

CORPUS DER ENTSCHEIDUNGEN DES
BUNDESVERWALTUNGSGERICHTS
(CE-BVerwG)

CODEBOOK

Version 2020-06-23

DOI: [10.5281/zenodo.3911068](https://doi.org/10.5281/zenodo.3911068)

Titel	Corpus der Entscheidungen des Bundesverwaltungsgerichts
Abkürzung	CE-BVerwG
Autor	Sean Fobbe
Version	2020-06-23
Download	https://doi.org/10.5281/zenodo.3911068
Lizenz	CC0 1.0 Universal

Zitiervorschlag

Zitieren Sie diesen Datensatz bitte, wenn Sie ihn in einem akademischen Kontext benutzen. Beispielsweise mit diesem Zitiervorschlag:

Fobbe, Sean, Corpus der Entscheidungen des Bundesverwaltungsgerichts (CE-BVerwG), Version 2020-06-23, Zenodo, DOI: <https://doi.org/10.5281/zenodo.3911068>

Urheberrecht

Der Datensatz und dieses Dokument sind unter einer **Creative Commons CC0 1.0 Universal (CC0 1.0) Public Domain Dedication** Lizenz veröffentlicht. Ich stelle den Datensatz und das Codebook vollständig gemeinfrei und verzichte weltweit auf alle damit verbundenen Urheberrechte, einschließlich aller ähnlichen Rechte, soweit dies gesetzlich möglich ist.

Sie können die Werke kopieren, modifizieren, verteilen und aufführen ohne um Erlaubnis bitten zu müssen, selbst für kommerzielle Zwecke. Patente und Markenrechte bleiben von CC0 unberührt. CC0 hat auch keine Auswirkungen auf etwaige Datenschutz- oder Persönlichkeitsrechte. Jegliche Haftung für die Benutzung dieses Werkes ist ausgeschlossen, bis zu dem maximalen Umfang in dem dies gesetzlich möglich ist.

Wenn Sie diese Werke nutzen oder zitieren sollten Sie nicht den Eindruck erwecken, der Autor unterstütze ihre Nutzung.

Dies ist nur eine unverbindliche deutsche Zusammenfassung der Lizenz, den vollständigen und rechtsverbindlichen Lizenztext finden Sie hier:

<https://creativecommons.org/publicdomain/zero/1.0/legalcode>

Inhaltsverzeichnis

Inhaltsverzeichnis	3
1 Einführung	4
2 Datenquelle	4
3 Hinweis zum Urheberrecht an den Rohdaten	5
4 Metadaten	5
4.1 Schema	5
4.2 Beispiel	5
4.3 Nutzung mit R	6
4.4 Details zu den Variablen	6
4.5 Achtung: Abweichende Codierung von Wehrdienstsenaten	7
5 Verfahrensarten vor dem BVerwG	8
6 Inhalt	9
6.1 Nach Senat	9
6.2 Nach Verfahrensart	9
6.3 Nach Jahr	10
7 Disclaimer	11

1 Einführung

Die quantitative Analyse von juristischen Texten ist in den deutschen Rechtswissenschaften ein noch junges und kaum bearbeitetes Feld. Zu einem nicht unerheblichen Teil liegt dies auch daran, dass die Anzahl an frei nutzbaren Datensätzen außerordentlich gering ist.

Die meisten hochwertigen Datensätze lagern (fast) unerreichbar in kommerziellen Datenbanken und sind wissenschaftlich gar nicht oder nur gegen Entgelt zu nutzen. Frei verfügbare Datenbanken wie *Opinio Iuris*¹ und *openJur*² verbieten ausdrücklich das maschinelle Auslesen der Rohdaten.³ Wissenschaftliche Initiativen wie der Juristische Referenzkorpus (JuReKo) sind nach jahrelanger Arbeit hinter verschlossenen Türen verschwunden.

In einem funktionierenden Rechtsstaat muss die Rechtsprechung öffentlich, transparent und nachvollziehbar sein. Im 21. Jahrhundert bedeutet das, dass sie auch quantitativen Analysen zugänglich sein muss. Der Erstellung und Aufbereitung des Datensatzes liegen daher die Gedanken der Transparenz und wissenschaftlichen Reproduzierbarkeit zugrunde.

