BBTalk: An online service for collaborative and transparent thesaurus curation

by Christos Georgis, George Bruseker, Eleni Tsouloucha (ICS-FORTH)

BBTalk is an online service designed to support collaborative interdisciplinary development and extension of thesauri. At present, it serves to support the curation of the BackBone Thesaurus (BBT), a meta-thesaurus for the humanities. This service allows for the transparent, community development of the BackBone Thesaurus by enabling users to submit suggestions for changes and additions to the terminology, as well as link specialist thesauri to the meta-thesaurus terms, while enabling the thesauri curators to jointly edit, add and delete terminology. This model of cooperative editing is linked to an online discussion system that allows thesauri curators to confer with one another, exchange views and ideas and finally determine any necessary changes to the BBT.

The BBTalk service [L1] addresses the need for a means to cooperatively and transparently build and curate a meta-thesaurus for the humanities, the **BackBone Thesaurus** (BBT) [L2]. The BBT is the research outcome of work undertaken by the Thesaurus Maintenance Working Group (TMWG) [L3] - established in the framework of DARIAH-EU. The BBT offers top-level-concepts (facets and hierarchies) that should serve as common ground for generic thesaurus building. The technique adopted for the BBT of faceted classification is, moreover, considered valid and consistent from a cross-disciplinary perspective. A major advantage of this approach is the potential expansion of the BBT into new scientific domains in a sustainable and manageable fashion. The BBT explicitly does not require domain experts to abandon their specialized terminology in the name of some universally accepted terminology. Rather, the core feature of the BBT is that it promotes alignment of cutting-edge terminology to well-formed general terms of the meta-thesaurus. This enables both collaboration and cross-disciplinary resource discovery while ensuring compatibility with thesauri that cover highly specific and developing areas of knowledge. In order to support this vision of a collaborative and gradually expanded top level meta-thesaurus, it is necessary to have the tools to curate and evolve the standard, adjusting it in an objective and cooperative manner to accommodate new areas of research and thought.

To meet this need, BBTalk has been developed by ICS FORTH [L4] and is offered as an online service that enables the basic functions necessary to support the open development of the BBT. These functions include, *inter alia*, a) proposing new terms and changes to the meta-thesaurus itself and b) connecting specialist thesauri to the federated system.

Key to the BBTalk service is the offer of a built-in communication system supporting discussions between the different user groups of the BBT regarding submissions for changes of the BBT and/or any connections proposed. Such proposed actions are then considered and managed in a formal but open editorial process. Moreover, BBTalk keeps a record of the different versions of the BBT and the history of the submissions in the form of discussions relating to the evolution of the meta-thesaurus. This ensures the level of transparency which is the *conditio sine qua non* for a meta-thesaurus to serve as central resource for academic research, where decisions must be reached openly and on principle, not by fiat. The ability to communicate in a structured way in relation to the evolution of the thesaurus, and the ability to trace back the decision process allows for a reasoned and criticizable evolution of the thesaurus.

With regard to the functionality for connecting specialist thesauri to the overall BBT schema, BBTalk functions as an alignment tool. Through this alignment functionality, specialist thesauri maintainers are able to align their thesaurus to a high-level and well structured thesaurus. This alignment process

creates a critical feedback relation between the specialist thesauri providers and the BBT maintainers. On the one hand, the proposed links offer to the BBT curatorial team use cases that either corroborate or provide disconfirmations to the top level classifications. On the other hand, specialist thesauri maintainers can test their terms against the BBT classifications to see whether they follow generically robust higher level classifications. The alignments performed in BBTalk can be exported as RDF and then used by researchers in their data transparently.

Taken together the functionalities of BBTalk allow it to work as a general maintenance system for the BBT, supporting the implementation and documentation of proposed changes and, by storing, the connections of specialist thesauri terms to the BBT together with the contact information of their maintainers, enabling the automated notification of changes to BBTalk users. The entire process is supported by a user-friendly browsing facility that allows users of BBTalk to view the connected terms of specialist thesauri that are linked to the meta-thesaurus and the type of relation they stand in, with respect to BBT's relevant terms, i.e. an exact or a broader match (Figure 1).

