



HAL
open science

Deconstructing for reconstructing: The use of the BackBone Thesaurus for reorganising the PACTOLS Thesaurus

Blandine Nouvel, Evelyne Sinigaglia, Véronique Humbert

► To cite this version:

Blandine Nouvel, Evelyne Sinigaglia, Véronique Humbert. Deconstructing for reconstructing: The use of the BackBone Thesaurus for reorganising the PACTOLS Thesaurus. FAIR Heritage: Digital Methods, Scholarly Editing, and Tools for Cultural and Natural Heritage, Consortium MASA, mémoire des archéologues et des sites archéologiques; Programme Intelligence des Patrimoines (ARD 2020), Jun 2020, En ligne, France. halshs-02890892

HAL Id: halshs-02890892

<https://shs.hal.science/halshs-02890892v1>

Submitted on 6 Jul 2020

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.

Deconstructing for reconstructing : the use of the BackBone Thesaurus for reorganising the PACTOLS Thesaurus

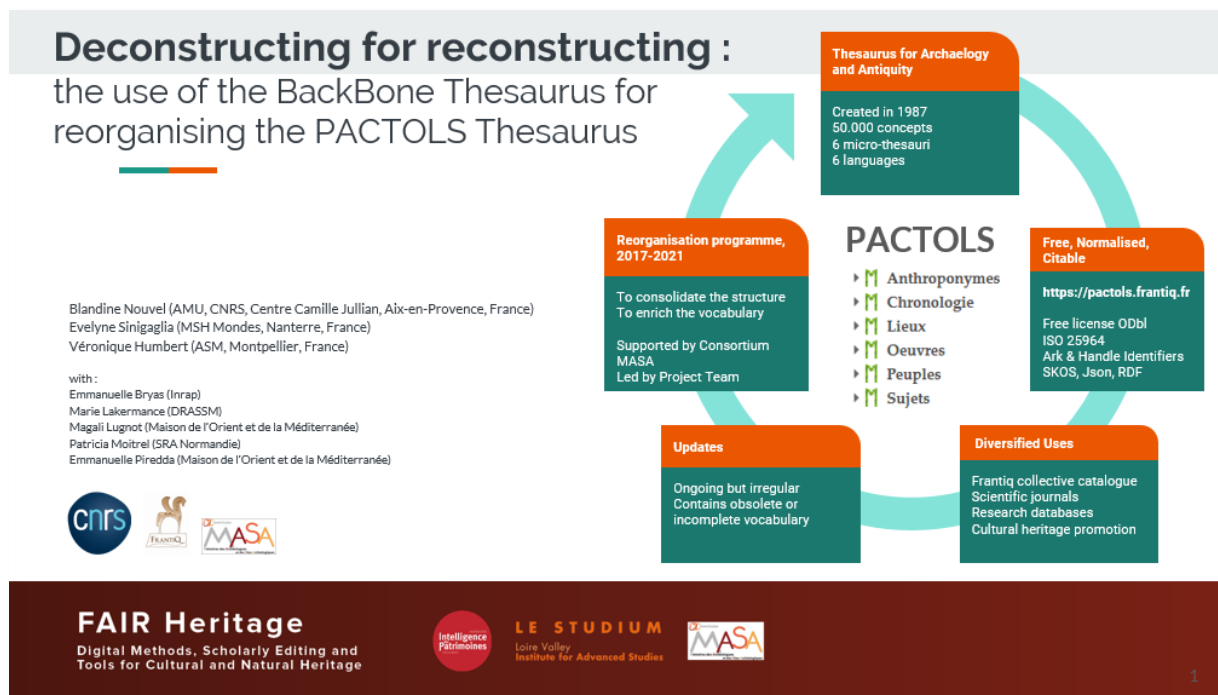
Blandine Nouvel (Aix Marseille Université, CNRS, CCJ, Aix-en Provence, France)

Evelyne Sinigaglia (MSH Mondes, Nanterre, France)

Véronique Humbert (ASM, Montpellier, France)

[FAIR Heritage](#), 17-18 juin 2020

Speaker : B. Nouvel



Good afternoon everybody,

My name is Blandine Nouvel and I am the leader of the PACTOLS reorganisation project team.