Zusätzlich zu den einfach maschinenlesbaren Formaten (TXT und CSV) sind die PDF-Rohdaten enthalten, damit Analysten gegebenenfalls ihre eigene Konvertierung vornehmen können. Die PDF-Rohdaten wurden inhaltlich nicht verändert und die Dateinamen nur geringfügig angepasst um die Lesbarkeit für Mensch und Maschine zu verbessern. Die TXT Dateien wurden mit **pdftotext 0.73.0** erstellt. Diese wurden mit **readtext 0.7.6** in **R 3.6.3** eingelesen und das CSV file mit dem Standard-Befehl **write.csv()** erstellt.

2 Datenquelle

Die Datenquelle ist: <https://www.bverwg.de/>

Der Datensatz enthält alle Entscheidungen die zum jeweiligen Stichtag (=Versionsnummer) auf der offiziellen Webseite des Bundesverwaltungsgerichts unter der Rubrik »Rechtsprechung | Urteile & Beschlüsse« abrufbar waren. Die Daten wurden unter Beachtung des Robot Exclusion Standard (RES) gesammelt.

Nutzer sollten zwei wichtige Einschränkungen beachten:

1. Der Datensatz enthält nur das, was das Gericht auch tatsächlich veröffentlicht (*publication bias*)
2. Erst ab dem Jahr 2002 sind die Daten einigermaßen umfangreich, auch wenn frühere Entscheidungen vereinzelt enthalten sind

¹ <https://opiniojuris.de/>

² <https://openjur.de/>

³ Openjur beabsichtigt eine API anzubieten, diese war aber im Juni 2020 immernoch nicht verfügbar. Openjur ist seit 2008 in Betrieb.

3 Hinweis zum Urheberrecht an den Rohdaten

An den Entscheidungstexten und amtlichen Leitsätzen besteht gem. § 5 I UrhG kein Urheberrecht, da sie amtliche Werke sind. Alle eigenen Beiträge (z.B. durch Zusammenstellung und Anpassung der Metadaten) stelle ich gemäß einer *CC0 1.0 Universal Public Domain Lizenz* vollständig urheberrechtsfrei.

4 Metadaten

Die Metadaten in den Dateinamen sind größtenteils unverändert von den Original-Dateien des Bundesverwaltungsgerichts übernommen. Hinzugefügt wurden von mir der Gerichtsname, sowie Unter- und Trennstriche um die Maschinenlesbarkeit zu erleichtern.

Einige Urteile enthalten das Suffix »D«, welches Verzögerungsrügen zum jeweiligen Aktenzeichen anzeigt. Die weit überwiegende Zahl der Urteile wurde daher mit »NA« markiert um eine weitere stimmige machinenlesbare Variable zu schaffen. Ein einziges Urteil enthält das Suffix »S«, dessen Herkunft unklar ist, wahrscheinlich aber aus dem Aktenzeichen des Vorverfahrens herrührt. Ich habe es daher entfernt.

Die Dateinamen enthalten Gerichtsname, Datum (nach ISO-8601, d.h. YYYY-MM-DD) und in der Regel das offizielle Aktenzeichen. Wichtige Ausnahme: abweichende Codierung der Wehrdienstsenate (siehe unten).

Die Typen der Variablen wurden mit regular expressions strikt validiert.

4.1 Schema

```
[Gericht]_[Datum]_[Entscheidungsart]_[Senat]_[Verfahrensart]  
_[Ordinalzahl]_[Eingangsjahr]_[Suffix]
```

4.2 Beispiel

```
BVerwG_2016-11-21_B_5_C_21_16_D.txt
```

4.3 Nutzung mit R

Die **TEXT** files inklusive Metadaten können zum Beispiel mit **R** und dem package **readtext** (auf CRAN verfügbar) folgendermaßen eingelesen werden:

```
txt.bverwg <- readtext("./CE-BVerwG/*.txt",
                      docvarsfrom = "filenames",
                      docvarnames = c("Gericht",
                                      "Datum",
                                      "Entscheidungsart",
                                      "Senat",
                                      "Verfahrensart",
                                      "Ordinalzahl",
                                      "Eingangsjahr",
                                      "Suffix"),
                      dvsep = "_",
                      encoding = "UTF-8")
```

Das **CSV** file ist mit **R** folgendermaßen einlesbar:

```
csv.bverwg <- read.csv("./Dateiname.csv")
```

Selbstverständlich lassen sich die Daten auch mit anderen Programmiersprachen (z.B. Python) und Programmen nutzen.

