

DBIS Data Workshop - Report

5th March 2021, University Library of Regensburg

Methodology	2
Aim of the measure	2
Concept	2
Participants	2
Procedure	3
Evaluation	3
Usage scenarios	4
GOKb	4
ZDB	5
LAS:eR	6
FOLIO	6
FIDELIO	7
Open discussion	8

Methodology

On March 5, 2021, a five-hour virtual workshop was held with representatives of possible data exchange institutions for DBIS. The goal was to uncover needs and opportunities, as well as to identify starting points for future data exchange. The following section begins by describing the goals, structure, and participants of the activity.

Aim of the measure

The goal was, as described in the DFG application, to find ways to better integrate DBIS into the library information infrastructure. For this purpose, relevant partners were invited to present usage scenarios, which led to an open discussion.

The focus was on how databases should be treated as e-resources in the library context, how data flows can be established and what purposes they can serve.

Concept

Representatives from different institutions were invited to present different usage scenarios. Each group prepared a presentation for which a template was distributed in advance and which was then openly discussed. At the end of the workshop, another open discussion was planned on further open points from the group presentations, as well as overarching topics.

In the context of the COVID-19 pandemic, the workshop could only take place as a virtual event. The teleconference system "Zoom" was used for this purpose.

Participants

The DBIS team invited suitable representatives who presented the following systems and their possibilities and needs:

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

"**ZDB**" (German Union Catalogue of Serials)

"**LAS:eR**" (Electronic Resource Management System (ERM))

"**FOLIO**" (library system "The Future Of Libraries Is Open")

"**FIDELIO**" (verification tool for FID licenses)

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 (Regensburg University Library, DBIS)
Felix Riedl (UB Regensburg, DBIS)
Constantin Lehenmeier (Regensburg University Library, DBIS)
Claudia Reisinger (Regensburg University Library, DBIS)
Cornelia Krellner (Regensburg University Library, DBIS)
Luis Moßburger (Regensburg University Library, DBIS)

Procedure

After a welcome by Dr. Schüller-Zwierlein, the director of the University Library Regensburg, the participants briefly introduced themselves.

The core event was divided into 3 blocks. Each group was allotted 30 minutes, 15 minutes for the presentation and 15 minutes for the subsequent discussion. In block I GOKb and ZDB presented their scenarios, in block II LAS:eR and FOLIO, block III was for FIDELIO and the open discussion.

Finally, the results were summarized and an outlook on further collaboration was given and the participants were bid farewell.

Evaluation

During the event, detailed minutes were taken. This was condensed and the statements were taken into account by the team in the requirements specification.

Usage scenarios

GOKb

The presentation can be viewed at https://dbis.ur.de/doc/Datenworkshop_GOKb.pdf.

Usage scenario

Currently there is no exchange between GOKb and DBIS. The GOKb has a global claim and should merge different IDs for the same e-resources for unique referencing. Databases could be included as "titles" or "packages", which can then contain titles, and flow from the GOKb to further systems; some detailed questions are unresolved. A prerequisite is fully automated exchange, possibly with "review requests" to be processed manually. ZDB IDs or ZDB product ids, if the database has been identified as a package, would be desirable in DBIS records.

Discussion

What data format would a database in the GOKb have? (Riedl)

A database would be a "single title", possibly packages could be modeled, currently both exist. Changes within a database could be a problem, for which either a new record would have to be created or the existing one would have to be updated, here a decision is missing (Horn). Difficult to solve would be multiple levels under a "single title", because most bibliographic systems outside of GOKb cannot map that, or a database that only contains articles but not their journals (Rolschewski).

GOKb is not built for a specific use case, but to make the mass of packages/titles flexibly accessible (Steilen).

Would data exchange between DBIS and GOKb have potential? (Riedl)

Not all systems can be synchronized, a "master system" is needed. So the question is rather if DBIS can use GOKB data, not the other way around. Göttingen e.g. uses the GOKb as master system. (Steilen)

ZDB is e.g. supraregional master system for journals, for e-books it is missing, which is a problem for GOKb. (Rolschewski)

Advantage of GOKb is unique assignment of IDs, which e.g. makes historical assignment of predecessors/successors etc. possible (Steilen, Rolschewski).

