

Graphentechnologien 2020

University of Vienna, 21 – 22 February 2020

Francesco Beretta

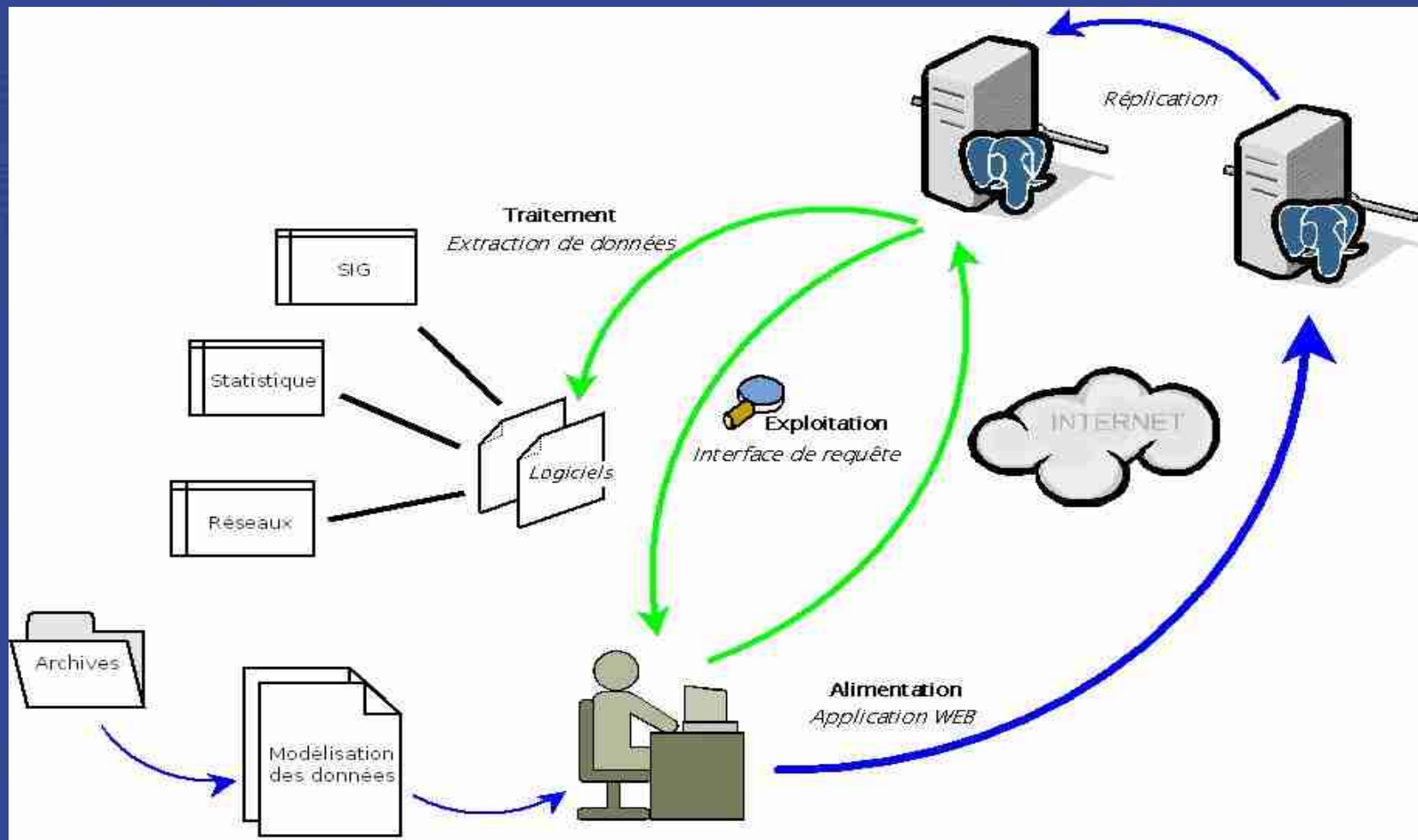
CNRS UMR5190 LARHRA – Université de Lyon

**Axe de recherche
en histoire numérique**

Modelling historical knowledge
in the context of
graph technologies and the semantic web

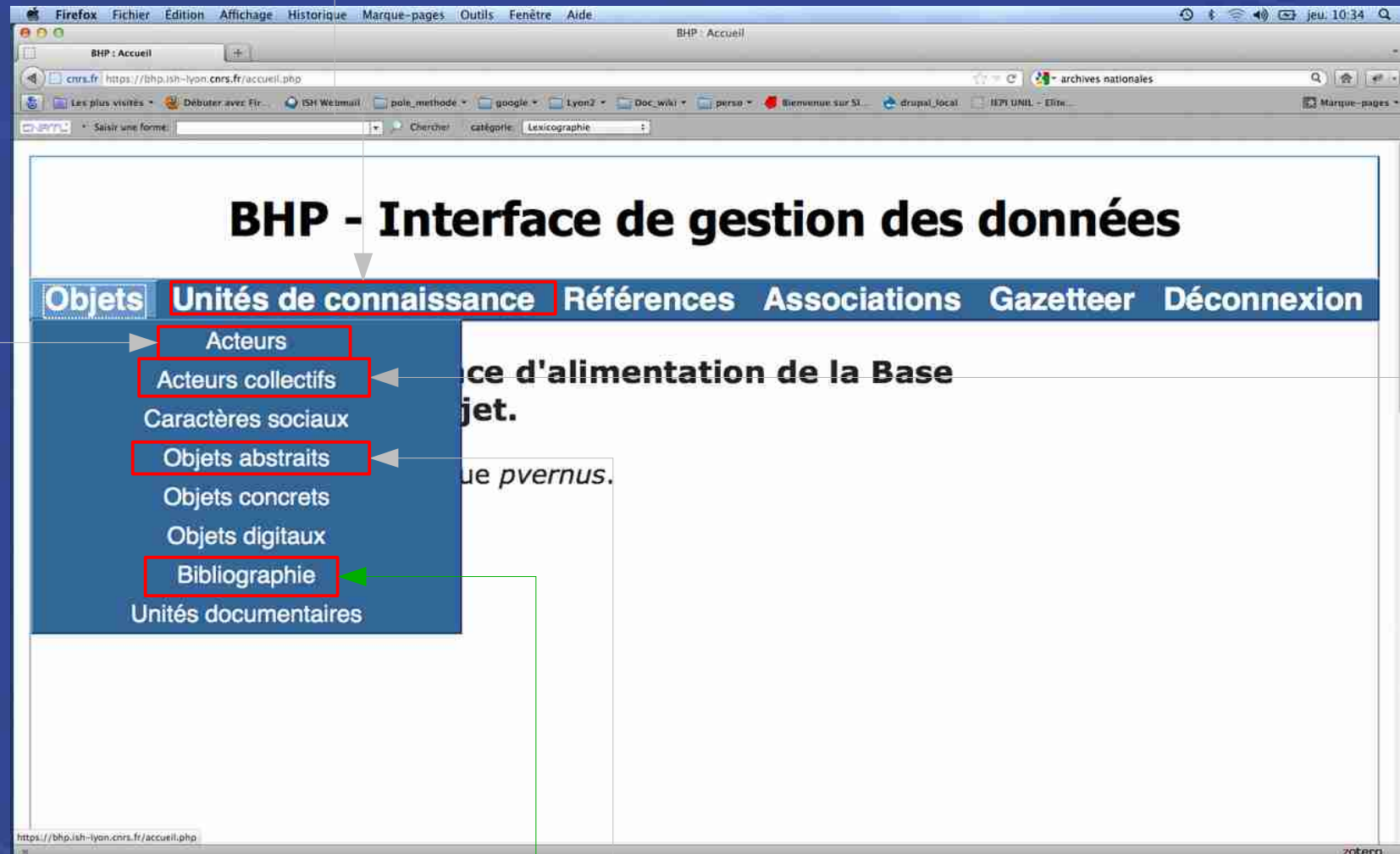
- * The *symogih.org* project's collaborative virtual research environment
- * Modelling historical knowledge in the context of the semantic web
- * Collaboratively modelling information : *dataforhistory.org*
- * Information extraction from sources : factoids or states of affairs ?

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 structured historical data

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



Actr : Galilei, Galileo

AbOb: Mathématiques

CoAc: Université de Padoue

Source : Dizionario biografico degli italiani, vol. 51

```

SELECT DISTINCT as1.cle_objet_associe 'idActeur', as1.libelle_calcule_objet_associe 'nom',
  as2.cle_objet_associe 'idLieu', as2.libelle_calcule_objet_associe as 'lieu', as2.longitude as 'longi',
as2.latitude as 'lat'
FROM web_talker_symogih.vue_association as1, web_talker_symogih.vue_association as2,
vues_bhp.acteurs_scholasticon acsc,
web_talker_symogih.vue_association as3, web_talker_symogih.vue_information inf
WHERE
/* jointure*/
as1.cle_objet_associe = acsc.CFAC AND as2.cle_etran_Info = as1.cle_etran_Info
AND as3.cle_objet_associe = as1.cle_objet_associe AND inf.cle_Info = as3.cle_etran_Info
/*selection*/
AND as1.cle_etran_TyRo = 'TyRo40'
AND as3.cle_etran_TyRo = 'TyRo12'

```

SQL queries to
extract data
corresponding to
the **research**
agenda of each
project participant

SQL Workbench/J - vues_fberetta - Default.wksp

User=epg@localhost, Database=vues_fberetta, URL=jdbc:mysql://localhost:8889/vues_fberetta

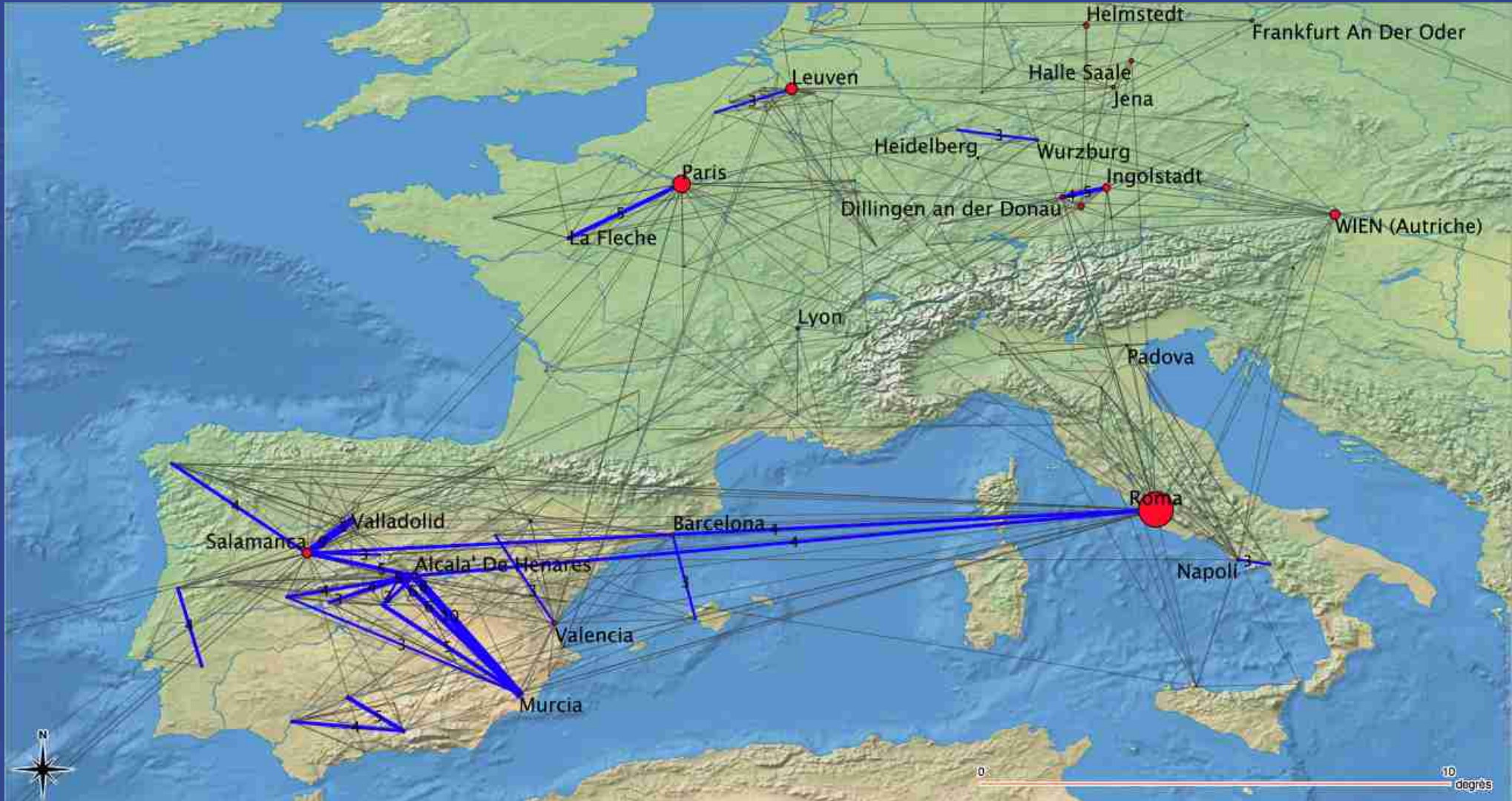
Statement 1 Database Explorer 2 Statement 3

11 /*selection*/
12 AND as1.cle_etran_TyRo = 'TyRo40' AND as2.cle_objet_associe LIKE 'NaPl%'
13 AND as3.cle_etran_TyRo = 'TyRo12' AND inf.cle_etran_TyIn = 'TyIn97'

Result Result Messages

cle_acteur	nom	cle_Info	date_debut	cle_lieu	lieu	long	lat
Actr38	Becanus, Martin	Info31818		NaPl2501	Köln	6,9347222222222222	50,942222222222222
Actr88	Caramuel Lobkowitz, Juan	Info31980		NaPl90258	Palazuelos	-2,6833333333333333	41,083333333333333
Actr88	Caramuel Lobkowitz, Juan	Info31981		NaPl10027	Salamanca	-5,6563888888888889	40,966666666666667
Actr164	Gassendi, Pierre	Info32430		NaPl71971	Aix	3,30596	50,4988
Actr187	Hodierna, Giovanni Battista	Info33117		NaPl3680	Ragusa	14,731944444444444	36,925833333333333
Actr272	Nifo, Agostino	Info34073	1514	NaPl3697	Roma	12,490277777777778	41,895277777777778
Actr314	Raynaud, Théophile	Info34673		NaPl1998	Lyon	4,8358333333333333	45,766666666666667
Actr349	Spina, Bartolomeo	Info35036		NaPl90265	Italie	12,60	42,50
Actr354	Suárez, Francisco	Info35079	1570	NaPl48	Espagne (Royaume d')	-4	40
Actr400	Fracastoro, Gerolamo	Info32381		NaPl90265	Italie	12,60	42,50
Actr405	Arriaga, Rodrigo de	Info31727		NaPl10124	Valladolid	-4,7213888888888889	41,650277777777778
Actr405	Arriaga, Rodrigo de	Info31728		NaPl1253	PRAHA (PRAGUE)	14,456388888888889	50,105833333333333
Actr457	Spinola, Stefano	Info35039		NaPl12229	Genova	8,9336111111111111	44,406111111111111
Actr537	Du Hamel, Jean Baptiste	Info32249		NaPl2085	Paris	2,3486111111111111	48,853333333333333
Actr562	Bagot, Jean	Info31761		NaPl90265	Italie	12,60	42,50
Actr624	Wittich, Christoph	Info35546		NaPl4251	Nijmegen	5,8580555555555556	51,8225
Actr809	Comenius, Johann Amos	Info32115	1628	NaPl90262	Lezno	18,433333333333333	54,35
Actr809	Comenius, Johann Amos	Info32116		NaPl4724	Elblag	19,405277777777778	54,158888888888889
Actr809	Comenius, Johann Amos	Info32117		NaPl3152	Sarospatak	21,571111111111111	48,321944444444444
Actr810	Major, John	Info33643	1518	NaPl11889	Glasgow	-4,2697222222222222	55,862777777777778
Actr810	Major, John	Info33644		NaPl90085	St. Andrews	-2,7988888888888889	56,338611111111111
Actr1505	Eschenbach, Johann Christian	Info32297		NaPl2497	Kiel	10,120555555555556	54,325277777777778

Ready, if you are L:13 C:68 17.1s Timeout: 0 Max. Rows: 0 1-22/1604



QGIS

symogih.org project website : documentation and data



The screenshot shows the SYMOGIH project website. The header is orange with the text 'SYMOGIH' and 'Références'. Below the header is a navigation bar with 'Accueil', 'Documentation', and 'Membres'. The main content area is white. On the left, there is a sidebar with three sections: 'Références' (listing 'Arborescence des classes de types d'unités de connaissances', 'Types d'informations', and 'Types de contenus'), 'Objets' (listing 'Acteurs', 'Acteurs collectifs', 'Objets abstraits', and 'Caractères sociaux'), and 'Sites propulsés par SyMoGIH' (listing 'GEO-LARHRA'). The main content area has a title 'Système Modulaire de Gestion de l'Information Historique (SyMoGIH)' and a section 'Le projet'. The 'Le projet' section contains two paragraphs and a list of features. The first paragraph describes the project's goal of developing a generic model for historical data storage and selective publication, leading to a collaborative platform for historical research. The second paragraph describes the platform's capabilities for storing primary data on human activity (social, economic, intellectual, etc.), XML-coded texts (following the 'Texte Encoding Initiative' standard), images, and their metadata, allowing for the association of spatial 'footprints' to objects. The list of features includes 'la modélisation progressive et évolutive de l'information historique grâce à un dictionnaire de types d'unités de connaissance ;' and a partially visible 'la plateforme permet de gérer les données historiques de manière structurée et normalisée ;'. A large blue box at the bottom right contains the URL 'http://symogih.org'.

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

Sites propulsés par SyMoGIH

- GEO-LARHRA

Système Modulaire de Gestion de l'Information Historique (SyMoGIH)

Le projet

Le projet SyMoGIH a développé un modèle générique de stockage des données historiques permettant leur interopérabilité et leur publication sélective. A partir de ce modèle, une **plateforme collaborative** pour la recherche en histoire a été mise en place, utilisée par plusieurs chercheurs et projets.

Cette plateforme permet le stockage de données primaires concernant toute activité humaine (sociale, économique, intellectuelle, ...), de textes codés en XML (traités selon le standard proposé par la **Texte Encoding Initiative**), d'images et de leur métadonnées, tout en permettant d'associer à ces différents objets leur 'empreinte spatiale'. La réalisation d'un **système d'information géographique** (SIG) joue un rôle essentiel dans le projet.

La plateforme permet :

- la modélisation progressive et évolutive de l'information historique grâce à un dictionnaire de **types d'unités de connaissance** ;
- la plateforme permet de gérer les données historiques de manière structurée et normalisée ;

<http://symogih.org>

Project specific websites : Patrons de France

SYMOGIIH

Références

PATRONS
DE
FRANCE



Système d'Information
Patrons et Patronat
Français

XIXe-XXe siècles

Accueil Le corpus Aide à la consultation Consultation Sources dépouillées Contributeurs

Accès rapide à la base

- Patrons
- Institutions
- Caractères sociaux
- Lieux

Contributions et contact

Les détenteurs d'exemplaires de bulletins ou d'annuaires indiqués comme manquants dans les collections dépouillées jusqu'à maintenant sont invités à nous contacter en envoyant un message à patronsdefrance@ish-lyon.cnrs.fr.

Consultation : liste des patrons

Affichage de 1 à 10 sur 3042

Nom, prénom

Contient

Année de naissance

Est égal à

Lieu de naissance

Contient

Appliquer

Réinitialiser

Nom, prénom	Genre	Année de naissance	Lieu de naissance	Année de décès
Acher, Maximilien	Homme	1862	Le Havre	1929
Acolas, Prosper	Homme	1838	Saint-Bonnet-Tronçais	1928
Adam, Alcide	Homme	1864	Ferrières	1931
Adam, Désiré	Homme	1859	Saint-Paul-du-Vernay	1929
Adenot, Henri	Homme	1904		1947
Ader, Clément	Homme	1841	Muret	1925
Adher, Pierre	Homme	1884		1955
Adnet, Aloïse	Homme	1848	Châlons-en-Champagne	1927
Agache, Donat	Homme	1882	Lille	1929
Agache, Edouard Donat Louis Joseph	Homme	1841	Lille	1923

<http://patronsdefrance.fr/>

Project specific websites : Professeurs de droit

SYMOGIH

Références

SIPROJURIS

Système d'information des professeurs de droit (1804-1950)

[Le corpus »](#) [Sources dépouillées](#) [Contributeurs](#) [Statuts](#)



LES ENSEIGNANTS

[VOIR LA LISTE](#)

<http://siprojuris.symogih.org>

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)

SIPROJURIS

Bienvenue sur le site du projet SIPROJURIS.



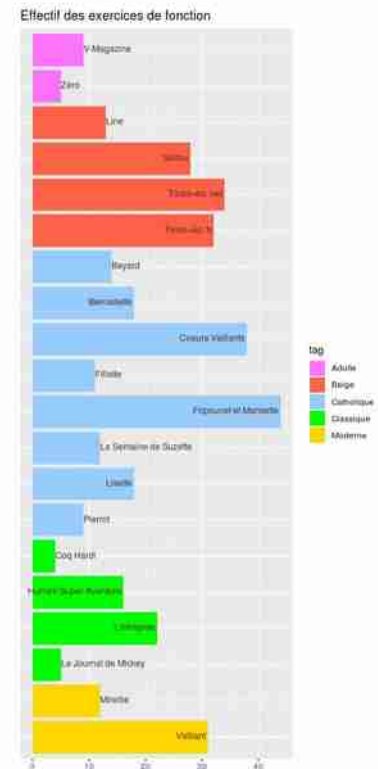
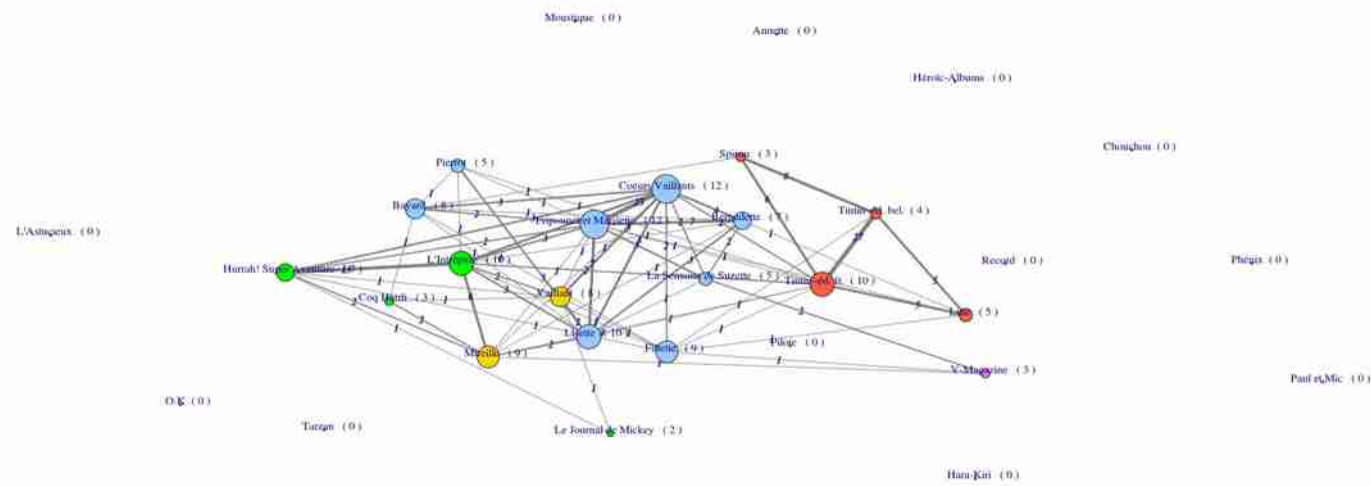
CODE
UNIVERSITAIRE
OU
LOIS ET STATUTS
DE L'UNIVERSITÉ ROYALE DE FRANCE



<http://siprojuris.symogih.org>

Huma-Num RStudio-Shiny
 tive and collaborative data an
<https://tinyurl.com/phn-shiny>

Relations entre périodiques (collaborations entre dessinateurs).



