

„Auf dem Weg zur eigenen Suche“ – Ein Tutorial zur Inbetriebnahme einer VuFind-Demoversion

Stefan Niesner

Einführung

Als Stand der Dinge eines modernen Online-Bibliothekskatalogs darf heute ein sogenanntes Discovery System angesehen werden.¹ Ein Discovery System basiert auf Suchmaschinentechologie und integriert Metadaten aus heterogenen Quellen in einem zentralen Suchindex. Neben dem Suchindex sind wesentliche Elemente eines Discovery Systems eine Suchmaske mit nur einem Suchschlitz, die Sortierung der Suchergebnisse mittels Relevanzranking sowie eine nachträgliche Einschränkung des Suchergebnisses mittels Facetten.²

Discovery Systeme gibt es als kostenpflichtige Produkte³ und auf Basis von Open Source-Software. Die entscheidenden Aspekte der Definition von Open Source sind ein frei verfügbarer Quelltext sowie die Erlaubnis zur beliebigen Nutzung, Veränderung und Verbreitung der Software.⁴ Für Bibliotheken besteht durch den Einsatz von Open Source-Software eine Möglichkeit zur Anpassung an lokale Gegebenheiten und Nutzerwünsche. Für diese Anpassungen kann eine Bibliothek auf den wiederum kostenpflichtigen Service externer Dienstleister setzen oder sie kann selbst das Wissen zur Durchführung der Anpassungen aufbauen. Als Beitrag zu diesem Wissen hat der Verfasser in einem Projekt⁵ ein Einsteiger-Tutorial zur Inbetriebnahme einer Demoversion des weithin bekannten Open Source Discovery Systems VuFind erstellt. Das Tutorial steht seit dem 27.02.2016 bei GitHub⁶ zur Verfügung und wurde am 14.03.2016

Neben kommerziellen Discovery Systemen gibt es Discovery Systeme auf Basis von Open Source-Software. Letztere geben Bibliotheken eine Möglichkeit zur Anpassung ihres Bibliothekskatalogs an lokale Gegebenheiten und Nutzerwünsche. Beschrieben wird ein Einsteiger-Tutorial für die Inbetriebnahme einer Demoversion des Open Source Discovery Systems VuFind. Im Verlauf des Tutorials wird eine lokal verfügbare VuFind-Instanz ohne Anbindung an ein Lokalsystem installiert, Metadaten in diese importiert und verschiedene Elemente des Systems konfiguriert. Die einzelnen Kapitel des Tutorials werden vorgestellt, bevor abschließend ein Ausblick auf eine mögliche Version für Fortgeschrittene gegeben wird.

Discovery Systems can be based on commercial software or Open Source software. When choosing the latter one libraries are in the position to customize their library catalog according to their own and their patrons needs. Described is a beginners' tutorial for implementing a demo version of the Open Source Discovery System VuFind. During the tutorial a local instance of VuFind without a ILS connection will be installed, metadata will be imported into it and several elements of the system will be configured. The individual chapters of the tutorial are presented, before eventually a future perspective on an advanced version is given.

beim 6. Bibliothekskongress in Leipzig im Rahmen eines Clips⁷ erstmals der breiten Fachöffentlichkeit vorgestellt.

In diesem Artikel wird nach einer Kurzeinführung in VuFind zunächst auf die Zielgruppen eingegangen, bevor die Inhalte des Tutorials anhand der einzelnen Kapitel näher beschrieben werden.

VuFind

VuFind wurde im Jahr 2007 an der Falvey Memorial Library der Villanova University (USA) entwickelt. Die aktuelle Version ist die Version 2.5.2 vom 03.02.2016. Mit VuFind können Metadaten aus unterschiedlichen Quellen in einer Suche vereinigt werden. Dies geschieht über den Suchindex Apache Solr.⁸ Vu-

1 Herrlich, Bernhard; Ledl, Andreas; Tréfás, David: Editorial: Die Farbe des Katalogs, in: 027.7 Zeitschrift für Bibliothekskultur 1 (2013) S. 43-46, hier S. 44, URL: http://www.0277.ch/ojs/index.php/cdrs_0277/article/view/25/75 [07.03.2016].

