

DSPACE STATISTICS: GOOGLE ANALYTICS, SOLR

Mukesh Pund

Principal Scientist

CSIR-NISCAIR

Outline:

- **Google Analytic**
- **Solr Statistics**
- **Dspace Community/Collection/Item Statistics**
- **Dspace Discover**

Dspace Google Analytics

- **To record User Interface traffic and downloads from Dspace**
- **Until Dspace 5.0 only User Interface activity could be recorded , means downloads initiated straight from a Google search were not recorded.**
- **In Dspace 5.0, it has also become possible to expose that recorded Google Analytics data within Dspace in XMLUI Mirage2 but not in JSPUI**

Dspace Google Analytics

- Go to your Google Analytics dashboard <http://www.google.com/analytics/>
- Login into with your Google credentials
- Sign up for Google analytics
- Add new account and provide the website/Dspace address which you want to track
- Get the tracking ID and accept Google agreement

Note: You have to provide the public address of website/Dspace



New Account

What would you like to track? _____

Website Mobile app

Tracking Method _____

This property works using Universal Analytics. Click *Get Tracking ID* and implement the Universal Analytics tracking code snippet to complete your set up.

Setting up your account _____

Account Name required

Accounts are the top-most level of organisation and contain one or more tracking IDs.

Setting up your property _____

Website Name required

Website URL required

Industry Category ?

Reporting Time Zone

Download bar showing files: DSpace-1.8-Databas..., ds.rtf, DSpace_3.x-Manual.pdf, Basic DSpace Tutoria...

Show all downloads...

Administration
testdspace / niscair

PROPERTY
niscair

- Property Settings
- User Management
- Tracking Info**
 - Tracking Code
 - User-ID
 - Session Settings
 - Organic Search Sources
 - Referral Exclusion List
 - Search Term Exclusion List

PRODUCT LINKING

- AdWords Linking
- AdSense Linking
- All Products
- Remarketing
- Custom Definitions

ANALYTICS EDUCATION

Find your basic tracking code, and additional data-collection settings.

The following settings are available only in Universal Analytics.

Click **Tracking Code** to find the basic code snippet for a website or to download platform SDKs for an app (websites & apps).

Click **Session Settings** to configure the length of a session and of a campaign (websites & apps).

An incorrect session-length setting can have a negative impact on visit-based analysis; for example, if your users would normally have more than 30 minutes of inactivity during a session, a shorter timeout would divide single-session data into multiple sessions and provide an inaccurate picture of activity.

Set campaign timeout to correspond to the length of time you want to be able to attribute a visit or conversion to one of your campaigns. For example, if you run seasonal campaigns, set this value to 3 Months; if you run weekly campaigns, set this value to 7 Days.

Tracking ID
UA-59145654-1

Website tracking

This is the Universal Analytics tracking code for this property. To get all the benefits of Universal Analytics for this property, copy and paste this code into every web page that you want to track.

This is your tracking code. Copy and paste it into the code of every page that you want to track.

```
<script>
(function(i,s,o,g,r,a,m){[i['GoogleAnalyticsObject']=r,i[r]=i[r]||function(){
  (i[r].q=i[r].q||[]).push(arguments)},i[r].l=1*new Date(),a=s.createElement(o),
  m=s.getElementsByTagName(o)[0];a.async=1;a.src=g;m.parentNode.insertBefore(a,m)
})(window,document,'script','/www.google-analytics.com/analytics.js','ga');
```

Dspace Google Analytics

- **Copy the tracking Id looking like UA-XXXXXXXX-X**
- **Open the dspace.cfg file in your Dspace directory**
- **Find for Google analytics code :**
jspui.google.analytics.key
- **Uncomment this line**
- **Paste the Google tracking Id in front of this code**
- **Restart Tomcat**
- **In Google analytics account , you can view the traffic in different forms (charts, reports etc.) after some time**
- **Now it's done!!**

Dspace Google Analytics

Screen shot of /dspace/config/dspace.cfg

```
1935 #
1936 # webui.session.invalidate = true
1937
1938 # If you would like to use Google Analytics to track general website statistics then
1939 # use the following parameter to provide your Analytics key. First sign up for an
1940 # account at http://analytics.google.com, then create an entry for your repository
1941 # website. Analytics will give you a snippet of JavaScript code to place on your site,
1942 # inside that snippet is your Google Analytics key usually found in this line:
1943 # _uacct = "UA-XXXXXXX-X"
1944 # Take this key (just the UA-XXXXXXX-X part) and place it here in this parameter.
1945 # jspui.google.analytics.key=UA-XXXXXX-X
1946 jspui.google.analytics.key=UA-591445654-1
1947
1948 #-----#
1949 #-----XMLUI SPECIFIC CONFIGURATIONS-----#
1950 #-----#
1951 # These configs are only used by the XML User Interface #
1952 #-----#
```

