

**9. Sitzung der Kommission für Erschließung und Metadaten
Sitzungsperiode 2019/2021 - Protokoll**

Datum: 5. Oktober 2021
Ort: Videokonferenz
Zeit: 12:30 Uhr bis 15:00 Uhr

Teilnehmer:

Michael Beer	BSB	(Vorsitz)
Nicole Heidingsfelder	KOBV-Verbundzentrale	(Ständiger Gast)
Alexandra Heiss	UB Frankfurt (Oder)	(Protokoll)
Dr. Ulrich Hohoff	UB Augsburg	
Werner Holbach	BSB	(Ständiger Gast)
Katja Koralewski	Bibliothek der Hochschule München	
Manfred Müller	BSB	
Viola Taylor	UB FU Berlin	
Peter Thiessen	BVB-Verbundzentrale	(Ständiger Gast)

Entschuldigt:

Dr. Rüdiger Hoyer	Bibl. des Zentralinst. für Kunstgeschichte (Ständiger Gast)
Sonja Kümmer	UB LMU München

Tagesordnung

1. Begrüßung, Organisatorisches und Feststellen der Tagesordnung
2. Kennzeichnung von „last copies“ im B3Kat
3. Bericht der AG Metadaten für Forschungsdaten
4. Verbundzentrale
 - 4.1. Aktionenliste
 - 4.2. Einspeicherung von unselbständigen Titelaufnahmen der MGH
5. Bericht aus der AGFE
6. Bericht aus der AGSE
7. Sonstiges
 - 7.1. Verlängerung der Amtszeit der KEM
 - 7.2. Entzug des Doktorgrades (neu)

1. Begrüßung, Organisatorisches und Feststellen der Tagesordnung

Herr Beer begrüßt die Teilnehmer der Sitzung. Frau Heidingsfelder aus der KOBV-Verbundzentrale stellt sich vor. Sie tritt die Nachfolge von Frau Rusch als ständiger Gast in der KEM an.

Auf Wunsch von Frau Taylor wird die Tagesordnung um Punkt 7.2. Entzug des Doktorgrades ergänzt.

2. Kennzeichnung von „last copies“ im B3Kat

Das von Herrn Beer und Herrn Dr. Hohoff entworfene Papier zur Kennzeichnung von „last copies“ im B3Kat (s.a. [KEM-Protokoll der 8. Sitzung, TOP 2](#)) wurde dem BVB-Verbundrat zugeleitet und auf der in Woche 39 abgehaltenen Verbundratssitzung kurz besprochen. Da die Kennzeichnung von „last copies“ im Zusammenhang mit Aussonderungen und Bestandserhaltungsmaßnahmen gesehen werden muss, wurde das Thema als Ganzes auf die nächste Sitzung des Verbundrates, die im März 2022 stattfinden wird, vertagt. Das Papier wird über Frau Heidingsfelder an die zuständigen Gremien im KOBV weitergeleitet.

3. Bericht der AG Metadaten für Forschungsdaten

Die AG Metadaten für Forschungsdaten hat sich, nachdem es eine grundsätzliche Zustimmung zum Nachweis von Forschungsdaten im B3Kat gegeben hat, mit der Datenverwaltung und ihrer technischen Umsetzung beschäftigt.

Da es unter den B3Kat-Bibliotheken unterschiedliche Bedürfnisse gibt, hat man sich darauf verständigt, zwei Varianten zur freien Wahl zu stellen:

Variante 1:

Die Primärererschließung findet im lokalen Repository statt, die Daten werden hieraus als Stufe 8-Sätze in den B3Kat exportiert.

Variante 2:

Die Primärererschließung findet im B3Kat statt. Die Daten können in regelmäßigen Routinen von den Bibliotheken selbständig über OAI aus Aleph geharvestet werden. Um die Daten sowohl im B3Kat als auch in den Repositorien vollständig abbilden zu können, muss ein Mapping zwischen MARC (Austauschformat des B3Kat) und DataCite (Standardformat für die Erschließung von Forschungsdaten) erstellt werden und ein Mindestdatenset (für DataCite und Aleph) festgelegt werden.

Als Identifikatoren sollen die DOIs genutzt werden.