2 Jansen, Heiko: Discovery-Services – Einführung, Marktübersicht und Trends, in: Bibliotheksdienst 48 (2014), S. 773-783.

3 Eine Übersicht des Marktes in Deutschland findet sich bei Jansen 2014 (wie Anm. 2), S. 776 ff.

4 URL: <http://opensource.org/osd> [07.03.2016].

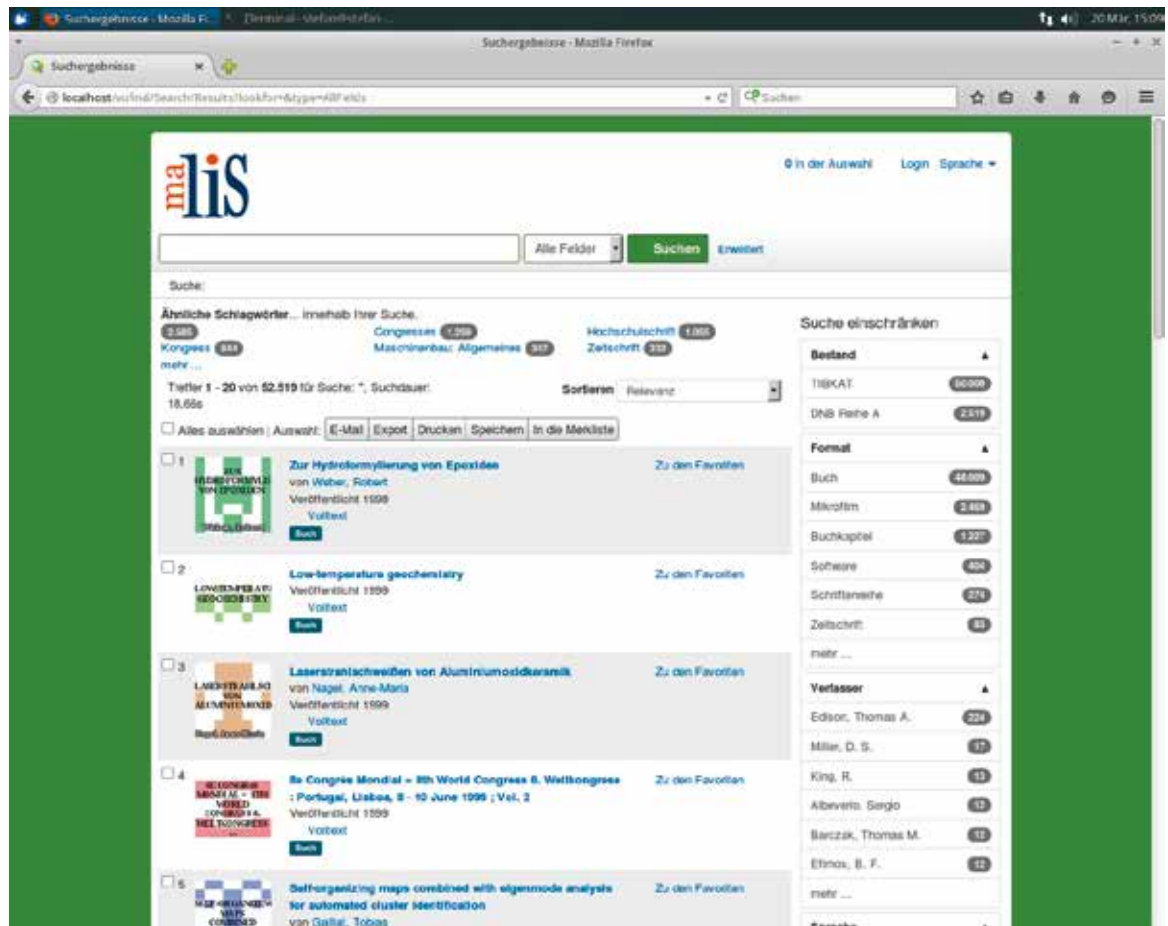
5 Das Projekt fand von April bis Juli 2015 im Rahmen des berufsbegleitenden Masterstudiengangs „Bibliotheks- und Informationswissenschaft“ der TH Köln (vormals FH Köln) statt und wurde von Herrn Dr. Peter Kostädt, USB Köln, betreut.

