



<http://flickr.com/photos/thepma/640151954/>

THE RISING SUN:

Making the most of Solr power

code4lib '09, Providence, RI
Erik Hatcher
erik.hatcher@lucidimagination.com

#9 Performance

Measure

#8 Memory

- omitNorms
- omitTf (1.4)

#7 Query Parsing

• "It Depends"

More Like This:

[http://www.lucidimagination.com/blog/2009/02/22/
exploring-query-parsers/](http://www.lucidimagination.com/blog/2009/02/22/exploring-query-parsers/)

#6 Data Import

- DIH, Solr Cell,
CSV, LuSql, APIs,...
- solr-ruby Mapper

Obligatory XML

```
<root>
  <parent>
    <child attribute="attribute1">text1</child>
    <child attribute="attribute2">text2</child>
  </parent>
</root>

def test_simple_xpath
  mapping = { :solr_field1 => '/root/parent/child',
              :solr_field2 => '/root/parent/child/@attribute' }

  mapper = Solr::Importer::XPathMapper.new(mapping)
  mapped_data = mapper.map(@doc)

  assert_equal ['text1', 'text2'], mapped_data[:solr_field1]
  assert_equal ['attribute1', 'attribute2'], mapped_data[:solr_field2]
end
```

#5 Request Handlers

- Leverage Solr's configurability

http://localhost:8983/solr/ timeline.xml

```
<requestHandler class="solr.SearchHandler" name="/timeline.xml">
  <lst name="defaults">
    <str name="defType">dismax</str>
    <str name="q.alt">*:*</str>
    <str name="rows">10</str>
    <str name="fl">*,score</str>
    <str name="qf">
      content subject^2 title^5
    </str>
  </lst>
  <lst name="invariants">
    <str name="wt">xslt</str>
    <str name="tr">timeline.xsl</str>
  </lst>
</requestHandler>
```

#4 Solr as IR toolkit

- TermVectorRequestHandler
- StatsComponent
- Luke

#3 LocalSolr

geo searching

#2 Faceting

- Solr 1.4 performance!
(*bows* to Yonik)
- Multi-select

Multi-select facetting



IMAGINATION

Home Community Downloads Documentation How We Can Help About

Options

PROJECT [clear facet](#)

Lucene (200)

Solr (185)

Nutch (28)

Mahout (1)

Lucene.Net (1)

SOURCE [clear facet](#)

Lucid (3)

Wiki (5)

Apache Lucene Web (1)

Search within: [Lucid site](#) [Lucene wikis and web](#) [Lucene issues, email, and code](#) [Lucid site, Lucene site, Lucene issues, email, and code](#)

marc



BETA

[Start new search](#)

Search Results for marc

Found 385 results in 0.036 seconds. Displaying page 1 of 39, sorted by [relevancy](#)

[\[WIKI\] Solr4 Lib](#)

Besides the great pointers on the [IntegratingSolr](#) page, libraries integrating Solr have their own unique challenges that the rest of the computer programming world is advised to steer clear of. How to index [MaRC](#) records into solr -----> * [solrmarc](#) is a java solution used by both Blacklight and VuFind, available at <http://code.google.com/p/solrmarc/> * Terry Reese has a video of how to use MarcEdit in this way: [<http://oregonstate.edu/%7Ereeset/blog/archives/418> [MARC](#) -> Solr via MarcEdit] * You can also use [<http://marc4j.tigris.org/> [marc4j](#)]s command-line utilities to convert binary [marc](#)

<http://wiki.apache.org/solr/Solr4Lib>

<http://www.lucidimagination.com/search/p:lucene,solr/s:email,wiki/l:user?q=marc>

#1 User Interface

• The Interface is
the APP

SolrJS

SolrJS Demonstration

Browse REUTERS business news from 1987

Current Selection

(x) allText:"iraq"

Search

(press ESC to close suggestions)

Top Topics

carcass corn cotton cpi **crude** earn fuel gas gnp grain heat
interest i-cattle livestock lumber meal-feed money-fx naphtha oilseed
pet-chem plywood propane reserves **ship** soy-meal soy-oil soybean
sugar veg-oil wheat

Top Organisations

ec fao gcc iea **opec**

Top Exchanges

not items found in current selection

By Country

[view the World](#)



<< < > >> displaying 1 to 10 of 134

IRAN ANNOUNCES NEW OFFENSIVE IN GULF WAR

LONDON, March 4 - Iran announced it had launched a new offensive against Iraq in the north-west area of the Gulf war front. Tehran radio, monitored by the British Broadcasting Corporation, said the attack was launched last night in the Haj Omran area of northern Iraq. It had already achieved "considerable vic [more](#)

IRAQ REJECTS NEW OPEC OIL QUOTA

crude

BAGHDAD, June 29 - Iraq has rejected its new OPEC production quota set in Vienna for the second half of this year, the official Iraqi news agency INA reported. It quoted Oil Minister Issam Abdul-Rahim al-Chalabi, who returned Sunday, as saying "Iraq will continue to adhere to its position considering itself as not [more](#)

IRAQ-TURKEY OIL PIPELINE CUT BY LANDSLIDE

crude

ANKARA, March 8 - Engineers are working to repair the Iraq- Turkey oil pipeline near the southern town of Adana after it was cut by a landslide, the Hurriyet and Anatolian news agencies said. Little oil was lost in the landslide Friday night because taps on the one mln bpd line were switched off after the accident, they said. The pipeline, which carr [more](#)

BRAZIL TO EXPORT POULTRY, MEAT TO IRAQ FOR OIL

crude livestock carcass

RIO DE JANEIRO, April 2 - Brazil will export 6,000 tonnes of poultry and 10,000 tonnes of frozen meat to Iraq in exchange for oil, Petrobras Commercial Director Carlos Sant'Anna said. Brazil has a barter deal with Iraq and currently imports 215,000 barrels per day of oil, of which 170,000 bpd are paid for with exports of [more](#)

IRAQ SAYS ITS FORCES SINK THREE IRANIAN VESSELS

crude

BAGHDAD, April 3 - Iraq said its forces sank three Iranian boats which tried to approach the disputed deep water oil terminal in the northern Gulf today. A military spokesman

Solritas

- Velocity templated response

#0 Community

• solr-user@lucene.apache.org

"LucidFind"

<http://www.lucidimagination.com/search>
find@lucidimagination.com





Thinking Lucene ▶ Think Lucid.

Lucid Supports Solr

lucidimagination.com

Lucid Articles

- Grant Ingersoll
 - Getting Started with Lucene
 - Debugging Relevance Issues in Search
 - Optimizing Findability in Lucene and Solr
- Yonik Seeley
 - Faceted Search with Solr
- Erik Hatcher
 - Getting Started with Solr (includes screencast), co-authored with Jonathan Knudsen
 - Sami Siren
 - Content Extraction with Tika
- Mark Miller
 - Scaling Lucene and Solr

Lucid Podcasts

- Interviews with:
 - Doug Cutting (creator of Lucene)
 - Ryan McKinley (Solr committer)
 - Chris Hostetter (Solr committer)
- More coming!

Lucid Blogging

- Mark Miller: "Exploring Query Parsers", "Highlighting Highlighter Thoughts", "Investigating OOM and other JVM issues", "Looking forward to new features in Solr 1.4"
- Erik Hatcher: "acts_as_solr with rich document indexing"
- Grant Ingersoll: "Sorting, Faceting, and Schema Design in Solr"

SECOND EDITION

Lucene in Action



Pre-release ebook available now!

The Good Ol' Blues Brothers Boys

