

Enjeux de la qualité des référentiels et des métadonnées pour la communauté scientifique

Vocabulaires contrôlés, alignements pour une meilleure FAIRisation et répliquabilité d'un modèle de données - au travers de l'exemple des outils de Frantiq



Véronique HUMBERT, Archéologie des Sociétés Méditerranéennes (Université Paul Valéry-Montpellier 3, CNRS)
Blandine NOUVEL, Centre Camille Jullian (Aix Marseille Univ, CNRS)

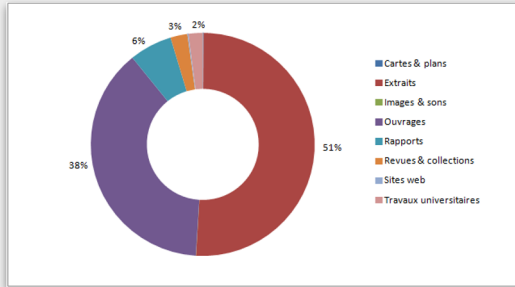
CNRS-MiTi, GT Atelier Données - Webinaire "Qualité des données", 5 juillet 2021

Frantiq, producteur de métadonnées

Catalogue collectif indexé CCI

<https://catalogue.frantiq.fr/>

40 bibliothèques de recherche
600000 notices



SIGB Koha
Unimarc
Identifiant ARK

Mots-clés 6XX
idNotice

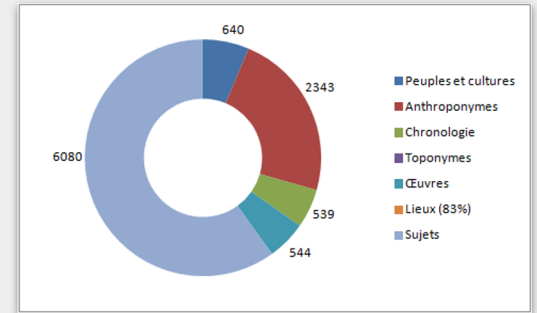
API de
Openetheso

PrefLabel
idConcept

Thésaurus PACTOLS

<https://pactols.frantiq.fr>

60000 concepts
multilingues
6 collections
thématiques



Programme de réorganisation 2017-2022

Gestionnaire Openetheso
SKOS, ISO 25964
Identifiant ARK

Catalogue collectif pour le confort du lecteur /1

1. Viser l'homogénéité des métadonnées pour la compréhension des notices par l'homme/la machine

- Cohérence et pertinence des données descriptives
 - Notices
 - Respect de la norme UNIMARC avec une indispensable harmonisation des pratiques
 - Contrôles/corrections réciproques au fil du catalogage
 - Autorités
 - Amélioration de l'indexation documentaire grâce au lien dynamique avec le thésaurus
 - Préparation des imports : homogénéité des autorités auteurs et autorités sujets (mots-clés)
 - Faciliter l'interopérabilité => vers Zotero
- Actions qualité
 - Interface de consultation : ergonomie de l'OPAC
 - Notices
 - Dédoublonnage récurrent
 - Programme ciblé Collections
 - Formations des catalogueurs + supports : flash et guides de catalogage

Catalogue collectif pour le confort du lecteur /2

2. Enrichir les métadonnées pour élargir l'accès à l'information

- passer de la description d'une ressource à la ressource elle-même
 - depuis 1984 :
 - la cote comme localisation physique du document dans la bibliothèque
 - PEB Frantiq
 - depuis 1990 : intégration de l'URL puis création des notices de documents numériques ([RDA-FR](#))
 - 2009 : essai infructueux de récupération des liens vers le document dans Persée
 - 30% de recouvrement
 - défaut d'enregistrement des titres
 - 2021 : intégration du lien vers le portail collaboratif de revues [Mir@bel](#)



Antiquités africaines [Revue & Collections] : l'Afrique du Nord de la Protohistoire à la conquête arabe / Université de Provence, Aix-en-Provence

