

John Doros (VC1A)
john.doros10@bcmail.cuny.edu
Project Title : Censorship Detection in Censored Planet

Time Logs 3/3/2025 - 3/9/2025

14 Hours Accumulated in this period

Total Hours Accumulated Since Beginning:

80 Hours Accumulated in total

Date	Duration (hours)	Category	Description of completed task	Challenges and/or next steps	Reflection
3/3/25	3	Design	Thoughtfully strategized a measurement point score strategy	Implement class into software	Ask Supervisor if this is a reliable approach
3/4/25	1	Coding	Thoughtfully strategized a filtering approach	Implement class into software	Can modify this filtering approach by adding more conditions so we need to make the class more flexible and adaptable
3/5/25	2	Coding	Implemented both filtering and point measurement approaches into software	Need to follow the conditions met in the strategies carefully since this both strategies can be adaptable to change	Show Supervisor my implemented strategies in the software and give updates to Supervisor
3/6/25	2	Documentation	Worked on designing diagrams for my thought process for the software so far	Be sure to upload to Github and update dashboard when designing diagrams	I want to be able to show how different approaches are utilized in the software and questions to be asked
3/7/25	1	Designing	Designing a new UI feature	Implemented a new UI feature into QT Designer in which visualizes the measurement point score that correlates to a color	This feature makes it easy to show whether a dataset is either a non censored event or censored event, we can modify this UI later as it is in development
3/8/25	1	Coding	Implemented UI feature into software		Update supervisor on this implementation

3/9/25	4	Documentation Coding	Fixed bugs in software, reorganized classes, and remodified classes so data processing is a lot cleaner, made new diagrams		Next steps will be to implement decision trees so that classifications can become more refined
--------	---	-------------------------	--	--	--

Reflection

What were your main goals in this time period?

The main goals in this time period were to come up with a strategy to classify an event (or a particular dataset) as a censored vs non censored as I want to really focus and align with on my hypothesis which is “How can censorship data be effectively aggregated and preprocessed to improve censorship detection accuracy while reducing false positives?”.

What were the main challenges during this phase? Were you able to meet the challenge, if so, what helped? If not, what could help?

The main challenge during this phase was to really focus on coming up with an effective/high-probability approach in order to classify each dataset as an event or non-event. This was a difficult task since I was only working with the dataset itself so I had to investigate what exactly each of the metadata fields meant and how it was correlated to a censorship event vs a non censorship event. I thought about different approaches such as only using specific conditions if the metadata fields were present or not but this was not enough to differentiate between events. So my final and thoughtful approach was to implement a score for each dataset as we view each metadata field and if some conditions had specific values then we can increase the score or decrease a score, this approach becomes effective in a sense that if a particular metadata field contained a value that suggested a censorship event we can increase the score, this score is a measurement value between 0 to 1, the closer to 1 the more likely this dataset was a censorship event. The next task for this approach is to answer if this approach is reliable by measuring against other known censorship reports during a specific time period, but I cannot use the censored planet itself. So I would need to study the paper in more detail in order to fix this issue. The next steps for the software itself is to work on the backend and implement useful features to the user of the software, for example such as implementing an informative visual within the software such as a graph to measure trends, and we can focus on refining our measurement point approach by incorporating decision trees as well.