

Exposer des données archéologiques dans le web sémantique selon les *FAIR principes*

Xavier Rodier & Olivier Marlet

Consortium MASA
Maison de Sciences de l'Homme Val de Loire
Laboratoire Archéologie et Territoires UMR 7324 CITERES



Le consortium MASA

« Mémoire des archéologues et des sites archéologiques »



- la Maison des Sciences de l'Homme du Val de Loire
- la Maison de la Recherche en Sciences Humaines de Caen
- la Maison de l'Orient et de la Méditerranée de Lyon
- la Maison Méditerranéenne des Sciences de l'Homme, d'Aix-en-Provence
- la Maison Archéologie & Ethnologie de Nanterre
- le Musée d'archéologie nationale de Saint Germain-en-Laye
- le Groupement de Services Frantiq
- l'INRAP (depuis le 1er janvier 2017)

Les données archéologiques

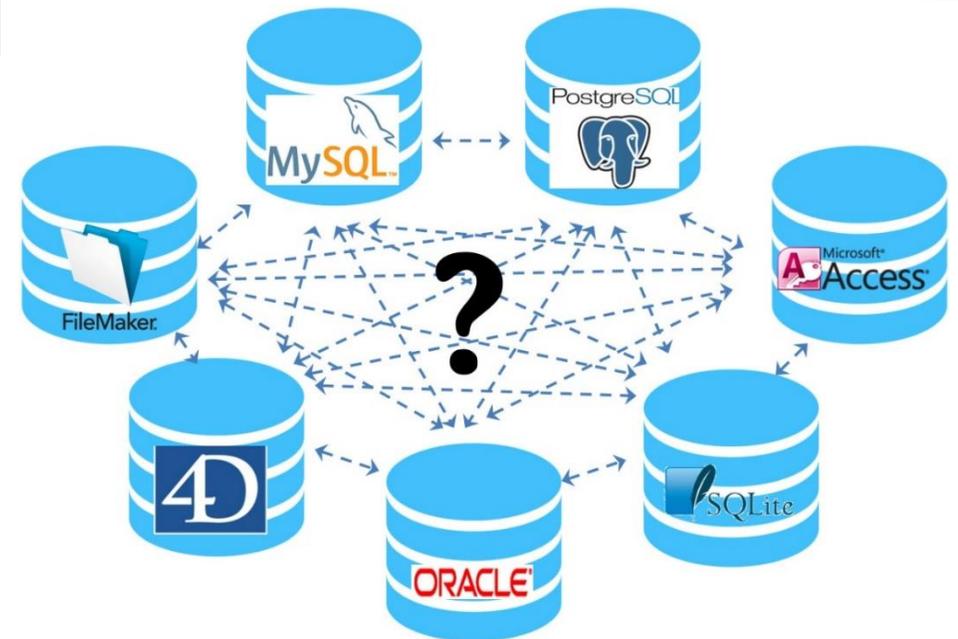
Accumulation des archives

Volonté de structurer et de publier ces données

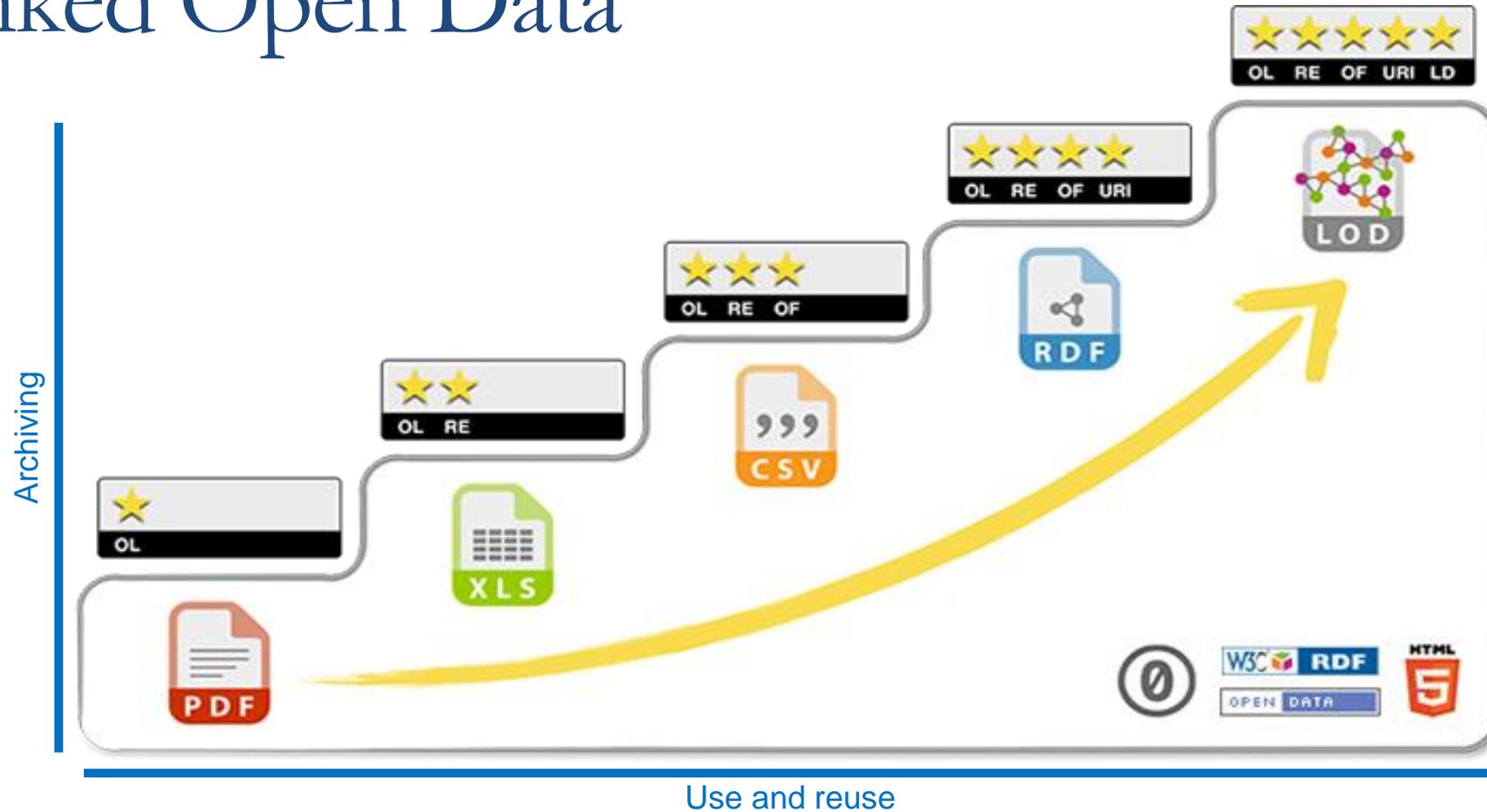
Multiplication des bases de données et hétérogénéité des formats

Nécessité de mettre en place des bonnes pratiques :

- métadonnées normalisées et archivage pérenne
- mise en ligne des données
- alignement sur des vocabulaires normalisés
- interopérabilité des données
- alimentation des plateformes



