

Information Retrieval – begleitendes Tutorium V

Thomas Schmidt

Rückblick

schema.xml

solrconfig.xml

Datei-Import (z.B. post.jar, DIH)



Velocity-Templates

AJAX Solr

Heute: Solr – Frontend

Solritas/Velocity-
Templates

Find:

☐ Boost by Cena

Field Facets

nazov_kategorie

- [lety_balónom](#) (47)
- [Romantika](#) (43)
- [Zážitky vo vzduchu](#) (39)
- [Pre firmy](#) (4)
- [Hobby a zábava](#) (1)
- [Zlavo Adina](#) (1)
- [Adina oblačenie](#) (0)
- [Aktívny oddych](#) (0)
- [Auto moto](#) (0)
- [Gurmán](#) (0)
- [Historické autá](#) (0)
- [Hmotné darčeky](#) (0)
- [Hotely](#) (0)
- [Jazda na Formule](#) (0)
- [Jazda po bežnej ceste](#) (0)

53 results found in 78 ms Page 1 of 3

LET BALÓNOM - Západné Slovensko [More Like This](#)

súkromne pre 16 osôb

Cena: 1968,07 €

Okres: Tmavský kraj Trenčianský kraj Nitrianský kraj Bratislavský kraj

Nazov kategoríe: lety balónom Romantika Zážitky vo vzduchu

Datum konca zľavy: nie je zľava

O produkte:

[\[+\]](#) Zažite vzdušné dobrodružstvo so všetkými, ktorých máte radi. Podelte sa s nim...

Aktívny: true

Hmotný: false

Hodnotenie: 0

→ traditionelle UIs (z.B.
Twitter, Websuche...)

AJAX Solr

Nächste Woche!!!

AJAX Solr Demonstration
Browse Reuters business news from 1987

< 1 2 3 ... 2157 2158 > displaying 1 to 10 of 21578

Current Selection
Viewing all documents!

Search
(press ESC to close suggestions)

Top Topics

Top Organisations

Top Exchanges

By Country
view the World

VALERO ENERGY DEBT UPGRADED BY MOODY'S
NEW YORK, March 17 - Moody's Investors Service Inc said it upgraded Valero Energy Corp's 120 mln dls of debt. Moody's cited an improved leverage position that will result from Valero's sale of its Valero Natural Gas subsidiary to Valero Natural Gas Partners L.P., the partial sale of units of that partnership to the [more](#)

NL FILES SUITS AGAINST UNITED CATALYSTS
HIGHTSTOWN, N.J., March 17 - NL Industries Inc's NL Chemical Inc subsidiary said it filed two complaints against for patent infringements and misappropriation of confidential information. NL said the patent infringement suit alleges that United's product, Theogel D55, infringes on NL's patent protec [more](#)

IRVING TRUST BUYS GULF/WESTERN UNIT
NEW YORK, March 17 - Irving Bank Corp said it bought the factoring division of Associates Commercial Corp, a unit of Gulf and Western Co Inc's Associates Corp of North America. The terms of the previously announced deal were not disclosed. It said the assets were transferred to Irving Commercial Corp. Reuter

DITTLER BROTHERS SEEK LOTTERY INVESTIGATIONS
ATLANTA, March 17 - Dittler Brothers called for investigations by the attorneys general of 24 states in connection with possible violation of state lottery laws by Bally Manufacturing Corp and its subsidiary, Scientific Games. Dittler Brothers said it requested the states' law enforcement chiefs to investigat [more](#)

AMRE INC 3RD QTR JAN 31 NET
DALLAS, March 17 - Shr five cts vs one ct Net 196,986 vs 37,968 Revs 15.3 mln vs 8,900,000 Nine mths Shr 52 cts vs 22 cts Net two mln vs 874,000 Revs 53.7 mln vs

→ komplexere, spezielle
UIs (z.B. Reuters...)

Installation (nochmal...)

- Solr-4.8.0.zip runterladen (Link im Grips oder googeln) und entpacken
- über Ausführen – cmd: Konsole öffnen
- Navigation zum solr-4.8.0 – Ordner über cd-Befehl
- zu example-Ordner navigieren, also cd example
- „java -jar start.jar“ eingeben
- Solr nun erreichbar unter: <http://localhost:8983/solr>

Gliederung

- Allgemeines
- Quick Tour – Solritas
- Velocity-Templates Basics
- Velocity-Templates in Solr nutzen
- Search-Features → Beispiel: Facetten, Highlighting, More like this
- Handlungsempfehlung für Projekt

Beispieldatei: tweets.xml
(GRIPS) – nächste Woche, bei
Bedarf



Solritas/VelocityResponseWriter

- erreichbar über <http://localhost:8983/solr/browse>
- zeigt einige Searchfeatures von Solr auf
- basiert auf Velocity-Templates
- Rapid UI Prototyping für die schnelle Präsentation eines neuen Suchalgorithmus oder eines aufbereiteten Datensets

Find:

☐ Boost by Cena

Field Facets

nazov_kategorie

- [lety balóňom](#) (47)
- [Romantika](#) (43)
- [Zážitky vo vzduchu](#) (39)
- [Pre firmy](#) (4)
- [Hobby a zábava](#) (1)
- [ZLAVO Adina](#) (1)
- [Adina oblečenie](#) (0)
- [Aktívny oddych](#) (0)
- [Auto moto](#) (0)
- [Gurmán](#) (0)
- [Historické autá](#) (0)
- [Hmotné darčeky](#) (0)
- [Hotely](#) (0)
- [Jazda na Formule](#) (0)
- [Jazda po bežnej ceste](#) (0)

53 results found in 78 ms Page 1 of 3

LET BALÓŇOM - Západné Slovensko [More Like This](#)

súkromne pre 16 osôb

Cena: 1968,07 €

Okres: Trnavský kraj Trenčianský kraj Nitrianský kraj Bratislavský kraj

Nazov kategorie: lety balóňom Romantika Zážitky vo vzduchu

Datum konca zlavý: nie je zľava

O produkte:
[\[+\]](#) Zážite vzdušné dobrodružstvo so všetkými, ktorých máte radi. Podelte sa s nim...