Eine Anleitung zur Erfassung von Forschungsdaten soll als Ergebnis auch auf den KKB-Online-Seiten eingestellt werden. (Vollständiger Bericht s. Anlage)

4. Verbundzentrale

4.1. Aktionenliste

K84	B3Kat-interne Anreicherung von Titelaufnahmen mit Sacherschließungsdaten auf Basis aktueller CultureGraph-Bündel	Eine erste Anreicherung mit GND-Schlagwörtern und Formbegriffen ist im Juli/August erfolgt (s. TOP 6).
K86	Umwandlung aller RDA-"Hybridaufnahmen" des B3Kat=BVB01 in sog. "angereicherte Aufnahmen" (Setup-Teil)	In Arbeit
K90 K91	Anreicherungen von Titelaufnahmen von Büchern in kyrillischer und griechischer Schrift mit OS-Feldern durch Re transliteration, Entwicklung einer Fix-Routine	In Arbeit. Momentan werden Tests vorbereitet.
K93	Migration von ca. 42.000 RAK-Altbestandstiteln der Amtsbibliotheken Staatlicher Archive	Nachdem ein erster Dublettencheck nur eine Trefferquote von 40% ergeben hatte, wurde erneut und dieses mal nur auf ISBNs geprüft. Dieser Check ergab eine Trefferquote von 50–60%. Problem ist, dass viele Titel gar keine ISBNs besitzen. Bei der Einspielung sind daher einige Dubletten zu erwarten.
K94 K95	Modifikation des Setups der VD16-Library BVB05 in Richtung RDA Umsetzung des VD16 in Richtung RDA	In Arbeit
K97	Aufsetzen eines Verfahrens zur jährlichen Übernahme von in RISM neu erfassten oder geänderten Handschriftenaufnahmen	In Arbeit
K100	Anreicherung mit 280.378 RSWK-konformen Verschlagwortungen des alten Sachkataloges des Zentralinstituts für Kunstgeschichte	Erledigt
K101	Aufsetzen eines Verfahrens zur vierteljährlichen Übernahme neuer Titelaufnahmen in einem kunstgeschichtlich relevanten Teilbestand der UB Heidelberg aus dem K10plus	Ein Verfahren wurde aufgesetzt, begonnen wurde mit den Aufnahmen für Zeitschriftenaufsätze, die zur bibliothekarischen Begutachtung ins Testsystem geladen wurden.

K103	Soweit möglich, maschinelle Generierung von 596a in Aufsatzaufnahmen, die bislang nur 596_ enthalten	Eine erste Umsetzung von ca. 1 Mio. Datensätzen ist erfolgt. Die noch offenen ca. 1,3 Mio. Sätze werden sukzessive umgesetzt.
K105	Ermittlung der Tp6-Sätze in BVB18 mit denen mindestens ein Titel verknüpft ist und Sortierung dieser Sätze absteigend nach Anzahl damit verknüpfter Titel.	165.000 Tp6-Sätze mit B3Kat-Titelverknüpfungen wurden ermittelt, 120 von diesen Personensätzen sind mit 50 und mehr Titeln verknüpft. Die Liste wird begutachtet und es soll geprüft werden, ob die Liste kooperativ abgearbeitet werden kann. Frau Taylor hat bereits die Bereitschaft der FU erklärt, bei der Bearbeitung mitzuarbeiten.

4.2. Einspeicherung von unselbständigen Titelaufnahmen der MGH

Die Monumenta Germaniae Historica (MGH) bittet die KEM einer Einspielung von 230.000 selbsterstellten Aufsatzaufnahmen zuzustimmen und die Verbundzentrale damit zu beauftragen. Die Aufsätze stammen aus Zeitschriften und Sammelwerken, sind nach Autopsie selbst erstellt und liegen bereits im ASEQ-Format vor. Der Wunsch ist, dass die Einspielung noch vor der Migration der MGH nach Alma im Frühjahr 2022 stattfinden soll, daher ist der Zeitrahmen für die Verbundzentrale sehr eng. Da es erfahrungsgemäß recht aufwändig ist, Aufsatzdaten einzupflegen (es müssen hier u.a. immer zwei Dinge zueinanderfinden: Aufsatz und Quelle), sieht die Verbundzentrale die Komplett einspielung in dieser Kürze der Zeit nur dann als realisierbar an, wenn die gelieferten Daten vollständig und weitgehend fehlerfrei sind.

