E*C*RC

ENVIRONMENTAL CHANGE RESEARCH CENTRE

University College London

RESEARCH REPORT

No. 81

European Diatom Database (EDDI).
An Information System for Palaeoenvironmental Reconstruction

R.W. Battarbee[1], S. Juggins[2], F. Gasse[3], N.J. Anderson[4,5],
H. Bennion[1], N.G. Cameron[1], D.B. Ryves[5], C. Pailles[3],
F. Chalié[3] and R. Telford[2]

[1]Environmental Change Research Centre, University College London, 26 Bedford Way, London, WC1H 0AP, UK.

[2]Department of Geography, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK.

[3]Centre Department of Européen de Recherche et d'Enseignement de GéoSciences de l'Environment (CEREGE), Europôle Méditerranéen de l'Arbois - BP 80 - 13545 Aix en Provence cedex 04, France.

[4]Department of Geography, University of Copenhagen, Oester Voldgade 10, DK-1350 Copenhagen K, Denmark.

[5]Geological Survey of Denmark & Greenland, Thoravej 8, DK-2400 Copenhagen NV, Denmark.

http://craticula.ncl.ac.uk:8080/Eddi/jsp/index.jsp

CONTENTS

Executive summary

Annexes

1. EDDI EXECUTIVE SUMMARY

1. Diatoms

Diatoms are algae that occur in almost all aquatic environments, usually in great abundance and diversity, and most taxa have a fully cosmopolitan geographical distribution. In addition, because they are siliceous, diatom valves are well preserved in lake sediments and are exceptionally powerful indicators of environmental change. They are particularly suitable for reconstructing past changes in lake water quality, especially pH, nutrient status and salinity.

2. The EDDI system

EDDI is a web-based information system for diatoms designed to enhance the application of diatom analysis to problems of surface water acidification, eutrophication and climate change. It has been produced by combining and harmonising data from a series of smaller datasets from across Europe (and parts of Africa and Asia), and it includes electronic images of diatoms, new training sets for environmental reconstruction, a diatom slide archive, and applications software for manipulating data (Figure 1). It is the result of a three-year collaboration between over 40 diatom taxonomists, palaeolimnologists, statisticians and database experts from 13 countries. The system is available on the web or can be run on a local computer from a CDROM.
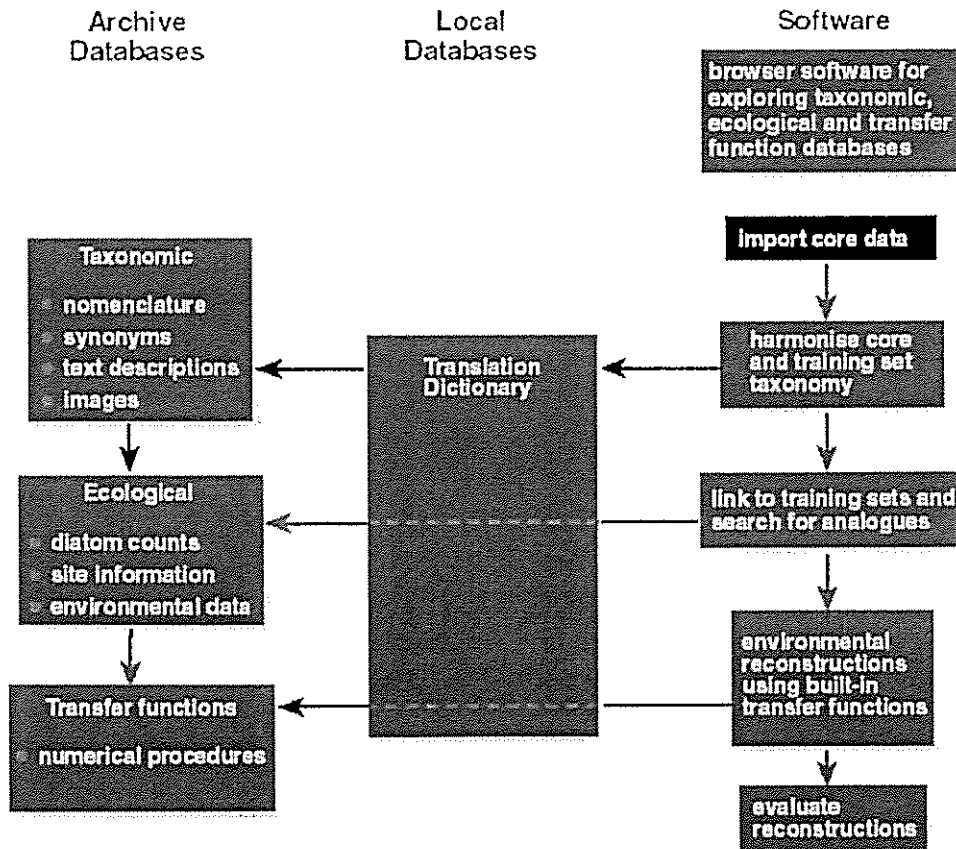
DIATOM INFORMATION SYSTEM

Figure 1: The EDDI Diatom Information System

3. EDDI and the EU

EDDI is a project funded by the European Union through its Framework V programme, and was designed to contribute to three priority research tasks: the climate system in the past, water resources, and the functioning of ecosystems.

4. EDDI personnel

EDDI was coordinated by Rick Battarbee and Helen Bennion of the Environmental Change Research Centre, University College London in partnership with Steve Juggins (University of Newcastle), Françoise Gasse (CEREGE, Aix-en-Provence) and John Anderson (GEUS, Copenhagen).

Diatom harmonisation was carried out by Nigel Cameron (pH), Dave Ryves, John Anderson and Helen Bennion (TP) and Christine Paillès, Françoise Gasse and Françoise Chalié (salinity), with the help of Jorunn Larsen, Jan Weckström, Peter Rosen, Nadia Solovieva, Viv Jones, Roger Flower, Phil Barker, Jane Reed, Laurence Carvalho, Sonja Hausmann, Patrick Rioual and Sybille Wunsam.

Numerical techniques, database development and data re-formatting and checking were the responsibility of Steve Juggins, assisted by Richard Telford, Anne-Marie Clarke, Kathryn Lyttle, and Emma Pearson. The Newcastle group also coordinated the integration of the raw diatom and environmental datasets, diatom images and taxonomic information, and were responsible for the statistical analysis of the new merged EDDI datasets and transfer functions. Dave Ryves also took responsibility for the final merging of the combined pH, salinity and TP datasets.

The three harmonisation centres (UCL, CNRS-CEREGE and GEUS) collected over 2000 digital images to document taxonomic concepts used in EDDI. A large number of these specimens were re-scanned by Shirin Rezai at the Royal Botanic Garden, Edinburgh under the supervision of David Mann and Micha Bayer. Stephen Droop and Micha Bayer also provided invaluable advice on microscopy and image capture protocols, and David Mann provided guidance on taxonomic and nomenclatural issues, while Eileen Cox of the Natural History Museum is responsible for archiving and curating the EDDI diatom slide collection. John Birks, Cajo ter Braak, Joel Guiot, Andy Lotter, Atte Korhola and Hannu Toivonen provided expert advice and guidance on statistical issues and transfer function development, and Don Charles (Academy of Natural Sciences, Philadelphia) provided comments on the web system and discussed issues of compatibility with the US Diatom Paleolimnological Data Cooperative. Gerard Bégni and colleagues at Medias France provided additional web-based support for the salinity datasets.

The taxonomic, distributional, ecological and palaeoecological information contained in the EDDI system is ultimately derived from individual diatom training sets that have been collected by diatomists working in laboratories across Europe. EDDI gratefully acknowledges the following people for generously donating their datasets to the project as follows:

3

John Anderson	Northern Irish, Danish and Northwest European TP datasets and SWAP pH dataset
Leila Ben Khelifa	North African Salinity dataset
Helen Bennion	Welsh CCW, Shropshire / Cheshire Meres, Southern England, and Northwest European TP datasets
Frode Berge	Norwegian and SWAP pH datasets
John Birks	Norwegian and SWAP pH datasets
John Boyle	Norwegian pH dataset
Nigel Cameron	UCL and combined ALPE mountain lake pH datasets
Jordi Catalan	Spanish mountain lake pH dataset
Roger Flower	SWAP pH dataset
Joan Garcia	Spanish mountain lake pH dataset
Francoise Gasse	North, East and combined African salinity datasets
Liz Haworth	SWAP pH dataset
Vivienne Jones	Svalbard pH dataset
Steve Juggins	Caspian salinity dataset and Northwest European TP dataset
Atte Korhola	Finnish pH dataset
Tom Korsman	Swedish pH dataset
Andy Lotter	Swiss TP dataset
Aldo Marchetto	Italian mountain lake pH dataset
Jane Reed	Spanish and Caspian salinity datasets
Sergi Pla	Spanish mountain lake dataset
Ingemar Renberg	SWAP pH dataset
Patrick Rioual	French Massif Central TP dataset
Peter Rosén	Swedish pH dataset
Roland Schmidt	Central European TP dataset
Nadia Solovieva	Kola pH dataset
Jan Weckström	Finnish pH dataset
Sybille Wunsam	Central European TP dataset

5. Harmonisation

The datasets used in EDDI to create a single harmonised database each consist of modern (largely surface sediment, but with a small number of benthic or planktonic) diatom samples and associated environmental data. Each is relatively small (mainly less than 100 lakes), regionally based and concerned with only a single environmental gradient (pH, TP, or salinity). They were created over the last 20 years or so by different laboratories using different methodologies, and many are not publicly available for use by scientists outside the specialist laboratories involved in their creation. These datasets have now been harmonised with respect to both diatom taxonomy and environmental data.

In EDDI taxonomic harmonisation was a two step process, first within the pH, TP and salinity datasets, and second, between them. The harmonisation process involved the standardisation of taxon nomenclature and codes, the screening of slides from the datasets to assess consistency between analysts, full documentation of decisions supported by hard copy micrographs and stored electronic images of all taxa and the archiving of slides for future inspection. No re-counting of slides was carried out and, therefore, harmonisation was largely the result of synonymy identification and merging of entities to the lowest common taxonomic denominator used during the original counting. The original data are preserved in the system.

In more detail, the taxonomic harmonisation procedure was as follows. Lists of all taxa used within each of the pH, TP and salinity datasets were created by Steve Juggins. These were then manipulated by the diatom harmonisation coordinators to produce ordered lists of the most numerically important names, and were further stratified into taxonomic groups which were either known to be synonymous with, or that could be confused with, the main taxon of each group. Subsequently, taxonomic groups were selected for discussion during a series of workshops for pH, TP and salinity dataset harmonisation, independently. The taxonomists who had counted the original dataset slides, or who had worked closely with the original taxonomist, participated in these workshops. Electronic images were captured using a video link to the microscope. A set of minutes was produced, documenting the decisions reached, which then formed the basis for the merging process. A further three workshops were held to harmonise between the pH, TP and salinity datasets. Again, taxa were ranked according to numerical importance as for the first merging exercise and a cut-off of 4% in any single sample was selected as the criterion for inclusion of a taxon. Of the 2163 original names used among all datasets within EDDI (pH, TP and Salinity), 1303 did not feature above 4% in any sample, and a further 54 were considered "unknown".

Merging both within and between the pH, TP and salinity datasets involved a six level taxonomic coding system, as shown in Table 1.

Table 1. EDDI taxonomic coding system used in the harmonisation procedure. Merge levels increase from 0 to 5.

Code	Type	Description
0	No change	Original name and code are unchanged
1	Code change	Original name is valid, code change only
2	Name / code error	Technical error - miscode/synonym/old name/different name used between workers for same discreet, unique entity
2A	Misidentification	Corrects misidentification
3A	Sub-specific merging	Upgrade status of "aff." Codes, merge varieties etc. into higher level, where the different varieties are not consistently identified across different datasets, and where the different taxa are not considered morphologically different
3B	Sub-specific merging	As 3A but where the different varieties etc. are considered morphologically different
3C	Sub-specific merging	As 3B but where the different varieties etc. have been identified across different datasets and are merged to rationalise data analysis
4	Taxonomic concepts start to differ	Medium level merging. Some differences in taxonomic concept here, across datasets, and there is some overlap in usage between workers. The same name may be used for different taxonomic entities across datasets: there is at least a systematic offset between datasets
5	Lowest	High level merging. Some considerable confusion here, both between workers and often in the literature at large. A large range of variation is covered as different taxonomists have used the same name to cover different entities, with often large differences in concepts, and often without the major splits coinciding. As a result, many forms may be included in this umbrella name & code Treat with caution & refer to images for different workers' concepts!
X	Indeterminate species	9999 code used, valves not identified below generic (or higher) level

Existing codes and names were supplied to merged taxa where these could be reconciled with unique concepts from the literature, generally with taxonomic level 3A and below. New codes were allocated in cases where merging created broader taxonomic units than fitted into single taxonomic concepts in the general literature, generally at level 3B and above. Within each of the pH, TP and salinity datasets, the new codes reflected thelaboratory in charge of the harmonisation: XXUnnn for pH at UCL, XXGnnn for TP at GEUS and XXCnnn for Salinity taxa at CEREGE. Numbers (nnn) began at 999 and

decreased by one for each new taxon. For the between-dataset merges, similar guidelines were followed, such that taxa adopted the least specific code found within the pH, TP and salinity datasets, or if new amalgamations of taxa were involved, new XXAnnn codes were introduced.

The combination of EDDI taxon code and taxonomy class can then be used to determine the degree of merging involved in any particular EDDI taxon. Together with the linked digital images and a description of what has been merged (or reference to a published description), this allows users to decide for themselves how appropriate any particular EDDI taxonomy is for their purpose. Sites can thus be selected or excluded on the basis of the taxonomic implications of resultant merges, and the effect this has on net ecological information gain or loss, whether for model building and down-core reconstruction of parameters, or present-day species distributions across gradients, for example.

Environmental data were harmonised simply by ensuring that the numerical data were expressed in common units of measurement. The number of samples on which the mean data are based is recorded in EDDI for each site so that the user has information on data frequency (quality). For the African sites in the salinity datasets, where there is great intra- and inter-annual climatic variability, all available published data (largely pH and conductivity) were compiled to derive the best estimates of modern environmental conditions. Table 2 shows a summary of the environmental data for the combined dataset.

Table 2. Summary of environmental data for the combined dataset.

Variable	Units	N	Min	Max	Mean
Alkalinity	ueq/l	1207	-51	968000	6315
Aluminium (labile)	ug/l	179	0	330	56
Aluminium (monomeric)	ug/l	150	0.1	470	45
Aluminium (total)	ug/l	312	10	800	114
Ammonium	ug/l	215	0	6120	504
Calcium	ueq/l	1092	6.17	3152200	13900
Chloride	ueq/l	922	0	6219200	79600
Chlorophyll	ug/l	247	0.1	350	24
Colour	mg Pt/l	143	3	240	51
Conductivity	uS/cm	1233	4.4	400000	5300
Iron	ug/l	396	1.8	2920	139
Magnesium	ueq/l	1092	4	2169900	23200
Manganese	ug/l	254	0.5	600	39
Maximum depth of lake	M	1239	0.02	410	15
Nitrate	ug/l	475	0.5	41000	2300
Nitrite	ug/l	114	20	4320	435
PH		1332	4.32	10.9	7.1
Potassium	ueq/l	1081	0.8	776300	6000
Salinity	mg/l	181	133	333000	29700
Secchi disk depth	M	288	0.1	11.8	3.7
Silica	mg/l	591	0.04	320	11.0
Sodium	ueq/l	1052	4.57	3872400	46100
Soluble reactive phosphate	ug/l	235	1	1016	65
Sulphate	ueq/l	895	4.4	1765300	24200
Total organic carbon	mg/l	458	0.12	20.2	4
Total organic nitrogen	ug/l	117	10	982	272
Total organic nitrogen (2)	ug/l	209	1.12	8360	675
Total phosphorus	ug/l	664	0.5	1190	58
Water depth of diatom sample	M	1216	0.02	191	12
Zinc	ug/l	0.36	730	23	232

6. Training datasets

Following harmonisation data were combined to create a single dataset of 1350 modern samples covering most of Europe and parts of Africa (Figure 2). The combined dataset includes 23 regional training datasets, plus three training datasets for pH, TP and salinity respectively (Table 3).

The database also allows the distribution of taxa to be plotted. Here (Figure 3) we show an example for four common species, *Achnanthes minutissima*, *Stephanodiscus parvus*, *Eunotia incisa* and *Cymbella pusilla*.

Table 3. The training datasets included in EDDI. N= number of samples in each dataset.

Type	Id	Name	N
pH	ALPE	ALPE mountain lake dataset	118
pH	ALPI	Italian mountain lake dataset	31
pH	ALPS	Spanish mountain lake dataset	28
pH	ALPU	UCL mountain lake dataset	30
pH	Bergen	Norwegian pH dataset	96
pH	Finland	Finnish pH dataset	98
pH	Kola	Kola penninsula pH dataset	25
pH	pH	Combined pH dataset	627
pH	SWAP	SWAP pH dataset	178
pH	Sweden	Swedish pH dataset	118
Salinity	Africa	African dataset	284
Salinity	AfricaE	East African dataset	187
Salinity	AfricaN	North African dataset	97
Salinity	Caspian	Caspian saline lake dataset	29
Salinity	Salinity	Combined salinity dataset	387
Salinity	Spain	Spanish saline lake dataset	74
TP	CCW	Welsh TP dataset	11
TP	CEuro	Central European TP dataset	86
TP	DK	Danish TP dataset	28
TP	French	French Massif Central TP dataset	28
TP	NI	Northern Irish TP dataset	54
TP	NWEuro	NW Europe TP dataset	164
TP	SCM	UK meres TP dataset	33
TP	SEng	Southern England TP dataset	26
TP	Swiss	Swiss TP dataset	69
TP	TP	Combined TP dataset	347

Distribution of EDDI
samples.

Red - pH training set
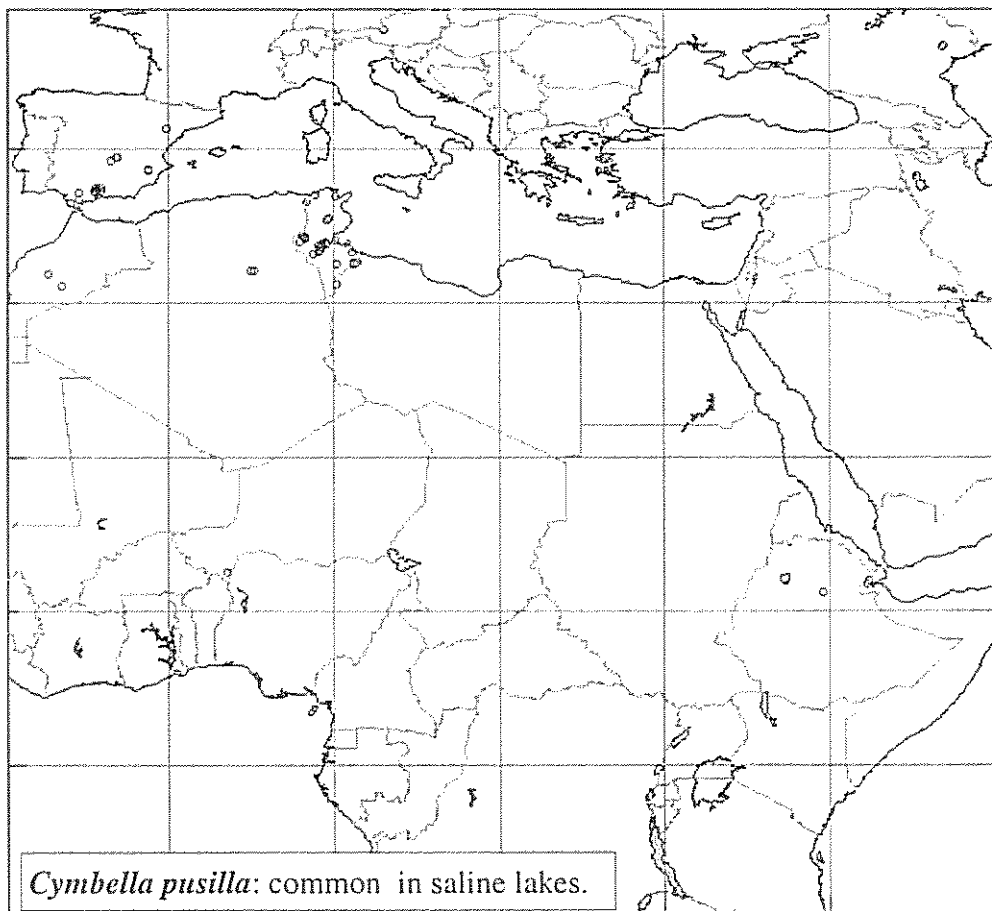
Green - TP training set

Blue - Salinity training set
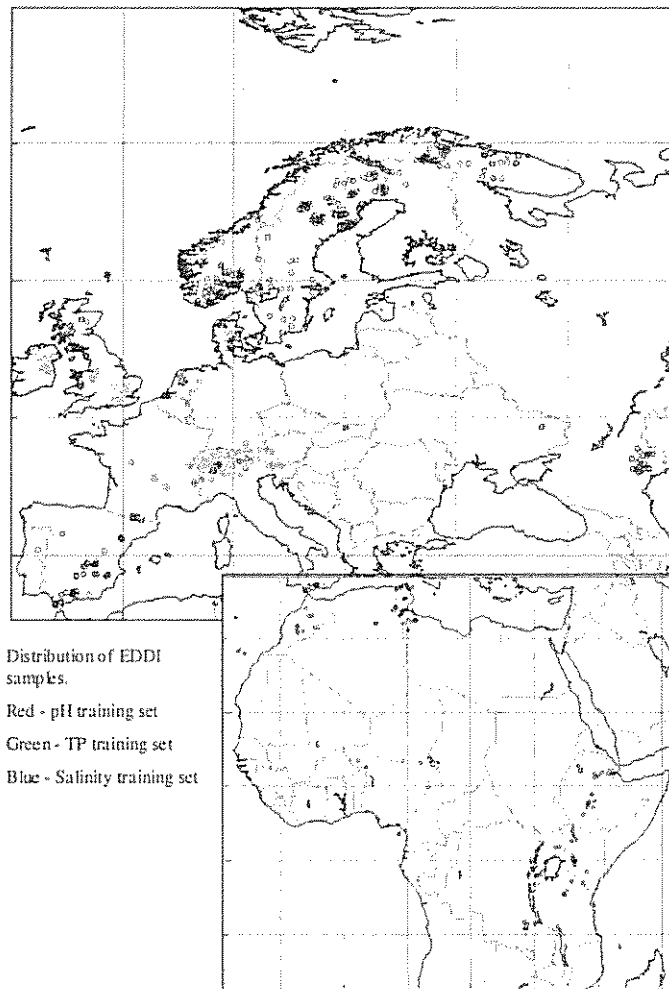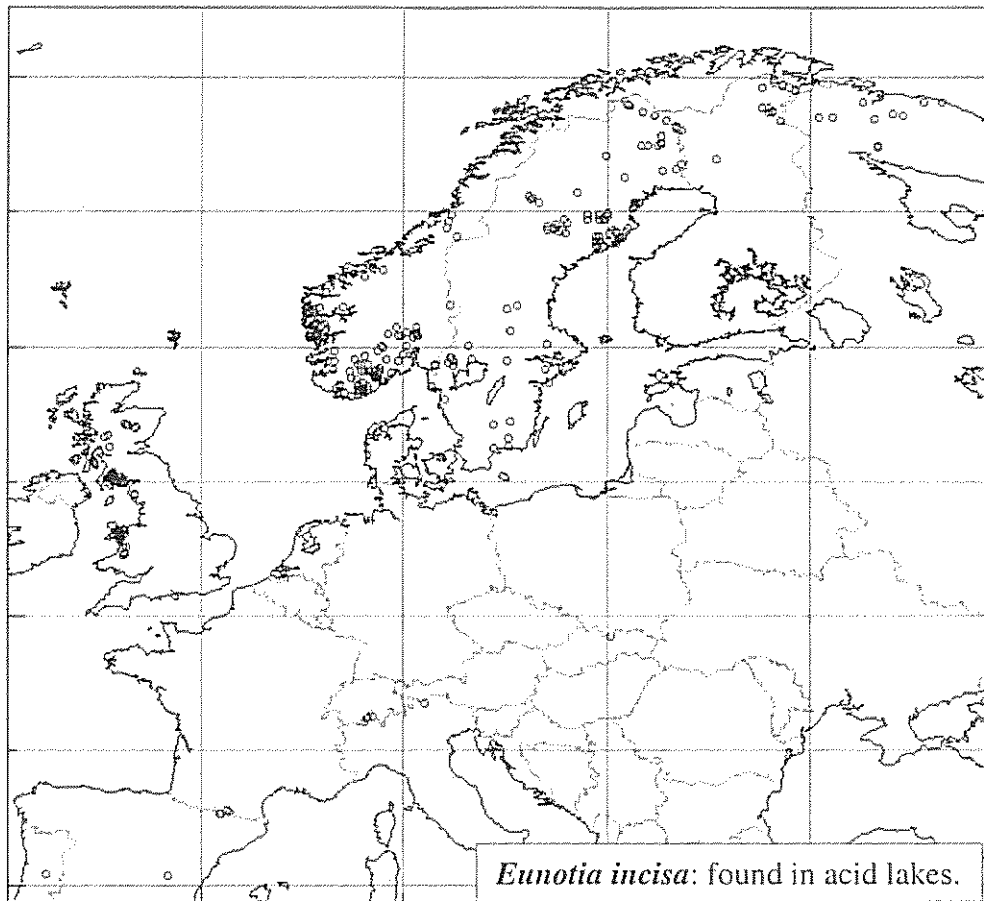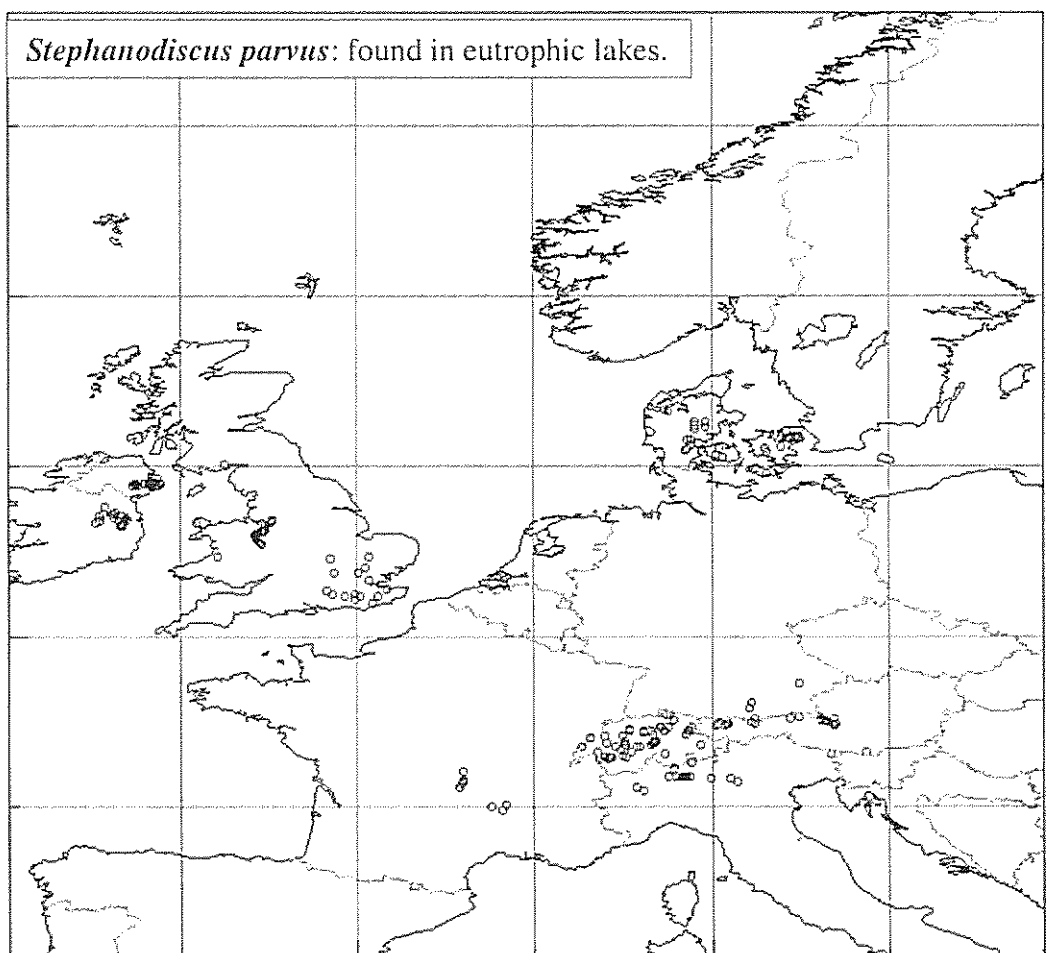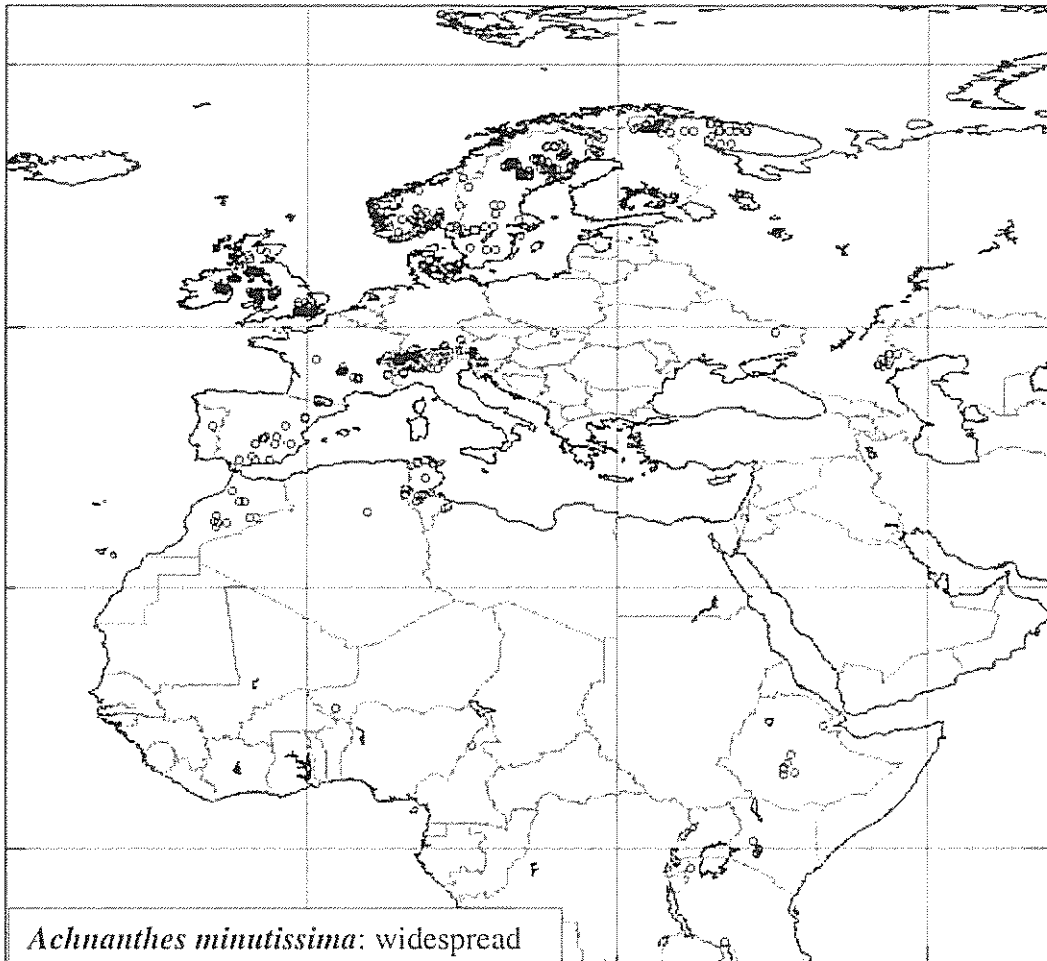
Figure 2. Map of sample distribution in EDDI datasets.

