

Thema: "Erarbeitung eines Standards zur OPC4 - Indexierung und - Anzeige"

Teilnehmer:

Frau Klappstein, Frau Thoms, Frau Wegelt, Herr Aschenbach, Herr Hapke, Herr Lörzer, Herr Marbach

Als Gäste: Frau Junger, Frau Kemner, Frau Niemann, Hr. Haverich (alle GBV Verbundzentrale)

Frau Blödorn-Meyer, Hr. Finger, Hr. Tannhof (alle FAG Sacherschließung)

Frau Hitzler (SUB Hamburg), Frau Rajski (TU Hamburg-Harburg), Herr Baum (TIB Hannover), Herr Weichert (SUB Hamburg)

Protokoll: Herr Hapke

Tagesordnung

TOP 1: OPC4 - Überblick über den aktuellen Stand und Ausblick

TOP 2: Indexierung

TOP 3: Syntax (Anzeige)

TOP 1: OPC4 - Überblick über den aktuellen Stand und Ausblick

Zwei Pilot-Projekte von OPC4 laufen für die Staatsbibliothek zu Berlin (<http://e450-043.gbv.de:8080>) und für die TIB Hannover (<http://sun31.tib.uni-hannover.de/cgi-bin/wwwopc4menu>). Da die aktuelle OPC4-Version nur auf einer SUN läuft, steht das Berliner System kurz vor der Produktionsreife, das Hannoveraner System ist nicht ganz aktuell. Die neue Version für die DEC-Hardware soll Ende November geliefert werden. Zunächst sollen alle SUNs im Verbund mit OPC4 ausgerüstet werden. Ein erstes Testsystem steht schon an der Universität der Bundeswehr Hamburg, danach wird die TU Hamburg-Harburg an der Reihe sein.

Die Schnittstelle zur Ausleihe muß mit der Produktionsreife funktionieren. Zur Zeit sind hier noch kleinere Probleme zu überwinden. Das PICA-Projekt LBS4 der Freien Universität Amsterdam, wo als fester Termin der 1.10.2000 vorgesehen war, hat sich bis zum Ende des Jahres verschoben, so dass daraus resultierende Verzögerungen auf den GBV durchschlagen.

OPC4 läuft parallel zu allen LBS3-Modulen, auch mit OPC3 und zum OWC3. Die OPC4-Indexierung wird auch Grundlage der Indexierung für den Dienstkatalog innerhalb LBS4 sein.

Ziel dieses Workshops ist die Erarbeitung einer Muster-Indexierung und -Syntax, die an allen GBV-Standorten von der Verbundzentrale installiert werden wird. Abweichungen vom Standard und Pflege dieser Änderungen bei Versionswechsel sollen nur von den einzelnen Bibliotheken erfolgen. Bibliotheken, die daran interessiert wären, wird ein Workshop in Göttingen angeboten werden (1. Quartal 2001). Eine Um-Indexierung ist relativ leicht möglich. An der SBB konnte dies innerhalb von 12 Stunden in parallelem Betrieb erfolgen. Auch für die Syntax/Anzeige ist ein Workshop geplant, da hierfür keine Dokumentationen vorliegen.

Ein wichtiger allgemeiner Gesichtspunkt bei der Indexierung wird das Verhältnis zur GSO-Oberfläche sein (Search&Order-Oberfläche des Gemeinsamen Verbundkatalogs GVK). Beide Indexierungen sollen so ähnlich wie möglich erfolgen.

Grundlagenpapiere der Diskussion:

- Dokumentation auf dem GBV-Server (<http://www.gbv.de/lbs/picadoc/lbs.shtml#opc4>) mit Muster-Index-Tabellen und Papier "Die Indexierung im OPAC" und Index Generation Manual
- Stellungnahmen von Herrn Baum (Mail vom 15.11.2000) und Herrn Lörzer (Mail vom 20.11.2000)
- Papier der FAG Sacherschließung (Hr. Müller-Dreier, 17.11.2000)

TOP 2: Indexierung:

Änderungen/Verbesserungen gegenüber der alten Indexierung:

- fast identisch mit CBS-Möglichkeiten
- viel flexibler: Verschiedene Suchschlüssel können in einem Index angeboten werden, diese können aber als "Unter-Indices" auch extra angeboten werden
- keine Begrenzung auf 1000 bzw. 15.000 Treffer
- All-Index; viele ADIs können durch Primärschlüssel ersetzt werden)
- Phrasensuche stark verbessert (Phrasen länger und über mehr als 2 Subfields programmierbar
- automatische Rechtstrunkierung
- erweitertes Angebot an Indexroutinen

Die automatische Rechts-Trunkierung bei der String-Indexierung ist leider nicht abstellbar. Es ist also beim Nutzer "Wissen" notwendig (Einschluß in " als Phrasensuche), um z.B. kurze Titelanfänge effektiv zu suchen. Sinnvoll wäre hier ein automatische Springen in die Index-Anzeige !

