

**Eastern Illinois University**

---

**From the Selected Works 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/](https://works.bepress.com/stacey_knight-davis/62/)

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

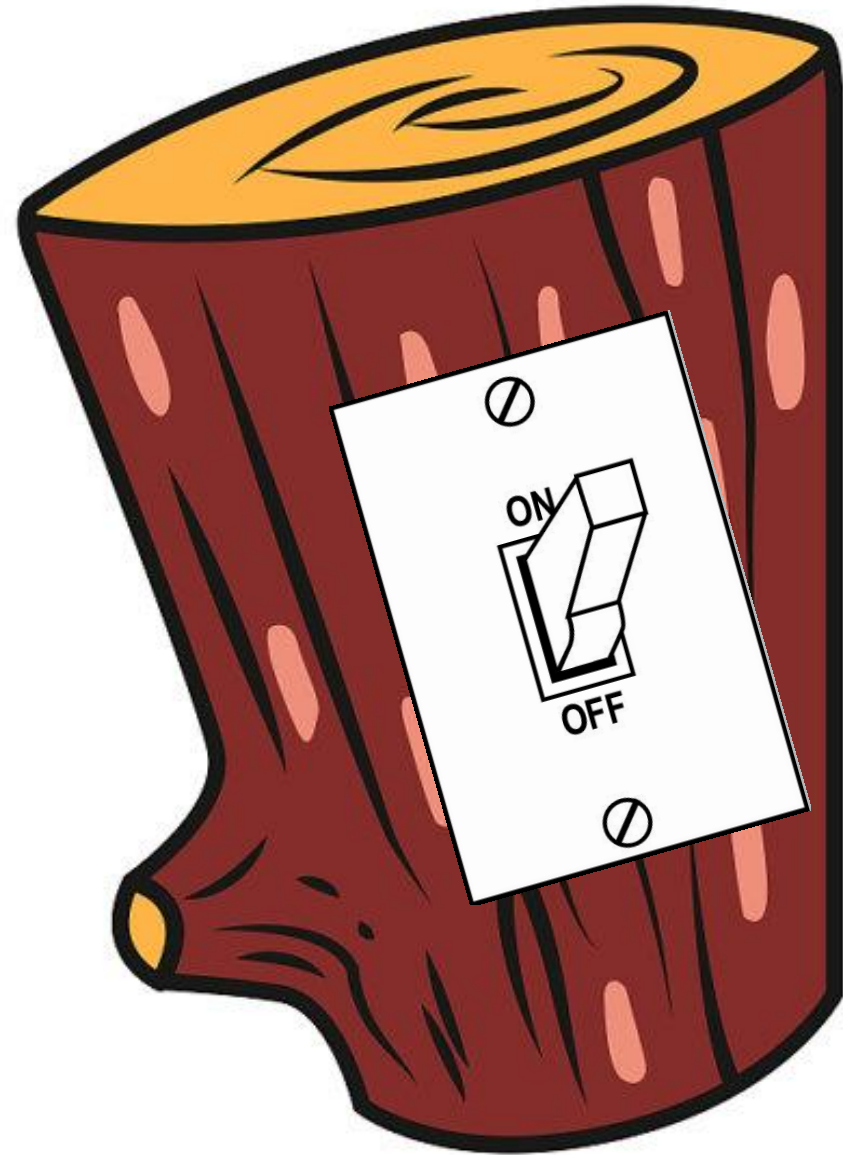
Stacey Knight-Davis  
Eastern Illinois University

# 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



Personally Identifying Information – Protect it!

# 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/>

### I-Share Log Files

vufind LogFormat:  
%v: %p %h %l %u %t \"%r\" %>s %O \"%{Referer}\" \"%v%\"

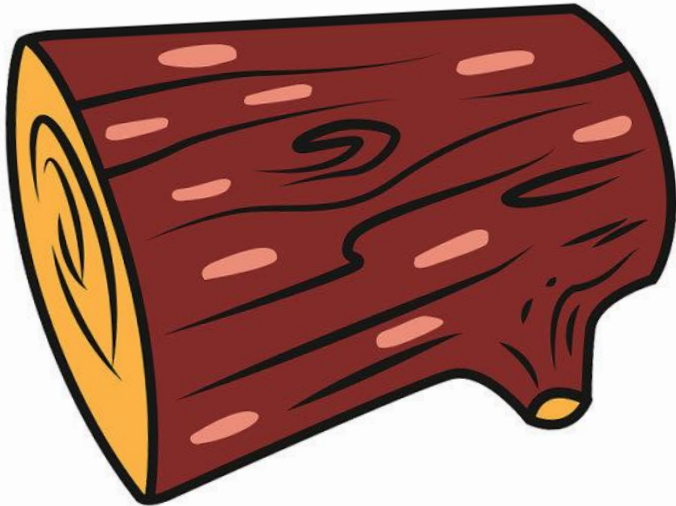
webvoyage LogFormat:  
%v: %p %h %l %u %t \"%r\" %>s %b \"%v%\"