6 URL: <https://github.com/stefanniesner/vufindtutorialde> [07.03.2016].

7 URL: <https://www.youtube.com/watch?v=nSIFdYjEZso> [20.03.2016]

8 Eine ausführliche Darstellung der Systemarchitekturen von Solr und VuFind findet sich in Katz, Demian/Nagy, Andrew: VuFind: Solr Power in the Library, In: Piedad Garrido (Hrsg.): Library Automation and OPAC 2.0: Information Access and Services in the 2.0 Landscape, Hershey 2013, S. 73-99. Die Versionsgeschichte findet sich im Changelog von VuFind, URL: <https://vufind.org/wiki/changelog> [07.03.2016].

Abbildung 1:
Oberfläche
der im Tuto-
rial erstellten
Demoversion
(nach Durch-
führung einer
leeren Suche)



Find ist unter der GNU General Public License (GPL) veröffentlicht.⁹

Die offizielle Dokumentation von VuFind ist frei verfügbar. Sie unterteilt sich in ein User Manual und ein Developer Manual. Das User Manual für VuFind 2.x enthält Anleitungen zur Installation, Konfiguration und Wartung des Systems.¹⁰ Der Stil des User Manuals setzt gewisse IT-Kenntnisse wie etwa die Bedienung der Terminal-Anwendung unter Linux-Betriebssystemen voraus. Die Installationsanleitung für VuFind liegt für die Linux-Distributionen Fedora und Ubuntu sowie Windows vor.¹¹

Zielgruppen

Mit dem Tutorial sollen sich insbesondere Bibliotheken ohne eigene IT-Abteilung auf den Weg zur eigenen Suche machen können und bei der Realisierung einer zeitgemäßen Suchumgebung auf Basis von Open Source-Software unterstützt werden. Die erste Zielgruppe des Tutorials sind daher Mitarbeiterinnen

und Mitarbeiter in Spezialbibliotheken und kleinen Öffentlichen Bibliotheken, welche sich mit der Einführung eines Discovery Systems auf Basis von VuFind näher beschäftigen möchten. Die zweite Zielgruppe des Tutorials sind Studierende der Bibliotheks- und Informationswissenschaft mit Interessenschwerpunkt im IT-Bereich. Alle Nutzer des Tutorials sollen sich einen Einblick in die Inbetriebnahme von VuFind verschaffen können, indem sie eine eigene lokal verfügbare VuFind-Instanz ohne Anbindung an ein Lokalsystem installieren, Metadaten in diese importieren und verschiedene Elemente des Systems konfigurieren. Abbildung 1 zeigt die fertige Demoversion. Es war bewusst nicht geplant, den kompletten Umfang der offiziellen Demoversion¹² abzubilden oder ähnlich wie Schmitt und Stehle¹³ oder Pohl¹⁴ in ihren Bachelorarbeiten einen prototypischen Bibliothekskatalog

9 URL: <https://vufind-org.github.io/vufind/about.html> [07.03.2016]. Die GPL findet sich unter der URL: <http://www.gnu.org/licenses/gpl-3.0.en.html> [07.03.2016].

10 URL: <https://vufind.org/wiki/start> [07.03.2016].

11 URL: <https://vufind.org/wiki/installation> [07.03.2016].

12 Die offizielle Demoversion findet sich unter der URL <http://vufind.org/demo/> [07.03.2016].

13 Schmitt, Jörg/ Stehle, Marcel: Der OPAC aus dem Baukasten. Realisierung eines Katalog 2.0 unter Einbeziehung der Community, überarb. Version 1.0, Hamburg 2010, URL: <http://edoc.sub.uni-hamburg.de/haw/volltexte/2011/1143/> [07.03.2016]

14 Pohl, Oliver: Konzept und prototypische Erstellung eines Informationssystems auf VuFind-Basis für die Bibliotheks- und Informationswissenschaft, Berlin 2012, URL: <http://nbn-resolving.de/urn:nbn:de:kobv:11-100207940> [07.03.2016].

zu erstellen. Die Nutzer des Tutorials sollen Anregungen und einen Überblick über die Möglichkeiten von VuFind erhalten. Was vor Ort daraus entsteht, hängt sehr von der Situation der jeweiligen Bibliothek ab und kann nicht vorausgesehen werden.

