

Deliverable D1.1

Version: 1
Date: 10/07/20
Author: RBINS
Dissemination status: PU
Document reference: Deliverable_D1.1



Freshwater data portal enabling querying of freshwater biodiversity data

STATUS: FINAL VERSION

Project acronym: BIOFRESH
Project name: Biodiversity of Freshwater Ecosystems: Status, Trends, Pressures, and Conservation Priorities
Call and Contract: FP7-ENV-2008-1
Grant agreement no.: 226874
Project Duration: 01/11/2009 – 30.04.2014 (54 months)
Co-ordinator: Leibniz-Institute of Freshwater Ecology and Inland Fisheries at Forschungsverbund Berlin e.V., Germany

Partners: RBINS, Royal Belgian Institute of Natural Sciences, Belgium
BOKU, Universität für Bodenkultur Wien, Austria
ICLARM, International Center for Living Aquatic Resources Management, Malaysia
IRD, Institut de Recherche pour le Développement, France
UDE, Universität Duisburg-Essen, Germany
IUCN, International Union for Conservation of Nature, Switzerland
UOXF.AC, Oxford University, UK
UB, Universitat de Barcelona, Spain
UFZ, Helmholtz Zentrum für Umweltforschung, Germany
UCL, University College of London, UK
EAWAG, Eidgenössische Anstalt für Wasserversorgung, Abwasserreinigung und Gewässerschutz, Switzerland
UCBL, Université Claude Bernard - Lyon 1, France
UPS, Université Paul Sabatier- Toulouse 3, France
ECOLOGIC, Ecologic GmbH Institut für Internationale und Europäische Umweltpolitik, Germany
EC-ERC, Commission of the European Communities - Directorate General Joint Research Centre, Italy
UD, University of Debrecin, Hungary
NRM, Naturhistoriska riksmuseet, Sweden
CFKK, Center za kartografijo favne in flore, Slovenia

This project is supported by funding from the specific programme 'Cooperation', theme 'Environment (including Climate Change)' under the 7th Research Framework Programme of the European Union.



BIOFRESH

Biodiversity of Freshwater Ecosystems: Status, Trends, Pressures, and Conservation Priorities


Project no. 226874

Large scale collaborative project

Freshwater data portal enabling querying of freshwater biodiversity data

Deliverable number	D1.1
Deliverable name	Freshwater data portal enabling querying of freshwater biodiversity data
WP no.	WP 1
Lead Beneficiary (full name and Acronym)	Royal Belgian Institute of Natural Sciences , RBINS
Nature	Others: Database
Delivery date from Annex I (proj. month)	M9
Delivered	yes
Actual forecast delivery date	2010-07-20
Comments	

Project funded by the European Commission within the Seventh Framework Programme Dissemination Level		
PU	Public	✓
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Name of the Authors	Name of the Partner	Logo of the Partner
Aaike De Wever	RBINS	
Hendrik Segers	RBINS	"
Koen Martens	RBINS	"

In case the report consists of the delivery of materials (guidelines, manuscripts, etc)

Delivery name	Delivery name	From Partner	To Partner

Introduction

WP1 (lead RBINS) is in charge of the creation of the BioFresh biodiversity data portal. This web-based science portal will be designed for providing users free and universal access to information on freshwater biodiversity, contained in a distributed network of interoperable databases. The data portal will integrate tools for web-based geospatial visualisation and analysis tools (incl. predictive models developed in the science workpackages) and will contribute to the development of a collaborative, global infrastructure for biodiversity assessment. As such, we will enable water managers to incorporate biodiversity science into effective conservation strategy, environmental management or related policy instruments.

State-of-the-art

Development of the BioFresh data portal started in March 2010. The progress of the RBINS team is documented on on-line wiki < <http://trac.bebif.be/wiki/BioFresh/>>, which covers technical aspects, the different elements to be considered for the portal (portal content), database structure, the web interface and planning of different tasks.

Technical underpinnings

The databases for the BioFresh data portal are built using PostgreSQL. Web applications using this database are built using the Grails web application framework and the Groovy and Java programming languages. For deployment, we rely on the Tomcat web application container. The construction benefited from experience gained using these technologies for the construction of the Antarctic SCAR-MarBIN data portal < <http://www.scarmarbin.be/>>.

Database structure

The database structure for the BioFresh data portal is built around a central species registry. The idea behind this is that all data sources will be accessed and linked through this central "species registry". Starting from the registry, information available for the specie(s) searched for will be resolved through the species names (and synonyms) or using the ID used in the foreign database (Fig. 1).

