

La mue d'un thésaurus documentaire ou comment Pactols est devenu un référentiel FAIR, avec et pour les archéologues

Blandine Nouvel

▶ **To cite this version:**

Blandine Nouvel. La mue d'un thésaurus documentaire ou comment Pactols est devenu un référentiel FAIR, avec et pour les archéologues. 9e Café Renatis. Accompagnement à l'élaboration de vocabulaires contrôlés dans une démarche de FAIRisation des projets de recherche, Renatis, Dec 2022, En ligne, France. hal-03886987

HAL Id: hal-03886987
https://amu.hal.science/hal-03886987

Submitted on 6 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La mue d'un thésaurus documentaire

ou comment Pactols est devenu un référentiel FAIR, avec et pour les archéologues

Blandine Nouvel

« Accompagnement à l'élaboration de vocabulaires contrôlés dans une démarche de FAIRisation des projets de recherche»
9e C@fé Renatis, 6 décembre 2022

Question de vocabulaire

- Référentiel ou *controlled vocabulary* (vocabulaire contrôlé)
  - Liste de termes qui « fait autorité dans son domaine »
  - Eviter les ambiguïtés du langage courant et faciliter la recherche documentaire comme l'interopérabilité des (méta)données
  - Types
    - Classifications, listes/fichiers d'autorité (taxonomies), thésaurus => indexer et enrichir les données
    - Ontologies => organiser les données
- Thésaurus
  - Recense le vocabulaire utile à la description de contenus propres à un domaine de la connaissance (descripteurs, non-descripteurs)
  - Organisé hiérarchiquement
  - Relations sémantiques (de dépendance, d'associations, alignements)
  - Norme ISO25964 – parties 1, recherche d'information et 2, interopérabilité entre les vocabulaires

Le thésaurus Pactols

- Créé en 1985 pour l'indexation du Catalogue collectif de Frantiq (Fédération et Ressources sur l'Antiquité, GDS3378 du CNRS)
- 62000 concepts multilingues pour l'archéologie et les sciences de l'Antiquité

- Libre d'accès
  - https://pactols.frantiq.fr
- Géré avec Opentheso
- Mis à jour en continu



Un thésaurus +/- FAIR

2017

Facile à trouver

Id pérenne	
Métadonnées riches	SKOS, Dublin Core
Référencé	
Interrogation standardisée	Recherche Opentheso, API

Accessible

Protocoles	https://pactols.frantiq.fr, API REST
Standards	ISO 25964-1

Interopérable

Langage normalisé	XML-SKOS, Json
Export	CSV, Json, Json-LD, SKOS-RDF
Alignements	AAT, Wikidata

