

CORPUS DER
AMTLICHEN ENTSCHEIDUNGSSAMMLUNG
DES
BUNDESVERFASSUNGSGERICHTS
(C-BVerfGE)

CODEBOOK

Version 1.1.0



DOI: [10.5281/zenodo.3831111](https://doi.org/10.5281/zenodo.3831111)

Titel	Corpus der amtlichen Entscheidungssammlung des Bundesverfassungsgerichts
Abkürzung	C-BVerfGE
Autor	Sean Fobbe
Version	1.1.0
Download	https://doi.org/10.5281/zenodo.3831111
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 amtlichen Entscheidungssammlung des Bundesverfassungsgerichts (C-BVerfGE), Version 1.1.0, Zenodo, DOI: 10.5281/zenodo.3831111

Digital Object Identifiers: Concept DOI und Version DOI

Die in diesem Codebook verwendete DOI ist eine »concept DOI« und bezieht sich auf das Gesamtkonzept des Datensatzes. Sie verlinkt daher immer automatisch die aktuellste Version. Für jede spezifische Version steht eine weitere »version DOI« zur Verfügung, die auf der Zenodo-Seite der jeweiligen Version abgerufen werden kann. Diese Seiten sind auch über Links von der aktuellsten Version erreichbar.

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 Markenschutzrechte 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	7
5 Registerzeichen am Bundesverfassungsgericht	8
6 Inhalt	9
6.1 Nach Spruchkörper	9
6.2 Nach Registerzeichen	9
6.3 Nach Jahr	10
7 Disclaimer	11
8 Changelog	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 0.76** in **R 3.6.3** eingelesen und das CSV file mit dem Standard-Befehl **write.csv()** erstellt.

2 Datenquelle

Die Datenquelle ist: `www.bundesverfassungsgericht.de`

Das Corpus der amtlichen Entscheidungssammlung des Bundesverfassungsgerichts (C-BVerfGE) ist eine digitale Zusammenstellung von möglichst vielen Entscheidungen, die in der amtlichen Entscheidungssammlung des Bundesverfassungsgerichts (BVerfGE) veröffentlicht sind. Es enthält alle Entscheidungen, die auf der offiziellen Webseite des Bundesverfassungsgerichts am jeweiligen Stichtag in der Auflistung der Entscheidungen der BVerfGE verlinkt waren. Die Stichtage für jede Version sind im Changelog unter Punkt 8 festgehalten.

Die Daten wurden unter Beachtung des Robot Exclusion Standard (RES) gesammelt. Die Entscheidungen sind laut dem Gericht anonymisiert, aber ungekürzt.

¹ <https://opinioiusuris.de/>

² <https://openjur.de/>

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

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 1998 sind die Daten einigermaßen umfangreich, auch wenn frühere Entscheidungen vereinzelt enthalten sind

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. § 5 UrhG ist auf amtliche Datenbanken analog anzuwenden (BGH, Beschluss vom 28.09.2006, I ZR 261/03, »Sächsischer Ausschreibungsdienst«). 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 Hyperlinks zur jeweiligen Datei übernommen. Hinzugefügt wurden von mir nur der Gerichtsname, sowie Unter- und Trennstriche um die Maschinenlesbarkeit zu erleichtern.

Die Dateinamen enthalten Gerichtsname, Datum (nach ISO-8601, d.h. YYYY-MM-DD), das offizielle Aktenzeichen und ein Suffix.

Die Typen der Variablen wurden mit *regular expressions* strikt validiert. Die möglichen Werte der jeweiligen Variablen wurden zudem durch Frequenztabellen auf ihre Plausibilität geprüft.

4.1 Schema

[Gericht]_[Datum]_[SpruchkoerperAZ]_[Registerzeichen]_[Ordinalzahl]_[Eingangsjahr]_[Suffix]

4.2 Beispiel

BVerfG_1997-12-03_2_BvR_0882_97_NA

4.3 Nutzung mit R

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

```
txt.bverfg <- readtext("./C-BVerfGE/*.txt",
                      docvarsfrom = "filenames",
                      docvarnames = c("Gericht",
                                      "Datum",
                                      "SpruchkoerperAZ",
                                      "Registerzeichen",
                                      "Ordinalzahl",
                                      "Eingangsjahr",
                                      "Suffix"),
                      dvsep = "_",
                      encoding = "UTF-8")
```

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

```
csv.bverfg <- read.csv("./Dateiname.csv", header = TRUE)
```

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	In diesem Datensatz ist nur der Wert »BVerfG« vergeben. Dies ist der ECLI-Code für »Bundesverfassungsgericht«. 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).
SpruchkoerperAZ	Numerisch	Der im Aktenzeichen angegebene Spruchkörper. Es sind nur die Werte 1 und 2 vergeben. Die Werte stehen für den 1. oder 2. Senat des Gerichts.
Registerzeichen	Alphabetisch	Das amtliche Registerzeichen. Es gibt die Verfahrensart an, in der die Entscheidung ergangen ist. Eine Erläuterung der Abkürzungen findet sich unter Punkt 5.
Ordinalzahl	Natürliche Zahl	Verfahren des gleichen Eingangsjahres erhalten vom Gericht eine Ordinalzahl in der Reihenfolge ihres Eingangs, um sie voneinander abzugrenzen. Die Zahl ist in den Dateinamen mit führenden Nullen (falls <1000) codiert.
Eingangsjahr	Natürliche Zahl	Das Jahr in dem das Verfahren beim Bundesverfassungsgericht anhängig wurde. Das Format ist eine zweistellige Jahreszahl (YY).
Suffix	Alphabetisch	In wenigen Fällen sind am gleichen Tag mehrere Entscheidungen zum gleichen Aktenzeichen ergangen. Diese werden durch ein Suffix mit einem Kleinbuchstaben ausdifferenziert. Für die meisten Entscheidungen ist der Wert der Variable »NA«, also nicht vorhanden.

