

DBIS Datenworkshop - Bericht

05.03.2021, Universitätsbibliothek Regensburg

Methodik	2
Ziel der Maßnahme	2
Konzept	2
Teilnehmende	2
Ablauf	3
Auswertung	3
Nutzungsszenarien	4
GOKb	4
ZDB	5
LAS:eR	6
FOLIO	7
FIDELIO	8
Offene Diskussion	9

Methodik

Am 05.03.2021 wurde ein fünfstündiger, virtueller Workshop mit Vertreter:innen möglicher Datenaustausch-Institutionen für DBIS durchgeführt. Ziel war, Bedarfe und Möglichkeiten aufzudecken sowie Startpunkte für zukünftigen Datenaustausch zu finden. Der folgende Abschnitt beschreibt zunächst Ziele, Struktur und Teilnehmende der Maßnahme.

Ziel der Maßnahme

Das Ziel war, wie im DFG-Antrag beschrieben, Möglichkeiten zu finden, wie sich DBIS besser in die bibliothekarische Informationsinfrastruktur einfügen kann. Dafür wurden relevante Partner:innen eingeladen, die Nutzungsszenarien vorstellten, die zu einer offene Diskussion führten.

Im Mittelpunkt stand, wie Datenbanken als E-Ressourcen im bibliothekarischen Kontext behandelt werden sollten, wie Datenflüsse etabliert werden können und welche Zwecke diese erfüllen können.

Konzept

Vertreter:innen verschiedener Institutionen wurden eingeladen, um unterschiedliche Nutzungsszenarien vorzustellen. Jeder Gruppe bereitete dazu eine Präsentation vor, zu der vorab ein Template verteilt wurde, und die anschließend offen diskutiert wurde. Am Ende des Workshops war nochmals eine offene Diskussion zu weiteren offenen Punkten aus den Gruppenpräsentationen, sowie übergreifenden Themen geplant.

Im Kontext der COVID-19-Pandemie konnte der Workshop nur als virtuelle Veranstaltung stattfinden. Hierfür wurde das Telekonferenzsystem "Zoom" genutzt.

Teilnehmende

Das DBIS-Team lud geeignete Vertreter:innen ein, die die folgenden Systeme und deren Möglichkeiten und Bedarfe vorstellten:

"**GOKb**" (Global Open Knowledgebase)

"**ZDB**" (Zeitschriftendatenbank)

"**LAS:eR**" (Lizenz-Administrations-System für elektronische Ressourcen)

"**FOLIO**" (Bibliothekssystem "The Future Of Libraries Is Open")

"**FIDELIO**" (Nachweisinstrument für FID-Lizenzen)

Moritz Horn (VZG, GOKb)
Martina Schildt (VZG, FOLIO)
Gerald Steilen (VZG, GOKb)

Hans-Jörg Lieder (SBB, ZDB)
Johann Rolschewski (SBB, ZDB)
Andrea Stei (SBB, ZDB)
Renate Polak-Bennemann (DNB, ZDB)

Michaela Selbach (hbz, LAS:eR)
Daniel Rupp (hbz, GOKb)

Friederike Glaab-Kühn (KfL, FIDELIO)
Patricia Otto (KfL, FIDELIO)

Dr. André Schüller-Zwierlein (UB Regensburg, DBIS)
Dr. Gernot Deinzer (UB Regensburg, DBIS)
Silke Weisheit (UB Regensburg, DBIS)
Dr. Brigitte Doß (UB Regensburg, DBIS)
Felix Riedl (UB Regensburg, DBIS)
Constantin Lehenmeier (UB Regensburg, DBIS)
Claudia Reisinger (UB Regensburg, DBIS)
Cornelia Krellner (UB Regensburg, DBIS)
Luis Moßburger (UB Regensburg, DBIS)

Ablauf

Nach einer Begrüßung durch Herrn Dr. Schüller-Zwierlein, dem Direktor der Universitätsbibliothek Regensburg, stellten sich die Teilnehmenden kurz vor.

