

The Diversity Workbench framework as data repository for biological data

Markus Weiss¹, Dagmar Triebel¹, Gregor Hagedorn², Stefan Jablonski³, Gerhard Rambold⁴,
Tobias Schneider³, Bernhard Volz³

¹IT Center of the Bavarian Natural History Collections, Munich

²Julius Kühn-Institut Berlin

³Applied Computer Science IV, University of Bayreuth

⁴DNA Analytics and Ecoinformatics Laboratory, University of Bayreuth

The database framework Diversity Workbench consists of databases and client components interacting via interfaces. Because of this, each component of the Workbench can be directly used as a stand-alone application but can also be called from other clients, without exhibiting details about the internal design and implementation (encapsulation principle). This results in an increased flexibility concerning the technical design, allowing a differentiated user administration and the rapid setup of user-adapted entry forms for specific projects. It facilitates the direct access to web services and external data resources. Among the more than ten components, one is providing GIS (Geographic Information System) functionality and another is set up for mobile platforms with synchronisation mechanisms. The framework is appropriate to store different kinds of data about interorganismal interactions and facilitates the management of institutional collections. The Diversity Workbench therefore represents a modularized system for multiple scientific purposes concerning analysis and management of biological data, e. g. collection data, observation data, descriptive and ecological data, taxonomy, taxon-mediated checklist data, ecological plots, images and literature data. It is also flexible enough to build the software backbone for long-term institutional data repositories for biological data.

A large number of biodiversity research projects depend on field mapping and ecological data of high quality. Therefore it is necessary to link data sets gathered in the field to a verified backbone information from major biological, taxonomic or environmental data sources, e.g., lists of taxonomic names. Further on, there is a need to link additional information like Global Positioning System (GPS) coordinates and multimedia information, such as images or sound, at the time of data gathering. A seamless and transparent flow of data from the field to a central data storage system which may simultaneously be used by several participants is a necessity. “Seamless” in this sense means that data are available as stored datasets shortly after gathering; and “transparent” in the sense that every operation applied to data can be traced back (“data provenance”). For these core requirements, the Diversity Workbench includes a mobile application used to gather biological research data in the field that enables the dataflow to the data repository. The application DiversityMobile is currently designed as a mobile system for entering ecological and biological monitoring data already in the field. It is capable of accomplishing the core requirements mentioned above. For reasons of model consistency the mobile client uses a subset of the database model of the central database of the Diversity Workbench framework. See <http://www.diversityworkbench.net> and <http://www.diversitymobile.net> and the contribution “The Diversity Workbench Framework: Data retrieval with DiversityMobile and Dataflow from DiversityMobile to GBIF” for further information.