



## CALL FOR PROPOSALS

### GBIF Nodes activities to enrich metadata and promote the publication of data papers

Deadline for submissions:  
17th September 2012

Main Contact:  
Vishwas Chavan  
Senior Programme Officer for Content Mobilization  
GBIF Secretariat  
[vchavan@gbif.org](mailto:vchavan@gbif.org), +45 28751475

#### Background:

GBIF is committed to ensuring that its network provides access to high-quality data that will support biodiversity research and appropriate action in areas related to biodiversity conservation, management and use. This goal can only be met if the network includes rich, up-to-date metadata describing each data resource and enabling users to assess the fitness for use of these resources to meet their needs.

Significance of Metadata: Metadata are data elements that describe data records and the context in which they were developed. Good metadata should include: how the data were collected; the taxonomic, spatial and temporal scope of the dataset; information on the ownership and access arrangements for the digital dataset; data quality assessment metrics and information on any processing applied to enhance the data. Such elements enable users to assess the characteristics of each dataset and its suitability for their specific uses. Metadata can thus greatly enhance the usefulness, and consequently increase the real use, of the data.

However, comprehensive metadata descriptions exist for few biodiversity datasets. One reason for this is that there have been few incentives to justify the additional work for data publishers involved in creating appropriate metadata. The GBIF Data Publishing Framework Task Group therefore recommended that GBIF should introduce approaches that would provide such incentives (Moritz. et. al., 2011)<sup>1</sup>. One response has been the promotion of 'data papers' as a visible product of metadata for biodiversity datasets. A data paper is a scholarly publication derived from metadata and serving as an easily-cited representation of one or more datasets as a scientific resource. Data papers can be referenced via DOIs and consequently fit readily into the established models for citing scientific publications and recognizing their impact. Data papers are published in accordance with standard academic practices (Chavan and Penev, 2011)<sup>2</sup>.

This process brings together several interconnected elements, all of which support a culture of free and open publication of data and bring benefits to GBIF, to data publishers and to users of biodiversity data, for example:

---

<sup>1</sup> Moritz et. al.: Towards mainstreaming of biodiversity data publishing: recommendations of the GBIF Data Publishing Framework Task Group. BMC Bioinformatics 2011 12(Suppl 15): S1.

<sup>2</sup> Chavan and Penev: The data paper: a mechanism to incentivize data publishing in biodiversity science. BMC Bioinformatics 2011 12 (Suppl 15): S2.

- Data papers provide a natural way to bring the work of publishing datasets into the accepted framework for referencing scientific publications and measuring the impact of research activities;
- Effort directed at improving metadata can deliver a publication and simultaneously greatly enhance the value and usefulness of the associated data, which in itself is the best way to encourage users to find, use and reference the data;
- Moving towards DOIs as the standard identifiers for biodiversity datasets will contribute to stable, long-term management of datasets and of the data records contained within them;
- Effort aimed at providing interactive views for exploring and visualizing the contents of a data set, including maps, timelines, summary charts, statistical analyses, etc. can be integrated into standard metadata authoring and incorporated into styling for data papers submitted to peer-reviewed journals.

The following table summarizes some of the key milestones already reached in the use of data papers within the GBIF network:

Timeline	Activity
March 2011	The GBIF Integrated Publishing Toolkit v.2.0 was released. This implemented the 'GBIF Metadata Profile' for authoring metadata.
May 2011	The IPT v.2.0.1 release facilitated auto-generation of data paper manuscripts (.RTF format) from authored metadata. Pensoft Data Publishing Policies and Guidelines for Biodiversity Data were released <sup>3</sup>
June 2011	Six of the Pensoft electronic journal titles (BioRisks, MycoKeys, Nature Conservation, NeoBiota, PhytoKeys, and ZooKeys) started accepting data paper submissions.
November 2011	First data paper published in ZooKeys (Narwade et.al., 2011) <sup>4</sup>
April 2012	Second data paper published in ZooKeys (Danis et.al., 2012) <sup>5</sup>
May 2012	First data paper published in PhytoKeys (Landuyt et. al., 2012) <sup>6</sup>
	Third data paper published in ZooKeys (Shao et. al., 2012) <sup>7</sup>
June 2012	Fourth data paper published in ZooKeys (Pierrat et.al., 2012) <sup>8</sup>

Following these experiences and insights gained from the data paper publication process, the GBIF Secretariat is drafting a best-practice guide for authoring and publishing data papers. Release of this document is expected in the next few weeks.

GBIF Nodes have a vital role to play in promoting an effective culture of data publishing with high-quality data and metadata, and in the use of data papers as a tool to achieve this goal. The GBIF Secretariat is therefore issuing this **call for GBIF Nodes activities to enrich metadata and promote the publication of data papers.**

<sup>3</sup> Penev et. al.: Pensoft Data Publishing Policies and Guidelines for Biodiversity Data. Pensoft Publishers, [http://www.pensoft.net/J\\_FILES/Pensoft\\_Data\\_Publishing\\_Policies\\_and\\_Guidelines.pdf](http://www.pensoft.net/J_FILES/Pensoft_Data_Publishing_Policies_and_Guidelines.pdf).

<sup>4</sup> Narwade et. al.: Literature based species occurrence data of birds of northeast India. ZooKeys 150: 407-417.