Réutilisable

Licence	
Décrit, documenté	Présentation

L’objectif 2025

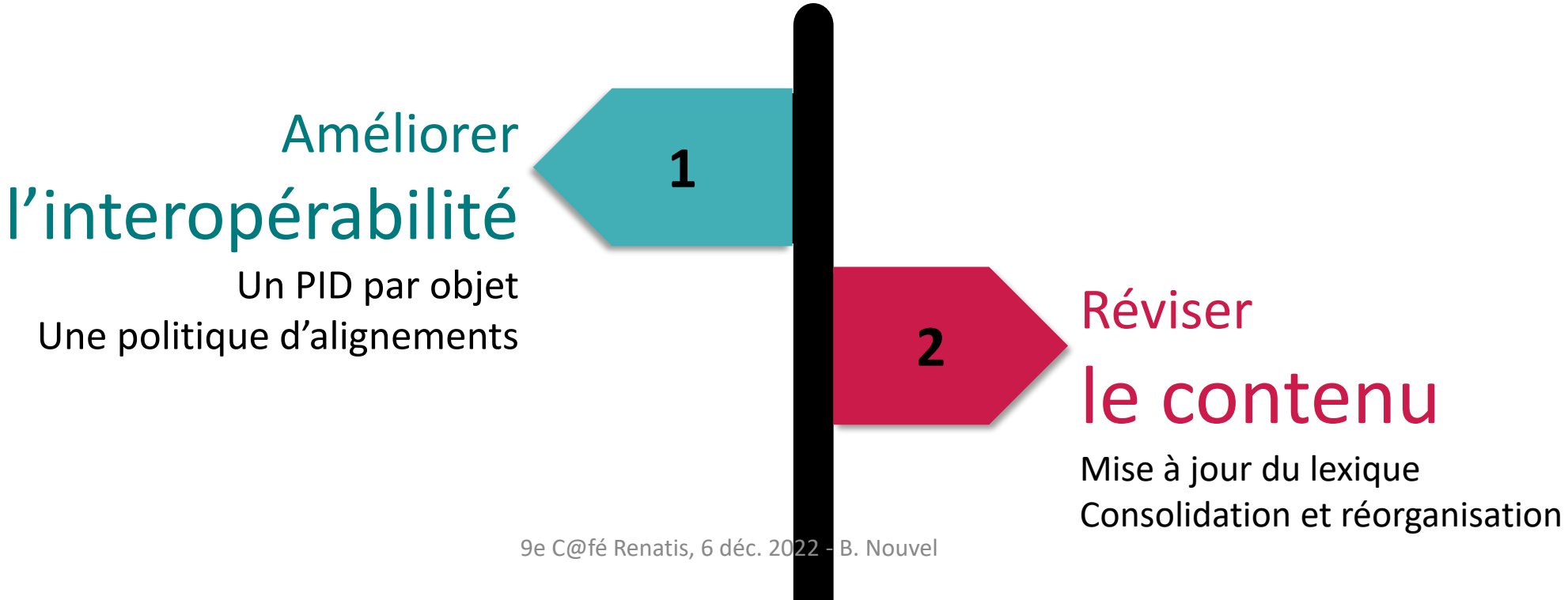


Fournir un référentiel disciplinaire
« clés en mains »

Transformer un thésaurus documentaire
en référentiel du web

Transformer Pactols

- Exploiter le potentiel du thésaurus (lexique, normalisation)
- Améliorer ses qualités FAIR



Un thésaurus de + en + FAIR

	2017	2022
Facile à trouver		
Id pérenne		Id ARK (thésaurus et concept)
Métadonnées riches	SKOS, Dublin Core	
Référencé		Opidor, écosystème numérique MASA
Interrogation standardisée	Recherche Opentheso, API	
Accessible		
Protocoles	https://pactols.frantiq.fr, API REST	évolution de l'API (filtre sur une·plusieurs collection·s)
Standards	ISO 25964-1	ISO 25964-2
Interopérable		
Langage normalisé (CIDOC-CRM)	XML-SKOS, Json	SKOS-RDF, modèle conceptuel du BackBone Thesaurus
Export	CSV, Json, Json-LD, SKOS-RDF	Github (https://github.com/frantiq/PACTOLS)
Alignements		Wikidata, GeoNames, IdRef, AAT, Pleiades, Periodo
Réutilisable		
Licence		Licence Open Database License (ODbL)
Décrit, documenté	Présentation	Guide, FAQ, sources Pactols sur Zotero

Comment conduire le chercheur à adopter un thésaurus + FAIR?



Communiquer

- Objectif : faire connaître le thésaurus au-delà du cercle des bibliothèques, informer sur l'avancement du projet (colloques MASA, AG Frantiq, Actu Frantiq

1. Communications dans des colloques
   - Journées MASA (2017-2022)
   - 18e congrès international de l'Union internationale des sciences pré et protohistoriques (UISPP) (Paris, juin 2018)
   - 5th Linked Past (Bordeaux, novembre 2018)
   - 11e Rencontres du GAAF, Groupe d'anthropologie et archéologie du funéraire (Tours, 2019)
2. Ateliers d'alignements Wikidata
   - *Magna Grecia* Open data (Aix-en-Provence, mars 2018)
   - 3 Ateliers (2019, 2020, Nanterre & En ligne)

