THE CESAB THESAUFORM

A COLLABORATIVE TOOL FOR THE DOMAIN SCIENTIST

Alison Specht (CESAB-FRB), Baptiste Laporte (CESAB-FRB),
Marie-Angelique Laporte (Biodiversity International), Eric Garnier (CEFE-CNRS)

Common and understandable terminology is essential for sharing data and information. The lack of clear and accessible terminology and definitions of those terms is very evident in a synthesis centre, where our business is to assist researchers to aggregate existing data. In some fields, such as in the identification and definition of plant traits, we are well progressed, but in others much less so, and in order to advance the world of open data it is increasingly urgent that we improve the semantic capture of domain-specific data.

This web-based tool allows a group of scientists to collaborate in the identification of the main concepts used in their domain, to propose definitions for the terms related to the concepts, and establish a list of synonyms for these terms.

The CESAB Thesauform allows :

- Concept categorisation,
- Enrichment : add / delete terms,
- Modification of concepts (terms and their associated information),
- Coordination of different points of view : voting protocol, and
- Graphic information display.

IMPLEMENTATION

The Thesauform is designed to allow domain specialists to define the terminological aspects of specialised domain concepts. The definition of the key concepts is accomplished collaboratively.

These collaboratively-defined concepts are added to a SKOS thesaurus with metadata information, conforming to the Dublin Core Abstract Model and structured according to property and value (respectively P and V in Figure 2).
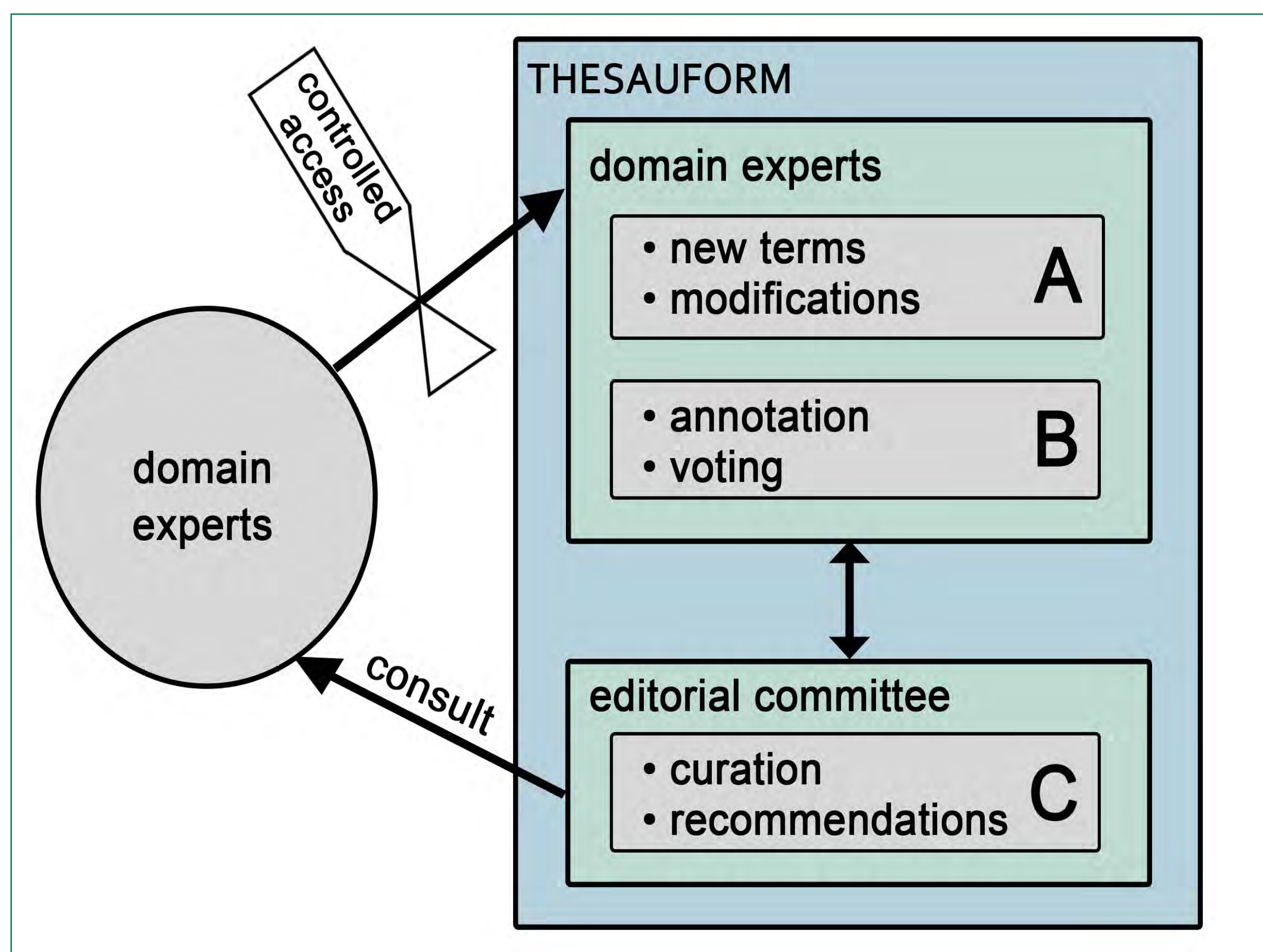