The BBTalk tool is currently used by a DARIAH-supported community of scholars, who have undertaken the task of curating the BBT. Members of this community include all members of the DARIAH TMWG; FORTH-ICS, AA [L5], DAI [L6] and FRANTIQ [L7]. As more institutions join, we envisage that the federation of more and more thesauri under a common banner will generate a marketplace of extant thesauri that can be reused by researchers working in the same discipline, thus supporting comparable research and reduce the need for harmonization and alignment of datasets.
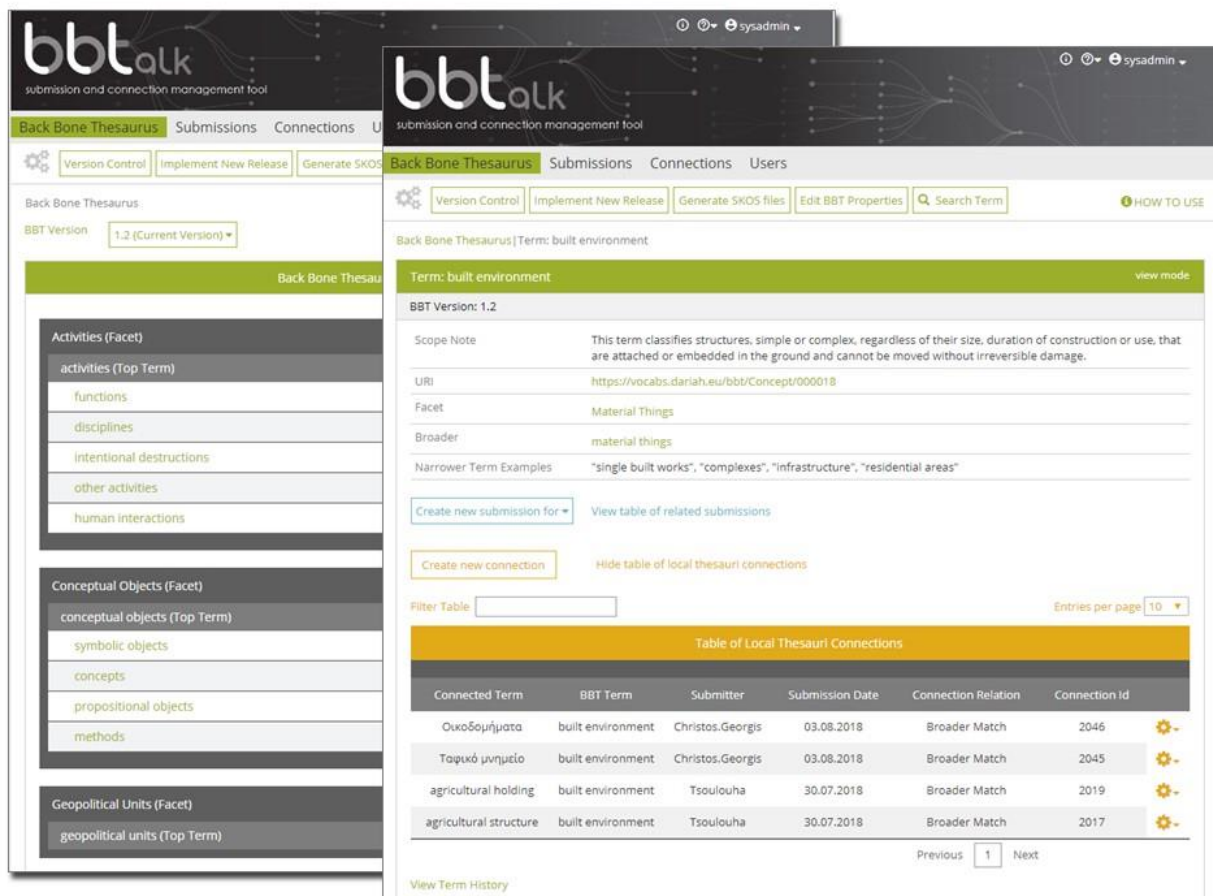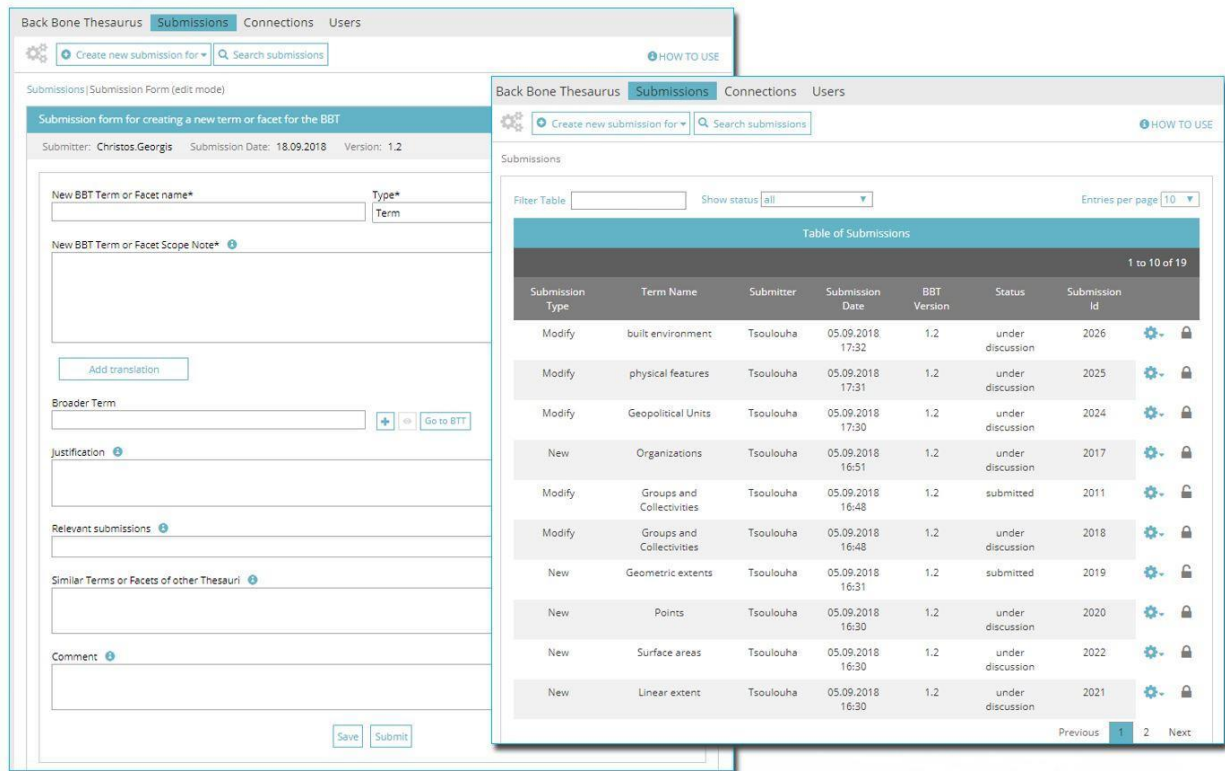