First I would like to thank the organisers of this session for giving me the opportunity to present our methodology for reorganising the PACTOLS thesaurus.

* PACTOLS is a thesaurus that covers all fields of archaeology and sciences related to Antiquity. Its 50 000 (fifty thousand) concepts are organised in 6 micro thesauri whose names form the acronym PACTOLS : Peuples/Peoples, Anthroponymes/Anthroponyms, Chronologie/Chronology, Œuvres/Works, Lieux/Places, Sujets/Subjects. Toponyms are a separate list which does not appear in the thesaurus.

- The PACTOLS thesaurus is managed using the free thesaurus management software Opentheso. It is multilingual (6 languages),
- normalised (in ISO and SKOS), polyhierarchical, citable (using ARK) and under free license.
- It meets the FAIR principles for findability, accessibility, interoperability and reusability.

- Its thematic coverage and technical qualities make it the only metadata repository for archaeology in France.

It is used for indexing the bibliographic records of the Frantiq Collective Catalogue and more recently has served to enrich the contents of scientific journals, databases for archaeological research, and cultural heritage programmes.

- Yet in spite of regular updates since its creation in 1987, the thesaurus sometimes presents significant gaps and outdated vocabulary.
- Our desire to increase its visibility and use among new users, editors and researchers for example, has led us to launch an extensive reorganisation programme aiming at consolidating the structure of PACTOLS and enriching its vocabulary.

The reorganisation process started at the end of 2017 supported by the MASA Consortium. It employs a team of 8 volunteers of the Frantiq network including the authors of this presentation.

The structural support of the BBT

DARIAH Meta-thesaurus for the Humanities,
Thesaurus Maintenance Working Group

https://vocabs.dariah.eu/backbone_thesaurus/

- Conceptual, simplified structure
- Inspired by the CIDOC-CRM
- Collaborative and evolving

How PACTOLS uses the BBT

- By creating a temporary BBT tree with additional lower levels
- By distributing PACTOLS concepts in the BBT categories, field by field
- By reorganising the contents systematically
- By correcting structural inconsistencies and confusion of categories in the branches
- By revising the entire vocabulary

Backbone Thesaurus

Alphabetical	Hierarchy	Groups
+ 000001 activities		
- 000010 disciplines		
- 000011 human interactions		
- 000012 intentional destructions		
- 000013 functions		
+ 000002 natural processes		
- 000015 natural disasters		
- 000016 geneses		
- 000003 materials		
+ 000004 material things		
- 000017 mobile objects		
- 000018 built environment		
- 000019 physical features		
- 000020 structural parts of material objects		
- 000005 types of epochs		
+ 000006 conceptual objects		
- 000021 symbolic objects		
- 000022 propositional objects		
- 000023 methods		
- 000024 concepts		
- 000007 groups and collectivities		
+ 000008 roles		
- 000025 offices		
- 000026 roles of interpersonal relations		
- 000009 geopolitical units		
+ 0000049 geometric extents		
- 0000050 points		
- 0000051 surface areas		
- 0000052 3d-volumes		
- 0000053 linear extent		

A first attempt at reorganising the fields in a traditional thematic way concentrated on the micro thesaurus Subjects. This attempt eventually appeared to be only partial and to have little effect on the general structure of the thesaurus.

So we chose to follow the structure of the BackBone Thesaurus (BBT), the meta-thesaurus developed by DARIAH the European infrastructure for the Humanities. The authors of this talk then joined the Steering Committee for the BBT.