Année

Pour définir une période,
utiliser la durée des relations

Année



Intensité du lien

Durée de la collaboration



Effectif des collaborations



Centralité

Centralité de :

Degré

Degré des sommets



Choix des personnes

☐ **Active**

NULL

Choix d'un ou plusieurs dessinateurs

SYMOGIH

Références

GEO-LARHRA

Partage de ressources géo-historiques

Se connecter

Accueil

Présentation

Gazetteer

Géocatalogue

Atlas historique

Consulter l'Atlas Historique

- Territoires historiques de l'Europe
- Présentation de l'atlas



Cette œuvre est mise à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions 4.0 International.



<http://geo-larhra.ish-lyon.cnrs.fr/>

Use the semantic web for historical research

Evolution des territoires en Italie

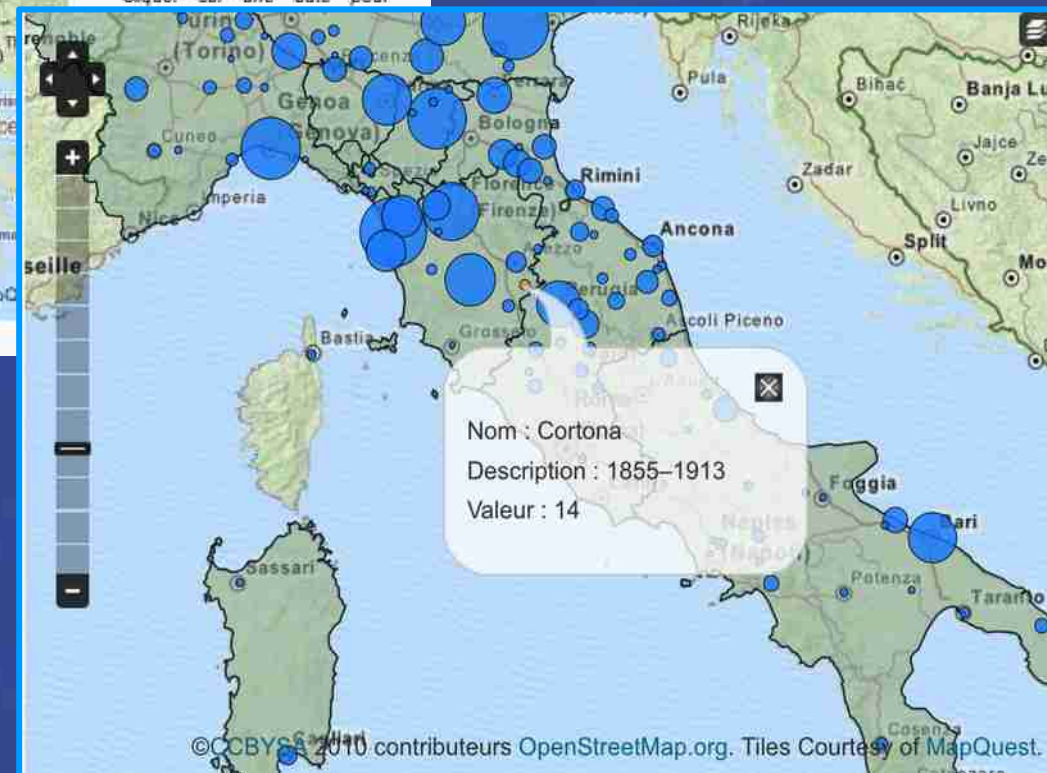
Site expérimental. Données non exhaustives, en cours de production.



Dates significatives

1815-06-09
1829-12
1847-12
1859-11-10
1860-03-24
1860-11-05
1861-03-17
1866-10-03
1870-10-02
1920-11-12
1929-02-11
1947-02-10

Cliquer sur une date pour

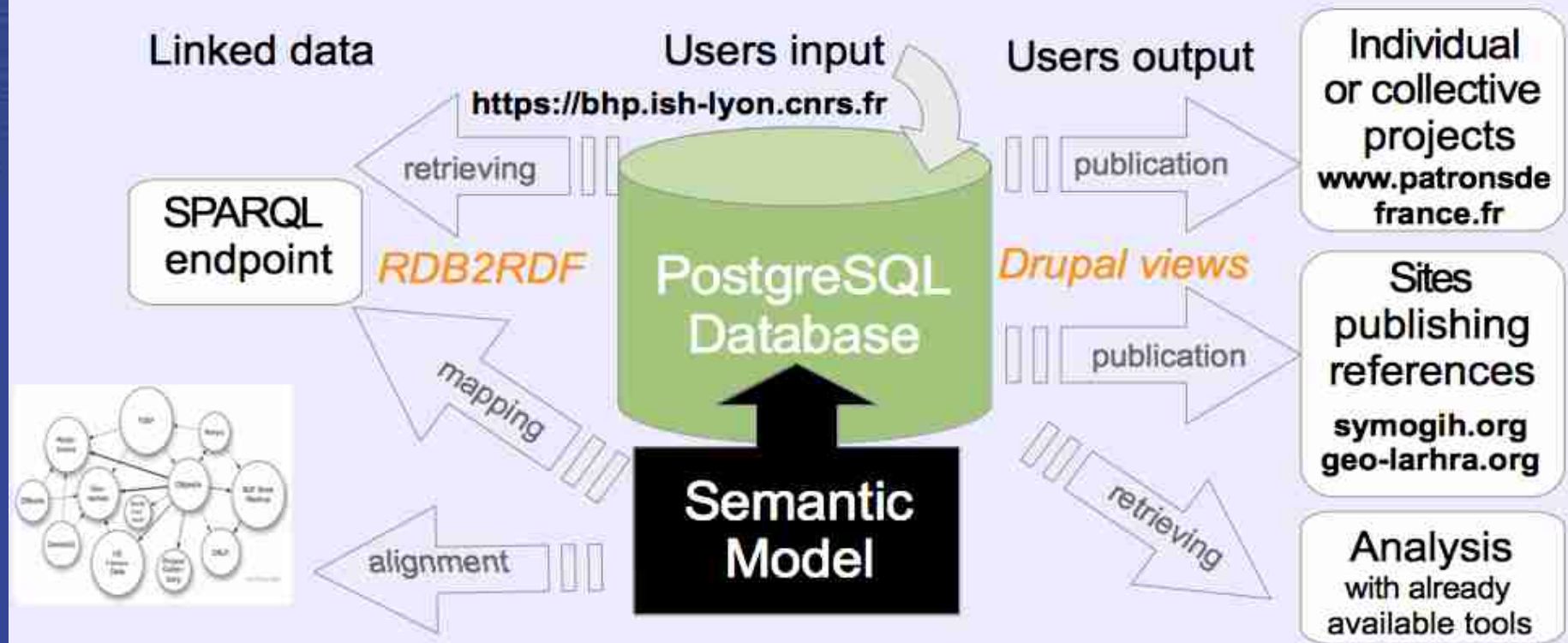


SPARQL – endpoint B3Kat

Bayerische Staatsbibliothek, Bibliotheksverbund Bayern, Kooperative Bibliotheksverbund Berlin-Brandenburg

<http://lod.b3kat.de/sparql>

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



The *symogih.org* project was started in 2008.

About 50 scholars and students, and 15 research project, are currently using the collaborative database to store and share historical information



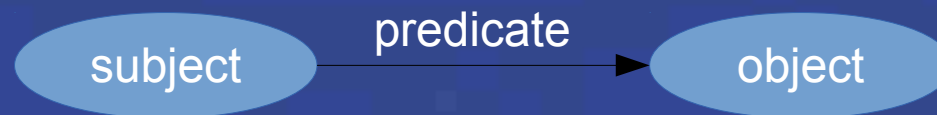
Geovistory : a VRE for historical research

- * The *symogih.org* project's collaborative virtual research environment
- * Modelling historical knowledge in the context of the semantic web
- * Collaboratively modelling information : *dataforhistory.org*
- * Information extraction from sources : factoids or states of affairs ?

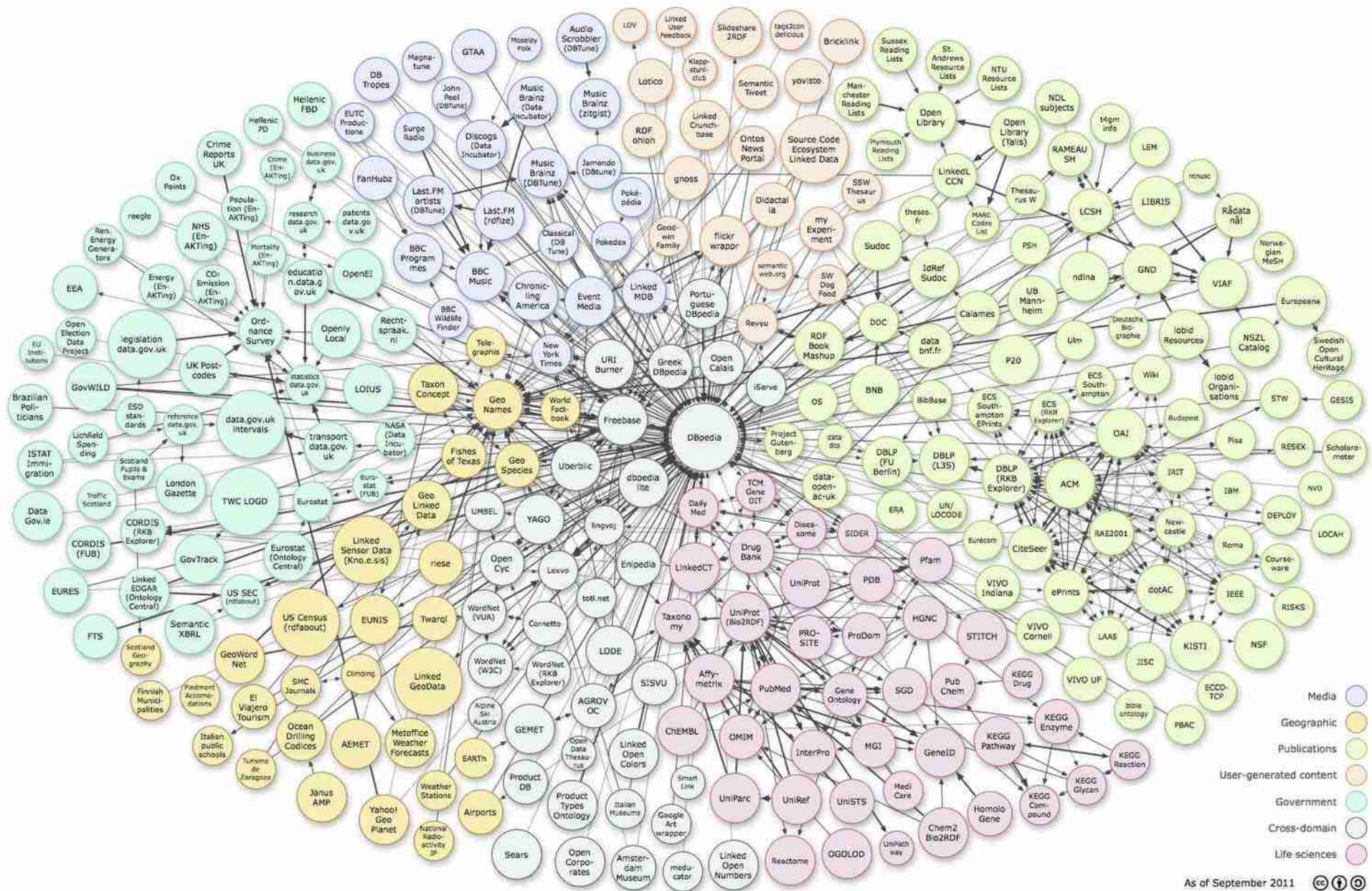
The semantic web

(<https://www.w3.org/TR/rdf11-concepts/>)

- « The Resource Description Framework (RDF) is a framework for **representing information in the Web.** »
- « A graph-based data model »



RDF triple



LOD diagram by Anja Jentzsch - Own work, CC BY-SA 3.0

<https://commons.wikimedia.org/w/index.php?curid=16574061>

Knowledge Graphs for Natural Language Processing

Leibniz wrote to Caroline of Ansbach that Newton's physics was detrimental to natural theology. However, eager to defend the Newtonian view, it was Clarke who responded and the correspondence between both continued until the death of Leibniz.

text

knowledge
base



dbr:Gottfried_Willhelm_Leibniz

rdf:type

dbo:Philosopher

rdfs:subClassOf

dbo:Person

dbo:doctoralAdvisor

foaf:name

"Gottfried Wilhelm Leibniz"@de

dbo:birthDate

"1646-07-01"^^xsd:date

dbo:deathDate

"1716-11-14"^^xsd:date

dbr:Ontology

dbo:notableIdea

dbr:Christian_Wolff

rdf:type

dbo:Philosopher

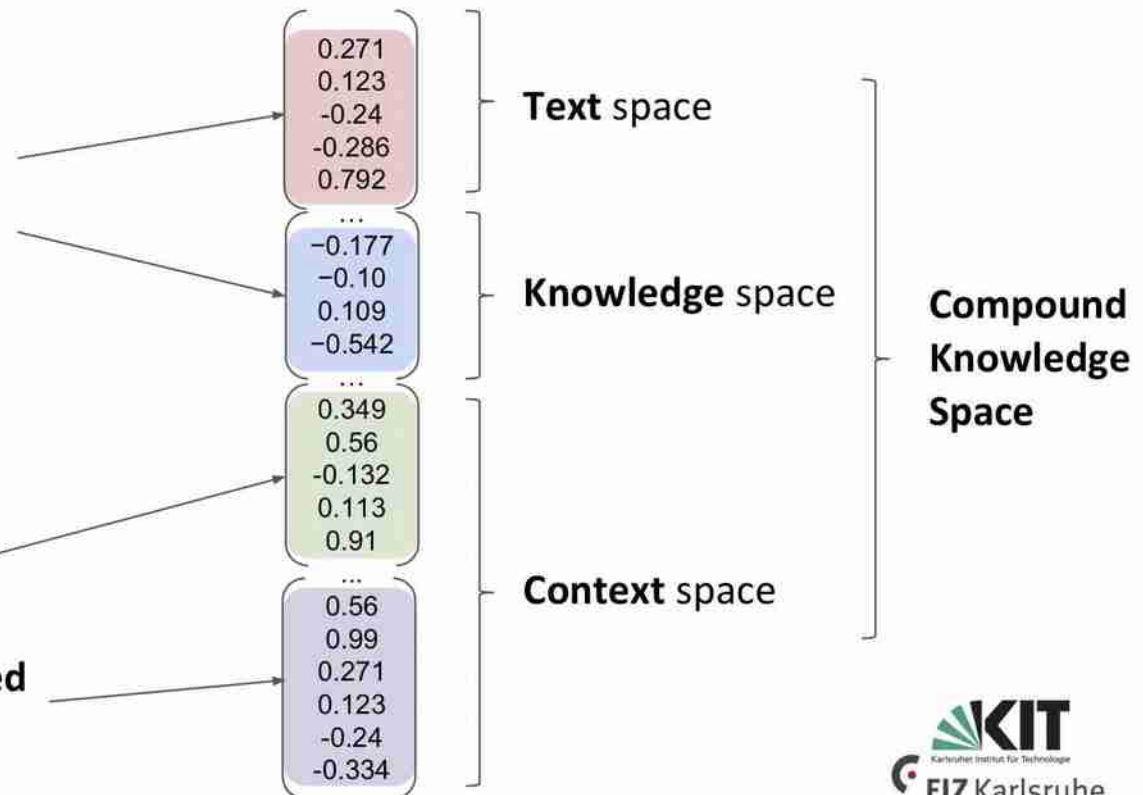
Karlsruher Institut für Technologie (29. Novembre 2017) – Antrittsvorlesung von

Prof. Dr. Harald Sack

Combining Semantics and Deep Learning for Intelligent Information Services

Combined Feature Embeddings for a Compound Knowledge Space

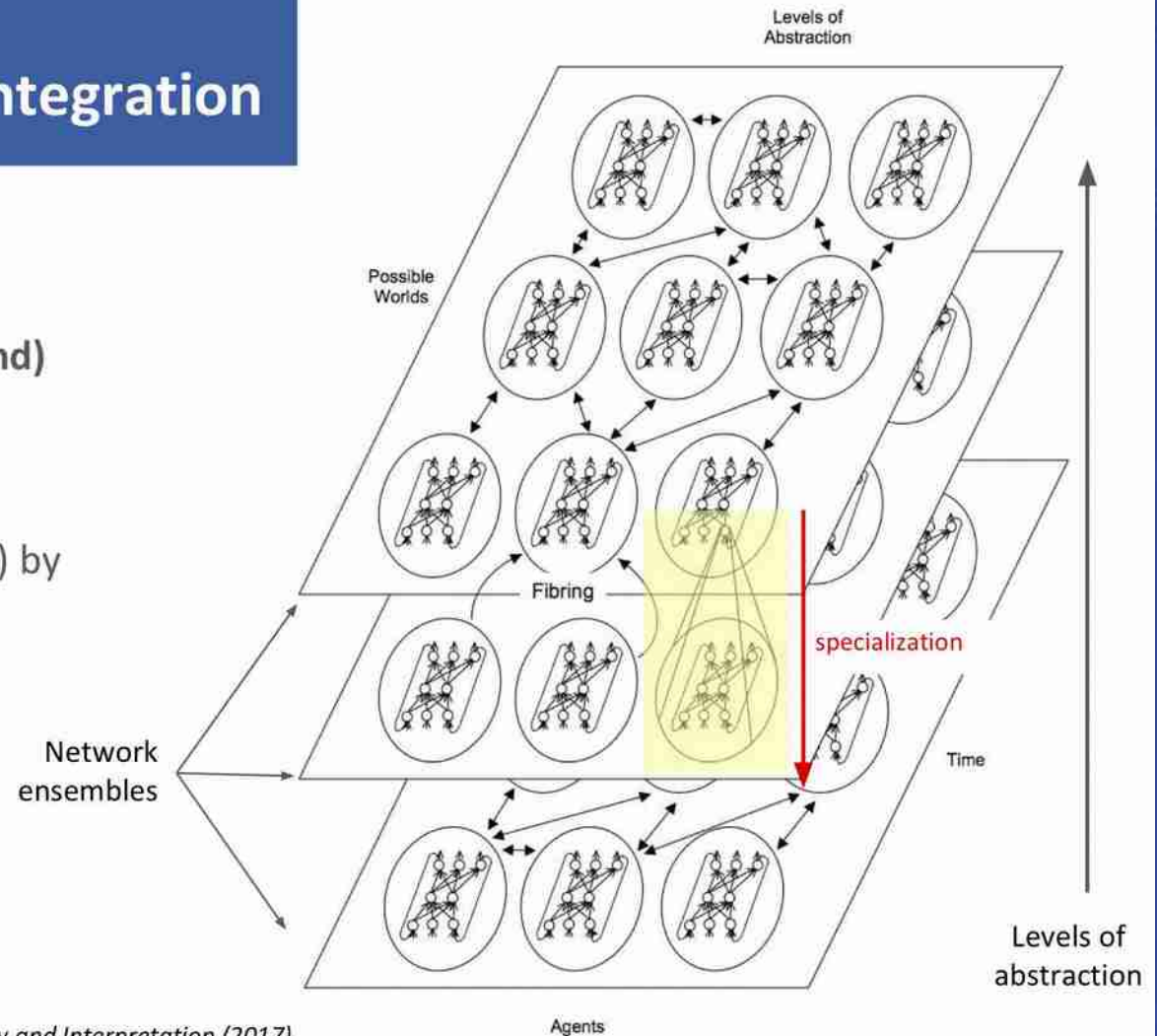
- Various feature vectors
 - **Word** embeddings
 - **Knowledge Graph** embeddings
 - Instances
 - Ontologies
 - Embeddings for **semantically enriched texts**
 - **Metadata** and **aggregated features**



Towards Neuro-Symbolic Integration

Neuro-Symbolic Systems

1. Translation of **symbolic (background) knowledge** into the network
2. **Learning of additional knowledge** from examples (and generalisation) by the network
3. **Executing the network** (i.e. reasoning), and
4. **Symbolic knowledge extraction** from the network.



