Eastern Illinois University

From the SelectedWorks of Stacey Knight-Davis

April 28, 2017

Using AWStats to Analyze Logs from EZProxy and from the Public OPAC Logs

Stacey Knight-Davis

Available at: https://works.bepress.com/stacey_knight-davis/62/

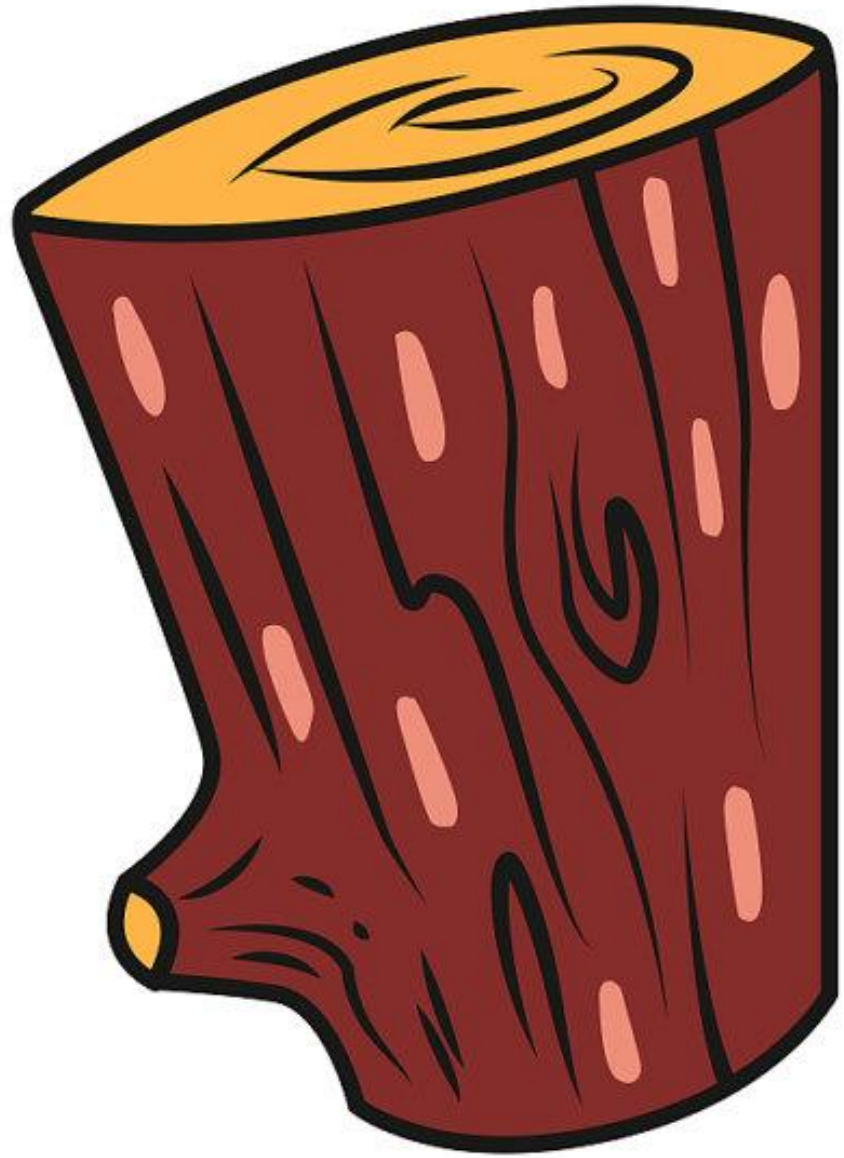
Using AWStats to Analyze Logs from EZProxy and from the Public OPAC Logs

