

Protokoll_HeBIS-AG 2016_06_15

Protokoll

Workshop der HeBIS-AG Sacherschließung und Recherche

„Sacherschließung im HeBIS-Verbund" am 15.06.2016

Ort: Universitätsbibliothek Johann Christian Senckenberg, Frankfurt am Main, Sitzungszimmer

Zeit: 10:00 -16:00 Uhr

Sitzungsleitung: Dr. Diana Müller (LBS Marburg/Vorsitz)

Protokoll: Karen Stuckert (LBS Rheinhessen)

Anwesende Mitglieder: Ursula Bernhardt (UB Frankfurt/Verbundredaktion SE), Manja Biermann (LBS Gießen/Fulda), Sabine Kroll (LBS Rheinhessen), Sabina Lüdemann (LBS Kassel), Dr. Diana Müller (LBS Marburg/Vorsitz), Dr. Eve Picard (LBS Frankfurt), Stephan Rosenke (LBS Darmstadt/Wiesbaden), Dr. Olaf Schneider (LBS Gießen/Fulda), Karen Stuckert (LBS Rheinhessen)

Ständige Gäste: Renate Koch-Krug (UB Frankfurt/Verbundredaktion SE)

eingeladene Referendare: Michael Ammon (LBS Darmstadt/Wiesbaden), Fee-Saskia Fricke (LBS Darmstadt/Wiesbaden), Björn Gebert (LBS Darmstadt/Wiesbaden), Verena Grunewald (LBS Gießen/Fulda), Max Kroehling (LBS Gießen/Fulda), Nadja Krüll (UB Frankfurt), Ingeborg Lederer-Brüchner (UB Frankfurt)

Referenten: Aline Deicke (ADW Mainz), Dr. Timo Glaser (UB Marburg), Dominik Kasper (ADW Mainz), Elisabeth Mödden (DNB Frankfurt)

Entschuldigt: Cindy Kathke (LBS Darmstadt/Wiesbaden), Elmar Reith (Verbundzentrale), Jörg Siedler (LBS Marburg), Brigitte Störch (LBS Frankfurt), Andrea Thiel (LBS Gießen/Fulda)

Begrüßung und Einführung

Frau Müller begrüßt alle Teilnehmer, die sich aus den Mitgliedern der AG, Referendaren aus HeBIS-Bibliotheken und den Referenten zusammensetzen. Sie führt ein in aktuelle Entwicklungen sowie sich verändernde Rahmenbedingungen für die Sacherschließung, die einen eigenen Workshop zum Thema ratsam erscheinen ließen.

Themenblock 1: Erschließungskonzepte und -strategien

Frau Müller stellt den ersten Referenten Herrn Timo Glaser (UB Marburg) vor, der als Fachreferent, OA-Beauftragter sowie Verantwortlicher für das Elektronische Publizieren tätig ist und zum Thema „Sacherschließung trotz Information Behavior? Beobachtungen am Marburger Publikationsserver" spricht. Timo Glaser hinterfragt in seinem Vortrag das Zusammenspiel von Rechercheverhalten unserer Nutzer und Sacherschließung. Mit Verweis auf die Untersuchung „How Readers Discover Content in Scholarly Publications" (2016) wird zunächst festgestellt, dass das Suchverhalten von Bibliothekaren und das von Nutzern weit auseinander liegt. Am Beispiel des Marburger Publikationsservers (11.000 Dokumente) untersucht Timo Glaser, wer die Sacherschließung der Dokumente vornimmt (Nutzer, Bibliothek, beide gemeinsam), welche Dokumente auf dem Publikationsserver liegen (Artikel, Zeitschriften, Dissertationen, andere Hochschulschriften, Daten, Bücher, Partituren, Videos) und welche Art der Erschließung erfolgt (GND, DDC, freie dt./engl. Schlagworte, dt./engl. Abstracts, keine). Außerdem wird analysiert, woher die Aufrufe kommen (Google oder andere Suchmaschine, Google Scholar, Direktaufrufe des Marburger Publikationsservers, Bibliothekskataloge) und es wird die Suchstrategie von Nutzern des Publikationsservers untersucht (Einwort-, Mehrwort-, Namen-, Anklick-Suche). Provokativ endet der Vortrag mit drei Fragen: Ist Google aus Nutzersicht der effizientere Bibliothekskatalog? Welchen Mehrwert hat die Doppelkatalogisierung im CBS und im Publikationsserver? Welchen Mehrwert bieten Normdaten?