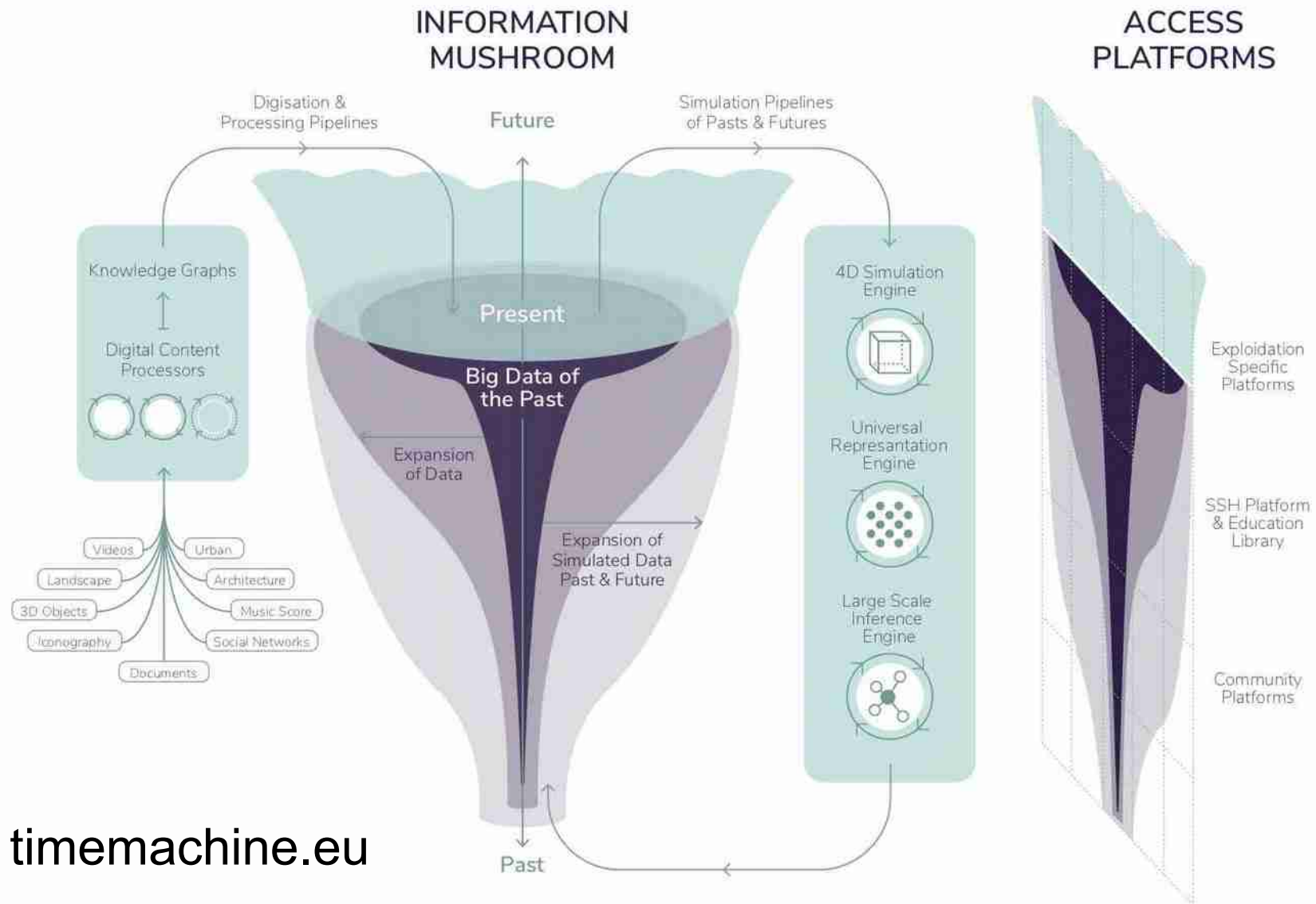
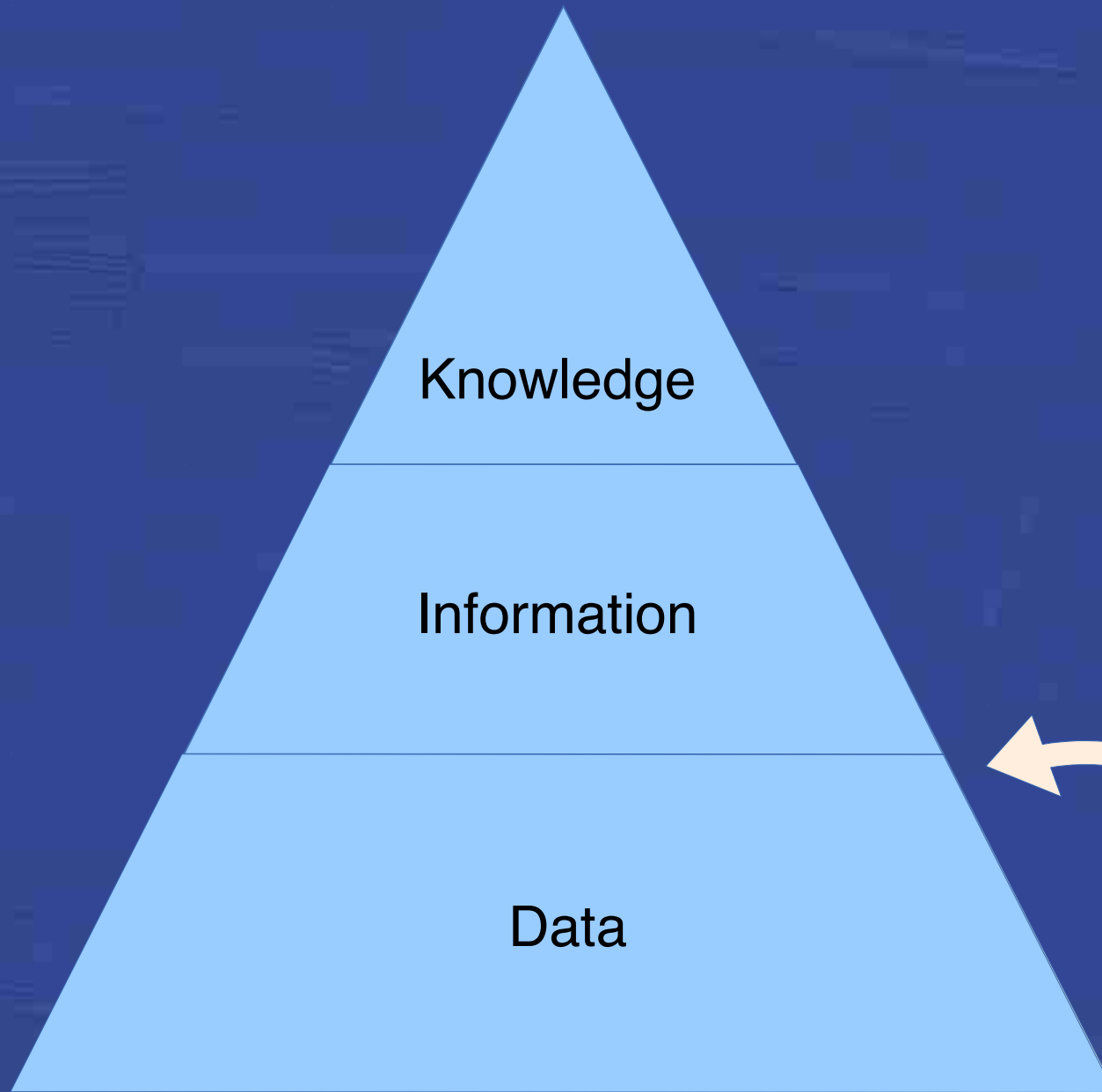
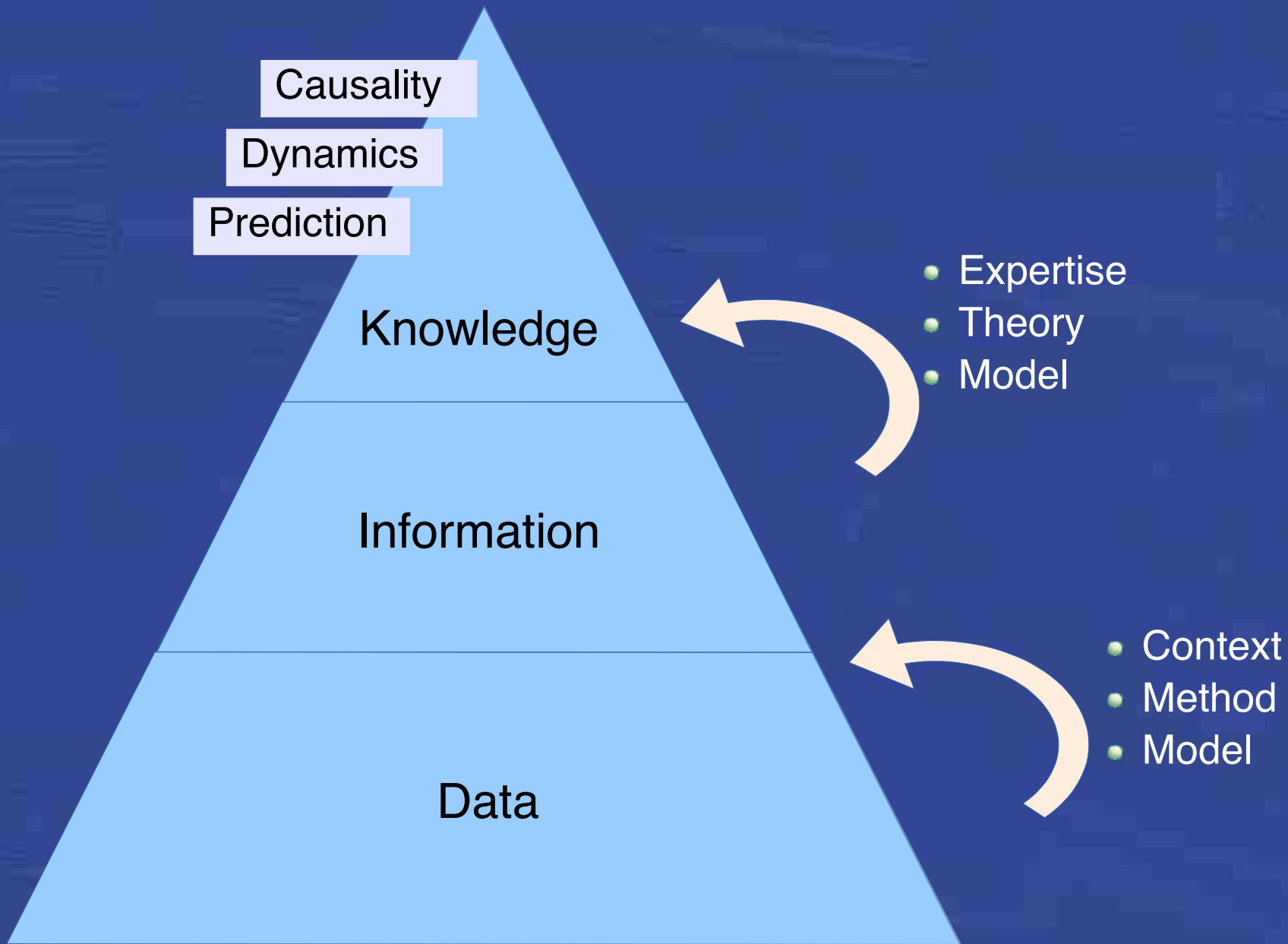
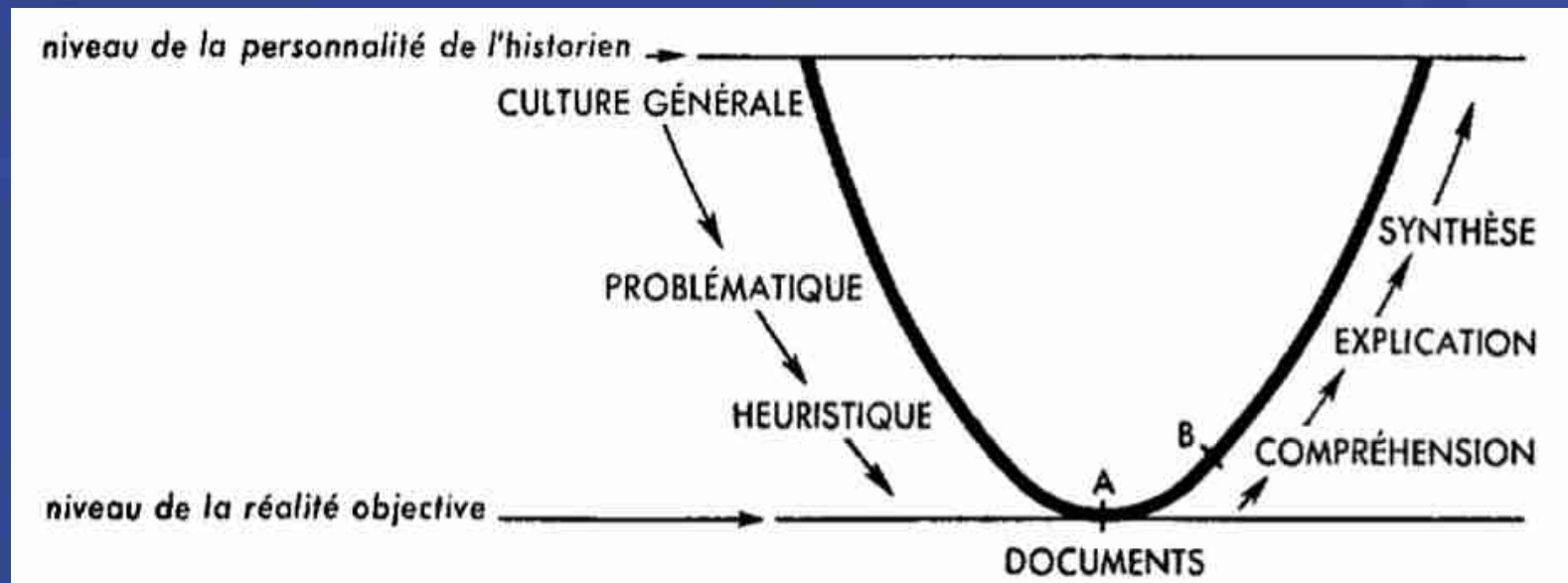


FIGURE 3: Schematic representation of the TM Digital Content Processor and the three simulation engines







Henri-Irénée Marrou, « Comment comprendre le métier de l'historien », *L'histoire et ses méthodes*, C. Samaran (éd.), Paris, Editions Gallimard, 1961, 1502.

General research
agenda



Research
questions



Collecting sources

General research
agenda



Research
questions



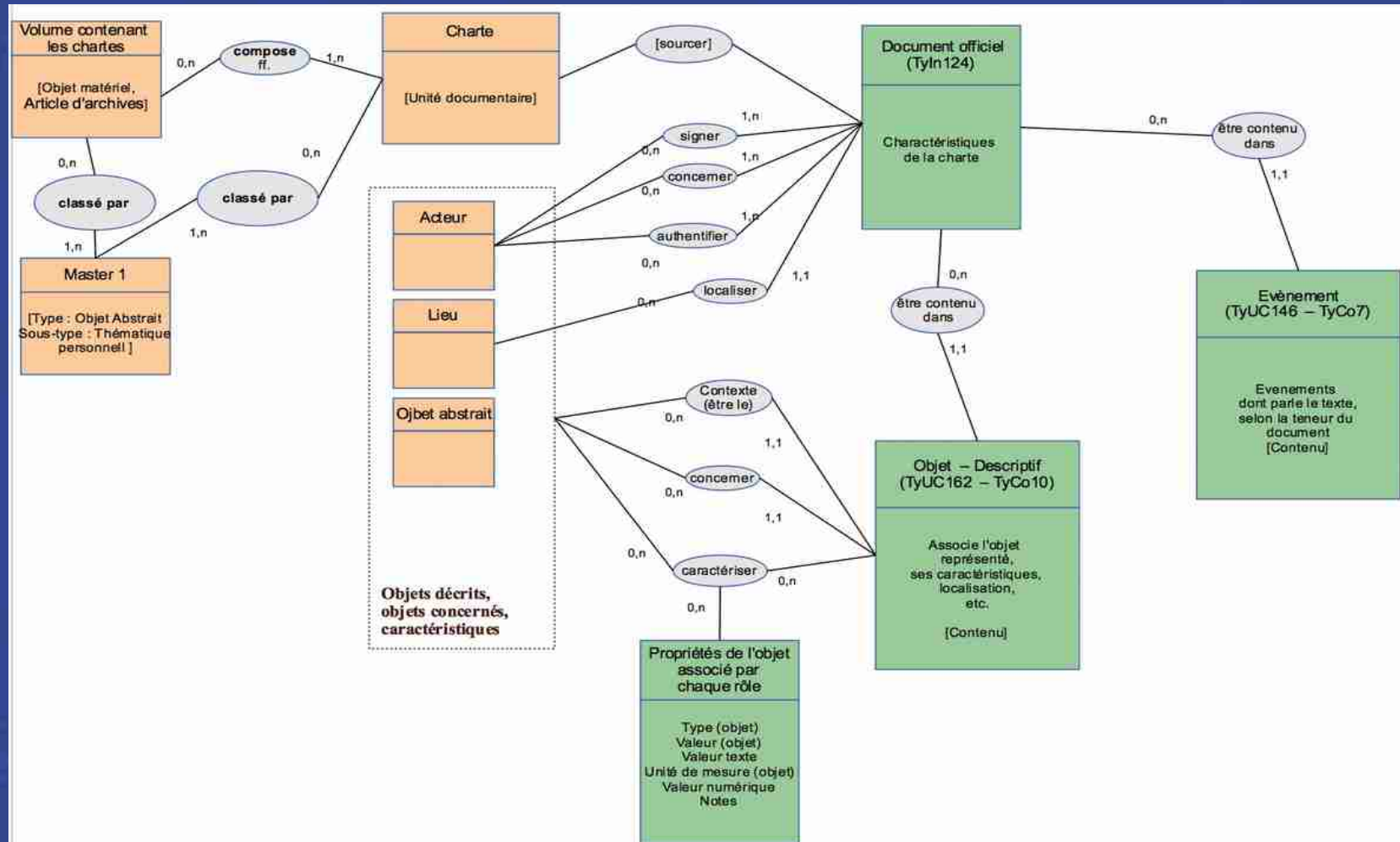
Collecting sources



Model

Data production

(Mémoire de master, Université de Lyon 3, 2015)



General research
agenda



Research
questions



Collecting sources



Model

Data production



Data cleansing
Coding of
variables



Model

Importing
existing data

General research
agenda



Research
questions



Collecting sources



Model

Data production



Model

Data cleansing
Coding of
variables



Model

Importing
existing data

General research
agenda



Research
questions



Collecting sources



Model

Data production



Importing
existing data

Model

Model
Data cleansing
Coding of
variables



Data analysis,
visualization,
mining

Research questions

R – RStudio / QGIS

RStudio File Edit Code View Plots Session Build Debug Tools Window Help

/Volumes/Fortitudo/Documents/SyMoGIH/cours_lyon_3/cours_2014/etudiants/taillandier_clement_de/taillandier_refuge - RStudio

taillandier_refuge

Environment History

Global Environment

- lpa 2238 obs. of 9 variables
- lpmp 2411 obs. of 8 variables
- lpmp_1685_1695 2273 obs. of 8 variables
- tst 180 obs. of 1 variable
- tstf 180 obs. of 1 variable

Values

- etiq chr [1:18] "NA" "Ba1" "Br1" "Brn" "Bra" "Fra" "Gen" ...
- p List of 9
- palette_man chr [1:18] "#C0C0C0" "#87CEFA" "#FF4500" "#00CED1" ...
- s int [1:9] 1 2 3 4 5 6 7 8 9
- v List of 7

Files Plots Packages Help Viewer

Zoom Export Publish

```
77
78
79
80
81
82
83
84 # première assistance
85 lpa <- read.csv("lieux_premiere_assistance_liste_20150622.csv", stringsAsFactors=FALSE)
86 s <- seq(as.Date("1678-01-01"), as.Date("1733-12-01"), by = "month")
87 x <- hist(as.Date(lpa$date_premiere_assistance), s, freq = TRUE, main='Présences dans le Refu
88         xlab = "Mois - Une mention par individu et par mois est comptée. Effectif total (blar
89         ylab = 'Effectif', plot = TRUE, right = FALSE)
90
91 plot(as.Date(lpa$date_premiere_assistance), as.factor(lpa$premier_lieu_assistance))
92
93
94
95
96
97
98
99
100
```

85:1 (Top Level) R Script

Console

```
/Volumes/Fortitudo/Documents/SyMoGIH/cours_lyon_3/cours_2014/etudiants/taillandier_clement_de/taillandier_refuge
> lpmp <- read.csv("lieux_premiere_assistance_liste_20150622.csv", stringsAsFactors=TRUE)
>
> lpmp$type_lieu <- as.factor(as.character(lpmp$ori_nature))
> lpmp$type_met <- as.factor(as.character(lpmp$met_nettoyee))
> lpmp$semestre <- as.Date(lpmp$trimestre)
> s <- seq.int(1,9)
> lpmp_1685_1695 <- subset(lpmp, lpmp$semestre > as.Date("1684-12-01")
+ & lpmp$semestre < as.Date("1696-01-01")
+ & as.integer(lpmp$met_nettoyee) %in% s)
> class(s)
[1] "integer"
> boxplot(semestre-met_nettoyee, data=lpmp_1685_1695, varwidth=TRUE, horizontal=T)
> ### premières explorations - éliminer ?
> plot(presences_annuelles_refuge$debut_annee, presences_annuelles_refuge$nombre, type = "p", col = "re
d",
+ main = "Effectifs par lieu et année",
+ xlab = "Année", ylab = "Effectif par lieu et année"
+ # , log = "y" # logarithme sur la valeur y
```


