

Thesaurus, alignement et enrichissement : cas d'usage et stratégie à l'INHA

Webinaire wiki, data et GLAM 2021

https://meta.wikimedia.org/wiki/Webinaire_wiki,_data_et_GLAM_2021/Programme

Les 3 principes fondamentaux d'AGORHA

AGORHA une méta-base "mutualisée"

AGORHA permet de consulter en ligne des bases de données patrimoniales et de recherche en histoire de l'art et archéologie produites par l'Institut national de l'histoire de l'art (INHA) et ses partenaires. Œuvre collective, AGORHA s'enrichit continuellement de nouvelles notices, de nouvelles bases de données témoignant à la fois de la diversité et de la complémentarité des problématiques abordées à l'INHA.

AGORHA des notices de différents "type"

Les notices d'AGORHA décrivent des "œuvres", "personnes et organismes", "références bibliographiques", "fonds d'archives", "événements" et "collections". Toutes les notices sont interconnectées et permettent des renvois dynamiques. Des ouvrages numérisés en lien avec les bases de données enrichissent ces ressources numériques.

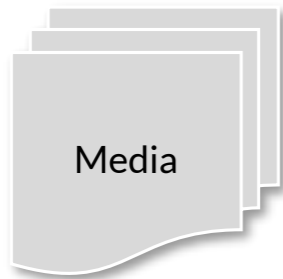
AGORHA un projet de "redocumentarisation"

Les programmes de recherche donnent lieu à la production de ressources documentaires disponibles en ligne pour la communauté scientifique et le grand public via AGORHA. Cette production originale, qui enrichit la connaissance d'objets conservés dans des institutions culturelles, est constitué de notices d'autorité, des instruments de recherche ou des bibliographies spécifiques.

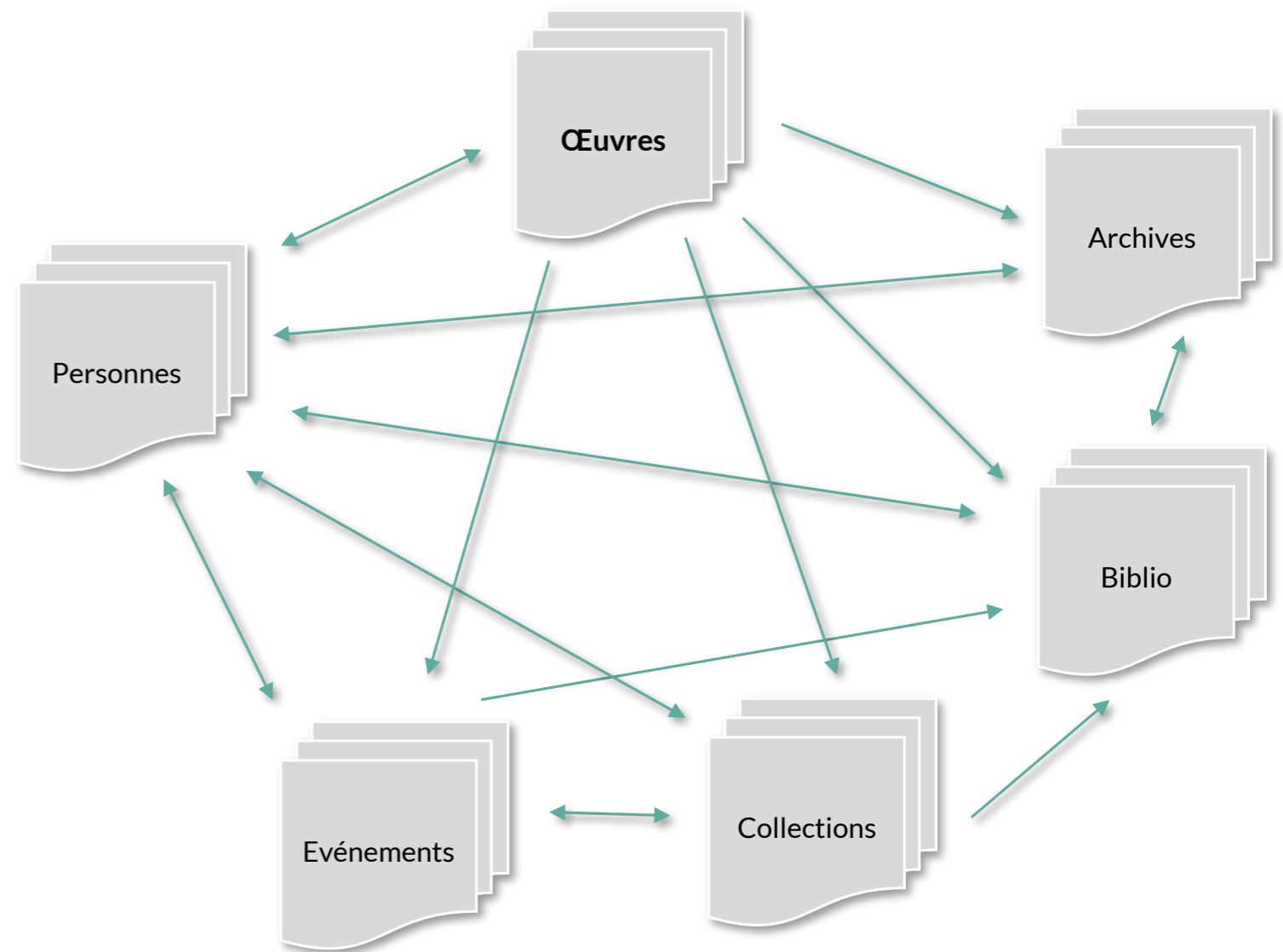
Des éléments techniques...