Figure 3. Diatom distribution using data from the combined dataset in EDDI
(see over)

Eunotia incisa: found in acid lakes.

Cymbella pusilla: common in saline lakes.

Achnanthes minutissima: widespread

Stephanodiscus parvus: found in eutrophic lakes.

A full description of the individual datasets in EDDI is available on the web-site. For each dataset a full taxon list, a sample list and a distribution map are available along with a list of relevant publications. Here (Table 4) we present the ALPE dataset as an example to illustrate the type, range and format of information provided.

Table 4. ALPE mountain lake dataset

EDDI dataset code	ALPE
Dataset type	pH
Number of samples	118
Collection date	1986 - 1993
Contributor	The ALPE Participants
Contacts	John Birks, Nigel Cameron

The ALPE diatom-pH calibration data-set consists of surface-sediment diatom assemblages from 118 lakes and contains 530 taxa. The ALPE training set is from high-altitude or high-latitude lakes in the Alps, Norway, Svalbard, Kola Peninsula, UK, Slovenia, Slovakia, Poland, Portugal, and Spain. Gravity or piston corers were used to collect surface-sediment samples of 0.25 or 0.5 cm thickness, usually from the deepest point in each lake and a total of at least 500 valves were counted from each sample. A large number of possible sites in existing calibration sets, plus new ALPE project sites, were first screened carefully to select a set of lakes meeting the criteria of an alpine or remote location and an undisturbed catchment. The resulting ALPE data set of 118 lakes is derived from the whole or parts of 5 data-sets. These are 31 lakes from a Central Alps data set (Marchetto & Schmidt, 1993), 28 lakes from a Pyrenean data set (Garcia & Catalan, unpub.), 30 AL:PE 1 and AL:PE 2 lakes in various countries, 9 lakes from a Norwegian diatom-pH data set (Birks, Boyle & Berge, unpub.), and 20 Norwegian, Welsh, and Scottish lakes from the SWAP calibration set (see above). These 118 lakes have an altitudinal range of 20 m (northern Svalbard 79o40'N, 10o45'E) to 3050 m (mean = 1762 m median = 2091m).

Variable	Units	N	Min	Max	Mean
Alkalinity	μeq/l	104	-31.75	630	67.6
Aluminium (labile)	μg/l	22	3.51	176	48.2
Aluminium (monomeric)	μg/l	24	3.0	147	42.5
Aluminium (total)	μg/l	43	10	256	61.9
Ammonium	μg/l	80	0	93	10.2
Calcium	μeq/l	118	6.17	519	88.8
Chloride	μeq/l	118	0	350	34.3
Conductivity	μS cm^{-1}	118	4.4	74.4	19.4
Iron	μg/l	20	0.01	471	49.7
Magnesium	μeq/l	118	4	222	24
Manganese	μg/l	15	1.5	17	7.2
Maximum depth of lake	m	118	1	73	15.5
Nitrate	μg/l	102	0.7	41	10.4
Nitrite	μg/l	28	0.4	10	5.8
pH		118	4.48	8.04	6.15
Potassium	μeq/l	118	0.89	48	6.09
Silica	mg/l	52	0.19	2.22	0.76
Sodium	μeq/l	118	1	309	39.8
Sulphate	μeq/l	118	13.8	198	50.5
Total organic carbon	μg/l	50	0.2	4.87	1.32
Total nitrogen	μg/l	37	3.22	770	202
Total phosphorus	μg/l	91	0.5	43	7.18
Water depth of diatom sample	m	118	1	73	15.5
Zinc	μg/l	21	0.36	14	6.75

7. The web-based diatom iconograph

During the harmonisation process diatom images were captured by each of the diatom coordinators. UCL used a Leica DMLB microscope and digital images were captured using a Neotech Image Grabber PC software system. GEUS used a Leitz DMR microscope and Kappa ImageBase software for capturing digital images. CEREGE used a Nikon ECLIPSE E-600 microscope and digital images were captured using the Image Vision Builder system (National Instrument, version 1.0). All images were captured as tagged image file format (TIFF).

Many of the images in EDDI are "working images" and are shown to define the range of forms that have been included for each of the main taxa. In each case, however, a single image has been selected to represent the form regarded as being most representative and these images were produced using a higher quality capture system at the Royal Botanic Gardens Edinburgh (RBGE).

Specimens were captured at RBGE on a Zeiss Axiophot light microscope with 63x and 100x oil immersion lenses (both with a numeric aperture of 1.4). All diatoms were photographed using brightfield optics only. A number of protocols were introduced to maximize resolution. The specimens were then digitally captured using a Kodak MegaPlus ES1.0 digital camera attached to the microscope. The camera was interfaced with a personal computer running Optimas imaging software version 6.2 (MediaCybernetics, Silver Spring, MD 20910, USA). The MegaPlus ES1.0 has a resolution of 1008 x 1018 pixels, and allows image resolutions between 7 and 18 pixels/μm. A purpose-written macro was developed that allows the placement of a scale bar as an image overlay. Following this, images were manually contrast adjusted using Adobe Photoshop image editing software.

8. The slide archive

armonisation of taxonomy amongst datasets is necessary to generate new training sets required by the EDDI project. However, this often leads to the loss of taxonomic resolution in the harmonised dataset, and, in some cases, entails subjective decision making on behalf of the participating diatomists. Consequently to enable the system to be as transparent as possible the original code and count data for each sample is preserved within the web system and slides for almost all samples are archived in the London Natural History Museum diatom collection to enable microscopic inspection in future. In many cases new slides have been made from old samples, but unfortunately in some cases the original material could not be located.

9. New training sets

One of the major goals of EDDI was to use the harmonised datasets to develop new training sets for the purposes of environmental reconstruction. The EDDI system contains a total of 20 training sets – 13 of these were pre-existing regional datasets , 2 were pre-existing merged datasets (ALPE, NW European and African) four are new merged datasets (combined conductivity, combined TP, combined software pH, combined TOC, Combined saline-lake pH (Table 5).

It was originally envisaged that the EDDI system would allow users to select any combination of samples to produce customised training sets, for example, according to the similarity between the core assemblage and the training set assemblage. In the final system this manual selection of customised training sets has been replaced by a new method, locally-weighted weighted averaging, that dynamically selects a training set optimised for each fossil sample. These transfer functions are thus not built into the system as static lists of samples by generated dynamically when the user performs a reconstruction.

Table 5. List of EDDI transfer functions

Variable	Training set	No. samples	No. taxa
Anion ratio (Alk / (Cl + SO4))	Combined salinity dataset	333	356
Conductivity	African dataset	270	479
Conductivity	East African dataset	179	332
Conductivity	North African dataset	91	289
Conductivity	Combined salinity dataset	370	604
Conductivity	Spanish saline lake dataset	71	203
PH	ALPE mountain lake dataset	118	332
PH	Norwegian dataset	96	277
PH	Finnish dataset	67	284
PH	Combined pH dataset	622	652
PH	Combined salinity dataset	366	603
PH	SWAP dataset	174	381
PH	Swedish dataset	118	214
Total organic carbon	Combined pH dataset	388	536
Total phosphorus	Central European dataset	85	207
Total phosphorus	Northern Irish dataset	54	163
Total phosphorus	NW Europe dataset	163	333
Total phosphorus	Southern England dataset	26	201
Total phosphorus	Swiss dataset	69	181
Total phosphorus	Combined TP dataset	477	345

10. Reconstructing environmental variables

In addition to pH, TP, and conductivity (as a surrogate for salinity) EDDI also contains transfer functions for reconstructing total organic carbon TOC and anion ratio (alkalinity / (Cl + SO4)). In software lakes TOC is a useful indicator of changes in catchment vegetation and distance to tree-line and in saline lakes changes in anion ratio indicate movement along evaporation trajectories. Both proxies are thus a useful indirect proxy of climate change. These transfer functions can be applied in two ways in EDDI – either on-line or using downloaded software. To perform reconstructions on-line the user simply uploads their core-data to the EDDI server, selects the transfer function and downloads the results. This is the easiest solution for the casual user with just a few cores to reconstruct. For the more intensive user we have written custom software that can be installed and run on a local. However, the process of applying the transfer functions is the same in both cases and the software for on-line and local reconstructions produces exactly the same output, so a user can easily swap between approaches. The details of the whole EDDI system, including the operation of the reconstruction software an interpretation of output is documented in a user guide that is available on the CDROM or for download from the web.

11. Numerical methods

EDDI includes a range of software tools to enable the user to verify their core taxonomy against that of the training set, to perform reconstructions using numerical methods most appropriate for each training set, namely weighted averaging (WA), weighted averaging partial least squares (WAPLS), modern analog technique (MAT), and the new technique of locally-weighted weighted averaging. Each reconstruction method yields environmental estimates and sample-specific standard errors of prediction, estimated using a robust method of cross-validation and Monte-Carlo simulation. In addition the software also calculates analog measures and provides an indication to the reliability of the reconstructed values for each fossil samples. We compared the performance of each reconstruction method on each dataset. Results show that no single numerical method is superior and that different methods are appropriate for different datasets. These results have been used to provide guidance on the choice of method in the user guide.

ANNEX 1 Achievements against objectives

The main objectives of this project were (i) to develop a web-based diatom information system to enhance the use of diatom analysis as a technique for environmental reconstruction with respect especially to problems of climate change and pollution; (ii) to make it possible for more laboratories to use the technique to a very high standard, and (iii) to develop a system that allows diatom-based environmental reconstruction within Europe to be carried out using a harmonised methodology.

Here we briefly summarise the achievements of the project in relation to the specific objectives set out in the proposal:

(i) bring together a series of small datasets from the whole of Europe (and parts of Africa & Asia) to produce a single high quality, integrated and harmonised training set of diatoms, with site information and environmental data.

This was successfully achieved. The EDDI system contains harmonised diatom counts and associated site and environmental information for over 2000 taxa in 1350 modern samples from 23 regional training datasets.

(ii) combine the training set data with taxonomic information, electronic images of diatoms, and software for data analysis on a CD-ROM;

This was successfully achieved. All details of the training sets, diatom samples and taxa can be listed or explored graphically via web pages and software linked to the EDDI database. Taxonomic conventions used in merging the EDDI datasets are fully documented with over 2000 digital images, and environmental reconstructions using EDDI transfer functions are available on-line. Identical material to that on the web site is also available on CD-ROM.

(iii) derive new, more accurate and more widely applicable transfer functions for key hydrochemical variables;

A number of new transfer functions are currently under development for publication in the international literature over the coming months. The new merged pH, TP and salinity datasets have far greater environmental and taxonomic diversity than their regional constituents. They are thus far more widely applicable and, using the new method of LWWA, are more accurate than many of the previously published transfer functions. A case study for TP comparing the use of the combined EDDI TP dataset with original, smaller regional datasets was presented at the 8th Palaeolimnology Symposium held in Kingston, Ontario in August 2000.

(iv) explore the relationship of diatoms with other environmental (especially climate) gradients on a European scale;

A study, based on the new dataset, to explore patterns of diatom distribution potentially linked to climate and geography is currently underway. The extent to which this objective can be fully achieved for the whole of Europe is limited by the number of samples that have the full range of chemical and other environmental variables associated with them. For example, only the pH datasets include values for dissolved organic carbon (TOC) a variable of interest in terms of climate change and its impact, and few samples have accurate water or air temperature data.

(v) evaluate a range of techniques for environmental reconstruction and evaluation, and develop a set of guidelines for "best practice" in developing transfer functions;

Guidelines for using transfer functions are provided in the user guide based on a comparison of the performance of each numerical method / dataset under cross-validation. Results to date indicate the new dynamic training set method of LWWA is robust and performs at least as well as traditional methods for smaller datasets and better for the large merged datasets. Publications detailing these comparisons and describing the new method of LWWA are in preparation.

(vi) map the geographical distribution of surface sediment diatoms and explore the extent to which distributions are most strongly influenced by water chemical, climatic, biogeographic or other factors;

Geographical distributions of the taxa can be mapped via the web pages. Preliminary exploration of the variables explaining the diatom distributions has been undertaken but this will be developed more fully in a forthcoming paper.

(vii) distribute all data and numerical procedures as a complete information system using the World Wide Web (WWW) and on CD-ROM.

This has been successfully achieved. The EDDI system is currently being tested on a server at Newcastle University at http://Craticula.ncl.ac.uk:8080/Eddi/jsp/index.jsp. When testing is complete the system will be transferred to the World Data Center A for paleoclimatology at Boulder, Co, USA where the system will be hosted independently. A CD-ROM version is also available on request.

ANNEX 2 Partner reports

Partner 1 ECRC-UCL

Prof. R.W. Battarbee, Dr. H. Bennion and Dr. N.G. Cameron.
ECRC, University College London, 26 Bedford Way, London, WC1H 0AP, England.

Prof. Rick Battarbee and Dr. Helen Bennion were responsible for central co-ordination and management of the project. Dr. Nigel Cameron co-ordinated work package 1a and had shared responsibility for leading work package 2 with partners 2 and 3. Other ECRC staff involved in EDDI were Drs. Roger Flower, Viv Jones and Nadia Solovieva who contributed data and taxonomic advice to work package 1a, and Dr. Patrick Rioual who contributed data and taxonomic expertise to work package 1b. These staff attended the work package 1 taxonomic workshops. Partner 1 contributed a number of datasets to EDDI including, for pH, SWAP, ALPE, KOLA and SVALBARD, and for TP, NWEURO and FRANCE. Dr. Paula Maliphant, and latterly David Seamark, in their role as ECRC database developer, ensured compatibility between the ECRC diatom database (AMPHORA) and the EDDI system.

ECRC-UCL was also responsible for the contributions of NHM-LOND (Dr. E. Cox), UHEL-LPG (Jan Weckstrom), UBERG-BOT (Jorunn Larssen), and UMEA-DEH (Dr. Tom Korsman, Peter Rosen).

Workshops and meetings
During the course of the project Partner 1 organised, and in most cases hosted, a number of meetings and workshops as follows:

21-22 May 1998: Initial workshop for all EDDI participants at UCL to formulate protocols for taxonomic harmonisation, slide archiving, image capture and transfer, and data formatting and transfer.

14-15 Jan 1999: pH taxonomic workshop (WP1) led by Dr. Nigel Cameron at UCL.

20-21 Jan 1999: Salinity taxonomic workshop (WP1) led by Dr. Christine Paillès at CEREGE.

25-26 Jan 1999: TP taxonomic workshop (WP1) led by Dr. David Ryves at GEUS.

The workshops brought together all the contributing diatomists to that particular group who worked through a "hit-list" of the most common taxa, documenting any synonyms, mis-identifications etc. Images were captured to illustrate species concepts and all decisions were documented in workshop minutes (attached). Any taxa requiring further work and clarification were noted. Helen Bennion attended all workshops to ensure that they were conducted effectively and in a standard manner.

1-3 March 2000: Taxonomic harmonisation and image capture workshop (WP2) at UCL. Issues relating to taxonomic harmonisation methodology for WP2 and the development of the high quality image database and image capture techniques were discussed.

12-14 June 2000: Taxonomic harmonisation workshop (WP2) at GEUS to finalise the harmonised training sets for pH, TP and salinity, respectively, and to further progress on the integration of all three datasets.

23-24 July 2001: Final EDDI workshop for all EDDI participants at UCL to discuss feedback on EDDI web pages, draft report, issues of data access and publication plans.

Partner 1 organised and chaired a series of steering group meetings attended by the key partner personnel (Anderson, Battarbee, Bennion, Cameron. Gasse, Juggins, Paillès, Ryves):

21 May 1998 to discuss the first phase of the project and ways in which EDDI could be publicised via posters, fliers and web sites.

12 February 1999 to review progress, refine protocols for data exchange and image transfer, discuss publications and conference presentations of EDDI, outline the timetable and deliverables for the next six month phase of the project and to discuss a proposal for a pilot study.

23 November 1999, Mid-term Review meeting. The project milestones for Work Package 1 were reviewed and a strategy for taxonomic harmonisation was agreed. Other outputs were the identification of gaps in the environmental datasets, a summary of slide availability, feedback on image capture protocols, progress and image quality.

19 March 2001 Final Review meeting to finalise WP2 and WP3, plan structure/content of final report and agree timetable for deliverables in final 6 months of the project.

23 July 2001 Final Meeting to discuss feedback on EDDI web pages, draft report, issues of data access and publication plans.

General co-ordination

In addition to workshops and meetings, Partner 1 co-ordinated the project via circulation of relevant minutes, newsletters and regular communication by email.

Partner 1 compiled a spreadsheet detailing the availability of slides, suspensions and sediment material from each sample to be included in EDDI and co-ordinated slide preparation at UCL.

Partner 1 took responsibility for structuring, editing and delivering the final reports.

Work packages 1 and 2

Dr. Nigel Cameron was responsible for co-ordinating the harmonisation of the pH datasets in work package 1. Following a pilot study of a single diatom group to agree a methodology for harmonisation and documentation of decisions, Dr. Cameron collated and merged existing pH training sets of taxonomy, chemistry, and site data. Images were captured of each relevant taxon, where necessary including an example from each of the different regional training sets. Documentation of taxonomic decisions was undertaken with emphasis on the problem taxa, involving a short text description to explain harmonisation decisions. Subsequently Dr. Cameron merged datasets in work package 2 in collaboration with Partners 2 and 3, under the leadership of Dr. Bennion.

Partner 2 CEREGE

Dr. F. Gasse, Dr. C. Paillès and Dr. F. Chalié
Centre Europeen de Recherche et d'Enseignement de GeoSciences de l'Environnement, Europole Mediterraneen de l'Arbois, B.P. 80, Université de Aix-Marseille III, 13545 Aix en Provence, Cedex 4, France.

Dr. Francoise Gasse was the senior partner at CEREGE responsible for overseeing work package 1c and contributing diatom training sets for salinity. Dr. Christine Paillès was appointed to co-ordinate this work package and to carry out the harmonisation of the salinity datasets within work packages 1c and 2. Francoise Chalié, a diatomist at CEREGE, assisted during the later stages of the project with final merging and documentation. Partner 2 contributed the AFRICA datasets to EDDI. CEREGE was also responsible for P. Barker (University of Lancaster, UK), J. Reed (University of Hull, UK), E. Cubero-Castan (Medias -France).

Partner 2 was responsible for the harmonisation of the salinity dataset (including initial datasets from North and East Africa, Russia and Spain). The contribution in WP1 mainly included the standardization of taxon nomenclature and code, the design and testing of protocols for recording taxonomic concepts of problematic taxa, taxonomic descriptions, image capture, and the data integration and harmonisation of the four datasets into a single harmonised salinity dataset.

A taxonomic workshop was organized by CEREGE on January, 20-21, 1999. At this stage of the project, the meeting provided the basis of the merging procedure and taxonomic harmonisation for the salinity datasets (WP1). Among the original list of 878 taxa, we produced a final merging list of 690 salinity EDDI taxa, Salinity dataset taxonomy and merging were illustrated by description and/or references for 485 EDDI salinity taxa. We also produced 663 digital images from 156 specimens, a large number of which were sent to RBGE for high-quality recapturing of key specimens.

In parallel with the taxonomic harmonisation, the compilation of geographical and chemical data was performed for the sites in the salinity datsets. In WP2, partner 2 shared responsibility for merging TP, pH and salinity datasets with Partners 1 and 3 and organised a WP2 taxonomic workshop at CEREGE in November 2000.

Partner 3 GEUS-DK

Prof N.J. Anderson and Dr. D. Ryves
Geological Survey of Denmark & Greenland, Environmental History & Climate Department, Thoravej 8, Copenhagen, Denmark.

Prof. John Anderson was the senior partner at GEUS responsible for overseeing work package 1b and contributing diatom training sets for TP. Dr. Dave Ryves was appointed to co-ordinate this work package and to carry out the harmonisation of the TP datasets within work packages 1b and 2. Partner 3 contributed a number of the NWEURO datasets to EDDI. GEUS was also responsible for the contributions of Prof. R. Schmidt (Mondsee) and Dr. Andy Lotter (University of Bern) and Dr. Patrick Rioual (ECRC-UCL).

Workshops were held in January 1999 (WP1) and in June 2000 (WP2). All taxonomic and harmonisation decisions made were recorded. Partner 3 entered diatom taxonomic information, diatom counts, environmental, sample and site data for the TP datasets into the EDDI database

For WP1, a list of all taxa used amongst all 8 TP datasets was created by Dr. Steve Juggins, University of Newcastle, containing 975 taxa from 343 sites. This was then manipulated at GEUS to produce an ordered list of the 56 most numerically important names used by the combined TP co-workers. Each taxon formed the nucleus for a group of synonymous or morphologically similar taxa for further consideration. By this method, 223 names, comprising about 86% of the total species data, were considered in WP1. After WP1 merging, the original list of 975 TP "taxa" used by EDDI taxonomists was reduced to around 720 TP EDDI taxa before merging with the pH and salinity datasets.

For WP2, Partner 3 shared responsibility for merging datasets with Partners 1 and 2, and followed a similar system of deciding which taxa should be the focus of attention (according to the overall numerical importance among all three datasets). Partner 3 wrote brief taxonomic descriptions or gave taxonomic references from the literature for 200 TP taxa, and wrote descriptions for 49 taxa merged at the WP2 stage.

In all, 1093 digital images were taken by Partner 3 at GEUS from 380 specimens. Approximately 130 diatom slides were sent to RBGE for high-quality image collection of selected specimens.

Partner 4 UNEW-DGEOG

Dr. S. Juggins
Dept of Geography, University of Newcastle, Claremont Road, Daysh Building, Newcastle Upon Tyne, NE1 7RU, England.

Partner 4, under the leadership of Dr. Steve Juggins, was responsible for leading Work Packages 3 and 4. Other University of Newcastle staff involved in EDDI were Dr. Richard Telford who was responsible for developing some of the web-based Java applets and, along with Annie Clarke, Kathryn Lyttle and Emma Pearson for screening the diatom and environmental data. Steve Juggins also contributed the Caspian salinity dataset to EDDI.

In the early stage of the project Partner 4 provided database and computing support to the training set co-ordinators in Work Packages 1 and 2, organising merged taxon lists and setting up a web-based image viewing system to aid inter-laboratory communication and taxonomic harmonisation.

In WP3 Partner 4 was responsible for creating the final EDDI database from Excel file supplied by partners 1-3, for statistical analysis of the EDDI datasets, and for all software development. This work included development of the new transfer functions for pH, TP, conductivity, TOC and anion ratio and comparisons of different numerical methods, and the development on the new technique of locally-weighted weighted averaging.

In WP4 Partner 4 developed the interface to the EDDI database and reconstruction software and implemented this as a web-based system and on CDROM. A stand-alone

version of the reconstruction software was also developed. Partner 4 also wrote the user guide and developed the guidelines describing 'best-practice' methods.

Partner 4 organised a database workshop in Arles, 17-20th March 1999, to discuss the structure and web-based implementation of the final EDDI database with representatives of MEDIAS-France and the World Data Centre-A, Boulder, Colorado.

Partner 4 was also responsible for the contributions of Prof. John Birks (University of Bergen), Dr Cajo ter Braak (Agricultural Mathematics Group-DLO, Wageningen) and Prof. David Mann (Royal Botanic Gardens, Edinburgh).

Partner 4 also organised a database workshop on 16-19th May 1999, which was attended by Stephen Juggins, John Birks, Cajo ter Braak, Joel Guiot, Atte Korhola, Hannu Toivonen and Franck Torre. The workshop discussed the procedures for evaluating the various transfer function methodologies. Methods for data screening, performance and reliability assessment were agreed and the appropriate software and responsible personnel were identified.

Battarbee, R.W, Juggins, S., Gasse, F., Anderson, N.J., Bennion, H., Cameron, N.G., Ryves, D.B. and Pailles, C. (1999). **European Diatom Database (EDDI): An Information System For Palaeoenvironmental Reconstruction.** Abstracts of the INQUA conference, August 1999.

Battarbee, R.W, Juggins, S., Gasse, F., Anderson, N.J., Bennion, H. and Cameron, N.G. (2000). **European Diatom Database (EDDI): An Information System For Palaeoenvironmental Reconstruction.** European Climate Science Conference, Vienna City Hall, Vienna, Austria, 19-23 October, 1998, pp. 1-10 (reprint attached).

Bennion, H., Battarbee, R.W, Juggins, S., Anderson, N.J., Cameron, N.G., Gasse, F. and the EDDI consortium, (1998). **European Diatom Database (EDDI): An Information System For Palaeoenvironmental Reconstruction.** Abstracts of the 15th International Diatom Symposium, Perth, Western Australia.

Bennion, H., Battarbee, R.W, Juggins, S., Anderson, N.J., Cameron, N.G., Gasse, F. and the EDDI consortium, (1999). **European Diatom Database (EDDI): An Information System For Palaeoenvironmental Reconstruction.** Abstracts of the Lake 99 Conference, Copenhagen 16-20th May 1999.

Bennion, H., Battarbee, R.W, Juggins, S., Anderson, N.J., Cameron, N.G., Gasse, F. and the EDDI consortium, (1999). **European Diatom Database (EDDI): An Information System For Palaeoenvironmental Reconstruction.** Abstracts of the Symposium for European Freshwater Sciences (SEFS), University of Antwerp.

Juggins, S., Anderson, N.J., Bennion, H., Lotter, A.F., Hausmann, S., Rioual, P., Ryves, D., Schmidt, R. & Wunsam, S. (2000). **Does merging regional diatom-based nutrient transfer functions and training sets lead to more accurate and precise hydrochemical reconstructions.** Abstracts of the 8th International Paleolimnology Symposium, Kingston, Canada, August 20-21, 2000, pp. 41-42.

Juggins, S. and the EDDI consortium. (2001). **The European Diatom Database (EDDI): A new tool for palaeoenvironmental reconstruction.** Abstracts of the International Conference on Past Climate Variability Through Europe and Africa, Aix-en-Provence, France, 27th-31st August 2001.

Planned papers:
The following papers are currently being prepared for submission:

1. Global level species distributions/biogeography of European diatoms – Lead author, Richard Telford.

2. Global level transfer functions – value added by merging, with comparison of different methods and an example application – Lead author, Steve Juggins.

3. Application of EDDI to reconstruct trends in TP in European lakes – Lead-authors, Helen Bennion and John Anderson.

4. The EDDI project, a brief description – Lead author, Rick Battarbee.

Members of the consortium and other scientists will be encouraged to make use of the information system to explore new diatom-environment relationships.

European Diatom Database (EDDI). An Information System for Palaeoenvironmental Reconstruction

Battarbee, R.W. Juggins, S., Gasse, F., Anderson, N.J., Bennion H, & Cameron, N.G. 2000 European Diatom Database (EDDI). An Information System for Palaeoenvironmental Reconstruction. *European Climate Science Conference*, Vienna City Hall, Vienna, Austria, 19-23 October 1998, pp. 1-10.

Abstract The European Diatom Database (EDDI) is an information system that w allow state of the art techniques for diatom-based environmental reconstruction for pH, total phosphorus and salinity, to be available to a range of users. The system aims to collate and harmonise existing Eurc training sets for diatoms and water chemistry and to make the combine available on CD-ROM and the WWW together with images of the main diatom taxa and software needed for data analysis. Copies of microsco slides used in the training set will be held for reference in the Natural H Museum, London, and the information system will be managed by MEDIAS-FRANCE in Toulouse. Further information about the project ca found at our web site: http://medias.meteo.fr/eddi/.

Introduction Diatom-water chemistry transfer functions have become one of the mos widely used and reliable means of environmental reconstruction, espec in relation to water quality problems such as lake acidification (Birks et 1990), lake eutrophication (Anderson 1997, Bennion et al. 1996) and tc problem of climate variability (Fritz et al. 1991, Gasse et al. 1987). The derived from a regional training, or calibration, dataset consisting of mc diatom samples and associated environmental data collected from a nu of lakes spanning the full range of the environmental gradient, or gradie of interest. These data are used to define a mathematical response fun or transfer function, that relates taxon distribution and abundance to contemporary limnological conditions. Once derived, the transfer functi can be applied to reconstruct values of the environmental variable from sediment-core assemblages.

Considerable progress has been made over the last decade in generati these regional trainings sets and in the development of transfer functio a range of environmental and palaeoenvironmental applications, espec with respect to pH, total phosphorus (TP) and salinity.

Diatom-pH transfer functions

Diatoms have been used traditionally to reconstruct the long-term acidification of low alkalinity waters, and the more recent acidification o lakes affected by "acid rain". Diatom-based pH reconstruction has now become a standard technique, used for example in national acid rain monitoring programmes within Europe (e.g. Juggins et al. 1996) and in European research programmes (e.g. Cameron et al. in press). There i:

12/07/00 16:

considerable interest in their potential role in reconstructing climate cha
At low alkalinity sites where the impact of acid deposition

is absent or is reduced in significance, pH changes through time are dr
by natural processes. Over long time-scales, base cation leaching of
catchment soils may cause progressive acidification (Renberg 1993). C
shorter time-scales, however, there is increasing evidence that pH varie
with temperature and with increased duration of the ice-free season
(Psenner & Schmidt 1992). High resolution diatom analysis and pH
reconstruction can then be used to reveal decadal and centennial varia
in temperature during the Holocene.

