

Inhaltliche Erschließung im Österreichischen Bibliothekenverbund OBV

Status quo – Daten und Fakten

Fragen zur inhaltlichen Erschließung im OBV:

- Wieviel % aller Datensätze (Verbundkatalog) sind (verbal/GND) inhaltlich erschlossen?
- Wieviel % der Neuzugänge werden mit GND erschlossen?
- Welche Klassifikationssystematiken bzw. andere Systeme werden komplementär eingesetzt? (DDC, RVK, BK, Eurothesaurus, ...)
- Was wird derzeit nicht / nicht ausreichend erschlossen?
- Wie werden Fachreferent*innen und andere Sacherschließende geschult?
- Was sind die Pläne für automatische Inhaltserschließung?

Umfrage zur Sacherschließungspraxis im OBV (Ergebnisse)

Umfrage 2017

IST-Analyse der Inhaltlichen Erschließung (IE) im OBV

Fragebogen – Umfrage im März / April 2017 (Ausschnitt)

Umfrage zu Umfang und Art der Inhaltlichen Erschließung an den Bibliotheken des Österreichischen Bibliothekenverbundes (OBV)

bitte ankreuzen!

Wird an Ihrer Bibliothek aktiv inhaltliche Erschließung durchgeführt?

Wenn nein – warum nicht?

Wenn ja

→ **Verbale inhaltliche Erschließung**

Nach welchen Regelwerken erschließt Ihre Bibliothek?

Regeln für den Schlagwortkatalog **RSWK** (Gemeinsame Normdatei **GND**)

Zusatzfrage: Werden im Bedarfsfall Neuansetzungen in der GND durchgeführt? (oja/onein)

Anderes Regelwerk – wenn ja, welche(s)?

z. B. lokale inhaltliche Erschließung nach eigenen Regeln...

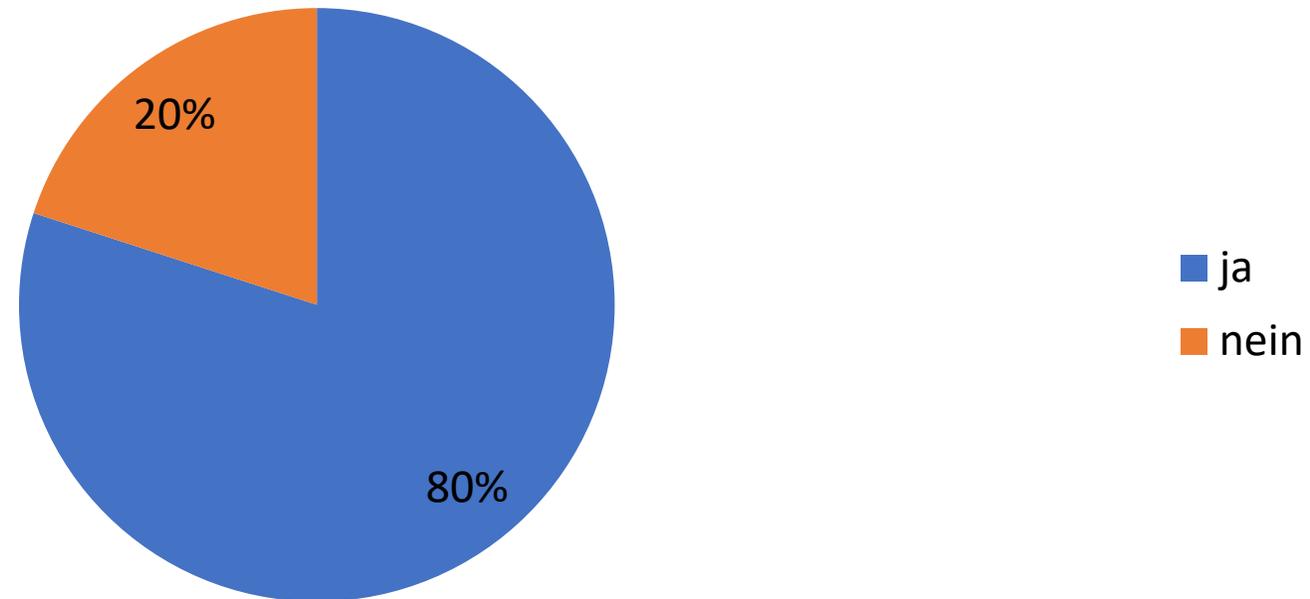
→ **Klassifikatorische inhaltliche Erschließung**

