OPEN ACCESS

Conference Abstract

Setup of BIOfid, a new Specialised Information Service for Biodiversity Research

Markus Koch[‡,§], Gerwin Kasperek[|], Thomas Hörnschemeyer[‡], Alexander Mehler[¶], Claus Weiland[‡], Angela Hausinger[|]

‡ Senckenberg Gesellschaft für Naturforschung, Dept. Information Technology and Biodiversity Informatics, Frankfurt am Main, Germany
§ Institute of Evolutionary Biology and Ecology, University of Bonn, Bonn, Germany
| University Library J. C. Senckenberg, Frankfurt am Main, Germany
¶ Text Technology Lab, Goethe University, Frankfurt am Main, Germany

Corresponding author: Markus Koch (mkoch@evolution.uni-bonn.de)

Abstract

In order to promote the accessibility of biodiversity data in historic and contemporary literature, we introduce a new interdisciplinary project called BIOfid (FID= *Fachinformationsdienst*, a service for providing specialized information). The project aims at a mobilization of data available in print only by combining digitization of scientific biodiversity literature with the development of innovative text mining tools for complex, eventually semantic searches throughout the complete text corpus. A major prerequisite for the development of such search tools is the provision of sophisticated anatomy ontologies on the one hand, and of complete lists of species names (currently considered valid as well as all synonyms) at a global scale on the other hand. In the initial stage, we chose examples from German publications of the past 250 years dealing with the geographic distribution and ecology of vascular plants (Tracheophyta), birds (Aves), as well as moths and butterflies (Lepidoptera) in Germany. These taxa have been prioritized according to current demands of German research groups (about 50 sites) aiming at analyses and modeling of distribution patterns and their changes through time. In the long term, we aim at providing data and open source software applicable for any taxon and geographic region. For this purpose, a platform for open access journals for long-term availability of

professional e-journals will be established. All generated data will also be made accessible through GFBio (German Federation for Biological Data). BIOfid is supported by the LIS-Scientific Library Services and Information Systems program of the German Research Foundation (DFG).

Keywords

anatomy ontologies, text mining tools, literature digitization, open access, non-commercial publishing

Presenting author

Markus Koch

Presented at

TDWG 2017 Annual Conference, Oct 1-6 in Ottawa, Canada