

# **K10plus** **der Katalog für zehn Bundesländer,** **die Stiftung Preußischer Kulturbesitz** **und weitere Einrichtungen**

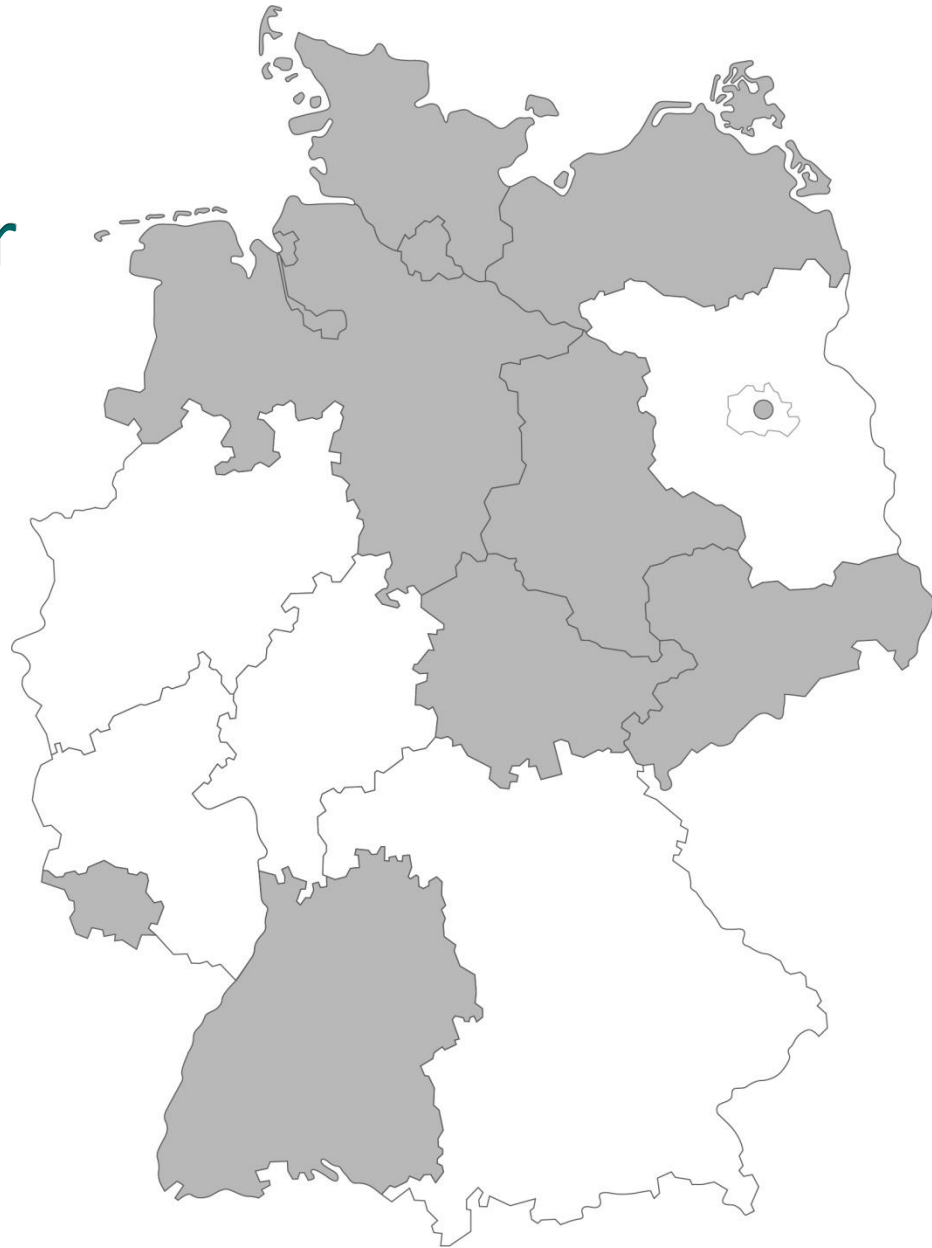
---

**BSZ  GBV**

18. BSZ-Kolloquium am 19.09.2017

---

