

Biodiversity needs data: Rescue natural history information!

Anton Güntsch & Walter Berendsohn, Botanic Garden and Botanical Museum Berlin-Dahlem, Freie Universität Berlin, Königin-Luise-Str.6-8, D-14195 Berlin, Germany, BiodiversityInformatics[at]bgbm.org

Primary biodiversity data are traditionally stored in natural history collections documenting the occurrence of organisms at a certain place and time. This information is usually associated with a physical specimen or an observation documented on a file card for example. Recent estimations suggest that between two and three billion collection units are held within biological collections of which about 200 millions have been databased and linked to international biodiversity information networks such as BioCASE (Biological Collection Access Service for Europe, www.biocase.org) and GBIF (Global Biodiversity Information Facility, www.gbif.org) in a standardized and open way. The vast majority of records are still in a pre-digital form and many data collections lack a sufficient maintenance and archiving plan so that their long-term availability is seriously threatened.

To realize the enormous potential of Natural History information in an increasingly networked information space calling for instant, efficient, and standardized access to primary biodiversity data, two fundamental problems have to be addressed. On the one hand we need more efficient ways for bringing pre-digital information in the digital domain. This has to be achieved by introducing highly efficient (industrialized) digitization hardware and workflows into the collections as well as improving costly Meta data capture processes. Within the last years, promising approaches to high-speed digitization in particular in herbaria have been developed.

On the other hand, the problem of "databases at risk" will have to receive more attention. At the very least, we will need a rescue plan and workflows for databases to transfer them into standardized biodiversity data schemes, link them up to the global biodiversity data infrastructures, and store them in a proper archive.

The benefit from the additional infrastructures for high-speed digitization and data archiving investment will easily outweigh the necessary investment by opening up

300 years of natural history to biodiversity sciences tackling pressing questions in a rapidly changing environment.