

# Corpus der Entscheidungen des Bundesverwaltungsgerichts (CE-BVerwG-Source)

ROBUSTNESS CHECKS

Version 2024-03-13

License MIT-0

DOI: 10.5281/zenodo.10809041

<b>Titel</b>	Source Code des »Corpus der Entscheidungen des Bundesverwaltungsgerichts«
<b>Abkürzung</b>	CE-BVerwG-Source
<b>Autor</b>	Seán Fobbe
<b>Version</b>	2024-03-13
<b>Download</b>	<a href="https://doi.org/10.5281/zenodo.10809041">https://doi.org/10.5281/zenodo.10809041</a>
<b>Lizenz</b>	MIT No Attribution (MIT-0)

### Zitiervorschlag

*Seán Fobbe* (2024). Source Code des »Corpus der Entscheidungen des Bundesverwaltungsgerichts« (CE-BVerwG-Source). Version 2024-03-13. Zenodo. DOI: 10.5281/zenodo.10809041.

### Digital Object Identifier (DOI): Concept DOI und Version DOI

Soweit nicht anders angegeben ist die DOI immer eine »Version DOI« und bezieht sich nur auf eine bestimmte Version der Software. Sie verlinkt daher nur Version 2024-03-13. Für das Gesamtkonzept der Software steht eine »Concept DOI« zur Verfügung, die auf der Zenodo-Seite jeder Version unter »Cite all versions?« zu finden ist. Sie lautet 10.5281/zenodo.4625134. Die »Concept DOI« verlinkt immer die aktuellste Version.

### Lizenz: MIT No Attribution (MIT-0)

Copyright — 2024 — Seán Fobbe

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the »Software«), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so.

THE SOFTWARE IS PROVIDED »AS IS«, WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

### Disclaimer

Dieser Datensatz ist eine private wissenschaftliche Initiative und steht in keiner Verbindung zu Behörden, Gerichten oder anderen amtlichen Stellen der Bundesrepublik Deutschland.

# Inhaltsverzeichnis

<b>1</b>	<b>Ergebnisse laden</b>	<b>4</b>
<b>2</b>	<b>Pipeline visualisieren</b>	<b>5</b>
<b>3</b>	<b>Vollständigkeit der Texte</b>	<b>7</b>
3.1	Download: Fehlende PDF-Dateien . . . . .	7
3.2	Konvertierung: Fehlende TXT-Dateien . . . . .	7
<b>4</b>	<b>Variablen</b>	<b>8</b>
4.1	Anzahl Variablen der Datensätze . . . . .	8
4.2	Alle Variablen-Namen der Datensätze . . . . .	8
4.3	Beispiel-Werte . . . . .	8
<b>5</b>	<b>Frequenztabellen erstellen</b>	<b>10</b>
5.1	Ignorierte Variablen . . . . .	10
5.2	Vorbereitung . . . . .	10
5.3	Frequenztabellen berechnen . . . . .	10
<b>6</b>	<b>Changelog</b>	<b>23</b>
6.1	Version 2024-03-13 . . . . .	23
6.2	Version 2023-03-21 . . . . .	23
6.3	Version 2022-08-07 . . . . .	23
6.4	Version 2021-10-19 . . . . .	24
6.5	Version 2021-04-15 . . . . .	24
6.6	Version 2020-06-23 . . . . .	24
<b>7</b>	<b>Parameter für strenge Replikationen</b>	<b>25</b>
	<b>Literaturverzeichnis</b>	<b>27</b>

## 1 Ergebnisse laden

```
tar_load(latexdefs)
tar_load(dt.download.final)
tar_load(dt.bverwg.full)
tar_load(files.pdf)
tar_load(files.txt)
```

## 2 Pipeline visualisieren

```
edgelist <- tar_network(targets_only = TRUE)$edges
setDT(edgelist)

g <- igraph::graph.data.frame(edgelist,
                              directed = TRUE)

ggraph(g,
       'sugiyama') +
  geom_edge_diagonal(colour = "#7e0731")+
  geom_node_point(size = 2,
                 color = "white")+
  geom_node_text(aes(label = name),
                color = "white",
                size = 2,
                repel = TRUE)+
  theme_void()+
  labs(
    title = paste(prefix.figuretitle,
                  "| Vollständiger Prozess der Datensatz-Kompilierung"),
    caption = caption
  )+
  theme(
    plot.title = element_text(size = 14,
                              face = "bold",
                              color = "white"),
    plot.background = element_rect(fill = "black"),
    plot.caption = element_text(color = "white"),
    plot.margin = margin(10, 20, 10, 10)
  )
)
```

```
## Warning: Using the `size` aesthetic in this geom was deprecated in ggplot2
3.4.0.
## i Please use `linewidth` in the `default_aes` field and elsewhere instead.
```

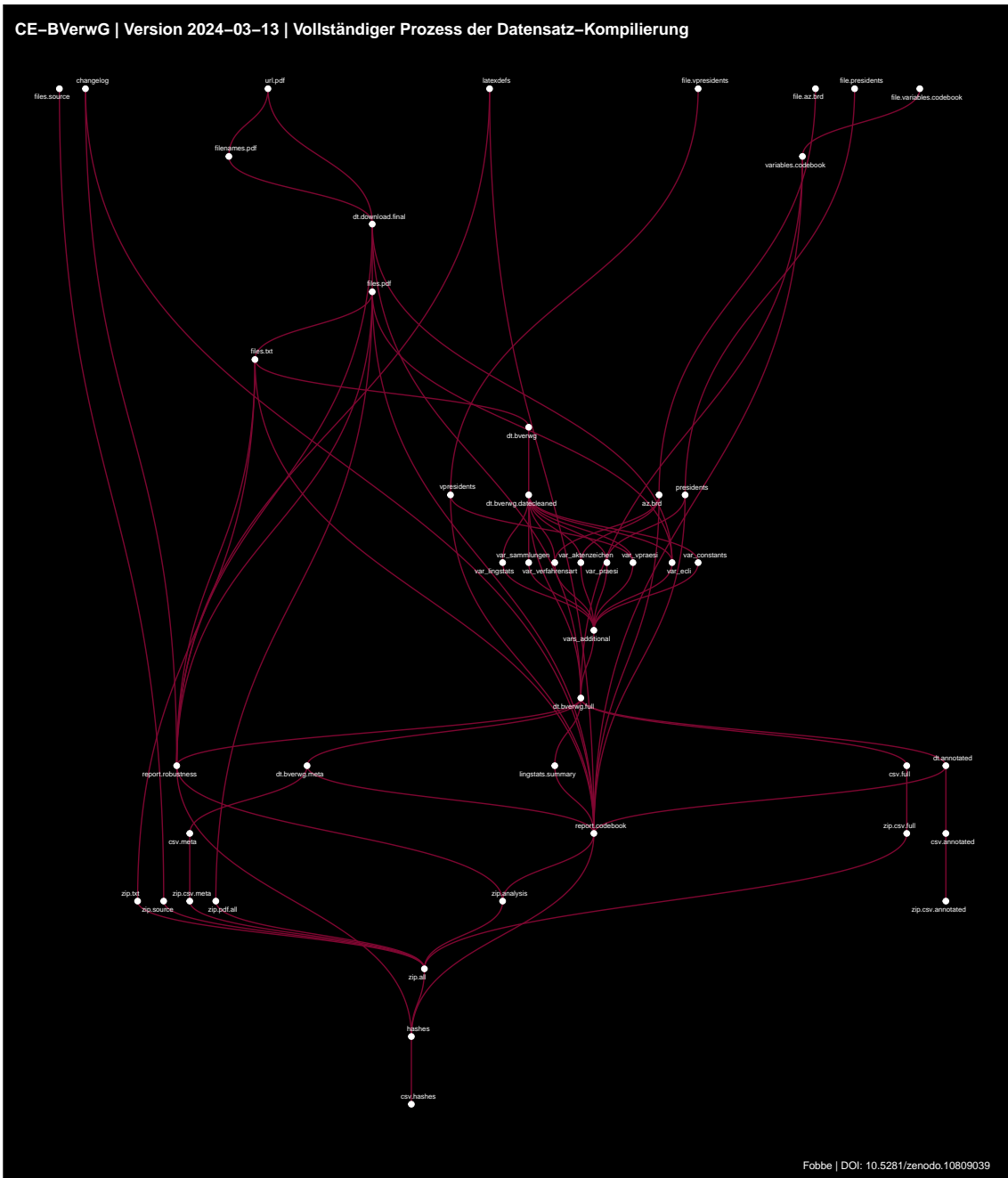


Abbildung 1: Der vollständige Prozess der Datensatz-Kompilierung.

### 3 Vollständigkeit der Texte

#### 3.1 Download: Fehlende PDF-Dateien

```
setdiff(dt.download.final$doc_id,  
        basename(files.pdf))
```

```
## [1] "BVerwG_2004-08-03_U_1_C_30_02_NA_0.pdf"  
## [2] "BVerwG_2010-06-08_B_1_WB_49_09_NA_0.pdf"
```

#### 3.2 Konvertierung: Fehlende TXT-Dateien

```
setdiff(tools::file_path_sans_ext(basename(files.pdf)),  
        tools::file_path_sans_ext(basename(files.txt)))
```

```
## character(0)
```

## 4 Variablen

### 4.1 Anzahl Variablen der Datensätze

```
length(dt.bverwg.full)
```

```
## [1] 31
```

### 4.2 Alle Variablen-Namen der Datensätze

```
names(dt.bverwg.full)
```

```
## [1] "doc_id"      "text"      "url"
## [4] "gericht"    "datum"    "entscheidung_typ"
## [7] "entscheidungs_jahr" "spruchkoerper_az" "registerzeichen"
## [10] "verfahrensart" "eingangsnummer" "eingangsjahr_az"
## [13] "eingangsjahr_iso" "verzoegerung" "kollision"
## [16] "aktenzeichen" "ecli"      "praesi"
## [19] "v_praesi"    "bverwge"  "leitsatz"
## [22] "fachpresse"  "uebersetzung" "zeichen"
## [25] "tokens"     "typen"    "saetze"
## [28] "version"    "doi_concept" "doi_version"
## [31] "lizenz"
```

