CSCC01 PHASE 3 REPORT

Team Chihuahuas (#16)

Table of Contents

State of the Project 2

Release Plan 3

Iteration Plan 3

Task Board 3

Burndown Chart 3

# State of the Project

We are currently on week 6 of our release plan; this is noted on our release plan/task board located on Trello. So far we have successfully completed the crawler (Scrapy) for the website which shows the articles in the table on the News Investigator website. In weeks two – five, we worked on getting the crawler, saving information to the database (CouchDB), and login working. When the user opens the News Investigator website, he/she is prompted to login or signup if it’s the first time. After successfully logging in, the user is redirected to the index page where the user can add/see news sources and add/see the keywords. An important point to mention here is that the user will only be able to enter a valid URL to the news source. We ensured this with the use of regex before sending the information to the database. Same criteria holds for the keywords where the user would only be able to enter an alphanumeric keyword without any spaces. The user would then start crawl from the navigation bar, which will kick off the scrapy crawler. The crawler currently parses the news sources from the database making sure that it saves those articles whose source contains the keywords. The results table in the database saves the id, title, link, date, and the source of the article. This information is then made available to the website through Flask and the user can view the results in the table by clicking on “View Sources Table” in the website to see the title, link, hyperlink(s), quotes, and mentions for each article on a separate row.

This week we started on using a python-based web framework, Flask, to ease the process of development and also be able to run scrapy from the website. We also started testing the backend with unittest in Python and the frontend using Selenium IDE. In addition to that we have also started working on the Twitter handles and showing that information on the website.

There couple changes we made to the release plan. First of we didn’t have any backend until Phase II. We just had the frontend and the database. Looking at the remaining user stories that we need to finish, we figured a framework would make it much easier to complete the remaining tasks while ensuring our code remains modular. Also, we had to move the user stories pertaining to twitter from week 5 to week 6. This was done primarily because the tasks from previous iterations were not completely done.

# Release Plan

To be filled in

# Iteration Plan

Legend:

* BP stands for Business Priority (The smaller the number, the higher the priority.)
* DD stands for Developer Days

**Iteration Plan – Week 3 - Oct 17 - Oct 23**

**Iteration Plan – Week 4 - Oct 24 - Oct 30**

**Iteration Plan – Week 5 - Oct 31 - Nov 6**

**Iteration Plan – Week 6 – Nov 7 - Nov 13**

# Task Board

Please see our Trello page (“CSCC01 Team 16 Task Board”) for the Task Board. If you do not have access to it, please contact one of the team members.

# Burndown Chart

To be filled in