Die Kern-Veranstaltung war in 3 Blöcke aufgeteilt. Jeder Gruppe wurden 30 Minuten Zeit eingeräumt, davon 15 Minuten für die Präsentation und 15 Minuten für die anschließende Diskussion. In Block I stellten GOKb und ZDB ihre Szenarien vor, in Block II LAS:eR und FOLIO, Block III war für FIDELIO und die offene Diskussion vorgesehen.

Abschließend wurden die Ergebnisse zusammengefasst und ein Ausblick zur weiteren Zusammenarbeit gegeben und die Teilnehmenden verabschiedet.

Auswertung

Während der Veranstaltung wurde ein ausführliches Protokoll erstellt. Dieses wurde kondensiert und die Aussagen in der Anforderungsspezifikation vom Team berücksichtigt.

Nutzungsszenarien

GOKb

Die Präsentation ist unter https://dbis.ur.de/doc/Datenworkshop_GOKb.pdf einsehbar.

Nutzungsszenario

Aktuell gibt es keinen Austausch zwischen GOKb und DBIS. Die GOKb hat einen globalen Anspruch und soll für eine eindeutige Referenzierung verschiedene IDs für gleiche E-Ressourcen zusammenfügen. Datenbanken könnten als "Titel" oder "Pakete", die dann Titel enthalten können, aufgenommen werden und von der GOKb aus in weitere Systeme abfließen, einige Detailfragen sind ungeklärt. Voraussetzung ist vollautomatischer Austausch, ggf. mit manuell zu bearbeiteten "Review Requests". Wünschenswert wären ZDB-IDs bzw. ZDB-Produktsigel, falls die Datenbank als Paket identifiziert wurde, in den DBIS-Aufnahmen.

Diskussion

Welches Datenformat hätte eine Datenbank in der GOKb? (Riedl)

Eine Datenbank wäre ein "Einzeltitel", ggf. könnte man auch Pakete modellieren, momentan existiert beides. Problem wären Änderungen innerhalb einer Datenbank, wofür entweder eine neue Aufnahme angelegt oder die bestehende aktualisiert werden müsste, hier fehlt eine Entscheidung (Horn).

Schwer lösbar wären mehrere Ebenen unter einem "Einzeltitel", denn das können die meisten bibliographischen Systeme außerhalb der GOKb nicht abbilden oder eine Datenbank, die nur Artikel, aber nicht deren Journals enthalten (Rolschewski).

GOKb ist nicht für bestimmten Use Case gebaut, sondern soll Masse von Paketen/Titeln flexibel zugänglich machen (Steilen).

Hätte Datenaustausch zwischen DBIS und GOKb Potential? (Riedl)

Es können nicht alle Systeme synchronisiert werden, es braucht ein "Master-System". Frage ist also eher, ob DBIS Daten der GOKb nutzen kann, nicht anders herum. Göttingen z.B. nutzt die GOKb als Master-System. (Steilen)

ZDB ist z.B. überregionales Leitsystem für Zeitschriften, für E-Books fehlt das, was ein Problem für die GOKb darstellt. (Rolschewski)

Vorteil der GOKb ist eindeutige Zuordnung von IDs, die z.B. historische Zuordnung von Vorläufern/Nachgängern etc. möglich macht (Steilen, Rolschewski).

Was sind Mindestqualitätsanforderungen für GOKb-Inhalte? (Weisheit)

Vieles ist Konvention, z.B. genießen Konsortien höheres Vertrauen als Einzelinstitutionen. Gut wäre, wenn es je eine ZDB-Aufnahme gäbe, mit der in der GOKb verknüpft werden könnte. (Steilen)

Ob es eine ZDB-Aufnahme gibt, hängt auch von der importierten Quelle ab. DOAJ z.B. enthält viele Zeitschriften aus Indien, die kaum in der ZDB enthalten sind. Grundfrage: Wer trägt diese Aufnahmen nach? Kooperatives Vorgehen zur Lastverteilung wäre gut. (Rolschewski)

