Grading Rubric – Project 1 – Proposal

Name(s):

Project Topic:

CATEGORY	Excellent - 2	Satisfactory - 1	Unsatisfactory - 0	Score
Introduction:	Proposal introduction describes what question about the dataset is being answered, along with why it is important/relevant to the broad public interest.	Proposal introduction describes what question is being answered, but only provides a shallow motivation for why it is important/interesting.	No introduction is provided.	
Algorithm:	Proposal describes in detail the specific algorithm(s) you will use to answer the question described in the introduction.	Proposal provides a shallow description of the algorithm(s) to be used in answering the proposed questions.	Proposal provides no discussion of how the project questions will be answered (i.e. the algorithm(s) to be used).	
Infrastructure: Data Set	Proposal specifies how much of CommonCrawl data will be processed (in GB/TB), and provides a rationale for how this number was determined.	Proposal specifies how much of the CommonCrawl data will be processed (in GB/TB), but provides a shallow discussion of how this number was determined.	Proposal provides no discussion of how much data will be processed.	
Infrastructure: Computation Resources	Proposal estimates how many EC2 nodes (and what size nodes) will be required to complete parallel processing of the MapReduce algorithm. Proposal estimates total execution time for algorithm to complete. **Estimate is based on prototyping or first-hand testing**, not simply a random guess that happens to nicely match your Gantt chart!	Proposal provides an estimate of EC2 resources required for project, but the method used in estimating resources is weak and incompletely described.	Proposal provides no estimate of EC2 resources and/or computation time, or an estimate is given but is completely unsupported.	
Infrastructure: Cost	Proposal estimates how much it will cost to develop/implement the algorithm and then apply it in a full production run to the entire data set (or as large a dataset as was proposed above). The basis for this computation is fully described.	Proposal provides a cost estimate, but the method used in estimating cost is weak and incompletely described.	Proposal provides no cost estimate, or an estimate is given but is completely unsupported or justified.	
Analysis:	Proposal clearly describes what will be the result of the program, and how the data will be presented to the user in the form of graphs, tables, lists, or discussion.	Proposal provides a shallow discussion of how the output of the MapReduce program will be converted into human readable results.	Proposal gives no discussion about how the output of the MapReduce program will be converted into human readable results.	
			Total Score (out of 12)	

Comments: