

## Fachbeitrag

Friederike Glaab-Kühn und Maike Kittelmann

# FID-Lizenzen sichtbar und suchbar machen: Der Zentrale Nachweis für FID-Lizenzen – Konzept, Umsetzung und Datenmanagement

<http://doi.org/10.1515/abitech-2018-3004>

**Zusammenfassung:** Mit der wachsenden Zahl an zugänglichen FID-Lizenzen wurde es immer wichtiger, die FID-Lizenzen zentral nachzuweisen. Zu diesem Zweck hat das Kompetenzzentrum für Lizenzierung elektronischer Ressourcen im DFG-geförderten System der „Fachinformationsdienste für die Wissenschaft“ einen Zentralen Nachweis mit Informationen zu allen verfügbaren FID-Lizenzen aufgebaut. Der Artikel präsentiert die damit verbundenen konzeptionellen Überlegungen sowie die Festlegungen zur Umsetzung. Anhand von Workflows, Dokumentation und Qualitätsmanagement wird veranschaulicht, wie das Datenmanagement für den Zentralen Nachweis umgesetzt wurde.

**Schlüsselwörter:** FID-Lizenz, Katalog, Datenmanagement

### Making FID-licenses visible and searchable: The catalogue for FID-licenses – concept, implementation and data management

**Abstract:** The increasing number of resources licensed under the new FID-license scheme resulted in a growing demand for a central catalogue. In response, the “Kompetenzzentrum für Lizenzierung” (KfL) created a catalogue that records all FID-licenses in one place. The project was funded by the DFG within the frame of “Specialist Information Services for Academic Purposes”. This article outlines the conceptual design and implementation of this catalogue. Using selected examples, it takes a closer look at the catalogue’s data management, its workflows, documentation, and quality management.

**Keywords:** FID-License, Catalogue, Data Management

## 1 Ausgangslage

Mit den seit 2014 im Rahmen der DFG-geförderten Fachinformationsdienste für die Wissenschaft (FID) abgeschlossenen FID-Lizenzen ist ein neuer Lizenztypus entstanden, der quer zu den bisherigen Lizenzierungskontexten (lokal, (über)regional) liegt. Die Lizenzen werden für einen vom jeweiligen FID fachlich definierten Nutzerkreis verhandelt, der meist aus Einzelnutzern besteht, die verschiedenen Einrichtungen, verteilt in ganz Deutschland, zuzuordnen sind. Für die Verhandlung und überregionale Bereitstellung dieser FID-Lizenzen werden die FID vom ebenfalls durch die DFG geförderten Kompetenzzentrum für die Lizenzierung elektronischer Ressourcen (KfL) unterstützt. Aktuell werden insgesamt 214 Produkte – E-Books, E-Journals und Datenbanken – als FID-Lizenzen den jeweiligen FID-Nutzern zur Verfügung gestellt, hiervon wurden 88 Produkte vom KfL verhandelt.<sup>1</sup> Über diese Produkte sind zum jetzigen Stand ca. 1751 E-Books, 1676 E-Journals sowie 90 Datenbanken zugänglich. Mit der zunehmenden Anzahl bereitgestellter FID-Lizenzen wurde es immer wichtiger, auch die Nachweissituation dieser Lizenzen zu klären, um sie zum einen in die etablierten bibliothekarischen Nachweisstrukturen zu integrieren, zum anderen um ihren Nachweis an zentraler Stelle zu bündeln und sie damit als FID-Lizenzen erkennbar zu machen.

Um künftig einen Überblick über alle sowohl vom KfL als auch von den FID selbst verhandelten FID-Lizenzen zu bieten, hat es sich das KfL zur Aufgabe gemacht, einen Zentralen Nachweis für FID-Lizenzen aufzubauen. Damit verbunden sind konzeptionelle Überlegungen zur Struktur und Benutzbarkeit des Nachweises, damit einhergehende Festlegungen zu ihrer Umsetzung sowie die Erarbeitung von Workflows. Der Artikel wird im Folgenden die konzeptionellen Schritte auf dem Weg zum Zentralen Nachweis für

<sup>1</sup> Als ein „Produkt“ wird hier gemäß der KfL-Statistik – je nach Zugschnitt des Verhandlungsauftrags – sowohl ein Einzeltitel als auch ein Ressourcenpaket angesehen.

FID-Lizenzen darlegen sowie einen Einblick in das damit verbundene Datenmanagement geben, indem gezeigt wird, wie Workflow, Dokumentation und Qualitätsmanagement im Rahmen dieses Projektes umgesetzt wurden.

## 2 Der Zentrale Nachweis für FID-Lizenzen – Herausforderungen und Ziele

Mit den FID-Lizenzen gibt es einen Lizenztypus, der nicht für einen einrichtungsbezogenen Nutzerkreis (z. B. Nutzer einer Universitätsbibliothek), sondern für einen fachlich definierten, meist jedoch überregionalen Nutzerkreis gemäß den Vorgaben des FID verhandelt wird. Als Folge erhalten im Regelfall nur Einzelnutzer einer bestimmten FID-Fachcommunity nach einem vorherigen Anmeldeverfahren Zugriff auf die Lizenzen.<sup>2</sup> Eine Ausnahme bildet der wesentlich geringere Anteil an FID-Lizenzen, die als Campus- oder Nationallizenz verhandelt wurden.<sup>3</sup> Generell erfolgt seitens des FID die Information über die Bereitstellung solcher Lizenzen direkt an die jeweilige Fachcommunity, d. h. direkt an die einzelnen Wissenschaftler und Wissenschaftlerinnen bzw. Institute oder Fachgesellschaften, für die ein Zugriff verhandelt wurde. Dies entspricht zwar dem FID-Konzept, hat jedoch zur Folge, dass Bibliotheken und Konsortien meist keine Information darüber erhalten, dass für eine bestimmte Ressource eine FID-Lizenz verhandelt wurde.

Die Tatsache, dass Bibliotheken grundsätzlich nicht in die Kommunikation zwischen FID und Fachcommunity integriert sind, drückt sich daher in einer mangelnden Kenntnis über die Verfügbarkeit dieser Lizenzen und damit einhergehend auch in deren mangelnder Sichtbarkeit aus. Für die Erwerbungscoordination eines Konsortiums oder einer Bibliothek sowie für den Nachweis in lokalen Systemen spielen diese Lizenzinformationen jedoch eine

wichtige Rolle. Bei den FID-Lizenzen handelt es sich oft um fachlich sehr spezielle Ressourcen, deren Erwerb für die lokalen Etats vieler Bibliotheken meist nicht vertretbar ist. Umso wichtiger ist die Information über die Verfügbarkeit solcher Ressourcen, die über eine andere – überregionale – Lizenzierungsebene gewährleistet ist, wie das bei FID-Lizenzen der Fall ist. Auch wenn sie nicht im lokalen Lizenzangebot einer Bibliothek enthalten sind, erhalten zumindest Angehörige der vom FID definierten Fachcommunity Zugriff auf diese fachspezifischen Ressourcen. Somit kann es für Bibliotheken durchaus von Interesse sein, FID-Lizenzen in ihrem lokalen Katalog nachzuweisen, um die berechtigten Angehörigen einer FID-Fachcommunity unter ihren Nutzerinnen und Nutzern auf die FID-Angebote aufmerksam zu machen und den Zugang zu erleichtern.

Im Laufe der ersten Projektphase des KfL (2014–2016) zeigte sich erwartungsgemäß immer deutlicher, dass nach einer erfolgreichen Verhandlung und Bereitstellung einer FID-Lizenz Informationen zu ihrer Zugänglichkeit auch zentral nachgewiesen werden müssen, wie das in ähnlicher Weise bereits für laufende Konsortialabschlüsse (GASCO-Liste) und National- und Allianz-Lizenzen der Fall ist.<sup>4</sup> Zwar bietet auch die Webseite des KfL eine allgemeine Übersicht, welche FID-Lizenzen abgeschlossen wurden.<sup>5</sup> Dort sind aber nur solche Lizenzen nachgewiesen, deren Zugang über die Infrastruktur des KfL bereitgestellt wird. Jedoch gibt es FID, die für die Bereitstellung ihrer Lizenzen nicht das KfL beauftragen, sondern eigene Systeme nutzen. Bislang gab es keine Möglichkeit, an einer Stelle gebündelt Informationen zu allen verhandelten FID-Lizenzen zu finden. Mit der zunehmenden Anzahl bereitgestellter FID-Lizenzen wurde der Wunsch nach einem zentralen Nachweis für diese seitens der FID und der DFG, insbesondere aber auch seitens der Bibliotheken und Konsortien als den Akteuren der etablierten Beschaffungsstrukturen für elektronische Medien in Deutschland immer dringlicher. Daher hat es sich das KfL in seiner zweiten Förderphase (2017–2019) zur Aufgabe gemacht, diesem Desiderat als Dienstleistung für die FID sowie für Bibliotheken und Konsortien außerhalb des FID-Kontextes durch eine Optimierung der Nachweissituation zu begegnen und einen zentralen Nachweis für die wachsende Zahl von FID-Lizenzen aufzubauen.<sup>6</sup>

<sup>2</sup> Vgl. Hillenkötter, Kristine, Kathrin Behrens, Friederike Glaab-Kühn et al. „Kompetenzzentrum für Lizenzierung im FID-Kontext: Es geht weiter! Verhandlung, Modellbildung und Erschließung im Fokus.“ *Zeitschrift für Bibliothekswesen und Bibliographie* 64,5 (2017): 249–260, hier 252.

<sup>3</sup> Für eine prozentuale Verteilung der Lizenzmodelle siehe Hillenkötter 2017: 254–255, sowie Hillenkötter, Kristine, Maria Huber, Ursula Stanek et al. „Das Kompetenzzentrum für die Lizenzierung elektronischer Ressourcen im DFG-geförderten System der ‚Fachinformationsdienste für die Wissenschaft‘ (FID): Betriebsorganisation, Verhandlung und Bereitstellung von FID-Lizenzen – ein Statusbericht.“ *Bibliothek – Forschung und Praxis* 40,1 (2016): 33–49, hier 39–40.

<sup>4</sup> Vgl. die GASCO-Produktübersicht unter: [https://www.erwerbungsdb.hbz-nrw.de/cgi-bin/GascoListe\\_ext.pl](https://www.erwerbungsdb.hbz-nrw.de/cgi-bin/GascoListe_ext.pl) (12.06.2018) sowie die Produktliste auf <https://www.nationallizenzen.de/angebote> (12.06.2018).

<sup>5</sup> <https://www.fid-lizenzen.de/angebote> (12.06.2018).

<sup>6</sup> Hillenkötter 2017: 258–259.

## 2.1 Das Konzept

Das Konzept des Zentralen Nachweises zielt darauf ab, typenübergreifend (E-Journals, E-Books, Datenbanken) alle Produkte nachzuweisen, für die FID-Lizenzen abgeschlossen wurden, unabhängig davon, ob sie durch das KfL oder die FID selbst verhandelt wurden. Zu diesem Zweck baut das KfL einen FID-Lizenzen-Katalog innerhalb des Verbundsystems des Gemeinsamen Bibliotheksverbunds (GBV) auf. Somit ist der Zentrale Nachweis kein solitäres System, sondern in die bestehenden Strukturen integriert. Das KfL fungiert in diesem Kontext – analog zu einer Bibliothek – als katalogisierende Einrichtung mit eigenem Siegel. Nachgewiesen werden sowohl bibliographisch selbstständige Medien als auch die zugehörigen Artikelmetadaten. Zu diesem Zweck nutzt das KfL die mit den FID-Lizenzen erworbenen Metadaten nach, zu deren Bereitstellung sich die Anbieter gemäß den „FID-Grundsätzen“ in der Regel vertraglich verpflichten (z. B. enthaltene Werke bei E-Book-Paketen oder Datenbanken, Artikelmetadaten bei Zeitschriften).<sup>7</sup> Die Metadaten werden hierfür vom KfL eingeholt, geprüft und konvertiert. Dies gilt auch für die Metadaten von FID-Lizenzen, die nicht vom KfL verhandelt wurden, um einen vollständigen Nachweis aller FID-Lizenzen zu gewährleisten.

Da FID-Lizenzen für alle denkbaren Produkttypen verhandelt werden, wäre es nicht ausreichend gewesen, die FID-Lizenzen allein in die bereits bestehenden bibliothekarischen Nachweissysteme (ZDB, EZB, DBIS, Verbundkataloge) zu integrieren, da dies keinen gebündelten Nachweis aller Ressourcen gewährleisten hätte. Der Zentrale Nachweis hingegen ermöglicht eine Übersicht *aller* FID-Ressourcen, da diese unabhängig vom Medientyp (z. B. Zeitschrift, E-Book, Datenbank), vom Lizenzmodell (FID Nutzerkreis, FID Campus, FID Nationallizenz bzw. weitere künftige Modelle)<sup>8</sup> oder vom Verhandlungsführer (KfL, FID) nachgewiesen werden.

Das Konzept des Zentralen Nachweises sieht außerdem vor, dass die Daten den verschiedenen FID und Verbänden über bereits etablierte Schnittstellen und Verteilungsmechanismen zur Nachnutzung bereitgestellt werden können. Über eine Präsentationsoberfläche in Form eines FID-Lizenzen-OPAC (FIDELIO)<sup>9</sup> werden die Daten zu denjenigen elektronischen Ressourcen, für die

FID-Lizenzen bestehen, nach außen sichtbar gemacht. FIDELIO ermöglicht aufgrund verschiedener Suchfunktionen und Filtermöglichkeiten eine passgenaue Sicht auf die jeweiligen FID-Lizenzangebote. Für jede FID-Lizenz stehen Informationen zur Verfügung, welcher Nutzerkreis auf die jeweilige Ressource zugreifen darf sowie ein Link zur Anmeldeseite. Damit stellt die Präsentationsoberfläche einen zentralen Sucheinstieg für alle abgeschlossenen FID-Lizenzen dar. Mit FIDELIO wird der bisherigen geringen Sichtbarkeit der FID-Lizenzen in den aktuellen Nachweissystemen durch eine zentrale Verzeichnung entgegengewirkt, wodurch perspektivisch auch die Nutzung der FID-Lizenzen befördert wird.<sup>10</sup>

## 2.2 Die Umsetzung

Dem eigentlichen Aufbau des Zentralen Nachweises gingen diverse Vorarbeiten voraus, in deren Rahmen seine Struktur und die zugrundeliegenden Workflows sowie die Ausgestaltung seiner Präsentationsebene zu klären waren. Parallel dazu wurden im Dialog mit den FID der einen, mit GBV und ZDB auf der anderen Seite Absprachen getroffen, um auf der inhaltlichen Ebene die für die Einrichtung des Zentralen Nachweises nötigen Rahmenbedingungen und Erfordernisse zu definieren und auf der technischen Ebene einen problemlosen Datenaustausch zu gewährleisten.