Stacey Knight-Davis

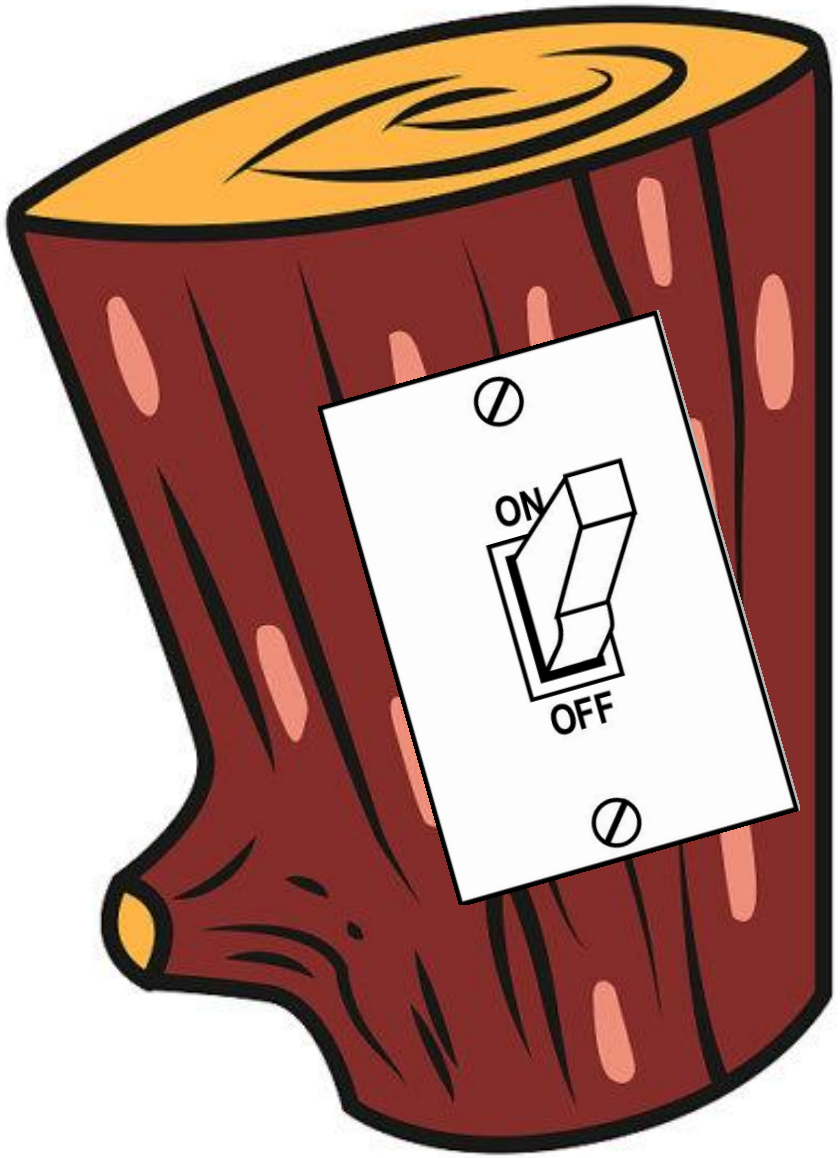
Eastern Illinois University

Log Analysis

With Google Analytics Query String Added Extras!



Log only exists when turned on!



Why might the access log be Off?

- It’s big
- It grows bigger every day
- Balancing patron privacy and data driven decision making

ILA best practice is to delete old logs or anonymize them

What might be in a log?

- Images served
- Pages served
- Search terms, if configured to do so
- Date
- Time
- User's Browser
- User's IP
- User's login

Where can I get logs?

