First Progress Report

1 February – 31 July 2001

Editor : B. Le Dantec – ERCIM
Remi Ronchaud
2004, route des Lucioles
B.P. 93
06902 Sophia Antipolis
France
Bruno.le_dantec@ercim.org
Remi.ronchaud@ercim.org

Covered period: From 1 February to 31 July 2000

Project Scientific Manager
Name: Umberto Straccia
Address: Istituto di Elaborazione dell'Informazione
Consiglio Nazionale delle Ricerche   Via Moruzzi,1 , 56124, Pisa (Italy)
Phone: 39. 050 315 2894   Fax: 39. 050 315 2810
e-mail: straccia@iei.pi.cnr.it

Management report editor
Name: Bruno Le Dantec
Address : 2004 Route des Lucioles - BP 93 - 06902 Sophia Antipolis cedex - France
Phone: 33 4 92 38 50 13   Fax : 33 4 92 38 50 11
e-mail: bruno.le_dantec@ercim.org

1.   EXECUTIVE SUMMARY

We report on the initial six months of the Cyclades project. The work accomplished and the accompanying activities concerning this project during this period are reported.

We report on the scientific progress of the ideas started out in this project and pursued by the partners. The initial work has used the project web pages to display and communicate the ideas exchanged which lead to the first reports. Management activities are reported concerning the dissemination activities pursued by the partners.

1.1 Project structure

Project organisation and synergy have been initiated through several meetings attended by all partners. All advances of the projects and results are recorded in the common project web pages, http://www.ercim.org/cyclades, which contains all reports, working documents and deliverables as any useful information. Such a system facilitates communication between partners, each of them being aware at any time of the state of the project.

The project structure is very well documented in the project programme. Both the scientific and management structures of the project are described in every detailed as required

1.2 Scientific Progress

The goal of the first six months of the project was to :
- specify the end users requirements : needs, expectations and
- Design the Global system Architecture
- Define a Preliminary exploitation Plan
- Coordinate Cyclades with the Open Archive Initiative

These interdependent tasks which are now achieved enable the project to move forward and towards the implementation phase. A lot of concertation within the consortium was necessary to provide choices compatible with the end users interest, the technical constraints and the system potentialities.

1.3 Results

Deliverable produced:

- D1.1 User Requirements Report
- D2.2.1 Global System Architecture Report
- D7.1.1 Web Site
- D7.2.1 Dissemination and Use Plan
- D8.1.1 Co-ordination Report and Preliminary Version
- D9.1.1 Quality Assurance Plan

2. MANAGEMENT

There have been various activities at both the consortium and local level of each institute.

2.1 Meetings

The project management has organised the kick-off meeting which took place in Pisa, Italy on the 19 February 2001. The meeting was very constructive and well organised from the point of view of work achieved. The emphasis was on defining precisely all the tasks assigned with a particular attention for the definition of the end user requirements that stand for the basis on which will rely all the work afterwards.
At the same time, partners involved in the system design and in the collaboration with the Open archive initiative were also given short and early deadlines to give the project a certain reactivity. The agenda of the Kick Off Meeting is to be found below:

Agenda of the CYCLADES Kick-off Meeting – Pisa, Italy – 19 February 2001

Participants
Donatella Castelli (IEI-CNR)
Gudrun Fischer (UniDo)
Norbert Fuhr (UniDo)
Tom Gross (GMD)
Lino Pagano (IEI-CNR)
Dimitris Plexousakis (FORTH)
Rémi Ronchaud (ERCIM)
Fabrizio Sebastiani (IEI-CNR)
Umberto Straccia (IEI-CNR)
Costantino Thanos (IEI-CNR)

9.00 – 9.05	Welcome (C. Thanos)
9.05 – 9.45	Partner Presentations
9.45 – 10.15	Managerial and Administrative procedures (ERCIM, CNR)
10.15 – 10.45	Project presentation (objectives, activities, responsibilities, schedule) (CNR)
10.45 – 11.15	Coffee Break
11.15 – 12.00	The Open Archives Initiative (OAi) and coordination with Cyclades (CNR) Discussion
12.00 – 12.45	User Requirements (FORTH) • Plan of Actions Discussion
12.45 – 14.00	Lunch
14.00 – 14.45	System Specification (UniDo) • Plan of Actions Discussion
14.45 – 15.15	Dissemination and Exploitation (CNR) • Plan of Actions Discussion
15.15 – 15.45	Coffee Break
15.45 – 17.30	Each partner should briefly explain their preliminary ideas on the functionality to be provided by the system: Mediator Service (FORTH) Access Service (UniDo) Collection & Personalization Service (CNR) Recommender & Collaborative Service (GMD) Search & Browse Service (UniDo) Discussion
17.30	End of the meeting

During this meeting, the partners had their first opportunity to investigate precisely the complementarity of their respective tasks and exchange ideas with each other about the system design and about Cyclades's cooperation with OAI.

