VuFind
Explain Addon

By Dennis Schrittenlocher
Why do we need it?
How are titles found and what is relevance ranking?

Search: vufind

1st result: 2x VuFind in title
- 1x VuFind in description

2nd result: 1x VuFind in title
How are titles found and what is relevance ranking?

Search : VuFInd

Result 10 – 13?
How are titles found and what is relevance ranking?

Sorted by relevance?
What is relevance

Solr:

debug-explainOther

23.08.2023 | Figuring out what relevance is
What is relevance

4617.9854 = max of:
7.1759204 = weight(allfields:vufind in 25554) [SchemaSimilarity], result of:
7.1759204 = score(freq=7.0), computed as boost * idf * tf from:
10.550706 = idf, computed as log(1 + (N - n + 0.5) / (n + 0.5)) from:
17 = n, number of documents containing term
668576 = N, total number of documents with field
0.6801375 = tf, computed as freq / (freq + k1 * (1 - b + b * dl / avgdl)) from:
7.0 = freq, occurrences of term within document
1.2 = k1, term saturation parameter
0.75 = b, length normalization parameter
344.0 = dl, length of field (approximate)
103.47469 = avgdl, average length of field

4617.9854 = weight(title_full:vufind in 25554) [SchemaSimilarity], result of:
4617.9854 = score(freq=2.0), computed as boost * idf * tf from:
600.0 = boost
12.496601 = idf, computed as log(1 + (N - n + 0.5) / (n + 0.5)) from:
2 = n, number of documents containing term
668576 = N, total number of documents with field
0.6158983 = tf, computed as freq / (freq + k1 * (1 - b + b * dl / avgdl)) from:
2.0 = freq, occurrences of term within document
1.2 = k1, term saturation parameter
0.75 = b, length normalization parameter
13.0 = dl, length of field
12.351727 = avgdl, average length of field

71.75931 = weight(series2:vufind in 25554) [SchemaSimilarity], result of:
71.75931 = score(freq=1.0), computed as boost * idf * tf from:
30.0 = boost
10.785799 = idf, computed as log(1 + (N - n + 0.5) / (n + 0.5)) from:
3 = n, number of documents containing term
169183 = N, total number of documents with field
0.2446985 = tf, computed as freq / (freq + k1 * (1 - b + b * dl / avgdl)) from:
1.0 = freq, occurrences of term within document
1.2 = k1, term saturation parameter
0.75 = b, length normalization parameter
16.0 = dl, length of field
5.167475 = avgdl, average length of field

3345.6562 = weight(title:vufind in 25554) [SchemaSimilarity], result of:
3345.6562 = score(freq=2.0), computed as boost * idf * tf from:
500.0 = boost
12.496601 = idf, computed as log(1 + (N - n + 0.5) / (n + 0.5)) from:
2 = n, number of documents containing term
668576 = N, total number of documents with field
0.5354502 = tf, computed as freq / (freq + k1 * (1 - b + b * dl / avgdl)) from:
2.0 = freq, occurrences of term within document
1.2 = k1, term saturation parameter
0.75 = b, length normalization parameter
13.0 = dl, length of field
8.15232 = avgdl, average length of field
Alternative visualization?
Overview of relevance-score
With explain:

General idea of relevance-score
**Most common course**

<table>
<thead>
<tr>
<th>#</th>
<th>Title not available</th>
<th>Call Number</th>
<th>Loading...</th>
<th>Located: Loading...</th>
<th><strong>Relevance</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Loading...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>2</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Loading...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>3</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>4</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>5</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>6</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>7</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>8</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>9</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>10</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>11</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>12</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>13</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>14</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>15</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>16</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>17</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>18</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>19</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
<tr>
<td>20</td>
<td>Title not available</td>
<td>Call Number</td>
<td>Loading...</td>
<td>Located: Learning...</td>
<td><strong>Relevance</strong></td>
</tr>
</tbody>
</table>

**Showing 1 - 20 results of 4,684 for search: 'wörter', query time: 5.34s**
Detailed Explanation
Three way explanation

- Explanation in text-form
  - Brakedown of relevance score

- Pie-chart
  - Graph of text explanation

- Overall relevance
  - Relevance-score in relation to max relevancescore
Detailed example

Explanation for search: hamburg berlin china

Record id: 211536474 found with a relevance value of 108,144.01
Product of 5,586.53 (relevance value) * 19.00 (boost)
(BOOST = const(19))

33.25% "china" in: title_full_unstemmed (explainFieldName_title_full_unstemmed), with a modifier of 600
19.37% "hamburg" in: title_short (explainFieldName_title_short), with a modifier of 750
16.60% "berlin" in: title_short (explainFieldName_title_short), with a modifier of 750
8.67% "china" in: title_full (explainFieldName_title_full), with a modifier of 400
6.77% "hamburg" in: title_full (explainFieldName_title_full), with a modifier of 400
5.82% "china" in: geographic (explainFieldName_geographic), with a modifier of 300
5.54% "berlin" in: title_full_unstemmed (explainFieldName_title_full_unstemmed), with a modifier of 600
3.60% "berlin" in: title_full (explainFieldName_title_full), with a modifier of 400
0.30% "china" in: allfields_unstemmed (explainFieldName_allfields_unstemmed), with a modifier of 10
0.08% "berlin" in: allfields_unstemmed (explainFieldName_allfields_unstemmed), with a modifier of 10
0.00% "china" in: title (explainFieldName_title), with a modifier of 0.1
0.00% "hamburg" in: title (explainFieldName_title), with a modifier of 0.1
0.00% "berlin" in: title (explainFieldName_title), with a modifier of 0.1
0.00% "china" in: allfields (), with a modifier of 0.0
0.00% "berlin" in: allfields (), with a modifier of 0.0
0.00% "hamburg" in: allfields (), with a modifier of 0.0
Detailed example

Explanation for search: hamburg berlin china

Record id: 192563482 found with a relevance value of 90.431.62
Product of 4.759.56 (relevance_value) * 19.00 (boost)
(boost) = const(19)

40.31% "china" in: title_full_unstemmed (explain_fieldName_title_full_unstemmed), with a modifier of 600
20.52% "hamburg" in: title_full (explain_fieldName_title_full), with a modifier of 400
18.94% "berlin" in: title_full_unstemmed (explain_fieldName_title_full_unstemmed), with a modifier of 600
10.51% "china" in: title_full (explain_fieldName_title_full), with a modifier of 400
6.83% "china" in: geographic (explain_fieldName_geographic), with a modifier of 300
4.40% "berlin" in: title_full (explain_fieldName_title_full), with a modifier of 400
0.37% "china" in: allfields_unstemmed (explain_fieldName_allfields_unstemmed), with a modifier of 10
0.10% "berlin" in: allfields_unstemmed (explain_fieldName_allfields_unstemmed), with a modifier of 10
0.00% "china" in: title (explain_fieldName_title), with a modifier of 0.1
0.00% "hamburg" in: title (explain_fieldName_title), with a modifier of 0.1
0.00% "berlin" in: title (explain_fieldName_title), with a modifier of 0.1
0.00% "china" in: allfields (), with a modifier of 0.0
0.00% "berlin" in: allfields (), with a modifier of 0.0
0.00% "hamburg" in: allfields (), with a modifier of 0.0

Relevance in comparison to the relevance of the top result
Detailed example

Explanation for search: hamburg berlin china

Record Id: 211705074 found with a relevance value of 85,662.78
Product of 4,476.99 (relevance value) * 19.00 (boost)
(boost) = const(19)

72.81% "china" in: title_short (explain_fieldName_title_short) , with a modifier of 750
15.81% "china" in: title_full_unstemmed (explain_fieldName_title_full_unstemmed) , with a modifier of 600
10.30% "china" in: title_full (explain_fieldName_title_full) , with a modifier of 400
0.55% "berlin" in: allfields_unstemmed (explain_fieldName_allfields_unstemmed) , with a modifier of 10
0.33% "china" in: allfields_unstemmed (explain_fieldName_allfields_unstemmed) , with a modifier of 10
0.19% "berlin" in: allfields_unstemmed (explain_fieldName_allfields_unstemmed) , with a modifier of 10
0.00% "china" in: title (explain_fieldName_title) , with a modifier of 0.1
0.00% "china" in: allfields ( ) , with a modifier of 0.0
0.00% "berlin" in: allfields ( ) , with a modifier of 0.0
0.00% "hamburg" in: allfields ( ) , with a modifier of 0.0
Detailed example

Match 3

Match 4
**Detailed example**

Explanation for search: hamburg berlin china

Record Id: 451613511 found with a relevance value of 13,353.04
Product of 702.79 (relevance value) * 19.00 (boost)
(boos) - const 19

53.93% "china" in: geographic (explain_fieldName_geographic) , with a modifier of 300
42.88% "berlin" in: geographic (explain_fieldName_geographic) , with a modifier of 300
1.33% "hamburg" in: allfields_unstemmed (explain_fieldName_allfields_unstemmed) , with a modifier of 10
1.32% "china" in: allfields_unstemmed (explain_fieldName_allfields_unstemmed) , with a modifier of 10
0.54% "berlin" in: allfields_unstemmed (explain_fieldName_allfields_unstemmed) , with a modifier of 10
0.00% "china" in: allfields () , with a modifier of 0.0
0.00% "berlin" in: allfields () , with a modifier of 0.0
0.00% "hamburg" in: allfields () , with a modifier of 0.0
Detailed example