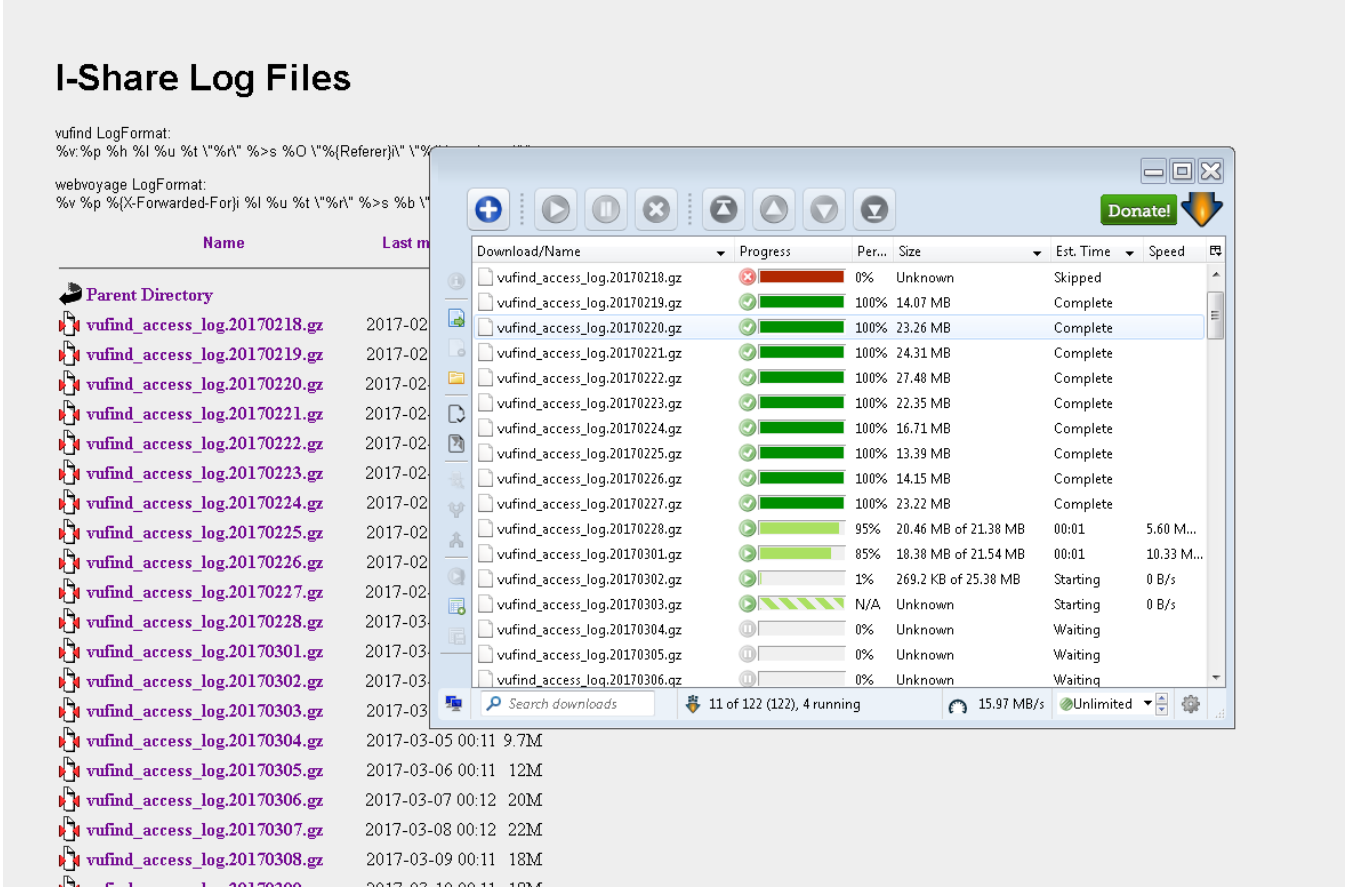
- CARLI search logs
- Pre-anonymized for your convivence