What are minimum quality requirements for GOKb content? (Weisheit)

Much is convention, e.g., consortia are more trusted than individual institutions. It would be good if there was ever a ZDB record to link to in GOKb. (Steilen)

Whether there is a ZDB record also depends on the imported source. DOAJ, for example, contains many journals from India that are hardly in the ZDB. Basic question: who carries these records? Cooperative approach to load sharing would be good. (Rolschewski)

Is mapping with extra-library IDs, e.g. discovery service, planned? (Moßburger)

The question is which ID is the "best". For example, the GOKb contains "uuid" to the origin of the package. (Steilen)

There are different reliability levels for IDs, such as nationally recognized (ISSN), which are preferred. Other namespaces are possible and desired, but automated filling is difficult. (Horn)

ZDB

Presentation can be viewed at https://dbis.ur.de/doc/Datenworkshop_ZDB.pdf.

Usage scenario

Currently there is no exchange between ZDB and DBIS. The ZDB lists continuing resources, which may include databases, websites, and blogs. Some databases are already included in the ZDB, sometimes with a link to DBIS, through manual work. Further evaluation on potential and already existing overlaps was suggested. Depending on the results, a script for automatic transfer from one to the other system analogous to the "EZB button" in the ZDB or semi-automatic transfer on a larger scale would be conceivable.

Discussion

Would automatic exchange scenario "First DBIS, then ZDB" be possible? (Riedl)

The problem here is that the ZDB IDs would only be created in the 2nd step, so if necessary an internal DBIS ID would have to be supplied in order to match DBIS and ZDB records internally. Semi-automatic transfer would probably be best for the time being. Also questionable whether automatic transfer is worth the effort, depending on how much monthly the amount of data in question grows, "flag" would need to be set to indicate suitable databases/websites for ZDB. (Rolschewski)

How are CC0 licenses handled on title data and library data? (Doß)

Both under CC0, library data/license data can be protected by libraries ("opt-out"). (Rolschewski)

How (controversial) was this decision made? (Weisheit, Moßburger)

Mainly internal considerations about Open Data, communicated to user:s who went along with it, 1-2 handful chose opt-out. (Rolschewski)

The consensus was that there was no level of creation. (Lieder)

Automatic transfer to catalogs only possible in this way, mass otherwise not manageable. (Rolschewski)

Scaling important, with millions of data in union catalogs only fully automated possible, also because librarians no longer see most of the data. (Steilen)

LAS:eR

The presentation can be viewed at https://dbis.ur.de/doc/Datenworkshop_LASER.pdf.

Usage scenario

Currently there is no exchange between LAS:eR and DBIS. LAS:eR provides management of e-resources and consortia structures. There is interest in IDs from DBIS to institutions and databases. In addition, LAS:eR could be used to report expiring licenses, for example. Mapping of IDs or institutions would be a prerequisite.

Discussion

Where does the metadata come from, from vendors or the GOKb? (Weisheit)

Here from the vendor, due to time pressure from "E-Books NRW", for which a "pick & choose" system needs to be implemented quickly. (Selbach)

Is the takeover via the GOKb in perspective more reasonable? (Doß)

For the time being, the need must be satisfied; in perspective, the work on metadata in particular should be shifted to the providers themselves. (Selbach)

Structured database-descriptions like in LAS:eR would also be interesting for DBIS! (Doß)

Also for other projects, e.g. GASCO monitor. Problem is lack of knowledge of what exactly is contained in a package between consortia, only providers know in each case, which would require one ID per package per consortium. (Selbach)

What would exchange look like in concrete terms? (Riedl, after FOLIO presentation)

DBIS ID would have to be added for all relevant recordings, as well as the DBIS internal institution IDs, for matching. Consent for passing on to DBIS necessary, then DBIS could e.g. automatically know when a license expires. Interface currently not yet public, procedure is tested e.g. with OA-Monitor. (Selbach)

FOLIO

The presentation can be viewed at https://dbis.ur.de/doc/Datenworkshop_FOLIO.pdf.

Usage scenario

Currently there is no exchange between FOLIO and DBIS. FOLIO is a library system that enables management of e-resources and licenses, among other things. DBIS could act as an "external data source" for FOLIO and thus bring in metadata on databases, possibly also via the GOKb. To date, much of the data for FOLIO comes from GOKb. Data from FOLIO is not very relevant for DBIS, since e.g. licenses are not recorded in a structured way.