Welche Suchschlüssel/ADIs soll die OPC4-Indexierung für den GBV enthalten?

Anhand der "Index definition table" (Muster TIB in der GBV-Dokumentation) werden nun alle Such-Indices (siehe unter "ikt definitions") nacheinander durchgesprochen. Die angestrebte Muster-Indexierung, Basis ist die GSO-Indexierung, soll möglichst alle Wünsche berücksichtigen. Die Bibliotheken können dann lokal selbst entscheiden, welche Einstiege sie ihren Kunden bieten wollen.

KOS (Körperschaften - String)

KNS (Kongresse - String)

Phrasenindices machten bisher Probleme in LBS, da nicht mehr als 2 subfields berücksichtigt werden konnten.

TIT (Titelstichwörter)

SER (Serien)

"Überindex", der SHA und VSE zusammenfaßt

ISB (ISBN)

ISS (ISSN)

PPN

KLS (Klassifikationen)

Auch Klassenbeschreibungen der BK sind indexiert. Dewey-Klassifikation ist Teil des Index davon, Schrägstriche und Hochkomma bei Dewey werden weggelassen (bei Neueinspielungen der Dewey-Daten der LoC-Version werden in Zukunft bis auf den Punkt alle Sonderzeichen

herausgenommen).

5308 und 5309 enthalten keine Normdaten-Verknüpfungen (Nutzung für "ZDB-Systematik") und sollen nicht mit indexiert werden. Siehe auch-Verweisungen (270 = 045B) bei der BK sollen nicht mit indexiert werden.

DTM (Datum)

Aus 70XX !

JAH (Erscheinungsjahr)

Kein ADI ! ADI bleibt notwendig um Größer-Kleiner-Beziehungen suchen zu können, z.B. alle Publikationen zwischen 1992 und 1995.

SWW (Schlagwort - String, Ketten)

Für feldübergreifende Suche zusätzlicher Index-Eintrag (analog ANS für Kategorien 801 bis 805 im CBS)

SGN (Signaturen - mit Blanks)

Sonderzeichen sind durch Leerzeichen ersetzt, enthält Einträge der Signatur, des Standortes extra und Standort mit Signatur.

SGB (Signaturen - ohne Blanks)

Sonderzeichen und Leerzeichen werden eliminiert (Standard-Index für Signaturen der SBB !)

PER (Personen)

Diskutiert wird hier das Problem der Weichheitszeichen in Namen. Die echten verschwinden bei der Indexierung im CBS. Das Ersatzzeichen Apostroph hat bei arabischen Namen z.B. auch trennenden Charakter. Daher wird z.B. Gor"ki indexiert als Gorki, gor und ki. Ein generelles Problem ist die Spatienempfindlichkeit der SWD (Goethe,j? führt zu keinem Treffer im Schlagwort-Index !). Das Weglassen von Spatien zwischen Nachname und Vorname ist nicht möglich, da die einzelnen Namensteile nicht in eigenen Subfields abgelegt sind und zur Zeit außer Personen auch andere Begriffe mit zwei oder mehr Wörtern in der SWD enthalten sind. (Mit der Personen-Normdatei kommt dieses Problem auch für den Personen-Index.)

KOR (Körperschaften - Stichwort)

KON (Kongresse - Stichwort)

NUM (Nummern)

Signatur ist bei den Nummern mit drin. Diskutiert wird hier kurz das Problem der Vermischung von Ebene 0 (Bibliographische Daten) und Ebene 2 (Exemplar-Daten). Signaturen sollen im Kunden-Interesse auch über den Nummern-Index gefunden werden können.

PUS (Erscheinungsvermerk - String)

Einträge unter Ort, Ort + Verlag und Verlag.

ALL (All-Index)

Gleiche Lösung wie in GSO ! Die Hannoveraner Lösung, Verlag und Ort nicht mit zu indexieren (wegen Problematik, wenn ein Verlagsort als Titelstichwort gesucht wird), wird nicht für den Standard übernommen, da ein Index "Alle Wörter" wirklich alles enthalten sollte. Lediglich die klassifikatorischen Beschreibungen und Notationen (?) sind nicht mit im Index. Es wird darauf hingewiesen, dass der All-Index in der Regel nicht für Ein-Wort-Anfragen genutzt werden wird.