Les données archéologiques dans le Linked Open Data

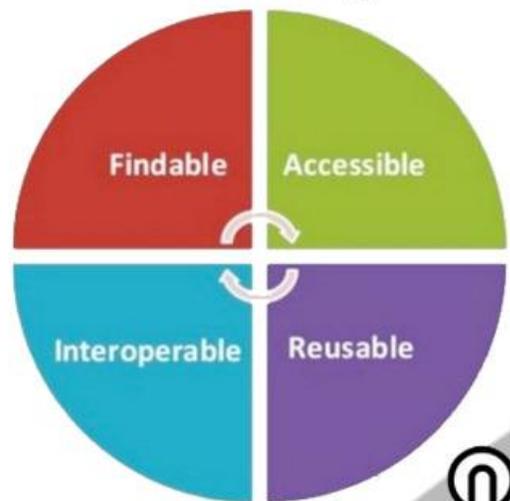


Vocabulary alignment
Publications

Interoperability
Ontology

The FAIR Data Principles

Make your data :



Findable

- ▶ Descriptive metadata
- ▶ Persistent identifiers



Accessible

- ▶ Standardized protocol
- ▶ Metadata must survive over the data's lifespan
- ▶ Access status



Interoperable

- ▶ XML standards
- ▶ Standardized vocabularies
- ▶ References to other data



Reusable

- ▶ Rights and licence models
- ▶ Permitted and non-permitted use
- ▶ Domain-relevant community standards

FAIR
DATA



Métadonnées et archivage



- Identification avec numéro ARK

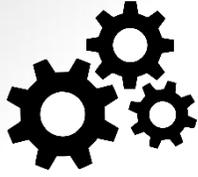
The screenshot shows the ARKéo website. The left column, titled 'Présentation générale du projet', describes the service's purpose and provides a URL: <http://ark.mom.fr/ark/76009ca657800>. The right column, titled 'Démonstration', shows a form for creating an ARK identifier with fields for URL, Title, Origin, and Date, and a 'Créer l'identifiant ARK' button.

- Mise en place de métadonnées en EAD

```
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE ead PUBLIC "-//Society of American Archivists/DTD ead.dtd (Encoded Archival Description (EAD) Version 2002)//EN" "ead.dtd">
<ead>
  <!-- Métadonnées de l'ag-arsol -->
  <eadheader langencoding="iso639-2b" scriptencoding="iso15924" dateencoding="iso8601" countryencoding="iso3166-1" repositoryencoding="iso15111">
    <!-- Id unique du fichier -->
    <eadid countrycode="FR" identifier="372612218-LAT37_arsol" mainagencycode="FR-372612218">FR-372612218-LAT37_arsol</eadid>
    <!-- description du fichier -->
    <filedesc>
      <titlestmt>
        <titleproper>Archives de fouilles du Laboratoire Archéologie et Territoires depuis 1969</titleproper>
        <subtitlen>Répertoire numérique détaillé</subtitlen>
        <author>la direction de Olivier Marlet, ingénieur CHRS</author>
      </titlestmt>
      <publicationstat>
        <publisher>Université de Tours, IHR 7324 CITERES</publisher>
        <date>2018</date>
        <address>
          <addressline>IHR 7324 - CITERES, MH Valles et Territoires</addressline>
          <addressline>BP 60449</addressline>
          <addressline>37204 TOURS cedex 03</addressline>
        </address>
        <date type="version initiale" normal="2018">2018</date>
      </publicationstat>
    </filedesc>
    <!-- description du profil -->
    <profiledesc>
      <creation>Encodé conformément à la DTD EAD 2002 réalisé par Olivier Marlet et connecté à Arsol (pour être généré dynamiquement dans sa
      <language>Instrument de recherche rédigé en <language langcode="fr">français</language></language>
    </profiledesc>
    <!-- description des révisions -->
    <revisiondesc>
      <change>
        <date normal="2018-09">septembre 2018</date>
        <item>Mise en ligne de la version initiale du fichier EAD pour Arsol</item>
      </change>
    </revisiondesc>
  </eadheader>
  <!-- titre, log, sous-titre, note explicative, avant propos -->
  <frontmatter>
    <titlepage>
      <publisher>
        <extptr actuate="onload" show="embed" href="logo_CITERES_LAT.jpg" />
      </publisher>
    </titlepage>
  </frontmatter>
  <!-- description archivistique -->
  <archdesc level="collection">
    <did>
      <origination>
        <corpname role="producteur">Laboratoire Archéologie et Territoires (IHR 7324 CITERES), Université de Tours (37)</corpname>
      </origination>
      <repository>
        <corpname>Laboratoire Archéologie et Territoires</corpname>
        <address>
          <addressline>40 rue James Watt</addressline>
          <addressline>41100 Amboise 2</addressline>
          <addressline>37200 Tours</addressline>
        </address>
      </repository>
    </did>
  </archdesc>
</ead>
```

- Référencement avec Nakala (Huma-Num)





Vocabulaire normalisé : les PACTOLS

Interoperable

The screenshot displays the PACTOLS interface for a normalized vocabulary. At the top, there is a user account section for 'miled.rousset', a thesaurus selection dropdown set to 'PACTOLS (TH_1)', and a language dropdown set to 'français'. Navigation tabs include 'Thésaurus', 'Candidats', 'Configurations', 'Edition', and 'Statistiques'. The main content area is divided into several panels:

- Terme(s) générique(s):** Lists 'bijou (BT)' and 'fixation de vêtement (BT)'.
- Terme(s) associé(s):** Lists 'broche' with a 'Caché:' status.
- Terme(s) synonyme(s):** Lists 'broche' with a 'Caché:' status.
- Notes:** Contains a definition: 'definition: Agrafe ou broche articulée par un ressort ou une charnière et dont la pointe (ardillon) se fixait à l'extrémité d'un arc de forme variée, servant à attacher les vêtements flottants et non cousus. (Musée d'Argentomagus) (fr)'. It also includes a 'Total de la branche' button.
- Terme(s) spécifique(s):** A section for specific terms.
- Alignement:** Shows 'exactMatch : Wikidata - Fibule'.
- Coordonnées Gps:** A section for geographical coordinates.
- Traduction:** Lists translations for 'Fibule' in German (de), Arabic (ar), English (en), Italian (it), Dutch (nl), and Spanish (es).
- Facette:** A section for facets.
- Image:** Displays an image of a fibula with a 'Wikimedia' watermark.

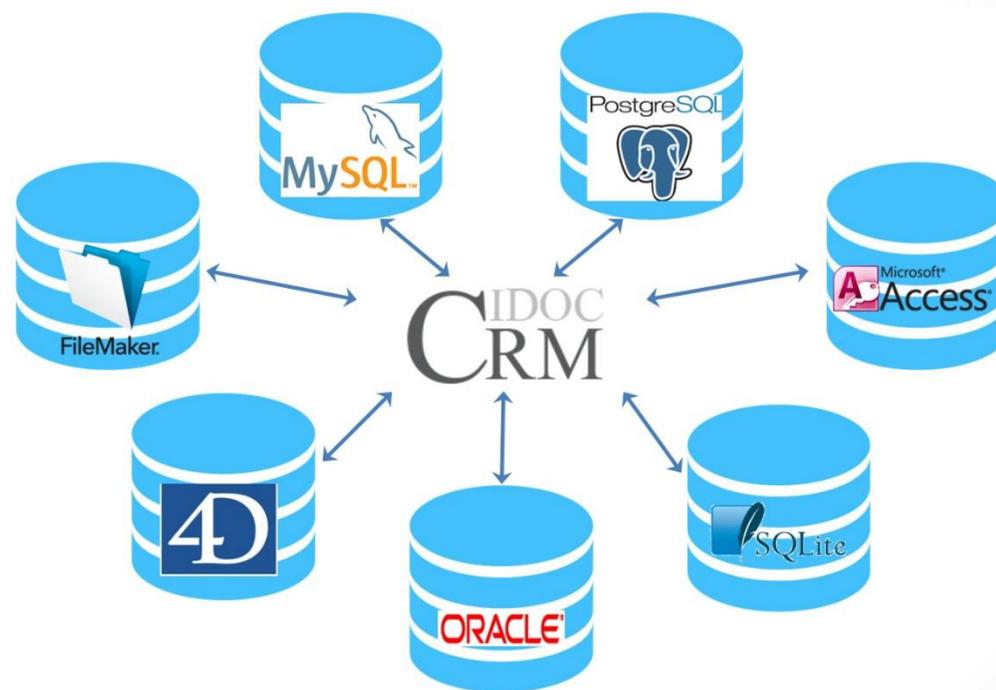
On the left side, there is a hierarchical tree structure under the 'index' search bar, with 'fibule' selected. The tree includes categories like 'chaussure', 'couvre-chef', 'vêtement', 'mobilier', 'parure', and 'bijou', with 'fibule' being a sub-entry under 'bijou'.