4.4 Details zu den Variablen

Variable	Typ	Erläuterung
Gericht	Alphabetisch	Es ist nur der Wert »BVerwG« vergeben. Dies ist der ECLI-Code für »Bundesverwaltungsgericht«. Diese Variable dient vor allem zur einfachen und transparenten Verbindung der Daten mit anderen Datensätzen.
Datum	Datum	Das Datum der Entscheidung im Format YYYY-MM-DD (ISO-8601).
Entscheidungsart	Alphabetisch	Die Art der Entscheidung: »B« für Beschluss (ohne mündliche Verhandlung), »U« für Urteil (nach mündlicher Verhandlung) oder »G« für Gerichtsbescheid.

Senat	Natürliche Zahl	Die Werte 1 bis 10 stehen für die Nummern der regulären Revisionsenate. Der Wert 20 steht für den Fachsenat. Die Werte 21 und 22 stehen für den 1. und 2. Wehrdienstsenat. Die Wehrdienstsenate sind abweichend vom normalen Aktenzeichen codiert, bitte beachten Sie den Hinweis unten.
Verfahrensart	Alphabetisch	Die Abkürzung für die jeweilige Verfahrensart in der die Entscheidung ergangen ist. Eine Erläuterung der Abkürzungen findet sich weiter unten. Manche Verfahrensarten enthalten einen Bindestrich. Als Teil des Aktenzeichens wird diese Variable »Registerzeichen« genannt.
Ordinalzahl	Natürliche Zahl	Verfahren des gleichen Eingangsjahres erhalten vom Gericht eine fortlaufene Nummer (Ordinalzahl) in der Reihenfolge ihres Eingangs, um sie voneinander abzugrenzen.
Eingangsjahr	Natürliche Zahl	Das Jahr in dem das Verfahren beim Bundesverwaltungsgericht anhängig wurde. Das Format ist eine zweistellige Jahreszahl (YY).
Suffix	Alphabetisch	Die meisten Urteile enthalten kein Suffix und sind mit »NA« (fehlender Wert) markiert. Entschädigungsverfahren wegen überlanger Verfahrensdauer erhalten den Suffix »D«.

4.5 Achtung: Abweichende Codierung von Wehrdienstsenaten

Die Wehrdienstsenate haben in den offiziellen Aktenzeichen die Senatsnummern 1 und 2. Erst im Zusammenspiel mit der Verfahrensart wird klar, dass es sich hier nicht um die regulären Revisionsenate handelt. Ich habe daher die Wehrdienstsenate als »21« (1. Wehrdienstsenat) und »22« (2. Wehrdienstsenat) codiert um statistische Analysen der Variable »Senat« zu vereinfachen. Um das ursprüngliche Aktenzeichen zu rekonstruieren muss der Wert 21 auf 1 und der Wert 22 auf 2 zurückgesetzt werden.