Stuff you Need

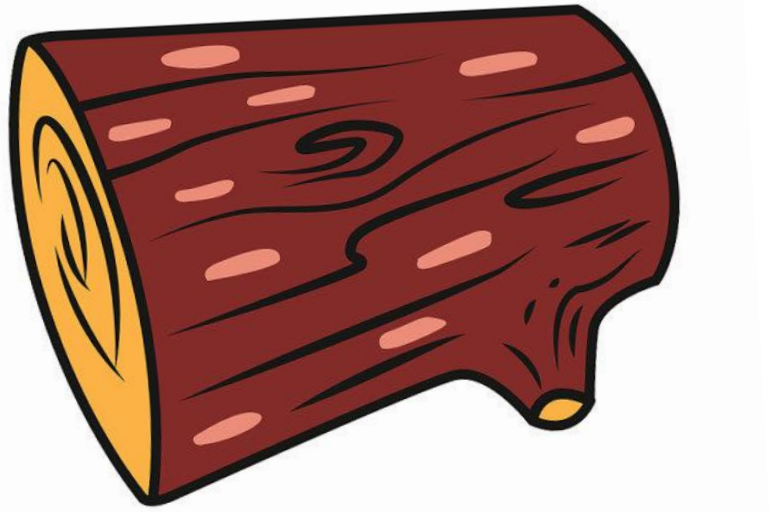
- Download Them All – Firefox Plugin
  https://addons.mozilla.org/en-US/firefox/addon/downthemall