THM (Thematische Suche - neu!)

Die FAG Sacherschließung schlägt als zusätzlichen Index einen eigenen Basic-Index für die thematische Suche vor (s. Papier von Herrn Müller-Dreier), der Titelstichwörter (nur Hauptsachtitel !) und alle Schlagwörter wortweise indexiert enthält. Der Index soll in den Standard aufgenommen werden und für GSO soll er ebenfalls vorgeschlagen werden. Alle anderen Register zur Sacherschließung sollen aber erhalten bleiben.

BKL (Basisklassifikation)

Bei der Suche im OPC4 werden Normsätze anscheinend nicht mehr mit angezeigt, auch nicht wenn man über das Linking sucht. Es ist schade, das man dies nicht mehr einstellen kann.

In den Pilotversionen erhält man nach Auswahl "Sachgebiete" neben den BK-Notationen z. Zt. einen Link "Verwandte Sachgebiete", der auf eine alphabetische Liste der mit dieser BK-Notation verbundenen Schlagwörter führt. Schön wäre an dieser Stelle eine Sprungmöglichkeit an eine bestimmte Stelle des Alphabets. Herr Weichert zeigt, dass dies durchaus möglich wäre, da man "im angezeigten Link durch Eingabe eines Buchstabens nach TRM= an die betreffende Stelle in der Liste springt".

An dieser Stelle wurde darauf hingewiesen, dass unter "Sachgebiete" im linken Auswahlmenü auch der Einbau einer hauseigenen Systematik möglich wäre (SBB plant dies). Ein neuer Button könnte dann zu HTML-Dateien mit einer anderen Systematik führen. Anklicken auf Notation in dieser System löst dann eine Recherche mit dieser Notation im Opac aus. (Die TUHH hat etwas Ähnliches z.B. unter der URL <http://www.tu-harburg.de/b/ls/lssys-e.htm> realisiert.)

SST (Standort)

SLW (Schlagwörter wortweise)

EPN

SWT (Swets-Nummern)

TAF (Titelanfänge - String)

Dieser Index wird auch für GSO gewünscht (entspricht TIS aus dem CBS ohne Einschränkung auf drei Wörter plus Kongresse). Nicht nur Hauptsachtitel !

Herr Hapke weist auf einen vorhandenen Fehler bei der Funktionalität von TAF hin (Blättern im Index und Auswahl bestimmter Strings führt nicht zum gewünschten Ergebnis). Dieses Problem bestand vor der Verbundkonferenz auch im GSO, konnte aber über Frau Hachmann schnell beseitigt werden.

GTI (Über-Index für ZTI und STI)

ZDB

SNR

BAR

LPN

(alle über NUM suchbar)

STI (Serien-Stücktitel)

Aus GSO

VSE (Serienstichwörter)

(enthält den Inhalt der Kategorien 411X und 417X)

SHA (Serien-Hauptaufnahme)

Aus GSO

ZTI

URL

PUB (Erscheinungsvermerk - wortweise)

MAK (Materialcode aus 0500)

Auch über den Link "Suchoptionen" rechts von der Standard-Eingabezeile steuerbar. Bei allen Codes wäre als Zwischenschritt bei der Recherche ein Auflösen und Umsetzen in Klartext notwendig (z.B. über eine zwischengeschaltete HTML-Seite ("kontextsensitives Aufklicken")). Generell problematisch ist ausserdem, dass bei der Recherche mit anderen Codes oft nur noch Teilbestände gefunden werden, da nicht alle Dokumente mit Codes versehen sind. **Bei den Suchoptionen wäre ein Button "Alles Demarkieren" sinnvoll, um dann z.B. nur Zeitschriften zu markieren.**

SPR (Sprachcode aus 1500)

LAN (Ländercode aus 1700)

ZTC (Zeitcode aus 1111)

DKT (Dokumenttyp aus 1110)

GEO (Geographischer Code aus SWD Normdatensatz 811*)

ERJ (Erscheinungsjahr, ADI) Notwendig für "Größer-Kleiner-Recherchen"

MAT (Materialcode) Notwendig zum Anbieten der Suchoptionen (siehe MAK)

KEY (aus 70XX)

Notwendig um auf Teilbestände einschränken zu können (logische Teilsichten innerhalb einer ILN, z.B. SUB-HH). Fr. Klappstein fragt nach der Bildung von Teilbeständen, die in der 70XX unterschiedliche Eintragungen enthalten. Das ist nicht möglich. Die ULB Halle verwendet für die Zweigbibliothek unterschiedliche Selektionscodes. Möchte man nun einen Teilbestand im OPAC über mehrere Bibliotheken eines Fachbereichs einrichten, ist das nicht möglich.

