

## Senior Design Bi-Weekly Status Report

**Team:** sdmay21-35

**Team Email:** [sdmay21-35@iastate.edu](mailto:sdmay21-35@iastate.edu)

**Team Members:**

Ahmad Alramahi - Lead Developer

Austin Boling - Meeting Facilitator

Joseph Naberhaus - Project Lead

Ekene Okeke - Report Coordinator

Ethan Ruchotzke - Documentation Manager

James Taylor - Linguistics SME

**Project:** POS Tagger for Software Documentation

**Report Period:** Aug 21 - Sep 4

# Summary of Progress in this Period

This week we began exploring the Stanford NLP. This includes:

- Reading and summarizing research papers about NLP and the Stanford NLP model
  - Found a paper detailing how the Cyclical Dependency Network of the Stanford NLP functions.
- Running the Stanford NLP on Python
  - Setting up Anaconda/Miniconda environments with dependencies
  - Passing test samples into Python Stanford NLP
- Running the Stanford NLP on Java
  - Setting up Maven dependencies
  - Configuring JVM to support resource-heavy Stanford NLP library
  - Passing test samples into Java Stanford NLP

# Pending Issues

- Java eating up RAM
  - Ahmad reports 7+ GB of ram necessary to run the NLP algorithm in Java
  - Future Goals:
    - Reduce the amount of ram to a more approachable level
- Java installation documentation
  - Create a detailed document explaining how to use maven to install Stanford NLP in a java environment
- Python installation documentation
  - Research the installation techniques between Windows / Mac / Linux
- Map out errors between Java and Python
  - Consistency. Are both processors returning the same results?
  - Error Consistency
    - Are both processors failing on the same words?
    - Are both processors returning the same errors on the same words?

# Plans for upcoming Reporting Period

Our plans for this upcoming reporting period (**Sep 18 - Oct 2**) are broken down as follows:

Who	What	Due When
Ahmad	debug the Java implementation of the Stanford NLP and discover why excessive RAM is being used, and potentially solve the issue	Oct. 2
Austin, James	find and/or define PoS tags of software code and documentation	Oct. 2
Austin, James	read the paper on Cyclical Dependency Networks, summarize their findings, and report this to the other group members	Oct. 2
Ethan, Joseph	check the availability of code for training for PoS in Python	Oct. 2
Ahmad, Ekene	check the availability of code for training for PoS in Java	Oct. 2
Ahmad, Ekene, Ethan, Joseph	compare and check consistencies and inconsistencies between the Java and Python implementations of the Stanford NLP	Oct. 2
Ahmad, Joseph	complete the lightning talk on the Project Plan	Sept. 23
All	familiarize yourself with the PoS documentation created in the previous reporting period	Oct. 2