Computer Programming II Final Project Report

short line

Noah Richardson and Joshua Conrad  
December 12, 2019

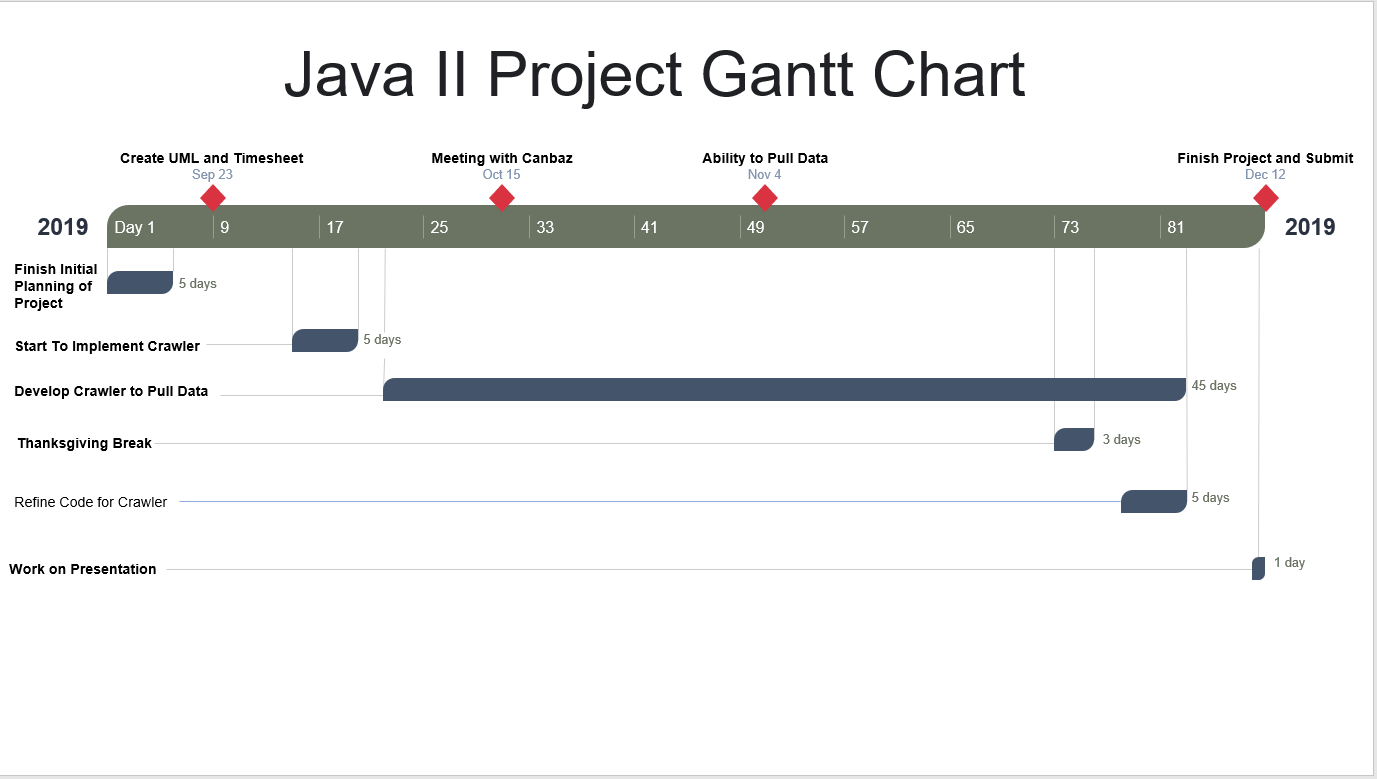
# Introduction

This report is for the final project in Computer Programming II. For this project, we have designed and created a program that pulls data from web news sources so it can be analyzed. The goal for this project was to be able to pull data and analyze it to determine whether or not some news sources were changing anything about various news articles they were providing to the public.

