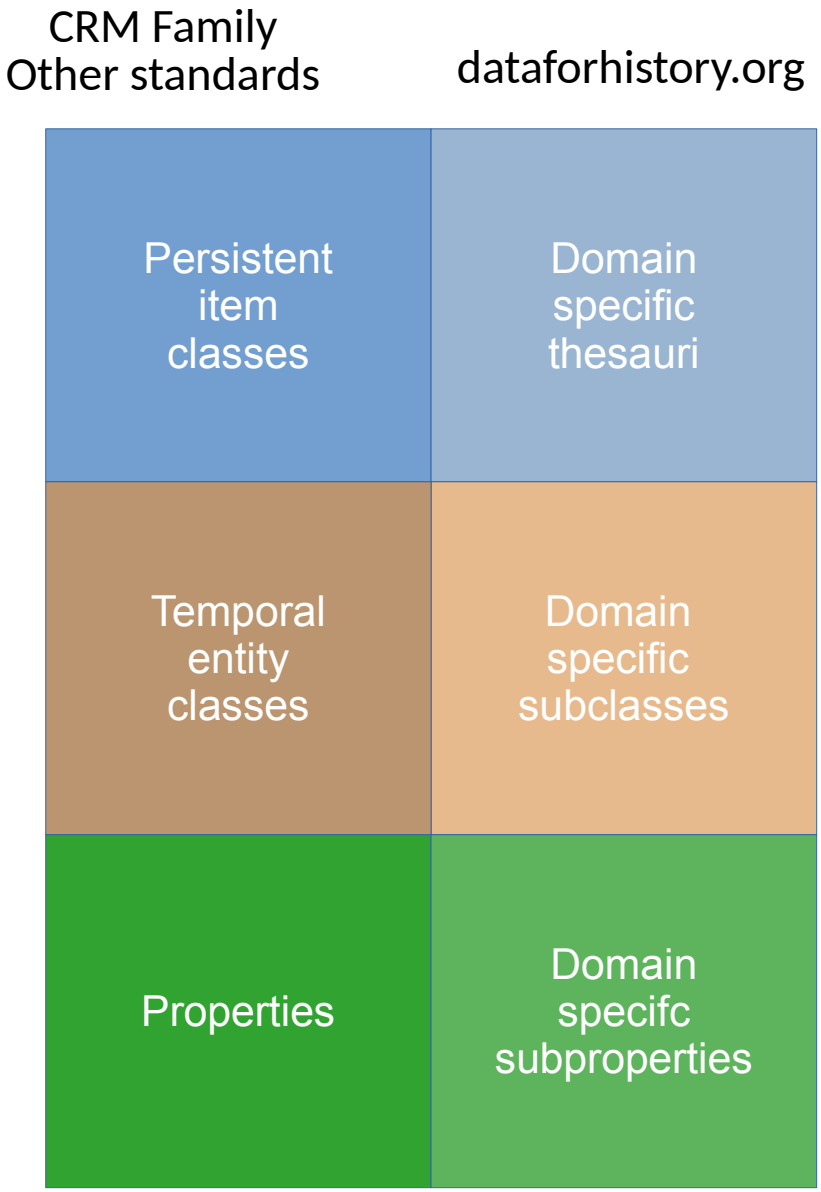
Ontology Management Environment

dataforhistory.org



Ontology Management Environment

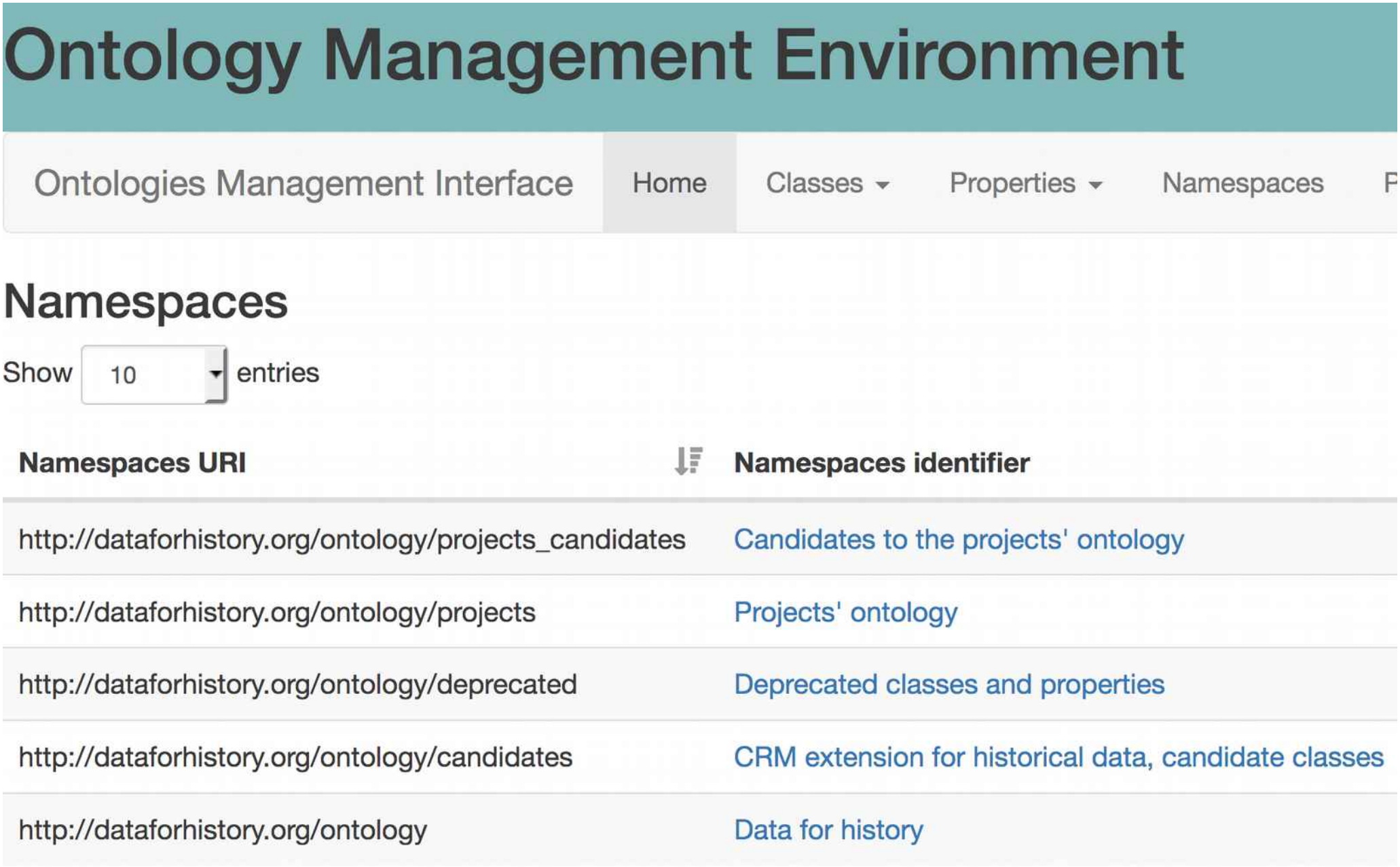
dataforhistory.org



- **Unique identifiers**
  - **histT1**
  - **histC1**
  - **histP1**
- **Namespaces**

Ontology Management Environment

dataforhistory.org



dataforhistory.org
http://phn-dev.ish-lyon.cnrs.fr/ontologies/web/

Ontology Management Environment

Ontologies Management Interface | Home | Classes | Properties | Namespaces | Profiles