BCL (ADI zur Basisklassifikation)

Der Vorschlag aus der TIB, einen weiteren Index (wortweise) für den Verlag anzubieten, wird nicht in den Standard übernommen.

Herr Baum schlägt vor, die Materialart "T" =Zeitschriften und Serien" so umzudefinieren, dass nur noch Zeitschriften (Kat. 0500=Ab) erfasst werden. Es wird auf die Grauzone bei der Definition Zeitschrift/Serie hingewiesen aber auch auf die Vorteile der besseren Zeitschriftensuche. Da eine Entscheidung auch Auswirkungen auf den GSO haben wird, wird Fr. Thoms eine Umfrage über Pica-Mailing-Listen starten und danach das Problem in den Fachbeirat bringen.

Die TIB schlägt einen eigenen Schlüssel für den Dissertationsvermerk vor, obwohl dieser im ALL-Index enthalten ist. Ein solcher Index wird in den Standard übernommen.

Diskutiert wird die Erstellung einer Dokumentation der Muster- bzw. Standard-Indexierung durch die Verbundzentrale. Es wäre für alle Beteiligten eine große Hilfe eine lesbare Übersicht über die Register mit den darin enthaltenen Kategorien zu bekommen (Beispiel für einen Teil: Papier von Hr. Müller-Dreier). Die TIB Hannover wird ihren vorhandenen Entwurf der Verbundzentrale zur Verfügung stellen und die Verbundzentrale die gewünschte Übersicht erstellen.

Hr. Finger fragt nach einem Schlüssel für die Suche nach dem Maßstab von Karten. Er wird in den Standard übernommen werden. Der Ausgabevermerk bei Musikalien ist sowohl in TIT als auch in ALL enthalten (Frage Hr. Weichert).

Ein Index ATS (analog dem CBS) wird in den Standard übernommen.

Ein Index für mehrbändige Werke wird nicht in den Standard übernommen.

Welche Suchschlüssel sollen dem Benutzer direkt im OPC4-Pull-Down-Menue angeboten werden?

(Reihenfolge, deutsche und englische Bezeichnung)

Die Steuerung, welche Suchschlüssel angezeigt werden, erfolgt über eine Datei (class-Profile) und kann von den Bibliotheken leicht selbst erfolgen. Für den Standard wird der Index THE nach dem ALL-Index aufgeführt. Eine Aufführung analog GSO ist ansonsten natürlich sinnvoll. Es sollte nicht zuviel aufgeführt werden (10, maximal 20), weil man je nach Software und Hardware eventuell sonst im Auswahlménú scrollen muss. Die Datei soll in zwei Versionen ausgeliefert werden, eine mit (analog ZDB-Oberfläche) und eine ohne den Abkürzungen der Schlüssel hinter dem Klartext.

Eine Diskussion über den Wegfall des Expertenménús ergab kein klares Ergebnis für ein Expertenménú. Aufgrund von Hamburger Untersuchungen (Fr. Blödorn-Meyer, Hr. Weichert) laufen maximal 10 % der Anfragen über dieses Ménú. Eine Verknüpfung von Suchanfragen in unterschiedlichen Feldern ist über die Suchgeschichte leicht möglich (natürlich auch direkt durch Verwendung der "Kommandosprache" (d.h. der Abkürzungen der Suchindices: per schmidt and tit verfahrenstechnik)).

An dieser Stelle wird auf die Problematik hingewiesen, dass alle Suchschlüssel automatisch als Stopwörter fungieren. Daher wird entschieden, für den Index der thematische Suche statt THE die Abkürzung THM zu wählen. Für ALL wäre vielleicht BSC ("Basic Index") sinnvoller (Abstimmungsnotwendigkeit mit GSO !), da das Suchwort "all" durchaus Teil einer sinnvollen Suche sein kann.

TOP 3: Syntax (Anzeige)

Das grundsätzlich z.Zt. vorhandene Problem der Anzeige ist nur sehr schwer zu lösen: Einerseits sollte alles angezeigt werden, was suchbar ist, um den Kunden eine Kontrolle seiner Suchbegriffe zu ermöglichen. Dies ist allerdings schon bei der Verwendung von Synonymen aus den Normdateien, die mit gesucht aber nicht angezeigt werden, nicht möglich. Andererseits schiebt eine umfangreiche Anzeige die Angabe der Signatur und den möglichen Bestellbutton immer weiter nach unten.