* 3 essential reasons led us to this choice :

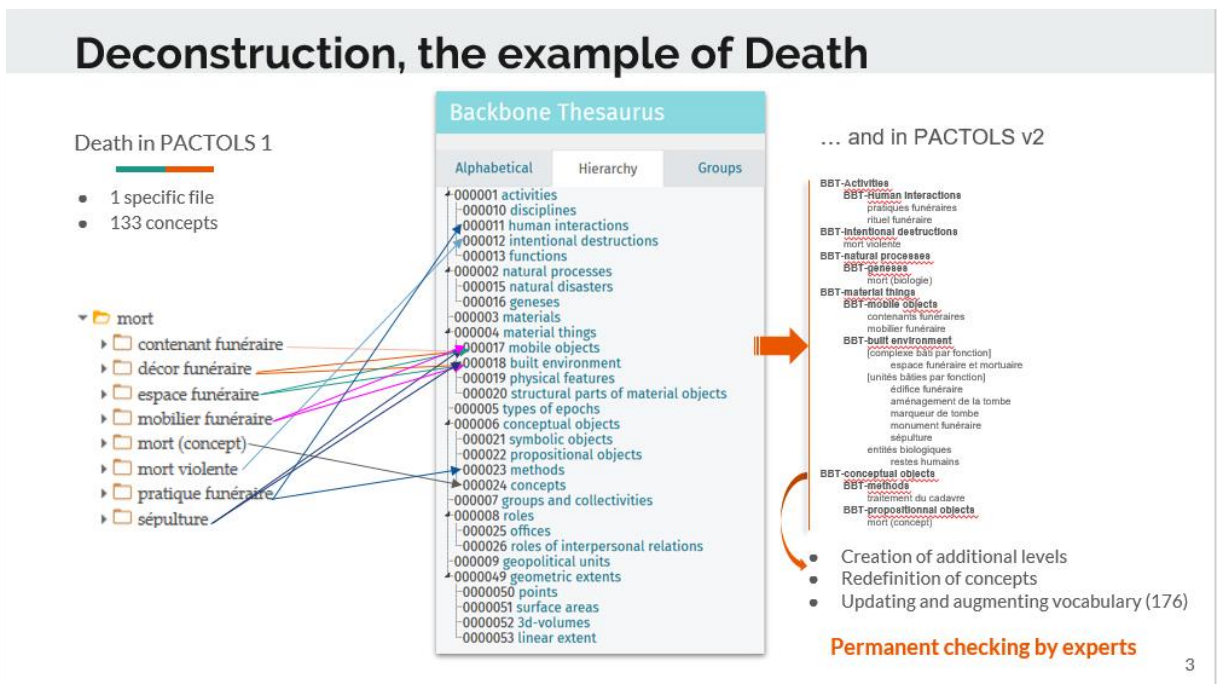
- First : its structure is conceptual and simplified to 2 levels (potentially 3). It facilitates comprehension and exchanges with other disciplines (architecture, art, biology, physics). The BBT is already a platform for alignment with partners like the AAT and the DAI

- Second : its structure is based on the CIDOC-CRM and is thus compatible with open information systems
- And finally : it is evolving (we are working on a 3rd level, the principle of which has been approved by the steering committee). It is also collaborative (we have supplied the French translation and the scope notes of the first two levels and submitted the creation of new facets : living beings and supranatural beings)

* This brings us to how PACTOLS uses the BBT :

- We created a temporary tree with the BBT categories, completed by new levels. We are progressively redistributing all the PACTOLS concepts, field by field starting with the Subjects thesaurus
- It has led to the dispersion of the contents and their reorganisation using a different approach
- This process has allowed us to correct structural inconsistencies and confusion of categories in the branches

- It has given us the opportunity to update the whole vocabulary, which is one of the goals of the programme. The process focused in particular on the definitions and translations of concepts making for an in-depth work on the terminology