BHP new data model v. 0.1

symogih.org

Profile identification | Classes | Properties | Projects

Associated classes

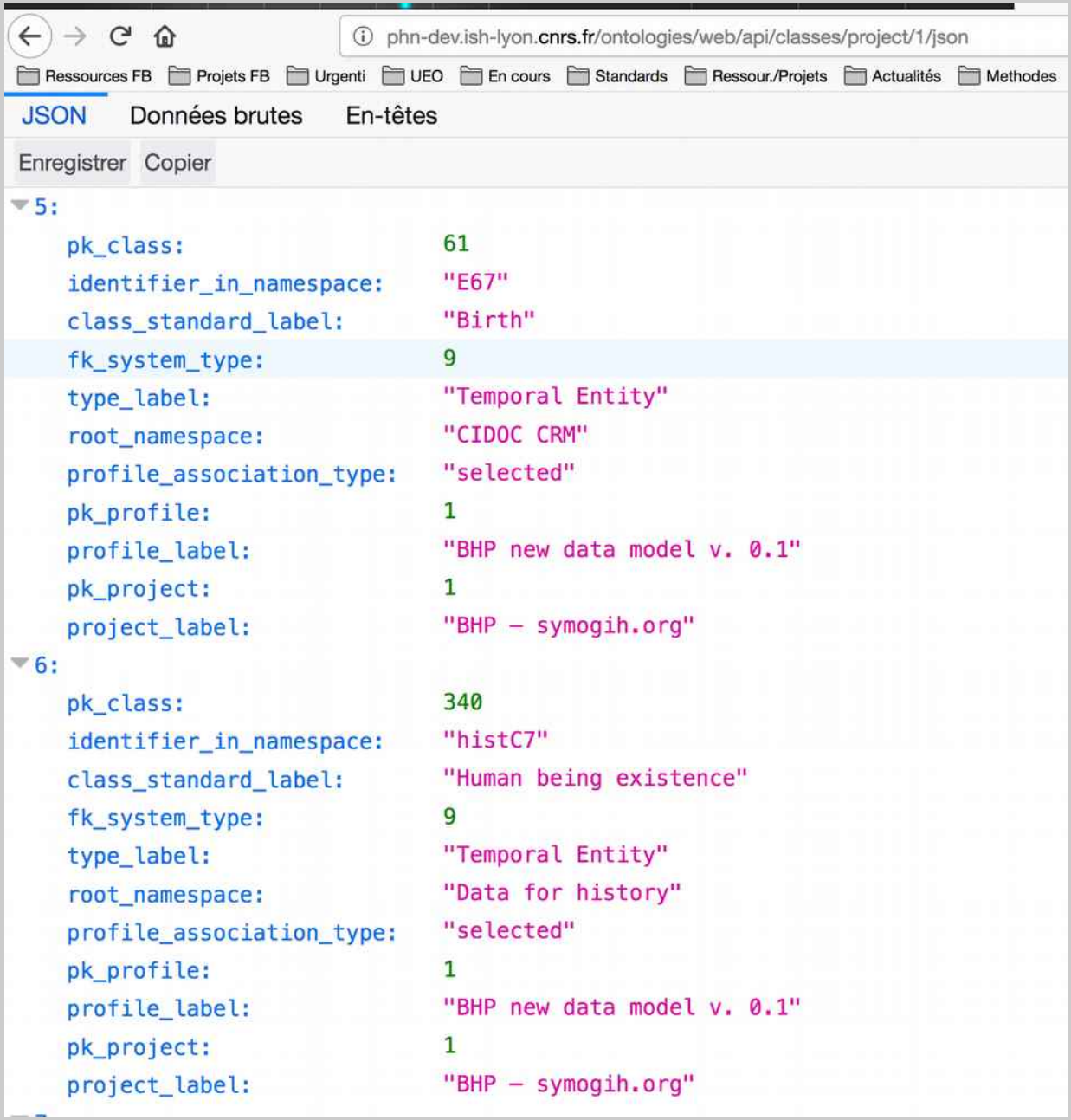
Show 10 entries

Class identifier	Association type	Namespace URI
E7	selected	CIDOC CRM
E70	associated	CIDOC CRM
E71	associated	CIDOC CRM
histC7	selected	Data for history

dataforhistory.org
http://phn-dev.ish-lyon.cnrs.fr/ontologies/web/

Retrieve your project's application profiles from an API

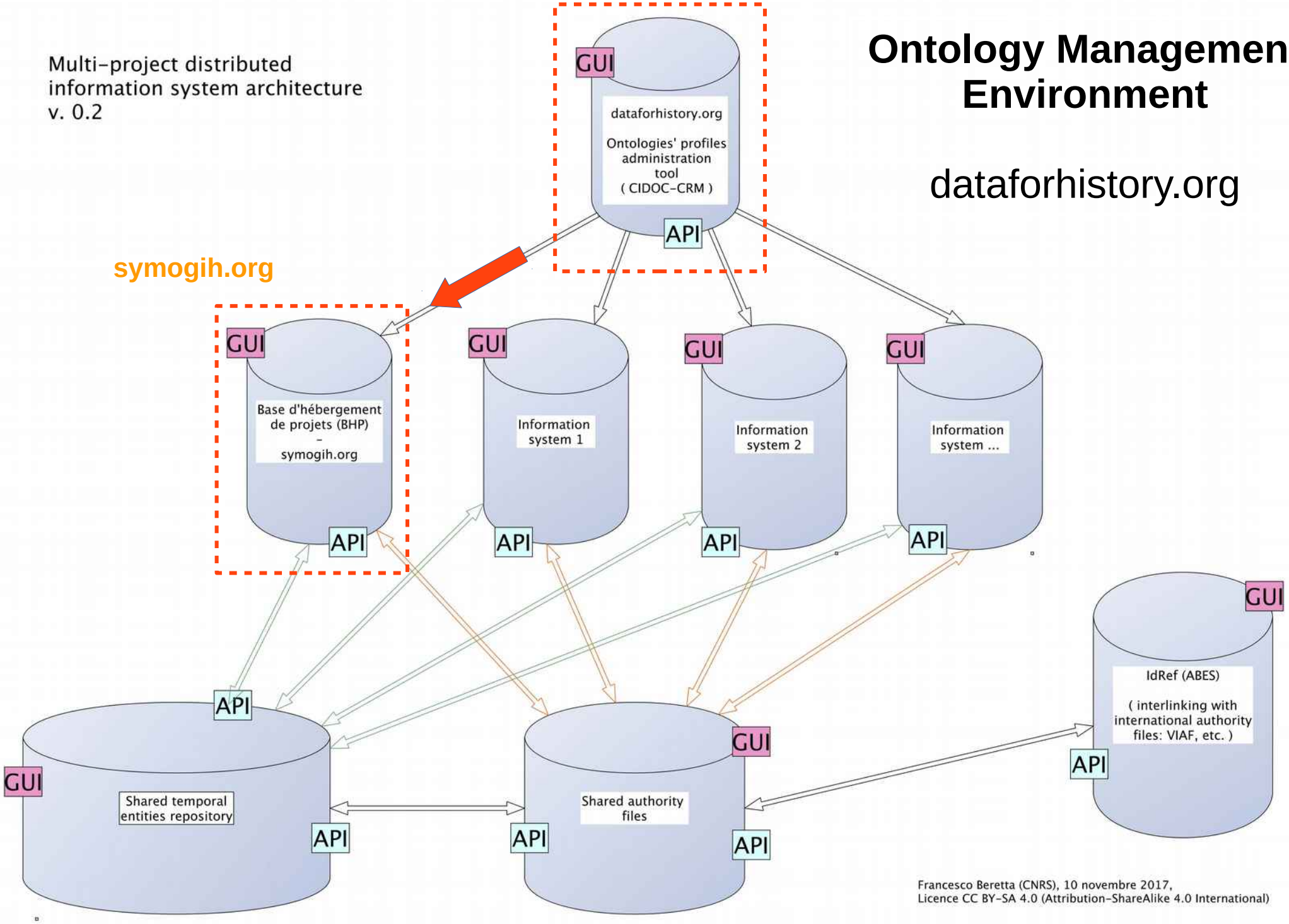
http://phn-dev.ish-lyon.cnrs.fr/ontologies/web/api/classes/project/1/json



Ontology Management Environment

dataforhistory.org

Multi-project distributed
information system architecture
v. 0.2



Francesco Beretta (CNRS), 10 novembre 2017,
Licence CC BY-SA 4.0 (Attribution-ShareAlike 4.0 International)

The Research Infrastructure Challenge

Horizon 2020 Work Programme 2016 - 2017

"The EU framework for Research and Innovation, **Horizon 2020, gives high importance to promoting world-class-research infrastructures**."

