

WEEKLY STATUS REPORT
Team #07
Project Data Mining of Digital Library Usage Data
Week #06

Progress

Our progress was in the areas of design and implementation. The algorithm for data mining was started, but how the data is stored internally must be changed to allow the algorithm to run on over 100,000 log instances. We split up the team into two groups, one group will work on the algorithm and the other will take on the task of clustering the output of that mined data.

We also had a meeting with Naomi, and presented our project to her. She was at USC for a metadata and information visualization lecture.

Problems

There were no notable problems for this week.

Risk Items & Problem Reports

<http://seacliff.usc.edu/~team7b/>

Immediate Plans

- Decide on algorithm storage choice, using a sparse matrix
- Work on clustering algorithm, research on whether to use MCL or create a new clustering algorithm
- Look into suggestions made by Naomi

Comments

No comments.

Client Meeting Notes:

A client meeting was held to discuss our plans for implementation and specific ideas for design plans. The meeting was about 2 hours long and the client approved of the development team's approach to the implementation. We also requested from the client specifications on how much processor power and memory (primary and secondary) is available for us in terms of being able to run the program on the target machine. We have to consider all these things when implementing so that we may maximize our resources and have the system run efficiently. We also brought up a possibility of running the complete program on some mainframe computer, because the mining is computationally expensive. Jewel said that she will look into this and get back to the development team.