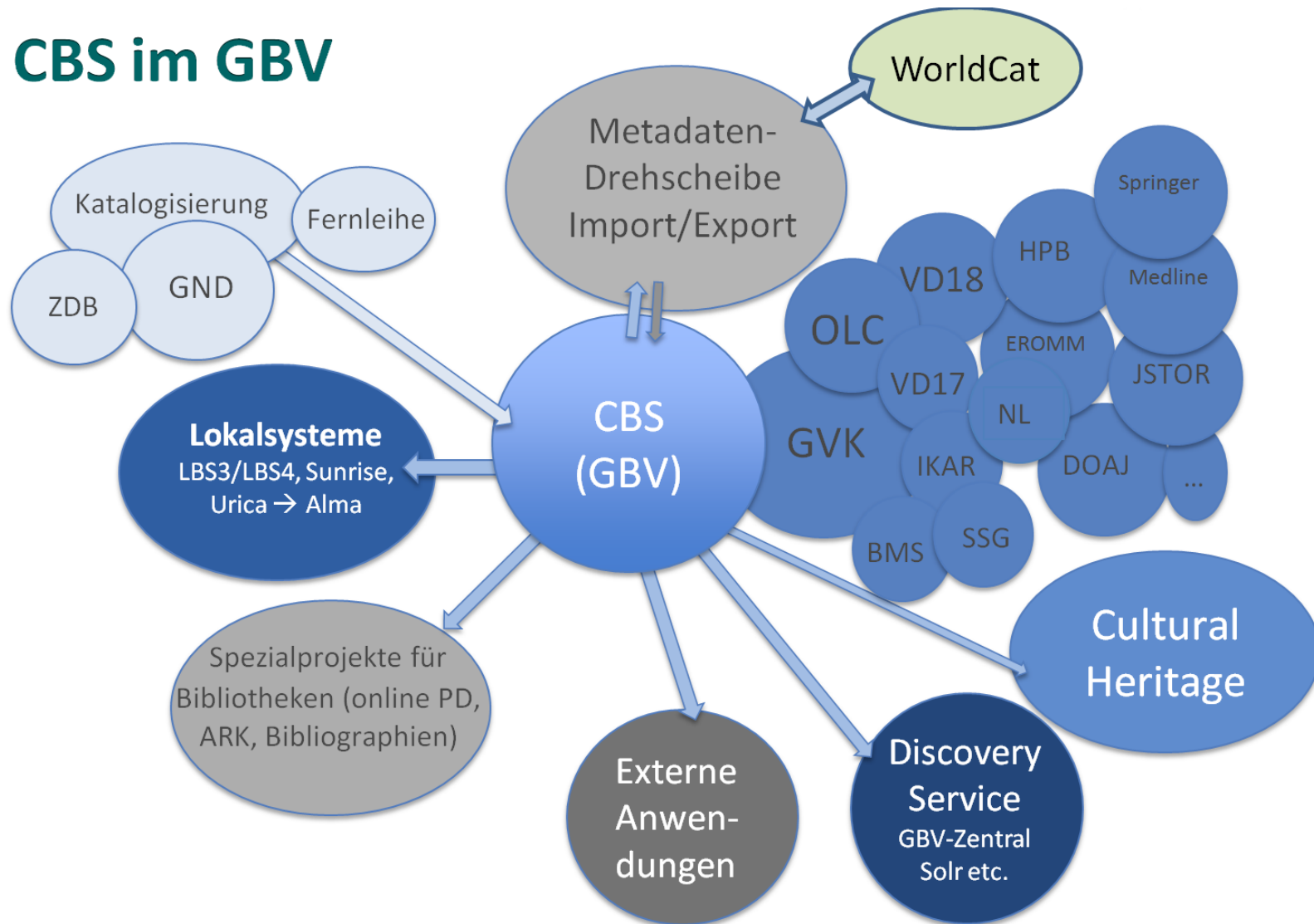
10 Bundesländer,  
Stiftung Preußischer  
Kulturbesitz,  
sowie weitere  
Einrichtungen



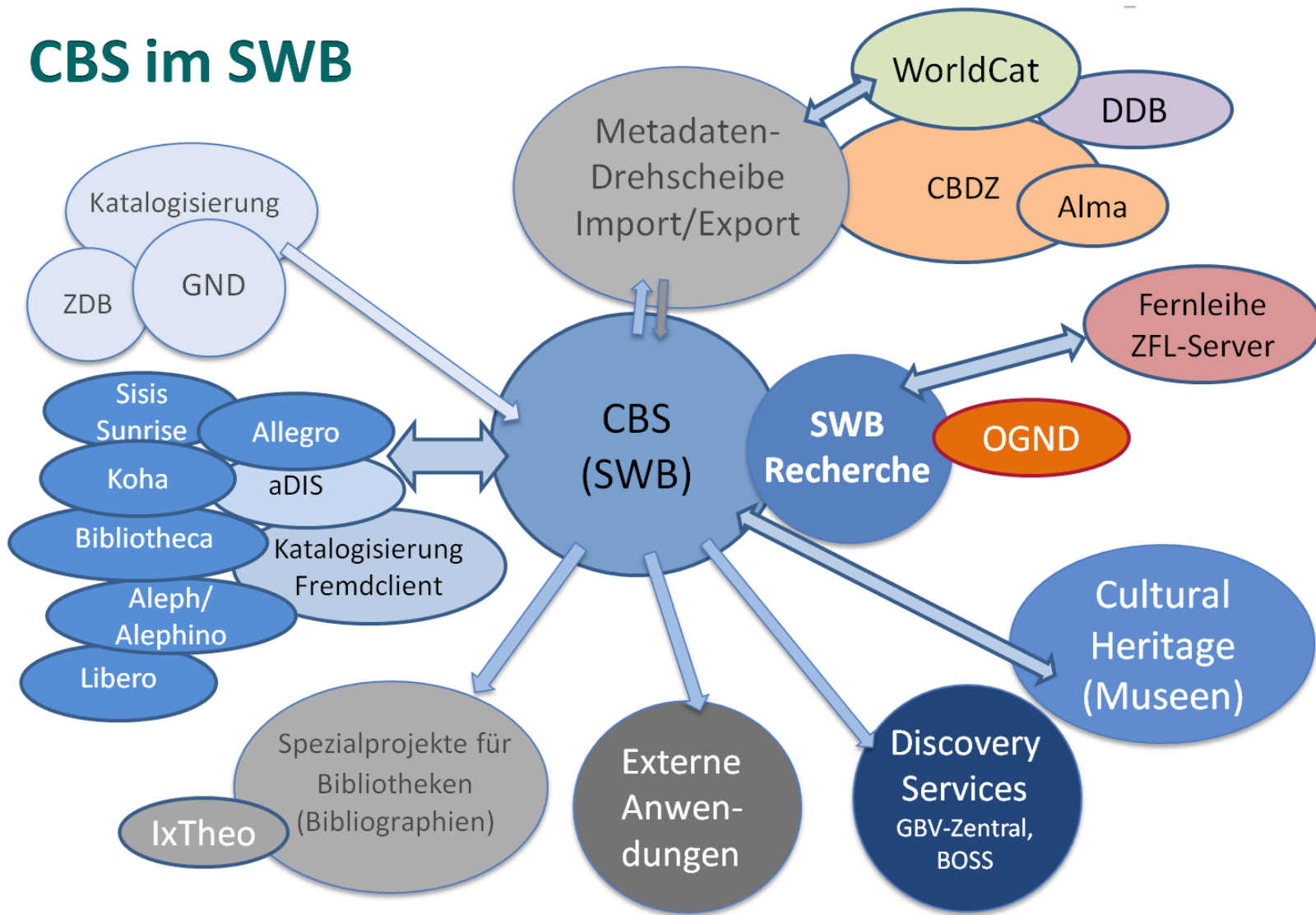
# Zusammenführung zweier „ähnlicher“ Verbundsysteme

- CBS-Software von OCLC
  - Gleicher Programmcode
  - Abweichende Parametrierung
  - Nutzung unterschiedlicher Funktionsumfänge (z.B. Fernleihe)
  - Unterschiedliche Anbindung der Lokalsysteme
- Katalogisierungsformat
  - Deskriptionszeichen (GBV) versus Unterfelder (SWB)
  - RDA-Umstieg: Verwendung Unterfelder im GBV
  - Abweichungen im Internformat
- Unterschiedliche „Katalogisierungskulturen“
  - BSZ: „Regelwerkstreu“, wenig Ausnahmen
  - GBV: viele über das Regelwerk hinausgehende Ergänzungen
  - Unterschiedliche Philosophien bei Normdaten

# CBS im GBV



# CBS im SWB

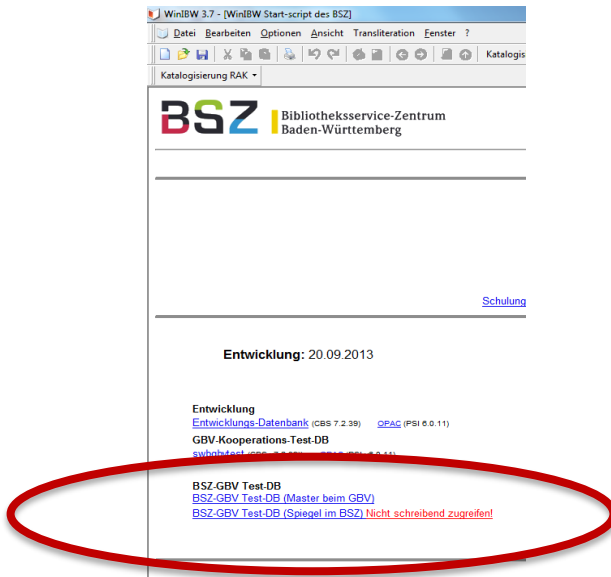
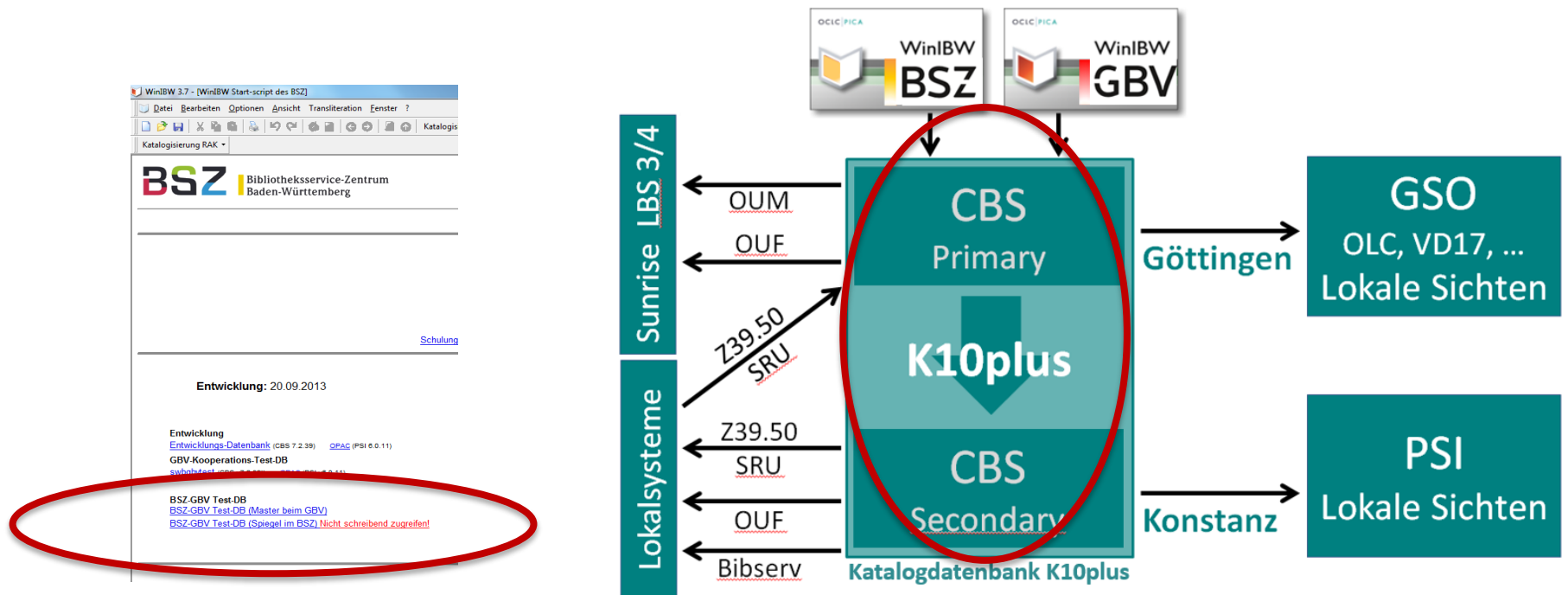


# Aktivitäten in 2017

## Technische Implementierung K10plus

- Synchronisiertes Testsystem in Göttingen und Konstanz **erledigt**
- Gemeinsame Entwicklungsumgebung und erste Konfigurationen z.B. Formattabelle **in Arbeit**

### CBS in K10plus: technisches Konzept



# Aktivitäten in 2017

- Analyse und Konzeption der administrativen Daten mit Rechtevergabe (wer darf was in K10plus?) **erledigt**
- Festlegung der Filesets (Datenbankbereiche) **erledigt**
- Deduplizierung: Grobkonzept zur Zusammenführung gleicher Titel- und Normdaten **erledigt**
- Migrationsszenario **erledigt**
- Identnummernkonzept **erledigt**
- Analyse und Konzept für die Indexierung **in Arbeit**
- Abgleich, Analyse, Definition wichtiger Systemtabellen in Arbeit (Workshop zur Systemkonfiguration Juni 2017) **in Arbeit**

# Deduplizierung

Grobkonzept zur Zusammenführung gleicher Titel- und Normdaten

- Deduplizierung von Daten, die über Identifier wie ISBN, DNB-Idn, ZDB-Idn, GND-Idn verfügen.
- Matching nicht nur über Identifier, sondern mit weiteren Kriterien wie Titel, Personen, Körperschaften, Auflage, Jahr usw.
- Merge von gleichen Sätzen kann sehr granular erfolgen (z.B. Ergänzung von Unterfeldern)
- Für alle nicht beim Match&Merge-Verfahren verarbeiteten Titeldaten wird das Verfahren des Master-Record-Modells angewandt



# Migrationsszenario

Ziel: möglichst kurze Schließungszeit für die Katalogisierung

- Export der Daten aus dem jeweiligen System in Pica+, Grundbestand + 1-n Updates aus dem jeweiligen System
- Konvertierung in das künftige K10plus Pica-Internformat
- Laden der Daten in die gemeinsame Anwendung  
Reihenfolge:  
GBV-Daten (Beibehaltung der PPNs),  
dann SWB-Daten
- Deduplizierung für bestimmte Datenmengen
- Aufsetzen eines einmaligen Prozesses zur Clusterung der nicht zusammengeführten Daten entsprechend dem Verfahren des Master Record Models
- Öffnung der gemeinsamen Datenbank für den Routinebetrieb

# Identnummernkonzept

- Erhalt der PPNs und EPNs aus dem Import der Daten aus der GBV-Datenbank (Grund: PPN ist Primärschlüssel sowohl im Verbund- als auch in den LBS-Systemen)
- SWB-Daten werden entweder
  - als neue Datensätze in K10plus angelegt
  - oder über Match&Merge mit dem GBV-Satz zusammen gespielt

Pica+	Pica3	Benennung	Bereich	Anmerkung
003@ \$0	0100	GBV und K10plus PPN	Titeldaten	
003S \$0	0110	SWB-PPN	Titeldaten	
003@ \$0	797	GBV und K10plus PPN	Normdaten	
003S \$0	798	SWB-PPN	Normdaten	
003D \$0	2119	PPN des umgelenkten gelöschten GBV- bzw. K10plus-Satzes	Titeldaten	
003D \$0	796	PPN des umgelenkten gelöschten GBV- bzw. K10plus-Satzes	Normdaten	
006S \$0	2111	SWB-PPN des umgelenkten gelöschten SWB-Satzes	Titeldaten	
006S \$0	799	SWB-PPN des umgelenkten gelöschten SWB-Satzes	Normdaten	
203@ \$0	7800	EPN des GBV und K10plus Exemplarsatzes	Exemplardaten	
203S \$0	7801	SWB-EPN des Exemplarsatzes	Exemplardaten	

Titel stimmt überein

SWB	GBV	K10plus
<ul style="list-style-type: none"> <li>•Titel 003@ \$0485937565 007G \$2GBV\$0882685260</li> <li>•Exemplarsatz 203@ \$0962659215</li> </ul>	<ul style="list-style-type: none"> <li>•Titel 003@ \$0882685260</li> <li>•Exemplarsatz 203@ \$01675704376</li> </ul>	<ul style="list-style-type: none"> <li>•Titel 003@ \$0882685260 003S \$0485937565</li> <li>•Exemplarsatz 203@ \$01675704376</li> <li>•Exemplarsatz 203@ \$0xxxxxxxxxxxx 203S \$0962659215</li> </ul>

Titel stimmt nicht überein

Titel aus dem SWB, die mit keinem GBV-Titel übereinstimmen, werden als neue Titelsätze in K10plus abgespeichert.

SWB	GBV	K10plus
<ul style="list-style-type: none"> <li>•Titel 003@ \$0485937565</li> <li>•Exemplarsatz 203@ \$0962659215</li> </ul>		<ul style="list-style-type: none"> <li>•Titel 003@ \$0xxxxxxxxxxxx 003S \$0485937565</li> <li>•Exemplarsatz 203@ \$0xxxxxxxxxxxx 203S \$0962659215</li> </ul>

# Datenbereinigungen

Datenbereinigungen und Anpassungen in den „Altsystemen“

- Entfernen fehlerhafter Kategorien *laufend*
- Bereinigung von Codierungen, Abkürzungen usw. *laufend*
- Erzeugen der IMD-Felder im Rahmen der RDA-Anpassungen für alle Titeldaten in der SWB-Datenbank *geplant zum Jahresende*
- Bereinigung von Exemplardatenfeldern *in Absprache mit Bibliothek*
- Rückbau von Verknüpfungen zu Verlegerserien *in Diskussion*
- Rückbau Verknüpfungen zu Namenssätzen *in Diskussion*
- **Bitte an die Bibliotheken:** Abarbeiten der **Dublettenkandidaten** (derzeit ca. 177.000) aus Bibliotheksdatenimporten (0500 Pos. 3 = „B“)

# Formatabstimmung in 2017

## Künftiges Format:

- Titeldaten: 269 Felder
- Normdaten: 91 Felder inklusive GND
- Exemplardaten: 81 Felder
- Aktivitäten dazu:
  - Dreitägiger Workshop im Januar
  - bis heute ca. 40 Telefonkonferenzen zur weiteren Abstimmung
  - Dokumentation des Zielformates im Wiki

031@	-	4025	Zählung von fortlaufenden Ressourcen						JG	
031@	\$a	-	4025	ohne	Zählung von fortlaufenden Ressourcen					
035E		*	4026		Maßstab				JG	TN
035E	\$T	-	4026	\$T	Feldzuordnung					
035E	\$U	-	4026	\$U, %%	Schriftcode					
015@	-	1800	Code für Erscheinungsfrequenz						DW	TN
015@	\$a	*	1800	\$a	Code für Erscheinungsfrequenz					
										erster Code ohne Unterfeld zu erfassen Zugelassene Codes
004A		*	2000		ISBN			E-Book-Pool	DW	MK
004A	\$T	-	2000	\$T	Feldzuordnung					
004A	\$U	-	2000	\$U, %%	Schriftcode					
004A	\$0	-	2000		ohne	ISBN				
004A	\$f	-	2000	\$f	Kommentar zur ISBN, Einbandart, Lieferbedingungen und/oder Preis					
004P		*	2003		ISBN einer Manifestation in anderer physischer Form			E-Book-Pool	DW	MK
004P	\$T	-	2003	\$T	Feldzuordnung					
004P	\$T	-	2003	\$U, %%	Schriftcode					
004P	\$8	-	2003	..	Indikator					
										Zugelassene Codes/Indikatoren:
										<ul style="list-style-type: none"> <li>•  x  ISBN der parallelen Ausgabe auf einem anderen Datenträger</li> <li>•  y  ISBN der parallelen Ausgabe im Fernzugriff</li> <li>•  z  ISBN der parallelen Druckausgabe</li> <li>•  n  ISBN für parallele Ausgabe in einer anderen physischen Form (wenn nicht einem der o.g. Indikatoren zuzuordnen)</li> </ul>

# Erzeugung Zielformat

- Umsetzung bisheriges Format in Zielformat
- Dokumentation in Excel-Tabellen als Konkordanz „alt/neu“

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
	Pica3 Alt	Pica+ Alt	Beschreibung Alt	P3 UF Alt	P4 UF Alt	WF Alt	Beschreibung UF Alt	Pica 3 Neu	Pica+ Neu	Beschreibung Neu	P3 UF Neu	P4 UF Neu	WF Neu	Beschreibung UF Neu	Bemerkung (braunmarkierte -> Besonderheiten / Klärungsbedarf)	FCV mit Häkchen versehen, wenn bearbeitet	Stand		
1																			
2	0000	001@	ILNs der Bibliotheken, die Lokaldaten am Titelsatz haben					0000	001@	ILNs der Bibliotheken, mit Exemplarsatz			Nein				07.08.2017		
3	0000	001@			SO			0000	001@		ohne	SO	Nein	ILN-Liste			07.08.2017		
4	0000	001@						0000	001@		5a	--	Nein	Useflag (0,9, A-Z)	Bei Erzeugung 002@ 50 Pos. 24-25 hier in 5a "u" für Inhaltsverzeichnis anzeigen	✓	07.08.2017		
5																			
6	000A	001X	Titel-Besitzer					000A	001X	Titel owner			Nein				22.06.2017		
7	000A	001X						000A	001X		ohne	SO	Nein	Titel owner			22.06.2017		
8																			
9	000K	001U	Unicode-Kennzeichen			Nein		000K	001U	Unicode-Kennzeichen			Nein				22.06.2017		
10	000K	001U						000K	001U		ohne	SO	Nein	Unicode-Kennzeichen			22.06.2017		
11								000Q	001Q	VAL-Warnungen			Nein				22.06.2017		
12								000Q	001Q		ohne	5a	Nein	Erzeugungsdatum des Feldes			22.06.2017		
13								000Q	001Q		:	5b	Nein	ELN der Bibliothek, die die Warnung verursacht hat			22.06.2017		
14								000Q	001Q		-	5c	Nein	Warnungsmeldung - Nummer			22.06.2017		
15								000Q	001Q		--	5d	Nein	Warnungsmeldung - Text			22.06.2017		
16																			
17																			
18																			
19	0100	003@	Pica-Produktionsnummer					0110	003S	SWB-Pica-Produktionsnummer			Nein		(In K10plus wird neue 003@ vergeben!)		22.06.2017		
20	0100	003@		ohne	SO	Nein	Pica-Produktionsnummer	0110	003S		ohne	SO	Nein	SWB-Pica-Produktionsnummer		✓	22.06.2017		
21																			
22	0200	001A	Kennung und Datum der Ersterfassung					0200	001A	Kennung und Datum der Ersterfassung			Nein				22.06.2017		
23	0200	001A		ohne	SO	Nein	Kennung und Datum der Ersterfassung	0200	001A		ohne	SO	Nein	Kennung und Datum der Ersterfassung			22.06.2017		
24																			
25	0210	001B	Kennung und Datum der letzten Änderung					0210	001B	Kennung und Datum der letzten Änderung			Nein				22.06.2017		
26	0210	001B		ohne	SO	Nein	Kennung und Datum der letzten Änderung	0210	001B		ohne	SO	Nein	Kennung und Datum			22.06.2017		
27				--	St	Nein	Uhrzeit	0210	001B		--	St	Nein	Uhrzeit			22.06.2017		
28																			
29	0230	001D	Kennung und Datum der Statusänderung					0230	001D	Kennung und Datum der Statusänderung			Nein				22.06.2017		
30	0230	001D		ohne	SO	Nein	Kennung und Datum der Statusänderung	0230	001D		ohne	SO	Nein	Kennung und Datum			22.06.2017		
31																			
32	0240	001E	Quelle und Datum Löschen allgemeine Daten					0240	001E	Kennung und Datum der Löschung			Nein				22.06.2017		
33	0240	001E		ohne	SO	Nein	Quelle und Datum	0240	001E		ohne	SO	Nein	Kennung und Datum			22.06.2017		
34	0240	001E		--	St	Nein	Uhrzeit	0240	001E		--	St	Nein	Uhrzeit			22.06.2017		
35																			
36								0320	002V	MIR: Datum/Uhrzeit Korrektur der Master-TA							22.06.2017		
37								0320	002V		ohne	SO	Nein	MIR: Datum/Uhrzeit der Master-TA			22.06.2017		
38																			
39								0321	001L	MIR: UTC-Zeitstempel der letzten Änderung							22.06.2017		
40								0321	001L		ohne	SO		UTC-Zeitstempel der letzten Änderung			22.06.2017		

# Online-Hilfe

- Abstimmung und Erstellung der gemeinsamen Online-Hilfe

Anzeigen:  Standard  VD17

K10plus Format-Dokumentation (Online-Hilfe) BSZ GBV

[0100](#) [0200](#)

## 0110 SWB-Pica-Produktionsnummer

Pica3	Pica+	wiederholbar	letzte Änderung
0110	003S	Nein	2017-06-03 09:39:39

Das Feld enthält die Pica-Produktionsnummer (PPN) aus dem SWB. Alle Datensätze die aus der SWB-Datenbank in die gemeinsame K10plus-Datenbank migriert wurden, enthalten das Feld 0110. Künftig wird das Feld maschinell belegt, falls SWB-Datensätze im K10plus zusammengeführt werden und im Zielsatz noch kein Feld 0110 belegt ist. Enthält der Zielsatz bereits ein Feld 0110 so wird die umgekehrte SWB-PPN maschinell in das Feld [2111](#) gespeichert.

Unterfelder (Übersicht):

Pica3	Pica+	Inhalt	Inhalt Nicht-RDA	wiederholbar	letzte Änderung
0110	003S	SWB-Prod-Messung...		Nein	2017-06-03 09:39:39

Anzeigen:  Standard  VD17

K10plus Format-Dokumentation (Online-Hilfe) BSZ GBV

[0370](#) [0501](#)

## 0500 Bibliografische Gattung und Status

Pica3	Pica+	wiederholbar	letzte Änderung
0500	002@	Nein	2017-05-14 09:49:02

Das Feld 0500 enthält in codierter Form Angaben zur physischen Form, zur Erscheinungsweise der vorliegenden Ressource und zum Status der Aufnahme. Die Positionen 1-3 von Feld 0500 sind verpflichtend, die Position 4 wird im Allgemeinen nur maschinell besetzt.

Folgende Codes sind auf den jeweiligen Positionen möglich:

**Position 1: Physische Form**

Code	Bedeutung
A	Gedruckte Ressource
B	Tonträger, Videodatenträger, Bildliche Darstellungen
C	Blindenschriftträger
E	Mikroform
H	Handschriftliches Material
L	Lokales Katalogisat
O	Elektronische Ressource im Fernzugriff
S	Elektronische Ressource auf Datenträger
V	Objekt
Z	Medienkombination
a	Mailboxsatz

**Position 2: Bibliografische Erscheinungsweise**

Code	Bedeutung
a	Einzelne Einheit
b	Zeitschrift/Zeitung
c	Gesamtaufnahme einer mehrteiligen Monografie
d	Monografische Reihe
f	Teil einer mehrteiligen Monografie mit abhängigem Titel

# Aktivitäten 2017: Format

- Erstellung gemeinsame Beispielsammlung

The screenshot shows a web application interface for K10plus Wiki. The main content area displays the page title "1.1 Personen als geistige Schöpfer" and a list of RDA fields and their values:

- 0500 Auz
- 0501 Text\$bbt
- 0502 ohne Hilfsmittel zu benutzen\$tn
- 0503 Band\$bnc
- 1100 2006\$ln[2006]
- 1108 \$n@ 2006
- 1500 ger
- 1505 \$erda
- 1700 XA-DE
- 2000 3-549-07204-X
- 2000 978-3-549-07204-2
- 3000 IPPN\Fischer, Frank \*1968-; /ID. gnd/. \$BVerfasserID4aut
- 4000 Danzig\$die zerbrochene Stadt\$lnFrank Fischer
- 4030 Berlin\$Propyläen
- 4060 416 Seiten
- 4061 Illustrationen, Karten

Below the list, there is a section titled "Erläuterungen" with a list of bullet points:

- Da es sich bei dieser Ressource um eine gedruckte Monografie handelt, die nur geschriebenes Wort enthält, werden die IMO-Typen (Felder 0501 - 0503) jeweils nur einmal erfasst.
- Das Erscheinungsdatum wurde aus dem Copyrightdatum ermittelt und wird deshalb in Feld 1100 im Unterfeld \$n in eckigen Klammern erfasst. Das Copyrightdatum kann fakultativ in Feld 1108 erfasst werden.
- Im Feld 1500 wird in codelter Form wiedergegeben, dass es sich bei diesem Titeldatensatz um eine RDA-gerechte Aufnahme handelt.
- Die Vergabe von Feld 1700 ist fakultativ.
- Mehrere ISBNs der Vorlage werden in getrennten Feldern (Feld 2000) und ohne Steuerzeichen erfasst.
- Der erste geistige Schöpfer (RDA 19.2) wird im Feld 3000 bei Vorliegen von Lebensdaten in Form einer Verknüpfung zum Normsatz erfasst. Bei Belegung der Felder 3000 und 3010 wird im K10plus stets eine Beziehungskennzeichnung aus RDA Anhang I erfasst.
- Nach RDA 2.8.4.2 wird der Verlagsname so übernommen, wie er in der Quelle erscheint, aus der der Hauptteil stammt. Die abweichende Angabe "Propyläen Verlag" von der Rückseite der Titelseite wird nicht berücksichtigt.
- In der Umfangsangabe wird die letzte gezählte Seite erfasst. Dabei wird die Form der Zählung (arabische oder römische Ziffern) übernommen und die Bezeichnung "Seiten" ausgeschieden.
- Die Ressource enthält wenige Abbildungen. Im Feld 4061 (illustrierender Inhalt) wird zu deren Beschreibung der normierte Begriff "Illustrationen" verwendet (s. RDA 7.15). Zusätzlich können auch spezifischere Begriffe aus der Liste in RDA 7.15 verwendet werden, z. B. wie hier "Karten". Ist die Anzahl der Abbildungen oder Karten leicht zu ermitteln, z. B. durch eine Abbildungszählung, kann diese ebenfalls angegeben werden (Beispiel: 15 Illustrationen).

The interface also shows a sidebar with a navigation menu and a top bar with user information and search options.

# Einbeziehung Gremien Katalogisierung

## Workshops der Katalogisierungs-AGs

- März 2017 in Stuttgart:  
Projektstand, Gremien, Arbeitsplattform Wiki, Dokumentation (u.a. Abstimmung über kooperative Erstellung von Handbücher), Schulungskonzept, Konzept Korrekturrecht
- Oktober 2017 in Göttingen:  
geplante Themen:  
Projektstand, K10plus-Format, -Handbücher, -Online-Hilfe
- Weiterer Workshop im Frühjahr 2018 geplant

## Workshop zur Sacherschließung in Planung



# Probleme der Formatabstimmung

## Umgang mit Personensätzen in K10plus

- Unterschiedliches Arbeiten in der Normdatei GND  
*Empfehlung (in Abstimmung mit Koordinierungsausschuss):*  
Bei Vorhandensein individualisierender Angaben in der Vorlage möglichst in der GND arbeiten
- Anlage von Namenssätzen in der SWB-Datenbank  
*Empfehlung (in Abstimmung mit Koordinierungsausschuss):*
  - Rückbau der Namenssätze
  - Verzicht auf die Anlage eines Namenssatzes
  - Anzahl von Namenssätzen mit abweichenden Namensformenmöglichst niedrig halten

# Formatänderungen und Lokalsysteme

## BSZ - Lokalsysteme

- Abfangen der Formatänderungen im Export-Konverter
  - für MAB-Bezieher
  - für MARC-Bezieher
- Für Pica+-Bezieher (aDIS-Systeme) geplant
  - Pica+-Änderungen müssen nachgezogen werden
  - Austausch der Identnummern in den aDIS-Systemen (BSZ stellt Konkordanzen bereit)

## GBV - Lokalsysteme

- Tabelle „cbs2lbs“ bei der Übertragung CBS → LBS (OUM, OUF), die „neues Pica+“ in „altes Pica+“ zurückverwandelt
- Nach Katalogisierungsbeginn im K10plus werden alle LBS-Systeme neu geladen (dauert ein paar Jahre)

# Geplanter zeitlicher Ablauf

- Bis Ende September 2017
  - Formatdefinition
  - Aufbau Testsystem mit administrativen Daten und manuell erfassten Beispieldaten
- Oktober – Dezember 2017: Aufbau erste vollständige DB
  - Erster Gesamtabzug der beiden Verbunddatenbanken
  - Einspielen der GVK-Daten
  - Deduplizieren BSZ Daten
  - Laden Restdaten
- Januar 2018
  - Evaluierung des ersten Aufbaus von K10plus
  - Planung der weiteren Schritte, insbesondere der Schulungen

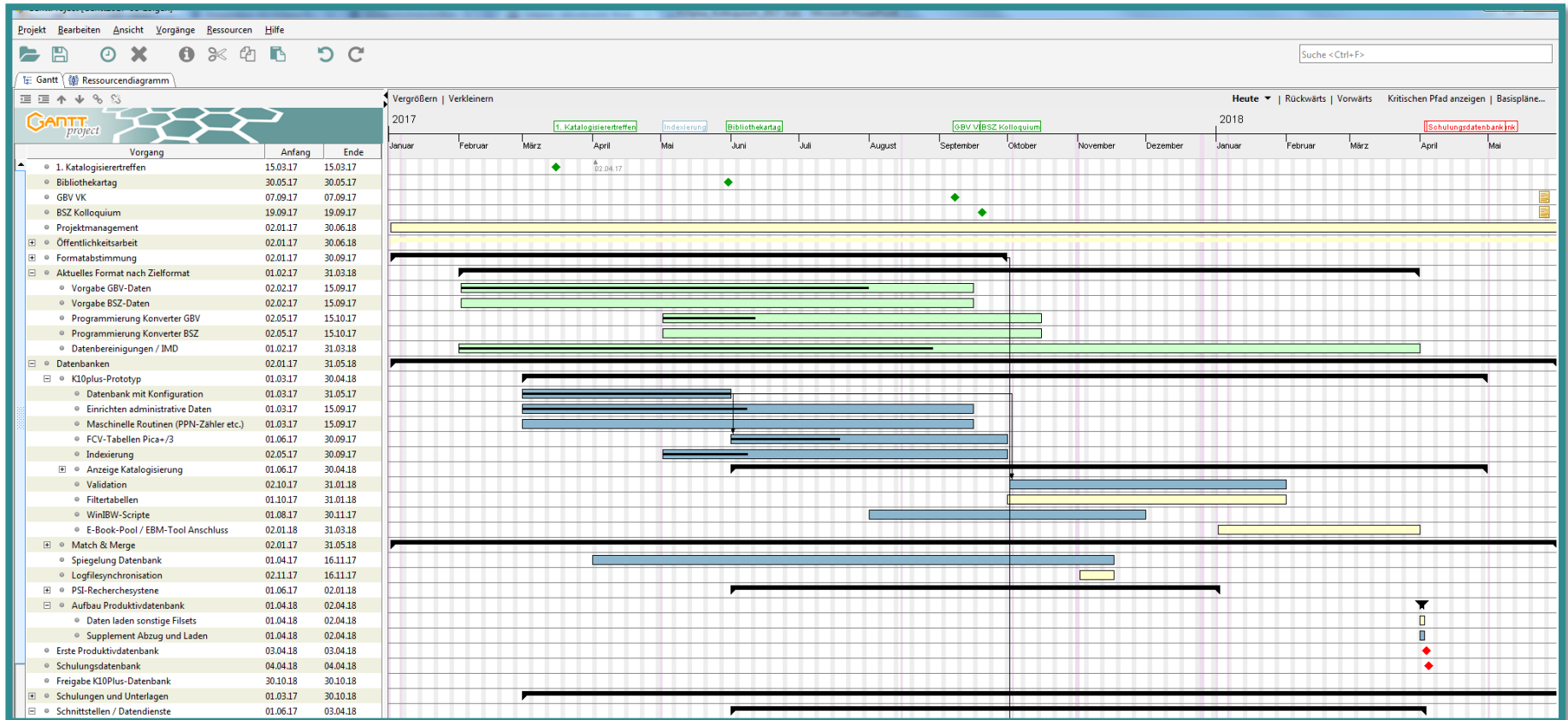
# Geplanter zeitlicher Ablauf

- Februar – April 2018: K10plus zweiter Versuch
  - Durchführung der erkannten Anpassungen
  - Basis für die anschließenden Schulungen
- Juni – September 2018
  - Multiplikatorenschulungen
  - Schulungen in den Bibliotheken
  - Updates aus den noch laufenden „Altsystemen“
  - Anpassung der Web-Umgebungen
  - Anpassungen sonstiger Abläufe: Import, Exporte, etc.
- Oktober 2018: Start K10plus

# Arbeitsorganisation im Projekt

- Vierteljährliche Projektmeetings
- Weitere Meetings und Telefonkonferenzen zu speziellen Themen (z.B. Konfigurationstabellen)
- Telefonkonferenzen im ein- bis zweiwöchigen Rhythmus zur Klärung bibliothekarischer Sachverhalte
- Dokumentation und Arbeiten im gemeinsamen Wiki
- Klärung von Sachverhalten, bei denen es zu keiner Einigung auf der Fachebene kommt, durch den Lenkungsausschuss bzw. den Koordinierungsausschuss
- Projektplan dient zu groben Zeitüberwachung

# Zeit- und Ressourcenplan (Stand August 2017)





**K10plus – eine Herausforderung!**

**Herzlichen Dank!**

**BSZ  GBV**

**Cornelia Katz (BSZ)**

**[cornelia.katz@bsz-bw.de](mailto:cornelia.katz@bsz-bw.de)**