Welche Verbundklassifikation wird an Ihrer Bibliothek

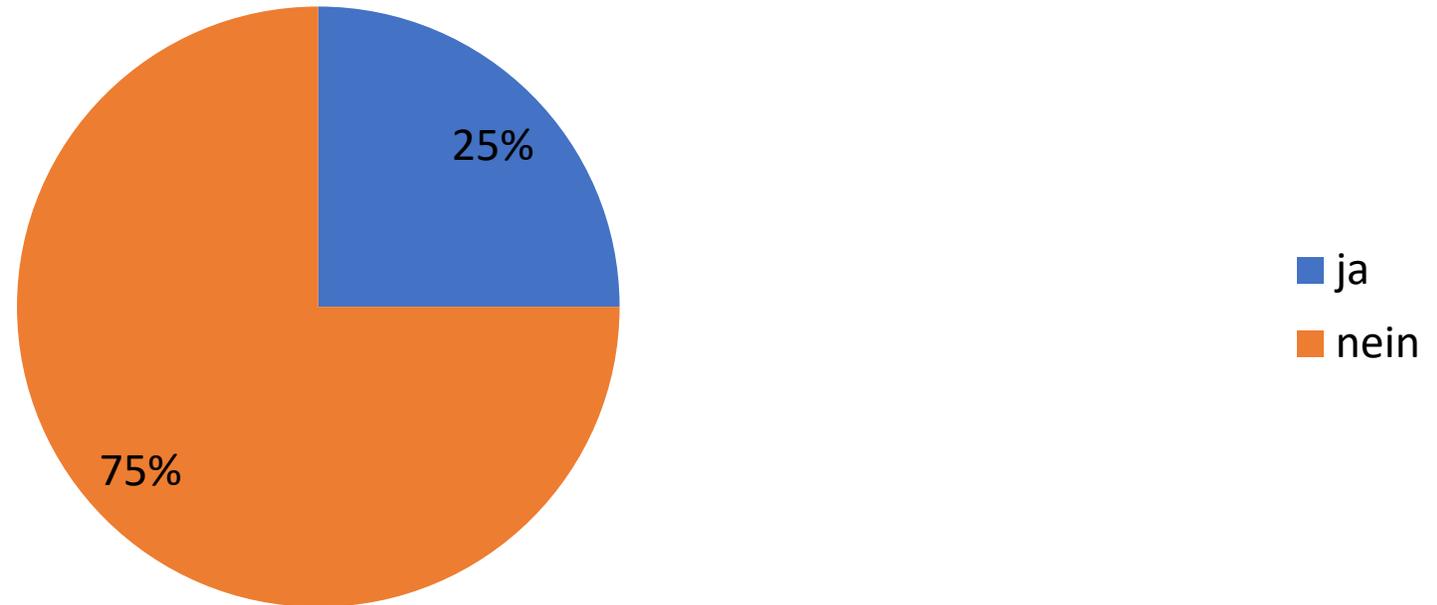
übernommen
(Fremddatennutzung)?

aktiv eingegeben?