5 Registerzeichen am Bundesverfassungsgericht

Registerzeichen	Erläuterung
AR	Vorverfahren oder sonstige Verfahrensarten
BvA	Verwirkung von Grundrechten
BvB	Verfassungswidrigkeit von Parteien, Art. 21 II GG
BvC	Wahlprüfungsverfahren, Art. 41 II GG
BvD	Anklage des Bundespräsidenten, Art. 61 GG
BvE	Organstreitverfahren, Art. 93 I Nr. 1 GG
BvF	Abstrakte Normenkontrolle, Art. 93 I Nr. 2 GG
BvG	Bund-Länder-Streitigkeiten, Art. 93 Abs. 1 Nr. 3 oder Art. 84 Abs. 4 Satz 2 GG
BvK	Landesverfassungsstreitigkeiten, Art. 99 GG
BvL	Konkrete Normenkontrolle, Art. 100 I GG
BvM	Feststellung der Anwendbarkeit einer Regel des Völkerge- wohnheitsrechts, Art. 100 II GG iVm Art 25 GG
BvN	Divergenzvorlagen eines Landesverfassungsgerichts zur Ausle- gung des Grundgesetzes, Art. 100 III GG
BvO	Fortgeltung vorkonstitutionellen Rechts als Bundesrecht, Art. 126 GG
BvP	Sonstige durch Bundesrecht zugewiesene Verfahren, Art. 93 III GG
BvQ	Einstweilige Anordnungen, § 32 BVerfGG
BvR	Verfassungsbeschwerden, Art. 93 Abs. 1 Nr. 4a, und Kommu- nalverfassungsbeschwerden, Art. 93 Abs. 1 Nr. 4b GG
BvT	Sonstige Verfahren
PBvS	Beendigung des Richteramtes am Bundesverfassungsgericht, § 105 BVerfGG
PBvU	Plenarentscheidungen, § 16 BVerfGG
PBvV	Rechtsgutachten, § 97 BVerfGG a.F.
PKH	Prozesskostenhilfe
Vz	Verzögerungsrügen

6 Inhalt

6.1 Nach Spruchkörper

Senat	Entscheidungen	% Gesamt	% Kumulativ
1	310	45.72	45.72
2	368	54.28	100
<NA>	0	0	100
Total	678	100	100

6.2 Nach Registerzeichen

Register- zeichen	Entscheidungen	% Gesamt	% Kumulativ
BvB	15	2.21	2.21
BvC	18	2.65	4.87
BvE	65	9.59	14.45
BvF	42	6.19	20.65
BvG	8	1.18	21.83
BvH	4	0.59	22.42
BvK	8	1.18	23.6
BvL	138	20.35	43.95
BvM	4	0.59	44.54
BvN	1	0.15	44.69
BvP	1	0.15	44.84
BvQ	18	2.65	47.49
BvR	355	52.36	99.85
PBvU	1	0.15	100
<NA>	0	0	100
Total	678	100	100

6.3 Nach Jahr

Jahr	Entscheidungen	% Gesamt	% Kumulativ
1952	2	0.29	0.29
1958	1	0.15	0.44
1983	2	0.29	0.74
1994	2	0.29	1.03
1996	1	0.15	1.18
1997	9	1.33	2.51
1998	51	7.52	10.03
1999	31	4.57	14.6
2000	20	2.95	17.55
2001	37	5.46	23.01
2002	23	3.39	26.4
2003	34	5.01	31.42
2004	40	5.9	37.32
2005	38	5.6	42.92
2006	25	3.69	46.61
2007	26	3.83	50.44
2008	29	4.28	54.72
2009	29	4.28	59
2010	29	4.28	63.27
2011	24	3.54	66.81
2012	35	5.16	71.98
2013	29	4.28	76.25
2014	37	5.46	81.71
2015	28	4.13	85.84
2016	29	4.28	90.12
2017	27	3.98	94.1
2018	30	4.42	98.53
2019	10	1.47	100
<NA>	0	0	100
Total	678	100	100

7 Disclaimer

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

8 Changelog

Version	Stichtag	Details
1.1.0	2020-08-09	<ul style="list-style-type: none">- Vollständige Aktualisierung der Daten- Angleichung der Variablen-Namen an andere Datensätze der CE-Serie von https://zenodo.org/communities/sean-fobbe-data/- Einführung der Variable »Suffix« um weitere Entscheidungen korrekt erfassen zu können; aufgrund der fehlenden Berücksichtigung des Suffix wurden einige wenige Entscheidungen in Version 1.0.0 irrtümlich von der Sammlung ausgeschlossen.
1.0.0	2020-05-16	Erstveröffentlichung