Progiel Flora :
Base de données Oracle SQL
(blob XML) + interface Java
interface



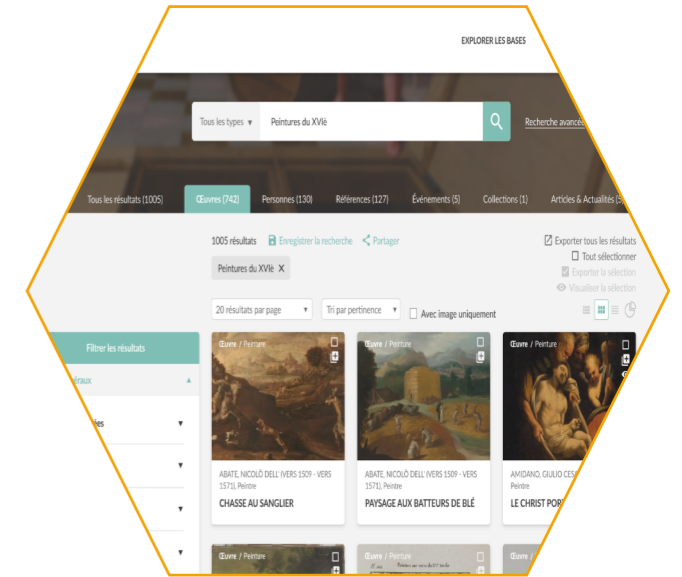
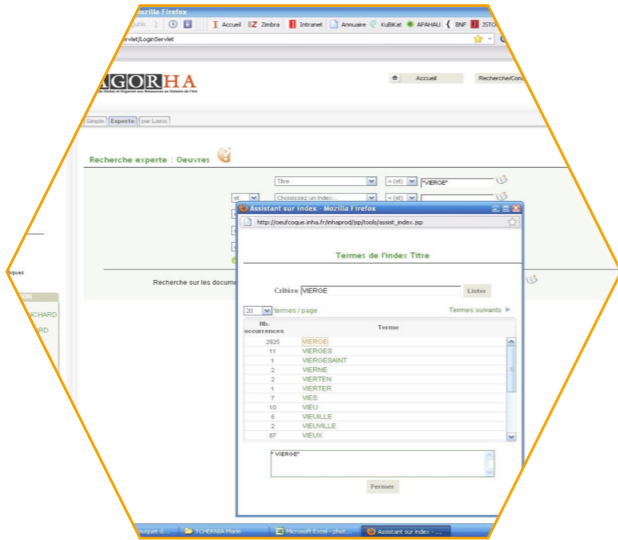
Stockés dans un webservice
et accessibles par un servlet
Java ou en OAI-PMH



2011

2018

2021 - bientôt



- 1^{ère} version d'AGORHA
- débutée en 2006, publiée en 2011

- montée de version de l'appliatif (progiciel Flora) en mai 2018

- Le Futur AGORHA 2
- Dialogue compétitif lancé en août 2018
- Début du projet accompagné par Sword : mai 2019



<https://numrha.hypotheses.org/2164>



Mais avant ce projet, des
expérimentations pour s'inscrire
dans le LOD

Mix'n'match

français



agorha

1/2



Se connecter à Wikidata pour les actions

Rechercher



AGORHA person

Action

identifieur for a person or institution in the Agorha database (INHA)

Importé par Magnus Manske | [Rafraichir](#)