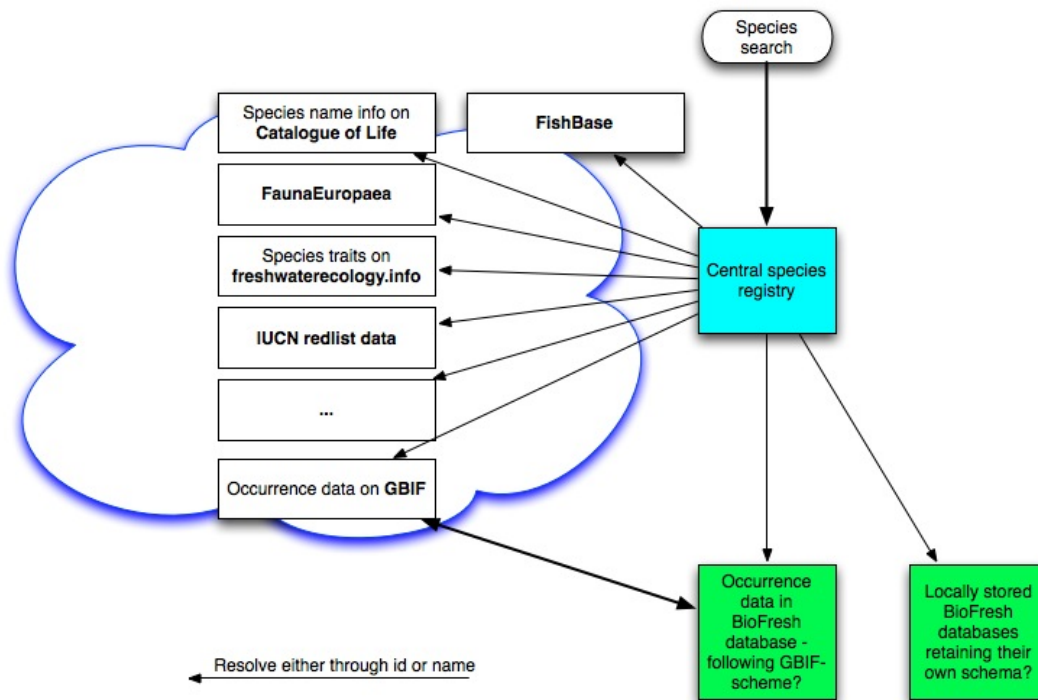
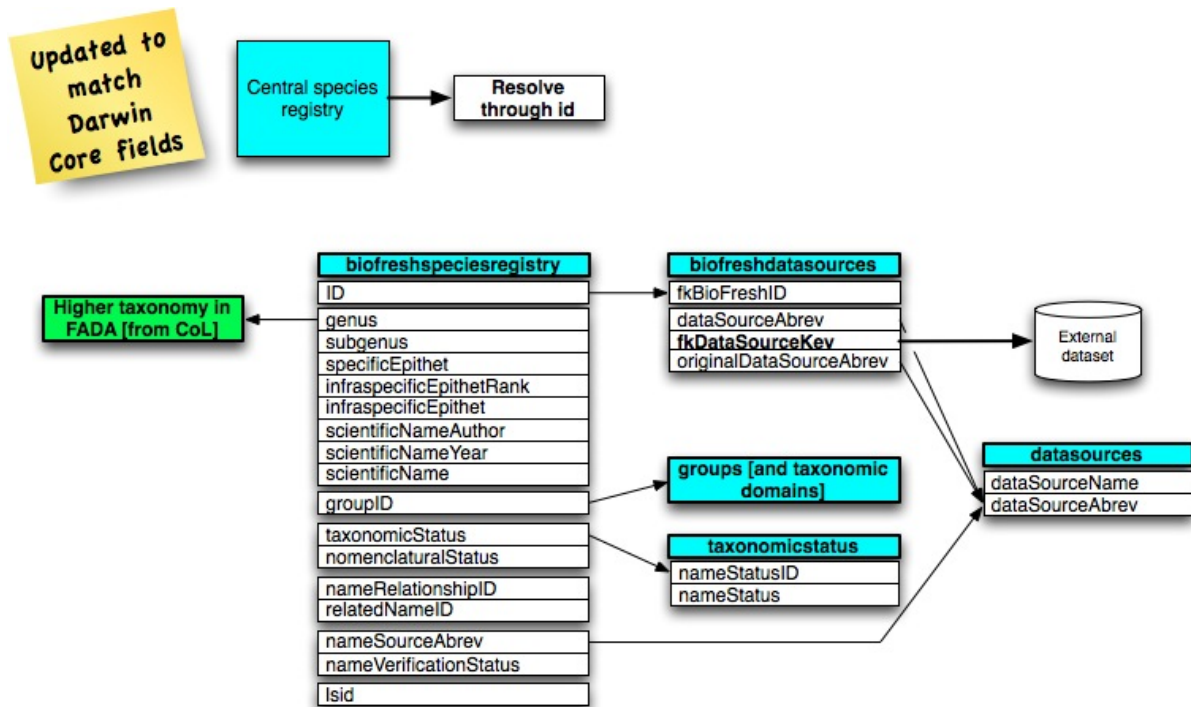


