Spring 2015 Research Plan

1.) Write a brief outline of your research plan:

The majority of topic modeling algorithms available today returns a list of keywords associated with each topic, while it is left to the user to analyze how meaningful this output is. We point our efforts towards providing more useful representations of topics (Topic Summarization), while integrating contrastive opinions on each topic.

The first phase of the project will involve a thorough literature review of topic modeling and graph mining methods, as later on we will proceed with implementing a graph-related topic summarization technique. The basic idea is to compare different topics by using a graph representation of the document collections and a weighted scoring function that will leverage the term distributions produced by topic modeling algorithms, such as LDA.

We will explore several graph similarity metrics as well as walks on graphs to develop a novel topical phrase mining method.

2.) What do you hope to learn and achieve by the end of the semester?

By the end of the semester, all members of the group will be able to learn what is topic modeling, graphical models and probabilistic methods for text mining, as well as popular topic modeling algorithms and topic summarization approaches. The group will also become familiar with graph mining theory and will be able to formulate a research question that will hopefully produce publishable results.