### 4.3 Beispiel-Werte

```
str(dt.bverwg.full)
```

```
## Classes 'data.table' and 'data.frame': 27200 obs. of 31 variables:
## $ doc_id : chr "BVerwG_1997-02-26_U_6_C_3_96_NA_0.txt" "BVerwG
## _2000-12-14_B_1_WB_107_00_NA_0.txt" "BVerwG_2002-01-10_B_9_A_9_02_NA_0.txt" "
## BVerwG_2002-01-14_B_4_BN_1_02_NA_0.txt" ...
## $ text : chr " BUNDESVERWALTUNGSGERICHT\n\n
## IM NAMEN DES VOLKES\n\n URTEI"| __truncated__
## "
## BUNDESVERWALTUNGSGERICHT\n
## Beschluss\n\n\nBVerwG 1 WB 107.00, 1 WB "| __truncated__ "
## BUNDESVERWALTUNGSGERICHT\n\n BESCHLUSS\n\nBVerwG 9 A
## 9.02 (vormals 4 A 74.01)\n\n"| __truncated__ " BUNDESVERWALTUNGSGERICHT
## \n\n BESCHLUSS\nBVerwG 4 BN 1.02\nVGH 8 S 1853/01\n\n\n
## n "| __truncated__ ...
## $ url : chr "https://www.bverwg.de/entscheidungen/pdf/260297
## U6C3.96.0.pdf" "https://www.bverwg.de/entscheidungen/pdf/141200B1WB107.00.0.
## pdf" "https://www.bverwg.de/entscheidungen/pdf/100102B9A9.02.0.pdf" "https://
## www.bverwg.de/entscheidungen/pdf/140102B4BN1.02.0.pdf" ...
```



```

## $ gericht      : chr "BVerwG" "BVerwG" "BVerwG" "BVerwG" ...
## $ datum       : IDate, format: "1997-02-26" "2000-12-14" ...
## $ entscheidung_typ : chr "U" "B" "B" "B" ...
## $ entscheidungsjahr: int 1997 2000 2002 2002 2002 2002 2002 2002 2002 2002
...
## $ spruchkoerper_az : int 6 1 9 4 1 1 1 1 8 8 ...
## $ registerzeichen : chr "C" "WB" "A" "BN" ...
## $ verfahrensart   : chr "Revision (Verwaltungsstreitsachen)" "Verfahren
nach Wehrbeschwerdeordnung" "Erstinstanzliche Klage, inklusive
Wiederaufnahmeverfahren (Verwaltungsstreitsachen)" "Nichtzulassungsbeschwerde
(Normenkontrollsachen)" ...
## $ eingangsnummer  : int 3 107 9 1 12 3 4 5 1 2 ...
## $ eingangsjahr_az : int 96 0 2 2 2 2 2 2 2 ...
## $ eingangsjahr_iso : num 1996 2000 2002 2002 2002 ...
## $ verzoegerung    : chr NA NA NA NA ...
## $ kollision       : int 0 0 0 0 0 0 0 0 0 ...
## $ aktenzeichen    : chr "6 C 3.96" "1 WB 107.00" "9 A 9.02" "4 BN 1.02" ...
## $ ecli            : chr "ECLI:DE:BVerwG:1997:260297U6C3.96.0" "ECLI:DE:
BVerwG:2000:141200B1WB107.00.0" "ECLI:DE:BVerwG:2002:100102B9A9.02.0" "ECLI:
DE:BVerwG:2002:140102B4BN1.02.0" ...
## $ praesi         : chr "Franßen" "Franßen" "Franßen" "Franßen" ...
## $ v_praesi       : chr "Franke" "Hien" "Hien" "Hien" ...
## $ bverwge        : num 0 0 0 0 0 0 0 0 0 ...
## $ leitsatz       : num 0 0 0 0 0 0 0 0 0 ...
## $ fachpresse     : num 0 0 0 0 0 0 0 0 0 ...
## $ uebersetzung   : num 0 0 0 0 0 0 0 0 0 ...
## $ zeichen        : int 37496 13855 826 3646 1401 1395 1402 1404 2470 2760
...
## $ tokens         : int 5479 2336 128 557 213 213 213 213 377 418 ...
## $ typen          : int 1465 721 84 277 128 128 126 128 224 227 ...
## $ saetze         : int 266 143 15 48 18 18 18 18 21 27 ...
## $ version        : chr "2024-03-13" "2024-03-13" "2024-03-13" "2024-03-13"
...
## $ doi_concept    : chr "10.5281/zenodo.3911067" "10.5281/zenodo.3911067"
"10.5281/zenodo.3911067" "10.5281/zenodo.3911067" ...
## $ doi_version    : chr "10.5281/zenodo.10809039" "10.5281/zenodo.10809039"
"10.5281/zenodo.10809039" "10.5281/zenodo.10809039" ...
## $ lizenz         : chr "Creative Commons Zero 1.0 Universal" "Creative
Commons Zero 1.0 Universal" "Creative Commons Zero 1.0 Universal" "Creative
Commons Zero 1.0 Universal" ...
## - attr(*, ".internal.selfref")=<externalptr>
## - attr(*, "sorted")= chr "doc_id"

```

## 5 Frequenztabellen erstellen

### 5.1 Ignorierte Variablen

**Hinweis:** Diese Variablen werden bei der Erstellung der Frequenztabellen nicht berücksichtigt.

```
freqtable.ignore <- c("text",  
                      "eingangsnummer",  
                      "datum",  
                      "doc_id",  
                      "ecli",  
                      "aktenzeichen",  
                      "tokens",  
                      "typen",  
                      "saetze",  
                      "zeichen",  
                      "url")
```

### 5.2 Vorbereitung

```
## Liste zu prüfender Variablen  
  
varlist <- setdiff(names(dt.bverwg.full), freqtable.ignore)  
print(varlist)
```

```
## [1] "gericht"           "entscheidung_typ" "entscheidungsjahr"  
## [4] "spruchkoerper_az" "registerzeichen"  "verfahrensart"  
## [7] "eingangsjahr_az"  "eingangsjahr_iso" "verzoegerung"  
## [10] "kollision"        "praesi"           "v_praesi"  
## [13] "bverwge"          "leitsatz"         "fachpresse"  
## [16] "uebersetzung"    "version"          "doi_concept"  
## [19] "doi_version"     "lizenz"
```

```
## Präfix definieren  
  
prefix <- paste0(config$project$shortname,  
                  "_01_Frequenztafel_var-")
```

### 5.3 Frequenztabellen berechnen

```
f.fast.freqtable(dt.bverwg.full,  
                 varlist = varlist,  
                 sumrow = TRUE,  
                 output.list = FALSE,
```

```

output.kable = TRUE,
output.csv = TRUE,
outputdir = dir.analysis,
prefix = prefix,
align = c("p{5cm}",
          rep("r", 4))

```

---

Frequency Table for Variable: gericht

---

1 unique value(s) detected.

gericht	N	exactpercent	roundedpercent	cumulpercent
BVerwG	27200	100	100	100
Total	27200	100	100	100