Für den Aufbau des Zentralen Nachweises benötigte das KfL zunächst Informationen zu den FID-Lizenzen, die nicht über das KfL verhandelt wurden, wodurch es auf die Zulieferung der entsprechenden Informationen durch die FID angewiesen war. Im Juli und August 2017 waren die FID über den Mail-Verteiler der AG-FID erstmals dazu aufgerufen, die von ihnen selbst verhandelten FID-Lizenzen inkl. der erforderlichen Eckdaten sowie der Angaben zu den Metadaten in einem entsprechenden Formular an das KfL zu melden. Es gab von insgesamt sieben FID-Rückmeldungen zu den von ihnen verhandelten Produkten. Diese Daten, zusammen mit den Informationen über die vom KfL selbst verhandelten FID-Lizenzen, bilden die Grundlage für die Aufnahme der Katalogisierung im Zentralen Nachweis. Um den Zentralen Nachweis aktuell zu halten, sind entsprechende Update-Routinen zu entwickeln. Besonders wichtig ist es dabei, zeitnah die Metadaten zu den FID-Lizenzen zu erhalten, die nicht vom KfL verhandelt wurden. Dafür werden in regelmäßigen Abständen durch das KfL bei den FID die selbst verhandelten Lizen-

<sup>7</sup> Siehe Grundsätze für den Erwerb von Publikationen in den DFG-geförderten Fachinformationsdiensten für die Wissenschaft: [http://www.dfg.de/formulare/12\\_101](http://www.dfg.de/formulare/12_101) (12.06.2018).

<sup>8</sup> Für eine Übersicht der Lizenzmodelle siehe <https://www.fid-lizenzen.de/dateien/nutzerkreismodelle-fuer-fid-lizenzen> (12.06.2018).

<sup>9</sup> <http://fidelio.fid-lizenzen.de> (12.06.2018).

<sup>10</sup> Hillenkötter 2017: 257–259.

zen abgefragt, da dies für die Vollständigkeit des Nachweises unabdingbar ist.

Da das KfL beim „Zentralen Nachweis für FID-Lizenzen“ als katalogisierende Bibliothek auftritt, musste als einer der ersten Schritte die Grundlage dafür geschaffen werden, die Katalogisierungsarbeiten in der Praxis aufnehmen zu können. Es wurden ein Siegel für das KfL beantragt und eigene Zugangsdaten für die Katalogisierungsinstanz im zentralen Bibliothekssystem (CBS) des GBV und in der ZDB eingerichtet. Des Weiteren mussten grundlegende Festlegungen zur Struktur des Zentralen Nachweises getroffen werden: So wurde mit der Verbundzentrale des GBV (VZG) beschlossen, dass der Zentrale Nachweis für FID-Lizenzen auf einem Solr-Index aufgebaut wird, der die Metadaten der bibliographisch selbstständigen Ressourcen mit den Artikelmetadaten zusammenführt.

Auch die Ausgestaltung der Datensätze auf bibliographischer und auf Exemplar-Ebene galt es zu definieren. Auf bibliographischer Ebene werden bei den Datensätzen das FID-Kennzeichen sowie ein ZDB-Produktsiegel hinterlegt, sofern die Lizenz mehr als einen Titel umfasst.<sup>11</sup> Die Beantragung eines Siegels erfolgt durch das KfL im Auftrag der FID. Über die Produktsiegel können Paketzusammenhänge abgebildet werden. Dies ist insbesondere deswegen wichtig, weil üblicherweise keine FID-Lizenzen für marktgängige Verlagspakete, sondern für individuell zusammengestellte fachliche Zuschnitte verhandelt werden. Bei Lizenzverlängerung kann sich die Paketzusammenstellung zudem verändern, wenn der FID z. B. nicht mehr für alle Zeitschriften eines bestimmten Paketes die FID-Lizenz erneuern möchte.

Auf der Exemplarebene wurde festgelegt, in welcher Form Informationen zu FID, Zugriffsrechten und Nutzergruppe hinterlegt werden, da dies wichtige Hinweise für die Nutzer der FID-Lizenzen sind. So werden in den Exemplarsätzen ebenfalls die FID-Kennzeichen mit einer entsprechenden Ausschrift des FID-Namens hinterlegt. Diese zeigen nicht nur den Nutzerinnen und Nutzern an, von welchem FID die Ressource lizenziert wurde, sondern fungieren ebenfalls als Klammer, um alle FID-Lizenzen eines FID zusammenzuhalten. Zudem finden sich im Exemplarsatz ein Link zur Beschreibung der Nutzergruppe sowie eine direkte Verlinkung zu der Anmeldeseite des FID, über die sich berechnete Nutzerinnen und Nutzer für FID-Lizenzen registrieren können. Im Falle von Zeitschriften erhält der Exemplarsatz zusätzlich die Angabe, für welche Jahrgänge die Zugriffsrechte gelten. Für die interne Verwaltung wird zudem in einer nicht sichtbaren Kategorie

eine vom KfL vergebene Produkt-ID und Verhandlungs-ID hinterlegt, um die Informationen aus dem Katalog mit den Informationen aus den Verhandlungen sowie den Vertragsunterlagen, die in einem KfL-internen System gepflegt werden, zusammenzuführen.

Der Zentrale Nachweis für FID-Lizenzen bietet mit dem FID-Lizenzen-OPAC FIDELIO einen zentralen Sucheinstieg. FIDELIO fungiert dabei als Katalog aller zugänglichen FID-Lizenzen. Dabei wird das KfL als besitzende Bibliothek angezeigt, deren Bestand die FID-Lizenzen sind. Über diese Struktur ist gewährleistet, dass die FID-Lizenzen verbundübergreifend an einer zentralen Stelle nachgewiesen sind. Jeder Datensatz enthält Informationen dazu, welcher FID die elektronische Ressource lizenziert hat sowie eine Verlinkung zur Anmeldeseite des FID, über die direkt überprüft werden kann, ob ein Zugriff für die Angehörigen der jeweiligen Fachcommunity des FID möglich ist. Dafür wurden im Vorfeld Layout und Ausschriften so angepasst, dass die Nutzerinnen und Nutzer ohne größere Probleme zur FID-Ressource weitergeleitet werden. FIDELIO bietet auch eine Suche über die enthaltenen Artikelmetadaten. Alle Metadaten werden indiziert und können entsprechend durchsucht werden. Verschiedene Filtermöglichkeiten ermöglichen passgenaue Sichten, so können z. B. die Ressourcen nach FID gefiltert werden.