General research
agenda



Research
questions



Collecting sources



Model

Data production



Importing
existing data

Model

Model
Data cleansing
Coding of
variables

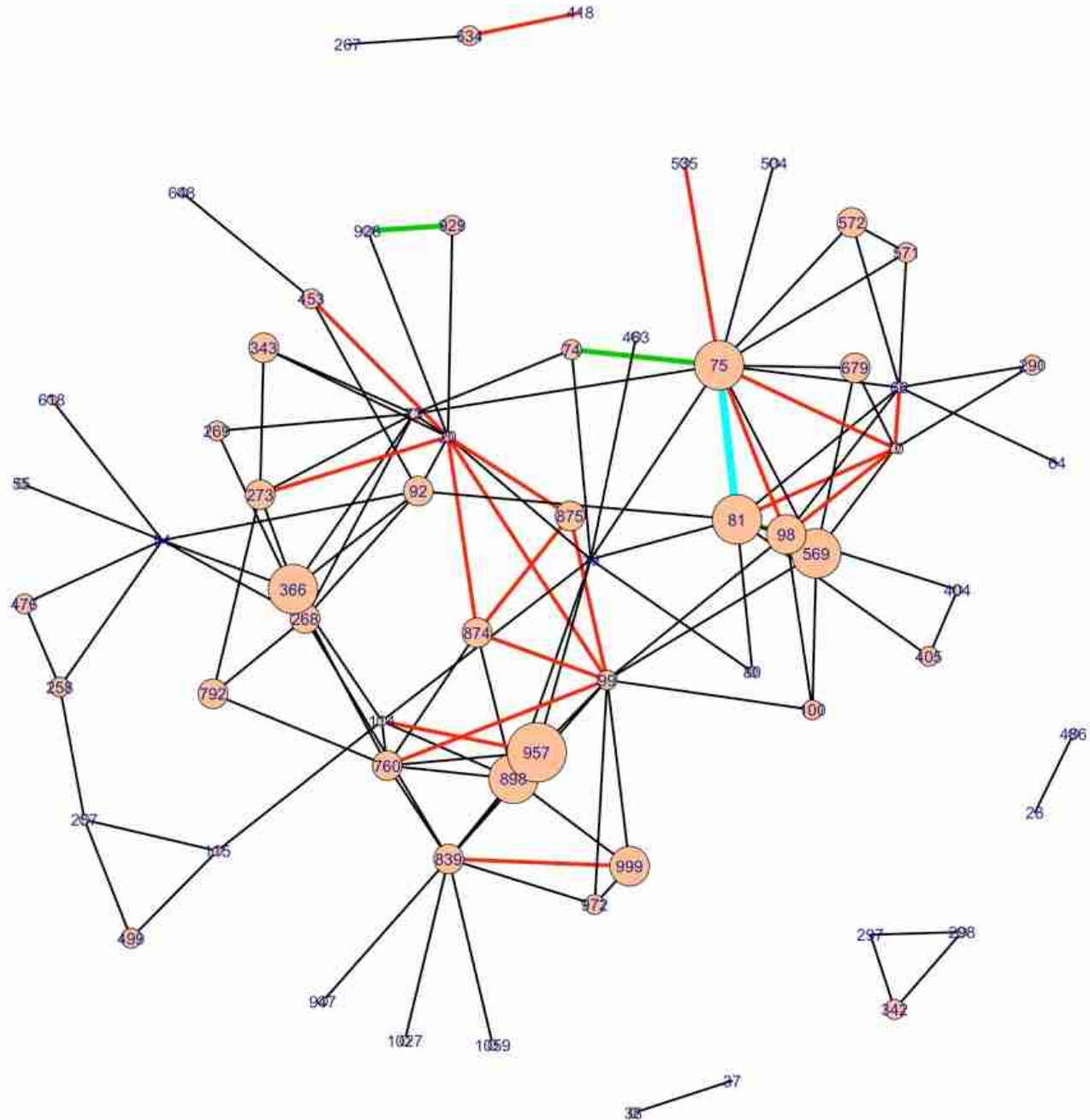


Data analysis,
visualization,
mining

Research questions

Model

Centralité de degré et nombre de liens entre signataires



General research
agenda

Research
questions

Collecting sources

Data production

Importing
existing data

Model
Data cleansing
Coding of
variables

Model
Data analysis,
visualization,
mining
Research questions

Interpretation
Verification

General research
agenda

Research
questions

Collecting sources

Data production

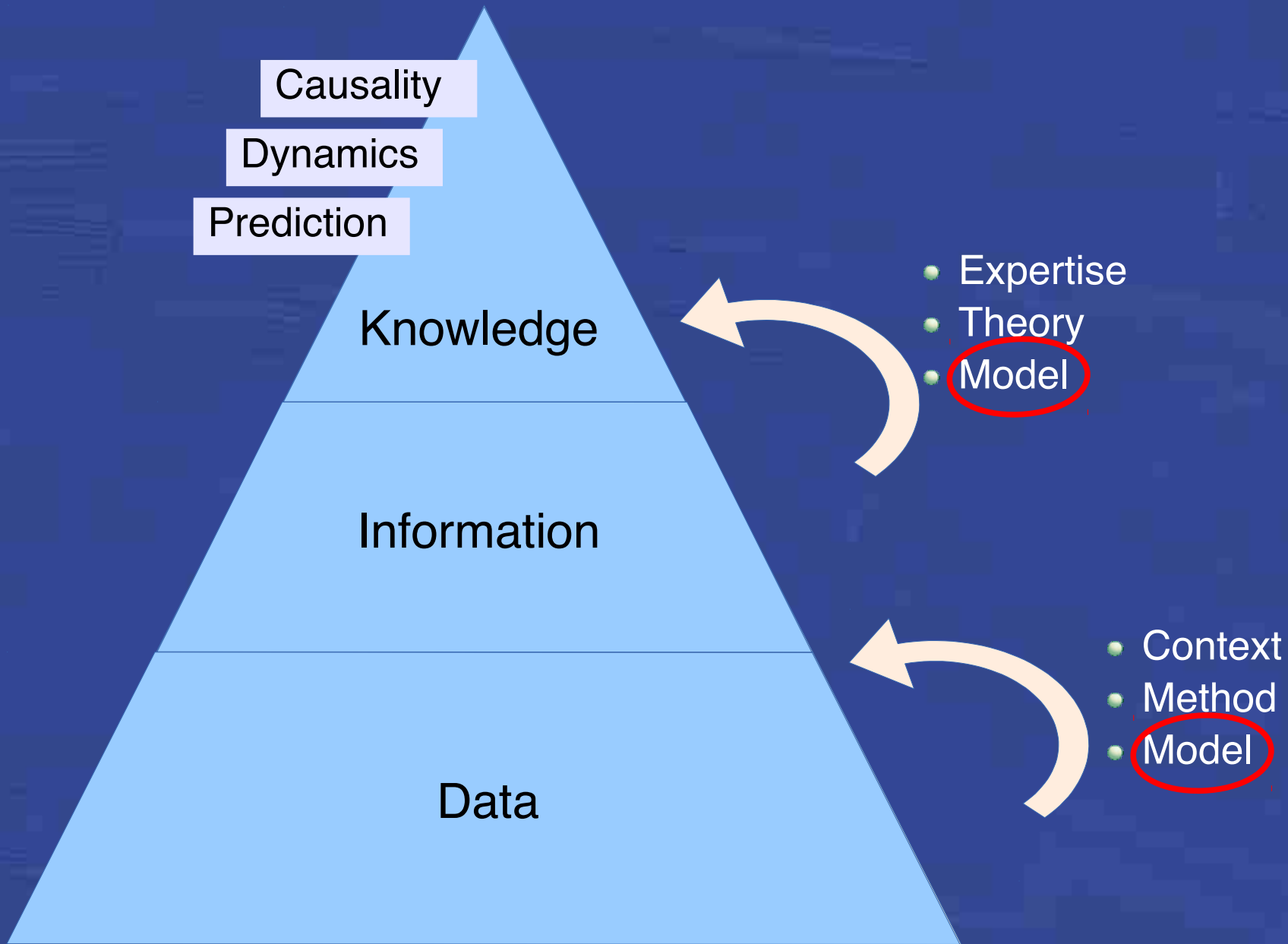
Importing
existing data

Model
Data cleansing
Coding of
variables

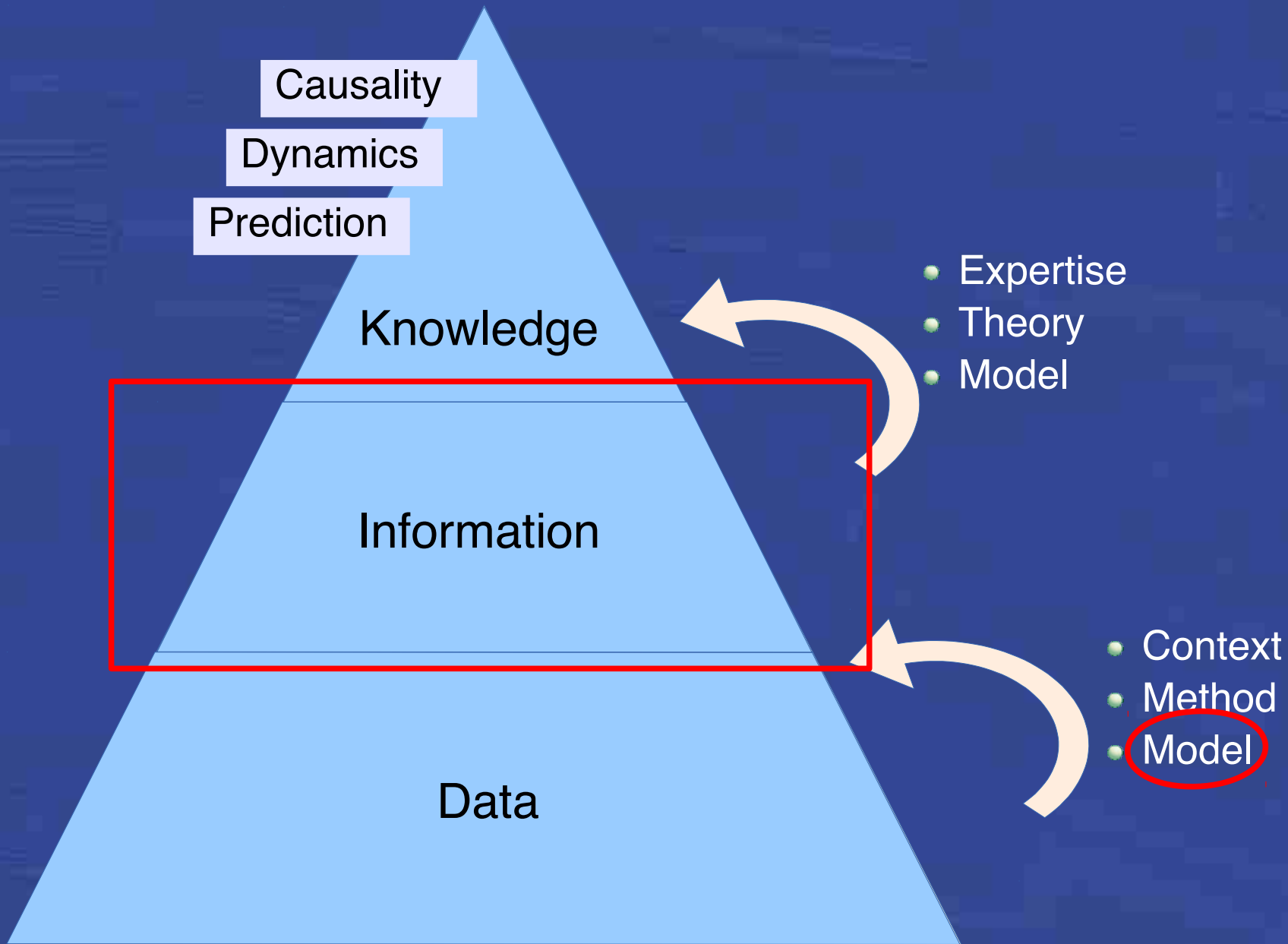
Model
Data analysis,
visualization,
mining
Research questions

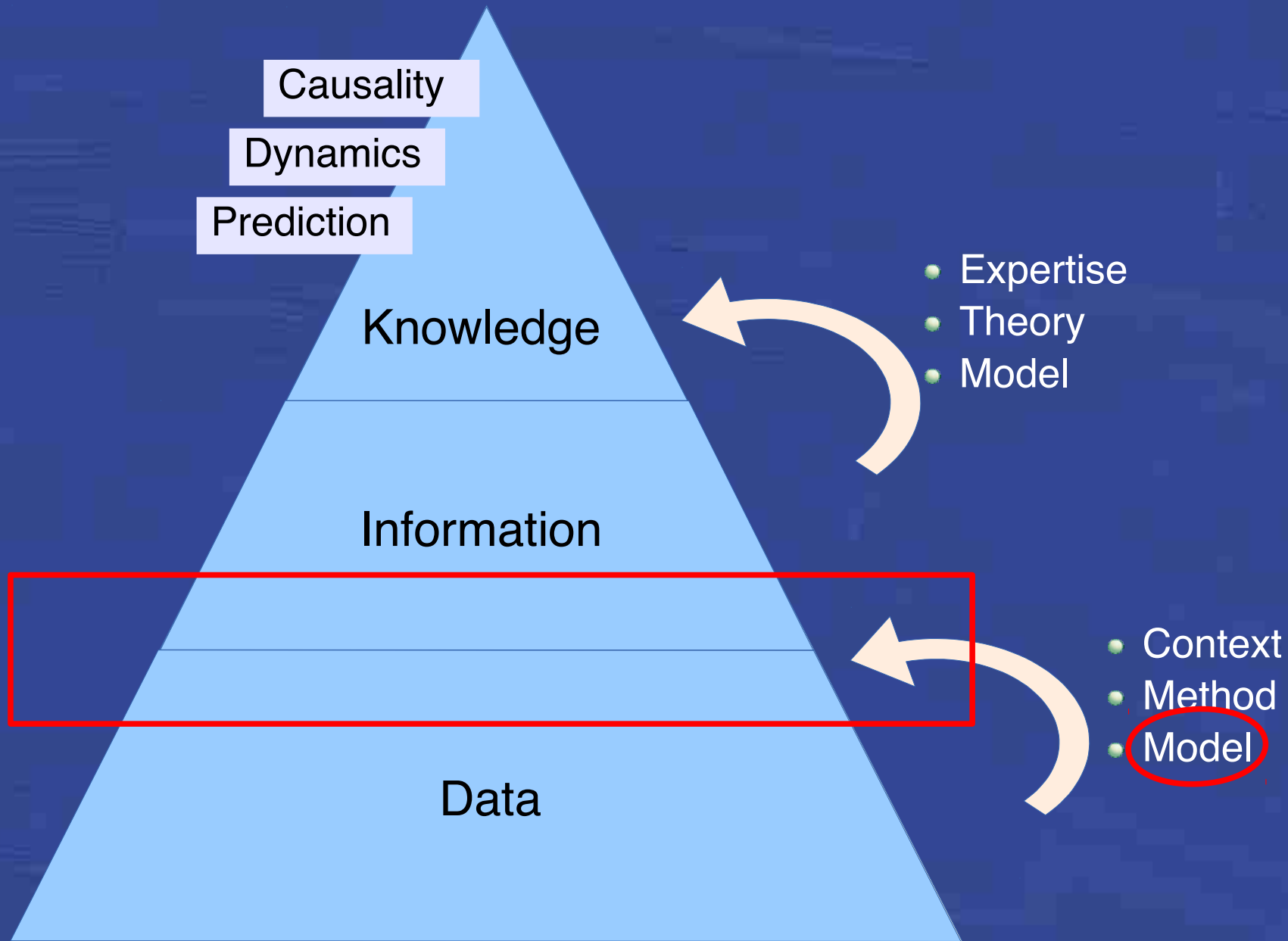
Interpretation
Verification











- * The *symogih.org* project's collaborative virtual research environment
- * Modelling historical knowledge in the context of the semantic web
- * Collaboratively modelling information : *dataforhistory.org*
- * Information extraction from sources : factoids or states of affairs ?

From a resource centered model...

Wikipedia Infobox

Cardinal de l'Église catholique	
Créé cardinal	11 septembre 1606 par le pape Paul V
Titre cardinalice	Cardinal-prêtre de Ss. Quattri Coronati Cardinal-prêtre de S. Lorenzo in Lucina Cardinal-évêque de Frascati

DBPedia.org

http://fr.wikipedia.org/wiki/Giovanni_Garzia_Millini

http://fr.dbpedia.org/page/Giovanni_Garzia_Millini



.. to a statement centered model...

Giovanni Garzia Millini (Q1079973)

Italian catholic cardinal (1562-1629)

> In more languages

Statements

position held

cardinal

start time 11 September 1606 *Gregorian*

1 reference

Catholic Hierarchy person ID [milligg](#)

wikidata.org

Millini,
Giovanni
Garzia

<https://www.wikidata.org/wiki/Q1079973>

statement

ps:P39
(position held)

wd:Q45722
(cardinal)

prov:wasDerivedFrom

<http://www.catholic-hierarchy.org/bishop/bmilligg.html>

pq:P580
(start time)

1606-09-11T00:00:00Z

... to an event centered model

Mellini, Gian Garsia - Nomination: Cardinal

Info93265
Type d'information: Nomination - TyIn6
Date: 1606-09-11

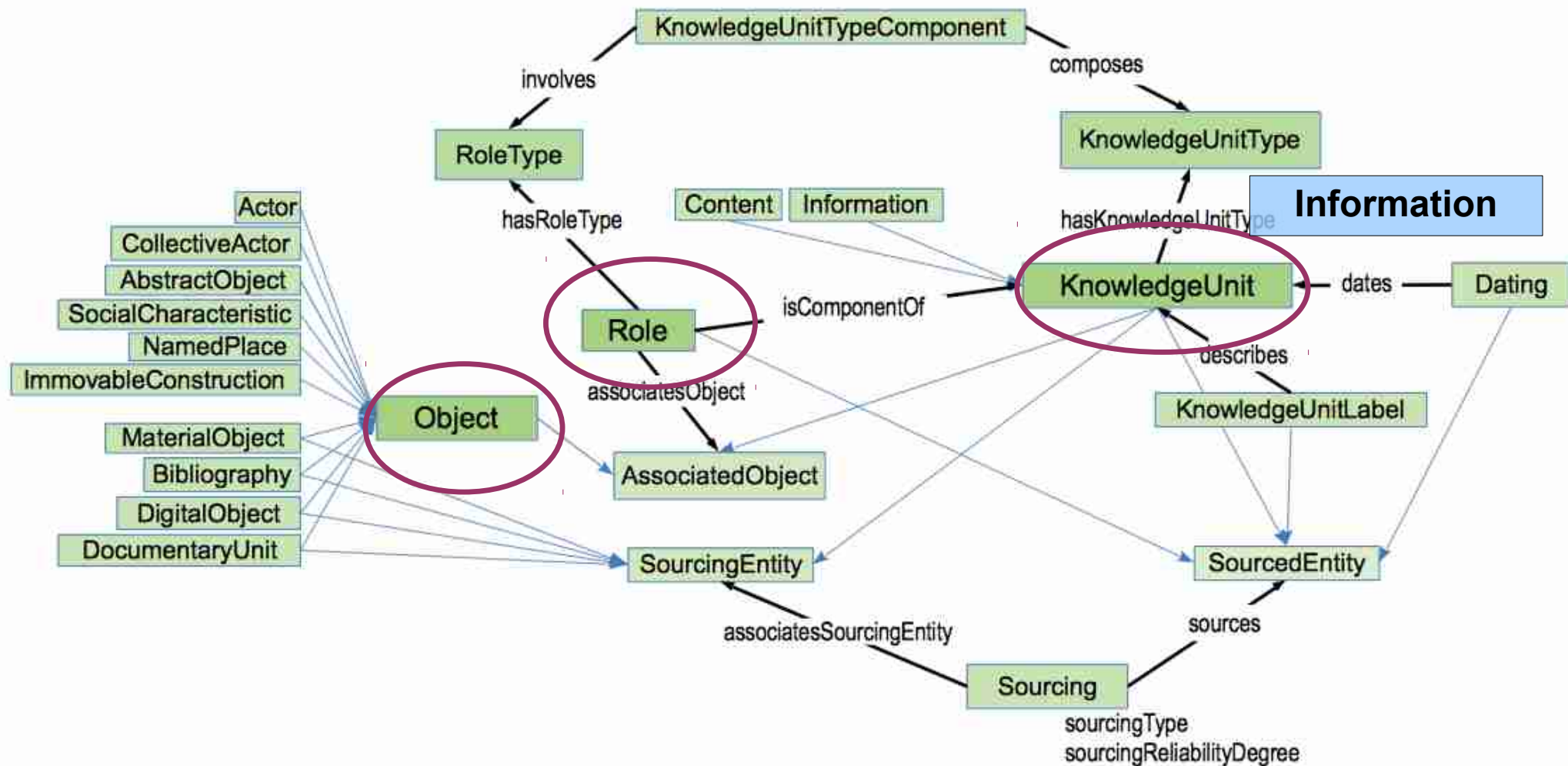
Composantes de l'information

Rôles Textes Sources

Libellé de l'objet	Type de rôle	Clé du rôle
Paulus V	nommer	InRo257887
Mellini, Gian Garsia	nommé (être)	InRo257885
Cardinal	destination (être la)	InRo257886

symogih.org





The symogih.org ontology

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

SYMOGIIH

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 types d'

Chercher une classe

Biographie

Enseignement -

Exercice d'une f

Fin de la vie

Liens acteurs - i

Localisation d'un

Rites sociaux

Vie militaire

Vie professionne

http://symogih.org

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

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

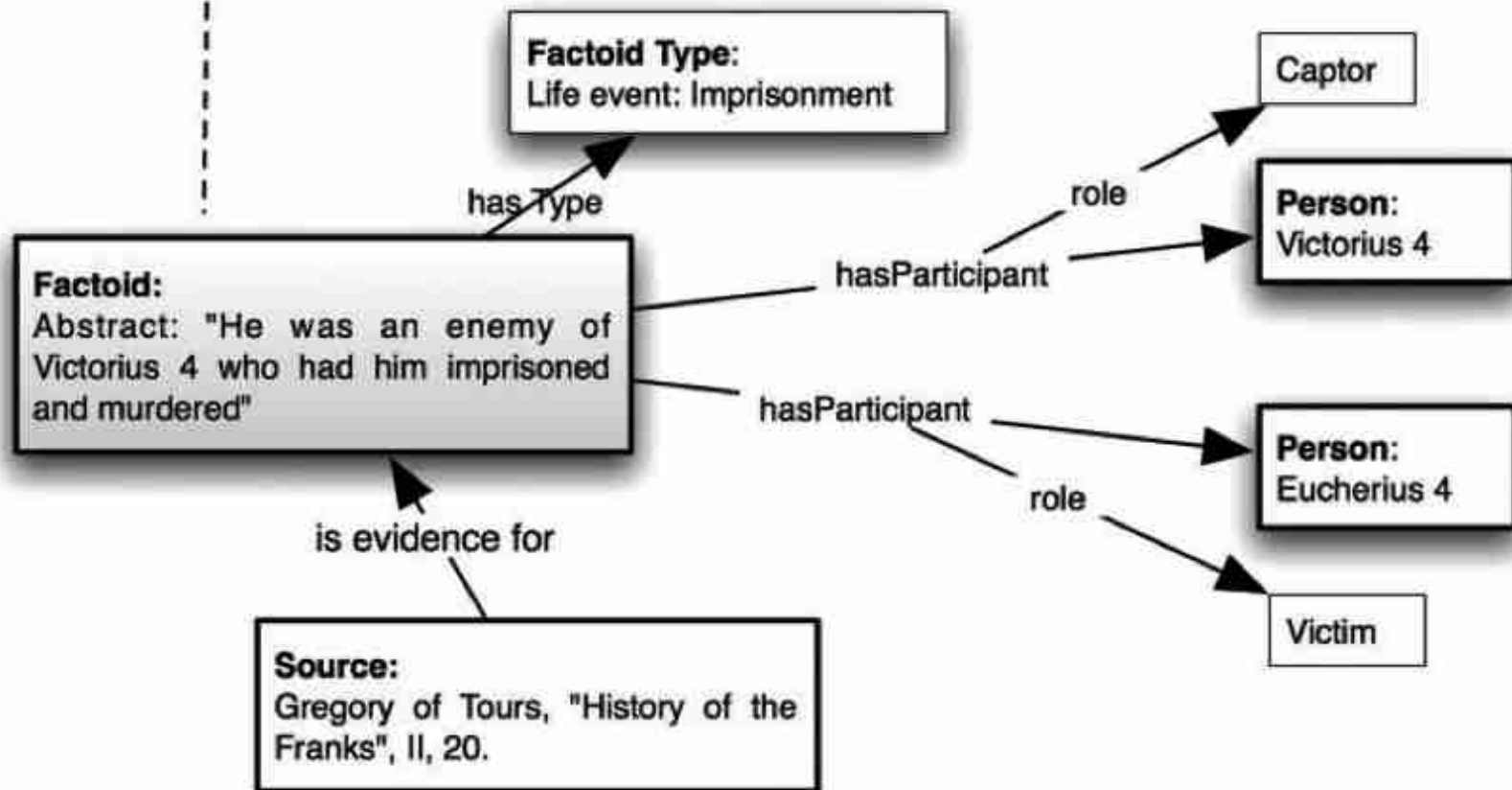
Rôles

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 and 1610



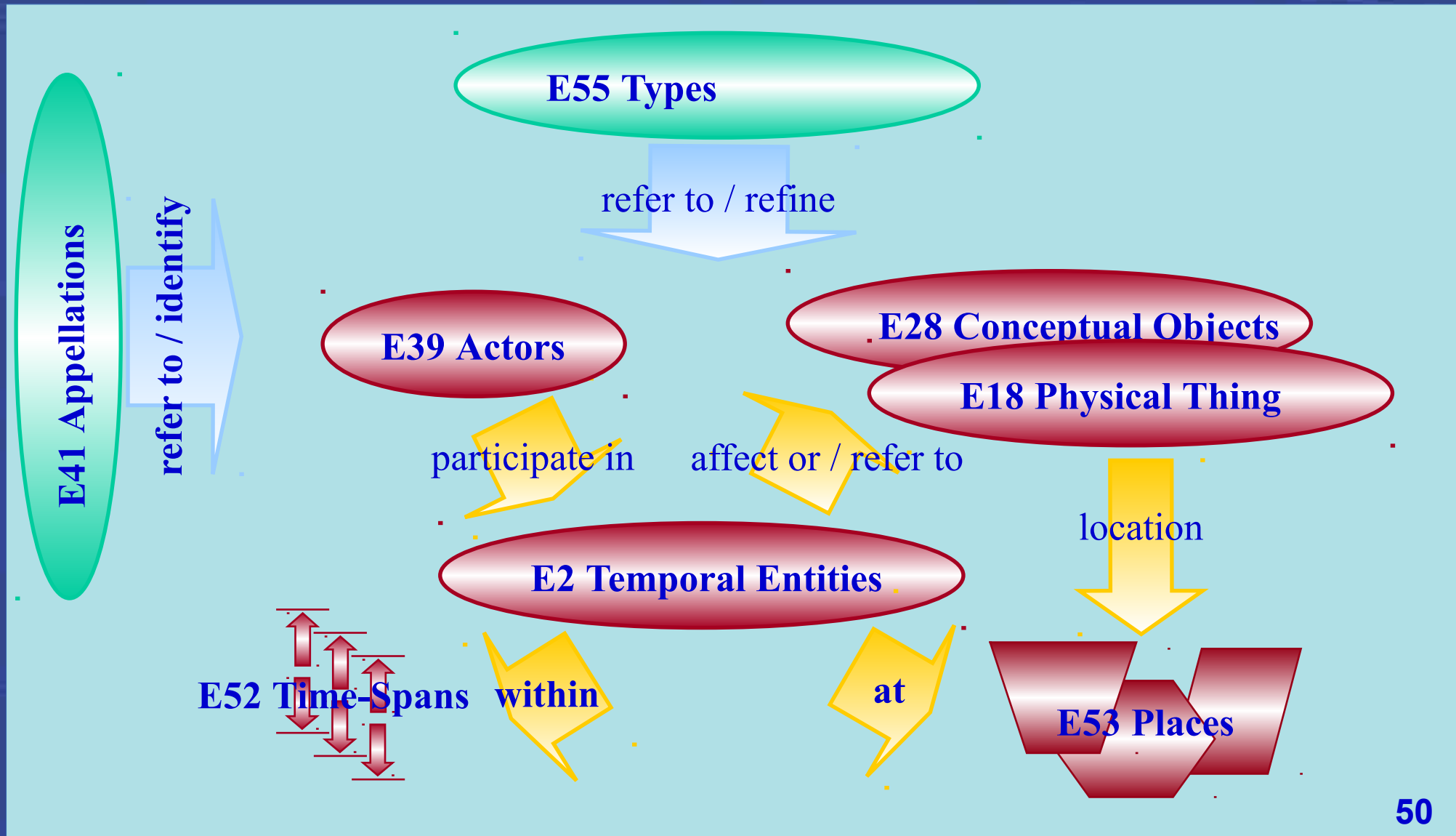
Factoid ontology

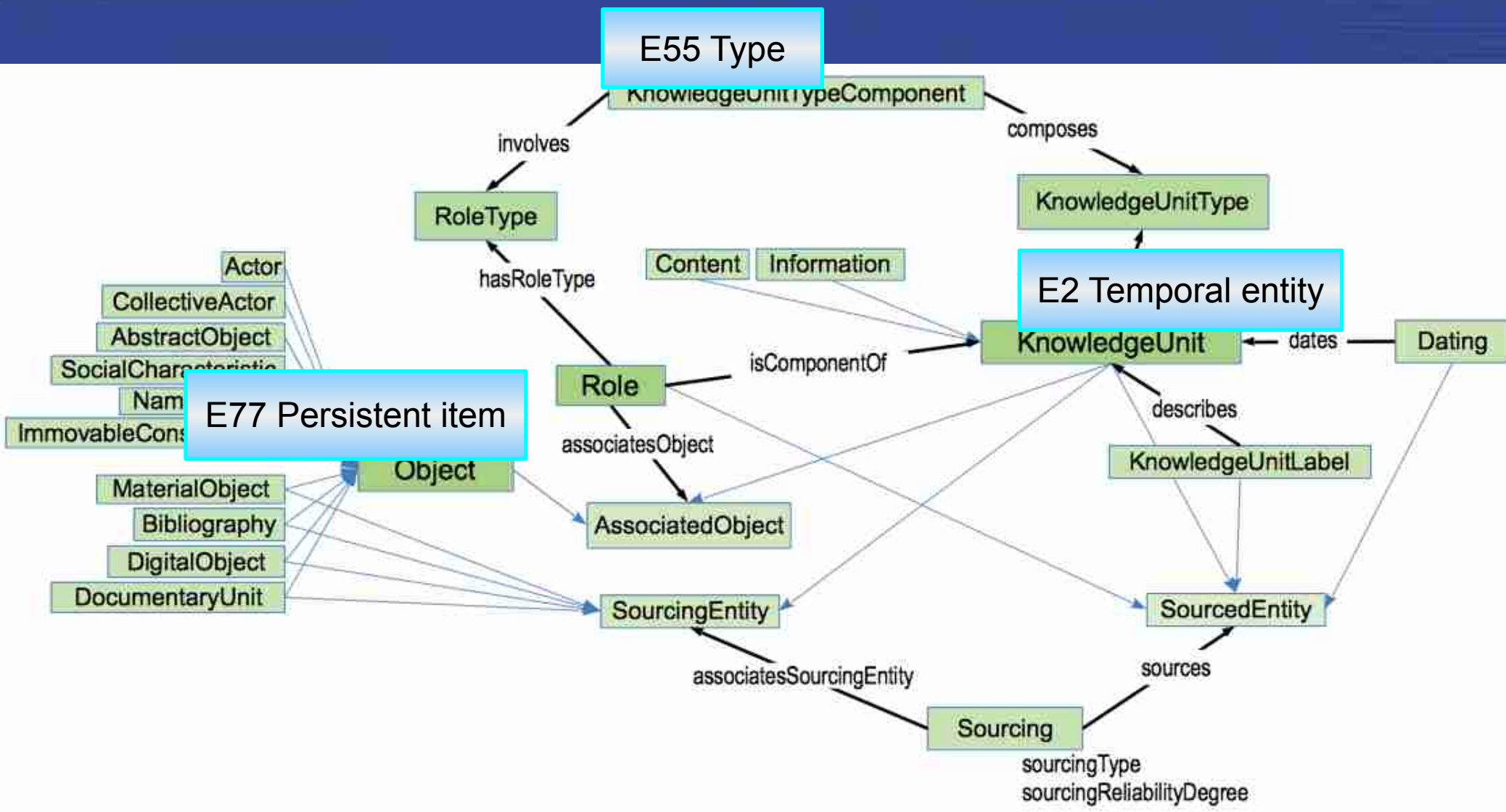
KCL's Department of Digital Humanities (DDH : Prosopographies of the Byzantine World (PBE and PBW), Anglo-Saxon England (PASE), Medieval Scotland (PoMS), Anglo-Scottish cross-border society ('Breaking of Britain': BoB)