Explanation for search: hamburg berlin china

Record Id: 011038953 found with a relevance value of 606.08
Product of 31.90 (relevance value) * 19.90 (boost)
(boom) – const(19)

45.79% "china" in: allfields_unstemmed (explain_fieldName_allfields_unstemmed), with a modifier of 10
35.39% "hamburg" in: allfields_unstemmed (explain_fieldName_allfields_unstemmed), with a modifier of 10
18.82% "berlin" in: allfields_unstemmed (explain_fieldName_allfields_unstemmed), with a modifier of 10
0.00% "china" in: allfields (), with a modifier of 0.0
0.00% "berlin" in: allfields (), with a modifier of 0.0
0.00% "hamburg" in: allfields (), with a modifier of 0.0
Adjusting searchspecs
AllFields:
DismaxFields:
- title_short^750
- title_full_unstemmed^600
- title_full^400
- title^500
- title_alt^200
- title_new^100
- series^50
- series2^30
- author^300
- contents^10
- topic_unstemmed^550
- topic^500
- geographic^300
- genre^300

DismaxHandler: edismax
- allfields_unstemmed^10
- fulltext_unstemmed^10
- allfields
- fulltext
- description
- isbn
- issn
- long_lat_display
Example move result

   Published 2007
   An error has occurred
   Call Number: Loading...
   Located: Loading...
   kostenlos
   kostenlos
   kostenlos
   thesis
   ebook

   by Kopp, Eduard 1960
   An error has occurred
   Call Number: Loading...
   Located: Loading...
   Inhaltsverzeichnis
   thesis
   ebook

3. Neue deutsche Außenpolitik: Netzwerke statt Allianzen
   by Fenn, Daniel 1974
   Published 2016
   An error has occurred
   Call Number: Loading...
   Located: Loading...
   % Get full text
   kostenlos
   ebook

4. China aktuell journal of current Chinese affairs
   Published 1972
   An error has occurred
   Call Number: Loading...
   Located: Loading...
   kostenlos
   Get full text
   electronic journal

5. Drache und Kreuz in China Nachrichtenblatt des Chinamissionsbundes
   Published 1910
   An error has occurred
   Call Number: Loading...
   Located: Loading...
   kostenlos
   electronic journal

6. Wirtschaftswocche Pflichtblatt der Wertpapierbörse in Frankfurt und Düsseldorf
   Published 1973
   Call Number: Loading...
   Located: Loading...
   %Get full text
   electronic newspaper
   Loading...
### Changing searchspecs

<table>
<thead>
<tr>
<th>#</th>
<th>Title</th>
<th>Author(s)</th>
<th>Year</th>
<th>Edition</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Kampf ums Dasein. Darwin-Diskurse und die politische Literatur bis 1900</td>
<td>Bühmann, Daniel; Schirmer, Daniel</td>
<td>2015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Bloomsbury scientists, science and art in the wake of Darwin</td>
<td>Studer, Michael Charles 1942; Studer, Michael Charles 1962</td>
<td>2017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Die Grosses Leben und Leistungen der 800 bedeutendsten Persönlichkeiten unserer Welt</td>
<td></td>
<td>1977</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Brown'shill Princess - A Study of the &quot;Faux Conscience&quot;</td>
<td>Gourley, Lloyd</td>
<td>2013</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Search: Darwin
Changing searchspecs

Explanation for search: darwin
Record id: 85459350 found with a relevance value of 110,545.01
Product of 5,818.16 (relevance value) * 19.00 (boost)
(boost) = const(19)

62.79% "darwin" in: title_full_unstemmed (explain_fieldName_title_full_unstemmed), with a modifier of 600
20.17% "darwin" in: title (explain_fieldName_title), with a modifier of 500
16.74% "darwin" in: title_full (explain_fieldName_title_full), with a modifier of 400
0.30% "darwin" in: allfields_unstemmed (explain_fieldName_allfields_unstemmed), with a modifier of 10
0.00% "darwin" in: allfields ( ), with a modifier of 0.0

Explanation for search: darwin
Record id: 235501793 found with a relevance value of 49,657.93
Product of 2,569.94 (relevance value) * 19.00 (boost)
(boost) = const(19)