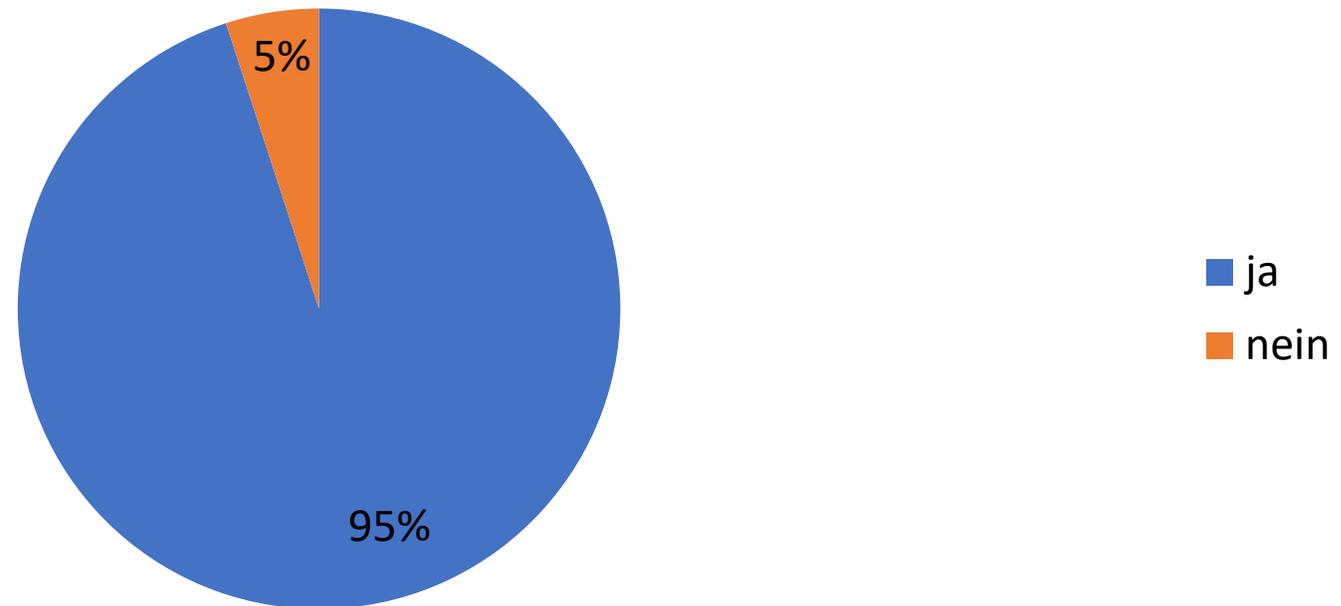
IE - RSWK



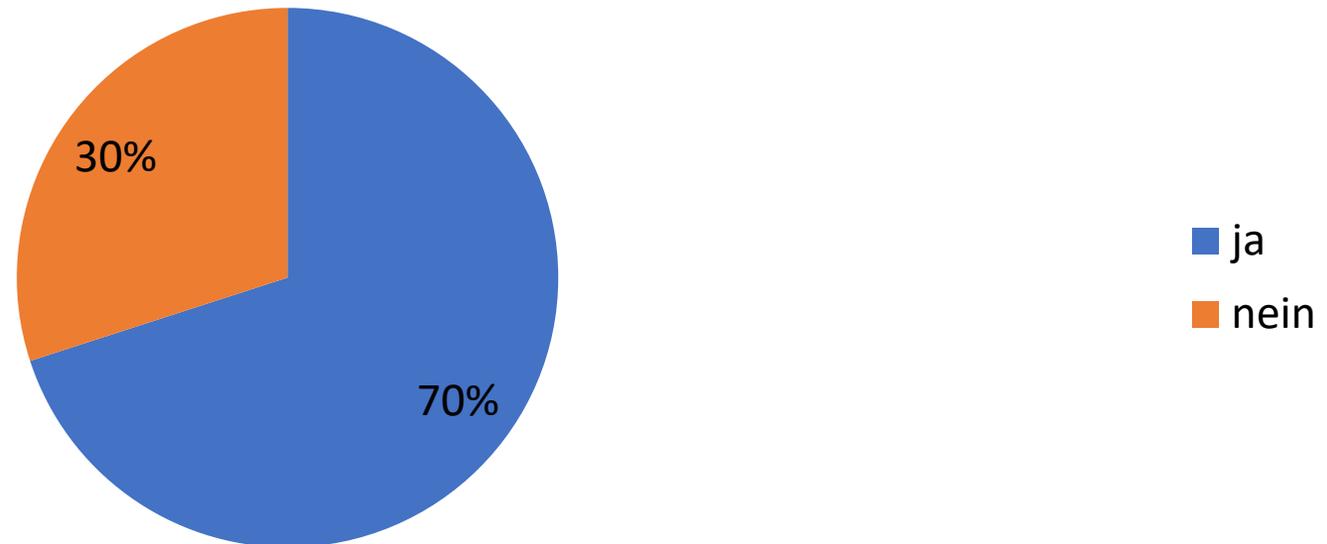
GND Neuansetzung



Übernahme von Klassifikationen



Interesse an Fremddaten / automationsunterstützte Datengenerierung



Zusammenfassend

Eine überwiegende **Mehrheit** der Verbundbibliotheken erschließt **aktiv** nach **RSWK / GND**.

Ein **Viertel** der Verbundbibliotheken betreibt aktiv **Normdatenarbeit** in der **GND**.

Eine überwiegende **Mehrheit** der Verbundbibliotheken beteiligt sich an der **Übernahme von Klassifikationsdaten**.

Bei Fremddatenübernahme – möglichst **keine IE-Daten löschen**
(-> mögliche Grundlage für allfällige Datenabgleiche zur späteren
Kataloganreicherung)

Eine Mehrheit der Verbund-Bibliotheken ist an
Datenanreicherung / Kataloganreicherung durch
automationsunterstützte Verfahren **interessiert**

Aber: ohne eigenen Dateneintrag möglichst vieler Institutionen
wird es auf Dauer **keine Fremddatennutzung** geben... - eigene
Daten, die zu Fremddaten werden, Fremddaten, die zu eigenen
mutieren – **Geben und Nehmen...**

Wieviel % der Titeldaten im Verbundkatalog sind verbal / systematisch inhaltlich erschlossen?

Ca. **50%** aller Titeldaten sind nach **RSWK / GND verbal** inhaltlich erschlossen.