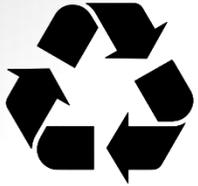


Interopérabilité des données

Interopérable

- Réaliser les appariements de chaque source de données avec le CIDOC-CRM et ses extensions.
- Mettre en place une application unique pour interroger ces sources de données hétérogènes.





Réutilisation des données

Reusable

The screenshot shows the ARIADNE website interface. At the top, there are navigation tabs for "Catalog", "Services", and "About". The main header features the ARIADNE logo, which consists of a stylized maze icon followed by the word "ARIADNE" in large, bold, red letters. Below the logo is a search bar with a dropdown menu set to "All fields" and a search button. The main content area is divided into three sections: "Welcome", "Browse the Catalog", and "What".

Welcome
Explore the digital resources and services that ARIADNE has brought together from across Europe for archaeological research, learning and teaching.

Browse the Catalog

Where
A map of Europe showing the geographical distribution of resources.

When
A line graph showing the number of resources over time, with a significant peak around 1900.

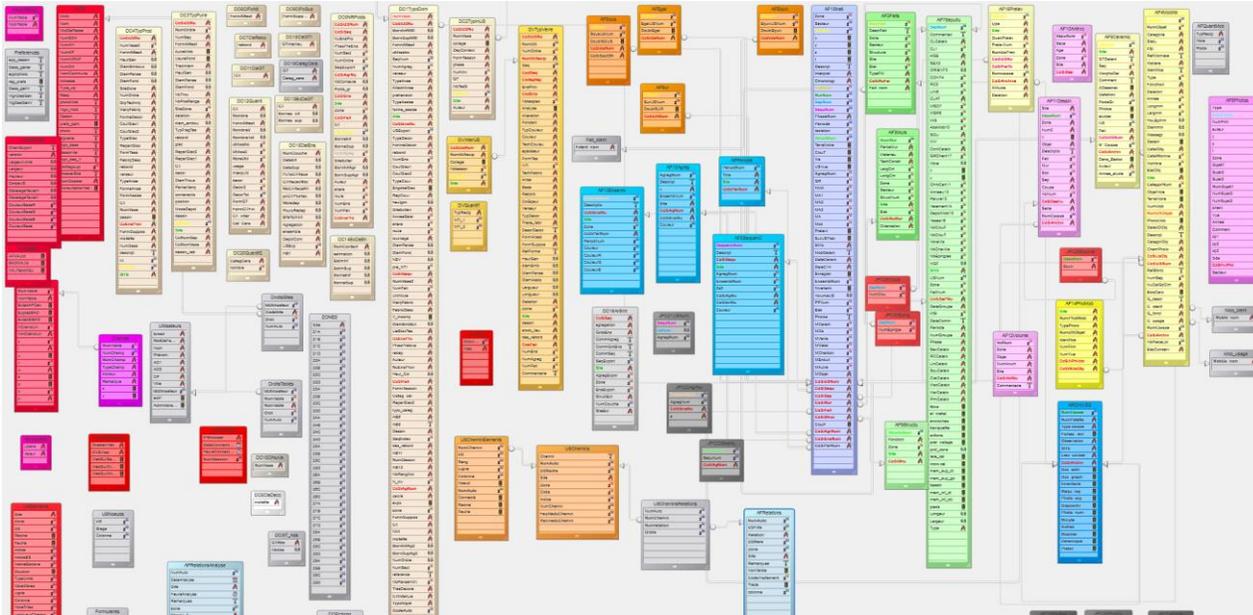
What
A word cloud of archaeological terms, with "houses" being the most prominent. Other terms include "lime kilns", "cist graves", "forts", "ditches", "cairns", "farms", "graves", "kilns", "cargo", and "vessels".

Evaluation des jeux de données (CITERES-LAT)

Data set	Format	5 stars Linked Open Data	FAIR				Start date	Web
ArSol	4D	*****	F indable ID Metadata	A ccessible HTTP	I nteroperable Standards Vocabularies Linked data	R eusable Licence Community	1990	http://arsol.univ-tours.fr/
Atlas Archéologique de Touraine	postgreSQL	*****	F indable ID Metadata	A ccessible HTTP	I nteroperable Standards Vocabularies Linked data	R eusable Licence Community	2006	http://a2t.univ-tours.fr/
Iceramm	mySQL	*****	F indable ID Metadata	A ccessible HTTP	I nteroperable Standards Vocabularies Linked data	R eusable Licence Community	2007	http://iceramm.univ-tours.fr/
Toposources	mySQL	*****	F indable ID Metadata	A ccessible HTTP	I nteroperable Standards Vocabularies Linked data	R eusable Licence Community	2012	http://toposources.univ-tours.fr/
Outillage de la Gaule antique	XML	*****	F indable ID Metadata	A ccessible HTTP	I nteroperable Standards Vocabularies Linked data	R eusable Licence Community	2016	http://outagr.huma-num.fr
Rigny	XML	*****	F indable ID Metadata	A ccessible HTTP	I nteroperable Standards Vocabularies Linked data	R eusable Licence Community	2016	(work in progress)
AERBA	XML	*****	F indable ID Metadata	A ccessible HTTP	I nteroperable Standards Vocabularies Linked data	R eusable Licence Community	2017	(work in progress)

ArSol : « Archives du Sol »

- Enregistrement archéologique de terrain 1990
- 4D (logiciel propriétaire)



- Données en ligne : <http://arsol.univ-tours.fr>
- Utilisation de ce cas particulier pour évaluer l'interopérabilité au sein de MASA

ArSol
Archives du Sol - © 1996 - 2014 UMR 7324 CITERES - Laboratoire Archéologie et Territoires

RECHERCHE CONSULTATION STRATIGRAPHIQUE

+ Recherche
+ Résultats [Retour à la Fiche du Départ](#)