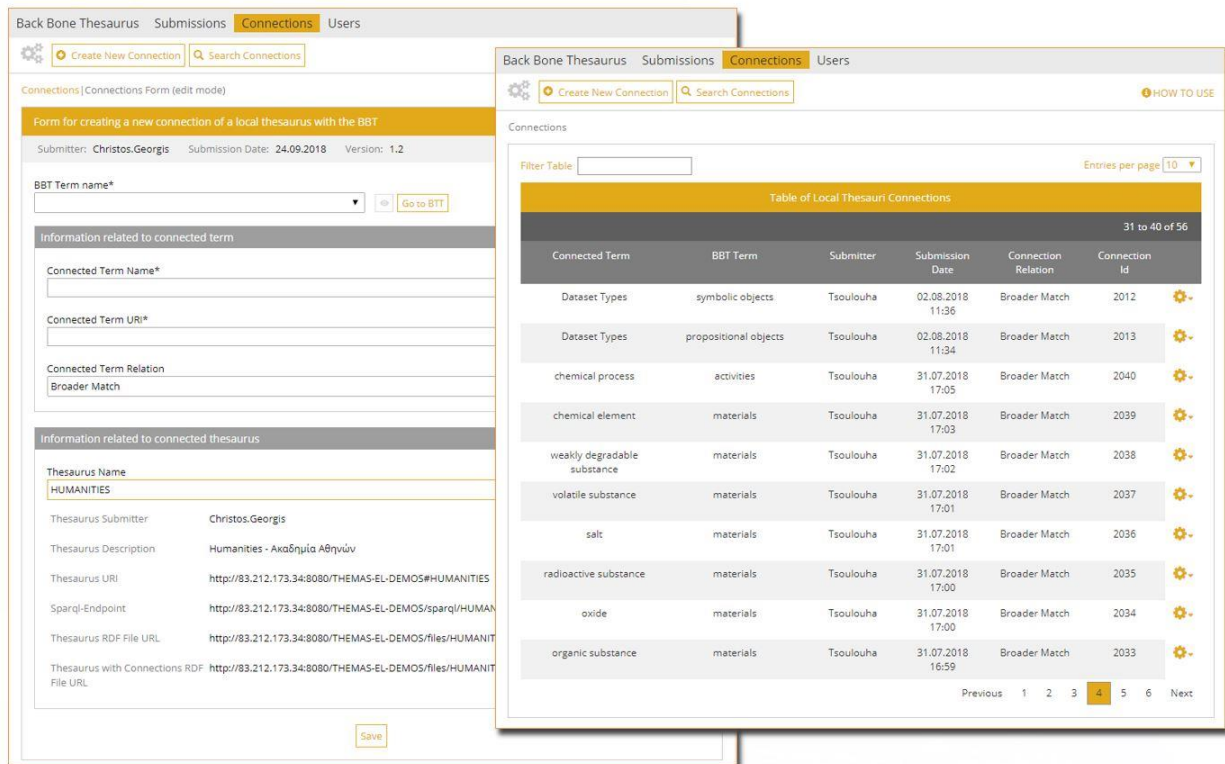*Figure 1: BBT overview interface*

Figure 2: Submission Interface

Figure 3: Connections interface

Links:

[L1] http://www.backbonethesaurus.eu/BBTalk

[L2] http://www.backbonethesaurus.eu/

[L3] https://www.dariah.eu/activities/working-groups/thesaurus-maintenance/

[L4] http://www.ics.forth.gr

[L5] www.academyofathens.gr

[L6] www.dainst.org

[L7] https://www.frantiq.fr

References:
[1] M. Daskalaki, L. Charami: "A backbone Thesaurus for Digital Humanities", in ERCIM News, Issue 111, October 2017, Special theme: "Digital Humanities", p. 18.

Please contact:
Christos Georgis
ICS-FORTH, Greece
georgis@ics.forth.gr