78.49% "darwin" in: title_full_unstemmed (explain_fieldName_title_full_unstemmed), with a modifier of 600
29.91% "darwin" in: title_full (explain_fieldName_title_full), with a modifier of 400
0.70% "darwin" in: allfields_unstemmed (explain_fieldName_allfields_unstemmed), with a modifier of 10
0.00% "darwin" in: allfields ( ), with a modifier of 0.0
Changing searchspecs
Changing searchspecs

Explanation for search: darwin
Record id: 2304415050 found with a relevance value of 88,256.51
Product of 4,645.08 (relevance value) * 19.00 (boost)
(boost) = const(19)

78.05% "darwin" in: title_full_unstemmed (explain_fieldName_title_full_unstemmed), with a modifier of 600
20.97% "darwin" in: title_full (explain_fieldName_title_full), with a modifier of 400
0.37% "darwin" in: allfields_unstemmed (explain_fieldName_allfields_unstemmed), with a modifier of 10
0.01% "darwin" in: title (explain_fieldName_title), with a modifier of 0.1
0.00% "darwin" in: allfields (), with a modifier of 0.0

Relevance in comparison to the relevance of the top result

Explanation for search: darwin
Record id: 2304417972 found with a relevance value of 48,657.93
Product of 2,556.94 (relevance value) * 19.00 (boost)
(boost) = const(19)

78.40% "darwin" in: title_full_unstemmed (explain_fieldName_title_full_unstemmed), with a modifier of 600
20.91% "darwin" in: title_full (explain_fieldName_title_full), with a modifier of 400
0.70% "darwin" in: allfields_unstemmed (explain_fieldName_allfields_unstemmed), with a modifier of 10
0.00% "darwin" in: allfields (), with a modifier of 0.0

Relevance in comparison to the relevance of the top result
**Config: min_percent and max_fields**

- **Top Result Relevance:** 1.43
- **Explanation for search:** hamburg berlin china
  - Record ID: HEBE11586474 found with a relevance value of 1.31
  - 20.39% "xina" in topic
  - 20.39% "tSina" in topic
  - **Rest**
    - 12.96% "hamburk" in title (explain_fieldName_title)
    - 12.97% "hanburk" in title (explain_fieldName_title)
    - 12.97% "amburk" in title (explain_fieldName_title)
    - 5.44% "berlin" in title (explain_fieldName_title), with a modifier of 500
    - 0.94% "xina" in title (explain_fieldName_title)
    - 0.94% "tSina" in title (explain_fieldName_title)
    - 0.00% "china" in allfields
    - 0.00% "berlin" in allfields
  - **Sum:** 59.22%

**Addendum:**
- Adds a rest which contains all fields with percentages lower than min_percent or if there are more fields the max_fields, then it contains the extra ones.
Why „Helix Joe Biden“ => „Brehms Tierleben“

Top Result Relevance: 0.00

Explanation for search: helix joe biden

Record Id: HEB048748900 found with a relevance value of 0.00
Product of 0.00 (relevance value) * 1.00 (boost)
(boost) =
product(if(exists(query(id:HEBr*,def=0.0)=0.0),const(0.4),const(13)),const(1)),const(6.5)),const(1))

48.13% "biden" in: misc (), with a modifier of 0.01
32.20% "helix" in: misc (), with a modifier of 0.01
19.67% "jo" in: misc (), with a modifier of 0.01

4. Sippe: Saatkäfer, Pschinen
   Psyche helix
   Rohreutopo (Psyche unicolor)

3. Sippe: Nymphalinen
   Silberstrich, Kaisermantel (Argynnis phia)
   Großer Perlmutterfalter (Argynnis Aglaja)
   Scheckenfalter (Melitaea)
   Pauwenpiegel (Vanessa Jo)
   Admiral (Vanessa Atalanta)

4. Sippe: Lampinen
   Omatus auratus, Elampus aeneus, biden-
tulus
References:

Example VuFind system hosted by hebis (w/o explain):  [https://ubffm.hds.hebis.de/](https://ubffm.hds.hebis.de/)

Demo VuFind9 with explain:  
[https://hds.hebis.de/vufind-demo/Search/Results?lookfor=hamburg+berlin+china](https://hds.hebis.de/vufind-demo/Search/Results?lookfor=hamburg+berlin+china)

Inspiration, presentation , Rafal Kuc, Solr.pl, Eurocon 2011:  
[https://www.youtube.com/watch?v=PwJNFHRDpVQ](https://www.youtube.com/watch?v=PwJNFHRDpVQ)

Source of graphic library:  [https://www.chartjs.org/](https://www.chartjs.org/)

Butterfly image:  [https://interlinking.bbf.dipf.de/index.php/Ad99998_04_029c](https://interlinking.bbf.dipf.de/index.php/Ad99998_04_029c)

Contact

Questions: p2_fragen@hebis.de