Entrées

| | | |
|------------------------------|--------------|-------|
| Associations validées | 12578 | 35.4% |
| Préassociées | 8961 | 25% |
| Pas de Wikidata | 0 | 0% |
| Non applicable pour Wikidata | 10 | 0% |
| Non associées | 13968 | 39.3% |
| Total | 35517 | |



```

1 PREFIX bd: <http://www.bigdata.com/rdf#>
2 PREFIX wikibase: <http://wikiba.se/ontology#>
3 PREFIX wd: <http://www.wikidata.org/entity/>
4 PREFIX wdt: <http://www.wikidata.org/prop/direct/>
5 select distinct ?personne ?personneLabel ?idAgorha ?idBnF ?ulan ?ISNI ?SUDOC ?VIAF
6 where {
7 ?personne wdt:P31 wd:Q5 . #pour chercher des personnes humaines
8 ?personne wdt:P106 wd:Q42973 . #qui occupe la fonction d'architecte
9 ?personne wdt:P27 wd:Q142 .
10 ?personne wdt:P2342 ?idAgorha . #on veut l'id d'AGORHA de l'INHA
11
12 SERVICE wikibase:label { bd:serviceParam wikibase:language "fr, en". }
13 OPTIONAL {?personne wdt:P268 ?idBnF . #et plein d'autres id
14 ?personne wdt:P245 ?ulan .
15 ?personne wdt:P213 ?ISNI .
16 ?personne wdt:P269 ?SUDOC .
17 ?personne wdt:P214 ?VIAF .
18 }
19 }
20

```



1195 résultats en 2988ms

Code

Télécharger

Lien

Search

| personne | personneLabel | idAgorha | idBnF | ulan | ISNI | SUDOC | VIAF |
|-----------------------------|----------------------------|----------|-----------|-----------|---------------------|-----------|----------|
| wd:Q1871779 | Louis-Jean Desprez | 10266 | 122667294 | 500013852 | 0000 0000 6644 6014 | 031470718 | 14829822 |
| wd:Q924704 | Victor Laloux | 81327 | 12098140r | 500023967 | 0000 0000 7140 7773 | 029332591 | 34723705 |
| wd:Q2396270 | Maximilien Joseph Hurtault | 78501 | 151079695 | 500033645 | 0000 0000 6663 8569 | 129535427 | 69226784 |
| wd:Q2825101 | Adrien-Louis Lusson | 78620 | 10530100x | 500033638 | 0000 0001 1466 0534 | 139098526 | 2459734 |

96 matching lignes (152 tota) Extensions: Wikidata

Le schéma Wikidata ci-dessous spécifie comment vos données sont transformées en modifications sur Wikidata. Vous pouvez glisser-déposer les noms des colonnes ci-dessous dans la plupart des champs: pour chaque ligne, des modifications seront générées avec les valeurs dans ces colonnes.

Enregistrer le schéma Annuler les modifications

| | | | | | | | | | | | |
|-----------------------------|------------------------|----------------------------|-----------------------|------------------------|-------------------|--------------------|------------------------|-----------------------|-------------|----------------------------|------------------|
| Reference_notice | URL | Niveau | Label | Lieu_conservation | Cote | Cote_origine | Cotes_extremes | Dates | Notes_dates | Note_importance_materielle | Nombre_de_pieces |
| Métrage_lineaire | Supports | Nature_contenants | Nombre_contenants | Producteurs | Histoire_admin | Histoire_conserv | Modalites_entree | Note_modalites_entree | | | |
| Auteur_instrument_recherche | Analyse | Evaluation_tri_elimination | Accroissements | Mode_de_classement | Noms_de_personnes | Mots_des_sujet | Mots_des_geographiques | | | | |
| Statut_juridique | Notes_statut_juridique | Communicabilite | Notes_communicabilite | Informations_pratiques | Reproduction | Notes_reproduction | Langues | | | | |

pas de notes, descriptions ou sites ajoutés

Déclarations

archivé par

numéro d'invent

niveau de descript

descript à PURL

▼ 1 références

URL de la référe

date de consulte

Paul Deschamps

Biographie

| | |
|-------------------------|-----------------------------|
| Naissance | 19 septembre 1888 |
| | Paris |
| Décès | 25 février 1974 (à 85 ans) |
| | Paris |
| Nom de naissance | Paul Deschamps |
| Nationalité | Français |
| Formation | École des chartes |
| Activité | Historien |

Autres informations

| | |
|--------------------------------|---|
| Membre de | Académie des inscriptions et belles-lettres |
| Distinctions | Chevalier de la Légion d'honneur (1931) Commandeur de la Légion d'honneur (1968) |
| Archives conservées par | Institut national d'histoire de l'art (Archives 001) |

