

PANGAEA – Database improvement for de.NBI services

Melanie Sattler, Uwe Schindler, Robert Huber, Anusuriya Devaraju, Andree Behnken, Michael Diepenbroek & Janine Felden,
 MARUM, University of Bremen, Bremen, Germany

Short description of the project

PANGAEA is a Data Publisher for Earth & Environmental Science and safeguards georeferenced and multidisciplinary data. The services for de.NBI are currently improved by content-related enhancement especially in respect to biomolecular data.

Therefore PANGAEA performs:

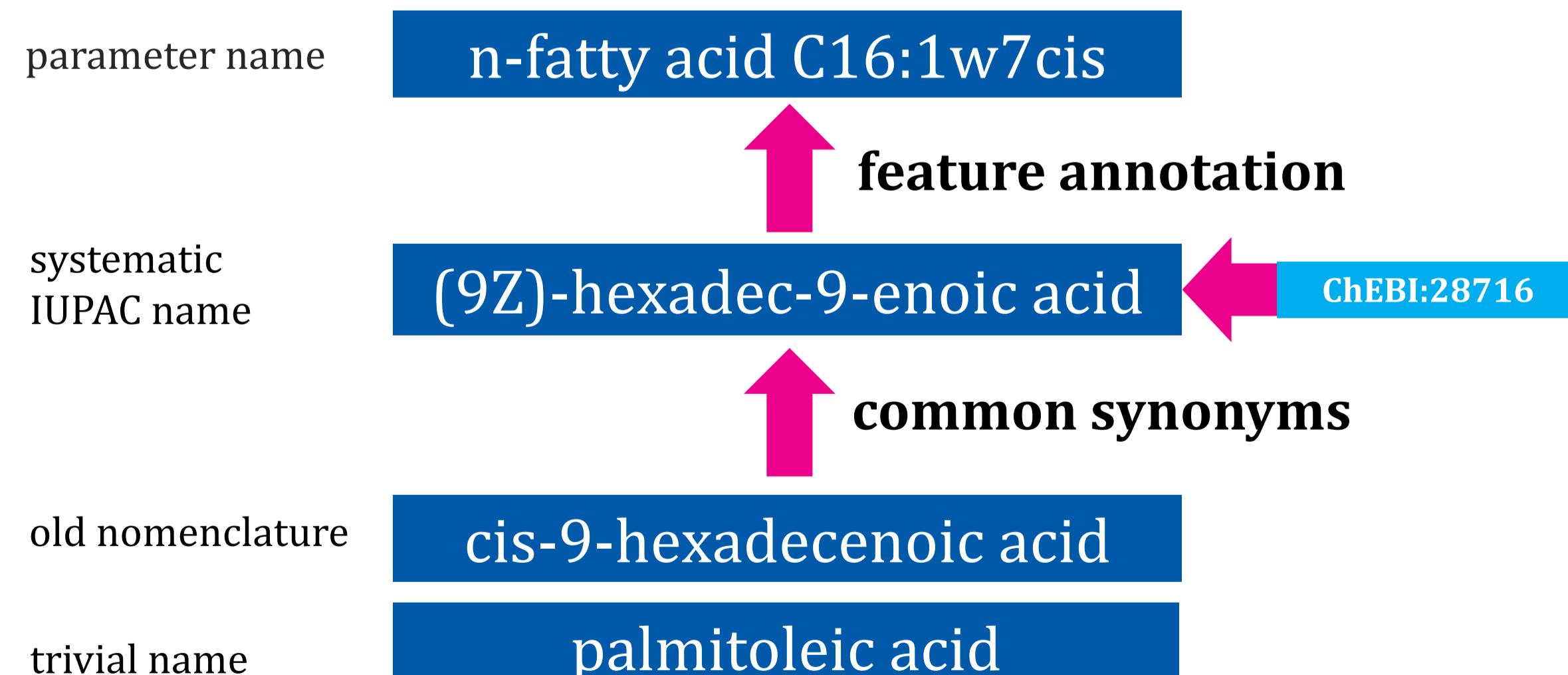
- Exploration of the PANGAEA biological content
- Mapping of PANGAEA ontologies to de.NBI partners
- Cross linking of environmental data to other de.NBI databases
- Definition and training of new curatorial workflows & strategies
- Adoption of data publication services for de.NBI

