

Jens Ambrasat | Christophe Heger | Annegret Rucker

DZHW- Wissenschaftsbefragung 2019

Daten- und Methodenbericht zur DZHW-Wissenschaftsbefragung
2019

Daten- und Methodenbericht

Februar 2022

Autor*innen:

Dr. Jens Ambrasat (Projektleiter)

Christophe Heger

Annegret Rucker

Unter Mitarbeit von (bei Kapitel 4):

Gregor Fabian

Anne Weber

Andreas Daniel

Henrike Schmidtchen

Herausgeber:

Deutsches Zentrum für Hochschul- und Wissenschaftsforschung GmbH (DZHW)

Lange Laube 12 | 30159 Hannover | www.dzhw.eu

Tel.: +49 511 450670-0 | Fax: +49 511 450670-960 | info@dzhw.eu

Geschäftsführung:

Prof. Dr. Monika Jungbauer-Gans

Karen Schlüter

Vorsitzender des Aufsichtsrats:

Ministerialdirigent Peter Greisler

Registergericht:

Amtsgericht Hannover | B 210251

Dieses Werk steht unter der Creative Commons "Namensnennung – Nicht kommerziell – Weitergabe unter gleichen Bedingungen 3.0 Deutschland Lizenz" (CC-BY-NC-SA)

<https://creativecommons.org/licenses/by-nc-sa/3.0/>



Inhaltsverzeichnis

Inhaltsverzeichnis	III
1 Studienüberblick	1
2 Methoden.....	3
2.1 Stichprobe und Grundgesamtheit.....	3
2.2 Feldphase und Rücklauf	4
2.3 Samplebeschreibung.....	7
2.4 Gewichtung	9
3 Zum Fragebogen	11
3.1 Themenübersicht	11
3.2 Modularisierung.....	12
4 Hinweise zum Scientific Use File	14
4.1 Allgemeine Datennutzungshinweise.....	14
4.2 Anonymisierung	15
Literaturverzeichnis.....	25
Anhang A - Einladungsschreiben	26
Anhang B - Übersicht der Trendfragen.....	27

1 Studienüberblick

Die DZHW-Wissenschaftsbefragung ist eine bundesweite, repräsentative Trendstudie zur Erforschung der Arbeits- und Forschungsbedingungen an deutschen Universitäten und gleichgestellten Hochschulen. Die ersten beiden Wissenschaftsbefragungen fanden in den Jahren 2009 und 2016 statt. Danach wurde beschlossen, die DZHW-Wissenschaftsbefragung als regelmäßige Trendstudie alle drei Jahre durchzuführen. Die aktuelle Welle wurde im Wintersemester 2019/20 von November 2019 bis Februar 2020 durchgeführt. Die COVID19 Pandemie ab März 2020 hatte dementsprechend noch keine Auswirkungen auf die Befragung der Wissenschaftler*innen, die sich folglich auf ihren regulären Forschungsalltag beziehen konnten.

Die DZHW-Wissenschaftsbefragung versteht sich als langfristiges Barometer für die Wissenschaft. Sie ist ein grundlegendes Werkzeug zur Bestandsaufnahme und Reflektion des Zustands an deutschen Universitäten und liefert ein umfassendes Meinungsbild zur Lage und Entwicklung des deutschen Wissenschaftssystems. Zeitliche Veränderungen von Einstellungen und Meinungen zu wissenschaftspolitischen Themen werden im Rahmen von Trendanalysen untersucht. Die Daten sind darüber hinaus Grundlage für Fachpublikationen und dienen so auch als empirische Grundlage für öffentliche Diskussionen und politische Entscheidungsprozesse. Zudem werden die aufbereiteten, anonymisierten Daten dauerhaft und dokumentiert über das DZHW-Forschungsdatenzentrum (FDZ) bereitgestellt und somit anderen Interessent*innen in Forschung und Wissenschaftspolitik zugänglich gemacht.

