



## Milestone M4.19

### Canvassing glossary-oriented potential contributors

**Leading partner:** JKI

**Compiled by:** Gregor Hagedorn

**Date:** Delayed submission July 7, 2012 (original planned May 15, 2012)

For the development of the MediaWiki-based ontology system as a task in WP4 we were seeking addition content providers and collaborators.

Background: Small-scale experimental work showed that defining ontologies on a semantic MediaWiki, with html-form based support to shorten the learning curve, is doable. However, it is not likely that by providing a platform with little examples, many biologists will become attracted to the proposed method in working on biological and especially descriptive (structural, morphological, anatomical, physiological, behavioural, etc.) ontologies. Widespread interaction, interest, and examples to follow are, however, essential for this plan to succeed. The goal of using a collaborative, community based tool like MediaWiki rather than conventional knowledge engineering tools is to foster involvement of biological knowledge domain experts.

Existing small demos before a call for further contributors were:

- \* <http://species-id.net/wiki/TaxPub> small-scale demo highlighting semantic interactions based on skos vocabulary, use of html forms (in the individual term pages like <http://species-id.net/wiki/TaxPub/collecting-event>, click on "edit with form" on top-tab row).
- \* [http://offene-naturfuehrer.de/web/Kategorie:Botanischer Fachausdruck](http://offene-naturfuehrer.de/web/Kategorie:Botanischer_Fachausdruck) German-language Glossary of the Open Nature Guides (= Offene Naturführer).

The present Milestone documents the canvassing for further interested contributors and collaborators. Both individual communications based on pre-existing collaborations and a general broad call on the TAXACOM list (see Appendix for text) resulted in a number of opportunities. The most promising collaboration opportunities are listed:

#### 1. ViBRANT internal collaborations

- \* GBIF (a ViBRANT partner) is exploring Semantic MediaWiki for their purposes, lead by Dag Endresen and Éamonn Ó Tuama. A technical meeting in August is scheduled to exchange experiences gained so far and decide which vocabulary to focus on as a test case. One vocabulary the availability of has already been assured is the bionomenclature glossary (<http://bionomenclature-glossary.gbif.org/>).
- \* Together with the BGBM we will explore the conversion of CDM content (e.g. an existing exemplar group) into a published wiki version, in parallel to the CDM to PDF process that is currently implemented. The primary interest is that the wiki process can offer semantic links, i.e. links that are not only anonymous hyperlinks but also links specifying the subject and property, and which are automatically exposed as RDF for machine-readability. Further features of interest are full version control, automatic indexing and cross-referencing, combined with a long-term (>10 yr) maintainable software environment. Another major advantage over PDF is that the output is optimized for viewing on digital devices and is more scalable to current and future mobile devices than the fixed PDF format, which is ideal for printing. The functionality is

considered to be particularly useful when the CDM will be further exploited to develop eFloras and eFaunas.

## 2. External collaborations

- \* We are collaborating with the WikiData project (<http://meta.wikimedia.org/wiki/Wikidata>) that is presently building the foundation to supplement the text-based Wikipedias with a semantic data facility. Dr. G. Hagedorn is a member of the Advisory Board for this project.
- \* Scott L. Gardner is collaborating with ViBRANT to re-publish the "Dictionary of Invertebrate Zoology". The discussion is very positive and the issue of legal licensing and copyright seems to have been cleared with all involved authors.
- \* James Macklin and collaborators working on the Flora of North America (FNA) project pursue a very similar idea in presenting the FNA glossary and ultimately the descriptions by means of a MediaWiki-based presentation. The terms of collaboration with ViBRANT are still under discussion.
- \* The UFZ Leipzig has agreed to support the exposure of the extensive BioFlor data set, a German language plant trait ontology, on a semantic MediaWiki. The ViBRANT partner JKI has already extracted and converted the data from their more complex legacy relational database editing system into a local database and is in the process of planning the presentation on a Semantic MediaWiki (which is not trivial as a result of modifier, annotation and source complexity added to the simple character).
- \* French botanist Laure Civeyre is offering an interesting set of plant bud and twig terminology together with description that allow identification of woody plants in the winter without leaves or flowers.

## APPENDIX

In addition to several individual requests, we used the TAXACOM mailing list and published the following request on 20 Apr 2012:

### **Seeking collaboration on illustrated biological dictionary/glossary for semantic MediaWiki with ontological relations**

Within the context of the EU-funded ViBRANT project ([vbrant.eu](http://vbrant.eu)) we are exploring, together with GBIF, semantic MediaWiki. The goal is to combine the ease of definition of biological terms in a Wikipedia-style, human-accessible, and possible richly illustrated manner with an RDF/OWL based ontology. Imagine it as a dedicated work with pages that can do everything that Wikipedia can (but with controlled access).

We have resources and the technical foundation to bring such a glossary into a long-term digital publishing framework, but we are lacking a good content for a morphological or taxonomic example. We

realize that most attractive illustrated taxon lists, dictionaries or glossaries are owned by major publishers. If you have created or point us to a quality illustrated glossary that the authors wish to be freely available, we appreciate your collaboration.

The publication would technically be implemented on the biowikifarm.net, a community-run integrated hosting site biological wiki project that has a long-term stability guarantee backed by Munich SNSB and Berlin BGBM. The created glossary would thus be available and editable for the foreseeable future far beyond the project duration.

If you are interested in a collaboration and wish to discuss this further, please contact us until the end of April.

Many thanks!

Gregor Hagedorn (JKI, Berlin), Yde de Jong (University of Amsterdam), Dag Endresen (GBIF Copenhagen)