FIDELIO fungiert mit dem von ihm gewährleisteten umfassenden Überblick über alle FID-Lizenzen als zentrale Drehscheibe für den Zugang zu den FID-Ressourcen. Ziel von FIDELIO ist es, zum einen sowohl für Einzelnutzer als auch für Bibliotheken einen ersten Sucheinstieg für die vorhandenen FID-Lizenzen zu bieten, und zum anderen die Nutzer zu den jeweiligen FID-Seiten weiterzuleiten. Anmeldung und Zugriff erfolgen dann direkt über die FID gemäß der jeweiligen Definition der zugangsberechtigten Fachcommunity.

Durch den Zentralen Nachweis sind die FID-Lizenzen im Gemeinsamen Verbundkatalog (GVK) an einer zentralen Stelle verzeichnet, und die Daten können über den verbundübergreifenden Datenaustausch für Bibliotheken nachgenutzt werden. Dazu steht das KfL mit der AG KVA und der AG Verbundsysteme über die Frage im Austausch, welche Workflows aktuell bestehen, um einen reibungslosen Datenaustausch zu gewährleisten. Da die FID immer öfter auch Campus- oder Nationallizenzen anbieten, erhält der verbundübergreifende Datenaustausch besondere Relevanz. Um die Nutzung und Sichtbarkeit der FID-Lizenzen zu erhöhen, müssen die Metadaten zeitnah in den Verbunddatenbanken verfügbar sein. Eine besondere Herausforderung besteht dabei in der großen Heterogenität der verfügbaren Metadaten zu den lizenzierten

<sup>11</sup> Eine Übersicht der FID-Kennzeichen ist unter <http://sigel.staatsbibliothek-berlin.de/fid-kennzeichen/> (12.06.2018) zu finden.

Produkten. Dies gilt sowohl in Bezug auf die gelieferten Datenformate als auch auf die Datenqualität. Daher war es von besonderer Bedeutung, ein eigenes Datenmanagement für den Zentralen Nachweis zu etablieren.

### 3 Einblick in die Praxis: Aufbau einer Datenmanagementinfrastruktur

Im Rahmen der Datenkonversion für den Zentralen Nachweis muss eine große Menge heterogener Daten verwaltet werden. Während es für alle Druckerzeugnisse in Bibliotheken fest etablierte Geschäftsgänge gibt, können die Bearbeitungsworkflows von Datenlieferungen von Projekt zu Projekt sehr unterschiedlich sein. So mussten für den Zentralen Nachweis Workflow, Dokumentation und Qualitätsmanagement der Metadatenbearbeitung entsprechend definiert werden. Die damit einhergehenden Erfahrungen und Entscheidungen aus dem Projekt sollen im folgenden Einblick in die Praxis transparent gemacht werden, auch um den Erfahrungsaustausch über Good Practices im Datenmanagement anzuregen.

#### 3.1 Von der Datenlieferung zum Zentralen Nachweis

##### 3.1.1 Datenlieferungen

Grundlage für den Zentralen Nachweis der FID-Lizenzen sind die Metadaten der via FID-Lizenzen zugänglichen Ressourcen.

Bereits bei Abschluss eines Lizenzvertrags wird mit dem Lizenzgeber geklärt, ob die Metadaten zu den lizenzierten digitalen Objekten und ggf. auch diese Objekte selbst (Volltexte, AV-Material, etc.) Teil des Lieferumfangs sind. Es wird darauf hingewirkt, dass – gemäß den DFG-Anforderungen für FID-Lizenzen – Metadaten immer Teil des Vertrages sein sollen.<sup>12</sup> Dies kann jedoch nicht immer umgesetzt werden, z. B. wenn der Lizenzgeber dies nicht anbietet. Dies gilt ebenso für die lizenzierten digitalen Objekte selbst (z. B. Volltexte). Die Datenbearbeitung im KfL beginnt mit der Prüfung, welche Daten laut Vertrag bereitgestellt werden, und der Anfrage dieser Daten durch

das jeweils zuständige KfL-Serviceteam beim Anbieter (Verlage und Aggregatoren, national und international). Mit dem Anbieter werden Details und Lieferwege abgesprochen, wobei das KfL-Datenmanagement als Service für die Datenlieferanten mehrere Lieferwege anbietet (FTP, WebDAV, Harvesten via OAI, Datenträger). Ein einheitliches Lieferverfahren anzustreben ist im KfL-Kontext derzeit unrealistisch, da sich die Möglichkeiten der Anbieter stark unterscheiden. Da die lizenzierten Produkte der FID vornehmlich den Spitzenbedarf abdecken, befinden sich unter den Anbietern zum Teil auch kleinere Nischenverlage oder Fachgesellschaften mit geringer technischer Infrastruktur und Erfahrung. Dies wirkt sich auch auf die gelieferten Datenformate aus.

Die gelieferten Metadaten werden nach einer mehrstufigen Serie von Tests (s. dazu unten) für den Zentralen Nachweis konvertiert und an den Dienstleister VZG zur Einspielung in das zentrale Bibliothekssystem (CBS) des GBV weitergegeben. Die gelieferten digitalen Objekte werden für eine weitere Bearbeitung zur Verfügung gehalten und für einen begrenzten Zeitraum archiviert. Alle Daten werden natürlich auch dem FID zur Verfügung gestellt, sowohl im Lieferzustand als auch konvertiert.

##### 3.1.2 Lieferformate

Das bevorzugte Lieferformat des KfL für E-Book-Metadaten ist MARC21<sup>13</sup>, für Zeitschriftenartikel XML nach der Journal Article Tag Suite (JATS)<sup>14</sup>. Auch andere Standardformate sind willkommen, im E-Book-Bereich ist dies vor allem das Buchhandelsformat ONIX. Die benötigte Zeichenkodierung ist UTF-8. Insbesondere für kleinere Verlage ist es jedoch oft nicht möglich, MARC21, ONIX<sup>15</sup> oder ein anderes Standarddatenformat zu liefern. In vielen Fällen besteht eine Metadatenlieferung daher z. B. aus einem XML-Export im internen Schema der anbietereigenen Datenbank, aus HTML-Header-Daten der Verlags-Webseiten oder auch aus verschiedenen Excel-Dokumenten. Aber auch die Lieferung von Standardformaten ermöglicht leider nicht immer die reibungslose automatisierte Weiterbearbeitung, da in den seltensten Fällen Anwendungsprofile beachtet werden und sich daher auch die Verwendung eines standardisierten Formates je nach Anbieter stark unterscheidet.

Ein häufiges Problem sind in den Metadaten fehlende, für den Lizenznehmer aber zentrale Angaben (z. B. ISSN,

<sup>12</sup> Siehe: Grundsätze für den Erwerb von Publikationen in den DFG-geförderten Fachinformationsdiensten für die Wissenschaft [http://www.dfg.de/formulare/12\\_101](http://www.dfg.de/formulare/12_101) (12.06.2018).

<sup>13</sup> <http://www.loc.gov/marc> (12.06.2018).

<sup>14</sup> <https://jats.nlm.nih.gov> (12.06.2018).

<sup>15</sup> <http://www.editeur.org/8/ONIX> (12.06.2018).

DOI, ggf. weitere fachspezifische Identifier oder Klassifikationen, Sprach-/Ländercodes). Dies kann in Einzelfällen durch Datenanreicherung gelöst werden, wenn entsprechende Datenquellen zur Verfügung stehen. Es ist ein wichtiges Desiderat, dass Datenlieferanten alle Angaben, die ihnen vorliegen und die für den Lizenznehmer von Interesse sind, auch in die Datenlieferung einschließen. Die Granularität der gelieferten Formate unterscheidet sich stark. Dies betrifft vor allem die Daten im Format OAI-DC, die sich nur eingeschränkt als Grundlage für die Erstellung von Katalogdaten eignen.

Die Lieferformate der digitalen Objekte richten sich nach den Vereinbarungen zwischen Anbieter und FID. Das KfL kann hierzu bei Bedarf beraten.

Die Heterogenität der Datenformate stellt hohe Anforderungen an das Datenqualitätsmanagement. Mit Blick auf die Entwicklung eines Geschäftsmodells für das KfL muss der durch die Heterogenität entstehende Aufwand langfristig quantifizierbar gemacht werden. Als Grundlage dazu dient u. a. die Dokumentation, welche Formate die unterschiedlichen Anbieter zu liefern in der Lage sind.

**Tab. 1:** Gelieferte Metadatenformate

Metadatenformat	Zahl der bisherigen Lieferungen in diesem Format
XML (JATS)	14
XML (OAI DC)	7
XML (Eigenformat Anbieter)	38
MARC 21	30
ONIX	5
UNIMARC	1
PICA+ (Konversion FID)	1
Excel-Tabelle	7
SGML (HTML)	2
Noch in Bearbeitung	30

### 3.1.3 Workflows

Eine Besonderheit der im KfL entwickelten Workflows ergibt sich aus der Projektarchitektur. FID-Lizenzen werden an drei Standorten (Staatsbibliothek zu Berlin, SUB Göttingen, Bayerische Staatsbibliothek) durch die dortigen KfL-Serviceteams verhandelt. Im zuständigen Serviceteam werden daher auch die Datenlieferungen entgegengenommen und für die zentrale Konversion an der SUB Göttingen bereitgestellt. Aus dieser Aufgabenteilung ergeben sich Anforderungen, wie z. B. die Organisation des Datentransfers durch für alle Mitarbeitenden zugäng-

liche Systeme, aber auch Chancen, da die mit der Datenannahme betrauten Mitarbeitenden in den Serviceteams über eine genaue Kenntnis der lizenzierten Produkte verfügen und bereits mit den Anbietern in Kontakt stehen.

### 3.1.4 Datenqualitätsmanagement

Aufgrund der Erfahrungen aus anderen Projekten und Arbeitsbereichen war von Anfang an klar, dass die Sicherung der Datenqualität besondere Aufmerksamkeit erfordert. Die problemlose Weiterarbeit mit gelieferten Daten ist im KfL-Projekt umso wichtiger, als die Kontakte zu den Lizenzgebern dezentral an den einzelnen KfL-Standorten bestehen. Reklamationen können also einfacher und schneller erfolgen, wenn sie direkt von den zuständigen Bearbeiterinnen und Bearbeitern vorgenommen werden können.

### 3.1.5 Dezentrale mehrstufige Datenprüfung

Die an das KfL gelieferten Daten durchlaufen eine mehrstufige, teilweise dezentral durchgeführte Serie von Datenqualitätsprüfungen. Eine vollumfängliche Datenprüfung direkt bei der Annahme der Datenlieferung ist nicht möglich, da Probleme der Datenstruktur oder fehlende Angaben erst bei genauerer Prüfung einer Datenlieferung sichtbar werden, die wiederum Spezialkenntnisse voraussetzt. Der erforderliche Zeitaufwand und die Komplexität der Erstprüfung dürfen hingegen nicht zu hoch sein. Die Datenqualitätstests des KfL werden daher in drei Stufen durchgeführt:

#### 1. Stufe: Sofortprüfung bei Annahme einer Datenlieferung

Es ist wichtig, sehr grundsätzliche Eigenschaften der gelieferten Daten bereits bei der Annahme der Lieferung zu prüfen. Dies erleichtert Reklamationen, da zu diesem Zeitpunkt ein aktiver Kontakt zum Datenlieferanten besteht. Frühe Datenprüfungen optimieren die Workflows, denn sie verhindern, dass offensichtlich fehlerhafte Daten erst weitere Geschäftsgänge durchlaufen und dann doch mit großem Zeitverlust reklamiert werden müssen. Zudem sind die überprüften Merkmale die Voraussetzung dafür, dass überhaupt weitergehende Tests durchgeführt werden können.

Im Kontext des KfL-Projekts werden derzeit die Daten unmittelbar nach Erhalt auf die folgenden Merkmale hin geprüft:

- automatisierte Zählung und Auflistung des Lieferumfangs,
- automatisierte Prüfung, ob das Dateiformat lesbar ist (ISO2709, XML, PDF),
- automatisierte Prüfung auf Wohlgeformtheit (XML) (schließt Prüfung auf korrekte Zeichenkodierung ein).

Die beschriebenen Prüfrouinen wurden mit Skripten umgesetzt. Als Skriptsprache der Tests wurde aus mehreren Gründen Perl gewählt. Zum einen waren einige entsprechende Perl-Testskripte bereits vorhanden und konnten für das KfL-Projekt angepasst und erweitert werden. Zum anderen müssen die Tests in jedem Serviceteam ausführbar sein, und Perl ist an vielen Arbeitsplätzen bereits vorhanden. Wichtig war vor allem, dass die Ausführung der Tests leicht handhabbar ist. Durch entsprechende Einstellungen (Registrierung der Perl-Umgebung im PATH) kann das Testskript einfach durch Doppelklick auf die Skriptdatei ausgeführt werden. Als Ergebnis wird eine Log-Datei geschrieben. Ist diese Log-Datei leer, ist alles in Ordnung. Tritt ein Fehler auf, enthält die Log-Datei die Beschreibung des Fehlers in einer Form, die direkt einer Reklamation angefügt werden kann.

Aufgrund der großen Heterogenität der Daten werden fortlaufend bei Bedarf weitere Tests hinzugefügt. Sollten bei Prüfungen Fragen auftreten, können die KfL-Service-teams den Arbeitsbereich Konversion direkt ansprechen.

## 2. Stufe: Pretest vor der Konversion

Die so geprüften Daten werden an das zentrale Datenmanagement des Projekts an der SUB Göttingen weitergeleitet. Hier finden vor der Konversion weitere Pretests statt. Diese Tests erfordern tiefere Fachkenntnisse als die Soforttests, wobei dies sowohl die Ausführung des Tests als auch die Interpretation des Testergebnisses betrifft. Durch die zentrale Durchführung dieser Tests kann der Zeitaufwand in den KfL-Service-teams gering gehalten werden.

