

## EE/CprE/SE 492 - sdddec22-19

### Text extractions from documents into Elasticsearch

### Report 13

10/11/22 - 10/25/22

Client: Kayla Gilleland, Buildertrend

Faculty Advisor: Goce Trajcevski

#### Team Members:

Bruce Bitwayiki

Jared Hayashi

Rushal Sohal

Tiffany Mayberry

#### Weekly Summary

The focus of the past two weeks has been beginning implementation of the last two major UI components (upload dialog & filter bar), continued efforts to access Elasticsearch, and deploying the text extraction component to our server. We also revised our schedule for the remainder of the semester and began planning our second status presentation for class.

#### Past Week Accomplishments

- Developed milestone schedule for the rest of the semester
- Reconfigured Elasticsearch to enable additional port access for clients
- Basic upload functionality achieved (for multiple files)
- Implementation of the Filter Bar Component

#### Individuals Contributions

Name	Individual Contributions	BiWeekly Hours	Hours Cumulative
Bruce Bitwayiki	Troubleshooting Elasticsearch port/certificates issue. Sync up regarding match case search use case on frontend	8	43
Jared Hayashi	Deployed Tika to the server and continued work on json extraction	5	45
Rushal Sohal	Worked on the upload functionality and some integration	13	45
Tiffany Mayberry	Worked on UI design / implementation of the filter bar component	6	63

## Plans for the Upcoming Weeks

- Finish Tika and Elasticsearch communication
- File specific metadata extraction
- Looking into Tesseract for images
- Complete upload UI
- Complete setting up the spring server to handle uploaded files
- Complete integration of the filter bar with the basic search query
- Work on PIRM presentation

## Summary of Weekly Advisor Meeting

- Revised PIRM meeting presentation
- Thinking about making revisions to remaining milestone schedule