

Oct 5th, 11:00 AM - 11:20 AM

## Detailed Search Stats from DSpace Solr

Eric Johnson

*Miami University - Oxford*, johnsoeo@miamioh.edu

Follow this and additional works at: <https://corescholar.libraries.wright.edu/irday>

 Part of the [Cataloging and Metadata Commons](#), [Scholarly Communication Commons](#), and the [Scholarly Publishing Commons](#)

---

### Repository Citation

Johnson, Eric, "Detailed Search Stats from DSpace Solr" (2016). *Ohio IR Day*. 3.  
[https://corescholar.libraries.wright.edu/irday/fall\\_2016/october5/3](https://corescholar.libraries.wright.edu/irday/fall_2016/october5/3)

This Event is brought to you for free and open access by the University Libraries' Events at CORE Scholar. It has been accepted for inclusion in Ohio IR Day by an authorized administrator of CORE Scholar. For more information, please contact [corescholar@www.libraries.wright.edu](mailto:corescholar@www.libraries.wright.edu), [library-corescholar@wright.edu](mailto:library-corescholar@wright.edu).

# Detailed search stats from DSpace Solr

Ohio IR day – 2016 October 5  
Eric Johnson  
Data Librarian, Miami University

# Repository Statistics

- How often is a resource being viewed?
- Where are the views coming from?  
(who is viewing the resource?)
- What is the trend? (increasing, decreasing)
- Statistics are used for funding requests.
- Statistics help us improve our selection of resources. (What types of items are people looking for?)

# Statistics

## Dspace statistics

### Total views for this community

Scripps Gerontology Center 2489

	April 2016	May 2016	June 2016	July 2016	August 2016	September 2016	October 2016
Scripps Gerontology Center	288	369	331	196	37	49	6

### Top country visits for this community

Japan 1169

United States 480

China 364

France 55

Germany 53

Canada 24

United Kingdom 16

Russia 15

Denmark 10

Austria 7

# Statistics

## Total views for this community

Scripps Gerontology Center 2489

Half year history

April 2016	May 2016	June 2016	July 2016	August 2016	September 2016	October 2016
8	369	331	196	37	49	6

## Top country visits for this community

Japan 1169

United States 480

China 364

France 55

Germany 53

Canada 24

United Kingdom 16

Russia 15

Denmark 10

Austria 7

# Statistics

## Total views for this community

Scripps Gerontology Center 2489

	April 2016	May 2016	June 2016	July 2016	August 2016	September 2016	October 2016
Scripps Gerontology Center	288	369	331	196	37	49	6

## Top country visits for this community

Japan	1169
United States	480
China	364
France	55
Germany	53
Canada	24
United Kingdom	16
Russia	15
Denmark	10
Austria	7

Popular in Japan!

# Search Statistics for Scripps Gerontology Center

## Top Search Terms

Overall

	Search Term	Searches	% of Total	Pageviews / Search
1	author_keyword:Applebaum, Robert	3789	12.67%	0.00
2	subject_keyword:PASSPORT	3425	11.46%	0.00
3	dateIssued_keyword:[2000 TO 2009]	3282	10.98%	0.00
4	author_keyword:Straker, Jane	3078	10.30%	0.00
5	author_keyword:Mehdizadeh, Shahla	2869	9.60%	0.00
6	subject_keyword:long-term care	2527	8.45%	0.00
7	subject_keyword:nursing homes	2487	8.32%	0.00
8	author_keyword:Kunkel, Suzanne	2466	8.25%	0.00
9	dateIssued_keyword:[2010 TO 2015]	2429	8.13%	0.00
10	subject_keyword:program evaluation	2346	7.85%	0.00

Over the  
last 6  
months

## Total

Searches	% of Total	Pageviews / Search
29894	100.00%	0.00

# Client's need

- Wanted detailed statistics not just an “overview”.
- Statistics for particular time periods
- Statistics for particular publications
- Needed for grant application submissions



# SOLR

- DSpace uses SOLR to index search and download actions.
- SOLR is a search platform developed by Apache and integrated into the Dspace repository software.