Fig. 1: Concept of database design

The design of the central species registry, which is organized around two main tables ("BioFreshSpecies" and "BioFreshInfoSources"), is shown in Fig. 2. The first table is the actual registry, the second table is intended to link up info sources to the species. Full documentation for the fields is given at < <http://trac.bebif.be/wiki/BioFresh/DatabaseStructure>>.



Example data

[not all fields are shown]

biofreshspeciesregistry						
ID	genus	species	taxonomicStatus	nameRelationship[ID]	relatedNameID	nameVerificationStatus
1	<i>Pristis</i>	<i>pristis</i>	accepted			verified by ...
2	<i>Pristis</i>	<i>antiquorum</i>	synonym	accepted synonym of	1	verified by ...

biofreshdatasources			
fkBioFreshID	dataSourceAbrev	fkDataSourceKey	originalDataSourceAbrev
1	FB	8940	FB
1	CoL	7599043	FB

datasources	
dataSourceName	dataSourceAbrev
Fishbase	FB
CoL	Catalogue of Life

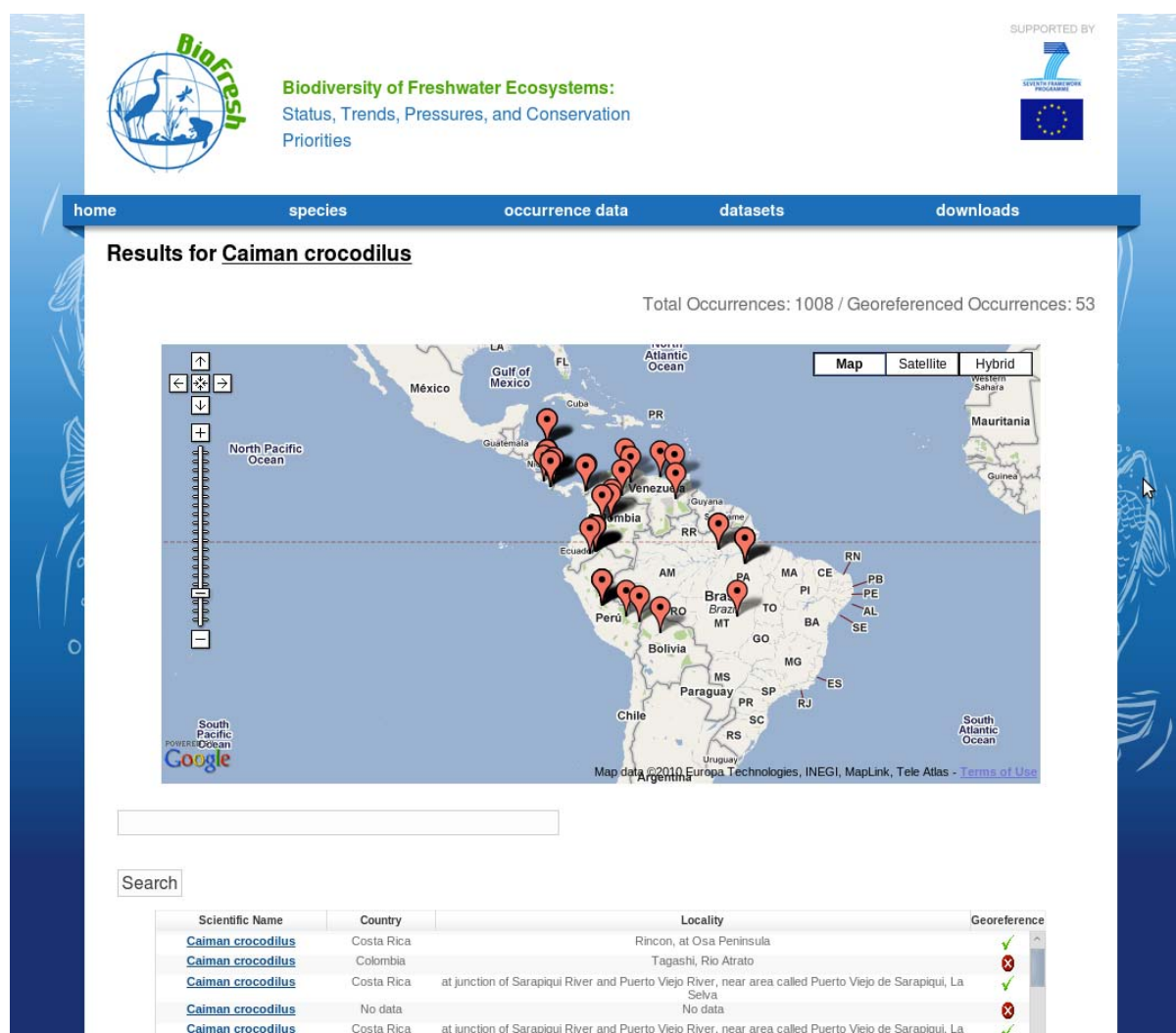
Fig. 2: Database structure for the central species registry.