---

Frequency Table for Variable: entscheidung\_typ

---

3 unique value(s) detected.

entscheidung_typ	N	exactpercent	roundedpercent	cumulpercent
B	22370	82.2426471	82.24	82.24
G	22	0.0808824	0.08	82.32
U	4808	17.6764706	17.68	100.00
Total	27200	100.0000000	100.00	100.00

---

Frequency Table for Variable: entscheidungsjahr

---

25 unique value(s) detected.

entscheidungsjahr	N	exactpercent	roundedpercent	cumulpercent
1997	1	0.0036765	0.00	0.00
2000	1	0.0036765	0.00	0.01
2002	1502	5.5220588	5.52	5.53
2003	2009	7.3860294	7.39	12.92

(continued)

entscheidungsjahr	N	exactpercent	roundedpercent	cumulpercent
2004	1817	6.6801471	6.68	19.60
2005	1619	5.9522059	5.95	25.55
2006	1684	6.1911765	6.19	31.74
2007	1690	6.2132353	6.21	37.95
2008	1549	5.6948529	5.69	43.65
2009	1353	4.9742647	4.97	48.62
2010	1474	5.4191176	5.42	54.04
2011	1553	5.7095588	5.71	59.75
2012	1281	4.7095588	4.71	64.46
2013	1072	3.9411765	3.94	68.40
2014	889	3.2683824	3.27	71.67
2015	845	3.1066176	3.11	74.78
2016	995	3.6580882	3.66	78.43
2017	960	3.5294118	3.53	81.96
2018	836	3.0735294	3.07	85.04
2019	887	3.2610294	3.26	88.30
2020	871	3.2022059	3.20	91.50
2021	780	2.8676471	2.87	94.37
2022	731	2.6875000	2.69	97.06
2023	747	2.7463235	2.75	99.80
2024	54	0.1985294	0.20	100.00
Total	27200	100.0000000	100.00	100.00

Frequency Table for Variable: spruchkoerper\_az

14 unique value(s) detected.

spruchkoerper_az	N	exactpercent	roundedpercent	cumulpercent
NA	1	0.0036765	0.00	0.00
1	4029	14.8125000	14.81	14.82
2	3697	13.5919118	13.59	28.41
3	2662	9.7867647	9.79	38.19
4	3173	11.6654412	11.67	49.86
5	2848	10.4705882	10.47	60.33
6	2677	9.8419118	9.84	70.17
7	1942	7.1397059	7.14	77.31
8	2402	8.8308824	8.83	86.14
9	2358	8.6691176	8.67	94.81
10	1079	3.9669118	3.97	98.78
11	10	0.0367647	0.04	98.82
20	320	1.1764706	1.18	99.99
30	2	0.0073529	0.01	100.00
Total	27200	100.0000000	100.00	100.00

---

Frequency Table for Variable: registerzeichen

---

31 unique value(s) detected.

registerzeichen	N	exactpercent	roundedpercent	cumulpercent
A	1423	5.2316176	5.23	5.23
AV	229	0.8419118	0.84	6.07
B	14727	54.1433824	54.14	60.22
BN	1397	5.1360294	5.14	65.35
C	4973	18.2830882	18.28	83.64
CN	216	0.7941176	0.79	84.43
D	99	0.3639706	0.36	84.79
D-PKH	1	0.0036765	0.00	84.80
DB	53	0.1948529	0.19	84.99

(continued)

registerzeichen	N	exactpercent	roundedpercent	cumulpercent
DW	3	0.0110294	0.01	85.00
F	313	1.1507353	1.15	86.15
GS	2	0.0073529	0.01	86.16
GrSen	1	0.0036765	0.00	86.17
KSt	206	0.7573529	0.76	86.92
P	266	0.9779412	0.98	87.90
PB	413	1.5183824	1.52	89.42
PKH	609	2.2389706	2.24	91.66
VR	559	2.0551471	2.06	93.71
WA	6	0.0220588	0.02	93.74
WB	807	2.9669118	2.97	96.70
WD	386	1.4191176	1.42	98.12
WDB	121	0.4448529	0.44	98.57
WDS-AV	4	0.0147059	0.01	98.58
WDS-KSt	4	0.0147059	0.01	98.60
WDS-PKH	1	0.0036765	0.00	98.60
WDS-VR	142	0.5220588	0.52	99.12
WDW	1	0.0036765	0.00	99.12
WKSt	2	0.0073529	0.01	99.13
WNB	148	0.5441176	0.54	99.68
WRB	39	0.1433824	0.14	99.82
WVR	49	0.1801471	0.18	100.00
Total	27200	100.0000000	100.00	100.00

Frequency Table for Variable: verfahrensart

29 unique value(s) detected.

verfahrensart	N	exactpercent	roundedpercent	cumulpercent
NA	53	0.1948529	0.19	0.19
Allgemeine Verfahren (Wehrdienstsenate)	4	0.0147059	0.01	0.21
Allgemeine Verfahren (sonstige Anträge außerhalb eines schwebenden Verfahrens)	229	0.8419118	0.84	1.05
Berufung (Wehrdisziplinarsachen)	386	1.4191176	1.42	2.47
Beschwerde (Disziplinarsachen)	53	0.1948529	0.19	2.67
Erstinstanzliche Klage, inklusive Wiederaufnahmeverfahren (Verwaltungsstreitsachen)	1423	5.2316176	5.23	7.90
Erstinstanzliche Klage, inklusive Wiederaufnahmeverfahren (Wehrdienstsenat)	6	0.0220588	0.02	7.92
Fachsenat (Verwaltungsstreitsachen)	313	1.1507353	1.15	9.07
Großer Senat des Bundesverwaltungsgerichts	1	0.0036765	0.00	9.07
Kostensachen	206	0.7573529	0.76	9.83
Kostensachen (Wehrdienstsenate)	4	0.0147059	0.01	9.85
Nichtzulassungsbeschwerde (Normenkontrollsachen)	1397	5.1360294	5.14	14.98
Nichtzulassungsbeschwerde (Personalvertretungs- und Richtervertretungssachen)	413	1.5183824	1.52	16.50
Nichtzulassungsbeschwerde nach Wehrbeschwerdeordnung	148	0.5441176	0.54	17.04
Nichtzulassungsbeschwerde oder Beschwerde (Verwaltungsstreitsachen)	14727	54.1433824	54.14	71.19
Prozesskostenhilfe	609	2.2389706	2.24	73.43
Prozesskostenhilfe (Disziplinarsachen)	1	0.0036765	0.00	73.43