Die anschließende Diskussion fragt nach den Ursachen für eine vermeintliche Nichtnutzung bibliothekarischer Sacherschließung. Sind bspw. die DDC-Sachgruppen zu grob? Sollten die Kataloge weiter durch Suchmaschinentechologie verbessert werden? Ist die Untersuchungsgrundlage eines Hochschulschriftenservers aussagefähig genug, da die Inhalte eher zufallsbestimmt sind und eine Parallelaufnahme im CBS nicht vollständig erfolgt, was wiederum eine Suche über Bibliothekskataloge nur bedingt sinnvoll macht? Sollte das Augenmerk verstärkt auf Discovery-Systeme gelenkt werden, die über den Umweg BASE die Inhalte einsammeln? Dadurch, dass die Inhalte auf den Publikationsservern derzeit von der Ablieferung der jeweiligen Autoren abhängig sind, wurde das Gespräch auf den HeFIS-Verbund (Forschungsinformationssysteme Hessischer Hochschulen) gelenkt, welcher zum Ziel hat, den wissenschaftlichen Output in seiner Gesamtheit zu dokumentieren und die Forschungsdaten besser zu vernetzen. Diskutiert wurde ebenfalls, welche Rolle eine Qualitätskontrolle zu spielen habe, da teilweise hessische Bibliotheken die Aufnahmen in den Publikationsservern gar nicht mehr kontrollieren.

Nach Abschluss des Themenfeldes ‚Online-Publikationen auf Hochschulschriftenservern‘ wird das Blickfeld auf ‚eBooks‘ erweitert. Frau Bernhardt stellt die bisherigen Maßnahmen vor, die in der UAG eBooks erarbeitet und z. T. bereits umgesetzt wurden: Pauschale RVK-Anreicherungen, Ausweitung des Springer-Verfahrens auf die übrigen Pakete, Konkordanz DNB-Sachgruppen HDS-Fachfacetten, in Kooperation mit der HeBIS AG E-Medien: Verknüpfung der ISBNs von Print- und E-Ausgaben (ausführlicher siehe Protokoll vom 25.04.2016).

In der anschließenden Diskussion wird die Wichtigkeit von Normdaten insbesondere Normdaten als normiertes Vokabular betont. Auch wenn sie

von Nutzerseite nicht bewusst wahrgenommen werden, tragen sie doch maßgeblich zu einer erfolgreichen Suche gerade bei fremdsprachiger Literatur bei. Ist nachgewiesen, dass Nutzer wenig direkt über die Schlagwortsuche recherchieren, fehlt der Nachweis, inwieweit erst die Integration der Schlagwörter in den all-Index die relevanten Treffer zu Tage fördert. Dies wäre ein spannendes Forschungsfeld bspw. im Rahmen einer Referendarsabschlussarbeit. Herr Rosenke weist darauf hin, dass es technisch durchaus möglich wäre, umgekehrt die Schlagwörter aus der searchspecs.yaml zu nehmen, um hier testweise und vergleichend den Einfluss testen zu können. Generell wird der Wunsch nach mehr empirischen Nachweisen zur Auswirkung von Sacherschließung auf relevante Treffer hervorgehoben. Kurz gestreift wird das „Modell Utrecht“, d. h. die Abschaffung eines eigenen Bibliothekskatalogs. Einigkeit besteht darüber, dass Bibliotheken die Kompetenz und die Verantwortung auf diesem Feld nicht grundsätzlich abgeben dürfen, da nicht ohne Weiteres davon ausgegangen werden kann, dass ein Unternehmen wie Google diese Ausgabe dauerhaft und verlässlich wahrnehmen kann.

Frau Mödden berichtet davon, dass auch Google die gesamte GND abgezogen hat und betont den Mehrwert v. a. der Synonyme und semantischen Beziehungen. Die Nutzungszusammenhänge der GND bzw. von Normdaten allgemein müssten noch viel stärker bekannt gemacht werden. Auch Firmen interessieren sich für den Einsatz von Normdaten in ihre Datenbestände. Frau Bernhardt berichtet, dass auch die neuen FIDs großes Interesse an Normdaten haben und betont die Perspektiven, die sich aus den Verknüpfungsmöglichkeiten ergeben. Beklagt wird die Hürde von GND-Neuansetzungen außerhalb der Bibliotheken. Eine stärkere spartenübergreifende Öffnung der GND wird insbesondere auch von Anwendern aus Wissenschaft und Forschung gewünscht.

