Clickthrough Tracking:
a successful way to track electronic government documents usage @ Randall Library

Jason Fleming, Elliot Simpson, and Beth Thompson

Why we needed the Clickthrough tracking

In Randall Library, we have an extensive Government documents collection. We keep statistics of how often the print resources are used. Within the last few years, many of these are either now available online or are born as a digital resource. We bring in hundreds of records for these electronic resources and websites every month. We wanted to find a way to count how often these resources are accessed through our catalog. Use of these resources and the government collection in general needed to be demonstrated for its value and purpose and how it supports our users needs

Creation Process / Research

- A number of webinars were viewed of other libraries implementing this process to record the use of their electronic documents. None of them used EZproxy.

- Gov Docs Cataloger met with Library IT Project Manager and aligned our environment with this process

- Library IT developed a potential solution using EZproxy log files to record and then display clickthroughs

Making EZProxy work

- An unrestricted Stanza was created to allow public access

```
#Allowing everyone through to GPO Purls
AutoLoginIP 0.0.0.0-255.255.255.255
T Government Documents
U http://purl.fdlp.gov
HJ https://purl.access.gpo.gov
HJ https://purl.fdlp.gov
H purl.fdlp.gov
DJ fdlp.gov
DJ access.gpo.gov
DJ gpo.gov
IncludeIP  0.0.0.0-255.255.255.255
```

- A cron job was set up to run a script every morning to copy instances of purl links in EZProxy logs, and write them to an alternate text file on the server that was then accessed by a web application.
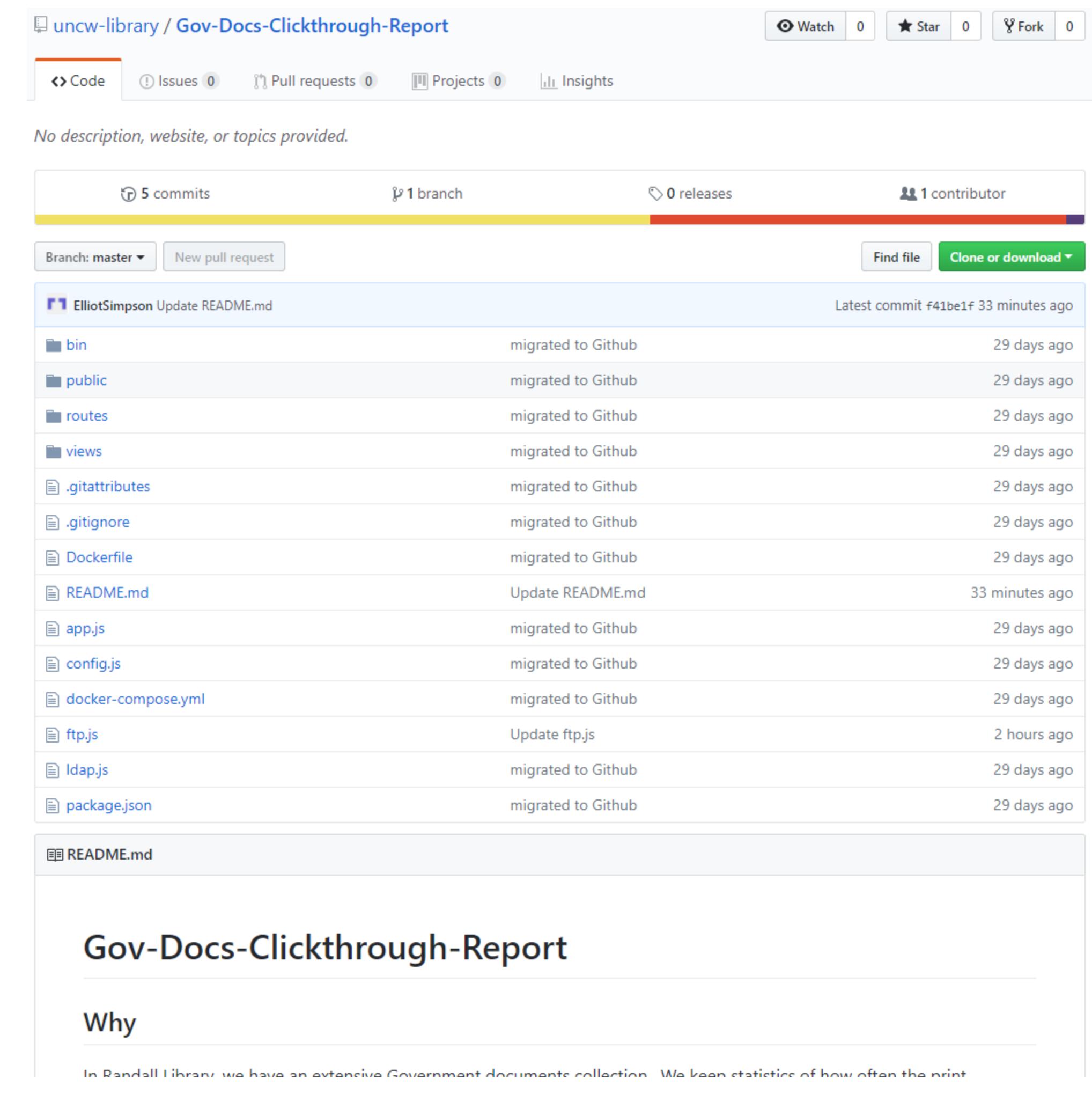
Connecting Gov Doc records to EZproxy

- All previously cataloged records were modified to include the proxy: *.liblink.uncw.edu

- The Gov Docs cataloger was instructed to use MarcEdit to modify URLs (856) in new records before loading them into the ILS

