

GND

Gemeinsame Normdatei

KofSE-Sitzung, 17.6.2010, Wien

Esther Scheven
DNB, Frankfurt am Main

Übersicht

- 1. Einführung**
- 2. GND-Format**
- 3. Nummernkonzept**
- 4. Datenanalysen**
- 5. Übergangsregeln**
- 6. Umsetzungsregeln**
- 7. Übergangsszenario mit Zeitplan**

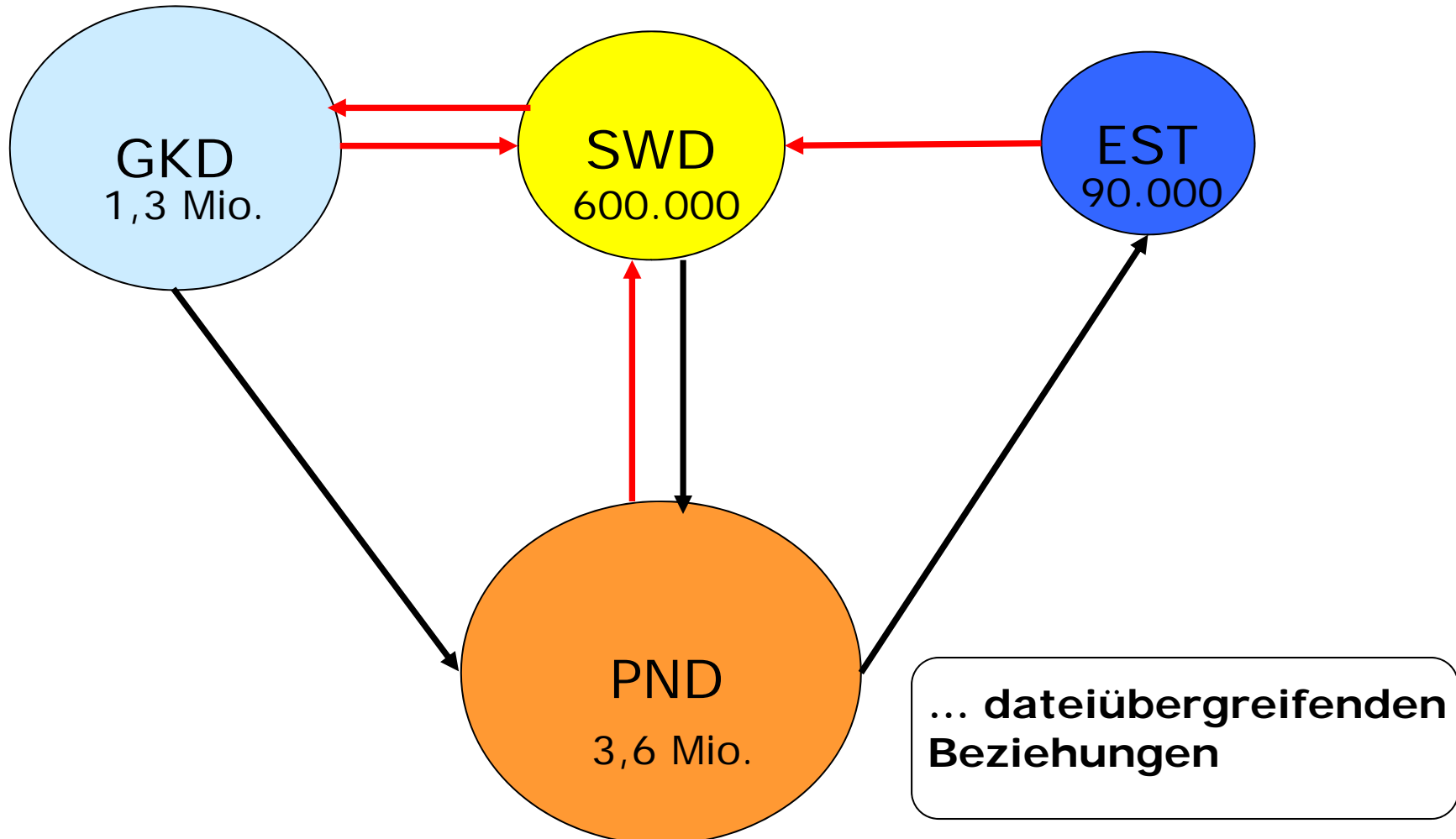
1. Einführung

- drei große Normdateien:
 - PND – Personennamendatei
 - GKD – Körperschaftsnamendatei
 - SWD – Schlagwortnormdatei
 - EST – Einsachtiteldatei der Musik

- unterschiedliche Regelwerke: RAK - RSWK
- unterschiedliche Datenmodelle
- unterschiedliche Datenformate
- redundante Datenhaltung

1. Einführung

heutige Normdateien mit ...



1. Einführung

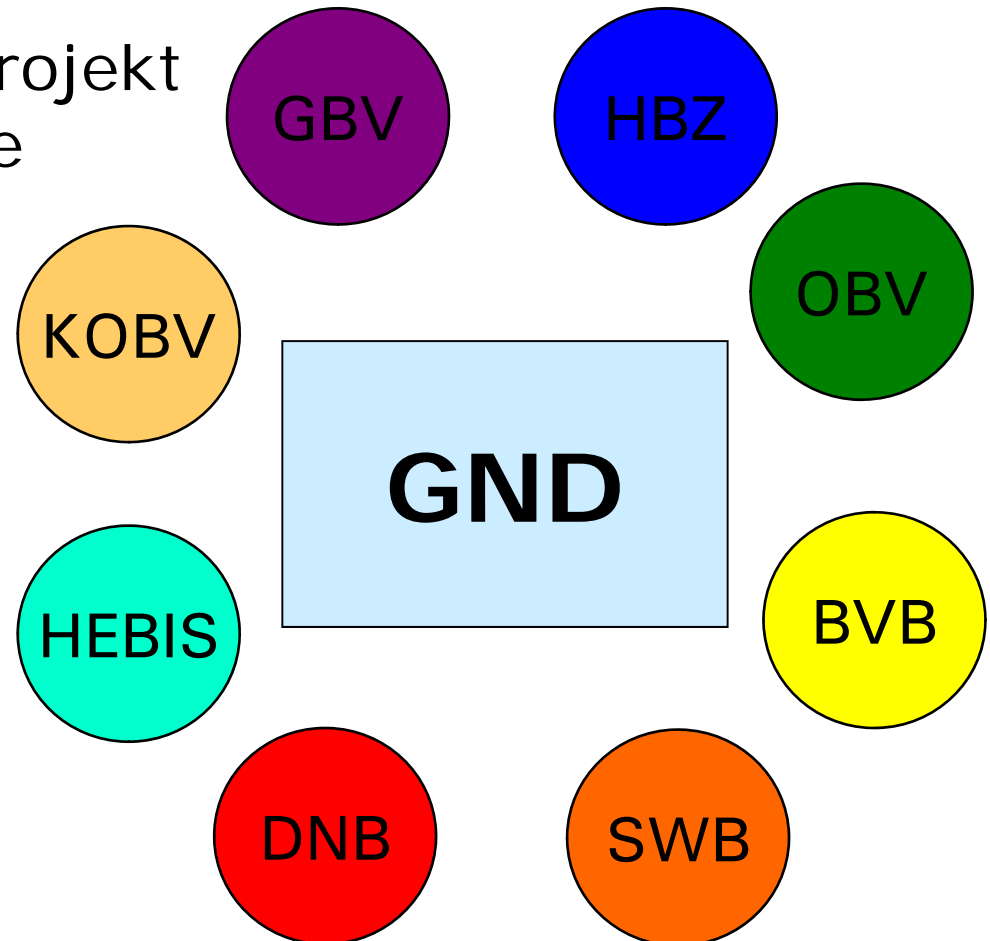
2003: Kooperationspartner fordern Zusammenführung

2005: Gemeinsames Projekt
DNB + Verbünde

2009: (im Juli)
DNB:
Start des
GND-Projekts
für eine

Gemeinsame

NormDatei



1. Einführung

Ziele einer Gemeinsamen NormDatei:

- Integration der Normdateien
 - Bestandswahrung
 - Aufhebung von Redundanzen
 - Einheitliches Datenformat und Regelwerk
 - Bereinigung ungeeigneter Datenstrukturen
- Angleichung an internationale Standards
 - Austauschformat MARC 21 - Authority
 - GND-Erfassungsformat nahe an MARC 21
 - Berücksichtigung von RDA, FRBR und FRAD
- Zukunftsfähig
 - Semantic-WEB fähig
 - Für automatische Verfahren geeignet

2. GND-Format

- Austauschformat MARC 21 (Authority)
- Katalogisierungsformat GND PICA 3
- Internformat GND PICA +
- Expansionsformat, Anzeigeformat

Formatentwurf im Wiki <https://wiki.d-nb.de/x/TYLZAO>

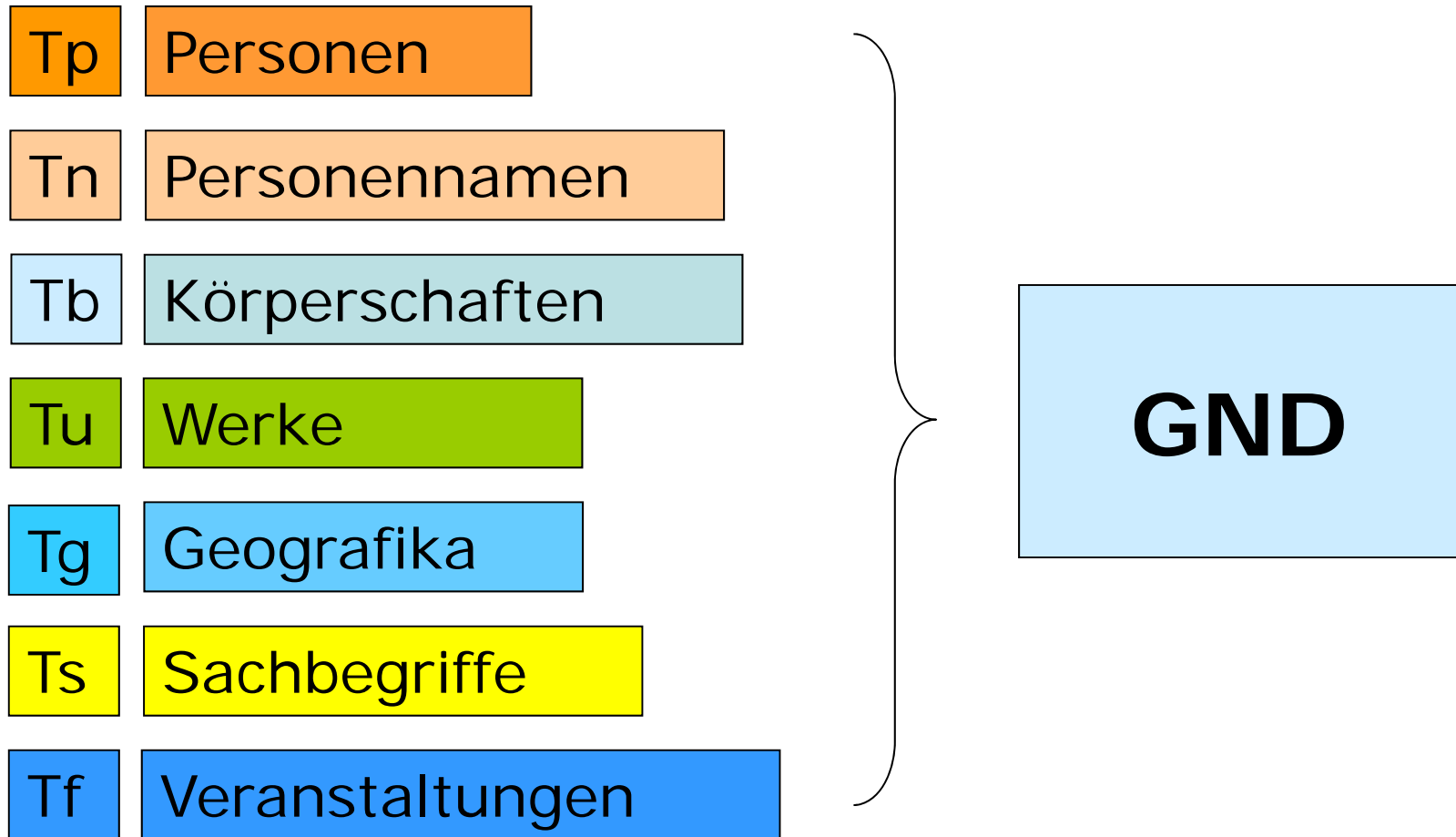
Modulare Datenstruktur



GND in sich ein semantisches Netz

2. GND-Format

Satzarten für verschiedene Entitäten



2. GND-Format

Entitäten werden analog zu MARC 21 durch Felder repräsentiert

Personen	X00
Personennamen	X00
Körperschaften	X10
Veranstaltungen	X11
Werke	X30 X00
Sachbegriffe	X50
Geografika	X51

Vorzugsbezeichnung
1XX

Äquival. Bezeichnung
4XX

Relationen
5XX

Alternativen
7XX

2. GND-Format

Eigenschaften der Entitäten: Attribute und Relationen

Codierte Angaben *Ländercode, Systematiknummern etc.*

GND PICA 3 / MARC 21: 0XX

RDA-Attributfelder *Content type, Carrier type, Media type ...*

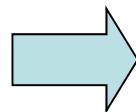
GND PICA 3 / MARC 21: 3XX

Relationsfelder *Oberbegriff, Geburtsort, Sitz, Architekt ...*

GND PICA 3 / MARC 21: 5XX

Spezifizierung der Relationen durch Codierungen

Relationstyp:
GND PICA 3: \$4 *Code*
MARC 21: \$4 *Code* \$wr \$i *Text*



Relationstyp:
GND PICA 3: \$4 *XXX*
MARC 21: \$4 *XXX* \$wr \$i *Sitz*

3. Nummernkonzept

Bisheriges Nummernkonzept

- Identifikationsnummer (IDN/PPN)
 - interner Datensatz-Identifizier (Pica 797/003@) in DNB-Datenbank
 - **MAB**: die IDN/PPN wird **nicht** ausgeliefert

- Normnummer (NID)
 - gilt als Normdaten-Identifizier in der Außenwelt

	PND	GKD	SWD
Pica3/+	797/003@	020/007P	021/007Q
MAB	001 + 020a	001 + 028b	001 + 028c

3. Nummernkonzept

Geplantes Nummernkonzept

- Normsätze, die in die GND **überführt** werden
 - behalten ihren internen Identifier (IDN/PPN)
 - bisherige Normnummer wird zur GND-Nummer
 - bisherige Normnummer wird zusätzlich aufbewahrt

	PND	GKD	SWD
GND-Nr. Pica3/+	035/007K mit Präfix g	035 \$agnd\$0100020836	
GND-Nr. MARC	035 \$a + ISIL (DE-5	035 ##\$a(DE-588)gnd/100020836	
Alte NID Pica3/+	039/007N Präfix pnd	039 \$apnd\$0100020836	
Alte NID MARC	035 \$z + ISIL (DE-588a)	035 ##\$z(DE-588a)100020836	

3. Nummernkonzept

Geplantes Nummernkonzept

- Normsätze, die **neu** in der GND erfasst werden
 - erhalten eine fortlaufende IDN/PPN
 - erhalten eine GND-Nummer;
 - ist eine fortlaufende Nummer, die aus der IDN/PPN des Datensatzes gebildet wird mit + Präfix „gnd“

	PND	GKD	SWD
GND-Nr. Pica3/+	035/007K mit Präfix gnd	035 \$agnd\$01XXXXXXXX	
GND-Nr. MARC	035 \$a + ISIL (DE-588)	035 ##\$a(DE-588)gnd/1XXXXXXXX	

- URI des Normdatensatzes
 - Pica3/+: 006/003U
 - MARC: 024 (z.B. [\\$2uri](http://d-nb.info/gnd/5540156-9/about))

4. Datenanalysen

Warum notwendig?

Zur Vorbereitung der Zusammenführung

- Kennzeichnung von Gebietskörperschaften in SWD + GKD
- Kennzeichnung von Organen in SWD + GKD
- Kennzeichnung von Kongressen in SWD + GKD
- Überprüfung der Personennamen (bei Werken oder in Verweisungen) in der SWD an der PND
- u.ä.

Zur Vorbereitung der Umsetzung der heutigen Datensätze in die neue GND-Datenstruktur

- Auflösung der Mehrgliedrigkeit soweit möglich
- u.ä.

4. Datenanalysen

Datenanalyse: Übersicht



Datenanalyse

Hinzugefügt von [Pfeifer, Barbara](#), zuletzt bearbeitet von [Kubbernuss, Antje](#) am 2009-10-23 ([Änderung anzeigen](#))

Datenanalyse

Einzelne Analysen:

[→ GKD](#)
 [→ SWD](#)
 [→ PND](#)
 [→ DMA-EST](#)

Zusammenstellung der [Ergebnisse und Übersicht über nötige Codierungen u. ä.](#) (dateiübergreifend)

[Stichwörter hinzufügen](#)

Untergeordnete Seiten (7) [Untergeordnete ausblenden](#) | [Hierarchisch anzeigen](#) | [Untergeordnete Seite hinzufügen](#)

-  [Abgleich 800 - 805](#)
-  [Abgleich 805-800](#)
-  [Datenanalyse DMA-EST](#)
-  [Datenanalyse - Ergebnisse, Codierungen](#)
-  [Datenanalyse GKD](#)
-  [Datenanalyse PND](#)
-  [Datenanalyse SWD](#)




Datenanalyse: GKD

Datenanalyse GKD

Hinzugefügt von [Pfeifer, Barbara](#), zuletzt bearbeitet von [Kubbernuss, Antje](#) am 2009-12-01 ([Änderung anzeigen](#))

Übersicht: Datenanalyse GKD

Legende:

-  = erledigt
-  = in Bearbeitung
-  = noch nicht bearbeitet

Analyse	Status
Kennzeichnung von Gebietskörperschaften in GKD und SWD	
Kennzeichnung von Organen von Gebietskörperschaften	
Kennzeichnung von Kongressen	
Kennzeichnung von ortsgebundenen Körperschaften	
Analyse: Verwaltungsgemeinschaften	
Analyse: Mehrgliedrige Ansetzungen	
Analyse: Reziproke Verknüpfungen	
Analyse: Hierarchische Beziehungen	

4. Datenanalysen

Datenanalyse: SWD



Datenanalyse SWD

1 Hinzugefügt von [Pfeifer, Barbara](#), zuletzt bearbeitet von [Kubbernuss, Antje](#) am 2010-01-19 ([Änderung anzeigen](#))

Übersicht: Datenanalyse SWD

Legende:

- = erledigt
- = in Bearbeitung
- = noch nicht bearbeitet

Analyse	Status
Kennzeichnung von Gebietskörperschaften in GKD und SWD (Link zur GKD-Seite)	
Kennzeichnung von Kongressen in der SWD	
8X0-Analysen	
80X: Mehrgliedrige Ansetzungen in der SWD - Statistik und Stichproben	
80X-Analyse	
Zusammenführung von Verwaltungseinheiten SWD/GKD (Link zur GKD-Seite)	

5. Übergangsregeln

1. Warum notwendig?

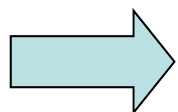
- Zusammenführung von 4 verschiedenen Normdateien mit unterschiedlichen Regelwerken
 - > Neues Ziel-Regelwerk ist noch nicht fertig
 - > Berücksichtigung der neuen internationalen Regelwerksentwicklungen (RDA/FRBR)
- Zielformat MARC 21 bedingt Regeländerungen
- Datenmodell der GND bedingt Regeländerungen

5. Übergangsregeln

2. Inhalte:

2.1. Gemeinsame Ansetzungsregeln für Formal- und Sacherschließung

*Berücksichtigung der Ergebnisse der beiden Projekten
GKR (Gemeinsame Körperschaftsregeln, 2005-2007)
und Perso (Angleichung der Ansetzung von
Personennamen nach RAK-WB und RSWK, 2003-2004)*



*Bestätigung der Ergebnisse dieser beiden Projekte
für die GND-Übergangsregeln*

5. Übergangsregeln

2.1. Gemeinsame Ansetzungsregeln für Formal- und Sacherschließung

Körperschaften:

- originalsprachlicher Name
- Splitregel
- Wegfall der Ortsbindung
- neue Entität „Kongresse“

Geografika:

- im Deutschen gebräuchlicher Name
- Was sind Namensbestandteile (Bad, Sankt u.ä.)?
- Was sind erläuternde Bestandteile?
- Umgang mit Gattungsbegriffen bei Verwaltungseinheiten

Werktitel: EST der Musik + SWD-Ansetzungen zusammenführen
(Vorgaben von RDA beachten)

5. Übergangsregeln

2.2. RSWK-Übergangsregeln

- Auflösen der Mehrgliedrigkeit soweit möglich
- Umwandlung der Hinweissätze für individuelle Objekte in Schlagwortsätze

➡ pro Entität ein Datensatz

- Zeitangaben als Relation im Feld 548 als Text
- Anderer Entitätenzuschnitt in MARC 21:
 Sprachen + Ethnografika zukünftig Sachschlagwörter
 Extraterrestrika zukünftig geografische Schlagwörter

5. Übergangsregeln

3. Darstellung des Verfahrens

- Abstimmung über das Prozedere mit dem Standardisierungsausschuss (erfolgte im September 2009)
- Inhaltliche Abstimmung über die Regeln durch Expertengruppen-Sitzungen
- Genehmigung der EG-Beschlüsse durch den StA
- Erarbeiten eines Veröffentlichungskonzepts
- Schulungen der Regeln
- Anwendung der Regeln mit GND-Produktionsaufnahme (voraussichtlich 2011)
- GND-Übergangsregeln: Grundlage für das neue Zielregelwerk

6. Umsetzungsregeln

1. Für die Zusammenführung von gleichen Entitäten aus zwei verschiedenen Normdateien:

- Festlegen des Gewinnerdatensatzes

Beispiele:

Gebietskörperschaften: Gewinnerdatensatz ist aus der SWD



Körperschaften: Gewinnerdatensatz ist aus der GKD



6. Umsetzungsregeln

2. Für datenmodell-bedingte strukturelle Änderungen im Datensatz

- Auflösen der Mehrgliedrigkeit

Beispiele:

Auflösen der Ortsbindung bei Körperschaften, SWD
(Entitätencodierung „kis“)

SWD:
800|c|Berlin
801|x|Berliner Philharmonie

GND (MARC 21):
110 \$a Berliner Philharmonie
551 \$a Berlin \$4 Code \$wr \$i *Sitz der Körperschaft*

6. Umsetzungsregeln

Bauwerke: Keine Ansetzung mehr unter dem Ort, sondern unter dem Namen an sich, SWD (Entitätencodierung „kib“)

➔ Anzahl der Elemente bei der mehrgliedrigen Ansetzung berücksichtigen

➔ Bei Normierungen wie „Burg“ einen individuellen Namen erzeugen

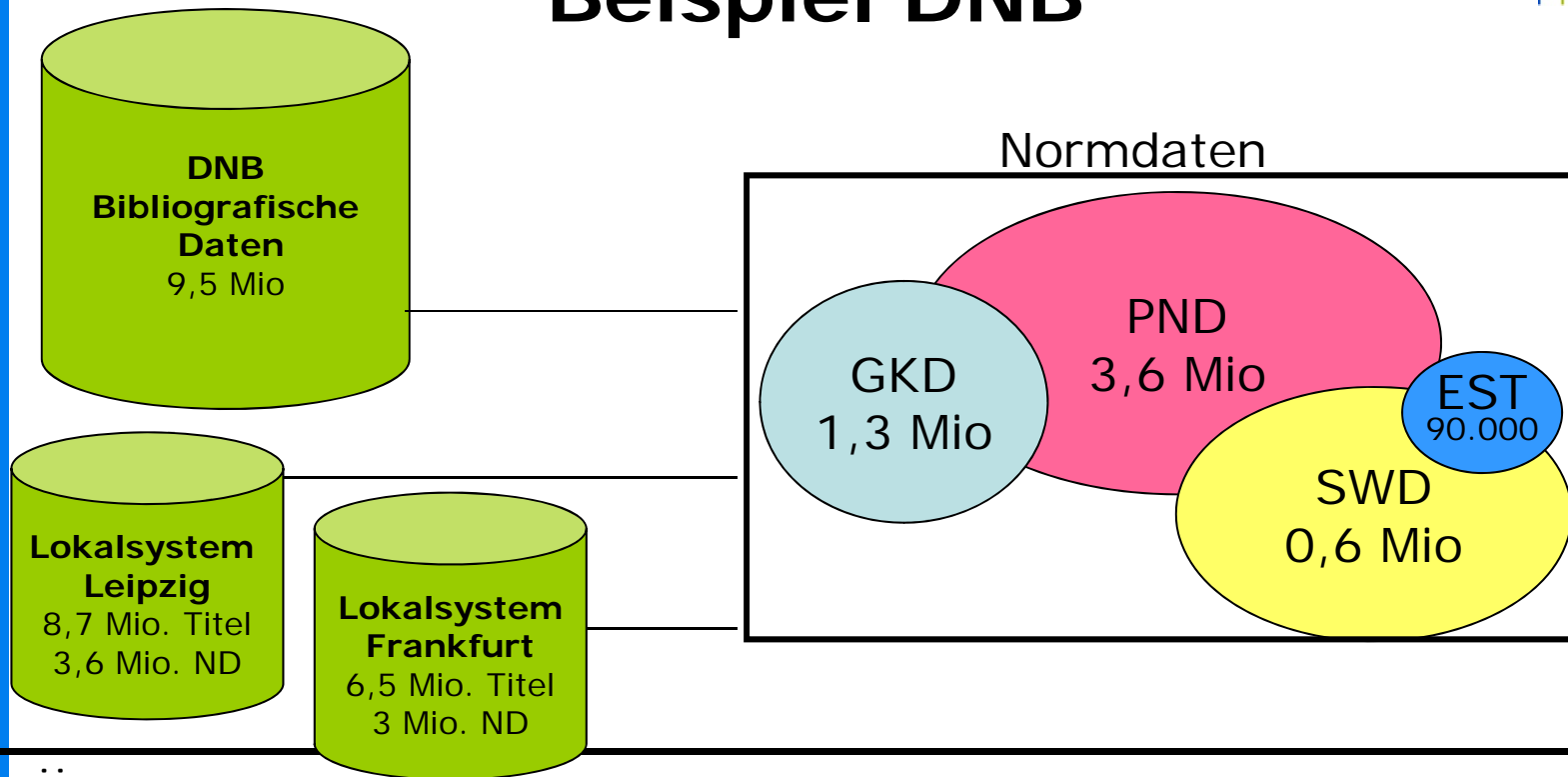
SWD:
800|g|Eppstein
801|x|Burg

GND (MARC 21):
151 \$aBurg Eppstein
551 \$a Eppstein \$4 *Code* \$wr \$i *Sitz des Bauwerks*

7. Umgangsszenario und Zeitplan

Folien von Frau Feilhauer, DNB

Beispiel DNB



Schnittstellen:

Portal

Batch-Datendienste

OAI:

- VIAF
- BVB
- HEBIS
- BSZ
- GBV

ONS:

- BVB
- HBZ
- OBV

Externe WinIBW:

- GBV
- BSZ
- HEBIS

Datenvolumen in DNB

- Zentrale Normdaten: ca. 5,5 Mio.
 - davon ca. 6 % der Normdaten von Zusammenführung betroffen,
 - ca. 94% der Normdaten werden 'nur' in ein neues Datenformat überführt
- Verknüpfungen zu ca. 10,3 Mio. bibliografischen Daten
- (nur!) zwei Lokalsysteme: Frankfurt und Leipzig

Verbünde haben ein wesentliches höheres Datenvolumen, wegen:

- z.T. Zusätzlich lokale Normdateien, die umgestellt werden müssen
- erheblich mehr bibliografische Datensätze
- erheblich mehr Lokalsysteme

Ablauf

Schritt 1:

Projekt-Datenbank
als Schattendatenbank bei DNB

Schritt 2:

Zusammenführung der Entitäten



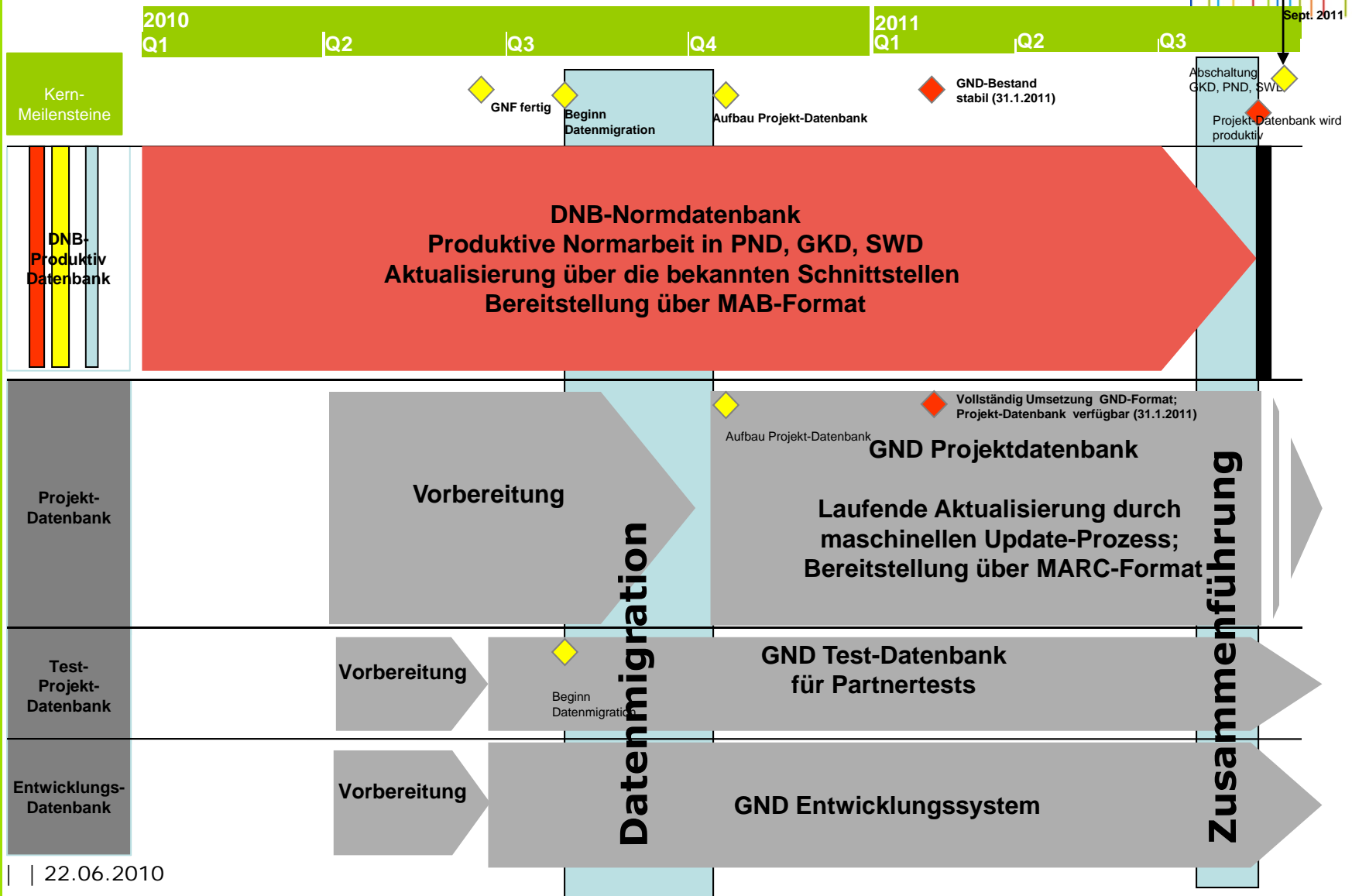
Schritt 3:

Produktiv-Schaltung der GND

Schritt 4:

Abschaltung der alten Normdateien

Projekt-Datenbank(en) (geplant)



7. Umgangsszenario und Zeitplan

Zeitplan

04 2010	GND-Testsystem aufgebaut
06 2010	GND-Format abgestimmt
3. Quartal 2010	Datenanalysen und Vorbereitungen in PND, GKD, SWD abgeschlossen Start: Aufbau der GND Migration und Zusammenführung der redundanten Entitäten (Match & Merge)
4. Quartal 2010	Testauslieferung an Pilotpartner
1. Quartal 2011	Start Partnertest für ONS GND
3. Quartal 2011	voraussichtlich GND in Echtbetrieb

*Vielen Dank für Ihre
Aufmerksamkeit!*

Esther Scheven e.scheven@d-nb.de

Weitere Informationen:

Im Wiki: <https://wiki.d-nb.de/x/TYLZAO>