Aktívny: true

Hmotny: false

Hodnotenie: 0

Apache Velocity



- Java-basierte Template-Engine, welche eine Template-Sprache bereitstellt, mit der in Java definierte Objekte referenziert werden können
- Ziel: klare Trennung von View und Model um das MVC-Pattern zu erfüllen → Trennung des Java-Codes (Model, Controller) von der Webseite (View)
- dynamische Erzeugung von HTML – und XML – Ausgaben
- Apache Lizenz



The Apache Velocity Project

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

Grobkonzept

- erlaubt es Webdesignern Markup-Ausdrücke, sog. Referenzen, zu integrieren
- die Referenzen werden von einem Context-Objekt geholt (Hashtabelle mit Setter und Getter Methoden um Objekte zu holen)
- Velocity bietet grundlegende Kontrollausdrücke (for each, if/else), um über einen bestimmten Wertebereich zu iterieren oder je nach Bedingung Textpassagen zu integrieren
- für webspezifische Funktionalität wird ein Framework benötigt → Velocity Tools



Vorteile

- Einfachheit → Seiten können von Webdesignern geschrieben und verwaltet werden, kaum höhere Programmierkenntnisse nötig
- sowohl Methoden als auch Eigenschaften können angesprochen werden (SolrQueryRequest, SolrQueryResponse)
- Konsistenz → kann auch für andere Texterstellungsaufgabe verwendet werden durch einheitliche Textauszeichnung

Velocity Templates in Solr

solr-4.4.0 ▶ example ▶ solr ▶ collection1 ▶ conf ▶ velocity

browse.vm

layout.vm

```
1  /**
2   * Overall HTML page layout
3   */
4
5  <html>
6  <head>
7      #parse("head.vm")
8  </head>
9  <body>
10     <div id="header">
11         #parse("header.vm")
12     </div>
13     <div id="content">
14         $content
15     </div>
16     <div id="footer">
17         #parse("footer.vm")
18     </div>
19 </body>
20 </html>
21
```

```
/**
 * Main entry point into the /browse templates
 */

#set($searcher = $request.searcher)
#set($params = $request.params)
#set($clusters = $response.response.clusters)
#set($mltResults = $response.response.get("moreLikeThis"))
#set($annotate = $params.get("annotateBrowse"))
#parse('query_form.vm')

<div class="navigators">
    #parse("quicklinks.vm")
</div>

## Show Error Message, if any
<div class="error">
    #parse("error.vm")
</div>

## Render Results, actual matching docs
<div class="results">
    #parse("results_list.vm")
</div>

<div class="pagination">
    #parse("pagination_bottom.vm")
</div>
```

▼ velocity

- browse.vm
- cluster.vm
- cluster_results.vm
- debug.vm
- did_you_mean.vm
- error.vm
- facet_fields.vm
- facet_pivot.vm
- facet_queries.vm
- facet_ranges.vm
- facets.vm
- footer.vm
- head.vm
- header.vm
- hintergrundbild_startseite.jpg
- hit.vm
- hit_grouped.vm
- hit_plain.vm
- join_doc.vm
- jquery.autocomplete.css
- jquery.autocomplete.js
- layout.vm
- main.css

... und das Ergebnis

→ Daten einfügen: im Ordner example docs → `java -jar post.jar *.xml`

[Solr Admin](#)

Apache
Solr

Examples: Simple [Spatial Group By](#)

Find:

☐ Boost by Price
> `cat:"electronics"` > `cat:"connector"`

Field Facets

cat

[Connector](#) (2)

[Electronics](#) (2)

manu_exact

[Belkin](#) (2)

Query Facets

[Ipod](#) (2)

Range Facets

price

0.0 - 50.0 (2)

popularity

0 - 3 (2)

manufactured_date_dt

Less Than

2001-01-01T00:00:00Z (0)

2006-01-01T00:00:00Z -

2006-01-01T00:00:00Z+1YEAR

(1)

2006-01-01T00:00:00Z -

2006-01-01T00:00:00Z+1YEAR

(1)

Pivot Facets

[cat.connector](#) (2)

[inStock=false](#) (2)

[cat.electronics](#) (2)

[inStock=false](#) (2)

Clusters

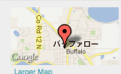
2 results found in 8 ms Page 1 of 1

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

Price: ¥20

Features: car power adapter, white

In Stock: false




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

Price: ¥12

Features: car power adapter for iPod, white

In Stock: false



Options: [enable debug](#) [enable annotation XML](#)

Generated by [VelocityResponseWriter](#)

Documentation: [Solr Home Page](#), [Solr Wiki](#)

Disclaimer: The locations displayed in this demonstration are purely fictional. It is more than likely that no store with the items listed actually exists at that location!

Leichte Anpassung...



Test with some GB18030 encoded characters [More Like This](#)

Id: GB18030TEST

Price: 0,USD

Features: No accents here 这是一个功能 This is a feature (translated) 这份文件是很有光泽 This document is very shiny (translated)

In Stock: true

Samsung SpinPoint P120 SP2514N - hard drive - 250 GB - ATA-133 [More](#)

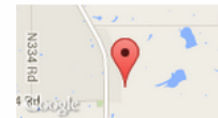
[Like This](#)

Id: SP2514N

Price: 92,USD

Features: 7200RPM, 8MB cache, IDE Ultra ATA-133 NoiseGuard, SilentSeek technology, Fluid Dynamic Bearing (FDB) motor

In Stock: true



[Larger Map](#)

Quick-Links

[Rauminfo](#) [BAFÖG](#) [LSF](#)

[Prüfungsamt](#) [Bewerbung](#)

[E-Mail](#) [Lageplan](#) [Vpn-Client](#) [K-](#)

[Laufwerk](#) [Lehramt](#) [Flex-Now](#)

[Kalender](#) [Uni-Sport](#) [Rechenzentrum](#)

[Drucken](#)

Eins nach dem Anderen

- Velocity Basics



Hello Velocity World

```
<html>
<body>
#set( $foo = "Velocity" )
Hello $foo World!
</body>
</html>
```

“References begin with \$ and are used to get something. Directives begin with # and are used to do something.”

Kommentare...

```
## This is a single line comment.
```

```
This is text that is outside the multi-line comment.  
Online visitors can see it.
```

```
##  
Thus begins a multi-line comment. Online visitors won't  
see this text because the Velocity Templating Engine will  
ignore it.  
*#
```

```
Here is text outside the multi-line comment; it is visible.
```

References – Variablen, Eigenschaften und Methoden

```
$foo  
$mudSlinger  
$mud-slinger  
$mud_slinger  
$mudSlinger1
```

```
$customer.getAddress()  
$purchase.getTotal()  
$page.setTitle( "My Home Page" )  
$person.setAttributes( ["Strange", "Weird", "Excited"] )
```

```
$customer.Address  
$purchase.Total
```


Output

Der finale Wert jeder Referenz muss ein String sein um ausgegeben werden zu können. Ist der Wert kein String, wird automatisch die toString() – Methode aufgerufen.

Formale Notationsmöglichkeit:

```
Jack is a $vicemaniac.
```

```
Jack is a ${vice}maniac.
```

Gut gemeint, aber problematisch in der Entwicklung

\$customer.address kann Verschiedenes bedeuten:

1. `getaddress()`
2. `getAddress()`
3. `get("address")`
4. `isAddress()`

Mehr davon...

```
$foo
```

```
$foo.getBar()
```

```
## is the same as
```

```
$foo.Bar
```

```
$data.setUser("jon")
```

```
## is the same as
```

```
#set( $data.User = "jon" )
```

```
$data.getRequest().getServerName()
```

```
## is the same as
```

```
$data.Request.ServerName
```

```
## is the same as
```

```
${data.Request.ServerName}
```

Programmierung

```
#if( $foo )  
  <strong>Velocity!</strong>  
#end
```

- *\$foo* is a boolean (true/false) which has a true value
- *\$foo* is a string or a collection which is not null **and** not empty
- *\$foo* is an object (other than a string or a collection) which is not null

```
<ul>  
#foreach( $product in $allProducts )  
  <li>$product</li>  
#end  
</ul>
```

Parsen von weiteren Templates

```
#parse( "me.vm" )
```

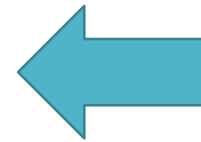
zentraler Vorteil von Velocity-Templates

Context – Objects

eine Möglichkeit ein API – ähnliches Konstrukt zu erzeugen, das den Zugriff auf Java-Elemente über Velocity-Templates ermöglicht

Zugriff auf Context-Objekte auch über \$-Notation → Möglichkeit, Felder, Methoden etc. aufzurufen

```
VelocityEngine engine = getEngine(request); // T  
  
Template template = getTemplate(engine, request);  
  
VelocityContext context = new VelocityContext();  
  
context.put("request", request);
```



VelocityResponseWriter

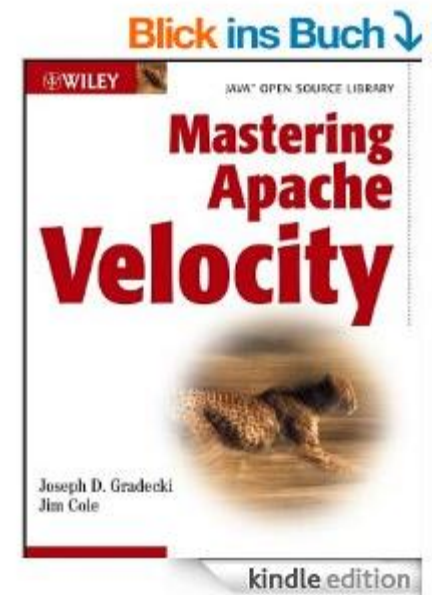
Mehr dazu...

<https://velocity.apache.org/engine/releases/velocity-1.5/user-guide.html>

<http://velocity.apache.org/engine/devel/developer-guide.html>

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

noch mehr Features: Listener, EventHandlers, Exceptions etc.



Velocity Templates in Solr – Ablauf

Anfrage



Solr Response

Solr Request

Velocity
Templates
mit Context-
Objekten

Apache Velocity Templating
Engine

VelocityResponse
Writer

Definition in solrconfig.xml

```
<lib dir="../../../contrib/velocity/lib" regex=".*\.jar" />  
<lib dir="../../../dist/" regex="solr-velocity-\d.*\.jar" />
```

```
<queryResponseWriter name="velocity" class="solr.VelocityResponseWriter" startup="lazy"/>
```

```
<requestHandler name="/browse" class="solr.SearchHandler">  
  <lst name="defaults">  
    <str name="echoParams">explicit</str>  
  
    <!-- VelocityResponseWriter settings -->  
    <str name="wt">velocity</str>  
    <str name="v.template">browse</str>  
    <str name="v.layout">layout</str>  
    <str name="title">Solritas</str>  
  </lst>  
</requestHandler>
```

Writer Type – hier: Velocity Response

v.Template: Template, das für das Rendering der Views zuständig ist (browse.vm)

v.layout: Wrapper für v.template
→ über \$content Zugriff auf v.template in v.layout

Rendering

layout.vm

```
1  ***
2  * Overall HTML page layout
3  *#
4
5  <html>
6  <head>
7    #parse("head.vm")
8  </head>
9  <body>
10    <div id="header">
11      #parse("header.vm")
12    </div>
13    <div id="content">
14      $content
15    </div>
16    <div id="footer">
17      #parse("footer.vm")
18    </div>
19  </body>
20 </html>
21
```

browse.vm

```
***
* Main entry point into the /browse templates
*#

#set($searcher = $request.searcher)
#set($params = $request.params)
#set($clusters = $response.response.clusters)
#set($mltResults = $response.response.get("moreLikeThis"))
#set($annotate = $params.get("annotateBrowse"))
#parse('query_form.vm')

<div class="navigators">
  #parse("quicklinks.vm")
</div>

## Show Error Message, if any
<div class="error">
  #parse("error.vm")
</div>

## Render Results, actual matching docs
<div class="results">
  #parse("results_list.vm")
</div>

<div class="pagination">
  #parse("pagination_bottom.vm")
</div>
```

Zugriff auf Solr

- über Objekte \$response und \$request
- \$request = SolrQueryRequest http://lucene.apache.org/solr/4_8_1/solr-core/org/apache/solr/request/SolrQueryRequest.html
- \$response = QueryResponse http://lucene.apache.org/solr/4_8_1/solr-solrj/org/apache/solr/client/solrj/response/QueryResponse.html
- Zugriff auf Methoden und Felder der Objekte über Velocity-Syntax
- Methoden und Felder können über obige Links eingesehen werden

- !!! Kann sehr problematisch werden, da Methodennamen und Felder vererbt werden und oft nicht klar ist, was die Objekte alles können. Auch werden Hash-Tabellen verwendet, deren Inhalt schwer einsehbar ist (NamedList)
- !!!Velocity-Syntax-Probleme

Beispiel: \$response

```
15  #foreach($doc in $response.results)
16    #parse("hit.vm")
17    ## Can get an extremely simple view of the doc
18    ## which might be nicer for debugging
19    ##parse("hit_plain.vm")
20  #end
```

result_list.vm:

wie: `$response.getResults();`

hit_plain.vm:

`$response.results` liefert ein Objekt
des Typs `SolrDocumentList` zurück
(hier Iteration)

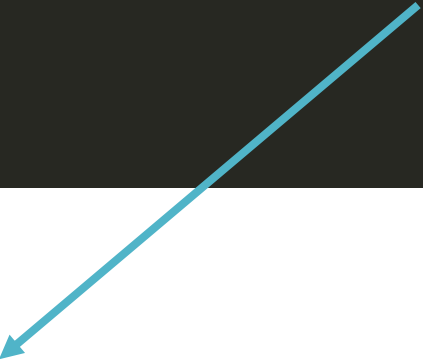
`$doc.fieldNames` = Methode
`getFieldNames` auf jedes
`SolrDocument` der
`SolrDocumentList`

auch hier: Methode
`getFieldValues(String)` von
`SolrDocument`

```
5  <table>
6    ## For each field
7    #foreach( $fieldName in $doc.fieldNames )
8      ## For each value
9      #foreach( $value in $doc.getFieldValues($fieldName) )
10        <tr>
11          ## Field Name
12          <th align="right" valign="top">
13            #if( $foreach.count == 1 )
14              $fieldName:
15            #end
16          </th>
17          ## Field Value(s)
18          <td align="left" valign="top">
19            $esc.html($value) <br/>
20          </td>
21        </tr>
22      #end      ## end for each value
23    #end      ## end for each field
24  </table>
25  <hr/>
26
```

Beispiel \$response (results_list.vm)

```
#if($response.response.get('grouped'))  
  #foreach($grouping in $response.response.get('grouped'))  
    #parse("hit_grouped.vm")  
  #end  
#else
```



