

# Updates und Roadmap

## Bericht zum Stand der Technik

Ines Häusler und Albert Schröder  
Universitätsbibliothek Regensburg  
**RVK-Anwendertreffen 2019**



Universität Regensburg

## Überblick

- 1) Zeitschlagwörter Geschichte
- 2) Löschung der Verweisungen
- 3) Registersuche in RVK-Online
- 4) Suche in RVK-Online: UTF-8
- 5) Unicode: Normalisierung
- 6) Fehlende Registerbegriffe
- 7) Dazu kommt...

# 1) Zeitschlagwörter Geschichte

Ausgangslage:

- Registerbegriffe sind mit den preferred names in der GND verknüpft.
- Antragsteller können bei einer Neumeldung nur noch GND-IDs eingeben.
- Alle Registerbegriffe, die nicht mit der GND verknüpft werden konnten, wurden ausgeblendet

Problem:

Informationsgehalt des Zeitschlagworts Geschichte und Komposita mit Zusatz (Jahresangaben) ging verloren, da keine GND-ID vorhanden

## Zeitschlagwörter Geschichte

Ziele:

- Zeitschlagwörter z Geschichte und bestimmte Komposita sollen unverknüpft und gegebenenfalls ergänzt um eine Jahresangabe ins Register eingetragen werden können
- Die z Geschichte mit Jahresangaben aus dem „alten“ Register sollen wieder im RVK-Register sichtbar sein
- Sortierung der z Geschichte an das Ende der Schlagwörter
- Aber: Alle anderen Registerbegriffe sollen GND-gerecht sein

Vorgehen:

Die Redaktionsdatenbank und Neumeldungsformular wurden angepasst (Grundlage RSWK, 4. Aufl. 2017), RVK-Online wurde aktualisiert

## Zeitschlagwörter Geschichte

Nach RSWK (4. Aufl. 2017) §402 gibt es folgende  
Zeitschlagwörter z:

Geschichte, Geistesgeschichte, Sozialgeschichte, Ideengeschichte,  
Vor- und Frühgeschichte, Kirchengeschichte, Weltgeschichte

Nach RSWK sind die Zeitangaben folgendermaßen einzugeben:

Geschichte 1989

Geschichte 1929-1989

Geschichte 700 v. Chr.-338 v. Chr.

Geschichte 122 v. Chr.-163

Geschichte Anfänge-500

## 2) Löschung der Verweisungen

Ausgangspunkt:

Verweisungen mit Kleinbuchstaben, die aus Zeiten der Druckversion stammten, wurden in RVK-Online gelegentlich mit Notationen verwechselt.

Zudem: In Aleph unklare Sortierung, z.B.:



— a|iiiiiii  
b txt

Indexeinträge in Bibliothek: BVB13 - Code: NOI

# Dok.	Eintrag	Norm. Info.
1	IV 48281   Sekundärliteratur	BVB13, NOI, , UPD=Y
1	IV 48320 --- IV 48321 (NEIN!)   Tozzi, Federigo	BVB13, NOI, , UPD=Y
1	IV 48320   Gesammelte Werke und Einzelwerke	BVB13, NOI, , UPD=Y
1	IV 48320a   Tranquilli, Secondo s. IV 45920 Silone, Ignazio	BVB13, NOI, , UPD=Y
1	IV 48321   Sekundärliteratur	BVB13, NOI, , UPD=Y
1	IV 48360 --- IV 48361 (NEIN!)   Travaini, Eugenio	BVB13, NOI, . UPD=Y

OK  
Gehe zu  
Details  
Anzeigen

## Löschung der Verweisungen

Entscheidung:

Die Verweisungen mit Kleinbuchstaben werden endgültig aus der RVK entfernt

Vorgehen:

- Zuerst Ausblenden in RVK Online und Abwarten eventueller Reaktionen seitens der Anwender
- Dann Löschung aus der Redaktionsdatenbank, zuerst die einzelnen Verweisungen, dann die die Bereichsangaben mit Verweisungen

### 3) Registersuche in RVK-Online

Ausgangspunkt:

Regel fürs Register: nur Einzelschlagwörter, keine Ketten

In bestimmten Fällen wurden doch SW-Ketten gebildet:

z.B.: USA / Umweltschutz (Geografika + SachSW)

Entscheidung:

Bei der Registersuche soll durch die UND-Suche eine Schnittmenge erzeugt werden.

Vorgehen:

Alle SW-Ketten werden in Schlagwörter zerlegt

## 4) Suche in RVK-Online: UTF-8

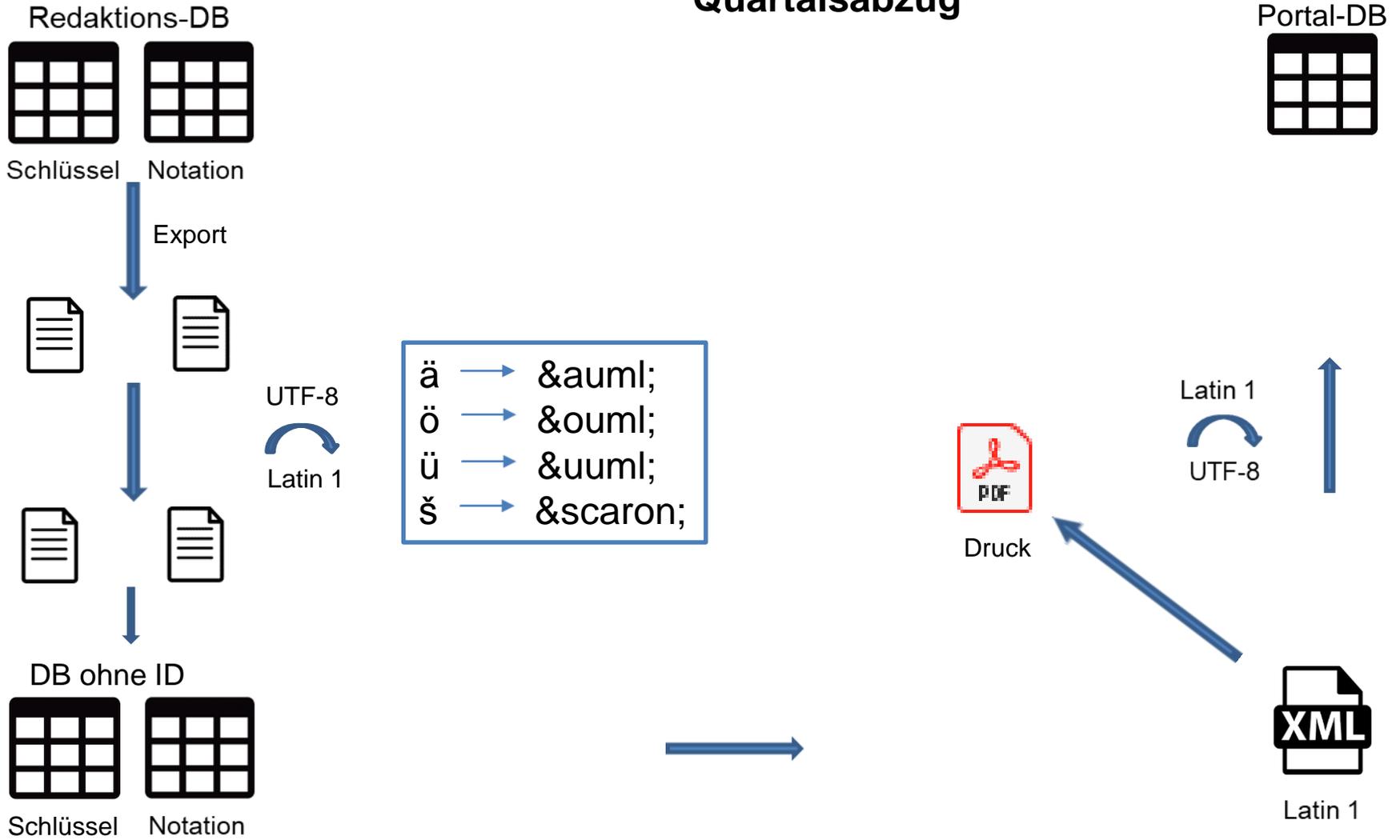
Ausgangslage:

Die Redaktionsdatenbank und die ausgeschriebene Datenbank (Portal-DB) sind UTF-8-fähig.

