

FIRST REVIEW REPORT
of
CYCLADES
by
Birte Christensen-Dalsgaard

A. Objectives, work plan and resources

The main idea behind the project, to combine services related to identifying and working with literature, in this case mainly the grey literature with a collaborative environment, is very sound and is certainly within the scope of the programme. If successful the project will provide a good demonstrator for creating services based on the Open Archive Initiative protocol. Further, the test of the relevance of integrating services like a collaborative environment and the recommender service into a resource discovery situation will provide insight into new services within the digital library community.

There partners seem committed to the project, which among other things is seen in the fact, that the project seems to be on schedule and is meeting its targets. Many projects seem to have a slow start and may catch up later. Also the interplay between to partners seems well established.

The workplan has been followed and no changes introduced.

The project gave no indication of resources as part of the project report and at the review numbers were presented for the first semester of the project. Of these only the number for WP1 and WP2 could be evaluated as we had no indication of % completed work for the others. Assuming both these two WPs to be finished indicated an over consumption. Since the numbers are quite small it is not enough to raise general concern, however if this trend continues, the project faces severe problems.

B. Approach, methods and results

The generally impression is, that the approach and the work done are sound and the consortium seems competent to do the envisioned tasks.

The scientific work being reviewed here fall in two categories:

- User survey
- Architecture

First I will address the user survey, which is where I have my main negative comments.

My concern relates to the user involvement and the identification of user needs. The questionnaire done is in my opinion inappropriate for the target group, which, according to the deliverables, include groups such as the mathematical research community. As an example it is stated (and I absolutely agree), that the questionnaire should be made simple. However, I doubt that a “normal” researcher will understand e.g. the term “Dublin core” and therefore I think the conclusion that Dublin core is sufficient, is at best misleading and at worst – wrong. Another example is the question related to client-side software. Exactly what is meant? Does this exclude doing the client side as a plug-in to the browser? If so, how can 73 % say they want to use an unmodified browser, considering they will have to install a different plug-in to read the articles, namely Acrobat reader.

Another concern is the identified categories. To me there seem to be two categories of users:

- The researcher, who wants to use the cyclades service
- Institutions such as libraries, which might want to install the software as part of their service towards e.g. the university environment.

The questionnaire seems solely to focus on the first group (somewhat in contrast to the envisioned exploitation through using the open source approach, where revenue normally comes through consulting).

Had the latter been included in a proper way, the project would no doubt have identified issues such as, that these communities wanted to have control over their users. A method to satisfy this requirement would be to support e.g. the use of a LDAP server for the user authentication. This surely would fit into the modular structure envisioned for the system.

Had the former been properly included the project would most likely have identified the importance of actually having access to the documents.

Now let me turn to the architecture:

The specification report seems sound, more worked through in some places than others, but generally sound. It is clear, that a lot of the functionality is taken from other services or projects. I think this a very good, but it would be nice to have an idea about the level of reuse of code, i.e. what is based on existing services/software components and what will be developed from scratch.

The modular approach to the system development seems quite sound. However I have a slight concern regarding the role of the Mediator. The Mediator is described as the centre for the environment. It is the Mediator, which controls the information flow and new services will have to be registered here. My concern is the scalability of an architecture, where all existing and future information and service exchange will be mediated via a single function. I think it will be

necessary to define API for parameter exchange between other parts e.g. as described in the Global System Architecture report, where there seem to be communication directly between e.g. the Collaborative work service and the Filtering and Recommendation service. (e.g. page 40).

Also, I have a slight concern about the use of two different development environments – Java and Python. These of course will function together, but it does put constraints on interoperability and slightly limits the usefulness of an open source product. I do understand the rationale behind this decision and favour in general reuse of existing code.

I assume, that the interface to the underlying databases will adhere to standards – and any of the databases used in the project in principle could be behind the whole complex of services.

Finally I am not sure it is a good idea to fully develop the system before it is tested with users. The concept of XP-development (eXtreme Programming), that is sort of rapid prototyping, is gaining momentum. The project might consider if the development can be done through several meaning phases – and then involve the users during part of or the whole development process.

