

ARPHA-BioDIV

A toolbox for scholarly publishing and dissemination of biodiversity data



Lyubomir Penev, Teodor Georgiev, Pavel Stoev, Viktor Senderov, Seyhan Demirov, Petar Geshev

Pensoft Publishers, Sofia, Bulgaria

Toolbox for scholarly publishing and dissemination of biodiversity data (ARPHA - BioDiv)



Overview

ARPHA-BioDiv – EU BON's Toolbox for scholarly publishing and dissemination of biodiversity data (Fig. 1) – is a set of standards, guidelines, recommendations, tools, workflows, services, and journals, based on Pensoft's ARPHA Journal Publishing Platform, designed to ease scholarly publishing of biodiversity and biodiversity-related data that are of primary interest to the EU BON and GEO BON networks.

Expected advantages

Figure 1. EU BON's Toolbox for scholarly publishing and dissemination of biodiversity data (ARPHA-BioDiv).

- Increased efficiency of data publication through novel tools and workflows.
- Mobilisation of non-conventional research outputs via novel article formats.
- Innovative forms of community engagement (data papers, open science collections).
- Machine-readable content available for harvesting and re-use.
- Streamlined creation and import of complex data-rich manuscripts via an Application Programming Interface (API).
- Automated export facilitating data dissemination and re-use.
- Data publishing strategies and guidelines supporting data interoperability.
- Additional incentives for researchers (permanent scientific record, citations) for their effort to collect, maintain and publish biodiversity data.

Applicability

The core element of EU BON's ARPHA-BioDiv Toolbox is the ARPHA Journal Publishing Platform (http://arphahub.com/) developed by Pensoft, along with the associated journals: Biodiversity Data Journal (BDJ), Research Ideas and Outcomes (RIO), and One Ecosystem. ARPHA is an innovative publishing solution that supports the full life cycle of a manuscript, from authoring and reviewing to publishing and dissemination. The data publishing strategy of ARPHA aims at increasing the proportion of structured text and data within the article content, so as to allow for both human use and machine readability to the maximum possible extent.

ARPHA stands for Authoring, Reviewing, Publishing, Hosting and Archiving, all in one place. ARPHA consists of two interconnected workflows. A journal can use either of the two or a combination of both (Fig. 2): 1) ARPHA-XML web-based authoring, peer review and publishing, and 2) ARPHA-DOC - Document-based peer review and publishing. The XML-based workflow is currently used by three journals of Pensoft – BDJ, RIO and One Ecosystem. The second, file-based submission workflow, is currently used by 20 other journals published by Pensoft (ZooKeys, PhytoKeys, MycoKeys, and others).



conversion ARPHA - DOC File(s) Submit a manuscript as document file(s)

Figure 2. ARPHA consists of two integrated workflows: in ARPHA-XML, the manuscript is written and processed via the ARPHA Writing Tool, and ARPHA-DOC, the manuscript is submitted and processed as document file(s).

ARPHA-BioDiv provides:

- Support through the full life cycle of a manuscript, from writing through submission, revisions and re-submission within a single online collaborative platform.
- Easy online collaborative editing by co-authors and peers.
- Novel, community-based and public, pre-submission, pre-publication and post-publication peer review processes.
- Import of Darwin Core compliant primary biodiversity data from spreadsheet templates or via a manual Darwin Core editor and consequent publication in structured downloadable format.
- Direct online import of Darwin Core compliant primary biodiversity data from GBIF, Barcode of Life, iDigBio, and PlutoF into manuscripts.
- Multiple import of voucher specimen records associated with a particular Barcode Index Number (BIN) from Barcode of Life.
- Automated generation of data paper manuscripts from Ecological Metadata Language (EML) metadata files stored at GBIF Integrated Publishing Toolkit (GBIF IPT), DataONE, and the Long Term Ecological Research Network (LTER).
- Automated export of the occurrence data published in BDJ into Darwin Core Archive (DwC-A) format and its consequent ingestion by GBIF. The DwC-A is freely available for download from each article's webpage that contains occurrence data.
- Automated export of the taxonomic treatments published in BDJ into Darwin Core Archive. The DwC-A is freely available for download from each article's webpage that contains taxonomic treatments data.

Potential users

Researchers and institutions from all scientific disciplines who want to publish their biodiversity-related data in scholarly articles.

Case studies

Fauna Europaea Special Issue

One of the major data mobilisation initiatives realised by ARPHA and the Biodiversity Data Journal is the publication of data papers on the largest European animal database 'Fauna Europaea'. A new series 'Contributions on Fauna Europaea' was launched in the beginning of 2014. This novel publication model aimed to assemble within a single collection 57 data-papers on different taxonomic groups covered by the Fauna Europaea project, as well as a range of accompanying papers highlighting various aspects of this project (gap-analysis, design, taxonomic assessments, etc.). The first two papers were published on 17 September 2014 and until the end of 2015, 11 articles altogether

- Novel article types in the ARPHA Writing Tool: Monitoring Schema, IUCN Red List compliant Species Conservation Profile, IUCN Global Invasive Species Database (GISD) compliant Alien Species Profile, Ecosystem Inventory, Ecosystem Services Mapping, Policy Brief, Workshop Report, Project Report, and others.
- Nomenclatural acts modelled and developed in BDJ as different types of taxonomic treatments for plant taxonomy.
- Markup and display of biological collection codes against the Global Registry of Biological Repositories (GRBIO) vocabulary.
- Workflow integration with the GBIF Integrated Publishing Toolkit (IPT) for deposition, publication, and permanent linking between data and articles, of primary biodiversity data (species-by-occurrence records), checklists and their associated metadata.
- Workflow integration with the Dryad Data Repository for deposition, publication, and permanent linking between data and articles, of datasets other than primary biodiversity data (e.g., ecological observations, environmental data, genome data and other data types).
- Automated archiving of all articles, published in Pensoft's journals in the Biodiversity Literature Repository (BLR) of ZENODO on the day of publication.

have been added in BDJ (de Jong et al., 2014, Biodiversity Data Journal, 2: e4034).

EU BON Open Science Publication Pilot

The journal Research Ideas and Outcomes (RIO) was designed to publish all outputs of the research cycle, from research ideas and grant proposals to data, software, research articles and research collateral, such as workshop and project reports, guidelines, policy briefs, and Wikipedia articles, to name just a few. Within RIO Journal, EU BON realised one of the first ever open science collections of project publications entitled: Building the European Biodiversity Observation Network (EU BON) Project Outcomes. Up to date, the collection contains 13 publications.

Contact

Lyubomir Penev: penev@pensoft.net