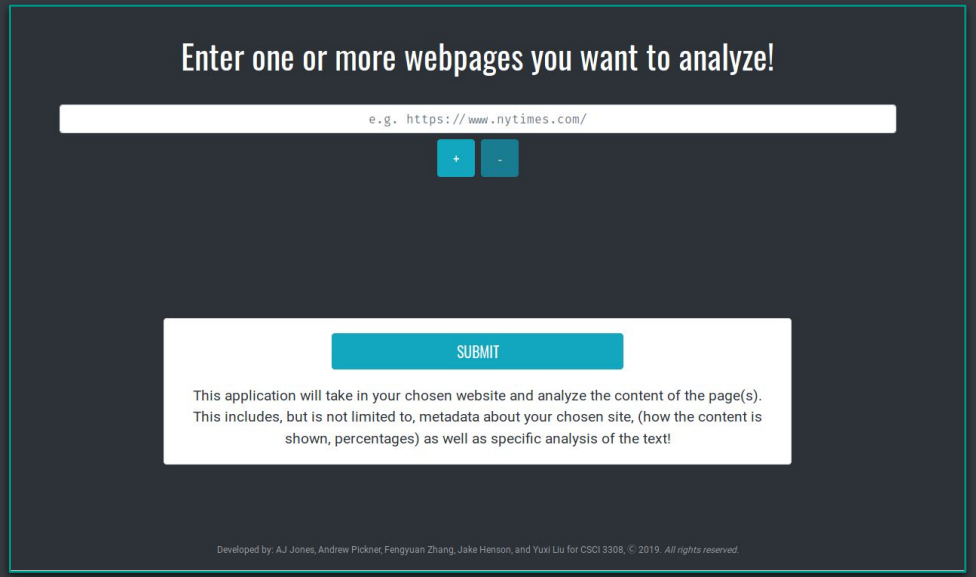
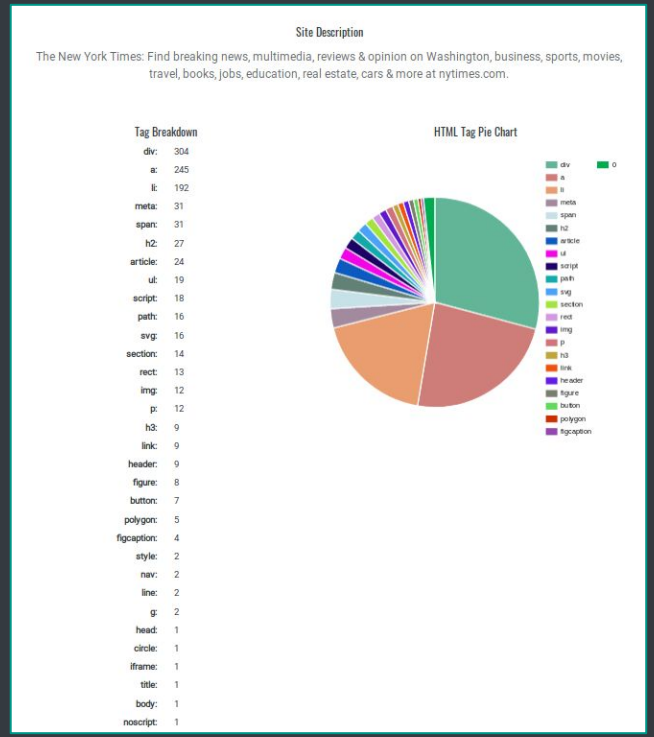
Web Page Analyzer

Andrew Pickner
AJ Jones
Jake Henson
Yuxi Liu
Fengyuan Zhang

Webpage Analyzer: What is it?





Who is it useful for?

Initially, we designed this with data scientists in mind

However, it can be useful for anyone who wants to know more about the sites they visit and the info about them. Displaying metadata and text analysis of a page can also be useful for just about anyone who's curious about the sites they visit!

How does it work?

