# Data sharing tools for Biodiversity Observation Networks

Larissa Smirnova [1], Patricia Mergen [1,2], Quentin John Groom [2], Hannu Saarenmaa [3], Pavel Stoev[4,5], Lyubomir Penev [5], Israel Pe'er [6], Veljo Runnel [7], Antonio García Camacho[8], Donat Agosti [9], Aaike De Wever [10], Timothy Vincent[11], Christos Arvanitidis [12]

*1 Royal Museum for Central Africa, Tervuren, Belgium; 2 Botanic Garden Meise, Meise, Belgium; 3 University of Eastern Finland, Joensuu, Finland; 4 National Museum of Natural History; 5 Pensoft Publishers & Bulgarian Academy of Sciences, Sofia, Bulgaria; 6 GlueCAD, Haifa, Israel; 7 University of Tartu, Tartu, Estonia; 8 CSIC, Spanish Council for Scientific Research, Seville, Spain; 9 Plazi, Bern, Switzerland; 10 Royal Belgian Institute of Natural Sciences, Brussels, Belgium; 11 INPA - National Institute for Amazonian Research, Manaus, Brazil; 12 Hellenic Centre for Marine Research (HCMR), Heraklion Crete, Greece*

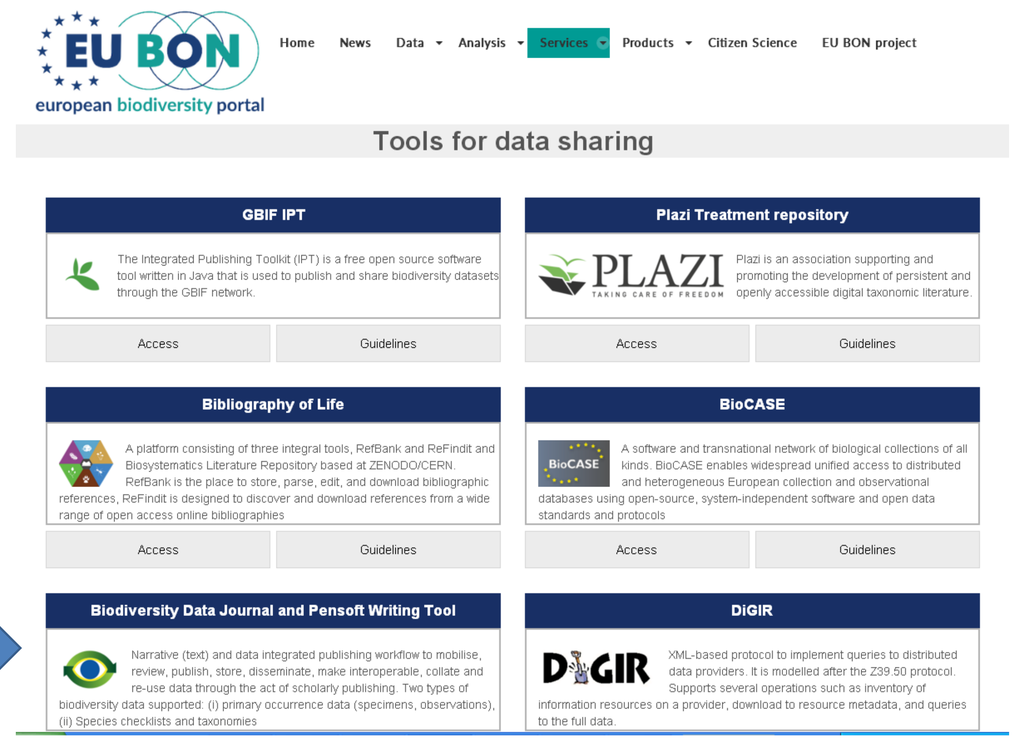
www.eubon.eu

It has been the policy of the EU BON project that it is better to promote and continue development of preexisting tools, rather than creating new one. This approach limits the fragmentation of the infrastructural environment and leverages former investments in software and training.

The recent study done by the EU BON project on available data sharing and data publishing tools used in the natural history domain, resulted in comprehensive report D2.2 which will be published soon in RIO journal.

This is done from the perspective of the needs of the biodiversity observation community with an eye on the development of a unified user interface to these data – the European Biodiversity Portal (EBP).

About 30 data sharing tools have been evaluated and the results of these assessments are presented in the report and is also available online.



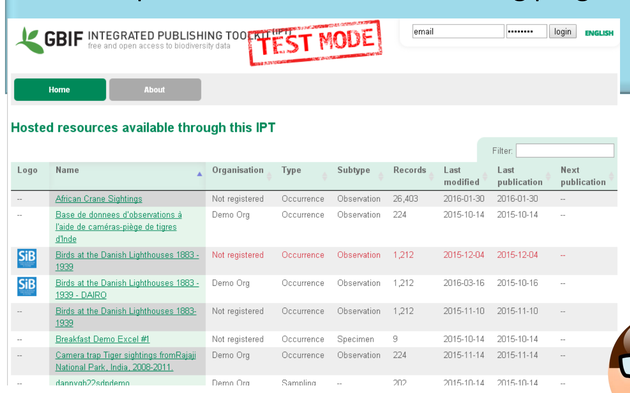

To accelerate data mobilization towards more comprehensive data coverage the EU BON has focused its efforts on citizen science recording and monitoring schemes, but also collection-based data, taxonomic data from scientific publications (including historic data) by mounting the targeted approaches to fill gaps in temporal, taxonomic, and geographic coverage.