Diatom-total phosphorus (TP) transfer functions

In well-buffered surface waters, diatom floras are more strongly influen(
lake-water nutrient concentrations than by pH. In such cases diatoms c
used to reconstruct past levels of total phosphorus, that may vary eithe
because of nutrient enrichment or because of climate change. This
technique is being increasingly used in water quality management to
establish the extent to which lakes have been enriched in comparison t
earlier (e.g. mid-nineteenth century) baselines. At times in the past whe
nutrient-enriching impact of humans can be discounted, diatom based 1
reconstruction can be used as a proxy for climate change, as warmer
conditions accelerate the rate of nutrient cycling and increase rates of
primary productivity (e.g. Ryves et al. 1996).

Diatom-salinity transfer functions

At high ion concentrations the composition of diatom assemblages bec(
influenced most strongly by the salt content of the water. As for pH and
the changing salinity of lake water can be due to human impact (e.g.
salinisation) or climate change, or both. Diatoms respond sensitively to
changes in lake water salinity and chemical facies, especially when che
changes occur close to the boundary between fresh and brackish watei
g TDS 1-1). Training sets have been developed for African (Gasse et a
1995), North America (e.g. Wilson et al. 1996, Fritz et al. 1993) and
European (Reed 1998) lakes and there has already been harmonisatioi
datasets between North America and Africa within the CASPIA Project
(Juggins et al. 1994).

▲ go to top

Work Program Harmonisation of data sets

The first goal of EDDI is to harmonise the small datasets generated for
different regions and for different environmental variables from the who
Europe (and parts of Africa & Asia) to produce a single high quality,
integrated training set of diatoms, with site information and environmen
data. Table 1 lists the principal existing training sets that are being
combined, and Figure 1 shows their approximate geographical coverag

Table 2 lists the diatomists, statisticians and database experts involved
the project, and their affiliations. Harmonisation is a two step procedure
harmonisation within the pH, TP and salinity training sets taking place k
harmonisation between the data sets.

The methodology of the diatom harmonisation involves the standardisa
taxon nomenclature and codes, the screening of slides from the training
samples to assess consistency between analysts, and the full documen
of decisions supported by hard copy micrographs and stored electronic
images of all taxa.
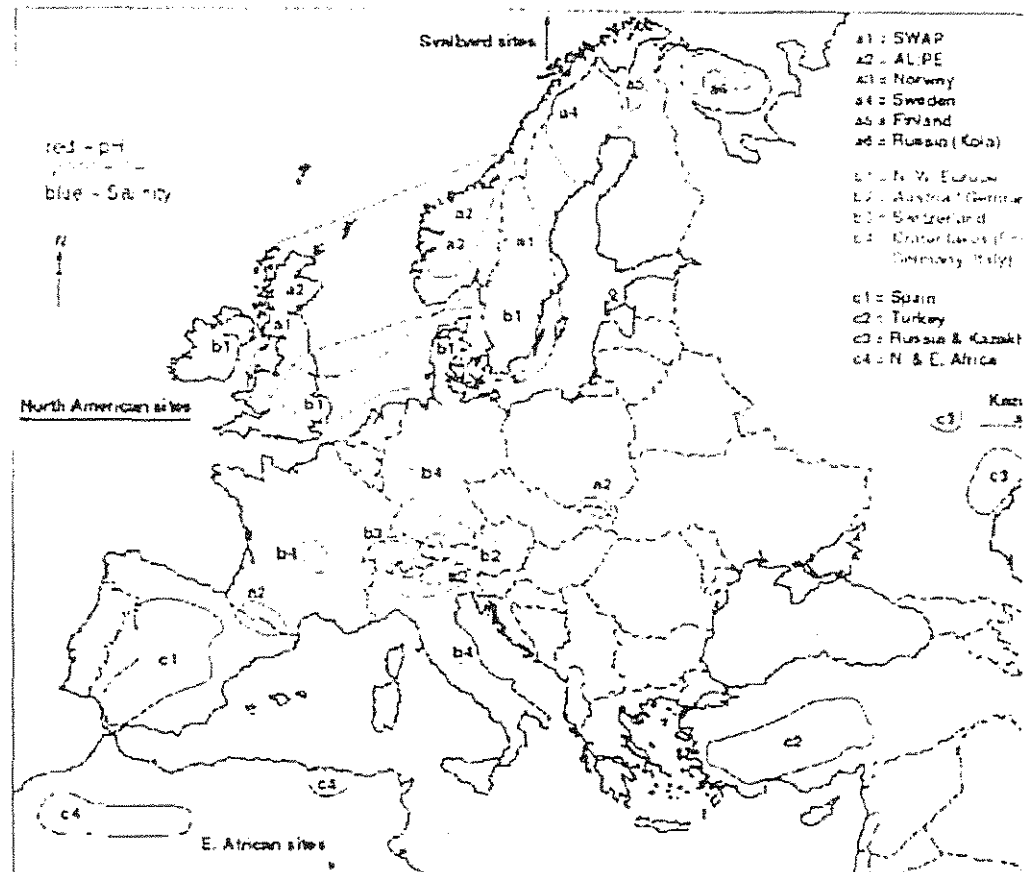

Figure 1: Map of European diatom training datasets

In addition to data on diatom composition, training sets include informa:
on sites (latitude and longitude, lake area, maximum depth etc.) and on
water chemistry (major ions, nutrients, pH etc.). As for the diatom data t
will be standardised, in particular to ensure that the numerical data are
expressed in common units of measurement. These data will then be lir
with the harmonised diatom data to produce a single relational databas
containing all taxonomic and training set data.

To enable use of the diatom data for environmental reconstruction in th
future, diatomists will be able to ensure taxonomic compatibility betwee
core assemblages and the integrated training set by using the electroni
images of taxa in the training set on CD-ROM or by using original slide:

the training set samples that will be archived in the Natural History Mus
in London.

Numerical procedures

The second major goal of EDDI is to develop software evaluate the
performance of a range of statistical and numerical techniques for
quantitative palaeoenvironmental reconstruction. Inferring environment
variables from diatom assemblages is a difficult multivariate calibration
problem and a variety of numerical techniques have been proposed for
solution. In palaeolimnology, the method of weighted averaging (WA) (I
Braak & Barendregt 1986) has gained support and appears to be partic
suited to the noisy, species-rich, compositional data that characterise d
training sets (ter Braak & Van Dam 1989).

More recently other techniques have also been proposed. These includ
weighted averaging partial least squares (WAPLS) (ter Braak & Juggin:
1993) and modern analogue techniques (Guiot 1990). Applications of b
these methods show that they can result in a significant reduction in
prediction error for some datasets. In addition, other statistical methods
Bayesian analysis) or data-based approaches (e.g. neural networks an
taxon response surfaces) show great promise in related areas of multiv
calibration. However, it is clear from these studies that there is no singl
of numerical procedures that will be optimal, that is, give the lowest
prediction error and minimum bias, for all training sets and fossil data. (
our current poor understanding of the performance and properties of
numerical techniques used for environmental reconstruction, the new
datasets being assembled in EDDI will be used to make a detailed
evaluation of a number of different reconstruction methods, including th
error estimation. This will involve the modification of existing and
development of new software for their implementation, and their thorou(
evaluation using real and simulated data.

Taphonomic problems, including the selective dissolution of weakly sili(
taxa, may bias assemblage composition, especially in saline and high p
lakes, and reduce the accuracy of reconstructions. This problem will be
addressed in EDDI by (i) including images of poorly preserved material
allowing diatomists to follow conventions for counting dissolved or brok
diatoms, and (ii) using information about the species composition and
preservation profile of an assemblage to calculate dissolution indices, v
in turn will provide an estimate of potential bias.

EDDI seeks to avoid adopting a single "standard" method and aims to
produce a set of guidelines for "best practice" in the numerical analysis
diatom training-sets. These guidelines will then allow users to make an
informed choice as to the best set of procedures to use with a particula
training and fossil dataset. These results will ultimately determine whicl
reconstruction procedures are included in the diatom information syster
(see below), and will help to guide the diatomist through the process of

12/07/00 16:

26

environmental reconstruction contained in the system.

A diatom information system

The final aim of EDDI is to develop a diatom information system that wil
allow diatomists to perform quantitative environmental reconstructions 1
sediment-core data (Figure 2). The system will be managed by
MEDIAS-FRANCE, and distributed via the WWW and on CD-ROM. It w
incorporate:

(i) a taxonomic database that contains the taxonomic information, code
checklist, and identification criteria,

(ii) an ecological database that contains the raw diatom counts, the
environmental information associated with the sample, and site informa

(iii) a database of transfer functions that contains information on the
individual transfer functions derived from the numerical analysis of the
dataset, and

(iv) a series of software tools that will allow diatomists to interrogate an
interact with the databases.

In addition, "browser software" will be developed to let users search an
the databases for taxonomic or ecological information, and to plot this
information as either geographical distributions or distributions along
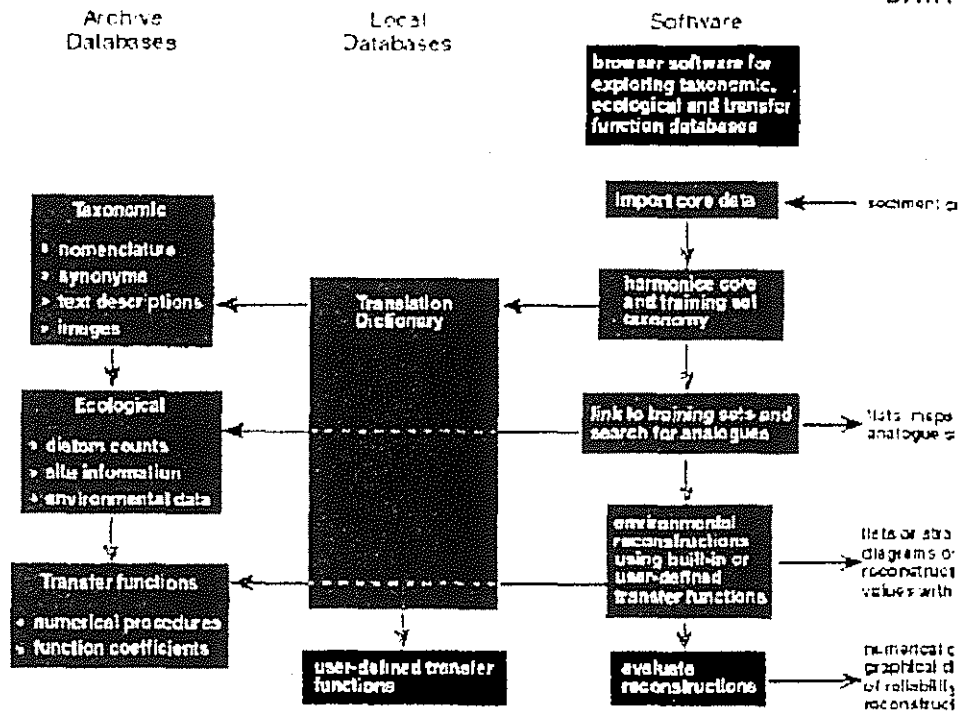selected environmental gradients.

Figure 2: The Diatom Information System

go to top

Discussion The primary purpose of the information system is to enable the harmon training set to provide a new set of robust transfer functions for pH, TP salinity. However, it is hoped that the new integrated training set will als allow additional transfer functions to be derived e.g. for inferring dissol\ organic carbon (DOC) and labile aluminium in soft waters (cf. Birks et a 1990), and ionic ratios in saline environments (cf. Gasse et al. 1995). Ir addition, given that the new training set will cover most of the geograph regions of Europe, there is scope for mapping the biogeographical distribution of diatoms within Europe, as well as for exploring the relatic between diatom assemblages and climate, and for developing and testi diatom-temperature transfer functions (cf. Pienitz et al. 1995).

Ultimately we hope that the ready availability of the information system, useful for both studies of water quality and of climate change, will be accessible to a greater range of research scientists, and that, by using standard taxonomic and numerical methods, diatom analysis will becon increasingly robust technique for inferring environmental change.

Further information about the project may be found on our web site at http://medias.meteo.fr/eddi/

go to top

R. W. Battarbee H. Bennion and N. Cameron are at the Environmental Change Research Centre, University College London, 26 Bedford Way London WC1H OAP, UK. Email: rbattarb@geog.ucl.ac.uk, hbennion@geog.ucl.ac.uk and ncameron@geog.ucl.ac.uk.

S. Juggins is at the Department of Geography, university of Newcastle, Newcastle upon Tyne, NE1 7RU, UK. Email: Stephen.Juggins@ncl.ac.

F. Gasse is at the Centre Européen de Recherche et d'Enseignement c GéoScience de l'Environment, CNRS-universite de Aix-Marseille m, UM 6635, B.P. 80, 13545, Aix en Provence, Cedex 4, France. Email: gasse@cerege.fr.

N. J. Anderson is at the Geological Survey of Denmark & Greenland, Environmental History & Climate Department, Thoravej 8, Copenhagen Denmark. Email: nja@geus.dk.

References Anderson, N. J. (1997) Historical changes in epilimnetic phosphorus concentrations in six rural lakes in Northern Ireland. *Freshwater Biolog)*

Bennion, H., S. Juggins and N. Anderson (1996) Predicting epilimnetic phosphorus concentrations using an improved diatom-based transfer function, and its application to lake eutrophication management. *Environmental Science and Technology 30:* 2004-2007.

Birks, H. J. B., J. M. Line, S. Juggins, A. C. Stevenson and C. J. F. ter E (1990) Diatoms and pH reconstruction. *Philosophical Transactions of th Royal Society of London, B 327: 263-278.*

Cameron, N. G., H. J. B. Birks, V. J. Jones, F. Berge, J. Catalan, R. J. Flower, J. Garcia, B. Kawecka, K.A. Koinig, A. Marchetto, P. Sánchez-Castillo, R. Schmidt, M. Sisko, N. Solovieva, E. Stefkovál & M (in press) Surface sediment and epilithic diatom pH calibration sets for remote European mountain lakes (AL:PE Project) and their comparison the Surface Waters Acidification Programme (SWAP) calibration set. *J(of Paleolimnology.*

Fritz, S. C., S. Juggins, R. W. Battarbee and D. R. Engstrom (1991) Reconstruction of past changes in salinity and climate using a diatom-b transfer function. *Nature* 352: 706-708.

Gasse, F., J. C. Fontes, J. C. Plaziat, P. Carbonel, I. Kaczmarska, P. D Dekker, I. Soulie-Marsche, Y. Callot and P. A. Dupeuble (1987) Biologi remains, geochemistry and stable isotopes for the reconstruction of

environmental and hydrological changes in the Holocene lakes from Nc Sahara. *Palaeogeography, Palaeoolimatology, Palaeoecology 60:* 1-46

Gasse, F., S. Juggins and L. Ben Khelifa (1995) Diatom-based transfer functions for inferring past hydrochemical characteristics of African lake *Palaeogrography, Palaeoclimatology, Palaeoecology* 117: 3-54.

Guiot, J. (1990) Methodology of the last climatic cycle reconstruction in France from pollen data. *Palneogeography, Palaeeclimatology, Palneoecology* 80: 49-69.

Juggins, S., R. W. Battarbee, S. C. Fritz and F. Gasse (1994) The CAS project: diatoms, salt lakes, and environmental change. *Journal of Paleolimnology* 12: 191-196.

Juggins, S., R. Flower and R. Battarbee (1996) Palaeolimnological evic for recent chemical and biological changes in UK Acid Waters Monitorii Network sites. *Freshwater Biology 36:* 203-219.

Pienitz, R., J. Smol and H. Birks (1995) Assessment of freshwater diato quantitative indicators of past climatic change in the Yukon and Northw Territories, Canada. *Journal of Paleolimnology 13:* 21-49.

Psenner, R. and R. Schmidt (1992) Climate-driven pH control depositio *Nature 356:* 7X1-783.

Reed, J. M. (1998) A diatom-conductivity transfer fun *Paleolimnology 1!* 399-416.

Renberg, I., T. Korsman and N. J. Anderson (1993) A temporal perspec of lake acidification in Sweden. *Ambio* 22: 264-271.

Ryves, D., V. J. Jones, P. Guilizzoni, A. Lami, A. Marchetto, R. W. Batt; R. Bettinetti and E. C. Devoy (1996) Late Pleistocene and Holocene environmental changes at Lake Albano and Lake Nemi (central Italy) a: indicated by algal remains. *Palaeoenvironmental Analysis of Italian Cra Lake and Adriatic Sediments. Mem. Ist. ital. Idrobiol.* (eds) P. Guilizzoni F. Oldfield. 55: 119-148.

ter Braak, C. J. F. and L. G. Barendregt (1986) Weighted averaging of species indicator values: its efficiency in environmental calibration. *Mathematical Biosciences* 78: 57-72.

ter Braak, C. J. F. and S. Juggins (1993) Weighted averaging partial le; squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages. *Hydrobiologia* 269/ 485-502.

ter Braak, C. J. F. and H. van Dam (1989) Inferring pH from diatoms: a

12/07/00 16:

comparison of old and new calibration methods. *Hydrobiologia* 178: 20?

Wilson, S. E., B. F. Cumming and J. P. Smol (1996) Assessing the reli?
of salinity inference models from diatom assemblages: an examination
219-lake data set from western North America. *Canadian Journal of Fis*
and Aquatic Sciences 53: 1580-1594.

Table 1: List of Europaen diatom training datasets to be harmonised in ?

pH datasets

Taxonomic Co-ordinator: N.G. Cameron	Region	number of samples	pH range
A1: SWAP	Scotland, Wales, NW England, S & W Norway and SW Sweden	178	4.3 - 7.3
A2: AL:PE	Spain, Italy, Austria, Slovakia and Poland	104	4.5 - 8.0
A3: BERGEN 1	S and W Norway	96	4.3 - 7.9
A3: BERGEN 2	W Norway	78	4.5 - 7.9
A3: SVALBARD	Svalbard	27	5.2 - 8.0
A4: N. SWEDEN	Northern Sweden	151	5.0 - 8.3
A5: LAPLAND	Finnish Lapland	101	5.0 - 7.8
A6: KOLA	Kola Peninsula, Russia	22	5.0 - 7.4

TP datasets

Taxonomic Co-ordinator: N.J. Anderson	Region	number of samples	TP range (µg L⁻¹)
B1: NW EUROPE	Denmark, Sweden, England, Scotland, Wales and Northern Ireland	164	5 - 1200
B2: CENTRAL EUROPE	Austria, Italy and Germany	86	4 - 270
B3: SWISS	Switzerland, France, Germany	75	5 - 300
B4: CRATER LAKE	France, Germany and Italy	30	4 - 260

Salinity datasets

Taxonomic Co-ordinator: F. Gasse	Region	number of samples	Conduct. range (mS cm⁻¹)
C1: SPAIN	Spain	74	0.2 - 340
C2: TURKEY	C. Turkey	30	0.1 - 120
C3:RUSSIA/ KAZAKHSTAN	Caspian Lowlands, Aral Sea & Balkash Depression	50	0.2 - 48
C4: AFRICA	N. Africa, E. Africa and Niger	282	0.04 - 100

31

Table 2: List of diatomists, statisticians and database experts involved in EDDL

Diatomists	
Prof. N.J. Anderson	Geological Survey of Denmark & Greenland, Denmark.
Dr. P. Barker	Department of Geography, University of Lancaster, UK.
Prof. R. Battarbee	ECRC, University College London, UK.
Dr. H. Bennion	ECRC, University College London, UK.
Dr. N. Cameron	ECRC, University College London, UK.
Dr. E. Cox	Natural History Museum, London, UK.
Dr. S. Droop	Royal Botanic Gardens, Edinburgh, UK.
Dr. R. Flower	ECRC, University College London, UK.
Dr. F. Gasse	CEREGE, Marseille, France.
Dr. V. Jones	ECRC, University College London, UK.
Dr. A. Korhola	Department of Geography, University of Helsinki, Finland.
Dr. T. Korsman	Department of Biology, University of Umea, Sweden.
J. Larssen	Botanical Institute, University of Bergen, Norway.
Dr. A. Lotter	Institute of Geobotany, University of Bern, Switzerland.
Prof. D. Mann	Royal Botanic Gardens, Edinburgh, UK.
Dr. J. Reed	Department of Geography, University of Newcastle, UK.
Prof. I. Renberg	Department of Environmental Health, University of Umea, Sweden.
P. Rioual	ECRC, University College London, UK.
Dr. D. Ryves	Geological Survey of Denmark & Greenland, Denmark.
Prof. R. Schmidt	Institute of Limnology, Mondsee, Austria.
N. Solovieva	ECRC, University College London, UK.
J. Weckstrom	Department of Geography, University of Helsinki, Finland.
Dr. S. Wunsam	Institute of Limnology, Mondsee, Austria.
Statisticians and Database experts	
Dr. G. Begni	MEDIAS, Toulouse, France.
Prof. H.J.B. Birks	Botanical Institute, University of Bergen, Norway.
Dr. F. Chalie	CEREGE, Marseille, France.
Dr. S. Juggins	Dept. Geography, University of Newcastle, UK.
Dr. P. Maliphant	ECRC, University College London, UK.
Dr. C. ter Braak	Centre for Biometry, Wageningen, The Netherlands.

▲ go to top

12/07/00 16

32

Annex 4 Workshop and meeting reports

Minutes of First EDDI Workshop: 21-22 May 1998, UCL.

Thursday 21st May

The steering group (Helen Bennion, Nigel Cameron, Rick Battarbee, Steve Juggins, Francoise Gasse) met prior to the meeting. John Anderson was unable to attend the meeting. The need to publicise the EDDI project was discussed.

ACTION: to create an EDDI web page and to submit a poster to the International Diatom Symposium in Australia this autumn (deadline for abstracts is 30-June-98). A flier could be produced from the poster. Steve Juggins and Helen Bennion will take the lead on this.

14:00 Start of Meeting.

Chair: Rick Battarbee.

Present: Dr Philip Barker, Prof Rick Battarbee (RWB), Dr Gerard Begni, Dr Helen Bennion (HB), Emily Bradshaw, Dr Nigel Cameron (NC), Dr Francoise Chalie, Dr Don Charles (DC), Dr Eileen Cox (EC), Dr Stephen Droop (SD), Dr Roger Flower, Dr Francoise Gasse (FG), Dr Steve Juggins (SJ), Dr Atte Kohola, Dr Tom Korsman (after 3:30pm), Dr Jorunn Larsen, Dr Anson Mackay, Dr Paula Maliphant, Prof David Mann (DM), Dr Jane Reed (JR), Patrick Rioual, Dr Dave Ryves, Dr Simon Patrick.

Apologies: Prof John Anderson, Prof John Birks, Dr Joel Guiot, Dr Roland Hall, Dr Andy Lotter, Dr Aldo Marchetto, Prof Reinhard Pienitz, Dr Ingemer Renberg, Dr Peter Rosen, Prof Roland Schmidt, Prof Cajo ter Braak, Dr Jan Weckstrom, Dr Sybille Wunsam.

Afternoon session

Helen Bennion provided an **overview of the EDDI proposal** - content, structure, personnel and timescales, and outlined the aims of this first EDDI workshop.

Gerard Begni from Medias introduced the **Medias data centre and web site,** and described it's potential contributions within EDDI. The MEDIAS team can supply a database co-ordinator, a system engineer, a senior engineer and a software engineer. They could contribute the following: i) development of a user interface to allow the selection of data subsets, ii) development of the web interface, iii) production of CDROMS, and iv) development of a GIS interface.

ACTION: Francoise Gasse to arrange a workshop (in Toulouse?) for those involved in work packages 3 and 4 to assess the exact roles of MEDIAS and others, and to decide on how to proceed.

Steve Juggins provided an overview of the proposed **EDDI Diatom Information System (DIS).** Steve stressed the importance of keeping an open mind throughout EDDI with the possibility of delivering more than that stated in the proposal itself. The aim is to create a tool for taxonomists, ecologists, biogeographers, and teachers as well as being a 'flora' ID system. The DIS will include a flexible taxonomic database, an ecological database, a database of transfer functions and supporting software.

The diatom taxonomic database should include: the taxon name, list of synonyms, a textural description with warning notes (eg. don't confuse with....; note feature.... etc), an ID guide, bibliography, a range of images that cover the variability of each taxon, enlarged drawings to illustrate key features and harmonization tables.

The ecological database should include information from the literature on distribution of each taxon, habitat preferences, lifeform and ecology (using existing coding systems eg. van Dam). Eileen Cox stressed the need to issue health warnings with the use of literature and floras for

describing taxa because of the uncertainty over whether our species concepts were the same as those in the published descriptions. She advised the use of type material where possible though this will probably not be feasible given the scale and timetable for EDDI.

The DIS should be viewed as a series of databases rather than as a single structure, enabling us to use aspects of the databases to address different issues in the future.

Steve described the links to the Automated Diatom Identification and Classification project (ADIAC). There is a lot of overlap between ADIAC and EDDI with the potential to bolt on extra pieces of software such as image analysis and ADIAC extensions to the EDDI DIS. For example we may want to develop applications that allow different levels of taxonomic resolution, such as genus level for teachers and lake managers.

The EDDI DIS should also be dynamic with the opportunity to add new datasets in future. This would need close supervision.

Steve indicated that there was no plan to incorporate plotting facilities into the DIS and that output would most likely be as .TIL and .XLS files to allowing plotting elsewhere.

Don Charles talked of his experiences with developing two **North American databases**. The Diatom Paleolimnology Data Cooperative (DPDC) is a NOAA funded project within the Paleoclimate Program, developed by Don, Roger Sweets, Tim Sullivan and Kellie Vache, with Platt-Bradbury, Fritz, and Smol acting as advisors.
(see http://www.indiana.edu/~diatom/dpdc/dpdc.html).

It aims to make paleoclimate inference data readily available to the climate change community and to improve the ability of diatom paleolimnology to infer climate-related characteristics. All data will be available on the Web. The database is an ACCESS system and includes stratigraphic and calibration datasets, the raw and diatom-inferred data and so far includes c. 450 sites. Don pointed out that the most significant problems had been formatting the different datasets, acquiring/entering all the environmental data and harmonising the taxonomy. These will also be the main issues to address in EDDI. In the future, the DPDC aims to add more datasets, to streamline and improve the taxonomy, to promote the database and to link with the EDDI project.

The second database described by Don was the North American Diatom Ecological Database (NADED), developed by Don Charles, Frank Acker and Pat Cotter at the Academy of Natural Sciences of Philadelphia. This is an ACCESS database, including diatom counts and accompanying count information and harmonisation dictionaries. The system records taxonomic amalgamations and a has browser facility to scroll taxa lists as well as applications for sample selection by location, habitat, environmental range etc. The database calculates the weighted average of the samples and is used more for river water quality assessments than for palaeolimnogical investigations. Data can be output in a range of formats. The system also includes images.

Contractual matters were discussed and it became evident that the sub-contracting structure within EDDI is far from clear!
ACTION: Anson Mackay (financial co-ordinator, UCL) to send out coherent statement of contractual issues to all participants.

There was also some concern raised by Francoise about the losses made by having to transfer between currencies.
ACTION: Francoise should check with the EU in Brussels and with her own institute on the possibility of dealing with ECU. This is how our French partners in another EU project handle financial exchange and it appears to work well.

Simon Patrick reported that the EDDI contract runs from 1 April 1998 to 31 March 2001. A pro-forma progress report form will be circulated to the main 4 partners (UCL, John Anderson, Steve Juggins & Francoise Gasse) by Helen Bennion (scientific co-ordinator) at the end of March each year, to be completed and signed and returned to Helen by mid-April. Helen will then collate the forms and send to Brussels by 1 May each year. Sub-contractors do not have to complete these forms.

The 4 main partners must also complete cost statements via their university accounts departments and send to Anson Mackay (financial co-ordinator), who will compile these and forward to Brussels. Partners need to manage their budgets and are responsible for their sub-contractors. The statements will need to be with Anson by the end of March each year in order to allow time to go via the university system. These cost statements should summarise monies spent in that year so that by the end of year 3, the full 100% of the costs are accounted for. Note that 40% of the contract finances have already been paid and 60% is in arrears. Please note that Brussels will not release the final 10% until the final report is submitted. The report is expected 3-6 months after the contract end date ie. summer 2001.

Friday 22nd May
09:30 Morning Session.
Chair: Helen Bennion
Present: Dr Philip Barker, Prof Rick Battarbee, Dr Helen Bennion, Emily Bradshaw, Dr Nigel Cameron, Dr Francoise Chalie, Dr Don Charles, Dr Eileen Cox, Dr Stephen Droop, Dr Roger Flower, Dr Francoise Gasse, Dr Viv Jones, Dr Steve Juggins, Dr Atte Korhola, Dr Tom Korsman, Dr Jorunn Larsen, Dr Paula Maliphant, Prof David Mann, Dr Jane Reed, Patrick Rioual, Dr Dave Ryves, Dr Carl Sayer, Nadia Solevevia

Apologies: as above and Dr Gerard Begni.

Dataset harmonization
The morning was dedicated to the introduction of the **individual training sets** that will be incorporated into the EDDI DIS. The first session was devoted to the **pH datasets**.

Steve Juggins described the **SWAP** set which is well documented and published. It is a very heterogeneous dataset with no rigorous AQC on the chemistry data. There are no TP data. The raw data are available (with the exception of Swedish data) and therefore data screening will be possible.

Nigel Cameron described the **AL:PE** set. This involved a large number of taxonomists and so harmonisation was carried out within the project. The AL:PE set is not so skewed towards acid sites as the SWAP dataset. The AQC chemistry was not focused on and there are no consistent suites of chemistry except pH, DOC and physical details. A spreadsheet of the raw data is available. AL:PE involved the development of a separate transfer function based only on the epilithic diatoms, which worked well.

Jorunn Larsen introduced the **BERGEN** set. This includes 96 lakes and 169 taxa. Twenty chemical determinands were measured on 4 occasions, including TP data. The lakes range from pH 4.3-8.3 but the mean is only 5.8. One potential problem for EDDI is that Frode Berge's slides are lost although John Birks should have a taxa list with notes.
Viv Jones described the **Svalbard** dataset. This includes 23 lakes covering a broad range of latitudes. The chemistry data are based on a single summer measurement and include the major ions, conductivity, TP, NO_3 and chlorophyll a. Physical data are also available and all raw data are in spreadsheet format. The diatom assemblages are dominated by benthic forms with 184 taxa, 25 of which are problematic (small *Navicula* spp. & *Achnanthes* spp. in particular).