Figure 1: Workflow between domain experts and the Thesauform. Letters refer to stages in the process illustrated accordingly
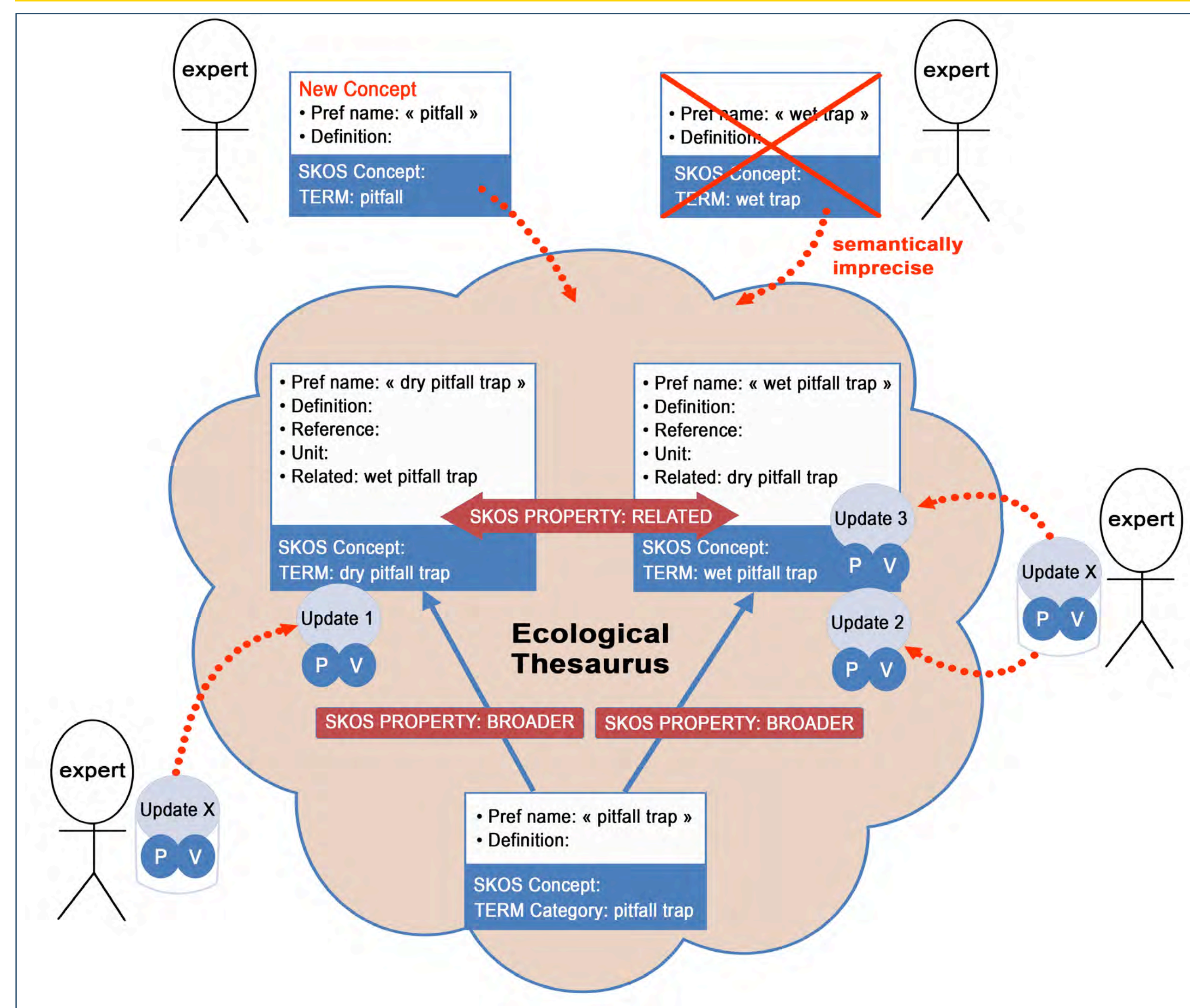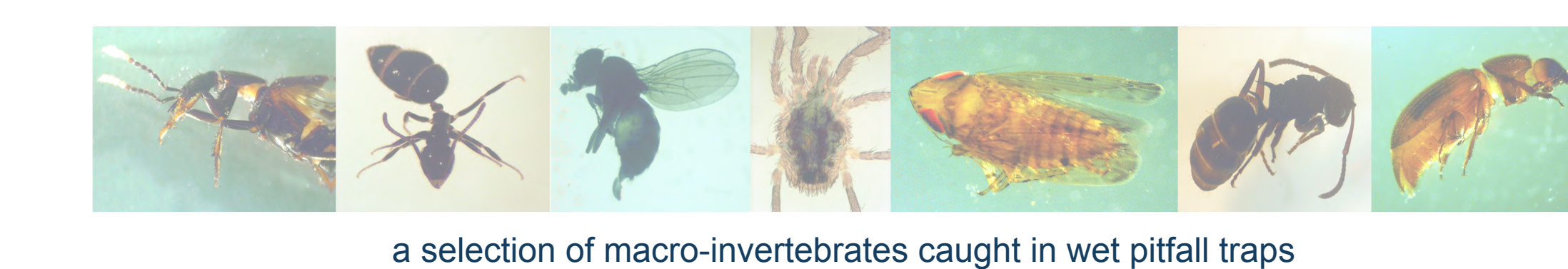

Figure 2: Thesauform structure. P = Property, V = Value

a selection of macro-invertebrates caught in wet pitfall traps

WHERE CAN I SEE IT?

The Thesauform source code is available online at : https://github.com/CESAB-FRB/Thesauform/wiki

Examples of utilisation can be found at:
- http://thesaurus.cesab.org/ and
- http://t-sita.cesab.org/BETSI_vizIndex.jsp

NEXT STEPS

The THESAUFORM will be used to:

- deliver concepts (terms, definitions, sources, related terms etc) to ontologies (e.g.ENVO, PCO) to be validated by the community.
- provide a framework for use by different domains.
- publish the resulting thesauri on-line.