

K10plus – Technik und Entwicklung

Autoren:

<p><u>Reiner Diedrichs (Hauptautor)</u> Verbundzentrale des GBV (VZG) Platz der Göttinger Sieben 1 37073 Göttingen sekretariat@gbv.de</p>	
<p>Dr. Ralf Goebel Bibliotheksservice-Zentrum Baden- Württemberg (BSZ) Felix-Wankel-Straße 4 78467 Konstanz sekretariat@bsz-bw.de</p>	

Bildrechte: BSZ und VZG

Schlüsselwörter: Verbundsystem, K10plus, Funktion

Abstract: Anregungen der Forschung aufgreifend wurde mit K10plus eine Plattform für Metadaten entwickelt, die mehr ist als nur eine Verbunddatenbank. Über die Hälfte der staatlichen Universitäten Deutschlands nutzen K10plus. Der Artikel beleuchtet die Entstehung und die Funktionsweise von K10plus unter technischen Aspekten.

Entstehung von K10plus

Die Entstehung der Kooperation zwischen dem Bibliotheksservice-Zentrum Baden-Württemberg (BSZ) und der Verbundzentrale des Gemeinsamen Bibliotheksverbundes GBV (VZG) im Jahr 2015 ist vor dem Hintergrund der damals geführten Diskussionen zur Neuausrichtung der Informationsinfrastrukturen zu sehen. Die Deutsche Forschungsgemeinschaft (DFG), der Wissenschaftsrat und der Rat für Informationsinfrastrukturen hatten sich dafür ausgesprochen, forschungsunterstützende Infrastrukturen zusammenzulegen, zu modernisieren und zu internationalisieren. Davon betroffen waren auch die deutschen Bibliotheksverbände. Die Verbundzentralen BSZ und VZG haben die Empfehlungen aufgegriffen und unter anderem die Zusammenlegung der Verbundsysteme (CBS¹, Central Bibliographic System) zu einem gemeinsamen System für zehn Bundesländer, für die Stiftung Preußischer Kulturbesitz und für weitere Einrichtungen der Forschung vereinbart. Die neue gemeinsame Verbunddatenbank ging im März 2019 unter dem Namen K10plus in den Produktivbetrieb. Sie enthält heute mehr als 200 Mio.

¹ Das CBS / LBS wurde ursprünglich von Pica (Project for integrated catalogue automation), einer Abteilung des Niederländischen Wissenschaftsministeriums, entwickelt. Später wurde Pica in eine Stiftung mit Sitz in Leiden umgewandelt. Die Softwareentwicklung und der Betrieb des Niederländischen Verbundsystems wurden dann in eine Niederländische BV überführt, bei der OCLC mittlerweile alleiniger Gesellschafter ist. Die Softwareentwicklung erfolgt weiterhin durch die BV am Standort Leiden.

Nachweise zu Büchern, E-Books, Zeitschriften, Aufsätzen, Kongressberichten, Mikroformen, elektronischen Dokumenten, Datenträgern, Musikalien und Karten. Zur Entstehung von K10plus und den Kooperationszielen verweisen wir auf den frei verfügbaren Aufsatz „Veränderungen sind nötig“² aus dem Jahr 2018. Im vorliegenden Beitrag sollen hingegen stärker die technischen Aspekte zum Aufbau und zum Betrieb von K10plus erörtert werden.

Verbundsystem und Bibliothekssystem

Zunächst ist wichtig zu verstehen, dass BSZ und VZG zwar die gleiche Verbundsoftware CBS einsetzen, die Versorgung der heterogenen Lokalsysteme aber unterschiedlich realisiert wird.

Im GBV wird entsprechend den DFG-Förderrichtlinien aus dem Jahr 1990 das integrierte lokale Bibliothekssystem (LBS: Local Bibliothek System), das unmittelbar mit dem Verbundsystem (CBS) verknüpft ist, von nahezu allen Staats-, Landes-, Universitäts- und Hochschulbibliotheken sowie von einer Vielzahl anderer Einrichtungen genutzt. Diese Einheitlichkeit hat sich als sehr effizient und kostengünstig erwiesen, so dass sie im Verwaltungsabkommen der GBV-Länder 1996 festgeschrieben wurde. In der Folge wurden nur wenige andere Lokalsysteme, u.a. SISIS-SunRise, Ex Libris Alma und BiblioMondo an die Verbundkatalogisierung angeschlossen worden. Weitere Systeme übernehmen Daten via Z39.50/SRU. Das BSZ bedient mit seinem CBS-Verbundsystem hingegen eine Vielzahl unterschiedlicher Lokalsysteme. Auch diese Lösung hat sich als überaus effizient und flexibel erwiesen. Während im GBV das LBS-System nicht ohne das CBS-Verbundsystem betrieben werden kann, funktionieren die durch das BSZ betreuten Lokalsysteme technisch vollkommen unabhängig und eigenständig. Jeder der beiden Lösungen hat Vor- und Nachteile.

Der Vorteil der CBS-LBS-Lösung besteht darin, dass die im CBS-Verbundsystem katalogisierten Daten in Echtzeit und ohne umständliche Datenmanagementprozesse durch das System selbst ins LBS übertragen werden. Daraus ergibt sich für die betreuende Verbundzentrale (VZG) der Vorteil, die eigenen Kompetenzen auf das Handling LBS-CBS konzentrieren zu können. Hinsichtlich des Personaleinsatzes ist das maximal synergetisch und effizient.

Will man als nutzende Bibliothek aber ein anderes Lokalsystem in Verbindung mit dem CBS-Verbundsystem einsetzen, kann sich der Vorteil auch als Nachteil herausstellen. Denn wenn die Unterstützung unterschiedlicher Lokalsysteme mehr die Ausnahme als die Regel darstellt, fehlen entsprechende Angebote, Kompetenzen und Workflows. Daher baut die VZG bereits seit einigen Jahren erfolgreich Kompetenzen mit dem Open-Source-System FOLIO auf und beteiligt sich maßgeblich an der Entwicklung des Systems. Erste Installationen des FOLIO-ERM-Moduls sind an der SUB Bremen und der ZBW Kiel erfolgt.