Progress report

Exploration of PANGAEA content:

- 23,500 parameters describing “biochemistry”- data in over 110,000 datasets have been analysed
- Chemical compound and enzyme names have been spell checked according to the latest IUPAC and IUBMB nomenclatures
- Compound terms have been crosslinked to existing ChEBI-terms (Chemical Entities of Biological Interest)

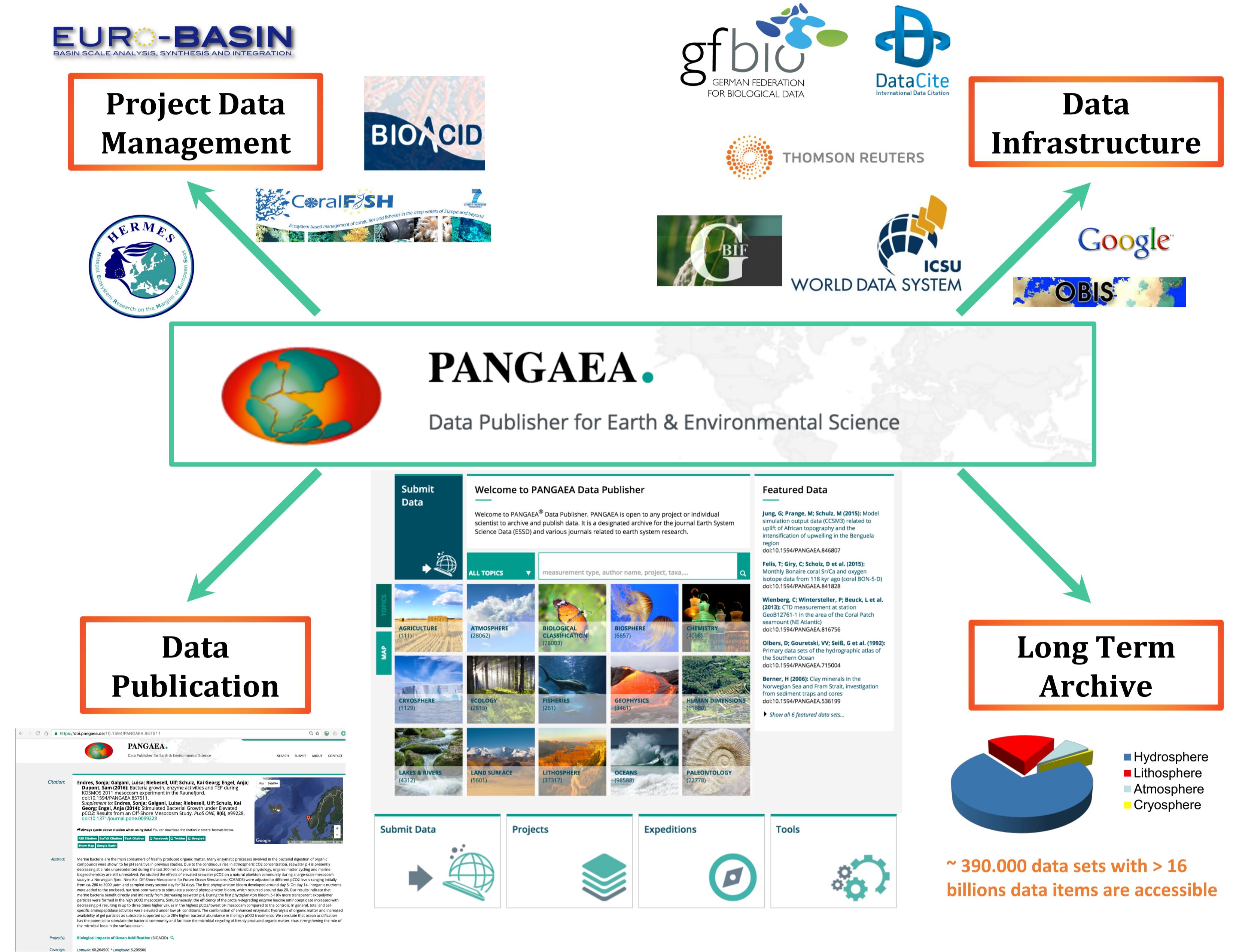
Example:



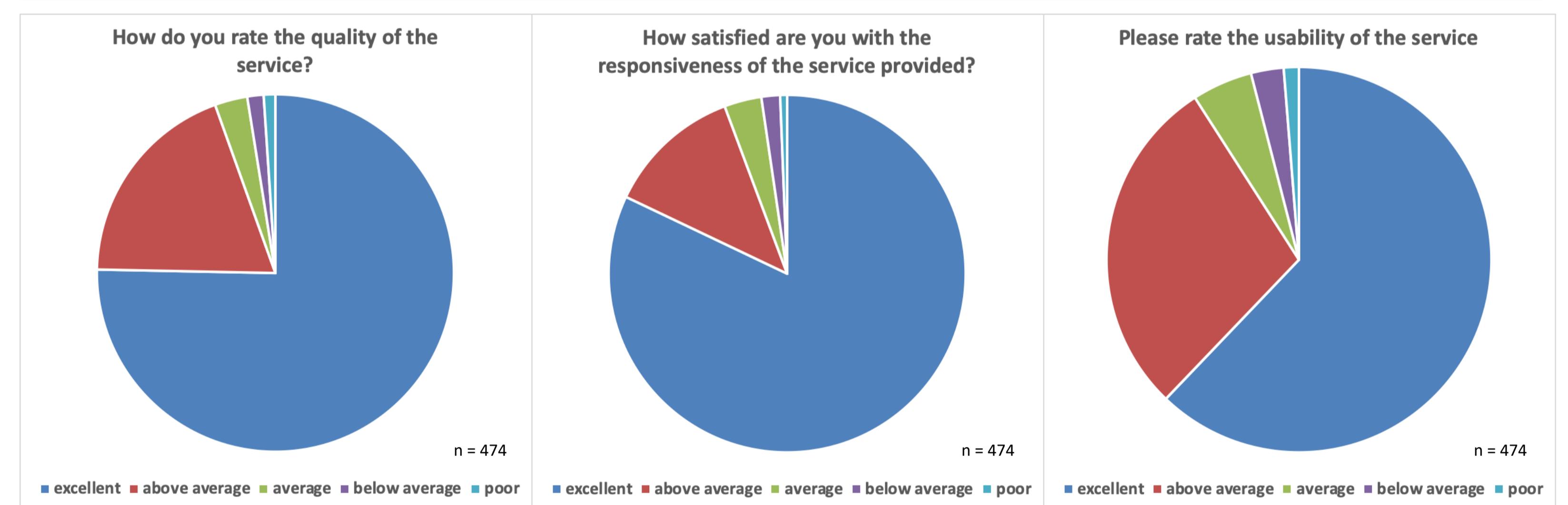
The current steps are:

- Development of curator workflows and guidelines for the precise naming of biomolecules like enzymes
- A parameter construction and measurement quantity recommender tool for curators.

de.NBI services



de.NBI Service Evaluation (2017-2019)



Publications

- M. Diepenbroek, U. Schindler, R. Huber, S. Pesant, M. Stocker, J. Felden, M. Buss, M. Weinrebe (2017): Terminology Supported Archiving and Publication of Environmental Science Data in PANGAEA. *Journal of Biotechnology*, 261:177-186.
<https://doi.org/10.1016/j.jbiotec.2017.07.016>
- G. Mayer, C. Quast, J. Felden, M. Lange, M. Prinz, A. Pühler, C. Lawerenz, U. Scholz, F. O. Glöckner, W. Müller, K. Marcus, M. Eisenache (2017): A generally applicable lightweight method for calculating a value structure for tools and services in bioinformatics infrastructure projects. *Briefings in Bioinformatics*, 20(4):1215-1221,
<https://doi.org/10.1093/bib/bbx140>
- N. Karam, A. Khiat, A. Algergawy, M. Sattler, C. Weiland (accepted): Matching Biodiversity and Ecology Ontologies: Challenges & Evaluation Results. *The Knowledge Engineering Review*.

General information on the project

No. of staff paid from de.NBI grant: 1 FTE

Other staff involved: 2 FTE