Dashboard

Logging

Core Admin

Java Properties

Thread Dump

statistics

Overview

Analysis

Dataimport

Documents

Files

Ping

Plugins / Stats

Query

Replication

Schema Browser



# SOLR dashboard

- Comes with Dspace
- Allows advanced queries



Apache Solr Documentation

6.2 Ref Guide (PDF Download)

Solr Tutorial

Solr Community Wiki

Older Versions of this Guide (PDF)

6.3 Draft Ref Guide Topics

- **About This Guide**

- › Getting Started

- Upgrading Solr

- › Using the Solr Administration User Interface

- › Documents, Fields, and Schema Design

- › Understanding Analyzers, Tokenizers, and Filters

- › Indexing and Basic Data Operations

- › Searching

- › The Well-Configured Solr Instance

- › Managing Solr

- › SolrCloud

- › Legacy Scaling and Distribution

- › Client APIs



# About This Guide

Created by Cassandra Targett, last modified on Feb 20, 2015

This guide describes all of the important features and functions of Apache Solr. download from <http://lucene.apache.org/solr/>.

Designed to provide high-level documentation, this guide is intended to be more less of a cookbook. It is structured to address a broad spectrum of needs, ranging from developers getting started to well-experienced developers extending their applications to troubleshooting. It will be of use at any point in the application life cycle, for where authoritative information about Solr.

The material as presented assumes that you are familiar with some basic search you can read XML. It does not assume that you are a Java programmer, although Java is helpful when working directly with Lucene or when developing custom extensions to a Lucene/Solr installation.

## Special Inline Notes

Special notes are included throughout these pages.

Note Type	Look & Description

# SOLR API

- If you want to programmatically access the SOLR data, you create URLs and parse the results.

```
http://MyDSpace/solr/search/select/?q=location.comm  
:25+AND+search.resourcetype:2&start=0&rows=3
```

- Data is returned in XML format

```
<doc>
```

```
<arr name="dc.title">
```

```
<str>
```

```
Implementation of the 2006 Ohio nursing home family  
satisfaction survey: Final report
```

```
</str>
```

# The Request – list all metadata for items in a particular community

- <http://myDSpace/solr/search/select/?> — The URL of your repository (SOLR server)
- `q=location.comm:25` — Community ID#
- `+AND+search.resourcetype:2` — Type = item (vs. collection, community, bitstream, etc.)
- `&start=0` — Which rows of the response should be sent with this request. Useful for long responses.
- `&rows=3`

```

▼<response>
  ▼<lst name="responseHeader">
    <int name="status">0</int>
    <int name="QTime">1</int>
    ▼<lst name="params">
      <str name="start">0</str>
      <str name="q">location.comm:25 AND search.resourcetype:2</str>
      <str name="rows">3</str>
      <str name="version">2.2</str>
    </lst>
  </lst>
▼<result name="response" numFound="243" sta
  ►<doc>...</doc>
  ▼<doc>
    ▼<arr name="dc.title_mlt">
      ▼<str>
        Implementation of the 2006 Ohio nursing home family satisfa
      </str>
    </arr>
    ▼<arr name="author">
      <str>Straker, Jane</str>
      <str>Ejaz, Farida</str>
    </arr>
    ▼<arr name="author_keyword">
      <str>Straker, Jane</str>
      <str>Ejaz, Farida</str>
    </arr>