owned by



Edmé-Antoine Durand

edit

start time

unknown value

end time

25 April 1836 *Gregorian*

▶ 1 reference



César-Emmanuel-Flavien Henrion de Staal de Magnoncour

edit

start time

25 April 1836 *Gregorian*

end time

18 April 1839 *Gregorian*

▼ 1 reference

statement supported by

Institut National d'Histoire de l'Art

source website for the property

<http://www.purl.org/inha/agorha/009/104>

reference URL

<http://www.purl.org/inha/agorha/009/2167>

retrieved

18 October 2017

+ add reference



Jarry

edit

start time

18 April 1839 *Gregorian*

end time

unknown value

▶ 1 reference

+ add value

Revenons au futur

AGORHA

- Affirmer ce qu'est AGORHA et ce que n'est pas AGORHA !
 - Méta-base mutualisée des bases (comme acte de mise en donnée de l'information) réalisées dans le cadre de projet de recherche
 - Pas une bibliothèque numérique ni portail d'instrument de recherche archivistique ni pour faire de l'édition numérique > choix d'outils dédiés (ATOM / ExistDB avec TEIPublisher)
 - Pas un portail fédéré à l'échelle de l'établissement
- **Différencier le modèle de travail, le modèle de saisie, le modèle de stockage, le modèle d'échange et le modèle conceptuel cible**
- Externaliser (consciemment) les objets documentaires
 - Les thesaurus dans une application tierce à l'applicatif AGORHA
 - Appeler des référentiels extérieurs (wikidata, ulan, idref)
 - Appeler des médias externes (via IIF ou via API de la RMN)



Interface de saisie des thesaurus
(permettant les alignements,
description des libellés, etc.)



Portail d'accès
aux thésaurus
thesaurus.inha.fr



Interface de saisie/acquisition des
données pour les bases de données

Portail d'accès
global
agorha.inha.fr



Interface de gestion et de saisie pour
la partie éditoriale : création des pages
d'accueil de bdd, article, actualités

GeoNames

```
moi@monserveur:~$
```

```
http://api.geonames.org/search?q=&name=Sardaigne&maxRows=1&continentCode=EU&country=IT  
&lang=fr&style=short&username=*****&token=*****
```



- **15.421** concepts de lieux
- **15.420** ont désormais une couple de coordonnées
- **1** concept « localisation inconnue » LAT : NULL; LON : NULL
- **1.854** termes (12%) ne matchaient pas ni avec l'API Geonames, ni en réconciliation WikiData
- nous avons choisi d'utiliser les coordonnées du concept « broader » (noté pour le reprendre ensuite).
- La géolocalisation est donc complète au 100% pour l'import dans GINCO.



Créer une notice

Base de données

Sélectionnez la base de données dans laquelle s'inscrit la notice

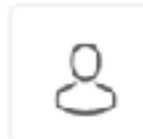
Entrez une base de données

Type de notice

Sélectionnez un type de notice dans la liste ci-dessous



Œuvre



Personne



Événement





Référence



Collection

Créer une notice


 Création d'une notice de type "Œuvre" dans le programme "Répertoire des tableaux italiens dans les collections publiques françaises (XIIIe-XIXe siècles), RETIF - INHA"

Identification Localisation Description Création-Exécution 
Historique Protection patrimoniale Œuvres imprimées et manuscrites 

Ajouter un champ / bloc de champs

Quitter sans enregistrer

Enregistrer et quitter

 Création d'une notice de type "Œuvre" dans le programme "Répertoire des tableaux italiens dans les collections publiques françaises (XIIIe-XIXe siècles), RETIF - INHA"

[Ajouter un champ / bloc de champs](#)[Quitter sans enregistrer](#)[Enregistrer et quitter](#)

Identification

Type d'oeuvre *

Commentaire Type d'oeuvre *

I **U**

Saisir et mettre en forme du texte...

Nombre d'objet

Bloc Titres


Titres 1

[+ Ajouter](#)

Type de titre

Titre

fre 

 Création d'une notice de type "Œuvre" dans le programme "Répertoire des tableaux italiens dans les collections publiques françaises (XIIIe-XIXe siècles), RETIF - INHA"

Identification

Type d'oeuvre *

peinture

peinture

A pour générique : peinture, arts graphiques et des média

peinture monumentale

A pour générique : peinture

Ajouter

peinture murale

A pour générique : peinture

Ajouter un champ / bloc de champs

Quitter sans enregistrer

Enregistrer et quitter

THÉSAURUS ET VOCABULAIRES CONTRÔLÉS DE L'INSTITUT NATIONAL D'HISTOIRE DE L'ART

L'Institut national d'histoire de l'art (INHA) donne ici accès à l'ensemble des vocabulaires scientifiques et techniques qu'il produit. Cette plate-forme est une instance de GINCO, développé par le ministère de la Culture, pour donner accès, de manière homogène et centralisée, à l'ensemble des vocabulaires scientifiques et techniques produits par le Ministère et ses partenaires.

De nouveaux vocabulaires y prendront progressivement place, tous disponibles au format SKOS. Les vocabulaires peuvent être consultés en ligne ou téléchargés et interrogés par les machines au moyen du langage de requête SPARQL.

La rubrique En savoir plus propose des informations complémentaires sur les conditions de production, de mise à disposition et d'utilisation des vocabulaires.

