



PLANKTON*NET workshop on long-term perspectives for online taxonomic resources, Bremerhaven April 20th-April 21st

1. Workshop Report

Workshop Aims

The principal aim was to place the EU PLANKTON*NET project in the context of existing funded initiatives and to establish areas for co-operation with these initiatives and projects (e.g. ICOMM, Marbef), that will help to focus work efforts and will facilitate the emergence of a resource that can be maintained in the long-term.

Workshop Activities

Day 1:

1. The workshop started with a brief summary of the main outcome and agreed items of the preceding PLANKTON*NET workshop. The following points were mentioned:

- AWI has updated its existing PLANKTON*NET node to the latest version of the MICRO*SCOPE software. Together with the MBL and PLANKTON*NET partners, AWI will adapt the current PLANKTON*NET to accommodate environmental data
- 3 PLANKTON*NET nodes will be installed, based on MICRO*SCOPE version 6
- The nodes will be linked via OAI-PMH and web services
- We will use Darwin core as metadata schema
- Xid keys will be created e.g. for: dinoflagellates, coccolithophores, HABs and a general key to the different groups
- The search functionality of the existing PLANKTON*NET will be extended.
- A portal for the entry of taxonomic descriptions will be created
- Data archaeology (non-digitized collections): Incorporation of such collections into PLANKTON*NET will become one component of PLANKTON*NET activities.

Edward vanden Berghe (VLIZ/Marbef) commented that Darwin Core was not really developed for use with taxonomic data although it would be a useful extension of its

application. Extensive testing of the new PLANKTON*NET system underpinned with Darwin Core was recommended.

2. A talk outlining the latest IT changes in MICRO*SCOPE was then presented by David Patterson of the Marine Biological Laboratory at Woods Hole. One of the main features is the concept of 'taxonomic intelligence'. This term refers to software tools that attempt to overcome the problems associated with taxonomic names (e.g. the same name for different species, one species known under different names). Several components of taxonomic intelligence were mentioned

- i. Lists of all species names: These form the basis of taxonomically intelligent software tools.
- ii. Reconciliation: If one organism has several names the alternative names are mapped against each other
- iii. Disambiguation: This is necessary when the same name is used for different organisms
- iv. Hierarchical classifications

The 'Biodiversity Heritage Library Project' was also outlined. This project aims at scanning and publishing online extremely large volumes of information (mostly older journals and books) at several library sites. The resulting 'online library' can be scanned using taxonomically intelligent tools. A demonstration of this was given.

3. In a series of talks invited representatives of European biodiversity initiatives then presented their work (see also Annex 1):

Edward vandenBerghe (VLIZ/Marbef) gave an overview of Marbef data management activities. Tools for the management of biogeographical data management and taxonomic data were introduced e.g.:

European Marine Gazetteer (part of the VLIZ Marine gazetteer): This integrates the different geographic names in the Marbef data pages

EUROBIS: a system for simultaneous searches for biogeographic information on organisms in multiple databases

ERMS (European Register of Marine species): ERMS is an authoritative list of species that is updated regularly by a board of editors. The register can be updated and edited online.

ALGATERRA A further talk explained the main features of the online taxonomic database ALGATERRA is a database providing the names of algae (freshwater, marine and terrestrial) and evaluated information on synonyms and taxonomic concepts. It is co-ordinated at the Botanical Garden Berlin (BGBM) which also co-ordinates the contribution of the botanical node of GBIF-Germany. ALGATERRA is concentrating on collections of type material, but also contains ecological information. The database also contains some taxonomic descriptions but this is not a main focus. The data model used is the 'Berlin Taxonomic Information Model' with project specific extensions (<http://www.bgbm.org/biodivinf/docs/bgbm-model/>)

Fishbase: Fishbase is an online encyclopaedia with information on various aspects of fish biology, ecology and taxonomy. Rather than containing all data in one central resource, Fishbase has several mirror sites around the world, information in which is exchanged monthly by DVD. Database content is also available in a database on DVD. Importantly, Fishbase provides multilingual access with machine translation of site content. Implementation of the multilingual facilities caused a considerable increase in the number of hits on the Fishbase website.