Tom Korsman described the **Northern Sweden** dataset of 118 boreal forest, headwater lakes. The mean pH of the dataset is 6.5 and so the species pH optima are likely to be higher than SWAP. There is also work in progress on 36 lakes in mountain regions which could be contributed at a later stage. The main chemical variables are pH, alkalinity, colour, altitude and maximum water depth. There are no nutrient data. The taxonomy largely follows SWAP protocols and there are not too many problems with the major taxa.

Atte Korhola described the **Lapland** dataset of 98 lakes from two different regions. The dataset includes 19 environmental variables (mean pH 7.0). The East Lapland set includes NO_3 and NO_2 but not TP and the northwest set has no nutrient data. The main taxonomic problems are small *Achnanthes* and *Navicula* taxa, *Nitzschia* spp. in general, and *Fragilaria* girdle views.

Nadia Solovieva described the **KOLA** dataset of 24 lakes spanning the pH range 5.0-7.5. The lakes are comparable to those in Sweden and Finland with similar species pH optima.

The second session focused on the **TP datasets**.

Helen Bennion described the **Northwest European** dataset of 152 lakes from England, Wales, Northern Ireland, Denmark and southern Sweden. A full suite of environmental data is available for all lakes including mean pH, TP, conductivity, alkalinity, the major ions, SiO_2 and chlorophyll *a*. The raw data could be made available. The main taxonomic problems are the small *Cyclotella* spp., *Fragilaria* girdle views, the medium-sized *Stephanodiscus* spp. and the long, fine *Synedra* spp.

Helen Bennion presented the **Central Europe** dataset on **Sybille Wunsam** and **Roland Schmidt's** behalf. This includes 86 lakes of the Alps and pre-Alpine regions of Austria, Germany & Italy. A comprehensive range of environmental variables are available for the lakes including mean pH, TP, conductivity, secchi depth, temperature, nitrate and ammonium. The raw data, however may not be available. The main taxonomic issues concern the *Cyclotella* taxa. The data are published.

Helen Bennion presented the **Swiss** dataset on **Andy Lotter's** behalf. The dataset includes 72 lakes with a suite of environmental data, namely catchment, climate and limnological parameters such as pH, alkalinity, DOC, TN and TP. The mean data are published but the status of the raw data is not known. There appears to be overlap with the Northwest European and Central European datasets in terms of taxonomic problems.

Patrick Rioual presented the **French Crater Lakes** dataset which includes 8 true crater lakes and additional lakes from the Massif Central region of France. Patrick has made an application to add 8 Italian and 9 German lakes to the training set but this awaits funding. The lakes were sampled on 8 occasions over a two year period and a suite of analyses has been carried out including pH, conductivity and TP. The main taxonomic problems are the *Cyclotella* taxa and the long fine *Fragilaria/Synedra* complex.

The final morning session was dedicated to the **salinity datasets**. **Francoise Gasse** introduced some of the problems in relation to diatom-salinity transfer functions, namely diatom dissolution causing bias in the fossil record (this needs to be documented in some way- index?) and large interannual variability in both diatom floras and chemistry of small semi-arid lakes.

Francoise Gasse described the **African** dataset which includes 282 lakes distributed throughout Africa. CASPIA has already addressed some of the taxonomic harmonisation

issues in these data. The environmental data focus on the major ions but pH is also available. There are no nutrient data in most cases.

Jane Reed introduced the **Spanish** and **Turkish** datasets. The Spanish work based on 71 lakes is published in Journal of Paleolimnology. The Turkish work is not published but Jane has just been successful in getting a 2 year funded project to develop this work further. There are 38 Turkish lakes samples so far. There are a lack of freshwater sites but this could be solved by extending into the Balkans. The dataset will benefit from merging it with the African data. Key taxonomic problem areas include the *Cymbella* taxa and *Cyclotella ocellata* but some issues have been addressed already by CASPIA.

Steve Juggins introduced the **Russian/Kazakhstan** dataset, which includes 25 lakes forming part of a Holocene climate change project. Steve anticipates that there may be some problems with the water chemistry. The diatomists were counted by a Russian diatomist and so careful harmonisation will be necessary.

The possibility of including Roland Schmidt's Adriatic lake dataset was discussed.
ACTION: Helen to contact Roland to see if he is interested in contributing these data.

Atte Korhola briefly introduced the developments that are being in multivariate statistical techniques with colleagues at the University of Helsinki. He spoke of the **Bayesian approach** and its application to ecological data. They have compared its performance with more commonly used weighted averaging techniques and are optimistic that this may provide a useful alternative method for developing transfer functions. The results of these comparative studies will hopefully be published in the near future.
ACTION: It was decided that it will be important to discuss the potential of these new techniques in future EDDI workshops.

The afternoon session focused on the **protocols for taxonomic harmonization** within EDDI. The first session introduced examples of previous projects where taxonomic harmonization was performed. **Don Charles** spoke of his experience with **PIRLA**. They decided on key taxa to work on prior to each workshop to keep the workshops focused. Photographs and documentation were brought along to the PIRLA workshops. Working plates were then collated and produced by one co-ordinator.

Nigel Cameron described the harmonization procedure within **SWAP**, which is published in the SWAP red book (Stevenson *et al.*, 1991). The main problems were nomenclature, splitting and amalgamation consistency and identification. Nigel also described the **AL:PE** procedure which involved the harmonisation of pre-existing training sets (similarly to EDDI). Diatom material was exchanged between diatomists and names or 'working names' were agreed and documented.

Laurence Carvalho presented the taxonomic harmonisation procedure adopted in the **CASPIA** project. This was somewhat different from the strategies adopted in SWAP and AL:PE. Within CASPIA detailed studies were carried out on just a few key taxa and then a flora with accompanying documentation was produced of the other common taxa. Detailed LM and SEM work was undertaken and reference to the type material was made. Laurence pointed out that this was not likely to be practical within EDDI.

Nigel Cameron and **Steve Juggins** then proposed a **strategy for taxonomic harmonization within EDDI.** This was discussed and the **final guidelines** were produced:

1. All dataset owners to send their data to Steve Juggins in the first instance in Excel spreadsheet format (.XLS). The file should include full taxon names and the raw counts. The authority and any other comments can be included and a code if you have your own current coding system. See Table 1 attached for recommended format for diatom counts.

2. Steve to compile a combined spreadsheet of taxa for each of the pH, TP and salinity working groups.

3. The three co-ordinators of work package 1 (a, b & c) will inspect the combined table for synonyms and correct any obvious inconsistencies. They will circulate a list of the common taxa to all dataset contributors in their work package.

4. Participants will then be asked to provide images of those common taxa for the co-ordinators to confirm common species concepts. Images must be accompanied by information on slide number and the co-ordinates of the specimen on their own microscope. This will allow images to be retaken if necessary.

5. Any remaining problematic taxa will be noted and the co-ordinators will list the problem areas, namely concept differences between analysts. These taxa will form the basis for the first taxonomic workshops (Jan/Feb 1999). A dialogue will be maintained between the three co-ordinators and all other participants to agree the agenda for the workshops.

6. The three co-ordinators will then require microscope slides representing the key problematic taxa for screening etc in preparation for the workshop discussions.

7. The workshops will agree definitions of the problematic taxa.

8. Following the first workshop, individuals will spend time with their own material and material from other training sets documenting problem taxa allocated to them (need to document across all training sets). Some slide swapping may be necessary at this stage.

Eileen Cox outlined the proposed system for **slide archiving of EDDI slides** at the Natural History Museum in London. Eileen suggested that the EDDI collection could stand alone as their own catalogue series (separate from the BM series). In normal circumstances the slides are accompanied by a database of information on locality, habitat, collector, date, co-ordinates, mountant and so on, but given that this information will be held in the EDDI DIS anyway this will not be necessary in our case. However, it will be important to maintain a link between the EDDI and BM databases so that all these data are available. Eileen suggested that 2 slides from every sample were archived.

ACTION: slide owners to email Helen Bennion indicating the ease or difficulty of satisfying these archiving requirements. Are the slides readily available; if not then can some be prepared relatively easily and cheaply? etc......

The second half of the afternoon session was devoted to **protocols for image capture and transfer** and was led by **David Mann, Stephen Droop and Steve Juggins.**

David Mann introduced the session, pointing out that one of the great opportunities within EDDI was to replace the system of Type slides in future. David stressed the importance of producing a consensus taxonomy and the need to have a lot of images of any problematic taxa to capture the full variation; circumscribe to convey a concept.

Stephen Droop gave some **advice on strategies for image capture and transfer** within EDDI:
- It is important that all participants standardise focal depth, light type, orientation of specimens as best as possible. A perpendicular (N-S or E-W orientation) image is

preferable to a diagonal image, as the latter causes loss of detail. Monochrome is sufficient for diatom images.

- Magnification changers on the microscope may be useful as they can greatly effect the resolution of the output image.

- Stephen recommends a minimum camera resolution of 1000x1000 pixels. Need to achieve a resolution of 2.5 pixels per resolution unit which would allow striae densities of 40-50/ 10μm to be seen in the captured images of small to medium sized diatoms.

ACTION: Participants should calculate the resolution of their systems and send images to Stephen Droop in Edinburgh for his comments on resolution and quality.

Within EDDI there will be requirements for two types of image:
i) rough images for working needs and
ii) high quality stored images for the EDDI DIS.
One issue for EDDI is who will take these images. If everyone is involved then we must have an AQC procedure and clear EDDI protocol. Hence.....

....the following **image capture and transfer protocol** was proposed:

1. Capture and store images unmodified in "loss-less" TIFF format. These can be compressed or uncompressed files. Use bright field light type or phase or DIC if necessary.

2. Send images to Steve Juggins where they will be converted into JPEG files and will be loaded onto the EDDI Web page for all participants to view and comment.

3. The final EDDI DIS will contain both JPEG files for Web view and TIFF files for high quality hard copy.

4. Distribution and transfer of images can be either on CDROM (c. 650 Mb) or on a Zip disc (c. 100 Mb). Steve recommends the purchase of Zip drives (c. £100) as the Zip discs are only £8 each.

5. Steve recommends the use of Paint Shop Pro 4 software to clean up images, which is available free for 30 days from the web site -   http://www.hensa.ac.uk/

The final discussion topic was **environmental data harmonization. Steve Juggins** proposed data formats for the water chemistry data, as well as for the site and sample information. These are shown in **Tables 2 and 3** attached.

The full list of **data types required for the training sets** is summarized below:

- Diatom counts
- Site information (location, type etc)
- Sample information
- Environmental information (chemistry data)
- Images and taxonomic text descriptions
- General description of study area (eg. vegetation, soils, geology, use, climate)
- Maps
- List of published papers
- Description of laboratory methods for water chemistry

ACTION: It is important that all participants comment on the data formats in Tables 1, 2 and 3 and on the list above and inform Helen if they anticipate any difficulties in supplying these kinds of data in the suggested formats.

The **receipt of data** can be prioritised as follows;
1. The **diatom data** are top priority as the leaders of the 3 training set activities will need to identify the common taxa and those that are problematic. Co-workers then need to have time to capture images, and a list of key taxa must be compiled in time for the January/February taxonomic workshops.

2. The **water chemistry (environmental) data** take next priority.

3. All **remaining data** (eg. site information, sample information, study area etc) can follow during the course of the EDDI project.

The meeting closed with a brief discussion of the **timetable for future workshops**. According to the timetable in the proposal, the taxonomic (work packages 1a, 1b, & 1c) and statistical (work package 3) workshops should take place in month 9-10 of the project, that is around January 1999. It was suggested that the taxonomic workshops all be held back-to-back with the statistical workshop added to the end of one of these meetings, and that perhaps Helen could attend all workshops to maintain an overview.

Finally, publicising EDDI was discussed. It was agreed that we should submit a poster to the International Diatom Symposium in Perth, Australia in September 1998. An electronic copy of the poster could be made available to all EDDI participants for their own use. In addition, a flier could be produced and a Web Site should be developed.

Table 1 FORMAT FOR DIATOM COUNTS

	A	B	C	D	E	F	G	H
					Sample Code/Number			
1	Code (optional)	Name	Authority	Comments				
2					1	2	3	4
3	3011	**Achnanthes minutissima**		Sensu KLB	1	23	0	45
4	5767	**Navicula gregaria**		KLB Fig 23	23			21
5	7465	**Nitzschia sp. 12**		Image 25	4	5		12
6	8765	**Nitzschia cf. frustulum**		Image 46	3		12	
7								
8								
9								
10		Bold = required data		Blanks or 0 =				
11				zero count				

Table 2 FORMAT FOR MEAN WATER CHEMISTRY DATA

	A	B	C	D	E	F	G
				Sample Code/Number			
1	Determinand	Units	Comments				
2				1	2	3	4
3	pH			6.4			
4	Conductivity	uS / cm		112			
5	Ca	mg/l		1.2			
6	Al^{3+}	ug/l	Total monomeric	30			
7	etc.						
8							
9	No. samples in mean			5			
10	Date of first sample			01/04/88			
11	Date of last sample			06/06/92			
12							
13							
14	Comments on individual samples						
15	Details of averaging method						
16							
17							
18							
19	[Blank cell implies no data						
20	Zero implies 0						
21	Use "< n" to indicate below detection limit]						

Table 3 FORMAT FOR SITE AND SAMPLE INFORMATION

	A	B	C	D	E
1	*Site Information*	Units	Sample Code/Number		
2			1	2	3
3	Site Code				
4	**Site Name**				
5	Notes on sample location				
6	**Lat.**				
7	**Long.**				
8	**UTM Zone**				
9	**UTM (E)**				
10	**UTM (N)**				
11					
12	**Site Altitude**	m			
13	Max Altitude of catchment	m			
14	Lake area	km^2			
15	Max depth	m			
16	Mean depth	m			
17					
18					
19	*Sample Information*				
20					
21	**Sample Type**		**Surface sediment**		
22	**Depth interval for core sample**	**cm**	(0-1)		
23	Corer type		Renberg		
24	**Estimate of time period**	yrs	5		
25	**Water depth**	m	15		
26	Date collected		15/05/91		
27	Location in lake				
28	Comments on sampling				
29					
30	Diatomist				
31	Slide Number				
32	Concentration data	cells / g dw			
33	Slide preparation method				
34	Comments on diatom count				

Essential information shown in **BOLD** font

Minutes of the EDDI Steering Group meetings

1. First Steering Group meeting: 12th February 1999, UCL.
Present: John Anderson, Rick Battarbee, Helen Bennion, Nigel Cameron, Francoise Gasse, Steve Juggins.

1.The role of the contributing diatomists
The role of the diatomists receiving consultancy fees was clarified. The main tasks that they will be asked to be involved in during the remainder of the EDDI project are:
- collation of any outstanding environmental and site data for their training sets (see below).
- collation of information concerning availability of sediment, suspensions, or slides and the location of material (see below)
- responsibility for sending material for preparation of duplicate slides
- intellectual input throughout the project and continuous consultation on taxonomic decisions being made in EDDI
- attendance at the final workshop.

The participating diatomists will not be expected to prepare duplicate slides, capture images (except on rare occasions) or write taxonomic descriptions. These tasks will be the responsibility of the three co-ordinating groups at UCL, GEUS and CEREGE.

2. Preparation of diatom slides for EDDI.
It is now urgent for slides to be made available to the three co-ordinating groups, so that the next phase of the project involving image capture, can progress.
ACTION: Helen will produce a spreadsheet of information detailing the location and status of the material, and availability of slides for each sample in the three training sets. All participating diatomists need to send the following information for each of their samples to Helen as soon as possible, so that the scale of the task can be assessed.

- What material is available?: answer SED for sediment; SUSP for cleaned suspension; DUPS for duplicate slides (2 sets?; please say how many); NONE for no material available.

- Where is the material?: please give details of the location of material for each sample eg. ECRC sediment archive; a slide box in my office; GEUS cold store etc.

- Remarks?: please add any other information that may be relevant eg. diatom preservation poor in this sample; slides may be too concentrated for EDDI etc. This will help us with slide preparation.

It is expected that most slides can be prepared at UCL, with possibly some assistance from Newcastle. This will be decided once the job has been costed up. If slides are prepared by research technical staff then it will be necessary for the three co-ordinating diatomists to screen the slides for quality and suitability for capturing EDDI images and for inclusion in the EDDI archive. Slides need to be of high quality where possible.

3.Handling of taxonomy - synonyms, groupings etc.
Decisions need to be made very soon on how to record and handle synonyms, amalgamations and so on within EDDI, as there are implications here for what information will be collected and recorded by the three co-ordinating diatomists in the next phase of the project.

ACTION: The three co-ordinating diatomists will conduct a pilot study whereby each centre will focus on one taxonomic group. They will capture the relevant images, write taxonomic descriptions, catalogue amalgamations, synonyms, cross-reference to the literature, and make accompanying notes. This should be completed by the end of February (although CEREGE may not be able to capture images as they are awaiting installation of their system).

The procedure will be reviewed and feedback will be requested from the participating diatomists. In the meantime, Steve Juggins will compile a wish-list of what the group would like to see included in the taxonomic database and what kind of questions we want to be able to ask of it. Steve would welcome any thoughts that you may have on this.
ACTION: Please send any thoughts that you have to Steve Juggins on what you think the database should do.

Furthermore, it was considered necessary to include an Illustrated Glossary to taxonomic terminology in EDDI.
ACTION: Need to consult with David Mann and Stephen Droop for advice on the correct way to handle and express synonyms.

4. Image capture protocols
There was some discussion about the merits of standardising to bright field for all images, in addition to phase and/or DIC, although this may not be helpful for small taxa. It was decided to take advice from Stephen Droop at Royal Botanic Gardens.
ACTION: Helen to contact Stephen Droop and circulate guidelines to the three co-ordinating diatomists.

Scale bars will need to be added to the images.
ACTION: The three co-ordinating diatomists need to calibrate their microscopes in terms of pixels per micron and send the information to Steve Juggins at Newcastle.

5.Distribution of images:
ACTION: The three co-ordinating diatomists should save all images as uncompressed TIFF format and send to Steve Juggins in winzipped form via ftp. Steve will them put all the images on to the EDDI web site for all diatomists to view and feedback.

6.Recording information associated with each image:
The following fields were suggested, discussed and agreed:

Slide	Unique number or text to identify the slide
SampleId	EDDI Sample ID
Dataset	EDDI Dataset
TaxonCode	EDDI taxon code as defined by original diatomist
Taxon Name	Optional, as it is defined by the TaxonCode
Microscope	Microscpoe code (we need to produce a list of microscopes used in EDDI)
Coord-N	Specimen coordinates
Coord-E	Specimen coordinates
ImageFile	File name of captured image (note - all images should be captured in uncompressed tif format)
Illumination	Phase, DIC, brightfield (where possible always collect an image in brightfield, as well as other illuminations)
Orientation	Horizontal or vertical
Photographer	Initials of diatomist who captured image
SpecimenNotes	e.g. "Raphe valve", "centre focus", "end focus", etc.
TaxonomicNotes	e.g. "possibly var. intermedia"

All image filenames should follow the format: 8 characters in length. The first character denotes the co-ordinating centre where the images were taken (ie. U=UCL; C=CEREGE; G=GEUS), and the other 7 characters represent a sequence of numbers starting at 1 eg. U0000001.TIF is image no. 1 captured at UCL; C0000126.TIF is image number 126 captured at CEREGE; G0002345.TIF is image no. 2345 captured at GEUS).

ACTION: The three co-ordinating diatomists should record the above information for each image in an Excel spreadsheet at this stage. These data can be uploaded to an Access 97 database at a later date, once everyone is agreed that we are collecting the right information in the most informative way!

7.Database structure:
Once the pilot study is completed and all the associated information is collated, Steve Juggins will formulate a prototype version of the database structure for discussion (Access97).

8.Gaps in Environmental Data: water chemistry data and associated site information data
Steve has produced a status report on what he has received and what is still missing for each dataset in EDDI:

pH:
SWAP: Almost complete. All data that is in Red Book has been sent. A few sites lack lat/longs that need to be looked up.

ALPE: Only pH data has been sent. Need other site & chemical information (except for SWAP lakes included above). Note also that diatom count data is missing for the Spanish ALPE samples.

Bergen1: All catchment & chemical data has been sent. Grid refs are in UTM only so need converting to Lat/longs.

Svalbard: All catchment & chemical data has been sent. Need site names and core & chemistry date information

Sweden: Mean pH, alkalinity & colour data have been sent. Need remaining chemistry and site information plus core & chemistry dates.

Finland: Location (Lat / Longs), catchment data + all mean chemistry has been sent. Need ful site names if possible plus core & chemistry dates.

Kola: No site or chemical data received so far!

TP:
NW Euro: Data complete except for some chemistry for Danish sites. Grid refs need converting to lat/longs. Need core dates.

Central Europe: data for pH, Cond, TP, NO3, NH4, Secchi and site & core information have been sent.

Switzerland: All site & chemical data sent. Need core & chemistry sample dates.

Crater Lakes: All site and coring dates sent. Need all chemistry plus sampling dates.

Salinity:
N & E Africa: All available site and environmental data has been sent, but there are many errors spotted since the 1995 paper. These will need to be checked.

Caspian: All chemical data sent. Need to look up grid refs.

Spain: No site or chemical data sent.

Turkey: No site or chemical data sent.

ACTION: All diatomists need to look through the above list and contact Helen regarding the missing data with information on whether the data are available and when we can expect to receive them. Files containing the missing data should be sent to Steve Juggins at Newcastle, who will convert them into the correct Excel format. The Excel files will then be sent to the 3 co-ordinating diatomists who will collate all the site and chemical data for their group. It would be helpful if all diatomists could provide any relevant accompanying notes with the data regarding possible erroneous values or reservations over data quality.

DEADLINE: All of the above data need to be collated by the end of work package 1, which will mean by October this year.

9. Web page development
The various roles of the EDDI web pages were discussed. They essentially serve 3 purposes:
1) as an information site to advertise the project,
2) as a research tool for transferring images and ideas between participants, and
3) as a platform for distributing the final EDDI information system.

It seems that there has been some confusion and misunderstandings regarding who is responsible for designing and managing the different sites and pages. Following discussion at the meeting, it was agreed that the main EDDI information site (1 above) should be developed and updated by UCL but once the content had reached a level of stability, it could be transfered to Medias for hosting and possibly additional technical input.

It was agreed for practical reasons that initially the diatom images will be distributed via the Newcastle web site, as it is easier for Steve Juggins to make them available on-line from the database he is constructing, and which will expand as the project develops (2 above). The advice of Medias will be sought on their possible role in distributing images and other information once the format of the taxonomic database has been finialised.

The final EDDI information system will be developed by Newcastle, and hosted by Medias on their web server in year 3 of the project (3 above). Discussions will be held with Medias regarding the extent to which they might also contribute to the design and programming of the Information System.
ACTION: Helen to write a letter to MEDIAS to put forward these suggestions and to clarify everyone's role and try to smooth out any misunderstandings. Steve and Francoise will be able to discuss these matters with MEDIAS at a meeting in Arles in March.

Links to other web sites were discussed. Steve Juggins has set up a link from his home page to the EDDI web page developed by Cathy Stickley at UCL. Cathy has also added a link between the EDDI page and the PEP3 page.
ACTION: We need to add a link to the EU Fourth Framework page. Helen to discuss with Cathy Stickley.

46

10.Publications
We need to make progress on the JOPL paper, which builds on the manuscript that was submitted for the Vienna meeting.
ACTION: Rick to contact John Smol to ask his advice on whether to submit as a short note or as a full paper.

Abstracts/Posters etc -
Abstracts have already been submitted by Rick & Francoise for the INQUA meeting (Aug 99) and by Helen for the Lake99 conference in Copenhagen (May 99) in order to publicise EDDI. Other forthcoming opportunities for presenting EDDI are at the EPS2 meeting to be attended by John Anderson in Florence (Sept 99), the European Freshwater Scientists Meeting to be attended by John Anderson in Antwerp (Aug 99) and the Second International Congress of Limnogeology to be attended by Francoise in Brest (May 99). **ACTION: Please let us know of any other opportunities that you can think of.**

Two or three updated versions of the EDDI poster will, therefore, need to be produced by May. The text needs some updating including addition of a few names and acknowledgements. Steve Juggins currently has the poster in electronic format in a Freehand file which is very large and a little awkward to work with because the images are linked rather than embedded.
ACTION: Steve to ftp the file to John Anderson at GEUS who will enquire about producing it at his institute. John will also enquire about the possibility of producing a glossy EDDI flier.

11.Timetable and deliverables for next 6 month phase of EDDI
The project is running to schedule fairly well. The forthcoming deadlines are :

    i) collation of information on availability and location of material/slides by end of February 1999. This is urgent because the co-ordinating diatomists need access to all EDDI samples for capturing the full range of images and for writing descriptions.

    ii) results of pilot study by end of February 1999.

    iii) comments on the pilot study from all EDDI partcipants by early March 1999.

    iv) prototype version of the Access97 database for recording information associated with the images to be developed by end of March 1999.

    v) annual report to EU is due in April 1999 . Helen will be asking participants to fill in their report forms in mid March 1999.

    vi) collation of all environmental and site data by October 1999.
END

2. Second Steering Group meeting: 23rd November 1999, UCL.
Present: John Anderson, Rick Battarbee, Helen Bennion, Nigel Cameron, Francoise Gasse, Steve Juggins, Christine Pailles.

1. Financial matters
(Anson Mackay present)
The matter of payment of consultancy fees to participating diatomists was raised. It was agreed that payment would be made on completion of the work, i.e. receipt of all environmental data and available material for slide preparation.

ACTION: Helen to compile an Excel spreadsheet of outstanding work for circulation to all participants (see below). When all work has been completed, Jorunn, Eileen, Jan, Sybille, Ingemar, Patrick and Sonja should invoice ENSIS Ltd. at UCL, and Jane & Phil should invoice CEREGE, for the full agreed fee of 3000 ecu.

2. The minutes and key action points of the February Steering Group Meeting were reviewed.
No major outstanding issues or action points.

3. Review of project milestones for Work Package 1 " Harmonisation of pH, TP and salinity training sets" at month 18:
Dave Ryves, Christine Pailles and Nigel Cameron produced short progress reports summarising progress on the following deliverables. These were discussed as follows:
i) Checklist of common taxa
Status: Complete.
This was produced by Steve Juggins prior to the first taxonomic workshops in January 1999 and the list was used at the workshops for identifying the key taxa requiring work.

ii) Database of sites and chemistry data
Status: This is still in progress but must be completed by the end of the year.
The major gaps in the data were identified and are indicated on the attached Excel spreadsheet.
ACTION: IMPORTANT FOR ALL - All data contributors to look at the spreadsheet and send any missing data in Excel format to Steve Juggins at Newcastle by the end of January. Steve is continuing to collate the data, update the files and redistribute them to the dataset co-ordinators. Please note that the fields marked "E" indicate where the data are essential for the project. Those marked "P" are preferable but less important than the essential data types. One essential field for which we still have very few details are dates of diatom samples and dates for the chemical analyses. Where mean chemical values have been provided, it is essential that the period of time represented by these data is also provided. If data holders are able to provide seasonal water chemistry with dates, could they please also send these to Steve. This will help to improve the final database.

ACTION: Nigel to provide missing lake names for the ALPE dataset. Send updated Excel file to Steve.

ACTION: As a matter of urgency, Nadia to provide missing data for Kola and Tom to provide missing data for Sweden.

Steve Juggins agreed to take responsibility for harmonising the environmental data, including checking units of measurements, data structure and formats. If all outstanding data is receivedin January, then Steve aims to complete this by end of February 2000.

iii) Database of diatom counts data
Status: Complete except for Caspian Sea samples. It was agreed that no further samples will be added to the EDDI counts database.
ACTION: Jane Reed will re-count the Caspian Sea samples for the Salinity dataset as soon as possible and will pass the data to Steve for incorporation into the counts database.

iv) Dataset of medium quality images with associated information
Status: In progress.
All diatom co-ordinators have captured a large number of images but of varying quality. There appear to be problems with the quality of images from UCL (poor contrast and

48

fuzziness). Also Nigel and Christine still unable to rotate microscopes to ensure horizontal and vertical specimens. Each co-ordinator has recorded associated information for each image in a standard Word table according to protocol agreed at a previous meeting.

ACTION: Helen to distribute one test slide to all three diatom co-ordinators with 5 to 6 selected specimens for each to locate and capture. This requires Stephen Droop to run his inter-calibration software to provide a translation of co-ordinates for the three different microscopes. Helen to approach Stephen to request his assistance. Nigel and Christine to report back on whether rotation of microscopes is possible.

The need to always take brightfield images in addition to any taken with other contrast-enhancing techniques was reiterated.

The extent to which the images captured by the three co-ordinators should be incorporated into the final image database was discussed. Many may only be of sufficient quality as working copies and for further harmonisation workshop discussions. Some, however, may be suitable for inclusion in the final product. The need for this also depends on how many images can be re-captured using the high resolution system at RBG, Edinburgh.

ACTION: Steve Juggins to contact Stephen Droop at RBG to discuss the above.

v) Report documenting taxonomic decisions

Status: In progress. The three co-ordinators have made some progress with documenting taxonomic decisions but following a pilot study on a selected group of problem taxa by each co-ordinator earlier in the year, a way forward has yet to be agreed. It was agreed, however, that detailed taxonomic descriptions would only be required for problematic taxa. Non-problematic taxa should be described simply with a reference to a published description.

ACTION: Steve Juggins to compile a template for taxonomic documentation so that all three co-ordinators record the information in a standard way. This will not be developed, however, until further progress has been made with harmonisation of the TP and pH training sets. See Item 5 on Work Package 2 below.

vi) Status database of confidence in taxonomic harmonisation

Status: No progress. This now becomes a priority in Work Package 2 as part of the bigger taxonomic harmonisation exercise.

vii) Annual report

Status: Completed and submitted to EU in April 1999.

4. Report on Work Package 3 Progress " Data analysis and evaluation of methods" (months 1-18)

Steve Juggins circulated a summary report on progress with methodology evaluations, software development, guidelines etc. The aim is to hold a 2nd workshop in Jan 2000 to discuss results to date.

5. How to proceed with Work Package 2 " Data integration and harmonisation of all training sets" (months 18-30)?

The strategy for the way forward was discussed and the priorities and deliverables for the next 6 month phase of EDDI were agreed as follows:

i) The TP and pH groups are ready to go ahead with harmonisation within their training sets. This will be the priority over the next month so that a system for recording the taxonomic harmonisations can be developed. This will be an iterative process, with problems and types of amalgamations being identified as the work takes place.

