



PLANKTON*NET Kick-off meeting, Bremerhaven, April 18th – April 19th Final Report

Introduction:

The aim of PLANKTON*NET project is the creation of a communal resource that is flexible (i.e. easily extensible) and will allow the maximum number of potential users to publish information online and retrieve taxonomic data. The architecture will be comprised initially of 3 OAI-PMH compliant data providers and one service provider dedicated to content harvesting, aggregation and dissemination. PLANKTON*NET will enable stakeholders of many different backgrounds to benefit from this resource for research and monitoring studies. As many of the issues to be discussed should not be debated in isolation but in the wider context of existing initiatives, the meeting was split into two events: the kick-off meeting proper and a 2 day workshop which aimed at the long-term integration/ co-ordination with existing initiatives and options for long-term financing of such resources.

Two phases for the development of PLANKTON*NET were identified: **1.** Installation of core services currently available, ensuring system stability of individual nodes and the initial build-up of database content, **2.** Development of additional tools. Phase 2 will go beyond the currently funded project and will include further joint funding applications to extend PLANKTON*NET funding

The aims of the kick-off meeting were threefold:

- **To introduce all partners.** All partners gave short introductions to their institutes, their research activities and the taxonomic data to be added to PLANKTON*NET (see summary of data in Table 2). Data are available from 4 principal group so far: Diatoms, Dinoflagellates, coccolithophores and Picoplankton taxa. These are to some extent already digitized. In addition images from culture collections e.g. from Picoplankton and heterotrophic dinoflagellates will be added.
- **To plan the practical components of the work** in the context of the deliverables stated in the PLANKTON*NET work description and to decide on ‘background parameters’ e.g. the metadata standards for interoperability, additional database functionality etc. In addition the new features of MICRO*SCOPE, as it forms the

technical basis for PLANKTON*NET and plan the co-operation between the PLANKTON*NET group and the team at Woods Hole

- **To prepare for the workshop following the kick-off meeting** where similar issues were to be discussed in a larger setting including a number of external participants from existing initiatives and

Table 1: Participants of the kick-off meeting

Name	Institute	E-mail
Alexandra Kraberg	AWI	akraberg@awi-bremerhaven.de
Karen Wiltshire	AWI	kwiltshire@awi-bremerhaven.de
Ana Macario (IT)	AWI	amacario@awi-bremerhaven.de
Bastian Onken (IT)	AWI	bonken@awi-bremerhaven.de
Hans Pfeiffenberger (IT)	AWI	hpfeiffenberger@awi-bremerhaven.de
Richard Crawford	AWI	rcrawford@awi-bremerhaven.de
Daniel Vaultot	Roscoff	vaultot@sb-roscoff.fr
Fernando Gomez	Roscoff	fernando.gomez@fitoplancton.com
Ian Probert	Caen	algobank@unicaen.fr
Ana Amorim	FFCUL	aaferreira@fc.ul.pt
Bruno Jesus (IT)	FFCUL	bmjesus@fc.ul.pt
Teresa Moita	IPIMAR	tmoita@ipimar.pt
Jeremy Young	NHM	j.young@nhm.ac.uk
David Patterson	MBL	dpatterson@mbl.edu
Adorian Ardelean	MBL	aardelean@mbl.edu
Thomas Raabe	AQUAECOLOGY	raabe@aquaeecology.de

Table 2: Taxonomic Data available for input into PLANKTON*NET (AM3, deliverable 10-12)

Partner	Source	Taxon Groups	Number of records	Comments	Completed by month 8
AWI	MPI-Ploen, Germany	Dinos, diatoms, cyanos, green algae	Up to 10.000	Images (mostly not digitized), short taxonomic descriptions, approximately 300 images have now been put into PLANKTON*NET	1000 (estimated)
AWI	University of Liverpool	Harmful algae, Coastal Planktonic Ciliates	App 500?	Images, short taxonomic descriptions, glossaries	yes
AWI		Dinoflagellates veg+cyst	Several hundred	Digital images	
AWI	University of Madeira	(tropical) Diatoms, dinoflagellates	tba	Digital Images	yes
Roscoff	English Channel	Dinos, diatoms	350	Images (digital), databased – SYBASE SQL can be exported to any SQL database.	
	Roscoff Culture Collection	Picoplankton mostly	1300	Images (digital), databased – SYBASE SQL can be exported to any SQL database.	
	Literature	Picoplankton species, diatoms	400	Taxonomic information from the literature, collected as an Access database	
Caen	University of Caen	Haptophytes	App 1000	SEM images (digitized), LM images (ca. 1/3 digitized), TEM images (ca. 1/5 digitized). Excel database exists for culture strains, but images not yet linked.	
Lisbon		HAB species mostly: Dinoflagellates, Diatoms, Haptophytes, Raphidophytes	100	LM images and a few SEM (digitised and paper), not in data base	
NHM		Coccolithophores	1000	Images	
			250	Taxon descriptions	
IPIMAR	IPIMAR, Lisbon	harmful algae, diatoms, dinoflagellates, coccolithophores	100-200	Digital images (LM and SEM) and images not yet digitized	