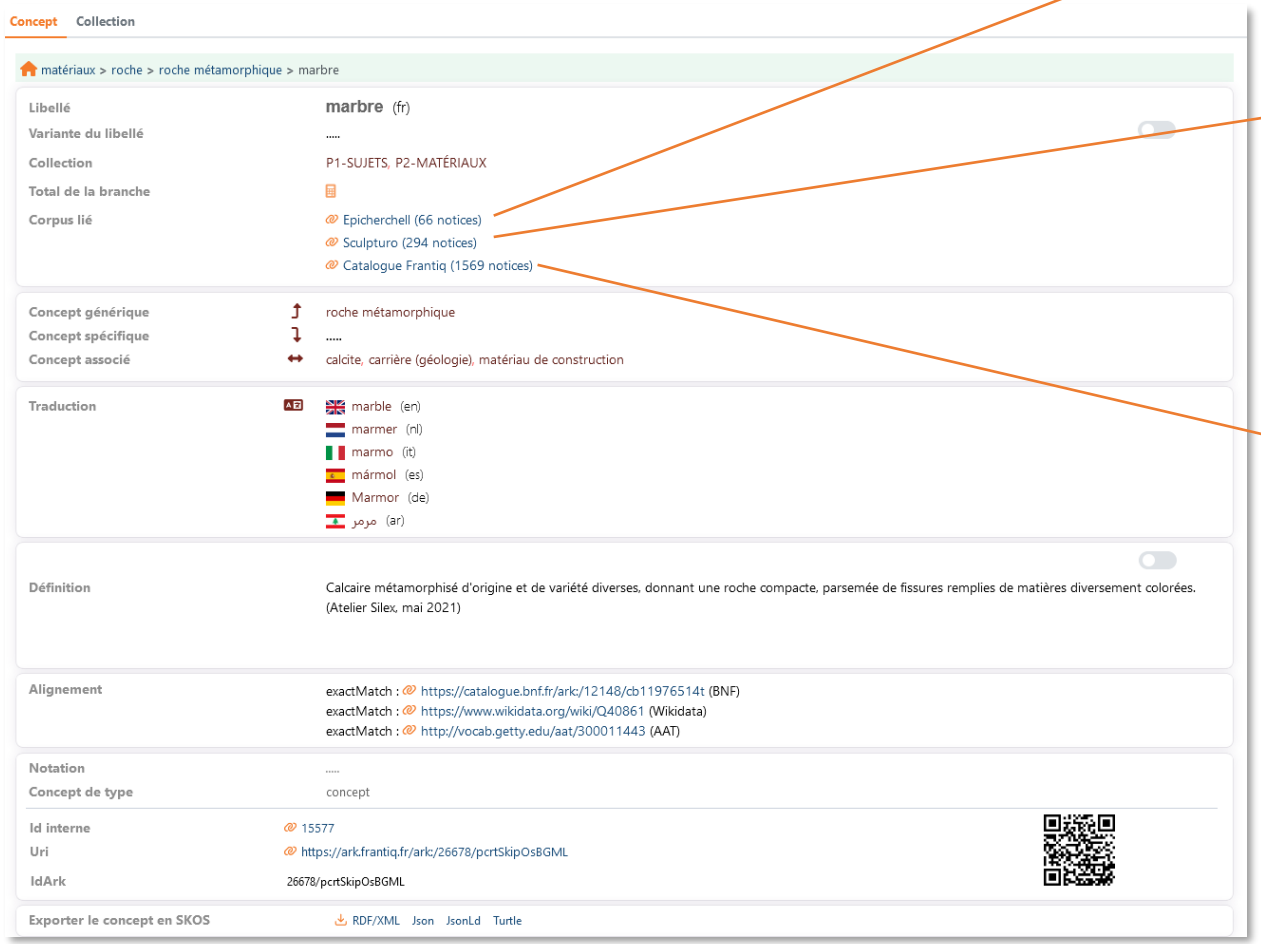


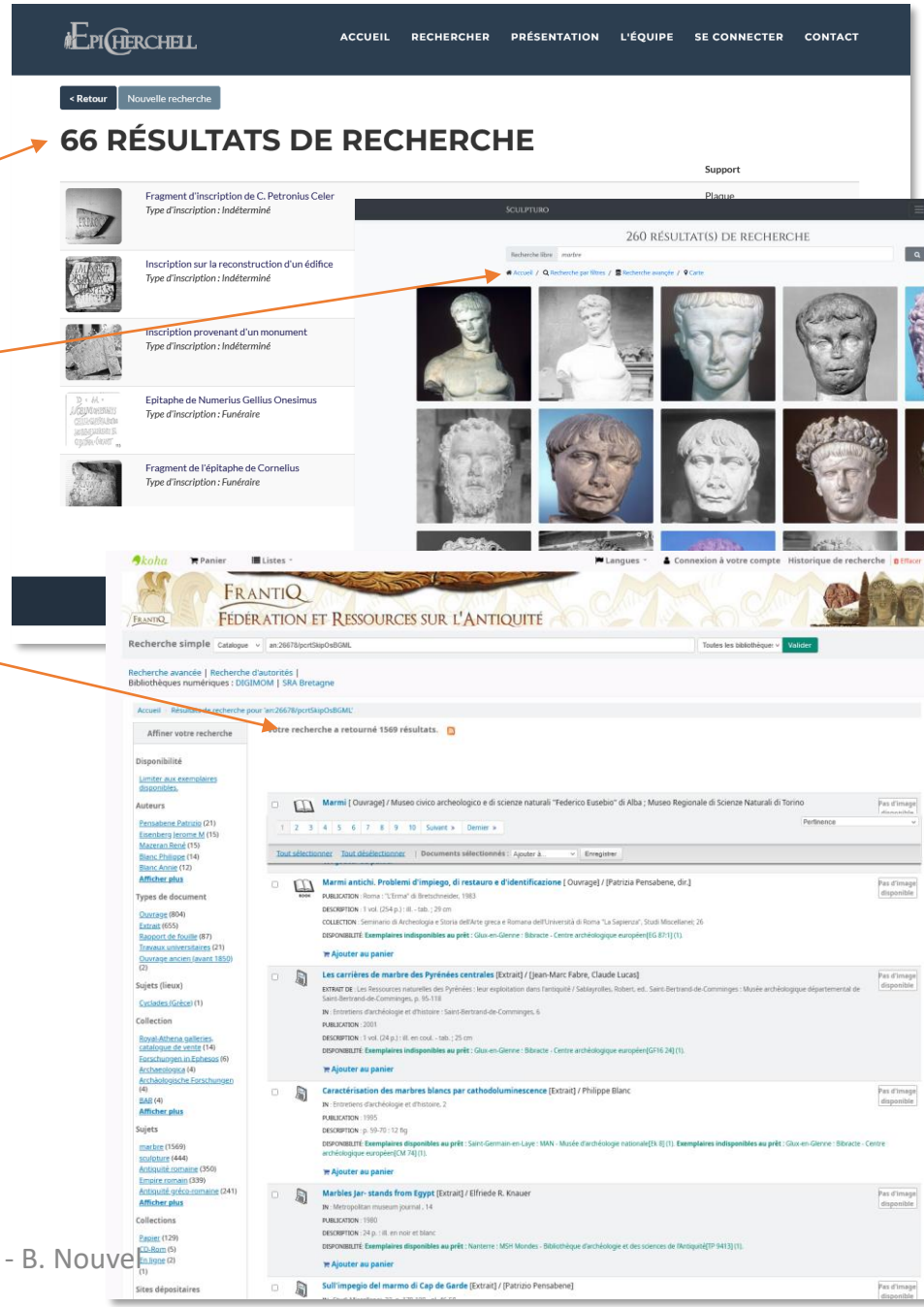
Associer

- Objectif : s'appuyer sur les métiers d'appui à la recherche, familiers des questions d'interopérabilité et de qualité des méta-données

