



Whitepaper zur Literaturverwaltung

Liabulo

„Ausblick“
von Maxim Bauer

erstellt im Rahmen der Projektgruppe
„**Virtuelle Organisation - Open Source**“

unter der Leitung von
Prof. Uwe Schneidewind und Hendrik Eggers (Lehrbeauftragter)

Inhaltsverzeichnis

1	Ausblick	1
1.1	Einleitung	1
1.2	Anlegen von Duplikaten	2
1.3	Überprüfung auf Duplikate	4
1.4	Nicht-sortier-Kennzeichen	6

Kapitel 1

Ausblick

1.1 Einleitung

Das Softwareprodukt „Liabolo“ unterstützt inzwischen sehr viele Funktionen, die für eine Literaturverwaltungssoftware von Bedeutung sind. Diese Funktionen sind ausführlich im Whitepaper „Gui-Architektur“ beschrieben. Es ist jedoch möglich und auch wünschenswert, dass noch mehr Funktionen, die „Liabolo“ noch besser und komfortabler machen, in dieses Softwareprodukt eingebunden werden. Einige der Funktionen können „Liabolo“ noch mächtiger im Bezug auf die Literaturverwaltungssoftware machen, andere könnten z.B. die Arbeit mit „Liabolo“ wesentlich erleichtern. Ein Softwareprodukt, welches die Arbeit der Nutzer beschleunigt und erleichtert, findet bei vielen Nutzern Verwendung und Unterstützung. Einige mögliche Funktionen, die für Liabolo wichtig werden können, werden in diesem Whitepaper genauer erläutert.

Die Funktionen, die in „Liabolo“ eingebunden werden können und hier beschrieben werden, sind das „Anlegen von Duplikaten“, „Überprüfung auf Duplikate“ und die „Unterstützung von Nicht-sortier-Kennzeichen“. Was beinhalten aber diese Funktionen im einzelnen und wie können sie realisiert werden? Die Beantwortung dieser Fragen wird im wesentlichen den Inhalt dieses Whitepapers bilden.

1.2 Anlegen von Duplikaten

Das Anlegen von Duplikaten beinhaltet in sich die Möglichkeit zum schnellen Anlegen von gleichen Datensätzen. Dies ist besonders dann sinnvoll, wenn mehrere gleiche Objekte existieren, die erfasst werden sollen¹. Zur Zeit ist dies nur als „Anlegen von neuen Datensätzen“ möglich. Dies ist jedoch beim Anlegen mehrerer gleicher Objekte sehr ineffizient und für den Benutzer langweilig, weil immer die gleichen Daten eingegeben werden müssen. Mit Hilfe der Funktion „Anlegen von Duplikaten“ werden die gleichen Daten bereits in das Eingabe-Formular mit übernommen, was viel Zeit und Routinearbeit erspart. Wie genau diese Funktion in „Liabolo“ eingebunden werden kann, wird weiter unten genauer beschrieben.

Umsetzung der Funktion in „Liabolo“

Um diese Funktion überhaupt aufrufen zu können, muss zunächst mindestens ein bereits definiertes Metadaten-Objekt² vorhanden sein. Der Aufruf der Funktion kann aus dem Ergebnis-Fenster entweder durch den entsprechenden Button oder das Kontext-Menü erfolgen, wobei als Vorlage immer das ausgewählte Metadaten-Objekt verwendet wird.

Wenn die Funktion gestartet wird, erscheint als erstes das Editier-Formular mit bereits ausgefüllten Eingabefeldern. An dieser Stelle wären dann generell zwei Varianten möglich:

Die erste Möglichkeit wäre, dass die Eingabefelder alle nicht editierbar sind, was für das Anlegen von Duplikaten durchaus Sinn macht. Denn bei einer Änderung der Daten des Metadaten-Objekts wäre das kein Duplizieren mehr (Die einzige Änderung, die zulässig wäre ist die Änderung des Standortes). Wenn der Benutzer anschließend die Eingaben bestätigt, wird das Metadaten-Objekt mit vorhandenen Daten und einer neuen Signatur in der Datenbank gespeichert. Zur Zeit gibt es in „Liabolo“ noch keine Möglichkeit zur Bildung einer für die Duplikate passenden Signatur (z.B. die Übernahme der Signatur von dem zu duplizierenden Objekt mit dem Hinzufügen eines Buchstaben am Ende der Signatur. Wenn der Buchstabe bereits vorhanden ist, wird dieser abgeschnitten und durch den nächstfolgenden ersetzt.), wird eine komplett neue Signatur generiert.

¹In einer Bibliothek existieren in der Regel öfters mehrere Exemplare der gleichen Bücher, Artikeln, etc.

²Metadaten-Objekt : Ein in „Liabolo“ definiertes Metadaten-Objekt wie z.B. Buch, Artikel, Zeitschrift und etc.

Als Ergebnis wird das duplizierte Metadaten-Objekt mit der modifizierten Signatur und eventuell neuen Standort in der Datenbank abgelegt.