- 7zip
  http://www.7-zip.org/

- Log URL
  https://www.carli.illinois.edu/logs/



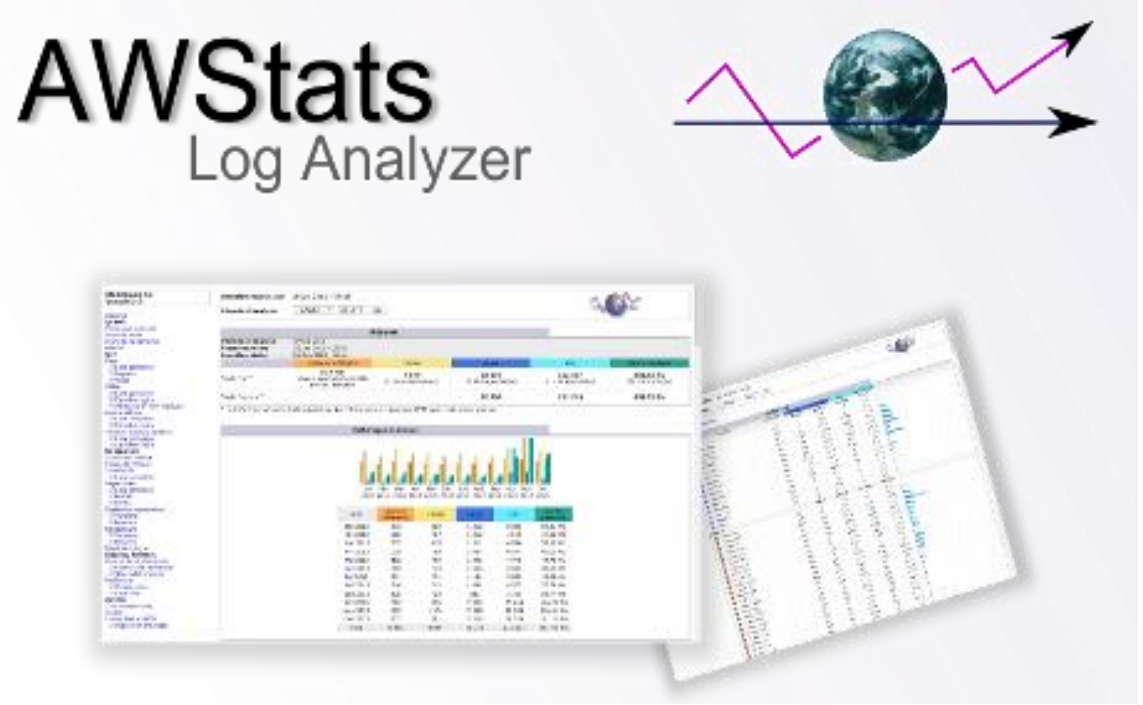
Without analysis, logs are about as useful as a chunk of wood



```
vufind.carli.illinois.edu:80 0.0.0.0 - - [26/Jul/2016:00:00:00 -
0500] "GET /vf-wrh/Record/isl_641878/favicon.ico HTTP/1.1" 200
1320 "-" "Mozilla/5.0 (compatible; YandexBot/3.0;
+http://yandex.com/bots)"
vufind.carli.illinois.edu:80 0.0.0.0 - - [26/Jul/2016:00:00:00 -
0500] "GET /vf-uiu/Record/uiu_7337077/Email HTTP/1.1" 200 10066
"-" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1)"
vufind.carli.illinois.edu:80 0.0.0.0 - - [26/Jul/2016:00:00:01 -
0500] "GET /vf-uiu/Record/uiu_7337077/Save HTTP/1.1" 500 297 "-"
"Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1; SV1)"
vufind.carli.illinois.edu:80 0.0.0.0 - - [26/Jul/2016:00:00:01 -
0500] "GET /vf-wrh/Record/uiu_1883620/favicon.ico HTTP/1.1" 200
1321 "-" "Mozilla/5.0 (compatible; YandexBot/3.0;
+http://yandex.com/bots)"
vufind.carli.illinois.edu:80 0.0.0.0 - - [26/Jul/2016:00:00:01 -
0500] "GET /vf-uiu/Record/uiu_7337077/Details HTTP/1.1" 200
29052 "-" "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.1;
SV1)"
vufind.carli.illinois.edu:80 0.0.0.0 - - [26/Jul/2016:00:00:01 -
0500] "GET /all/vf-dom/Record/3819286 HTTP/1.1" 200 6370 "-"
"Mozilla/5.0 (compatible; bingbot/2.0;
+http://www.bing.com/bingbot.htm)"
vufind.carli.illinois.edu:80 0.0.0.0 - - [26/Jul/2016:00:00:01 -
0500] "GET /vf-wrh/Record/wrh_326558/favicon.ico HTTP/1.1" 200
1318 "-" "Mozilla/5.0 (compatible; YandexBot/3.0;
+http://yandex.com/bots)"
vufind.carli.illinois.edu:80 0.0.0.0 - - [26/Jul/2016:00:00:02 -
0500] "GET /all/vf-rou/Record/10132272 HTTP/1.1" 200 6487 "-"
"Mozilla/5.0 (compatible; bingbot/2.0;
+http://www.bing.com/bingbot.htm)"
vufind.carli.illinois.edu:80 0.0.0.0 - - [26/Jul/2016:00:00:02 -
0500] "GET /vf-wrh/Record/wrh_326558/favicon.ico HTTP/1.1" 200
1318 "-" "Mozilla/5.0 (compatible; YandexBot/3.0;
+http://yandex.com/bots)"
vufind.carli.illinois.edu:80 0.0.0.0 - - [26/Jul/2016:00:00:03 -
0500] "GET /vf-wrh/Record/wrh_287760/favicon.ico HTTP/1.1" 200
1319 "-" "Mozilla/5.0 (compatible; YandexBot/3.0;
+http://yandex.com/bots)"
vufind.carli.illinois.edu:80 0.0.0.0 - - [26/Jul/2016:00:00:04 -
0500] "GET /vf-uiu/Record/uiu_944857 HTTP/1.1" 200 5747 "-"
"Mozilla/5.0 (compatible; Yahoo! Slurp;
```

Analyzing logs to make them Useful

- AWStats
- Unpleasant to install
- Requires full admin rights
- Requires IIS or Apache
- Requires Perl
- Can be run on Windows
- Can be run on a desktop
- http://www.awstats.org/



What does AWStats do?