Discussions also took place and every effort was made from all partners to inform the rest about their work, experience and expectations within this project.

All the presentations are available for download from the dedicated web pages (using Member / Cyclades) at http://www.ercim.org/cyclades/vip/.

Then from the 7th to the 11th the of May 2001, a Technical Meeting took place in Crete. The core of the discussion was to determine more precisely the System Functional Specifications. This meeting proved very useful since a general agreement was found among partners and the general motivation of the Cyclades team was boosted up.

The minutes are also to be found on the web site but the real conclusions of this meeting are described in D2.2.1, the Global System Architecture Report.

2.2 Web Server

The official Web server was set up at the ERCIM address.

http://www.ercim.org/cyclades,

This web sites have been created to dispatch various information related to the project. It includes a restricted area for the Consortium, mainly used as document repository, and a public section. The site has been opened for public access in January 2001 and is regularly updated.

Particular effort was made by the management to provide the consortium with the best work environment possible. On top of this regularly updated web site, all the partners has also access to the BSCW workspace. The BSCW (Basic Support for Cooperative Work) enables collaboration over the Web. BSCW is a 'shared workspace' system which supports document upload, event notification, group management and much more.

This has proven extremely useful when it came to exchange ideas, representations of the system architecture or any other document and deliverable in progress. The final version of these documents is put on the regular web page as the BSCW is a mere workspace.

2.3 Presentations in Conferences and Exhibitions and Public Relations

Partners were made very sensitive to the importance for the project to be represented in various public events. In that perspective the management has been informing all the partners about the coming conferences, workshops or meetings related to the field of Digital Libraries. Below are some of the events in which Cyclades was being represented:

- ERCIM's and CNR's contacts with the DELOS network of Excellence were a real support to bring Cyclades to take part to major DL events.

- EU-DL All Projects meeting, Luxemburg, February 2001

- DELOS, IPI RAS and IITE Workshop on European/Russian Cooperation in Digital Libraries, Moscow, June 2001

- ECIR, Darmstadt, April 2001

- Participation on a panel on 'e-Democracy held on the 25 and 26 June 2001: A Panacea or Pandora's Box?' at the 14th Bled Electronic Commerce Conference 'e-Everything: e-Commerce, e-Government, e-Household, e-Democracy'.

- OAI Open Meeting, Germany, 26 February 2001

These meetings and conference were the opportunity for a representative of Cyclades to discuss and promote their project in the context of DL research and virtual communities.

Effort was also made by the management to tighten the links between Cyclades and DELOS. Indeed the aim of the DELOS Network of Excellence for Digital Libraries (DLs) is to provide an open context in which an international research agenda for future research activities in the digital libraries domain can be developed.

This Network constitutes a reference point for all DL projects funded by the 5th Framework Programme, stimulating the exchange of experiences and know-how in this multidisciplinary domain, and will also establish close contacts with relevant application communities.

Therefore it appeared as a necessity for the management to integrate Cyclades into this framework DELOS provides.

2.4 Publications

Several publications were made to promote the project within the European and international scientific DL research communities. Among those publications, the **ERCIM news** magazines offer an valuable opportunity to disseminate information to the targeted scientific communities.

Advertising through ERCIM news (7000 copies) or through other specialised bulletins provides an excellent vector for dissemination. ERCIM News reports on joint actions of the ERCIM partners, and aims to reflect the contribution made by ERCIM to the European Community in Information Technology.

Through short articles and news items, it provides a forum for the exchange of information between the institutes and also with the wider scientific community. ERCIM News is published quarterly. The printed edition has a circulation of over 7000 copies.

Cyclades was mentioned in ERCIM news in a dedicated article written by Donatella Castelli, "Scholnet and Cyclades: two new digital Library projects", ERCIM News, April 2001;

Other publications were also enlisted:

- An early version of the questionnaire devised for the requirements collection was communicated to the Open Archives Initiative working group who provided input that helped shape the questionnaire's final version.