Jeweils **etwa 20%** sind nach **RVK** bzw. nach **BK klassifikatorisch** inhaltlich erschlossen.

Etwa **5%** sind nach **DDC klassifikatorisch** inhaltlich erschlossen.

Dabei ist zu berücksichtigen, dass aus zahlreichen **Retroprojekten** lediglich **formale Metadaten** Teil des Katalogs sind (ohne inhaltliche Erschließung). Dazu kommen etwa rund **2 Mio. Aufsatzdaten** (davon 1,5 Mio. IVSCAN), die in der Regel ebenfalls **nicht inhaltlich erschlossen** sind.

Daher ist der (vergleichbare) Anteil inhaltlich erschlossener Literatur wesentlich höher und liegt bei **über 60%**.

An der **ÖNB** ist der Anteil sogar noch höher, weil historische Schlagwort-Kataloge vorhanden sind und diese Metadaten mit den Normdaten / SWD -> GND abgeglichen wurden: fast **75%** der Titeldatensätze sind verbal inhaltlich erschlossen ... 😊

Aktuelle Daten aus der ALMA-Statistik zur Inhaltlichen Erschließung (IE):

(aktuelle Datenabfrage und –analyse von Ing. Josef Labner und Mag. Stefan Majewski, OBVSG, vom 9.11.2021)

14,551.272 Datensätze gesamt -> minus unselbständige Literatur (Aufsätze,...), IE nur marginal (ARIADNE)
2,000.033

12,551.239 Datensätze -> minus E-Ressourcen, Migrationen / Retro-Projekte ohne IE (v. a. Altbestände, Handschriften, Sonderbestände nur mit lokaler IE, ...)
1,500.000 (geschätzt – Einzelprojekte nicht genau quantifizierbar)

11,051.239 Datensätze relevant (prozentuelle Basis für die Anteile der verbalen / klassifikatorischen IE)



Anteile verbaler / klassifikatorischer Inhaltlicher Erschließung:

GND-Normdaten (MARC 6XX, außer 655)	-> 6,410.627	58%
Formangaben und Gattungen IE (655 \$\$22 bellobov)	-> 157.393	1,4%
Verbund-Klassifikationen (RVK, BK, DDC, NLM, MSC) (überlappend mit GND)	-> 3,835.941	34,7%
Gesamt (zumindest eins der o. a. Kriterien) „verodert“☺	-> 7,201.180	65,2%

Einbezug der Fremddaten ...