Die im KfL durchgeführten Pretests umfassen z. B.:

- automatisierte Prüfung auf eindeutige Identifier und Vorhandensein bestimmter Angaben, z. B. Titel (für MARC-Dateien),
- automatisierte Prüfung der Zeichenkodierung (für MARC-Dateien),
- automatisierte Validierung (für XML-Dateien),
- automatisierte Prüfung auf Verdachtsfälle für DRM (für PDF-Dateien),
- automatisierte Prüfung des Codecs (Videos).

Auch die Pretests werden konstant erweitert, wo immer dies sinnvoll erscheint und der Workflow optimiert werden kann.

## 3. Stufe: Detaillierte Prüfung der Daten im Zuge der Konversion

Sollte eine Auslagerung in vorgeschaltete Pretests nicht erfolgen, werden diese Merkmale im Rahmen der Konversion geprüft.<sup>16</sup> Die spezifischen Tests und Abfragen im Zuge der Konversion sind daher ebenfalls Teil des Datenqualitätsmanagements. Viele Arten von Problemen können erst bei eingehender Prüfung erkannt werden. Dies ist umso mehr der Fall, wenn keine Standarddateiformate geliefert wurden. Ein letztes Mal durchlaufen KfL-Daten automatisierte Tests, bevor sie an die Verbundzentrale gegeben werden. Dies stellt sicher, dass valide Daten ohne Verzögerung in den Zentralen Nachweis eingespielt werden können, denn bei Prüfung der konvertierten Daten können auch noch einmal Probleme in den gelieferten Daten auffallen.

Ein systematisches, den jeweiligen Anforderungen angepasstes Datenqualitätsmanagement kann viele Probleme bereits in einem frühen Stadium der Bearbeitung erkennen und ist ein unabdingbarer Bestandteil von Datenworkflows. Ein systematisches Datenqualitätsmanagement kann Problemquellen zwar nicht hundertprozentig ausschließen, aber doch ganz entscheidend verringern.

Alle genannten Tests der gelieferten Daten beziehen sich auf deren formale Korrektheit, Validität und automatisierte Bearbeitbarkeit. Sollte ein FID selbst zusätzliche Qualitätstests durchführen wollen, ist dies jederzeit möglich.<sup>17</sup>

Die häufigsten bisher beobachteten Reklamationsgründe sind Zeichensatzfehler oder anderweitig nicht wohlgeformtes XML.

### 3.1.6 Spezifikation des Zielformats

Für die Bearbeitung der Daten ist es von besonderer Bedeutung, das Zielformat zu spezifizieren. Ohne genaue Spezifikation des Zielformats kann die Konversionsarbeit nicht aufgenommen werden. Spezifikationsänderungen nach Beginn der Konversionen führen zu zeit- und kostenintensivem Mehraufwand wie z. B. der Neuprogrammierung.

<sup>16</sup> Dies kann der Fall sein, wenn ein Format zum ersten Mal geliefert wird und es noch keinen Pretest gibt oder wenn ein Format vermutlich sehr selten geliefert wird und das Konzipieren und Entwickeln von Pretests (noch) nicht lohnt. Des Weiteren kann es vorkommen, dass die zu prüfenden Merkmale nicht oder nicht mit vertretbarem Aufwand mittels automatischer Routinen geprüft werden können.

<sup>17</sup> Dies können insbesondere weitere Vollständigkeitsprüfungen sein, falls diese sehr aufwendig sein sollten, oder eine Prüfung auf fachspezifische inhaltliche Aspekte (z. B. Korrektheit von Transliterationen), sofern diese nicht an das KfL delegiert werden können.

nung bereits programmierter Konversionsroutinen, wiederholten Konversionen derselben Daten, aufwändigen Nachkonversionen von Daten, deren Bearbeitung bereits abgeschlossen war u. a. m.

Angaben wie „MARC21“ reichen als Spezifikation nicht aus, es müssen auch die notwendigen Anwendungsprofile beachtet werden. Daten müssen beispielsweise im Zentralen Nachweis nicht nur auffindbar sein, sondern auch aktualisiert werden können. Datensatzupdates erfordern eindeutige und stabile Identifier. In einigen Fällen bieten Datenformate auch alternative Möglichkeiten, Merkmale abzubilden. Wenn der Empfänger der Datenlieferung nur eine dieser Möglichkeiten weiterverarbeiten kann, dann muss dies frühzeitig bekannt sein.

Die KfL-Metadaten werden vor dem Einspielen in das zentrale Bibliothekssystem (CBS) des GBV durch die Verbundzentrale in das Internformat Pica+ konvertiert. Die Artikelmetadaten werden in einen SolR-Index eingespielt. Die bei diesen Prozeduren verwendeten Standardroutinen greifen nicht nur auf Standardinhalte des verarbeiteten Datenformats zu, sondern stützen sich u. U. auf in sogenannten lokalen Feldern vorgehaltene Inhalte, die nicht Teil des standardisierten Datenformats sind. Solche Besonderheiten müssen vor der Konversion mitgeteilt werden.

Unklarheiten in der Spezifikation des Zielformates sind einer der Gründe für die Kostenintensität von Konversionsarbeiten. Zeitaufwändige Neu- und Nachkonversionen behindern den Projektablauf erheblich und sind i. d. R. vermeidbar.

### 3.2 Workflowmanagement und Dokumentation

Datenqualitätsmanagement und Datenbearbeitung werden von einer systematischen Dokumentation der vorhandenen Datenlieferungen sowie deren Attributen und Lifecycles begleitet. Im KfL-Projekt wird die Wiki-Software Confluence für die Dokumentation verwendet, der Workflow ist in der Projektmanagementsoftware Jira modelliert. Im Folgenden wird vertieft auf die Anforderungen an die Dokumentation eingegangen.<sup>18</sup>

Die einzelnen Datenlieferungen und die Datenkonversionen werden separat dokumentiert. Zum einen gibt es Einzellieferungen, deren Daten aufgrund von Mängeln

nicht konvertiert werden können. Zum anderen können in manchen Fällen mehrere Datenlieferungen in einer Konversion zusammen bearbeitet werden.

Zur Beschreibung von Datasets wurden verschiedentlich Vokabulare entwickelt.<sup>19</sup> Existierende Vokabulare wurden zur Optimierung des Retrievals von Forschungsdatasets entwickelt und sind daher zu allgemein, um direkt für die Verwaltung von Datenlieferungen nachgenutzt werden zu können. Das Datenmodell des KfL beschreibt drei Ebenen. Angaben, die sich nicht nur auf eine einzelne Datenlieferung beziehen, werden datenlieferungsübergreifend festgehalten. Jede einzelne Datenlieferung wird separat dokumentiert. Ebenfalls separat werden Angaben zur Konversion erfasst. Da die digitalen Objekte im KfL über die Vollständigkeits- und Qualitätsprüfung hinaus nicht weiter bearbeitet werden, erfolgt ihre Dokumentation analog zur Dokumentation der Metadaten, aber entsprechend weniger umfangreich, diese wird daher hier nicht eigens aufgeführt. Für die Bearbeitung der Dokumentation werden durchgehend Vorlagen und Ausfüllhilfen zur Verfügung gestellt. Wenn später quantitative Auswertungen vorgenommen werden sollen, sollte dies bereits in die Planung der Dokumentation einbezogen werden.

Die Angaben in den Tabellen 2 und 3 beziehen sich auf die Dokumentation von Einzellieferungen. Einzellieferungen können sowohl Metadaten als auch digitale Objekte (einzeln oder zusammen) umfassen.

In Tabelle 4 findet sich eine Übersicht zur Konversionsdokumentation der Metadaten.

Da alle verwendeten Konversionsdateien derzeit mit der Konversion abgelegt werden, ist die Konversion dadurch immer vollständig dokumentiert, und die Punkte 17–19 entfallen. Sollte sich dies ändern, würden Ablageort und Version der verwendeten Konversion in die Dokumentation aufgenommen. Der Ablageort der Konversionsergebnisse ist einheitlich und daher hier nicht eigens aufgeführt.

Die entwickelten Datenbearbeitungsworkflows und Datenqualitätsstandards wurden im Projekt ausgiebig erprobt und gewährleisten eine effiziente Zusammenarbeit bei gleichzeitiger Minimierung etwaiger Wartezeiten. Die Verwendung einer aktuellen Projektmanagementsoftware sowie die Einbettung des Qualitätsmanagements in die einzelnen Bearbeitungsschritte ermöglichen einen umfassenden Überblick und tragen entscheidend zu einem optimalen Konversionsergebnis bei.

Die so erstellten Daten bilden die Basis für FIDELIO und damit für einen zentralen Nachweis der FID-Lizenzen.

<sup>18</sup> Auf den im Projekt erarbeiteten Jira-Workflow wird an dieser Stelle nicht näher eingegangen. Informationen zur Arbeit mit selbst erstellten Jira-Workflows finden Sie unter <https://confluence.atlassian.com/jira> (12.06.2018).

<sup>19</sup> Z. B. <https://developers.google.com/search/docs/data-types/dataset> (12.06.2018).



**Tab. 2:** Lieferungsübergreifende Angaben

Feld	Wertebereich	Hinweis
1. Datentransfer	Freitext	Angaben zum vereinbarten Datentransfer mit dem Lizenzgeber sowie Zugangsdaten.
2. ZDB-Paketsigel	String	FID-Lizenzen werden mit Paketsigeln aus dem Namensraum „ZDB-1“ erfasst.
3. ZDB-ID	String	bei Zeitschriften
4. Zugang	Tabelle	Genauere Angaben zum Umfang des lizenzierten Materials und den erworbenen Zugriffsrechten in an das KBART-Format angelehnten Tabellen

**Tab. 3:** Kernindikatoren Datenlieferung (Einzellieferung)

Feld	Wertebereich	Hinweis
1. Verhandlungs-ID	alpha-numerisch	Die Verhandlungs-ID ermöglicht die Zuordnung der Datenlieferung zum zugehörigen Lizenzvertrag.
2. Datenlieferungs-ID	alpha-numerisch	Da es zu jedem Lizenzvertrag mehrere Datenlieferungen geben kann, bekommt jede einzelne Lieferung einen internen Identifier, um sie eindeutig zu referenzieren.
3. Zuständiges KfL-Serviceteam	definierte Terme	Im KfL-Projekt werden an drei Standorten Datenlieferungen entgegengenommen. Auch die grundlegende Qualitätsprüfung bei Datenerhalt findet dort statt.
4. Bearbeiterin/Bearbeiter	String	Für Rückfragen ist die Ansprechpartnerin oder der Ansprechpartner unmittelbar ersichtlich.
5. Datum der Lieferung	Datum	
6. Status	definierte Terme	Details zum Bearbeitungsstatus gehen immer aktuell aus dem zugehörigen JIRA-Vorgang hervor. Die Dokumentation hält die wesentlichen Ergebnisse fest. In Bezug auf Datenlieferung sind die möglichen Werte: „Daten erhalten“, „Daten geprüft“ und ggf. „Reklamation“.
7. Kommentar zum Status	Freitext	Da das Feld „Status“ Einträge in kontrolliertem Vokabular erwartet, wird es durch ein Kommentarfeld für etwaige Anmerkungen ergänzt.
8. Zugehörige Lieferungen	Link	Es kommt vor, dass verschiedene Datenlieferungen zusammengehören, z. B. Ersatzlieferungen bei Reklamation. Hier kann bei Bedarf direkt zu anderen Lieferungen verlinkt werden.
9. Datentransfer	Freitext	Angaben zum Datentransfer werden i. d. R. auf einer übergeordneten Ebene für alle Datenlieferungen erfasst. Sollte es Besonderheiten geben oder wechselt die Transfermethode häufig, werden Angaben zum Datentransfer bei der jeweiligen Lieferung vermerkt.
10. Zeitraum der Lieferung	Datum	Datenlieferungen decken i. d. R. bestimmte Veröffentlichungszeiträume/Lizenzzeiträume ab. Die Angabe wird u. a. für die Vollständigkeitsprüfung benötigt. Dokumentiert wird der erwartete Zeitraum.
11. Umfang der Lieferung	Freitext	Dokumentiert wird der tatsächliche Umfang der Datenlieferung, d. h. Anzahl, Typ und Größe der gelieferten Dateien.
12. Fehlende Dateien	definierte Terme	Sollte die Vollständigkeitsprüfung Diskrepanzen zwischen den erwarteten Daten und den erhaltenen Daten offensichtlich machen, wird dies hier dokumentiert.
13. Metadatenformat	definierte Terme	Das Metadatenformat der Lieferung wird so genau wie möglich angegeben, z. B. XML, MARCXML, MARC-ISO2709, JATS, JSON. Der Eintrag „NICHT ANWENDBAR“ wird vergeben, wenn die Datenlieferung keine Metadaten enthält.
14. Metadatenformat formal korrekt?	definierte Terme	Dokumentation der ersten Stufe des Qualitätsmanagements („Soforttests“).
15. Inhaltsdatenformat	definierte Terme	Als „Inhaltsdaten“ werden projektintern die gelieferten digitalen Objekte bezeichnet, z. B. Volltexte, Bildmaterial, AV-Medien, etc.
16. Inhaltsdatenformat formal korrekt?	definierte Terme	Dokumentation der ersten Stufe des Qualitätsmanagements („Soforttests“).
17. Reklamation/Grund	Freitext	Sollte eine Datenreklamation erforderlich sein, wird dies hier vermerkt.
18. Reklamation/Datum	Datum	Das Reklamationsdatum wird zur leichteren Auswertung in einem separaten Feld angegeben.

Feld	Wertebereich	Hinweis
19. Ablageort	Pfad	Der Ablageort der Datenlieferung im Dateisystem. Im KFL-Projekt werden standardisierte Dateipfade verwendet.
20. Versand an FID	definierte Terme	Notiz, ob der lizenznehmende Fachinformationsdienst die gelieferten Daten erhalten hat. Alle weiteren workflowbezogenen Angaben sind aus JIRA ersichtlich. Diese Angabe kann als Beispiel für weitere erforderliche Angaben dienen.
21. Archivierungsdatum	Datum	Nach Abschluss der Bearbeitung werden die Datenlieferungen archiviert. Datum und Pfad im Archiv werden in separaten Feldern erfasst.
22. Archivierungsort	Pfad	Nach Abschluss der Bearbeitung werden die Datenlieferungen archiviert. Datum und Pfad im Archiv werden in separaten Feldern erfasst.