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

The CIDOC CRM (ISO21127:2006)

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





Compatible semantic structure of *symogih.org* and the CRM :
modelling states of affairs using an event centered model

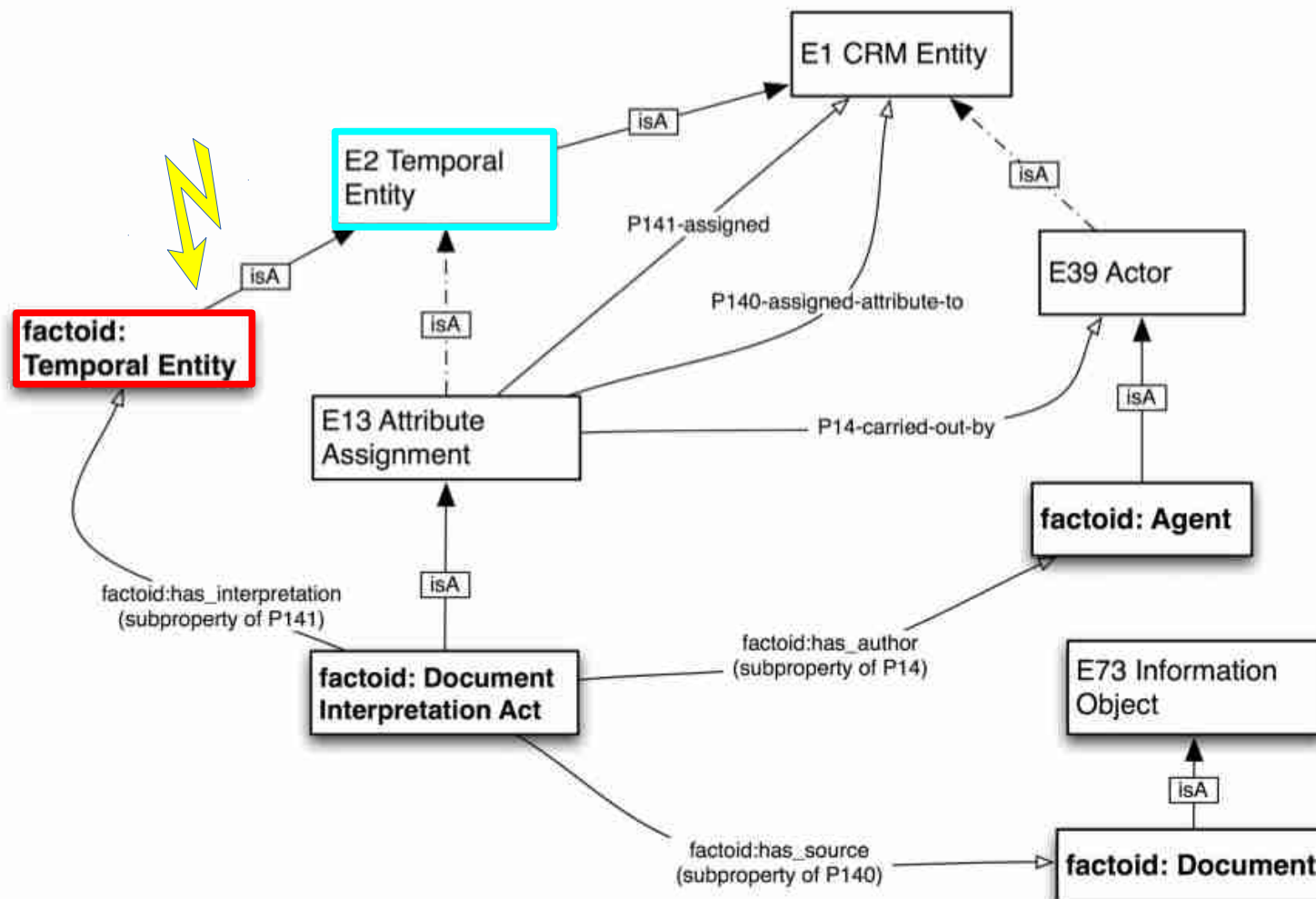
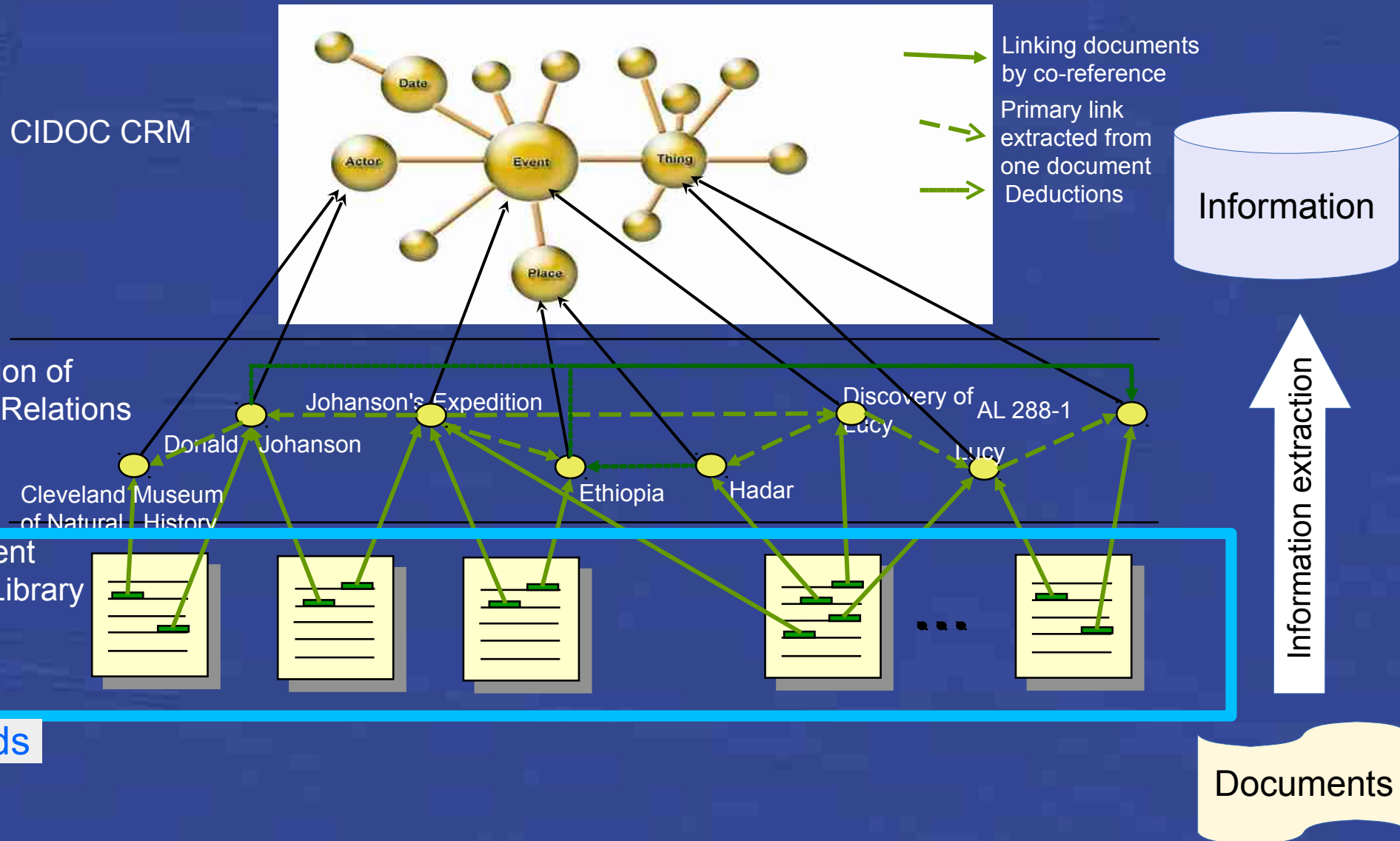


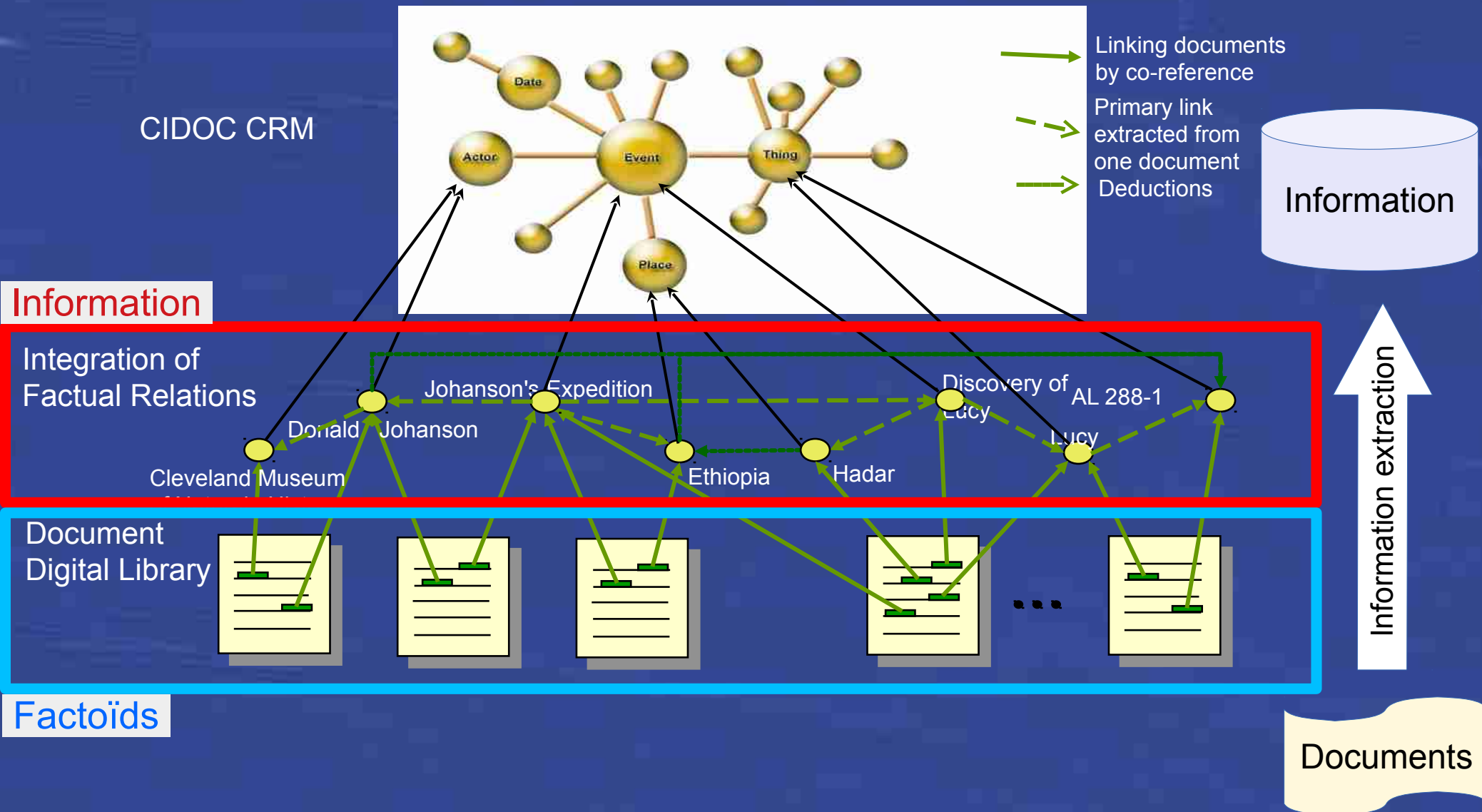
Fig. 4 Mapping the factoid model to CIDOC-CRM primitives

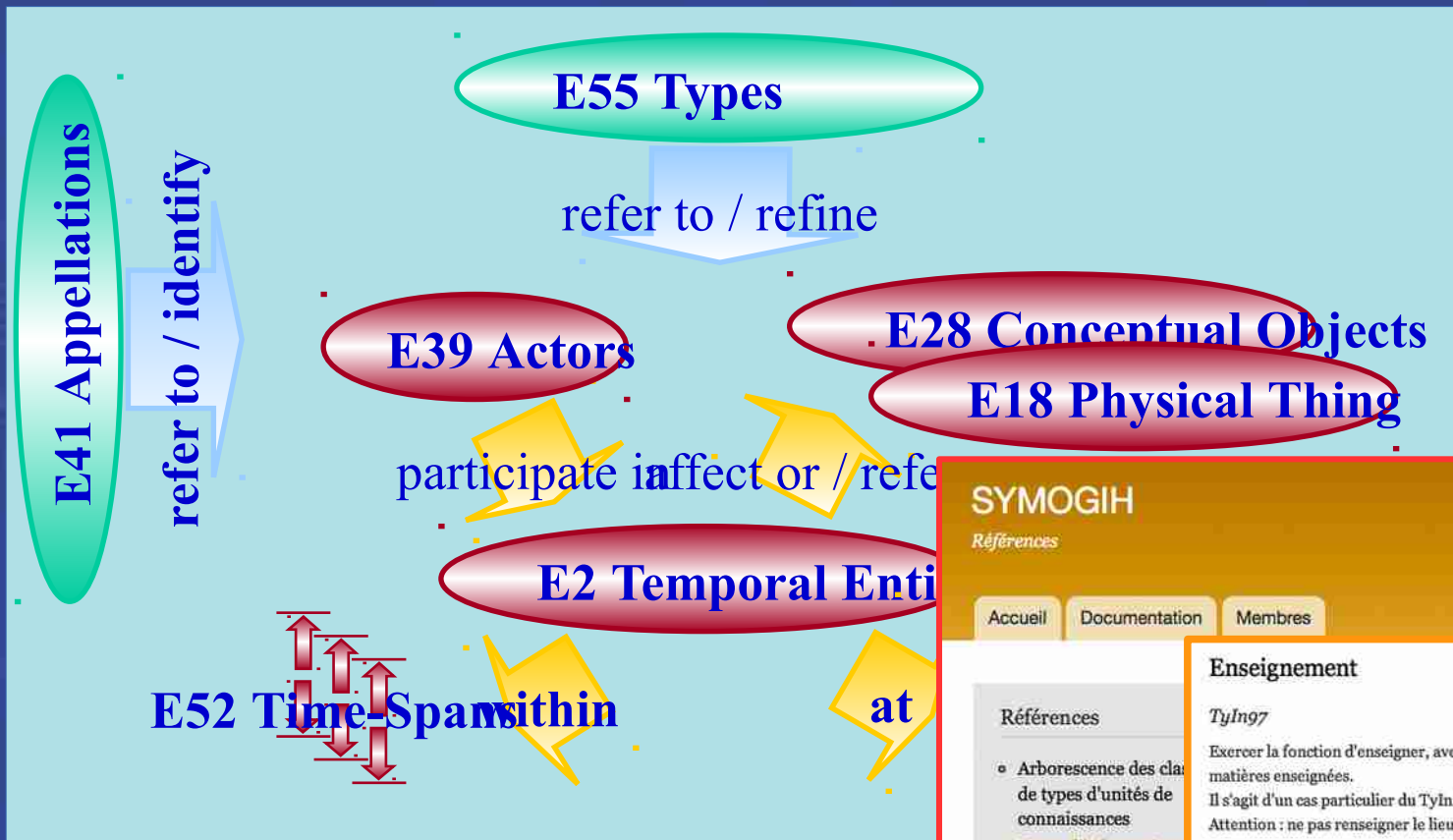
« The factoid approach prioritizes the sources, rather than our historians' reading of them » (Bradley / Pasin).

Integration of information extracted from documents using the CIDOC CRM



Integration of information extracted from documents using the CIDOC CRM





<http://symogih.org>

symogih.org and the CRM :
extending the standard
for geo-historical projects'
data production

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

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.

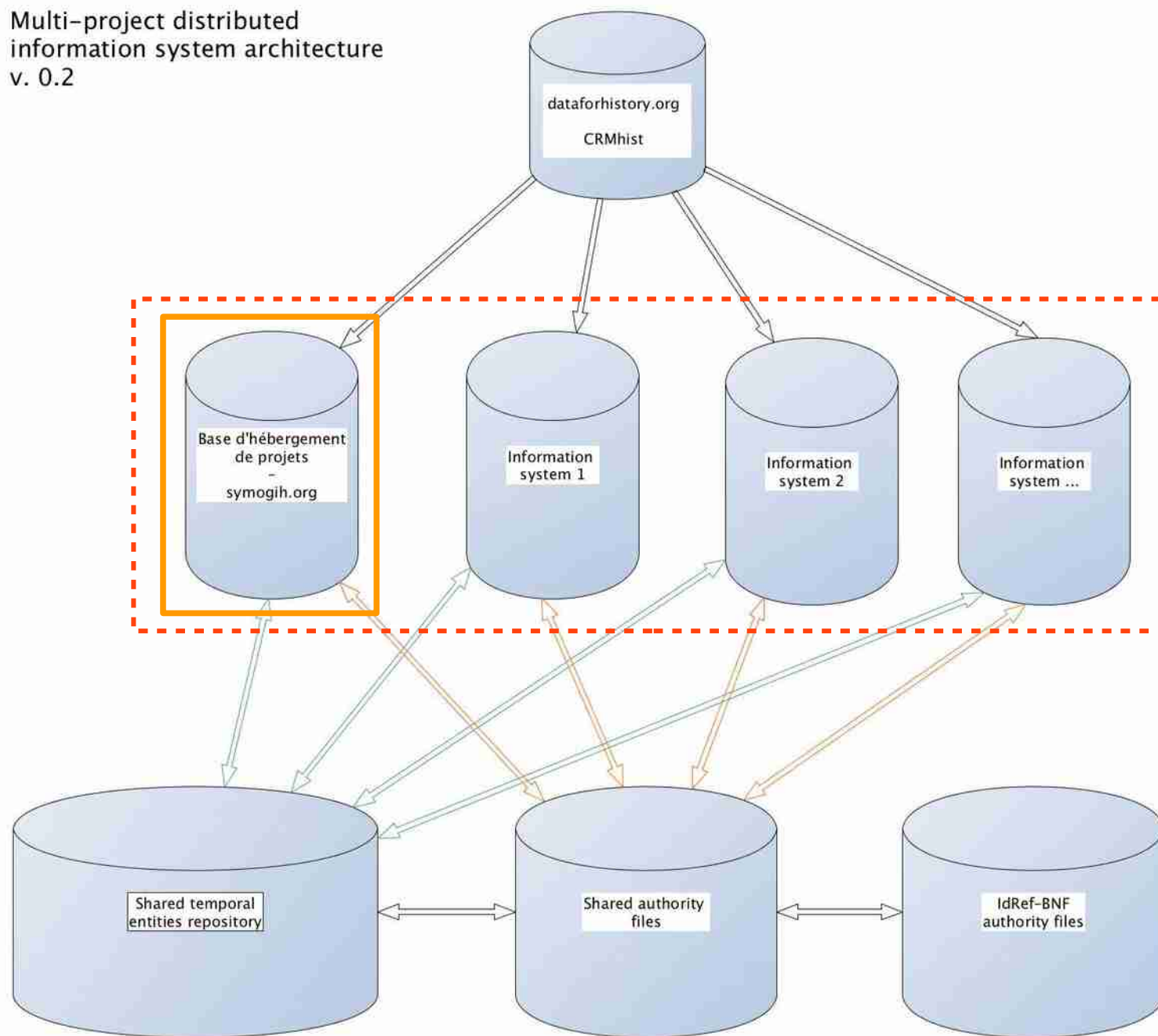
Liste des types de rôles associés

Libellé du type de rôle	Cle 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