For external datasets, we will either retain the original database structure (e.g. FADA) or store the data in a standard format for occurrence records (conforming Global Biodiversity Information Facility [GBIF] data). The basic database structure for the later was obtained from the Belgian GBIF-node.

User interface and portal content

The main design for the data portal was copied from the project website < <http://www.freshwaterbiodiversity.eu/>> in order to retain a consistent look. Different elements for the data portal considered for a first phase include; access to relevant freshwater GIS-layers, species occurrence and distribution data, species information from FishBase, authoritative species lists, example maps and tools, access to the BioFresh metadatabase, documentation including workflow example, usage agreement, call for providers and site statistics. More information on these can be found on < <http://trac.bebif.be/wiki/BioFresh/PortalContent>>.

For the first version of the data portal, we implemented a selection of the planned content and we will gradually add more elements, data and tools. The content for the version to be launched in July 2010 includes all static content, basic authoritative species information and visualisation of the occurrence records (see Fig. 3).



First interoperable databases

For the development of the first version of the data portal, we mainly focussed on the design of the underlying database structure and technology. The datasets that are currently included in the portal should therefore be regarded as datasets to illustrate the current system, which we hope to improve

based on feedback on this early version. For filling the central species registry, we mainly referred to the Freshwater Animal Diversity Assessment (FADA)-database, while occurrence records were retrieved from GBIF (www.gbif.org).

The main source of authoritative species lists of freshwater organisms is the Freshwater Animal Diversity Assessment (FADA) Project (<http://fada.biodiversity.be/>), which currently holds data on Halacaridae, Hydroida, Cladocera, Copepoda, Mysidacea, Ephemeroptera, Plecoptera, Bivalvia, Rotifera, and Fish. For the current release, we integrated FADA-data for the groups for which we have occurrence data available.

The first set of occurrence data was obtained and processed for demonstration purposes. It consists of the result of a query on a list of freshwater megafauna (>45 kg) on the global network of databases connected through GBIF. In total, we obtained over 28'000 occurrence records for 116 species. We currently have the tools available to quickly retrieve, cache and show GBIF data on the portal.

Tools

For the first version of the data portal, we included basic mapping tools to show the occurrence records. Future tools and models to be integrated in the portal will be evaluated based on the priorities set by and the developments from the science workpackages.

Launch

The data portal was deployed on August 16th 2010 and is be accessible using the URL: <http://www.freshwaterbiodiversity.eu/dataportal>. After a review period by the BioFresh project members, we will configure the subdomain <http://data.freshwaterbiodiversity.eu> for this purpose.

Future developments

Obviously the first version of the data portal to be released in July 2010 is an early preview and there's much more to come. As mentioned in the state-of-the-art, we will gradually add more of the elements, data and tools mentioned under "User interface and portal content". Below we list some issues that will be dealt with in the near future:

- Integrate all FADA data
- Offer possibility for data download
- Provide access to freshwater data available in FishBase
- Move maps to geoserver or mapserver
- Integrate shapefile tools on maps
- Show IUCN redlist status and link to the info at that site
- Link to taxonomic data at Catalogue of Life
- Start working on some databases available within BioFresh