Un outil de recherche simple permet de rechercher des concepts ou des termes présents dans les différents vocabulaires disponibles. La recherche s'effectue à la fois sur les entrées des vocabulaires et dans les notes et définitions qui leur sont associées.

Les différents concepts sont accessibles via le formulaire de recherche suivant :

Chercher le concept

Accéder à la liste des vocabulaires :

- [Vocabulaires par service producteur](#)
- [Vocabulaires par sujet](#)
- [Tous les vocabulaires](#)

Enduit

ENDUIT

Vocabulaire source

[Thésaurus Caractéristiques de description matérielle](#)

Descripteur

- (fr-FR) enduit
- (en-US) plaster

Note d'application

- (fr-FR) Couche protectrice et/ou décorative, appliquée sur les surfaces architecturales. Elle peut servir de support secondaire aux peintures murales.

Concept(s) équivalent(s) dans d'autres vocabulaires

- <https://vocab.getty.edu/page/aat/300014922>
- <https://www.wikidata.org/wiki/Q47496>

Statut

- Validé

Date de création

10 mai 2021

Date de dernière modification

26 mai 2021



Données brutes en : [En savoir plus](#) [SPARQL endpoint](#) [Site du producteur](#)

- RDF/XML
- Notation3
- Turtle

Blocs répétables...

Matières et techniques 3

Matières et techniques 2

Matières et techniques 1

Type de description matérielle

Caractéristiques

Matériaux

Techniques

Le « Bloc » est répétable autant de fois qu'on nécessite...

Sources du bloc 'Matières et techniques'

Liste de "Source"

Source 1 + Ajouter

Source bibliographique / Mention d'archive

Rechercher ou ajouter un(e) référence

Source externe

Source base de données

Entrez une base de données

Précisions source

Supprimer

Fermer

...et « sourçables »

Terme

Guido Reni

 Lieu Titre Personne Organisation

wikidata
Guido Reni
Bolognese painter (1575-1642)
<http://www.wikidata.org/entity/Q109061>

dbpedia
Guido Reni
Guido Reni (Italian pronunciation:
[ɡwido ˈreni]; 4 November 1575 – 18
August 1642) was an Italian painter of
the Baroque period, although his
works showed a classical manner. a

getty
Reni, Guido
Italian painter, draftsman, 1575-1642
<http://vocab.getty.edu/ulan/500030334>

wikidata
**Guido Reni (1575-1642),
Italian artist, painter**
painting by unknown artist
<http://www.wikidata.org/entity/Q4322030>

wikidata
Guido Reni - 16 Bologna
primary school in Bologna in the
province of Bologna (Italy) [school Id:
BCMM87301D]
<http://www.wikidata.org/entity/Q52645389>

Terme

Musée des augustins

 Lieu Titre Personne Organisation

wikidata
Musée des Augustins
fine arts museum of Toulouse, France
<http://www.wikidata.org/entity/Q2711480>

dbpedia
Musée des Augustins
The Musée des Augustins de Toulouse is a fine
arts museum in Toulouse, France which
conserves a collection of sculpture and paintings
from the Middle Ages to the early 20th century...
http://dbpedia.org/resource/Musée_des_Augustins

getty
Freshley Shoals
No description available
<http://vocab.getty.edu/gn/2040735>

Gineco MCC
musée des Augustins
[musées non nationaux > Liste d'autorités
Localisation - Joconde](#)
<http://data.culture.fr/thesaurus/source/work/627477515-1...>

wikidata
Musée des Augustins
Wikimedia disambiguation page
<http://www.wikidata.org/entity/Q330118>

getty
Des Lacs
No description available
<http://vocab.getty.edu/gn/2077537>

wikidata
**Musée des Augustins
d'Hazebrouck**
museum in France
<http://www.wikidata.org/entity/Q330112>

getty
Des Lacs
No description available
<http://vocab.getty.edu/gn/1121551>

getty
Des Arc
No description available
<http://vocab.getty.edu/gn/2056550>

getty
Des Plaines

Annuler

Rechercher

- Utilisé soit :
 - Aligner nos entités documentaires
 - Alimenter des champs de saisie d'une entité documentaire en s'appuyant sur une entité externe
- Récupération du libellé (value) pour l'URI de l'entité

- 2003 - la première norme du JSON est ECMA-404
- Ensembles de paires « nom » (alias « clé ») / « valeur »
- Les valeurs représentent trois types de données :
 - objets
 - tableaux
 - valeurs génériques

```

{
  "objet": {
    "clé": "valeur",
    "autre clé": 1,
    "clé qui contient un objet": {
      "clé qui contient un tableau": [
        { "clé": "chaîne de caractères" },
        { "clé": null },
        { "clé": true }
      ]
    }
  }
}

