Databases

An all-round database system for biological information

By Tim Deprez

ALL BIOLOGICAL SPECIES are characterised by very different kinds of data: written species descriptions, drawings, pictures, movies, geographical information, molecular and genetic information, systematic positions, phylogenetic relationships, ecological information, etc. Until now, there has been no all-round system available which would help in opening up this biological information for a given taxon.

NeMys is a tool created for unravelling all kinds of biological information using an integrated database approach, but based and linked (!) on information which is available in the literature (although found in very different formats).

Seven years ago, the Marine Biology Section of Ghent University (Belgium) started a project aiming to collate biological data on marine nematodes digitally. This small project has evolved into the online generic biological information system NeMys (http://www.nemys.ugent.be), which is now already used for seven different kinds of taxa (such as nematodes, mysids and amphibians). As the system is set up in a generic way, the database is and can be used for many different taxa grouped in separate datasets.

The bits and bytes

NeMys is a web-based tool combining several types of biological information. A systematical backbone is therefore linked to a digital library and the individual scientist responsible for the data.

All data in NeMys are linked to taxa based upon a data-source, more often than not an electronic reference in the digital library. This link is one of the main features which distinguishes NeMys from other information systems.

All datasets have a set of fixed data types: distribution data, morphological data, pictorial data, collection data, literature data and molecular data. In addition, an unlimited set of data fields can be added for each of the different datasets (e.g. habitat characteristics, feeding type, sounds, etc).

All data is displayed to the users through a number of easy-to-use web interfaces. Morphological data can be displayed in tabular format, and can be at the basis of interactive polytomous identification keys, which are easily created through this system. In addition to this, phylogenetic analyses can be performed starting with the raw data information.

Geographical data is used to make online maps which can be exported to desktop applications such as Google Earth™ as well. Links to a set of biodiversity portals (EurObis, GBIF) are easily created.

Some of the technical characteristics of NeMys are:

- A true online-managed system: users can (according to their rights) enter, edit and consult data online. General integrated scientific tasks and management tasks (such as creation of identification keys, user management, interface layout, etc) are performed through web-based modules.
- A fully digital system: all data is stored digitally (pictures, movies, personal notes, etc).
- A genuine scientific research tool: all data is connected to the original data-source.
- An online work-tool: a private section enables registered users to enter unpublished research data, which could be shared with other users of the system (if required).
- Online mapping tools, which allow true biogeographic scalable analyses and are available for all users; these maps can be saved in a user's own private section.
- A generic system: the whole system (including identification keys, glossaries, mapping tools, etc) can be used for any group of biological organisms.

Marine data

At the moment, two main datasets on marine taxa are available, dealing with marine nematodes (6,900 taxa) and mysid shrimps (1,675 taxa).
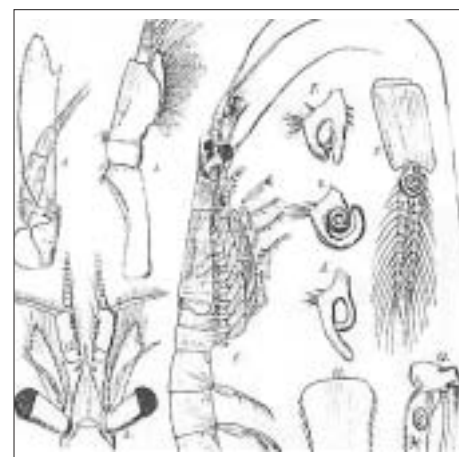
Siriella armata (top), *Molgolaimus turgofrons* (© Marine Biology Section, Ghent University).

About 15,000 biogeographical records are available, mainly based upon published literature. Most taxa are linked to illustrations and data from the original species descriptions.

All documents are searchable through an interactive literature index. At the moment, the majority of the mysid species and about one-third of the nematoda dataset are linked to distribution data.

NeMys in MarBEF

The marine nematode dataset is considered to be the backbone for meiobenthic research in the MarBEF MANUELA RMP. Within MANUELA, NeMys acts as a tool for creating identification keys for free-living marine nematodes. Species information (taxonomical and ecological) for most of the European marine nematodes is added systematically in order to facilitate meiobenthic research within the MarBEF community and worldwide.

The private user section through which data can be shared within a small set of users will be intensively used in trying to create consistency between unidentified species in different datasets (e.g. *Chromadora* spec a will become unique *Chromadora* sp. a,b,…,n in the integrated dataset). Emphasis will be given in the first phase to congeneric species with a so-

called 'cosmopolitan' distribution. All available information resulting from investigations into these species will be facilitated throughout the database structure presented in NeMys.

During the kick-off meeting of the MarBEF RMP MANUELA consortium (Ghent, September 2005), NeMys was accepted as the online identification tool for marine nematodes by the specialists present.

NeMys will also be used for updating the European Register of Marine Species (ERMS) through the MarBEF RMP PROPE-taxon project (decided in Crete, October 2005).

The Marine Biology Section of Ghent University contributes to EurOBIS through the mysid and nematode datasets available in NeMys.

Dracognomus sp. (© Marine Biology Section, Ghent University, Belgium).

Some statistics about NeMys...

- A recent analysis of the user community revealed that the combination of different kinds of biological data in one system attracts a wide range of users. About 800 registered users (originating from 60 countries), half of them academic, half with a non-professional background, are using the different aspects of the system on a regular basis.

- Future developments will include extra data-input, data-analysis and data-exchange tools, integration of the system in educational projects, and the start-up of new marine and non-marine datasets.

Tim Deprez, Marine Biology Section, Biology Department, Ghent University, Krijgslaan, 281-S8, B-9000 Ghent, Belgium. Email: Tim.deprez@ugent.be http://www.nemys.ugent.be

FEATURES

Taxonomy

European Distributed Institute of Taxonomy (EDIT)
– a new Network of Excellence

By Geoff Boxshall

ON 1ST MARCH THIS YEAR another Network of Excellence was signed into existence, this one dedicated to taxonomy. EDIT is a consortium of 26 organisations under the leadership of Simon Tillier (Muséum National d'Histoire Naturelle, Paris) that are working to change the ways that taxonomy is done. So, stripping away the Eurospeak, what is in this for participants of MarBEF?

EDIT aims to do nothing less than revolutionise the way taxonomists work, first by encouraging collaboration and cooperation and by providing a computing environment with suitable tools to make the whole task easier, and second by changing the way taxonomists publish their work, putting the information directly onto the web and providing context-sensitive portals to return the information users actually need.

The heart and soul of EDIT is in its desire to reduce the fragmentation in European taxonomic research and expertise. We want to co-ordinate the European contribution to the global taxonomic effort, in particular the Global Taxonomy Initiative, through an integrated initiative aimed at improving society's capacity for biodiversity conservation. EDIT focuses on improvement of both production and delivery of taxonomy through:

1. coordination of the research policies of its member institutions, which employ altogether around 1,500 researchers and doctoral students in taxonomy;

2. progression toward integration of their scientific expertise and infrastructures to improve both production and access to taxonomic knowledge and information, within the network as well as in the framework of international structures and initiatives, such as the GBIF and the GTI;