

WEEKLY STATUS REPORT
Team #07
Project Data Mining of Digital Library Usage Data
Week #01

Progress

Our progress includes organization and preparation for the restart of the project since the start of the spring 2005 csci577b semester. Our first task was an initial team meeting for introductions and a quick run through of the project. This was necessary because only one out of the four members of the team is a returning member from team 7. The team has been having at least 2 meetings a week since the semester which include the clients.

In terms of preparation for the RLCA, almost all documents have been updated and all changes accounted for. The SSRD was changed to meet updates on certain requirements, and the LCP was changed to update new cost estimates and new plans. The SSAD has also been modified to correct several UML diagrams to match our current system. The risk assessment was updated in the FRD and the DART tool for the team has been set up so that the team may identify and analyze risks.

The team has also read 9 research papers that were suggested to give some relevant information on data mining by the client. All papers were read by the team and also presented to the client during a client-team meeting.

Problems

One problem we had was that our 5th team member dropped the class after having been assigned to the OCD and other tasks as well. This resulted in bit of a delay as well as a redistribution of work and the taking on of roles.

Another problem was that last semester the client assumed that the resource click count would be available from the web server with respect to the digital archive. However, the client brought to our attention that this was not the case, but at the same time provided software that was already implemented that accomplishes this task instead. A sample print out of the raw data was distributed to the team by the client. This software need not be incorporated into our system, since it already makes the data we need to mine available. However, this causes some delay because the documents need to be updated to account for this minor change.

Risk Items & Problem Reports

This section will be updated since the team website is not yet operational.

Immediate Plans

- Implement team website
- Decided on type of algorithm (frequency or time) to mine data on
- RLCA draft due February 9th
- ARB is on February 16th

Comments

We need to decided on the algorithm, then decide what people will work on the algorithm and what people will work on the visualization portion of the project. The client also supplied us with the raw data of which the system will mine, and this should be used to think of more ideas on that mining algorithm and possible concerns/risks.

Client Meeting Notes:

Client meeting on 1/28/05 for 3 hours. This consisted of members presenting to the client and the team the papers that were assigned for reading. The data mining algorithm was also discussed as well as the overview and the purpose of the system.