Tours site 17 - Marmoutier ✕
sol carrelé
Fait 14

Nom : sol carrelé Secteur : 1 Structure : Zone : 1

Description :
Dallage de carreaux de terre cuite

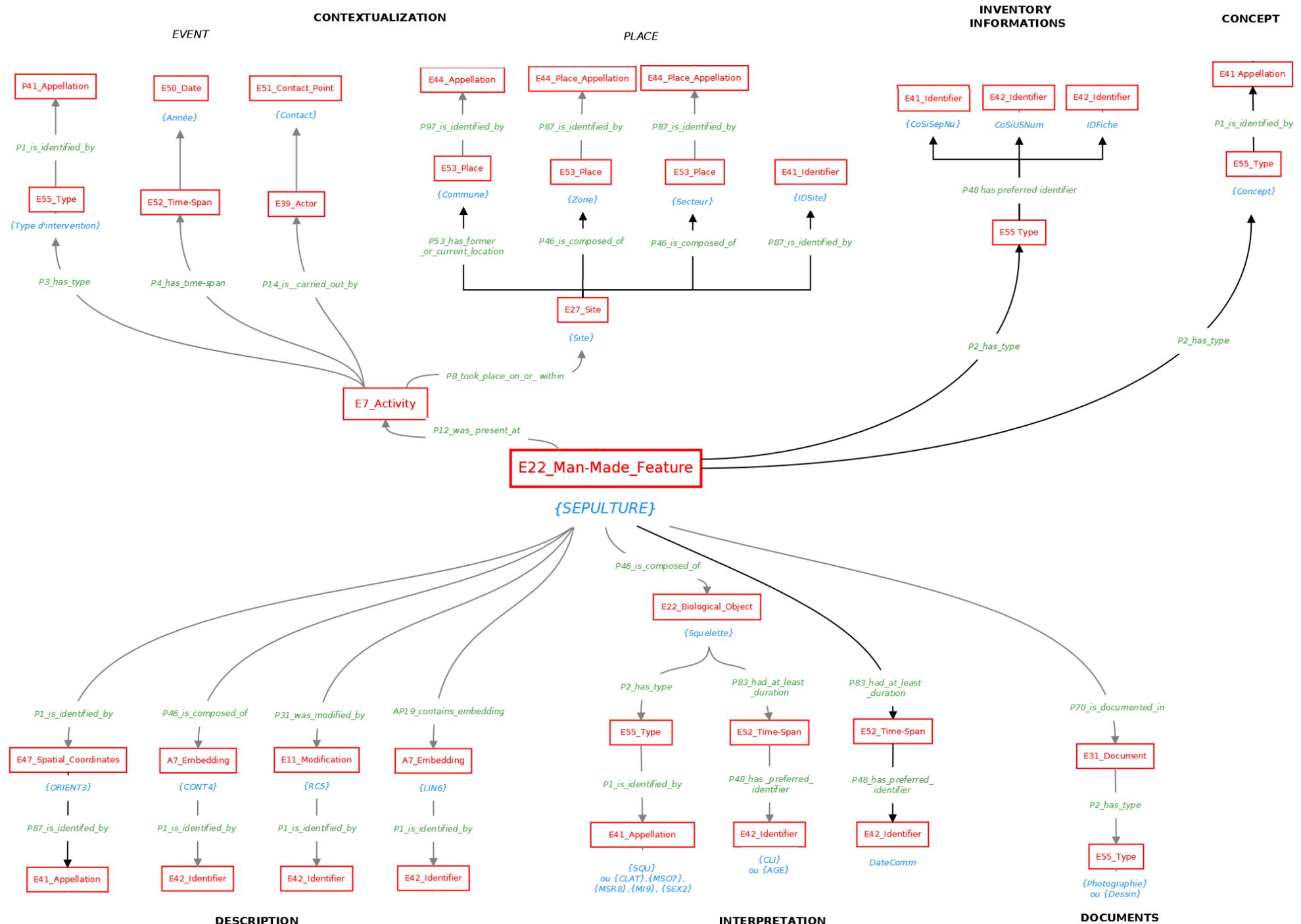
Le Fait contient 6 US.

Numéro	Description
10172	Carreaux de terre cuite. Le carrelage occupait une surface limitée à l'intérieur du bâtiment 4 dans un espace délimité.
10239	Niveau très friable de mortier de chaux de couleur blanche comprenant des gravillons fin. Niveau de préparation...
10240	Mortier de chaux de couleur blanche contenant des inclusions très fines. Réfection du sol de carreaux de terre cuite...
10252	terre limoneuse de couleur brune avec des inclusions très fines de tuffeau, très peu de chaux et de sable servant de...
10258	Mortier de chaux avec des inclusions de galets (1cm), des fragments de tuffeau et de TCA. Il s'agit d'une réfection du...
10291	niveau de terre limono-argileuse compact de couleur brune, peu épais.

[Aide à la datation par la céramique](#)
[Affichage du fichier RDF](#)

36 PHOTOS [1](#) [2](#)

Appariement des champs de la table "Sépulture"



Utiliser Ontop

- Protégé, logiciel d'ontology
- CIDOC au format OWL
- Charger les extensions CRM
- Explorer l'arborescence
- Connecter la base de données
- Ecrire les mappings
- Générer un fichier RDF

The screenshot displays the Ontop web interface for the CIDOC CRM ontology. The browser address bar shows the URL `http://www.cidoc-crm.org/cidoc-crm/`. The interface includes a menu bar (File, Edit, View, Reasoner, Tools, Refactor, Window, ontop, Help) and a search bar for entities. The main content area is divided into several panels:

- Class hierarchy:** Shows a tree structure starting with 'Thing', including 'Entité CRM', 'I5_Inference_Making', and 'S4_Observasion'.
- Object property hierarchy:** Lists various object properties such as 'a actuellement pour lo', 'a affecté un attribut à', 'a attribué', 'a changé de détenteur', 'a changé de mains du f', 'a des caractéristiques', 'a duré au moins', 'a duré au plus', 'a déplacé vers', 'a employé', 'a employé un objet du', 'a eu lieu dans', 'a eu lieu sur ou dans', 'a fait passer le droit de', and 'a influencé'.
- Mapping editor:** Displays a list of mappings for the 'testarchoe' datasource. Each mapping is shown in a separate pane with its SPARQL query. For example, 'MOB-G_usage-01' has the query: `arsol:imobi{CoSiUSObj} :P2_has_type arsol:imobi{CoSiUSObj}#usage .
SELECT CoSiUSObj FROM mobilier`

At the bottom, the interface shows 'Mapping count: 46', a search bar, and an 'Enable filter' checkbox. A footer note states: 'To use the reasoner click Reasoner [] Start reasoner [x] Show Inferences'.

Requête SPARQL

- Interrogation dynamique d'une base de données en ligne, à partir des entités du CIDOC
- Qui serait prêt à utiliser une interface aussi peu conviviale ?

The screenshot shows the 'ontop query editor' interface. At the top, there's a browser address bar with 'current (http://erlangen-crm.org/current/)' and a search box. Below are navigation tabs: OWLViz, DL Query, ontop SPARQL (selected), OntoGraf, ontop Mappings, Ontology Differences, and SPARQL Query. The main area is titled 'ontop query editor:' and contains a 'Query Editor' window with the following SPARQL query:

```
PREFIX : <http://erlangen-crm.org/current/>

SELECT ?artefact ?use
WHERE {
  ?artefact a :E22_Man-Made_Object; :P101_had_as_general_use ?usage .
  ?usage a :E55_Type; :P1_is_identified_by ?usageid .
  ?usageid a :E41_Appellation; :P3_has_note ?use .
  FILTER (?use != "")
}
```

Below the query editor, it shows 'Execution time: 1.442 sec - Number of rows retrieved: 3 850' and buttons for 'Attach Prefixes', 'Execute', and 'Save Changes'. The results are displayed in a table with two columns: 'artefact' and 'use'.