Ist Mapping mit außerbibliothekarischen IDs, z.B. Discovery Service, geplant? (Moßburger)

Die Frage ist welche ID die “beste” ist. Die GOKb enthält z.B. “uuid” zum Ursprung des Pakets. (Steilen)

Es gibt verschiedene Zuverlässigkeitssstufen für IDs, wie z.B. überregional anerkannte (ISSN), die bevorzugt werden. Weitere Namensräume sind möglich und erwünscht, automatisierte Füllung aber schwierig. (Horn)

ZDB

Die Präsentation ist unter https://dbis.ur.de/doc/Datenworkshop_ZDB.pdf einsehbar.

Nutzungsszenario

Aktuell gibt es keinen Austausch zwischen ZDB und DBIS. Die ZDB verzeichnet fortlaufende Ressourcen, zu denen unter Umständen auch Datenbanken, Websites und Blogs gehören. Manche Datenbanken sind bereits in der ZDB enthalten, teilweise auch mit Link zu DBIS, durch händische Arbeit. Vorgeschlagen wurden weitere Auswertungen über Potential und bereits vorhandene Überlappungen. Je nach Ergebnis wäre ein Skript zur automatischen Übernahme von einem in das andere System analog zum “EZB-Button” in der ZDB oder semiautomatische Übernahme in größerem Stil denkbar.

Diskussion

Wäre automatisches Austausch-Szenario “Erst DBIS, dann ZDB” möglich? (Riedl)

Das Problem hierbei ist, dass die ZDB-IDs erst im 2. Schritt angelegt werden würden, ggf. müsste also eine interne DBIS-ID mitgeliefert werden, um intern DBIS- und ZDB-Aufnahme zu matchen. Semi-automatische Übernahme wäre wohl vorerst am besten geeignet. Fraglich auch, ob automatische Übernahme den Aufwand wert ist, je nachdem wie sehr monatlich die in Frage kommende Datenmenge wächst, “Flag” müsste gesetzt werden, um geeignete Datenbanken/Websites für ZDB zu kennzeichnen. (Rolschewski)

Wie werden CC0-Lizenzen auf Titeldaten und Bibliotheksdaten gehandhabt? (Doß)

Beides unter CC0, Bibliotheksdaten/Lizenzen können durch Bibliotheken geschützt werden (“Opt-Out”). (Rolschewski)

Wie (umstritten) wurde diese Entscheidung getroffen? (Weisheit, Moßburger)

Vor allem interne Überlegungen zu Open Data, kommuniziert an Anwender:innen, die das mitgetragen haben, 1-2 handvoll haben Opt-Out gewählt. (Rolschewski)

Dass keine Schöpfungshöhe besteht war Konsens. (Lieder)

Automatische Übernahme in Kataloge nur so möglich, Masse sonst nicht stemmbar. (Rolschewski)

Skalierung wichtig, bei Millionen von Daten in Verbundkatalogen nur vollautomatisiert möglich, auch weil Bibliothekar:innen Großteil der Daten nicht mehr sehen. (Steil)

LAS:eR

Die Präsentation ist unter https://dbis.ur.de/doc/Datenworkshop_LASER.pdf einsehbar.

Nutzungsszenario

Aktuell gibt es keinen Austausch zwischen LAS:eR und DBIS. LAS:eR bietet Verwaltung von E-Ressourcen und Konsortialstrukturen. Interesse besteht an IDs aus DBIS zu Institutionen und Datenbanken. Außerdem könnten über LAS:eR z.B. auslaufende Lizenzen gemeldet werden. Voraussetzung wären ein Mapping der IDs bzw. der Institutionen.

