
**2. Sitzung der Kommission für Erschließung und Metadaten
Sitzungsperiode 2019/2021 - Protokollentwurf**

Datum: 5. September 2019
Ort: Berlin, Universitätsbibliothek der FU Berlin
Zeit: 9:00 Uhr bis 15:10 Uhr

Teilnehmer:

Michael Beer	BSB	(Vorsitz)
Anne-Kathrin Brandau	UB FU Berlin	(Gast, Protokoll)
Petra Buchholz	UB FU Berlin	(Gast, ab TOP 7.2)
Michael Franke-Maier	UB FU Berlin	(Gast)
Kerstin Helbig	CMS ¹ HU Berlin	(Gast, ab TOP 7.2)
Werner Holbach	BSB	(Ständiger Gast)
Katja Koralewski	Bibliothek der Hochschule München	
Sonja Kümmer	UB München	
Dr. Mathias Kratzer	BVB-Verbundzentrale	(Ständiger Gast)
Manfred Müller	BSB	
Dr. Carolin Odebrecht	CMS HU Berlin	(Gast, ab TOP 7.2)
Ute Ristau	UB HU Berlin	(Gast)
Beate Rusch	KOBV-Verbundzentrale	(Ständiger Gast)
Birgit Schlegel	UB FU Berlin	(Gast, ab TOP 7.2)
Arnost Stanzel	BSB	(Gast, ab TOP 7.2)
Viola Taylor	UB FU Berlin	

Entschuldigt:

Alexandra Heiss	UB Frankfurt/Oder
Dr. Ulrich Hohoff	UB Augsburg
Dr. Rüdiger Hoyer	Bibl. des Zentralinst. für Kunstgeschichte (Ständiger Gast)

Tagesordnung

1. Begrüßung, Organisatorisches und Feststellen der Tagesordnung
2. Digitaler Assistent DA-3 – Bericht von den Tests
3. RDA-Toolkit Beta-Version – Stand der Vorbereitungen
4. Verbundzentrale
 - 4.1. LoC-Titeldaten
 - 4.2. ADAM-Anreicherungen mit ToCs der DNB
 - 4.3. B3Kat-Aktionenliste
5. Bericht aus der AGFE
6. Bericht aus der AGSE
7. Erschließung und Nachweis von Forschungsdaten
 - 7.1. Einführung
 - 7.2. Impulsvorträge aus der Forschungsdatenpraxis
 - 7.3. Diskussion

¹ CMS = Computer- und Medienservice

8. Anhang

1. Begrüßung, Organisatorisches und Feststellen der Tagesordnung

Herr Beer begrüßt die Teilnehmer der Sitzung. Er wird für die nächste Sitzung eine Terminumfrage für den Zeitraum 13.-16. Januar 2020 starten.

Die Tagesordnung wird in der vorliegenden Form genehmigt.

2. Digitaler Assistent DA-3 – Bericht von den Tests

Herr Holbach stellt den Bericht über die Tests zum Digitalen Assistenten DA-3 vor. Das Produkt der Schweizer Firma Eurospider Technology ist schon seit einiger Zeit im SWB im Einsatz und der OBV hat das Produkt vor kurzem lizenziert.

Der Verbundrat beauftragte die AGSE, sich mit dem Produkt auseinanderzusetzen und einen Bericht zu erstellen.

Herr Holbach beschreibt kurz die Vorgehensweise und die Testerfahrungen der Teilnehmer aus BSB, UBs, Fachhochschulen und Spezialbibliotheken.

Die Funktionalität des Tools wurde im Allgemeinen als positiv bewertet, vor allem die Choose & Click-Methode, die Antwortzeiten, die guten Hilfeseiten und dass das Produkt webbasiert ist.

Die qualitative Einschätzung wurde eher ambivalent betrachtet. Die Stärken des Assistenten liegen bei allgemeinen Titeln, die sehr häufig in den verschiedenen Bibliotheken erworben werden. Bei Titeln mit anderen Sprachen bzw. Spezialliteratur war die Qualität nicht so gut. Das Fehlen von verknüpften Bibliografien und Katalogen aus anderen Ländern (z.B. SUDOC) wurde bemängelt. Zudem wurden Vorschläge aus Konkordanzen als zu weit, allgemein und unspezifisch angesehen.