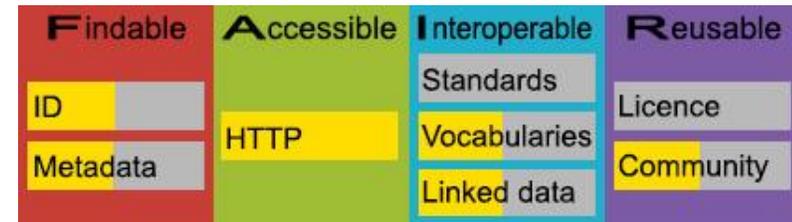
artefact	use
<http://arsol.univ-tours.fr/4DACTION/WFICHEWEB/imobiAA2574.1>	"accessoire vestimentaire"
<http://arsol.univ-tours.fr/4DACTION/WFICHEWEB/imobiAA5161.1>	"parure"
<http://arsol.univ-tours.fr/4DACTION/WFICHEWEB/imobiAA3067.7>	"accessoire vestimentaire"
<http://arsol.univ-tours.fr/4DACTION/WFICHEWEB/imobiAA1258.3>	"mobilier"

At the bottom, there's a 'Hint: Try to continue scrolling down the table to retrieve more results.' and an 'Export to CSV...' button.

Progression pour ArSol



- En ligne
- Données structurées
- ~~Format libre~~
- ~~standard W3C~~
- Interopérable



- Faut-il migrer les données dans un format ouvert et standard, alors que les données sont interopérables ?

Atlas des établissements ruraux de la Beauce antique

Inventaire des sites ruraux de Beauce connus par prospections aériennes et pédestres

Atlas des établissements ruraux de la Beauce antique

Accueil Carte de l'atlas Liste des sites Classification sites Bibliographie Mentions légales

Cette carte interactive vous donne accès à la fiche descriptive propre à chacun des sites, chaque point correspond à un site archéologique en passant le curseur sur un point vous obtiendrez des informations sur celui-ci ainsi qu'un lien vers la fiche détaillée. Plusieurs options vous sont proposées pour l'exploitation de cette carte, notamment le fait de pouvoir rechercher un site par le biais du champ de recherche **Q**, vous pouvez également mesurer des distances et des aires directement sur la carte avec l'outil **M** et enfin faire apparaître les différentes couches de la carte avec l'outil **C**. Enfin, d'affiner votre recherche vous pouvez limiter le nombre de sites en fonction de critères grâce aux éléments ci-dessous. Si vous faites plusieurs recherches les unes à la suite des autres, chacune d'entre elle sera faite sur l'ensemble des sites qui répondaient à vos précédents critères de recherche. Le bouton **R** vous permet de réinitialiser la carte.

Rang **>>** **<<**

Q Période exclusive : **Q** Période inclusive : **Q**

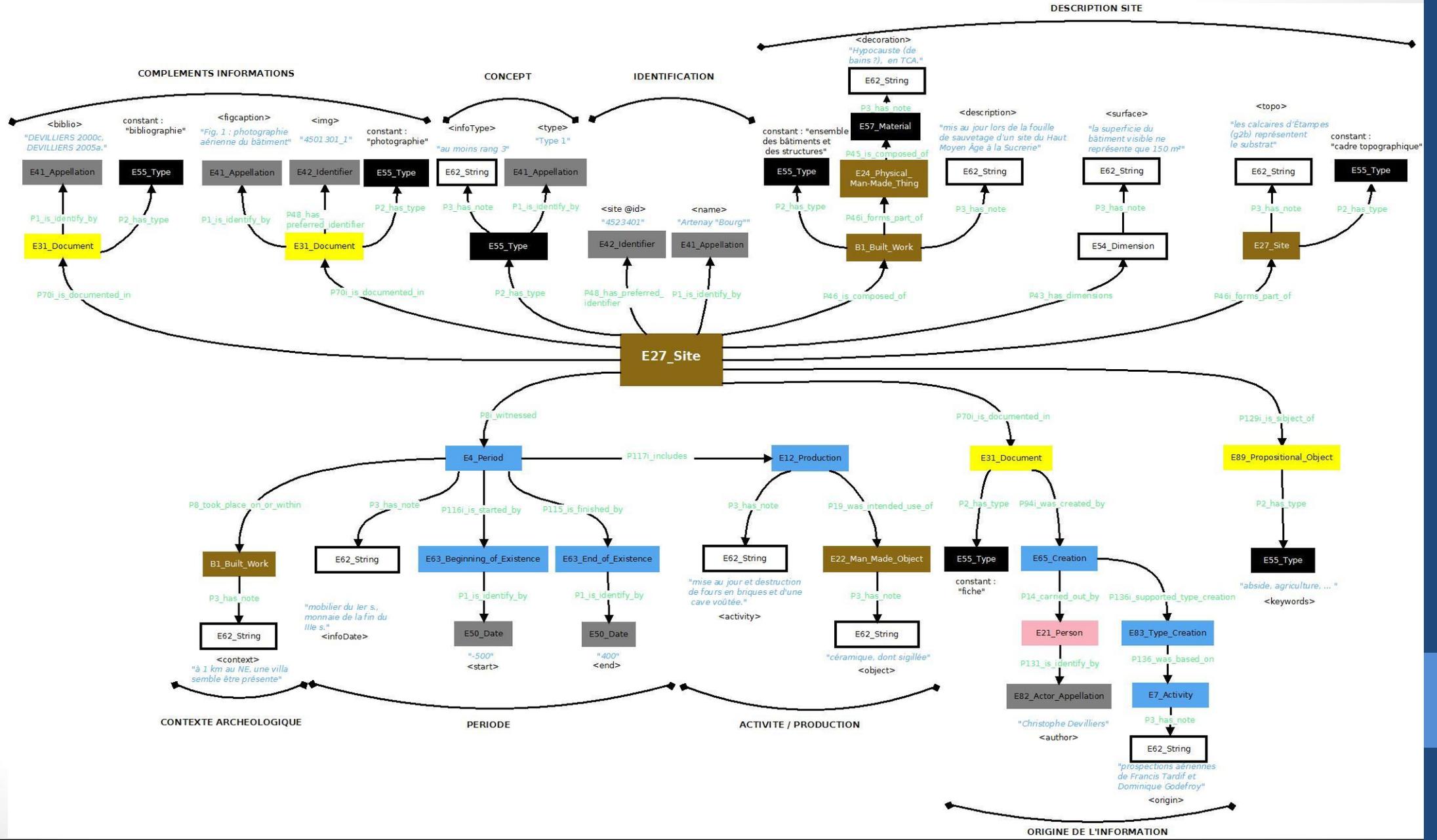
Fermer l'outil de recherche

Choix mot-clé

- Choix mot-clé
- abside
- acqueduc
- adobe
- agriculture
- alimentation
- antéfixe
- appareil
- architecture
- artisanat
- bain
- base
- bassin (structure)
- bâtiment
- bauge
- bergerie
- bloc
- bois
- brique
- calcaire

Map showing the Beauce region with numerous blue circular markers indicating the locations of rural establishments. The map includes a scale bar (0 to 1000m) and a 'Rang' dropdown menu.