Beschluss der KEM: Der Auftrag zur Einspielung der Aufsatzdaten wird erteilt. Die Einspielung soll in Absprache mit der MGH im Januar/Februar 2022 erfolgen. Falls es für die Einspeicherung problematische Daten gibt, wird über das Vorgehen in diesen Fällen in der nächsten KEM-Sitzung im Januar entschieden.

5. Bericht aus der AGFE

- RDA-3R-Projekt: In der letzten Sitzung wurde über das RDA-Erschließungshandbuch gesprochen. Hier laufen die Arbeiten planmäßig.
Für die Erfassung müssen auch die ASEQ-Felder beachtet werden. Dafür wird es auch dieses Mal wieder eine Gruppe Datenformate geben. Da mittlerweile jedoch nicht mehr so viele Verbünde mit dem ASEQ-Format arbeiten wie bei der Umstellung von RAK auf RDA, wird diese Gruppe wahrscheinlich kleiner sein. Aus dem BVB gibt es Kolleg:innen, die sich für eine Mitarbeit in der Gruppe Datenformate bereit erklärt haben. Herr Beer wird diese Kolleg:innen melden.

Die Elementbeschreibungen sollen bis November fertig sein. Im Anschluss daran müssen die Texte zu Allgemeines und Ressourcentypen erstellt werden und die Fachgruppe Datenformate sollte mit ihrer Arbeit beginnen.

- Erfassung von mehrteiligen Aufsätzen: Für die AGFE wurde eine Anleitung für die Erfassung mehrteiliger Aufsätze im B3Kat erstellt. Sie orientiert sich an der Erschließung dieser Ressourcen in anderen DACH-Verbänden und wird auf den KKB-Online-Seiten veröffentlicht werden.
- Springer-Pakete: Auf die verteilte Bearbeitung der Metadaten von Springer-E-Book-Paketen wird in Zukunft verzichtet. Grund: Es wird immer schwerer Bibliotheken zu finden, die sich beteiligen, außerdem werden Sacherschließungsdaten durch die Culturegraph-Anreicherungen automatisch ergänzt.

6. Bericht aus der AGSE

- Seit der letzten KEM-Sitzung am 1. Juli 2021 fand keine Sitzung der AG Sacherschließung statt.
- Ende Juli/Anfang August erfolgte die in der letzten Sitzung angekündigte Anreicherung von einzelnen GND-Schlagwörtern und Schlagwortfolgen nach RSWK innerhalb des B3Kat-Bestandes (Phase 1 des Sacherschließungsanreicherungsprojekts mithilfe von Culturegraph-Bündeln).
Hier die Zahlen für die Phase 1 (vgl. auch Mail an BVB_INFO vom 23.08.2021):
 - Datensätze, bei denen Schlagwörter und/oder Schlagwortfolgen angereichert wurden: ca. 2,2 Millionen.
Davon sind:
 - Print-Datensätze: ca. 1,7 Millionen
 - E-Books: ca. 465.000
 - Andere: ca. 35.000.
 - Insgesamt wurden ca. 4,1 Millionen Schlagwörter oder Schlagwortfolgen angereichert.
 - Durchschnittlich angereicherte Folgen pro Datensatz: 1,86.
 - Angereicherte Datensätze, die vorher keine Schlagwörter oder Schlagwortfolgen hatten: ca. 1 Million.
 - Datensätze, bei denen Formangaben angereichert wurden: ca. 900.000.Die Rückmeldungen aus den B3Kat-Bibliotheken waren sehr positiv.
Die AGSE veranlasste nach der Aktion einige kleinere maschinelle Korrekturen bei Inkunabeln und Patentschriften.
- Die AG bereitet gegenwärtig insbesondere die in der nächsten Sitzung zu behandelnden Aktionen zur Migration der RVK- und sonstigen Notationen aus Feld ASEQ 700 nach ASEQ 701 durch Datenauswertungen sowie die Phase 2 der Sacherschließungsanreicherungsaktion (RVK-Notationen innerhalb des B3Kat) vor.
- Die nächste Sitzung der AGSE findet am 7. Oktober (als Videokonferenz) statt.

7. Sonstiges

7.1. Verlängerung der Amtszeit der KEM

Die Amtszeit der KEM, wie auch die aller anderen Gremien, wird um ein Jahr verlängert. Die nächste KEM-Sitzung findet am 19.01.2022 statt.