| Name                          | Last m                |
|-------------------------------|-----------------------|
| Parent Directory              |                       |
| vufind_access_log.20170218.gz | 2017-02               |
| vufind_access_log.20170219.gz | 2017-02               |
| vufind_access_log.20170220.gz | 2017-02               |
| vufind_access_log.20170221.gz | 2017-02               |
| vufind_access_log.20170222.gz | 2017-02               |
| vufind_access_log.20170223.gz | 2017-02               |
| vufind_access_log.20170224.gz | 2017-02               |
| vufind_access_log.20170225.gz | 2017-02               |
| vufind_access_log.20170226.gz | 2017-02               |
| vufind_access_log.20170227.gz | 2017-02               |
| vufind_access_log.20170228.gz | 2017-02               |
| vufind_access_log.20170301.gz | 2017-03               |
| vufind_access_log.20170302.gz | 2017-03               |
| vufind_access_log.20170303.gz | 2017-03               |
| vufind_access_log.20170304.gz | 2017-03-05 00:11 9.7M |
| vufind_access_log.20170305.gz | 2017-03-06 00:11 12M  |
| vufind_access_log.20170306.gz | 2017-03-07 00:12 20M  |
| vufind_access_log.20170307.gz | 2017-03-08 00:12 22M  |
| vufind_access_log.20170308.gz | 2017-03-09 00:11 18M  |

| Download/Name                 | Progress | Per...               | Size     | Est. Time  | Speed |
|-------------------------------|----------|----------------------|----------|------------|-------|
| vufind_access_log.20170218.gz | 0%       | Unknown              |          | Skipped    |       |
| vufind_access_log.20170219.gz | 100%     | 14.07 MB             | Complete |            |       |
| vufind_access_log.20170220.gz | 100%     | 23.26 MB             | Complete |            |       |
| vufind_access_log.20170221.gz | 100%     | 24.31 MB             | Complete |            |       |
| vufind_access_log.20170222.gz | 100%     | 27.48 MB             | Complete |            |       |
| vufind_access_log.20170223.gz | 100%     | 22.35 MB             | Complete |            |       |
| vufind_access_log.20170224.gz | 100%     | 16.71 MB             | Complete |            |       |
| vufind_access_log.20170225.gz | 100%     | 13.39 MB             | Complete |            |       |
| vufind_access_log.20170226.gz | 100%     | 14.15 MB             | Complete |            |       |
| vufind_access_log.20170227.gz | 100%     | 23.22 MB             | Complete |            |       |
| vufind_access_log.20170228.gz | 95%      | 20.46 MB of 21.38 MB | 00:01    | 5.60 M...  |       |
| vufind_access_log.20170301.gz | 85%      | 18.38 MB of 21.54 MB | 00:01    | 10.33 M... |       |
| vufind_access_log.20170302.gz | 1%       | 269.2 KB of 25.38 MB | Starting | 0 B/s      |       |
| vufind_access_log.20170303.gz | N/A      | Unknown              | Starting | 0 B/s      |       |
| vufind_access_log.20170304.gz | 0%       | Unknown              | Waiting  |            |       |
| vufind_access_log.20170305.gz | 0%       | Unknown              | Waiting  |            |       |
| vufind_access_log.20170306.gz | 0%       | Unknown              | Waiting  |            |       |