\$response.response = \$response.getResponse()
getResponse() ist Methode von Oberklasse SolrResponseBase und liefert eine NamedList<Object> aus name/value - Paaren zurück
→ enthält Objekte je nach aktivierten Suchfeatures (hier: grouping)

Beispiel: \$request

browse.vm

```
5 #set($searcher = $request.searcher)
6 #set($params = $request.params)
```

`SolrIndexSearcher`

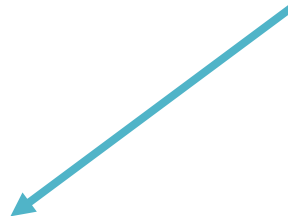
`getSearcher()`

The index searcher associated with this request

`SolrParams`

`getParams()`

returns the current request parameters



z.B. Zugriff auf Query

Das Methoden vs Felder – Problem

Properties

The second flavor of VTL references are properties, and properties have a distinctive format. The shorthand notation consists of a leading *\$* character followed a VTL Identifier, followed by a dot character (".") and another VTL Identifier. These are examples of valid property references in the VTL:

```
$customer.Address  
$purchase.Total
```

Take the first example, *\$customer.Address*. It can have two meanings. It can mean, Look in the hashtable identified as *customer* and return the value associated with the key *Address*. But *\$customer.Address* can also be referring to a method (references that refer to methods will be discussed in the next section); *\$customer.Address* could be an abbreviated way of writing *\$customer.getAddress()*. When your page is requested, Velocity will determine which of these two possibilities makes sense, and then return the appropriate value.

Exkurs

also wenn man es mal ganz genau wissen will in den Sourcecode schauen...

<https://svn.apache.org/repos/asf/lucene/dev/trunk/solr/solrj/src/java/org/apache/solr/client/solrj/response/QueryResponse.java>

<http://grepcode.com/file/repo1.maven.org/maven2/org.apache.solr/solr-core/4.0.0/org/apache/solr/request/SolrQueryRequest.java>

Velocity-Context – Objekte

Objektname	Funktion
\$esc	EscapeTool – Output escapen
\$date	ComparisonDateTool – zum Vergleich von Zeitangaben
\$list	ListTool – zur Arbeit mit Listen und Arrays
\$math	MathTool – für mathematische Operationen mit Fließkommazahlen
\$number	NumberTool – Operationen auf Zahlen
\$page	PageTool – Seitenangabe bei Antworten
\$request	SolrQueryRequest
\$response	QueryResponse
\$sort	SortTool – Sortieroperationen

VelocityResponseWriter - Sourcecode

<http://svn.apache.org/repos/asf/lucene/dev/branches/preflexfixes/solr/src/java/org/apache/solr/response/VelocityResponseWriter.java>

```
context.put("response", rsp);

// Velocity context tools - TODO: make these pluggable
context.put("esc", new EscapeTool());
context.put("date", new ComparisonDateTool());
context.put("list", new ListTool());
context.put("math", new MathTool());
context.put("number", new NumberTool());
context.put("sort", new SortTool());
```

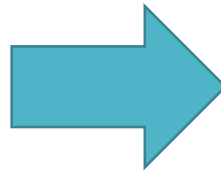
theoretisch Anpassung möglich: Solr source holen, anpassen, neu kompilieren, einfügen

VM_global_library.vm

Datei, in der Makros definiert werden, auf die jedes Template zugreifen kann
VM_global_library.vm Name muss so bleiben

```
#macro(url_root)/solr#end
```

```
<hr/>  
<div id="impressum">  
  <a href="#{url_root}/impressum.html">Impressum</a>  
</div>
```



```
<div id="impressum">  
  <a href="/solr/impressum.html">Impressum</a>  
</div>
```

Zugriff auf Makros über #

Einbindung von css – und js – Dateien

```
<script type="text/javascript" src="#{url_root}/js/lib/jquery-1.7.2.min.js"></script>
<script type="text/javascript" src="#{url_root}/js/uniSearch.js"></script>
<link rel="stylesheet" type="text/css" href="/solr/bootstrap/css/bootstrap.css">
<link rel="stylesheet" type="text/css" href="#{url_for_solr}/admin/file?file=/velocity/main.css&
contentType=text/css"/>
<link rel="stylesheet" href="#{url_for_solr}/admin/file?file=/velocity/jquery.autocomplete.css&
contentType=text/css" type="text/css" />
<link rel="shortcut icon" type="image/x-icon" href="faviconUr.ico" />
<script type="text/javascript" src="#{url_for_solr}/admin/file?file=/velocity/jquery.autocomplete.js&
contentType=text/javascript"></script>
```

Mehr dazu...

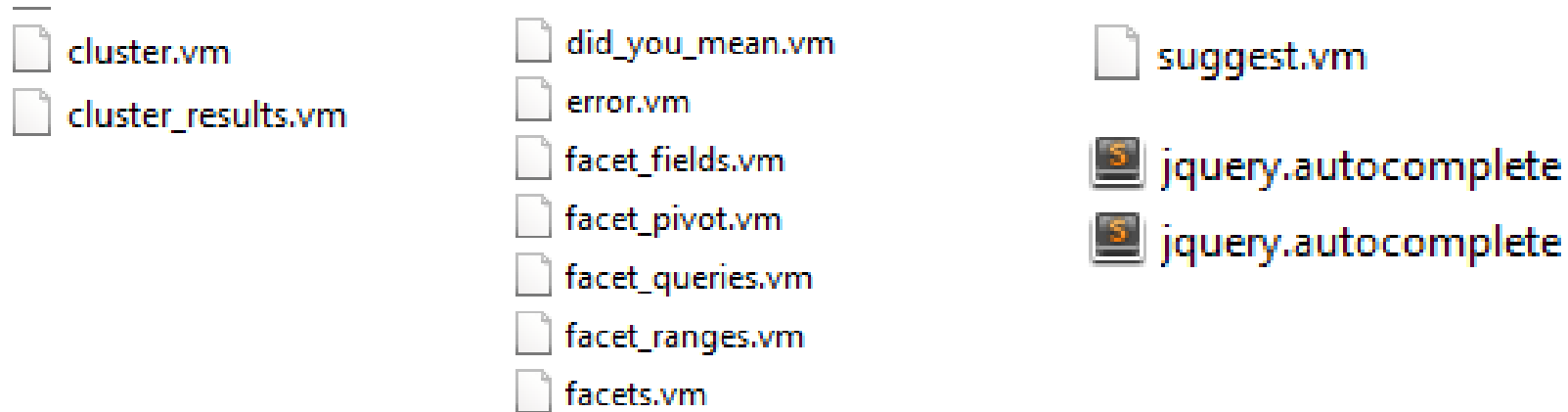
<https://wiki.apache.org/solr/VelocityResponseWriter>