```



<https://agorha.inha.fr/ark:/54721/3152a4f5-7250-4dac-927f-09bf06dd586d>

← les uuid ne sont pas encore actifs... :)

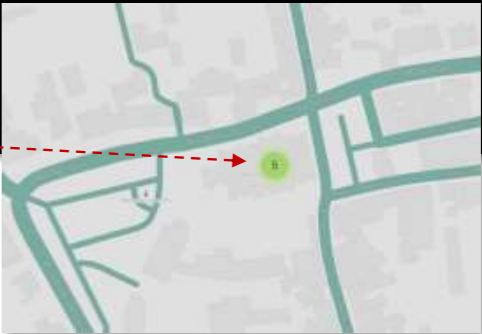
```

"localizationInformation": {
  "localization": [
    {
      "place": {
        "thesaurus": {
          "prefLabels": [
            {
              "value": "Eglise Saint-Trémeur de Carhaix (Carhaix-Plouguer)",
              "language": "fre"
            }
          ]
        },
        "ref": "https://thesaurus.inha.fr/c004ad47-b259-4b09-bd03-123800199ee3",
        "conceptPath": "/Europe/France/Bretagne/Finistère/Carhaix-Plouguer/Eglise Saint-Trémeur de Carhaix (Carhaix-Plouguer)",
        "geoPoint": "48.277338,-3.575036"
      }
    }
  ]
}

```

Localisation

Lieu de conservation :



```

{
  "@context": {
    "dc": "http://purl.org/dc/terms/",
    "crm": "http://www.cidoc-crm.org/cidoc-crm/"
  },
  "@id": "https://agorha.inha.fr/ark:/54721/3152a4f5-7250-4dac-927f-09bf06dd586d",
  "@type": "crm:E22_Man-Made_object",
  "crm:P1_is_identified_by": {
    "crm:P102_has_title": {
      "@language": "fre",
      "@value": "Carhaix-Plouguer, Eglise Saint-Trémeur de Carhaix : élévation"
    },
    "crm:P67_refers_to": {
      "@id": "https://agorha.inha.fr/ark:/54721/7a78775c-842e-4bb3-8ee7-cd8981b75751",
      "@type": "crm:E22_Man-Made_object"
    },
    "dc:available": "Publiée",
    "dc:creator": "system",
    "dc:date": "2018-07-27T16:23:23",
    "dc:dateSubmitted": "2021-03-01T21:59:40.532781"
  }
}

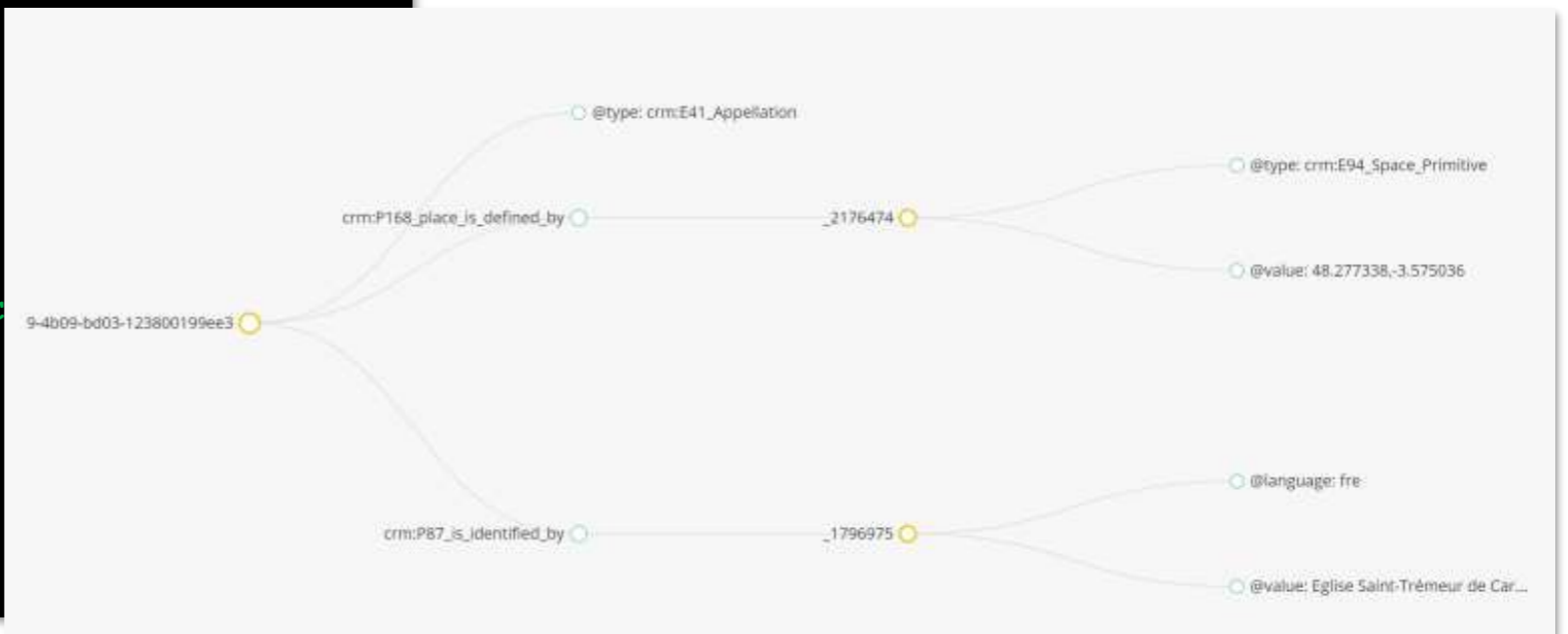
```

