Arbeitsgruppe Sacherschließung (AG SE) 10. Sitzung am 15.06.2021

Teilnehmende (inkl. Gäste)

- Hertel, Sabine (Gast; bis 15 Uhr)
- Holbach, Werner (Vorsitz)
- Dr. Kobold, Sonja
- Dr. Magen, Antonie (Protokoll)
- Merker, Kathrin (Gast)
- Christian Rüter
- Dr. Schweikl, Gabriele
- Thiessen, Peter (Ständiger Gast)
- Dr. Völkl, Martin
- Wolf-Dahm, Barbara
- Dr. Ziegler-Stryczek, Birgit

entschuldigt:
- Franke-Maier, Michael

Zeit & Ort

Datum: Di., 15.06.2021
Zeit: 13.30-17.00 Uhr
Ort: Videokonferenz

TOP 1: Begrüßung, Formalia, Protokollführung, Protokoll der letzten Sitzung, Tagesordnung

Das Protokoll der letzten Sitzung, das bereits auf den Seiten der AG SE des BVB veröffentlicht ist, wird offiziell bestätigt.

TOP 2: Nachnutzung der Sacherschließung der Culturegraph-Bündel

Im Vorfeld der Sitzung hat Herr Thiessen Dateien mit Testumsetzungen für Phase 1 (vgl. die Protokolle der 7. und 9. Sitzung der AG SE am 29.10.20 und 15.4.21) der Culturegraph-Anreicherung zur Verfügung gestellt. Auf dieser Grundlage stellen er und Herr Völkl den Stand der Dinge von Phase 1 folgendermaßen dar: Bereits umgesetzt hat Herr Thiessen, dass

1. Permutationen innerhalb der für einen Satz jeweils anzureichernden Schlagwortfolgen ausgeschlossen sind und

2. Schlagwortfolgen, die als Teilfolgen bereits innerhalb der für einen Satz jeweils anzureichernden Schlagwortfolgen enthalten sind, ebenfalls ausgeschlossen sind.

Gesprächsbedarf besteht hingegen noch in der Frage, wie mit Datensätzen umgegangen wird, die Einzelschlagwörter aufweisen, die bereits vollständig oder in Teilen in zu übernehmenden Schlagwortfolgen vorhanden sind. Herr Völkl und Frau Wolf-Dahm sehen hier aufgrund des größeren Sinnzusammenhangs in den Schlagwortfolgen gegenüber den Einzelschlagwörtern einen Mehrwert und machen den Vorschlag, Einzelschlagwörter zu löschen, sofern sie bereits in einer oder mehreren Schlagwortfolgen vorhanden sind. Da es nach Aussage von Herrn Thiessen technisch mit vertretbarem Mehraufwand möglich wäre, im Zieldatensatz Einzelschlagwörter zu löschen, wenn sie bereits in der zu übernehmenden Schlagwortfolge vorhanden sind, wird dieses Vorgehen beschlossen.

Mit diesem Beschluss kann Herr Thiessen mit der Umsetzung von Phase 1 der Culturegraph-Anreicherung beginnen. Allerdings möchte er vorher noch eine weitere Testumsetzung initiieren und an die AG SE verschicken. Wie beim letzten Mal wird Herr Rüter sie in die besser zu bearbeitende Form einer Word-Datei bringen. Etwaige Korrekturen, die dann noch auffallen, sollen aber nicht mehr in die nächste Sitzung vertagt werden. Vielmehr wird eine Verschweigefrist bis Mitte Juli vereinbart, so dass der Zeitplan für die Umsetzung von Phase 1 somit folgendermaßen aussieht: Herr Thiessen stellt der AG SE bis Ende Juni per Mail eine weitere Testumsetzung zur Verfügung. Etwaige Hinweise und Korrekturwünsche, die bis zum Ende der Verschweigefrist Mitte Juli bei ihm eingehen, berücksichtigt er noch. In der zweiten Julihälfte folgt dann die produktive Umsetzung von Phase 1. Dabei sollen für etwaige spätere Berichte über die Culturegraph-Anreicherung folgende statistische Daten erhoben werden:

- Wie viele Datensätze wurden bearbeitet?
- Wie viele Schlagwortfolgen wurden insgesamt ergänzt?
- Wie viele Datensätze weisen nach der Culturegraph-Anreicherung eine Sacherschließung auf, die vorher keine hatten?
- Wie viele Schlagwortfolgen wurden durchschnittlich pro Datensatz vergeben?
- Wie hat sich die prozentuale Verteilung von Schlagwortfolgen pro Datensatz nach der Culturegraph-Aktion in der betroffenen Datenmenge verändert?