ETI Bioinformatics: The Linnaeus II software was introduced. This is a freely available data management tool for taxonomists. The software includes a taxonomically oriented database, literature and glossary databases and identification keys. Several Linnaeus II tools are available including a builder and web publisher tool for creating biodiversity information either offline or online. A portal kit for creating dynamic links with other databases is also available (Buffie) and ETI is working on the development of interoperability tools for supporting GBIF. In addition, ETI develops the 'Catalogue of Life' Annual checklist software.

Several areas for concrete collaborations were established as the result of these presentations:

1. The Flanders Marine Institute (also a member of the Network of Excellence Marbef) is already offering mapping tools for species data and a collaboration with VLIZ will be established to add this functionality to the PLANKTON*NET resource.
2. During the kick-off meeting it was decided that Multilanguage facilities should be set-up for PLANKTON*NET. The online resource FISHBASE is already employing machine translation tools on its site and the feasibility of establishing similar tools on PLANKTON*NET will be explored

3. The PLANKTON*NET project will see the construction of a small library of taxonomic keys. Links will be established with the PROPETAXON responsive mode project in Marbef to jointly foster the development of further web based taxonomy resources. ETI Bioinformatics will also be involved in the process.

Additionally 2 talks presented the latest IT developments and future plans for the MICRO*SCOPE resource at the MBL and PLANKTON*NET respectively

The strength of using the ubio (<http://www.ubio.org>) biological name server for organising information about organisms was demonstrated. The potential for assembling information on a given taxon from multiple resources via the uBio portal was explained. By indexing the different names referring to the same taxon in different web sites, information for all these different names will be included in the search results. Remote applications can use ubio SOAP services to access the ubio data

In the 2. talk the activities to be carried out during the PLANKTON*NET EU proposal were outlined. All PLANKTON*NET nodes (AWI, Roscoff and Lisbon) will be made compliant with the Open Archives initiative – Protocol for metadata harvesting (OAI-PMH, www.openarchives.org). All nodes will be data providers with data being harvested into a central repository. In addition to OAI-PMH, the central repository will also be exposed via SOAP. The advantages of using OAI-PMH were described e.g. the possibility of harvesting of the repository by other OAI compliant resources (creates redundancy which aids long-term preservation)

Day 2:

To conclude the series of talks from day 1 an update was given by Robert Ptacnik of NIVA on the EU project Rebecca (Relationships between ecological and chemical status of surface waters)

More short talks were held, this time concentrating on user requirements (what is missing in present resources, what additional data might we have to accommodate?) These talks were given by two potential endusers Claus Dürselen of AQUAECOLOGY (Germany) and Teresa Moita of IPIMAR, Portugal.

Talk by AQUAECOLOGY: Several examples were given of information not yet easily obtainable from existing online resources: **a.** several images of the same species e.g. different orientations of the same dinoflagellate cell showing the plate patterns, **b.** illustrations showing taxonomic detail. Biovolumes were mentioned as an additional parameter that would be very useful in online taxonomic databases, possibly with tools for calculating them for different geometric shapes

A further item that emerged was the need for properly quality controlled data. A concept for this still has to be developed fully for PLANKTON*NET. Different degrees of quality control could be adopted including full editorial boards, very loose quality control regimes (similar to Wikipedia) or ad hoc quality control of items that have been identified by administrators as doubtful. This would mean however, that a degree of quality control has to be carried out by site owners rather than providers of data to these sites.

Discussion of financing issues:

Prior to the open discussion, two talks were held considering very different models for the long-term financing of these resources and the different options for the involvement of private companies (SMEs) in such projects. At the AWI and many other institutes 'open access' is one of the principles underpinning the online publication of data irrespective of whether one is dealing with images, literature or biological data and the importance of keeping these principles in mind when developing financing strategies for online taxonomic resources was stressed.

Details of talks:

Thomas Raabe stressed the basic premise that core online data have to be accessible free of charge but also that a permanent management/ maintenance budget is needed to maintain a given resource online in the long-term and to ensure quality control beyond fixed term funded projects. Several options to achieve this were outlined:

- Financing by advertising (Google etc) (a fast but in terms of magnitude unreliable stream of income)
- Offering/selling special services & tools
- Contracts with bigger institutions for special services

Hans Pfeiffenberger (AWI) gave an introduction to the concept of world data centres and the Berlin declaration on open access to scholarly publications

He stressed that: “Business models of publishers or providers of other services must be formulated without resorting to legal or factual (technical) ownership of basic scientific information - published texts and data”.

