

# E E/Cpr E/S E 492 Bi-Weekly Report 6

## Intelligent Code Editor

Client & Advisor: Ali Jannesari

sdmay20-46

John Jago – Software Lead

Keaton Johnson – Systems Lead

Jon Novak – Machine Learning Lead

Matthew Orth – Meeting Facilitator

Garet Phelps – Report Manager

Isaac Spanier – Test Lead

## Weekly Summary

During this sprint, now that the dataset is done it will be used to train the model. We will see how accurate our model is and make changes as necessary. All the parts of the system (plugin, Python scripts, AWS, translation engine) should be finished up so we can see how well the end-to-end system performs.

## Past week accomplishments

John Jago

- Python services on AWS
  - Helped Isaac finish setting up our preprocessing and translation engine services on AWS
    - Deployed the OpenNMT-py REST server along with our latest trained model on an EC2 instance; it's accepting requests for translation at <http://18.221.191.233/translator/translate>. See the *aws-setup* directory in the GitLab project for details.
    - Deployed sentence preprocessing script as a Lambda function
  - Documented the setup of both services in GitLab
- IntelliJ plugin text preprocessing
  - Assisted Garet in getting the plugin set up to use the preprocessing script on AWS

## Keaton Johnson

- Analyzed training data
  - Looking at ways to improve the dataset accuracy.
  - Mainly looked at ways to expand the dataset utilizing various javadocs.

## Jon Novak

- Looking into possible ways to improve dataset accuracy

## Matthew Orth

- Java Method Invocation Dataset Training and Evaluation:
  - Trained a model on the 1,500 sample Java method invocation dataset we created
    - Identified and documented potential optimizations that could be made to the system
      - Determined that reducing the beam size to 5 in translations improves the results to around 60% accuracy
- Research Java Speech to Text Functionality:
  - Created a list of possible Java speech to text functionality libraries and APIs future groups could use to implement speech to text
  - Created implementation of Google speech API based on the following project <https://github.com/goxr3plus/java-google-speech-api>
- Final Report, Poster, and Presentation Updates:
  - Made various updates and revisions to the final report, poster, and presentation to include our more updated progress and work

## Garet Phelps

- Finished setting up the IntelliJ plugin interface, and connecting it to the lambda function and the EC2 server.

## Isaac Spanier

- Setting up Lambda function with a zip file of the project, and giving the correct permissions to all users.
- Setting up the AWS account to be passed to the next group as a legacy.

## Individual contributions

Name	Contributions	Bi-Weekly Hours	Cumulative Semester Hours
John Jago	<ul style="list-style-type: none"><li>• Python services on AWS</li><li>• IntelliJ plugin text preprocessing</li></ul>	13	58
Keaton Johnson	<ul style="list-style-type: none"><li>• Analyzed training data</li><li>• Researched ways to enumerate javadoc dataset</li></ul>	7	43
Jon Novak	<ul style="list-style-type: none"><li>• Researching improving dataset accuracy</li></ul>	10	46
Matthew Orth	<ul style="list-style-type: none"><li>• Java Method Invocation Dataset Training and Evaluation</li><li>• Research Java Speech to Text Functionality</li><li>• Final Report, Poster, and Presentation Updates</li></ul>	13	80
Garet Phelps	<ul style="list-style-type: none"><li>• Finished setting up the IntelliJ plugin interface, and connecting it to the lambda function and the EC2 server.</li></ul>	15	54
Isaac Spanier	<ul style="list-style-type: none"><li>• AWS Configuration with Lambda and nltk python scripts</li><li>• Setting up AWS to be transferred</li></ul>	15	48

## Plans for the upcoming sprint

John Jago

- Prepare and practice final presentation
- Update final report, poster, and remaining deliverables

Keaton Johnson

- Prepare and practice final presentation
- Update final report, poster, and remaining deliverables

Jon Novak

- Prepare and practice final presentation
- Update final report, poster, and remaining deliverables

Matthew Orth

- Prepare and practice final presentation
- Update final report, poster, and remaining deliverables

Garet Phelps:

- Prepare and practice final presentation
- Update final report, poster, and remaining deliverables

Isaac Spanier

- Prepare and practice final presentation
- Update final report, poster, and remaining deliverables

## Summary of weekly client/advisor meeting

### **Meeting with Professor Jannesari and Hung Phan on 2020-04-06 at 4:30 pm**

During this meeting, we discussed our updates on the 1,500 sample Java method invocation dataset we labeled and trained a model on. We determined that the 50-60% accuracy that we achieve is very good as it is about 15-25% higher than our AnyCode, which we used as a benchmark for our project. We also discussed that during the next week, we will complete our end-to-end system and the final sprint will be used to write and finalize our final report, poster, and presentation.

### **Meeting with Hung Phan on 2020-04-13 at 4:30 pm**

During this meeting, we discussed the progress on our system integration and research. Our system is almost fully integrated with just a few preprocessing things to consider. We will have our entire system fully integrated and a rough draft of our final report, presentation, and poster for next week's meeting.