**Presenters:**

Theodore LaGrow

Jacob Bieker

**Abstract Title:**

Do You Know Where Your Research Is Being Used?

An exploration of the National Science Foundation’s databases using Natural Language Processing.

**Abstract Body:**

We have developed software to decipher the uses of previously funded technology used and cited in other scholars’ research. In such a complex and dynamic field as computer science, it is of interest to understand what resources are available, how much the resources are used, and for what the resources are used. We demonstrated the feasibility of automatically identifying resource names on a large-scale from scientific literature in the National Science Foundation’s database and showed that the generated data can be used for exploration of software usage. While scholarly literature surveys can provide some insights, large-scale computer-based approaches to identify mentions of previously funded technology from primary literature is needed to automate systematic cataloguing and facilitate the monitoring of usage in a more effective method. We developed a pdf parser to extract text from articles in the database that we then trained using Natural Language Processing to evaluate if the article relates to the technology in question. The articles were given scores based on tokenization of nouns, titles, mentions, “strong” clues, “weak” clues, and references compared to a dictionary composed of important items related to the technology. From the article’s scores, we were then able to evaluate a trend of when and where the funded technology is being used. As we continue to expand this software, we will analyze the researchers’ sentiment about the technology.