



HAL
open science

Réviser les Lieux du thésaurus Pactols, entre pragmatisme et logique sémantique

Clara Lelièvre, Blandine Nouvel

► To cite this version:

Clara Lelièvre, Blandine Nouvel. Réviser les Lieux du thésaurus Pactols, entre pragmatisme et logique sémantique. Atelier du GT du consortium DISTAM (DIGital STudies Africa, Asia, Middle East), DISTAM, Oct 2023, En ligne, France. hal-04261204

HAL Id: hal-04261204

<https://amu.hal.science/hal-04261204>

Submitted on 26 Oct 2023

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.

Réviser les Lieux du thésaurus Pactols, entre pragmatisme et logique sémantique



Clara Lelièvre (Consortium MASAPlus & Frantiq)
Blandine Nouvel (Centre Camille Jullian, CNRS-Aix Marseille Univ & Frantiq)

DISTAM, Séminaire du GT3, Jeudi 26 octobre 2023





Le thésaurus Pactols

- Né en 1987 pour rationaliser les mots-clés du tout jeune réseau des bibliothèques d'archéologie Frantiq
 - normalisé
 - multilingue
- Années 2010 : le contexte scientifique et numérique (web sémantique, SO) impose une révision
- 2017-2023 : programme de réorganisation du Consortium 
 - nouvelle version des Sujets en septembre 2022
 - réorganisation des Lieux en 2023



Les spécificités d'un thésaurus

Outil ou langage documentaire

- composé de mots-clés/concepts relatifs à un domaine de la connaissance,
- contrôlés, pour contourner les ambiguïtés du langage naturel et constituer un socle terminologique partagé
- utilisé pour permettre la recherche documentaire dans des systèmes informatiques

Caractérisé par des relations sémantiques riches

- au sein du thésaurus
 - des relations de **dépendance** entre un concept “père” et ses “fils” qui marquent l'organisation hiérarchique
 - des **associations** de concepts entre les branches
 - des groupes d'**appartenance** : collections et facettes
- en dehors du thésaurus
 - des **alignements** qui rapprochent les concepts avec les items d'autres référentiels du web

Les Lieux comme mots-clés

Indispensables pour l'indexation
des ressources en archéologie



Sujet

Temps

Espace

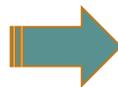
Pourquoi Pactols
quand il existe Geonames et Pleiades ?

- un seul référentiel pour renseigner Sujet/Temps/Espace
- évolutif et participatif
- interopérable (Webservices)
- visibilité des ressources indexées (corpus liés, alignements)

Les concepts de Lieux dans Pactols

PACTOLS v.1

- > Anthroponymes
- > Chronologie
- ✓ Lieux
 - > Afrique
 - > Afrique hydrographie
 - > Afrique physique
 - > Amérique
 - > Anciennes nomenclatures
 - > Asie
 - > Asie hydrographie
 - > Asie physique
 - > Australie
 - > Europe
 - > Europe hydrographie
 - > Europe physique
 - > Lieux imaginaires
 - > Monde antique
 - > Monde médiéval
 - > Océans
- > Oeuvres
- > Peuples
- > Sujets



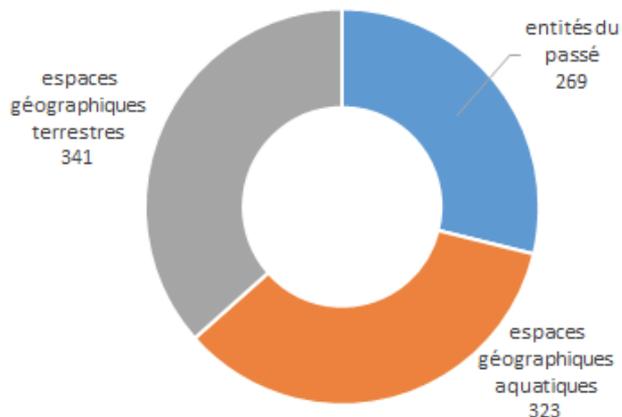
Pactols v. 2 - Lieux

- ✓ entités actuelles
 - > Afrique
 - > Amérique
 - > Asie
 - > Australie
 - > Europe
 - > Océanie
 - ✓ entités du passé
 - > entités du passé récent
 - > Monde antique
 - > Monde médiéval
 - > espaces aquatiques
 - > espaces terrestres
 - > sites archéologiques
 - > ~[termes dépréciés]
- 50283 concepts
- 269 concepts
- 323 concepts
- 341 concepts