Aber auch im Fall, dass keine Schlagwörter angeboten werden, kann man das Tool gut verwenden, da sowohl die RVK, wenn auch in abgespeckter Form, und die GND angeboten werden und nutzbar sind.

Fazit des Berichts: Das Tool wird von den Kolleg*innen, die nur wenig in Aleph arbeiten und wenig Erfahrung mit Sacherschließung haben, als sinnvoll erachtet. Experten hingegen sehen für sich in der Regel keine Zeitersparnis.

Die Firma sollte aber noch Nacharbeiten wie das Einbinden neuer Datenquellen vornehmen. Eine Schnittstelle zur GND zum Anlegen neuer Sätze ist angedacht.

Ein Einsatz des Assistenten wird nur dann als sinnvoll erachtet, wenn es eine Schnittstelle zum Verbund gibt. Der OBV arbeitet derzeit an einer Alma-Schnittstelle.

Empfehlung im Bericht: Die AGSE empfiehlt die Einführung im Rahmen einer Pilotphase von 1-2 Jahren und die Einrichtung einer Projektstelle zur Begleitung und Evaluation des Projekts.

Laut Herrn Kratzer und Herrn Brecheisen ist die Einrichtung einer Schnittstelle zu Aleph möglich, würde jedoch auf dem aktuell verfügbaren Informationsstand geschätzt ca. sechs Personenmonate in Anspruch nehmen. Die einzige Methode zum Reimport angereicherter Titelaufnahmen, die mit Aleph funktionieren würde, wäre das „Polling“ von Daten aus dem DA-3.

Unbekannt ist derzeit noch, ob es sich bei den von DA-3 abzuholenden Daten nur um die vorgenommenen Ergänzungen handelt oder um komplette Datensätze. Je nachdem kann sich der Reimport, bei dem in Aleph schon vorhandene Felder nicht überschrieben werden sollen, unterschiedlich kompliziert gestalten. Eventuell ist sogar die Beauftragung von Ex Libris notwendig.

Bei der anschließenden Diskussion stellt sich heraus, dass die Finanzierung ein Problem sein könnte. Frau Rusch bittet deswegen darum, noch einmal mit dem KOBV gesondert zu sprechen (Finanzierungsvorbehalt).

Der Bericht wird von den Teilnehmern als sehr gut bewertet. Es wird beschlossen, dass der Vorsitzende der KEM, Herr Beer, diesen mit einem positiven Votum an den Vorstand/Vorsitzenden des Verbundrates weiterleitet. Die AGSE wird gebeten, zuvor noch die Zusammenfassung und das Fazit zu überarbeiten und ein etwa einseitiges Management-Summary an den Anfang des Berichts zu stellen.

Folgende Punkte sollen noch einmal herausgearbeitet werden:

- Für kleinere Bibliotheken ist DA-3 von Vorteil / Für Experten der Sacherschließung aber nicht
- Welche Kosten treten auf? Wie hoch ist der Aufwand?
 - Projektstelle
 - Lizenzierung und Entwicklung
 - Verbundzentrale und Ex Libris
- Dauer des Projekts
- Zeitpunkt, an dem begonnen werden soll
- Wo liegt die Rolle der AGSE?
- Was sind die Vorteile und Nachteile des Tools?

3. RDA-Toolkit

Herr Beer berichtet, dass in der FG-Erschließung vier Beispiele (Einverfasserschrift, Aufsatzband/Zusammenstellung, Zeitschrift, Normdaten) exemplarisch benutzt wurden, um konkrete Ressourcen zu katalogisieren. Themenschwerpunkte sind:

Sprache/Verständlichkeit, Konzepte, Navigation, Recording Methods und wie/ob sich das Standardelemente-Set im neuen Toolkit anwenden lässt.

Kurz gefasst ist zu sagen, dass es oft sehr lange dauert um an die richtige Stelle zu gelangen und die Sprache häufig schwer verständlich ist, so dass „normale Katalogisierer“ sicher nicht nur mit dem Toolkit arbeiten können. Deshalb müssen andere Lösungen erarbeitet werden, damit die Endanwender gut nach dem Regelwerk katalogisieren können. Der entstehende Aufwand ist zeit- und ressourcenmäßig momentan nicht abschätzbar. Auch eine Alternative ist nicht in Sicht.