<https://cwiki.apache.org/confluence/display/solr/Velocity+Search+UI>

Search-Features

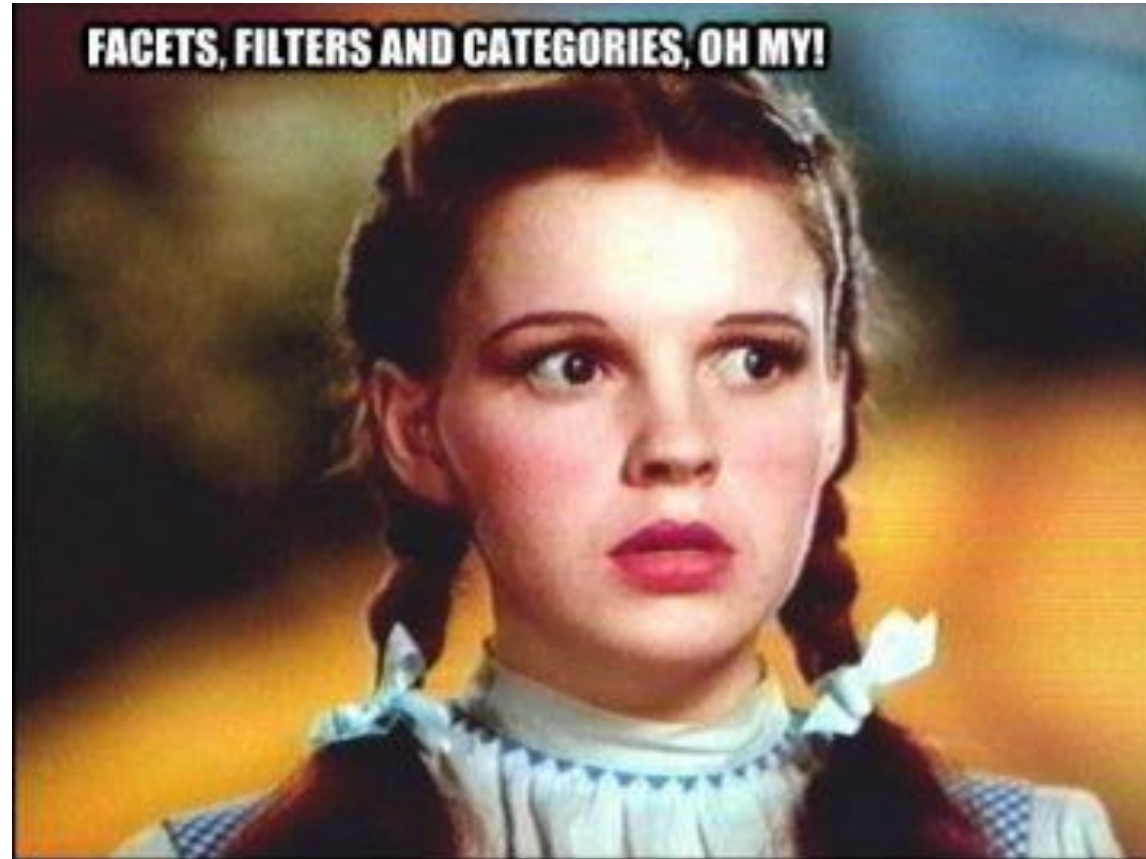


Search-Features in Velocity-Templates



einige Search-Features vorgefertigt in velocity-templates → anpassbar an eigenes Projekt

Facetten



Initialisierung

Standardkomponente
– muss nicht extra
referenziert werden

Facetten in
solrconfig.xml
aktivieren und
konfigurieren → Solr-
Response gibt neues
Element facet_counts
aus

auch als
Aufrufparameter bei
der Suche

```
<!-- Faceting defaults -->
<str name="facet">on</str>
<str name="facet.field">cat</str>
<str name="facet.field">manu_exact</str>
<str name="facet.field">content_type</str>
<str name="facet.field">author_s</str>
<str name="facet.query">ipod</str>
<str name="facet.query">GB</str>
<str name="facet.mincount">1</str>
<str name="facet.pivot">cat,inStock</str>
<str name="facet.range.other">after</str>
<str name="facet.range">price</str>
<int name="f.price.facet.range.start">0</int>
<int name="f.price.facet.range.end">600</int>
<int name="f.price.facet.range.gap">50</int>
<str name="facet.range">popularity</str>
<int name="f.popularity.facet.range.start">0</int>
<int name="f.popularity.facet.range.end">10</int>
<int name="f.popularity.facet.range.gap">3</int>
<str name="facet.range">manufacturedate_dt</str>
<str name="f.manufacturedate_dt.facet.range.start">NOW/YEAR-10YEARS</str>
<str name="f.manufacturedate_dt.facet.range.end">NOW</str>
<str name="f.manufacturedate_dt.facet.range.gap">+1YEAR</str>
<str name="f.manufacturedate_dt.facet.range.other">before</str>
<str name="f.manufacturedate_dt.facet.range.other">after</str>
```

Allgemeine Konfigurationsparameter

Name	Funktion
facet	auf true setzen, um Facetten zu aktivieren
facet.mincount	Spezifiziert, wie oft Term vorkommen muss, dass Eintrag gebildet wird (meist auf 1)
facet.limit	Einschränkung, wie viele Einträge eine Facette haben kann
facet.sort	Festlegung der Sortierreihenfolge
facet.prefix	Facette nimmt nur Werte mit entsprechenden Präfix auf

Facettentypen in Solr

Feld-Facetten

Range-Facetten

Date-Facetten

Query-Facetten

Pivot-Facetten

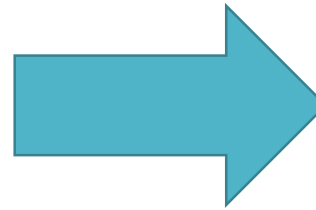
<code>FacetField</code>	<code>getFacetDate(String name)</code>
<code>List<FacetField></code>	<code>getFacetDates()</code>
<code>FacetField</code>	<code>getFacetField(String name)</code>
	<code>get</code>
<code>List<FacetField></code>	<code>getFacetFields()</code>
	See also: <code>getLimitingFacets()</code>
<code>Map<String, Integer></code>	<code>getFacetQuery()</code>
<code>List<RangeFacet></code>	<code>getFacetRanges()</code>



Ausschnitt der Methoden der QueryResponse und damit von \$response

Feld-Facetten

```
<str name="facet.field">cat</str>  
<str name="facet.field">manu_exact</str>  
<str name="facet.field">content_type</str>  
<str name="facet.field">author_s</str>
```



Field Facets

cat

- [electronics](#) (12)
- [currency](#) (4)
- [memory](#) (3)
- [connector](#) (2)
- [graphics card](#) (2)
- [hard drive](#) (2)
- [search](#) (2)
- [software](#) (2)
- [camera](#) (1)
- [copier](#) (1)
- [electronics and computer1](#) (1)
- [electronics and stuff2](#) (1)
- [multifunction printer](#) (1)
- [music](#) (1)
- [printer](#) (1)
- [scanner](#) (1)

manu_exact

- [Apache Software Foundation](#) (2)
- [Belkin](#) (2)

browse.vm

```
<div class="navigators">  
  #parse("facets.vm")  
</div>
```

```
#parse('facet_fields.vm')  
#parse('facet_queries.vm')  
#parse('facet_ranges.vm')  
#parse('facet_pivot.vm')  
#parse('cluster.vm')
```

```
#if($response.facetFields)  
  <h2 #annTitle("Facets generated by adding &facet.field= to the request")>  
    Field Facets  
  </h2>  
  #foreach($field in $response.facetFields)  
    ## Hide facets without value  
    #if($field.values.size() > 0)  
      <span class="facet-field">$field.name</span>  
      <ul>  
        #foreach($facet in $field.values)  
          <li>  
            <a href="#url_for_facet_filter($field.name, $facet.name)">$facet.name</a> ($facet.count)  
          </li>  
        #end  
      </ul>  
    #end  
  #end  
#end  
## end if response has facet fields
```