Multi-project distributed
information system architecture
v. 0.2



dataforhistory.org

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>



← Précédent

Suivant →

026947676

Lien permanent

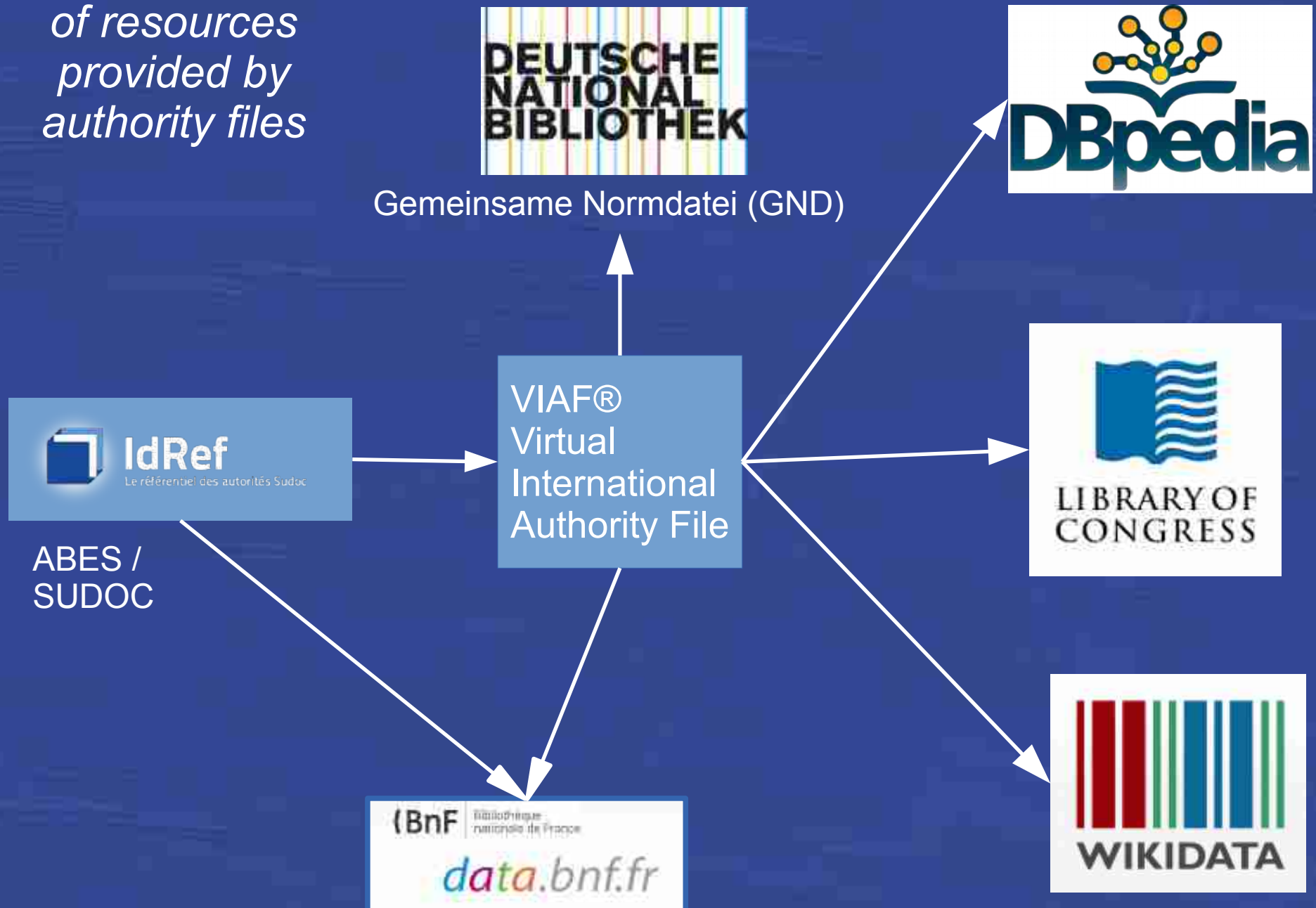
Notice de type
Personne

Forme retenue

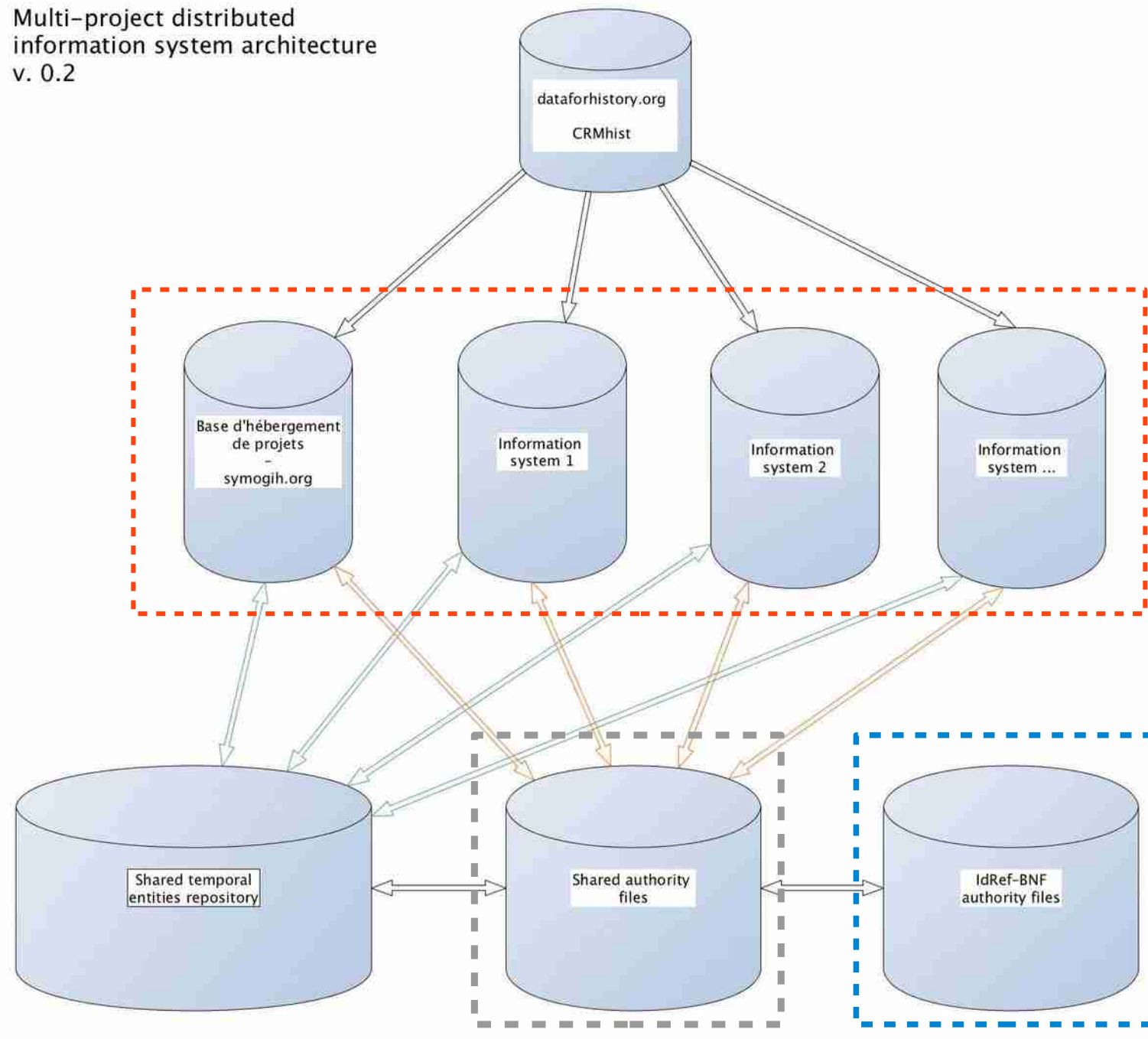
Kepler, Johannes (1571-1630)

<http://www.idref.fr/026947676>

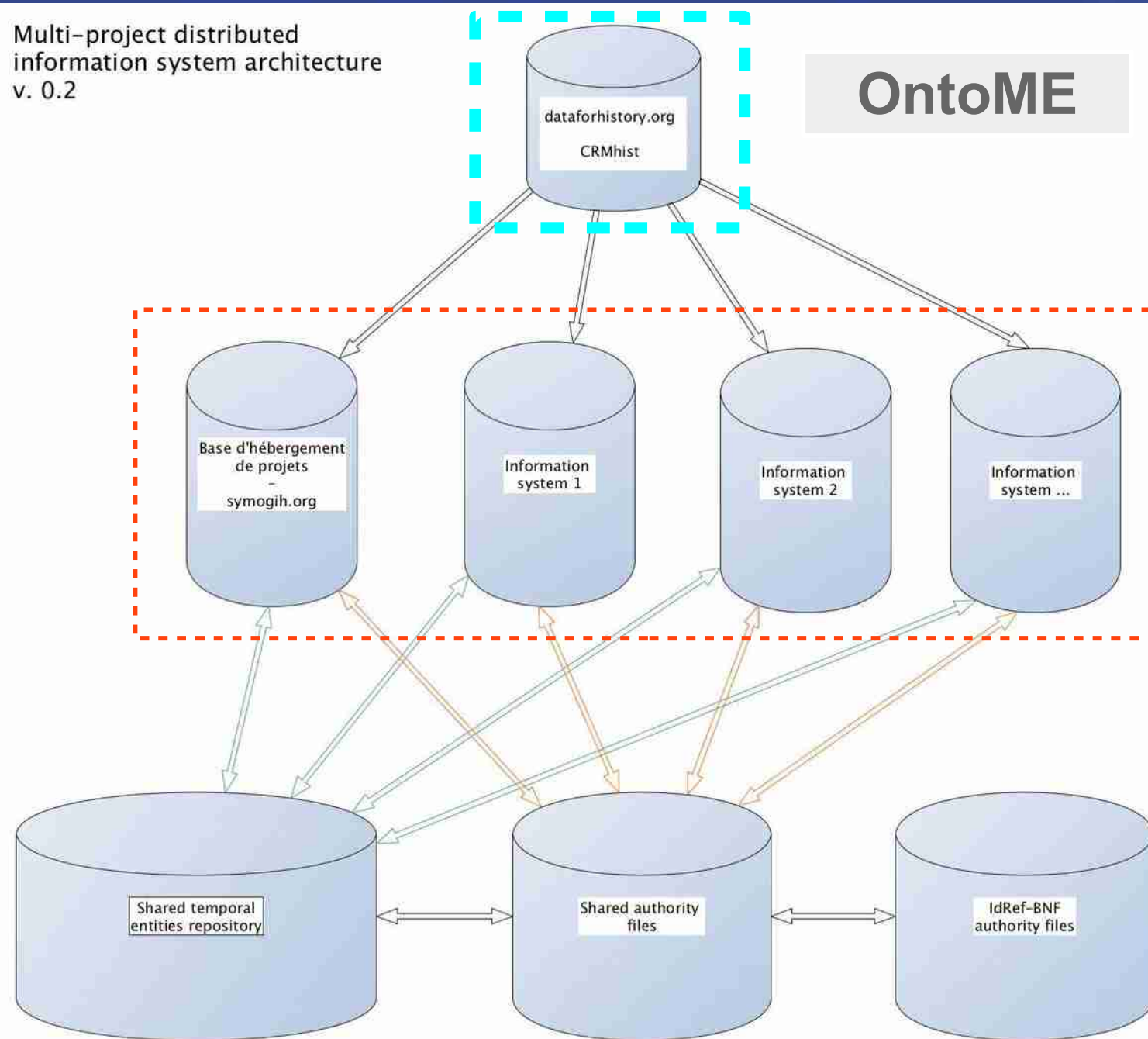
*Interlinking the URIs
of resources
provided by
authority files*



Multi-project distributed
information system architecture
v. 0.2



dataforhistory.org



<http://ontome.dataforhistory.org>

Home Classes Properties Namespaces Projects Profiles Dashboard

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 : Ma

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: $C2(x) \supset E22(x)$

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

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

CIDOC CRM

CIDOC CRM



CRM hist extension

CIDOC CRM



CRM hist extension



Projects' ontology

CIDOC CRM



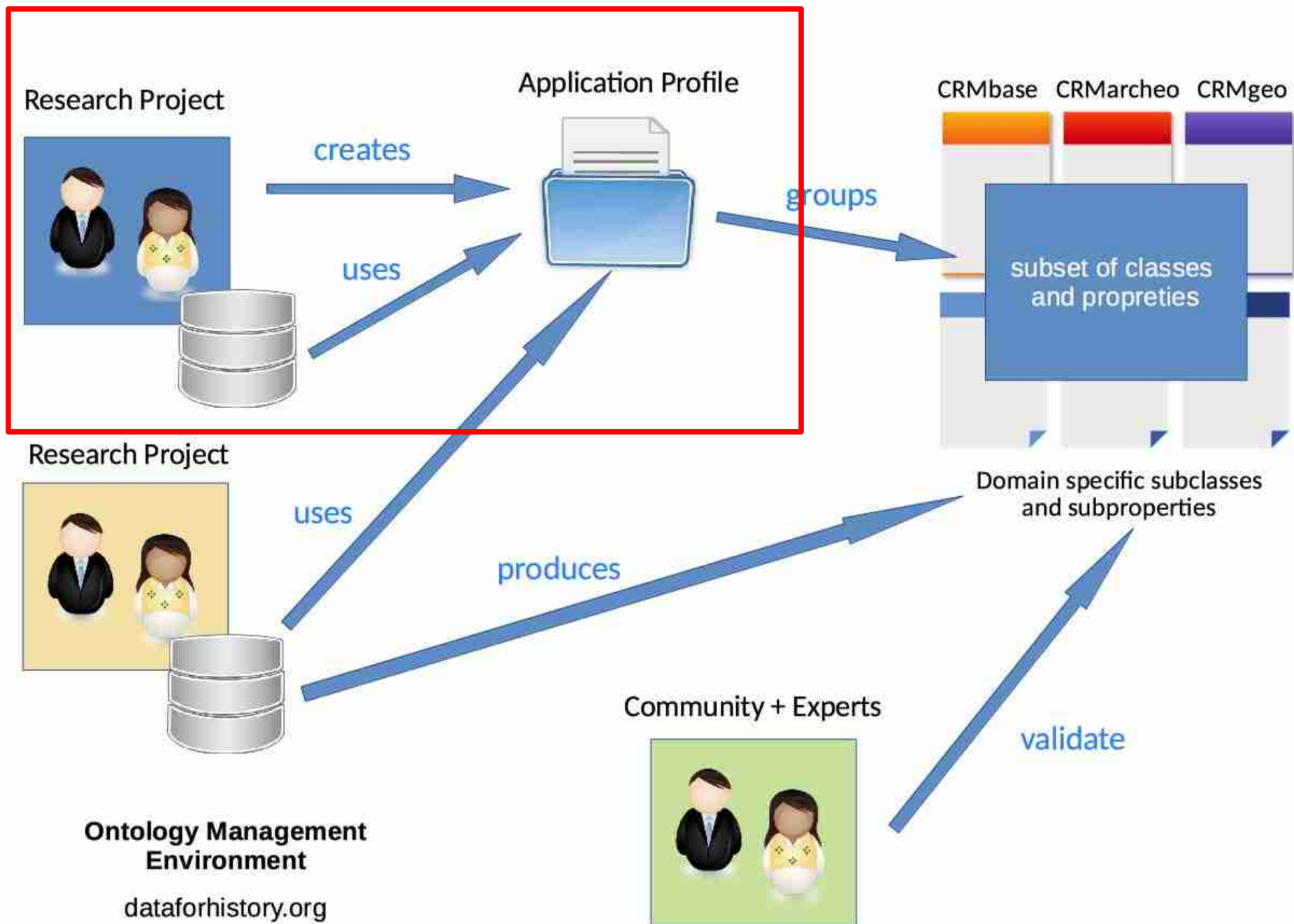
CRM hist extension



Projects' ontology



Information systems



Retrieve your project's application profiles from an API

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

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

Ressources FB Projets FB Urgenti UEO En cours Standards Ressour./Projets Actualités Methodes

JSON Données brutes En-têtes

Enregistrer Copier

▼ 5:

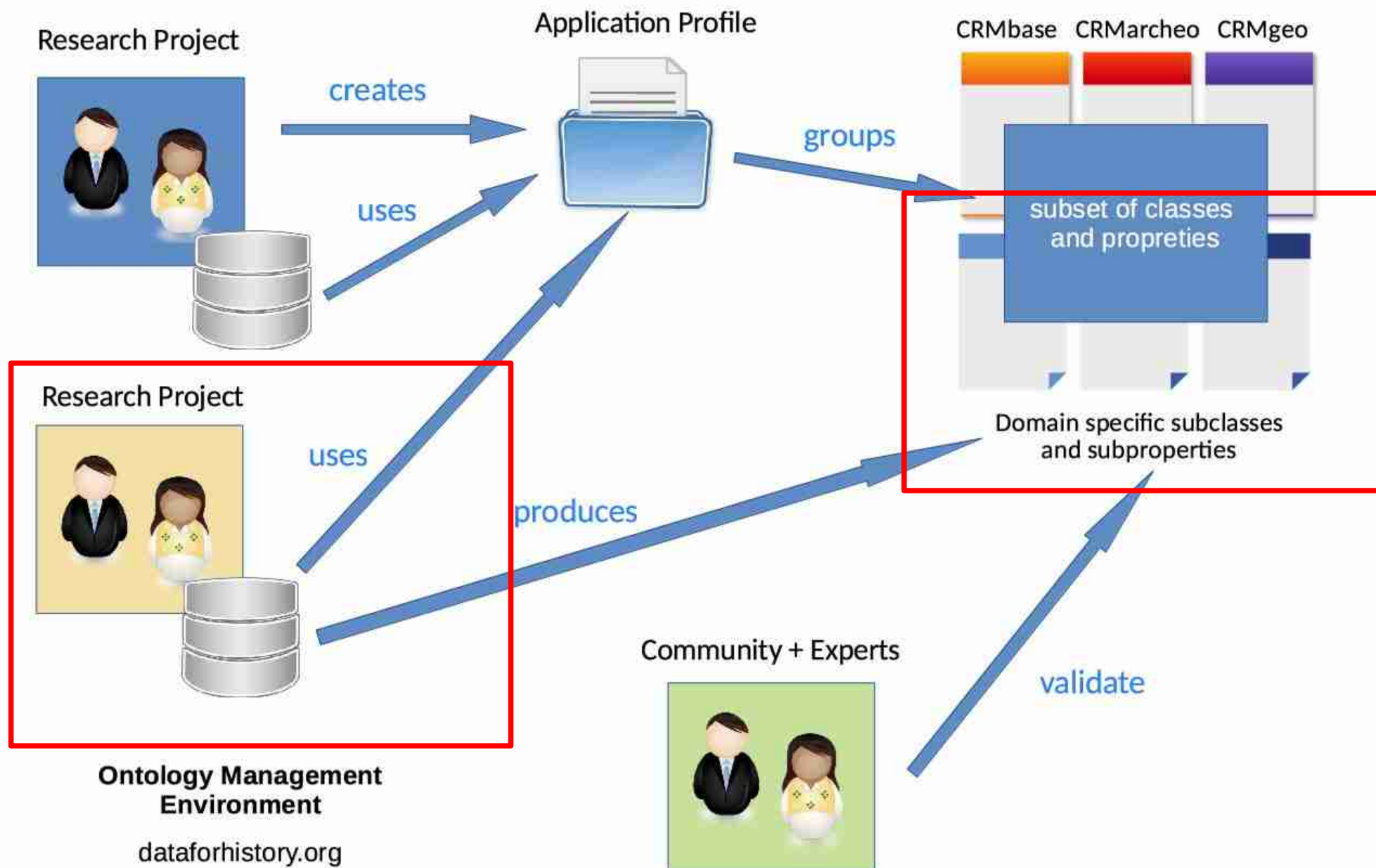
pk_class:	61
identifiant_in_namespace:	"E67"
class_standard_label:	"Birth"
fk_system_type:	9
type_label:	"Temporal Entity"
root_namespace:	"CIDOC CRM"
profile_association_type:	"selected"
pk_profile:	1
profile_label:	"BHP new data model v. 0.1"
pk_project:	1
project_label:	"BHP – symogih.org"

▼ 6:

pk_class:	340
identifiant_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 : a VRE for historical research



Design the classes and properties of your project's namespace

<https://ontome.dataforhistory.org/api/namespaces-rdf-owl.rdf?namespace=3>

OntoME
Ontology Management Environment - beta version

Logout

Home Classes Properties Namespaces Projects Profiles Dashboard User guide

Active project : Maritime history

CIDOC CRM Generic Extension for Historical Data Management and Interoperability, ongoing

Summary Definition Identification Hierarchy Classes Properties Graph Comments

Classes

Show 25 entries

Class

- Appellation for language – histC10
- Argument – histC15
- Argument's method – histC22
- Built work type – histC13

Class hierarchy: histC6

- owl:Thing
 - D1
 - E1
 - E13
 - E2
 - histC1
 - histC10
 - histC2
 - histC3
 - histC7
 - histC6
 - E21
 - E24
 - histC11
 - E26
 - histC8
 - E39
 - E41
 - E5
 - E54

Annotation property hierarchy: Datatypes

Object property hierarchy: Data property hierarchy: Individuals by type

Object property hierarchy: Asserted

- owl:topObjectProperty

Annotations: histC8

Annotations

- rdfs:label [language: en] Geographical Place
- rdfs:comment [type: xsd:string] This class refers to portions of the surface of the Earth intended as constellations of matter which can be represented by photographs, paintings and maps. The relevant portion of the surface of the Earth can be covered by water (river, sea, ...). The more specific identity of instances of this class is provided by a controlled vocabulary of geographical place types.
- skos:notation [type: xsd:string] histC8

Description: histC8

Equivalent To

SubClass Of

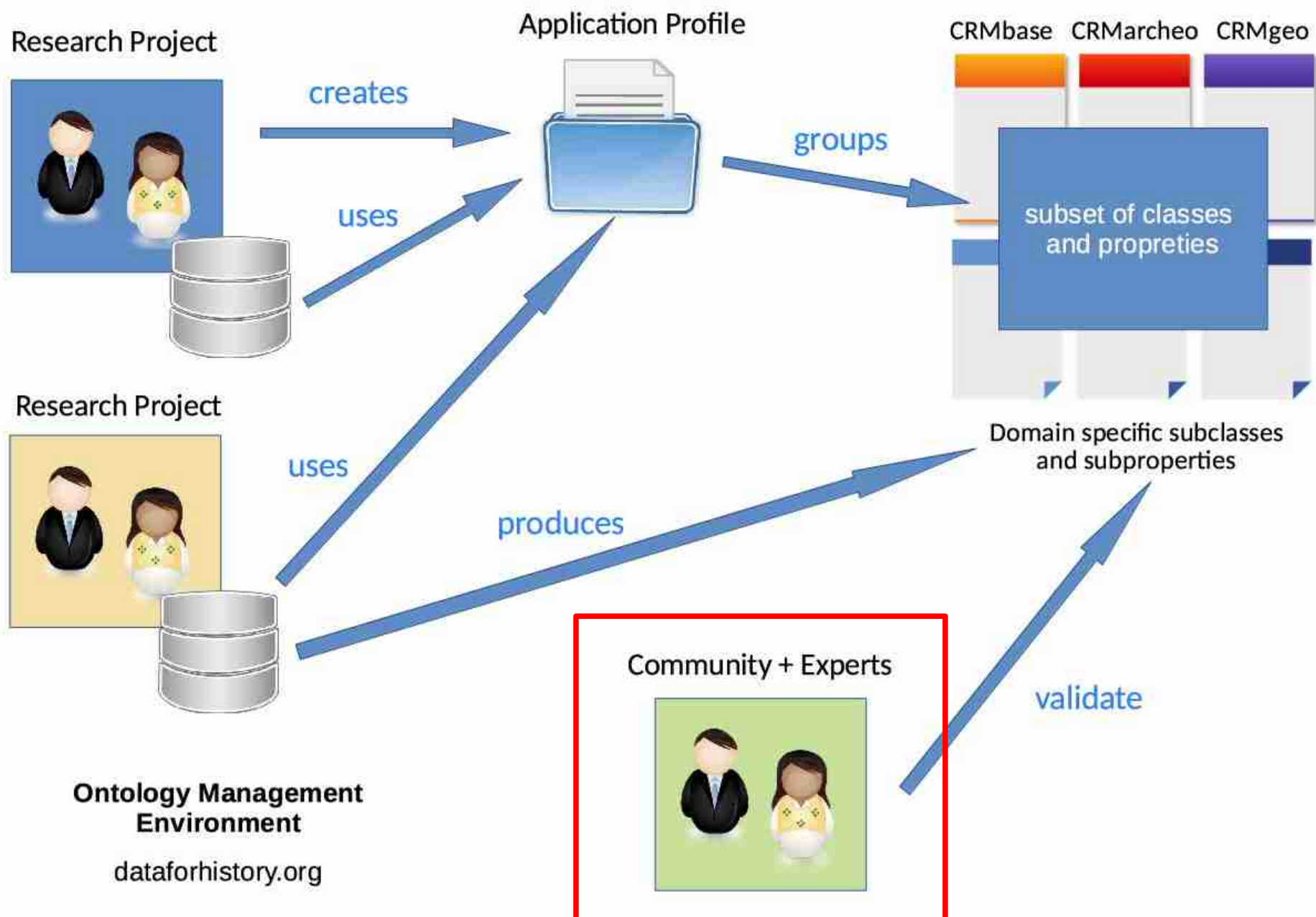
- E26

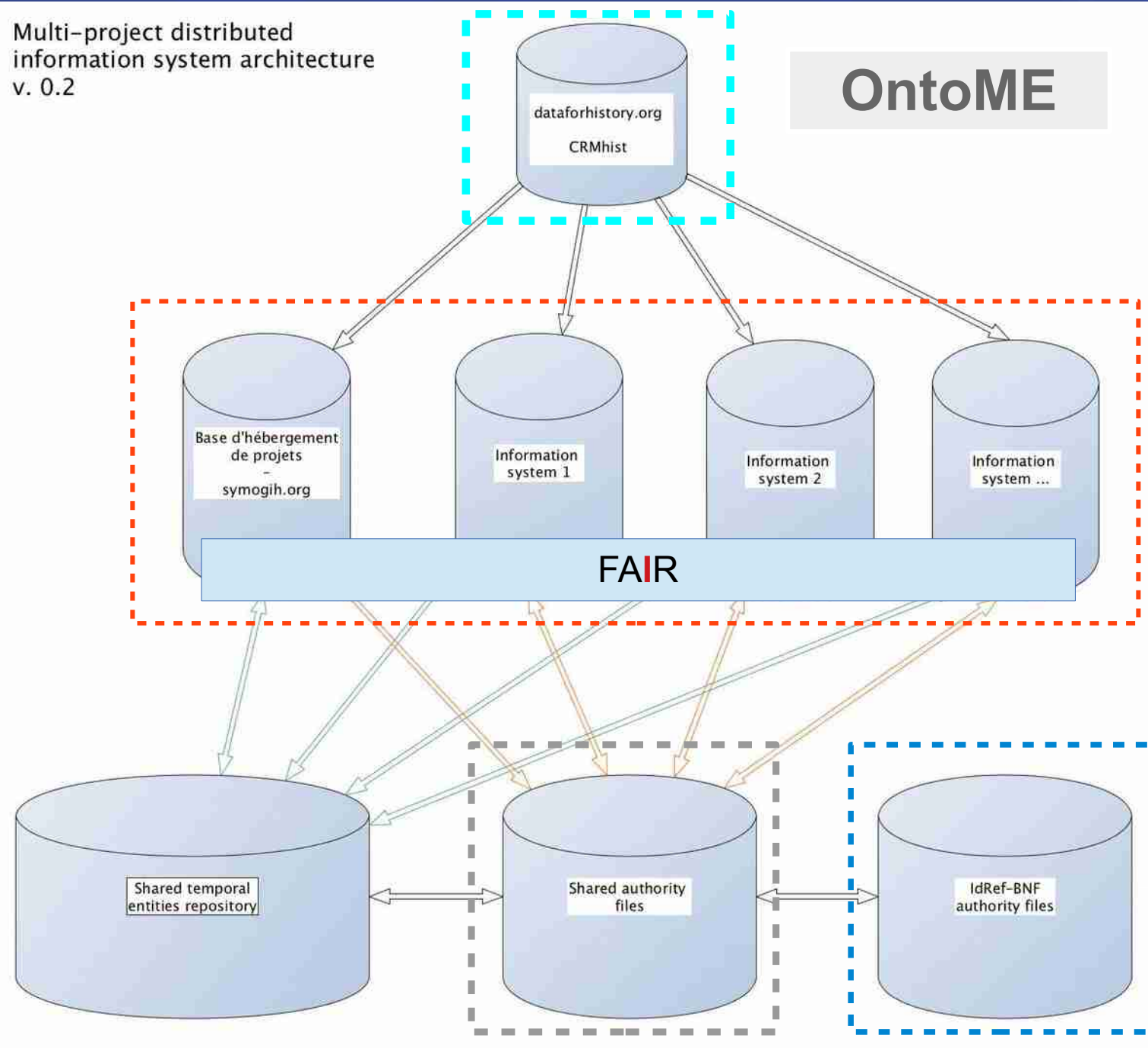
General class axioms

SubClass Of (Anonymous Ancestor)