Schulungsunterlagen sind kaum nachnutzbar und müssen, auch weil es sich ja um Umsteigerschulungen handelt, neugestaltet werden. Die Zuordnung der DACH-AWRs zu den korrekten Stellen wurde schon getestet. Oft gelingt es, es ist aber nicht immer möglich.

4. Verbundzentrale

4.1. LoC-Titeldaten

Herr Kratzer berichtet, dass die BVB-Verbundzentrale seit Herbst 2018 keinen Zugriff mehr auf den FTP-Server der LoC hat. Auf Anfragen dazu wurde von Seiten der LoC zunächst nicht reagiert. Nur über Herrn Heuvelmann von der DNB ist es schließlich gelungen, Kontakt herzustellen. Dabei wurde mitgeteilt, dass das Formular zur Bestellung des Cataloging Distribution Service noch einmal ausgefüllt und die zugehörige Zahlungsaufforderung abgewartet werden soll. Aufgrund fehlender Rechnungen wurde an die LoC auch keine Zahlungen geleistet und der FTP-Zugang ist deshalb gesperrt worden.

Es stellt sich die Frage, ob der B3Kat die aktuellsten Daten der LoC braucht. Zumal es auf der Webseite <https://www.loc.gov/cds/mdsftp.html> einen Copyrighthinweis gibt, demzufolge Titelaufnahmen der LoC außerhalb der USA nur eingeschränkt weitergegeben werden dürfen. Open Data sind derzeit anscheinend nur die bis 2016 erzeugten Titeldaten.

Herr Kratzer hat der LoC den im B3Kat klassischen Use-Case der Fremddatennachnutzung und Veröffentlichung unter CC0 beschrieben, aber daraufhin noch keine Rückmeldung erhalten.

Herr Beer erklärt, dass für die Katalogisierung die Titeldaten eher aus den deutschen Verbänden und dem Worldcat übernommen werden und dass es bisher keine Beschwerden gibt. Es könnte aber zu Mehrarbeit bei der Bestellautomatisierung kommen. Die Aufnahmen der Buchhändler sind oft unvollständig, so dass sie, wenn möglich, schon bei der Einspeicherung durch die besseren Aufnahmen der LoC ersetzt werden. Es wird beschlossen, dass abgewartet wird, wie die Antwort der LoC auf den Use-Case ausfällt. Ist sie positiv, spricht nichts gegen eine Nutzung der LoC-Daten. Fällt sie negativ aus, muss das Thema noch einmal in der KEM besprochen werden.

4.2. ADAM-Anreicherungen mit ToCs der DNB

Teilweise wurden Inhaltsverzeichnisse der DNB falschen Titelaufnahmen zugeordnet. In vielen Fällen war der Grund dafür, dass die OCLC- und DNB-IDs bis Ende 2018 im gleichen Index gespeichert wurden und sich die entsprechenden Nummernkreise begonnen haben zu überschneiden. Deshalb wurde die Anreicherung der DNB-ToCs vorübergehend ausgesetzt. Seit Ende Januar 2019 werden nun alle IDs in jeweils einem eigenen Index indexiert.

Trotzdem werden weiterhin aus noch nicht abschließend geklärter Ursache DNB-ToCs sporadisch falschen Titeln zugeordnet. Das Debugging der relativ komplexen bestehenden Zuordnungsroutine stellt jedoch einen weitaus höheren Aufwand dar als die komplette Neuimplementierung eines vereinfachten Zuordnungsworkflows, bei dem das Matching nunmehr über die DNB-ID oder die ISBN und nicht länger auch über eine ggf. übereinstimmende Titelansetzung erfolgen würde.

Herr Beer ergänzt, dass die nach wie vor zu beobachtenden falschen Anreicherungen in der Regel auf zufällig übereinstimmende Titelansetzungen zurückzuführen sind. Titelaufnahmen, die weder von der DNB übernommen wurden noch eine ISBN besitzen, würden durch den vereinfachten Zuordnungsworkflow nicht mehr angereichert werden können. Für Hochschulschriften, graue Literatur oder Amtsschriften bereitet das aber vermutlich wenig Probleme. Sie liegen entweder sowieso nur als Online-Ausgabe vor oder werden von der DNB oft gar nicht mehr erschlossen.

