

Tagung DigitalSpecimen 2014, MfN Berlin 8.9.-12.9.2014

Poster:

The Biodiversity Network of the Humboldt-Ring (BiNHum): Joint Data Portal for Natural History Museums

Joachim Holstein (State Museum of Natural History Stuttgart), Peter Grobe (Zoological Research Museum Alexander Koenig, Bonn), Dagmar Triebel (Bavarian Natural History Collections, Munich), and Anton Güntsch (Botanic Garden and Botanical Museum Berlin-Dahlem).

BiNHum (<http://wiki.binhum.net>) is a project of five natural history museums and research collections representing the Humboldt-Ring: the State Museum of Natural History Karlsruhe (SMNK), the State Museum of Natural History Stuttgart (SMNS), the Zoological Research Museum Alexander Koenig in Bonn (ZFMK), the Bavarian Natural History Collections in Munich (SNSB), and the Botanic Garden and Botanical Museum Berlin-Dahlem (BGBM).

The three-year project is funded by the German Research Foundation and deals with:

- data recovery and data mobilisation (lead by SMNS, SMNK and the University of Ulm),
- implementation of modern search technologies and development of the data portal (lead by ZFMK),
- data harvesting and data quality assurance (lead by BGBM),
- deployment of Diversity Workbench (DWB) as a virtual environment for BiNHum (lead by SNSB).

The BiNHum data portal will provide rich metadata content and multimedia data types like 3D images and audio files: The digitized fossil leaves database MORPHYLL (at SMNS) with ecophysiologicaly relevant morphometry as well as ZFMK's acoustic multimedia databases will be accessible. For these purposes, the BiNHum portal will offer search options for complex data features (such as sound frequencies or geometric shapes) in addition to autocorrection, autocomplete, faceting, "shopping cart" and other harvesting features. A first test version of the portal is already available (<http://demo.binhum.net/>) using a scalable full-text search indexing server (SOLR 4.4.0) and a relational database (MySQL). Currently the portal needs a logon to detain webcrawlers. The access code is available by request ([joachim.holstein \[at\] smns-bw.de](mailto:joachim.holstein[at]smns-bw.de) or [c.koehler \[at\] zfmk.de](mailto:c.koehler[at]zfmk.de)). Reviews and/or suggestions would be highly appreciated.

All partners use the Biological Collection Access Service (BioCASE) to provide their data to GBIF and to the BiNHum portal. For BiNHum the transformed data sources are harvested by GBIF's Harvesting and Indexing Toolkit (HIT), which has been extended to fulfill project-specific purposes (e.g. identification history, organismic associations).

The Diversity Workbench (DWB), available at <http://www.diversityworkbench.net>, is a framework with more than 10 independent relational databases (MS SQL) with rich clients. DWB is used as the core database management system by most of the partners. The applications are being expanded to support the transformation of existing data to current IT standards in the BiNHum context. Independent DWB installations for the BiNHum partners SMNK, SMNS, SNSB and ZFMK allow to curate, quality-control and manage the original data before being processed by BioCASE and the BiNHum portal. The database DiversityDescriptions (V. 3) with a triple store architecture for the management of descriptive and measurement data is newly implemented.

After a first phase of BiNHum, when high-quality object data of Humboldt-Ring institutes have been made available and the data portal is fully operable, further museums and collections outside of the Humboldt-Ring are invited to participate.