The Problems with indexed_search and how to fix 'em

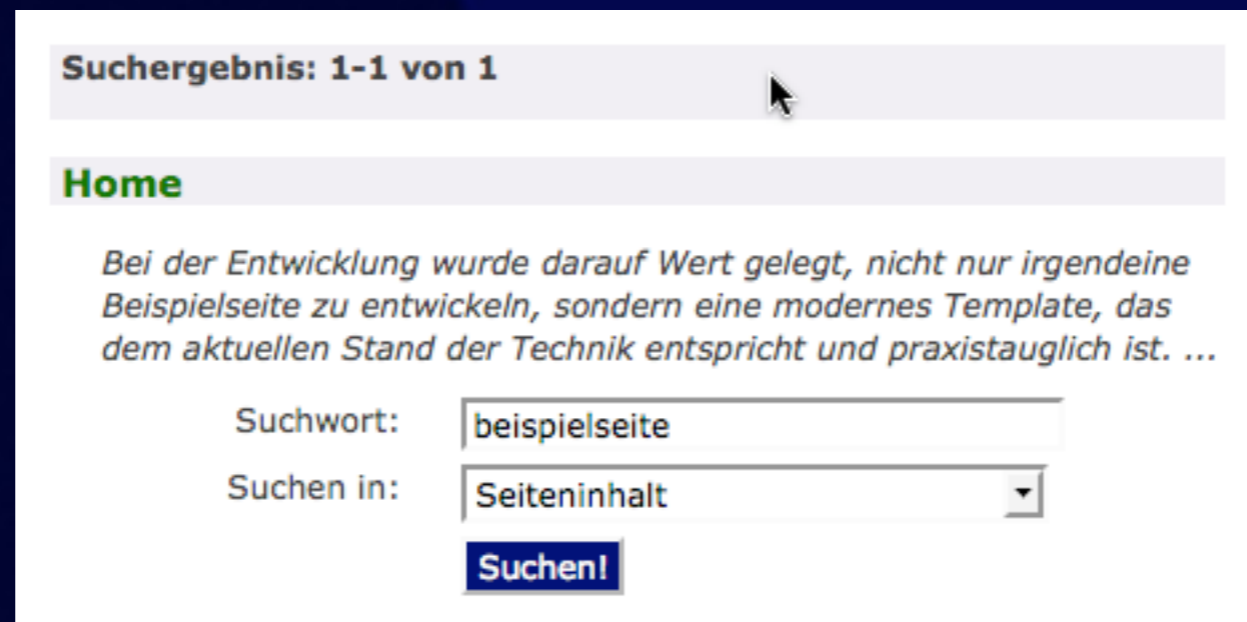Jochen Weiland
TYPO3 Conference Dallas, 2010

jweiland.net

3 Search Functions

- Standard Search

- solr

- indexed_search

jweiland.net

Standard Search

- Slow

- Headlines and text from tt_content

- No advanced search

- No external documents

jweiland.net

TYPO3 solr

- "Google for TYPO3"

- Fast, powerful

- Requires separate Index-Server (Apache Tomcat)

jweiland.net

indexed_search

- Fast (if configured right)

- external Files (pdf, doc, xls, txt, ...)

- Problems with re-indexing

jweiland.net

Delete Search Index

- TRUNCATE `index_fulltext`;

- TRUNCATE `index_grlist`;

- TRUNCATE `index_phash`;

- TRUNCATE `index_rel`;

- TRUNCATE `index_section`;

- TRUNCATE `index_words`;

jweiland.net

```
<body>
  <div id="fullheightcontainer">
    <div id="content">
      <div id="contentWrap"> </div>
      <div id="leftmenu"> </div>
    </div>
    <div id="header"> </div>
    <div id="logo"> </div>
    <div id="rootline"> </div>
    <div id="search"> </div>
    <div id="printlink"> </div>
    <div id="servicemenu"> </div>
    <div id="language"> </div>
  </div>
  <div id="footer"> </div>
</body>
```