Es wird beschlossen, dass der vereinfachte Workflow von der BVB-Verbundzentrale implementiert und ggf. zu einem späteren Zeitpunkt nach Bedarf erweitert werden soll.

4.3. B3Kat-Aktionenliste

K96 (Neuaufbau des B3Kat-OpenData-Repositorys) wurde storniert.

K98 (Maschinelle Ergänzung von VD17-Nummern in ASEQ 580_ und Dokumentation der Aktion in ASEQ 501) wurde erledigt.

K99 (Auflösung der Parkfelder ASEQ OSC bei ZDB-Aufnahmen mit Originalschrift, die vor Oktober 2018 importiert wurden) musste noch einmal aufgenommen werden, da die Aktion auch auf BVB03 noch durchgeführt werden muss

K90 (Anreicherungen von Titelaufnahmen von Büchern in kyrillischer und griechischer Schrift mit OS-Feldern durch Retransliteration)

Frau Taylor erkundigt sich nach dem Stand der Tests und ihrer diesbezüglichen Anfrage an Herrn Brecheisen. Herr Kratzer wird Herrn Brecheisen bitten, sich dazu mit Frau Taylor in Kontakt zu setzen.

5. Bericht aus AGFE

Das Protokoll der 1. Sitzung ist veröffentlicht.

Thema waren u.a. die GND Namenssätze, mit denen nicht mehr verknüpft werden soll. Die Satzschablone für Namenssätze wurde entfernt. Es ist geplant, die Namenssätze auch aus dem Index zu entfernen, damit mit ihnen nicht versehentlich verknüpft werden kann. [Anm. nach der Sitzung: Diese Aktion ist inzwischen gelaufen]

Die drei Punkte, die bei Nicht-RDA-Aufnahmen in untergeordneten Sätzen teilweise anstelle des Titels der Reihe in ASEQ-Feld 451 stehen, sollen durch den Reihentitel ersetzt werden, da die Frontendsysteme oft den Titel aus der ÜG nicht anzeigen. Es soll die Möglichkeit geschaffen werden, geschätzte Erscheinungs- oder Entstehungsdaten in ASEQ 425 anzugeben. Die AGFE hat einen Vorschlag zur Umsetzung an die Verbundzentrale weitergegeben.

Es wurde ein [Mindeststandard](#) für die Hochkatalogisierung von E-Books festgelegt, der bereits auf den KKB-Online-Seiten veröffentlicht wurde.

6. Bericht aus AGSE

Herr Holbach berichtet von der 14. Sitzung am 7. Juni 2019. Das Protokoll ist veröffentlicht. Es ging neben den Tests zum Digitalen Assistenten DA-3 auch um die Festlegung der zukünftigen Arbeitsschwerpunkte der AG, die in der nächsten Sitzung priorisiert werden sollen.

Herr Franke-Maier plant eine Umfrage zu dem Thema „Was ist in der RVK nicht gut bzw. gar nicht geregelt?“ Die „Sacherschließer“ sollen dabei u.a. mitteilen, ob es für alle Fälle Regelungen gibt oder ob Nacharbeiten in der RVK notwendig sind.

Es wird daran gearbeitet, die falschcodierten Forms Schlagwörter zu bereinigen. Dabei fällt immer wieder auf, dass die Forms Schlagwörter in der ZDB nicht korrigiert werden. Herr Franke-Maier hat sich deswegen an Herrn Lohrum vom KOBV gewendet um das Thema in die AG-KVA zu tragen und um auf diesen Weg auf die ZDB einzuwirken.

Herr Baumgartner (Gast in AGSE) hat die beiden neuen ASEQ-Felder 950 und 951 erläutert. Er hat den Vorschlag geäußert, dass in den neuen Feldern auch die Daten aus ASEQ 700, 711, 720 und 740 umgesetzt werden. Sie bieten zusätzliche Möglichkeiten, z.B. bei der Hinterlegung von Provenienzen. Das Thema muss noch abschließend besprochen werden.

