

## PRESS RELEASE

For immediate release: 19 December 2011

### German research team targets ‘at risk’ data on biodiversity

#### Project example of data hosting infrastructure promoted by GBIF

Copenhagen, Denmark - A new German-based project is setting out to rescue biodiversity data at risk of being lost, because they are not integrated in institutional databases, are kept in outdated digital storage systems, or are not properly documented.

The project, run by the Botanic Garden and Botanical Museum Berlin-Dahlem, provides a good example for a GBIF recommendation to establish hosting centres for biodiversity data. This is one of a [set of data management recommendations](#) just published by GBIF.

The team behind the German project called [reBiND](#), or Biodiversity Needs Data, has started identifying threatened databases for archiving, and will make them accessible via the GBIF network.

The focus is initially on specimen and observational data that are already digitized but that are not part of the documentation process of a museum or other institution. Examples include data from diplomas or PhD theses, generally stored on a computer hard drive or a disc and often in danger of getting lost because of lack of documentation.

Examples of data being ‘rescued’ by the reBiND project are:

- A private collection of observation data on beetles from meadow orchards in Southern Germany. The data had been stored by biologist Andreas Kohlbecker on a 1986 Mac 512 computer, using a Mac OS 6.8 operating system and making use of Filemaker II software from 1989. The data were rescued by running the operating system and software in Basilisk II Mac Emulator.
- Extensive primary data from a PhD thesis on epiphytic moss vegetation in the Canary Islands. They had been stored in 1997 on obsolete 3.5-inch floppy discs using Excel files. They have been made readable using an external floppy drive and will be converted to XML format. These data are especially valuable because the study was the first to document moss communities on the islands taking microclimates and human impacts into account.

The workflow developed by the reBiND project uses the Biological Collections Access Services (BioCAsE) provider software package to transform data into XML files. BioCAsE is one of the publishing tools through which data are published to the GBIF network. Repair software detects and corrects any errors introduced during the conversion process. ReBiND aims to enable users with a minimum of technical background knowledge to transform and archive their biodiversity data.

At present, the XML files containing the rescued data are stored on the project’s own server, in a database specifically designed for the purpose. The intention is to make the data discoverable and accessible globally through the GBIF network, and the team is working with GBIF Germany to bring this about.

Anton Güntsch, the Principle Investigator of reBiND says, “The project benefits from experiences, standards, and infrastructure developments of GBIF and BioCAsE. These provide a solid foundation for data rescue workflows.”

Maren Gleisberg, the node manager of GBIF Germany adds, “We are looking forward to cooperating with reBiND. The project enables contact with data providers such as diploma and PhD theses, which have not been a focus for GBIF. GBIF Germany is promoting the project among the eight German Nodes.”

Although the project is specialized in saving zoological and botanical data, the reBiND team believes the workflow it has developed is generic enough to be used in any field.

The project expects to take on data rescue work globally. The team is also working on a best practice handbook on the rescue and storage of threatened digital data.

The project is also illustrated in an entertaining five-minute video, which is available at [http://rebind.bgbm.org/rebind\\_movie](http://rebind.bgbm.org/rebind_movie)

The three-year project is funded by the [German Research Foundation \(Deutsche Forschungsgemeinschaft\)](#).

The [GBIF position paper on data hosting infrastructure for primary biodiversity data](#) looks at the rescuing and re-hosting of data stored in formats that are difficult to access. It emphasizes that the biodiversity community must adopt standards and develop tools to enable data discovery and thus help preserve data.

**For more information on the project, please contact:**

Anton Güntsch  
reBiND  
[rebind@bgbm.org](mailto:rebind@bgbm.org)

**For more information on the GBIF position paper on Data Hosting Infrastructure please contact:**

Vishwas Chavan  
[vchavan@gbif.org](mailto:vchavan@gbif.org)

**For information on GBIF and for press and media enquiries, please contact:**

Tim Hirsch  
GBIF Secretariat  
[thirsch@gbif.org](mailto:thirsch@gbif.org)

***Notes to editors:***

The Global Biodiversity Information Facility (GBIF) was established by governments in 2001 to encourage free and open access to biodiversity data, via the Internet. Through a global network of 57 countries and 47 organizations, GBIF promotes and facilitates the mobilization, access, discovery and use of information about the occurrence of organisms over time and across the planet.

***Relevant links:***

- reBiND project website: <http://rebind.bgbm.org/>
- Video on reBiND project: [http://rebind.bgbm.org/rebind\\_movie](http://rebind.bgbm.org/rebind_movie)
- GBIF Position Paper on Data Hosting Infrastructure for Primary Biodiversity data (pdf file, 333 KB): [http://imgbif.gbif.org/CMS\\_ORC/?doc\\_id=4386&download=1](http://imgbif.gbif.org/CMS_ORC/?doc_id=4386&download=1)