Frau Lüdemann greift die Frage nach dem Sinn von Sacherschließung auf und verweist darauf, dass auch kommerzielle Datenbankanbieter und der Online-Buchhandel Schlagwörter und Sacherschließungen verwenden. Dies lässt im Umkehrschluss durchaus auf die Nützlichkeit und Sinnhaftigkeit von Sacherschließungsdaten schließen. Die Diskussion sollte daher eher in Richtung einer besseren intuitiven Nutzung gehen.

Themenblock 2: Automatisierte Sacherschließung

Elisabeth Mödden (DNB Frankfurt/Main) spricht in ihrem Impulsreferat über „Maschinelle Klassifikation und Beschlagwortung der Reihe O in der Deutschen Nationalbibliothek Einblicke in die Praxis“ [Folien siehe Anhang]. Die komplette Reihe O wird nicht mehr bibliothekarisch bearbeitet, was eine nicht-RDA-gerechte Formalschließung einschließt. Für deutschsprachige Online-Hochschulschriften erfolgt seit April 2014 eine maschinelle Beschlagwortung mit dem Vokabular der GND. Die manuelle stichprobenhafte Fehleranalyse lässt erkennen, dass ca. 20% als falsch eingestuft werden. Die richtige Vergabe schwankt je nach Fachgebiet enorm. Dies liegt u. a. daran, dass je nach Fach mehr oder weniger eindeutiges Material vorhanden ist. So klappt die automatische Vergabe bei Medizin auch deshalb so gut, weil ca. 50% des Trainingsmaterials aus medizinischen Dissertationen stammt. Besonders betont Frau Mödden, dass sie lieber kein Schlagwort als ein falsches vergeben. Außerdem erfolgt eine gesonderte Kennzeichnung, ob ein Schlagwort maschinell vergeben wurde oder nicht, da diese Meta-Informationen für weitere Analysen genutzt werden.

Zukünftig möchte die DNB dieses Verfahren auf eBooks und Artikel erweitern. Für diese sollen Inhaltsverzeichnisse und Abstracts zur Gewinnung maschinell erzeugter Schlagwörter herangezogen werden. Des Weiteren soll eine Cross-Konkordanz zwischen GND und subject headings der Library of Congress entwickelt werden, um englischsprachige Publikationen ebenfalls auf diese Art und Weise erschließen zu können. Ebenso ist ein Vorschlagstool zur GND-Pflege in Planung.

Auf Nachfrage erklärt die Referentin, dass in dem Projekt PETRUS die Software von vier Firmen getestet wurde und man sich daraufhin für die Averbis-Software [siehe <https://averbis.com>] entschieden habe.

Frau Bernhardt betont, dass eine Öffnung der GND auf englische (Fach-)Begriffe und die Möglichkeit Fachthesauri aufzunehmen, eine große Bereicherung darstellen würde.

Themenblock 3: Normdaten und Forschungsdaten

Im letzten Impulsreferat sprechen Aline Deicke und Dominik Kasper (Akademie der Wissenschaften Mainz) zu „Der GND-Beaconizer in Online-Applikationen der ADW Mainz“ [siehe auch: <https://typo3.org/extensions/repository/view/beaconizer> (18.07.2016)]. Auch im Akademiebereich, wo man sich mehr im Forschungs- als im Bibliothekskontext bewegt, haben verschiedene Vorhaben mit Normdaten zu tun. Meist handelt es sich um Personennormdaten. Beacon als Dateiformat wird derzeit v. a. dafür genutzt, Personen mittels ihrer GND-Nummer zu identifizieren und verschiedene Inhalte, die als gemeinsames Element die GND-Nummer haben, zu verlinken. [siehe <https://de.wikipedia.org/wiki/Wikipedia:BEACON> (18.07.2016)] Beispielhaft wird dies anhand des Mainzer Professorenkatalogs gezeigt. [siehe <http://gutenberg-biographics.uni-mainz.de> (18.07.2016)]

Ein Vorhaben, das für den Einsatz eines Tools wie den Beaconizer geeignet wäre, wären die Senckenberg-Tagebücher, die zahlreiche Querverweise zu Personen enthalten. Ebenso sind solche Linkverknüpfungen losgelöst von Personen v. a. im Bereich der Geographika ein großes Thema in den Digital Humanities. Hier schlägt jedoch geonames.org derzeit die GND durch die Einbindung von Karten als Visualisierungstool.