Répartition des 51216 concepts Lieux

(sept. 2023)

Répartition des branches Lieux,
sauf entités géopolitiques actuelles



entités du passé

- entités du passé récent
- Monde antique
- Monde médiéval

espaces aquatiques

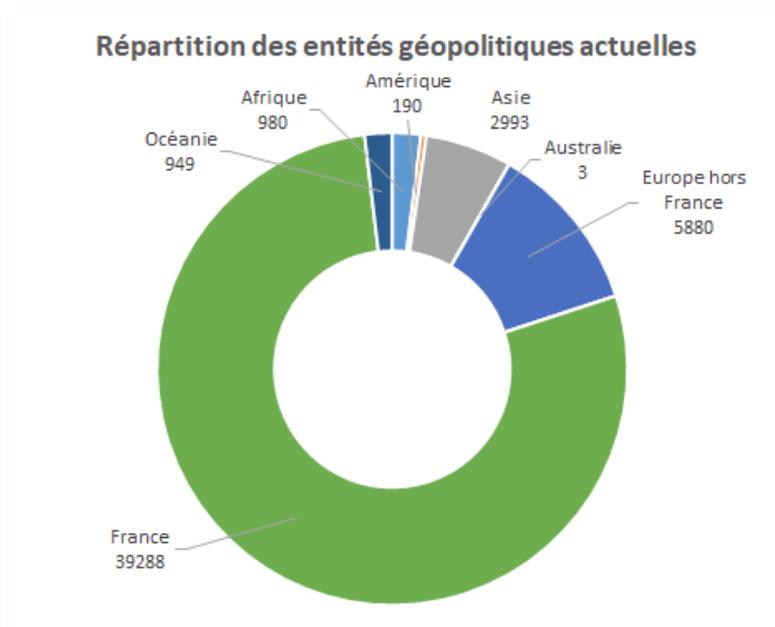
- bassins hydrographiques
- deltas
- détroits
- fleuves
- lacs
- mers
- océans
- rivières
- Afrique hydrographique
- Amérique hydrographique
- Asie hydrographique
- Europe hydrographique

espaces terrestres

- archipels
- déserts
- espaces immergés
- îles
- montagnes et massifs montagneux
- péninsules
- plages et littoraux
- plaines et plateaux
- vallées
- Afrique géographique
- Amérique géographique
- Asie géographique
- bassin méditerranéen
- Europe géographique
- régions polaires

Répartition des 51216 concepts Lieux

(sept. 2023)