Search downloads 11 of 122 (122), 4 running 15.97 MB/s Unlimited

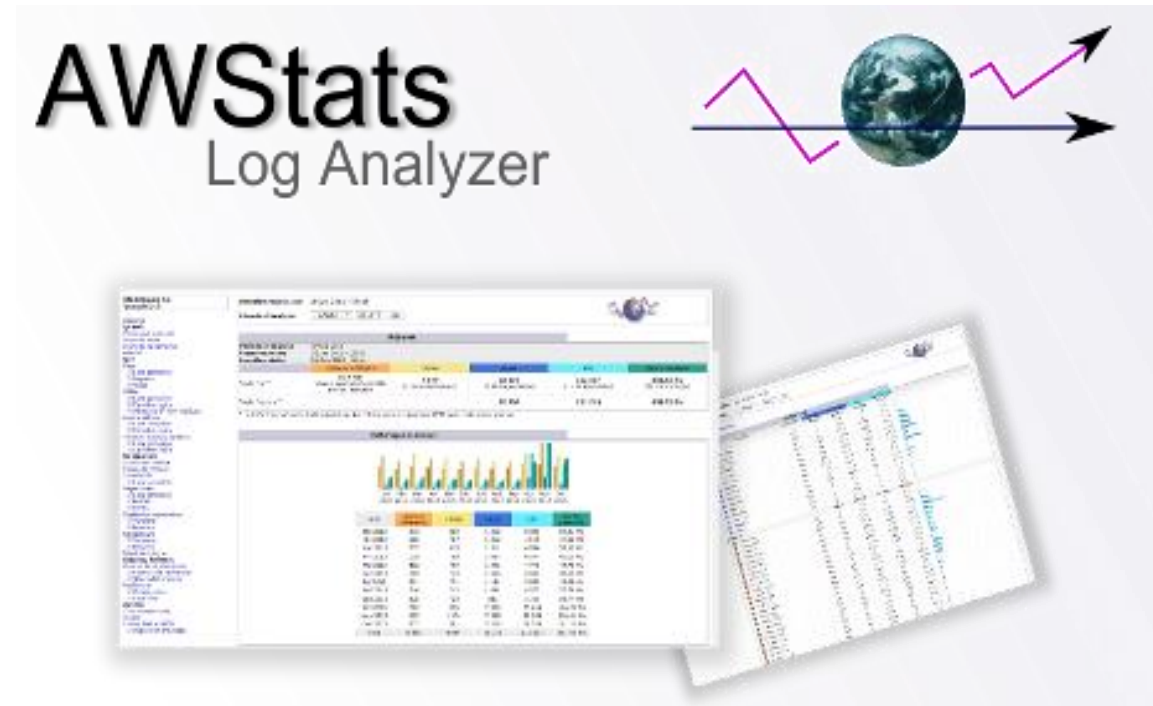
Without analysis, logs  
are about as useful as  
a chunk of wood



```
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 : <input type="text" value="/vf-eiu/Record/"/>   |  | Exclude filter : <input type="text" value="/all/vf-eiu/"/> |              |
|---|--|--|--------------|
| Pages-URL   |  |  |              |
| Filter <b>/vf-eiu/Record/</b> - Exclude Filter <b>/all/vf-eiu/</b> : 23,804 different pages-url |  | Viewed   | Average size |
| <a href="#">/vf-eiu/Record/AJAX</a>   |  | 74,192   | 675 Bytes    |
| <a href="#">/vf-eiu/Record/eiu_1309074</a>  |  | 156  | 6.13 KB      |
| <a href="#">/vf-eiu/Record/eiu_92052</a>  |  | 129  | 6.04 KB      |
| <a href="#">/vf-eiu/Record/ebl_1441586</a>  |  | 43   | 1.30 KB      |
| <a href="#">/vf-eiu/Record/ebl_1441789</a>  |  | 43   | 1.27 KB      |
| <a href="#">/vf-eiu/Record/eiu_1330918</a>  |  | 43   | 1.28 KB      |
| <a href="#">/vf-eiu/Record/eiu_1035353</a>  |  | 41   | 1.05 KB      |
| <a href="#">/vf-eiu/Record/eiu_1073776/Holdings</a>   |  | 41   | 1.03 KB      |
| <a href="#">/vf-eiu/Record/eiu_1361864</a>  |  | 41   | 1.05 KB      |
| <a href="#">/vf-eiu/Record/eiu_81112/AJAX</a>   |  | 34   | 674 Bytes    |
| <a href="#">/vf-eiu/Record/eiu_1262785</a>  |  | 31   | 6.13 KB      |
| <a href="#">/vf-eiu/Record/eiu_719621/AJAX</a>  |  | 27   | 592 Bytes    |
| <a href="#">/vf-eiu/Record/sms.cgi</a>  |  | 25   | 1.27 KB      |
| <a href="#">/vf-eiu/Record/eiu_691160/AJAX</a>  |  | 23   | 581 Bytes    |
| <a href="#">/vf-eiu/Record/eiu_1136915/Cite</a>   |  | 23   | 6.00 KB      |
| <a href="#">/vf-eiu/Record/eiu_1354543/AJAX</a>   |  | 22   | 509 Bytes    |
| <a href="#">/vf-eiu/Record/niu_756493/TOC</a>   |  | 20   | 5.89 KB      |
| <a href="#">/vf-eiu/Record/niu_756493/AJAX</a>  |  | 19   | 449 Bytes    |
| <a href="#">/vf-eiu/Record/eiu_1038843</a>  |  | 18   | 6.21 KB      |