jweiland.net

```html
<body>
  <div id="fullheightcontainer">
    <div id="content">
      <div id="contentWrap"> </div>
      <div id="leftmenu"> </div>
    </div>
    <div id="header"> </div>
    <div id="logo"> </div>
    <div id="rootline"> </div>
    <div id="search"> </div>
    <div id="printlink"> </div>
    <div id="servicemenu"> </div>
    <div id="language"> </div>
  </div>
  <div id="footer"> </div>
</body>
```

```html
<body>
  <div id="fullheightcontainer">
    <div id="content">
  <!--TYPO3SEARCH_begin-->
      <div id="contentWrap"> </div>
  <!--TYPO3SEARCH_end-->
      <div id="leftmenu"> </div>
    </div>
    <div id="header"> </div>
    <div id="logo"> </div>
    <div id="rootline"> </div>
    <div id="search"> </div>
    <div id="printlink"> </div>
    <div id="servicemenu"> </div>
    <div id="language"> </div>
  </div>
  <div id="footer"> </div>
</body>
```

jweiland.net

```html
<body>
  <div id="fullheightcontainer">
    <div id="content">
    <!--TYPO3SEARCH_begin-->
       <div id="contentWrap"> </div>
       <div id="leftmenu"> </div>
     </div>
     <div id="header"> </div>
     <div id="logo"> </div>
     <div id="rootline"> </div>
     <div id="search"> </div>
     <div id="printlink"> </div>
     <div id="servicemenu"> </div>
     <div id="language"> </div>
   </div>
   <div id="footer"> </div>
</body>
```

jweiland.net

```html
<body>
  <div id="fullheightcontainer">
    <div id="content">
      <div id="contentWrap"> </div>
      <div id="leftmenu"> </div>
    </div>
    <div id="header"> </div>
    <div id="logo"> </div>
    <div id="rootline"> </div>
    <!--TYPO3SEARCH_end-->
    <div id="search"> </div>
    <div id="printlink"> </div>
    <div id="servicemenu"> </div>
    <div id="language"> </div>
  </div>
  <div id="footer"> </div>
</body>
```

jweiland.net

```
<body>
  <div id="fullheightcontainer">
    <div id="content">
    <!--TYPO3SEARCH_begin-->
       <div id="contentWrap"> </div>
    <!--TYPO3SEARCH_end-->
       <div id="leftmenu"> </div>
    </div>
    <div id="header"> </div>
    <div id="logo"> </div>
    <div id="rootline"> </div>
    <!--TYPO3SEARCH_begin-->
    <div id="search"> </div>
    <div id="printlink"> </div>
    <!--TYPO3SEARCH_end-->
    <div id="servicemenu"> </div>
    <div id="language"> </div>
   </div>
   <div id="footer"> </div>
</body>
```

jweiland.net

Speed

```
plugin.tx_indexedsearch._DEFAULT_PI_VARS.type = 0
```

jweiland.net

Page Browser not working?

- In Template replace

  <form action="###ACTION_URL###" method="post" id="tx_indexedsearch">

  with

  <form action="###ACTION_URL###" method="post" name="tx_indexedsearch">

Crawler Re-Index

- ```
  tx_crawler.crawlerCfg.paramSets.index =
  &contentId=[_TABLE:tt_content]
  tx_crawler.crawlerCfg.paramSets.index {
      procInstrFilter = tx_indexedsearch_reindex
      baseUrl = http://user60.t3sem.de/
  }
  ```

jweiland.net

Crawler Re-Index

- tx_crawler.crawlerCfg.paramSets.index-news =
  &tx_ttnews[backPid]=14&tx_ttnews[pointer]=1
  &tx_ttnews[pointer]=1
  &tx_ttnews[tt_news]=[_TABLE:tt_news;_PID:11]
  # PID 11 is "news storage folder"

  tx_crawler.crawlerCfg.paramSets.index-news {
      procInstrFilter = tx_indexedsearch_reindex
      cHash = 1
      pidsOnly = 12 # PID 12 is "news single view"
      baseUrl = http://user60.t3sem.de/
  }

jweiland.net

Invoke Crawler

- typo3/cli_dispatch.phpsh crawler_im 1
  -d=10 -conf index -n 10000 -o queue

  cli_dispatch.phpsh crawler_im 1
  -d=10 -conf index-news -n 10000 -o queue

  /typo3/cli_dispatch.phpsh crawler

jweiland.net

- Download Presentation:

  http://jweiland.net/t3con10

- Vote and comment:

  http://joind.in/talk/view/1678

jweiland.net

Questions?

jweiland.net