- Finds all the matching needles in the haystack

We need the needles
that match our CARLI code

Most viewed records at EIU, 2/19/2017-4/19

Filter : /vf-eiu/Record/ Exclude filter : /all/vf-eiu/

Pages-URL

Filter **/vf-eiu/Record/** - Exclude Filter **/all/vf-eiu/**: 23,804 different pages-url	Viewed	Average size	Er
/vf-eiu/Record/AJAX	74,192	675 Bytes	
/vf-eiu/Record/eiu_1309074	156	6.13 KB	
/vf-eiu/Record/eiu_92052	129	6.04 KB	
/vf-eiu/Record/ebl_1441586	43	1.30 KB	
/vf-eiu/Record/ebl_1441789	43	1.27 KB	
/vf-eiu/Record/eiu_1330918	43	1.28 KB	
/vf-eiu/Record/eiu_1035353	41	1.05 KB	
/vf-eiu/Record/eiu_1073776/Holdings	41	1.03 KB	
/vf-eiu/Record/eiu_1361864	41	1.05 KB	
/vf-eiu/Record/eiu_81112/AJAX	34	674 Bytes	
/vf-eiu/Record/eiu_1262785	31	6.13 KB	
/vf-eiu/Record/eiu_719621/AJAX	27	592 Bytes	
/vf-eiu/Record/sms.cgi	25	1.27 KB	
/vf-eiu/Record/eiu_691160/AJAX	23	581 Bytes	
/vf-eiu/Record/eiu_1136915/Cite	23	6.00 KB	
/vf-eiu/Record/eiu_1354543/AJAX	22	509 Bytes	
/vf-eiu/Record/niu_756493/TOC	20	5.89 KB	
/vf-eiu/Record/niu_756493/AJAX	19	449 Bytes	
/vf-eiu/Record/eiu_1038843	18	6.21 KB	
/vf-eiu/Record/eiu_872005	18	6.35 KB	
/vf-eiu/Record/eiu_77338/AJAX	18	496 Bytes	
/vf-eiu/Record/eiu_860690	18	6.26 KB	

Click the link to see what it is

- **Encyclopedia of public relations (156)**
- **Business regulation in the Common Market nations (129)**
- **Hispanic-serving institutions : advancing research and transformative practice (43)**
- **Higher education in America (43 CARLI EBL)**
- **Higher education in the digital age (43 CARLI EBL)**

Other Log Sources

- EZ Proxy Access Log
- This will not be sanitized unless you sanitize it yourself
- Contains IP address