Damit wäre Phase 1 der Culturegraph-Anreicherung abgeschlossen. Phase 2, in der RVK-Notationen angereichert werden sollen (s. Protokoll der 9. Sitzung der AG SE vom 15.04.2021), wird zurückgestellt, bis die Migration der RVK-Notationen von AESQ 700g nach ASEQ 701g/i/y (s. TOP 3) erfolgt ist.

TOP 3: Migration der RVK-Notationen von ASEQ 700g nach ASEQ 701

Gegenüber dem in der 9. Sitzung der AG SE festgelegten Fahrplan hat sich in puncto Migration der RVK-Notationen von ASEQ 700g nach ASEQ 701g/i/y bisher nichts geändert: Sobald Herrn Thiessen die „Positivliste“ (Notationen, deren Bedeutung sich in den letzten 10 Jahren nicht geändert hat) aus Regensburg vorliegt, kann er einen Testlauf durchführen. Frau Schweikl hält es für wahrscheinlich, dass Herr Thiessen die entsprechende Liste noch vor den Sommerferien bekommen wird und dementsprechend nach der Sommerpause die Migration in der Produktivphase umsetzen kann. Die Freigabe der Produktivphase nach Erhalt der Positivliste wird einstimmig beschlossen.

Frau Schweikl berichtet, dass ein erster Test zeigt, dass die Anzahl der Datensätze mit Konzeptänderung gering sein wird (35 Datensätze seit der Einführung der IDs in 2016/4.Quartal) und regt – unabhängig von dem beschlossenen Beginn der Produktivphase – an, dass durch die Auswertung der Rundbriefe vor 2009 die Dokumentation der Konzeptänderung noch komplettiert, mithin die „Positivliste“ ergänzt werden könnte.

Frau Schweikl und Herr Thiessen berichten, dass die Problematik ungültiger RVK-Notationen/-ID im Hinblick auf den verbundübergreifenden Austausch für die AG KVA zusammengestellt wird. Frau Schweikl bietet an, das einschlägige Papier der AG SE zukommen zu lassen. Der wesentliche Punkt ist hierbei die Frage, ob im Austauschformat ein neues MARC-Unterfeld benötigt wird, das ungültige IDs kennzeichnet.

Die endgültige Entscheidung, wie im B3Kat mit dem Import von RVK-Notationen umzugehen ist, wird solange zurück gestellt, bis bekannt ist, wie andere Bibliotheksverbünde dieses Thema handhaben.

TOP 4: Migration der sonstigen Altdaten von ASEQ 700# nach ASEQ 701#

Besprochen werden die von Herrn Thiessen im Vorfeld zur Verfügung gestellten Dateien „mapping 700_701“ (Übersicht) sowie die zugehörigen Beispieldateien. Eine erste Einschätzung führt zu der Vermutung, dass vor allem der Fall ASEQ 700_ ohne das Unterfeld 2 intensiver zu prüfen ist, alle andern Fälle können aller Wahrscheinlichkeit nach ohne Probleme übernommen werden. Auf die Frage, wie diese Listen weiter zu bearbeiten sind, schlägt Herr Rüter vor, dass sich die Mitglieder der AG SE die Liste aufteilen und bis zur nächsten Sitzung die Datenlage analysieren. In diesem Zusammenhang sind vor allem die Fragen zentral, ob Auffälligkeiten in der Belegung erkennbar sind oder (gerade bei 700_) einzelne Notationssysteme aufgrund ihrer Syntax identifizierbar sind und eine systematische Behandlung durch einen Algorithmus ermöglichen.

Herr Rüter erklärt sich bereit, die Beispieldateien in Excel-Dateien zu verwandeln, eine Kommentarspalte einzufügen und zur gemeinsamen Bearbeitung in der HU-Box zu hinterlegen.

In der nächsten Sitzung soll eine erneute Diskussion aufgrund der oben beschriebenen Datenanalyse stattfinden.