Die SUB-HH möchte nur normiertes Vokabular anzeigen, die TUHH legt werden für ihr internationales Publikum auf die internationale Sacherschliessung bei der Anzeige.

Eine Anzeige der Signatur ist laut Aussage von Fr. Kemner nicht oberhalb der Aufnahme möglich. Prinzipiell muß bei der Anzeige die bibliographische Ebene eingehalten werden. Noch besser wäre übrigens eine Anzeige von Signatur und Ausleihstatus schon in der Kurztitel-Liste nach einer Recherche ! Hr. Haverich schlägt vor zu versuchen, oberhalb der Titelbeschreibung einen Link anzubringen, der direkt zu den Bestandsinformationen führt.

Gut wäre eine Differenzierung der Anzeige, so wie sie im Gesamtkatalog praktiziert wird, mit Kurzaufnahme und Verfügbarkeitsinformation und mit Titeldaten. Durch einen "blinden" Link zwischen dem linken Auswahlménú und der Beschreibung der Suche ist übrigens das PICA+-Format anzeigbar.

Es wird entschieden, dass von der Verbundzentrale zwei Versionen für die Anzeige angeboten werden sollen: eine Standard Version für kleine Bibliotheken mit einer deutlichen Einschränkung bei der Anzeige, eine maximale Version für größere Bibliotheken, die dann durch "Auskommentieren" an die lokalen Bedürfnisse angepasst werden muss (Achtung: Pflege bei Versionswechsel!). Leider ist zur Zeit hier keine Dokumentation vorhanden. Hr. Haverich erwähnt allerdings, dass eventuell die Pflege/Verwaltung der Anzeige vom Systemverwalter auch über eine Browser-Oberfläche möglich sein soll.

Die **Formaldaten** werden in der Standard-Version analog wie im GSO angezeigt. Hr. Baum bemängelt die vollständige Anzeige der Kategorie 4000, da Verfassernamen dann zweimal und direkt untereinander auftauchen. (Hier ist eine Abstimmung mit GSO nötig.)

Der Vorschlag von Hr. Baum für die **Online-Titel (elektronischen Dokumente)** wird übernommen:

Volltext: Die Kat. 4083 / 009P/03 sollte nicht angezeigt werden (da die Links nicht immer zum richtigen Server führen), sondern nur (zugleich als Link) die 7133 / 209R (mit dem Text aus 4083). 009P/09 soll aber weiterhin angezeigt werden.

Die "Physikalische Beschreibung" bei den Mikroformen soll weggelassen werden.

Im "kleinen Standard" erfolgt die Anzeige für die **Sacherschließung** so:

Schlagwörter 51XX

Schlagwörter 55XX

Sachgruppen 530X

Weitere lokale Sacherschließungselemente

Für die "große Version" der Standard-Anzeige werden die Vorschläge der FAG Sacherschließung übernommen:

Klassifikationen evtl. unter der Bezeichnung "ähnliche Titel": 5010 (DDC), 5030 (LCC), 5040 (NLM) in einer Zeile

Schlagwörter: 5500 (LCSH), evtl. 5540 (BLSH), falls 5500 nicht vorhanden, 5510 (MESH)

Hr. Baum weist daraufhin, dass die BK nicht mit den anderen Klassifikationen zusammen angezeigt werden sollte, da die anderen wesentlich spezieller seien und eher einem Link "Ähnliche Titel" (Amazon!) entsprächen.

Linking in der Anzeige

Ein Linking der ISBN erscheint nicht notwendig.

Der Vorschlag beim Sachtitel nur den Hauptsachtitel zu linken wird abgelehnt, da bei den meisten Titel nur ein Treffer das Ergebnis der Recherche wäre (nur bei gleichen Auflagen wäre dieser Link sinnvoll).

Bei mehrgliedrigen Reihen sollte der Link auf alle Teile erfolgen (Subfields mit berücksichtigen).

Beim Linking bei Kongressen sollte das Linking ohne die Subfields für Zählung und Ort erfolgen, bei gleichzeitiger automatischer Und-Verknüpfung der Kongresstitel-Stichwörter (geht das überhaupt ?) oder unter Verwendung des INDEX KOS.

Die Verknüpfungen 039B-039E müssen über die ZDB-Nummer als String gelinkt werden (Beispiel ppn 129597503/Bauwelt oder ppn 315041005) (Mail Hr. Baum). Auch bei GSO berücksichtigen !

Für das Protokoll: Hapke, 23.11.2000, Korrekturen: 14.12.2000

[Schließen](#)