Das Tutorial

Wer das Tutorial nutzen möchte, sollte die Neugierde mitbringen, sich mit bisher möglicherweise wenig bekannten Dingen aus der IT-Welt näher zu beschäftigen. Diese wenig bekannten Dinge manifestieren sich in Begriffen wie „Virtuelle Maschine“, „VirtualBox“ oder „Ubuntu“. Was es damit auf sich hat, wird im nächsten Abschnitt kurz erläutert.

Es wird davon ausgegangen, dass der überwiegende Teil der Zielgruppen Windows als einziges Betriebssystem nutzt. Dennoch wird VuFind im Tutorial unter einem Linux-Betriebssystem installiert und benutzt, weil VuFind unter Linux entwickelt wurde und aus diesem Grund die Installation und der Betrieb einfacher sind.¹⁵ Linux als Betriebssystem gibt es in den verschiedensten Varianten, den sogenannten Linux-Distributionen. Die Wahl fiel hier auf Xubuntu, eine Abwandlung der Linux-Distribution Ubuntu. Die Benutzeroberfläche von Xubuntu ist näher am Look and Feel von Windows, womit der Einstieg in die Bedienung für die Nutzer des Tutorials erleichtert wird. Damit diese weiter mit ihrem gewohnten Betriebssystem arbeiten können, wird Xubuntu als sogenannte Virtuelle Maschine installiert. Eine Virtuelle Maschine bildet mit Hilfe einer Software einen Computer und dessen Betriebssystem nach. Eine gängige Software dafür ist die Freeware VirtualBox von Oracle.

Aufbau des Tutorials

Das Tutorial ist in diese zwölf Kapitel unterteilt:

1. Installation von VirtualBox
2. Installation von Linux
3. Installation von VuFind
4. Durchführung eines Testimports
5. Überblick zur Konfiguration
6. Allgemeine Einstellungen
7. Anpassung der Katalogoberfläche
8. Konfiguration der Facetten
9. Konfiguration von Kataloganreicherungen
10. Konfiguration der Suche, Teil I
11. Datenimport MARC21 MARCXML
12. Konfiguration der Suche, Teil II

Jedes Kapitel besteht aus einer Einleitung, den Arbeitsschritten, der Nennung des nächsten Kapitels

und einer Auflistung der verwendeten Quellen. Die Arbeitsschritte enthalten Arbeitsanweisungen in Form von Text und Screenshots. Die Screenshots sind dabei fester Bestandteil der Arbeitsanweisungen und sollen den Benutzer bei der Wiedererkennung von Bildschirmdialogen unterstützen. Weitere wichtige Elemente wie Terminalbefehle und Programmcode sind in einer anderen Schriftart gesetzt und lassen sich so leicht vom übrigen Text unterscheiden. Besonders wichtige Hinweise werden durch drei Ausrufezeichen hervorgehoben. Einen Ausschnitt des Tutorials zeigt Abbildung 2. Im Folgenden werden die einzelnen Kapitel und ihre jeweiligen Inhalte näher beschrieben.

Installation (Kapitel 1 bis 4)

Im Kapitel 1 wird die Version 4.3.28 von VirtualBox¹⁶ unter Windows 7 installiert. Anschließend wird im Kapitel 2 innerhalb von VirtualBox eine Virtuelle Maschine mit dem Betriebssystem Xubuntu 14.04.2 LTS¹⁷ erstellt. Bei der Installation der Virtuellen Maschine wird für die Gesamtgröße der virtuellen Festplatte ein Wert von 25 Gigabyte eingestellt, um genügend Festplattenkapazität für die spätere Durchführung des Datenimportes zur Verfügung zu haben. Für den Hauptspeicher der Virtuellen Maschine wird im Tutorial die maximal von VirtualBox empfohlene Größe eingestellt. Bei einem Rechner mit vier Gigabyte Hauptspeicher sind dies 1,5 Gigabyte und damit weniger als die von VuFind empfohlenen zwei Gigabyte¹⁸. Dies kann jedoch vernachlässigt werden, da die Performanz von VuFind bei der angestrebten Demoversion nicht entscheidend ist.