Die zweite Variante wird von der ersten nur durch die Möglichkeit des Editierens im Editier-Formular aller möglichen Eingabefelder des Metadaten-Objekts unterschieden. Dies macht Sinn, weil damit nicht nur genaue Duplikate erfasst werden können, sondern auch Objekte, die sich geringfügig von einander unterscheiden (z.B. andere Auflage). In diesen Fall sollte aber genau unterschieden werden, welche Signatur generiert werden muss. Bei Duplikaten sollte der Signaturalgorithmus nur den Buchstaben an das Ende der Signatur anhängen, bei allen anderen sollte eine neue Signatur generiert werden.

Als Ergebnis wird entweder ein modifiziertes Metadaten-Objekt mit einer neuen Signatur in der Datenbank abgelegt oder ein Duplikat mit der modifizierten Signatur und eventuell neuen Standort.

Für welche der beiden Varianten die Entwickler sich auch entscheiden, es sollte bedacht werden, dass die erste Variante zwar ziemlich einfach zu implementieren wäre, die zweite jedoch mehr Raum für weitere mögliche Erweiterungen bietet.

1.3 Überprüfung auf Duplikate

Das Überprüfen auf Duplikate beim Anlegen eines neuen Metadaten-Objekts bietet die Möglichkeit zum schnellen finden der gleichen Metadaten-Objekte. Wenn ein neues Objekt angelegt wird, wird gleichzeitig nach allen Objekten in derselben Bibliothek gesucht, die die gleiche ISBN-Nummer aufweisen. Danach kann der Benutzer entscheiden, ob die in der Bibliothek enthaltene Objekte Duplikate des gerade neu angelegten Objekts sind oder nicht. Abschließend wird der Benutzer vor die Wahl gestellt ob er das neu angelegte Metadaten-Objekt trotzdem in der Datenbank speichern möchte, oder ob er dieses Objekt verwerfen möchte.

Diese Funktion ist nur dann sinnvoll, wenn mehrere Nutzer die gleiche Bibliothek pflegen, was in „Liabolo“ häufig der Fall sein kann. Durch diese Mehrbenutzerfähigkeit können unter Umständen viele redundante Daten entstehen, die mit der obenbeschriebenen Funktion minimiert werden können. Wichtig ist diese Funktion auch bei der Signaturbildung (wenn der Signaturalgorithmus erstmals angepasst wurde), denn somit bekommen die gleiche Metadaten-Objekte nicht eine komplett neue Signatur, sondern eine modifizierte Signatur, die dem Umgang mit den Duplikaten erleichtert.

Umsetzung der Funktion in „Liabolo“

Umgesetzt werden kann diese Funktion auf folgende Art und Weise: Zunächst wird das Formular zum Anlegen der neuen Metadaten-Objekte aufgerufen. Wenn der Benutzer seine Eingaben getätigt hat und bestätigt hat, wird eine Suche in der Datenbank angestoßen, welche alle Metadaten-Objekte mit der ISBN-Nummer des neu eingegebenen Objektes in der Datenbank findet³. Es wird ausschließlich nur nach der gleichen ISBN-Nummer gesucht. Danach bekommt der Benutzer ein Ergebnis-Formular eingeblendet, wo die gesamten Ergebnisse der Suche aufgelistet sind.

An dieser Stelle hat der Benutzer die Möglichkeit die aufgelisteten Ergebnisse durchzusehen und zu entscheiden, was mit dem von ihm eingegebenen Metadaten-Objekt geschehen soll. Soll das Objekt trotzdem in die Datenbank aufgenommen werden, weil dieses Metadaten-Objekt-Duplikat auf ein anderes Objekt verweist, oder nicht, weil die Metadaten in Wirklichkeit zum einen und denselben Objekt gehören.

³Bei Objekten, die keine ISBN-Nummer (oder ISSN-Nummer) haben, wird die Funktion nicht aufgerufen.

Je nachdem, wie der Benutzer sich entscheidet, wird das Metadaten-Objekt in die Datenbank aufgenommen oder nicht. Somit kann erreicht werden, dass die Datenbank sauber in Bezug auf unerwünschte redundante Informationen bleibt.

1.4 Nicht-sortier-Kennzeichen

Nicht-sortier-Kennzeichen sind Zeichen mit deren Hilfe bestimmte Wörter markiert werden können, um von der Sortierung ausgeschlossen zu werden. Als ein Nicht-sortier-Kennzeichen wird meistens ein Zeichen gewählt, welches in der Regel nicht in der Zeichenkette, die sortiert werden soll, vorkommt. Die Markierung der Wörter erfolgt durch das setzen des Nicht-sortier-Kennzeichens vor dem jeweiligen Wort. Somit wird dieses Wort von der Sortierung ausgeschlossen. Die Nicht-sortier-Kennzeichen können überall dort verwendet werden, wo bestimmte Zeichenketten sortiert werden können. Die Eingabe von Nicht-sortier-Kennzeichen erfolgt ausschließlich durch den Benutzer.