Problem:

- Dazwischen erfolgt eine Umsetzung auf HTML-Entities → Codierung in IsoLatin, nötig für PDF und XML
- Im letzten Schritt erfolgt dann die erneute Umkodierung für RVK-Online
- Darstellungsproblemen im Portal. Damit das PDF erzeugt werden kann, sind beide Wege nach wie vor nötig

# Quartalsabzug



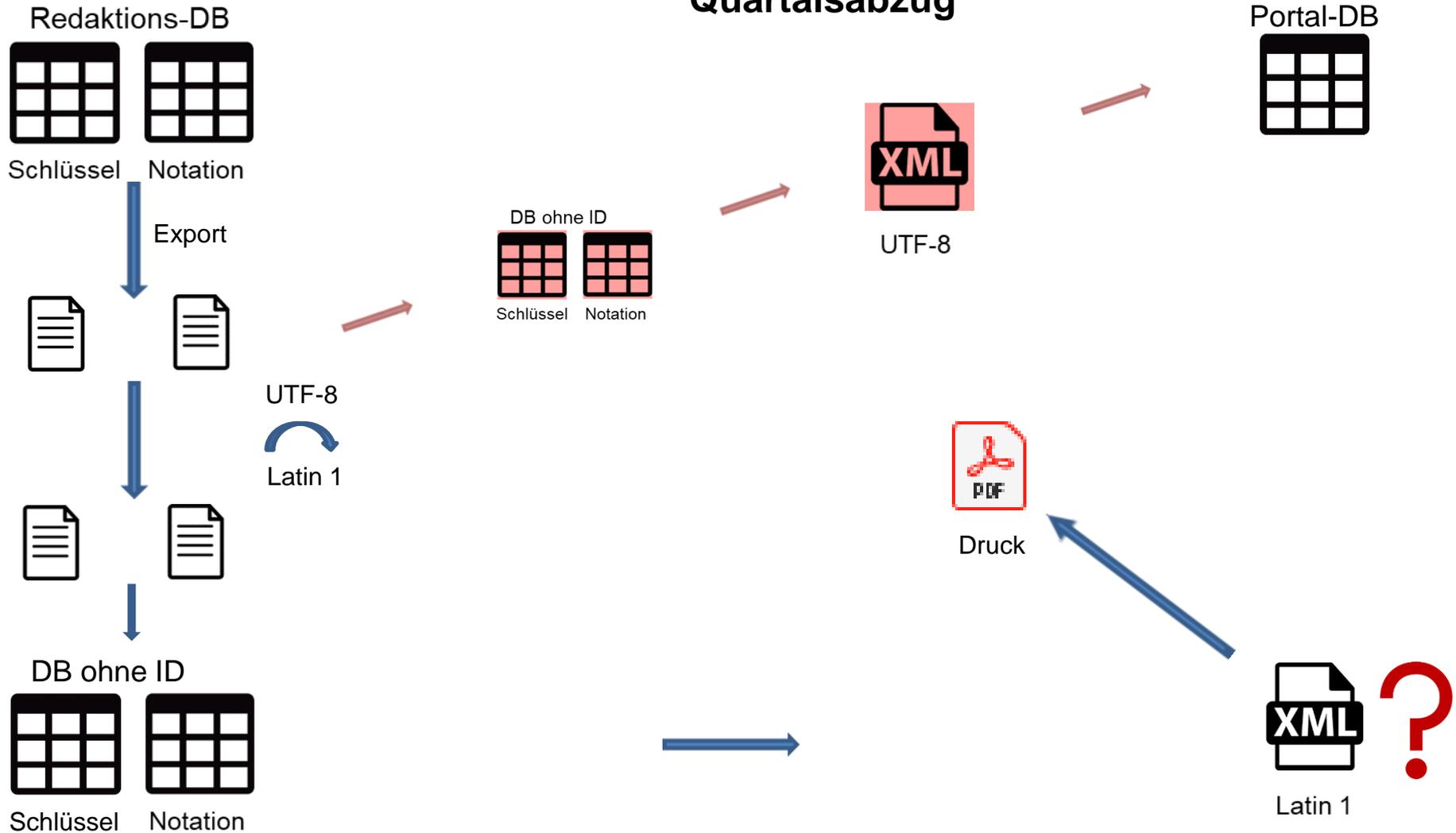


Universität Regensburg

Ines Häusler und Albert Schröder  
Universitätsbibliothek Regensburg

# UTF8

Lösung:

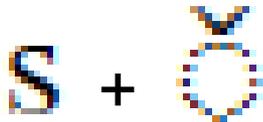


## 5) Unicode: Normalisierung

Buchstaben mit Diakritika können verschieden dargestellt werden



LATIN SMALL LETTER S WITH CARON



LATIN SMALL LETTER S + COMBINING CARON

Für Suche und Sortierung Normalisierung notwendig:

- NFC: kanonische Komposition
- NFD: kanonische Dekomposition

## Unicode: Normalisierung

Problem:

Datenbank war teilweise in UTF8-NFD (zerlegt), Register war in UTF8-NFC (kombiniert)

Entscheidung:

Die Zeichen, die noch nicht in NFC vorliegen, wurden ausgespeichert, umgeändert und wieder eingespeichert

## Beispiele aus dem RVK-Portal

**EN 3357 - EN 3358** Suyūṭī, Ġalāl-ad-Dīn ‘Abd-ar-Raḥmān Ibn-Abī-Bakr as-

Bemerkung: Islamischer Theologe und Philosoph; 1445-1505  
Register:

- Suyūṭī, Ġalāl-ad-Dīn ‘Abd-ar-Raḥmān Ibn-Abī-Bakr as-

**LH 14161** Mittel- und Südamerika

Register:

- Bibliografie
- Kunst
- Mittelamerika
- Südamerika



## 6) Fehlende Registerbegriffe

Problem: Fehlende Registerbegriffe, die wichtig sind, aber ausgeblendet sind

z.B. fehlen bei NN 1597 die Registerbegriffe Doria (Familie) und Bentivoglio (Familie)

N	Geschichte
NN	Neuere Geschichte
NN 1500 - NN 1740	Zeitalter der Entdeckungen und der Glaubensspaltung (bis 1648)
NN 1560 - NN 1695	Renaissance und Humanismus
NN 1590 - NN 1645	Italien
M239	
NN 1597 - NN 1645	Persönlichkeiten

**NN 1597** Sammelbiografien und Geschichte bedeutender Familien

## Fehlende Registerbegriffe

Das Fehlen hat meistens die 4 Gründe:

- Die Namen gab es 2012 noch nicht exakt in der Schreibweise in der GND (weder bei den bevorzugten Namen noch bei den Variationen).
- Namen gab es mehr als 1x in der GND und konnten deshalb maschinell nicht zugeordnet werden.
- runde Klammern oder <> unterschieden die Begriffe bei RVK und GND, z.B. Doria (Familie) bzw. Doria <Familie>
- Die GND hatte 2012 knapp 10 Millionen Begriffe, jetzt mehr als 15 Millionen

## Fehlende Registerbegriffe

Entscheidung:

Es werden keine Listen erstellt, die durchgearbeitet und eingepflegt werden müssen.

Entscheidung:

Sinnvoller ist es, eine Möglichkeit zu schaffen, wie interne und externe Kolleginnen und Kollegen in die Redaktionsdatenbank eigenständig Registerbegriffe einbringen können.

Dazu wird ein Zugriff auf die Redaktionsdatenbank nötig sein  
→ Zukunft?

## 7) Dazu kommt...

- Aufwändiges Einpflegen umfangreicher Vorschläge, z.B. Indologie
- Änderungen im Lieferformat der GND müssen nachgezogen werden
- Konvertierung nach RDF: URI wurde festgelegt
- Korrektur einer fehlerhaften Funktion des „CutterJo“
- Vorbereitung und Durchführung der Updates (viermal/Jahr)
- Konkordanz/Normdatenmodell: Erstellung eines Testmodells

## CutterJo: Vorher

-

## Nachher

CutterJo	
<input type="text" value="smith, marc"/> <input type="button" value="Suchen"/> 	
Name	Nummer
Smar	S636
Sme	S637
Smel	S638
Smi	S639
Smil	S641
Smit	S642
SmithB	S643
SmithC	S644
SmithD	S645
SmithE	S646
SmithF	S647

CutterJo	
<input type="text" value="smith, marc"/> <input type="button" value="Suchen"/> 	
Name	Nummer
SmithH	S649
SmithJ	S651
SmithJohn	S652
SmithJos	S653
SmithL	S654
SmithM	S655
SmithO	S656
SmithR	S657
SmithRobert	S658
SmithS	S659
SmithSol	S661