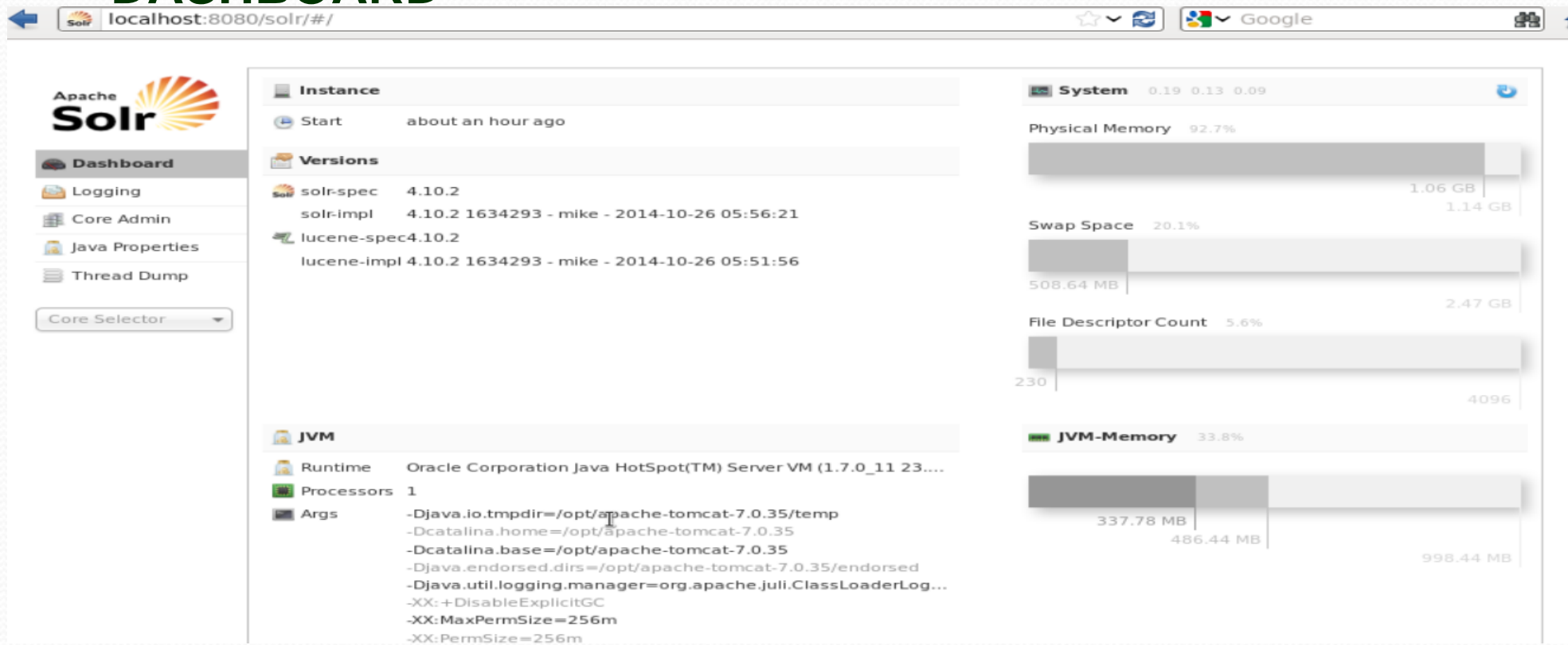

Solr Statistics

- For Solr UI access, type in browser

`http://server_ip_or_dns:port_number/solr`

(default permission solr access on localhost)

DASHBOARD



Solr Statistics

- e. g of Plugin/Statistics for current running server in cache

The screenshot shows the Apache Solr Admin UI. The browser address bar displays `localhost:8080/solr/#/statistics/plugins/cache?entry=documentCache`. The left sidebar contains the Solr logo and a navigation menu with items like Dashboard, Logging, Core Admin, Java Properties, Thread Dump, and a dropdown menu for 'statistics' which is currently open to show Overview, Analysis, Dataimport, Documents, Files, Ping, Plugins / Stats (selected), Query, Replication, and Schema Browser.

The main content area is divided into two sections. The left section, titled 'CACHE', lists various plugins: CORE, HIGHLIGHTING, OTHER, QUERYHANDLER, QUERYPARSER, UPDATEHANDLER, Watch Changes, and Refresh Values. The right section, titled 'documentCache', displays the following details:

```
class: org.apache.solr.search.LRUCache
version: 1.0
description: LRU Cache(maxSize=512, initialSize=512)
src: null
```

```
stats:
  lookups: 0
  hits: 0
  hitratio: 0
  inserts: 0
  evictions: 0
  size: 0
  warmupTime: 0
  cumulative_lookups: 7
  cumulative_hits: 0
  cumulative_hitratio: 0
  cumulative_inserts: 7
  cumulative_evictions: 0
```

Below the documentCache section, there are four checked checkboxes for other cache types: fieldCache, fieldValueCache, filterCache, and queryResultCache.

The footer of the page contains several links: Documentation, Issue Tracker, IRC Channel, Community forum, and Solr Query Syntax.

Solr Statistics

- e. g of Analysis Report in Solr

The screenshot shows the Apache Solr Administration interface. The browser address bar displays the URL: localhost:8080/solr/#/statistics/analysis?analysis.fieldvalue=India&analysis.fieldname=city&. The page features a sidebar on the left with navigation options: Dashboard, Logging, Core Admin, Java Properties, Thread Dump, and a 'statistics' dropdown menu. The 'Analysis' option is selected. The main content area is divided into several sections:

- Field Value (Index):** A text input field containing 'India'.
- Field Value (Query):** An empty text input field.
- Analyse Fieldname / FieldType:** A dropdown menu set to 'city'.
- Verbose Output:** A checked checkbox.
- Analyse Values:** A blue button to execute the analysis.
- Analysis Results:** A table showing the analysis of the field 'city' with the value 'India'.

DA	text	India
	raw_bytes	[49 6e 64 69 61]
	start	0
	end	5
	positionLength	1
	type	word
	position	1

At the bottom of the page, there are links to Documentation, Issue Tracker, IRC Channel, Community forum, and Solr Query Syntax.

DSpace Statistics

- Default statistics details of Communities/Collections/Items visible public (for non restricted items, i.e. Item Metadata and bitstream accessible by Anonymous user or specified group)

The screenshot shows the DSpace JSPUI interface. At the top, there is a navigation bar with links for Home, Browse, and Help, a search box labeled 'Search DSpace', and options for 'Sign on to:' and 'Language'. Below the navigation bar, the main content area displays the DSpace logo and the text 'DSpace JSPUI' followed by a description: 'Test5 कमल घर ज DSpace preserves and enables easy and open access to all types of digital content including text, images, moving images, mpegs and data sets'. A 'Learn More' button is present. A green breadcrumb trail indicates the current location: 'DSpace Authorization / Information Sciences / Cloud Department'. The main content area features a collection entry for 'Enterprise solutions : [1] Collection Home page'. A small blue icon with a bar chart, representing statistics, is circled in red. Below this, there is a 'Browse' section with buttons for 'Issue Date', 'Author', 'Title', and 'Subject'. On the right side, there is a 'Discover' section with input fields for 'Author' (containing 'author, lst') and 'Date issued' (containing '1992'), both with a '1' in a circle next to them.

DSpace Statistics

- Statistics details display as

DSpace JSPUI navigation bar with Home, Browse, Help, Search DSpace, Sign on to, and Language options.

DSpace JSPUI

Test5 कमल घर ज DSpace preserves and enables easy and open access to all types of digital content including text, images, moving images, mpegs and data sets

[Learn More](#)



DSpace Authorization / Information Sciences / Cloud Department / Enterprise solutions

Statistics

Total Visits

Views

Enterprise solutions 4

Total Visits per Month

August 2014 September 2014 October 2014 November 2014 December 2014 January 2015 February 2015

Enterprise solutions 0 0 0 0 0 4 0

DSpace Statistics

- For customizing default dspace statistics (e.g.: from start date of statistics) in /dspace/config/dstat.cfg
- Dspace indexing (set in cron-job)
 - Also need to re-index discovery
/dspace/bin/dspace index-discovery
 - Indexing optimize
/dspace/bin/dspace index-discovery -o

```
# DSpace information

# the year and month to start creating reports from
# - year as four digits (e.g. 2005)
# - month as a number (e.g. January is 1, December is 12)
start.year = 2005
start.month = 1

# actions to place in the general summary
general.summary=archive_item
general.summary=view_bitstream
general.summary=view_item
```

DSpace Statistics

- Other Statistics command (set in cron-job)

```
# Cleanup Web Spiders from DSpace Statistics Solr  
/dspace/bin/dspace stats-util -i
```

```
# Re-Optimize DSpace Statistics Solr Index  
/dspace/bin/dspace stats-util -o
```

```
# This ensures each year has its own Solr index, which  
improves performance.
```

```
/bin/dspace stats-util -s
```

DSpace Discovery

After indexing Discover (set in cron-job), display discovery in Home page and left side of search

/dspace/bin/dspace index-discovery

Indexing optimize

/dspace/bin/dspace index-discovery -o

Navigation bar with Home, Browse, Help, Search DSpace, Sign on to, and Language options.

More exciting news to appear here.

This is a default installation of DSpace!

It can be extensively configured by installing modified JSPs, and altering the site configuration.

Communities in DSpace

Choose a community to browse its collections.

Information Sciences 1

Discover

Author

author, lst 1

Date issued

1992 1

Enable DSpace Statistics

Statistics tab is present under Administer in JSPUI and Administrative in XMLUI

The following scripts must be run (in this order) generate the statistics:

```
/dspace/bin/dspace stat-initial  
/dspace/bin/dspace stat-general  
/dspace/bin/dspace stat-monthly  
/dspace/bin/dspace stat-report-initial  
/dspace/bin/dspace stat-report-general  
/dspace/bin/dspace stat-report-monthly
```

Thanks