- 2014 – standard W3C
- Concept de « contexte »
- Associations depuis du JSON vers un modèle RDF
- @id spécifiés comme IRI

Les ontologies

```
{
  "@id" : "https://agorha.inha.fr/ark:/54721/3152a4f5-7250-4dac-927f-09bf06dd586d",
  "@type" : "crm:E22_Man-Made_object",
  [...]
  "crm:P54_has_current_permanent_location" : {
    "crm:E53_Place" : {
      "crm:P1_is_identified_by" : {
        "@id" : "https://thesaurus.inha.fr/c004ad47-b259-4b09-bd03-123800199ee3",
        "@type" : "crm:E41_Appellation",
        "crm:P168_place_is_defined_by" : {
          "@type" : "crm:E94_Space_Primitive",
          "@value" : "48.277338,-3.575036"
        },
        "crm:P87_is_identified_by" : {
          "@language" : "fre",
          "@value" : "Eglise Saint-Trémeur de Car...
Plouguer) "
        }
      }
    },
    [...]
  }
}
```

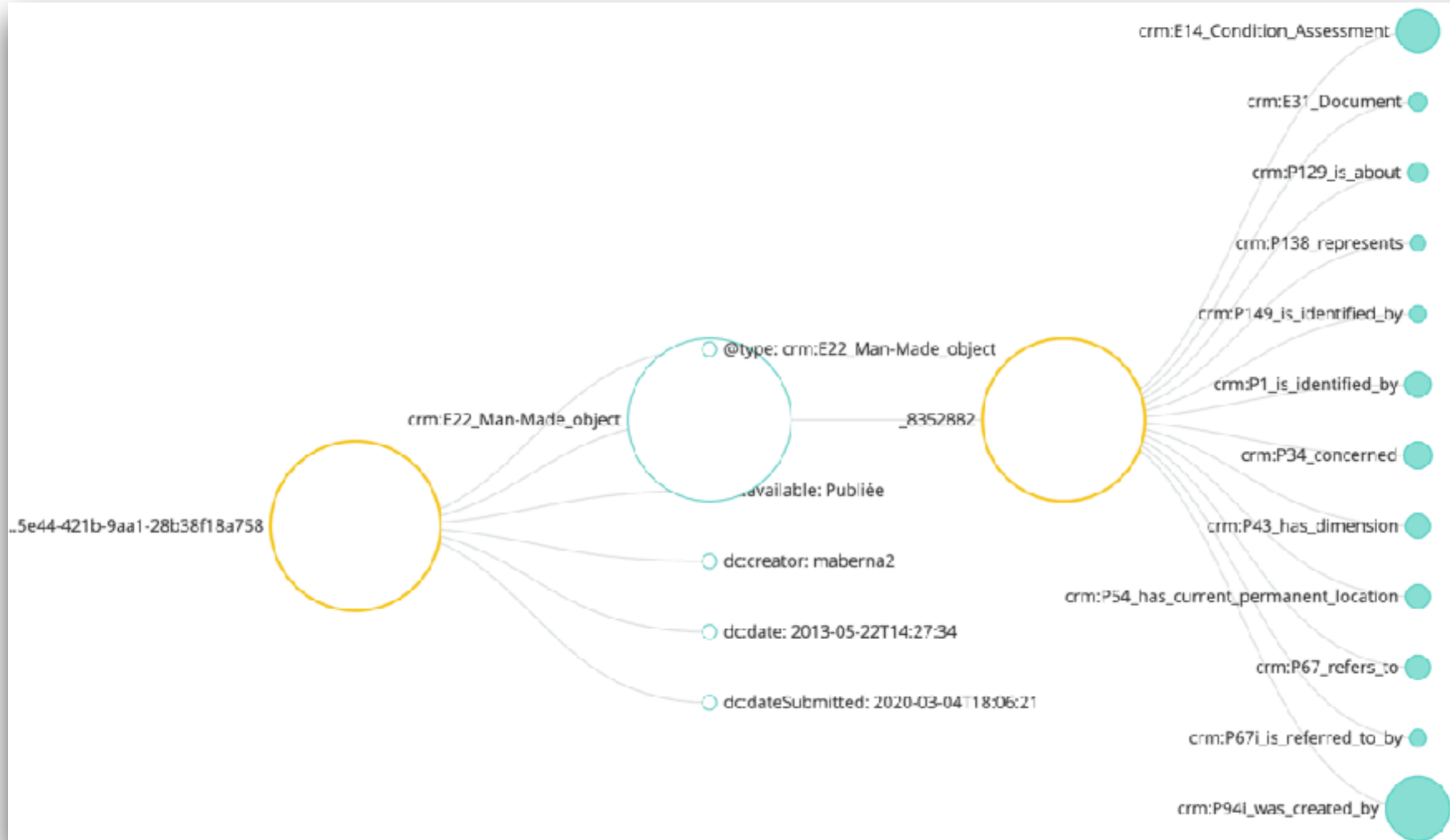
```
"@context" : {
  "schema" : "http://schema.org/",
  "rdfs" : "https://www.w3.org/TR/rdf-schema/",
  "prov" : "http://www.w3.org/ns/prov#",
  "dc" : "http://purl.org/dc/terms/",
  "crm" : "http://www.cidoc-crm.org/cidoc-crm/"
}
```



| | | |
|--|-------------------------|------------------------------------|
| internal.uuid | @id | id (à enlever) |
| internal.noticeType | | crm:E22_Human-Made_object |
| internal.digest.title | rdfs:label | |
| content.identificationInformation | crm:P1_is_identified_by | crm:E33_E42_Linguistic_Appellation |
| content.identificationInformation.title | crm:P102_has_title | crm:E35_Title |
| content.identificationInformation.title.type | crm:P2_has_type | crm:E55_Type |
| content.identificationInformation.title.type.thesaurus | | |
| content.identificationInformation.title.type.thesaurus.ref | @id | |
| content.identificationInformation.title.type.thesaurus.conceptPath | | |
| content.identificationInformation.title.type.thesaurus.prefLabels | | |
| content.identificationInformation.title.type.thesaurus.prefLabels.value | rdfs:label | |
| content.identificationInformation.title.type.thesaurus.prefLabels.language | | |
| content.identificationInformation.title.type.thesaurus.altLabels | | |
| content.identificationInformation.title.type.thesaurus.hiddenLabels | | |
| content.identificationInformation.title.label | | |
| content.identificationInformation.title.label.value | crm:P190_has | |
| content.identificationInformation.title.label.language | @language | |
| content.identificationInformation.title.comment | crm:P3_has_r | |

```

object ▶ crm:P1_is_identified_by ▶ 0 ▶ crm:P72_has_language ▶ 0 ▶
  ▼ object {5}
    @context : agorha.inha.fr/modele/agorha_templating.json
    id : https://agorhatest.inha.fr/ark:/54721/0ed5b7e5-5dd8-4cfa-90a7-903e78736083
    type : crm:E22_Human-Made_object
    rdfs:label : Saint Abel découvrant la terre
    ▼ crm:P1_is_identified_by [1]
      ▼ 0 {5}
        type : crm:E33_E41_Linguistic_Appellation
        ▼ crm:P2_has_type [1]
          ▼ 0 {3}
            @id : https://referentiel.inha.fr/resource/ark:/54721/8d62bef5-a3af-4c8f-bc09-9139f67657d9
            @type : crm:E55_Type
            rdfs:label : titre figurant sur le dessin
            crm:P190_has_symbolic_event : Saint Abel découvrant la terre
          ▼ crm:P72_has_language [1]
            ▼ 0 {2}
              type : crm:E56_language
              rdfs:langage : fre
            crm:P3_has_note : Le mot terre est barré sur le dessin
  
```



Kyathos attique à figures noires sur fond blanc

Groupe du Vatican G 57



- Exports Json
- Format Json LD**
- Exports RDF
- Format RDF/XML
- Format N3
- Format N-Triples

A row of action buttons: "Ajouter à une sélection", "Imprimer la notice", "Modifier la notice", "Partager la notice", and "Citer la notice". Below these is a red "Supprimer la notice" button. An "Exports" dropdown menu is open, showing the list of export formats.

Pour le moment, uniquement accès aux données RDF (avec x sérialisation) via négociation de contenu (suffixe de l'url)

Ce qui me pose question sur le long terme

- Quid des mises-à-jours des alignements et notamment pour des alignements déjà fait pour des entités qui « peuvent bouger » et qui influerait sur le type d'alignement ?

Merci !

der-snr@inha.fr