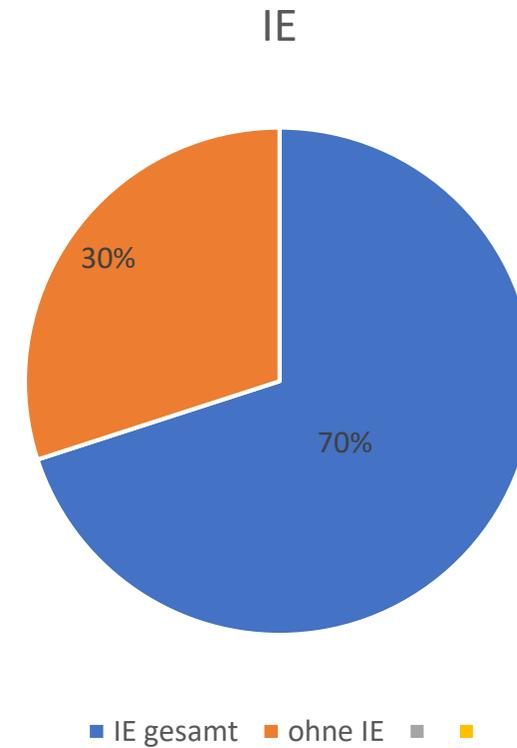
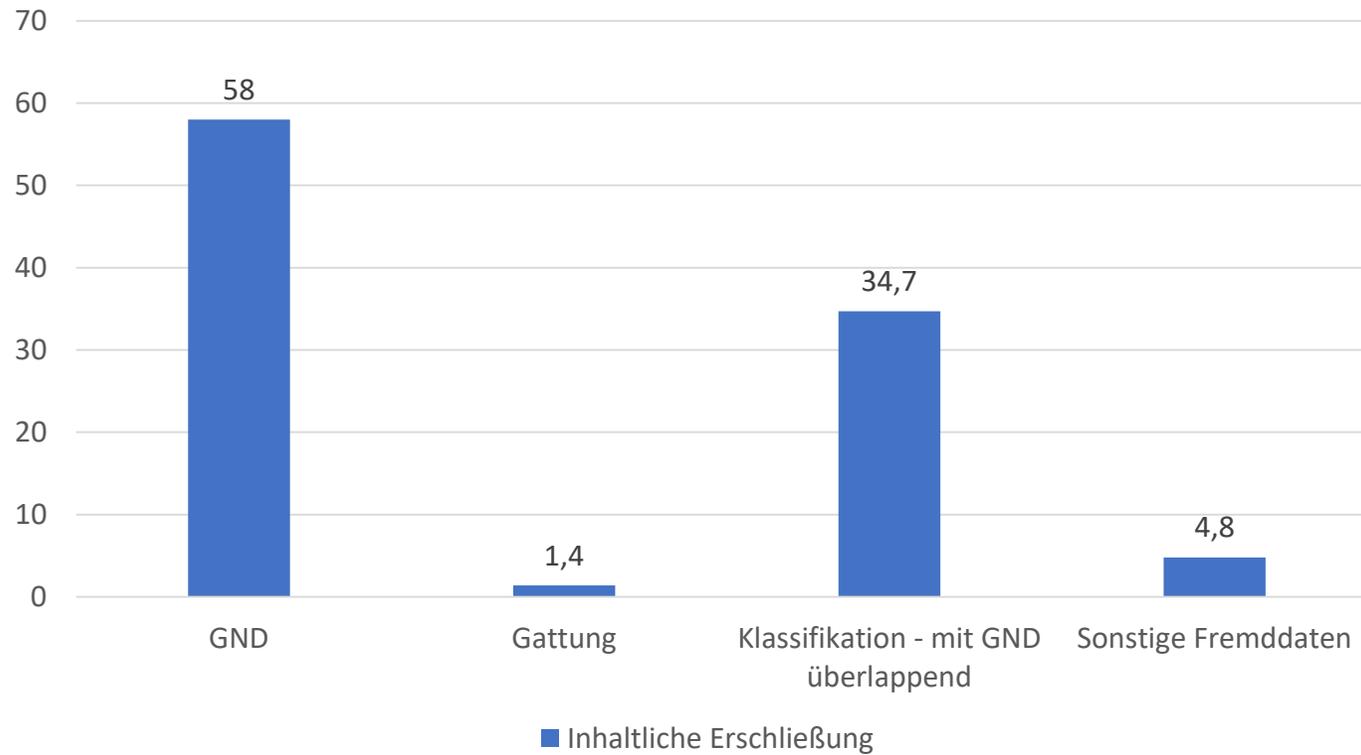
Freie Schlagwörter

Schlagwörter diverser Provenienzen (z. B. LCSH, MeSH, FAST)