(continued)

verfahrensart	N	exactpercent	roundedpercent	cumulpercent
Prozesskostenhilfe (Wehrdienstsenate)	1	0.0036765	0.00	73.43
Rechtsbeschwerde (Personalvertretungs- und Richtervertretungssachen)	266	0.9779412	0.98	74.41
Rechtsbeschwerde nach Wehrbeschwerdeordnung	39	0.1433824	0.14	74.56
Rechtsmittelverfahren über Anträge, Beschwerden und Vorlagen nach der Wehrdisziplinarordnung	121	0.4448529	0.44	75.00
Revision (Disziplinarsachen)	99	0.3639706	0.36	75.36
Revision (Normenkontrollsa- chen)	216	0.7941176	0.79	76.16
Revision (Verwaltungsstreitsa- chen)	4973	18.2830882	18.28	94.44
Verfahren nach Wehrbeschwerdeordnung	807	2.9669118	2.97	97.41
Vorläufiger Rechtsschutz	559	2.0551471	2.06	99.46
Vorläufiger Rechtsschutz (Wehrdienstsenate)	142	0.5220588	0.52	99.99
Wiederaufnahme (Disziplinarsachen)	3	0.0110294	0.01	100.00
Wiederaufnahme (Wehrdisziplinarsachen)	1	0.0036765	0.00	100.00
Total	27200	100.0000000	100.00	100.00

Frequency Table for Variable: eingangsjahr\_az

26 unique value(s) detected.

ingangsjahr_az	N	exactpercent	roundedpercent	cumulpercent
0	1	0.0036765	0.00	0.00
1	2	0.0073529	0.01	0.01



(continued)

eingangsjahr_az	N	exactpercent	roundedpercent	cumulpercent
2	2269	8.3419118	8.34	8.35
3	1948	7.1617647	7.16	15.51
4	1822	6.6985294	6.70	22.21
5	1575	5.7904412	5.79	28.00
6	1670	6.1397059	6.14	34.14
7	1794	6.5955882	6.60	40.74
8	1451	5.3345588	5.33	46.07
9	1440	5.2941176	5.29	51.37
10	1445	5.3125000	5.31	56.68
11	1493	5.4889706	5.49	62.17
12	1293	4.7536765	4.75	66.92
13	964	3.5441176	3.54	70.47
14	864	3.1764706	3.18	73.64
15	936	3.4411765	3.44	77.08
16	1020	3.7500000	3.75	80.83
17	869	3.1948529	3.19	84.03
18	836	3.0735294	3.07	87.10
19	864	3.1764706	3.18	90.28
20	808	2.9705882	2.97	93.25
21	809	2.9742647	2.97	96.22
22	646	2.3750000	2.38	98.60
23	374	1.3750000	1.38	99.97
24	6	0.0220588	0.02	100.00
96	1	0.0036765	0.00	100.00
Total	27200	100.0000000	100.00	100.00

Frequency Table for Variable: eingangsjahr\_iso

26 unique value(s) detected.

eingangsjahr_iso	N	exactpercent	roundedpercent	cumulpercent
1996	1	0.0036765	0.00	0.00
2000	1	0.0036765	0.00	0.01
2001	2	0.0073529	0.01	0.01
2002	2269	8.3419118	8.34	8.36
2003	1948	7.1617647	7.16	15.52
2004	1822	6.6985294	6.70	22.22
2005	1575	5.7904412	5.79	28.01
2006	1670	6.1397059	6.14	34.15
2007	1794	6.5955882	6.60	40.74
2008	1451	5.3345588	5.33	46.08
2009	1440	5.2941176	5.29	51.37
2010	1445	5.3125000	5.31	56.68
2011	1493	5.4889706	5.49	62.17
2012	1293	4.7536765	4.75	66.93
2013	964	3.5441176	3.54	70.47
2014	864	3.1764706	3.18	73.65
2015	936	3.4411765	3.44	77.09
2016	1020	3.7500000	3.75	80.84
2017	869	3.1948529	3.19	84.03
2018	836	3.0735294	3.07	87.11
2019	864	3.1764706	3.18	90.28
2020	808	2.9705882	2.97	93.25
2021	809	2.9742647	2.97	96.23
2022	646	2.3750000	2.38	98.60
2023	374	1.3750000	1.38	99.98
2024	6	0.0220588	0.02	100.00
Total	27200	100.0000000	100.00	100.00

Frequency Table for Variable: verzoegerung

2 unique value(s) detected.

verzoegerung	N	exactpercent	roundedpercent	cumulpercent
NA	27130	99.7426471	99.74	99.74
D	70	0.2573529	0.26	100.00
Total	27200	100.0000000	100.00	100.00

Frequency Table for Variable: kollision

2 unique value(s) detected.

kollision	N	exactpercent	roundedpercent	cumulpercent
0	27176	99.9117647	99.91	99.91
1	24	0.0882353	0.09	100.00
Total	27200	100.0000000	100.00	100.00

Frequency Table for Variable: praesi

7 unique value(s) detected.