ACTION: Dave (TP) and Nigel (pH) to begin harmonisation with a pilot study of one or two groups of taxa (e.g. *Stephanodiscus* or *Brachysira*). Preliminary results to be forwarded to Helen for circulation to the EDDI steering group and all three diatom co-ordinators. Deadline: 22nd December 1999.

ii) Following email feedback and discussions in January 2000, all the types of problem areas and types of merges and amalgamations will be identified. Steve will then develop and circulate a prototype template for recording the taxonomic merges, which will be trialled by all three groups. Feedback will continue until a final version of the template is agreed by early February 2000.

iii) Documentation of taxonomic decisions will be ongoing with the emphasis on the problem taxa.

iv) Image capture will also be ongoing with emphasis on problem groups, followed by common taxa.

v) Christine (Salinity) is not quite ready for harmonisation yet so her priorities will be to continue to concentrate on the problem taxa, documenting the taxonomic decisions in some detail where necessary. Christine to ask Phil Barker for assistance with this. Secondly, Christine will scan the Spanish (and Turkish) slides from Jane Reed and will identify any new taxonomic problems. As for Dave and Nigel, image capture should be ongoing with emphasis on problem groups followed by common taxa.

vi) Deadline for harmonisation of the individual pH, TP and salinity training sets (i.e. completion of WP1) is end of February 2000, prior to the Work Package 2 workshop (see Item 6 below).

6. Future Workshops
Under Work Package 2, a workshop was proposed for the week 28 February to 3 March, involving Dave, Nigel, Christine and Helen. It was agreed that Stephen Droop should be invited along to the workshop for advice on image capture etc and to discuss a strategy for the way forward.
ACTION: Helen to invite Stephen and agree a date will all workshop participants.

A second Work Package 2 workshop was proposed for July/August 2000. This should bring about the completion of Work Package 2 and facilitate the merging of all three training sets into one.

7. High Quality Image Database
See notes above on item 3 iv) -Dataset of medium quality images with associated information.

8. Preparation of duplicate slides for EDDI archive
Helen reported on availability of material so far. We are making good progress but we are still waiting to hear from some diatomists.
ACTION: All diatomists to look at the attached Excel spreadsheet which details those datasets where information on availability of material for slide preparation is still uncertain. Those concerned to contact Helen as soon as possible with details please.

9. Web page development
Nothing new to report.
The EDDI web site, developed by Cathy Stickley at ECRC, UCL can be found at *http://www.geog.ucl.ac.uk/ecrc/eddi/*
Steve Juggins also has an EDDI section on his home page, incorporating the EDDI Sample Finder facility. This can be found at *http://www.staff.ncl.ac.uk/stephen.juggins/eddi.htm*

10. Abstracts/Publications
The idea of submitting a short paper to Journal of Paleolimnology to announce EDDI has been dropped because the project is now too far advanced.

Helen is keeping note of all EDDI related presentations/abstracts etc presented at conferences. Please continue to pass on any relevant information.

11. Posters
There are currently two A0-size copies of the latest version of the EDDI poster, one with Dave at GEUS and one with Helen at ECRC, UCL. Steve Juggins will put a downloadable version of the poster on the Web site so that EDDI participants can print their own copies if they wish. Details to follow shortly.
END

3. Third Steering Group Meeting: 19th March 2001, UCL.
Present: Rick, Helen, Steve, John, Nigel, Francoise   Apologies: Dave, Christine.

1. Database of sites and chemistry data - Helen reported on the outstanding gaps:
* *Sitecode* - **complete** for all.
* *Lake names* - **complete** for all. (Note that KOLA and FINLAND do not have lake names.)
* *Lats/longs* - **complete**.
* *Catchment data* – complete except for DENMARK in NWEURO. (Note that data are not available for TURKEY and CASPIAN SEA.)
* *Chemistry data* – **complete**.
* *Units of measurement* -missing for SWAP, JCEN, APEN, and ALEN.
* *Chemistry dates* - missing for SWAP, JCEN, APEN, ALEN, CASPIAN SEA and all Danish and Irish lakes in NWEURO.
* *Diatom dates* - missing for SWAP, JCEN, APEN, ALEN, SWEDEN, CASPIAN SEA and all Danish and Irish lakes in NWEURO.

ACTION:
* **John to provide missing NWEURO data identified above (ie. catchment data from Denmark , chemistry and diatom sampling dates from Denmark and Ireland).**
* **Nigel to chase up missing ALPE related data identified above (ie .units of measurement for chemistry, plus chemistry and diatom sampling dates).**
* **Steve to add Caspian Sea data and provide SWAP units of measurement, and SWAP chemistry and sampling dates.**
* **Helen has contacted Tom K re. Swedish diatom sampling dates but needs to chase again.**

2. Database of diatom counts data
Complete except for the following outstanding issues
* ALPE SPAIN raw diatom data in the database are still percentages.
* **ACTION: Nigel to get the raw counts from Sergi or at least the total count & back calculate.**

* Uncertainty over whether to include Jane's Turkish samples given the preliminary nature of the counts and lack of original slides to enable complete harmonisation.
* **ACTION: Steve to speak to Jane to discuss. Then Helen will contact Christine to enquire about the degree to which the Turkish samples were harmonised.**

3. Taxonomic harmonisation WP1 and WP2
* All agreed that revision of the taxonomy confidence coding system should be undertaken according to the new scheme outlined in Steve's document. A zero ("0") category should

be added to include taxa where the "original name is valid and no code change is required". Where a code is not assigned this indicates that a taxon has not been examined in EDDI. This will allow summary statistics to be produced for the final report of how many taxa have been merged at the different levels.

- **ACTION: Francoise, Nigel & Dave to recode using the new system once we have approval from Dave in early April. See Timetable below.**

- All agreed that there was still potential to further "tidy" the merge tables using a harsher merge strategy as suggested in Steve's discussion document. Dave has already addressed this and has placed his new tables on the Newcastle Geus directory site.
- **ACTION: Francoise and Nigel to refine their merge tables similarly.**

4. Taxonomic descriptions

- It was agreed that Francoise, Nigel & Dave will provide a short text description (a la SWAP Red Book) for all taxa which are coded with a 3, 4 or 5 status to explain what has been merged and why. For all other taxa, a KLB reference (or relevant other publication) should be given. A bibliography should be compiled of all floras/papers referred to in the text. It was agreed that a list of any other taxa that a particular EDDI taxon can be confused with should be provided in each case. There is no need for Francoise, Nigel and Dave to provide the authorities because Steve can link these automatically.
- **ACTION: Francoise, Nigel & Dave to provide the above for all important taxa in their training sets. The easiest way to do this would be to add a text field to the original species list (in Excel). This can then be imported into Word for formatting (e.g. italics for species names – please ensure that formatting is complete and that files are spell checked and error free to reduce Steve's work load!). A separate bibliography should be compiled by Francoise, Nige & Dave in a Word file (this should not be a lot of work given that most taxonomy follows KLB). See Timetable below.**

5. Images

- Helen has contacted Micha and he has agreed to provide some text on imaging and capture methods for the final report/web pages.

- Helen has a list of all images recaptured at RBGE (for UCL and CEREGE samples only) from Shirin.
- **ACTION: Steve will check this list against the list of working images to ascertain how many of the working images have not been recaptured. The quality of those working images which have not been recaptured will then we assessed and a decision will be made on which should be included in the EDDI image database.**

- Shirin requires a further 2 weeks salary in order to finish capturing all of the images requested by Dave in his final batch of slides. She cannot start this until early May. UCL have agreed to find £500 from the EDDI budget but a further £400 is required to cover the total RBGE costs.
- **ACTION: John to look into the possibility of funding this from the GEUS EDDI budget as a matter of urgency.**

- It was agreed that the final image database should include the initials of the taxonomist for each image with links to descriptions of original datasets, sample details etc.
- **ACTION: Steve will produce a prototype of the image database and will redesign according to feedback from the EDDI taxonomists.**

6. Transfer functions

- It was agreed that transfer functions generated from individual WP1 datasets, merged datasets, dynamic datasets based on analogue matching techniques to produce local training sets, and user-defined training sets (where certain criteria could be specified to select samples) would all be useful.
- Transfer functions should be developed for TP, pH and salinity, and for new variables such as Al, DOC, and ionic ratios. These are all listed in the deliverables! A temperature transfer function cannot be developed from the EDDI datasets.
- Steve will produce the text on evaluation of methods and guidelines for best practice (WP3 report) for the final report/web pages.

7. Access to raw diatom and environmental data.

- It was agreed that a (2 year?) moratorium should be placed on the EDDI database to restrict use to EDDI-contributors only. We need the views of all EDDI contributors on this and ultimately a memo of agreement will need to be written.
- **ACTION: Helen to produce a matrix of all EDDI datasets and their respective levels of access/restriction based on feedback from all EDDI data contributors.**

8. The EDDI system

- Translation dictionaries – it was agreed that a user ID (password protected) would be useful so that users could save and re-use translation dictionaries (i.e. a shopping basket).

- **ACTION: Steve agreed to produce a draft outline of the structure of the web pages and will indicate where text contributions are needed by others.**

- According to the proposal, we are contractually obliged to produce both a Web and CD-ROM version of the final system. As these require different software, it will be a significant task to produce a CD-ROM version of the complete EDDI system in the timeframe of the project. However, it was agreed that we cannot ask Brussels for further extensions to the project.
- **ACTION: Steve to further consider the implications of producing both formats.**

9. Production of text/Report writing

The following will be required for the Final Report and can be edited accordingly for the Web pages. Helen and Rick will circulate a draft outline of the report structure/content for comments by the steering group.

- Overall project description: this can be largely cut and pasted from the original proposal and the Vienna paper.
- **ACTION: Rick/Helen to write.**

- Text to explain the content and nature of the catchment and chemistry data (i.e what is missing and why) with details of quality and quantity.
- **ACTION: Rick/Helen to write.**

- Text to explain the content and nature of the diatom data (i.e surface sediments or plankton; counts/ % data etc)
- **ACTION: Rick/Helen to write.**

- Text on the taxonomic systems/nomenclature used and the merging strategy (its purpose, functions etc).
- **ACTION: Rick/Helen to write with contributions from others.**

53

- Text on image capture methods/approaches
- **ACTION: Micha to write.**

- Text on evaluation of numerical methods and guidelines for best practice in transfer function development
- **ACTION: Steve to write.**

- Users Guide production: this cannot be written until the basic system is up and running.
- **ACTION: Review again once the system has been developed and tested.**

- Bibliography of taxonomic floras and papers
- **ACTION: Francoise, Nigel and Dave to produce list of references whilst writing their taxonomic descriptions (as detailed in 4. above).**

- Bibliography of datasets in EDDI:
- **ACTION: Helen will contact all contributors for lists of references to their datasets.**

- List of contributors with contact details:
- **ACTION: Helen to ask all contributors to provide their up to date details.**

- Bibliography of transfer function papers (methods and applications). It was agreed that John Birks' review paper in JOPL would be a useful starting point as well as John's forthcoming methods for climate reconstruction paper.
- **ACTION: Helen to contact John.**

- Links to other relevant web-based resources such as partners' homepages, other diatom sites.
- **ACTION: Helen to contact all EDDI partners for suggestions and details of web sites.**

10. Timetable and deliverables
- A date for the Final EDDI workshop was agreed. It will be held **all day of Monday 23rd and the morning of 24th July 2001 at UCL.** The aim of this meeting will be for all EDDI partners to feedback with their views on the EDDI web pages and the draft Final Report, as well as to discuss issues of data access, paper production etc. It is, therefore, vital that the pages are up and running and that the draft report is circulated by early July at the latest. A key deliverable of the project is a series of papers and therefore it was suggested that the workshop could be extended to include a paper writing session.
- **ACTION: Helen to announce the date to all EDDI partners by email and to request thoughts on the format/scope of the workshop and to explore whether there is any enthusiasm for the paper writing session.**

- A final steering group meeting will be held on the **afternoon of Tuesday 24th July 2001** at UCL (after the workshop) to discuss any outstanding issues.

- In order to evaluate the EDDI system (a key deliverable), a series of test sets and cores will be needed.
- **ACTION: EDDI steering group members to start thinking about potential suitable test sets and 20 or so suitable cores.**

• REVISED TIMETABLE
It is critical that we adhere to these strict deadlines

IMMEDIATELY: Rick to contact Hans Brelen re timetable for reporting.

IMMEDIATELY: Helen to send out an email to all EDDI partners to announce the dates of the Final workshop and to request all the information listed above.

11th APRIL: Deadline for revised final WP1 and WP2 merge tables using the new scoring scheme and the harsher merge strategy. Francoise, Nigel & Dave to send new tables to Steve following approval of this strategy by Dave (who returns on April 3rd).

11th APRIL: Deadline for receipt of any outstanding environmental data (see item 1 above).

1st MAY: Deadline for test datasets and core data to be sent to Steve for model evaluation.

END OF MAY: Deadline for completion of taxonomic descriptions and bibliographies.
(NB John has offered expertise with Bibliographic downloads from the Web into Endnote).

END OF MAY : Deadline for completion of the RBGE final image list.

END OF MAY: Deadline for completion of WP3 report on method evaluations (following May meeting organised by Steve), and collation of transfer function literature (Steve).

END OF JUNE: Deadline for all Final Report/Web page text listed in item 9 above, except for the User Guide.

EARLY JULY: Deadline for EDDI web pages to go on line and for draft final report to be circulated to all EDDI partners. Feedback should follow.

23/24 JULY: Final Workshop

END OF JULY: Financial deadline

JULY/AUGUST: Production of all other non-Web based material such as additional Final Report data/text, mapping, download facilities.

JULY/AUGUST: Production of User Guide

JULY/AUGUST: Redesign of EDDI Web pages based on feedback from Workshop

END OF AUGUST: Completion of Final Draft Report

END OF SEPT: Completion of Final Report.

11. Financial and contractual matters

- There was some uncertainty over whether the standard annual report would need to be produced in March (as in previous years) and when cost statements would need to be submitted given that the project has been extended to end of July. We are assuming that the deadline for completion of the Final Report will be end of September 2001.

- **ACTION: Rick to contact Hans Brelen to clarify our obligations and timing of reports/deliverables given the new timescale.**

- Each partner needs to ensure that they have sufficient funds to cover the travel and subsistence costs of the following participants at the Final Workshop:

(i) ECRC-UCL

ECRC-UCL will pay T&S for the people detailed below to attend the following number of workshops.

Name	Year 1	Year 2	Year 3	Total workshops
Helen Bennion	1	2	3	6
Nigel Cameron	1	2	3	6
Paula Maliphant/Dave?	1	1	1	3
Don Charles	1	1	2	4
Patrick Rioual	1	1	1	3
Rick Battarbee	1	2	3	6
Jorunn Larssen	1	1	1	3
Ingemar Renberg	1	1	1	3
Atte Korhola	1	1	1	3
Eileen Cox	1	1	2	4
John Birks	1	1	2	4
Jan Weckstrom	1	1	1	3
Roland Schmidt	1	1	1	3
Sybille Wunsam	1	1	1	3

It was further agreed at the May 1998 meeting that UCL would also meet the costs of Jane Reed and Philip Barker to attend workshops, to make payment into their accounts more simple. UCL will invoice CEREGE for this amount at the end of each workshop.

(ii) CEREGE

All workshop costs for Francoise Gasse, Francoise Chalie, Gerard Begni + un-named will be met by CEREGE.

(iii) GEUS

GEUS will contribute to the costs of Andy Lotter for attending any workshops, as well as Dave Ryves or John Anderson.

Name	Year 1	Year 2	Year 3	Total workshops
John Anderson &/or Dave Ryves	2	2	2	6
Andy Lotter	1	1	1	3

(iv) Newcastle

Workshop costs for the following will be met by Newcastle.

Name	Year 1	Year 2	Year 3	Total workshops
Steve Juggins	2	2	2	6
David Mann &/or		1	1	2

Stephen Droop (Micha)				
Cajo ter Braak		1	1	2
un-named		1	1	2

END

Minutes of the EDDI Taxonomic Workshops

Work package 1
pH datasets: 14/15 January 1999, UCL.
Present: Nigel, Helen, Peter, Jan, Nadia, Viv, Roger.   Apologies: Jorunn.

Helen gave an update on the EDDI datasets received to date.
- Peter thinks that Steve has already received Tom's Swedish dataset - lat, long and altitude data. *ACTION: Helen to contact Steve to confirm.*
- Both Peter & Jan think that the environmental data files sent to Steve contain only a single value for each parameter (based on a sample size of one rather than a mean of several samples).
ACTION: Helen to confirm with Steve.

The issue of the possible use of **DIATCODE in EDDI** was raised by Helen. All were in agreement that this seemed a sensible way forward. There were no objections.

The following section documents the decisions made and any discussion points raised during the taxonomy workshop. The workshop was structured as follows:

1. Work through Nigel's hit-list of the top 100 taxa and their possible associations and confusions. Note obvious synonyms, miscodings etc.
2. Use Steve's web page sample finder to locate the slides that best represent these taxa and check the slides under the microscope to confirm whether all diatomists are in agreement.
3. Capture 'working' images as we go along of any taxa that are not straightforward ie. where there is a range of specimens to illustrate a species concept or where there is still some confusion and need for further work.
4. Save the images as JPG files in a directory, recording only the species code and sample code within the file name at this stage (eg. BR001ASwapLGR for Brachysira vitrea in Loch Grannoch, a Swap site). No associated information is recorded here.
5. Nigel plans to find good examples for each of the common taxa over the next few months. Any further taxonomic uncertainties will be discussed with the other diatomists and working images will be circulated via the internet or email before any final decisions are reached. Once agreed, images of the taxa will be captured by Nigel and the stage co-ordinates and any associated information will be recorded in an ACCESS database (to be generated by Steve following discussions at the steering group meeting?).

GROUP 1: BRACHYSIRA
1. Jan's Brachysira sp.1 (RLGH) is amalgamated into B.brebisonii (BR006A), but Tom and Swap split them into 2 separate taxa.

GROUP 2: TABELLARIA

57

1. Jan uses his own taxon T. flocculosa (Nord-Chill) which includes T. flocc v flocc (the short form used in SWAP) and T. flocc agg (as used in SWAP). However most of the taxa are T. flocc v flocc. Both Tom and Nigel follow the SWAP splits (see SWAP Red Book p. 77).
2. Need to check Viv's use of T.flocc var IV in Svalbard dataset.
3.T.binalis (TA003A) includes both the elliptic and panduric forms. Nigel and Peter do not split into varieties. Jan has just one sample where he splits into T.binalis var elliptica.
4. No image taken of T.quadrisceptata as well described (note high %s in Swap & Bergen).
5. T. fenestrata not discussed because overall M =1%

GROUP 3: FRAGILARIA
1. SF001A>FR005D = miscoded in Jan's dataset
2. FF001A>FR005A = miscoded in Jan's dataset

GROUP 4: FRUSTULIA
1. A range of images taken of F. rhomboides varieties - all agreed on concepts.
2. No images of F. vulgaris - non-problematic and less common.

GROUP 5: ACHNANTHERS MINUTISSIMA
1. A.minutissima (Nord-Chill) includes all minutissima vars, but pusilla/linearis are not included.
2. A (minutissima agg) used by Tom includes all minutissima vars, and equates to the A.minutissima (Nord-Chill).
3. A minutissima var minutissima used in Swap also includes several vars. Esssentially the same as 1 and 2 above.

GROUP 6: FRAGILARIA CONSTRUENS
1. PS002A>FR056A = miscoded in Jan's dataset
2. SR001A>FR002A = miscoded in Jan's dataset
3. Difficult to consistently split F. con v venter from fine forms of F.pinnata - some overlap agreed. Need to capture range of images and issue warning of confusion.
4. Difficult to consistently split F. con v venter from F.elliptica - some overlap agreed. Also Tom & Peter have different concepts than Nadia. Need to capture range of images and issue warning of confusion.
5. F. pseudoconstruens consistently split be all.

GROUP 7: ACHNANTHES MARGINULATA/SCOTICA
1. Nord-Chill (ie. Finland dataset) has a 11 μm cut off for spliting thse 2 taxa, where any > 11 is marginulata.

GROUP 8: PINULARIA BICEPS
1. No problems.

GROUP 9: FRAGLARIA PINNATA
1. SS001A>FR001A= miscoded in Jan's dataset
2. Only Viv has used varieties so Nigel to check with Viv.

GROUP 10: CYCLOTELLA
1.Tom uses C (kuetz agg) to include schumanii, krammeri and all rossii types. However, Nord-Chill splits C.rossii into C.rossii (which are with random punctae), rossii type 2 (with clear tri-feature) and type 3 (with larger processes forming tri feature). Although split by Jan, they can be amalgamated as C. rossii (would probably be best).
2. The Bergen dataset uses code CY006A rather than CY9991 (C. kuetz agg) so need to check if these codes describe the same range of taxa. Nigel to ask Jorunn.
3. Both the C.comensis group and the C.pseudostelligera/glomerata group proved difficult and no decisions made. Nigel to further investigate these groups.

GROUP 11: AULACOSEIRA

1. Tom's A.distans/subarctica and the A.subarctica type 2 (Nord-Chill) are the same taxon. NB type 2 descibes the shorter, wider form of this species.
2. Tom's A.subarctica v sub-borealis could be A. subarctica type 1 (the narrower, longer form). Need to clarify.
3. Nigel was unsure of what is included in the Swap code Aul. (subarctica agg.).
4. Nigel to work through the names used in Swap and Alpe.

GROUP 12: EUNOTIA

1. The split between incisa and rhomboides presented no problems.

GROUP 13: FRAGILARIA BREVISTRIATA

1. PS001A>FR006A= miscoded in Jan's dataset
2. The code FR9998 is probably a data entry error? Check.

GROUP 14: ACHNANTHES AUSTRIACA

1. The Alpe and Swap taxon AC9965 is very distinct and was not a problem
2. AC014B used in Swap, Alpe and Svalbard to separate the small form from the nominate AC004A.
3. NB there are none in Sweden! Check with Tom.

GROUP 15: NAVICULA LEPTOSTRIATA

1. No problems

GROUP 16: ACHNANTHES CURTISSIMA/SACCULA

1. Alpe doesn't use A.saccula. Only used by Tom and Jan.
2. Tom & Jan do not split levanderi and lacus-vulcani. They use A.levanderi only. However, Viv & Nadia do use both names in their Svalbard & Kola sites.
3. Swap has a taxon called A.(cf levanderi) (which may be synonymous with A. curtissima used in Nord-Chill & Alpe). This is used to separate from the larger nominate A. levanderi.

GROUP 17: CYMBELLA PERPUSILLA

1. The NordChill taxon C.amphicephala is a distinct taxon and is not syn. with C. gaeumanii.

GROUP 18: EUNOTIA

1. Both the Swap & Bergen sets split into var E. exigua v tridentula.
2. Only Bergen, Swap & Jan split into var undulata. All others amalgamate into E.exigua nominate.
3. There was no confusion over E.tenella and E.paludosa. These were split consistently by all. However, a taxon called E.tenella/paludosa is used in Alpe & Swap. Tom also recognises this taxon but there are none present in the Swedish dataset.
4. E. bilunaris and E.lunaris are synonomous with E.curvata. E. curvata is the currently accepted name. Also, E. curvata v. subarcuata is synonym to E. bilunaris v. mucophila. Jan has separated E. curvata from E. curvata v.subarcuata although this is difficult and maybe they should be amalgamated?
5. The Eunotia group needs more work but note that many are infrequent. One exception is E.(sp.10 minima) present in 44 samples, and used in Swap, Alpe & by Tom in Sweden.
6. F_EU058B>EU008D= miscoded in Jan's dataset
7. Jan uses E.pectinalis v minor fo impressa to include E. implicata and E.pectinalis v impressa.

GROUP 19: ACHNANTHES PUSILLA/LINEARIS

1. Consistently split by all.

GROUP 20: NAVICULA DIGITULUS
1. All agreed on this taxon. Only occurs in the northern and mountain sites so maybe a good climate indicator?
2. Nigel to check with Roger re N.(cf digitulus).

GROUP 21: NAVICULA HOEFLERI
1. All agreed that N. hoefleri sensu Ross et.Simms was synonomous with N.simsii.
2. N. cumbriensis (see Haworth working paper) - narrow, linear, wide apices, no change in striae density and no central area.
3. N. madumensis - no central area, no change in striae density but bowed sides.
4. Some confusion over what N. subtilissima var 1 (NJA & VJJ?) might be. Nigel to ask Viv.
5. Jan puts subtilissima variety (bow-tie type) into the nominate. However, Peter and Nigel and Swap split them but the exact name given was uncertain. Needs clarifying.
6. See Hustedt (1930-1966 III p89) for N.subtilissima, but not Lange-Bertalot.

GROUP 22: PERONIA FIBULA
1. No problems.

GROUP 23: PINNULARIA RUPESTRIS
1. No problems.
2. Nigel to check Roger's use of code PI9978.

GROUP 24: NAVICULA MEDIOCRIS
1. No problems.
2. var atomus has more tapered ends - all agreed.

GROUP 25: ASTERIONELLA
1. A.ralfsii var americana used only by Tom- has a finer, smaller and narrower foot pole than A.formosa. The valve is longer than A.ralfsii v ralfsii. Can have a flat top bulb shape.

GROUP 26: SEMIORBIS HEMICYCLUS
1. No problems

GROUP 27: EUNOTIA NAEGELII
1. No problem. All follow Swap.

GROUP 28: NAVICULA PUPULA
1. NordChill uses their own code for N.pupula.
2. Tom uses UME425 to describe N.(absoluta/pupula) in the Swedish lakes. Nigel to check.

GROUP 29 CYMBELLA LUNATA/GRACILIS
1. These are syn and most use C.lunata. NB Alpe has used both and needs to be harmonised.

GROUP 30: PINNULARIA MICROSTAURON/CAUDATA/SUBCAPITATA
1. NordChill lumps all vars of P.microstauron into the nominate.
2. The NordChill and NGC P.caudata codes could be merged as they are the same taxon.
3. P.subcapitata - elongated ends, often narrower than main valve.
4. Nigel will work on the P.subcapitata/braunii group

GROUP 31: NEIDIUM
1. N.iris v ampliatum is syn with N.impliatum.
2. N alpinum is only present in Alpe, Kola and one Swedish sample.
3. Peter notes that N.hercynicum can be confused with N alpinum. Needs checking.

GROUP 32: ACHNANTHES DETHA/SUBATOMOIDES

1. It was agreed in Alpe & Nordchill that A. detha would be merged into A.subatomoides
2. But Tom and Nadia use A. detha. These are therefore syn here.
3. Recode F_AC042A> AC136A.
4. Need to check Tom's taxon UME426 A. (detha,minor).

GROUP 33: CYMBELLA AEQUALIS
1. Need to look at C.subaequalis and C. incerta as these could be confused.

GROUP 34: NITZSCHIA FONTICOLA
1. Ok except that NI022A is missing from the Swedish dataset. Check with Tom if this is true. Synonym possibilities?

GROUP 35: NAVICULA RADIOSA
1.Tom appears to use N.radiosa v tenella, whereas in NordChill, N.cryptocephala is more frequent. Check if there is any confusion here.
2. Note that Tom's radiosa v tenella is quite long and narrow (not the short, fatter type).
3. In Swap, N. cari is syn with N radiosa v tenella.

GROUP 36: STAURONEIS
1. Need to check the use of S. anceps v hyalina in Alpe.
2. Need to check S. gracillima used in Swap and in one Finnish sample. This has been renamed Nupella teucephala.

GROUP 37: ACHNANTHES ALTAICA
1. A.altaica v minor is used by some and not others. It is mostly present in Swap, especially Loch Grannoch. It has less reflexed ends, more rounded outline and is smaller (normally <10um) than the nominate.
2. Jan warns that he may have called A.austrica v minor, A. marginulata in his Finnish dataset. Need to check.

GROUP 38: EUNOTIA VANHEURKII
1. No problems

GROUP 39: CYMBELLA HEBRIDICA
1. EY003A>CM017A= miscoded in Jan's dataset

GROUP 40: CYMBELLA MICROCEPHALA
1. No problems

GROUP 41: CYMBELLA MINUTA/SILESIACA
1. C. minuta can be separated from C. silesiaca as it is less broad, and the punctae are not visible on the striae.
2. Need to check whether Viv's taxon C. ventricosa in Svalbard is syn with C.minuta or C.silesiaca.
3. Note that C. silesiaca is syn also wiith C. minuta v silesiaca.
4. EY011A>CM031A= miscoded in Jan's dataset

GROUP 42: SURIRELLA
1. Need to check AM's use of A_ZZZ956 in Alpe. Could be syn with S.linearis?
2. Nigel to look at the common ones, especially S.delicatissima, biseriata and linearis (& v constricta).

GROUP 43: AMPHORA LIBYCA
1. Tom has no AM011A A. libyca is the Swedish dataset. Is this correct or does he use another code/name? Need to check.

GROUP 44: FRAGILARIA (cf OLDENBURGIANA)
1. There are 2 types in the datasets: FR9990 used only in Bergen and by Tom in Sweden, which is a very small form; and FR9991 (see PIRLA plates). They look slightly different.
2. Jan uses FR013A F.oldenburgiana v oldenburgiana too.

GROUP 45: FRAGILARIA CROTONENSIS
1. mostly present in the Bergen dataset.

GROUP 46: ACHNANTHES NODOSA
1. This is a clear taxon - bolder striae than A.pusilla and has non-parallel sides.

GROUP 47: NAVICULA SCHASSMANII
1. No problems.