| <a href="#">/vf-eiu/Record/eiu_872005</a>   |  | 18   | 6.35 KB      |
| <a href="#">/vf-eiu/Record/eiu_77338/AJAX</a>   |  | 18   | 496 Bytes    |
| <a href="#">/vf-eiu/Record/eiu_860690</a>   |  | 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

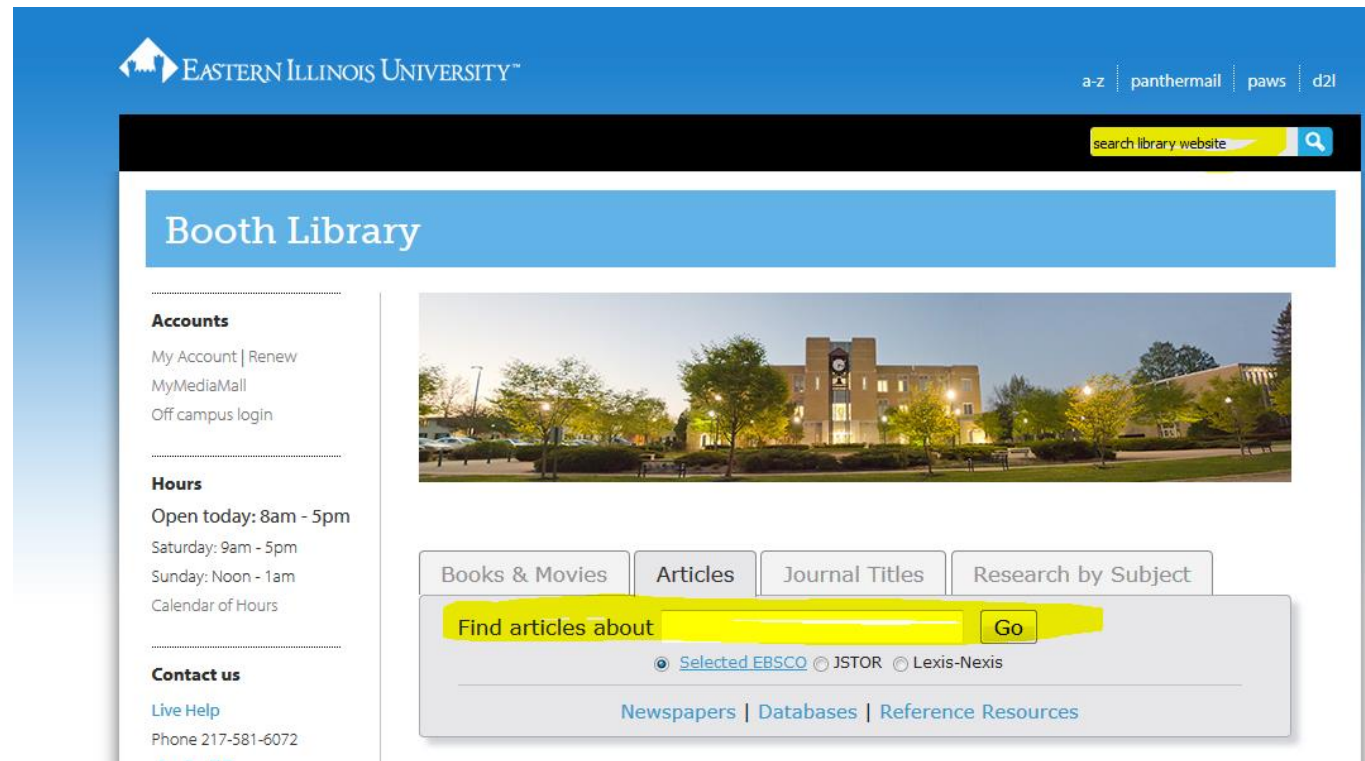
# EZ Proxy logs can tell us

- 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

Search reports and help

CUSTOMIZATION

Reports

REAL-TIME

AUDIENCE

ACQUISITION

BEHAVIOR

Overview

Behavior Flow

Site Content

All Pages

Content Drilldown

Landing Pages

Exit Pages

Plot Rows

Secondary dimension

Sort Type: Default

Page

Pageviews

Unique Pageviews

Avg. Time on Page

Entrances

Bounce Rate

% Exit

1,884  
% of Total:  
0.91%  
(207,220)

1,715  
% of Total:  
1.20%  
(143,498)

00:01:03  
Avg for View:  
00:03:12 (-67.28%)

130  
% of Total:  
0.14% (90,041)

44.53%  
Avg for View:  
43.76% (1.77%)

21.1%  
Avg for View:  
(-5.1%)

1. /search/?cx=016584991895908587232:jrfnm8i4fq8&q=search+library+website

28 (1.49%)

24 (1.40%)

00:04:22

0 (0.00%)

0.00%

2.1%

2. /search/?cx=016584991895908587232:jrfnm8i4fq8&q=bus+1000

19 (1.01%)

19 (1.11%)

00:00:06

1 (0.77%)

0.00%

1.1%

3. /search/?cx=016584991895908587232:jrfnm8i4fq8&q=reserve

16 (0.85%)

16 (0.93%)

00:00:04

0 (0.00%)

0.00%

0.8%

4. /search/?cx=016584991895908587232:jrfnm8i4fq8&q=sci+finder

16 (0.85%)

16 (0.93%)

00:00:20

0 (0.00%)

0.00%

0.8%

5. /search/?cx=016584991895908587232:jrfnm8i4fq8&q=benner

15 (0.80%)

15 (0.87%)

00:06:29

0 (0.00%)

0.00%

4.0%

6. /search/?cx=016584991895908587232:jrfnm8i4fq8&q=ereserve

12 (0.64%)

11 (0.64%)

00:01:28

3 (2.31%)

66.67%

1.1%

7. /search/?cx=016584991895908587232:jrfnm8i4fq8&q=medical+malpractice

12 (0.64%)

12 (0.70%)

00:00:00

11 (8.46%)

100.00%

10.0%

8. /search/?cx=016584991895908587232:jrfnm8i4fq8&q=bookflix

11 (0.58%)

11 (0.64%)

00:00:15

0 (0.00%)

0.00%

0.5%

9. /search/?cx=016584991895908587232:jrfnm8i4fq8&q=the+keep