- Donatella Castelli, Pasquale Pagano and Umberto Straccia, Scholnet and Cyclades: extending the role of digital libraries D-Lib Magazine, May 2001

- Finally, a specific Cyclades leaflet has been designed and edited as a handout to support the promotion of the project in various conferences

3. ACTIVITIES AND ACHIEVEMENTS

3.1 ERCIM

Activities

ERCIM has been paying a special attention to starting the project on optimal grounds. The partners were provided with various tools to help them in carrying out their tasks throughout the project. The tools rank from guidelines for writing reports and deliverables to the BSCW workspace on the internet.

The next important move was to prepare a efficient **Kick Off meeting**. From ERCIM's experience in other projects some partners are leaving this meeting with no clearly defined tasks

with less accurate deadlines. In Cyclades, the management made strict recommendations to the partners and by the end of the kick-off meeting, all partners had actions to fulfil within a short deadline. By getting the team to work rapidly and, most importantly, jointly on the project, the management has been able to trigger extensive collaboration among partners while giving the whole project the impulse to solve the pending issues rapidly: definition of the expected output, user requirements, interaction between the work-packages, information flow in the consortium, etc…

Having focused on improving the communication flow, ERCIM has also improved the self-assessment of the Project. The various documents due to the Commission are presented to all partner on the BSCW site for reviewing. Any document can be read, opened or modified. Once all the relevant modifications have been done, the BSCW keeps track of all actions for every document. If a Deliverable has been read by a partner, the BSCW platform can provide the name of the Partners, the date at which he read the document and if any modifications were made.
This is all the more useful as some partners are still not familiar with the Commission documents.

Regarding WP9, ERCIM has brought an additional degree of quality in the work achieved through the evaluation and review of all the strategic documents. Some of these document being necessary for some partners to move onwards in the project, it was WP9's priority to ensure that such documents fulfilled their role, especially internal roadmap documents that are the cornerstones to the project.

In that perspective, the quality of each document produced was reviewed not only by the work-package leaders and the co-ordinator, but also by the person using this document. If it was not regarded as satisfactory, the document was sent back to its author, precising what elements needed further details.

This system, operational from the beginning of the project, was applied to all documents, especially to Deliverable that needed to be reviewed and assessed by all key players of the project before being sent to the European Commission.

As for WP7, ERCIM has rapidly gathered all of the Cyclades partners to think about a coherent dissemination plan, and exploitation plan. Every partner was asked to hand in its contribution to the Preliminary Exploitation Plan, precising the expected results of the project and describing the way and means to exploit these results.

Achievements

ERCM has provided the project with extensive dissemination tools (ERCIM news, leaflets…).

All Deliverables and Reports have been achieved and reviewed by the partners. They will give all the required information on the project management, the main achievements in every WP and define more precisely the general course of action for Cyclades.

Regular efforts were devoted to the collection of every partner's contribution to the reports and deliverables. Final corrections to these documents were realised before being sent to the European Commission.

The management integrated several elements to realise its deliverables that were sent to the European Commission, only after a general partners' assessment of the reports. The management was also gathering the various detailed Cost Summary to hand in the Cost Statement consolidated figures. This task was made all the more difficult as some partners had had some internal changes (rates, internal rules…) and as some corrections had to be made to match the European commission's standards. Nevertheless, the elements were gathered and consolidated in time to produce the first Cost Statement.

ERCIM has also been involved in ensuring a close follow up of the scientific work by setting up a technical meeting for the partner to take a final decision of the system functional specifications.

Additional efforts were made by the Management to strengthen the links between the various Workpackages. In that sense, the Technical Co-ordination followed closely the achievements of each WP individually and brought every partner to further interact with the different WP leaders accordingly the work programme.

A precise list of Actions was set up after each meeting, gathering all the pending tasks and is updated regularly to avoid any delay. This procedure gives the co-ordination a better visibility of the project and allows every partner to spot the progress of every task and WP.

On top of this, a particular attention was put on the cooperation with other DL oriented projects such as Scholnet (IST-1999-20664). The outcome is a better follow up of their activities that allows improved synchronisation of efforts with European Partners. This initiative is followed by a certain implication in the DELOS network of excellence activities.

While supervising the realisation of the deliverables, ERCIM has produced the following :

- D7.1.1 Web Site
- D7.2.1 Dissemination and Use Plan
- D9.1.1 Quality Assurance Plan

ERCIM has also slightly contributed to D8.1.1 "Co-ordination Report and Preliminary Version" presenting the Cyclades interactions with the Open Archive Initiative.