7. Erschließung und Nachweis von Forschungsdaten

7.1. Einführung

Frau Kümmerer präsentiert eine Einführung in das Thema Forschungsdatenmanagement. (s. [Anhang](#))

7.2. Impulsvorträge aus der Forschungsdatenpraxis

Herr Stanzel berichtet über den geplanten Aufbau eines Forschungsdatendienstes für die Ost-, Ostmittel- und Südosteuropaforschung (OstData) an der Bayerischen Staatsbibliothek sowie vier Partnerinstitutionen. Die Partner sind das Collegium Carolinum – Forschungsinstitut für die Geschichte Tschechiens und der Slowakei (CC), das Leibniz-Institut für Geschichte und Kultur des östlichen Europas (GWZO), das Herder-Institut für historische Ostmitteleuropaforschung (HI) und das Leibniz-Institut für Ost- und Südosteuropaforschung (IOS). Für die Arbeit der KEM relevant ist zum einen das Arbeitspaket zur Entwicklung eines OstData-Metadaten-Schemas. Dieses soll die adäquate Beschreibung sozial- und geisteswissenschaftlicher Forschungsdaten aus der interdisziplinären Ost-, Ostmittel- und Südosteuropaforschung erlauben und an das Material angepasste Suchen ermöglichen. Zum anderen wird das einzurichtende OstData-Forschungsdatenrepositorium auf dem Dokumentenserver des ZEP an der BSB basieren, welcher für die Bereitstellung von digitalen Objekten eine BV-Nummer zur Voraussetzung hat. Hierfür wird ein Mapping vom OstData-Schema zum B3Kat von Nöten sein, wobei nicht alle Felder eine Entsprechung benötigen werden.

Frau Schlegel und Frau Buchholz stellen das Forschungsdatenmanagement an der FU Berlin vor. (s. [Anhang](#))

Frau Helbig präsentiert das Forschungsdatenmanagement an der HU Berlin. (s. [Anhang](#))

Frau Kümmer stellt Zwischenergebnisse aus dem Forschungsdaten-Modellprojekt „e-Humanities – interdisziplinär“ (<https://www.fdm-bayern.org/ehumanities-interdisziplinär/>) vor, dass u.a. an der UB der LMU angesiedelt ist und sich intensiv mit der Erschließung von Forschungsdaten beschäftigt. (s. [Anhang](#))

7.3. Diskussion

Die sich an die Impulsvorträge anschließende rege Diskussion dreht sich hauptsächlich um die Frage „Sollen Forschungsdaten auch im Bibliotheks- und/oder im Verbundkatalog oder einem Verbundkatalog speziell für Forschungsdaten nachgewiesen werden?“. Das Für und Wider wird überwiegend anhand eines konkreten Beispiels, des von Frau Kümmer bearbeiteten Pilotprojekts „VerbaAlpina“ (<https://www.verba-alpina.gwi.uni-muenchen.de/>), diskutiert. Charakteristisch für die Erschließung von VerbaAlpina sind zum einen die Granularität – Es wird nicht nur das Projekt als Ganzes, sondern auch jeder einzelne Datensatz aus VerbaAlpina mit Metadaten versehen –, und zum anderen die hierarchischen Beziehungen zwischen den Informationseinheiten.

Im Folgenden sollen einige Aussagen, Gedankengänge und Fragestellungen aus dem Diskussionsverlauf wiedergegeben werden:

