

Durchführung eines Testimports

Um VuFind und seine Funktionen erstmalig in Augenschein zu nehmen, empfiehlt sich die Durchführung eines Testimports. Dabei laden Sie Testdaten in den Suchindex von VuFind und lernen somit bereits, wie Sie einen ersten einfachen Datenimport durchführen. Gut geeignet für einen Testimport sind Testdaten der Deutschen Nationalbibliothek (DNB).

1 Download von Testdaten

Öffnen Sie die Seite des „Datendienstes Bibliografische Dienstleistungen“ auf der Webseite der DNB:

<http://www.dnb.de/datendienst>

Wählen Sie dort (ohne Abbildung) den Link „Testdaten“ aus.

Es öffnet sich eine Liste mit den verfügbaren Testdaten:

Datendienst "Bibliografische Dienstleistungen"

Bezieher: testdat

Die folgenden Dateien können Sie zurzeit abrufen; bitte wählen Sie die Datei, die Sie übertragen möchten, durch Klicken auf den entsprechenden Dateinamen aus.

Dateiname	Anzahl der Datensätze	Dateigröße (Bytes)	Erzeugt am
Atest.htm	2243 / Testdaten Reihe A im Format HTML	1961860	14.01.2015
Atest.mrc.xml	2519 / Testdaten Reihe A im Format MARC21xml	15267480	14.01.2015
Atest.pdf	2243 / Testdaten Reihe A im Format PDF	2576297	14.01.2015
Atest.utf8.mrc	2519 / Testdaten Reihe A im Format MARC21, Zeichensatz UTF8	4367177	14.01.2015
Btest.htm	519 / Testdaten Reihe B im Format HTML	395899	14.01.2015
Btest.mrc.xml	595 / Testdaten Reihe B im Format MARC21xml	2788495	14.01.2015
Btest.pdf	519 / Testdaten Reihe B im Format PDF	659963	14.01.2015
Btest.utf8.mrc	595 / Testdaten Reihe B im Format MARC21, Zeichensatz UTF8	808905	14.01.2015
Ctest.htm	630 / Testdaten Reihe C im Format HTML-Format	847517	12.05.2015
Ctest.mrc.xml	1141 / Testdaten Reihe C im Format MARC21xml	7308160	12.05.2015
Ctest.pdf	630 / Testdaten Reihe C im Format PDF	943890	12.05.2015
Ctest.utf8.mrc	1141 / Testdaten Reihe C im Format MARC21, Zeichensatz UTF8	2046688	12.05.2015

Laden Sie die Datei << Atest.utf8.mrc >> in das Verzeichnis << Downloads >> Ihres Homeverzeichnis herunter.

2 Start von VuFind

Starten Sie VuFind:

```
cd /usr/local/vufind2  
./vufind.sh start
```

3 Anpassung der Einstellungen für den Import

Öffnen Sie mit Mousepad die Datei << marc_local.properties >> im Verzeichnis << /usr/local/vufind2/import/ >>:

```
#####  
# marc_local.properties -- local overrides to standard VuFind MARC mappings. #  
#                               All settings in this file override equivalent #  
#                               settings in marc.properties. #  
#####  
  
# Uncomment the following settings to insert appropriate values for your site:  
#collection = "Catalog"  
#institution = "MyInstitution"  
#building = "Library A"
```

Editieren Sie die drei Werte „collection“, „institution“ und „building“ wie folgt:

```
#####  
# marc_local.properties -- local overrides to standard VuFind MARC mappings. #  
#                               All settings in this file override equivalent #  
#                               settings in marc.properties. #  
#####  
  
# Uncomment the following settings to insert appropriate values for your site:  
collection = "DNB Reihe A"  
institution = "DNB"  
building = ""
```

Speichern Sie die Änderungen.

4 Starten des Importes

Öffnen Sie ein weiteres Terminal:

```
cd /usr/local/vufind2
```

Starten Sie den Import:

```
./import-marc.sh /home/stefan/Downloads/Atest.utf8.mrc
```

(Achten Sie darauf, hierbei „stefan“ durch den Namen Ihres Benutzers zu ersetzen.)

Während des Importes wird für jeden Datensatz aus der Datei eine Statusmeldung in der Form „Added record 2519 from file:99840540X“ ausgegeben. Zum Abschluss meldet der Import unter anderem, wie viele Datensätze er dem Index hinzugefügt hat:

```
INFO [main] (MarcImporter.java:317) - Added record 2517 read from file: 996842713
INFO [main] (MarcImporter.java:317) - Added record 2518 read from file: 997431709
INFO [main] (MarcImporter.java:317) - Added record 2519 read from file: 99840540X
INFO [main] (MarcImporter.java:615) - Adding 2519 of 2519 documents to index
INFO [main] (MarcImporter.java:616) - Deleting 0 documents from index
INFO [main] (MarcImporter.java:490) - Calling commit (with optimize set to false)
INFO [main] (MarcImporter.java:502) - Done with the commit, closing Solr
INFO [main] (MarcImporter.java:505) - Setting Solr closed flag
INFO [main] (MarcImporter.java:625) - Finished indexing in 1:48,00
INFO [main] (MarcImporter.java:634) - Indexed 2519 at a rate of about 23.0 per sec
INFO [main] (MarcImporter.java:635) - Deleted 0 records
INFO [Thread-1] (MarcImporter.java:565) - Starting Shutdown hook
INFO [Thread-1] (MarcImporter.java:584) - Finished Shutdown hook
stefan@VuFind03v:/usr/local/vufind2$
```

