

LukeRequestHandler

<!-- Solr1.2

The [LukeRequestHandler](#) is modeled after [Luke](#), the Lucene Index Browser by Andrzej Bialecki.

This handler will display information for **any** lucene index - even if it was not made by solr or does not match the solr schema.xml.

To use the LukeRequestHandler, make sure it is defined in your solrconfig.xml:

```
<requestHandler name="/admin/luke" class="org.apache.solr.handler.admin.LukeRequestHandler" />
```

⚠ Solr3.6 ⚠ Solr4.0

There have been significant speed improvements in these two versions, see the discussion under the reportDocCount parameter

- [Example URLs](#)
- [Example Output](#)
- [Parameters](#)
 - [numTerms](#)
 - [fl](#)
 - [id](#)
 - [docId](#)
 - [show](#)
 - [reportDocCount](#)

Example URLs

Assuming you have this handler mapped to "/admin/luke" in solrconfig.xml and are running the example on port localhost:8983, visit:

- [/admin/luke](#)
- [/admin/luke?fl=cat](#)
- [/admin/luke?fl=id&numTerms=50](#)
- [/admin/luke?id=SOLR1000](#)
- [/admin/luke?docId=2](#)

To view results with more human readable output you have two options

- Click on *SCHEMA BROWSER* in Solr's admin area. This starts an AJAXified schema browser which also displays information from the [LukeRequestHandler](#).
- Try using the luke.xslt stylesheet: <http://localhost:8983/solr/admin/luke?wt=xslt&tr=luke.xsl>

Example Output

For the request /admin/luke, the following is example output from the Solr Example, with annotations marked as <!-- ANN: COMMENT -->

```
<?xml version="1.0" encoding="UTF-8"?>
<response>
  <lst name="responseHeader">
    <int name="status">0</int>
    <int name="QTime">90</int>
  </lst>
  <lst name="index"> <!-- ANN: Provides info about the state of the index -->
    <int name="numDocs">17</int>
    <int name="maxDoc">17</int>
    <int name="numTerms">1044</int>
    <long name="version">1297337332283</long>
    <bool name="optimized">true</bool>
    <bool name="current">true</bool>
    <bool name="hasDeletions">false</bool>
    <str name="directory"> <!-- ANN: The choice of Directory can sometimes effect performance. Lucene tries to
    -->
      org.apache.lucene.store.NIOFSDirectory:org.apache.lucene.store.NIOFSDirectory@[{PATH}/dev/trunk/solr
/example/solr/data/index
      lockFactory=org.apache.lucene.store.NativeFSLockFactory@349319d9
    </str>
    <date name="lastModified">2011-02-10T11:29:03Z</date>
  </lst>
```

```

<lst name="fields"> <!-- ANN: Provide stats and metadata about the fields in the index -->
  <lst name="store_0_coordinate">
    <str name="type">tdouble</str> <!-- ANN: The FieldType, see the schema.xml -->
    <str name="schema">IT-----OF-----</str> <!-- ANN: The key/legend at the bottom says what each entry in
this list is. This field is indexed, tokenized, Omits norms and sorts missing first -->
    <str name="dynamicBase">*_coordinate</str> <!-- ANN: This field is a "polyField" and actually corresponds
to multiple Lucene fields -->
    <str name="index">(unstored field)</str>
    <int name="docs">14</int> <!-- ANN: 14 documents have a value for this field -->
    <int name="distinct">64</int> <!-- ANN: There are 64 distinct terms -->
    <lst name="topTerms">
      <int name="2.0">14</int> <!-- ANN: The token "2.0" occurs 14 times in the field -->
      <int name="44.0">6</int>
      <int name="37.775177001953125">4</int>
      <int name="37.7734375">4</int>
      <int name="37.77519989013672">4</int>
      <int name="36.0">4</int>
      <int name="37.7752">4</int>
      <int name="37.77519999999822">4</int>
      <int name="37.775199999956712">4</int>
      <int name="40.71429991722107">2</int>
    </lst>
    <lst name="histogram"> <!-- ANN: The histogram measures the distribution of the distinct terms across the
documents -->
      <int name="1">46</int> <!-- ANN: 46 of the 64 terms occur in only 1 document -->
      <int name="2">9</int> <!-- ANN: 9 of the 64 terms occur in 2 or 3 (TODO: Check this) -->
      <int name="4">7</int>
      <int name="8">1</int>
      <int name="16">1</int>
    </lst>
  </lst>
  <lst name="text">
    <str name="type">text</str>
    <str name="schema">IT-M-----</str>
    <str name="index">(unstored field)</str>
    <int name="docs">17</int>
    <int name="distinct">389</int>
    <lst name="topTerms">
      <int name="electron">14</int>
      <int name="2">8</int>
      <int name="inc">8</int>
      <int name="x">8</int>
      <int name="1">8</int>
      <int name="gb">7</int>
      <int name="3">7</int>
      <int name="0">6</int>
      <int name="20">5</int>
      <int name="mb">5</int>
    </lst>
    <lst name="histogram">
      <int name="1">278</int>
      <int name="2">67</int>
      <int name="4">31</int>
      <int name="8">12</int>
      <int name="16">1</int>
    </lst>
  </lst>
  <lst name="cat">
    <str name="type">string</str>
    <str name="schema">I-SM---OF---l</str>
    <str name="index">I-S---0-----</str>
    <int name="docs">16</int>
    <int name="distinct">14</int>
    <lst name="topTerms">
      <int name="electronics">14</int>
      <int name="memory">3</int>
      <int name="graphics card">2</int>
      <int name="search">2</int>
      <int name="hard drive">2</int>
      <int name="connector">2</int>
      <int name="software">2</int>
    </lst>
  </lst>

