

Project Title: Analysis of Citation Network

Team Members: Aditya Garg, Aashutosh Trivedi, Ansa Mary Ephraim

Project Task List

	Task	Date	Team Member
1	Write a parser to read the DBLP xml data file (requires understanding of the file format)	4/6	Aditya, Aashutosh
2	Design development pipeline (eg MapReduce functions to extract information based on different keywords, input format, output format)	4/6	Ansa
3	Build a collaboration graph	4/10	Ansa
4	List all journals/conferences (publication platforms) mentioned in DBLP	4/10	Aditya, Aashutosh
5	Identify top 20-30 publication platforms, which would be a combination of popularity (based on the list of popular conferences) as well as which find frequent mention in the database	4/12	Ansa, Aditya, Aashutosh (8-10 each)
6	Find 5 journals/conferences (publication platforms) from the above list for which we have information about the panel members, and information about a sufficient number of papers	4/13	Ansa, Aditya, Aashutosh
7	Build a citation graph starting out from the papers published in the above identified journals (if DBLP is inadequate, might require access to Google Scholar and/or Compute Science Bibliography database)	4/21	Aditya, Aashutosh
8	Analyze the citation graph and collaboration graph with different metrics between the panel members and the paper authors	4/24	Ansa
9	Identify key results and tie them together	4/27	Ansa, Aditya, Aashutosh
10	Iterate through steps 7,8, 9 till results hold	5/1	
11	Write paper (in LaTeX)	5/6	Ansa
12	Prepare presentation (need to consider effective data visualization)	5/6	Aditya, Aashutosh