Im BSZ verhält es sich gerade anders herum. Durch die strikte Trennung der Systeme ist das BSZ hinsichtlich der Kompetenzen und Workflows optimal auf die Unterstützung einer Vielzahl unterschiedlicher Lokalsysteme eingerichtet. So betreibt das BSZ zum einen für die

² Diedrichs, Reiner / Goebel, Ralf: Veränderungen sind nötig. Kooperation des Bibliotheksservice-Zentrums Baden-Württemberg (BSZ) und der Verbundzentrale des GBV (VZG), in: Kooperative Informationsinfrastrukturen als Chance und Herausforderung. Festschrift für Thomas Bürger zum 65. Geburtstag. Hrsg. v. Bonte, Achim / Rehnolt, Juliane. DeGruyter, Berlin/Boston 2018, S. 211 – 223 (<https://www.degruyter.com/viewbooktoc/product/499349>)

Preprint

drei Bibliothekssysteme aDIS/BMS, Koha und Ex Libris Alma eigene Support-Strukturen. Zum anderen unterstützt das BSZ über entsprechend angepasste Schnittstellen eine Reihe weiterer Bibliotheks- und Dokumentationssysteme, die mit dem Verbundsystem CBS interagieren. Hierzu zählen Systeme wie Libero, Aleph, Alephino, SISIS-SunRise, Bibliotheca, Koha, Allegro, Star, Adlib, LIDOS, AUGIAS und weitere. So bleiben die Einrichtungen mit BSZ-Support bei der Wahl ihres lokal eingesetzten Systems unter technischen Gesichtspunkten frei.

Aufgrund der unterschiedlichen Architekturen der beiden Systemlandschaften von BSZ und GBV mussten bei der Migration auf K10plus an den Verbundzentralen auch unterschiedliche Schritte erfolgen.

CBS – die technische Grundlage von K10plus

Hinsichtlich der Möglichkeiten des Datenmanagements ist das in K10plus eingesetzte CBS-System sehr mächtig. So ist es möglich, forschungsunterstützende Services vergleichsweise leicht aufsetzen. Die ZBW Kiel pflegt den Standardthesaurus Wirtschaft und erschließt die Metadaten für das Fachportal EconBiz³ ausschließlich im CBS. Ein anderes Beispiel ist der FID Theologie in Tübingen. Für die Bibliographie Index Theologicus⁴ (IxTheo) werden innerhalb des CBS-Systems automatische Datenabgleiche mit Datenquellen von Casalini usw. nach Relevanz für den FID vorgenommen, diese Titel automatisch isoliert, normiert und in die Bibliographie eingespielt.

Der Kern des CBS ist das interne Datenformat, das zwar Kategorien und Unterfelder verwendet, im Gegensatz zu bibliothekarischen Formaten wie z.B. MARC oder UNIMARC aber keine Abhängigkeiten von bibliothekarischen Regelwerken kennt. Die Feldstrukturen sind nicht in die Datenbankstruktur implementiert, sondern werden durch ein konfigurierbares Regelwerk zur Präsentation der Inhalte der externen Darstellung definiert. Dies ermöglicht die Reduktion des Datenbankmanagementsystems (DBMS) auf einen „Recordmanager“ mit einer außerordentlichen Performance.

Die Gesamtindexierung aller Daten erfolgt mit der CBS-Software in wenigen Stunden, wobei pro Titel ca. 200-400 Indexeinträge anfallen können. Systeme mit relationalen Datenbanksystemen benötigen mehrere Tage!

Neben der Katalogisierungsdatenbank K10plus hält das CBS noch gehostete Datenbanken, wie z.B. VD17, VD18, IKAR, Hand Press Books, EROMM, Digitalisierte Zettelkataloge, Aufsatzdatenbanken OLC, Springer, Elsevier, Wiley, de Gruyter, JSTOR, Crossref – zusammen ebenfalls mehr als 200 Mio. Titel.

Nicht nur technisch, sondern auch wirtschaftlich hat das CBS-Verbundsystem noch eine lange Zukunft vor sich. Neuverträge für weitere CBS-Installationen konnten seitens OCLC in den vergangenen Jahren erfolgreich abgeschlossen werden, u.a. in Großbritannien JISC, mit weiteren Ländern und Regionen wird verhandelt. Auch hat der personelle Generationenwechsel bei OCLC bereits stattgefunden: heute steht ein junges und

³ Im Fachportal EconBiz ist die weltweit verfügbare Wirtschaftsliteratur nachgewiesen. Mithilfe modernster Suchtechnologien wird der meist kostenfreien Zugang zu Millionen von Online-Dokumenten ermöglicht.

⁴ IxTheo ist die anerkannte internationale wissenschaftliche Open-Access-Bibliographie für Theologie und Religionswissenschaft. Der Index umfasst wissenschaftlich Informationen zeit- und medienübergreifend vom Mittelalter bis zur Gegenwart und wird von der Universität Tübingen verantwortet, gepflegt und weiterentwickelt.

Preprint

dynamisches Team hinter der Weiterentwicklung des Systems. Schließlich setzt OCLC vermehrt auf Webtechnologien, nicht nur durch die Entwicklung zeitgemäßer Plattformdienste wie WMS und die Einbindung in die neuen Syndeo-Strukturen⁵, sondern auch durch die Ablösung tradierter Client-Server-Architekturen. So wird bspw. aktuell die WinIBW, der lokal installierte Katalogisierungsclient, durch ein zeitgemäßes Webinterface (CCWeb) ergänzt. Die ersten Tests fallen positiv aus. Auch die Integration des CBS in die neue von OCLC geplante Systemarchitektur Syndeo schreitet gut voran. Sicherheits- und Funktionsupdates werden mit Vorlauf angekündigt und zuverlässig ausgeliefert. Aus diesem Grund ist es den Verbundzentralen einfach möglich, die hinter K10plus stehende Software immer auf dem aktuellen Stand zu halten. Momentan ist die Version 8.3 in Vorbereitung, mit der CCWeb im Produktivbetrieb voll unterstützt wird.

Mitbestimmung gehört zur Kultur von OCLC. Einmal im Jahr treffen sich die CBS-Anwenderinnen und -Anwender auf Leitungsebene und diskutieren Entwicklungen mit Vertretern des Managements von OCLC, bringen Vorschläge ein und priorisieren diese. Technische Entwicklungen werden gemeinsam abgestimmt.

Abb. 1 verdeutlicht, dass das CBS-Verbundsystem die mächtige technische Grundlage für eine Reihe weiterer Verbundservices über das Katalogisieren hinaus bildet. So werden bspw. schon heute E-Book-Daten sowie die ausgelieferten Artikeldaten für den Discovery-Index K10plus-Zentral über das CBS-System prozessiert.

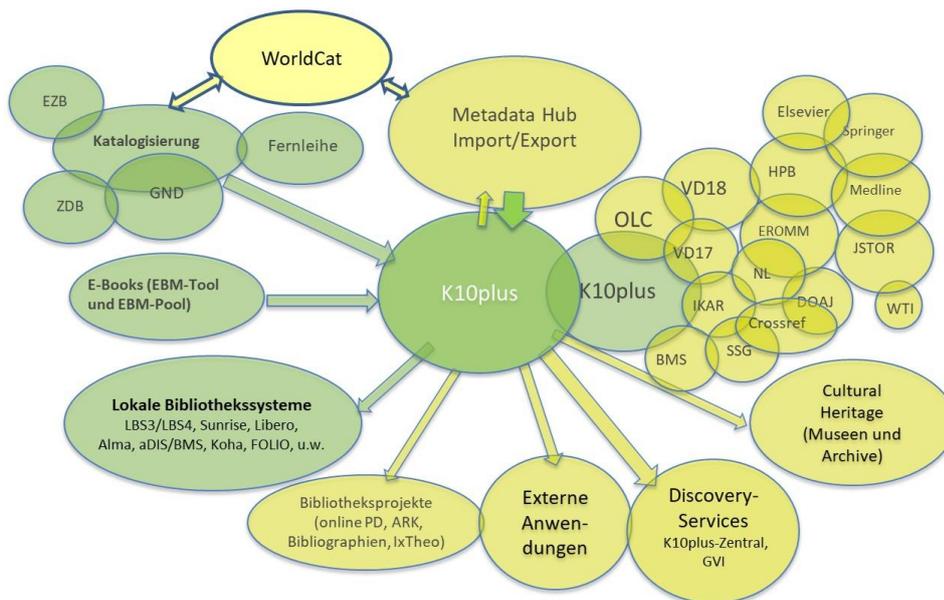


Abb. 1: Daten und Dienste K10plus

Die technische Zusammenlegung der Verbunddaten von BSZ und VZG bedeutet somit mehr als nur die Zusammenlegung von Katalogisierungsdatenbanken. Ganz im Sinn von DFG und Wissenschaftsrat geht es darum, Synergien zu schaffen und die eigenen technischen Strukturen nachhaltig und offen für Innovation und Zukunftsthemen aufzustellen.

Vorbereitung und Migration

Am Anfang der Migration auf K10plus stand die Einrichtung einer adäquaten Projektorganisation sowie die Analyse und Beschreibung der zu bearbeitenden

⁵ siehe <https://www.oclc.org/de/syndeo.html>

Aufgabenfelder. Zunächst wurde eine interne Projektleitung und ein Lenkungsausschuss für die einzelnen themenbezogenen Arbeitsgruppen eingerichtet. Später wurden dann mit dem Koordinierungsausschuss BSZ-GBV auch die Verbundgremien und Ministerien in einem gemeinsamen Gremium eingebunden.

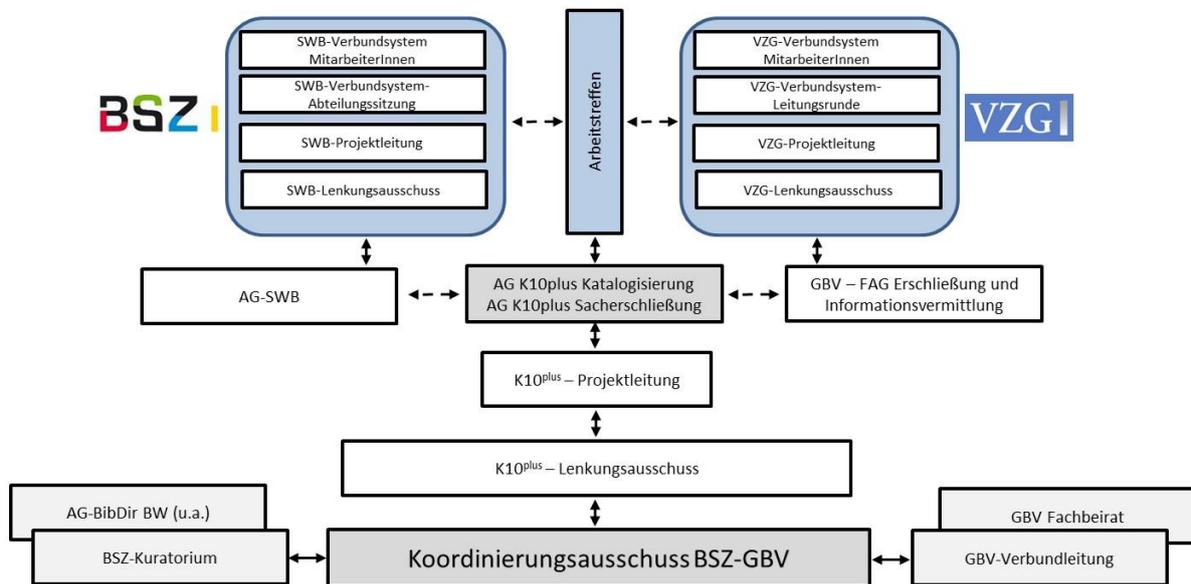


Abb. 2: K10plus Projektorganisation

Als wesentliche Arbeitsfelder wurden identifiziert:

1. Vereinheitlichung des Katalogisierungsformates

Unterschiede zwischen beiden Systemen gab es sowohl in der externen Darstellung des Katalogisierungsformates als auch bei der funktionalen Interpretation von Kategorien. Als gemeinsames externes Format wurde die am BSZ genutzte Unterfelddarstellung gewählt. Statt „4000 Titel : Zusatz“ im GBV wird in K10plus nun „4000 Titel \$dZusatz“ verwendet. Daneben gab es Abweichungen in zentralen Statuskategorien und widersprüchliche Kategorienbelegungen, die bereinigt werden mussten.

Ein umfangreiches und langwieriges Arbeitspaket war auch die Entwicklung des gemeinsamen Indexierungsschemas, da die Suchschlüssel sich über lange Zeit bei der Katalogisierung tief eingepreßt haben und Änderungen nur schwer durchsetzbar waren. Umfangreiche Analysen der Verwendung der einzelnen Suchschlüssel halfen, eine für alle Seiten tragfähige Lösung zu finden.

2. Vereinheitlichung der Anwendungen

Bei der Verwendung von Norm- und Fremddaten gab es sehr unterschiedliche Verfahrensweisen. Im GBV waren z.B. Fremddaten – beispielsweise DNB, Library of Congress (LoC), Casalini – immer in großem Umfang in der Verbunddatenbank integriert, so dass ein Rückgriff auf externe Datenquellen via Z39.50 nur in Ausnahmefällen notwendig war. Im BSZ war die Nutzung via Z39.50 dagegen der Regelfall. In K10plus werden daher nur die aktuellsten Daten aus der DNB, der LoC und von Casalini zur Verfügung gestellt. Das Angebot an externen Datenquellen wurde auf 17 Quellen einschließlich WorldCat erweitert.

3. Zusammenführung der administrativen Daten

Bei der Zusammenführung der administrativen Daten, d.h. der Benutzerkennungen, Bibliotheks- und Rechtfestlegungen, galt es, ein gemeinsames Schema zu finden, das eine möglichst weitgehende Weiternutzung der vorhandenen Daten erlaubte. Dies konnte über Prefix-Buchstaben bei den Benutzerkennungen und der Festlegung von Bereichen für die internen Bibliotheks-IDs erreicht werden.