5 Verfahrensarten vor dem BVerwG

Abkürzung	Verfahrensart
AR	Vorverfahren oder sonstige Verfahrensarten
A	Erstinstanzliche Klage (inklusive Wiederaufnahmeverfahren), Verwaltungsstreitsache
AV	Allgemeine Verfahren
B	Nichtzulassungsbeschwerde oder Beschwerde, Verwaltungs- streitsache
BN	Nichtzulassungsbeschwerde, Normenkontrollsachen
C	Revision, Verwaltungsstreitsache
CN	Revision, Normenkontrollsache
D	Berufung, Disziplinarsachen (nur bis 1. Juni 2015)
D-PKH	Prozesskostenhilfe, Disziplinarsachen (nur bis 1. Juni 2015)
DB	Beschwerde, Disziplinarsachen (nur bis 1. Juni 2015)
DW	Wiederaufnahme, Disziplinarsachen (nur bis 1. Juni 2015)
F	Fachsenat, Verwaltungsstreitsachen
KSt	Kostensachen
P	Rechtsbeschwerde, Personalvertretung/Richtervertretung
PB	Nichtzulassungsbeschwerde, Personalvertretung/Richterver- tretung
PKH	Prozesskostenhilfe
VR	Vorläufiger Rechtsschutz
WB	Verfahren nach Wehrbeschwerdeordnung
WD	Berufung, Wehrdisziplinarsachen
WDB	Rechtsmittel in Verfahren nach Wehrdisziplinarordnung
WDS-AV	Wehrdienstsenat, Allgemeine Verfahren
WDS-KSt	Wehrdienstsenat, Kostensachen
WDS-PKH	Wehrdienstsenat, Prozesskostenhilfe
WDS-VR	Wehrdienstsenat, Vorläufiger Rechtsschutz
WDW	Wiederaufnahme, Wehrdisziplinarsachen
WNB	Nichtzulassungsbeschwerde nach Wehrbeschwerdeordnung
WRB	Rechtsbeschwerde nach Wehrbeschwerdeordnung

6 Inhalt

6.1 Nach Senat

Senat	Entscheidungen	% Gesamt	% Kumulativ
1	2553	10.54	10.54
2	2795	11.54	22.08
3	2452	10.12	32.2
4	2779	11.47	43.67
5	2623	10.83	54.5
6	2440	10.07	64.57
7	1786	7.37	71.95
8	2120	8.75	80.7
9	2094	8.64	89.34
10	993	4.1	93.44
20	278	1.15	94.59
21	859	3.55	98.13
22	452	1.87	100
<NA>	0	0	100
Total	24224	100	100

6.2 Nach Verfahrensart

Verfahrensart	Entscheidungen	% Gesamt	% Kumulativ
A	1246	5.14	5.14
AV	148	0.61	5.75
B	13583	56.07	61.83
BN	1133	4.68	66.5
C	4401	18.17	84.67
CN	160	0.66	85.33
D	98	0.4	85.74
D-PKH	1	0	85.74
DB	53	0.22	85.96
DW	3	0.01	85.97
F	273	1.13	87.1
KSt	167	0.69	87.79
P	233	0.96	88.75
PB	366	1.51	90.26
PKH	570	2.35	92.61
VR	478	1.97	94.59

WA	2	0.01	94.6
WB	647	2.67	97.27
WD	328	1.35	98.62
WDB	73	0.3	98.92
WDS-AV	4	0.02	98.94
WDS-KSt	4	0.02	98.96
WDS-PKH	1	0	98.96
WDS-VR	130	0.54	99.5
WDW	1	0	99.5
WNB	96	0.4	99.9
WRB	25	0.1	100
<NA>	0	0	100
Total	24224	100	100

6.3 Nach Jahr

Jahr	Entscheidungen	% Gesamt	% Kumulativ
1997	1	0	0
2000	1	0	0.01
2002	1502	6.2	6.21
2003	2009	8.29	14.5
2004	1813	7.48	21.99
2005	1619	6.68	28.67
2006	1685	6.96	35.63
2007	1688	6.97	42.59
2008	1544	6.37	48.97
2009	1350	5.57	54.54
2010	1474	6.08	60.63
2011	1552	6.41	67.03
2012	1276	5.27	72.3
2013	1070	4.42	76.72
2014	887	3.66	80.38
2015	845	3.49	83.87
2016	995	4.11	87.97
2017	960	3.96	91.94
2018	835	3.45	95.38
2019	885	3.65	99.04
2020	233	0.96	100
<NA>	0	0	100
Total	24224	100	100

7 Disclaimer

Dieser Datensatz ist eine private wissenschaftliche Initiative und steht weder mit dem Bundesverwaltungsgericht noch mit den Herausgebern der BVerwGE in Verbindung.