**Tab. 4:** Kernindikatoren Datenkonversion

Feld	Wertebereich	Hinweis
1. Verhandlungs-ID	alpha-numerisch	Die Angabe dienen der Zuordnung. Zur Dokumentation siehe die Erläuterungen dieser Felder unter „Einzellieferung“.
2. Datenlieferungs-ID	alpha-numerisch	Die Angabe dienen der Zuordnung. Zur Dokumentation siehe die Erläuterungen dieser Felder unter „Einzellieferung“.
3. Zuständiges KFL-Serviceteam	definierte Terme	Die Angabe dienen der Zuordnung. Zur Dokumentation siehe die Erläuterungen dieser Felder unter „Einzellieferung“.
4. Bearbeiter/Bearbeiterin	String	Für Rückfragen ist die Ansprechpartnerin oder der Ansprechpartner unmittelbar ersichtlich.
5. Versand an FID	definierte Terme	Notiz, ob der lizenznehmende Fachinformationsdienst die konvertierten Daten erhalten hat. Alle weiteren workflowbezogenen Angaben sind aus JIRA ersichtlich. Diese Angabe kann als Beispiel für weitere erforderliche Angaben dienen.
6. betroffene Lieferungen	Links	Im Unterschied zur Lieferungsdokumentation geht es hier um die ggf. in einer Konversion zusammengefassten Lieferungen.
7. Status	definierte Terme	Details zum Bearbeitungsstatus gehen immer aktuell aus dem zugehörigen JIRA-Vorgang hervor. Die Dokumentation hält die wesentlichen Ergebnisse fest. In Bezug auf Datenlieferung sind die möglichen Werte: „Daten erhalten“, „Daten geprüft“ und ggf. „Reklamation“.
8. Kommentar zum Status	Freitext	Da das Feld „Status“ Einträge in kontrolliertem Vokabular erwartet, wird es durch ein Kommentarfeld für etwaige Anmerkungen ergänzt.
9. Verwendete Tests/Ergebnisse	Tabelle	An dieser Stelle werden Ergebnisse der Pretests und weiterführender Tests während der Konversion festgehalten.
10. Unique Identifier in den Datensätzen	Freitext	Da Datensätze eindeutig identifizierbar sein müssen (z. B. für Updates), wird hier festgehalten, welche eindeutigen Identifier die gelieferten Datensätze enthalten (z. B. DOI). Sollten keine eindeutigen Identifier in den Daten vorhanden sein, ist es wichtig, dies zu dokumentieren. Dokumentiert wird in diesem Fall auch, nach welchem Schema in der Konversion Identifier generiert wurden.
11. Umfang der Lieferung	Freitext	Während Umfang und Vollständigkeit der Lieferung bei der jeweiligen Einzellieferung formal geprüft wurden, werden im Rahmen der Konversion die Daten detailliert gesichtet. Die Angaben hier beziehen sich daher nicht auf die Dateizahl, sondern die Anzahl der Datensätze.
12. Vollständigkeit der Lieferung	definierte Terme	Sollte das Fehlen von Daten erst bei der Konversion auffallen, werden die Details zur Reklamation nicht hier, sondern anschließend bei der jeweiligen Einzellieferung dokumentiert. Die Konversion wird fortgeführt, wenn eine Ersatzlieferung vorhanden ist.
13. Fehlende Metadaten	definierte Terme/Tabelle	Genauere Auflistung der fehlenden Datensätze bzw. Zeiträume.
14. Anmerkungen zur Konversion	Freitext	Es werden alle Besonderheiten der Konversion festgehalten, die für die Bearbeitung weiterer Datenlieferungen zu einem Produkt relevant sind.
15. Dateiname der konvertierten Datei	Dateiname	Die Benennung der konvertierten Dateien erfolgt nach einem einheitlichen Schema. Darin enthalten sind Produkt- und Lieferungsidentifizier sowie ein UNIX-Timestamp.
16. Link auf Verlagsserver gültig?	definierte Terme	Um den Zugang zum lizenzierten Produkt sicherzustellen, werden Zugriffslinks exemplarisch getestet.

Feld	Wertebereich	Hinweis
17. Verwendete Konversion	definierte Terme	Da sehr unterschiedliche Formate geliefert werden, erfolgt die Konversion mit den jeweils geeigneten Tools.
18. Konversion – Version	definierte Terme	Alle Konversionsdateien werden versionskontrolliert.
19. Konversion – Ablageort	Pfad	Speicherort der für die Konversion benötigten Dateien im Dateibaum.

## 4 Ausblick

Zum Bibliothekartag 2018 in Berlin konnte FIDELIO – der OPAC des Zentralen Nachweises für FID-Lizenzen – in einer Betaversion vorgestellt werden. Damit ist ein wichtiger Meilenstein im Rahmen der zweiten Projektphase des KfL erreicht. Nach dem Einspielen der ersten Metadaten ausgewählter FID-Lizenzen zu Testzwecken wird FIDELIO in einem absehbaren Zeitraum alle FID-Lizenzen nachweisen, die vom KfL verhandelt oder von den FID als selbst verhandelt gemeldet wurden. Durch Update-Routinen wird das Lizenzangebot stets auf einem aktuellen Stand gehalten und bietet einen zentralen Sucheinstieg. Damit wird dem Desiderat begegnet, das Angebot der FID-Lizenzen besser an verschiedene Zielgruppen zu kommunizieren, nicht nur an interessierte Nutzerinnen und Nutzer, sondern auch an Bibliotheken. Dies wird als Folge auch die Nutzung der elektronischen Ressourcen mit FID-Lizenz erhöhen können. Das KfL hat sich als Serviceeinrichtung für die FID dieser Aufgabe angenommen und einen Nachweis aller FID-Lizenzen geschaffen sowie ein entsprechendes Datenmanagement aufgebaut. Da mit den FID-Lizenzen soweit möglich auch dauerhafte Zugriffsrechte auf die lizenzierten Ressourcen verhandelt wurden, bietet FIDELIO als Zentraler Nachweis für FID-Lizenzen einen nachhaltigen Überblick und schnellen Zugriff auf diese Ressourcen.

## Autoreninformationen



### Friederike Glaab-Kühn

Abteilung Bestandsaufbau  
Staatsbibliothek zu Berlin – Preußischer Kulturbesitz  
Potsdamer Straße 33  
10785 Berlin  
[friederike.glaab-kuehn@sbb.spk-berlin.de](mailto:friederike.glaab-kuehn@sbb.spk-berlin.de)  
[orcid.org/0000-0002-6421-9584](https://orcid.org/0000-0002-6421-9584)



### Maike Kittelmann

Niedersächsische Staats- und Universitätsbibliothek  
Metadaten und Datenkonversion  
Platz der Göttinger Sieben 1  
37073 Göttingen  
[kittelmann@sub.uni-goettingen.de](mailto:kittelmann@sub.uni-goettingen.de)  
[orcid.org/0000-0001-6999-3184](https://orcid.org/0000-0001-6999-3184)