Summary of agreed items/ work schedule

1. A compilation of PLANKTON*NET relevant taxons (species lists) will be created in order to form a taxonomic backbone.
2. 3 PLANKTON*NET data providers will be created based initially on existing software MICRO*SCOPE version 6. Every data provider will be free to configure its index page as wished as long as no changes to the core software are done. This is an important aspect in order to prevent problems in future upgrades.
 - a. [PLANKTON*NET@AWI](#) currently based on MV6
 - b. [PLANKTON*NET@Roscoff](#) will be remotely upgraded to MV6.
 - c. [PLANKTON*NET@Lisbon](#) data provider software to be installed.
3. Re-factoring of MICRO*SCOPE version 6 is needed at several levels:
 - a. extend the database structure so as to incorporate all relevant fields in the Darwin Core metadata format and to allow for multi-lingual approach (feasibility studies are necessary beforehand; consultancy with FishBase is planned); information related to the environmental conditions during sample collection will be initially added in a generic descriptive field; detailed environmental information might be linked to external provider of environmental data when available (e.g., WDC-Mare/PANGAEA)
 - b. changes to the ingestion and display web-interfaces accordingly; interaction with all partners will be needed to come up with user-friendly interfaces at the discovery and searching level
 - c. PID uniqueness, consistent collection definition, relationships between assets and collections, etc have to be formally controlled/assured in order to guarantee consistency in the OAI framework and the feasibility of federations with other information systems.
4. Extended/upgraded version of PLANKTON*NET data provider software will be distributed to all partners
5. An OAI-PMH plug-in module will be created so as to assure that all PLANKTON-NET data providers are protocol compliant; data providers will be registered in the official site for www.openarchives.org
6. Documentation on the database structure, data provider and OAI-PMH plugin software will be made available to all partners

7. Metadata from all PLANKTON*NET data providers will be harvested and indexed in a central repository to assure long-term preservation; harvesting of assets (images) might be done if requested.
8. A service provider (portal) will be created which aggregates taxonomic information from all PLANKTON*NET data providers together with existing plankton-related content (Biopedia, publication-type of repositories, etc). Standard mapping services (google-earth based) will be provided.
9. Data archaeology (non-digitized collections): This refers to the processing of collections that have as yet not been published online or even digitized, but might contain valuable data. It was agreed that this could in principle be an important process but within the 2 year PLANKTON*NET project there is only limited scope for carrying out intensive data discovery/ processing for external partners within the actual EU project. But we will consider choosing one partner (i.e. one collection) in each of the partner countries for whom we will process image material and publish the material in relevant PLANKTON*NET nodes. These would be in addition to the data already listed in Table 2.

Possible selection criteria: vulnerability, taxonomic and/ or ecological importance

10. We will create xid keys for: coccolithophores, dinoflagellates (possibly with several sub keys), habs and a general key pointing the user towards the larger phytoplankton groups. This will result in the emergence of a vocabulary of descriptive metadata for phytoplankton.
11. Commit to the development of a protocol for quality control
12. Commitment to PLANKTON*NET phase 2 with follow-up funding proposals in Framework 7 and possibly other funding organizations

Further discussions were held on functionality that would be desirable to be addressed within phase 2 of the project:

- Annotation/ comments for images
- Clustering of images, e.g. to display images for the complete life cycle of a species.
- Searches that result in a number of images that can be compared.