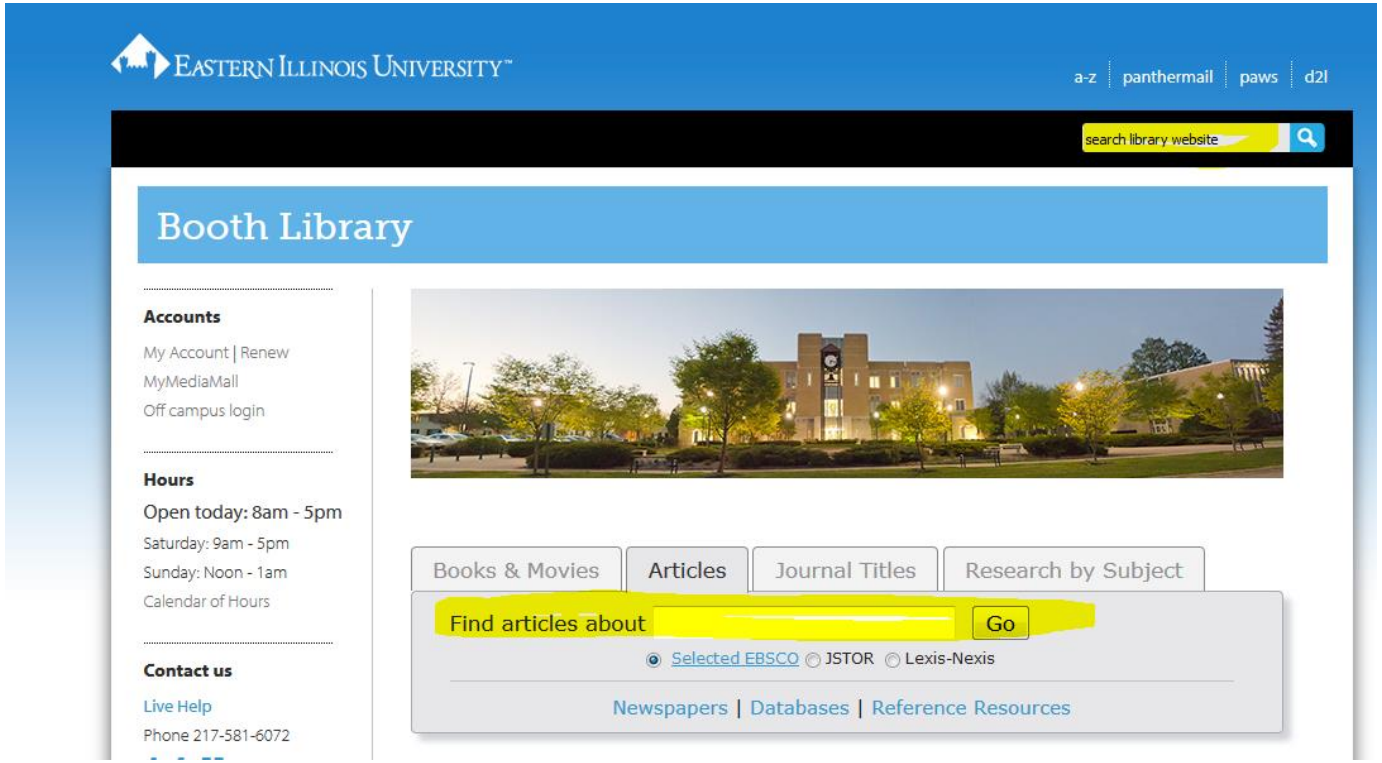
EZ Proxy logs can tell us

- What is used off campus
- What time of day users are online
- Top downloaded articles
- Usage by provider (requires some Excel work or several config files)

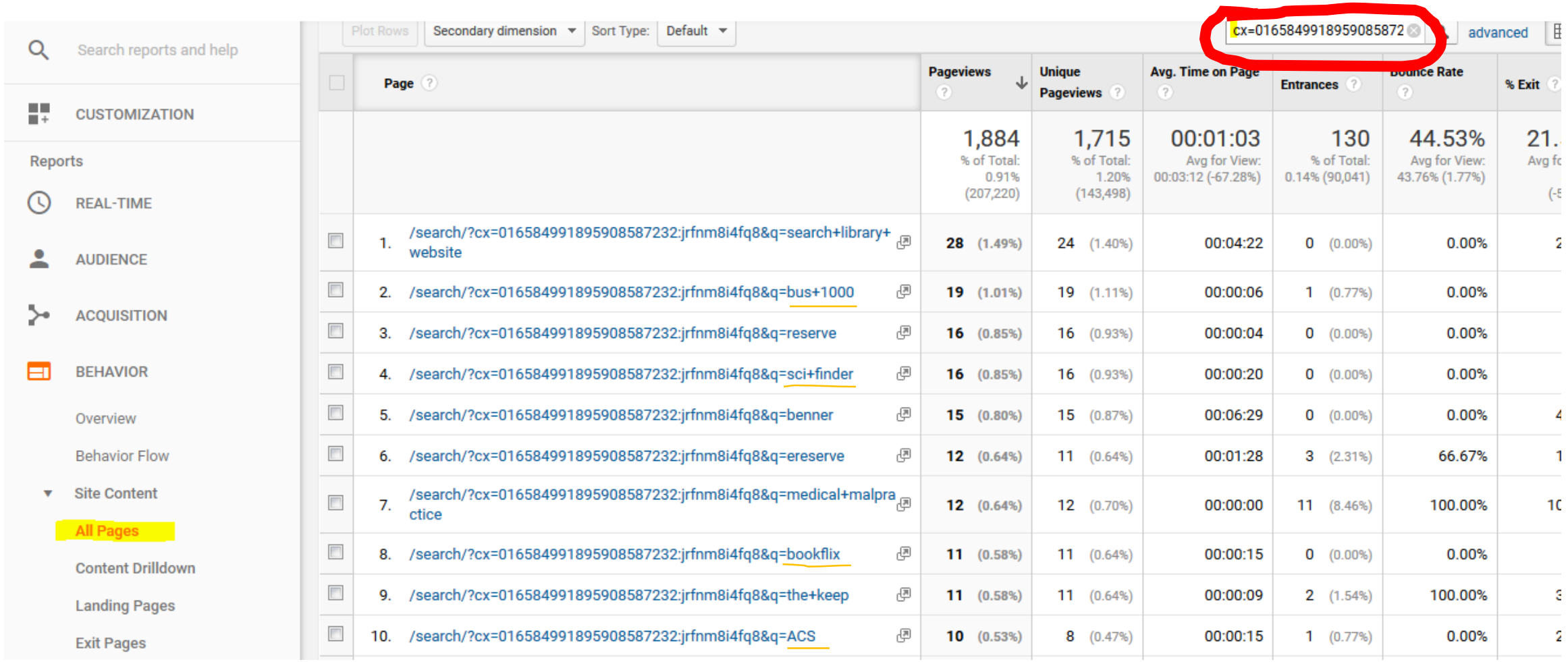
- AW Stats Config info
  - LogType=W
  - LogFormat=4