```

# The response

Each <doc> is a separate item

- XML format
- Search the tags to retrieve information.

```
▼<response>
  ▼<lst name="responseHeader">
    <int name="status">0</int>
    <int name="QTime">1</int>
    ▼<lst name="params">
      <str name="start">0</str>
      <str name="q">location.comm:25 AND search
      <str name="rows">3</str>
      <str name="version">2.2</str>
    </lst>
  </lst>
  ▼<result name="response" numFound="243" start="0">
    ►<doc>...</doc>
    ▼<doc>
      ▼<arr name="author">
        ▼<str>
          nursing home family satisfa
        </str>
      </arr>
      ▼<arr name="author_keyword">
        <str>Straker, Jane</str>
        <str>Ejaz, Farida</str>
      </arr>
    </doc>
  </result>
</response>
```

- Use “start” and “rows” to download the responses in batches

- “numFound” is the total number of responses

# Tracking of each search action

```
▼ <result name="response" numFound="38744" sta
  ▶ <doc>...</doc>
  ▼ <doc>
    <str name="ip">74.62.104.98</str>
    <str name="referrer">http://sc.lib.muohio.edu/handle/2374.MIA/161</str>
    <str name="dns">rrcs-74-62-104-98.west.biz.rr.com.</str>
    <str name="continent">NA</str>
    <str name="countryCode">US</str>
    <str name="city">Overland Park</str>
    <float name="latitude">38.982193</float>
    <float name="longitude">-94.6708</float>
    <bool name="isBot">>false</bool>
  ▼ <str name="userAgent">
    Mozilla/5.0 (Windows NT 6.1; WOW64; rv:35.0) Gecko/20100101 Firefox/35.0
    </str>
    <int name="id">1820</int>
    <int name="type">0</int>
  ▼ <arr name="owningItem">
    <int>653</int>
  </arr>
  ▼ <arr name="owningColl">
    <int>19</int>
  </arr>
  ▼ <arr name="owningComm">
```

Total # of searches during that time

Address of each searcher

Location of searcher:  
Continent, Country, city.  
Latitude & Longitude.

Was the searcher a robot or human?

ID # of this search

Item # that was returned



# Workflow

- Get parameters (date range, collection)
- Identify month periods
- Get a list of items in that collection
- Find the number of search requests for each item in each time period
- Output a TSV (Tab separated variables) file
- (CSV won't work because some titles contain commas)

# Interface

The screenshot shows a web application window titled "SOLR activity downloader". The interface includes two date input fields with calendar icons, a "To" label, a dropdown menu, a progress indicator, and two buttons.

09/04/2013 To 10/05/2016

Scripps Gerontology Center: community 25

0%

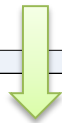
doing query # 21

# Result

Title	Author(s)	ID#	URL	2015-03-01 to 2015-04-01	2015-04-01 to 2015-05-01	2015-05-01 to 2015-06-01
2011 Title VI Survey Results	Carpio, Elizabeth   Mc	707	<a href="http://hdl.handle.net/10120/1.0/MIA/4524">l.MIA/4524</a>	9	40	34
2014 Title VI Survey Results	Reece, Heather R.   M	1794	<a href="http://hdl.handle.net/10120/1.0/MIA/4524">http://hdl.</a>	0	0	0
A Guide to quality in consumer	Applebaum, Robert	645	<a href="http://hdl.handle.net/10120/1.0/MIA/4524">http://hdl.</a>	20	46	30
A Review of nursing home resi	Mehdizadeh, Shahla	723	<a href="http://hdl.handle.net/10120/1.0/MIA/4524">http://hdl.</a>	9	34	23
A Study of Adult Guardianship	Reese, Heather R. Rol	1796	<a href="http://hdl.handle.net/10120/1.0/MIA/4524">http://hdl.</a>	0	0	0
A study of home care and nursi	Applebaum, Robert	588	<a href="http://hdl.handle.net/10120/1.0/MIA/4524">http://hdl.</a>	8	7	0
A survey of non-certified and n	Straker, JaneAppleba	602	<a href="http://hdl.handle.net/10120/1.0/MIA/4524">http://hdl.</a>	18	6	7
A ten year retrospective look a	Mehdizadeh, Shahla	612	<a href="http://hdl.handle.net/10120/1.0/MIA/4524">http://hdl.</a>	36	4	0
AAA Caregiving Survey Results	Straker, Jane   Lackme	684	<a href="http://hdl.handle.net/10120/1.0/MIA/4524">http://hdl.</a>	17	39	25
Adult Education and Training P	Cummins, Phyllis A.	1072	<a href="http://hdl.handle.net/10120/1.0/MIA/4524">http://hdl.</a>	40	59	121
Adult care homes in Ohio	Applebaum, RobertR	780	<a href="http://hdl.handle.net/10120/1.0/MIA/4524">http://hdl.</a>	7	1	0
African American elders' choice	Groger, Lisa	599	<a href="http://hdl.handle.net/10120/1.0/MIA/4524">http://hdl.</a>	15	10	2
African American elders' long-t	Groger, Lisa	770	<a href="http://hdl.handle.net/10120/1.0/MIA/4524">http://hdl.</a>	20	18	11
Aging and Long-Term Care Issu	Payne, Michael   Stral	626	<a href="http://hdl.handle.net/10120/1.0/MIA/4524">http://hdl.</a>	16	36	26
An Evaluation of the Assisted L	McGrew, Kathryn   Me	630	<a href="http://hdl.handle.net/10120/1.0/MIA/4524">http://hdl.</a>	19	43	27

# Much more informative now

April 2016	May 2016	June 2016	July 2016	August 2016	September 2016	October 2016
288	369	331	196	37	49	6



Title	Author(s)	ID#	URL	2015-03-01 to 2015-04-01	2015-04-01 to 2015-05-01	2015-05-01 to 2015-06-01
2011 Title VI Survey Results	Carpio, Elizabeth   Mc	707	<a href="http://hdl.handle.net/10115/10115.1.MIA/4524">10115.1.MIA/4524</a>	9	40	34
2014 Title VI Survey Results	Reece, Heather R.   M	1794	<a href="http://hdl.handle.net/10115/10115.1.MIA/4524">http://hdl.</a>	0	0	0
A Guide to quality in consumer	Applebaum, Robert	645	<a href="http://hdl.handle.net/10115/10115.1.MIA/4524">http://hdl.</a>	20	46	30
A Review of nursing home resi	Mehdizadeh, Shahla	723	<a href="http://hdl.handle.net/10115/10115.1.MIA/4524">http://hdl.</a>	9	34	23
A Study of Adult Guardianship	Reese, Heather R. Rol	1796	<a href="http://hdl.handle.net/10115/10115.1.MIA/4524">http://hdl.</a>	0	0	0
A study of home care and nursi	Applebaum, Robert	588	<a href="http://hdl.handle.net/10115/10115.1.MIA/4524">http://hdl.</a>	8	7	0
A survey of non-certified and n	Straker, Jane Appleba	602	<a href="http://hdl.handle.net/10115/10115.1.MIA/4524">http://hdl.</a>	18	6	7
A ten year retrospective look a	Mehdizadeh, Shahla	612	<a href="http://hdl.handle.net/10115/10115.1.MIA/4524">http://hdl.</a>	36	4	0
AAA Caregiving Survey Results	Straker, Jane   Lackme	684	<a href="http://hdl.handle.net/10115/10115.1.MIA/4524">http://hdl.</a>	17	39	25
Adult Education and Training P	Cummins, Phyllis A.	1072	<a href="http://hdl.handle.net/10115/10115.1.MIA/4524">http://hdl.</a>	40	59	121
Adult care homes in Ohio	Applebaum, Robert R	780	<a href="http://hdl.handle.net/10115/10115.1.MIA/4524">http://hdl.</a>	7	1	0
African American elders' choice	Groger, Lisa	599	<a href="http://hdl.handle.net/10115/10115.1.MIA/4524">http://hdl.</a>	15	10	2

# Questions?

Eric Johnson

Miami University

[johnsoeo@miamioh.edu](mailto:johnsoeo@miamioh.edu)