5 Neustart von VuFind

Wechseln Sie nun in das erste Terminal. Drücken Sie die Enter-Taste, um einen Cursor sichtbar zu machen. Starten Sie VuFind neu:

```
./vufind.sh restart
```

6 Suche in VuFind

Öffnen Sie VuFind in Firefox und starten Sie eine leere Suche, indem Sie den Suchschlitz leer lassen und einfach „Suchen“ anklicken.

The screenshot shows the VuFind search interface. At the top, there is a search bar with the text 'Suche:' and a search button labeled 'Suchen'. Below the search bar, there are filters for 'Ähnliche Schlagwörter...' (Literatur 19, Frau 15, Weihnachtsmann 19, Freundschaft 15, Deutsch 16, Junge 14) and 'Suche einschränken' (Institution: DNB 2.519, Format: Book 2.031, Serial 274, Journal 83, Software 76, Audio 56, Kit 40). The search results are sorted by 'Relevanz' and show 1 - 20 of 2.519 results. The first result is 'O du heiliger Weihnachtsbaum Vergnügliches für die Advents- und Weihnachtszeit' (Book) and the second is 'Küchen Themenmagazin für Planen, Bauen, Wohnen' (Zeitschrift).

Sie können unter anderem erkennen, dass alle 2.519 importierten Datensätze mit der leeren Suche gefunden wurden. Außerdem wird die Suchdauer mit 8,08 Sekunden angegeben. (Bei Ihnen kann dieser Wert abweichen.)

Die Suchdauer lässt sich verkürzen, indem Sie den Index optimieren.

7 Index optimieren

Wechseln Sie in das Terminal, in welchem Sie den Datenimport durchgeführt haben:

```
cd util
php optimize.php
```

Die Optimierung läuft ohne Anzeigen im Terminal ab und ist beendet, wenn wieder ein Cursor im Terminal erscheint.

8 Suche in VuFind

Führen Sie erneut eine leere Suche durch:

The screenshot shows the VuFind search interface. At the top, there is a search bar with the text "Suche:" and a search button labeled "Suchen". Below the search bar, there are several filters and options. On the left, there are "Ähnliche Schlagwörter..." (Similar keywords) including "Literatur" (19), "Frau" (15), "Weihnachtsmann" (19), "Freundschaft" (15), "Deutsch" (16), and "Junge" (14). In the center, there is a "Sortieren" (Sort) dropdown menu set to "Relevanz". On the right, there is a "Suche einschränken" (Filter search) section with two main categories: "Institution" and "Format". Under "Institution", "DNB" is selected with 2,519 results. Under "Format", "Book" is selected with 2,031 results. The search results list shows two items: 1. "O du heiliger Weihnachtsbaum Vergnügliches für die Advents- und Weihnachtszeit" (Published 2014, Book format) and 2. "Küchen Themenmagazin für Planen, Bauen, Wohnen" (Magazine format). The search duration is displayed as "Suchdauer: 2,23s".

Die Suchdauer ist nun geringer geworden (hier: 2,23 Sekunden).

9 Sicherungspunkt in VirtualBox setzen

Fahren Sie Xubuntu herunter und setzen Sie in VirtualBox einen weiteren Sicherungspunkt namens „mit geladenen Daten aus Testimport“.

10 Weitere Informationen zum Datenimport

Bei jedem Datenimport werden die einzelnen Felder eines Datensatzes einem Indexfeld im Solr-Index zugeordnet. Dies geschieht mit Hilfe von Zuordnungsdateien wie der oben verwendeten << marc_local.properties >>. Eine Zuordnungsdatei steuert, welche Felder aus den zu importierenden Daten in welches Indexfeld im Solr-Index importiert werden.

Direkt in VuFind importieren können Sie Daten in den Formaten:

- MARC21 (binäres MARC),
- MARCXML und
- XML.

Für den Import von Daten in XML-Formaten sind größere Anpassungen erforderlich. Dies betrifft ebenfalls das Harvesting von Daten über OAI-PMH ein. Innerhalb des Tutorials werden wir uns ausschließlich mit dem Datenimport in den beiden MARC-Formaten beschäftigen.

Nächstes Kapitel

Überblick zur Konfiguration

Quellen

Java Tuning. VuFind Documentation. <https://vufind.org/wiki/performance>

MARC Records. VuFind Documentation. https://vufind.org/wiki/importing_records

Starting and Stopping VuFind. VuFind Documentation. https://vufind.org/wiki/starting_and_stopping_vufind