

Technical Test

The purpose of this test is to evaluate your ability to solve problems, the use of best practices and your ability to review functional requirements. Please do not focus on the frontend, adding a nice look and feel is a nice to have but the evaluator will not focus on this. Last but not the least as a Company we pay extra attention to the readability of your solution, the testing that is performed, the completeness of the test suite, the usage of best practices and patterns, and the instructions for us to run the project.

Summary

This is going to be a simple application, where a user can add a link to a web page and the application will scrape all the information of that page and get a list of all of the links in that page.

Scope

You will only check for a tag, any other way that a link could form (i.e. via javascript you could ignore it).

Features

- As a user, I must be able to register on the platform. For this, it will only be necessary to enter a username and password.
- As a user, I must be able to log into the system using an email and a password.
- As a user, I should be able to see a list of all the pages that I have scrapped with the number of links that the scraper found.
- As a user, I should be able to see the details of all the links of a particular page, that means the url of a link and the “name” of a link.

- As a user I should be able to add a url and the system should be able to check for all the links and add it to the database. A link will have the following format ` Visit W3Schools.com! ` the href will be the link and the body will be the name of the link. Keep in mind that the body of a link sometimes is not only text and could be other html elements, in those cases you could save only a portion of the html. The title of the web page will be the page name. Keep in mind that some pages will take more time than others to scrape.

Nice to have

- Pagination in the list of pages and the list of links
- As a user, when I see the list of pages, I should be able to see the ones that are currently being processed.

Guidelines

We will expect to have something like this for the landing page:

Scrape

Name	Total links
Google product	20
Amazon DIY page	6
Facebook user anatoly	32
Blog posts 24	33
Sample page	13
page processing	in progress

[Previous](#) [1](#) [2](#) [3](#) [Next](#)

And for the internal page we will expect something like this:

Google Product

[< Back](#)

Name	Total links
Home page	www.google.com
This is a product	www.google.com/products/1
<img src="soemthing.png" ...	www.google.com/assets/image.png
Resume	www.amazon.com/offer

[Previous](#) [1](#) [2](#) [3](#) [Next](#)

Tools

- Please include all the tools that you used in your process in your readme file.