Instances

No Reasoner set. Select a reasoner from the Reasoner menu. Show Inferences





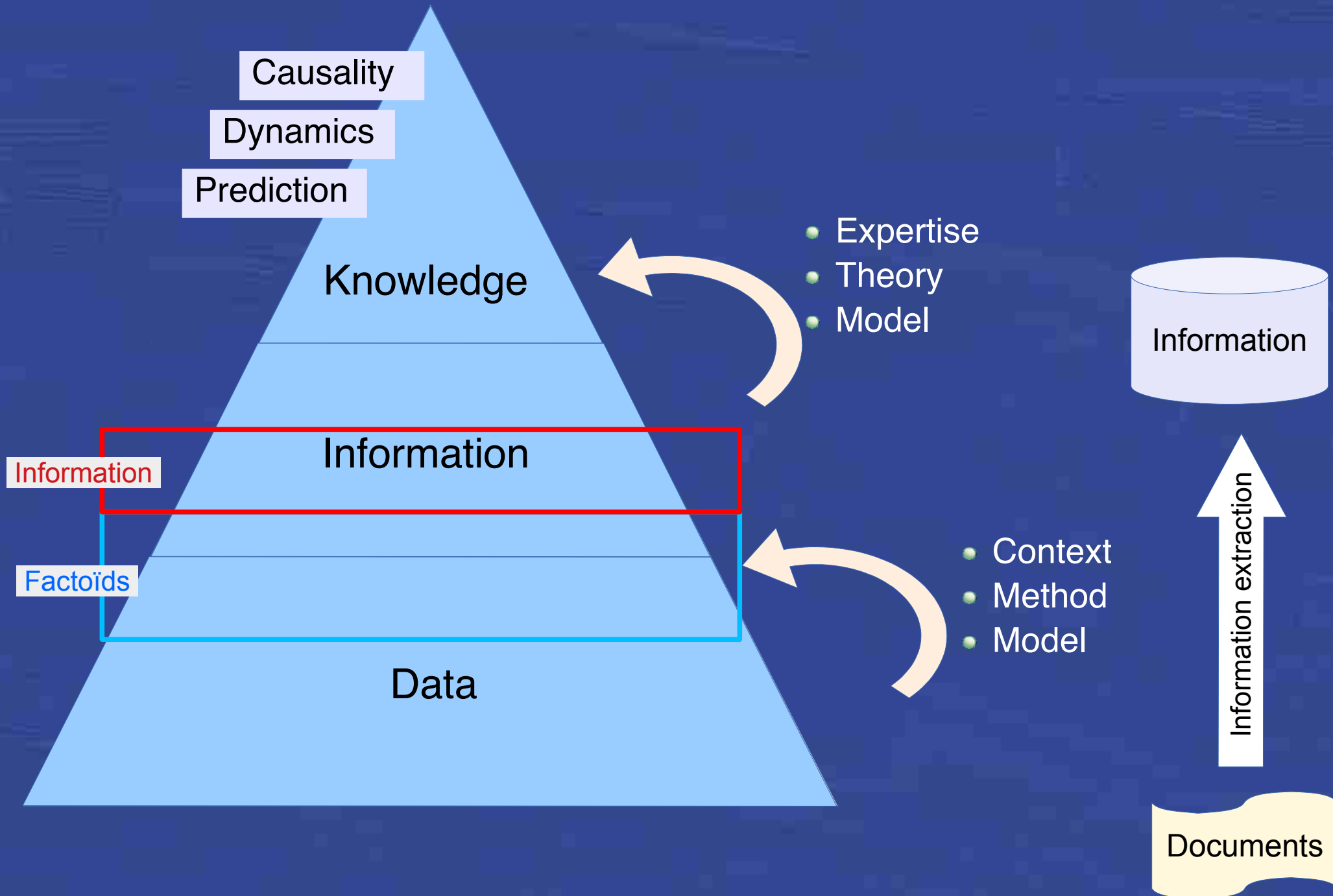
Data for History Consortium

Meetings

- Lyon, November 2017 (founding meeting)
- Lyon, Mai 2018
- Panel presenting the Data for History vision
at the EADH 2018 conference in Galway (December 2018)
- Leipzig, 4-5 April 2019
- Forthcoming : Berlin, 28-29 Mai 2020

The Data for History consortium
is in the process of being formally established
and is **open to all interested institutions and researchers**

- * The *symogih.org* project's collaborative virtual research environment
- * Modelling historical knowledge in the context of the semantic web
- * Collaboratively modelling information : *dataforhistory.org*
- * Information extraction from sources : factoids or states of affairs ?

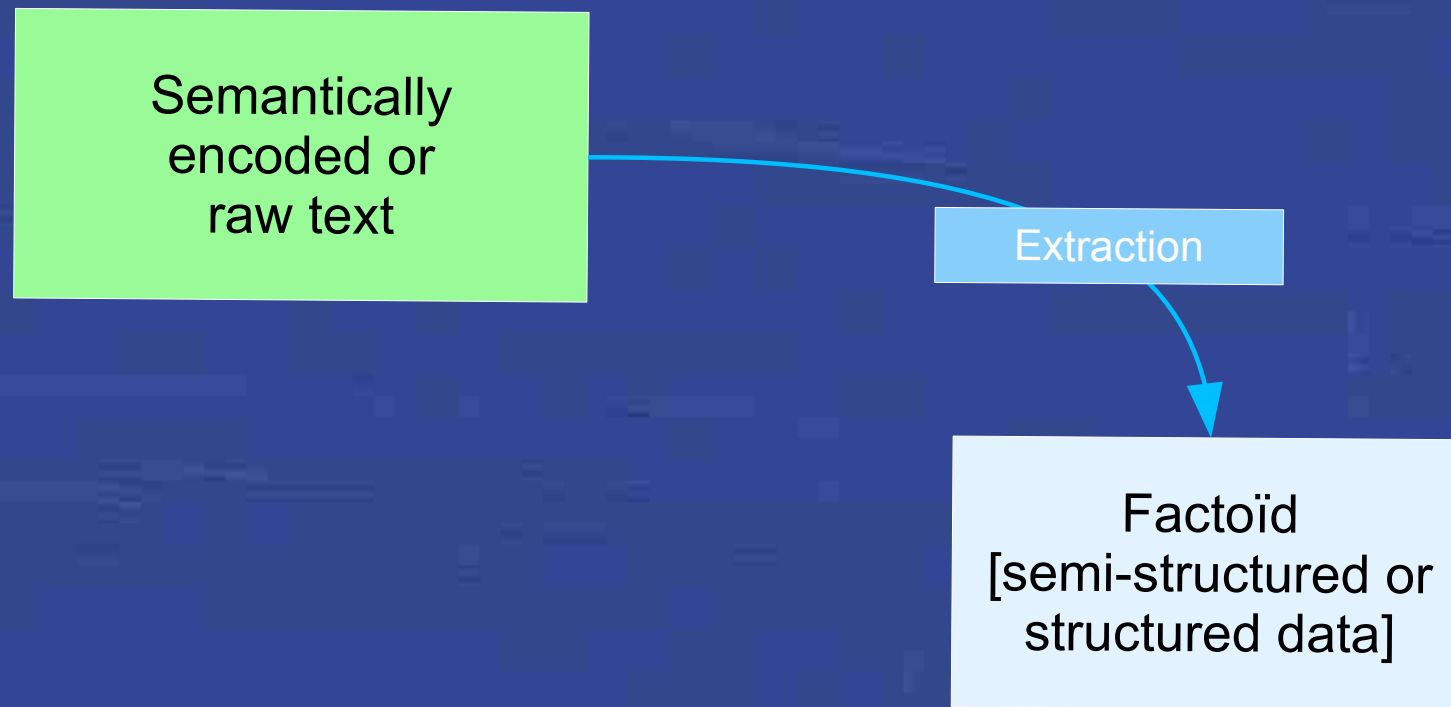


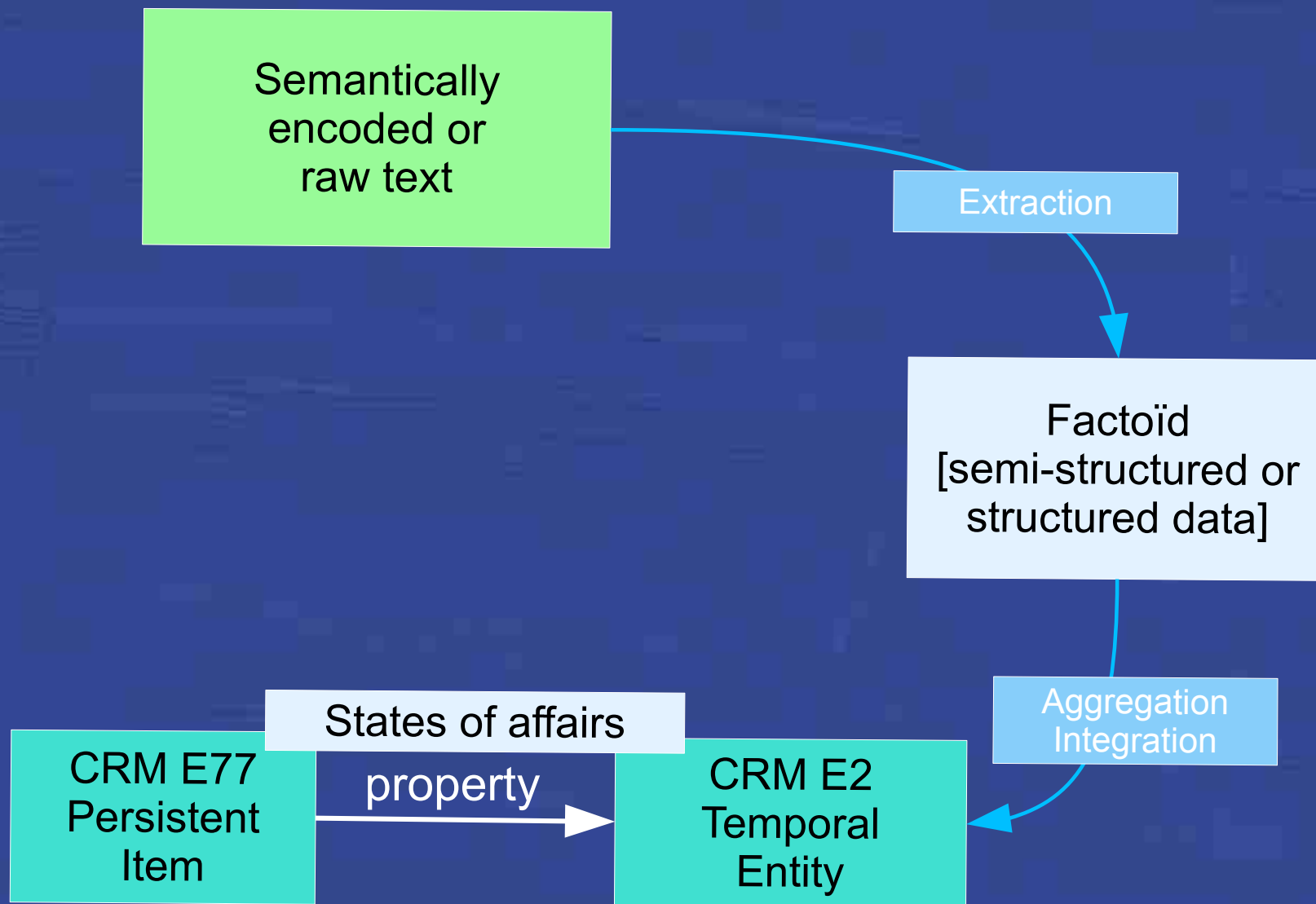
Manual encoding of the text or use of NLP techniques to identify information and produce structured data

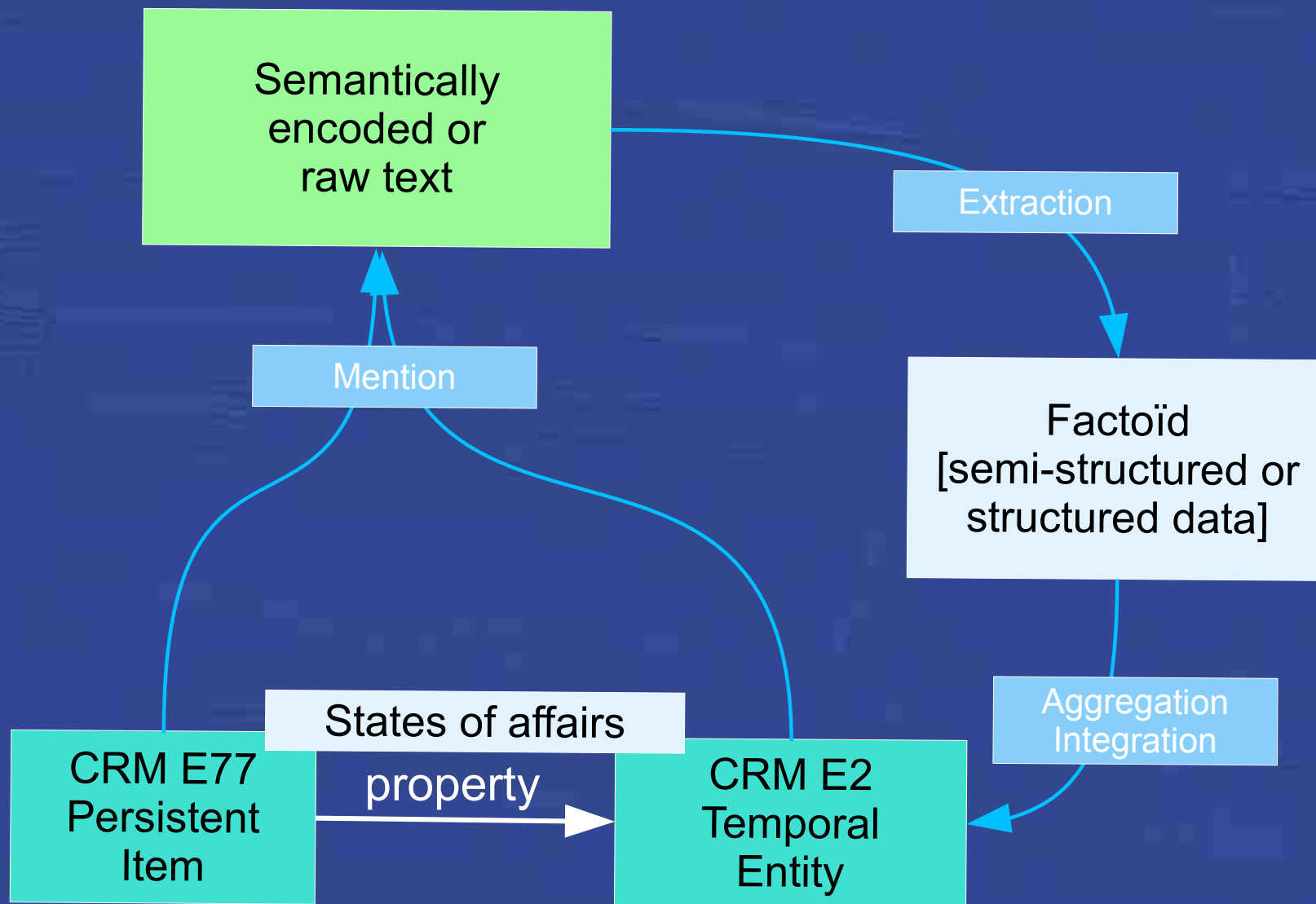
<p corresp="Info29659"><name ref="Actr195">Johannes Kepler</rs>
was born on <date when="1571-12-27">December 27, 1571</date> in
<name ref="NaPI90073">Weil der Stadt</name></p>

```
{"Tokens" : [  
  {"doc" : 1, "sentence" : 6, "token_id":1,  
"token" : Kepler}, was  
  {"doc" : 1, "sentence" : 6, "token_id":2,  
"token" : was},  
  {"doc" : 1, "sentence" : 6, "token_id":3,  
"token" : born},  
  {"doc" : 1, "sentence" : 6, "token_id":4,  
"token" : on}, was  
  {"doc" : 1, "sentence" : 6, "token_id":5,  
"token" : December},  
  {"doc" : 1, "sentence" : 6, "token_id":6,  
"token" : 27},  
  {"doc" : 1, "sentence" : 6, "token_id":7,  
"token" : 1571},  
  ... ]}
```

Chunk [Tokens ids]	Entity	Value
1	Actr195	
5,6,7		1571-12-27



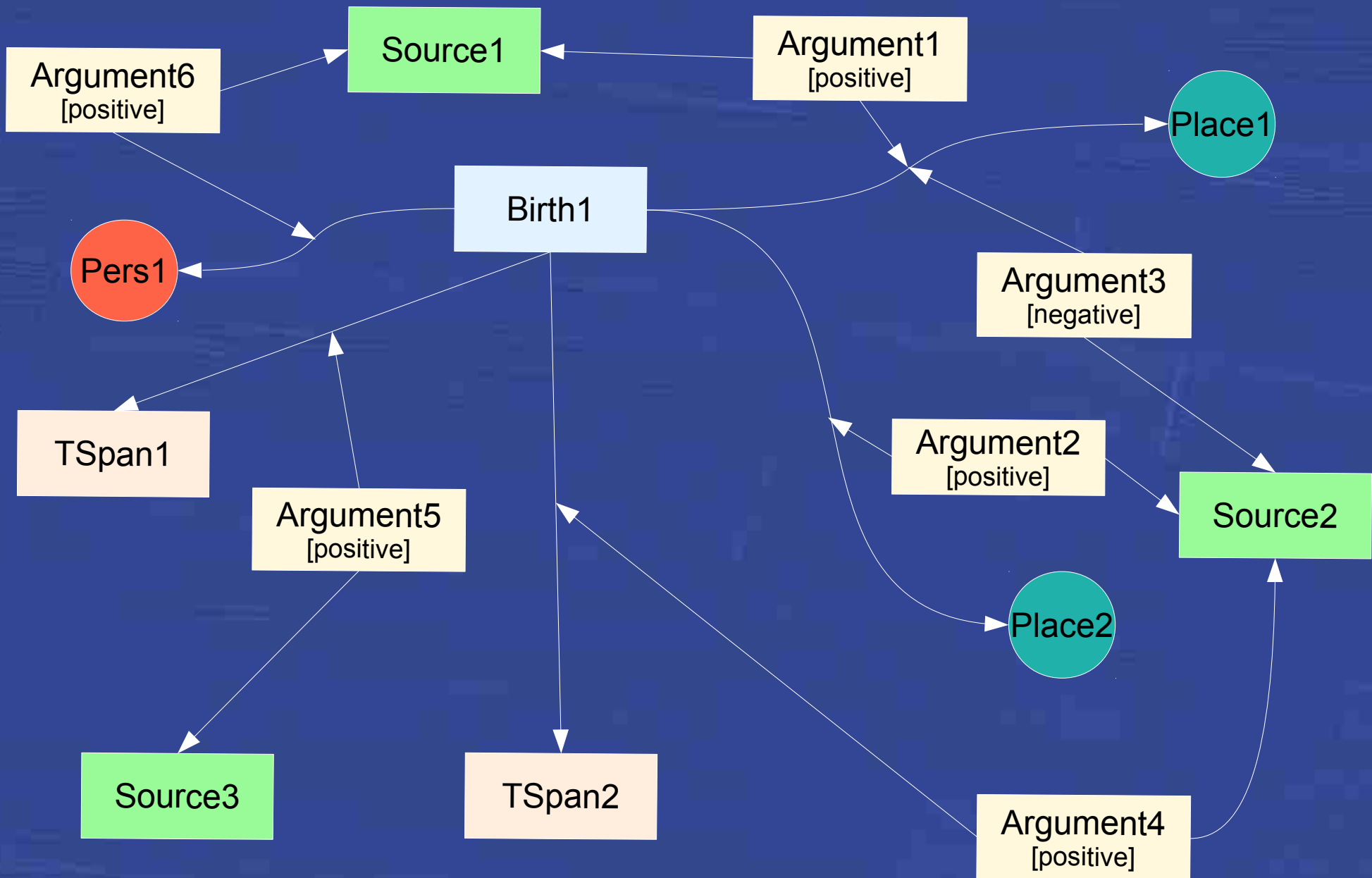


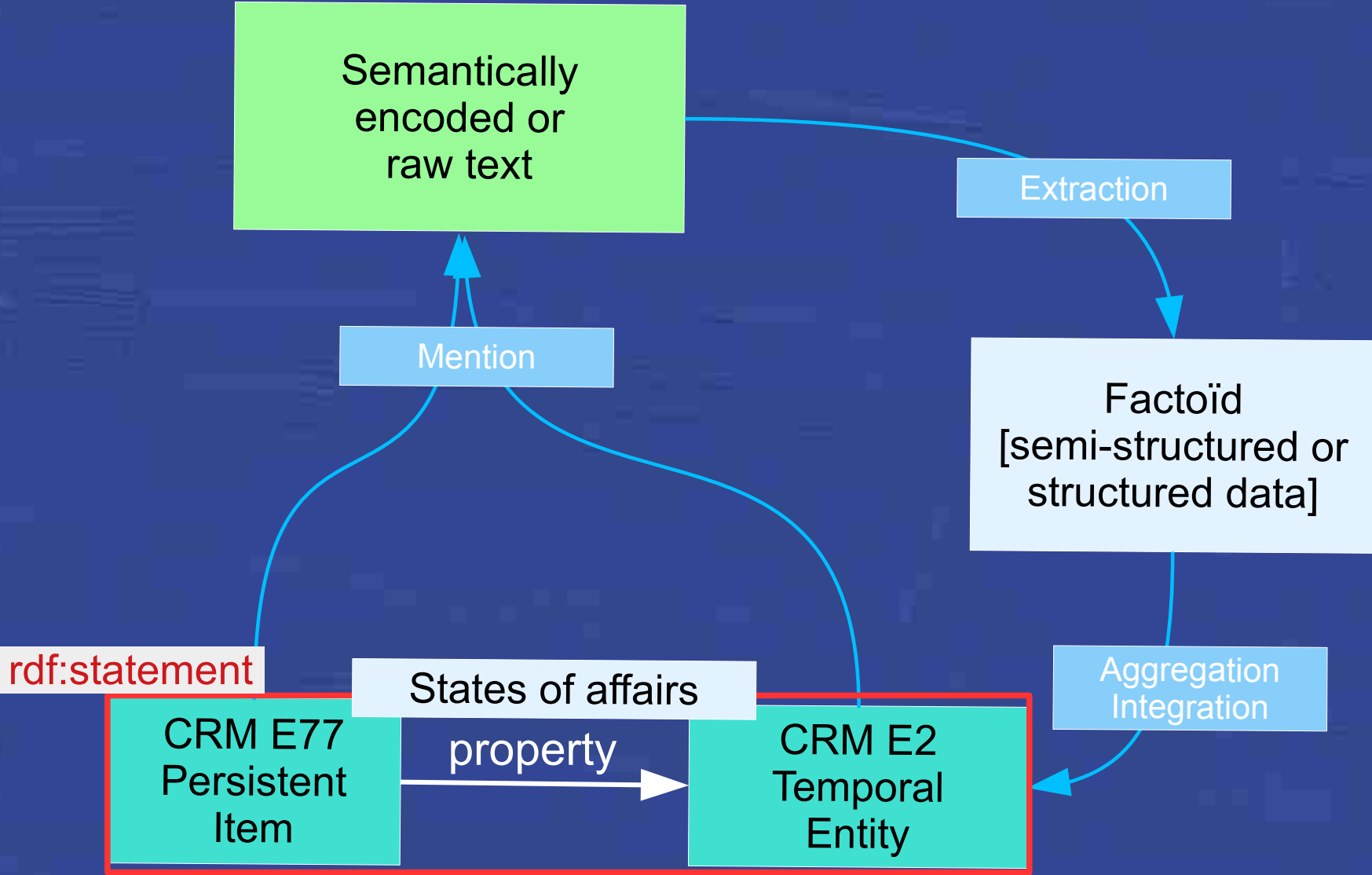


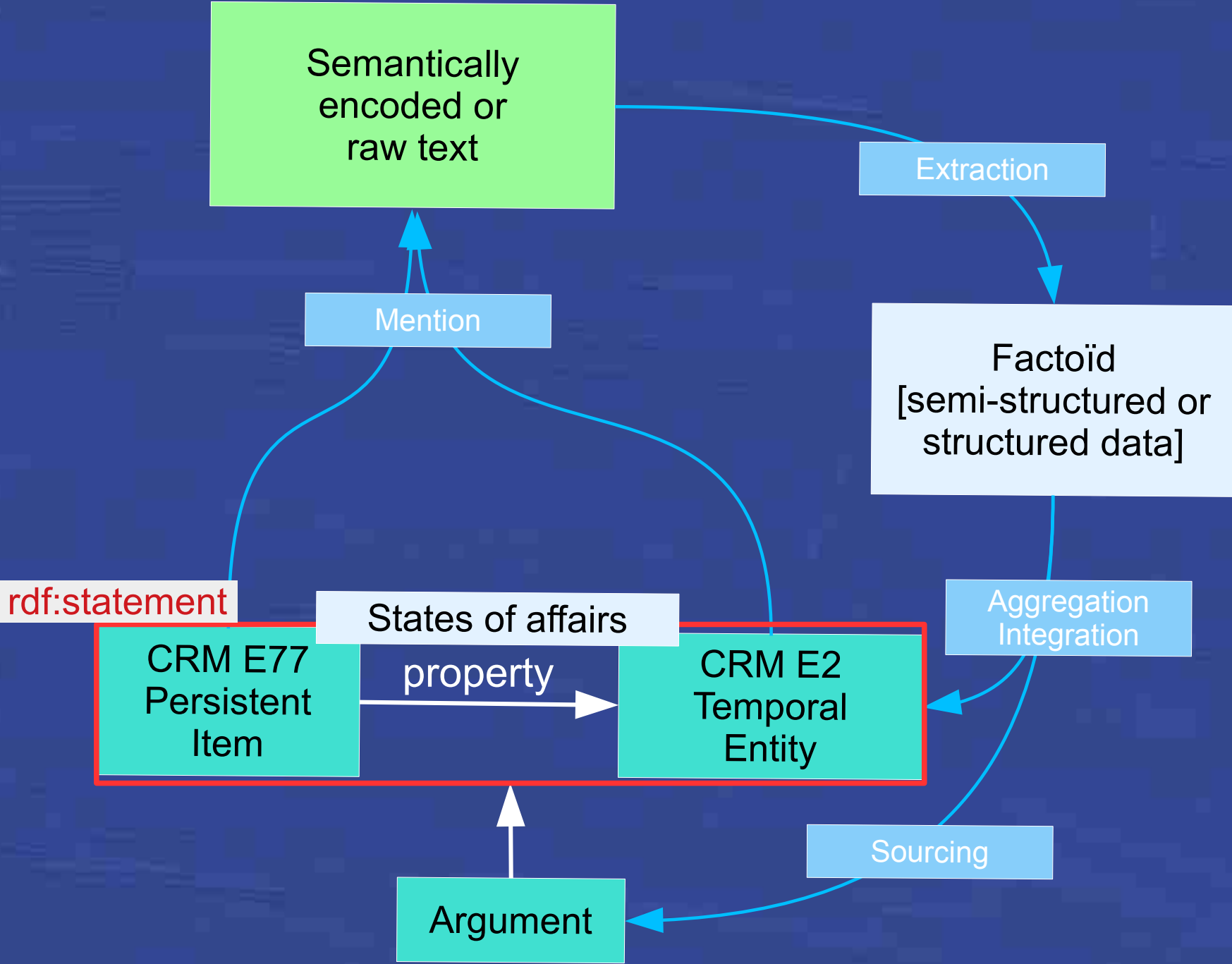
« Quantifiers for properties are provided for the purpose of semantic clarification only, and should **not** be treated as implementation recommendations.

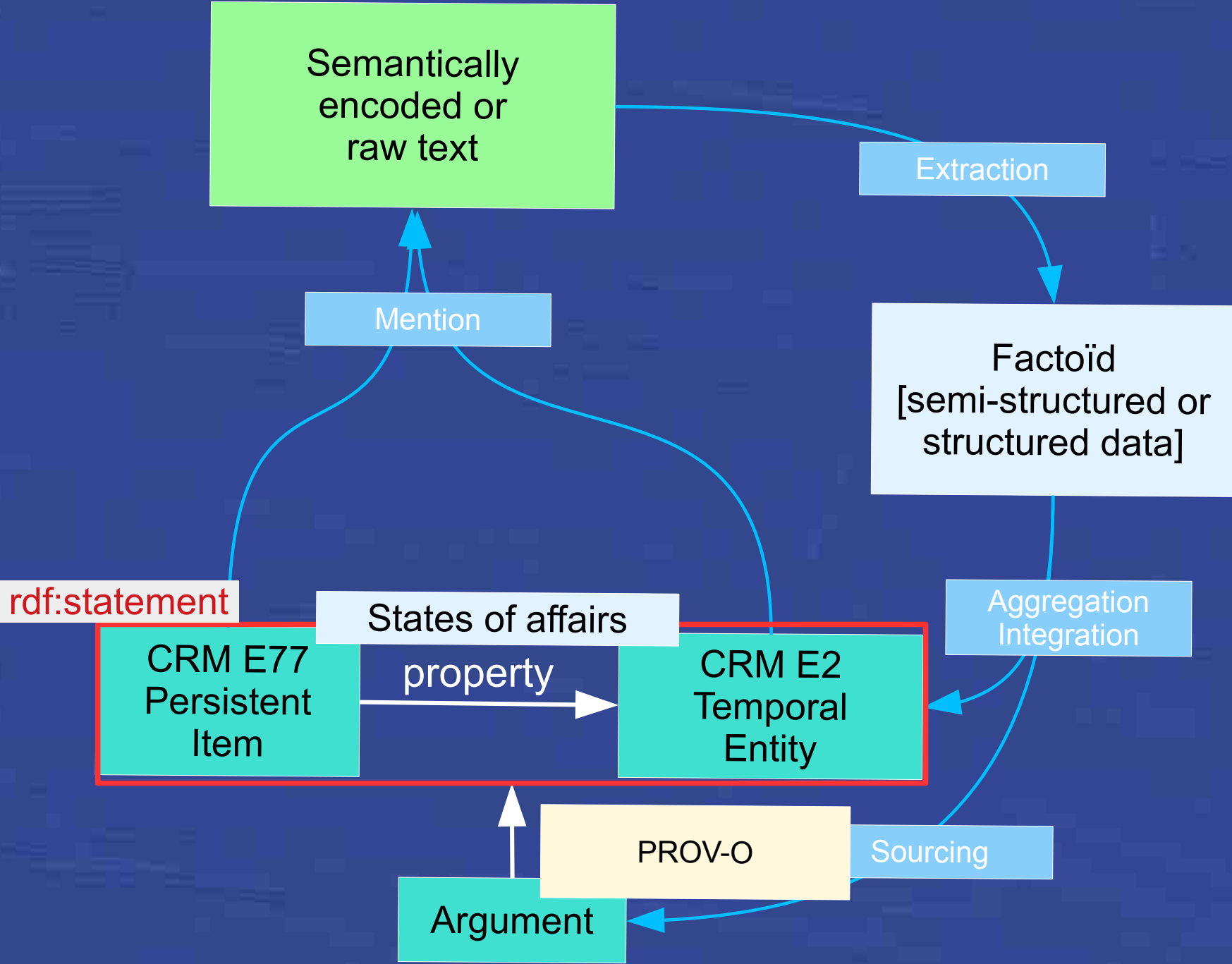
The CRM **has been designed to accommodate alternative opinions** and **incomplete information**, and therefore **all** properties should be implemented as optional and repeatable for their domain and range (“many to many (0,n:0,n)”).

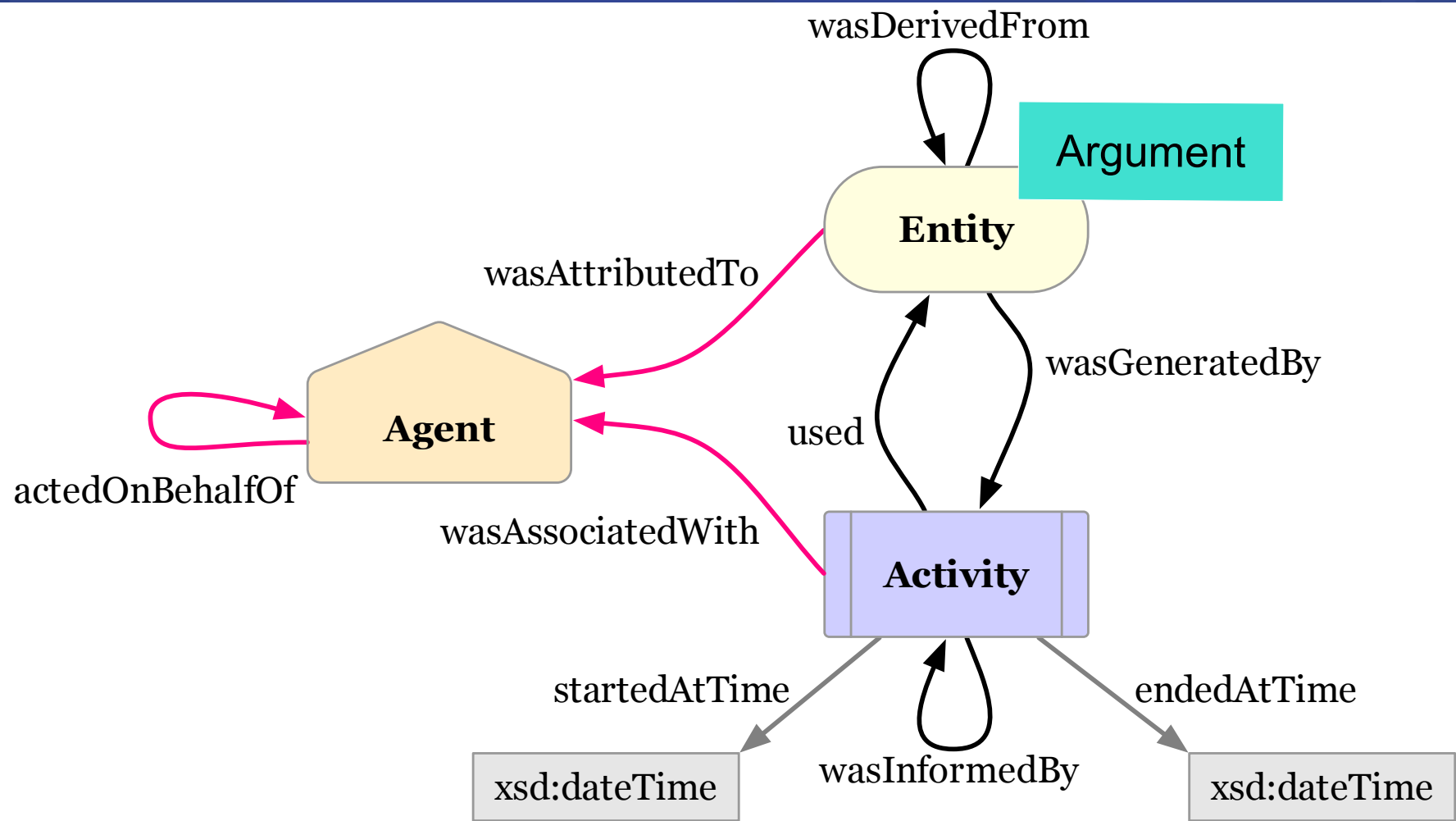
Therefore the term “cardinality constraints” is avoided here, as it typically pertains to implementations. » (CRM 6.2)





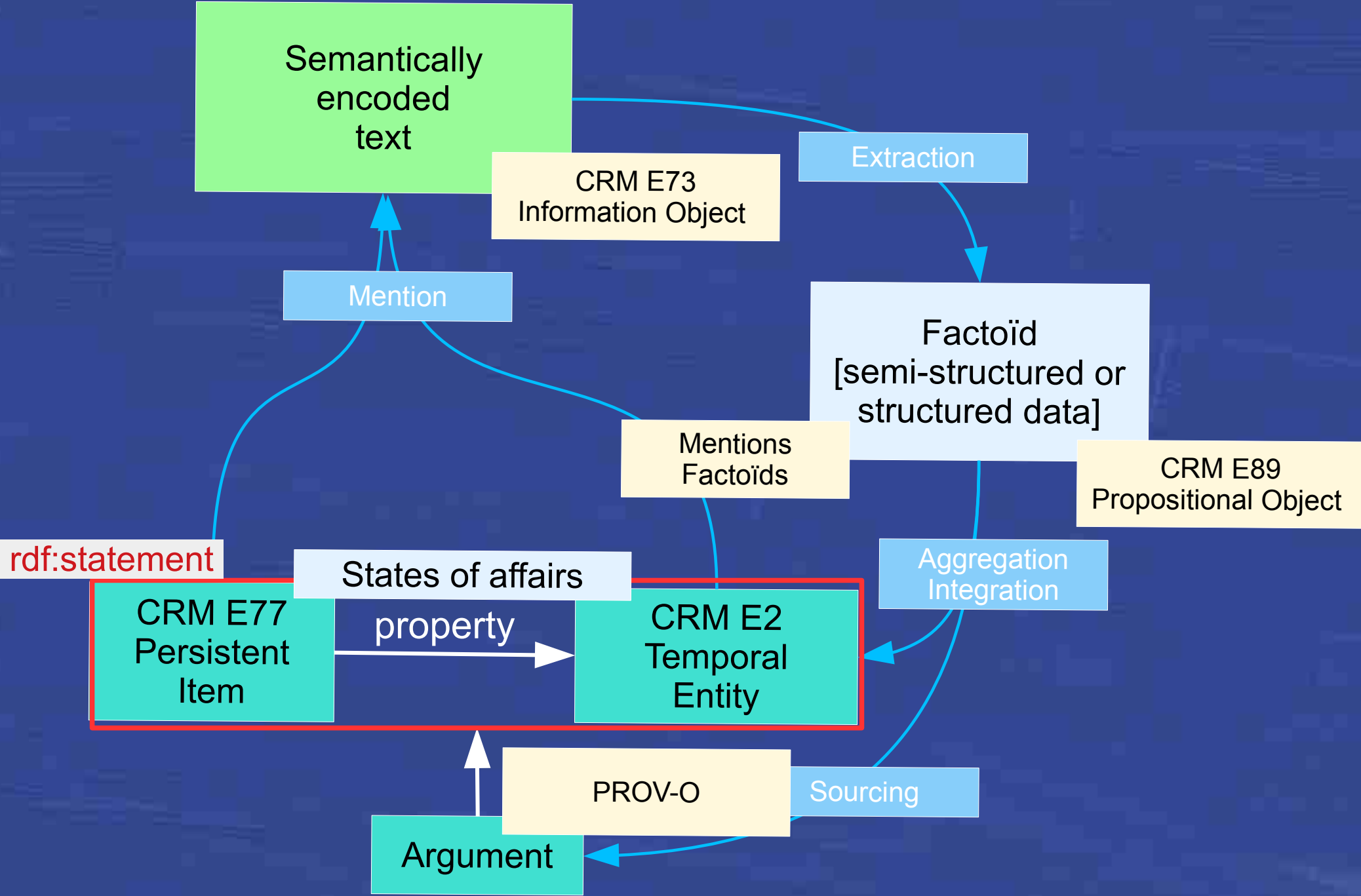






PROV-O

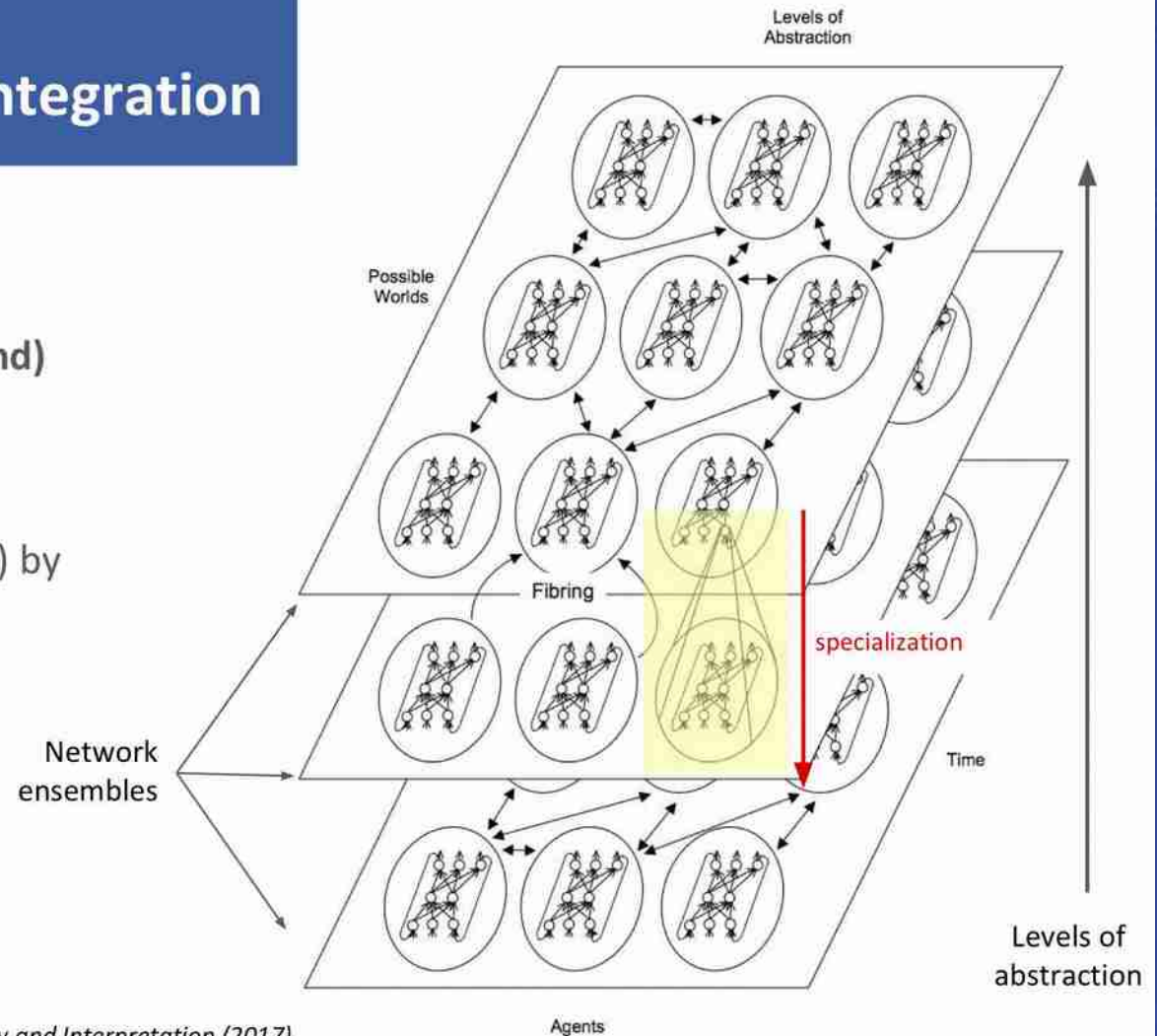
Provenance is information about entities, activities, and people involved in **producing a piece of data or thing**, which can be used to form assessments about its quality, reliability or trustworthiness.

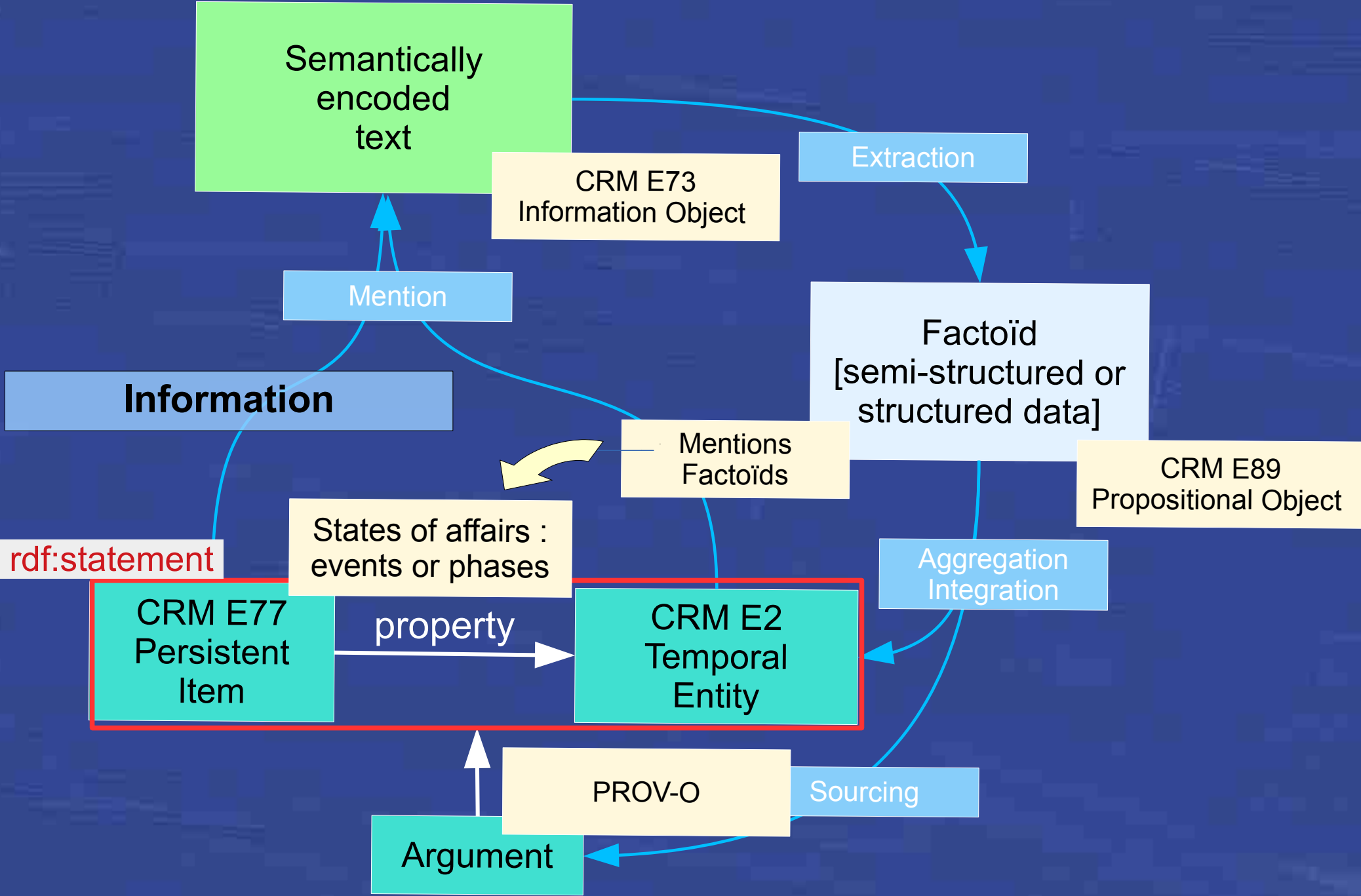


Towards Neuro-Symbolic Integration

Neuro-Symbolic Systems

1. Translation of **symbolic (background) knowledge** into the network
2. **Learning of additional knowledge** from examples (and generalisation) by the network
3. **Executing the network** (i.e. reasoning), and
4. **Symbolic knowledge extraction** from the network.





dataforhistory.org

develop methods and build infrastructure for
digital geo-historical research

FAIR

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 **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.

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*.