praesi	N	exactpercent	roundedpercent	cumulpercent
Eckertz-Höfer	9296	34.176471	34.18	34.18
Franßen	1049	3.856618	3.86	38.03
Hien	8346	30.683823	30.68	68.72
Korbmacher	1073	3.944853	3.94	72.66
Rennert	6198	22.786765	22.79	95.45
VACANCY-5	397	1.459559	1.46	96.91
VACANCY-6	841	3.091912	3.09	100.00
Total	27200	100.0000000	100.00	100.00

Frequency Table for Variable: v\_praesi

10 unique value(s) detected.

v_praesi	N	exactpercent	roundedpercent	cumulpercent
Christ	3117	11.4595588	11.46	11.46
Eckertz-Höfer	8346	30.6838235	30.68	42.14
Franke	1	0.0036765	0.00	42.15
Hien	1048	3.8529412	3.85	46.00
Hund	6599	24.2610294	24.26	70.26
Korbmacher	2638	9.6985294	9.70	79.96
Rennert	1696	6.2352941	6.24	86.19
Rublack	711	2.6139706	2.61	88.81
VACANCY-3	1398	5.1397059	5.14	93.95
VACANCY-4	1646	6.0514706	6.05	100.00
Total	27200	100.0000000	100.00	100.00

Frequency Table for Variable: bverwge

2 unique value(s) detected.

bverwge	N	exactpercent	roundedpercent	cumulpercent
0	26326	96.786765	96.79	96.79
1	874	3.213235	3.21	100.00
Total	27200	100.000000	100.00	100.00

Frequency Table for Variable: leitsatz

2 unique value(s) detected.

leitsatz	N	exactpercent	roundedpercent	cumulpercent
0	25007	91.9375	91.94	91.94
1	2193	8.0625	8.06	100.00
Total	27200	100.0000	100.00	100.00

---

---

Frequency Table for Variable: fachpresse

---

---

2 unique value(s) detected.

fachpresse	N	exactpercent	roundedpercent	cumulpercent
0	25618	94.183824	94.18	94.18
1	1582	5.816177	5.82	100.00
Total	27200	100.000000	100.00	100.00

---

---

Frequency Table for Variable: uebersetzung

---

---

2 unique value(s) detected.

uebersetzung	N	exactpercent	roundedpercent	cumulpercent
0	27152	99.8235294	99.82	99.82
1	48	0.1764706	0.18	100.00
Total	27200	100.0000000	100.00	100.00

---

---

Frequency Table for Variable: version

---

---

1 unique value(s) detected.

version	N	exactpercent	roundedpercent	cumulpercent
2024-03-13	27200	100	100	100
Total	27200	100	100	100

---

---

Frequency Table for Variable: doi\_concept

---

---

1 unique value(s) detected.

doi_concept	N	exactpercent	roundedpercent	cumulpercent
10.5281/zenodo.3911067	27200	100	100	100
Total	27200	100	100	100

Frequency Table for Variable: doi\_version

1 unique value(s) detected.

doi_version	N	exactpercent	roundedpercent	cumulpercent
10.5281/zenodo.10809039	27200	100	100	100
Total	27200	100	100	100

Frequency Table for Variable: lizenz

1 unique value(s) detected.

lizenz	N	exactpercent	roundedpercent	cumulpercent
Creative Commons Zero 1.0 Universal	27200	100	100	100
Total	27200	100	100	100

```
cat(readLines(tar_read(changelog)),
    sep = "\n")
```

## 6 Changelog

### 6.1 Version 2024-03-13

- Vollständige Aktualisierung der Daten
- Die Pipeline mit allen Zwischenergebnissen wird nun automatisch in “output/” archiviert
- Gruppierung von Dateinamen der Diagramme nach Typ
- Vereinfachung der Repository-Struktur
- Anpassung von Docker Compose an Debian 11
- Update GPG Public Key in Repository

### 6.2 Version 2023-03-21

- Vollständige Aktualisierung der Daten
- Gesamte Laufzeitumgebung mit Docker versionskontrolliert
- 4 neue Variablen: BVerwGE, Leitsatz, Fachpresse und Übersetzung (jeweils binär)
- Linguistisch annotierte Variante ist wieder verfügbar
- Aktenzeichen aus dem Eingangszeitraum 2000 bis 2009 nun korrekt mit führender Null formatiert (z.B. 1 BvR 44/02 statt 1 BvR 44/2)
- Aktenzeichen in Verzögerungsverfahren nun korrekt formatiert
- Aktenzeichen verwenden nun richtigerweise den Punkt als Trenner (statt den Schrägstrich wie bei anderen Gerichten)
- Berücksichtigung von Präsident Korbmacher und Vize-Präsidentin Rublack in den jeweiligen Variablen
- Update des Run-Skripts und des Delete-Skripts
- Proto-Package Mono-Repo entfernt, alle Funktionen nun fest projektbasiert versionskontrolliert
- Vereinfachung der Konfigurations-Datei
- Neue Funktion für automatischen clean run (Löschung aller Zwischenergebnisse)
- Update der Download-Funktion
- Überflüssige Warnung in `f.future_lingsummarize`-Funktion entfernt
- Alle Roh-Dateien werden nun im Unterordner “files” gespeichert
- Vom BVerwG nicht veröffentlichte Entscheidungstypen im Codebook genauer beschrieben
- Verbesserte Formatierung von Profiling, Warnungen und Fehlermeldungen im Compilation Report
- Zusätzliche Unit-Tests
- Verbesserung des Robustness Check Reports
- README im Hinblick auf Docker überarbeitet

