ILLINOIS INSTITUTE OF TECHNOLOGY

impact

# Nayar Prize II Phase I Quarterly Progress Report
**April 2017**

**Project:** **Cyberbullying Early Warning and Response System**
**Team:** **Libby Hemphill and Aron Culotta**

## Progress Summary of Nayar Prize II Phase I

The goal of the Cyberbullying Early Warning and Response System project is to develop software tools to forecast imminent cyberbullying threats and vulnerabilities in online social networks that individuals and communities can use to avoid escalation. Highlights of our work from January 2017 – March 2017 include:

- Implemented a baseline classification model
- Crawled Instagram to collect data relevant to local schools (~4M comments)
- Implemented a search engine to explore all crawled Instagram posts (~15M comments)
- Created Amazon Mechanical Turk HITs to collect data annotations

## Model Specification

This quarter, computer science doctoral student Ping Liu joined our team and is leading our model building efforts. He and Aron also implemented a search system for our existing data to facilitate selection, exploration, and annotation. He implemented a baseline classification model that predicts whether there is bullying in individual Instagram comments.

## Data Annotation

Josh Guberman leads our data annotation efforts and iterated on training and annotation HITs through Amazon Mechanical Turk. As of March 28, 2017, we had 500 Instagram conversations labeled in which 244 conversations exhibited cyberbullying. We chose conversations for annotation by searching the interface built by Aron and Ping for comments containing malicious or inflammatory phrases such as "say it to my face," "I will break both sides," and a slew of racial, ethnic, and gender slurs.

## Plans for Q3

Our efforts in Q3 will focus on training and validating the model using data labeled by Turkers and designing the user interface for the user-facing components of the warning system. We expect to have a robust model and working prototype of the interface near the end of Q3 or earlier in Q4 so that we can iterate and refine the model and collect user feedback during Q4.

Next steps:

- Train and validate model on labeled data
- Iterate to refine accuracy of model and forecasting ability
- Implement user interface
- Solicit feedback from system users