"Research infrastructures help to structure the scientific community and play **a key role in the construction of an efficient research and innovation environment**."

They have an ability to **generate** '**a critical mass of people, knowledge and investment**'.

"**e-Infrastructures will make every European researcher *digital***, increasing creativity and efficiency of research and bridging the divide between developed and less developed regions"

Information Integration as Key Aspect of RI Challenge

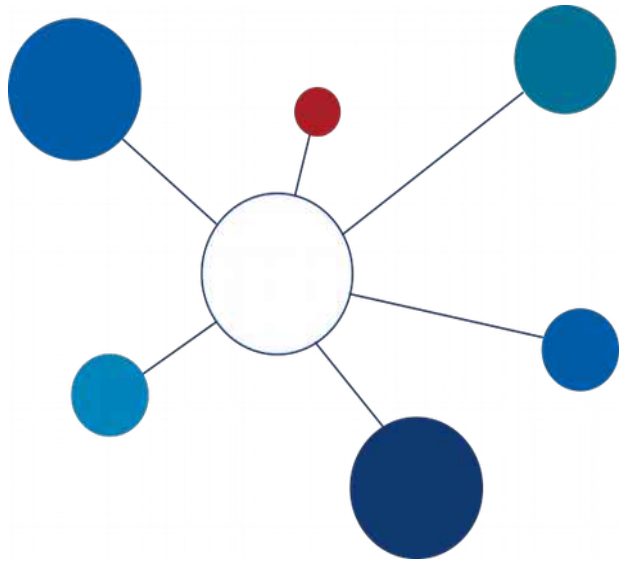
<u>What constitutes an RI?</u>

1. Major Equipment
2. **Knowledge Resources (collections, archives, scientific data)**
3. E-Infrastructure (data and computing systems)

Can be:

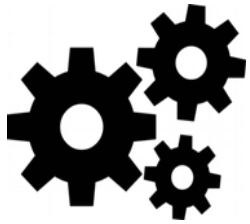
Single Sited, **Virtual**, Distributed

To move towards meeting the intention of the above goals, an RI must consider and develop a strategy towards data integration.

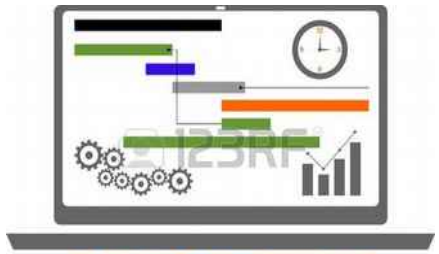


An effective strategy to create integrated data resources is the only means to meet the goals of efficiency and open access to research and resources that are hallmarks of an RI in a digital age.

Before Integration we need a picture of the world to be integrated



Services

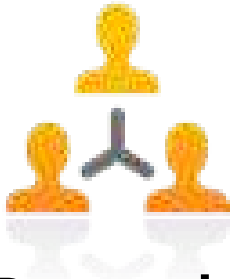


Projects

Datasets

Software

Actors



Research Infrastructures

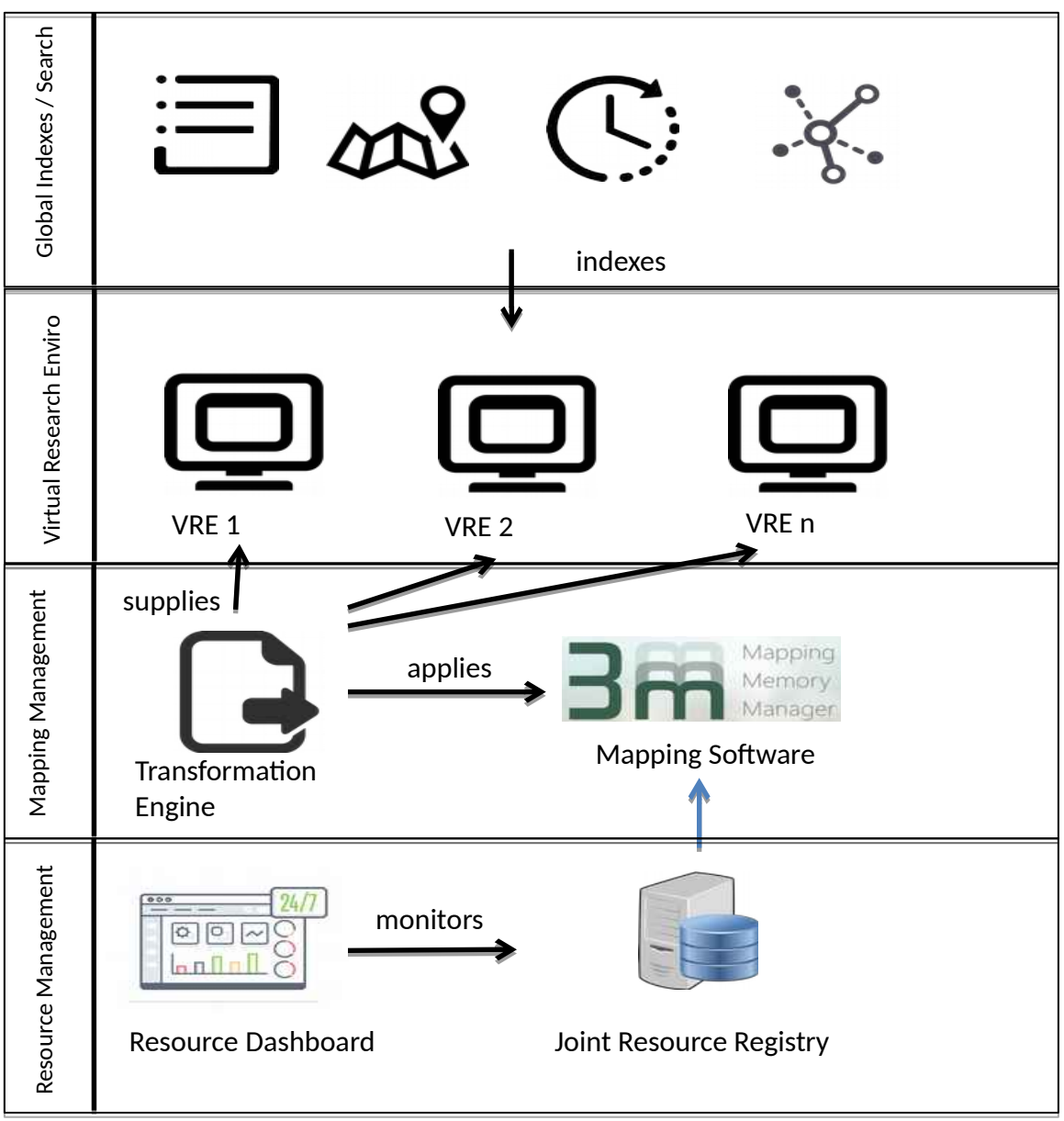
Meta-metadata and the database of databases…