1. Éditeurs
   - Projet MetaPHoRA (AAP MESRI 2019-2022), Métadonnées Plurilingues et Homogènes pour un collectif de Revues françaises en Archéologie
2. Gestionnaires de données et porteurs de projets de valorisation patrimoniale
   - Adopter Pactols comme vocabulaire commun
   - Optimiser le choix en intégrant les « Corpus liés » à Pactols

Lier les ressources





Intégrer

- Objectif : donner une place officielle aux scientifiques dans l'évolution du contenu

1. Révision de dossiers thématiques confiée à de jeunes chercheurs
   - Industries osseuse et lithique, pétrographie
   - Industrie céramique
   - Philosophie, Religion
2. Travailler sur des thématiques de spécialité
   - CollEx RefDivinités = contrôler et définir les figures mythologiques de la Méditerranée antique
   - Ateliers Mort = stabiliser un vocabulaire de spécialité
   - Dictionnaire de la pétrographie (réseau national des lithothèques) = étoffer et définir le lexique
3. Constitution d'une liste d'experts de référence



Répondre

- Objectif : sollicitations croissantes de chercheurs pour utiliser Pactols, avec des aménagements

1. Mettre en place un workflow spécifique
   - Soumission de listes personnelles
   - Rapprochement avec Pactols (80% des termes sont déjà dans Pactols)
   - Propositions pour utiliser les concepts Pactols identifiés comme exactMatch
   - Discussion sur les termes closeMatch
   - Création de nouveaux concepts

2. Permettre les interventions dans le thésaurus : nouvelles fonctions de Opentheso
   - Ouverture du module Candidats aux contributeurs
   - Possibilité de suggérer des améliorations sur un concept



Accompagner

- Objectif : ne pas laisser seul le chercheur face à la question des référentiels dans ses bases de données

1. Partie du « système numérique MASA », outils pour la gestion FAIR des données de l'archéologie et les bonnes pratiques
2. Former
   - Ateliers d'alignements Wikidata
   - Des ateliers MASA à l'offre de formations
3. Aider à la construction de thésaurus spécifiques
   - Pactols ne répond pas au besoin (granularité, haute spécialité)
   - Mais il constitue un point d'alignement : il joue un rôle de pivot disciplinaire

   - Exemple : Projet CollEx Bibracte ville ouverte
     - Un thésaurus pour Bibracte, aligné sur Pactols par des concepts-liaison
       (Ex. : céramique non tournée : dernier niveau dans Pactols et 16 TS dans Bibracte)
     - Un « Atelier thésaurus » (nov. 2022) pour la construction de thésaurus de projets scientifiques alignés sur Pactols

Point de réflexion

- Faut-il favoriser une « mode » thésaurus ?

Éparpillement des référentiels



Construction de solutions adaptées



Logique de structuration, normalisation, alignements

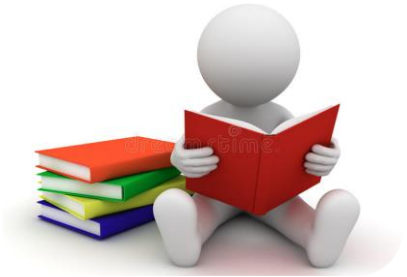


Enfin,

- Compter sur l'effet dissémination de bonnes pratiques et d'outils FAIR
- Jouer la transparence avec le collectif
  - Afficher les procédures
  - Préciser les calendriers
- Une répartition des rôles préservée



- expertise sur le vocabulaire et sur sa description
- participe à la logique hiérarchique



- contrôle la cohérence et la structuration,
- garantit la normalisation et l'interopérabilité,
- assure l'accessibilité et la pérennité



Merci de votre attention

Blandine Nouvel
Centre Camille Jullian | Frantiq

« Accompagnement à l'élaboration de vocabulaires contrôlés dans une démarche de FAIRisation des projets de recherche»
9e C@fé Renatis, 6 décembre 2022