# New biodiversity data publishing framework proposed

December 15 2011

A comprehensive framework to encourage and facilitate the sharing of biodiversity data has been published in a peer-reviewed journal.

The 24 recommendations of the GBIF Data Publishing Framework Task Group are included in a special supplement of *BMC Bioinformatics*.

Some of the measures proposed by the task group have already been taken up by GBIF and others, including the publication of 'data papers' to bring academic recognition to those publishing metadata to describe datasets.

Others are in the pipeline, such as new services to monitor and publish use of data in a Data Usage Index, and a mechanism to ensure that all those involved in collecting, adding value to and publishing data are acknowledged in a single citation.

Introducing the data publishing framework in the central paper of the supplement, Tom Moritz and co-authors write that the first recommendation is the primary one from which all the others follow: "All data relevant to the understanding of biodiversity and to biodiversity conservation should be made freely, openly and effectively available."

The authors define a data publishing framework as an environment conducive to ensuring free and open access to the world's primary biodiversity data: "The core purpose of the framework is to overcome barriers or impediments affecting access to data and the publishing of

data."

The paper goes on to argue that sharing of biodiversity data must be the expected norm, and that data should only be withheld in exceptional circumstances when precise localities need to be protected, for example in cases involving marketable plants or animals, or for species of special concern. "We emphasize that such data represent a small fraction of biodiversity data and should not be allowed to dictate normal practice," the authors add.

Other recommendations include:

- expanding and improving the metadata framework used by GBIF to help users establish whether data records are fit for use;
- strengthening the network of GBIF's mirror sites and 'trusted digital repositories' or data hosting centres;
- meeting the need for stable and proven persistent identifiers so that data records remain permanently discoverable and accessible;
- establishing a 'biodiversity informatics potential index' to demonstrate the potential and urgency of countries' investment in management of biodiversity data (GBIF will shortly release such an index);
- encouraging sponsors of biodiversity research to set mandatory requirements for free and open access to biodiversity data—the paper welcomes moves in this direction by the National Science Foundation in the United States; and
- encouraging professional societies and professional disciplines to require evidence of effective sharing of data in evaluations for hiring, promotion and tenure.

The authors of the paper conclude: "Implementation of these

recommendations should expedite the progress of archiving, curation, discovery and publishing of primary biodiversity data, because scientists and originators of data will realize the value and incentives for such efforts.

"We believe that implementation of our recommendations by the GBIF network, and its adoption by similar initiatives such as GEO-BON (Global Earth Observation – Biodiversity Observation Network), IPBES (Intergovernmental Science-Policy Interface on Biodiversity and Ecosystem Services) and CBD (Convention on Biological Diversity), will contribute to a much needed global research infrastructure and specifically to an open access regime in biodiversity and conservation science."

Separate papers in the BMC Biodiversity Informatics supplement cover: the 'data paper' mechanism; indicators for a Data Usage Index; an index for biodiversity informatics potential of countries; and proposals for a data hosting infrastructure.

Provided by Global Biodiversity Information Facility