Progress Report for team DivideByZero

Connor Brown (connorb8), Frank Salamone (frankns2), Jonathan Trusheim (jt38), Ray Butler (raylb2), Xinlei Huang (xhuang84)

Our team is in the process of constructing an electronic medical record search engine that allows the user to search through fabricated patient charts. We have 5 main tasks to complete, including construction of fabricated patient charts from data sets available on the internet, construction of the search engine itself, design of a user interface, implementation of the front end, and integration of these parts. We have made significant progress since the beginning of the project.

1. What tasks have been completed?

Construction of our dataset has largely been completed. This involved bringing together multiple data sources and constructing from this data 'patient charts' that are a reasonable facsimile to real charts. Our goal was to create several hundred charts. We now have around 10,000. We have constructed a search engine using elasticsearch, and have constructed an indexer to make an inverted index of each patient's chart. We have essentially completed the design of the front end.

2. What tasks are pending?

As we work on the individual components of our system, we have found that these components need to be adjusted to function together. This integration continues to be one of our most important pending tasks. We have made significant progress in construction of the back end including having a working search engine, but we need to integrate this into a functioning unit allowing for API calls. Although the front end has been designed, and we have made progress on this task, it does need to be completed.

3. What challenges do we face?

We have found that a big challenge is making each component of our system work on our various disparate computer systems. For example, the setup of Docker has been difficult on Windows machines. Moving the large files associated with the search engine has been difficult at times. In addition, packaging of these components into a usable product will be a continuing challenge.