TOP 5: B3Kat-Aktionen und B3Kat-Bereinigungen

Übernahme von Sacherschließungsdaten der Print-Ausgabe für E-Book-Metadateneinspeicherung von Verlagsdaten

Auf der Grundlage eines neu entwickelten Rahmenskripts für eine Anreicherungsroutine für E-Book-Aufnahmen schildert Herr Thiessen den Stand der Dinge: Probeweise hat er ein Springer-Paket herangezogen und festgestellt, dass mit Hilfe der in Feld 776 hinterlegten ISBNs parallele Printaufnahmen im B3Kat ermittelt werden können. Sofern Feld 776 entsprechend belegt ist, können somit beliebige Feldinhalte selektiert und in den E-Book-Aufnahmen angereichert werden. In Bezug auf das herangezogene Springer-Paket würde das – nach einer groben Schätzung von Herrn Rüter und Herrn Völkl – einen Sacherschließungs-Mehrwert (RVK) für ca. 20% der E-Books bedeuten. Beispielhaft wurden bisher RVK-Anreicherungen (sowohl aus Feld 701g als auch aus Feld 700g) für Testzwecke umgesetzt. Das Anreicherungsskript muss allerdings noch optimiert werden.

Im Anschluss wird deshalb die Frage diskutiert, wie eine Anreicherungs-Routine gestaltet werden könnte, bis eine regelmäßige Culturegraph-Update-Routine (s. o.) etabliert und somit auf diesem Weg auch eine automatische E-Book-Anreicherung garantiert ist. Vorstellbar wäre ein retrospektives Vorgehen ab dem dritten Quartal 2020, denkbar zudem die Priorisierung einzelner Pakete. Herr Rüter erklärt sich bereit, im KOBV-Rahmen geeignete Pakete zu eruieren.

Da eine gleichzeitige Anreicherung nicht nur mit RVK-Notationen, sondern auch mit Schlagwörtern komplizierter und dementsprechend weniger schnell umsetzbar wäre, wird beschlossen, sich vorläufig auf eine RVK-Anreicherung zu beschränken und eine Schlagwortanreicherung ggf. zu einem späteren Zeitpunkt vorzunehmen.

Löschung von nicht in der GND enthaltenen Schlagwörtern in den Feldern 902ff. mit Indikator f

In der Frage, wie mit Schlagwörtern umgegangen werden soll, die nicht in der GND enthalten sind, aber mit der Codierung f in Feld 902ff. abgelegt wurden, gibt es grundsätzlich zwei Lösungsvorschläge: Entweder werden diese falschen Formschlagwörter ersatzlos gelöscht oder in Feld 711 als freie Schlagwörter verschoben. Über die beiden Vorgehensweisen wird abgestimmt. Für die Verschiebung in Feld 711 stimmen 5 Mitglieder der AG SE, für die Löschung lediglich 3. Somit wird die Verschiebung in Feld 711 beschlossen.

TOP 6: Bericht aus der RVK

Frau Schweikl, die den Bericht im Vorfeld bereits an die AG SE versendet hatte, stellt ihn kurz vor, wobei sie besonderes Augenmerk auf den Punkt „Ausdifferenzierung von Schlüsseln“ legt. Der Bericht wird dankend zur Kenntnis genommen.

TOP 7: Verschiedenes

Personelle Veränderungen in der Verbundzentrale

Herr Kratzer verlässt zum 1.8.2021 die Verbundzentrale und übernimmt als Nachfolger von Herrn Strasser die Leitung des IT-Referats der BSB. Aufgrund der bis zur Neubesetzung der Stelle von Herrn Kratzer eingeschränkten personellen Ressourcen der Verbundzentrale wird u.U. eine Priorisierung der in diesem Zeitraum zu bearbeitenden Projekte der AG SE notwendig.

Anfragen an die AG SE für Gastteilnahme

Herr Holbach und Herr Rüter informieren darüber, dass sich eine Kollegin mit der Bitte an sie gewendet hat, zukünftig den Sitzungen der AG SE als Gast beizuwohnen. Diese Anfrage wird als Anlass dafür genommen, über die Frage zu diskutieren, wie die AG SE mit (ständigen) Gästen umgehen möchte. Grundsätzlich zeigt sie sich einer Erweiterung des Kreises um Gäste (entweder zu einzelnen oder mehreren Sitzungen) gegenüber aufgeschlossen; der Einzelfall müsse jedoch jeweils diskutiert werden. Im vorliegenden Fall wird beschlossen, dass Herr Rüter nochmals mit der Kollegin Kontakt aufnimmt, ihre Interessenlage klärt und sie gegebenenfalls zur nächsten Sitzung der AG SE einlädt.

Termin der nächsten Sitzung

Vor Ende der Amtsperiode zum Jahresende 2021 sollen noch drei weitere (virtuelle) Sitzungen der AG SE stattfinden und zwar am

7.10.2021 (9:30-13:00 Uhr)

4.11.2021 (9:30-13:00 Uhr)

16.12.2021 (9:30-13:00 Uhr)

München, 02.07.2021