Miscellaneous NAVICULA comments from Roger:
1. Navicula (cf.schadei) was used is Swap as a syn for N. carissima. Note that N.carissima was not described in Swap.
2. Roger can provide Swap concept descriptions of N.submuralis, N.muralis, N.minima and a few other taxa from old notes.
END

Salinity datasets: 20-21 January 1999, CNRS – CEREGE.
Participants: Jane Reed, Laurence Carvalho, Helen Bennion, Phil Barker, Francoise Gasse, Francoise Chalie, Christine Pailles

FRAGILARIA NANANA /CAPUCINA and var. / ACUS var. radians / ACUS var. angustissima GROUP

Fragilaria nanana (TU_0013):
- KLB 2/3, p 130, pl: 114 fig: 9-11, pl: 115 fig: 14-16
- L=40-90μm, w=1.5-2μm, 22-25 st/10μm

Shorter than **S.acus**, extremely delicate, L=70μm, w=1.5 μm . Visible striae (> 32 st/10μm) with a good microscope. According to KLB , the distinction between **F.nanana** and **F.tenera** is not valid. The species JR called **F.nanana** in the Caspian dataset is in fact **F.tenera**.
- Sample **PB-MA 19b**, from Scar Lake middle Atlas mountains, 640μS/cm, pH=8.5, plankton sample.
- *Phil B. will send raw material or slides*

F.capucina var. gracilis (FR009H) :
- KLB 2/3, p 123, pl: 110 fig: 8-12, pl: 111 fig: 1-3, pl:113 fig:22-26
- up to 20 st/10μm

Specimens observed are L=35.5 μm, w=3 μm, 18-19 st/10μm, striae interrupted in the centre, very narrow pseudo-raphe, central area hyaline slightly inflated (very clear in connective view), extremities very slightly capitated. Looks like specimens in KLB2/3, p123, pl:110 fig:9-11.
- Sample SJ-9P (37) plankton from lower Volga, falsely called *F.* **tenera** in this sample
- *Sample to be given by SJ*

F.capucina var. rumpens (FR009G) :
- KLB 2/3, p 122, pl: 108 fig: 16-21, pl: 110 fig: 1-6
- w=4μm, 18-20 st/10μm

62

Specimen are L=30-31μm, w=3-3.5 μm, 16-18 st / 10 μm, asymetrical axial area with strongly attenuated striae, inflated only on 1 side.
- Sample SJ-9P (37), plankton from lower Volga. The sample is also containing few S.acus
- *Sample to be given by SJ*

S.rumpens var. **neogena** (AF_5418)
- Huber-Pestalozzi 1942, p459, fig: 537
- L=27-70μm; w= 2-3μm, 19-20 st/10μm
S.rumpens var. **neogena** is probably equal to **F.tenera**
- Sample FG **Naivasha H1a**

S. acus var. **angustissima** (SY003C)
- KLB2/3, p144, pl114, fig:21, pl:122, fig:15-16
- L= 40-500μm, w=1-4μm, 12-18 st/10μm
S. acus var. **angustissima** is longer and more regular than **S.acus** var. **radians**
- Sample FG-**Naivasha H1a**

For the EDDI inter-group discussion, we leave open the question of distinction between **S.acus** var. **radians**, **S.rumpens** var. **neogena** and **F.tenera**. However, **S.acus** var. **radians** is longer and coarser.

AMPHORA GROUP

Amphora acutiuscula (AM002A):
- KLB 2/1, p 348, pl:151, fig: 6'
- L=13-60μm, w= 10-19μm, Mid Dorsal= 10-12 st/10μm
Specimen observed are L=26μm, w=10 μm, MidDorsal: 15 st / 10 μm.
Striae are clearly punctate
- Sample JR- **HPD1**(31%) from Spain
- Sample SJ-**12B** from lower Volga (22%),
- Sample SJ-**26B** (24%???) in this sample **A.acustiuscula** is wrongly identified as **A.coffeaeformis** and grouped with another Amphora (to be sorted out)
- *Samples HPD1, 12B & 26B to be given by JR & SJ*

Amphora coffeaeformis (AM006E):
- KLB 2/1, p 347, pl:151, fig: 1-6
- L=13-60μm, w= 10-19μm, Mid Dorsal= 16-24 st/10μm
Specimen observed are L=31-32 μm, w=10 μm, MD: 18-20 st / 10 μm. Compared with A. acutiuscula specimen are much narrower, the ventral side is slightly convex, striae are more delicate and striae punctuation is less visible.
- Sample JR-**HPD1**(12%) from Spain
- Sample SJ-**12B** from lower Volga (8%)
- *Samples HPD and 12B to be given by JR & SJ*

Amphora delicatissima (AM039A):
- KLB 2/1, p 351, pl:152, fig: 19-23
- L=5-20μm, w= 3-7μm, Mid Dorsal= up to 30 st/10μm
Specimen observed are L=20.5 μm, w= 4.2 μm; MD: 20-21 st / 10 μm. Striae punctate, roughly parallel, coarser in the middle than in the poles. Ventral margin linear, no visible striae in the ventral side, raphe very slightly arched toward the pole but not curved in the central area. Both valves are curved in the same direction.
- **Sample FG-Dji 87/8c**

63

- *According to L.C. the species called* **A.tenerrima** *in Dji 87/8c, Dji D121 & Dji D123 is in fact A.delicatissima because the striae are parallel.*

Amphora micrometra (AM122A) :
Specimen observed are L=17-24 µm, w=4.2 µm. Striae extremely difficult to observe, very delicate striae on the dorsal side. Ventral side linear, dorsal side strongly convex, extremities rostrate slighty directed toward the ventral side, raphe excentric. In connective view the ventral area appears rather hyaline & large.
- Sample JR-**Bol 96 /0-0.5 386** from Turkey (**A. micrometra** is associated **with N. salinicola**).

Amphora tenerrima (AM110A):
- Schoeman 1972, p241, fig:8-10
- L=9.5-17 µm, w=2.1-2.6 µm, MD=24st/10µm
Specimen observed are L=20 µm, w=4 -4.5 µm; MD: about 25 st./ 10 µm . Striae only on dorsal side, slightly punctate, parallel or slightly convergent compared to **A.delicatissima**.
- Sample FG-**Alg 85/01**

NAVICULA HALOPHILA GROUP

Craticula cuspidata (CI004A)& C.ambigua (AF_3616)
The two taxa are mixed in the NGP and Africa data sets. Check in new records if they could be separated or not. **C. ambigua** has been counted so far as **C. cuspidata**.
- KLB2/1, p126, pl:43, fig:1-8
- L=30-120µm, w=13-25µm, 11-19 st/10µm
C.ambigua (AF_3616):
- Sample P.B-**9918A admare**, there is another species present close to N.gregaria and it should be checked
- Sample FG-**H116** & **V24**
C.cuspidata (CI004A):
- Sample FG-**H134** (See LC for negatives # 34/9)

Navicula halophila (TU0007-NA9851-NA9850-NA9849-AF_3631-NA022A-NA022C)
- KLB2/1, p126, pl:44, fig:1-11 & 14-18
- L=7-140µm, w=4.5-18µm, 15-24 st/10µm
J.R distinguished **N.halophila** as having striae strongly convergent at the valve ends and agrees with F.G. on african material.
Sample SJ-**26B** was supposed to have 32% of N.halophila but we saw 1 valve and the rest was a small very fine **N.accomoda**.
Sample SJ-**16B** had some N.halophila but it is in fact a large variety of Navicula (WHAT???)
- Sample F.G.**Tun.EH93**
- Samples J.R. **EALM1 & ESLC1**.
- *2 Samples to be given by JR*

Navicula subhalophila (AF_3745)
- Hustedt 37-39, suppl5, p229, pl:17, fig:1
- L=28µm, w=7µm, 30-32 st/10µm
Specimen subcapitate, striae delicately punctate, more numerous and parallel than in **N.halophila**. Very narrow and linear central area.
- Sample from FG- **Kenya-H98-99**

Navicula salinicola (NA614 A)
- KLB2/1, p:111, pl:35, fig:9,10
- L=7-17µm, w=2-3 µm, 17-20 st/10µm

the number of striae is lower than figures given by LB and much lower than in population from Africa & Spain. So it doesn't seem to be N.salinicola. More rounded morphology.
- Sample SJ-**14B**
- *Sample to be given by JR*

Navicula sp2 af salinicola (NA9838)
Striae slightly radial but more similar to **N.salinicola** than Caspian examples
- Sample JR-**EMNJ1** from Spain
- *Sample to be given by JR*

N.gregaria (NA023A)
- KLB2/1, p:116, pl:38, fig:10-15
- L=13-42µm, w=5-10 µm, 13-22 st/10µm

N.gregaria is clear

NITZSCHIA FONTICOLA vs LACUUM vs FRUSTULUM vs BACILLUM

GENERAL COMMENTS - Dave Mann 's thesis points out that:
- *width is a good character (conservative parameter) whereas length and ends vary within a species. e.g. when larger the neds seems to be more protracted, when shorter the ends tends to be rounder.*
- *striae density is also a useful characteristic*
- *fibulae density can be widely ranging but may be useful when used with striae density ratio.*

Nitzschia fonticola (NI002A)
- KLB2/2, p:103, pl:75, fig:1-22
- L=10-65µm, w=2.5-5 µm, 9-16 fib/10µm, 23-33 st/10µm

Specimen observed have 28 str / 10 µm, clear central nodule, parallel sides but a bit constricted in the central nodule area on the ventral side only.

N.bacillum (no code nb) vs **N.lacuum** (AF_3946)

N.bacillum (no code nb)
- KLB2/2, p:108, pl:78, fig:7-12
- L=12-20µm, w=2-3.5 µm, 12-16 fib/10µm, 27-32 st/10µm

N.lacuum (AF_3946)
- KLB2/2, p:107, pl:78, fig:1-6
- L=10-20µm, w=2-3 µm, 13-18 fib/10µm, 35-40 st/10µm

From our observations, **N.lacuum** is always capitated or sharp ended, sometimes slightly strangled/ constricted in the middle, 27-40 striae / 10 µm. See notes from CASPIA workshop Orsay 92.
In KLB **N.bacillum** & **N.lacuum** are differentiated by the number of striae. But they both can have varying morphology from a lemon shape to parallel in the middle.
We note that the specimens of KLB2/2, Pl 78 fig:1&2 are much coarser than the one described by KLB for **N.lacuum** (>35 st). The type selected by KLB for **N.lacuum** from the plankton of Lake Edward can be also found by F.G.86 as **N. af.fonticola type 1**. The number of striae in **N.lacuum** from Lake Edward is of 34 st showing that there is really a continuum in the number of striae between **N.bacillum** & **N.lacuum.**
- Sample FG-**JT 25** for **N.lacuum**

FG's **N. af.fonticola type 1** (no code nb) is a typical planktonic form. Typical shape of a lemon , not fully lanceolate, clearly subcapitated at ends. No central nodule, striae parallel at the center and finely punctate. **N. af.fonticola type 1** is not **N.fonticola** . We decide to call it **N.lacuum**.

What is called **N. sp. af. fonticola type 2** by F.G.86 has the same shape variations from lanceolate to lemon shape than **type 1** and is similar in the nb of striae . However, it differs by coarser fibulae sometimes slightly elongated .We decide to call it **N. lacuum var.3**.
* Sample from F.G.-**V21**, **N. lacuum var.3**

N. lacuum var.2 is as **N.lacuum** but the ends are rounded rather than capitated (See notes from CASPIA workshop Orsay 92). The shape varies somewhat from linear-lanceolate (Australia, East Africa) to distinctly lanceolate (China, East Africa).

At the moment we propose to group **N.bacillum & N.lacuum** and we decide to call it **N.lacuum**. The shape varying from typical lanceolate to lemon shape with somewhat subcapitate ends, sometimes constricted in the middle. The nb of striae can vary from 27 to 40 st/10μm. But attention should be paid in future work and other samples to confirm the validity of grouping.

Nitzschia frustulum (NI008A)
* KLB2/2, p:94, pl:68, fig:1-19
* L=5-60μm, w=2-4.5 μm, 11-16 fib/10μm, 19-30 st/10μm
It has a clear central nodule and we adopt the KLB definition but the central nodule is not always very clear. The ends of N.frustulum are bluntly rounded.
* Sample FG-**Kenya H7**

N.liebetruthii (NI203A)
* KLB2/2, p:96, pl:69, fig:14-32
* L=5-40μm, w=2-4.5 μm, 11-16 fib/10μm, up to 20 st/10μm
Acute non-attenuated apices, lanceolate form. **N. liebetruthii** is basically **N.frustulum** but without central nodule
* Sample SJ-**15P** from lower Volga ???????
* *Sample to be given by SJ.*

NITZSCHIA INCONSPICUA GROUP

Nitzschia inconspicua (AF_4098 & NI043A)
* KLB2/2, p:95, pl:69, fig:1-13
* L=3-22μm, w=2.5-3.5 μm, 8-13 fib/10μm, 23-32 st/10μm
According to KLB, **N.inconspicua** has rather rounded ends but several specimen have sharp ends, relatively broad with a high width/ lenght ratio.
N.inconspicua is called **N. sp af. frustulum type.3** in F.G.86.
N.inconspicua was called **N.frustulum var. pusilla**. BY WHO???????
* Sample JR- **HPD1.**
* Sample F.G.-**V21.** The small specimens are **N.inconspicua** (looks like a small N.frustulum) with central fibulae apart.
* *Sample HPD1 to be given by JR*

N. pusilla (NI152A)
* KLB2/2, p:111, pl:79, fig:12-15

- L=8-33μm, w=2.5-5 μm, 14-20 fib/10μm, 43-55 st/10μm

Not linear, definately lanceolate with high fibulae density and striae not visible in LM.
- Sample FG-**N85/299** (to be checked)

NITZSCHIA PALEA / ACICULARIS GROUP

N.paleacea (NI033A)
- KLB2/2, p:114, pl:81, fig:1-7
- L=8-55μm, w=1.5-4 μm, 14-19 fib/10μm, 44-55 st/10μm

Specimen are narrow with a clear node.Striae extremely delicate & difficult to observe in LM.
- Sample FG-**F786**
- Sample SJ-**16P**
- *Sample to be given by SJ*

N.palea (NI009A)
- KLB2/2, p:85, pl:59, fig:1-24; pl:60, fig:1-7
- L=15-70μm, w=2.5-5 μm, 9-17 fib/10μm, 28-40 st/10μm

Specimen are clearly linear, lanceolate, clear inflexion between linear and tapening portion of the valve, no node, fibulae density irregular, gaps apparent.
- Sample FG-**NIG85/299**

N.palea var.debilis (NI009C)
- KLB2/2, p:86, pl:60, fig:1-7

Has no node, narrower than **N.palea** and denser fibulae. Looks the same than **N.acicularis**
Both species cohabitate in the same sample.
- Sample from F.G. **NIG / 254e**

N.acicularis (NI042A)
- KLB2/2, p:123, pl:85, fig:1-4
- L=30-150μm, w=2.2-5 μm, 15-22 fib/10μm, 60-72 st/10μm

Specimen are definately elongated, but narrow, no visible striae, not the same than the one in KLB2/2 (pl:85, fig:1-3) which is broader. **N.palea debilis** is still elongate but has shorter ends that can be capitated whereas **N.acicularis** has long, elongate ends not capitated at all.
In **NIG 85 / 298** what is called **N.acicularis** is perhaps **N.palea debilis**.
- Sample from F.G. **NIG / 254e**

Check if there is a gradient between N.acicularis and N.palea var.debilis (CP's job).

For the moment we differentiate **N.acicularis** from **N.palea debilis** in N.Af. dataset. Further work and discussion with other EDDI group (TP and pH) are needed to decide either grouping or splitting between **N.acicularis** and **N.palea debilis.**

N.subacicularis (NI171A)
- KLB2/2, p:118, pl:67, fig:4-10
- L=20-80μm, w=2-3 μm, 12-16 fib/10μm, 27-33 st/10μm

Specimen are L=25μm, w= 3μm, 24 -25 st / 10 μm, finely punctate.
- Sample from F.G.- **JT 13**

CYCLOTELLA CASPIA GROUP

Cyclotella caspia (CY012A)

- KLB2/3, p:46, pl:45, fig:1-8
- diam=10-50µm, 8-10 st/10µm

10-20 fultoportula only on raised side of tangential undulation (external view). Small specimens have very few fultoportula whereas bigger specimens have fultoportulae on both sides - but is this *C. caspia*?(Caspia dataset - **24B**). The smaller specimen don't seem to be **C.caspia or C. choctawhatcheeana** (it is not colliculate on one side but appear smoother in general) and we need to find a new name for it. In the African fossil dataset (F.G. **AB52b**), another small specimen called **C.caspia** is now called **C. choctawhatcheeana**. It has less than 4 (1-3) fultoportula on one side and both sides appear colliculate.

- Sample SJ-**24B**
- *Sample to be given by SJ*

Cyclotella choctawhatcheeana (no code number):
Specimen are Diam. =3- 9.5 µm, 20 - 26 st / 10 µm, 1 to 3 fultoportula on one side. Both sides appear colliculate.

- Sample F.G-**AB52b** **CHERCHER SAMPLE - LC has negatives and photos - see paper from Diatom Research.**

NAVICULA CRYPTOCEPHALA / CRYPTOTENELLA GROUP

Navicula cryptocephala var.exilis (NA007D)
- N.cryptocephala : KLB2/1, p:102, pl:31, fig:8-14   ;   L=20-40µm, w=5-7 µm, 14-17 st/10µm

Navicula cryptocephala var.**exilis** is smaller and has more delicate striae than **N.cryptocephala.**

According to LC **N. cryptocephala** var. **exilis** has marginally coarser striae, in general (14/10 µM) Valves are generally shorter than the nominate variety with sub-rostrate apices. The nominate is generally capitate/sub-capitate. It is the shortness that is critical. This is a weak charachter which is probably why it is best left as a variety.

- Sample FG-**GIX-3** (**LC**)

Navicula cryptocephala var. veneta (NA007B)
- KLB2/1, p:104, pl:32, fig:1-4
- L=13-30µm, w=5-6 µm, 13-15 st/10µm

Specimen have a square central area and less pulled out ends.
According to L.C. Valves linear-lanceolate with broadly sub-rostrate apices. Striae slightly radiate near the centre, where two or three are also shorter, otherwise striae are ± parallel and transverse. The central area of **Navicula cryptocephala** var. **veneta** is transapically widened, whereas the other species generally have rounded central areas. It also does not have the clearly protracted apices of the nominate and var. exilis. LC thinks Cox makes it a species in its own right and he would tend to agree.

- Sample FG-**LBKEH58** (**LC**)

Navicula phyllepta (NA058A)
- KLB2/1, p:104, pl:32, fig:5-11
- L=12-45µm, w=4-8 µm, 14-20 st/10µm

Specimen are smaller than **N.cryptocephala** and have a very small central rounded area. Striae length reduces progressively compared to the others, striae are more radiant and of even length, finer striae than **N.cryptocephala** (>20µm). The finer, less strongly radiate striae are the key distinguishing feature.

- Sample FG-**GXVIII 3+F** (**LC**)

Navicula cryptotenella (NA751A)

- KLB2/1, p:106, pl:33, fig:9-11 & 13-17
- L=14-40µm, w=5-7 µm, 14-16 st/10µm

Our **N. cryptotenella** is definately different from **N.cryptotenella** from pH group (see KLB 2/4, pl:31, fig:19-20, narrower than ours) and TP group. Full length striae are disposed at some angle with shorter striae.

- Sample from F.G. **KenH148**

Navicula tenelloides (NA675A)

- KLB2/1, p:117, pl:38, fig:16-20
- L=14-21µm, w=2.5-4 µm, 15-17 st/10µm

N.tenelloides from J.R. is in fact **N.salinicola** according to the above discussion.

For J.R. **N.tenelloides, N.sp 2cf.salinicola and N.salinicola** are in fact **N.salinicola.**

N.tenelloides from the Caspia set (sample **25B**) is smaller than the one from KLB2/1 but the number of striae is higher and the striae are subparallel. So this species differ from **N.tenelloides** by its dimension, striae density and direction of striae so it could be **N.perminuta** (KLB 2/1, pl: 35, fig: 14-20) TO BE CHECKED.

- Sample SJ-**25B** *(to be given by SJ)*
- Sample FG-**LBKEH89**

ACHNANTHES MINUTISSIMA GROUP

A.minutissima var affinis (AF_0101)

- KLB2/4, p:58, pl:33, fig:13-22
- L=8-30µm, w=3.5-5 µm, 22-24 st/10µm

In the African dataset **A.minutissima** var. **affinis** is separated from **A.minutissima** var.**minutissima** (AC013A) but is not recorded in the other datasets (e.g. Caspia, Spain & Turkey). JR did not separate them and PB has split them in the Mt Kenyan dataset (to be added later). **A.minutissima var affinis** can be separated from the nominate by the clearer break in striae in the central area of the raphe valve (stauros). FG thinks that **A.minutissima var. affinis** is generally found in waters with higher conductivity than the nominate.

- **Sample FG-GXVIII (is this OK?) LC does not have a good valve view picture from this sample. What about Alg 85/02 or EH89 as alternatives?**

CHAETOCEROS GROUP

C.wighamii (CH082A) :
No problems , JR and FG agree on this taxon.

NITZSCHIA GROUP

Nitzschia apiculata (NI016A) vs N. constricta (NI083A)

- KLB2/2, p:43, pl:35, fig:1-6
- L=20-58µm, w=4.5-8.5 µm, 15-20 st/10µm

In JR' Spanish dataset **Nitzschia apiculata** is the same as **N. constricta** (NI083A) in FG african dataset. **N. constricta** is the name adopted (although this has now been placed in the genus **Tryblionella**)

- Sample FG-**LBKA403** **(is this OK????) LC???**
- *Sample to be given by JR*

Nitzschia compressa v. vexans (NI200C)

- KLB2/2, p:46, pl:38, fig:5-8
- L=10-130µm, w=3.5-26 µm, 5-21 st/10µm (dimensions for **Nitzschia compressa**)

Appears to be equivalent to FG's **N. punctata** (NI004C). FG also had a smaller variety of **N.punctata**.

N. compressa is the current adopted name (although this has now been placed in the genus **Tryblionella**)

- Sample JR-**DSLD2**
- *Sample to be given by JR*

TABLE 1: List of species of interest and slides submitted

Species	Slides	Availability	from	Site
F.nanana	MA 19b	ok	P.B.	mid-Atlas
F.capucina gracilis	9P(37)	to be given	S.J.	L.V.
F.capucina var.rumpens	9P(37)	to be given	S.J.	L.V.
F.rumpens var.neogena	NaivH1a	ok	FG	Naivasha
S.acus var.angustissima	NaivH1a	ok	FG	Naivasha
A.acustiuscula	HPD1	to be given	J.R.	Spain
A.acustiuscula	12B	to be given	S.J.	L.V.
A. coffeaeformis	HPD1	to be given	J.R.	Spain
A. coffeaeformis	12B	to be given	S.J.	L.V.
A.delicatissima	Dji 87/8c	ok	F.G.	Djibouti
A.micrometra	Bol 96 / 0-0.5 386	ok	J.R.	Turkey
A.tenerrima	ALG85/ 01	ok	F.G.	Algeria
C.cuspidata	H134	ok	FG	Kenya
C.ambigua	H116 + V24	ok	FG	Kenya/ Abaitou
C.ambigua	9918A admre	ok	P.B.	Morocco
N.halophila	EH93	ok	LBK	Tunisia
N.halophila	EALM1	to be given	J.R.	Spain
N.halophila	ESLC1	to be given	J.R.	Spain
N.subhalophila	Ken. H98-99	ok	F.G.	Kenya
N.salinicola	14B	to be given	S.J.	Caspian Sea
N.sp2af.salinicola	EMNJ1	to be given	JR	Spain
Nitzschia lacuum	JT25	ok	F.G.	Uganda
N.fonticola typ.2/ N.lacuum var.3	V21	ok	F.G.	Djibouti
N.frustulum	H7	ok	F.G.	Kenya
N.liebetruthii	15P	to be given	SJ	L.V.
N.inconspicua	HPD1	to be given	J.R.	Spain
N.inconspicua	V21	ok	F.G.	Djibouti
N.pusilla	N85/299	ok	F.G.	Niger
N.paleacea	F786	ok	F.G.	
N.paleacea	16P	to be given	S.J.	Caspian Sea
N.palea	N85/299	ok	F.G.	Niger
N.palea debilis	NIG 85 / 254e	ok	F.G.	Niger
N.acicularis	NIG 85 / 254e	ok	F.G.	Niger
N.subacicularis	JT13	ok	F.G.	Ethiopie
C.caspia	Actual sample	to be given	F.C.	Caspian see
C.caspia	24B	to be given	S.J.	Lower Volga
C.choctawhatcheeana	AB52	to be given	F.G.	
Nav.cryptocephala exilis	GlX-3	ok	F.G.	
Nav.cryptocephala veneta	EH58	ok	LBK	Tunisia
N.phyllepta	GXVIII	ok	F.G.	
N.cryptotenella	H 148	ok	F.G.	Kenya
N.tenelloides	25B	to be given	S.J.	Lower Volga
N.tenelloides	EH89	ok	LBK	Tunisia
A.minutissima var.affinis	GXVIII	ok	F.G.	
Ni. apiculata	A403	ok	LBK	
Ni. apiculata	CTRJ2	to be given	JR	Spain
Ni.compressa	DSLD2	to be given	JR	Spain

END

TP datasets: 25‾26 January 1999, Jyllinge, Denmark.
Attended by:
John Anderson (NJA), Helen Bennion (HB), Sonja Hausmann (SH), Patrick Rioual (PR), David Ryves (DR) & Sybille Wunsam (SW)

Minutes taken by:
Emily Bradshaw

<u>Monday, January 25th</u>

GENERAL BUSINESS

Species and Environmental Data
1. Noted that there are no Swedish sites in TP EDDI dataset – on checking with Steve J., these sites have already been included in EDDI as part of the pH dataset. DR now has a copy of the diatom counts. **ACTION on DR to check for overlap with the other TP sites, and capture images where relevant for comparison with other datasets.**

2. Several other people (Sonja, Patrick) have more sites in progress or counted - but perhaps too late to add to EDDI? Some environmental data need to be sent to Steve too. **ACTION on those with outstanding data to send these in to Steve ASAP.**

Slides
1. **ACTION:** Sonja to check with Andy on availability and quality of suspensions or alternatively sediment, for making duplicate slides of the Swiss samples.
2. **ACTION:** Sybille to check with Roland re. availability of the Austrian sediment samples, with Aldo re. the Italian samples and Rolf re. the Bavarian samples. There may be a problem in obtaining some of the material.
3. Helen left a duplicate set of her south-east England (SEng) slides (the other set being in the ECRC archive) and the only set of the meres (SCM) slides with Dave. A few of the CCW slides were also left but **ACTION:** Helen will arrange for another set of all the CCW slides to be made and sent to Dave ASAP.

Microscope intercalibration
1. Dave will use the ECRC England Finder to record the stage co-ordinates of the GEUS image capture microscope according to Stephen Droop's protocol, to allow calibration with other microscopes being used in EDDI. Dave will pass these data onto Helen (done).

TAXONOMY SESSIONS
Note: Images captured during the workshop are coded <u>TP0001</u> to <u>TP0093</u>, and referenced in the appendix to these minutes.

Group 1 - Small *Stephanodiscus*
* *S. parvus* (ST010A)
Everyone happy - mostly the same, perhaps Patrick's are chunkier but NB some crossover to *S. minutulus* (e.g. AMME, Sybille's)
* *S. minutulus* (ST021A)
See VARE (Sybille's). This was not the typical "Mexican hat" *S. minutulus*, since only the central area was slightly domed.
Some crossover possible to *S. alpinus* - image captured from Mondsee (<u>TP0001</u>).
* *S. minutula* (ST004A)
Image captured from Bryrup Lang Sø (<u>TP0002</u>) - **DECISION: should be *S. minutulus*, merge with ST021A.**
* *S. hantzschii* (ST001A)

Several 'classics' seen from Bryrup Lang Sø (DK)
NB all datasets combine *S. hantzschii* v. *tenuis* EXCEPT Sybille's which splits them. All were happy with Sybille's v. *tenuis* but had grouped them in other sets.
- *S. binderanus* (ST015A)

Only in John's DK & NI sites. Good example image captured from Creeve (TP0003). Nobody else had ever seen it.

Group 2 - Larger *Stephanodiscus*
- *S. alpinus* (ST009A)

2 images captured from Fårup Sø (DK; TP0004 & TP0005) of what John calls *S.* cf. *alpinus*. 3rd image captured (TP0006) - not like Sybille's *alpinus* but not *minutulus*. Sybille's, real *alpinus* - good potential for image capture from Austria (ALTA). Image captured from White Lough (NI; TP0007) - both Helen and Patrick would call this *medius*. Further investigation of Fårup Sø – some specimens are thought to be 'real' *alpinus*. Some like *minutulus* but too big. Image captured (4th from Fårup Sø; TP0008) & from Patrick's (TP0009) - very like the White Lough images. Good image of 'classic' *alpinus* from Swiss set (TP0010).
- *S. medius* (ST014A)

3 images captured from Helen's SCM dataset (TP0011, TP0012 & TP0013). Little *alpinus*? Ravnso (DK; TP0014) - has *alpinus* and *medius*.
DECISION: Swiss/central European, bigger, typical *alpinus* stay as *alpinus*.
Other datasets, call them *alpinus/medius*. DECISION: all medius changes to *alpinus/medius* and John & Patrick's *alpinus* changes to *alpinus/medius*. DR to get SW's ALTA slide for photography of *S. alpinus*.
- *S. neoastrea* (ST022A)

All happy.

Group 3 - *Asterionella*
- *A. formosa* (AS001A)

All happy.
- *A. ralfsii* (AS003A)

Only Patrick has this. Didn't find one to capture image. **ACTION: Patrick to send material for photography.**

Group 4 – *Achnanthes minutissima*-types
- AC013FM (*A. minutissima (2)* - Sybille's) is synonymous with *A. straubiana* (AC178A). **DECISION: Maintain separation from other 'minutissima' types as AC178A.**
- *A. catenata* (AC165A)

Only seen by Patrick. Maintain name.
- *A. petersenii* (AC105A)

Maintain name.
- *A. linearis* (AC002A)

Not seen in central Europe / Swiss sets. Maintain name.
DECISION: all other *minutissima* varieties to be combined.

Groups 5, 6 & 7 – Small *Fragilaria*
Patrick follows K&LB
- *F. pinnata* v. *pinnata* (FR001A)

2 images from Patrick's lakes captured (TP0015 & TP0016). All happy with that one.
Image captured from Swiss slide WAN41 (TP0017) (*elliptica* PR / v. *venter* NJA & HB / *pinnata* SW & possibly Andy).
- *F. pinnata* v. *lancettula* (FR001B)

Only John has counted it. **DECISION: Merge with FR001A.**
- *F. construens* v. *venter* (FR002C)

Image (SCM; TP0018) has 2 v. *venter* and 1 *pinnata* agreed by all!
* *F. elliptica* (FR018A)

NJA and HB only use '*elliptica*' for the smaller v. *minor* form. All agreed these were *elliptica* (e.g. John's Irish site Heron: NI032). Sybille doesn't use *elliptica*. Image TP0019 taken from PR's dataset. Andy doesn't appear to use *elliptica* but possibly puts them in *pinnata*. **ACTION** Sonja to check this with Andy.

SEng images (TP0020 & TP0021) captured of 2 classic HB/NJA *elliptica*.
* *Martyana martyi* (MT001A)

DECISION: *Opephora martyi* (OP001A), *F. leptostauron* v. *martyi* (FR065A) & *F. leptostauron* v. *martyii* (FR014CM) are all synonyms for the above.

