**Summary**

The research project will involve the study of approaches from social web mining and social media analytics for recommender systems. The work will culminate in the development and evaluation of a system to recommend affinity groups to an individual user based on the user's profile and shared topical interests between the user and other users of a social network.

The primary data sources to be used will consist of data gathered from social annotation systems such as last.fm, delicious, flickr, and others. A final report will include the full description of the system and its evaluation, as well as a review of relevant scientific literature.

Contact hours for the project will include participation in weekly group meetings, as well as individual meetings with the project supervisor as needed.

**Data Sources**

* Linkedin
  + <http://helpnv10.qsrinternational.com/desktop/concepts/approaches_to_analyzing_linkedin_data.htm>
* Twitter
  + <http://www.infochimps.com/tags/twitter>
* Flickr
  + <http://www.flickr.com/services/api/flickr.photos.geo.setLocation.html>
* Wikipedia
  + <http://en.wikipedia.org/w/index.php?title=Recommender_system&action=history>
* Yelp
  + <https://www.yelp.com/academic_dataset>
* Instagram
  + <http://instagram.com/developer/endpoints/tags/>
* GoodReads
  + <http://packtlib.packtpub.com/library/9781847194442/ch02lvl1sec04>
* Citeulike
* Bibsonomy

**Scenarios**

**Recommendation Types:**

* A group is recommended to a user
  + They can either follow the group, like the group, or join the group
* A friend/user is recommended to the user
  + They can “friend” the recommendation

**S1**

* Recommendations are automatically generated and suggested to a user
* They are based off of the similarity between the user’s tagging of resources and a group’s tags
* This could be based off of the service’s interests or the groups interests
  + I.e.- Either the group is looking for people so it scans OR the service (Linkedin, Facebook, etc.) searches for ALL relevant groups for any user.
  + So, there is context provided by the user. The medium may vary, but the recommendation is formed on this.

**S2**

* Recommendations are made after a user requests a group recommendation
* The basis of the recommendation is the same as in S1
* ­­­

**S3**

**Project Components/Phases**

1. Chose Data source to work with and gather data (preferably > 10k)
2. Process Data
3. Write Software
4. Experiment
5. Evaluate
6. Analyze
7. Rinse/Repeat
8. Report

**Timeline and Schedule**

**Week 1(4/1/13 - 4/7/13)**

* Come up with top 3-4 data sources
* Write brief scenarios for each source
* Solidify project objectives

**Week 2 (4/8/13 - 4/14/13)**

* Finish anything not completed in Week 1
* Begin data gathering
* Have project outline

**Week 3 (4/15/13 - 4/21/13)**

* All data should be gathered at this point
* Start to process Data

**Week 4 (4/22/13 - 4/28/13)**

* Continue to process data
* Decide on methodologies for processing/experimenting/evaluating
* Look at other research papers?

**Week 5 (4/29/13 - 5/5/13)**

* Begin experimentation
* Have methods of evaluation decided upon

**Week 6 (5/6/13 – 5/12/13)**

* Finish experimentation
* Start Evaluation

**Week 7 (5/13/13 – 5/19/13)**

* Finish Evaluation
* Begin Analysis

**Week 8 (5/20/13 – 5/26/13)**

* Finish Analysis

**Week 9 (5/27/13 – 6/2/13)**

* Report

**Week 10 (6/3/13 – 6/9/13)**

* Report

**Week 11(6/10/- 6/14/13)[FINALS WEEK]**

* Report/fine tuning
* DONE