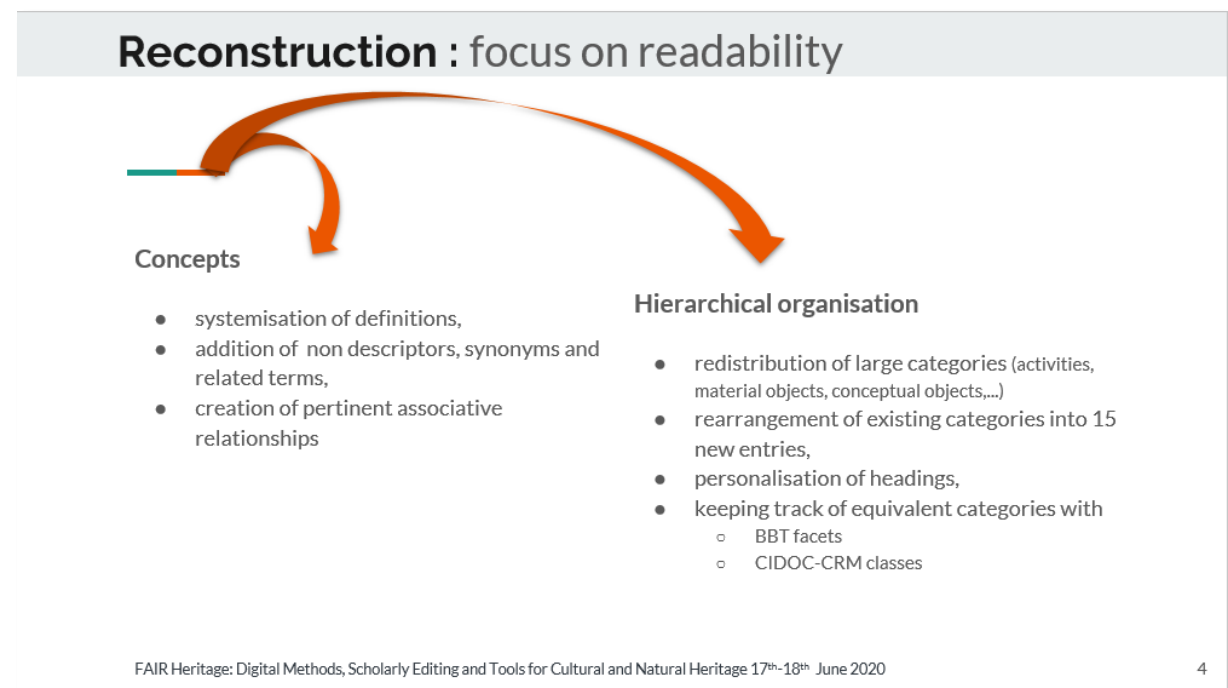
Let us take the example of the thematic field we have just finished ; the Death file. We have worked on a temporary structure containing a copy of the original PACTOLS, PACTOLS 1 and the BBT Facets

* In the original PACTOLS, the Death file included 133 concepts in 8 sub-files. Distribution of the terms was prepared on a shared spreadsheet with the BBT categories completed if necessary with additional hierarchic levels. Once validated, that tree was entered in the PACTOLS 2 until the Death file was emptied.

* The concepts of the Death hierarchy are now divided among 5 new BBT categories and 11 new subdivisions which are better integrated into the other domains of the thesaurus.

- The revision process put an emphasis on the vocabulary : outdated terms were either renamed or merged, and definitions were updated with recent bibliographic sources. The revision of certain terms will lead us in the future to correct the indexing of the Frantiq Collective Catalogue to keep its connection with the PACTOLS thesaurus.

We have worked under the control of funerary archaeologists and anthropologists eager to be involved in the creation of a terminology they had been trying to formalise for many years.



Reconstructing the thesaurus cannot be considered as long as all the categories of the micro Subjects thesaurus have not been filtered through the BBT.

