

# Thesaurus Maintenance Working Group

## DARIAH Open Meeting 2021<sup>1</sup>

18 November 2021

### Participants List:

**Aming Shi** (NTNU Universitetsbiblioteket); **Ana Salgado** (NOVA CLUNL); **Andrej Gogora** (IESA SAS); **Annika Kirscheneder** (DAI); **Beatriz Garrido** (UNED); **Blandine Nouvel** (CNRS); **Bruno Almeida** (ROSSIO Infrastructure); **Bruno Batllou** (Ville de Nîmes); **Camilla Colombi** (DAI); **Carolina Corral Trullench** (UNED); **Christelle Molinié** (Musée Saint-Raymond ; musée d'Archéologie de Toulouse); **Christos Georgis** (FORTH); **Chryssoula Bekiari** (FORTH-ICS); **Cristina Faria** (Teatro Nacional D. Maria II); **Dan Matei** (Institutul Național al Patrimoniului); **Dominic Oldman** (British Museum); **Eleni Tsouloucha** (FORTH); **Flora Sanz Calama** (UNED); **Gerald Hiebel** (University of Innsbruck); **Gerasimos Chrysovitsanos** (Academy of Athens); **Helen Goulis** (Academy of Athens); **Helen Katsiadakis** (Academy of Athens); **Ifigenia Dionissiadou** (Benaki Museum); **Ivana Dević** (University of Split Library); **Lida Harami** (FORTH – CCI); **Margarida Ramos** (NOVA CLUNL); **Maria Papadopoulou** (USMB); **Massimiliano Carloni** (Austrian Centre for Digital Humanities and Cultural Heritage ; ÖAW); **Patritsia Kalafata** (Academy of Athens); **Phil Carlisle** (Historic England); **Sofia Aleixo** (Universidade NOVA de Lisboa); **Ulrike Peter** (Berlin-Brandenburg Academy of Sciences and Humanities); **Véronique Humbert** (CNRS-FRANTIQU); **Wibke Kolbmann**

### Agenda Items:

- Introduction to the WG and the BBT meta-thesaurus (H. Goulis, Academy of Athens)
- The BBT federation of thesauri; a service for managing and connecting thesauri (E. Tsouloucha, ICS-FORTH)
- Using the BBT as a filter for the reorganisation of the PACTOLS thesaurus (B. Nouvel, FRANTIQU)
- iDAI.world Thesaurus and the BBT: shaping an archaeological classification system (C. Colombi, DAI)
- Mapping the ROSSIO Thesaurus to the BBT: challenges and possibilities (B. Almeida, ROSSIO Infrastructure)
- Ontological definition of terms: examples from Cultural Heritage terminology (M. Papadopoulou, USMB)
- Discussion, next steps

---

<sup>1</sup> <https://www.youtube.com/watch?v=qQJJf7eesqY>

## Presentations:

### Introduction to the WG and the BBT meta-thesaurus --H. Goulis (Academy of Athens)

Overview of the goals of the WG and the work it has undertaken

- the Backbone thesaurus as a model for semantic interoperability of thesauri used in various domains of the Humanities --resolves the problem of heterogeneous pools of data

Update on activities performed:

- established a curation committee within the TMWG,
- performed a lot of editorial work, run a curation meeting,
- ran virtual meetings in the post-covid era,
- translated the content of BBT,
- had a presence in the virtual DARIAH annual event 2020,
- new partners (Ligatus/Language of Bindings Thesaurus)

### The BBT federation of thesauri; a service for managing and connecting thesauri (E. Tsouloucha, ICS-FORTH)

An overview of

- how the Backbone Thesaurus is managed collectively (via the BBTalk, submission function)
- how local thesauri are integrated with the BBT (via BBTalk, connection function)
- how can the federated resources be accessed and used simultaneously (through BBT-viewer)

### Using the BBT as a filter for the reorganisation of the PACTOLS thesaurus (B. Nouvel, FRANTIQU)

Some features of PACTOLS (size, fields of archaeological research it covers, supporting multilingualism, major groupings of concepts-each corresponding to a microthesaurus, the software to maintain it (Opentheso), goals and diversified uses)

The need to update it:

- part of it was outdated, part of it was underdeveloped.
- MASA consortium for technical support (to skosify it)

The case of reorganising PACTOLS using the BBT.

- BBT used as a filter for the fields in the micro-thesaurus Sujets
- additional levels --hooks on the BBT
- associative relations abandoned in favor of hierarchical
- terms got merged/deprecated/introduced, definitions got updated
- very impressive statistics

## iDAI.world Thesaurus and the BBT: shaping an archaeological classification system (C. Colombi, DAI)

Unification of thesauri in use by the DAI based on the BBT.