4. Zusammenführung der bibliographischen Daten

Die wohl ambitionierteste und auch noch nicht abgeschlossene Aufgabe war und ist die Zusammenführung der Titeldaten aus beiden Systemen. Im ersten Schritt konnte dies nur für einen vergleichsweise kleinen Bereich der einbändigen Werke mit verlässlichen Identifikationsnummern durchgeführt werden. U.a. müssen für mehrbändige Werke und Serien zunächst die Überordnungen zusammengebracht werden, ehe die Teile abgeglichen werden können. Ansonsten besteht die Gefahr, dass zusammengehörende Familien auseinander gerissen werden. Die Schwierigkeiten beim Abgleich von Titeldaten ohne identifizierende Nummern sind hinreichend bekannt.

5. Anpassung der Schnittstellen und der Versorgung der Lokalsysteme

Das in weiten Bereichen gegenüber den Altsystemen geänderte Format erfordert die Anpassungen aller Import- und Exportschnittstellen. Da das MAB-Format schon seit längerem nicht mehr weiterentwickelt wird, wurde hier die Gelegenheit genutzt und die Unterstützung eingestellt.

Ein besonderes Problem stellte die Anpassung der Versorgungsschnittstellen für die Lokalsysteme im BSZ-Bereich dar, da sich die Identifikationsnummern der Titel (PPN = Pica Production Number) geändert hatten. Hier wurde ein Mechanismus entwickelt, der für Altdaten die aktuelle PPN gegen die ursprüngliche PPN aus dem BSZ-System austauscht. Im GBV gab es zwar keine Probleme mit den Identifikationsnummern, aber dafür Probleme mit dem parallelen Handling des alten und neuen Formates in den Lokalsystemen, da beide nicht zeitnah vollständig neu geladen werden konnten. Als Lösung wurde eine Rückumwandlung des Formats in den Updatemechanismus integriert.

Für die Bearbeitung der beiden zentralen Themen „Gemeinsames Katalogisierungsformat“ und „Indexierung K10plus“ wurden zwei mehrtägige Workshops in Tagungshotels weitab vom Alltag der beiden Verbundzentralen organisiert. Dieses Vorgehen hat sich als sehr erfolgreich herausgestellt. Die Ergebnisse dieser Arbeitsgruppen waren dann die Grundlage für die Erstellung der Richtlinien und Schulungsunterlagen. Ergänzend wurde so auch das Konzept für die technische Plattform K10plus entwickelt.

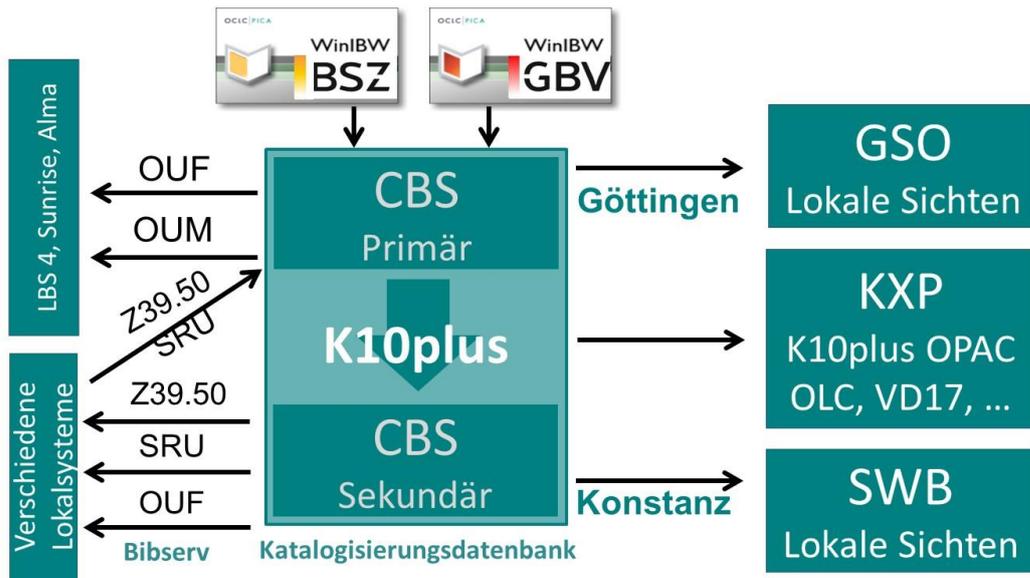


Abb. 3: Technisches Konzept K10plus

Im Kern besteht das System aus einer gespiegelten Katalogisierungsdatenbank an den Standorten Göttingen (VZG) und Konstanz (BSZ). Alle schreibenden Zugänge erfolgen primär in der VZG-Instanz und werden mittels eines neu entwickelten Synchronisierungsmechanismus nach Konstanz übermittelt. Die beiden Instanzen dienen darüber hinaus dazu, die Versorgung der lokalen Bibliothekssysteme jeweils lokal zu organisieren und die bewährten Auftritte im Netz sicherzustellen.

Während das Produktivsystem in einer VMware-Umgebung auf virtuellen Servern läuft, wurde für die Entwicklung und Zusammenführung eine extrem leistungsfähige eigenständige Hardwarekonfiguration verwendet. Diese bestand aus einem leistungsfähigen Server mit sehr großem Speicher, 4 TB SSDs und weiteren internen Platten. Insgesamt wurden 5 vollständige Probedurchgänge gefahren. Die endgültige Zusammenführung der beiden Verbunddatenbanken konnte dann technisch in lediglich drei Tagen durchgeführt werden, wobei für den gesamten Prozess aus Sicherheitsgründen eine Zeitspanne von zwei Wochen vorgesehen war.

Gesamtzugriff der GVK-Daten erstellen, kopieren	02:00 h
Vorbereitende Änderungen: 9,4 Mio.	08:00 h
Konvertierung GVK-Daten	
50,6 Mio. Titel	07:46 h
11,0 Mio. Normsätze	02:30 h
135,0 Mio. Exemplarsätze	18:48 h
Übernahme in Produktiv-Datenbank („Bulkload“)	03:00 h
Nachbearbeitung GVK-Daten (ca. 60 Mio. Änderungen)	15:00 h
Neuindexierung	04:00 h
Gesamtzugriff der SWB-Daten, konvertieren, kopieren, ca. 22 Mio. Titelsätze, 75 Mio. Exemplarsätze, RVK	11:00 h
Einlesen der SWB-Daten	04:30 h
Laden der SWB-Daten	05:00 h
Deduplizierung (8,9 Mio. Titel)	18:00 h

Nachbearbeitung
Indexierung

07:30 h
04:00 h

<u>Mat</u>	<u>Gesamt</u>	<u>Besitz</u>	<u>Nachweise</u>	
Ac	2.031.913	1.986.555	5.707.634	MBW Überordnungen
A*	39.508.174	38.051.354	104.550.853	Monographien Print
Ab	1.404.048	937.070	3.327.827	Zeitschriften Print (ZDB)
Ad	266.785	235.044	2.591.831	Serien Print (ZDB)
Ob	216.053	198.094	6.424.893	E-Journals (ZDB)
	1.886.886	1.370.208	12.344.551	
*v	2.802.598	2.769.114	4.367.215	Zeitschriften-Bandsätze
As	10.847.659	10.784.013	15.022.229	Aufsätze
Os	1.093.039	1.086.439	1.775.044	Aufsätze online
O*	8.880.200	6.795.543	30.705.170	E-Books
E*	1.022.820	1.019.181	2.014.425	Mikroformen
K*	826.072	820.498	1.168.520	Karten
M*	1.358.788	1.349.725	2.096.504	Musikalien
B*	1.485.122	1.407.273	2.003.669	Audio
...				
	-----	-----	-----	
	72.597.387	68.268.388	183.002.407	= ca. 210 Mio.

Abb. 4: K10plus am 25. März 2019

Die Katalogisierung in den beiden Altsystemen wurde am 15. März 2019 beendet und in K10plus am 25. März 2019 wieder aufgenommen. Damit war der erste Schritt zur Zusammenführung der Systeme von BSZ und GBV geschafft.

Direkt anschließend wurde die Online-Fernleihe für die beiden Leihverkehrsregionen wieder in Betrieb genommen. Eine besondere Herausforderung war dabei der Umstand, dass im BSZ als technische Basis der externe ZFL-Server, im GBV aber das in das CBS integrierte Fernleihsystem genutzt wird, d.h. in dem gemeinsamen Verbundsystem musste eine verbundübergreifende Fernleihe abgebildet werden.

Nach dem Start von K10plus standen und stehen noch eine ganze Reihe von Nacharbeiten an. Wichtige Vorhaben waren u.a. der Abgleich der Personen und Körperschaften bei zusammengeführten Titeln wegen der Anzahl von Kategorien und GND-Verknüpfungen, die Vereinheitlichung der Verwendung von Sacherschließungskategorien und die Verwendung der RVK-Normdatei. Daneben gab und gibt es noch eine Reihe von kleineren Anpassungen an das neue Format. Insgesamt ergaben sich in den ersten Monaten mehr als 50 Mio. Updates in der Verbunddatenbank.

Nachdem bereits frühzeitig der E-Book-Pool und das zugehörige E-Book-Management-Tool (EBM-Tool) als gemeinsamer Service erfolgreich angeboten werden konnte, wurde die Zusammenführung der Workflows für die Versorgung der Bibliotheken mit hoher Priorität angegangen.

Preprint

Nach der Umstellung der Katalogisierung waren noch einer Reihe weiterer Datenbanken nach K10plus zu überführen. Mit Priorität wurden VD17, IKAR und BMS⁶ umgestellt, da diese direkt Daten aus der K10plus-Katalogisierung übernehmen.

Nicht ganz so zeitkritisch war die Migration der Aufsatzdatenbanken, die im Wesentlichen für den Discovery-Index K10plus-Zentral bestimmt sind. Die Aktualisierung erfolgte zunächst weiterhin im Altsystem des GBV. Noch nicht gelöst ist die erforderliche neue Konzeption der Online Contents Datenbank (OLC) mit mehr als 50 Mio. Nachweisen von Aufsätzen, die direkt mit den Zeitschriftentiteln verknüpft und über die Fernleihe bestellbar sind.

Technik und Services an der VZG

Die VZG betreibt eine eigene technische Infrastruktur im Gebäude der SUB Göttingen und nutzt einen Volumenanteil am Gigabit-Anschluss der Universität Göttingen. Als ein davon unabhängiger „Notzugang“ wird ein 100 Mbit-Kabelanschluss vorgehalten. Die gesamte Infrastruktur ist redundant ausgelegt und auf zwei getrennte voneinander unabhängige Räume mit unabhängigen unterbrechungsfreien Stromversorgungen aufgeteilt.

Die für K10plus relevante Infrastruktur basiert auf einem VMWare-Cluster mit 8 HP BL460c und 260 aktiven virtuellen Servern.⁷ Daneben werden weitere Cluster für die GSO⁸-Web-Datenbanken mit 3 Instanzen für Online Recherche und Fernleihe sowie für schnittstellenbasierte Zugriffe (Z39.50/SRU) und für den Discovery-Index K10plus-Zentral mit über 200 Mio. Titelsätzen betrieben.

Für das Hosting von 23 Lokalsystemstandorten mit LBS4 werden noch Oracle T4-Server unter Solaris benötigt. Die erforderliche Speicherinfrastruktur ist weitgehend virtualisiert (DataCore) und wird zurzeit durch eine neue auf Ceph basierende Infrastruktur abgelöst. Neben dem eigentlichen Verbundsystem K10plus werden weitere CBS-Instanzen für Schulung, Entwicklung und Test vorgehalten.

Neben den auf CBS basierenden Services betreibt die VZG auf dieser Infrastruktur eine Reihe weiterer Services. Zu nennen sind hier die Discovery-Plattform K10plus-Zentral, das Hosting von Lokalsystemen, der LBS-Service für mehr als 80 Bibliotheken sowie das Hosting für extern betreute Anwendungen wie z.B. Goobi, Dandelon, Repositorien und Digitalisierungsanwendungen.

⁶ Bibliographie des Musikschritttums online

⁷ Eine detaillierte Aufstellung ist im Jahresbericht 2018 der VZG ab Seite 48 enthalten (https://www.gbv.de/Verbundzentrale/Publikationen/PDF/jahresbericht_2018)

⁸ GSO GBV Search & Order

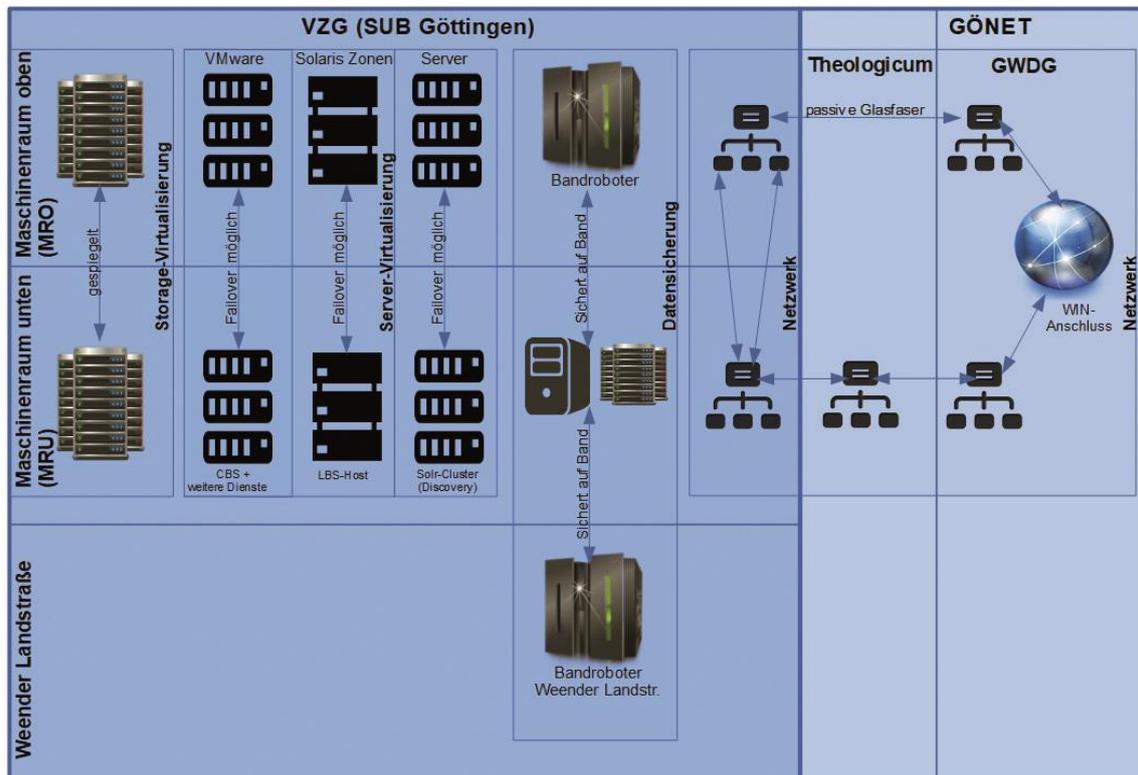


Abb. 5: Technische Infrastruktur der VZG

Technik und Services am BSZ

Das BSZ betreibt keine eigenen technischen Infrastrukturen. Stattdessen nutzt es Kapazitäten der universitären Rechenzentren Baden-Württembergs. Während also das SWB-Verbundsystem (bzw. der K10plus-Spiegel) am Rechenzentrum der Universität Konstanz (KIM) betrieben wird, läuft das landesweite Bibliothekssystem aDIS/BMS an der Universität Tübingen. Aufgrund der Trennung der Schichten von Verbundsystemebene und Bibliothekssystemebene ist das kein Problem. Handlungsfelder wie Backup, Speicherbedarf, Langzeitarchivierung usw. werden ebenso mit Unterstützung universitärer Rechenzentren des Landes Baden-Württemberg realisiert. Die durch das BSZ betriebene Alma-Network-Zone läuft technisch in einem Rechenzentrum des Anbieters in Amsterdam, K10plus läuft primär in Göttingen an der VZG.

Der K10plus-Spiegel in Konstanz versorgt aus Lastgründen die OPACs (lokale Sichten bzw. PSI-Sichten) im SWB-Einzugsbereich. Bei längerem Ausfall an einem Standort können OPACs auf den anderen Standort verlagert werden. Nach dem derzeitigen Ausbaustand könnte die Verlagerung zur einer leichten Performance-Einbuße bei den Reaktionszeiten der OPACs führen. Dies scheint derzeit unter Kosten-Nutzen-Betrachtungen ein angemessenes Notfall-Szenario. Zudem werden standardmäßig einige wenige allgemeine K10plus-OPAC-Installationen als Auffanglösung an beiden Standorten redundant aktiv vorgehalten. Diese kommen bspw. bei Wartungsfenstern zum Einsatz.

Das BSZ selbst fährt neben dem CBS-Spiegelsystem noch zwei weitere Systeme für die OPACs. Mit zwei Systemen kann bei Ausfall des einen sofort auf das zweite System

umgeschaltet werden. Auch bei Wartungsarbeiten steht damit immer mindestens ein System zur Verfügung.

Bei der Betrachtung der technischen Rolle des BSZ in K10plus kommt der Echtzeit-Spiegelung der K10plus-Daten von der VZG an das BSZ zentrale Bedeutung zu. Um die Abgleich-Prozeduren so schlank wie möglich zu halten, haben sich BSZ und VZG für ein Primary-Secondary-Modell bei der Synchronisation der beiden Standorte entschieden. Dies bedeutet, dass alle aktiven Katalogisierungsvorgänge sowie maschinelle Einspielungen primär in der CBS-Installation der VZG in Göttingen erfolgen. Dies betrifft im Bereich des SWB-Supports unter anderem den von den Bibliotheken benutzten Erfassungsclient WinIBW mit seinen sehr vielfältigen Skriptmöglichkeiten, schreibende Schnittstellen bei Titeldaten für die Systeme des Deutschen Literaturarchivs Marbach und des Germanischen Nationalmuseums Nürnberg, schreibende Schnittstellen bei Exemplardaten für unterschiedliche Systeme wie z. B. die Bibliothekssysteme Koha, aDIS/BMS oder Ex Libris Alma sowie die schreibenden OAI-Schnittstellen für Daten aus der ZDB oder der GND und Offline-Importe von Bibliotheksdaten.

Das Synchronisationsverfahren zwischen Göttingen und Konstanz erfolgt online. Die verfügbaren Bandbreiten des Forschungsnetzes reichen hierfür vollkommen aus. Änderungen an einem Datensatz in Göttingen werden somit zuverlässig im Zeitbereich bis zu einer Sekunde in Konstanz repliziert. Eine Änderung auf dem Server in Göttingen ist damit eine Sekunde später über den Server in Konstanz online abfragbar. Auch wenn die Synchronisationszeit bei manuellen Schreibvorgängen unter einer Sekunde liegt, so ist dies angesichts der vielen tausend Clients an den Bibliotheken doch vergleichsweise lang. Es kann passieren, dass an einem manuell geänderten Datensatz nur Bruchteile von Sekunden später weitere Änderungen vorgenommen werden, die ebenfalls synchronisiert werden müssen. Technisch gelöst wurde dies durch eine serielle Abarbeitung der entsprechenden Änderungen des Datensatzes. Der Spiegel in Konstanz ist zudem schreibgeschützt. Inkonsistenzen werden auf diese Weise logisch vermieden.

Erreicht wird die hundertprozentige Zuverlässigkeit bei der Synchronisation manueller bearbeiteter Datensätze auch durch die dynamische Zuweisung von Systemprioritäten. Sollten also zeitgleich maschinelle Einspielungen durch eine Verbundzentrale vorgenommen werden, erhalten manuelle Eingaben automatisch die höhere und die maschinellen Änderungen automatisch eine niedrigere Priorität. Aus diesem Grund nehmen BSZ und VZG maschinell durchgeführte Datenveränderungen sowie Einspielungen in der Regel zu Zeiten vor, in denen die manuellen Eingaben relativ gering sind, also an Wochenenden oder in der Nacht.

Neben den Daten selbst werden auch Logfiles, die für die Wiederherstellung früherer Systemzustände und für die Versorgung der Lokalsysteme mit relevanten Änderungen von Bedeutung sind (sowie Anpassungen an den Systemparametern) mit synchronisiert. Pro Tag werden ca. 600.000 Änderungen an der Datenbank vorgenommen! Hiervon entfällt etwa die Hälfte auf den Bereich der Titeldaten (Änderungen, Korrekturen, Ergänzungen, Löschungen). Die andere Hälfte betrifft Änderungen an den Lokal- bzw. Exemplardaten.

Aktuell wird auf dem Konstanzer Spiegel nicht nur der SWB-OPAC betrieben, sondern viele lokale Kataloge wie zum Beispiel der OPAC der Saarländischen Universitäts- und Landesbibliothek, einige regionale Kataloge und unterschiedliche SWB-Bibliographien und – Fachkataloge. Auch die deutschlandweit genutzte OGND (Zugriff auf die Gemeinsame Normdatei) wird über den Konstanzer Spiegel bereitgestellt.

Preprint

Das gespiegelte CBS-System in Konstanz läuft wie das System in Göttingen virtualisiert in einer ESX-Umgebung am KIM der Universität Konstanz. Daher wird das System wegen der sehr hohen Performanz-Anforderung bevorzugt im ESX-Cluster bedient. Derzeit stehen für K10plus 24 CPUs und 192 GB RAM zur Verfügung. Ein Plattenplatzbedarf bis zu 6 TB ist reserviert (4 TB werden derzeit genutzt).

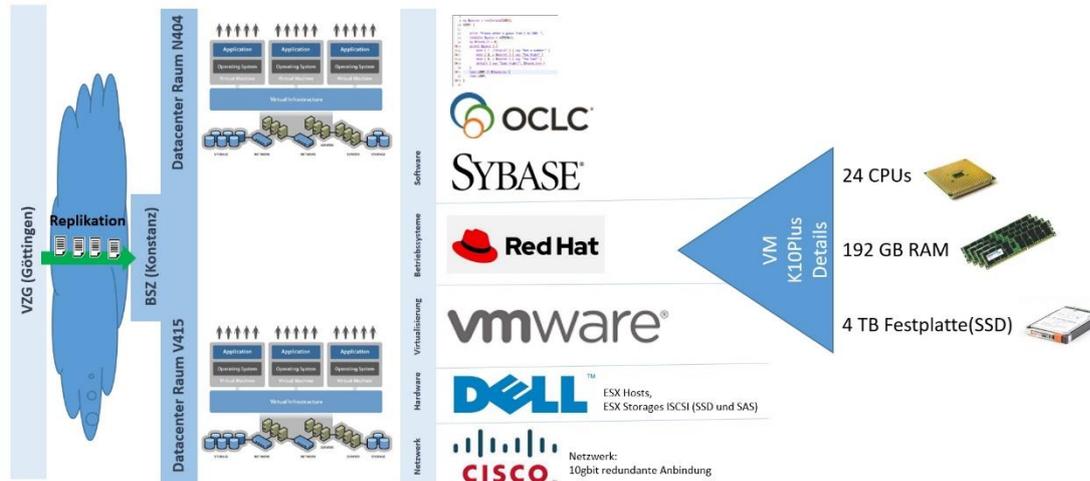


Abb. 6: Technische Infrastruktur des BSZ für K10plus

Zusammenarbeit

Für K10plus wurden bestehende Arbeitsgruppen der Verbundlandschaften BSZ und GBV reformiert und zu neuen gemeinsamen Arbeitsgruppen zusammengefasst. Eine gemeinsame neue Geschäftsordnung wurde für diese verabschiedet. Bisher haben zwei Arbeitsgruppen die Arbeit aufgenommen: die AG K10plus Katalogisierung und die AG K10plus Sacherschließung. Vertiefende Informationen zu den Aufgaben und Tätigkeiten finden sich im gemeinsamen K10plus-Wiki.⁹

Die zuständigen Gremien und Ministerien von BSZ und VZG sind in einem gemeinsamen Koordinierungsausschuss BSZ-GBV vertreten. Werden in den Arbeitsgruppen operative Fragen diskutiert, so stehen im Koordinierungsausschuss strategische Themenstellungen im Vordergrund.

In der alltäglichen Zusammenarbeit von BSZ und VZG kann man grob zwei Fälle unterscheiden:

Fall 1: Das BSZ oder die VZG übernehmen Services für den jeweils anderen Partner und für dessen Verbundregion. Es werden keine redundanten technischen oder gedoppelten personelle Strukturen vorgehalten. Ein Beispiel ist die Aufteilung der Unterstützung von Lokalsystemen. Die VZG bietet aktuell die Unterstützung von LBS4 der Firma OCLC sowie perspektivisch die Open Source Lösung FOLIO an, das BSZ bietet aktuell die Unterstützung von aDIS/BMS, Koha und Ex Libris Alma an. FOLIO wird durch das BSZ beobachtet. Auch Digishelf, die gemeinsam verantwortete digitale Bibliothek, ist ein gutes Beispiel: diese betreut die VZG alleinverantwortlich, bspw. auch den Auftritt der Bodenseebibliotheken im regionalen Umfeld des BSZ.¹⁰ Supportanfragen auch aus Baden-Württemberg werden durch

⁹ <https://wiki.k10plus.de/display/K10PLUS/Gremien>

¹⁰ <https://www.digishelf.de/bodensezeitschriften/>

Mitarbeitende der VZG in Göttingen bearbeitet. Dagegen wird bspw. das in K10plus genutzte EBM-Tool technisch durch das BSZ verantwortet.

Fall 2: Das BSZ und die VZG teilen sich die Verantwortung bei gemeinsamen Services. Das bedeutet, dass Services hinsichtlich der personellen Strukturen arbeitsteilig an beiden Standorten organisiert werden. Ein wichtiges Beispiel ist natürlich K10plus selbst. Hier ist die Bandbreite der geteilten Verantwortung groß: es geht um Datenbankstrukturen und Datenkorrekturprozesse, um den Einsatz und die Weiterentwicklung gemeinsamer Skripte, um die Planungen zu einer einheitlichen WinIBW, um den gemeinsamen K10plus-Katalog, um Anpassungen und Vereinheitlichung von Workflows, um den Aufbau von Schulungsmaterial für die gemeinsame Nutzung und um vieles mehr.

In beiden Verbundzentralen steht – wie bei vielen Bibliotheken auch – in den nächsten zehn Jahren ein umfangreicher Generationenwechsel bevor. Die Zusammenarbeit von BSZ und VZG ist auch unter diesem Gesichtspunkt zu verstehen: in welchen Bereichen bündeln wir künftig Arbeitsprozesse und Kompetenzen an einem Ort, und in welchen Bereichen verteilen wir diese auf unsere beiden Standorte? Zur Frage der künftigen Ausrichtung und Strategie finden hierfür getrennte Workshops mit dem Fokus auf die eigene Einrichtung statt. Die VZG organisiert zwei Workshops für ihre Leitungsebene, das BSZ zunächst einen Workshop für die Mitarbeitenden. Im Rahmen eines gemeinsamen Treffens wollen wir im Sommer 2020 die gefundenen Positionen miteinander abgleichen und synchronisieren. Selbstverständlich handelt es sich um einen andauernden Prozess, der mit diesen ersten Workshops nicht zu einem Ende kommen wird.

In der Rückschau auf die vergangenen Jahre wird deutlich, dass Themen wie das Zusammenwachsen der Mitarbeitenden, das Entwickeln von Verständnis entwickeln für den anderen oder das sich Bewusstmachen der je anderen „Hauskultur“ größere Bedeutung zukam als der Diskussion über rein technische Themen. Tatsächlich war und ist K10plus weniger ein technisches als ein organisatorisches Projekt, in dem Menschen und Organisationen mit unterschiedlichen Geschichten zusammenkommen um eine gemeinsame Erzählung zu entwickeln. Dieser Prozess ist heute noch in vollem Gang. Immer wieder gibt es Dissens zwischen den Partnern, den es auszuhalten und in tragfähige Kompromisse umzuwandeln gilt. Aber die Gemeinsamkeiten überwiegen. Neben vielen Telefon- und Videokonferenzen tauschen wir uns in regelmäßigen Treffen in Stuttgart aus, suchen nach Synergien und neuen gemeinsamen Workflows und verabreden Standards. Im Rahmen einer internen Fachtagung mit allen Mitarbeitenden beider Verbundzentralen konnten wir uns im Herbst 2019 besser kennenlernen. Dies wollen wir alle zwei Jahre durchführen.

Die Zusammenarbeit zwischen BSZ und VZG basiert auf einer Kooperationsvereinbarung vom September 2015. Grundlage der Vereinbarung ist es, unter Beibehaltung der Autonomie der Einrichtungen und ihrer Steuerungsstrukturen kollegial und zum Nutzen für die Wissenschaft zusammenzuarbeiten. Hierfür wurden Handlungsstränge festgelegt. In einigen Feldern konnten dann sehr rasch und erfolgreich gemeinsame Services aufgesetzt werden. In anderen Bereichen arbeiten wir noch intensiv. Der „größte Brocken“, also die Zusammenlegung der Verbundkataloge von BSZ und GBV, wurde im März 2019 realisiert, also etwa dreieinhalb Jahre nach dem Abschluss der Vereinbarung.

Unser Dank gilt natürlich den Bibliothekarinnen und Bibliothekaren, die die eine oder andere Unzulänglichkeit gerade in der Startphase von K10plus mit uns ertragen, gemeistert und Lösungen gestaltet haben. Der besondere Dank gilt aber klar den Mitarbeitenden von BSZ und VZG, die in der vergleichsweise kurzen Zeit zwei Jahrzehnte Verbundgeschichte mit

Preprint

den entsprechenden unterschiedlichen Strukturen homogenisiert und zu K10plus zusammengefügt haben. Auf die Kolleginnen und Kollegen in unseren Häusern sind wir überaus stolz! Ohne das unglaubliche Engagement der Beschäftigten an beiden Standorten wäre K10plus nicht möglich gewesen.

Zukunft

Nachdem auch OCLC Dublin (Ohio) die Bedeutung des CBS für die Einbindung der regionalen Serviceorganisationen und der damit verbundenen Bibliotheken erkannt hat, wurde unter dem Label „CBS Longevity“ ein Programm zur Modernisierung und langfristigen Stabilität der CBS-Plattform gestartet. Neben der bereits erwähnten Neuentwicklung einer browserbasierten Katalogisierungsumgebung war die Umstellung der Betriebssystemumgebung von Oracle-Solaris auf Red Hat Linux ein erster wichtiger Schritt. Mit der in Kürze anstehenden Migration der Oracle Sybase-Datenbank auf PostgreSQL wird dann auch die letzte Abhängigkeit von Softwarekomponenten, deren Supportende absehbar ist, beseitigt. Da PostgreSQL Open Source ist, besteht auch keine Gefahr überraschender Steigerungen von Lizenzkosten. Im Anschluss an die Migration des CBS wird auch das Datenbanksystem der LBS-Systeme in der GBV-Region nach PostgreSQL migriert.

Für den Migrationsprozess wirkt sich ebenfalls sehr positiv aus, dass für die CBS-Software die Datenbankabhängigkeiten in einer Programmbibliothek gekapselt sind. Damit ist im Prinzip nur der Austausch dieser Bibliothek ohne Änderung der eigentlichen Softwarekomponenten erforderlich. Diese Architektur hat sich in der Vergangenheit bereits bei der Migration von Tandem Enscribe¹¹, ein indexsequentielles Dateiensystem, nach Sybase bewährt. Da PostgreSQL bereits an beiden Verbundzentralen für andere Anwendungen im Einsatz ist, werden hier keine grundsätzlichen Schwierigkeiten erwartet.

Ein weiterer wichtiger Schritt für die Zukunft des CBS ist die Erweiterung des internen Datenmodells um moderne hierarchische Strukturen, wie sie in Formaten wie z.B. JATS (Journal Article Tag Suite) selbstverständlich sind. Mit dieser Entwicklung wird sichergestellt, dass künftig alle Arten von zu beschreibenden Metadaten, insbesondere auch aus dem Bereich der Forschung, aber auch von Archiven, Museen und anderen Wissens- und Kultureinrichtungen, adäquat mit den Werkzeugen des CBS prozessiert werden können.

Auch wenn K10plus und die Discovery-Komponente K10plus-Zentral (Index) nach außen noch getrennt erscheinen, so werden die Metadaten für beide Anwendungen bereits heute über dasselbe CBS in ein einheitliches Format umgesetzt. Die Daten werden normiert und für weitere Services wie eben K10plus-Zentral, aber auch für FIDs, lokale Discovery-Indices und andere Anwendungen bereitgestellt. Künftig sollen K10plus und K10plus-Zentral daher auch in einem einzigen gemeinsamen Frontend für die Recherche zur Verfügung stehen. Dies wird ein weiterer wichtiger Beitrag zur Konsolidierung der Katalogsituation in Deutschland und ein neues Feld der Zusammenarbeit der beiden Verbundzentralen sein.

BSZ und VZG sind Partner aus tiefster Überzeugung. Gemeinsam wollen wir unsere Services weiter professionalisieren, gemeinsam wollen wir neue innovative Dienste schaffen. Profitieren werden Bibliotheken, Wissenschaft und Forschung. Und wir sind offen für weitere Zusammenarbeit. Für Hochschulen, Forschung, Bibliotheken und Verbundzentralen ist K10plus ein attraktives Angebot!

¹¹ siehe <https://en.wikipedia.org/wiki/Enscribe>