Diskussion

Woher kommen die Metadaten, von Anbietern oder der GOKb? (Weisheit)

Hier vom Anbieter, wegen Zeitdruck durch "E-Books NRW", für das schnell ein "Pick & Choose"-System implementiert werden muss. (Selbach)

Ist die Übernahme über die GOKb perspektivisch sinnvoller? (Doß)

Vorerst muss Bedarf gestillt werden, perspektivisch soll v.a. die Arbeit an Metadaten an die Anbieter selbst verlagert werden. (Selbach)

Strukturierte Datenbank-Inhaltsbeschreibungen wie in LAS:eR wären auch für DBIS interessant! (Doß)

Auch für andere Projekte, z.B. GASCO-Monitor. Problem ist Unwissen, was genau in einem Paket enthalten ist zwischen Konsortialstellen, wissen jeweils nur Anbieter, womit eine ID pro Paket pro Konsortium nötig wäre. (Selbach)

Wie würde Austausch konkret aussehen? (Riedl, nach FOLIO-Vortrag)

DBIS-ID müsste für alle relevanten Aufnahmen, als auch die DBIS-internen Institutions-IDs, ergänzt werden für Matching. Zustimmung für Weitergabe an DBIS nötig, dann könnte DBIS z.B. automatisch erfahren, wenn eine Lizenz ausläuft. Schnittstelle aktuell noch nicht öffentlich, Vorgehen wird z.B. mit OA-Monitor getestet. (Selbach)

FOLIO

Die Präsentation ist unter https://dbis.ur.de/doc/Datenworkshop_FOLIO.pdf einsehbar.

Nutzungsszenario

Aktuell gibt es keinen Austausch zwischen FOLIO und DBIS. FOLIO ist ein Bibliothekssystem, das unter anderem Verwaltung von E-Ressourcen und Lizenzen möglich macht. DBIS könnte als "externe Datenquelle" für FOLIO fungieren und dadurch Metadaten zu Datenbanken einspielen, ggf. auch über die GOKb. Bisher kommt ein Großteil der Daten für FOLIO aus der GOKb. Daten aus FOLIO sind für DBIS wenig relevant, da z.B. Lizenzen nicht strukturiert festgehalten werden.

Diskussion

GOKb trotz angesprochener Defizite die beste Möglichkeit für den Datenimport? (Riedl)

Wir nehmen was wir kriegen, sollte in FOLIO verwaltet werden können, z.B. JSON-Job von DBIS nach FOLIO. (Schildt)

Wie werden Datenbanken aktuell in FOLIO verwaltet? (Moßburger)

Bisher gar nicht, nur E-Books und E-Journals, händischer Eintrag möglich. (Schildt)

Wären "freie Datenbanken" aus DBIS für FOLIO als "Paket" relevant? (Riedl)

Ja, "Vereinbarungen" auch ohne Lizenz(-merkmale) möglich. Frage ist eher, ob so etwas lokal in FOLIO überhaupt verwaltet werden muss? (Schildt)

Erfahrung mit freien Ressourcen der EZB zeigt, dass Lokalsysteme aufgeblättert werden.

Möglichkeit wären fachspezifische Ausschnitte. (Rolschewski)

Nur Gesamtabzüge mit JSON/KBART möglich oder auch Update-Lieferungen? (Weisheit)

Schnittstelle noch Entwicklung, Update-Lieferungen sind möglich. (Schildt)

Gibt es "Lizenz-Typen", wie in DBIS oder EZB? (Moßburger)

Nur Unterscheidung national/konsortial/lokal, ansonsten nur Einzelbeschreibungen. (Schildt)

FIDELIO

Die Präsentation ist unter https://dbis.ur.de/doc/Datenworkshop_FIDELIO.pdf einsehbar.

Nutzungsszenario

Aktuell gibt es keinen Austausch zwischen FIDELIO und DBIS. FIDELIO ist ein Nachweisinstrument für Produkte, die unter FID-Lizenzen verfügbar sind und dient z.B. als Datenquelle für Discovery Services über die ZDB-Produktsigel. Manche Produkte sind Datenbanken, in deren Aufnahmen händisch auch ein Link zu DBIS hinzugefügt wurde. Metadaten zu FID-Lizenzen werden verhandelt und sind i.d.R. nachnutzbar. Voraussetzung wäre auch hier ein Mapping der geeigneten Produkte.

