

Ulrike Junger

Inhaltliche Erschließung

Aktuelle Entwicklungen in der DNB

Kurzer Rückblick

Datenbezieher-Workshop 2017:

- Vorstellung des Konzepts zur Weiterentwicklung der
Inhaltserschließung
 - Einsatz maschineller Verfahren für Netzpublikationen und wo möglich für physische Medienwerke
 - Entwicklung eines Systems von DDC-Kurznotationen, die sich auch maschinell vergeben lassen
 - Intellektuelle Inhaltserschließung da, wo erforderlich

Kurzer Rückblick

– Erste Umsetzungsschritte

- Maschinelle Inhaltserschließung für gedruckte Publikationen außerhalb des Buchhandels und Hochschulschriften (Reihe B und H)
- Entwicklung eines Verfahrens zur Übernahme maschinell erzeugter Inhaltserschließungsdaten von Online- auf parallele Printausgaben
- Entwicklung eines Systems von DDC-Kurznotationen

Aktuelle Entwicklungen kurz gefasst

1. Maschinelle Beschlagwortung englischsprachiger Online-Dissertationen
2. Maschinelle Beschlagwortung von Print-Publikationen der Reihen B und H
3. Übernahme von maschinellen Erschließungsdaten der Online-Publikation in die Metadaten der parallelen Print-Publikation
4. Übergangsszenario für intellektuelle Inhaltserschließung ab 1. Juli 2019

Maschinelle Beschlagwortung

Start: 2014

- Deutschsprachige monografische Online-Publikationen der Reihe O
 - ohne Zeitschriftentitel
 - ohne Belletristik
- Deutschsprachige Online-Zeitschriftenartikel

Maschinelle Beschlagwortung

Seit 2017

- Deutschsprachige Print-Publikationen der Reihen B und H

Seit 2018

- Englischsprachige Online-Dissertationen der Reihe O

- Gesamtzahl maschinell beschlagwortet:
363.158 Publikationen
(Stand März 2019)

Maschinelle Beschlagwortung

- Methode: Computerlinguistisches Verfahren auf der Basis eines Wörterbuches
- Basis: Bibliografische Titeldaten und elektronische Texte (Volltexte oder digitalisierte Inhaltsverzeichnisse)
- Terminologie:
 - Gemeinsame Normdatei (GND)
 - **NEU seit Juni 2018: Library of Congress Subject Headings (LCSH)** in Kombination mit Nicht-Sachbegriffen der GND (Personen, Geografika, Körperschaften, Kongresse und Werke)

Maschinelle Beschlagwortung

Beispiel englischsprachige Dissertation

- 4000 Reconstructing European Copyright Law for the Digital Single Market :
Between Old Paradigms and Digital Challenges / Bernd Justin Jütte
- 4204 \$dDissertation\$eUni. Luxembourg\$f2016
- 5050 **340**\$Em\$Haep-sg\$K0,99414\$D2018-06-05
- 5540 [**GND**]!940431181!**Europäische Union** [Tb1]\$K0,00119\$D2018-06-05
- 5540 [**LCSH**]**Copyright**\$Lsh85032446\$uhttp://id.loc.gov/authorities/subjects/
sh85032446\$K0,20713\$D2018-06-05

Anzeige in MARC-21

- 610 2\$**81**\p\$(DE-588)5098525-5\$http://d-nb.info/gnd/5098525-5\$(DE-101)940431181\$a**Europäische Union**\$2gnd
- 650 7\$**82**\p\$a**Copyright**\$(DLC)sh85032446\$http://id.loc.gov/authorities/subjects/sh85032446\$2lcsch
- 883 0 \$**81**\p\$a**maschinell gebildet**\$c0,00119\$d20180605\$qDE-101
- 883 0 \$**82**\p\$a**maschinell gebildet**\$c0,20713\$d20180605\$qDE-101

Auslieferung der LCSH seit Februar 2019

Bibliografiejahrgang 2017 maschinell beschlagwortet



Bibliografiejahrgang 2018 maschinell beschlagwortet



Qualitätsanforderungen an die maschinelle Beschlagwortung

Die einzelnen GND-Terme sollen das Thema der Publikation richtig und sinnvoll wiedergeben durch:

- Nützliche, gebräuchliche Schlagwörter
- Angemessene Anzahl der relevanter Schlagwörter
- Aktuelle / Gültige Schlagwörter
- Spezifität / Genauigkeit der Schlagwörter im Hinblick auf das vorhandene Vokabular
- Konsistente Vergabe von Schlagwörtern

Bewertungskriterien - Schlagwort

Das einzelne Schlagwort beschreibt einen wesentlichen Aspekt des Textes ausreichend und trifft absolut zu.

Sehr nützlich 3

Das einzelne Schlagwort beschreibt einen wichtigen Aspekt des Textes aus einer etwas weiteren/engeren Perspektive zutreffend.

Nützlich 2

Das einzelne Schlagwort beschreibt einen wichtigen Aspekt des Textes nicht ausreichend, ist aber auch nicht völlig unzutreffend oder falsch.

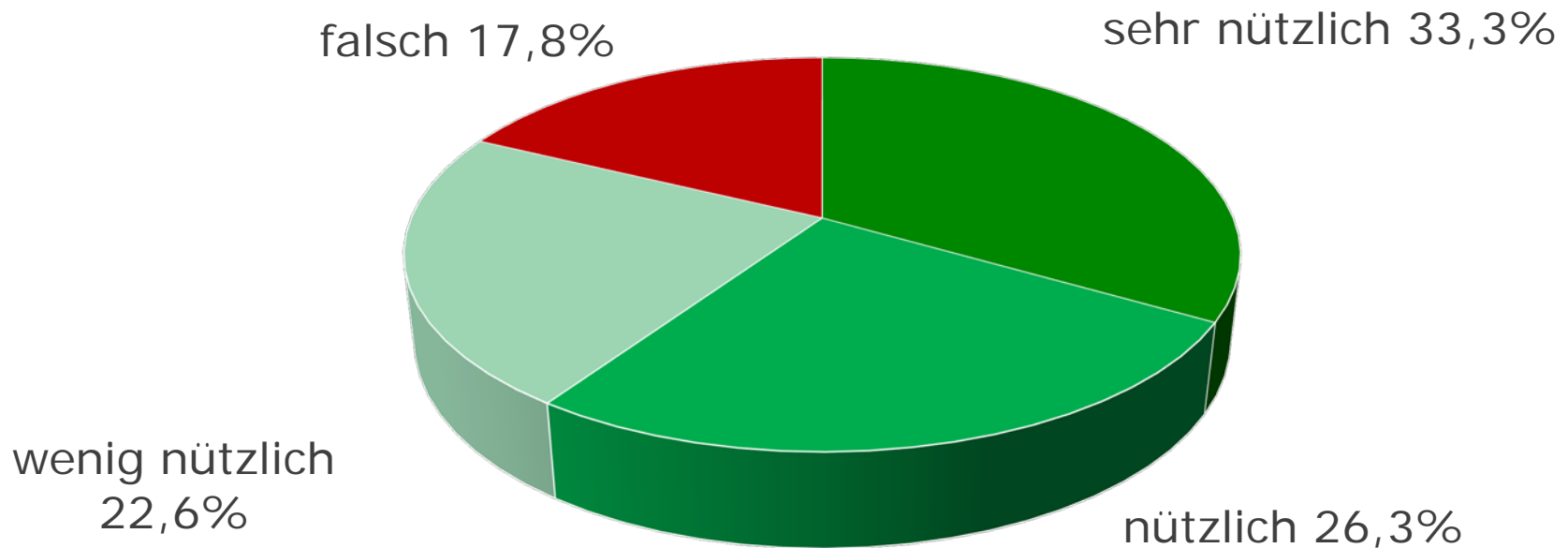
Wenig nützlich 1

Das einzelne Schlagwort beschreibt keinen wichtigen Aspekt des Textes und ist falsch oder nicht nützlich.

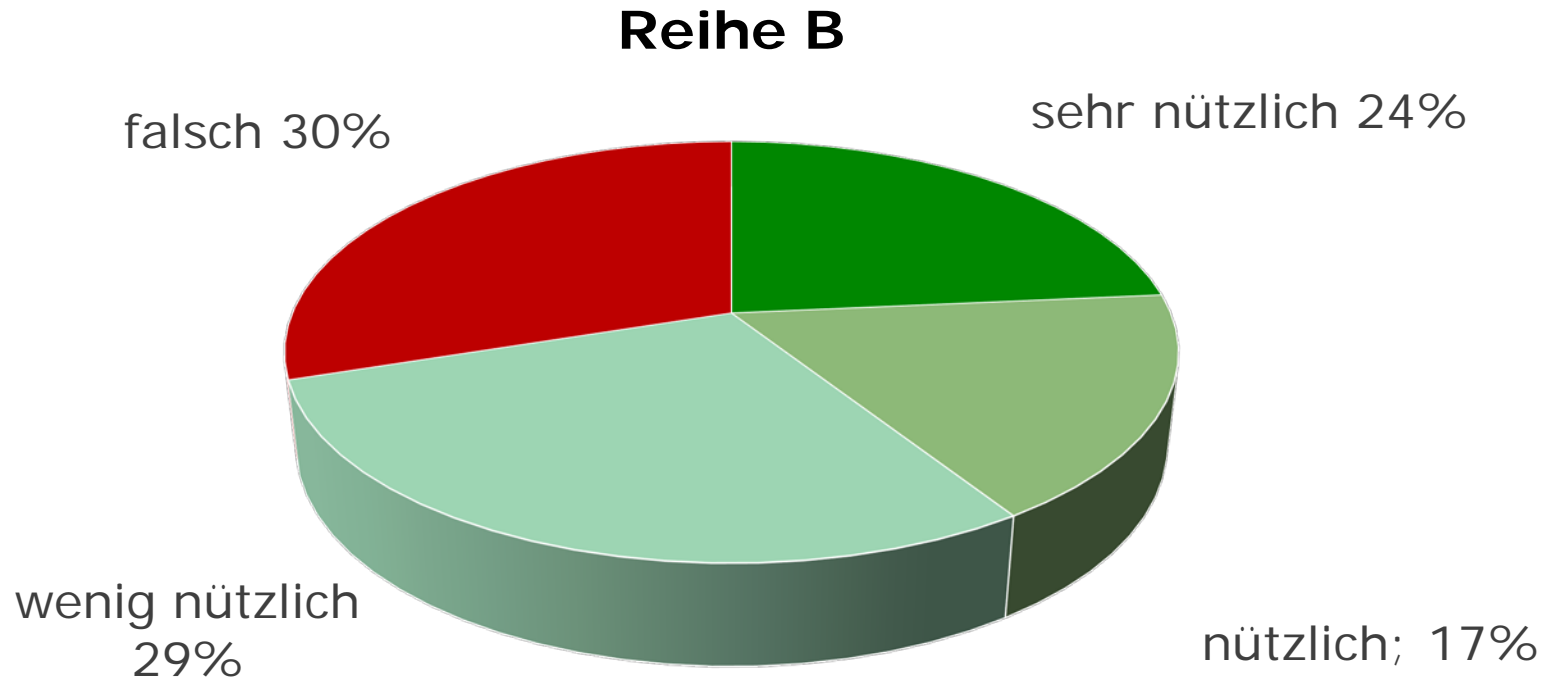
Falsch 0

Bewertungsergebnisse für maschinell vergebene Schlagwörter

Reihe H



Bewertungsergebnisse für maschinell vergebene Schlagwörter



Maschinelle Beschlagwortung der Reihe B und H

- Funktioniert für Reihe H relativ gut:
 - Kapitelüberschriften in den Inhaltsverzeichnissen mit inhaltsrelevanten Begriffen
- Maschinelle Beschlagwortung der Reihe H wird weitergeführt
- Bei Reihe B relativ hoher Anteil falscher Schlagwörter:
 - Publikationen überwiegend aus den Sozialwissenschaften, enthalten viele Individualnamen
- Ab 1. Juli 2019 Einstellung der maschinellen Beschlagwortung für Reihe B geplant

Parallelabgleich - Datenübernahme

Seit 2011:

- Automatische Verknüpfung von parallelen Print- und Onlineausgaben von Monografien
- Maschinelle Übernahme von intellektuell erstellten Inhaltserschließungsdaten und Normdatenverknüpfungen in die parallele Online-Publikation
- 25% der Online-Monografien und 10% der Print-Monografien haben eine Verknüpfung

Parallelabgleich - Datenübernahme

- Seit Februar 2019 Übertragung
 - maschinell vergebener Schlagwörter
 - maschinell vergebener DDC-Kurznotationenaus den Metadaten der Online-Publikation in die Metadaten der parallelen Print-Publikation
- Anreicherung der Print-Publikationen nur bei deutsch- und englischsprachigen Hochschulschriften, die als Verlagsausgabe erscheinen
- Übernommene Daten sind **zusätzlich** zu intellektueller Inhaltserschließung

Parallelabgleich - Datenübernahme

Beispiel Print-Publikation

- 4000 Rechtsrahmen und Regulierung privater Militär- und Sicherheitsunternehmen unter besonderer Berücksichtigung ihrer Einsatzmöglichkeiten in UN-Friedensmissionen / Tina Linti
- 4030 Berlin ; Bern ; Wien : Peter Lang
- 4243 Erscheint auch als Online-Ausgabe !1173656820!--
- 4700 |PE|*Parallelverknüpfung wurde automatisch erstellt
-
- 5050 **340\$Ei\$Hie-sg\$D2019-02-14**
- 5100 !043238637!**Friedensmission** [Ts1]
- 5101 !042011906!**Regulierung** [Ts1]
- 5102 !994812930!**Privates Militärunternehmen** [Ts1]
- 5103 !041127420!**Bewachungsgewerbe** [Ts1]
- 5104 !040636933!**Völkerrecht** [Ts1]

Aus der parallelen Online-Publikation übernommen:

- 5540 [GND]!!043238637!**Friedensmission** [Ts1]\$Ep\$HaeP-gnd\$K0,79204\$D2019-02-14
- 5540 [GND]!!042011906!**Regulierung** [Ts1]\$Ep\$HaeP-gnd\$K0,04665\$D2019-02-14
- 5540 [GND]!!040393054!**Militär** [Ts1]\$Ep\$HaeP-gnd\$K0,03712\$D2019-02-14
- 5540 [GND]!!041347919!**Privater** [Ts1]\$Ep\$HaeP-gnd\$K0,02774\$D2019-02-14

Anzeige in MARC-21

Schlagwort

650 7\$**81**\p\$(DE-588)4323863-4\$0<http://d-nb.info/gnd/4323863-4>\$(DE-101)043238637\$a**Friedensmission**\$2gnd

883 0 \$**81**\p\$a**maschinell gebildet**\$c0,79204\$d20190214\$qDE-101

DDC-Kurznotation

082 74\$**81**\p\$a**006.69**\$qDE-101\$223kdnb

883 0 \$**81**\p\$a**maschinell gebildet**\$c0,96497\$d20190313\$qDE-101

Anzeige in MARC-21

- Auslieferung der MARC21-Felder für maschinell gebildete Schlagwörter und DDC-Kurznotationen erfolgt für Netzpublikationen bereits
- **Neu:** Diese Felder können seit Anfang Februar 2019 auch in gedruckten Publikationen enthalten sein

Übergangsszenario Intellektuelle Inhaltsererschließung

- Ab **1.7.2019** Übergangsszenario für 3-5 Jahre
- Erarbeitet von der DNB und der Initiativgruppe
Inhaltsererschließung des Standardisierungsausschusses
- Orientiert sich am Bedarf buchaffiner Fächer
- Festlegung der Liste buchaffiner Fächer durch die DNB in
Abstimmung mit der Initiativgruppe

Übergangsszenario Intellektuelle Inhaltsererschließung

- Die DNB garantiert für diese buchaffinen Fächer die intellektuelle Inhaltsererschließung gedruckter Monografien der Reihe A, sofern es sich um wissenschaftliche Literatur handelt
- Bei nicht-buchaffinen Sachgruppen weiterhin intellektuelle Inhaltsererschließung von Monografien, die z. B. die jeweilige Geschichte des Fachs zum Thema haben

Vielen Dank
für Ihre Aufmerksamkeit!

Ulrike Junger
Fachbereich Erwerbung und Erschließung
Deutsche Nationalbibliothek
Telefon: +49-69-1525-1500
E-Mail: u.junger@dnb.de
www.dnb.de