11 (0.58%)

11 (0.64%)

00:00:09

2 (1.54%)

100.00%

3.0%

10. /search/?cx=016584991895908587232:jrfnm8i4fq8&q=ACS

10 (0.53%)

8 (0.47%)

00:00:15

1 (0.77%)

0.00%

2.0%

cx=0165849918959085872

advanced

# Finding Search Terms from Article Searches

|    |  |            |            |          |           |         |        |                |
|----|--|------------|------------|----------|-----------|---------|--------|----------------|
| 3. | <a href="#">/login?url=http://search.ebscohost.com/login.aspx?direct=true&amp;authtype=ip&amp;bQuery=critical thinking&amp;db=a9h&amp;db=tfh&amp;db=hlh&amp;db=bsh&amp;db=pdh&amp;db=ufh&amp;db=hjh&amp;db=hch</a>       | 20 (0.17%) | 17 (0.16%) | 00:15:53 | 0 (0.00%) | 0.00%   | 75.00% | \$0.00 (0.00%) |
| 4. | <a href="#">/login?url=http://search.ebscohost.com/login.aspx?direct=true&amp;authtype=ip&amp;bQuery=arthritis&amp;db=a9h&amp;db=tfh&amp;db=hlh&amp;db=bsh&amp;db=pdh&amp;db=ufh&amp;db=hjh&amp;db=hch</a>               | 17 (0.14%) | 15 (0.14%) | 00:06:32 | 0 (0.00%) | 0.00%   | 52.94% | \$0.00 (0.00%) |
| 5. | <a href="#">/login?url=http://search.ebscohost.com/login.aspx?direct=true&amp;authtype=ip&amp;bQuery=social media&amp;db=a9h&amp;db=tfh&amp;db=hlh&amp;db=bsh&amp;db=pdh&amp;db=ufh&amp;db=hjh&amp;db=hch</a>            | 16 (0.14%) | 15 (0.14%) | 00:07:07 | 2 (0.11%) | 100.00% | 81.25% | \$0.00 (0.00%) |
| 6. | <a href="#">/login?url=http://search.ebscohost.com/login.aspx?direct=true&amp;authtype=ip&amp;bQuery=leadership&amp;db=a9h&amp;db=tfh&amp;db=hlh&amp;db=bsh&amp;db=pdh&amp;db=ufh&amp;db=hjh&amp;db=hch</a>              | 15 (0.13%) | 12 (0.11%) | 00:08:58 | 0 (0.00%) | 0.00%   | 73.33% | \$0.00 (0.00%) |
| 7. | <a href="#">/login?url=https://search.ebscohost.com/login.aspx?direct=true&amp;authtype=ip&amp;bQuery=nurse to patient ratio&amp;db=a9h&amp;db=tfh&amp;db=hlh&amp;db=bsh&amp;db=pdh&amp;db=ufh&amp;db=hjh&amp;db=hch</a> | 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