Die DZHW-Wissenschaftsbefragung adressiert Wissenschaftler*innen an deutschen Universitäten und gleichgestellten Hochschulen. Dafür werden öffentlich verfügbare Adressdaten von Professor*innen und wissenschaftlichen Mitarbeiter*innen auf den Webseiten der Hochschulen recherchiert. In der aktuellen Befragungswelle konnte ein Netto-Analysesample von 8.822 Befragten realisiert werden. Diese Stichprobe ist groß genug, um Subgruppen detailliert zu untersuchen und fächervergleichende Analysen vornehmen zu können.

Die DZHW-Wissenschaftsbefragung ist als Mehrthemenbefragung konzipiert. Jede Befragung enthält ein Kerninstrument (ein über mehrere Wellen konstantes Erhebungsprogramm) mit zentralen Fragen zu den Arbeits- und Forschungsbedingungen und Module mit tiefergehenden Fragen zu forschungspolitisch aktuellen Themen und Forschungsschwerpunkten. Das Kerninstrument wird komplementiert durch Schwerpunktthemen, für die die Instrumente in Kooperation mit Forschergruppen innerhalb und außerhalb des DZHW entwickelt werden.

Die Schwerpunktthemen der 2019er Wissenschaftsbefragung waren Open Data, Forschungsinformationen, Autorschaft, Wissenstransfer und Fachkulturen.

Sowohl für die Wissenschaftsforschung als auch für die Wissenschaftspolitik ist ein möglichst umfassendes und genaues Verständnis der Veränderungen im Wissenschaftssystem von Interesse. Dies erfordert systematische Beobachtung und Forschung, die Veränderungen auch im Zeitverlauf erfassen kann. Abgesehen von den größeren Befragungen der 1970er und 1980er Jahre des Instituts für Demoskopie Allensbach (IFD) (Maier-Leibnitz 1985) sowie verschiedenen thematisch, disziplinär oder regional begrenzten Einzelstudien fehlte jedoch bis dato eine gezielte und

regelmäßige Erfassung der Sicht der Forschenden. Die DZHW-Wissenschaftsbefragung setzt an diesem Punkt an.

Das vom Stifterverband durchgeführte *Hochschulbarometer* befragt seit 2011 bundesweit Hochschulleitungen zur aktuellen Situation in der deutschen Hochschullandschaft. Das von Wissenschaft im Dialog durchgeführte *Wissenschaftsbarometer* befragt jährlich rund 1000 Bürger*innen zu ihren Wahrnehmungen, Einstellungen und dem Vertrauen gegenüber der Wissenschaft. Die DZHW **Wissenschaftsbefragung** komplementiert diese Perspektiven auf die Funktionalität des Wissenschaftssystems durch die Erfahrungen und Einschätzungen der Wissenschaftlerinnen und Wissenschaftler selbst. Ihre Praktiken, Wahrnehmungen und Einstellungen stehen im Mittelpunkt der Wissenschaftsbefragung.

Die zentralen Ergebnisse der Studie wurden als *Barometer (von und) für die Wissenschaft* im November 2020 veröffentlicht und auf der Projektseite öffentlich zugänglich gemacht.

Nachfolgend wird über die methodischen Details der Befragung (Stichprobendesign, Grundgesamtheit, Feldphase und Rücklaufstatistik) und über den Scientific Use File informiert.

2 Methoden

2.1 Stichprobe und Grundgesamtheit

Die DZHW-Wissenschaftsbefragung 2019 ist eine Onlinebefragung des hauptberuflichen, wissenschaftlich-künstlerischen Personals an deutschen Universitäten und gleichgestellten Hochschulen¹. Da für diese *Grundgesamtheit* (GG) keine Liste mit Kontaktinformationen existiert, wurde ein Auswahlrahmen festgelegt, der dieser Grundgesamtheit möglichst nahe kommt (vgl. Gabler und Häder 2015). Als *Auswahlrahmen* wurden die im Internet auf Webseiten von Universitäten und gleichgestellten Hochschulen auffindbaren Wissenschaftlerinnen und Wissenschaftler mit Adressinformationen bestimmt. Geht man davon aus, dass nahezu alle Wissenschaftler*innen auf den Internetseiten ihrer Einrichtungen präsent sind, kommt dieser Auswahlrahmen nächstmöglich an die Grundgesamtheit heran. Darüber hinaus sind keine systematischen Abweichungen bekannt.