```

```

        <int name="monitor">2</int>
        <int name="copier">1</int>
        <int name="music">1</int>
    </lst>
    <lst name="histogram">
        <int name="1">6</int>
        <int name="2">6</int>
        <int name="4">1</int>
        <int name="8">0</int>
        <int name="16">1</int>
    </lst>
</lst>
<!-- ... -->
</lst>
<lst name="info">
    <lst name="key">
        <str name="I">Indexed</str>
        <str name="T">Tokenized</str>
        <str name="S">Stored</str>
        <str name="M">Multivalued</str>
        <str name="V">TermVector Stored</str>
        <str name="o">Store Offset With TermVector</str>
        <str name="p">Store Position With TermVector</str>
        <str name="O">Omit Norms</str>
        <str name="L">Lazy</str>
        <str name="B">Binary</str>
        <str name="f">Sort Missing First</str>
        <str name="l">Sort Missing Last</str>
    </lst>
    <str name="NOTE">Document Frequency (df) is not updated when a document is marked for deletion. df values
include
    deleted documents.
    </str>
</lst>
</response>

```

Parameters

numTerms

How many top terms for each field. The default is 10.

fl

limit the returned values to a set of fields. This is useful if you want to increase the **numTerms** and don't want a massive response

- </admin/luke?fl=text&numTerms=100>

id

Get a document using the uniqueKeyField specified in schema.xml

- </admin/luke?id=SOLR1000>

docId

Get a document using a lucene documentID

- </admin/luke?docId=2>

show

 Solr1.3

'show=schema' Show the schema fields and properties

- </admin/luke?show=schema>

reportDocCount

! Solr3.6 and subsequent 3.x releases ONLY. Not relevant for 4.x

'reportDocCount=true | false'. Default value: false. THIS IS A CHANGE IN BEHAVIOR.

- /admin/luke?reportDocCount=true (http://localhost:8983/solr/admin/luke?reportDocCount='true' or 'false')

The Luke request handler was almost unusably slow for large indexes. It turns out that most of the time was spent gathering the number of documents that contained a field. As of Solr 3.6, the behavior is changed so that this statistic is NOT reported by default but can be reported if you specify 'true' for this parameter.

3.x will still require some patience. Even specifying 'reportDocCount=false' still took 53 seconds on my 89M sample data, down from 640 seconds or so.

Note that this parameter can also be specified on the solr/admin/schema browser link if document counts are desired when using that interface as:

- /admin/schema.jsp?reportDocCount=true (http://localhost:8983/solr/admin/schema.jsp?reportDocCount='true' or 'false')

! Solr4.0

reportDocCount is a very short-lived parameter. The changes to the underlying Lucene index structure in the trunk (4.x) code line make gathering the number of documents that a particular field occurs in very fast. A sample index with 89M documents took on the order of 450 seconds under the original structure (i.e. getting range queries for each field with field:[* TO *]). Using more efficient 4.x ways of collecting that data, it took 160 seconds. Using really efficient techniques available in 4.x, it is down around 16 seconds. So the reportDocCount parameter seems unnecessary and is removed in the 4.x code line.

Notes:

- In the 4.x code line, the document count returned will *include* deleted documents if there are any. If exact counts are required, you can optimize the index before calling this handler.
- numTerms is still supported in both versions and will cause only the definitions in schema.xml to be returned, no statistics from the index will be included.
- See SOLR-1931 for all the details.

