

Web Crawler Take Home Test:

We'd like you to write a simple web crawler in a programming language you're familiar with. Given a starting URL, the crawler should visit each URL it finds on the same domain. It should print each URL visited, and a list of links found on that page.

The crawler should be limited to one subdomain - so when you start with [anonymized url](#), it would crawl all pages on the that website, but not follow external links, for example to [facebook.com](#) .

We would like to see your own implementation of a web crawler. Please do not use frameworks like scrapy which handle all the crawling behind the scenes or someone else's code. You are welcome to use libraries to handle things like HTML parsing.

Ideally, write it as you would a production piece of code. This exercise is not meant to show us whether you can write code –

we are more interested in how you design software. This means that we care less about a fancy UI or sitemap format, and more about how your program is structured: the trade-offs you've made, what behavior the program exhibits, and your use of concurrency, test coverage, and so on.