Writing a web application

- We wrote a reporting application to track clickthrough statistics using NodeJS – a runtime environment for writing Javascript on the server side.
- The application works by connecting to the EZproxy FTP server and reading the contents of the text files generated in EZproxy – for a specified month
- The contents of the file are parsed and reported on in a readable format.
- In an effort to report valid and useful data we had to do the following:
  - Remove duplicate entries
  - Count the number of sessions per IP address
  - Count the number of sessions in general
- We implemented the Datatables library to create table filters and export options

Rewriting the web application

- It was discovered that the PURL links had a number of different URL formats and so we adjusted the EZproxy stanza, and in the application and started accounting for them when editing the records in MarcEdit

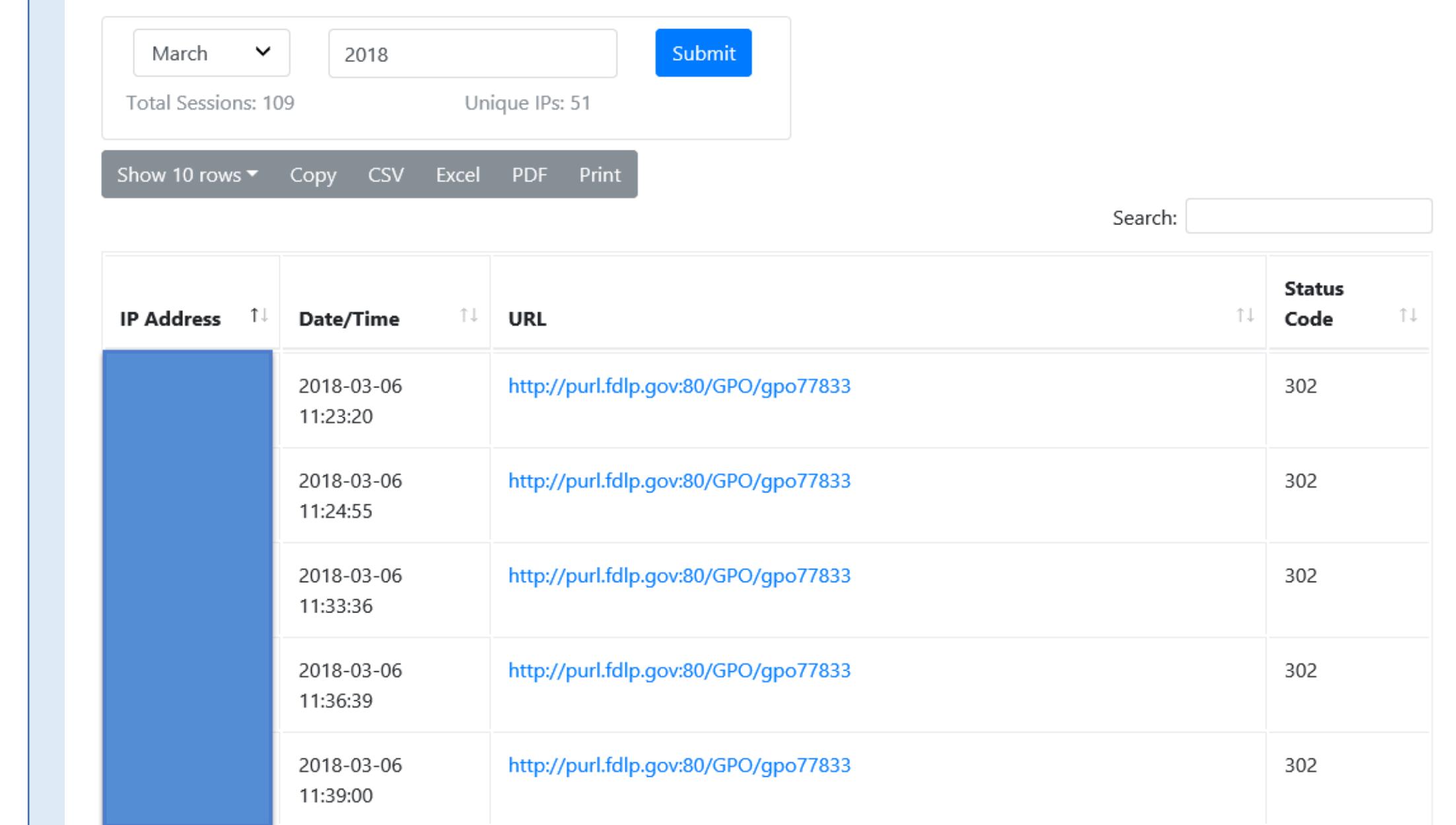- The Clickthrough tracking page was modified to show us when a link was broken and needed to be fixed (403)

Current Use

As you can see from the screenshot below, our web application contains all the data in one easy to read interface with multiple options for sorting and exporting

Gov. Documents Clickthrough Report

IP Address	Date/Time	URL	Status Code
	2018-03-06 11:23:20	http://purl.fdlp.gov:80/GPO/gpo77833	302
	2018-03-06 11:24:55	http://purl.fdlp.gov:80/GPO/gpo77833	302
	2018-03-06 11:33:36	http://purl.fdlp.gov:80/GPO/gpo77833	302
	2018-03-06 11:36:39	http://purl.fdlp.gov:80/GPO/gpo77833	302
	2018-03-06 11:39:00	http://purl.fdlp.gov:80/GPO/gpo77833	302

We are currently using this report for the purpose of counting monthly usage statistics.

Ideas for the Future

- Pushing the data to an external database for more efficient querying

- Implementing data visualization using the Google Charts API

- Retrieving more data points based off of individual links – possibly retrieving link information such as resource type.

- Provide on demand linking to record # for 400 error fixes

- Adding clickthrough tracking to our North Carolina born digital records