# Architecture, System Overview, etc.

To be filled by Nasir

# Back end

## Twitter Harvesting

Twitter provides free, but rate-limited, access to public data using their Developer API. Specifically, the Stream API was used to obtain new tweets as they are posted onto Twitter. To overcome rate-limitation, a harvesting program was run in each of the back-end servers for a span of three weeks, compiling approximately 250,000 tweets combined, in and around Adelaide, Australia from the period of 20/09/2013 to 10/10/2013.

The Java program harnesses the Twitter4J framework to simplify interactions with the Stream API.

### The Program

Lucas’ program goes here

### Reasoning

## CouchDB

For the type of data being processed, a document-based database allows for easier processing of a large amount of data (Big Data Analytics).