It was also stressed however, that private companies could still be involved in the above projects and two areas for business opportunities were named:

- Outsourced services** to achieve ingest and quality control of articles or datasets, that is: the editorial processing
- Discovery and aggregation services** (“Portals”, search engines,...) – which can compete on a level playing field

Additional options for long-term financing were suggested e.g. charging for bandwidth, i.e. the data would be free to access, but if very fast access was wished that would be at a cost to the user. The importance for maintaining mirror sites to maintain a resource in the long-term should one of the sites become unavailable, was also stressed

Final agreement

1. A press release will be prepared by AWI to publicise the most important PLANKTON*NET developments and workshop outcomes to the scientific community once PLANKTON*NET@AWI has completed the transition to MV6.
2. The PLANKTON*NET initiative will be maintained beyond the 2 year of funding currently provided by the European commission. The AWI resource PANGAEA is part of the world data centre and therefore committed to the long-term archival of data that are freely available. PLANKTON*NET will be integrated into this world data centre, initially as a ‘current project’. This is designed to attract new partners who will establish new PLANKTON*NET nodes and collections but also to open up new funding routes e.g. via UNESCO.
3. The formation of a consortium, initially including participants of the workshop, will be sought that collaborates on the development of new tools, will consider joint funding applications and will develop a long-term strategy for funding (both institutes, additional grants...)

Annex 1: Workshop Programme

Day 1	Introduction to resources/ User needs	Speakers
From 10:20	Workshop introduction+ summary of the PLANKTON*NET kick-off meeting Talks: Introduction to existing online taxonomic/ initiatives <ul style="list-style-type: none"> • German GBIF node: ALGATERRA • MARBEF Initiatives 	Alex Kraberg Wolf-Henning Kusber Edward vanden Berghe
11:30	Coffee	
12:00	Further talks: <ul style="list-style-type: none"> • New developments at MICRO*SCOPE • Larvalbase/FishBase • ETI 	Paddy Patterson Bernd Überschär Peter Schalk
13:00-14:00	Lunch	
From 14:00	Options for networking online databases/ practical demonstrations: PLANKTON*NET (OAI) Updates on plans for MICRO*SCOPE	Ana Macario (AWI) Adorian Ardelean (MBL)
Ca. 15:15	Coffee break	
From 15:30	Discussion: Future requirements by individual initiatives and researcher e.g what is already being developed and how much overlap is there, what taxonomic information should be included in these databases	
From 19:00	Workshop Dinner	
Day 2	Financing options	
09:00	Short talks: The REBECCA project AQUAECOLOGY (user needs) Enduser needs (IPIMAR)	NIVA Claus Dürselen Teresa Moita

	Summary of results from questionnaires re future needs for online taxonomic resources, preferred funding options?* (this might be done partly on day 1 and partly on day 2)	
09:20	Introduction to the funding situation at PLANKTON*NET/MICRO*SCOPE	Alex Kraberg
10:45	Talks: Factors to consider in designing long-term funding + introduction to the main features and management/ financing of the AWI database Pangaea	Hans Pfeiffenberger (AWI)
	Aquaecology: funding suggestions for PLANKTON*NET to start the discussion	Thomas Raabe
13:00	Lunch	
	Discussion: What would be viable scenarios for funding different resources, is there one solution for the different types of initiatives, mechanisms for co-operation between initiatives:	
App. 14:00 End	Final step: try to get agreement to present to respective institutes on some kind of memorandum of understanding for a structured long-term co-operation between the different initiatives	

List of URLs

URL	Comments
http://www.awi.de/PlanktonNet	PLANKTON*NET@AWI node
http://www.sb-roscoff.fr/baypaul/microscope/general/page_01.htm	Roscoff PLANKTON*NET node
http://www.pangaea.de	Homepage of Pangaea (Publishing network for geoscientific and environmental data)
http://www.openarchives.org	Homepage for the open archives initiatives with info on data standards and protocols
http://www.ubio.org	
http://www.marbef.org	Marbef (EU network of excellence homepage)
http://www.marbef.org/data/erms.php	The European Register of Marine Species
http://www.marbef.org/data/geoabout.php	The European Marine Gazetteer
http://microscope.mbl.edu	MICRO*SCOPE homepage
http://www.fishbase.org	Fishbase homepage
http://www.algaterra.net	Algaterra homepage (Information on marine, freshwater, terrestrial algae)
http://www.bgbm.org/biodivinf/docs/bgbm-model/	Berlin Taxonomic information model
http://www.eti.uva.nl	ETI Bioinformatics homepage: Linnaeus II software can be downloaded from here
http://www.nlbif.nl/Services/search_species	Go here to see a demonstration of the Linnaeus II web Publisher
http://www.tdwg.org/standards.html	List of data standards (approved and under development) on the webpage of the Taxonomic databases working group
http://www.sp2000.org/	
http://annual.sp2000.org/2006/search.php	The catalogue of Life annual checklist CD-ROM is available here
http://ec.europa.eu/research/fp6/ssp/rebecca_en.htm	Information on the Rebecca project (Relationships between ecological and chemical status of surface waters)
http://www.flickr.com	An online facility for managing collections of photos with facilities for formatting images, commenting on them and annotating particular features in images

Annex 3: Pre-workshop questionnaire results

Prior to the meeting a questionnaire was sent to the participants to results of a short questionnaire were presented that was posted to workshop participants prior to the event. Below are some of the main results of this survey, which also included questions about preferred management options an possible roles of private partners. Importantly most respondents seemed to prefer management options with full editorial boards for evaluating information in taxonomic online resources rather than more relaxed arrangements. In terms of financing preferences very different views were held. When asked whether they would prefer a resource financed by individual grants, a consortium of organizations pledging regular low level funding or through a user fee system, no clear majority for either option emerged. Several respondents named (mixed models, e.g. consortia with additional individual grants, individual grants with userfees, combinations of all 3 funding options). This survey represents a very small sample and was merely carried out to focus discussions during the workshop if necessary

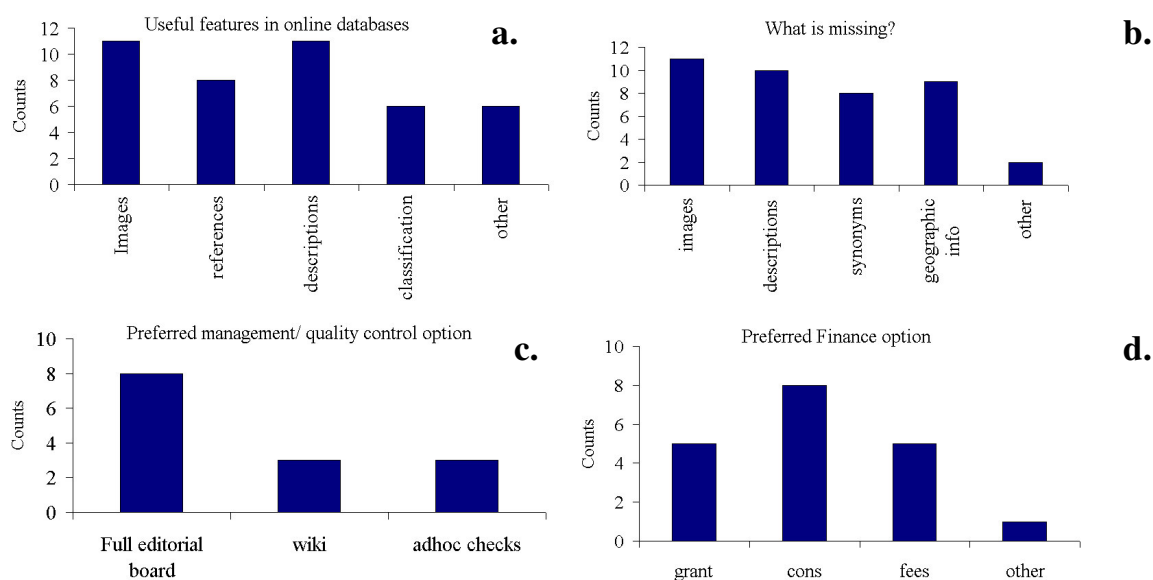


Fig 1 Outcome of a questionnaire distributed to workshop participants prior to the meeting. A. Features in existing online resources considered useful by users, B. Features that are considered lacking in existing resources, C. Preferred management/ quality control options for online databases, D. Preferred finance option for online biodiversity resources (individual grants vs. consortia (=cons) pledging low level but long-term funds vs. access fees to be paid by the user