Special Bonus Feature: Google Analytics

Google Analytics can capture search terms from a locally hosted search box.



Go to All Pages, Filter for query string



Finding Search Terms from Article Searches

3.	/login?url=http://search.ebscohost.com/login.aspx?direct=true&authtype=ip&bQuery=critical thinking&db=a9h&db=tfh&db=hlh&db=bsh&db=pdh&db=ufh&db=hjh&db=hch	**20** (0.17%)	17 (0.16%)	00:15:53	0 (0.00%)	0.00%	75.00%	$0.00 (0.00%)
4.	/login?url=http://search.ebscohost.com/login.aspx?direct=true&authtype=ip&bQuery=arthritis&db=a9h&db=tfh&db=hlh&db=bsh&db=pdh&db=ufh&db=hjh&db=hch	**17** (0.14%)	15 (0.14%)	00:06:32	0 (0.00%)	0.00%	52.94%	$0.00 (0.00%)
5.	/login?url=http://search.ebscohost.com/login.aspx?direct=true&authtype=ip&bQuery=social media&db=a9h&db=tfh&db=hlh&db=bsh&db=pdh&db=ufh&db=hjh&db=hch	**16** (0.14%)	15 (0.14%)	00:07:07	2 (0.11%)	100.00%	81.25%	$0.00 (0.00%)
6.	/login?url=http://search.ebscohost.com/login.aspx?direct=true&authtype=ip&bQuery=leadership&db=a9h&db=tfh&db=hlh&db=bsh&db=pdh&db=ufh&db=hjh&db=hch	**15** (0.13%)	12 (0.11%)	00:08:58	0 (0.00%)	0.00%	73.33%	$0.00 (0.00%)
7.	/login?url=https://search.ebscohost.com/login.aspx?direct=true&authtype=ip&bQuery=nurse to patient ratio&db=a9h&db=tfh&db=hlh&db=bsh&db=pdh&db=ufh&db=hjh&db=hch	**14** (0.12%)	8 (0.07%)	00:00:16	8 (0.43%)	62.50%	35.71%	$0.00 (0.00%)

The end

- Questions?
- Comments?
- Presentation is posted on CARLI website