### 6.3 Version 2022-08-07

- Vollständige Aktualisierung der Daten
- Neuentwurf des Source Codes im *targets* framework
- Einführung von separater Konfigurations-Datei
- Strenge Versionskontrolle aller R packages mit *renv*
- Variante mit linguistischen Annotationen ist temporär nicht mehr verfügbar
- Neue Visualisierung der gesamten Daten-Pipeline
- Vielzahl zusätzlicher Unit Tests (inklusive `type safety`) für wichtige Funktionen

## 6.4 Version 2021-10-19

- Vollständige Aktualisierung der Daten
- Neue Variante mit linguistische Annotationen
- Neue Variable für Lizenz
- Strenge Kontrolle und semantische Sortierung aller Variablen-Namen
- Source Code des Codebooks deutlich vereinfacht (insbes. Diagramme und Changelog)
- Erweiterung der Dokumentation
- In den linguistischen Kennzahlen werden jetzt auch die Anzahl Typen bezogen auf den Gesamtdatensatz berechnet
- Standardisierung der Diagramme auf 6:9 Zoll (Breite/Höhe)

## 6.5 Version 2021-04-15

- Vollständige Aktualisierung der Daten
- Veröffentlichung des vollständigen Source Codes
- Deutliche Erweiterung des inhaltlichen Umfangs des Codebooks
- Einführung der vollautomatischen Erstellung von Datensatz und Codebook
- Einführung von Compilation Reports um den Erstellungsprozess exakt zu dokumentieren
- Einführung von Variablen für Versionsnummer, Concept DOI, Version DOI, ECLI, Präsident:in, Vize-Präsident:in, Verfahrensart und linguistische Kennzahlen (Zeichen, Tokens, Typen, Sätze)
- Zusammenfügung von über Zeilengrenzen getrennten Wörtern
- Automatisierung und Erweiterung der Qualitätskontrolle
- Einführung von Diagrammen zur Visualisierung von Prüfergebnissen
- Einführung kryptographischer Signaturen
- Variable »Suffix« in »kollision« umbenannt.
- Variable »Ordinalzahl« in »eingangsnummer« umbenannt.
- Variable »Entscheidungsart« in »entscheidung\_typ« umbenannt.
- Alle übrigen Variablen sind nun in Kleinschreibung und Snake Case gehalten

## 6.6 Version 2020-06-23

- Erstveröffentlichung



## 7 Parameter für strenge Replikationen

```
system2("openssl", "version", stdout = TRUE)
```

```
## [1] "OpenSSL 3.0.2 15 Mar 2022 (Library: OpenSSL 3.0.2 15 Mar 2022)"
```

```
sessionInfo()
```

```
## R version 4.2.2 (2022-10-31)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 22.04.2 LTS
##
## Matrix products: default
## BLAS: /usr/lib/x86_64-linux-gnu/openblas-pthread/libblas.so.3
## LAPACK: /usr/lib/x86_64-linux-gnu/openblas-pthread/libopenblas-p-r0.3.20.so
##
## locale:
## [1] LC_CTYPE=en_US.UTF-8 LC_NUMERIC=C
## [3] LC_TIME=en_US.UTF-8 LC_COLLATE=en_US.UTF-8
## [5] LC_MONETARY=en_US.UTF-8 LC_MESSAGES=en_US.UTF-8
## [7] LC_PAPER=en_US.UTF-8 LC_NAME=C
## [9] LC_ADDRESS=C LC_TELEPHONE=C
## [11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] stats graphics grDevices utils datasets methods base
##
## other attached packages:
## [1] future.apply_1.10.0 future_1.32.0 spacyr_1.2.1
## [4] quanteda_3.2.4 readtext_0.81 data.table_1.14.8
## [7] scales_1.2.1 ggraph_2.1.0 ggplot2_3.4.1
## [10] pdftools_3.3.3 kableExtra_1.3.4 knitr_1.42
## [13] rvest_1.0.3 httr_1.4.5 mgsub_1.7.3
## [16] zip_2.2.2 fs_1.6.1 testthat_3.1.7
## [19] RcppTOML_0.2.2 tarchetypes_0.7.5 targets_0.14.3
##
## loaded via a namespace (and not attached):
## [1] webshot_0.5.4 rprojroot_2.0.3 future.callr_0.8.1
## [4] tools_4.2.2 backports_1.4.1 utf8_1.2.3
## [7] R6_2.5.1 colorspace_2.1-0 withr_2.5.0
## [10] tidyselect_1.2.0 gridExtra_2.3 processx_3.8.0
## [13] curl_5.0.0 compiler_4.2.2 cli_3.6.0
## [16] xml2_1.3.3 desc_1.4.2 labeling_0.4.2
## [19] stringfish_0.15.7 callr_3.7.3 askpass_1.1
## [22] systemfonts_1.0.4 stringr_1.5.0 digest_0.6.31
## [25] rmarkdown_2.20 svglite_2.1.1 pkgconfig_2.0.3
## [28] htmltools_0.5.4 parallelly_1.34.0 fastmap_1.1.1
## [31] rlang_1.0.6 rstudioapi_0.14 RApiSerialize_0.1.2
## [34] farver_2.1.1 generics_0.1.3 dplyr_1.1.0
```

```

## [37] magrittr_2.0.3      Matrix_1.5-1      Rcpp_1.0.10
## [40] munsell_0.5.0      fansi_1.0.4       viridis_0.6.2
## [43] lifecycle_1.0.3   furrr_0.3.1       stringi_1.7.12
## [46] yaml_2.3.7         MASS_7.3-58.1     brio_1.1.3
## [49] grid_4.2.2         parallel_4.2.2    listenv_0.9.0
## [52] ggrepel_0.9.3     lattice_0.20-45   graphlayouts_0.8.4
## [55] ps_1.7.2           pillar_1.8.1      igraph_1.4.1
## [58] base64url_1.4     codetools_0.2-18  stopwords_2.3
## [61] pkgload_1.3.2     fastmatch_1.1-3   glue_1.6.2
## [64] evaluate_0.20     qpdf_1.3.0        RcppParallel_5.1.7
## [67] vctrs_0.5.2       tweenr_2.0.2      selectr_0.4-2
## [70] gtable_0.3.1     purrr_1.0.1       polyclip_1.10-4
## [73] tidyr_1.3.0       qs_0.25.5         xfun_0.37
## [76] ggforce_0.4.1     tidygraph_1.2.3   viridisLite_0.4.1
## [79] tibble_3.2.0      globals_0.16.2