Good examples on Helen's SE057 (range of examples and possible crossover to *pinnata* - NJA). **ACTION: DR to capture images.**
* *F. brevistriata* v. *brevistriata* (FR006A)

ACTION: DR to take pictures of range and circulate.
* *F. construens* v. *construens* (FR002A)

All happy.
* *F. pseudoconstruens* (FR056A)

Only Patrick & Andy have it. Nobody else has counted it but believe it has not been confused with *construens*.
* *F. construens* v. *binodis* (FR002B), v. *subsalina* (FR002E), v. *exigua* (FR002D) all fine
* *F. lapponica* (FR011A) fine
* *F. leptostauron* v. *leptostauron* (FR014A) fine
* *F. robusta* (FR063A) fine
* *F. brevistriata* v. *capitata* (FR006E)

Only Helen has this, in one site. **ACTION: DR to capture image and check.** Possibly merge it with *brevistriata*.

Group 9 - *Cyclotella stelligera* group
* *C. pseudostelligera* (CY002A)

Image captured from Helen's SE76 (TP0022).

3 image captured from John's NI021 (TP0023, TP0024 & TP0025).

Big range!

Patrick's - similar to what John calls *glomerata* and Sybille calls *stelligeroides*.

Image from PR's CHAU (TP0026) 4 valves = *C. pseudostelligera* PR.
* *C.* cf. *stelligeroides* (1) (CY9972M)

2 images (with different focus) captured from Sybille's GOGG (TP0027 & TP0028) – type 1 is lower valve, type 2 upper.
* *C.* cf. *stelligeroides* (2) (CY9973M)

Image captured from GOGG (TP0029).
* *C. glomerata* (CY007A)

None in the Swiss set - probably call them *pseudostelligera*.

2 images (TP0030 & TP0031) captured from John's NI019 of different sized *glomerata* - all agreed it's not like the K&LB *glomerata*.

Image from PR's CHAU (TP0032): top valve = *C. glomerata*; bottom = *C. pseudostelligera* PR. John and Helen wouldn't call the bottom one *pseudostelligera*.

ACTION: Sonja to send image of *glomerata* 'type' material.
* *C. woltereckii* (CY048A)

DECISION: merge with *C. pseudostelligera* (CY002A).

All beginning to believe that there is a 'continuum' (as Liz Haworth argues) with different forms reflecting different environments.

Group 10 - *C. comensis*
* *C. comensis* (CY9983M)

Sybille uses this for 'true' *comensis* (images TP0033 & TP0034) and splits other types (images TP0035, TP0036, TP0037 & TP0038 are respectively types 1, 2, 3 and 4).

- *C. comensis* (CY010A)

Andy has used *C. comensis* for all the 'patterned' varieties (inc. *gordonensis*) and has used *Thalassiosira pseudonana* (TH031A) for non-patterned *gordonensis*.

John uses this for *comensis* and its varieties (image TP0039 from DK021: upper is (now) *C. rossi*, lower is (now) agg. *comensis*).

Patrick uses this for 'true' *comensis* and uses cf. *comensis* (CY9987) for 'others'.

DECISION: merge all the *comensis* types as *C. comensis* agg. but attach a health warning and capture images of the range.

Presentation: Sonja Hausmann - *Cyclotella comensis*

Core from Bachsee - lots of *C. comensis* but not a good fit to summer temp - evidence of several spp. in "*C. comensis*" (multimodal response to temp) - one cold and one warm form – transfer function is improved (in terms of fit to Bachsee expected recent last 100yrs increase in temp). PCA based on binary values for 15 valve features for 100 valves from 6 lakes: splits need to be tested for statistical significance (DISCRIM/CANVAR/DISKFN? – but latter 2 depend on *a priori* group membership) - what is best procedure? Also check on fit to temp from transfer function of old vs. new (good species fit) or HOF/in CALI/in CANODRAW - other methods?

Tuesday January 26th

Group 11 - *Cyclotella* cf. *gordonensis*
*Only recorded in Sybille's dataset but see note against *C. comensis* above.
Sybille comments that *gordonensis* only appears in the spring.

- *Cyclotella* cf. *gordonensis* (1) CY9974M

Almost featureless. Image captured from COMO (TP0040).

- *Cyclotella* cf. *gordonensis* (2) CY9981M

Stronger features. Image from Swiss site BUG taken (TP0041) but this is called *comensis* in Switzerland (as above).

- *Cyclotella* sp. 1 (F_ZZZ973) in Patrick's dataset is the same as Sybille's type 2 (CY9981M). Image from PR's TAZE taken (TP0042). **DECISION: Merge to CY9981M.**

Group 12 - *Cyclotella radiosa* group
DECISION: John's *C. comta* v. *comta* (CY001A) to be merged with all the *C. radiosa* (CY019A) - coding error

- *C. radiosa* (CY019A)

Image of three individuals captured from Swiss OBE (TP0043 & TP0044 in bright field. TP0045 in phase-contrast).

Fårup Sø (DK) - John's *C. comta* (*radiosa*) - Sybille recognises some *bodanica* and *praetermissa* also...

2 images captured from Fårup showing a range (TP0046 & TP0047).

Helen's SCM37 *radiosa* very flat. John & Patrick call this *radiosa*. Image captured TP0048.

Ireland site - NI018 (image TP0049 taken). Sybille would call this *C. bodanica* v. *lemanica*. Also some real *radiosa*.

Also thought (Sybille/Sonja) that Andy groups *bodanica* v. *lemanica* with *radiosa*.

Patrick's *radiosa* looks more like John and Helen's small *radiosa*. 2 images captured from PR's TAZE and FRON (TP0050 & TP0051).

- *C. radiosa* (2) (CY019BM)

Sybille's smaller *radiosa* (2) (CY019BM) which she finds in more eutrophic waters. John, Helen and Patrick would all call this *radiosa*. Image TP0052 taken.

- *Cyclotella bodanica* (CY022A)

Image captured from Sybille's ALTA (TP0053).
Nobody else has seen this, inc. we believe, Andy.
Image TP0054 (showing dissolution) and TP0055 taken, both from ALTA.
- *C. praetermissa/quadrijuncta* CY039AM
Image captured from Swiss LOC (TP0056) with large central area.
DECISION: merge the Swiss/Sybille's *praetermissa* CY039AM and SW_CPRA.

Group 13 - *Cyclotella distinguenda* group
- *Cyclotella distinguenda* v. *unipunctata* (CY028B)
DECISION: merge CY9977M, CY9976M, CY9978M & CY059A into this.
Image captured from Swiss lake OBE (TP0057).
More heavily silicified than *comensis* type.
Image taken from Patrick's Holzmaar site (TP0058) - top valve is PR's *cyclopuncta* and
bottom one is CY028B. Now to be merged.
- *C. unipunctata* (CY049A)
Only used by John (image TP0059). **DECISION: merge to *comensis* agg.**
- *Cyclotella distinguenda* (CY028A)
Image from Sybille's GOES (TP0060).

Group 14 - *C. meneghiniana/ C. kuetzingiana* group
- *C. meneghiniana* v. *meneghiniana* (CY003A)
Think this is okay but ACTION: DR to get image from Helen's SE065 - possibly smaller
type.
- *C. kuetzingiana* v. *kuetzingiana* (CY006A)
Only counted by John. **ACTION: John to recount DK021. All other CY006A to go into**
Cyclotella **sp.**
- *Cyclotella kuetzingiana* v. *minor* (CY006D)
DECISION: merge with *comensis* agg.

Group 15 - *Cyclotella ocellata*
- *Cyclotella ocellata* (CY009A)
Very different concepts HB/PR/NJA to Swiss set.
DECISION: merge in CY009BM, CY009CM, CY009DM, CY009FM and CY009EM
N.B. 3-4 ocelli are more like the classic *ocellata*; others are aff. *ocellata*.
Image from SEng (TP0061) of the 'classic' *C. ocellata*.
Image from Swiss BUG (TP0062 – 3-4 ocelli) with a much bigger range of types. Andy's
idea of *ocellata* is much broader. Andy splits away *rossii* and *krammeri*.
Sybille didn't distinguish *rossii*. Range of images captured from her ALSE site with 3 ocelli
(TP0063 & TP0064), 4 ocelli (TP0065) and 5 ocelli (TP0066). Also from this lake views of
both valves from 2 heterovalvar cells captured (TP0067 & TP0068; TP0069 & TP0070).
Image of 4 ocelli forms (2 valves) also taken from PR's TAZE (TP0071).

Group 16 fine (PR's Cyclotella sp.1 merged with CY9981M – see above).

Group 17 – finer *Cyclostephanos* **spp.**
- *Cyclostephanos* cf. *tholiformis* (CC9997)
Not present in Andy's set, or in Sybille's. Not thought to have been mis-identified. Image
TP0072 taken from John's NI030.
DECISION: merge cf. *tholiformis* to *tholiformis* as CC003A
Helen and John's are "exactly the same". Happy with PR's too.
- *Cyc. invisitatus*
Everyone happy. Image TP0073 taken from PR's TAZE.

Group 18 - coarser *Cyclostephanos* **spp**

- *Cyclostephanos dubius* (CC001A)

All happy! *Dubius* in name only...

- *Cyc. dubius* v. *minor* (CC001B)

Only in John's NI sites. Spines on every fascicle. Nobody else has seen this. Not sure it is *Cyclosteph.* but it is different to anything else.

ACTION: DR to take photos of this and *Thalassiosira guillardii.*

Group 19 - *Aulacoseira ambigua*

- *Aulacoseira ambigua* (AU002A)

Everyone happy (a pleasant irony given the specific epithet).

Group 20 - *Aulacoseira subarctica/islandica*

- *Aulacoseira subarctica* (AU020A)

Images from John's sites in Ireland (TP0074 & TP0075) and Denmark (TP0076) taken. DECISION: merge in AU001B - synonym.

Sybille only has type 1. All happy. **ACTION: DR to get new slide from Sybille and compare image to Danish material - very similar, squat form.**

- *Aulacoseira (subarctica* - type 2)(AU9986)

Only Patrick has used this (image from CHAU2: TP0077). **ACTION: if Swedish lakes come back in - John will check his type 2 and coding.**

- *Aulacoseira islandica* (AU009A)

Only John & Sybille. John hasn't split out v. *helvetica*. Decided that split to *helvetica* is rather subjective. Image captured from NI054 (TP0078).

Image from Sybille's GARD (TP0079). **DECISION: merge in v. *helvetica* (AU009B).**

Group 21 - *Aulacoseira granulata*

- *Aulacoseira granulata* v. *granulata* (AU003A)

ACTION: check all Swiss A. granulata (AU003A & morphotype *curvata* SW_AGCU) with Andy - Sonja did not have slide to check if v. *angustissima* was grouped in. DECISION: merge in AU003D additional code error.

Happy with Sybille's.

- *Aulacoseira granulata* v. *angustissima* (AU003B)

N.B. John has not split *curvata* and this is certainly what is in DK014 (image TP0080).

John & Sybille would call some of Helen's v. *angustissima* taxon *A. 'ambigua'*.

Confusion all round! *ambigua* without sulcus or *angustissima*?

Some range in John's NI042.

ACTION: DR to capture range of images and investigate a little...

ACTION: Helen to send DR slides of the Talley lakes to look at.

- *Aulacoseira granulata* morphotype *curvata* (SW_AGCU)

Only Andy has used this. ACTION: Material to be sent to DR for photography (see above).

Group 25 and 30 *Synedra* & *Fragilaria*

Appears that Andy splits by broadness/fineness....

ulna = coarse

acus = finer

nanana = very thin & fine

We think, but check.

DECISION: All the following are the v. long and fine, can't see striae:

FR9980M (*F.* aff *nanana*)

SW_FNAN (F. *nanana* – Swiss) – image TP0081 from Swiss NER

SY009A (*Synedra nana*) – image TP0082 from NI037

John's *Synedra tenera* (SY013A) also merges to *F. nanana* (image TP0083 from DK023)

FR9982M (Sybille's *Fragilaria nanana-, delicatissima-* group) is this too but may include some others.

**Can group as *F. nanana (currently no code for this taxon – use SY009A for present)*
Image <u>TP0084</u> – from SW's lake TURN - but which taxon code should this be?**

- *Synedra ulna* v. *ulna* (SY001A)
DECISION: merge in SW_FULN - synonyms.
- *Synedra delicatissima* v. *delicatissima* (SY011A)
Only Patrick (<u>TP0085</u>) has used this but
= *F. ulna - danica - tenera* (FR9981M) to Sybille
= *ulna* v. *acus* (SW_FUAC) to Andy (<u>TP0086</u>)
= *radians* (SY017A) to John (<u>TP0087</u>). "Not too coarse, not too long, capitate ends". This area needs looking into (**ACTION: DR to look at material from several datasets**).
- *S. ulna* v. *danica* (SY001C)
DECISION: John's is "what other people might call *acus*" - merge it!

ulna seems clear
acus seems clear
some confusion with *tenera, delicatissima, radians*. **ACTION: DR to try to take some images to circulate and sort into ~4 groups (see taxonomic sketch by PR).**

- *F. crotonensis* (FR008A)
N.B. Rostherne Mere (SCM039) good for pictures. Also MILL (Sybille's). Image grabbed from SCM39 (<u>TP0088</u>).
ACTION: Question mark in Swiss set - couldn't find any in GRD. Sonja to check with Andy.
Everyone else happy. N.B. Helen's look like they are a different form - don't connect at ends and so see colonies less often.

Group 35 - *Amphora*
- *Amphora pediculus* (AM012A)
All happy.
- *A. libyca* (AM011A)
DECISION: merge in *A. ovalis* v. *libyca* (AM001C).
- *A. ovalis* v. *ovalis* (AM001A)
All happy.
- *A. inariensis* (AM013A)
DECISION: merge to AM012A

Group 40 - *Tabellaria*
Some confusion here. These have not been split consistently except by John.

- *T. flocculosa* agg. (TA9996)
Image <u>TP0089</u> (TA001C) from Ireland. **DECISION: merge in TA001A, TA001C, TA9997, and TA9998.**
- *T. binalis* fo. *elliptica* (TA003B)
Fine.
- *T. fenestrata* (TA002A)
Fine.
- *T. quadriseptata* (TA004A)
Should have "shark's teeth" spines. In MILL (Sybille's) looks "close to *quadriseptata*" but also "something like" *flocculosa*. Images captured from SW's MILL (<u>TP0090</u> & <u>TP0091</u>).
ACTION: DR to send pictures to Rog?

Group 43, 44 - *Navicula*
- *N. cryptocephala* v. *cryptocephala* (NA007A)

Okay
- *N. veneta* (NA054A)

DECISION: merge in *N. cryptocephala* v. *veneta* (NA007B) - synonyms.
ACTION: DR to find good image and circulate.
- *N. exilis* (NA064A)

ACTION: DR to see Germain and grab good image
- *N. cryptotenella* (NA751A)

DECISION: merge in *N. cryptotenella* (NA9982M) and *N. radiosa* v. *tenella* (NA003B) -
synonyms.
Big range.
ACTION: DR to capture range of images.
- *N. gregaria* (NA023A)

All happy. Image from Denmark taken (DK003: TP0092).
- *N. radiosa* v. *radiosa* (NA003A)

All happy.

Group 45 - *Achnanthes lanceolata*
- *A. rostrata* (AC031A)

DECISION: merge in *A. lanceolata* v. *rostrata* (AC001B) - synonyms.
- *A. lanceolata* (AC001A)

DECISION: merge in subsp. *frequentissima* (AC001R)

Group 50, 51 - *Cocconeis*
- *Cocconeis placentula* v. *placentula* (CO001A)

DECISION: merge in other varieties CO001C, CO001B, CO001FM.
- *C. neodiminuta* (CO066A)

DECISION: merge in *C. neodiminuta* (CO9997M) and *C. diminuta* (CO006A).
- *C. neothumensis* (CO067A)

DECISION: merge in *C. neothumensis* (CO9998M) and *C. thumensis* (CO009A).

Group 55, 56 - *Nitzschia*
- *N. palea* v. *palea* (NI009A)

Image TP0093 (Ireland) captured.
DECISION: merge in cf. *palea* small (NI9971)
Happy with the rest but **ACTION: check *palea* v. *debilis* with Helen.**

Appendix – Image catalogue

The file names are derived from the SampleId (as seen in the Sample Finder on Steve's website) and the TaxonId as used by the taxonomist involved with that dataset, in the form <SampleId>_<TaxonId>.

Number	Image	File_name	Magnification
1	TP0001	MOND_ST021A	x1250
2	TP0002	DK008_ST004A	x1250
3	TP0003	NI054_ST015A	x1250
4	TP0004	DK025_ST009A1	x1250
5	TP0005	DK025_ST009A2	x1250

6	TP0006	DK025_ST009A3	x1250
7	TP0007	NI036_ST009A	x1250
8	TP0008	DK025_ST009A4	x1250
9	TP0009	TAZE_ST009A	x1250
10	TP0010	LOC_ST009A	x1250
11	TP0011	SCM026A_ST009A	x1250
12	TP0012	SCM028A_ST014A1	x1250
13	TP0013	SCM028A_ST014A2	x1250
14	TP0014	DK007_ST014	x1250
15	TP0015	GODB_FR001A	x1000
16	TP0016	BOUR1_FR001A	x1000
17	TP0017	WAN41_FR001A	x1000
18	TP0018	SCM043A_FR2SPP	x1000
19	TP0019	BOUR1_FR018A	x1000
20	TP0020	SE083_FR018Aa	x1000
21	TP0021	SE083_FR018Ab	x1000
22	TP0022	SE076_CY002A	x1250
23	TP0023	NI021_CY002A1	x1250
24	TP0024	NI021_CY002A2	x1250
25	TP0025	NI021_CY002A3	x1250
26	TP0026	CHAU_CY002A	x1000
27	TP0027	GOGG_CY9972&3Ma	x1000
28	TP0028	GOGG_CY9972&3Mb	x1000
29	TP0029	GOGG_CY9973M	x1000
30	TP0030	NI019_CY007A1	x1000
31	TP0031	NI019_CY007A2	x1000
32	TP0032	CHAU_CYSPP	x1000
33	TP0033	FARC_CY9983M	x1000
34	TP0034	ATTE_CY9983M	x1000
35	TP0035	WOER_CY9982M	x1000
36	TP0036	GRUN_CY9979M	x1000
37	TP0037	GARD_CY9988M	x1000
38	TP0038	FELD_CY9975M	x1000
39	TP0039	DK021_CYSP	x1000
40	TP0040	COMO_CY9974M	x1250
41	TP0041	BUG_CY010A	x1250
42	TP0042	TAZE_CYSP1	x1250
43	TP0043	OBE_CY019Aa	x1000
44	TP0044	OBE_CY019Ab	x1000
45	TP0045	OBE_CY019Ac	x1000
46	TP0046	DK025_CY001Aa	x1000
47	TP0047	DK025_CY001Ab	x1000
48	TP0048	SCM037_CY019A	x1000
49	TP0049	NI018_CY001A	x1000
50	TP0050	TAZE_CY019A	x1000

51	TP0051	FRON_CY019A	x1000
52	TP0052	FLAT_CY019BM	x1000
53	TP0053	ALTA_CY022A1	x1000
54	TP0054	ALTA_CY022A2	x1000
55	TP0055	ALTA_CY022A3	x1000
56	TP0056	LOC_SW_CPRA	x1000
57	TP0057	OBE_CY028B	x1000
58	TP0058	HOLZ_CY059A&28B	x1000
59	TP0059	NI048_CY049A	x1000
60	TP0060	GOES_CY028A	x1000
61	TP0061	SE113_CY009A	x1000
62	TP0062	BUG_CY009A	x1000
63	TP0063	ALSE_CY009BM1	x1000
64	TP0064	ALSE_CY009BM2	x1000
65	TP0065	ALSE_CY009CM	x1000
66	TP0066	ALSE_CY009DM	x1000
67	TP0067	ALSE_CY_HETEROa	x1000
68	TP0068	ALSE_CY_HETEROb	x1000
69	TP0069	ALSE_CY_HETEROa2	x1000
70	TP0070	ALSE_CY_HETEROb2	x1000
71	TP0071	TAZE2_CY009A	x1000
72	TP0072	NI030_CC9997	x1000
73	TP0073	TAZE1_CC002A	x1000
74	TP0074	NI014_AU001Ba	x1000
75	TP0075	NI014_AU001Bb	x1000
76	TP0076	DK007_AU001B	x1000
77	TP0077	CHAU2_AU9986	x1000
78	TP0078	NI054_AU009A	x1000
79	TP0079	GARD_AU009A	x1000
80	TP0080	DK014_AU003B	x1000
81	TP0081	NER_SW_FNAN	x1000
82	TP0082	NI037_SY009A	x1000
83	TP0083	DK023_SY013A	x1000
84	TP0084	TURN_FRAGSP	x1000
85	TP0085	PAVI_SY011A	x1000
86	TP0086	BUG_SW_FUAC	x1000
87	TP0087	NI037_SY017A	x1000
88	TP0088	SCM039_FR008A	x1000
89	TP0089	NI046_TA001C	x1000
90	TP0090	MILL_TA004A1	x1000
91	TP0091	MILL_TA004A2	x1000
92	TP0092	DK003_NA023A	x1000
93	TP0093	NI001_NI009A	x1000

END

Work package 2 -Taxonomic harmonisation workshops
First WP2 workshop: 1-3 March 2000, UCL.
Helen Bennion (HB), Nigel Cameron (NC), Christine Pailles (CP), Dave Ryves (DR), Micha Bayer (MB).

Financial matters
A reminder that all cost statements and the Technical annual reports must be submitted by end of March 2000. Owing to some problems with release of finances at CEREGE last year, it is vital that there is minimum delay with processing of this years cost statements.
ACTION: Christine/Francoise (CEREGE), Dave/John (GEUS), Helen/Anson/Nigel (UCL) and Steve (NCL) to ensure that their institutes send the signed cost statements to UCL, and that signed copies of the Technical Annual Report forms (as last year) are sent to Helen by the end of March.

Review of minutes/ key action points of November Steering Group Meeting re. WP2
No major outstanding issues.
Image capture has been ongoing although CP reported that she still feels a little behind schedule due to a late start.
WP1 harmonisation is progressing in parallel with WP2. Finalised harmonisation of WP1 datasets has been delayed awaiting a final steering group decision on how to structure the taxonomic merging tables.

Review of the aims of WP2
All agreed on the objectives. The issue of a hard copy taxonomic guide as one of the deliverables was raised.
ACTION: The format and scope of the Hard Copy Taxonomic Guide needs to be discussed at the next Steering Group meeting.

The Image Database
Various issues relating to the development of the EDDI image database and image capture technique were discussed with Micha Bayer (RBGE). The main points were:
1. Links with ADIAC. There is potential to use images from the ADIAC project within EDDI. These images are all in brightfield (BF), include various focal planes and are of excellent quality. Micha demonstrated the ADIAC database. The system uses imported PANDORA synonyms for each taxon and uses PANKEY to generate automated keys and text descriptions.
ACTION: Discuss the potential links between EDDI & ADIAC at the next Steering Group meeting. HB to send MB our EDDI hit-list of the most important taxa to see how many of these are included in the ADIAC database already.

2. Image capture protocols. We still need to standardise our microscopy methods as NC has used both BF and DIC, CP has used mostly DIC, and DR has used BF and phase. Also MB advised us to try not to select inverted valves or whole frustules if possible. If as you focus down, the valve outline is black and then turns to white with a ghost-like appearance when the striae are focused in black, then you have an inverted valve. Perhaps not a problem for the working images but may cause problem of poor consistency & difficulty of comparison if these images are to appear in the final EDDI image database. MB suggested that we delete a number of points from our initial image capture protocols, as this will open up our choice of valves.
ACTION: We can now, therefore, select valves that are not level, valves that have some debris, and valves where there is something over/underlying.

3. Image quality. MB thinks that the noisy UCL images are caused by a camera problem, such as electronic interference. However, he thinks that the quality of all of our EDDI images are

adequate for inclusion in a "working images database". MB suggested that the final web page could have a "Front page" high quality image captured at RBGE (maybe a single image only), linked to a larger set of "working images" taken by CP, NC & DR.
ACTION: Discuss this approach at the next Steering Group meeting.

4. Image filenames. MB suggests that we might rethink our protocol for naming our image files. Currently we have a system of 8 characters. The first character denotes the co-ordinating group (U, C or G) and the remaining 7 characters are a numerical sequence starting at 0000001. MB suggests that instead we could adopt the RBGE system of using only the first 6 characters for the group & numerical sequence, and reserve the 7th character as an alphabetic one to denote the various focal planes, and have the 8th character also as alphabetic to denote the various modifications to the image such as different illumination (eg. U00034ad would be image no. 34 taken at UCL with the first choice of focal plane and the fourth modification to that image). This allows for a distinction between ranges of images taken of the same specimen with those taken of different specimens.
ACTION: Steering Group to discuss and decide.

5. Microscope intercalibration exercise. It was decided that we would redo this. MB has the UCL England Finder and a test-slide, and will circulate these to DR, CP and back to NC asap.
ACTION: MB, NC, CP & DR to report co-ordinates back to HB.

6. High Quality Image Database. MB pointed out that we should think about focal plane choice for the images to be captured at RBGE. If only one focal plane is required, then this could reduce MB's workload by 20-40%. The option of a TifTag (saves the stage co-ordinates and slide number as part of the TIF file) must also be considered because if this is not required this could reduce MB's workload by a further 20%. MB can add scale bars automatically at RBGE. MB estimates that he can capture images of c. 40-50 specimens per day. Given that there is a total of 3 months salary allocated for this work, then approximately 2500 images could be captured for the high quality image database. However, because the microscope is not for the sole use of MB, it could probably not be dedicated on a full-time basis to EDDI purposes for a 3 month block of time.
ACTION: MB will look into this as it will clearly affect the timing of the image capturing and how soon specimens need to be selected. EDDI steering group also to discuss strategy.

Taxonomic harmonisation WP1
1. DR, CP & NC should continue harmonisation based on their individual taxa hit-lists (produced at the various WP1 taxonomic workshops). Images should be captured of each taxon and where necessary this should include an example from each of the different regional training sets or at least from each of the contributing taxonomists to ensure that the full range is covered. For each specimen, one BF image, and where appropriate one image using another illumination technique, should be captured. A variety of focal planes should be used, where required, to illustrate diagnostic features. It was agreed that the image capture protocols have to remain flexible to some extent because not all taxa require the same treatment and some of the less important or rare taxa may be difficult to find. The suggestions of MB above should be adhered to as best as possible.

Taxonomic harmonisation WP2
1. Taxonomic harmonisation template. Steve's template for recording taxa and merges was trialled by DR for a group of *Stephanodiscus/Cyclotella* taxa. All agreed that this was a sensible and workable approach.
ACTION: All continue to use this template for recording taxonomic merges.

2. Taxonomic descriptions. There is still no clear protocol for writing the taxonomic descriptions for each EDDI taxon. It was agreed that owing to time pressure that the minimum information should be provided. We proposed to provide only the authority and one literature reference for non-problematic taxa; and to provide a short text string describing size ranges and merging notes with caveats etc for problematic and merged taxa. i.e. something along the lines of the SWAP Red Book. These of course would not be easy to query in the final database. DR, CP & NC all agreed that this took low priority at the moment and that the harmonisation spreadsheets and image capture should take priority at this stage.
ACTION: Need confirmation of acceptance of this proposal/approach with the other Steering Group members.

3. Taxonomic harmonisation methodology.
i) A master sheet was produced at the workshop in Excel, listing all of the taxa in the full EDDI dataset. The taxa were ordered firstly according to their importance in all 3 datasets (ie present in pH, salinity & TP datasets). Importance was defined as the sum of the product of number of occurrences and maximum abundance for each of the individual regional training sets. The resulting number was labeled "BigOrder" in our spreadsheet. Secondly the same system was used to produce a list of those taxa occurring in the pH & TP datasets, then the pH & salinity datasets and the TP & salinity datasets. Finally those taxa occurring in the pH dataset only, the TP dataset only and the salinity dataset only were listed in order of importance.

ii) We then worked through the separate groupings and selected the 40 most important taxa from the list that occurred in all 3 datasets. For each taxon, it was decided whether CP, NC or DR should take overall responsibility for the harmonisation decision. The same was then done for the taxa that were identified as important in the various pairs of datasets. Interestingly, there were very few overlaps between the pH & salinity datasets, and it was the salinity dataset where taxa were often found in their highest abundance. Consideration of synonymous taxa and different codes applied to the same taxon name were made during our discussions. We decided to ignore "sp. " taxa as these are an amalgamation of many different species and are used as dumps by most taxonomists. However there were a number of alarmingly high % of "sp." in some cases and it was, therefore, suggested that it might be worth identifying the slides where these occur to see if they refer to a single unidentified taxon or many different species from unknown/broken/dissolved specimens.

iii) Over the next two months, the priority action will be for DR, NC & CP to all take at least one image of each of the important overlap taxa determined from the exercise above. Where necessary (e.g. where there is marked variability), a range of images should be captured to represent the variation. Images should be taken in BF only and the number of different focal planes should be kept to a minimum. These images will be written to Steve's EDDI FTP site for perusal by all three diatomists. A new subdirectory called WP2 should be set up so that the images are kept separate from the WP1 images already sent. The FTP address is 128.240.122.71
ACTION: Steve to set up a WP2 subdirectory in EDDI\UPLOAD.

Once all of the overlap images have been uploaded to the WP2 subdirectory, the responsible diatom co-ordinator (identified in ii above) will make the ultimate taxonomic harmonisation decision, in consultation with the others where necessary. This requires each co-ordinator to have an alphabetically sorted list of all EDDI taxa so that synonyms, mis-codings, possible amalgamations can be more easily identified. Any notes from WP1 workshops should be made available to all co-ordinators.
ACTION: DR to produce the alphabetically ordered Excel lists and email the new file to HB, NC & CP. DR, CP & NC to ensure that they have circulated all relevant harmonisation workshop notes to each other.

4. Hierarchies and the diatom coding system

This was discussed at some length at the workshop and it became clear that this is a complex issue. For instance we will need new codes when merging say varieties into an aggregate (eg. *Stephanodicus hantzschii* fo. *tenuis* merged with the nominate) at the level of the individual datasets in WP1. Then we will need another set of newcodes to describe the new merged taxa at the pairs of datasets level in WP2, and finally some kind of "supercode" to describe the lowest taxonomic resolution when merging across all of the EDDI datasets in WP2. Will this also link to some kind of hierarchical system for the images as there will potentially be a range of images from the separate WP1 datasets and then a bigger set of images for the pairs of datasets, and the full range of images for those "merged taxa" that occur in all 3 datasets?

ACTION: The allocation of new codes to the merged taxa concepts and the potential need for a hierarchical coding system (& image database system) needs discussion by the Steering Group.