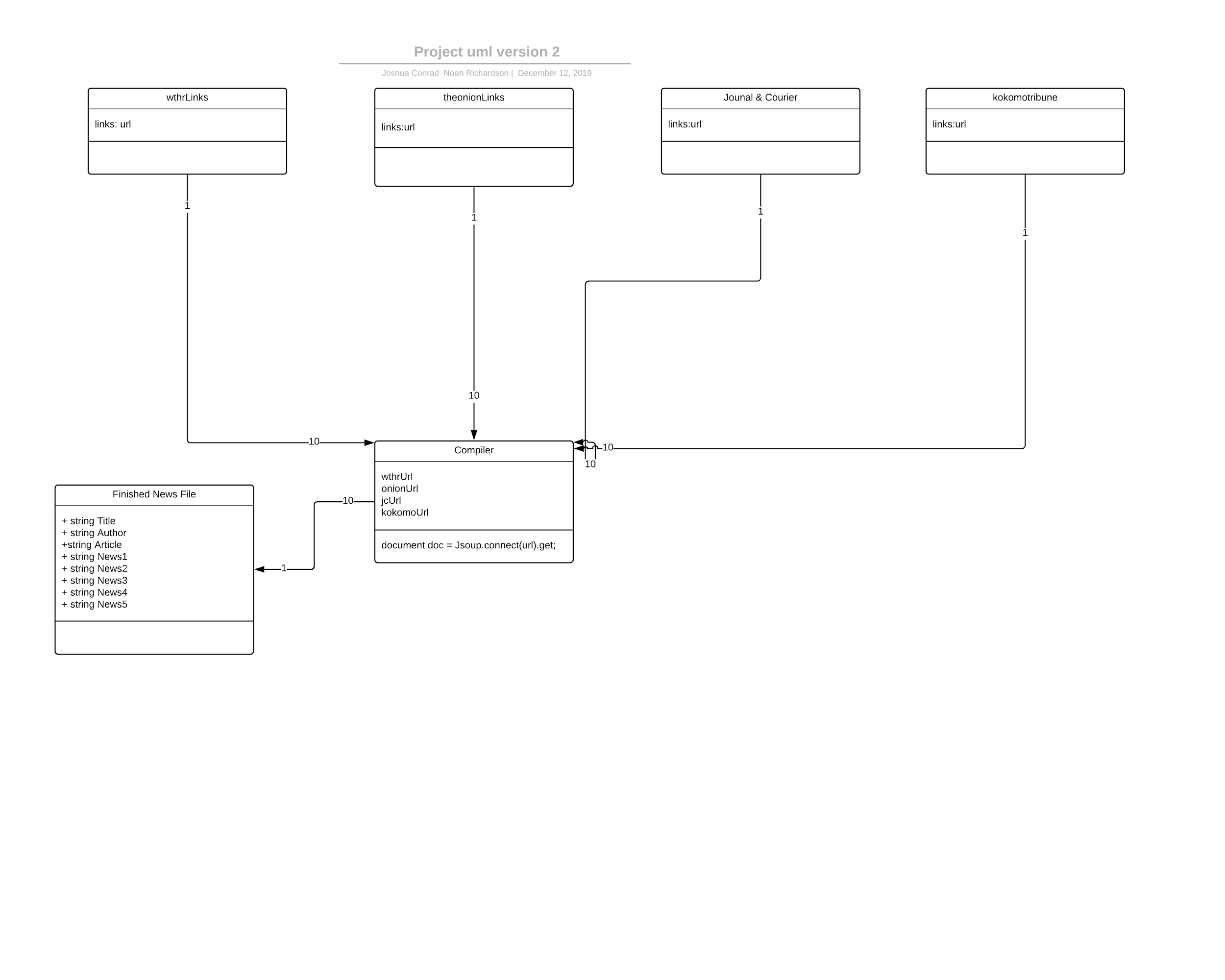
# 

# 

# Project Timesheet



# UML



This is what the final product should have been if the links would have worked.

# What We Learned

This project was very fascinating and challenging at the same time. Trying to pull data from web sources, in the grand scheme of things, does not sound like a challenging task. This was quickly proven wrong while working on this project. Being able to pull specific data from multiple sources in one program turned into a very difficult ordeal. We found ourselves racking our brains trying different methods to extract data from our sources and running into various issues along the way.

One of the biggest challenges we faced, which ultimately took us the longest time to complete, was just extracting data from multiple sources. All of our sources did not respond to our code the same way, forcing us to adapt and rewrite our code to try and fix the issues.

# 

# What we Had Issues With

# Project was no small feat for us and because of that, we ran into multiple problems during development. One of the biggest we ran into was writing a program that would allow us to pull the data we wanted from multiple sources simultaneously. This was the major issue because we want to pull the data as efficiently as possible.

One of the other major issues we ran into was the formatting of the websites. Not all of the websites we used had the same format which made it difficult to superclasses in our code. This was an issue because it limited us during our development and forced us to resort to other methods when developing our program.

This last problem was also increased due to the fact that when we were trying to pull specific data, the URLs for some of our sources were not wanting to work correctly. This forced us to use either more refined links, or just the regular link to the source home page.

One of the more strange issues we ran into, was dealing with security clearance. We would have to download and save the websites certification to run our code. This was a major problem for Conrad because his personal computer was unable to add the certifications to the right file.

During the course of the project we were using the Jsoup libraries to aid in our development and it initially took us a considerable amount of time to get it working. This caused quite a setback in our progress but ultimately was resolved and we were able to get back on track.

# 

# 

# Conclusion

Looking back at the project we can see that there is still a lot left to do. If we had more time the output documents would be better formatted. There is also the fact that we would have enjoyed being able to actually sort out our data. This project has allowed us to see the importance of not only time management, but also gave us the chance to see the difficulties of programming in the real world.

short dash