- Reorganization of iDAI.world concepts around the BBT facets, (basis for thesaurus building)
- Addition of facets for Animate beings and Chronologies
- Facets are not all equally densely populated
- Associative relations abandoned in favor of hierarchical
- Definitions were introduced/updated
- iDAI.thesaurus managed through iQvoc

Challenges:

- Huge task, also they were primed to put all object-related information under Objects; they have come to terms with it
- Compound terms
- Descriptive adjectives present in the various thesauri used in DAI, they need consultation wrt what to do with them

## Mapping the ROSSIO Thesaurus to the BBT: challenges and possibilities (B. Almeida, ROSSIO Infrastructure)

Introduction to the ROSSIO Infrastructure (partner institutions, aim-purpose, aggregation services)

Developing vocabulary services --ROSSIO Thesaurus: for searching and indexing

- Topics: social sciences, arts & humanities
- also contains Named Entities

Outline of top-level concepts:

- definitions rely on CIDOC CRM and Dolce
- mappings to BBT
  - BBT Geometric Extents not used at all,
  - Roles used only to some extent -overlapped by Agents in the sense of AAT
  - ROSSIO Places aggregates BBT built environment, physical features, and Geopolitical Units
  - ROSSIO Temporal Entities aggregates BBT Activities, Natural Processes and Types of Epochs (as well as social processes)
  - ROSSIO Physical Objects aggregates BBT Material Things and Materials
  - ROSSIO conceptual objects aggregates BBT Roles and disciplines and also includes qualities (properties of individuals)

Next steps: portuguese translation of BBT terms and scope-notes

## Ontological definition of terms: examples from Cultural Heritage terminology (M. Papadopoulou, USMB)

Thesaurus Management Platform:

- built using the OntoTerminology engine
- hierarchical display, preferred & non-preferred terms

The approach combines ontology, terminology and thesauri

- associates terms to their definitions (i.e. definitions are broken down to a finite set of formally defined characteristics, corresponding to concepts)
- Tedi: software to implement the formal definitions of concepts (cast as classes linked by properties in an aristotelian manner, namely, genus + differentia). Relations allowed: isA, partitive, associative

## Discussion

### Questions to thesauri maintainers:

- Have you (or someone in your network) identified hierarchies to replicate/use as a basis from the BBT federation of thesauri?
- Do you have quick statistics re. the number of terms that can be integrated with the BBT vs. the ones that don't?
  - PACTOLS: Anthroponymes, Lieux and Chronologie are not a very good fit. They also need to define qualities (adjectives in particular) and include named entities
  - iDAI.world: adjectives/participles and agents (as types of individuals or groups), places and other individuals
- Updating data on the dariah.vocabs platform: either ask directly the group (Matej Durco) in OÄW or bring it up with Christos Georgis.

Proposal directed to Bruno for ROSSIO: maybe make concrete proposals to be discussed/ not only in an in-person meeting, but also through BBTalk, because improving the model is part of this group's work.

Submitting a proposal through BBT does not resolve incompatibility issues automatically. But it can be documented there that a given issue needs to be addressed. Even if something is not made part of the BBT, this does not mean that it can't be brought up as an issue in a curation committee meeting.

Alternative proposals may be put forth (especially wrt named entities), we can discuss this offline and reach a concrete proposal before the next curation committee meeting.

## Next steps

Issues that still need to be tackled. The immediate next step is the organization of the 2<sup>nd</sup> BBT curation committee. It's not a given that we will be able to have an in-person meeting. Either start a doodle to determine when the next meeting will be or mark preferences right away.

Possible timeframes for the curation committee meeting include:

- end of January 2022 if we go for a virtual event
- in the week of the DARIAH Annual Event 2022 (to be held in Athens from May, 31st until June, 3rd 2022, if it all goes well)

### **How to proceed:**

- Most people in favor of having 2<sup>nd</sup> curation committee meeting in late January/beginning of February
- Coordinators will share a google doc to record the issues proposed by all WG open meeting participants to be addressed in the 2<sup>nd</sup> BBT Curation Committee meeting.
- The discussion on named entities will be conducted prior to the Curation Committee meeting, so that we form a concrete proposal. The initial discussion needn't be documented through BBTalk (it can take place on zoom and on google docs), but the proposal and a summary of the discussions should be made available there.
- Coordinators will circulate a doodle among Curation Committee members to fix the date and time of the meeting.
- Coordinators along with the curation committee members will finalize the topics to be discussed and who wants to get involved in what.

Last, DARIAH offers the possibility to run another open working group meeting in the same format, confirm with Francesca.