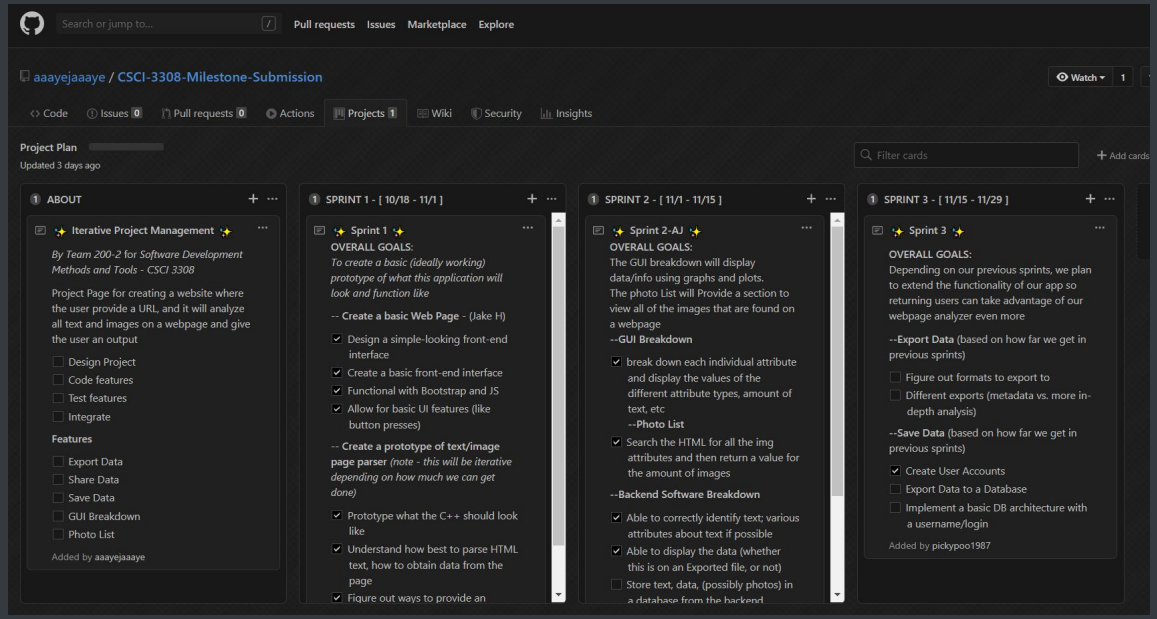
1. Front-end site to allow the user to enter website URLs
2. API calls to Heroku server using NodeJS
3. Python web-crawler gets data from the requested URL
4. Population of the page with python API call + ChartsJS
5. Changeable views of forms/data with buttons on the Results Page

Tools we used:

VCS Repository: GitHub Repository	5/5 ★★★★★
Project Tracker: Github Projects	4/5 ★★★★☆
IDE: Atom	5/5 ★★★★★
Deployment Achiever: Heroku	2/5 ★★☆☆☆
Database: PostgresSQL	2/5 ★★☆☆☆
FrameWork: NodeJS	4/5 ★★★★☆
Languages: Python, HTML, CSS, SQL	5/5 ★★★★★
Pug	2/5 ★★☆☆☆
Methodologies: Agile + Iterative	4/5 ★★★★☆

Workflow:

- GitHub Projects as a project planner to make sure we kept track of our sprints
- Discord to message each other about the project



Challenges

Setting up the Heroku backend: Heroku had to be created with the correct file structure and importing a ***ton*** of dependencies for python.

Integrating the individual components: Connecting the HTML to the Python script and connecting the database to the webpage.

Rewriting HTML: We ended up switching our HTML to Pug

SQL Database: We had issues creating/populating SQL databases both locally and on Heroku

Middle-Layer Integration: We had difficulty merging all of the different pieces together

Let’s see a live demo!

The project:

https://website-analysis-csci3308.herokuapp.com/

Our github:

https://github.com/aaayejaaaye/CSCI-3308-CodeNStuff