Das Einsetzen der Nicht-sortier-Kennzeichen in „Liabolo“ würde die Arbeit damit noch angenehmer gestalten. Im Moment werden in „Liabolo“ noch alle in den Zeichenketten enthaltenen Wörter bei der Sortierung berücksichtigt. Dies führt dazu, dass viele Datensätze, welche z.B. als erstes Zeichen ein Artikel enthalten, anders sortiert werden, als diejenigen, die keinen Artikel haben. Damit wird eine deutliche Verfälschung des Sortierergebnisses sichtbar. Um diese Verfälschung nicht zu erlauben, sollten Nicht-sortier-Kennzeichen eingeführt werden, die dann vom Sortieralgorithmus unterstützt werden müssen.

Alternativ zu Einführung der Nicht-sortier-Kennzeichen kann auch eine andere Möglichkeit in Betracht gezogen werden. Diese beinhaltet in sich die Anpassung des Sortieralgorithmus um Zusatzlisten mit Angaben über Artikel und andere Zusätze, welche dann vom Sortieralgorithmus nicht berücksichtigt werden. Diese Möglichkeit verlagert die Problematik des Sortierens vom Benutzer auf den Entwickler. Denn wenn beim Einsatz der Nicht-sortier-Kennzeichen die Benutzer selbst entscheiden konnten, welche Zeichen von der Sortierung ausgeschlossen werden sollen, müssen die zusätzliche Listen vom Entwickler gepflegt werden.

Umsetzung der Funktion in „Liabolo“

In „Liabolo“ kann die Unterstützung von Nicht-sortier-Kennzeichen auf folgende Weise erfolgen: Zuerst muss ein bestimmtes Zeichen ausgewählt werden, welchen dann die Funktion des Nicht-sortier-Kennzeichen übernimmt. Bei der Wahl des Zeichen ist zu berücksichtigen, dass dieses nicht in den zu sortierenden Zeichenketten vorkommt. Nachdem dies geschehen ist, müssen noch zwei weitere Elemente in „Liabolo“ geändert bzw. angepasst werden.

- Erstens muss der Sortieralgorithmus so angepasst werden, dass er die Nicht-sortier-Kennzeichen mit den dazugehörigen Wörtern aus den zu sortierenden Zeichenketten heraus filtert. Damit ist gemeint, dass diese Wörter nicht entfernt werden, sondern nur bei der Sortierung nicht berücksichtigt werden. Danach erfolgt die Sortierung und die Rückgabe der sortierten Zeichenketten.
- Zweitens müssen einige GUI-Formulare angepasst werden. Die Formulare zur Eingabe oder zum Editieren der Metadaten-Objekte müssen die Nicht-sortier-Kennzeichen anzeigen. Damit wird die Möglichkeit zum Eingeben und Ändern der Nicht-sortier-Kennzeichen gegeben. Beim Anzeigen der Metadaten-Objekte sollen die Nicht-sortier-Kennzeichen aber ausgeblendet werden, um das Lesen von Ergebnissen nicht zu behindern.

Im ersten Fall sollte im Sortieralgorithmus ein Filter eingebaut werden, welcher die mit Nicht-sortier-Kennzeichen markierten Wörtern bei der Sortierung ignoriert. Im zweiten Fall muss nur das Ergebnis-Formular geändert werden. Beim Anzeigen der Ergebnisse werden vorher aus allen Metadaten-Objekten die Nicht-sortier-Kennzeichen entfernt. Bei Anlegen oder Editieren ist keine Änderung notwendig, da die Nicht-sortier-Kennzeichen in den Metadaten gespeichert werden.

Alternativ zu Umsetzung der Nicht-sortier-Kennzeichen kann die zweite Möglichkeit viel einfacher umgesetzt werden. In diesem Fall muss nur der Sortieralgorithmus um eine Liste erweitert werden, wo alle Zeichen stehen, die vom Sortieralgorithmus ignoriert werden. Diese Liste kann vorerst nur vom Entwickler geändert werden. Um die Pflege der Liste auf den Benutzer zu übertragen, ist ein neues Formular-Fenster notwendig, wo die ganzen Nicht-sortier-Wörter enthalten sind, neu angelegt oder gelöscht werden können. Somit wäre eine weitere Möglichkeit zur Individualisierung gegeben.

Index

Überprüfen auf Duplikate, 4
Überprüfung auf Duplikate, 1

Anlegen von Duplikaten, 1, 2

Duplikate, 2, 4

Mehrbenutzerfähigkeit, 4

Metadaten-Objekt, 2–4

Metadaten-Objekte, 7

Nicht-sortier-Kennzeichen, 1, 6, 7

Signaturalgorithmus, 3, 4

Sortieralgorithmus, 7

sortierte Zeichenketten, 7

Zusatzlisten, 6