Appariement des données XML d'AERBA avec le CIDOC



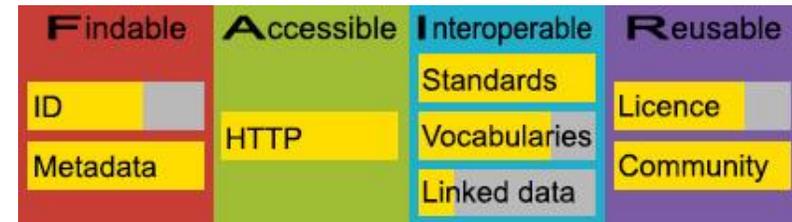
3M

#	SOURCE	TARGET	CONSTANT EXPRESSION	IF RULE	COMMENTS
1	D <input type="checkbox"/> ../site	<input type="checkbox"/> E27_Site			
1.1	P <input type="checkbox"/> @id	<input type="checkbox"/> P48_has_preferred_identifier			
	R <input type="checkbox"/> @id	<input type="checkbox"/> E42_Identifier			
1.2	P <input type="checkbox"/> ../name	<input type="checkbox"/> P1_is_identified_by			
	R <input type="checkbox"/> ../name	<input type="checkbox"/> E41_Appellation			
1.3	P <input type="checkbox"/> ../author	<input type="checkbox"/> P70i_is_documented_in			
		<input checked="" type="checkbox"/> E31_Document <input checked="" type="checkbox"/> [bibliography]	[P2_has_type]	[E55_Type = "fiche"]	
		<input type="checkbox"/> P94i_was_created_by			
R <input type="checkbox"/> ../author	<input checked="" type="checkbox"/> E65_Creation <input checked="" type="checkbox"/> [creation_bibliography]				
	<input type="checkbox"/> P14_carried_out_by				
	<input type="checkbox"/> E21_Person				
		<input type="checkbox"/> P131_is_identified_by			
		<input type="checkbox"/> E82_Actor_Appellation			
1.4	P <input type="checkbox"/> ../topo	<input type="checkbox"/> P46i_forms_part_of			
		<input checked="" type="checkbox"/> E27_Site	[P2_has_type]	[E55_Type = "cadre topographique"]	
		<input type="checkbox"/> P3_has_note			
R <input type="checkbox"/> ../topo	<input type="checkbox"/> rdf-schema#Literal				
1.5	P <input type="checkbox"/> ../origin	<input type="checkbox"/> P70i_is_documented_in			
		<input checked="" type="checkbox"/> E31_Document <input checked="" type="checkbox"/> [bibliography]			
		<input type="checkbox"/> P94i_was_created_by			
		<input checked="" type="checkbox"/> E65_Creation <input checked="" type="checkbox"/> [creation_bibliography]			
		<input type="checkbox"/> P136i_supported_type_creation			
		<input type="checkbox"/> E83_Type_Creation			
R <input type="checkbox"/> ../origin	<input type="checkbox"/> P136_was_based_on				
	<input type="checkbox"/> E7_Activity				
	<input type="checkbox"/> P3_has_note				
	<input type="checkbox"/> rdf-schema#Literal				
1.6	P <input type="checkbox"/> ../type	<input type="checkbox"/> P2_has_type			
		<input checked="" type="checkbox"/> E55_Type <input checked="" type="checkbox"/> [type_site]			
		<input type="checkbox"/> P1_is_identified_by			
R <input type="checkbox"/> ../type	<input type="checkbox"/> E41_Appellation				
1.7	P <input type="checkbox"/> ../infoType	<input type="checkbox"/> P2_has_type			
		<input checked="" type="checkbox"/> E55_Type <input checked="" type="checkbox"/> [type_site]			
		<input type="checkbox"/> P3_has_note			
R <input type="checkbox"/> ../infoType	<input type="checkbox"/> rdf-schema#Literal				
1.8	P <input type="checkbox"/> ../context	<input type="checkbox"/> P8i_witnessed			
		<input checked="" type="checkbox"/> E4_Period <input checked="" type="checkbox"/> [site_period]			
		<input type="checkbox"/> P8_took_place_on_or_within			
		<input type="checkbox"/> B1_Built_Work			
R <input type="checkbox"/> ../context	<input type="checkbox"/> rdf-schema#Literal				

Progression pour AERBA



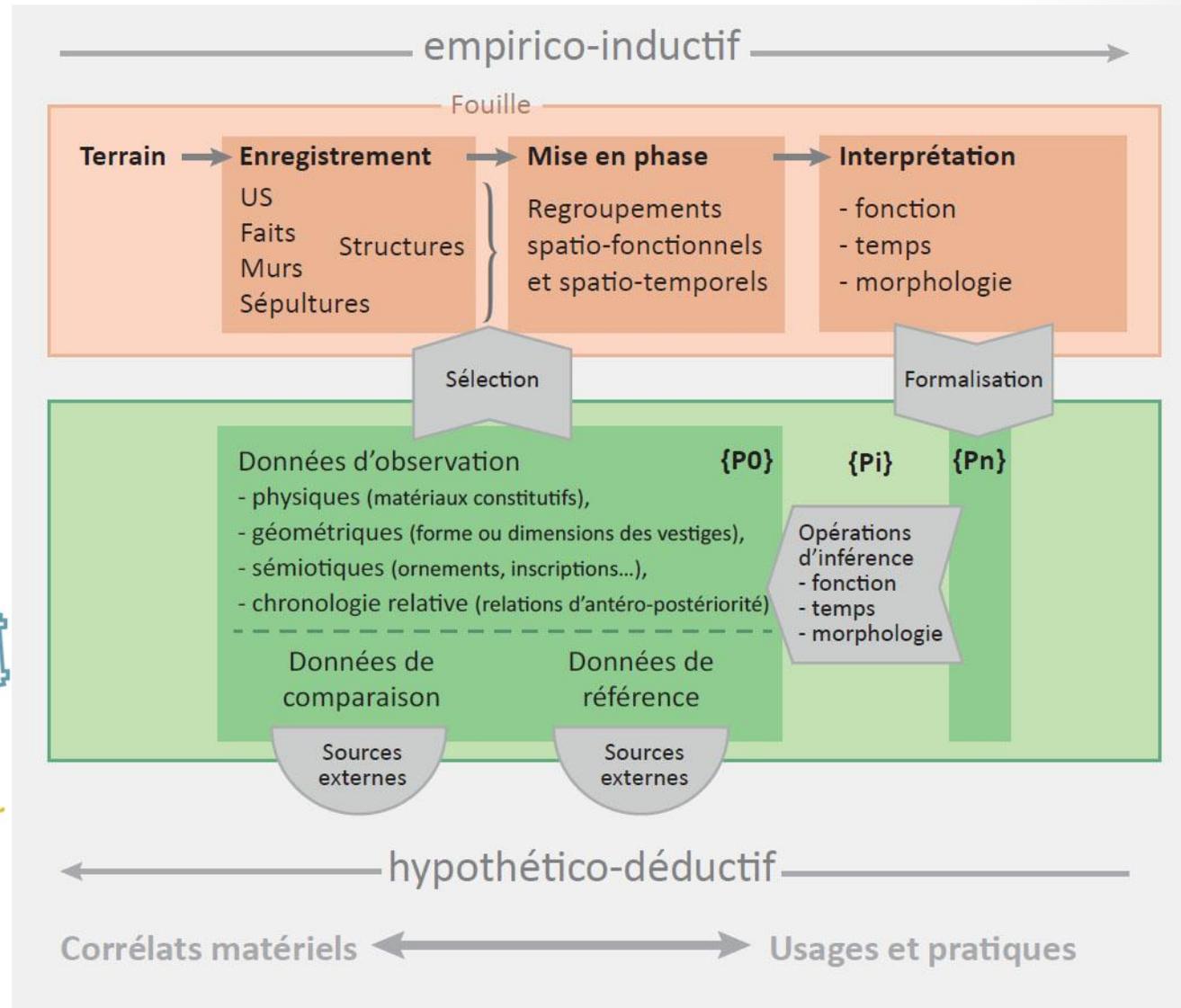
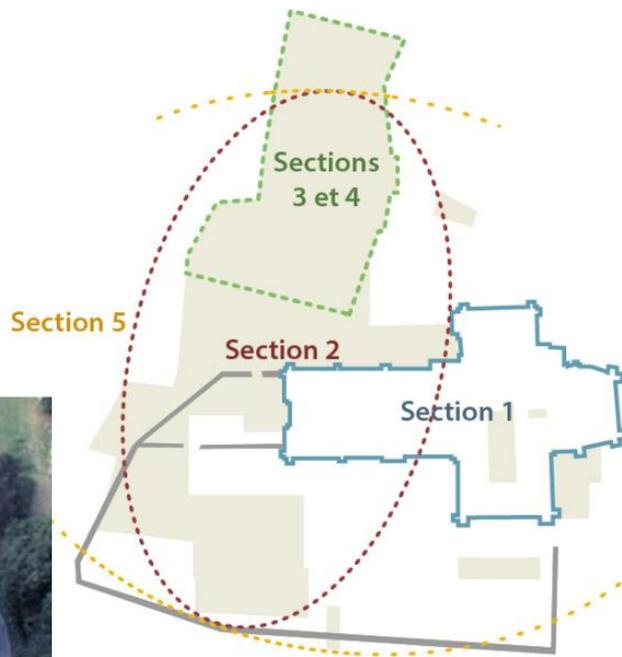
- En ligne
- Données structurées
- Format libre
- standard W3C
- Interopérable