Diverse andere durch Fremddaten eingespielte Klassifikationen

...dadurch steigt der Anteil der IE auf bis zu **70%**

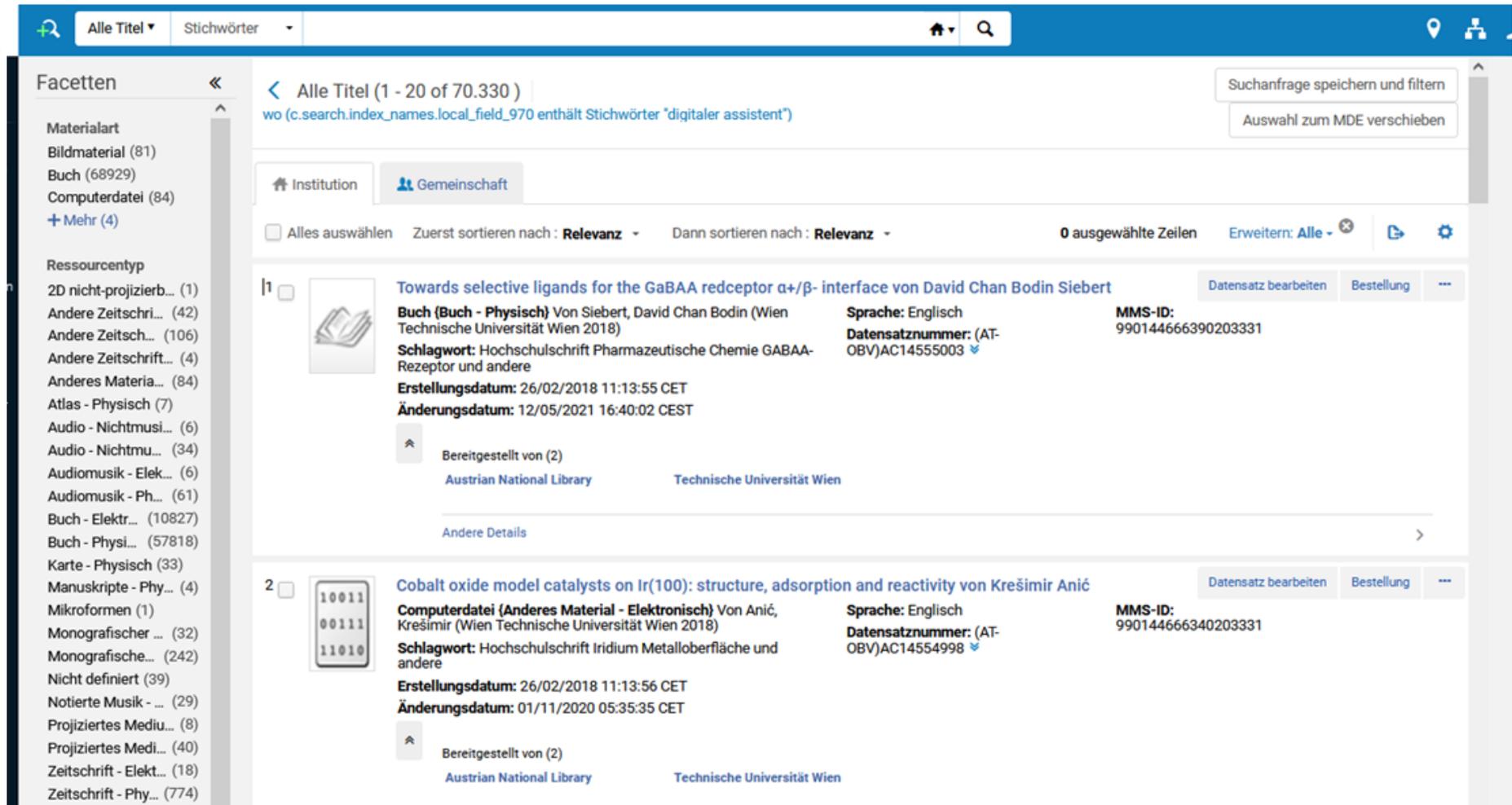
Inhaltliche Erschließung



Inhaltliche Erschließung der jährlichen Produktion ...

Darüber liegen keine genaue Daten vor – **aktuelle Erschließungsarbeit** und diverse **Migrationsprojekte (Wechsel des Bibliotheksverwaltungssystems – ALEPH auf ALMA** in über 9 Tranchen / „Kohorten“ / „Waves“) sind derzeit datentechnisch nur sehr aufwendig voneinander zu trennen. Daher müssen fein-granulare statistische Erhebungen auf einen späteren Zeitpunkt verschoben werden. Sowohl die Umfrage bei den teilnehmenden Verbundbibliotheken als auch die Analyse des Gesamtbestandes der Network Zone (des Verbundkatalogs) lassen jedoch Rückschlüsse auf den IST-Stand zu.

Erfreuliche Details
vom Einsatz des
DA-3:
Die Suche in der NZ
mit dem Index
"local_field_970" =
Digitaler Assistent
ergibt **70.330** Treffer...!



The screenshot shows a search results page with a blue header. The search bar contains 'Alle Titel' and 'Stichwörter'. The search results are displayed in a list format. The first result is a book titled 'Towards selective ligands for the GABA_A receptor α / β - interface' by David Chan Bodin Siebert. The second result is a computer file titled 'Cobalt oxide model catalysts on Ir(100): structure, adsorption and reactivity' by Krešimir Anić. The interface includes a left sidebar with facets, a top navigation bar, and a main content area with search filters and sorting options.

Facetten

- Materialart
 - Bildmaterial (81)
 - Buch (68929)
 - Computerdatei (84)
 - + Mehr (4)
- Ressourcentyp
 - 2D nicht-projizierb... (1)
 - Andere Zeitschri... (42)
 - Andere Zeitsch... (106)
 - Andere Zeitschrift... (4)
 - Anderes Materia... (84)
 - Atlas - Physisch (7)
 - Audio - Nichtmusi... (6)
 - Audio - Nichtmu... (34)
 - Audiomusik - Elek... (6)
 - Audiomusik - Ph... (61)
 - Buch - Elektr... (10827)
 - Buch - Physi... (57818)
 - Karte - Physisch (33)
 - Manuskripte - Phy... (4)
 - Mikroformen (1)
 - Monografischer ... (32)
 - Monografische... (242)
 - Nicht definiert (39)
 - Notierte Musik - ... (29)
 - Projiziertes Mediu... (8)
 - Projiziertes Medi... (40)
 - Zeitschrift - Elekt... (18)
 - Zeitschrift - Phy... (774)

Alle Titel (1 - 20 of 70.330)
wo (c.search.index_names.local_field_970 enthält Stichwörter "digitaler assistent")

Suchanfrage speichern und filtern
Auswahl zum MDE verschieben

Institution | Gemeinschaft

Alles auswählen | Zuerst sortieren nach: Relevanz | Dann sortieren nach: Relevanz | 0 ausgewählte Zeilen | Erweitern: Alle |

1  **Towards selective ligands for the GABA_A receptor α / β - interface** von David Chan Bodin Siebert Datensatz bearbeiten Bestellung ...

Buch (Buch - Physisch) Von Siebert, David Chan Bodin (Wien Technische Universität Wien 2018) **Sprache:** Englisch **MMS-ID:** 990144666390203331

Schlagwort: Hochschulschrift Pharmazeutische Chemie GABAA-Rezeptor und andere **Datensatznummer:** (AT-OBV)AC14555003

Erstellungsdatum: 26/02/2018 11:13:55 CET **Änderungsdatum:** 12/05/2021 16:40:02 CEST

Bereitgestellt von (2)
Austrian National Library Technische Universität Wien

Andere Details >

2  **Cobalt oxide model catalysts on Ir(100): structure, adsorption and reactivity** von Krešimir Anić Datensatz bearbeiten Bestellung ...

Computerdatei (Anderes Material - Elektronisch) Von Anić, Krešimir (Wien Technische Universität Wien 2018) **Sprache:** Englisch **MMS-ID:** 990144666340203331

Schlagwort: Hochschulschrift Iridium Metalloberfläche und andere **Datensatznummer:** (AT-OBV)AC14554998

Erstellungsdatum: 26/02/2018 11:13:56 CET **Änderungsdatum:** 01/11/2020 05:35:35 CET

Bereitgestellt von (2)
Austrian National Library Technische Universität Wien

... das sind rund **4000 Datensätze pro Monat**

bzw. (wenn man ca. 1,5 Jahre rechnet und ca. 220 Arbeitstage im Jahr),

über 200 Datensätze pro Tag,

die mit dem DA-3 angereichert werden!

...Tendenz steigend 😊



Komplementäre Systeme inhaltlicher Erschließung – Verbundklassifikationen u. a. m.:

(alle Systeme sind im DA-3-Client des OBV eingepflegt ☺)

- **RVK** v. a. im universitären Bereich (Aufstellungssystematik)
- **BK** ideal als zusätzlicher Sucheinstieg (Universitätsbibliotheken und andere Großbibliotheken)
- **DDC** einzelne Bibliotheken
- **bellobv** Gattungsbegriffe (Bibliotheken mit namhaftem Belletristik-Bestand – Pflichtexemplare)
- **MSc** Fachbibliotheken [Mathematics Subject Classification]
- **NLM** Fachbibliotheken [National Library of Medicine Classification]
- **OBV-Fachgruppen** (einzelne Bibliotheken – zu statistischen Zwecken – Neuzugänge nach Fachgebieten)

Was wird derzeit nicht / nicht ausreichend erschlossen?

- **Aufsätze** (z. B. IVSCAN) aus diversen Sammelwerken (Konferenz-, Festschriften, Jahrbücher, Ausstellungskataloge,...)
- außer: Dokumentation (z. B. ARIADNE, Parlament etc.)
- **E-Book-Pakete** (Jahresabos) – außer: einzelne E-Books
- **Retroprojekte** (Formalerfassung – kaum Inhaltserschließung)
- **Forschungsdaten** (nur in Auswahl – je nach Projekt)

Schulung von Mitarbeiter*innen für die Inhaltliche Erschließung:

- **Gremien des OBV** (GND-Level1-Redaktion, ZRSE / ZR) – **Normdaten-Schulung**
- **ULG** (Universitätslehrgang – Library and Information Studies LIS)
- ggfs. **Brain-Pool-Kurse** (Bibliothekarische Fortbildung – Österreichische Nationalbibliothek ÖNB)
- **Einschulung** in den einzelnen Institutionen (z. B. durch **Lokalredaktionen**)

Zukunftspläne – automatische Inhaltserschließung ...

- Erster Schritt: **DA-3** 
- **KI / AI ?** Volltexte – Textmining, distant reading, Kombination von vorhandenen Metadaten / Normdaten mit statistisch-linguistischen Verfahren ... ???

à la Petrus-Projekt (DNB) – ML-Verfahren oder
Vergabe von DDC-Sachgruppen mittels eines Schlagwort-Thesaurus (GND) – Semantic Web Technologie



Teilaspekt: **Forschungsdaten ...**

Beispiele:

- **UBW – Phaidra** (Universität Wien)
[Forschungsdatenmanagement \(univie.ac.at\)](https://univie.ac.at/research-data)
- **ÖNB – Literaturarchiv – Digitale Editionen** (Österreichische Nationalbibliothek)
(Metadaten noch nicht im Bibliothekssystem ALMA)

PHAIDRA ist ...

... das Repositorium zur dauerhaften Sicherung von digitalen Beständen an der Universität Wien.

Der Dienst ist unter folgender Adresse erreichbar:

phaidra.univie.ac.at

Die in PHAIDRA archivierten digitalen Objekte erhalten dauerhafte Adressen (z.B. handle, bei Büchern URN, für Publikationen und Forschungsdaten DOI), werden mit strukturierten Metadaten (fünf Pflichtfelder: Titel, Sprache, Beschreibung, Rolle/n, Lizenz) versehen und von Suchmaschinen gefunden (s. **Metadaten in Phaidra**)

Die Vorteile von PHAIDRA

- Offener und freier Zugang (alle Mitarbeiter*innen und alle Studierende der Universität Wien können PHAIDRA aktiv nutzen)
- Persistente Zitierbarkeit (handle-Link, DOI auf Anfrage)
- Differenziertes Zugriffskonzept
- Die Daten können lizenziert werden
- Verlässliche Langzeitverfügbarkeit ohne zeitliche Beschränkung
- auf Dublin Core und LOM basierende **Metadatenschemata mit fünf Pflichtfeldern (Titel, Sprache, Beschreibung, Rolle/n, Lizenz)**

In manchen Fachdisziplinen können Metadaten auch die eigentlichen Forschungsdaten sein. In jedem Fall sind Metadaten mehr als nur Beschreibungen, deshalb ist es uns ein großes Anliegen, dass Metadaten vor dem Upload in das Repository optimal vorbereitet und ständig gepflegt werden. Datenmanagementpläne helfen dabei, den Umgang mit Metadaten vorzubereiten.

In PHAIDRA wird [Dublin Core \(dc\)](#) und ein erweitertes LOM-Schema verwendet (s. [Phaidra University of Vienna Metadata \(PDF\)](#) by Lorisa Andreoli et al., University of Padua 2019).

Derzeit sind folgende fünf Metadatenfelder in PHAIDRA verpflichtend auszufüllen:

- Titel des Objekts (dc:title)
- Beschreibung (dc:description)
- Sprache des Objekts (dc:language)
- Rolle/n (dc:contributor) Hier können Sie auch Authority Files eintragen
- Lizenz (dc:rights): Wenn Sie eine Lizenz wählen, beachten Sie bitte, dass dies nicht mehr rückgängig gemacht werden kann

Zusätzlich stehen Ihnen weitere Metadatenfelder zur Verfügung, die die Auffindbarkeit und eventuell die Visualisierung ihrer Objekte unterstützen. Sie haben auch die Möglichkeit einen **DOI** (Digital Object Identifier; [DOI Service der Universität Wien](#)) bei den Metadaten in Phaidra einzutragen, und zwar im ersten Reiter "Allgemeines" des Metadateneditors, ganz unten unter "Identifikatoren".

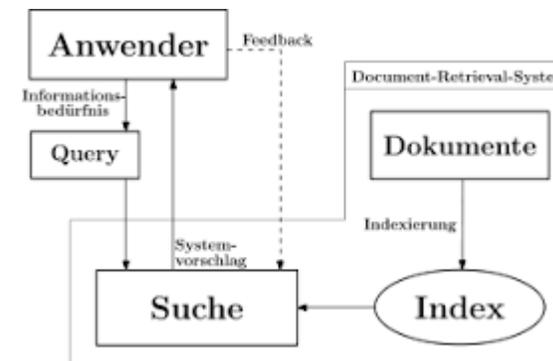
Es stehen zur Zeit folgende Klassifikationen in PHAIDRA zur Verfügung:

- ÖFOS 2002
- The ACM Computing Classification System [1998 Version]
- EuroVoc 4.2
- Physics and Astronomy Classification Scheme
- Getty Art & Architecture Thesaurus
- Getty Union List of Artist Names
- Basisklassifikation **Inhaltliche Erschließungselemente!**
- Dewey - Dezimalklassifikation
- ÖFOS 2012

Ausblick...

... in den nächsten Monaten ist eine durchgängige Recherchemöglichkeit für alle Forschungsdaten geplant – sozusagen ein

Forschungsretrieval



Danke für Ihre Aufmerksamkeit!

Ansprechpersonen bei Fragen:

Ing. Josef Labner OBVSG

Josef.labner@obvsg.at

Dr. Christoph Steiner ÖNB

Christoph.steiner@onb.ac.at