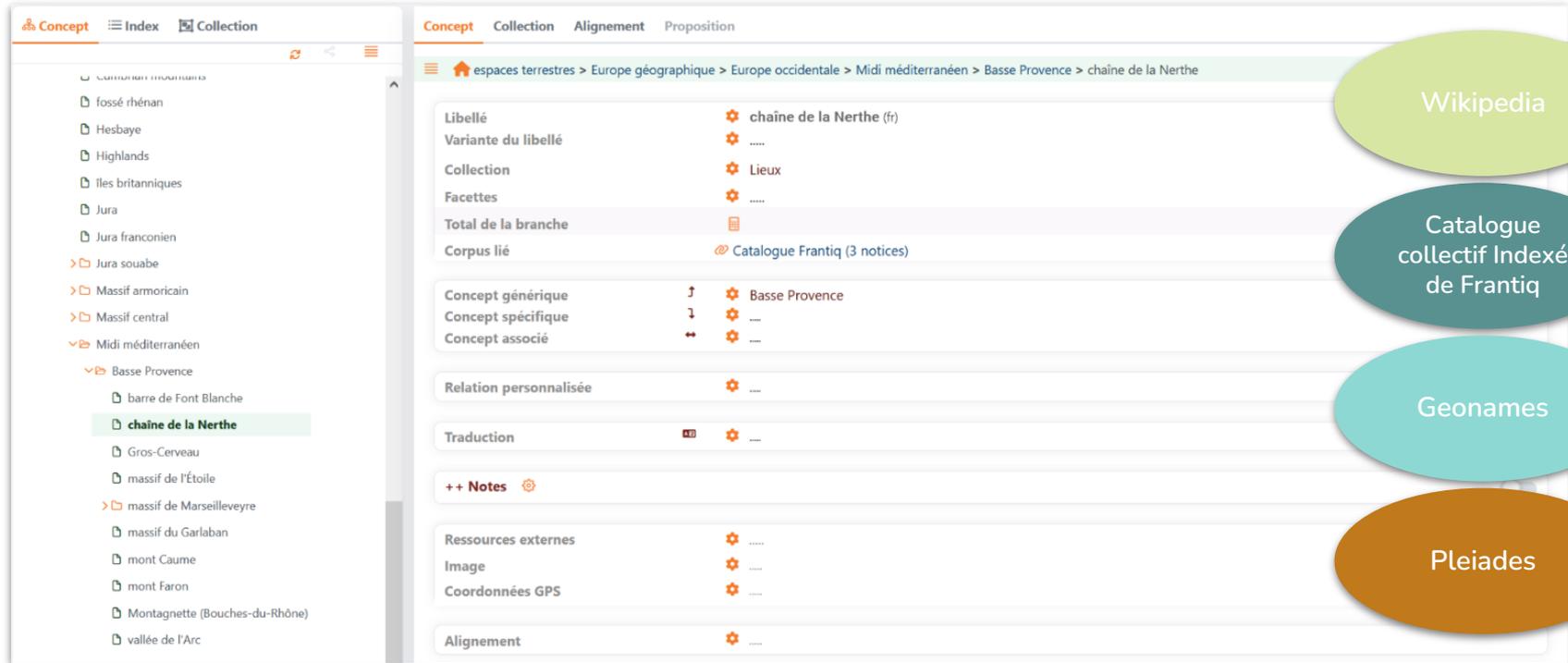
entités actuelles

- > Afrique
- > Amérique
- > Asie
- > Australie
- > Europe
- > Océanie

Europe

- > regroupements géopolitiques d'États européens
- > Albanie
- > Allemagne
- > Andorre
- > Autriche
- > Belgique
- ▣ Biélorussie
- > Bosnie-Herzégovine
- > Bulgarie
- > Chypre
- > Croatie
- > Danemark
- > Espagne
- ▣ Estonie
- > Finlande
- > France

Comment identifier les lieux ?



Concept Index Collection

- Communions montagnards
- fossé rhénan
- Hesbaye
- Highlands
- îles britanniques
- Jura
- Jura franconien
- Jura souabe
- Massif armoricain
- Massif central
- Midi méditerranéen
 - Basse Provence
 - barre de Font Blanche
 - chaîne de la Nerthe**
 - Gros-Cerveau
 - massif de l'Étoile
 - massif de Marseilleveyre
 - massif du Garlaban
 - mont Caume
 - mont Faron
 - Montagnette (Bouches-du-Rhône)
 - vallée de l'Arc

Concept Collection Alignement Proposition

espaces terrestres > Europe géographique > Europe occidentale > Midi méditerranéen > Basse Provence > chaîne de la Nerthe

Libellé chaîne de la Nerthe (fr)
Variante du libellé

Collection Lieux
Facettes

Total de la branche
Corpus lié Catalogue Frantiq (3 notices)

Concept générique ↑ Basse Provence
Concept spécifique ↓

Concept associé ↔

Relation personnalisée

Traduction

++ Notes

Ressources externes

Image

Coordonnées GPS

Alignement

Wikipedia

Catalogue collectif Indexé de Frantiq

Geonames

Pleiades



L'écriture des libellés

- **Caractérisation des espaces géographiques**
 - concept “Arc” renommé “vallée de l’Arc” après étude de son utilisation dans le CCI
 - concepts “bassin de la Loire”, “Loire (fleuve)” et “Maine (Loire)”
- **Renommage des concepts**
 - Précision chronologique
“Monde arabe” (branche “Monde médiéval”) renommé “Monde arabe médiéval”
 - Suppression des abréviations
I = île (Délös I = île de Délös et non Délös 1),
D = désert, Gov et Gouv, M = mer, Mt = mont, montagne, Val = vallée
- **Nom simple vs nom composé**
 - théâtre d’Orange | Colisée ; variante : amphithéâtre Flavien | Port de Délös | grotte de Lascaux
 - monument de granit (Délös) | Maison I de l’îlot des Bijoux (Délös)



