

Datenformat zum Import von CSV-Dateien

(Stand: FB 2011)

Allgemeines

Zur Vereinfachung der Dateneingabe für die Deutsche Bibliotheksstatistik (DBS) haben die **Fachstellen** die Möglichkeit, die Daten der von ihnen betreuten Bibliotheken direkt zu importieren. Die Dateien im **CSV-Format** (*Comma Separated Values*)¹ werden dabei über ein Web-Formular hochgeladen, geprüft und nach erfolgreicher Kontrolle in die Datenbank der DBS eingespielt. Dafür muss für jede Fachstelle (genauer: jede Fachstellen-ID) jeweils genau eine Import-Datei geliefert werden. Unmittelbar danach kann auf die importierten Daten zugegriffen werden.

Dieser Weg zum Import der DBS-Daten richtet sich ausschließlich an Fachstellen!!!

Format

Grundsätzlich erzeugt **Microsoft Excel** Export-Dateien in einem für den DBS-Import gültigen Format². Es sind dabei folgende Hinweise zu beachten:

1. Erwartet werden Dateien im **ASCII-Format**, die Zeichensatz-Kodierung sollte „Western ISO-8859-1“ entsprechen (also kein Unicode / UTF8 o.ä.).
2. Jede Zeile enthält mehrere Spalten, die durch **Semikolon** (»;«) voneinander getrennt sind. Die Werte der einzelnen Spalten können durch doppelte oder einfache Anführungszeichen (»"«,»'«) eingeschlossen sein, notwendig ist dies jedoch nicht.
3. Die erste Zeile kann **Überschriften** enthalten, dies ist aber nicht erforderlich. Leerzeilen werden beim Import übersprungen.
4. Für den Import der Daten sind in jeder Zeile (mindestens) **99 Spalten** notwendig, das heißt eine Spalte für die DBS-ID und 98 Spalten für die Daten des Grundfragebogens („Fragebogen für Öffentliche Bibliotheken“). Zum Format der Spalten:
 - a. In der ersten Spalte steht die **DBS-Identifikationsnummer** (DBS-ID), die jeweils aus genau fünf Zeichen besteht, zwei Buchstaben gefolgt von drei Ziffern (z.B. „AB123“).
 - b. Die übrigen Spalten enthalten die **Antworten** des Fragebogens. Dabei gibt es aktuell drei verschiedene Datentypen:
 - i. **Numerisch, ohne Nachkommastelle**: Hier wird jeweils eine (maximal 11-stellige) ganze Zahl erwartet, Nachkommastellen werden nicht berücksichtigt, Tausenderpunkte sollten nicht enthalten sein.

¹ vgl. Beispiel auf Seite 3

² unter Office 2000: Menüpunkt „Datei“: „Speichern unter...“, dann im Dialog die Auswahlliste „Dateityp“: „CSV (Trennzeichen getrennt) (*.csv)“ wählen

- ii. **Numerisch, zwei Nachkommastellen:** Wie unter i., jedoch können bis zu zwei Nachkommastellen mit Komma getrennt angegeben werden (weitere Stellen werden ignoriert).
 - iii. **Auswahlliste** (alphanumerisch): Dabei kann im Fragebogen per „Drop-Down-Box“ jeweils eine von mehreren Alternativen ausgewählt werden (meist „ja / nein“-Fragen). Hier sollte die Datei entweder den **Klartext** (exakt wie im Fragebogen, keine Abkürzung) oder die **Ziffer** der Auswahl (in der Reihenfolge des Fragebogens beginnend mit 1 enthalten).³
- c. Wichtig ist der Unterschied zwischen den Einträgen „**N**“ und „**0**“ (**Null**): Ein „N“ besagt „vorhanden, aber nicht erfasst“ (keine Zählung möglich, Wert unbekannt), eine Null hingegen bedeutet „der Wert beträgt Null“. Bei Auswahllisten kann das Feld auch leer bleiben.
- d. Die **Zahl der Spalten** ist variabel, je nachdem, welche Teile des Fragenkanons importiert werden sollen. Es gibt genau vier Möglichkeiten:
- i. **99 Spalten:** Damit wird ausschließlich der Hauptteil des Fragebogens importiert (Fragen 1-102).
 - ii. **110 Spalten:** Dann wird der Hauptteil und der Abschnitt für Patientenbibliotheken (Fragengruppe 200, Fragen 200-210) importiert.
 - iii. **118 Spalten:** Importiert wird wie unter (ii.), hinzu kommt der Abschnitt für Fahrbibliotheken (Fragengruppe 300, Fragen 300-307).
 - iv. **140 Spalten:** Hier wird wie unter (iii.) importiert, hinzu kommt der Abschnitt für Musikbibliotheken (Fragengruppe 400, Fragen 400-421).

Wichtig ist, dass jede Zeile der Import-Datei jeweils genau die oben angegebene Zahl von Spalten enthält.

5. **Bereits existierende Daten**, die von Bibliotheken evtl. schon von Hand eingegeben wurden, werden beim Import auf jeden Fall überschrieben. Dies gilt jedoch nur für die Bereiche, die im Import enthalten sind. Bei importierten 99 Spalten z.B. bleiben die Fragen 200-210 etc. unberührt.

Prüfungen

Beim Import werden verschiedene Prüfungen durchgeführt, die das Format und teilweise auch den Inhalt der Daten betreffen.

1. Es wird geprüft, ob die **DBS-ID** in der ersten Spalte das korrekte Format hat („AB123“), ob sie in der DBS existiert und ob die beim Import eingestellte Fachstelle überhaupt für diese Bibliothek zuständig ist (dies ist den DBS-Adressdaten zu entnehmen).

³ Beispiel: Bei einem „ja/nein“-Feld kann diese Spalte beim Import also entweder leer bleiben (= „nicht erfasst“) oder die Werte „ja“, „nein“ bzw. „1“ (= „ja“), „2“ (= „nein“) oder „0“ (= „nicht erfasst“ bzw. „keine Antwort“) enthalten.

2. **Numerische Felder** werden auf das korrekte Format hin untersucht (Zahlen mit/ohne Nachkommastellen). Bei Feldern mit **Auswahllisten** wird geprüft, ob der Inhalt tatsächlich einer möglichen Option entspricht.

Beispiel

Hier nun ein Beispiel für den Aufbau einer Import-Datei (in tabellarischer Form etwa wie in Excel):

Spalte1	2	3	4	...	62	63	64	...	98
DBS-ID	primäre Nutzer	Zahl Bibl.	Zweigstellen		Einnahmen	Haushaltsführung	Ben.-geb.		Sonst. vertr. Dienstl.
Frage-Nr. (*)	1	2	3	...	63	64	65	...	102
AB123	13538	1	0	...	12900	kameralistisch	nein	...	3
CD456	9877	1	0	...	5890	budgetiert		...	0
EF789	25987	2	1	...	23980	teil-budgetiert	ja	...	2

(*) Die Frage-Nr. dient hier nur als Orientierungshilfe und ist nicht Bestandteil einer Import-Datei

In der Import-Datei im **CSV-Format** würde dies dann so aussehen:

AB123;13538;1;0; ... ;12900;kameralistisch;nein; ... ;3
CD456;9877;1;0; ... ;5890;budgetiert;; ... ;0
EF789;25987;2;1; ... ;23980;teil-budgetiert;ja; ... ;2

Hinweis für den Export aus Microsoft Excel

Beim Export aus Excel heraus ist zu beachten, dass **ausgeblendete Spalten** (oder Zeilen) immer komplett mit exportiert werden. Hier ist die folgende Technik zu empfehlen: Markieren Sie in der Excel-Tabelle alle Zellen, die zum Import gehören sollen, kopieren Sie diese Zellen und fügen Sie diese auf einem neuen Blatt über „Inhalte einfügen...“ (im Menü „Bearbeiten“) ein, wobei Sie im angezeigten Dialog „Einfügen: Werte“ auswählen sollten. Auf diesem Weg können Sie in der ursprünglichen Tabelle auch berechnete Zellen (Formeln) beibehalten.

Webseite: <http://www.bibliotheksstatistik.de>

Kontakt per E-Mail: dbs@hbz-nrw.de

Stand: 09.09.2011