Name: Keanu Nichols
email:nichols.keanu9@gmail.com
Github: https://github.com/kmn5409
Phone: +1 868 7907076

Reporting of CHAOSS Metrics

Synopsis: Grimoire Lab includes a tool for reporting, using things such as ElasticSearch it is able to pull data from different data sources for example GitHub. It then uses Manuscripts to create things for reporting on things like GitHub to find out things such as the amount of new committers for a particular time frame. Using tools such as Python Pandas data visualization can take place by creating things such as charts to better understand the distribution of data.

Benefits to Community: In looking at what can be gained by this, it must be acknowledged that this is an opportunity to further to use of things such as Data Analytics and making it easier to better understand how different data sources are being interacted with. It also encourages the learning of new materials that will become vital in contributing to open source projects that involve things relating to Data Analytics and Data Science by allowing persons the opportunity to practice their skills in things such as Python and obtain new skills along the way.

Deliverables:

- **Goals:** In working on this project I hope to make it easier to analyse contributions made by the CHAOSS Community. By using things such as ElasticSearch to download information about the different data sources the CHAOSS Community is interacting with and using things such as Pandas to create CSV and even matplotlib to create charts it will be easier to understand how persons are contributing to the community. I also hope to learn to use Jupyter Notebooks to make it easier to run Python scripts in the browser.
- **Implementation:** A number of the data analytics that need to be documented and converted in charts and or CSV can be done using things such as Perceval to pull information from Git and Github and then using matplotlib to create different charts to showcase the data. Using things such as Pandas we can also calculate things like averages for the amount of time things are open. Sorting Hat can also be used to determine particular things about the users that are contributing.

- **Timeline:**

Week 1:Become familiar get the Chaoss Community and determine the tasks that are of highest urgency to complete.

Week 2,3,4:Begin creating scripts to query the data about the community.

Week 5,6: Create charts and CSV files for the different data.

Week 7,8: Begins using Manuscripts to begin creating reports on the different metrics created.

Week 9: Convert the Python code to Jupyter books to be able to run it in browser

Week 10,11: Beginning of cleaning of code and documentation

Week 12: Extra time needed to fix any bugs in the code and uploading of scripts.

Related Work:

-Kaggle: I have had experience with using things such as Pandas, SciKit Learn and Jupyter Notebooks in creating data models online, but it is still a work in progress.

https://www.kaggle.com/kmn5409/my-model

-Chaoss: Completed the different microtasks and contributed to the grimoire-perceval repository **(https://github.com/kmn5409/chaoss-microtasks) .**

Biographical Information:

I am a First year student at the University of the West Indies St.Augustine (Trinidad) and I am studying Computer Science. I have an interest in Data Science and Data Analytics.