PUBLICATION: Paris : CNRS Editions, 1967-
PÉRIODICITÉ: Annuel
ISSN: 2117-539X
LANGUE: Français.
PAYS: France (iso).
NOTE DE CONTENU : [Sommaires](#)
URL: [Accès libre \(Persée\)](#) - 1(1967) - (barrière mobile 2 ans)

Accès en ligne	Ressource	Accès	Numéros
Sommaire	Centre Camille julian	libre	1967 (Vol. 1) — ...
Intégral	OpenEdition Journals	libre	2016 (Vol. 52) — ...
Intégral	Persée	libre	1967 — 2015

[Voir cette revue dans](#)

Exemplaires (1) Notes Commentaires (0)

Site actuel	Collection	Cote	Info vol.	URL	Statut	Date de retour prévue
Accès en ligne Libre accès	En ligne	www (Parcourir l'étagère)	1(1967) - (barrière mobile 2 ans)	http://www.persee.fr/collection/antaf	Disponible	

Catalogue collectif pour le confort du lecteur /3

3. Encourager le lien dynamique avec le thésaurus pour rebondir sur d'autres ressources

<https://pactols.frantiq.fr/>

The screenshot shows the 'pactols.frantiq.fr' interface. On the left is a hierarchical tree of categories, with 'marbre' highlighted under 'roche métamorphique'. The main area displays the concept page for 'marbre (fr)'. It includes fields for 'Terme préféré', 'Collection', 'Concept générique', 'Concept spécifique', 'Concept associé', 'Synonyme', and 'Traduction'. The 'Corpus lié' section contains three links: 'Catalogue Frantiq (1521 notices)', 'Epicherhell (115 notices)', and 'Sculpturo (61 notices)'. The 'Identifiant / Lien permanent' section provides various URIs and ARK identifiers. At the bottom, it shows the creation and modification dates: 'Créé le: 2007-02-08 Dernière modification: 2014-05-12'.

<https://catalogue.frantiq.fr>

The screenshot shows the search results page on 'catalogue.frantiq.fr'. It features a search bar at the top, a list of search results with various filters, and a detailed view of a selected result. The interface is clean and professional, with a clear navigation menu.

<http://cçi-epicherhell.huma-num.fr/>


The screenshot shows the search results page on 'cçi-epicherhell.huma-num.fr/'. It displays '115 RÉSULTATS DE RECHERCHE' in a grid format. Each result includes a small thumbnail image and a title. The page is designed to be user-friendly and easy to navigate.

<http://sculpturo.huma-num.fr/>

The screenshot shows the search results page on 'sculpturo.huma-num.fr/'. It displays '61 RÉSULTATS DE RECHERCHE' in a grid format. Each result includes a small thumbnail image of a sculpture and a title. The page is designed to be user-friendly and easy to navigate.

L'évolution du thésaurus PACTOLS

Exploiter le potentiel du thésaurus pour construire un outil plus ouvert et mieux adapté aux besoins d'enrichissement des données de la recherche

➔ Programme de réorganisation 2017-2022 soutenu par le consortium 

Interopérabilité

Import/exports : CSV, SKOS, RDF, JSON-LD
API de type REST
1 Identifiant pérenne ARK par concept

Structure


Consolider la structuration SKOS (relations sémantiques hiérarchiques, associatives et d'équivalences)

Lexique

Réorganisation des domaines thématiques inspirée du [BackBoneThesaurus](#) de DARIAH (aligné sur le CIDOC-CRM)

Validation scientifique des concepts

Mettre à jour et faire valider un vocabulaire scientifique d'assise inégale selon les domaines, avec des lacunes, des termes vieillis



Les experts des domaines interviennent pour enrichir le vocabulaire, le structurer et définir les concepts

- 30 séances d'ateliers
- programmes MASA, CollEx Persée
- +40 chercheurs et ingénieurs impliqués

Domaines validés ou en cours

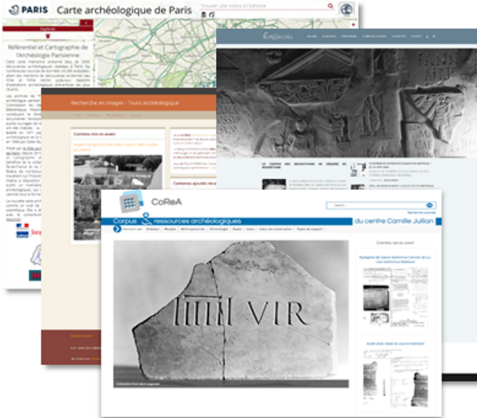
- archéologie et anthropologie funéraire
- industries lithique et osseuse
- religion et mythologie du monde méditerranéen antique
- céramologie
- pétrographie
- protohistoire européenne
- philosophie antique

PACTOLS, référentiel terminologique pour l'archéologie



Enregistrement
de fouilles
archéologiques

Bibliothèques

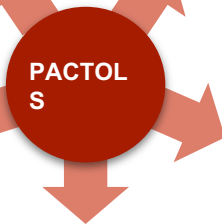


Corpus de
recherche et
patrimoine

Portails enrichis



Publications
scientifiques



Stratégie d'alignements

- un processus ancien, soutenu par l'évolution de Opentheso
 - concerne le thésaurus et plus récemment le CCI
 - des choix opportunistes selon les programmes et collaborations
- objectif affiché de s'inscrire dans le web des données en liant les métadonnées
- mais aussi :
 - enrichir réciproquement les référentiels
 - consolider ses propres métadonnées en les confrontant aux autres

2019 & 2020

IdRef 

PACTOLS : programmes CollEx [RefDivinités](#) avec l'Abes et la BIS et [ArchéoAI](#) avec l'IFAO, [Bibracte ville ouverte](#) avec le CAE-Bibracte
CCI : Autorités auteurs avec l'Abes

fin 2017

Wikidata 

[Projet PACTOLS/Wikidata](#), avec le support de Wikimedia France (Sylvain Boissel)

2017

Geonames  GeoNames

en prévision de l'abandon des Lieux, avec l'Inist

2015-2016

Art and Architecture Thesaurus - Getty 





Dans le cadre d'infrastructure européenne [ARIADNE](#), avec l'Inrap

2014

Fonction Alignements de Opentheso

Premiers essais


La FAIRisation progressive des métadonnées de Frantiq

	F _{indable} 	A _{ccessible} 	I _{nteroperable} 	R _{eusable} 
Applications	Minitel SUNIST Texto Application propriétaire	Sites web PHP-MySQL	CSD-ISIS Koha Opentheso (phase 1)	Koha Opentheso (phase 2)
Fonctionnalités	Recherche par mots clés Recherche par terme thésaurus		Format Unimarc Norme ISO-25964 SKOS / JsonLd / Turtle import /export libres	URI Ark Alignements Licences <u>ODbL v1.0</u> et <u>GNU GPL</u>

Donc d'après notre expérience,

- Qu'est-ce qu'une méta-donnée de qualité ?

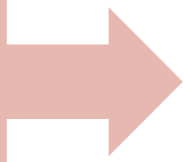
- sa structure est **conforme** à une norme/standard
=> elle est interopérable, circule d'un système à l'autre sans perdre de sa valeur formelle
- sa licence est **ouverte**
=> elle est libre d'usage dans le respect du droit d'auteur
- sa valeur sémantique est **fiable**
=> elle est contrôlée et approuvée par l'humain



elle est **réutilisable**,
techniquement et en
toute confiance

- Quelle organisation mettre en place ?

- seul et/ou ensemble,
 - se conformer à des règles,
 - utiliser des outils normalisés et ouverts,
 - documenter ses choix,
 - remettre sans cesse en jeu les expériences, dans des processus itératifs
- associer les compétences : documentaire-informatique-scientifique



elle devrait être
rigoureuse et **exigente**

Merci !

veronique.humbert@frantiq.fr

blandine.nouvel@frantiq.fr

<https://frantiq.fr>