Im Vorfeld der Befragung wurden zwischen Mai und August 2019 an 132 Universitäten und gleichgestellten Hochschulen Adressinformationen von wissenschaftlichen Angestellten recherchiert (Auswahlgesamtheit). Die Adressen wurden in Kooperation mit der Universität Zürich erhoben und umfassen Professor*innen, Postdocs (Wissenschaftler*innen mit abgeschlossener Promotion) und Prädocs (Wissenschaftler*innen ohne Promotion, unabhängig davon, ob sie an einer Promotion arbeiten oder nicht). Die Mailadressen und weitere für das Anschreiben notwendige Informationen wurden halbautomatisiert als Adressdaten aufbereitet und qualitätsgeprüft. Anschließend wurde aus diesem Pool durch ein Zufallsverfahren eine *Bruttoadressesstichprobe* von 60.002 Adressen gezogen. Dabei wurden Professorinnen und Professoren sowie Postdocs überproportional, d.h. mit größeren Wahrscheinlichkeiten gezogen als Prädocs. Eine solche disproportional geschichtete Stichprobe führt qua Design dazu, dass die Prädocs zunächst unterrepräsentiert sind. Dies muss in den Analysen berücksichtigt und, wo nötig, durch Gewichtung ausgeglichen werden. Die entsprechenden Designgewichte liefern wir mit (siehe Abschnitt zur Gewichtung).

¹ Lehrbeauftragte, Privatdozent*innen und Honorarprofessor*innen sind nicht Teil der Stichprobe.

2.2 Feldphase und Rücklauf

Am 26.11.2019 startete der initiale Versand der Einladungsmails (vgl. Anhang A - Einladungsschreiben) mit darin enthaltenen individualisierten Links zum Fragebogen an die 60.002 Adressen des Bruttosamples. Die gesamte Feldzeit erstreckte sich bis zum 24. Februar. Nach den Einladungen gab es im Abstand von ungefähr jeweils einem Monat zwei Erinnerungsschreiben, und zwar zwischen dem 8. und 10. Januar die ersten Reminder und zwischen dem 11. und 13. Februar die zweiten Reminder. Wenige Tage nach den Einladungen (bzw. Remindern) gab es Abbrecher-Reminder für Personen, die den Fragebogen schon begonnen, aber noch nicht zu Ende gebracht hatten. Diese Abbrecher-Reminder hatten den Zweck, den Personen den direkten Link zur Wiederaufnahme der Befragung zu schicken, so dass sie nicht erneut über die Projektwebseite und die Datenschutzerklärung gehen mussten.

Tabelle 1. Verlauf der Feldphase

	Datum	Effektivität
Einladungen		3774 (42,8%)
Softlaunch	26. November	
1. Drittel (Anredeexperiment)	28. November	
2. und 3. Drittel	03.- 05. Dezember	
Abbrecher-Reminder	09. Dezember	
1. Reminder		2435 (27,6 %)
Reminder	08.- 10. Januar	
Abbrecher-Reminder	13. Januar	
2. Reminder		2613 (29,6 %)
Reminder	11. – 13. Februar	
Abbrecher-Reminder	17. Februar	