Sample Output

```
<response>
  <lst name="responseHeader">
    <int name="status">0</int>
    <int name="QTime">150</int>
  </lst>
  <str name="WARNING">This response format is experimental. It is likely to change in the future.</str>
  <lst name="index">
    <int name="numDocs">5000</int>
    <int name="maxDoc">5000</int>
    <int name="numTerms">278687</int>
    <long name="version">1202410579307</long>
    <bool name="optimized">true</bool>
    <bool name="current">true</bool>
    <bool name="hasDeletions">false</bool>
    <str name="directory">
      org.apache.lucene.store.FSDirectory:org.apache.lucene.store.FSDirectory@<path>/solr/data/index
    </str>
    <date name="lastModified">2008-02-11T18:24:35Z</date>
  </lst>
  <lst name="schema">
    <lst name="fields">
      <lst name="title">
        <str name="type">text</str>
        <str name="flags">ITSM-----</str>
      </lst>
      <lst name="docid">
        <str name="type">string</str>
        <str name="flags">I-S----0----l</str>
      </lst>
      <lst name="articleDate">
        <str name="type">date</str>
        <str name="flags">I-S----0----l</str>
      </lst>
      <lst name="docname">
        <str name="type">string</str>
        <str name="flags">I-S----0----l</str>
      </lst>
    </lst>
  </lst>
  <lst name="body">
```

```

    <str name="type">text</str>
    <str name="flags">ITSM-----</str>
</lst>
<lst name="id">
    <str name="type">string</str>
    <str name="flags">I-S---0---l</str>
    <bool name="required">true</bool>
</lst>
</lst>
<lst name="types">
    <lst name="double">
        <null name="fields"/>
        <bool name="tokenized">>false</bool>
        <str name="className">org.apache.solr.schema.DoubleField</str>
        <str name="analyzer">org.apache.solr.schema.FieldType$DefaultAnalyzer</str>
    </lst>
    <lst name="text">
        <arr name="fields">
            <str>title</str>
            <str>body</str>
        </arr>
        <bool name="tokenized">>true</bool>
        <str name="className">org.apache.solr.schema.TextField</str>
        <str name="analyzer">org.apache.solr.analysis.TokenizerChain</str>
    </lst>
    <lst name="string">
        <arr name="fields">
            <str>docid</str>
            <str>docname</str>
            <str>id</str>
        </arr>
        <bool name="tokenized">>false</bool>
        <str name="className">org.apache.solr.schema.StrField</str>
        <str name="analyzer">org.apache.solr.schema.FieldType$DefaultAnalyzer</str>
    </lst>
    <lst name="sfloat">
        <null name="fields"/>
        <bool name="tokenized">>false</bool>
        <str name="className">org.apache.solr.schema.SortableFloatField</str>
        <str name="analyzer">org.apache.solr.schema.FieldType$DefaultAnalyzer</str>
    </lst>
    <lst name="integer">
        <null name="fields"/>
        <bool name="tokenized">>false</bool>
        <str name="className">org.apache.solr.schema.IntField</str>
        <str name="analyzer">org.apache.solr.schema.FieldType$DefaultAnalyzer</str>
    </lst>
    <lst name="float">
        <null name="fields"/>
        <bool name="tokenized">>false</bool>
        <str name="className">org.apache.solr.schema.FloatField</str>
        <str name="analyzer">org.apache.solr.schema.FieldType$DefaultAnalyzer</str>
    </lst>
    <lst name="slong">
        <null name="fields"/>
        <bool name="tokenized">>false</bool>
        <str name="className">org.apache.solr.schema.SortableLongField</str>
        <str name="analyzer">org.apache.solr.schema.FieldType$DefaultAnalyzer</str>
    </lst>
    <lst name="date">
        <arr name="fields">
            <str>articleDate</str>
        </arr>
        <bool name="tokenized">>false</bool>
        <str name="className">org.apache.solr.schema.DateField</str>
        <str name="analyzer">org.apache.solr.schema.FieldType$DefaultAnalyzer</str>
    </lst>
    <lst name="sint">
        <null name="fields"/>
        <bool name="tokenized">>false</bool>
        <str name="className">org.apache.solr.schema.SortableIntField</str>

```

```

    <str name="analyzer">org.apache.solr.schema.FieldType$DefaultAnalyzer</str>
  </lst>
  <lst name="boolean">
    <null name="fields"/>
    <bool name="tokenized">false</bool>
    <str name="className">org.apache.solr.schema.BoolField</str>
    <str name="analyzer">org.apache.solr.schema.BoolField$1</str>
  </lst>
  <lst name="sdouble">
    <null name="fields"/>
    <bool name="tokenized">false</bool>
    <str name="className">org.apache.solr.schema.SortableDoubleField</str>
    <str name="analyzer">org.apache.solr.schema.FieldType$DefaultAnalyzer</str>
  </lst>
  <lst name="long">
    <null name="fields"/>
    <bool name="tokenized">false</bool>
    <str name="className">org.apache.solr.schema.LongField</str>
    <str name="analyzer">org.apache.solr.schema.FieldType$DefaultAnalyzer</str>
  </lst>
</lst>
</lst>
<lst name="info">
  <lst name="key">
    <str name="I">Indexed</str>
    <str name="T">Tokenized</str>
    <str name="S">Stored</str>
    <str name="M">Multivalued</str>
    <str name="V">TermVector Stored</str>
    <str name="o">Store Offset With TermVector</str>
    <str name="p">Store Position With TermVector</str>
    <str name="O">Omit Norms</str>
    <str name="L">Lazy</str>
    <str name="B">Binary</str>
    <str name="C">Compressed</str>
    <str name="f">Sort Missing First</str>
    <str name="l">Sort Missing Last</str>
  </lst>
  <str name="NOTE">Document Frequency (df) is not updated when a document is marked for deletion. df values
include
  deleted documents.
</str>
</lst>
</response>

```