An adequate Data or Conceptual Model for Ris should provide a framework for describing and mapping basic relations that hold between:

Services, Actors, Projects, Datasets, Software, Research Infrastructures

This will **provide** not yet another aggregation but **a picture of the research infrastructure and information integration landscape itself to identify, support and build useful and sustainable aggregations**.

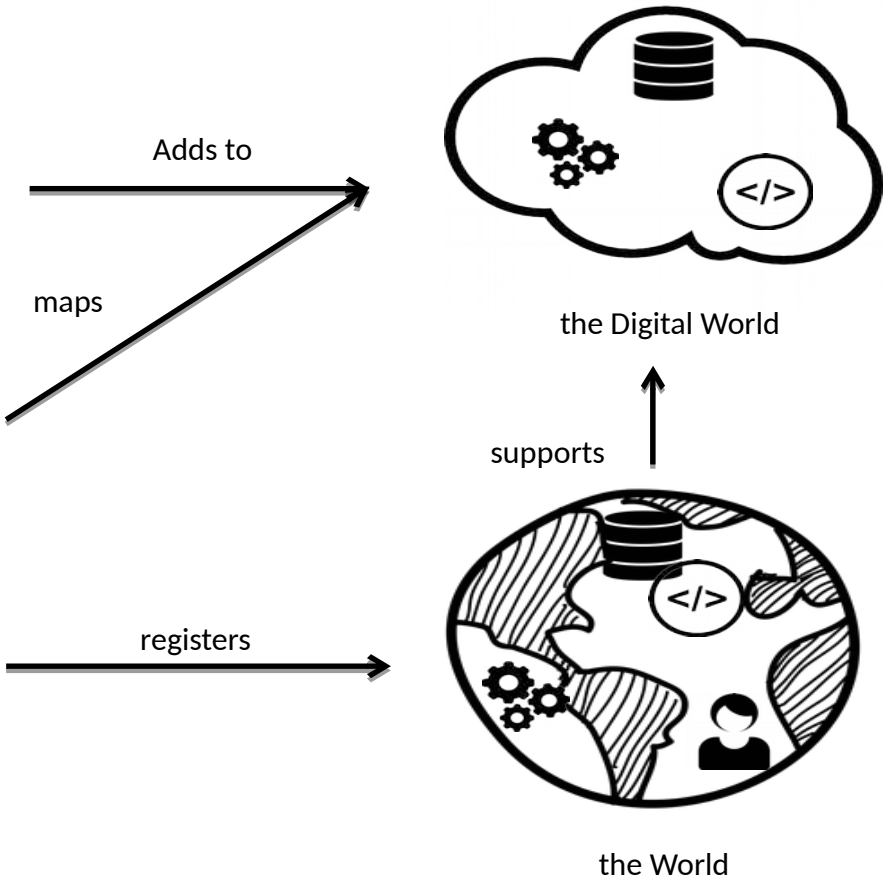
This picture serves the function of determining what has been integrated, where and by who and gives the ability to plan and managed the continuous activity of multiple integrations for different purposes



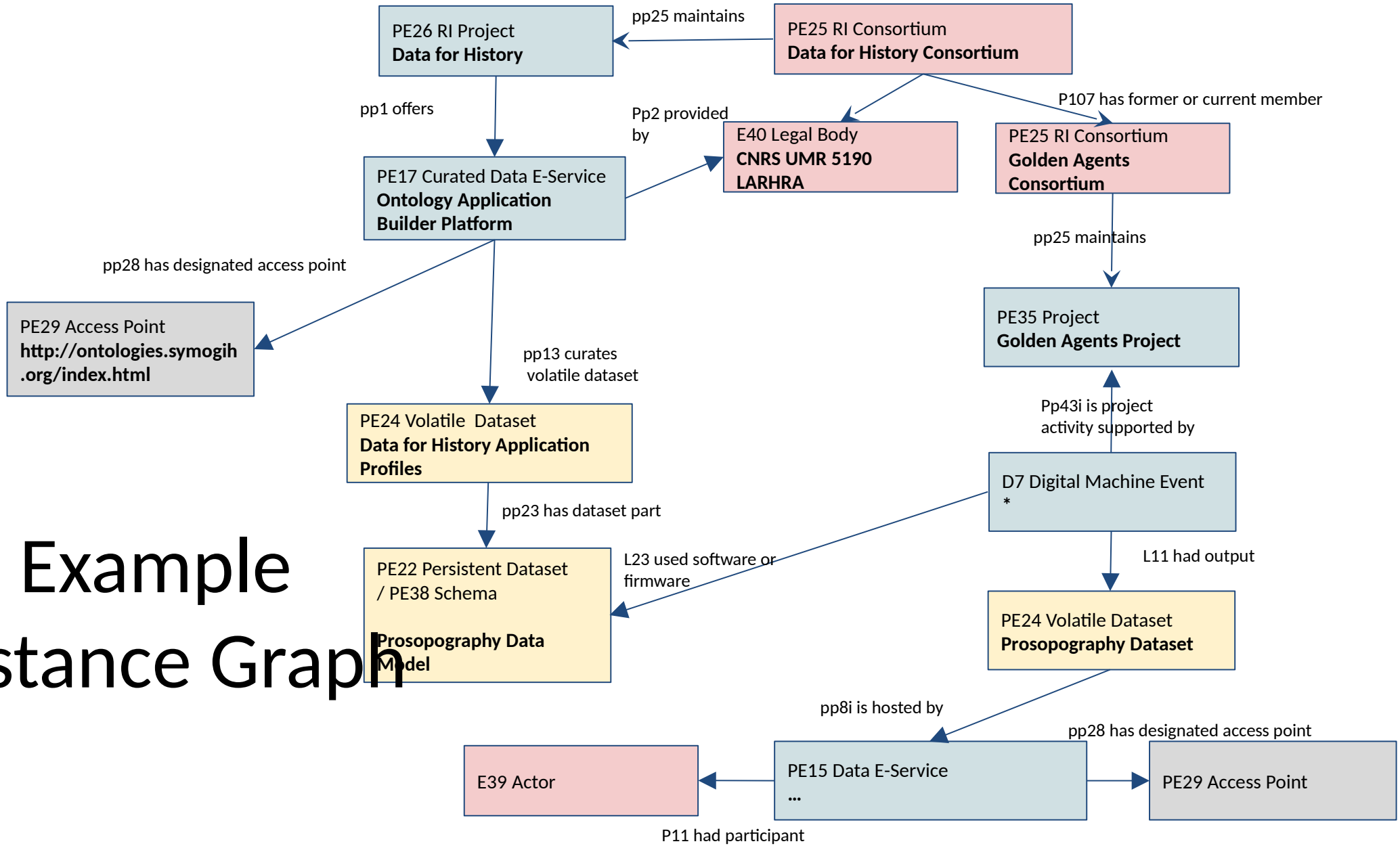
PARTHENOS

Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies

Semantic Infrastructure

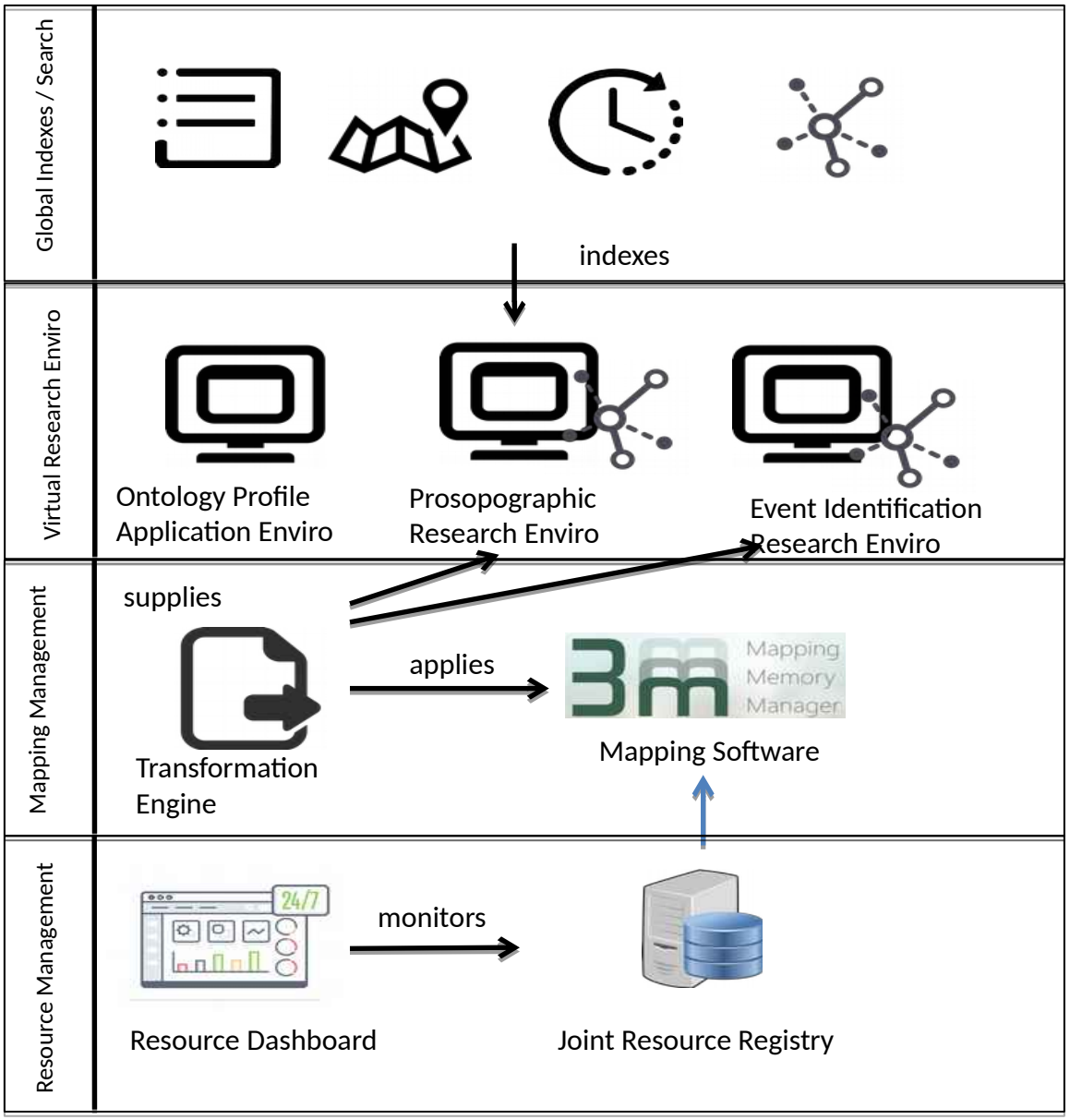


Example Instance Graph

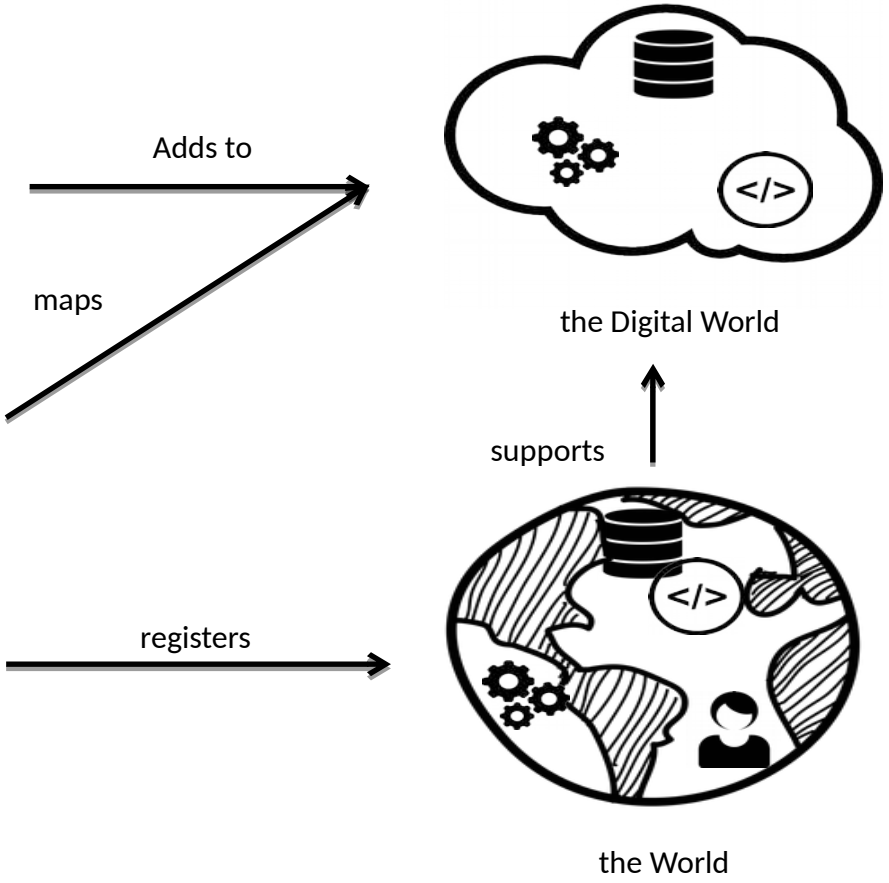


Top Level Picture of “Data for History” World

Entity	Instances
Projects	Data for History, APIS, Golden Agents, Henri Poincaré Papers, histHub, Histoire des paysages culturels industriels et maritimes, IMPRESSO, PANDORA Linked Open Data (LOD) Framework, symogih.org
Services	Data for History Platform (Forum, Mailing List, Website, Tools, ...)
Actors	Data for History Consortium, Agence bibliographique de l'Enseignement supérieur, Akademie der Wissenschaften zu Göttingen (Göttingen Centre for Digital Humanities), Archives de France, Centre François Viète d'épistémologie et d'histoire des sciences et des techniques, CNR (ISTI), CNRS, École normale supérieure de Lyon, ETH Zurich, Fondation des sources du droit de la Société suisse des juristes, FORTH (Center for Cultural Heritage - Institute of Computer Science), Hochschule für Technik, Wirtschaft und Kultur Leipzig (Agile Knowledge Engineering and Semantic Web), Huygens ING, INRAP, Laboratoire d'histoire des sciences et de philosophie – Archives Henri-Poincaré UMR 7117, Laboratoire de recherche historique Rhône-Alpes UMR 5190, Österreichische Akademie der Wissenschaften (Austrian Centre for Digital Humanities), Laboratoire TEMOS, Università degli studi di Firenze (Vast-Lab), Universität Graz (Zentrum für Informationsmodellierung), Université de Bretagne Occidentale, Université de Bretagne Sud, Université de Lorraine, Université du Luxembourg (CD2H Luxembourg), Université Jean-Moulin Lyon 3 (IFROSS), Université Lumière Lyon 2, Universiteit van Amsterdam
Datasets	Symogih.org (BHP)
Software	Ontology Application Profile Builder



