

Tentative 10 Week Schedule

Week 1 : Get the group together. Get the cluster operational

Week 2 : Start collecting data and start building the database. Learn how to query the database.

Week 3 : Generate keywords from tweets: tf-idf, chi-squared, ANOVA, ...

Week 4 : Generate clusters of tweets based on their keywords: Expectation Maximization algorithm, BIRCH algorithm, ...

Week 5 : Generate reasonable summaries using an "expert AI" system

Week 6 : Generate reasonable summaries using HMM (Hidden Markov Models)

Week 7 : Other research Interests. i.e. finding trending topics, finding which tweets carry a negative or positive message, ...

Week 8 : Catch-up on anything, or take a vacation?

Week 9 and 10 : Make a poster/paper