

Länder- und Institutionenkodierung Statusbericht

10.11.2021

Universität Bielefeld
AG Bibliometrie
PD Dr. Niels Taubert
Christopher Lenke
Postfach 10 01 31
33501 Bielefeld

E-mail:
niels.taubert@uni-bielefeld.de
christopher.lenke@uni-bielefeld.de

Länder- und Institutionenkodierung

Die Institutionenkodierung wird vom I²SOS (Universität Bielefeld) durchgeführt und liefert eine Zuordnung von deutschen Adressdatensätzen aus den vom Kompetenzzentrum Bibliometrie lizenzierten Datenbanken des Web of Science (WoS) und Scopus zu real existierenden Forschungseinrichtungen unter Berücksichtigung von strukturellen Veränderungen der Institutionenlandschaft über die Zeit in aktuell jeweils zwei Varianten.

Dabei werden entsprechende Ergebnistabellen für die verschiedenen Datenbanken (WoS, Scopus) erstellt und in die zugehörigen Bibliometriedatenbanken integriert. Eine Dokumentation und die Tabellenstruktur der jeweils zur Verfügung gestellten Ergebnistabellen sind im Wiki auf der Seite der Institutionenkodierung verfügbar.

Die Länderkodierung liefert eine Zuordnung der Inhalte der Länderfelder aus den Rohdaten (z.B. 'Germany') zu ISO-Codes (z.B. 'DEU'). Bei Erstellung einer neuen Bibliometriedatenbank wird vom FIZ jeweils eine Liste mit Länderbezeichnungen übermittelt, die über die bisherige Kodierung nicht zugeordnet werden können. Anhand dieser Liste wird die Länderkodierung entsprechend aktualisiert und anschließend vom FIZ in die Bibliometriedatenbank übernommen.

Erstellte Versionen der Institutionenkodierung

Es wurden/werden seit dem letzten Statusbericht folgende Versionen von Institutionenkodierungen für die genannten Datenbestände/Datenbanken durchgeführt:

Web of Science (WoS):

- Fortschreibung und Stand der WoS-Rohdatenbank ist die 37. KW 2021. Aktualisierung mit dem Suffix ‚R_XML‘ erfolgte am 21.07.2021
- Die erste Version der Institutionenkodierung für die Bibliometriedatenbank WOS_B_2021 erfolgte am 09.08.2021.

Scopus:

- Die erste Version der Institutionenkodierung für die Bibliometriedatenbank SCOPUS_B_2021 erfolgte am 19.08.2021.

In jeder bereitgestellten Version ist der jeweils aktuelle Stand der laufend vorgenommenen Arbeiten an den Basisdaten eingegangen (Aufnahme von weiteren Einheiten, Erfassung weiterer Zuordnungsregeln, Korrekturen).

Statistik: Änderungen der Institutionenkodierung

Veränderungen der WoS RDB-Kodierung:

	Kodierung 23.12.2020	Kodierung 21.07.2021
Anzahl distinkte Adressen	2.448.512	2.511.303
Anzahl distinkte Adressen, zugeordnet	2.107.932	2.162.342
Anzahl distinkte Adress-Dokument-Kombinationen	7.414.416	7.636.321
Anzahl distinkte Adress-Dokument-Kombinationen, zugeordnet	6.896.786	7.105.295
Anzahl distinkte Adress-Dokument-Kombinationen 2016- dato	1.700.614	1.921.812
Anzahl distinkte Adress-Dokument-Kombinationen, zugeordnet 2016- dato	1.559.293	1.764.742
Anzahl distinkte Adress-Dokument-Kombinationen 2006-2015	2.924.779	2.925.333
Anzahl distinkte Adress-Dokument-Kombinationen, zugeordnet 2006-2015	2.746.794	2.747.821
Anzahl distinkte Adress-Dokument-Kombinationen 1996-2005	1.704.758	1.704.687
Anzahl distinkte Adress-Dokument-Kombinationen, zugeordnet 1996-2005	1.603.965	1.604.421
Anzahl distinkte Adress-Dokument-Kombinationen 1986-1995	734.907	734.929
Anzahl distinkte Adress-Dokument-Kombinationen, zugeordnet	674.065	675.020
Anzahl distinkte Dokumente	3.997.628	4.097.435
Anzahl distinkte Dokumente mit mindestens einer Zuordnung	3.786.686	3.883.487

Veränderungen der WoS BDB-Kodierung:

	Kodierung 06.04.2021	Kodierung 09.08.2021
Anzahl distinkte Adressen	2.377.698	2.498.241
Anzahl distinkte Adressen, zugeordnet	2.049.695	2.151.318
Anzahl distinkte Adress-Dokument-Kombinationen	7.168.930	7.590.629
Anzahl distinkte Adress-Dokument-Kombinationen, zugeordnet	6.677.561	7.063.434
Anzahl distinkte Adress-Dokument-Kombinationen 2016- dato	1.455.770	1.876.357
Anzahl distinkte Adress-Dokument-Kombinationen, zugeordnet 2016- dato	1.338.468	1.723.086
Anzahl distinkte Adress-Dokument-Kombinationen 2006-2015	2.924.115	2.925.303
Anzahl distinkte Adress-Dokument-Kombinationen,	2.746.490	2.747.809

zugeordnet 2006-2015		
Anzahl distinkte Adress-Dokument-Kombinationen 1996-2005	1.704.800	1.704.687
Anzahl distinkte Adress-Dokument-Kombinationen, zugeordnet 1996-2005	1.604.526	1.604.420
Anzahl distinkte Adress-Dokument-Kombinationen 1986-1995	734.887	734.923
Anzahl distinkte Adress-Dokument-Kombinationen, zugeordnet	674.975	675.014
Anzahl distinkte Dokumente	3.884.486	4.076.963
Anzahl distinkte Dokumente mit mindestens einer Zuordnung	3.680.679	3.863.908

Veränderungen der Scopus BDB-Kodierung:

	Kodierung 18.03.2021	Kodierung 19.08.2021
Anzahl distinkte Adressen	3.259.935	3.344.110
Anzahl distinkte Adressen, zugeordnet	2.740.055	2.806.437
Anzahl distinkte Adress-Dokument-Kombinationen	5.800.180	6.061.346
Anzahl distinkte Adress-Dokument-Kombinationen, zugeordnet	5.097.642	5.326.738
Anzahl distinkte Adress-Dokument-Kombinationen 2016- dato	1.496.669	1.855.312
Anzahl distinkte Adress-Dokument-Kombinationen, zugeordnet 2016- dato	1.321.181	1.637.337
Anzahl distinkte Adress-Dokument-Kombinationen 2006-2015	2.670.566	2.612.826
Anzahl distinkte Adress-Dokument-Kombinationen, zugeordnet 2006-2015	2.361.540	2.308.955
Anzahl distinkte Adress-Dokument-Kombinationen 1996-2005	1.632.270	1.593.208
Anzahl distinkte Adress-Dokument-Kombinationen, zugeordnet 1996-2005	1.414.871	1.380.446
Anzahl distinkte Adress-Dokument-Kombinationen 1986-1995	41	0
Anzahl distinkte Adress-Dokument-Kombinationen, zugeordnet	31	0
Anzahl distinkte Dokumente	3.325.542	3.488.491
Anzahl distinkte Dokumente mit mindestens einer Zuordnung	3.020.942	3.169.666

Neu aufgenommene Institutionen (für alle Kodierungen):

- GeneWerk GmbH
- Hopp-Kindertumorzentrum Heidelberg (KITZ)
- Center for Systems Biology Dresden
- Leipzig Heart Institute GmbH
- Herzzentrum Leipzig