Contact details of kick-off meeting and workshop participants

Participants	Postal address	E-mail
Adorian Ardelean	Marine Biological Laboratory 7 MBL ST Woods Hole, MA 02543 USA	ardelean@mbi.edu
Ana Amorim	Ana Amorim Instituto de Oceanografia Faculdade de Ciências da Universidade de Lisboa Campo Grande 1749-016 Lisboa PORTUGAL Portugal	aaferreira@fc.ul.pt
Claus-Dieter Dürselen	Aquaecology Marie-Curie Straße 1 26129 Oldenburg Germany Tel. 0049 441 36116 250 Fax. 0049 44136116 255	duerselen@aquaeology.de
Bernd Überschär	IFM Geomar Leibniz-Institut für Meereswissenschaften 20, Düsternbrooker Weg, Kiel 24105	bueberschaer@ifm-geomar.de
Markus Geisen	Alfred Wegener Institute for Polar and Marine Research	mgeisen@awi-bremerhaven.de
Fernando Gomez	Station Biologique de Roscoff, CNRS-INSU-UPMC, UMR 7144 BP 74 Place G. Tessier 29682 Roscoff Cx	fernando.gomez@fitoplancton.com
Alexandra Kraberg	Biologische Anstalt Helgoland, Kurpromenade 201 27498 Helgoland Tel. 0049 4725 819246 Fax. 0049 4725 819283	akraberg@awi-bremerhaven.de
Wolf Henning Kusber	Botanischer Garten und Botanisches Museum Berlin- Dahlem, Freie Universität Berlin, Königin-Luise-Straße 6- 8, D-14191 Berlin, Germany.	w.h.kusber@bgbm.org
Ana Macario	Alfred Wegener Institute for Polar and Marine Research Am Handelshafen 12 D-27570 Bremerhaven (Building E-3045)	amacario@awi-bremerhaven.de

Teresa Moita	IPIMAR Portugal	tmoita@ipimar.pt
Robert Ptacnik	NIVA, Oslo Norway	robert.ptacnik@niva.no
Bastian Onken	Alfred Wegener Institute for Polar and Marine Research	bonken@awi-bremerhaven.de
David Patterson	Josephine Bay Paul Center for Comparative Molecular Biology and Evolution Marine Biological Laboratory 7 MBL Street, MA 02540 Woods Hole, USA	dpatterson@mbl.edu
Hans Pfeiffenberger	Alfred Wegener Institute for Polar and Marine Research	hpfeiffenberger@awi-bremerhaven.de
Ian Probert	Université Basse Normandie, France	algobank@unicaen.fr
Thomas Raabe	AQUAECOLOGY Oldenburg	raabe@aquaecology
Peter Schalk	Director ETI BioInformatics University of Amsterdam Mauritskade 61 1092 AD Amsterdam Phone: +31 20 5257239 Fax: +31 20 5257238 The Netherlands	peter@eti.uva.nl
Edward vanden Berghe	Manager, Flanders Marine Data and Information Centre Flanders Marine Institute Vismijn, Pakhuizen 45-52, B- 8400 Ostend, Belgium Tel +32 59 342130 Fax +32 59 342131	wardvdb@vliz.be
Karen Wiltshire	Biologische Anstalt Helgoland, Kurpromenade 201 27498 Helgoland Tel. 0049 4725 819238 Fax. 0049 4725 819283	kwiltshire@awi-bremerhaven.de
Daniel Vaultot	Station Biologique de Roscoff, CNRS-INSU-UPMC, UMR 7144 BP 74 Place G. Tessier 29682 Roscoff Cx France	vaulot@sb-roscoff.fr

Jeremy Young	Natural History Museum Head of Micropalaeontology Tel: +44 (0)20 7942 5286 Fax: +44 (0)20 7942 5546 LONDON, SW7 5BD, UK UK	jy@nhm.ac.uk
Tamar Zohary	Dr Tamar Zohary Acting Director Kinneret Limnological Laboratory Israel Oceanographic & Limnological Research Ltd PO Box 447 Migdal 14950 Phone: +972-4-6721444 ext 201 Fax: +972-4-6724627 Israel	tamarz@ocean.org.il