Gestion des noms de sites

- **Noms anciens (grecs, latins, ...) notés en variante du site actuel**
 - permet la recherche sur le nom ancien
 - évite la multiplication des concepts et simplifie l'indexation
ex. : Marseille + Massilia, Massalia
 - épargne les difficultés de positionnement géographique
- **Noms des sites archéologiques**
 - placés en spécifiques de l'entité administrative actuelle
 - Polyhiérarchie autorisée pour les sites étendus
ex. Bibracte TS de Glux-en-Glenne, Larochemillay, Saint-Léger-sous-Beuvray
 - Rattachés à des facettes spécifiques
 - par type de site archéologique
 - par pays
- Association entre les **noms de sites**, les noms **géographiques** et les noms d'**entités du passé** et d'**entités administratives actuelles**
 - ex. **Bibracte**
TG Glux-en-Glenne, Larochemillay, Saint-Léger-sous-Beuvray
TA Bibracte EPCC, **Gaule Lyonnaise**, **mont Beuvray**, **Morvan**

Articulation Sujets/Lieux

- Les termes dits “de regroupement”:
 - dépréciés au profit d’une facette. ex. Océans devient F-océans
 - transférés dans le thésaurus Sujets s’ils peuvent y trouver une place ex. “provinces romaines” et “colonies phéniciennes”

Lieux > Monde antique

✓ Provinces romaines

- > Gaule
- > Possessions républicaines
- > Provinces impériales
- > Provinces sénatoriales

Lieux > Monde antique

✓ Monde romain

- > provinces impériales
- > provinces républicaines
- > provinces sénatoriales

Sujets > Unités géopolitiques

- ✓ unités géopolitiques
 - > divisions ecclésiastiques
 - > unités territoriales administratives
 - ✓ région
 - ✓ province
 - éparchie
 - ✓ province romaine
 - province impériale
 - province républicaine
 - province sénatoriale

Articulation Sujets/Lieux

- La branche “sites archéologiques “ du thésaurus Lieux :

++ Notes 

Définition  Lieu ou unité bâtie identifié comme site archéologique ou ayant fait l'objet de fouilles archéologiques au cours desquelles ont potentiellement été mis au jour des vestiges de l'activité humaine, géolocalisable dans son emplacement d'origine.

Note d'application  Les vestiges déplacés, considérés comme des artefacts ou œuvres artistiques, sont nommés dans le thésaurus Sujets/Entités nommées/œuvres artistiques.
Les anciennes appellations de villes actuelles ne sont pas considérées comme des sites archéologiques et ne figurent pas dans cette branche.

Thésaurus Sujets

o uvres architecturales nommées

- ▢ Ara Pacis
- ▢ Arc d'Orange
- ▢ Arc de Trajan
- ▢ Autel de Pergame
- ▢ Colisée
- ▢ Colonne trajane
- ▢ Maison Carrée



Lieux > entités actuelles > ...

Rome

- ▢ Ara Pacis
- ▢ Aventin
- ▢ Caelius
- ▢ Capitole
- ▢ Colisée
- ▢ Colonne trajane



Lieux > F-sites archéologiques

lieux de culte et sanctuaires

- ▢ Ara Pacis
- ...

lieux de jeux (théâtres, cirques, arènes)

- ▢ Colisée
- ...

sites archéologiques d'Italie

- ▢ Ara Pacis
- ▢ Colisée

entités actuelles > Europe > France > France métropolitaine > Bourgogne-Franche-Comté > Département de Saône-et-Loire > Saint-Léger-sous-Beuvray > entités actuelles > Europe > France > France métropolitaine > Bourgogne-Franche-Comté > Département de la Nièvre > Glux-en-Glenne > Bibracte

Sites archéologiques placés en TS de la commune à laquelle ils sont rattachés "Bibracte" : un exemple de polyhiérarchie

Libellé Bibracte (fr)
Variante du libellé oppidum de Bibracte
Collection Lieux
Facettes oppida sites archéologiques de France
Total de la branche
Corpus lié Catalogue Frantiq (883 notices)

Création de facettes par types de site archéologique et par pays

Concept générique Glux-en-Glenne. Larochemillay Saint-Léger-sous-Beuvray
Concept spécifique
Concept associé Bibracte EPCC. Gaule Lyonnaise. mont Beuvray. Morvan

Ajout de concepts associés afin de lier les sites archéologiques aux espaces géographiques et aux entités du passé