Die Voraussetzung für den Betrieb von VuFind ist das Vorhandensein weiterer Softwarekomponenten (Apache, Solr, PHP, MySQL, Java).¹⁹ Gemäß der Installationsanleitung von VuFind werden im Kapitel 3 zunächst diese Komponenten und anschließend VuFind in der Version 2.4.²⁰ installiert. Anschließend wird VuFind erstmalig gestartet und die erforderlichen ersten Konfigurationsschritte in der Autokonfiguration durchgeführt.²¹

Um die Funktionalität von VuFind zu testen, werden im Kapitel 4 wie von Schmitt und Stehle empfohlen Testdaten der Deutschen Nationalbibliothek (DNB) in

¹⁶ Die Version 4.3.28 erschien am 13.05.2015, siehe URL: <https://de.wikipedia.org/wiki/VirtualBox> [07.03.2016].

¹⁷ LTS steht für Long Term Support, siehe URL: <http://xubuntu.org/news/xubuntu-14-04-2-released/> [07.03.2016].

¹⁸ URL: <https://vufind.org/wiki/installation:requirements> [07.03.2016].

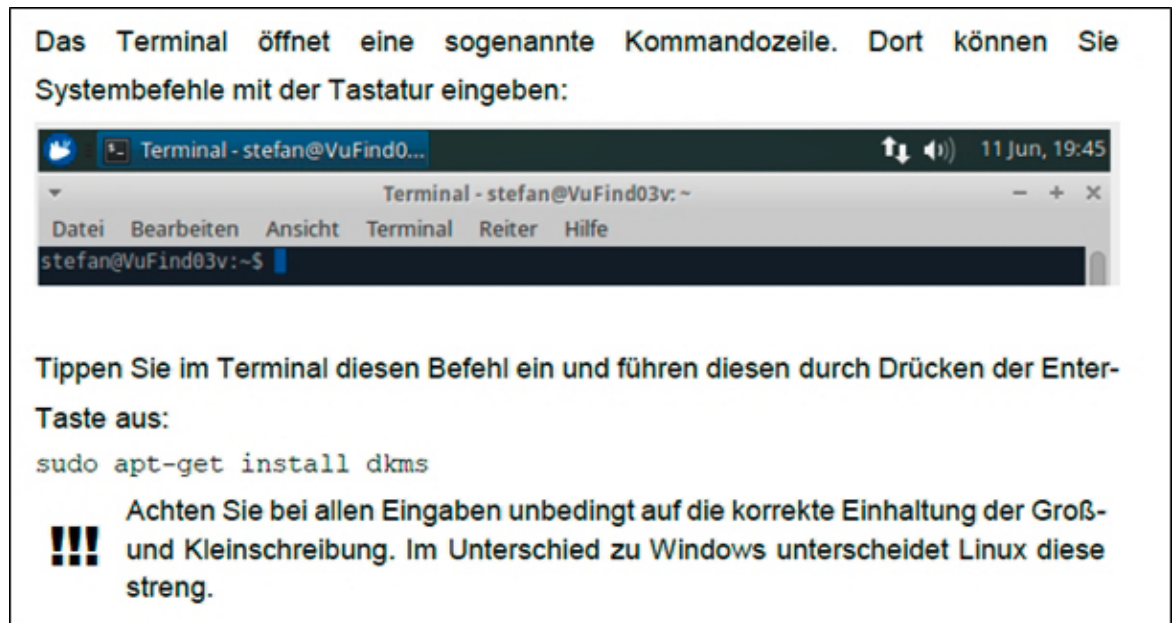
¹⁹ URL: <https://vufind.org/wiki/installation:notes> [07.03.2016].

²⁰ URL: <https://vufind.org/wiki/changelog> [07.03.2016].

²¹ URL: <https://vufind.org/wiki/installation:ubuntu> [07.03.2016].

¹⁵ URL: <https://vufind.org/wiki/installation:notes> [07.03.2016].

Abbildung 2:
Ausschnitt aus
dem Kapitel
„Installation von
Linux“



den Index importiert.²² Eingesetzt werden Testdaten aus der DNB-Reihe A. Diese enthalten mehr Datensätze (2.519 Stück) als die Testdaten der Reihe B (595 Stück) und sind daher besser für die Durchführung der im Folgenden beschriebenen Konfigurationsschritte geeignet.