```

## Literaturverzeichnis

- Allaire, JJ, Yihui Xie, Jonathan McPherson, Javier Luraschi, Kevin Ushey, Aron Atkins, Hadley Wickham, Joe Cheng, Winston Chang, and Richard Iannone. 2023. *Rmarkdown: Dynamic Documents for R*.
- Bengtsson, Henrik. 2021. “A Unifying Framework for Parallel and Distributed Processing in R Using Futures.” *The R Journal* 13 (2): 208–27. <https://doi.org/10.32614/RJ-2021-048>.
- . 2022. *Future.apply: Apply Function to Elements in Parallel Using Futures*.
- . 2023. *Future: Unified Parallel and Distributed Processing in R for Everyone*.
- Benoit, Kenneth, and Akitaka Matsuo. 2020. *Spacyr: Wrapper to the spaCy 'Nlp' Library*. <https://spacyr.quanteda.io>.
- Benoit, Kenneth, and Adam Obeng. 2021. *Readtext: Import and Handling for Plain and Formatted Text Files*. <https://github.com/quanteda/readtext>.
- Benoit, Kenneth, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Müller, and Akitaka Matsuo. 2018. “Quanteda: An R Package for the Quantitative Analysis of Textual Data.” *Journal of Open Source Software* 3 (30): 774. <https://doi.org/10.21105/joss.00774>.
- Benoit, Kenneth, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Müller, Akitaka Matsuo, and William Lowe. 2022. *Quanteda: Quantitative Analysis of Textual Data*. <https://quanteda.io>.
- Csardi, Gabor, and Tamas Nepusz. 2006. “The Igraph Software Package for Complex Network Research.” *InterJournal Complex Systems*: 1695. <https://igraph.org>.
- Csárdi, Gábor, Kuba Podgórski, and Rich Geldreich. 2022. *Zip: Cross-Platform Zip Compression*. <https://github.com/r-lib/zip#readme>.
- Dowle, Matt, and Arun Srinivasan. 2023. *Data.table: Extension of 'Data.frame'*.
- Eddelbuettel, Dirk. 2023. *RcppTOML: Rcpp Bindings to Parser for "Tom's Obvious Markup Language"*. <http://dirk.eddelbuettel.com/code/rcpp.toml.html>.
- Ewing, Mark. 2021. *Mgsub: Safe, Multiple, Simultaneous String Substitution*.
- file., See AUTHORS. 2023. *Igraph: Network Analysis and Visualization*.
- Gagolewski, Marek. 2022. “stringi: Fast and Portable Character String Processing in R.” *Journal of Statistical Software* 103 (2): 1–59. <https://doi.org/10.18637/jss.v103.i02>.
- Gagolewski, Marek, Bartek Tartanus, others; Unicode, Inc., and others. 2023. *Stringi: Fast and Portable Character String Processing Facilities*.
- Landau, William Michael. 2021a. *Tarchetypes: Archetypes for Targets*.
- . 2021b. “The Targets R Package: A Dynamic Make-Like Function-Oriented Pipeline Toolkit for Reproducibility and High-Performance Computing.” *Journal of Open Source Software* 6 (57): 2959. <https://doi.org/10.21105/joss.02959>.
- . 2023a. *Tarchetypes: Archetypes for Targets*.
- . 2023b. *Targets: Dynamic Function-Oriented Make-Like Declarative Pipelines*.

- Ooms, Jeroen. 2023a. *Magick: Advanced Graphics and Image-Processing in R*.
- . 2023b. *Pdftools: Text Extraction, Rendering and Converting of Pdf Documents*.
- Pedersen, Thomas Lin. 2022. *Ggraph: An Implementation of Grammar of Graphics for Graphs and Networks*.
- Ushey, Kevin. 2023. *Renv: Project Environments*. <https://rstudio.github.io/renv/>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- . 2022. *Rvest: Easily Harvest (Scrape) Web Pages*.
- . 2023. *Httr: Tools for Working with Urls and Http*.
- Wickham, Hadley, Winston Chang, Lionel Henry, Thomas Lin Pedersen, Kohske Takahashi, Claus Wilke, Kara Woo, Hiroaki Yutani, and Dewey Dunnington. 2023. *Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*.
- Wickham, Hadley, and Dana Seidel. 2022. *Scales: Scale Functions for Visualization*.
- Xie, Yihui. 2014. “Knitr: A Comprehensive Tool for Reproducible Research in R.” In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC.
- . 2015. *Dynamic Documents with R and Knitr*. 2nd ed. Boca Raton, Florida: Chapman; Hall/CRC. <https://yihui.org/knitr/>.
- . 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in R*. <https://yihui.org/knitr/>.
- Xie, Yihui, J. J. Allaire, and Garrett Golemund. 2018. *R Markdown: The Definitive Guide*. Boca Raton, Florida: Chapman; Hall/CRC. <https://bookdown.org/yihui/rmarkdown>.
- Xie, Yihui, Christophe Dervieux, and Emily Riederer. 2020. *R Markdown Cookbook*. Boca Raton, Florida: Chapman; Hall/CRC. <https://bookdown.org/yihui/rmarkdown-cookbook>.
- Zhu, Hao. 2021. *KableExtra: Construct Complex Table with Kable and Pipe Syntax*.