To fully meet the objectives of the project and the user requirements for different types of data, a combination of tools have been selected for deployment.

**The Integrated Publishing Toolkit (IPT)**:

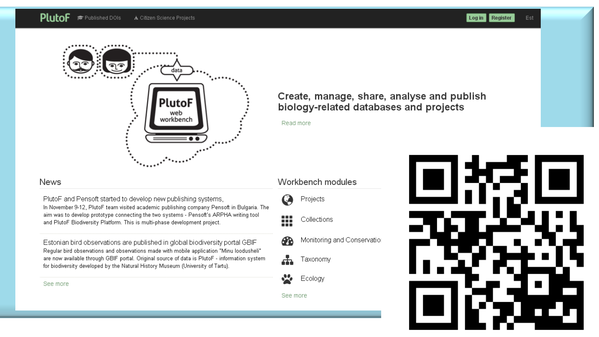
To publish and share biodiversity data sets and metadata through the GBIF network. It allows publication of three types of biodiversity data:

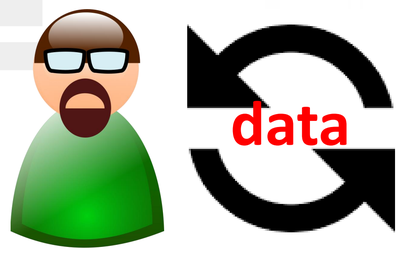
- primary occurrence data (specimens and observations);
- species checklists and taxonomies;
- sample-based data from monitoring programs.



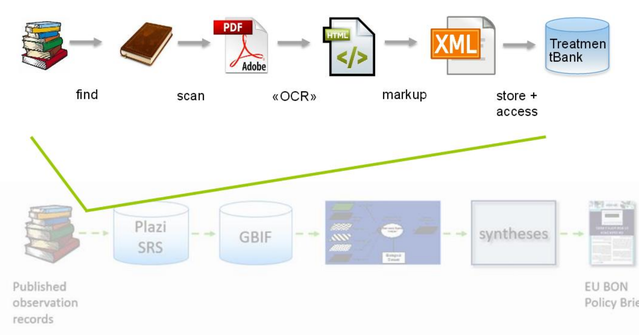

**PlutoF API:**

An online service to create, record, manage, share, analyze and mobilize biodiversity data. Data types include ecology, taxonomy, metagenomics, nature conservation and natural history collections including citizen science projects.





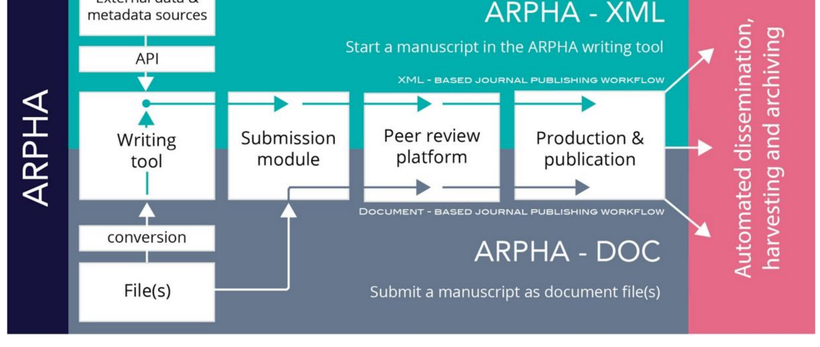




**The ARPHA Publishing Platform:**

Narrative (text) and data integrated publishing workflow to mobilize, review, publish, store, disseminate, make interoperable, collate and re-use data through the act of scholarly publishing.

**TreatmentBank:**

A platform to store, annotate, access and distribute taxonomic treatments and the data objects within them. It works with GoldenGate and XML schemas TaxonX and TaxPub, which are tools to convert unstructured text into semantically enhanced documents with an emphasis on taxonomic data such as treatments, scientific names, material observation, traits and bibliographic references.

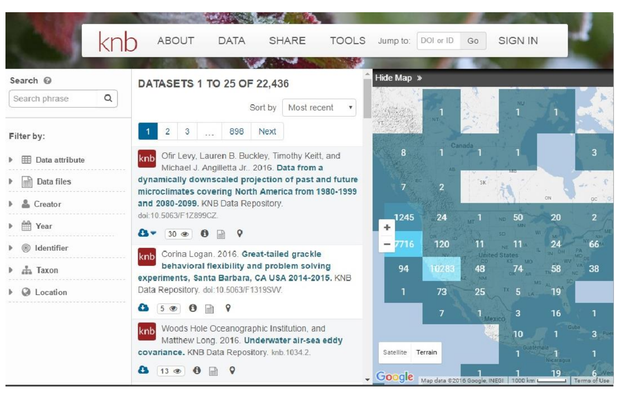



**Spreadsheet tools:**

1 - GBIF Spreadsheet processor is a web application that supports publication of biodiversity data to the GBIF network using pre-configured Microsoft Excel spreadsheet templates;

2 - DataUp is the tool developed by DataOne to help environmental scientists to upload files to a repository for data management;

3 - OpenRefine is recommended for data clean-up and transformation to other formats.




**Metacat and Morpho:**

Metacat is a repository that helps scientists store metadata and data, search, understand and effectively use the data sets they manage or those created by others. A data provider using Metacat can become a DataONE member node with a relatively simple configuration. Morpho is an application designed to facilitate the creation of metadata.

This project has received funding from the European Union's Seventh Programme for research, technological development and demonstration under grant agreement No 308454