Roadmap of activities for year 1:

1. Domain name registration: We will register the domain names Plankton-Net.org/Plankon-Net.eu and Planktonnet.org/ Planktonnet.eu immediately

2. Project website: A very preliminary version has been created and will be published online by the end of May. Pages will be created to describe the partner institutes and resources. This information will be provided by partners as soon as possible but by the end of June at the latest? An online forum where messages regarding PLANKTON*NET, MICRO*SCOPE and general issues related to online taxonomic databases can be posted, has been established as a subsection of the Marine Microbes Forum which is hosted at the Station Biologique de Roscoff

([http://www.sb-](http://www.sb-roscoff.fr/marine_microbes/index.php?option=com_simpleboard&Itemid=26&func=view&catid=20&id=330#330)

[roscoff.fr/marine_microbes/index.php?option=com_simpleboard&Itemid=26&func=view&catid=20&id=330#330](http://www.sb-roscoff.fr/marine_microbes/index.php?option=com_simpleboard&Itemid=26&func=view&catid=20&id=330#330)).

3. Software re-factoring + OAI-PMH plugin: Point (3) above should be finished by early Fall 2006; point (5) by end of fall 2006.

4. Portal prototype: Jan. 2007, (consolidation, scalability testing, reporting, etc to be finalized until end summer 2007).

5. Populating PLANKTON*NET nodes with existing partners' data:

This will be one of the most important activities during year 1. The tasks involved are the digitization of existing image material, beginning the process of producing additional images from culture collections and the assembly/ writing of taxonomic descriptions. An initial set of data also including short taxonomic descriptions is to be completed by month 8 (end of November 2006, Table 2) but the process of image acquisition will be continued throughout the project.

Taxonomic descriptions: These will be entered into the database using the BIOPEdia tool. At present descriptions are entered in one field. It was discussed whether the description field should be split into several fields. Partners agreed that in the short term descriptions should still be entered as one field but that this will be re-evaluated after more progress has been made with modifying the database structure (see below).

Suggestion for co-ordinating the tasks

- Many descriptions have already been written for the Helgoland phytoplankton data sets including dinoflagellates and diatoms. The assembly of further descriptions for the diatom

data could be co-ordinated at AWI (by A. Kraberg). For the freshwater data of the Plussee the data contributor will also write a number of short descriptions (Other potential contributors e.g. Malte Elbrächter can be approached).

- Fernando Gomez could co-ordinate the addition of dinoflagellate descriptions working with Lisbon (A Amorim who will work on harmful species) and AWI (A. Kraberg) where necessary.
- Caen, NHM and AWI will collaborate to add one or more coccolithophore collections to the Roscoff node on the basis of an existing guide coccolithophores (?).
- The University of Lisbon will collaborate with IPIMAR to add data to the emerging Lisbon node. Data sets completed before the establishment of the Lisbon node will be added as a temporary collection to the AWI PLANKTON*NET node but will be added to the Lisbon node as soon as possible. They will write descriptions for harmful algae.

6. Interoperability between PLANKTON*NET and external environmental data providers. Given the history of archiving water column data from BAH in WDC/Mare-PANGAEA (e.g Helgoland tie series dating back to 1962), cross-linking between this system and the taxonomic information archived in PLANKTON*NET is clearly wished. A “proof of concept” prototype will be set up until end of June 2006 where this interoperability is demonstrated.

7. Provision of documentation for the installation process: of Microscope v6 and for using the system (both tech and non-tech). This will also include a description of how to prepare image materials texts etc for input into the database. This documentation will enable new users to install and run their own PLANKTON*NET nodes

8. Publicity: By month 6-7 the new PLANKTON*NET site will be announced via e-mail lists such as ALGAE-L. A poster will also be presented at the Harmful algae conference in Copenhagen. This will take place in September 2006.

Comments on timing/ deliverables:

- All PLANKTON*NET nodes will be based on the MICROSCOPE 6 software which will include coverage at species level rather than genus level as in the oldest MICROSCOPE version. Therefore deliverable 9 is already covered. This new version now also contains a front end with detailed data entry pages

and a system of user privileges so that not only node owners but also collection owners can enter and modify records for which they have privileges.

- It was envisaged that installation of the new node at Lisbon would serve as a final quality control step of the modified/ extended PLANKTON*NET database system, accompanying documentation etc. This can be brought forward so that data already available at Lisbon and IPIMAR can be added to the node at an early stage. This team will still take part in this quality control step. At least one additional PLANKTON*NET node will be installed by the end of the two year project, which can be used for further functionality tests