- Ein verbundweiter, zentraler Nachweis von Forschungsdaten wird von Allen als wünschenswert angesehen.
- Die Diskussionsteilnehmer stimmen darin überein, dass ein Nachweis der Forschungsdaten im Verbundkatalog nicht die projektspezifische Präsentation der Daten ersetzen kann, v.a. auch im Hinblick auf kartographische Darstellungsformen o.ä.
- Laut Aussage von Herrn Kratzer, ist die Aufnahme einer größeren Menge von Datensätzen (im Falle von VerbaAlpina mehrere Zehntausende) in den Verbundkatalog möglich; Problematisch könnte jedoch die Abbildung der zwischen den Datensätzen bestehenden Relationen sein.
- Herr Kratzer gibt außerdem zu bedenken, dass speziell Metadatenätze zu Forschungsdaten auf sehr feingranularer Ebene nur nach Anreicherung von Information aus grobgranularer Ebene, also durch massive Erhöhung der Redundanz eine Chance haben würde, die Plausibilitätskontrollen des B3Kat zu passieren. Frau Kümmer entgegnet hierauf, dass sich die Redundanzen auf formale Kriterien, wie z.B. creator, contributor, publisher, resourceType, language usw., beschränken. Die inhaltsbeschreibenden Aspekte, d.h. title, description, subjects usw., sind im Falle von VerbaAlpina von Datensatz zu Datensatz verschieden. Inwieweit ein sehr starker Anstieg von Einträgen für diese Entitäten jedoch zu Problemen bei der Recherche führen konnte wurde in der Diskussion angesprochen, aber nicht abschließende geklärt.
- Frau Kümmer und Herr Duschner (UB LMU) haben im Rahmen des eHumanities-Projekts einen ersten Mapping-Entwurf DataCite -> MARC erarbeitet, der belegt, dass die beiden Formate grundsätzlich gut kompatibel sind.
- Sollte die Entscheidung zugunsten einer Aufnahme von Forschungsdaten-Metadaten in den Verbundkatalog fallen, wäre eine Verknüpfung zwischen den bereits im Katalog nachgewiesenen Publikationen und den ihnen zugrundeliegenden Forschungsdaten wünschenswert.
- Die Aufnahme von Forschungsdaten-Metadaten in den Verbundkatalog würde von vielen WissenschaftlerInnen sehr begrüßt werden, zum einen weil sie mit einem Katalognachweis Renommee verbinden und zum anderen weil sie die Dauerhaftigkeit von bibliothekarischen Diensten – im Vergleich zu projektfinanzierten Angeboten – schätzen.
- Herr Stanzel wünscht, dass sein Projekt in Form eines reduzierten Metadatenschemas im Verbundkatalog erschlossen wird, um die Daten auffindbar zu machen. Im Repository wäre dann eine spezifischere Abbildung möglich.

Fragen:

- Wichtigste Frage: Welche Rolle hat der Verbundkatalog?

- Müssen die Daten wirklich im Verbundkatalog gespeichert werden, von wo aus sie dann in die einzelnen Lokalsysteme geliefert werden? Oder genügt eine Einbindung der Daten in das Gateway Bayern?
- Geht es um eine kooperative Erschließung oder nur um den Nachweis?
- Was soll im Verbund möglich sein?
 - Beschreibung nur auf Projektebene (-> Sucheinstiege)?
 - Granulare Beschreibung, d.h. zusätzliche Erschließung von Informationseinheiten unterhalb der Projektebene?
 - Granulare Beschreibung mitsamt Relationen zwischen den Informationseinheiten?
- Ist eine domänenspezifische Erschließung und Präsentation der Daten im Verbundkatalog möglich / wünschenswert?
- Sollten die Daten im Verbundkatalog zusätzlich durch Taxonomien erschlossen werden?

Aufgrund der Komplexität der Thematik, wird beschlossen, eine temporäre Arbeitsgruppe für Forschungsdaten einzurichten, die die aufgeworfenen Fragen weiter diskutiert und der KEM Bericht erstattet. Herr Beer und Frau Kümmer werden mögliche AG-Teilnehmer kontaktieren und sich erkundigen, ob sie an einer Teilnahme interessiert sind. Die erste Aufgabe dieser neuen AG Forschungsdaten ist die Erstellung eines ersten Berichtes zur standardisierten Erschließung von Forschungsdaten bis zur nächsten KEM-Sitzung im Januar 2020. In dem Bericht soll u.a. festgehalten werden, was erschlossen wird und wie tief die Erschließung gehen soll.

Anhang (Links zu den Präsentationen)

- a) [Präsentation 'Forschungsdatenmanagement \(FDM\)' von Sonja Kümmer](#)

- b) [Präsentation 'Forschungsdatenpraxis an der Freien Universität Berlin' von Petra Buchholz und Birgit Schlegel](#)

- c) [Präsentation 'Erschließung von Forschungsdaten an der HU Berlin' von Niels Fromm und Kerstin Helbig](#)

- d) [Präsentation 'Forschungsdaten-"Praxis" an der LMU München' von Sonja Kümmer](#)