

## FAQs for Final Project

### CIS493/593 Big Data

Q:

Since my partner and I have decided to do the coronavirus twitter data for the project, and the json data is given to us on blackboard, doesn't that already do phase 1 for us since we are already given the data and don't have to stream tweets into a json file?

Also, could you explain to me a little more about what exactly you are looking for in the proposal.

Answer:

If you want to use precollected Twitter data for the final project, some part of the first phase is already done. The first phase is to process raw data to convert them to the structured data to store them in a database server to retrieve data in Phase2 for data analytics processes. You will have to store the data either in a SQL server or MongoDB. If you will use MongoDB, you can directly store in Json and retrieve them in Json in MongoDB queries in Phase 1. If you want to store in MySQL, you will need to convert Json Data to Table form.

Proposal should have your data description on content and format, how to collect, which APIs/Tools to use, Description of your Platform, more importantly what do you want to find out from your big data? Any Analytic goal, how to find the information, which algorithms/steps to perform, which visualization API/Tools you plan to use for visualization.