Konfiguration erster Teil (Kapitel 5 bis 10)

In die Kapitel zur Konfiguration wurden die Funktionalitäten aufgenommen, für die eine Nutzung und Anpassung durch die Zielgruppen angenommen wurde. Das Kapitel 5 bietet einen ersten Überblick und dient der Vorbereitung. Im Kapitel 6 wird die zentrale Konfigurationsdatei von VuFind (config.ini) vorgestellt und einige allgemeine Einstellungen wie etwa Sprache durchgeführt. Außerdem wird die Anpassung der Übersetzungen für Elemente von VuFind exemplarisch am Beispiel der Merkliste gezeigt.

Das Kapitel 7 führt in die Gestaltung der Katalogoberfläche von VuFind durch sogenannte Themes ein. Durch den exemplarischen Austausch des Logos sowie die Anpassung des Footers der lokalen Webseite von VuFind wird die Anpassung eines Themes demonstriert.

Im Kapitel 8 wird die Konfigurationsdatei facets.ini vorgestellt, die angezeigten Facetten angepasst und die Übersetzung der Inhalte einer Facette beispielhaft anhand der Facette language vorgestellt.

Im Kapitel 9 werden die unterschiedlichen von VuFind beherrschten Formen der Kataloganreicherung behandelt. VuFind kennt Kataloganreicherungen in den Bereichen Titelbilder, Rezensionen, Auszüge, Buchvorschau und Autorenbiographien. Diese können von

unterschiedlichen Anbietern eingebunden werden.²³ Der Bereich der Auszüge (Excerpts) wird nicht thematisiert, da diese nur vom kostenpflichtigen Anbieter Syndetic Solutions angeboten werden.

Im Kapitel 10 werden die Sucheinstellungen und die Definition sogenannter Suchtypen vorgestellt. Konkrete Einstellungen werden nicht behandelt, da diese vom Anwendungskontext des jeweiligen Nutzers abhängig sind. Das Kapitel schließt mit einer Anpassung der Stoppwortliste von VuFind. Diese Anpassung erfordert die Reindexierung aller Datensätze im Index. Um dies innerhalb des Tutorials für eine größere Datenmenge zu vermeiden, wurde bis zum Kapitel 10 des Tutorials nur mit den Daten aus dem Testimport gearbeitet. So müssen bei der Nutzung des Tutorials nur die Testdaten reindexiert werden.

Datenimport (Kapitel 11)

VuFind unterstützt den Import von Metadaten in den folgenden Formaten:

- Daten im binären MARC-Format nach ISO2709,
- Daten im MARCXML-Format und
- Daten in beliebigen XML-Formaten.²⁴

Die Bereitstellung der zu importierenden Daten kann manuell oder durch Harvesting mittels OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) erfolgen. Zur Durchführung des Importvorgangs und auch für das Harvesting stellt VuFind geeignete PHP-Skripte zur Verfügung. Für die Ver-

²² Schmitt/Stehle 2010 (wie Anm. 12), S. 48 f.

²³ Einen Überblick bieten URL:

https://vufind.org/wiki/configuration:book_previews [07.03.2016] und URL: https://vufind.org/wiki/configuration:external_content [07.03.2016].

²⁴ URL: <https://vufind.org/wiki/indexing> [07.03.2016].

wendung der Skripte müssen Konfigurationsdateien angepasst oder erstellt werden.

Der manuelle Import von Daten im binären MARC-Format wurde bereits im Kapitel 4 mit dem Import von Testdaten aus der DNB-Reihe A eingeübt. Im Kapitel 11 des Tutorials lernt der Nutzer nun den manuellen Datenimport von Daten im MARCXML-Format kennen. Dazu werden 50.000 Datensätze aus den als Open Data verfügbaren Daten des TIBKAT importiert.²⁵ Warum werden Open Data genutzt? Im Hinblick auf die Verfügbarkeit von Metadaten sind die beiden Zielgruppen sehr unterschiedlich. Die Zielgruppe der Bibliotheksmitarbeiter könnte alternativ die Katalogdaten der eigenen Bibliothek nutzen. Der Zielgruppe der Studierenden steht diese Möglichkeit vermutlich jedoch nicht zur Verfügung. Ähnliches gilt für den Verfasser bei der Erstellung des Tutorials. Daher sind als Open Data frei verfügbare Metadaten die beste Wahl. Die genutzten Daten aus dem TIBKAT werden in kleineren Einheiten angeboten. Dies bietet insbesondere den Vorteil, dass eine für die angestrebte Demoversion ausreichende Datenmenge zügig importiert werden kann.