7.2. Entzug des Doktorgrades

Frau Taylor erläutert das Verfahren der FU zum Umgang mit Dissertationen, wenn der Doktorgrad entzogen wurde:

- Der Hochschulschriftenvermerk bleibt.
- In die Printausgabe wird ein Vermerk eingetragen.
- Im Repository wird ein interner Vermerk gemacht.
- Die DNB wird informiert.

Dieses Vorgehen ist kompatibel mit dem Beschluss der AGFE vom 12. November 2014, der von der KEM noch einmal ausdrücklich bestätigt wird. Für das Handling im B3Kat bleibt alles wie bisher: Die Titelaufnahme bleibt unverändert, der Hochschulschriftenvermerk wird nicht entfernt. Diese Aussage gilt für den Verbundkatalog. Den Bibliotheken bleibt überlassen, was in den lokalen Repositorien und Lokalsystemen gemacht wird ([s.a. AGFE-Protokoll zur 1. Sitzung der Amtszeit 2014/2016 vom 16.7.2014, TOP 5.b.](#)).

Anlage zu TOP 3.

AG Metadaten für Forschungsdaten Bericht an die KEM 2

Nach der grundsätzlichen Zustimmung des Verbundrats zum Nachweis von Forschungsdaten im B3Kat, hat sich die AG Metadaten für Forschungsdaten mit der konkreten Umsetzung beschäftigt. Festzulegen sind die Datenflüsse zwischen den Repositorien und dem Verbundkatalog incl. dem Entwurf eines Mappings zwischen DataCite und MARC und die Art der Erfassung im B3Kat.

1. Datenverwaltung

- a) Bei der Diskussion in der AG Metadaten für Forschungsdaten wurde klar, dass die Bedürfnisse hierfür nicht einheitlich sind. Es gibt Bibliotheken, die Metadaten primär

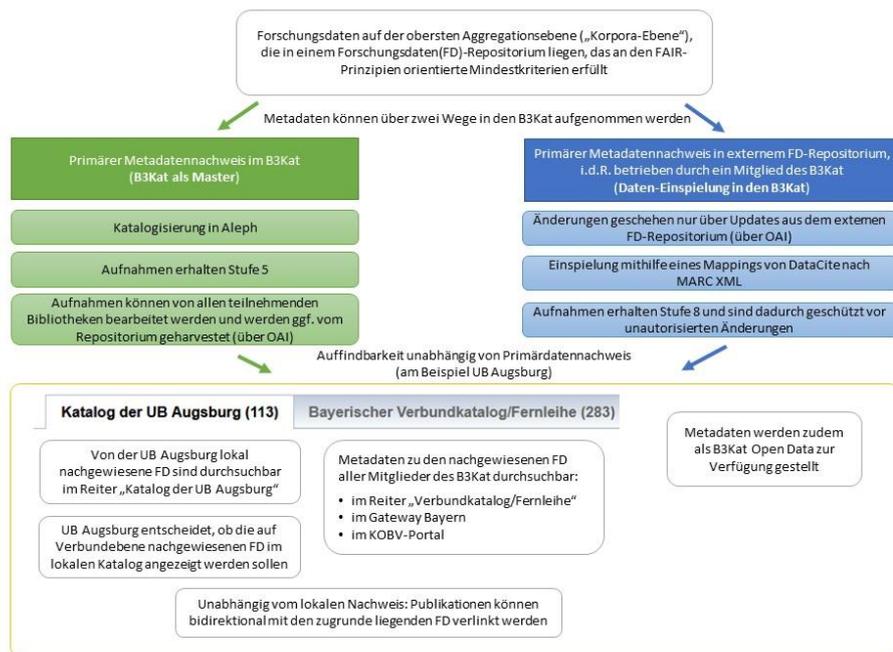
in den Repositorien erfassen und von dort in den B3Kat exportieren möchten, andere Bibliotheken wollen Forschungsdaten primär im B3Kat erschließen und von dort in die Repositorien exportieren.

b) Gründe für die Primärererschließung in den lokalen Repositorien

- i) Forschungsdaten werden z. T. schon seit über 10 Jahren in lokalen Forschungsdaten-Repositorien vorgehalten. Die Workflows für die Erschließung der Forschungsdaten und die Metadatenverwaltung sind entsprechend etabliert und lassen sich nicht ohne Weiteres auf eine Primärdatenerfassung im B3Kat umstellen. Eine weitere manuelle Datenpflege im B3Kat ist nicht zumutbar.
- ii) Die Kolleg*innen, die die Metadaten für Forschungsdaten kuratieren, sind nicht unbedingt mit den Regeln der Formalerschließung nach RDA und dem Arbeiten mit Aleph vertraut; dies wäre jedoch Voraussetzung für eine Primärdatenerfassung im B3Kat.
- iii) Forschungsdaten können spezifischer in Repositorien erschlossen werden. Manche Informationen, z.B. Angaben zur Förderung können zwar im B3Kat in ASEQ erfasst werden, aber nicht in einer so strukturierten Form, wie in den lokalen Repositorien.

c) Gründe für die Primärererschließung im B3Kat

- i) Bestimmte Forschungsdatenkorpora stehen in einem engen Verhältnis zu einer meist monografischen Veröffentlichung bzw. sind genau dafür erhoben (z.B. bei Dissertationen). Diese können unkompliziert im gleichen System verzeichnet und verknüpft werden.
 - ii) Die Katalogisierungsumgebung im B3Kat erlaubt die leichte Nutzung, Verknüpfung und Neuerstellung von Normdatensätzen (GND, RVK, usw.) ohne eine Arbeitsumgebung im Repositoryum dafür bereitstellen zu müssen.
 - iii) Bibliotheken erhalten die Möglichkeit, Metadaten zu Forschungsdaten der eigenen Einrichtung im B3Kat nachzuweisen auch wenn die Daten in einem Repositoryum liegen, dessen Metadaten nicht vom Verbund geharvestet werden.
- d) Da beide Modelle mit schon vorhandener Technik relativ unaufwändig durchzuführen sind, plädiert die AG Metadaten für Forschungsdaten dafür beide Modelle zu ermöglichen. Den Bibliotheken wäre es freigestellt, ob und nach welchem dieser beiden Modelle sie ihre Forschungsdaten im B3Kat nachweisen wollen.



2. Skizzierung der technischen Umsetzung

a) Primärschließung in den lokalen Repositorien

- i) Die Metadaten zu den in den lokalen Repositorien vorgehaltenen Forschungsdaten werden im DataCite-Format per OAI PMH von den Repositoriumsbetreibern bereitgestellt. Der Verbund harvestet die Daten, konvertiert sie nach MARC und spielt sie mit Katalogisierungsstufe 8 in den B3Kat ein. Voraussetzung ist ein Identifikator, der das beschriebene Forschungsdatum zuverlässig und dauerhaft eindeutig identifiziert. Im Unterschied zu b. werden Änderungen an den Metadaten im Herkunftsrepositorium vorgenommen und per Update in den B3Kat eingespielt. Manuelle Änderungen im B3Kat sind dagegen nicht vorgesehen, da diese Gefahr laufen durch Updates aus dem jeweiligen Repositorium überspielt zu werden.

b) Primärschließung im B3Kat

- i) Die Metadaten zu den Forschungsdaten werden direkt im B3Kat, genauer gesagt in Aleph, mit Katalogisierungsstufe 5 erfasst. Über die OAI-PMH-Schnittstelle "B3Kat Open Data" können sie geharvestet und in lokale Services eingebunden werden. Um z.B. die Metadaten-Forschungsdaten einer bestimmten Bibliothek aus dem Datenpool herauszufiltern, können entsprechende Eintragungen in 078q vorgenommen werden. Diese werden automatisch als Sets interpretiert. Änderungen an den Metadaten werden - im Gegensatz zu a. - ausschließlich in Aleph vorgenommen.

3. Erschließung der Forschungsdaten

Da die Metadaten für Forschungsdaten sowohl in den Repositorien als auch im B3Kat vollständig abgebildet werden sollen, muss ein Mapping zwischen MARC als Austauschformat des B3Kat und DataCite als Standardformat für die Erschließung von Forschungsdaten erstellt werden. Elemente, die nicht in beiden Systemen dargestellt werden können, wurden nicht in die empfohlenen Datensets aufgenommen. Bibliotheken können sich aber entscheiden im System der Primärschließung mehr zu erfassen, müssen dann aber bereit sein die Inkonsistenzen zum anderen System in Kauf zu nehmen.

a) Mapping MARC - DataCite

- b) Mindestdatenset festlegen (ausgehend von dem Mindestdatenset in DataCite und Aleph)