Relation personnalisée

Traduction Bibracte (de)
 Bibracte (en)
 Bibracte (es)
 Bibracte (it)
 Bibracte (nl)

++ Notes

Définition Oppidum du IIème et Ier siècle avant notre ère, ancienne capitale des Éduens, situé sur les communes de Saint-Léger-sous-Beuvray, Glux-en-Glenne et Larochemillay. (d'ap. article "Bibracte" de Wikipédia en français, <https://fr.wikipedia.org/wiki/Bibracte>)

Définition, note d'application

Ressources externes
Image
Coordonnées GPS

Alignement exactMatch : <https://ark.mom.fr/ark:/39676/bib133xc31wdx>
exactMatch : <https://fr.wikipedia.org/wiki/Bibracte>

Alignements : closeMatch pour les anciennes appellations (ex: Marseille - Massalia)

Les facettes des sites archéologiques

- ✓ sites archéologiques
 - > ateliers et usines
 - > camps et fortifications
 - > grottes et sites rupestres
 - > hospices, thermes et bains publics
 - > lieux d'échanges (forums, marchés, halles)
 - > lieux de culte et sanctuaires
 - > lieux de jeux (théâtres, cirques, arènes)
 - > maisons et villas
 - > monuments commémoratifs (arcs, colonnes, portes)
 - > nécropoles et tombes
 - > occupations anciennes
 - > oppida
 - > palais et châteaux
 - routes et voies

- > sites archéologiques d'Allemagne
- > sites archéologiques d'Angleterre
- > sites archéologiques d'Autriche
- > sites archéologiques d'Azerbaïdjan
- > sites archéologiques de Bulgarie
- > sites archéologiques de Crimée
- > sites archéologiques de France
- > sites archéologiques de Grèce
- > sites archéologiques de Hongrie
- > sites archéologiques de Jordanie
- > sites archéologiques de Libye
- > sites archéologiques de Palestine
- > sites archéologiques de Roumanie
- > sites archéologiques de Serbie
- > sites archéologiques de Slovénie
- > sites archéologiques d'Espagne

Bien indexer : le choix du concept précis

Lieux > espaces
terrestres

Anatolie

Définition : Zone géographique comprenant les territoires actuels situés à l'ouest d'une ligne Tchorkhi-Oronte, qui va de la Méditerranée à la mer Noire, et qui sont séparés de l'Europe par la mer de Marmara.

Note d'application : Utiliser le concept "Anatolie" pour les territoires actuels et le concept "Asie mineure" pour les entités du passé.

Lieux > entités du passé

Asie mineure

Définition : Zone géographique antique située à l'extrémité occidentale de la péninsule anatolienne, la façade égéenne située entre le Bosphore et la Lycie et par extension, englobe également l'Anatolie actuelle, la partie asiatique de la Turquie.

Note d'application : Utiliser le concept "Asie mineure" pour les entités du passé et le concept "Anatolie" pour les territoires actuels.

Lieux > entités actuelles

Péloponnèse (périphérie)

Définition : Région administrative de Grèce ou périphérie comprenant cinq des sept districts régionaux de la péninsule du Péloponnèse.

Note d'application : À différencier du concept "péninsule du Péloponnèse" qui est un espace géographique et qui ne couvre pas la même étendue que la périphérie du Péloponnèse.

Lieux > espaces
terrestres

péninsule du Péloponnèse

Définition : Espace géographique désignant la péninsule du Péloponnèse, bordée par la mer Ionienne et la mer Égée et reliée à la Grèce continentale par l'isthme de Corinthe.

Note d'application : À différencier du concept "Péloponnèse (périphérie)" qui est une région administrative de Grèce et qui ne couvre pas la même étendue géographique que la péninsule du Péloponnèse.



Le programme aujourd'hui et demain



- Réorganisation des branches
- Création des facettes
- Déplacement des concepts



- Remplacement des abréviations par des libellés clairs
- Enrichissement des branches hiérarchiques
 - mettre à jour des "entités géopolitiques actuelles"
 - continuer d'ajouter les sites archéologiques aux facettes correspondantes (types/pays)



- Réindexation du CCI
- Alignements et géolocalisations
 - Wikidata, sur le modèle semi-automatique mis en place pour les Sujets
 - incluera Geonames et Pleiades avec les géolocalisations