VM_global_library.vm

```
#macro(url_for_facet_filter $field $value)#url_for_home#lens&fq=$esc.url($field):%22$esc.url($value)%22#end
```

```
#macro(url_for_facet_filter $field $value)#url_for_home#lens&fq=$esc.url($field):%22$esc.url($value)%22#end
```



Query „Belkin“ mit Facette cat

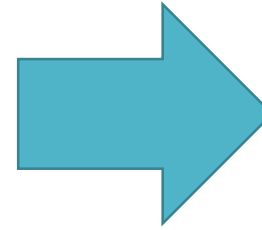


 localhost:8983/solr/collection1/browse?&q=belkin&fq=cat:"connector"

Range-Facetten → numerische Felder nötig

Name des Attributs	Funktion
Facet.range	Angabe des numerischen Feldes
f.feldname.facet.range.start	Startwert
f.feldname.facet.range.end	Endwerte
f.feldname.facet.range.gap	Größe der Bereiche, die generiert werden sollen
f.feldname.Facet.range.other	Zusätzliche Konfigurationsmöglichkeiten

```
<str name="facet.range">popularity</str>
<int name="f.popularity.facet.range.start">0</int>
<int name="f.popularity.facet.range.end">10</int>
<int name="f.popularity.facet.range.gap">3</int>
```



popularity

0 - 3 (3)

3 - 6 (1)

6 - 9 (9)

9 - 12 (2)

facet_ranges.vm

```
#foreach ($field in $response.response.facet_counts.facet_ranges)
  ## Hide facets without value
  #if($field.value.counts.size() > 0)
    #set($name = $field.key)
    #set($display = $name)
    #set($f = $field.value.counts)
    #set($start = $field.value.start)
    #set($end = $field.value.end)
    #set($gap = $field.value.gap)
    #set($before = $field.value.before)
    #set($after = $field.value.after)
    #display_facet_range($f, $display, $name, $start, $end, $gap, $before, $after)
  #end ## end if has any values
#end ## end for each facet range
```


Exkurs: Zugriff auf Suchparameter der Response

```
#foreach ($field in $response.response.facet_counts.facet_ranges)
```

Zugriff über NamedList-Objekt
von \$response und
facet_counts

facet_counts ist wieder eine
NamedList, unter anderem mit
facet_ranges, Zugriff auf Inhalt
über value

→ Orientierung an response-
Ausgabe („Trial and Error“)

```
<response>
...
<lst name="facet_counts">
  <lst name="facet_queries"/>
  <lst name="facet_fields"/>
  <lst name="facet_dates"/>
  <lst name="facet_ranges">
    <lst name="manufacturedate_dt">
      <lst name="counts">
        <int name="2003-01-01T00:00:00Z">0</int>
        <int name="2004-01-01T00:00:00Z">2</int>
        <int name="2005-01-01T00:00:00Z">9</int>
        <int name="2006-01-01T00:00:00Z">0</int>
        <int name="2007-01-01T00:00:00Z">0</int>
        <int name="2008-01-01T00:00:00Z">0</int>
        <int name="2009-01-01T00:00:00Z">0</int>
        <int name="2010-01-01T00:00:00Z">0</int>
        <int name="2011-01-01T00:00:00Z">0</int>
        <int name="2012-01-01T00:00:00Z">0</int>
      </lst>
      <str name="gap">+1YEARS</str>
      <date name="start">2003-01-01T00:00:00Z</date>
      <date name="end">2013-01-01T00:00:00Z</date>
    </lst>
  </lst>
</lst>
</response>
```

Nicht ideal... besser: in neuer Version QueryResponse mächtiger

besser über Methodenname, da Referenz
vorhanden

Namen der Namen/Wert – Paare in den NamedList
nicht immer sofort bekannt

Vorsicht: oft Anpassung der Display-Methode in
den Makros notwendig, also nur anpassen wenn
unbedingt notwendig

→ Beispiel später

Methods	
Modifier and Type	Method and Description
<T> List<T>	getBeans (Class<T> type)
Map<String, Object>	getDebugMap ()
Map<String, SolrDocumentList>	getExpandedResults ()
Map<String, String>	getExplainMap ()
FacetField	getFacetDate (String name)
List<FacetField>	getFacetDates ()
FacetField	getFacetField (String name)
	get
List<FacetField>	getFacetFields ()
	See also: getLimitingFacets ()
NamedList<List<PivotField>>	getFacetPivot ()
Map<String, Integer>	getFacetQuery ()
List<RangeFacet>	getFacetRanges ()
Map<String, FieldStatsInfo>	getFieldStatsInfo ()
GroupResponse	getGroupResponse ()
	Returns the GroupResponse containing the group commands.
NamedList<Object>	getHeader ()
Map<String, Map<String, List<String>>>>	getHighlighting ()
List<FacetField>	getLimitingFacets ()
String	getNextCursorMark ()
SolrDocumentList	getResults ()
NamedList<ArrayList>	getSortValues ()
SpellCheckResponse	getSpellCheckResponse ()
TermsResponse	getTermsResponse ()
protected List<PivotField>	readPivots (List<NamedList> list)
void	removeFacets ()
	Remove the field facet info
void	setResponse (NamedList<Object> res)

Date-Facetten

ähnlich zu Range-Facetten → Bezug auf Felder des Typs Date
unterstützt NOW und Datumsarithmetik

Example: `facet.date.gap=+1DAY,+2DAY,+3DAY,+10DAY` -- This creates 4+ buckets of size, 1 day, 2 days, 3 days and then 0 or more of 10 days each, depending on the start and end times

Query-Facetten

Nutzung eines beliebigen Querys um Facetten zu erzeugen
so lassen sich eigenständige Kategorien erzeugen, mit Hilfe von mehreren Querys
jede Ergebnissammlung für eine Query ist dann eine Kategorie

komplette Solr Query Syntax nutzbar
z.B. bei Range-Facetten, die nicht
gleich große gaps haben sollen

```
<response>
...
<lst name="facet_counts">
  <lst name="facet_queries">
    <int name="manufacturedate_dt:[NOW-1YEARS TO NOW]">0</int>
    <int name="manufacturedate_dt:[NOW-10YEARS TO NOW-1YEARS]">11</int>
    <int name="manufacturedate_dt:[NOW-100YEARS TO NOW-10YEARS]">0</int>
  </lst>
  <lst name="facet_fields"/>
  <lst name="facet_dates"/>
  <lst name="facet_ranges"/>
</lst>
</response>
```