<sup>5</sup> Danis et. al.: Antarctic Starfish (Echinodermata: asteroidean) from the ANDEEP3 expedition. ZooKeys 185: 73-78.

<sup>6</sup> Landuyt et.al.: Florabank 1: a grid-based database on vascular plant distribution in the northern part of Belgium (Flanders and the Brussels Capital region). PhytoKeys, 12: 59-67.

<sup>7</sup> Shao et. al.: A dataset from botton trawl survey around Taiwan. ZooKeyes 198: 103-109.

<sup>8</sup> Pierrat et. al.: Antarctic, Sub-Antarctic and cold temperate echinoid database. ZooKeys 204: 47-52.

### Key documents:

Before submitting a proposal, interested Nodes are encouraged to read following documents to familiarize themselves with concepts and processes associated with metadata authoring and data paper publishing.

What is a data paper?	Chavan and Penev (2011). <a href="#">The data paper: a mechanism to incentivize data publishing in biodiversity science. BMC Bioinformatics, 12 (Suppl 15): S2.</a>
How to publish a data paper	Penev et. al. (2011). <a href="#">Pensoft data publishing policies and guidelines for biodiversity data.</a>
Data papers describing faunal datasets	<a href="#">Narwade et. al.: Literature based species occurrence data of birds of northeast India. ZooKeys 150: 407-417.</a> <a href="#">Danis et. al.: Antarctic Starfish (Echinodermata: asteroidean) from the ANDEEP3 expedition. ZooKeys 185: 73-78.</a> <a href="#">Shao et. al.: A dataset from bottom trawl survey around Taiwan. ZooKeys 198: 103-109.S</a> <a href="#">Pierrat et. al.: Antarctic, Sub-Antarctic and cold temperate echinoid database. ZooKeys 204: 47-52.</a>
Data papers describing floral datasets	<a href="#">Landuyt et.al.: Florabank 1: a grid-based database on vascular plant distribution in the northern part of Belgium (Flanders and the Brussels Capital region). PhytoKeys, 12: 59-67.</a>

### The call

*Through this call, the GBIF Secretariat invites proposals from the GBIF Nodes for activities to engage with data publishers for (a) delivery of enriched metadata for existing and new datasets, and (b) promotion of data papers as an incentive for metadata publishing.*

**Purpose:** The objective of this call is to enable Nodes to promote the culture of rich metadata authoring and data paper publishing amongst their existing and potential data publishers.

**Potential approaches:** We foresee Nodes employing variety of approaches, which may include:

1. Developing promotional material (print, online, audio-video) emphasizing (a) the need for metadata, (b) how to author metadata and data papers, and (c) benefits of data paper publishing;
2. Organizing meetings/workshops/training of key data publishers to promote metadata authoring and data paper publishing;
3. Authoring metadata and data papers for select datasets already published by the Node itself or its existing data publishers;
4. Organizing national and/or thematic competitions for metadata and data papers derived from existing (already published) and new data;
5. Engaging with academic publishers to introduce a data paper section in their journal titles;
6. Acting as a 'champion' Node, supporting other GBIF participants and data publishers in authoring rich metadata and data paper publishing.

**Target audience:** GBIF Nodes, biodiversity data publishers, information managers, academic journal editors and publishers, and funding agencies.

**Anticipated impacts:**

- a. Increased awareness of the significance of metadata and of data papers;
- b. Accessibility to enriched metadata for datasets published through the GBIF network;
- c. Due recognition to the efforts of those involved in data publishing.

**Catalytic support:** In order to encourage Nodes to undertake these activities as part of their data discovery and mobilization efforts, catalytic support will be provided to successful proposals. This will be in the range of €2000 to €5000, depending on the nature of the activities and their anticipated impacts.

**Specifications and procedures:**

The GBIF Secretariat invites concise proposals from GBIF Nodes (voting and associate participants) endorsed and forwarded by the Node Manager. Individuals, data publishers and research groups experienced in metadata authoring and data publishing can approach relevant GBIF Nodes and team up with them to submit a proposal.

**Value of support:** Requested support for each proposal should not exceed € 5000.

**The proposal should include:**

- a. Description of proposed activities;
- b. Expected impact;
- c. List of potential beneficiaries;
- d. List of datasets for which enriched metadata will be authored, and for which data papers will be published;
- e. Ideas and plans for mainstreaming metadata authoring and data paper publishing beyond the catalytic support phase.
- f. Brief CVs of institutions/individuals/experts involved, and their specific roles.

**Collaborations:** Given the nature of the call and the activities involved, close long-term collaborations among Nodes and/or relevant agencies/networks are encouraged. Several Nodes may jointly present a single proposal.

Please note that no overheads can be charged.

Proposals should be submitted by email to Dr. Vishwas Chavan, Senior Programme Officer for Content Mobilization ([vchavan@gbif.org](mailto:vchavan@gbif.org)) and will be accepted up to 15th September 2012.

Deadline for submitting a proposals	17 <sup>th</sup> September 2012
Evaluation and selection of proposals	1 <sup>st</sup> October 2012
Notification of selected proposals	15 <sup>th</sup> October 2012
Release of award letters for selected proposals	30 <sup>th</sup> October 2012