

International Networking of Large Amounts of Primary Biodiversity Data

J. Holetschek, P. Kelbert, A. Müller, P. Ciardelli, A. Güntsch, W.G. Berendsohn

Dept. of Biodiversity Informatics and Laboratories
Botanic Garden and Botanical Museum Berlin-Dahlem
Königin-Luise-Str. 6-8
D-14195 Berlin-Dahlem

{j.holetschek, p.kelbert, a.mueller, p.ciardelli, a.guentsch, w.berendsohn}@bgbm.org

Primary Biodiversity Data is a general term for information documenting the planet's biodiversity, where each record represents the existence of a particular organism at a given location at a point in time. These data are scattered throughout numerous collections and databases worldwide, making it difficult to find all information available on, for instance, a certain species or a particular region. Several international networks and initiatives share a vision of free and open access to these resources and are working together to connect these heterogeneous data sources.

This article provides an overview of the basic system architecture of these networks and some of the efforts aimed at solving the issues arising from the explosion of the amount of biodiversity data in recent years. It discusses the concept of "special interest networks" aimed at facilitating the setup of thematic, user- or subject-specific biodiversity data portals; describes SYNTHESYS' data portal software and its query expansion mechanism; and finally, outlines the duplicate detection system for identifying possible duplicate records in biodiversity networks.