

Introduction to



erik.hatcher

@

lucidimagination.com



Abstract

Apache Solr serves search requests at the enterprises and the largest companies around the world. Built on top of the top-notch Apache Lucene library, Solr makes indexing and searching integration into your applications straightforward.

Solr provides faceted navigation, spell checking, highlighting, clustering, grouping, and other search features. Solr also scales query volume with replication and collection size with distributed capabilities. Solr can index rich documents such as PDF, Word, HTML, and other file types.

Come learn how you can get your content into Solr and integrate it into your applications!

About me...

lucid

IMAGINATION





<http://lucene.apache.org/solr/>



<http://lucene.apache.org/>

Solr

- **Simple**: easy to use
- **Powerful**: feature rich and scales
- **Open Source**
 - from the "Lucene people"
 - encapsulates Lucene best practices

Fire It Up

- `cd example`
- `java -jar start.jar`
- `[cd example/exampledocs;`
 - `java -jar post.jar *.xml]`

Indexing

- `/update[/csv|/json|/extract]`
- stream from local, remote, or POST data
- tutorial:
 - `cd example/exampledocs`
 - `java -jar post.jar *.xml`
 - Tip: `java -jar post.jar -help`

Indexing JSON

POST to /update/json

```
[  
  {"id" : "1", "title" : "Doc One"},  
  {"id" : "2", "title" : "Doc Two"}  
]
```

Indexing CSV

```
curl http://localhost:8983/solr/update/csv  
--data-binary @data.csv  
-H 'Content-type:text/plain; charset=utf-8'
```

Indexing Rich Documents

```
http://localhost:8983/solr/update/extract  
?stream.file=/path/to/file.doc  
&stream.contentType=application/msword  
&literal.id=ds1-file.doc"
```

Other conduits

- DataImportHandler (DIH)
- API's: SolrJ, RSolr, (py)solr(.py), etc
 - It's just data over HTTP
- "Enterprise"
 - LucidWorks: SharePoint, (split) crawling, S3, HDFS, etc; including access control

Searching

- `http://localhost:8983/solr/select?q=*.*`
- Typical looking Solr request - `http://localhost:8983/solr/select +`
 - `?q=ipod`
 - `&facet=on`
 - `&facet.field=cat`
 - `&fq=cat:electronics`
 - `[&rows=10&start=20]`
 - `[&fl=id,name,price&sort=price asc]`
 - `[&wt=xml|json|csv|ruby|python|php|xslt|velocity&indent=on]`
 - `[&debugQuery=true]`



/browse

Examples: [Simple](#) [Spatial](#)

Find:

[Submit Query](#)

[Reset](#)

Boost by Price

Location Filter:

Distance (KM):

Field Facets

cat

[Electronics](#) (3)

[Connector](#) (2)

[Music](#) (1)

manu_exact

[Belkin](#) (2)

[Apple Computer Inc.](#) (1)

Query Facets

[ipod](#) (3)

[GB](#) (1)

Range Facets

price

[0.0 - 50.0](#) (2)

[350.0 - 400.0](#) (1)

popularity

[0 - 3](#) (2)

[9 - 12](#) (1)

3 results found in 9 ms Page 1 of 1

iPod & iPod Mini USB 2.0 Cable [More Like This](#)

Price: \$11.50

Features: car power adapter for *iPod*, white

In Stock: false



[Larger Map](#)

Belkin Mobile Power Cord for iPod w/ Dock [More Like This](#)

Price: \$19.95

Features: car power adapter, white

In Stock: false



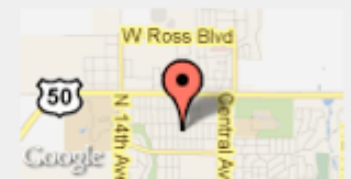
[Larger Map](#)

Apple 60 GB iPod with Video Playback Black [More Like This](#)

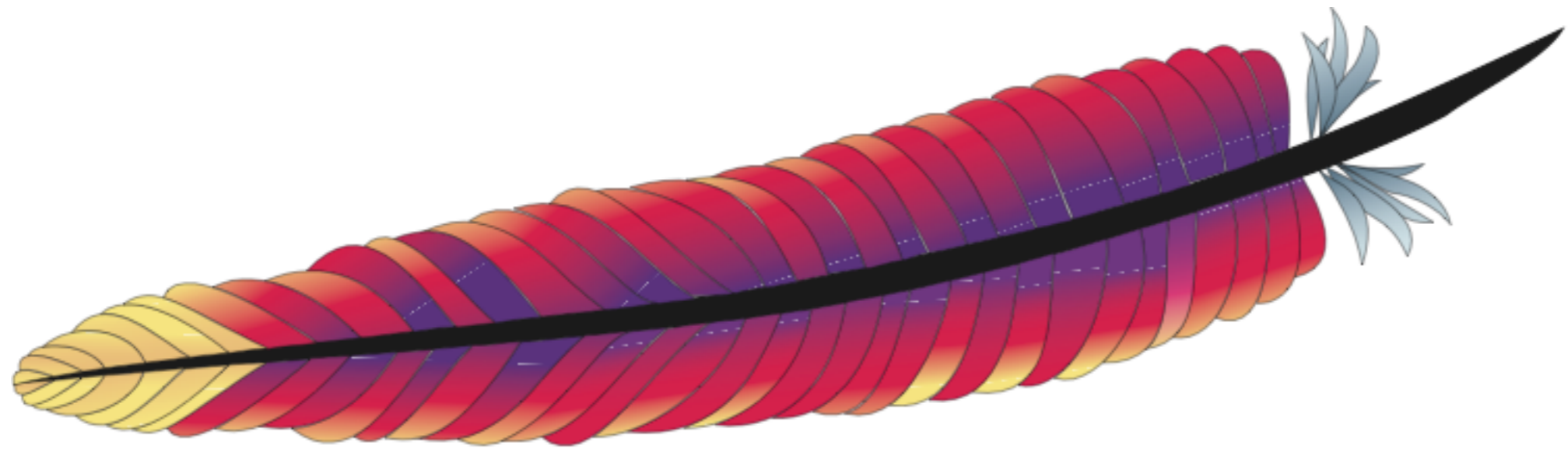
Price: \$399.00

Features: iTunes, Podcasts, Audiobooks Stores up to 15,000 songs, 25,000 photos, or 150 hours of video 2.5-inch, 320x240 color TFT LCD display with LED backlight Up to 20 hours of battery life Plays AAC, MP3, WAV, AIFF, Audible, Apple Lossless, H.264 video Notes, Calendar, Phone book, Hold button, Date display, Photo wallet, Built-in games, JPEG photo playback, Upgradeable firmware, USB 2.0 compatibility, Playback speed control, Rechargeable capability, Battery level indication

In Stock: true



[Larger Map](#)



"Heavy Committing"

<http://www.apache.org/>

lucid... works
IMAGINATION

lucidworks
cloud

lucidworks
search platform

lucidworks
enterprise

www.lucidimagination.com

Events



Presented by Lucid Imagination
17-18 October | **Training**
19-20 October | **Conference**
<http://2011.lucene-eurocon.org>