Pivot-Facetten

- hierarchische Facetten bzw. Entscheidungsbäume
- statt Facetten → Hierarchieebenen
- zum Beispiel Kleidungstyp „Schuhe“ für Kategorie „Herren“
- facet.pivot mit einer kommaseparierten Liste der Felder, die eine hierarchische Struktur bilden sollen

```
<str name="facet.pivot">cat,inStock</str>
```

facet_pivot.vm

```
<h2 #annTitle("Facets generated by adding &facet.pivot= to the request")>
  Pivot Facets
</h2>

#set($pivot = $response.response.facet_counts.facet_pivot)

#display_facet_pivot($pivot, "")
```

Pivot Facets

[cat::electronics](#) (3)

[inStock::false](#) (2)


[inStock::true](#) (1)

[cat::connector](#) (2)

[inStock::false](#) (2)

[cat::music](#) (1)

[inStock::true](#) (1)



besser: Methode von neuer
Version von
QueryResponse:
\$response.getFacetPivot()

Highlighting

```
<!-- Highlighting defaults -->
<str name="hl">on</str>
<str name="hl.fl">content features title name</str>
<str name="hl.encoder">html</str>
<str name="hl.simple.pre">&lt;b&gt;</str>
<str name="hl.simple.post">&lt;/b&gt;</str>
<str name="f.title.hl.fragsize">0</str>
<str name="f.title.hl.alternateField">title</str>
<str name="f.name.hl.fragsize">0</str>
<str name="f.name.hl.alternateField">name</str>
<str name="f.content.hl.snippets">3</str>
<str name="f.content.hl.fragsize">200</str>
<str name="f.content.hl.alternateField">content</str>
<str name="f.content.hl.maxAlternateFieldLength">750</str>
```



2 results

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

Id: F8V7067-APL-KIT

Price: 19.95,USD

Features: **car** power adapter, white

In Stock: false

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

Id: IW-02

Price: 11.50,USD

Features: **car** power adapter for iPod, white

Name	Funktion
hl	Auf true setzen um Highlighting zu aktivieren
hl.q	Parameter nach dem gesucht werden soll, per default q, also die Query
hl.fl	Felder, in denen der Begriff markiert werden soll
hl.snippets	Anzahl der dargestellten Fundstellen
hl.fragsize	Größe des Snippets (per Default 100 Zeichen)
hl.alternateField	Rückgabefeld, falls kein Snippe gefunden wird
hl.maxAlternateFieldLength	Maximal Größe des alternativen Feldes
hl.formatter	Spezifikation der Formatter-Implementierung
hl.simple.pre	Was vor der Fundstelle eingefügt wird
hl.simple.post	Was nach der Fundstelle eingefügt wird
hl.fragmenter	Spezifikation der Fragmenter-Implementierung
hl.useFastVectorHighlighting	termOffsets, termPositions und termVectors auf true → auf true setzen für schnellere Ausführungszeit

Zugriffsmöglichkeiten

#set(\$hl = \$response.getHighlighting()) oder vorgefertigten Mechanismus aus VM_global_library.vm verwenden

```
#macro(field $f)
  #if($response.response.highlighting.get($docId).get($f).get(0))
    #set($pad = "")
    #foreach($v in $response.response.highlighting.get($docId).get($f))
$pad$v##
    #set($pad = " ... ")
    #end
  #else
    #foreach($v in $doc.getFieldValues($f))
$pad$v##
    #end
  #end
#end
```

product_doc.vm

```
<div>Id: #field('id')</div>
<div>Price: #field('price_c')</div>
<div>Features: #field('features')</div>
<div>In Stock: #field('inStock')</div>
```

More like this


- thematisch relevante Dokumente identifizieren und dem Nutzer präsentieren
- Vergleich von Termvektoren notwendig, mathematisch naheliegende Dokumente im Vektorraum sind ähnlich
- Termvektoren müssen in solrconfig.xml aktiviert werden

```
<field name="category" ... termVectors="true" />
```

Name	Funktion
mlt	true, um MoreLikeThis als Searchfeature zu verwenden (normalerweise nicht sofort erwünscht)
mlt.count	Anzahl der vorgeschlagenen, ähnlichen Dokumente
mlt.fl	Liste der Felder, die für Analyse verwendet werden
mlt.mintf	gibt an, wie oft ein Term in einem Dok. vorkommen muss, um zum Vergleich verwendet zu werden
mlt.mindf	gibt an, in wie vielen Dokumenten ein Term enthalten sein muss, um zum Vergleich verwendet zu werden

Beispieldatei: solrconfig.xml

```
<str name="mlt.qf">  
  text^0.5 features^1.0 name^1.2 sku^1.5 id^10.0 manu^1.1 cat^1.4  
  title^10.0 description^5.0 keywords^5.0 author^2.0 resourcename^1.0  
</str>  
<str name="mlt.fl">text,features,name,sku,id,manu,cat,title,description,keywords,author,resourcename</str>  
<int name="mlt.count">3</int>
```



genutzte Query-Felder und ihr
jeweiliger Boost

→ Felder müssen auch in mlt.fl
spezifiziert werden

In den Velocity-Templates

Beispiel product_doc.vm

browse.vm

```
#set($mltResults = $response.response.get("moreLikeThis"))
```

```
<div class="result-title"><b>#field('name')</b><span class="mlt">  #if($params.getBool('mlt', false) == false)<a  
href="#lensNoQ&q=id:$docId&mlt=true">More Like This</a>#end</span></div>
```

Anfrage mit aktiviertem
MoreLikeThis

Prüfung ob MoreLikeThis –
Dokument vorhanden

wenn MoreLikeThis – Query
aufgerufen wurde, werden
entsprechende Doks angezeigt

```
<div class="mlt">
  #set($mlt = $mltResults.get($docId))
  #set($mltOn = $params.getBool('mlt'))
  #if($mltOn == true)<div class="field-name">Similar Items</div>#end
  #if ($mltOn && $mlt && $mlt.size() > 0)
  <ul>
    #foreach($mltHit in $mlt)
      #set($mltId = $mltHit.getFieldValue('id'))
      <li><div><a href="#url_for_home?q=id:$mltId">$mltId</a></div><div><span class="field-name">Name:</span>
        $mltHit.getFieldValue('name')</div>
        <div><span class="field-name">Price:</span> $!number.currency($mltHit.getFieldValue('price')) <span class="
        field-name">In Stock:</span> $mltHit.getFieldValue('inStock')</div>

      </li>
    #end
  </ul>
  #elseif($mltOn && $mlt.size() == 0)
    <div>No Similar Items Found</div>
  #end
</div>
```

Link zu Queryanfrage der
MoreLikeThis – Dokuments über die
id

Verschiedene Möglichkeiten für MLT-Search

→ extra MoreLikeThis – RequestHandler bauen

```
<requestHandler name="/mlt" class="solr.MoreLikeThisHandler">  
</requestHandler>
```

→ performanter

Beispiele im Netz

Mehr zu Facetten, Highlighting, More like this usw.