Nochmals wird festgestellt, dass die Verwendung von Normdaten in Datenbeständen vor allem geisteswissenschaftlicher Forschungsprojekte und in FID-Projekten immer wichtiger wird. So nennt Frau Bernhardt das Vorhaben eines Online-Theaterlexikons Deutschland analog zum Theaterlexikon Schweiz. Ein Personenabgleich hat bereits stattgefunden und sukzessive von der hessischen Verbundredaktion Sacherschließung in die GND eingearbeitet. Eine Erweiterung auf Körperschaften wird von der Community des FID Darstellende Kunst gewünscht.

Abschlussdiskussion

In einer abschließenden Diskussion werden die Ergebnisse und Vorschläge zu den vorangegangenen Themenblöcken zusammengefasst:

1. Es wird der Wunsch nach mehr Empirie formuliert. Statistische Auswertungen und auch qualitative Auswertungen von Nutzerverhalten

können dazu dienen, sinnvolle Schwerpunkte in der Sacherschließung zu setzen. Gleichzeitig wird darauf hingewiesen, dass die Auswertung solcher Daten oder spätere Rückschlüsse daraus immer eine Frage der strategischen Zielrichtung und der damit einhergehenden Interpretation ist. Dennoch ist es wichtig, mehr über die Recherchewege der Nutzer zu wissen und darüber, wie Nutzer an ihre Treffer kommen. Die Rolle der Sacherschließung für die „precision“ sollte ebenfalls genauer angeschaut werden.

2. Es soll einen zweiten Workshop geben, der das Thema „Sacherschließung und Nutzerverhalten“ ins Zentrum rückt, da diese Aspekte gerade in ihrem Zusammenspiel äußerst komplex sind.
3. Die Referenten der Akademie der Wissenschaften Mainz betonen die vielversprechenden Perspektiven, die mit der Normdatennutzung in wissenschaftlichen Datensammlungen verbunden sind und den hohen Stellenwert der GND als ein wichtiger Standard für die Digital Humanities.
4. Die Möglichkeiten der automatisierten Sacherschließung werden sehr interessiert zur Kenntnis genommen und grundsätzlich begrüßt. Kritisch wird jedoch die Frage aufgeworfen, ob wir uns mit einem Ergebnis „80% einigermaßen OK“ abfinden müssen.
5. Als letzter spannender Punkt wurde überlegt, inwiefern die maschinelle Sacherschließung bspw. für Retrodaten genutzt werden kann und ob das derzeit erprobte Verfahren an der DNB auch für andere Bibliotheken geöffnet werden kann. Leider war es nicht mehr möglich, Frau Mödden direkt zu fragen. Ebenso wurde die automatische Sacherschließung für die Online-Publikationen auf unseren Publikationsservern und die Frankfurter Katalogkarten als lohnende Aufgabe gesehen. Wichtig wäre hierfür u. a. eine Konkordanz Eppelsheimer - RVK.

Die Referendare regen eine Nutzerbeteiligung mittels Crowdsourcing an, wie sie bspw. in dem Projekt Transcribe Bentham [siehe <http://blogs.ucl.ac.uk/transcribe-bentham> (18.07.2016)] erfolgreich umgesetzt wurde. Als Plattform könnte Wikimedia genutzt werden. Allerdings befürchten sie, dass die Motivation, alte Katalogkarten zu transkribieren, sich eher in Grenzen halten wird. Nichtsdestotrotz könnte man einen Versuch starten.

6. Im Anschluss an den Workshop sollen in der AG folgende Aspekte und Fragen weiter verfolgt werden:

- Auflistung weiterer Anwendungsszenarien für automatische Sacherschließung
- RDA-Evaluierung: Was ist von einem neuen Erschließungskonzept der DNB zu erwarten? Was müssen die Verbünde ggf. an Erschließungsleistung auffangen? (Thema in der Herbstsitzung der AG)
- Betrachtung der Option, Daten zu kaufen. Wer aus der Gruppe erklärt sich bereit, dieses Thema eingehender vorzubereiten?
- Wie können vorhandene Sacherschließungsdaten besser genutzt werden? Wie können diese besser an den Nutzer gebracht werden? Welches Potential ist noch unentdeckt?
- Wollen wir Schwerpunkte setzen? Wo erreichen wir mit unserer Sacherschließung signifikante Effekte und Mehrwerte?

Frau Müller bedankt sich bei allen Teilnehmerinnen und Teilnehmern des Workshops für die anregenden Diskussionen und die interessanten Beiträge.