3.2   CNR

CNR is actively involved in all the Workpackages, with a particular attention to WP 3, 7 and 8.

The CNR team has been ensuring the scientific management of the entire project and has been led to take some strategic decisions. The consortium agreed on some basic issues on the software to be developed within CYCLADES. Indeed, we agreed that all software should rely on open source or freely available software, that is no pay software should be included with CYCLADES. Main programming languages will be Perl and Java. Concerning the background components provided by partners, they will be made available to all the partners for free.

CNR also contributed to the definition of the Global System Architecture (Deliverable D2.2.1).

CNR has also just started the implementation phase of the System Development of WP3. This Workpackage having to start at t+6, CNR has already initiated an internal roadmap of its activities in that respect.

As for WP8, CNR is establishing closer relations with the Open Archive Initiative. The user requirement questionnaire has been sent to the several communities, including the Open Archive Initiative (OAI) and, thus, making the Project goal publicly available. The Co-ordination Report Preliminary version has been produced and sheds light on the on-going interactions with the OAI.

Indeed, as the questionnaire of WP1 was sent to the Open Archives Initiative group, a valuable feedback was accommodated in the final version of the questionnaire. A first meeting has been held in July 2001 with Carl Lagoze, the representative of Cornell University and heavy involved in the OAI, discussing about the project that finally has been approved by NSF.

 We expect to closely collaborate and a human resource exchange program is under the study.
The current version of the OAI protocol, Version 1.1, has been presented to the partners. A metadata gathering activity has been started to determine some statistics about the to be gathered data from the OAI compliant archives The report of this gathering process is submitted to a OAI Workshop and maybe considered as a feedback activity. Subscription to the OAI mailing list has been performed.

3.3 FORTH

Achievements

The objective of Workpackage 1 was to identify relevant user communities, select a user requirements collection and analysis methodology and collect and analyse user requirements regarding the CYCLADES environment.

The first task was the identification of relevant user communities. The following scholarly communities were identified as such :

- the German Society of Physics
- the German Society of Mathematics
- the Italian Society of Mathematics
- the Delos Network of Excellence in Digital Libraries
- the PLANET Network of Excellence in Artificial Intelligence Planning
- the Open Archives Initiative group

Academic and research groups among the partners' institutions have also been consulted.

The next step consisted in creating the specific questionnaire for eliciting requirements from prospective users of a collaborative archive environment. All partners contributed to the formation of the questionnaire by specifying appropriate questions for the different functional modules of the CYCLADES environment. A goal-directed approach was followed in devising the questionnaire.

Considering the different disciplines of the user groups we intended to address with our questionnaire, an abstraction of technical details was made in order to produce a set of questions that sufficiently covered the entire system functionality, but which were presented in understandable terms by the scientific community at large. The questionnaire comprises a total of 34 questions distributed as evenly as possible among the different services. A particular effort was made to keep the questionnaire as succinct as possible.

The questionnaire also included a concise description of the CYCLADES objectives and of its different service components. Several versions of the questionnaire were produced, refined as well as modified by the CYCLADES partners. As mentioned above, the questionnaire was also disseminated to the Open Archives Initiative group, whose valuable feedback was accommodated in the final version of the questionnaire. The final version of the questionnaire was included in Appendix A of Deliverable D1.1 (User Requirements Report).

It was decided to electronically disseminate the questionnaire in order to avoid delays with mailing printed versions. The questionnaire was made accessible via the World-Wide Web on a site hosted by ICS-FORTH.
(URL: http://sappho.ics.forth.gr:5000/User_requirements/questionnaire.html). Solicitation letters were sent by the partners to the above mentioned communities. A total of 34 answers to the questionnaire were collected through the web site. The analysis of the results was presented in Deliverable D1.1 (User Requirements Report). The complete set of answers and related statistics were included in Appendix B of the deliverable.

The resulting requirements were used as input in Workpackage 2 for devising detailed functional specifications of the system while taking into account user needs.

The consortium agreed to give a short extension to the period initially assigned for the collection of responses to the questionnaire so that the number of responses could be increased. This extension did not result in any delays for the Workpackage completion.

Within WP2, FORTH undertook the task of specifying the Mediator Service component of Cyclades, following the agreed upon methodology. The task of this component is the provision of a homogeneous interface to the provided services and the integration and dispatching of services within Cyclades. FORTH proposed an extensible architecture for the Mediator component and a simple interservice communication protocol.

3.4   GMD

GMD is the leader of Workpackage 4, starting at t+19. Nevertheless, GMD has contributed to the on going tasks Cyclades, especially in the definition of the User Requirements and in the definition of the system functional specifications.

Achievements:

User requirements

Analysis of user requirements by means of a Web-based questionnaire; in particular, the feedback of future users brought insight concerning the recommendation service, the rating service, and the collaborative work service

Shared glossary

Elaboration of a common terminology in the form of a shared glossary for the CYCLADES project

Specification

Thorough discussion and specification of the system functionality; in particular, the functionality that will be provided by the rating management service and the collaborative work service concerning single-user activities (e.g., registering to the system, logging in, managing private folders and documents) and community activities (e.g., setting up a community, joining and leaving a community, awareness of the activities of the other community members, and communication with other community members). For these parts of system use cases were developed and activity diagrams were drawn. The overall architecture of the CYCLADES system was specified and the communication between the rating management service and the collaborative work service on the one hand and the other services on the other hand was clarified and documented by means of interaction diagrams.

But beyond this technical contribution to the project, GMD has offered a strong support to the dissemination activities by producing a specific leaflet in English and German for distribution at conferences as well as to visitors at GMD-FIT.

This institute is also promoting the project through the information found under the CYCLADES project GMD page linked to ERCIM (cf. http://fit.gmd.de/projects/cyclades)

GMD has also been active in its participation to a panel on 'e-Democracy: A Panacea or Pandora's Box?' at the 14th Bled Electronic Commerce Conference25-26/6/01. The leaflet produced were distributed and a specific target audience was successfully reached during this conference.

3.5   UNIDO

Achievements:

For work package WP2, the functionality of the proposed system was specified in more detail, taking into account the user requirements resulting from WP1.

Different user roles were identified and described, and for each role, the use cases were discussed in detail. Each partner developed activity diagrams to describe the user's interaction with the system. The use cases and diagrams will later form the basis for testing the functionality of the implemented system.

Based on the proposed behaviour of the system from the user perspective, the corresponding necessary system activity was identified. The resulting system actions were assigned to system components that will be implemented by the Cyclades services, and the most relevant classes were defined.

Then, the interaction between the resulting classes and services was specified in detail, also defining the necessary methods that will be called during service interaction. For each activity diagram, corresponding interaction diagrams were drawn. Thus, it was verified that the class definitions contain all the methods that have to be known between services in order to implement the functionality defined by the use cases and background activity descriptions.

The services thus identified are:

- **Access Service**
  This service will provide harvest-based gathering of metadata from different open archives, the possibility to register and unregister archives, and a basic search functionality (for internal use only).
- **Collaborative Work Service**

This service will implement a folder environment for the user to store, organise, rate, annotate and share metadata records and documents with her project partners and within a community.

- **Search and Browse Service**
  This service will provide a sophisticated environment for searching and browsing in virtual collections of metadata records.
- **Collection Service**
  The management of virtual collections of metadata records will be the purpose of this service.
- **Filtering and Recommendation Service**
  This service will provide personalised filtering both for queries and for query results, and recommendation of records, users, communities and collections to the user.
- **Mediator Service**
  The task of this service will be to provide a homogeneous outward appearance for the system, to integrate the other services, including any future service types yet to be defined, and to provide security.

For inter-service communication, different protocols were discussed. XMP-RPC was chosen, as it is already well established and tested. The communication will be implemented in a separate layer so that it might be changed in the future to some other protocol.

The suitability of various existing software tools as building blocks for the Cyclades system was discussed. BSCW will be adapted to function as Collaborative Work Service.

The GUI of the future system will consist of a graphical user interface for every service that has interaction with the user, and a general system GUI that will call and integrate the graphical user interfaces of the individual services.

The proposed core services of the Cyclades system are designed to run autonomously and work together via the Mediator service which will function as a registry. Replication is provided for in the current design, while distribution of services will not be supported yet. The possibility of service distribution was discussed at length, however, to make sure that the current design will be easily extendable to support distribution of services in the future.

4. OTHER

- GMD is merging with FRAUNHOFER and is officially changing its name (document 1)

- Within UNIDO, Sascha Kriewel was hired as a research assistant in July 2001 to reinforce the Cyclades team.

5. DEVIATION FROM PLAN

None.