- Le seul fait de choisir le format XML permet d'avancer vite vers l'interopérabilité.
- Les nouveaux projets demandent moins de travail que l'actualisation des anciens.

Publication logiciste des fouilles de Rigny

- formalisation de Jean-Claude Gardin
- structuration claire de la logique du raisonnement archéologique



SECTION 4 - LE PRESBYTÈRE DE RIGNY ET SES DÉPENDANCES (MILIEU 15^E-MILIEU 19^E S.)

[\[-\] Le bâtiment 11](#)

[\[+\] Etat 1, datation et fonction](#)

[\[+\] Etat 2, datation et fonction](#)

[\[+\] Etat 3, datation et fonction](#)

[\[-\] Le bâtiment 9](#)

[\[+\] Etat 1, datation et fonction](#)

[\[+\] Etat 2, datation et fonction](#)

[\[+\] Transformation du centre paroissial](#)

[\[+\] Le bâtiment 5](#)

[\[+\] Le bâtiment 8](#)

[Circuler dans les illustrations](#)

[Accéder au diagramme logicieste](#)

ETAT 3, DATATION ET FONCTION

P0_12. Dans l'Etat 3, le mur pignon ouest (M42) et le mur nord (M35) sont détruits : le bâtiment 11 devient plus long et plus étroit.



Commentaire :

Le nouveau mur nord (M36) du bâtiment 11 est construit en retrait de l'ancien, dans le prolongement du mur nord (M85) du bâtiment 22. Le four construit à l'ouest, sur l'emprise du bâtiment 22, prend appui sur le conduit des anciennes latrines du bâtiment 11 et scelle la tranchée de destruction de l'ancien mur pignon M42.

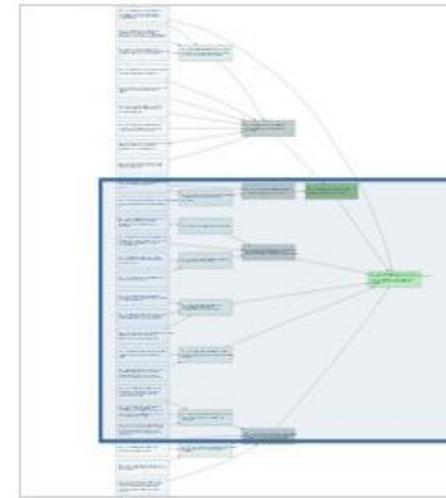
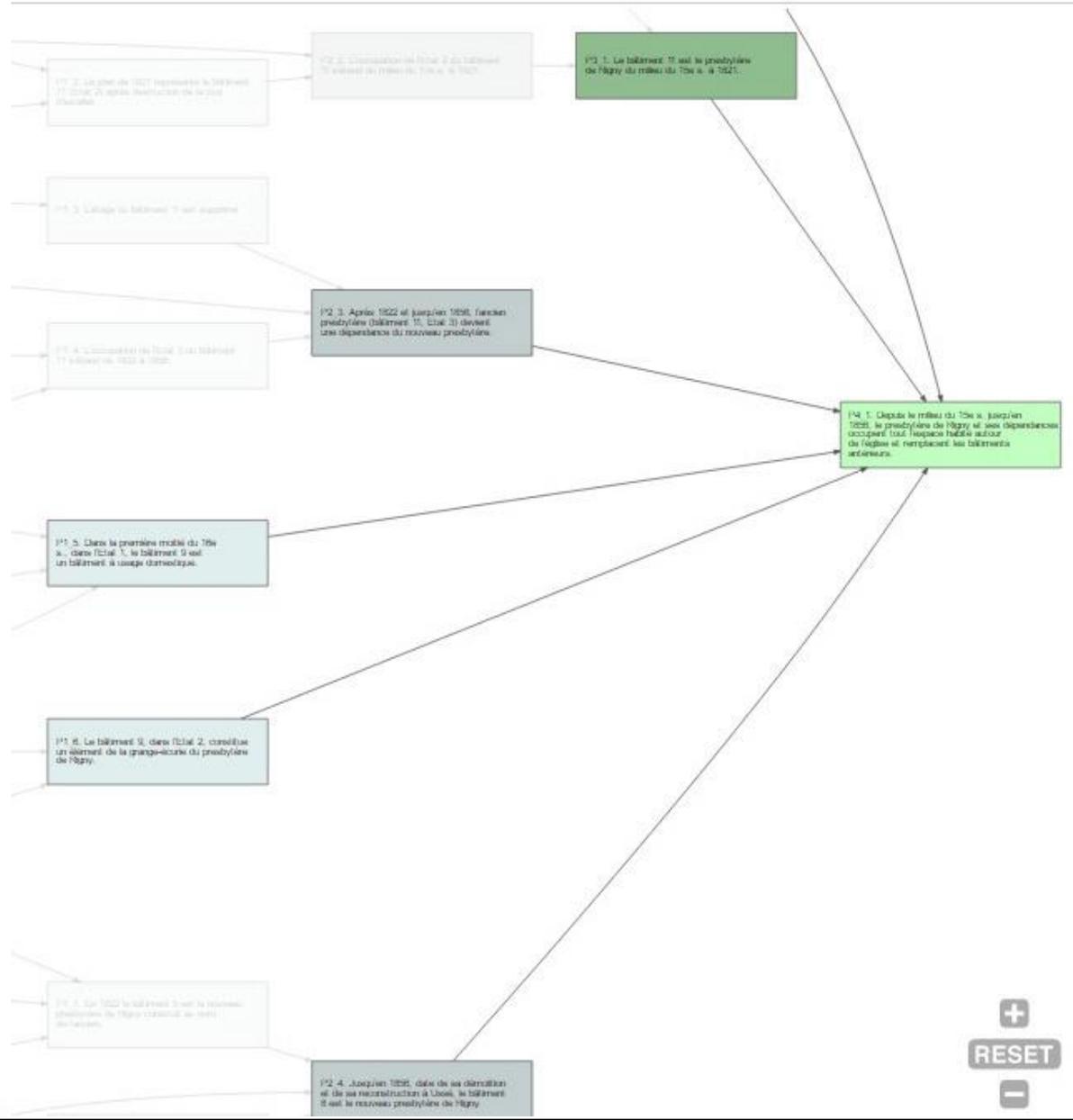
P0_13. D'après un procès-verbal de 1824, le bâtiment 11 sert de cuisine au nouveau presbytère construit plus au nord en 1822 (bâtiment 5).

Commentaire :

« Le conseil... a reconnu que la réclamation de M. le desservant était juste ... et a examiné que le four qui est en dehors est en très mauvais état sans pouvoir cuire le pain attendu qu'il reçoit toute l'eau de la toiture... Le conseil a aussi remarqué que le plancher de ladite cuisine est très vieux, qu'il fond de toutes parts et qu'il est urgent de la refaire à neuf » (A.D.I.L., série D, conseil municipal du 8 mai 1824).

P1_3. L'étage du bâtiment 11 est supprimé.

Section 4 - Le presbytère de Rigny et ses dépendances (milieu 15e-milieu 19e s.)



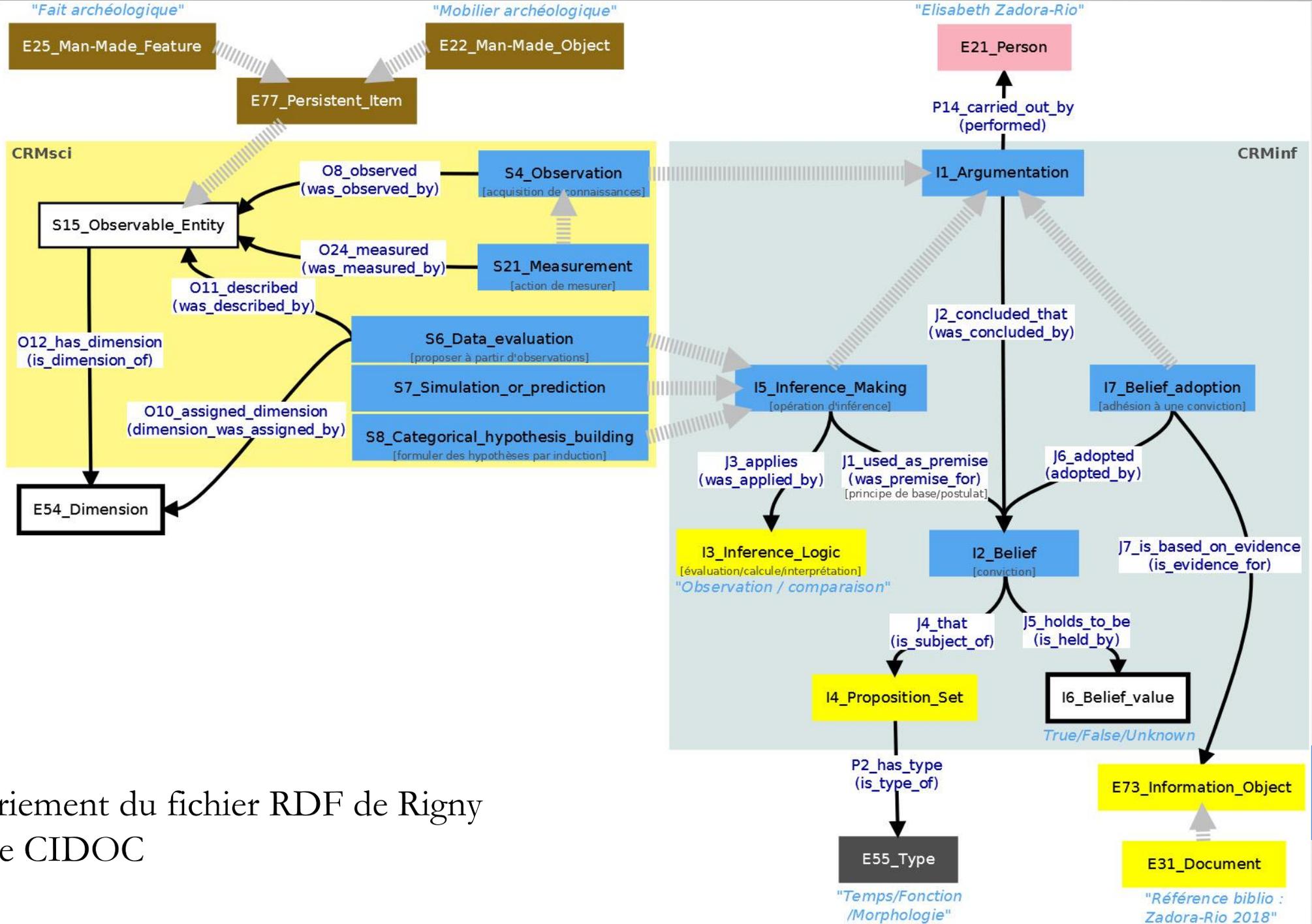
40 propositions / 5 niveaux

- 26 de niveau 0:
- 8 de niveau 1:
- 4 de niveau 2:
- 1 de niveau 3:
- 1 de niveau 4:

Proposition P4_1 / 6 antécédent(s)

P4_1. Depuis le milieu du 15e s. jusqu'en 18: le presbytère de Rigny et ses dépendances occupent tout l'espace habité autour de l'église et remplacent les bâtiments antérieurs.

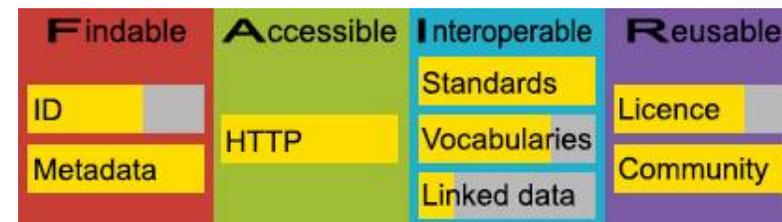




Appariement du fichier RDF de Rigny avec le CIDOC

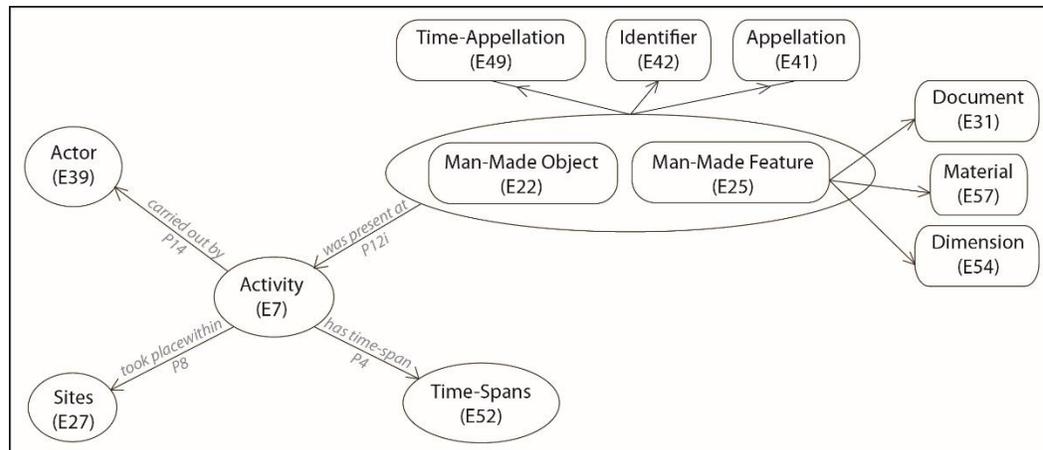
Progression pour Rigny

- En ligne
- Données structurées
- Format libre
- standard W3C
- Interopérable



Conclusion

- Recherche d'autres sources de données archéologiques en ligne pour tester l'interopérabilité des données que nous avons déjà appariées avec le CIDOC.
- Tester l'interopérabilité à différentes échelles pour des bases de granularités différentes (régionales, fouille).
- Formation des archéologues de MASA avec l'organisation d'une école thématique en 2019.



Merci