C. Exploitation and dissemination

The project has made a good start towards dissemination. It has identified a number of meetings and conferences as well as a number of relevant publications.

To me the project has two main target groups for the exploitation, which should be reached via the dissemination: the Information specialist, who might be responsible for installing the Cyclades system, and the researcher, who will use Cyclades. The information research community reached through e.g. DELOS, is interesting as they are also researchers and information specialist – however the interest is wider – and the dissemination should take this into account. This means that Cyclades, when it has a demonstrator, should try to visit e.g. educational conferences and target articles to this forum.

A first attempt towards the exploitation plan is provided in terms of the useplan for each individual partner. These give a good explanation of how each partner plans to exploit their participation in CYCLADES. An exploitation plan is scheduled for the future.

D. Management

The management seems sound and efficient. The synergy between the partners seems to be excellent and there seem to be a good division of work between the partners.

No figures for the performance of the project were not provided as part of the progress report and were not part of the planned presentation at the review. At our request these were provided for the first period.

The consortium behind is competent and the expertise of the individual partners supplements each other. The consortium seems to work well together.

My only concern here is the interest and ability of the project to address other relevant communities such as the education market and interact with work in areas such as the pedagogical ideas for discovery learning. The project is encouraged to identify relevant projects in this area in order to receive relevant information. I am not encouraging the project to attempt to integrate these ideas in the development plans – just to keep up to date with the developments and potentially be influenced by relevant ideas.

E. Modifications and developments since the last review (if applicable):

Not applicable

F. View on project status:

My overall assessment is, that this is a worthwhile project with an interesting approach to establishing a search services based on the Open Archive Initiative. The idea of combining search with add-on services such as the recommender system and a collaborative environment is interesting to get tested.

The project has described a number of potential services in their TA, which sounds interesting. It will be interesting to see, if the recommender system will work – to me there seems very much to be a chicken and egg problem – to be useful you need many participants – but will they start using the service before it is running satisfactory? In other words, what is the critical mass for this service to be useful.

Since the project aims at identifying relevant new services, I think it is problematic that so little has been devoted to involve the users. The study done seems more like an academic exercise to confirm existing notions than a piece of work done because the project really wanted to investigate new territory and get new ideas. At the review the idea came up, that the project could involve users in a test phase midway. I strongly support this idea.

Based on experience with users of “my” own library university users will be very displeased with a service, which does not bring them directly to the relevant full text. This point has received little attention in the Cyclades project.

Based on experience from a Danish project creating a combined access service to electronic journals I can say, that Libraries want to control the user authentication and authorisation themselves. To be useful in this environment it is therefore necessary to support external authorisation and authentication e.g. via an LDAP server.

G. Community added value:

The project addresses the interplay between access to literature via the Open Standard Initiative, collaborative services and added value services related to discovering interesting literature. Furthermore it will bring the European dimension to the primary American based OAi group.

As such the project has a clear European dimension, both in its direct goals – to develop the system, and in its spin-off – to influence the future development within a standardisation initiative.

The partners involved are involved in several other EU-funded projects. Many are parts of DELOS. Therefore it is expected to draw upon and integrate relevant development in Europe.

H. Contribution to social objectives (if relevant):

Not relevant

I. Recommendations for future work

My recommendations have been given above as part of the description. To summarise, the project should (listed in the order they appear in the text):

Provide more information in connection with progress reports allowing an assessment of resource usage compared to actual achievement. Including a statement of % fulfilled and % resource usage could e.g. do this.

It is recommended to consider a more iterative approach to the system development. This will allow earlier testing and will sooner provide proof of concept.

Repair the inappropriate user survey with introduction of usability runs during the development phase.

Consider how to ensure access to the full text documents (or video or simulation or....)

Consider using standards for user authentication and authorisation.

Consider dissemination and exploitation towards e.g. the educational market and the general library community,

J. Overall recommendation

The overall recommendation is:

- **Continue** (*without major modifications of the work programme*)

Århus 7/2 2002

Birte Christensen-Dalsgaard