Timetable/deliverables for the next 6 months.

1. MB to re-do microscope intercalibration and circulate the UCL England Finder and test slide to DR, CP and back to NC. **Deadline: End of March.**

2. Cost statements and Annual Technical reports to be completed and signed (Action on Dave, Christine, Steve, Helen & Nigel) . **Deadline: End of March. VERY IMPORTANT!!!**

3. NC, CP & DR to work on the image capture of the overlap taxa that occur in all 3 datasets as a first priority, followed by those occurring in the pairs of datasets. Upload images onto the FTP site as soon as possible. **Deadline: End of May.**

4. NC, CP & DR to continue taxonomic harmonisation (using Steve's template) and image capture of the WP1 individual hit-lists of taxa. **Deadline: ???**

5. Taxonomic descriptions should be ongoing but should take lower priority than the harmonisation work. Descriptions should continue to be text format (ie in Word) but cross-referenced to the image files.

6. **Next workshop: June 1999**, Denmark? HB, CP, DR, NC plus Rick, John, Francoise & Steve?

ACTION: HB to arrange a meeting with Rick & Anson at UCL asap to assess the EDDI workshop budget. If funds are adequate then we propose a WP2 Second Workshop in June 2000, along with a Steering Group meeting. Perhaps organised by DR in Denmark? Everyone to feed back to HB with their thoughts on this.

END

Second WP2 workshop: 12-14 June 2000, University of Copenhagen.

Participants: Helen Bennion, Christine Pailles, Dave Ryves, Steve Juggins, John Anderson, Rick Battarbee.

Apologies: Francoise Gasse (written apology and progress report sent to the steering group), Nigel Cameron (ill).

Database of sites and chemistry data - Helen reported on outstanding gaps as follows:

- *Sitecode* - complete for all.
- *Lake names* - complete for all. Note that KOLA and FINLAND do not have lake names.
- *Lats/longs* - complete except for NW-EURO where still no values for Danish sites.
- *Catchment data* - complete for all. Note that data are not available for TURKEY and CASPIAN SEA.
- *Chemistry data* - complete for all.
- *Units of measurement* -missing for SWAP, JCEN, APEN, and ALEN.
- *Chemistry dates* - missing for SWAP, JCEN, APEN, ALEN, CASPIAN SEA and all Danish and Irish lakes in NWEURO.

- *Diatom dates* - missing for SWAP, JCEN, APEN, ALEN, SWEDEN, CASPIAN SEA and all Danish and Irish lakes in NWEURO.

ACTION: John to provide missing NWEuro data identified above. Nigel to chase up missing pH related data identified above. Steve to provide Caspian Sea sampling dates. Helen to chase up Tom's Swedish diatom sampling dates.

Report on WP1 Progress - Dave reported that the WP1 merge tables for TP are almost complete and that WP1 image capture of the TP hit-list is around a third to a half complete. Christine reported that the WP1 merge tables for Salinity were almost complete except that the *Nitzschia* spp. were still problematic and were being harmonised by Francoise and Phil Barker. The WP1 image capture of the Salinity hit-list is around a third to a half complete. Francoise's letter and Christine reported delays owing to difficulties with working with Jane's datasets (e.g. lack of slides, sample coding inconsistencies, Caspian Sea recounts).
ACTION: Francoise to finalise the Nitzschia spp in time for the September 15th deadline for WP1 harmonisation tables completion. This is urgent !

WP1 and WP2 Merge tables
The protocol for generating new codes and for harmonising taxa between datasets was reviewed. It was agreed that the templates work well. However, an "Authority" field should be added to Table 2 to distinguish true taxa from merge concepts.
ACTION: Steve to use Dave's TP WP1 merge tables as a trial run and to feedback if any adjustments to the protocols are required.

A list of actions for Steve was compiled as follows:
- *EDDI Sample Finder*: Caspian diatom counts need updating. Also the East African counts are being re-checked by Francoise. Christine will send any changes to Steve so that the Sample Finder can be updated. The fact that the 12 Swedish samples are currently part of only the SWAP pH dataset in the EDDI Sample Finder was raised by Dave as these samples also form part of the NWEuro TP dataset. Steve to see whether these samples can be included in both datasets? Note that the same samples have different site codes in the NWEuro file than in the SWAP file.
- *EDDI Slide Query?*- A request was made to enable the EDDI counts database to be queried by slide. Steve to set up.
- *Image Viewer*-the field names in Christine's files have become scrambled. Steve/Christine to check and restructure.

High Quality Image Database- It was agreed at the previous WP2 meeting with Micha that he would require approximately 6 months at 50% time to re-capture the selected specimens for the high quality image database. The first set of slides will, therefore, need to be available to send to Micha in October 2000. This means releasing original slides, thus rendering them unavailable for further image capture by Dave, Nige & Christine for short spells of time. The process of slide exchange needs to be co-ordinated. It was agreed that Nige, Christine & Dave should be responsible for identifying specimens for the high quality image database for their respective WP1 hit-list of taxa. Specimens for the common overlapping taxa can be selected jointly at the October 2000 workshop.

There is still a question over Micha's time on ADIAC and how EDDI will be timetabled around this. It is possible that ADIAC time can be used for capturing EDDI images although Micha's first impression is that only c. 5% of the EDDI specimens that we have images for so far fulfil the stringent ADIAC criteria. The way forward is still to be decided and will depend on the outcome of the forthcoming ADIAC meeting on 25/26th June.
ACTION: Steve to feed back to the EDDI steering group on the outcome.

Hard Copy Taxonomic Guide

It was agreed that this should continue to take low priority. Any taxonomic documentation can be incorporated into the final EDDI package so we should not concentrate efforts on producing the deliverable of a Hard Copy Taxonomic Guide at this stage. The fields can be set up later. The final taxonomic guide will be downloadable from the Web site. It would be useful to provide an estimate of how many taxa can be described simply by their authorities and how many will require additional notes so that the workload can be assessed.

ACTION: Nige, Christine & Dave to provide the above information based on their WP1 merge tables.

EDDI slide archive -Helen reported that material for some samples are missing and therefore slides cannot be made for the EDDI slide archive. A list is shown below:

pH

SWAP - most samples are available in the ECRC solution archive, with the exception of the Swedish samples (check this with Nige?). Slides still need to be prepared.

JCEN (Jordi)- No. 83 is missing. Two sets of slides prepped of all others (both NGC).

APEN (Alpe)- complete set should be available in the ECRC solution archive. Slides still need to be prepared.

ALEN (Aldo) - Complete. Two sets of slides prepped (both NGC).

SVALBARD - complete set available in ECRC solution archive. Slides still need to be prepared.

FINLAND - complete set available. Two sets of slides prepped (HB/NGC).

SWEDEN (Tom)- all available material supplied but Nos. 407, 905, 1518 are missing. Two sets of slides prepped (HB/NGC).

BERGEN - all available material supplied but 21 samples are missing. Two sets of slides prepped (both NGC).

KOLA - one set of slides only (NGC).

TP

NW-EURO - Two sets of slides prepped for all SEng and Meres samples (both HB). Four CCW lakes missing from ECRC archive and the rest still to be preppped from ECRC archive suspensions. Danish and Irish material still at GEUS but some samples are missing.

C-EURO - The Italian set is complete and two sets of slides prepped (HB/DR). Roland has sent one set of slides from the Austrian lakes but 8 samples are missing (DR). All 17 of the German samples are missing.

FRENCH-CRATER -complete. Two sets of slides prepped (HB/DR).

SWISS- complete. Two sets of slides prepped (HB/DR).

Salinity

N&E AFRICA - Multiple copies of slides are available at CEREGE but some samples are missing.

CASPIAN SEA - no material supplied.

SPAIN- One set of slides for 6 samples available (CP). Material for the rest are missing.

TURKEY- all available material supplied but 25 out of 40 samples are missing. Two sets of slides prepped (HB/CP).

ACTION: Helen to arrange for all outstanding slides to be prepared at UCL over the next 6 months. Dave to send NWEURO Danish and Irish material to Helen.

Helen to approach Francoise to request whether African slides can be made available to the EDDI slide archive.

It was agreed that Helen and Rick arrange a meeting with Eileen Cox at UCL soon to discuss labelling protocols etc.

ACTION: Helen and Rick to agree a date with Eileen.

Work Package 4

Steve will appoint someone for 9 months on EDDI WP4. This leaves 3 months money to employ a second person (possibly Karina Weckstrom) to assist with issues such as revising and updating DIATCODE, linking EDDI with PANDORA, sorting synonyms, adding new taxa names & authorities etc.

ACTION: Steve to discuss with Dave Mann and Micha Bayer on how to proceed.

EDDI Users Guide
It was agreed that the co-ordinators should take responsibility for producing the Users Guide to EDDI. The system will be menu-driven with on-line help functions so the hardcopy guide can be relatively brief.

EDDI Poster Presentations - the original, multi-author EDDI poster has been accepted at the Palaeolimnology Symposium in Canada (Aug 20-24th). It was agreed that a second poster should be submitted to provide an update and present some results. Given that harmonisation of the TP datasets is nearing completion, it was agreed that the poster should focus on only the TP datasets and that following approval by all data contributors to these datasets, Steve and John would produce a multi-author abstract by Friday 23rd June. Approval has now been confirmed by all and Brian Cumming has given permission for the late abstract to be submitted.

ACTION: Dave to send all TP merge tables to Steve. Steve/John to submit an abstract with Steve as first author. Steve to run the models etc and John to provide text. Dave/John/Steve/Helen to prepare a poster with production at GEUS.

EDDI Publications- it was agreed that we submit a Note to JOPL as part of the Paleo 2000 Conference Proceedings to introduce EDDI (building on the Vienna paper) and to publicise the on-line reconstructions aspect of EDDI. It was agreed that we should arrange an informal discussion group with other members of the EDDI consortium at the Canada meeting in August 2000. A bigger paper building on the TP poster which compares optima from different training sets and applies the new models to cores will be led by Steve and John.

Contracts and finance
Nigel and Dave are both employed until the end of EDDI at the end of March 2001. However, Christine's contract expires at the end of October 2000. Given the agreed timetable for completion and need to continue image capture beyond October, a 2 month extension to Christine's contract was suggested by the Steering group.

ACTION: Rick & Christine to discuss the possibility of an extension to Christine's contract with Francoise. Christine to discuss finances in the CEREGE contract with Francoise.

The second annual payment appears not yet to have been made by Brussels.
ACTION: Helen to see Anson/Patrick about chasing up with UCL.

Post-EDDI plans and follow up ideas
Rick introduced the idea of the PEP3 Multiproxy database for climate change reconstruction proposal. It was agreed that EDDI community could service the diatom side of the project. The database could initially be comprised of EU funded data to avoid issues of data rights.
ACTION: Rick to pursue. Deadline September 2000.

Remaining workshops
1. *WP2* - A final WP2 workshop will be held in October to finalise the WP2 merge tables. Nige, Christine & Dave to attend. Participation by others is optional. Marseilles was proposed as the venue.
2. *WP3* - There will be a workshop between Steve and John Birks but no further WP3 large workshops.
3. *WP4* - Steve plans to visit colleagues at the World Data Centre in Boulder, Colorado to discuss the database structures etc.
4. *Final Review-* This was originally timetabled for month 30/31 (September 2000) but it was agreed that this should now be postponed until February 2001. All data contributors should be invited to attend. Steve suggests that the database goes on-line in advance so that all participants can test it and feed back at the review meeting.

Revised Timetable (June 2000-April 2001)
It was agreed that the project is behind schedule and therefore a new timetable was produced for the remaining time available. The deadlines below MUST be adhered to in order for the project to be successfully completed by end of March 2001.

1. *June 23rd 2000* - Dave to complete WP1 TP harmonisation tables and send to Steve. Steve to begin work on these in preparation for the Canada EDDI poster/paper. Steve/John to submit an abstract to the Paleo Symposium. Steve will use this as an example dataset and will advise if any adjustments are required to the merge tables template etc asap.
2. *July/August 2000* - Steve and John to produce text and figures for the TP EDDI poster. Helen and Dave et al to contribute. Dave to co-ordinate poster production at GEUS.
3. *Before end of August* - Francoise and Phil to have finalised the *Nitzschia* spp in the Salinity datasets.
4. *September 15th 2000* - Final deadline for completion of all three WP1 sets of merge tables, and for all WP2 overlap images to be uploaded to the FTP site. CRITICAL DEADLINE!!!
5. *Mid-Sept to mid-Oct 2000* - Nigel, Dave & Christine to finalise decisions on WP2 overlap taxa based on FTP site images and note any outstanding issues for discussion at the October workshop.
6. *Mid-late October 2000* - WP2 Final Workshop to complete the cross-datasets merge tables for WP2.
7. *October onwards???* (to be confirmed with Micha) - Capture all remaining WP1 images from hit-list. Select specimens for Micha to recapture and forward slides to Micha where necessary. Slide exchange to be co-ordinated by Helen.
8. *End of December 2000* - completion of first set of taxonomic descriptions and documentation for Steve to trial and design the database structure.
9. *January-February 2001* - completion of all taxonomic descriptions and documentation.
10. *April 1st 2001* - Third Year Report on standard EU forms PART A and B.
11. *July 1st 2001* - Final Report.

END

Third WP2 workshop: 6 November 2000, CEREGE.
Minutes not available.

Minutes of Final EDDI meeting – 23/24 July 2001, UCL.

Participants:
Rick Battarbee, Helen Bennion, Nigel Cameron, Viv Jones – UCL
Steve Juggins, Richard Telford – Univ Newcastle
John Anderson – Univ Copenhagen
Dave Ryves – GEUS
Francoise Gasse, Francoise Chalie– CEREGE
Andy Lotter - Univ Utrecht
Sonja Hausmann- Univ Bern
Don Charles- Academy of Natural Science, USA
Phil Barker- Univ Lancaster

General thoughts on content/structure of EDDI web pages

- Rick pointed out that a link was needed between the taxa and the samples and then ultimately to the slides so that users could locate specimens if necessary.

- A disclaimer is needed to explain that data are presented in the way that they were given to the EDDI project and that input data (especially environmental data) are of varying quality and quantity.

- A front page is needed that explains the purpose and scope of EDDI. We need to stress that this is not a taxonomic database.

- A paragraph of text is needed on the nature of surface sediment samples, depth intervals, coring methodology.

- Examples of the benefits and pitfalls of using the EDDI combined/new training sets need to be included.

Missing data

1. **Taxa with missing authorities**
Steve circulated a list of taxa that have no authority listed in Diatcode. These need to be completed.
Action: Helen to arrange for authorities to be assigned where possible using van Landingham and send list back to Steve.

2. **Taxa without taxonomic/merging descriptions**
Steve circulated a list of taxa where descriptions are still needed.
Action: Dave, Nige, Francoise to provide these as Word text files (in same format as for other taxa).

3. **Missing sample information**
- African sampling dates
Action: Francoise to chase up and send to Steve.
- Italian ALPE sampling dates and ALPE units for Aluminium and alkalinity data.
Action: Nige to chase up and send to Steve.
- Clarification of Finland and Kola water depth data. Are these actual max depths of the lakes or coring depths?
Action: Helen to email Atte/Jan and Nadia to clarify.

Images
Image selection – the best image of each taxon to appear as the front page needs to be chosen.

90

Action: Steve will send list of taxa with images to Francoise, Nige & Dave who should then go through their list and pick the best image for each using the Taxon search on the EDDI web page.

Scale bars –There are no scale bars on the UCL images.
Action: Nige to let Steve have pixels per micron information and to highlight those images taken at different magnifications (e.g. large Pinnularia spp.).

Outstanding tasks on WP1 and WP2 merges
WP1 Merge files
Steve noted inconsistencies in the use of the taxonomic confidence codes.
Action: Steve to change Dave's "0" codes to blanks and to change Christine's "1"s to "0"s.

Final check/tidy up needed of WP1 merge lists
Action: WP1 merge lists were given to Dave, Nige & Francoise for final checking. To be returned to Steve by Friday 27th July.

WP2 Merge files
These will need to be finalised following completion of WP1 files above.
Action: Steve to send new version of WP2 merge list to Dave who will check and tidy.

Feedback on system from EDDI contributors
* Feedback needed on how users would like the dataset merging options to function.

* Comments from contributors needed on Dataset ID, Dataset Title, Contributor, Taxonomist and Contact name for each EDDI Dataset.

* Comments also needed on the text descriptions of each dataset as written by the co-ordinators.

Action: Steve to announce the EDDI Web Page details to all EDDI participants. User trials to be conducted between now and mid August in preparation for launch of upgraded system at the Aix meeting on 23 August. All contributors to feedback with comments on the above.

Data access
Three-levels of data access were identified

1. Open access now = view all data on-line, and download diatom + env data files (as for SWAP + any others?).

2. View all data on line (all diatom percentages + mean & individual chem values). No download of data files for 2 years.

3. Use data for reconstructions but no viewing of raw diatom percentages or chemical data on-line or downloading of data for 2 years.

Each dataset needs a text introduction on restrictions.
Action: Helen to email EDDI contributors to ascertain at which level they wish their data to be accessed.

Compatibility with the US DPDC system

1. It was agreed that a link with DPDC and plans for future compatibility should be mentioned in the Final EDDI Report.
2. Issues such as Taxonomic harmonisation dictionaries, standard data formats for inputting and outputting data and common environmental data formats (names/units) should be addressed to maximize compatability and ease of use.
3. We also discussed the possibility of linking images and other aspects of the two datasets. The possibility of looking at cross-Atlantic differences and similarities in distribution and ecology of diatom taxa, making use of data in both the EDDI and the DPDC databases, was also discussed.

Action: Steve and Don to liase on image links, table structures, units etc once Steve has finished documenting the EDDI system.

Reporting

• The format and reporting requirements were unclear.

Action: Rick to contact Hans Brelen in Brussels to clarify reporting format, need for Technical Implementation Plan, and contractual/financial reporting guidelines/time schedules.

• The preferred option is a Brief Report to the EU to accompany the Web Pages and CD-ROM. The report would be a c.20 page document including an Executive Summary and extracts from the web page text, plus a few images and tables. The report needs to include a description of problems encountered.

Web page text to be produced

1. Taxonomic descriptions to be completed
2. Dataset titles, one sentence description and a paragraph description for each dataset. This should state whether the dataset has ever been published as a "training set" and if so should give the transfer function performance statistics. Steve will add summary table of diatoms (number of taxa etc) and summary chemistry (N, min, mean, max for each variable), plus N, min, mean max for lake depth & area automatically.

Action: John/Helen to write an example version of a typical dataset description for the Northern Ireland dataset. This will be circulated to the others so that the format can be followed by all.

3. The "About" section will provide an overview of EDDI and should include:
- Overview
- Source datasets (list, descriptions, map)
- Methods (taxonomic harmonisation workshop details, N taxa, worked example; use of surface sediments, coring)
- Image capture (working v RBGE images, N taxa, N images, microscopy, quality etc)
- Results (i. environmental data merges, ii. taxonomic merges, iii. numerical procedures).
- **Action: Rick to improve the original draft outline of the text. Others to fill out relevant sections.**
- **Dave to write section on taxonomic harmonisation methods (just a few paragraphs including nomenclature used).**

DEADLINE FOR ALL WEB TEXT TO STEVE IS FRIDAY 10th AUGUST.

Publications

1. Global level species distributions/biogeography of European diatoms – Richard Telford to take the lead on this.

2. Global level transfer functions – value added by merging, with comparison of different methods and an example application building on the TP paper presented in Canada for all variables. – Steve to take the lead on this.

3. Application of EDDI to reconstruct trends in TP in European lakes – multi-author with Helen & John to take the lead.

Other suggestions:
- Technical paper on the Local WA method – Steve (single author?)
- Taxa responses in relation to env gradients, physical factors and geographical distributions (ecologically meaningful?) – more detailed for fewer taxa than paper 1 above.
- European v USA species distributions.

Guidelines will be needed. It was suggested that 2-3 consortium papers are published (i.e. 1 and 2 above) and then others wishing to publish EDDI data should email the EDDI consortium for approval. An EDDI list server could be set up for such exchanges.
Action: Invite other EDDI contributors to participate in compiling a list of potential paper titles and authorship suggestions, as well as views on guidelines for EDDI publications.

Future additions to EDDI
- It was agreed that existing raw data in EDDI could be edited in the future where necessary. Update notes would then have to be distributed to notify of any changes.

- It was agreed that new datasets could be added but the onus would be on the contributor to ensure harmonisation with EDDI taxonomy.
END

ANNEX 5 Technology Implementation Plan 1

TECHNOLOGY IMPLEMENTATION PLAN

PART 1 – Project Identification

*A Framework for the further development and exploitation
of the results of EC RTD Projects*

DOCUMENT TITLE:		
Technology Implementation Plan		
Part 1 – Project Identification		
DATE: 01/04/00	VERSION: FP4 2.2	ORIGINATOR: European Commission

Part 1 : Project Identification

Mandatory

1.	EC Programme:	Framework V
2.	Project title & acronym:	European Diatom Database: An information system for palaeoenvironmental reconstruction (EDDI)
3.	Project number:	ENV4-CT97-0562
4.	Consortium details:	Environmental Change Research Centre University of Newcastle Centre Européen de Recherche et d'Enseignement de GéoSciences de l'Environment (CEREGE) Geological Survey of Denmark & Greenland
5.	Number of results submitted :	1

ANNEX 6 Technology Implementation Plan 2

TECHNOLOGY IMPLEMENTATION PLAN
PART 2 – Project Results
A Framework for the further development and exploitation
of the results of EC RTD Projects

DOCUMENT TITLE:		
Technology Implementation Plan		
PART 2 – Project Results		
DATE: 01/04/00	VERSION FP4 2.2	ORIGINATOR: European Commission

Please give information on each of the results chosen for a specific exploitation route. Refer to the guidelines for further details.

Table 6.   Summary of exploitable result

Mandatory

This information is for administration purposes only and will not be published.

Summarise exploitable result, identify the partners (result owners) involved and describe the exploitation intentions

Title of Result	Lake-water quality reconstruction using diatom-chemistry transfer functions
Partners involved	ECRC-UCL, UNEW-DGEOG, GEUS-DK, CEREGE-FR, plus contributing laboratories
Exploitation intention	The data and techniques needed to perform water quality reconstruction using diatoms will be freely available on the web-based information system developed by the project
Category	☐ Exploitable result used only within consortiums ☐ non exploitable result ☒ exploitable result of interest for third parties

(you can use free text in each table cell, but be as short and to the point as possible. In the **Category** cell tick the appropriate box, one box only)

7. Summary (200-300 words maximum)

Mandatory

CONFIDENTIAL
No
Select Yes/No from dropdown menu

Provide an overview of the result which gives the reader an immediate impression of the nature of the result and its relevance and potential!

The project has brought together many disparate diatom datasets from across Europe and Africa, harmonised the taxonomic and environmental information associated with them and generated a single, comprehensive training set that can be used to perform high quality, standardised pH, total phosphorus and salinity reconstructions for European and African lakes. The web-based system also provides diatom photographs with taxonomic and ecological information and the software tools needed to carry out water quality reconstruction. It is now possible for diatomists in any laboratory to use a methodology that was hitherto restricted to only a few specialist laboratories.

* - insert the number of the specific exploitable result

8. Description of result
Mandatory

The main product from the EDDI project is the web-based information system for diatom-based water quality reconstruction providing a service to palaeolimnologists world wide, and relevant to scientists interested in environmental monitoring (annex 5 code C08), ecosystem modelling (annex 5 code C04), pollution abatement (annex 5 code C16), and climatology (annex 5 code C06).

This product represents the state of the art in the field of palaeolimnology, and is a culmination of over 20 years of research by the partners and associated laboratories in diatom-based reconstruction linked to problems of surface water acidification, eutrophication and climate change.

Categorise subject description using codes from Annex 4.

Subject descriptor codes	C09	C10	C20	C06

9. Current stage of development

STAGE OF DEVELOPMENT	Select one category only (tick the box)
Basic research	☐
Applied research	☐
Experimental development stage (Laboratory prototype)	☐
Prototype/demonstrator available for testing	☒
Results of demonstration trials available	☐
Other: (Please specify!)	☐

Briefly describe the current status/applications of the result!

The web system is currently in place with all basic statistical options available. The system is being tested by EDDI diatomists and there are plans to update the system and provide additional options in future.

10. Documentation and information on exploitable result

Add here a list of the **most important and relevant** information and documentation, indicating the confidentiality status of each document. The "document status" box indicates whether the document being referred to is confidential (and might be made available to third parties only after the signing of a confidential disclosure agreement). The "confidentail" box indicates, whether the knowledge that a document exists is in itself confidential.
Add promotional material that can be used for illustrating the result in dissemination services such as photographs, items of artwork, video clip, interviews, piece of animation, etc.

Documentation type	Document Status **PU**=Public **CO**=Confidential	Details (Title, ref. number, general description, language)	CONFIDENTIAL Select Yes/No from dropdown menu
article	PU	Battarbee, R.W, Juggins, S., Gasse, F., Anderson, N.J., Bennion, H. and Cameron, N.G. (2000). European Diatom Database (EDDI): An Information System For Palaeoenvironmental Reconstruction. European Climate Science Conference, Vienna City Hall, Vienna, Austria, 19-23 October, 1998, pp. 1-10.	No
			No
			No
			No
			No
			No
			No
			No
			No
			No
			No
			No
			No
			No
			No
			No

Part 2b: Exploitation of result

11. Exploitation strategy for the specific result
Mandatory

11.1 Using the table below, indicate the intellectual and industrial property rights being exploited (all foreground and possible background rights)

CONFIDENTIAL

No

Select Yes/No from dropdown menu

Type of IPR		Details (what is covered, reference numbers, countries covered) for all IPRs indicated in the Foreground (FG) and/or Background (BG) fields.	Number Fore-ground IPR's	Number Back-ground IPR's
Patent applied for	FG			
	BG			
Patent search carried out	FG			
	BG			
Patent obtained	FG			
	BG			
Registered design	FG			
	BG			
Trademark Applications	FG			
	BG			
Copyrights	FG			
	BG			
Secret know-how	FG			
	BG			
Other – Please specify	FG			
	BG			

Please enter in the "Details" field the information for **all** the IPR's. If you have more than one IPR per type (e.g. more than one patent), indicate in the "Nr of Foreground IPR's" and/or in the "Nr of Background IPR's fields" the respective numbers.

11.2 Define the role of each partner and the co-operation between the partners involved in the exploitation

The web-system is designed to be free standing and self-explanatory enabling third parties full use of the system without the involvement of the consortium partners. However, where help and advice is needed UNEW-DGEOG will provide assistance with numerical methods, ECRC-UCL with the taxonomy of diatoms found in acid waters, GEUS-DK with the taxonomy of diatoms found in eutrophic lakes, and CEREGE-FR with the taxonomy of diatoms found in saline lakes.

11.3 Collaboration sought

If you are looking for support by third parties, please indicate by using the keys or boxes below

KEY "Collaboration Sought"

R&D	☐	: Further research or development	JV	☐	: Joint venture
LIC	☐	: Licence agreement	MKT	☐	: Marketing agreement
MAN	☐	: Manufacturing agreement	FIN	☐	: Financial support
C	☐	: Venture Capital/spin-off funding	PPP	☐	: Private-public partnership
INFO	☐	: Information exchange	Other	☐	: (Please specify below)

Other:

Describe the exploitation opportunity that you can offer your potential partner.

Not applicable - see above

12. Exploitation activities and timetable
Mandatory

Describe the exploitation activities, the milestones involved and give a timetable (what will be done by whom and when?)

Not applicable - third parties are free to exploit the product that is openly available

Timetable:

Activity	Partner(s) involved	starting from ... to ...
		to
		to
		to
		to
		to
		to
		to
		to
		to
		to

13. Exploitation potential*

When describing the exploitation potential, you might want to consider one or all of the following factors:
- What are the potential applications for this result?
- Who are the users of this result?
- What are the main innovative features and benefits (technical/commercial success factors)?
- Analysis of the market sector
- Potential barriers

for PROSOMA users and those providing commercially relevant results, please concentrate on describing the business opportunity of your result

The product can be used by environmental consultants interested in lake-water quality and by research scientists interested in the dynamics of freshwater ecosystems. To use the product such scientists need be highly trained diatomists with access to high quality light microscopes and with competence in the numerical methods used in environmental reconstruction.

Categorise market application sector using codes from Annex 5.

Market application sectors	C04	C08	C16	C21

14. Ability of partners to carry out the exploitation

When describing this part, you might want to consider one or all of the following factors:
- Estimate the investment and describe the skills which will be required for exploitation of the result
- How do you intend to finance these investments?
- What is the expected return on investment?
- What risks are involved?

If you seek additional partners, clearly describe your input and the expected input from the external partner(s)!

New users of this product need to invest in the appropriate computing and light microscopic equipment needed and should consider attending relevant training courses in diatom analysis and in the numerical analysis of environmental data. The possibility to attend such courses is offered each year by the Environmental Change Research Centre, University College London.

15. Contact person for this exploitable result
Mandatory

Name	Professor R.W. Battarbee
Position	Director
Organisation	Environmental Change Research Centre, University College London
Address	26 Bedford Way, London WC1H 0AP
Telephone	44 (0)20 7679 7582
Fax	44 (0)20 7679 7565
E-mail	r.battarbee@ucl.ac.uk

16. Organization information
Mandatory

Provide a short description of your organization and if necessary, provide contact details on persons who are more involved in the exploitation aspects and/or the technical aspects.

ECRC research focuses on studies of aquatic ecosystem change and climate change. In particular it uses palaeolimnological methods, especially diatom analysis to reconstruct past changes in lake-water quality. General enquiries and those concerning pH datasets should be directed to the ECRC, but other key contacts include:

1. Dr Stephen Juggins for numerical methods: Dept of Geography, University of Newcastle, Claremont Road, Daysh Building, Newcastle Upon Tyne, NE1 7RU, England.

2. Professor John Anderson: Department of Geography, University of Copenhagen, Øster Voldgade 10, DK-1350 Copenhagen K, Denmark,
and Dr. David Ryves: Geological Survey of Denmark & Greenland, Environmental History & Climate Department, Thoravej 8, Copenhagen, Denmark,
for TP datasets.

3. Dr Francoise Gasse for salinity datasets: Centre Europeen de Recherche et d'Enseignement de GeoSciences de l'Environnement, Europole Mediterraneen de l'Arbois, B.P. 80, Université de Aix-Marseille III, 13545 Aix en Provence, Cedex 4, France.

17. Authorisation
Mandatory

I confirm that the information contained in the Technology Implementation Plan which is marked **CONFIDENTIAL / NO** may be disseminated by the Commission :

Name: Date:

Organisation: