

Synergien nutzen - Maschinelle Unterstützung der intellektuellen Inhaltserschließung

Einsatz des DA-3 in der Deutschen Nationalbibliothek

Agenda

1. Maschinelle Unterstützung der intellektuellen
Inhaltserschließung (Helga Karg)
2. Technische Umsetzung (Hans-Jürgen Becker)
3. Exkurs: Vorschläge generieren mit der Plattform
Culturegraph (Angela Vorndran)

1. Maschinelle Unterstützung der intellektuellen Inhaltserschließung

Unsere Ziele

- Erschließung aller Publikationsformen mit hoher Qualität
- Erschließung auf verschiedenen Ebenen
Werk – Artikel – Kapitelebene
- Erschließungsergebnisse austauschen – Fremddaten nutzen –
Doppelarbeit vermeiden
- Erschließungsinstrumente pflegen – aktuell halten – weiterentwickeln
- Erschließungsinstrumente vernetzen und erweitern
 - ➔ Technische Unterstützung

Tools zur Erschließung

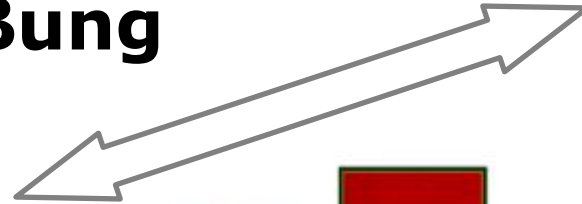
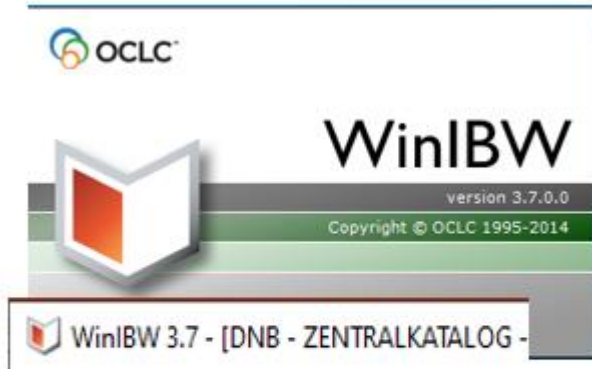
Was müssen sie leisten?

- Maschinelle Erschließung → Projekt *Erschließungsmaschine* (EMa)
- Unterstützung der intellektuellen Erschließung
 - Arbeiten unter einer Oberfläche
 - Präsentation der Erschließungsinstrumente
 - Vorschläge aus verschiedenen Quellen – Schlagwörter – Notationen
 - Beschleunigung des Erschließungsprozesses – Intuitive Menüführung

..... und gerne noch mehr



Recherchequellen bei der Inhaltserschließung





Der Digitale Assistent – DA-3

- Praxistest 2020 -2022
- Freischaltung für alle Inhaltserschließter*innen
- Nutzung für alle Fächer
- Parallele Nutzung des DA-3 und der WinIBW
- Vergabe der DNB-Sachgruppen, GND-Schlagwortfolgen und Gattungsbegriffe* im DA-3
- DDC-Vergabe und Freischaltung des Titels zur Anzeige in der NB in der WinIBW
- Recherche nach Vergleichstiteln, „Katalogsuche“, GND-Pflege in der WinIBW

*63 festgelegte Gattungsbegriffe zur Beschreibung literarischer Gattungen von Medienwerken;
in der DNB zur Erschließung von belletristischen Werken und Werken der Kinder- und Jugendliteratur eingesetzt

Alles unter einer Oberfläche

The screenshot shows the DNB interface with several callouts and annotations:

- Tools - Erschließungsinstrumente**: A box pointing to the left sidebar menu.
- Info**: A box pointing to the 'Info' tab in the top right corner.
- Kurztitel**: A box pointing to the 'Kurztitel' section on the right.
- Info**: A box pointing to the 'Info' section in the middle pane.
- Scratchpad**: A box pointing to the 'Scratchpad' section at the bottom left.
- Vorschläge**: A box pointing to the 'Vorschläge' section in the bottom middle pane.
- Erschließung**: A box pointing to the 'Erschließung' section at the bottom right.
- CBS Zentrales Katalogisierungssystem**: A large box on the right side, with arrows pointing to the 'GND' and 'DNB Sachgruppen' items in the 'Erschließung' list.

Annotations in the interface include:

- A red box around 'Gemeinsame Normdatei (GND)' in the Tools sidebar.
- A green box around 'DNB Sachgruppen' in the Tools sidebar.
- A red circle around 'GND' in the 'Erschließung' list.
- A green circle around 'DNB Sachgruppen' in the 'Erschließung' list.
- A red circle around the 'Exportieren' button in the top right.

Anpassungen für die DNB

- Fachliche Anpassungen:
 - Detailliertere Anzeige der Titeldaten und der GND-Normdaten
 - Anzeige der maschinellen Erschließungsdaten und der MVB*-Fremddaten als Vorschläge
 - Unterscheidung von Haupt- und Nebensachgruppen
 - Einbindung der Gattungsbegriffe
 - Nahezu synchrone Anbindung der GND
 - Fünf-sekündlicher Export der Erschließungsdaten in CBS
- Technische Anpassungen → zweiter Teil des Vortrags

GND-Normdatensatz

Wichtige Datensatzinformationen:

- Schlagwortart
- Katalogisierungslevel
- Entitätencode
- Systematiknummer
- DDC-Notation
- Teilbestandskennzeichen
- Quellenangabe
- Verweisungen

Info > GND

Erneuerbare Energien

Sachschlagwort

GND-Nr.: [4068598-6](#) [gnd1]

Entitäten: Allgemeinbegriff [saz]

Systematik: Energietechnik, Kerntechnik [31.5]

DDC: [333.794](#), [621.042](#)

Bestand: s [Sacherschließung]

Quellen: B 2006 (online)

Varianten: Sanfte Energie
Erneuerbare Energiequelle
Regenerative Energiequelle
Regenerative Energie
Erneuerbare Energie
Alternativenergie
Alternativenergien
Regenerative Energien
Alternative Energiequelle
Grüne Energie
Green energy

Oberbegriffe: Energiequelle
Erneuerbare Ressourcen

→ weitere Beziehungen
→ Hierarchie

WebGND

DEUTSCHE NATIONAL BIBLIOTHEK

GND – Recherche – Anzeige – Auswahl

Tools > GND

Ärztliches Honorar Präfix ✕

Ärztliches Honorar H

#

Alle

Sach

Geo

Person

Körp.

Kong.

Werk

Zeit

Form

Datentr.

Musik

Zielgr.

Bestand f

Info > GND Verlauf #

Zurück

Ärztliches Honorar [Hinweissatz]

Sachschlagwort

GND-Nr.: 7502811-6 [gnd1]

Systematik: Gebührenrecht [7.11c]
Handel einzelner Branchen, Dienstleistung einzelner Branchen [10.13b]

Bestand: s [Sacherschließung]

Quellen: analog

Für: Arzt Honorar

→ weitere Beziehungen

→ Hierarchie

Erschließung Prüfen | Rohdaten | Einstellungen #

DNB Sachgruppen

340 Recht

GND

Neue Folge anlegen

1: Arzt [Sach]

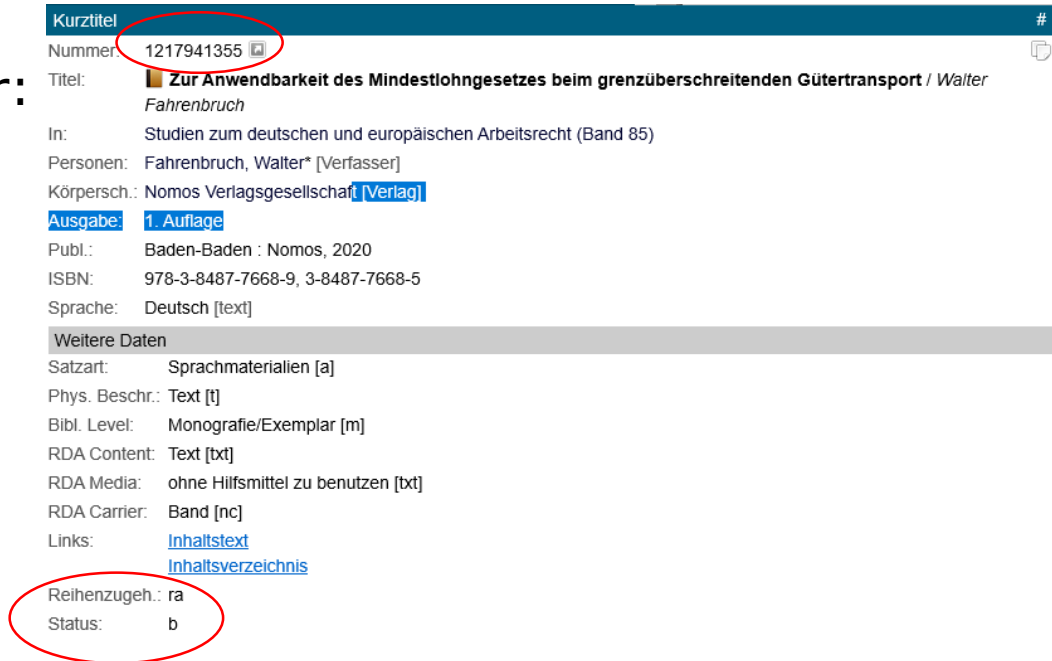
2: Honorar [Sach]

Suche getrennt nach Schlagwortart für eine eigene Erschließung

Titeldatensatz

Wichtige Infos für die Bearbeiter:

- IDN
- Link in die WinIBW
- Bibliografische Reihe
- Bearbeitungsstatus



Kurztitel #

Nummer: 1217941355

Titel: **Zur Anwendbarkeit des Mindestlohngesetzes beim grenzüberschreitenden Gütertransport** / Walter
Fahrenbruch

In: Studien zum deutschen und europäischen Arbeitsrecht (Band 85)

Personen: Fahrenbruch, Walter* [Verfasser]

Körpersch.: Nomos Verlagsgesellschaft [Verlag]

Ausgabe: 1. Auflage

Publ.: Baden-Baden : Nomos, 2020

ISBN: 978-3-8487-7668-9, 3-8487-7668-5

Sprache: Deutsch [text]

Weitere Daten

Satzart: Sprachmaterialien [a]

Phys. Beschr.: Text [t]

Bibl. Level: Monografie/Exemplar [m]

RDA Content: Text [txt]

RDA Media: ohne Hilfsmittel zu benutzen [txt]

RDA Carrier: Band [nc]

Links: [Inhaltstext](#)
[Inhaltsverzeichnis](#)

Reihenzugeh.: ra

Status: b

Vorschläge

- Erschließungsdaten der Verbünde
 - Fremddaten der DNB: MVB, Broadcastsearch, u. a.
 - Maschinelle Erschließungsdaten
- } **dnb/auto**
- Ergebnisse aus Konkordanzen → @
 - Herkunft deutlich gekennzeichnet
 - Individuelle Auswahl → **Einstellungen**

Vorschläge		Status	Rohdaten	Einstellungen
Filtern	Aktualisieren	Erweitern		
DNB Sachgruppen				
004 Informatik	@ddc, dnb/auto, obv			
300 Sozialwissenschaften, Soziologie, Anthropologie	dnb/auto, obv			
600 Technik	dnb/auto, obv			
GND				
Agent <Künstliche Intelligenz> [Sach]	@rvk			
Anwendung [Sach]	bvb, dnb, hebis			
Benutzeroberfläche [Sach]	@lcsb			
Chatbot [Sach]	bvb, dnb, hebis			
Fuzzy-Logik [Sach]	@rvk			
Kognitionswissenschaft [Sach]	bvb, dnb, hebis			
Kreativität [Sach]	bvb, dnb, hebis			
	itelligenz [Sach]	@lcsb, bvb, dnb, hebis		
	chine-Kommunikation [Sach]	bvb, dnb, hebis		
Neuronales Netz [Sach]	@rvk			
Soft Computing [Sach]	@rvk			
Unterstützungssystem <Informatik> [Sach]	bvb, dnb, hebis			
Aufsatzsammlung [Form]	bvb, dnb			
1: Künstliche Intelligenz [Sach]	k10plus			
2: Soft Computing [Sach]				
3: Mensch-Maschine-Kommunikation [Sach]				
LCSH				
User interfaces (Computer systems).	hbz, k10plus			
Computational intelligence.	hbz, k10plus			
Artificial intelligence.	hbz, k10plus			
RVK				
Ethik der Künstlichen Intelligenz, Maschinenethik, Roboterethik (CC 7270)	@gnd			
Intelligenz; Begabung; Kreativität (CR 5000)	@gnd			

Vorteile der Erschließung mit dem DA-3

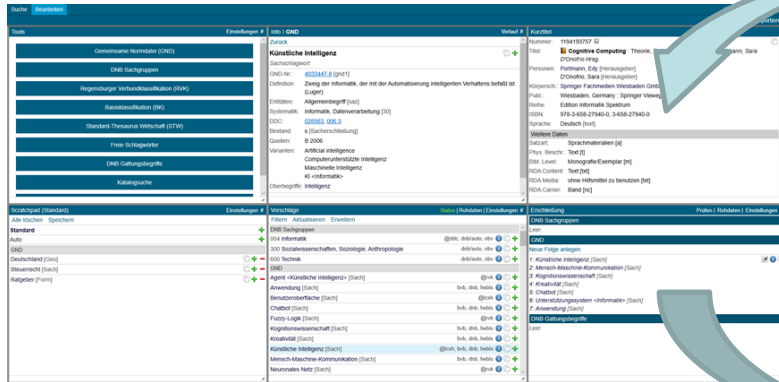
- Bearbeitung des Titels und gleichzeitig Recherche in den GND-Normdatensätzen
- Schlagwortvergabe aus allen Fenstern mit einem Klick (Tools, Info, Vorschläge oder Scratchpad)
- Individuelle Anpassung der Anzeige: z.B.: Tools mit den Erschließungsinstrumenten, Vorschläge
- Gezielte Suche in den Normdaten der GND (nach Schlagwortkategorien getrennt, Kennzeichnung der Formschlagwörter, Anzeige der Homonymenzusätze; Verknüpfung der Schlagwortkombination bei Hinweisdatensätzen)
- Stapelverarbeitung - Listenfunktion

Wünsche und Anregungen

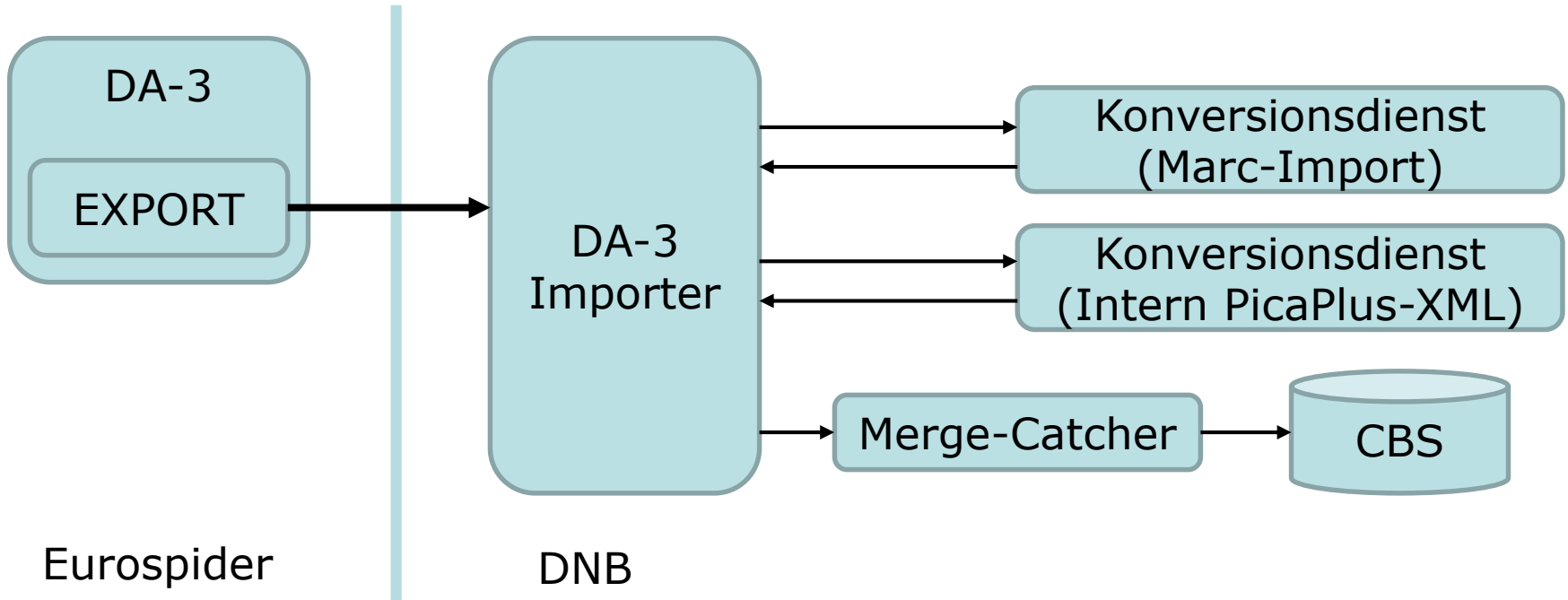
- Verbesserung der Katalogsuche:
 - Suche ausgehend vom GND-Schlagwort nach erschlossenen Titeln
 - Suche nach Titeln mit einer bestimmten Schlagwortfolge
 - Anzeige der Schlagwörter pro Titel innerhalb eines Titeltreffersets
- Einbindung von WebDewey zur DDC-Notations-Vergabe
- Qualitätsmanagement* für die maschinellen Erschließungsdaten der DNB (Sachgruppen, DDC-Kurznotationen, Schlagwörter)
- GND-Normdatenpflege – Mailboxverkehr
- Löschen von bereits vorhandener intellektueller Schlagwortfolgen

* In der DNB eingesetzte Qualitätskontrolle mit einem ordinal skalierten vierstufigen Bewertungssystem

2. Technische Umsetzung der Anbindung des DA-3 an das Katalogisierungssystem der DNB



Export-Schnittstelle DA-3 -> CBS



Abholen (= polling) der geänderten Daten

https://beispiel.url/da3/export/dnb?f=20201029170324

Abfrage wird fünf-sekündlich durchgeführt

alle Änderungen seit angegebenem Datum

Extraktion Marc-Datenfelder aus EXPORT

```
<?xml version="1.0"?>
<item queueid="1687" id="1206485094" date="20201105111656" export-id="dnb">
  <seq type="keyword" sourcetype="new" source="user" system="gnd">
    <prim type="keyword" subtype="subject" sourcetype="new" source="user" system="gnd">
      <datafield tag="689" ind1="0" ind2="0">
        <subfield code="D">s</subfield>
        <subfield code="a">Berufsausbildung</subfield>
        <subfield code="0">(DE-588)4005871-2</subfield>
        <subfield code="2">gnd</subfield>
      </datafield>
    </prim>
    <prim type="keyword" subtype="subject" sourcetype="new" source="user" system="gnd">
      <datafield tag="689" ind1="0" ind2="1">
        <subfield code="D">s</subfield>
        <subfield code="a">Internationaler Vergleich</subfield>
        <subfield code="0">(DE-588)4120509-1</subfield>
        <subfield code="2">gnd</subfield>
      </datafield>
    </prim>
  </seq>
</item>
```

Identifizier
(IDN)

Neu!

Marc-
Felder
(Schlagwort-
Folge)

Einbetten in Marc-Datensatz-“Hülle”

```
<?xml version="1.0" encoding="UTF-8"?>
<collection xmlns="http://www.loc.gov/MARC21/slim">
  <record type="Bibliographic">
    <leader>00000aam a2200000 c 4500</leader>
    <controlfield tag="001">1206485094</controlfield>
    <controlfield tag="003">DE-101</controlfield>
    <datafield tag="689" ind1="0" ind2="0">
      <subfield code="D">s</subfield>
      <subfield code="a">Berufsausbildung</subfield>
      <subfield code="0">(DE-588) 4005871-2</subfield>
      <subfield code="2">gnd</subfield>
    </datafield>
    <datafield tag="689" ind1="0" ind2="1">
      <subfield code="D">s</subfield>
      <subfield code="a">Internationaler Vergleich</subfield>
      <subfield code="0">(DE-588) 4120509-1</subfield>
      <subfield code="2">gnd</subfield>
    </datafield>
  </record>
</collection>
```

Identifizier
(IDN)

Marc-
Felder

-> an Konversionsdienst

Konversion nach PicaPlus-XML

```
<ppxml:record xmlns:ppxml="http://www.oclcpcica.org/xmlns/ppxml-1.0">
  <ppxml:global opacflag="" status="">
    <ppxml:tag id="003@" occ="">
      <ppxml:subf id="0">1206485094</ppxml:subf>
    </ppxml:tag>
    <ppxml:tag id="007I" occ="">
      <ppxml:subf id="S">a</ppxml:subf>
      <ppxml:subf id="0">1206485094</ppxml:subf>
    </ppxml:tag>
    <ppxml:tag id="011@" occ="">
      <ppxml:subf id="a">9999</ppxml:subf>
    </ppxml:tag>
    <ppxml:tag id="041A" occ="">
      <ppxml:subf id="6">gnd4005871-2</ppxml:subf>
    </ppxml:tag>
    <ppxml:tag id="041A" occ="1">
      <ppxml:subf id="6">gnd4120509-1</ppxml:subf>
    </ppxml:tag>
  </ppxml:global>
</ppxml:record>
```

Identifier
(IDN)

Schlagwortfolge
PicaPlus 041A/xx
Pica3 51XX

Verpacken in SRU-RU-Request

```
<?xml version="1.0" encoding="UTF-8"?>
<soap:Envelope xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/">
  <soap:Body>
    <ucp:updateRequest xmlns:ucp="http://www.loc.gov/zing/srw/update/" xmlns:srw="http://www.loc.gov/zing/srw/">
      <srw:version>1.0</srw:version>
      <ucp:action>info:srw/action/1/replace</ucp:action>
      <ucp:recordIdentifier>1206485094</ucp:recordIdentifier>
      <srw:record>
        <srw:recordPacking>xml</srw:recordPacking>
        <srw:recordSchema>info:srw/schema/5/picaxml-v1.0</srw:recordSchema>
        <srw:recordData>
          <ppxml:record xmlns:ppxml="http://www.oclc.org/xmlns/ppxml-1.0">
            <ppxml:global opacflag="" status="">
              <ppxml:tag id="003@" occ="">
                <ppxml:subf id="0">1206485094</ppxml:subf>
              </ppxml:tag>
              <ppxml:tag id="041A" occ="">
                <ppxml:subf id="6">gnd4005871-2</ppxml:subf>
              </ppxml:tag>
              <ppxml:tag id="041A" occ="1">
                <ppxml:subf id="6">gnd4120509-1</ppxml:subf>
              </ppxml:tag>
            </ppxml:global>
          </ppxml:record>
        </srw:recordData>
      </srw:record>
    </ucp:updateRequest>
  </soap:Body>
</soap:Envelope>
```

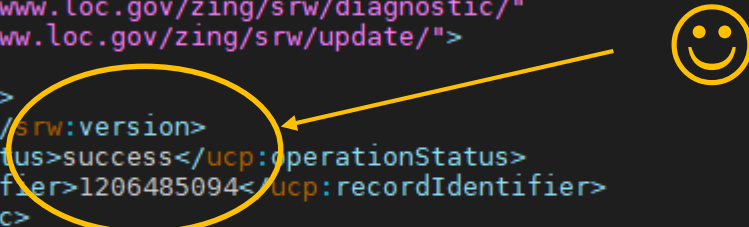
Änderung

Identifizier
(IDN)

Schlagwortfolge

Versenden an Merge-Catcher

```
<?xml version="1.0" encoding="UTF-8"?>
<soap:Envelope
  xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:srw="http://www.loc.gov/zing/srw/"
  xmlns:diag="http://www.loc.gov/zing/srw/ diagnostic/"
  xmlns:ucp="http://www.loc.gov/zing/srw/update/">
  <soap:Body>
    <ucp:updateResponse>
      <srw:version>1.0</srw:version>
      <ucp:operationStatus>success</ucp:operationStatus>
      <ucp:recordIdentifier>1206485094</ucp:recordIdentifier>
      <diag:diagnostic>
        <diag:uri>info:srw/diagnostic/12/0</diag:uri>
        <diag:details>Diagnostics only</diag:details>
        <diag:message>Changed PPN: 1206485094 main-level/bib-record.</diag:message>
      </diag:diagnostic>
    </srw:diagnostics>
  </ucp:updateResponse>
</soap:Body>
</soap:Envelope>
```



In der WinIBW

Schlagwortfolge wurde
zum Satz hinzugefügt
(Merge)

(+) 2020-11-05 11:37:29.050

(+) 2020-11-05 11:17:20.446

(-) 2020-11-05 11:17:02.420

- 001B \$0 1250:05-11-20 \$t 11:16:25.000

+ 001B \$0 1220:05-11-20 \$t 11:17:02.000

+ 041A \$9 040058719

+ 041A/01 \$9 04120509X

+ 041A/09 \$e DE-101 \$r DE-101

(+) 2020-11-05 11:16:25.950

(+) 2020-07-13 07:54:59.583

3010 !121270041!Pilz, Matthias [Tp1]\$BHerausgeber\$4edt

3010 Li, Junmin\$BHerausgeber\$4edt

3100 !1213311454!International G.R.E.A.T. Conference Lost in VET? S

3110 |m|!1043386068!Springer Fachmedien Wiesbaden GmbH [Tb1]\$

4000 Comparative Vocational Education Research : Enduring Challeng

4030 Wiesbaden : Springer VS ***5107608

4060 XV, 329 Seiten

4061 Illustrationen

4062 21 cm, 457 g

4190 Internationale Berufsbildungsforschung

4190 Research

4201 Third international conference of the German Research Center fo

4243 Erscheint auch als\$Online-Ausgabe!1212606825!-Oaf--: Comp

4700 |PE|*Parallelverknüpfung wurde automatisch erstellt

4700 |ME-F|

4700 |FE-F|

4715 =u <http://deposit.dnb.de/cgi-bin/dokserv?id=1f8d417cb9cf49d082>

4715 =u <http://www.springer.com/=d MVB=e 2>

4715 =u \$=c 04=d DNB=e 1

5050 370\$Ei\$Hi\$e-sg\$D2020-11-05

5100 !040058719!Berufsausbildung [Ts1]

5101 !04120509X!Internationaler Vergleich [Tsz]

5102 !041229266!Bildungsforschung [Ts1]

5109 (DE-101){DE-101}

5400 [DDC23ger]370.11309

5401 370.113

Fehlerbehandlung

```
<?xml version="1.0" encoding="UTF-8"?>
<soap:Envelope
  xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:srw="http://www.loc.gov/zing/srw/"
  xmlns:diag="http://www.loc.gov/zing/srw/diagnostic/"
  xmlns:ucp="http://www.loc.gov/zing/srw/update/">
  <soap:Body>
    <ucp:updateResponse>
      <srw:version>1.0</srw:version>
      <ucp:operationStatus>fail</ucp:operationStatus>
      <ucp:recordIdentifier>1173821767</ucp:recordIdentifier>
      <diag:diagnostic>
        <diag:uri>info:srw/diagnostic/12/12</diag:uri>
        <diag:details>Cataloguing/validation error.</diag:details>
        <diag:message>Beginn der Sachgruppe u. der synthetischen Notation stimmen nicht ueberein</diag:message>
      </diag:diagnostic>
    </srw:diagnostics>
  </ucp:updateResponse>
</soap:Body>
</soap:Envelope>
```

Technische Anpassungen

- Auslieferung der Statusangaben in Marc-Feld 925
- Verknüpfung zur WinIBW über Protokoll-Handler

`ibw://1219372293`

- Anpassung Konversion Marc-Felder 689 für Schlagwortfolgen
- Titelsuche über AKZ (Akzessionsnummer der DNB)

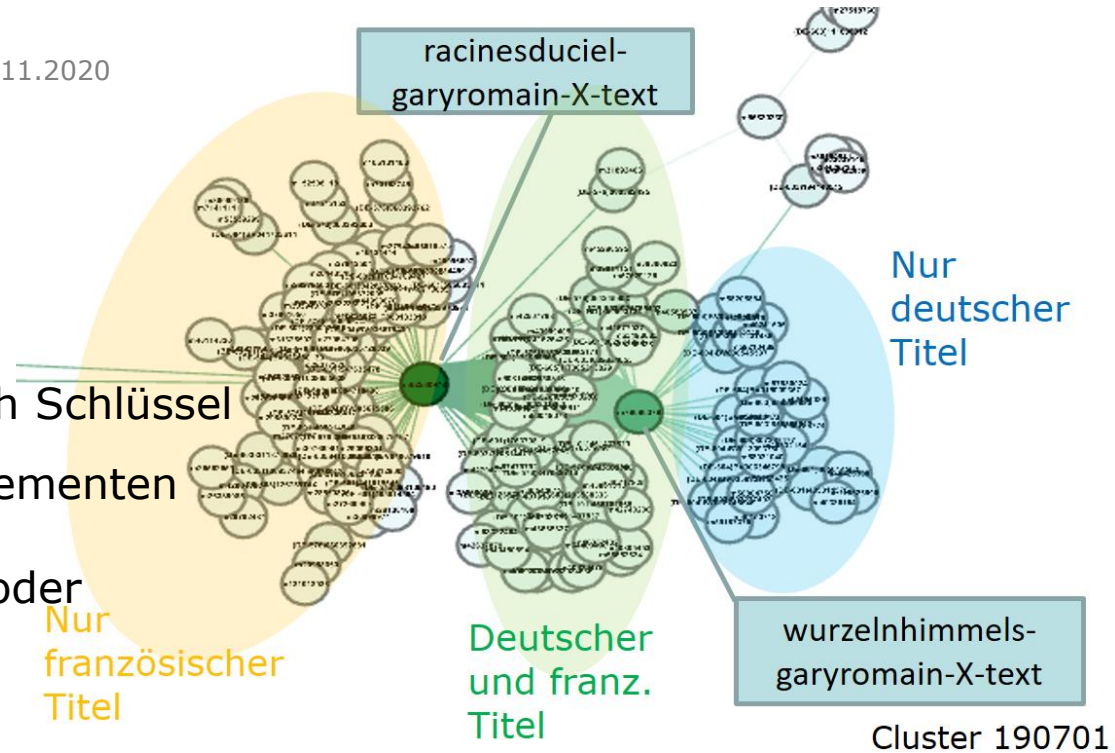
3. Vorschläge generieren mit der Plattform Culturegraph

Datengrundlage



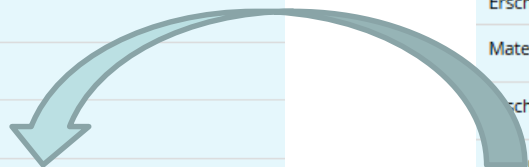
Werkbündel

- Abgleich der Publikationen durch Schlüssel aus verschiedenen Metadatenelementen
- Für jeden Datensatz liegen ein oder mehrere Schlüssel vor
- Bündel vereinen alle Datensätze, die mindestens einen gemeinsamen Schlüssel mit einem anderen Mitglied besitzen



Anreicherungen

Hauptsachtitel	Das Meer Als Versprechen
Zusatz	Bedeutung Und Funktion Von Seeherrschaft Bei Thukydides
Person	aut Kopp, Hans
Körperschaft	-
Umfang	1 Online-Ressource (303 Seiten)
Erscheinungsjahr	2017
Material	file
Erscheinungsort	Göttingen
Herausgeber	Vandenhoeck & Ruprecht
Standard-Identifizier	DOI 10.13109/9783666253249 EKI 1002269444
Verlags-Identifizier	-



Anreicherung

Hauptsachtitel	Das Meer Als Versprechen
Zusatz	Bedeutung Und Funktion Von Seeherrschaft Bei Thukydides
Inhalt	Inhaltsverzeichnis Inhaltstext Inhaltstext
Person	aut Kopp, Hans 1127785095
Körperschaft	-
Umfang	303 Seiten
Erscheinungsjahr	2017
Material	book
Erscheinungsort	Bristol, CT
Herausgeber	Vandenhoeck & Ruprecht
Schlagwort	655 Hochschulschrift 4113937-9 650 Seeherrschaft 4077302-4 600 Thucydides 4138063-0
Klassifikation	DDC 938.05 RVK FH 26175 RVK NH 5850
Sachgruppe	DNB 355 DNB 880 DNB 930
Standard-Identifizier	DNB 1115738097 OCLC 968670393 EKI DNB1115738097
Verlags-Identifizier	GTIN14 9783525253243 ISBN13 9783525253243

Mögliche Anbindung an DA-3

- Zurzeit retrospektive Anreicherung der Bestände mit inhaltserschließenden Elementen in der Umsetzung
- Wünschenswert auch Einbinden in den Erschließungsprozess von Neuzugängen, z.B. über DA-3
- Dafür muss zunächst Culturegraph-API erstellt werden

Herzlichen Dank für Ihre Aufmerksamkeit

Fragen beantworten wir gerne! Auch später jederzeit:

Helga Karg: h.karg@dnb.de

Hans-Jürgen Becker: h.becker@dnb.de

Dr. Angela Vorndran: a.vorndran@dnb.de