<https://cwiki.apache.org/confluence/display/solr/Faceting>

<https://cwiki.apache.org/confluence/display/solr/Highlighting>

<https://cwiki.apache.org/confluence/display/solr/MoreLikeThis>

Die üblichen Quellen:

<https://cwiki.apache.org/confluence/display/solr/Apache+Solr+Reference+Guide>

<https://wiki.apache.org/solr/>

Literatur

bereits genannte Bücher auch hier sehr nützlich!

- Grainger, T., Potter, T. (2014). Solr in Action. Shelter Island, NY: Manning Publications.
- Klose, M., Wrigley. (2014). Einführung in Apache Solr. Praxiseinstieg in die innovative Suchtechnologie. Köln: O'Reilly Verlag.

Genereller Tipp



Funktionalität von SearchFeatures erst im Solr-Admin-UI testen

→ Querys eingeben und Request und Response überprüfen

Trennung von Fehlerquelle – solrconfig.xml (Indexierung) vs .vm – Dateien (UI)

Query-Syntax

Wissen über Query-Syntax oft notwendig um Anfragen zu formulieren z.B. in Form von href-Attributen in a-Elementen.

Erweiterte Suche so realisierbar.

http://lucene.apache.org/core/2_9_4/queryparsersyntax.html

<http://www.solrtutorial.com/solr-query-syntax.html>

Velocity-Templates in der Projektpraxis

- meist kein kompletter Neustart, aber Aufräumen (Löschen nicht benötigter Features etc.)
- Anpassung der bestehenden Templates an eigene Bedürfnisse
- Erweiterung durch eigene Templates
- Erweiterung mit eigenen CSS-, Javascript-, und HTML – Dateien

Im „Real Life“

- Warum nicht Solritas als Live-System/ „in production“ ?
 - Direkter Zugang zur Solr-Instanz wird benötigt (Response Writer) ABER Solr sollte nicht direkt ansprechbar sein - es sollte ein separates Front-End-Layer verwendet werden(über HTTP mit PHP)
 - Generell wird davon abgeraten, falls man sich doch dafür entscheidet sollten folgende Punkte beachtet werden:
 - “if you are going to use a DIH (Data Import Handler) with a database do not let the database configuration plain in *data-config.xml*. Let *data-config.xml* references a jndi of your database config
 - in order to restrict access to the Solr admin interface, change authentication configuration in the solr war’s *web.xml*
 - have in mind that *Solritas* uses a template system ([Apache Velocity](#)). It is limited in the way you wont be able to build backend processing with it
 - fell free to customize **Solritas** layout using CSS3, HTML5, ajax and all client side stuff
 - consider using and [Apache Httpd](#) as a reverse proxy to restrict direct access to your solritas
 - if you have a high traffic landing on you **Solritas**, consider having a [Varnish](#) in front of it”

Quelle: <http://thoughtsasaservice.wordpress.com/2012/05/10/should-you-use-solritas-on-production/>

Nächste Woche

keine Evaluationssitzung! → Evaluation sowohl maschinell als auch nutzerzentriert am 7.7.

bei Nachfrage: Anwendung von Velocity-Templates auf eine Twitter-Suche, Schritt für Schritt (tweets.xml)

weitere Möglichkeit, um ein Web-UI von Solr zu gestalten

→ mehr Kontrolle, komplexere Gestaltungsmöglichkeiten, bessere Nachschlagemöglichkeiten, sichere Javascript und JQuery – Kenntnisse

AJAX Solr

Beispielseite AJAX Solr:

<http://evolvingweb.github.io/ajax-solr/examples/reuters/index.html>

AJAX Solr Demonstration

Browse Reuters business news from 1987

Current Selection

Viewing all documents!

Search

(press ESC to close suggestions)

Top Topics

[acq](#) [alum](#) [barley](#) [bop](#) [carcass](#) [cocoa](#) [coffee](#) [copper](#) [corn](#) [cotton](#) [cpi](#) [crude](#) [dlr](#) [earn](#) [fuel](#) [gas](#) [gnp](#) [gold](#) [grain](#) [interest](#) [ipi](#) [iron-steel](#) [jobs](#) [lead](#) [livestock](#) [meal-feed](#) [money-fx](#) [money-supply](#) [nat-gas](#) [oilseed](#) [orange](#) [palm-oil](#) [pet-chem](#) [rapeseed](#) [reserves](#) [rice](#) [rubber](#) [ship](#) [silver](#) [sorghum](#) [soybean](#) [strategic-metal](#) [sugar](#) [tin](#) [trade](#) [veg-oil](#) [wheat](#) [wpi](#) [yen](#) [zinc](#)

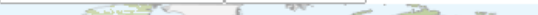
Top Organisations

[adb-africa](#) [adb-asia](#) [atpc](#) [ec](#) [eib](#) [fao](#) [gatt](#) [icco](#) [ico-coffee](#) [ida](#) [iea](#) [imf](#) [inro](#) [itc](#) [iwc-wheat](#) [oecd](#) [opec](#) [un](#) [unctad](#) [worldbank](#)

Top Exchanges

[amex](#) [ase](#) [cboe](#) [cbt](#) [cme](#) [comex](#) [fse](#) [hkse](#) [liffe](#) [lme](#) [lse](#) [nasdaq](#) [nyce](#) [nycse](#) [nymex](#) [nyse](#) [pse](#) [simex](#) [tose](#) [tse](#)

By Country



< 1 2 3 ... 2157 2158 > displaying 1 to 10 of 21578

VALERO ENERGY DEBT UPGRADED BY MOODY'S

NEW YORK, March 17 - Moody's Investors Service Inc said it upgraded Valero Energy Corp's 120 mln dlrs of debt. Moody's cited an improved leverage position that will result from Valero's sale of its Valero Natural Gas subsidiary to Valero Natural Gas Partners L.P., the partial sale of units of that partnership to the [more](#)

NL FILES SUITS AGAINST UNITED CATALYSTS

HIGHTSTOWN, N.J., March 17 - NL Industries Inc's NL Chemicals Inc subsidiary said it filed two complaints against for patent infringements and misappropriation of confidential information. NL said the patent infringement suit alleges that United's product, Thixogel DSS, infringes on NL's patent protec [more](#)

IRVING TRUST BUYS GULF/WESTERN UNIT

[acq](#)
NEW YORK, March 17 - Irving Bank Corp said it bought the factoring division of Associates Commercial Corp, a unit of Gulf and Western Co Inc's Associates Corp of North America. The terms of the previously announced deal were not disclosed. It said the assets were transferred to Irving Commercial Corp. Reuter

DITTLER BROTHERS SEEK LOTTERY INVESTIGATIONS

ATLANTA, March 17 - Dittler Brothers called for investigations by the attorneys general of 24 states in connection with possible violation of state lottery laws by Bally Manufacturing Corp and its subsidiary, Scientific Games. Dittler Brothers said it requested the states' law enforcement chiefs to investigat [more](#)

AMRE INC 3RD QTR JAN 31 NET

Danke für eure Aufmerksamkeit!