Data for History
Semantic Infrastructure



The Scholarly Process

Researcher,
Intepreter

Refer
interpret
present

Search,
correlate,
integrate

discover
collect
aggregate
update

reexamine

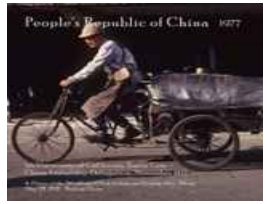
Publications
Stories
exhibitions

Layer of
"Latest
Knowledge"

"Evidence layer"
Things
Sources
Collections
Corpora

Curator,
Conservator,
Excavator

秦



Global Indexes / Search

indexes

Virtual Research Enviro

Ontology Profile Application Enviro

Prosopographic Research Enviro

Event Identification Research Enviro

Mapping Management

supplies

applies

Transformation Engine

Mapping Memory Manager

Mapping Software

Resource Management

monitors

Resource Dashboard

Joint Resource Registry

Data for History
Semantic Infrastructure

Adds to

maps

registers

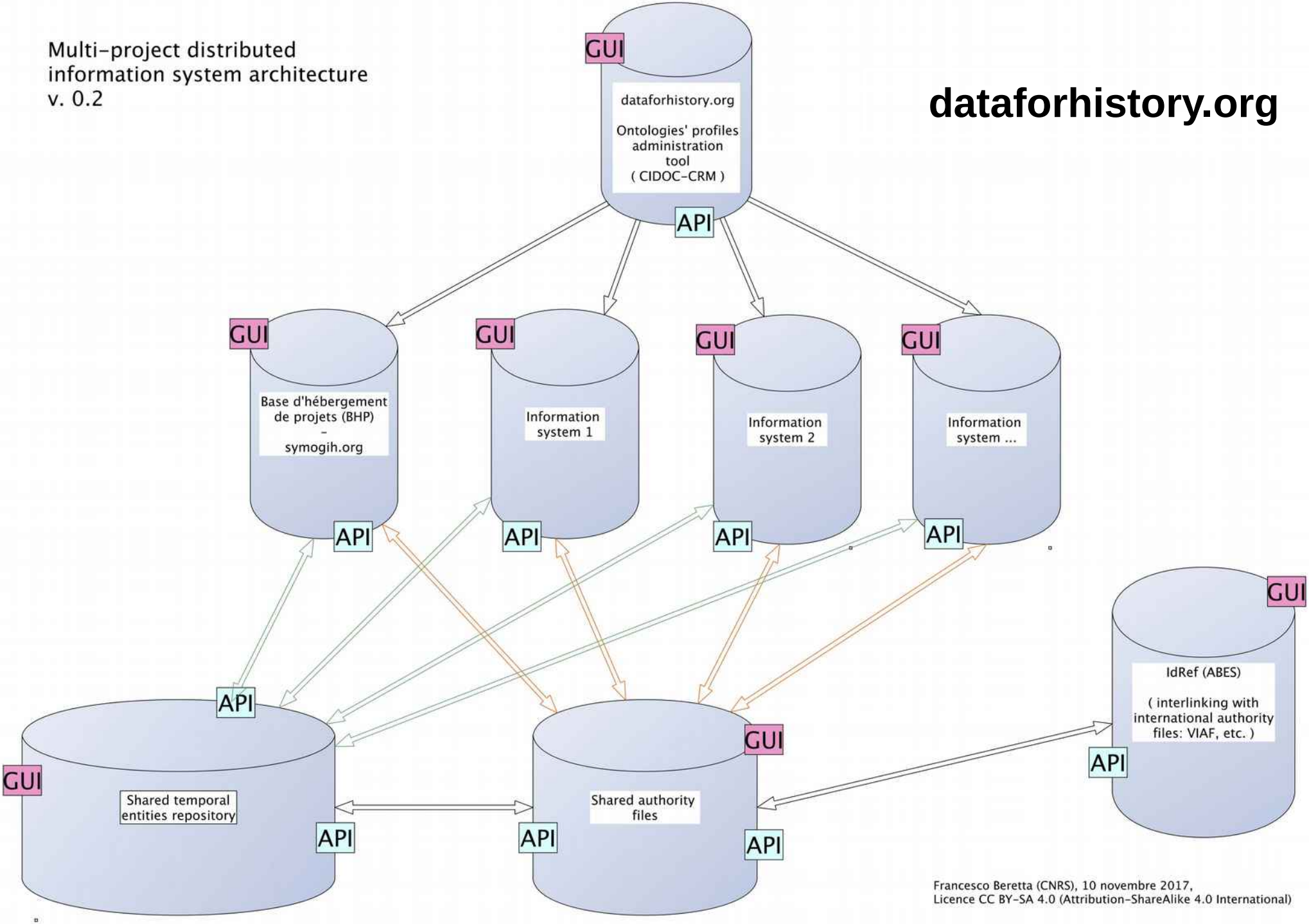
the Digital World

supports

the World

Multi-project distributed
information system architecture
v. 0.2

dataforhistory.org



Francesco Beretta (CNRS), 10 novembre 2017,
Licence CC BY-SA 4.0 (Attribution-ShareAlike 4.0 International)