Discussion

Is GOKb, despite the mentioned deficits, the best possibility for data import? (Riedl)

We take what we get and should be able to manage in FOLIO, e.g. JSON job from DBIS to FOLIO. (Schildt)

How are databases currently managed in FOLIO? (Moßburger)

So far not at all, only e-books and e-journals, manual entry possible. (Schildt)

Would "free databases" from DBIS be relevant for FOLIO as a "package"? (Riedl)

Yes, "agreements" are possible without license(-features). Question is rather, if something like this has to be managed locally in FOLIO at all? (Schildt)

Experience with free resources of the ECB shows that local systems are inflated. Possibility would be subject-specific excerpts. (Rolschewski)

Only total extracts in JSON/KBART possible or also update deliveries? (Weisheit)

Interface still under development, update deliveries are possible. (Schildt)

Are there "license types", like in DBIS or EZB? (Moßburger)

Only distinction national/consortia/local, otherwise only single descriptions. (Schildt)

FIDELIO

Presentation can be viewed at https://dbis.ur.de/doc/Datenworkshop_FIDELIO.pdf.

Usage scenario

Currently there is no exchange between FIDELIO and DBIS. FIDELIO is a discovery tool for products available under FID licenses and serves e.g. as a data source for discovery services via the ZDB-Produktsigel. Some products are databases in whose records a link to DBIS has also been added manually. Metadata for FID licenses are negotiated and can usually be reused. The prerequisite here would also be a mapping of the appropriate products.

Discussion

What is the central use case, who uses FIDELIO for what? (Moßburger)

FIDs and interested libraries for license verification. (Otto)

Acts as a "data hub", e.g. for discovery services via ZDB-Sigel. (Steilen)

Links 1) co-acquisition of metadata and 2) verification tool of FID licenses for FIDs, libraries and interested parties. (Glaab-Kühn)

Is the license type in FIDELIO, such as a national license, not visible? (Doß)

Only visible via the proxy link of the FIDs. (Otto)

About which IDs & license would exchange between DBIS and FIDELIO take place? (Riedl)

ECB ID, ZDB Sigel, DBIS ID all possible. (Otto)

FID license agreements cover further use of metadata. (Glaab-Kühn)

Open discussion

Results

All present see added value in further communication. As a first step, closer cooperation between DBIS and ZDB is appropriate, since ZDB has supraregional importance and the greatest commonalities exist. Data exchange should be automatic and synchronized where possible and worthwhile.

Product hedgehog most important ID, should always be used! (Rolschewski)

Application not through DBIS itself, but should be stored. (Doß)

What data should DBIS hold for consortia? (Doß)

Basic problem is that not every consortium maintains data in DBIS, e.g. also "DigiLink" of the hbz - overview is missing. (Selbach)

Some consortia create DBIS records, especially national licenses. Then, however, manual release of each library is necessary, information via mailing list. (Reisinger)

"Copying" of data is no solution, automatism and real synchronization necessary! (Steilen)

Comprehensive topic, not only DBIS? (Doß)

Also today many IDs mentioned, GOKb should help to solve this. (Steilen)

Is there a status quo for concrete implementation? (Riedl)

ZDB offers e.g. OAI-PMH interface, and also offers change overview by period. Large changes are a problem for local systems. Otherwise RSS or Atom, especially stick to standards.

Standard only for bibliographic data and holdings, so for some parts still need discussion. (Rolschewski)

OAI-PMH also good for GOKb. (Steilen)

Webhooks (data giving informs data taking) as an option? (Moßburger)

Established are daily runs and requests to data givers. (Steilen, Horn, Rupp, Polak-Bennemann).

No "pushing" updates, as e.g. updates can be lost in case of synchronization errors. (Steilen)

How is data updated in the GOKb? (Weisheit)

Cooperation with publishers considered, slow realization that visibility increases through proof (listing) in library systems. Data preparation is partly time-consuming. (Rolschewski)

Automation does not save work, but becomes necessary, stronger standardization makes exceptions more difficult. (Steilen)

From our point of view ZDB would be a good starting point for data exchange? (Doß)

Advantage would be further data flow from ZDB, e.g. to LAS:eR & catalogs. (Rolschewski)

With little effort, this could also achieve a lot. (Steil)