Diskussion

Was ist zentraler Use Case, wer nutzt FIDELIO wofür? (Moßburger)

FIDs und interessierte Bibliotheken für Lizenznachweise. (Otto)

Fungiert als "Datendrehscheibe", z.B. für Discovery Services über ZDB-Sigel. (Steilen)

Verknüpft 1) Miterwerb von Metadaten und 2) Nachweisinstrument von FID-Lizenzen für FIDs, Bibliotheken und Interessierte. (Glaab-Kühn)

Ist Lizenztyp in FIDELIO, wie z.B. Nationallizenz, nicht sichtbar? (Doß)

Nur über den Proxylink der FIDs sichtbar. (Otto)

Über welche IDs & Lizenz würde Austausch zwischen DBIS und FIDELIO stattfinden? (Riedl)

EZB-ID, ZDB-Sigel, DBIS-ID alle möglich. (Otto)

FID-Lizenzverträge decken weitere Nutzung der Metadaten ab. (Glaab-Kühn)

Offene Diskussion

Ergebnis

Alle Anwesenden sehen Mehrwert in weiterer Kommunikation. Als erster Schritt bietet sich engere Zusammenarbeit zwischen DBIS und ZDB an, da ZDB überregionale Bedeutung besitzt und die größten Gemeinsamkeiten bestehen. Datenaustausch soll wo möglich und lohnend automatisch und synchronisiert stattfinden.

Produktsigel wichtigste ID, sollte immer genutzt werden! (Rolschewski)

Beantragung nicht durch DBIS selbst, aber sollte hinterlegt sein. (Doß)

Welche Daten sollte DBIS für Konsortien bereithalten? (Doß)

Grundproblem ist, dass nicht jede Konsortialstelle Daten in DBIS pflegt, z.B. auch "DigiLink" des hbz - Überblick fehlt. (Selbach)

Manche Konsortien legen DBIS-Aufnahmen an, v.a. Nationalizenzen. Dann aber manuelle Freigabe jeder Bibliothek nötig, Information über Mailingliste. (Reisinger)

"Kopieren" von Daten keine Lösung, Automatismen und echte Synchronisierung nötig! (Steilen)

Übergreifendes Thema, nicht nur DBIS? (Doß)

Auch heute viele IDs genannt, GOKb soll helfen das zu lösen. (Steilen)

Gibt es Status Quo für konkrete Umsetzung? (Riedl)

ZDB bietet z.B. OAI-PMH-Schnittstelle, bietet auch Änderungsübersicht nach Zeitraum. Große Änderungen sind Problem für Lokalsysteme. Ansonsten RSS oder Atom, v.a. an Standards halten. Standard nur für bibliographische Daten und Holdings, für manche Teile also noch Diskussionsbedarf. (Rolschewski)
OAI-PMH auch gut für GOKb. (Steilen)

Webhooks (Datengebend informiert Datennehmend) als Option? (Moßburger)

Etabliert sind Tagesläufe und Anfragen bei Datengebenden. (Steilen, Horn, Rupp, Polak-Bennemann).

Keine "schiebenden" Updates, da z.B. bei Synchronisierungsfehlern Updates verloren gehen können. (Steilen)

Wie werden Daten in der GOKb geupdated? (Weisheit)

Kooperation mit Verlagen angedacht, langsame Einsicht, dass Sichtbarkeit durch Nachweis (Verzeichnen) in bibliothekarischen Systemen steigt. Datenaufbereitung teilweise aufwändig. (Rolschewski)

Automatisierung spart keine Arbeit, aber wird nötig, stärkere Standardisierung macht Ausnahmen schwieriger. (Steilen)

Aus unserer Sicht wäre ZDB ein guter Startpunkt für Datenaustausch? (Doß)

Vorteil wäre der weitere Datenfluss von ZDB aus, z.B. zu Katalogen. (Rolschewski)

Mit wenig Aufwand könnte dadurch auch viel erreicht werden. (Steil)