Konfiguration zweiter Teil (Kapitel 12)

Das zwölfte und letzte Kapitel des Tutorials beschäftigt sich mit der Definition von Synonymen für Suchterme und der Einstellung „Minimum Should Match“. Mit letzterer kann definiert werden, in welchem Umfang eingegebene Suchterme mit einem Datensatz im Index übereinstimmen müssen.

Ausblick

Denkbar ist die Erweiterung des Tutorials um eine Version für Fortgeschrittene. In dieser könnten insbesondere die folgenden Arbeitsschritte zusätzlich aufgenommen werden:

- Einrichtung einer Webseite;
- Anbindung eines Lokalsystems;
- Datenimport über OAI-PMH (Harvesting);
- Datenimport beliebiger XML-Formate;
- Indexieren von Volltexten und Webseiten;
- Einstellungen für die Suchsyntax, das Relevanzranking, die ähnlichen Einträge und das Boosting²⁶;
- weitere Möglichkeiten zur fehlertoleranten Gestaltung der Suche;
- Rechtschreibkorrektur und Stemming²⁷.

²⁵ Verwendet wurde die Datei „tibkat_full_20150305_1_of_4.tgz“, URL: <http://www.tib-hannover.de/de/dienstleistungen/opendata/download/4/> [04.08.2015].

²⁶ Schmitt/Stehle 2010 (wie Anm. 12), S. 76 ff.

²⁷ Schmitt/Stehle 2010 (wie Anm. 12), S. 85 ff.

Mitarbeit erwünscht

Da neben der freien Nutzung auch die Mitarbeit bei der Weiterentwicklung der Inhalte des Tutorials erwünscht sind, wurde die finale Version des Tutorials unter eine CC-BY-SA-Lizenz gestellt. Diese erlaubt jedwede Verbreitung, Verarbeitung und Verwendung des Materials bei Namensnennung des Autors sowie eine Weitergabe unter gleichen Bedingungen.* Letzteres ist wichtig, damit auch potenzielle Weiterentwicklungen der Inhalte ebenfalls zur freien Nutzung zur Verfügung stehen.

* Eine alltagssprachliche Zusammenfassung findet sich unter der URL: <http://creativecommons.org/licenses/by-sa/4.0/deed.de> [07.03.2016].

Über das Tutorial hinaus gibt es weiteren Forschungsbedarf insbesondere die Langzeitverfügbarkeit von Open Data-Angeboten betreffend. Existiert diese und wenn ja, wer stellt das sicher? Ohne das Open Data-Angebot der TIB wäre der Datenimport im Tutorial nicht mehr durchführbar. Ein Schritt in Richtung dieser Langzeitverfügbarkeit ist die Plattform D:SWARM, eine auf Open Source basierende Plattform für das Management bibliothekarischer Daten. Mit einem Blick auf die Kernaufgaben von Bibliotheken und die zu deren Erfüllung genutzten Werkzeuge plädiert Mittelbach für die Beteiligung von Bibliotheken an der Schaffung offener Inhalte, Daten und Infrastrukturen. Besonders kleinere Einrichtungen ohne eigene IT-Kapazitäten profitieren von der zunehmenden Verfügbarkeit offener Infrastrukturen.²⁸ Ein Teil dieser offenen Infrastrukturen sind Open Source Discovery Systeme wie VuFind. Damit besteht für Bibliotheken die Chance, wieder stärker bei der Erfüllung einer ihrer Kernaufgaben, dem Nachweis von Informationen, wahrgenommen zu werden. ■



Stefan Niesner

Dipl.-Bibl. (FH)
Privatanschrift:
Zollstockgürtel 19
50969 Köln
st.niesner@gmail.com