

## **Data Management Plan**

The main data generated as part of the project are the semantic knowledge about processes. There will be intermediate data acquired via web search and crawling. There will be a curation of the system outputs using annotators. The collected data will only include the annotations on this knowledge and will not include any information about the annotators themselves. The data collection will proceed only after obtaining the appropriate Institution Review Board approvals. The data will be stored in Stony Brook servers available to the PI and will be released to the community. We will also release all the software developed as part of the project for the wider community to use.