

Die technische Umsetzung des „Pfeffer-Verfahrens“ im B3Kat

Dr. Stefan Brecheisen

Agenda

- ▶ Ausgangspunkt
- ▶ Einschränkungen
- ▶ Vorbereitung
- ▶ Durchführung
- ▶ Ergebnis
- ▶ Ausblick

Ausgangspunkt

- ▶ Verbundkatalog enthält viele ähnliche Aufnahmen unterschiedlicher Erschließungstiefe
- ▶ Bündelung auf Werkebene
 - ▶ gleicher Titel
 - ▶ gleicher Autor / Herausgeber
- ▶ Übernahme von RSWK und RVK aus Quellsatz in Zielsätze des gleichen Bündels
- ▶ MARC-Exporte der Verbände für CultureGraph
- ▶ Verbundübergreifende Liste potentieller Quell- und Zielaufnahmen erstellt von Prof. Pfeffer

Einschränkungen

- ▶ **Zunächst nur Anreicherung innerhalb des B3Kat**
 - ▶ Quellaufnahmen aus B3Kat
- ▶ **Keine Anreicherung von Bänden**
 - ▶ Im MARC-Format keine Unterscheidung von H- und U-Sätzen
 - ▶ Bündelung aller Bände eines mehrbändigen Werks
- ▶ **Keine Übernahme von Notationen aus dem Bereich LP – LY**
 - ▶ Fehlklassifikation von Büchern, Noten, Ton- und Bildträgern vermeiden

Vorbereitung

- ▶ Perl- und Shell-Skripte unter Unix
 - ▶ Systemnummern von Quell- und Zielsätzen bestimmen und austauschen

Liste

```
BV026590780 HEB186598564 SWB264883829 BV022475468
```

```
015682888 001 L $$aBV022475468
015682888 700g L $$aNN 1400
015682888 700g L $$aFV 2410
```

Quellsatz

```
025952030 001 L $$aBV026590780
025952030 700g L $$aNN 1400
025952030 700g L $$aFV 2410
```

Zielsatz

Durchführung

- ▶ Anreicherungen in Datenbank laden
 - ▶ Am 7.2.2014, insgesamt 1.312.749 Datensätze
 - ▶ Dauer: ca. 24 Stunden („geht übers Wochenende“)
- ▶ Indexierung
 - ▶ Hintergrundprozess mit niedriger Priorität
 - ▶ Dauer: ca. 3 Wochen („läuft nebenbei“)
- ▶ Versorgung der Lokalsysteme
 - ▶ BV-Nummernlisten zur Nachversorgung
 - ▶ Hoher Arbeitsaufwand
 - ▶ Dauer: 2 Monate (abgeschlossen 6.4.2014)
(„mühsam“)

Ergebnis

- ▶ Bis jetzt größte Anreicherungsaktion im B3Kat
 - ▶ insgesamt 1.312.749 Titelaufnahmen,
 - ▶ davon 697.464 mit RSWK-Schlagwörtern
 - ▶ und 945.555 mit RVK-Notationen
- ▶ Deutliche Verbreiterung der inhaltlichen Erschließung
- ▶ Anteil RVK-notierter Titel steigt von 28,9% auf 33,4%

Anzahl Titel nach
Lokalsystem (Auszug)

286.211 Bayerische Staatsbibliothek
203.705 FU Berlin
184.312 UB München
159.297 UB Erlangen
153.864 HU Berlin
143.706 TU Berlin
103.905 Kubikat
71.066 UB Regensburg

Ausblick

- ▶ Verbundübergreifende Anreicherung
- ▶ Verbundkataloge als Open Data verfügbar
- ▶ Integration in CultureGraph
- ▶ Abfrage der Bündel aus CultureGraph

Ende

Vielen Dank!
Fragen?