We can however already distinguish several stages aiming at providing good readability of the contents. Although we need to continue working on the concepts (for example, systemise definitions, add synonyms, non-descriptors and increase the number of related terms),

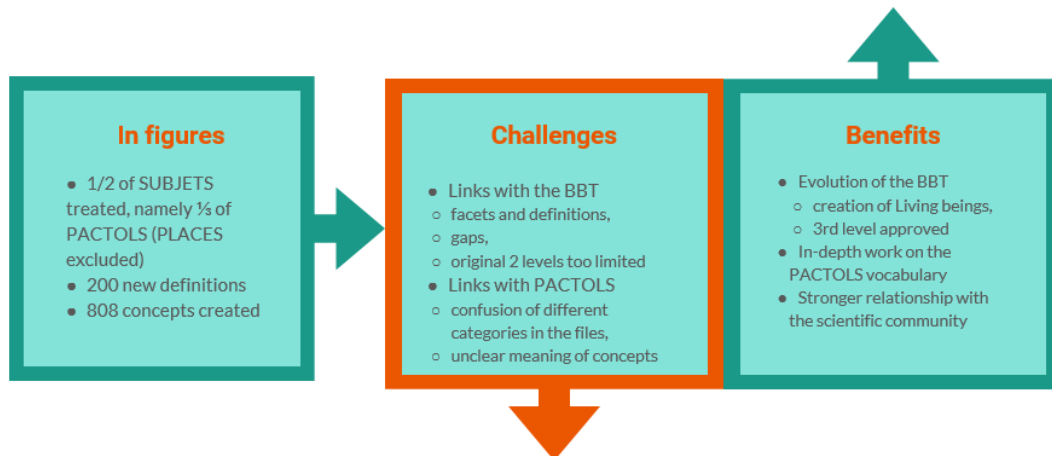
we will focus on the general organisation of the hierarchies.

We will :

- Disperse the largest categories among more accessible files thus limiting the depth and number of levels. This will provide a more direct access to the contents (1868 material entities > 759 built environment ; 1171 mobile objects)
- Add approximately 15 fifteen new top terms (that is terms on the 1st level) merging current PACTOLS micro thesauri and BBT facets. This will eliminate the 6 former PACTOLS micro thesauri and introduce the new categories : agents, built environment, chronology, disciplines, space and territory, living organisms (flora and fauna), languages, materials, mobile objects, works of the mind, society and techniques...

- Personalise headings in clear and understandable language for all users
- Ensure that we keep track of the equivalent files and sub-files of the BBT Facets and CIDOC CRM classes

A first assessment



A first assessment can be outlined after 18 months of solid work on the 2nd version of PACTOLS

Here are a few figures :

- 2800 concepts of the Subjects have been redistributed, that is one third of the PACTOLS micro thesauri
- More than 200 new definitions have been given, in addition to definitions that have been revised
- More than 808 concepts from the lists of candidate terms suggested by librarians, and from scientific vocabularies (anthropology and funerary archaeology, lithic and bone industries, architecture, languages...) have been added

Our main difficulties were :

- That it took us almost 6 months to understand in depth the BBT facets and their definitions. This caused work to start very slowly. Concepts were missing in the BBT (for example Living beings) and the number of levels was too limited to include the PACTOLS subdivisions. We had to create temporary facets (for example built unit by function, built unit by form)
- The other difficulties were linked to PACTOLS mixing different categories in the files. For instance in the Death file one could find actions, material objects and conceptual objects. And above all there was the ambiguity of some concepts whose meaning had to be clarified while keeping the indexation of the Frantiq catalogue bibliographic records

The results, though, are vastly positive :

- The BBT is evolving : our participation in the BBT working Group allowed us to suggest a new facet : « Living Beings », to propose the creation of a third level and to submit a French translation of the BBT Facets and definitions
- That work impacts all the PACTOLS categories with an in-depth revision of the whole vocabulary
 - our relationship with the scientific community has been strengthened through participating workshops and permanent exchanges on scientific vocabulary.

Thank you very much to all !