

Google Advanced Search Using the Custom API

Team Members: _____

Using the “Getting Started with the API – Custom Search – Google Developers” site, https://developers.google.com/custom-search/v1/getting_started create your own Google account and read through the documentation before arriving to class. In class, you will be provided with a short PHP program that uses CURL to access the API, but you should be familiar with the basics of what Google’s API can do.

The sample code you will be given in class will be used to scrape ISBNs from pages on the www.amazon.com site. This is legal but closely resembles what black hats would do to scrape account or social security numbers from sites with malicious intent. Using one of your developer API keys, use this code to search for ISBNs on a few pages to become clear on how the program works.

[Stetson Specific Assignment: Could be replaced with more general assignment] As we all have signed a Non-Disclosure Agreement with Stetson University allowing us access to penetration test the University’s site, you will refine your search to determine if any public-facing data is present that Stetson would clearly consider private. You may only search for such sensitive data on the Stetson site as results are to be made available to the CIO.

1. First, use the Google Search Directives by hand to experiment with finding sensitive data on the Stetson University web site
2. Using the code provided, refine the code to automate the search to find data that may be sensitive. As you will only receive 100 free hits per day, you are to continue this assignment out-of-class each day for a week using both team members’ keys and accumulate the results obtained. As a class, you may wish to work together to determine the best search queries and refinements so that each group may focus on different target points without repeating the same queries between groups.
3. Write up a short presentation that summarizes your findings. You should aim to provide the most visually logical means of presenting your data, be that tables, graphs, etc. Each group will present their results in class.