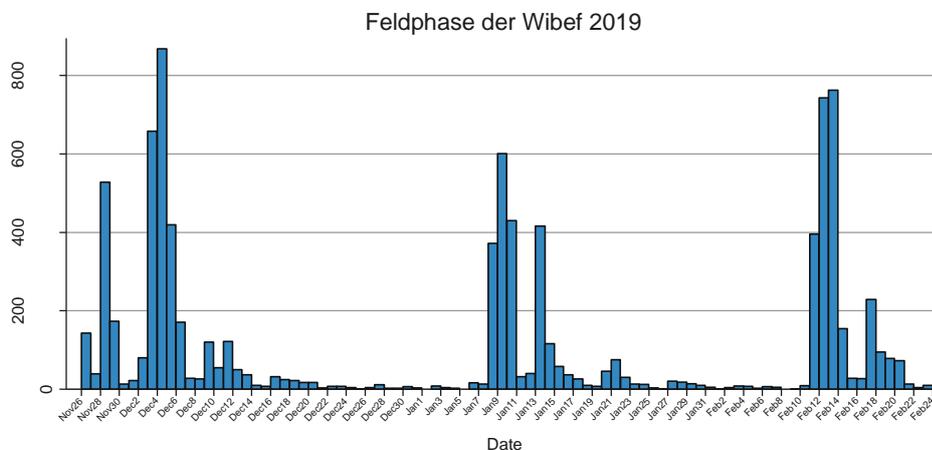


Abbildung 1. Verlauf der Feldphase

Abbildung 1 zeigt deutlich die positive Wirkung der Erinnerungsschreiben. Der erste Reminder generierte 28 % des gesamten Rücklaufs und der zweite Reminder mit 2.613 beendeten und auswertbaren Fragebögen sogar 30 % der insgesamt 8.822er Nettostichprobe.

Von den insgesamt 60.002 Bruttoadressdaten konnten 52.769 Kontakte realisiert werden. Über 3.511 der 60.002 verwendeten Adressen konnte kein Kontakt realisiert werden (Bounced mails), aufgrund veralteter oder fehlerhafter E-Mailadressen, die nicht korrigiert werden konnten. Ein Teil der anfänglichen Teilnehmer*innen gehörte nicht zur anvisierten Grundgesamtheit und wurde im Fragebogen ausgescreent. Hochgerechnet auf die Bruttoadressstichprobe entspricht das einem geschätzten Umfang von 3.722, die nicht zur GG zugehörig waren. Weitere 2.011 Befragte haben die Befragung zwar aufgerufen, aber während der Befragung abgebrochen. Die Beendigungsquote von 80 % ist für eine Befragung dieser Länge, ohne Vorakquise und ohne materielle Incentivierung äußerst zufriedenstellend (vgl. Galesic und Bosnjak 2009). Wir führen dieses Ergebnis auf die für die Situation der befragten Wissenschaftler*innen passenden Fragen und ein konsistentes Routing des Onlinefragebogens zurück. Durch die Datenbereinigung wurden weitere 211 Datensätze entfernt, die sich als übermäßig unvollständig und inkonsistent bzw. als so genannte „Durchklicker“ erwiesen.

Tabelle 2. Rücklaufstatistik

Adressstichprobe (brutto)	60.002
Davon:	
Bounced Mails	3.511
Nicht zur GG gehörig (Hochrechnung ²)	3.722
bereinigte Adressstichprobe	52.769
Initial Response ³ , davon:	11.044
Abbrecher ⁴	2.011
Unvollständig oder inkonsistent ⁵	211
Nettostichprobe (beendet, vollständig)	8.822
Ausschöpfungsquote (RR1)	14,7 %
Ausschöpfungsquote (RR3)	16,7 %
Beendigungsquote (1 - Abbruchquote)	79,9 %

² Die im Fragebogen ausgescreenten dienen als Grundlage für die Hochrechnung auf die Bruttoadressstichprobe, vgl. AAPOR 2016.

³ Befragte, die ihr datenschutzrechtliches Einverständnis gegeben haben und mindestens auf der ersten Seite des Fragebogens waren.

⁴ Als Abbrecher zählen alle Befragten, die nicht die letzte Frage (Zufriedenheit mit der Vereinbarkeit) erreicht haben.

⁵ Als unvollständig wurden alle herausgenommen, die im gesamten Fragebogen mehr als 30 % Item- NonResponse aufweisen. Als inkonsistent zählen zusätzlich Fälle, die Auffälligkeiten wie „Durchklicken“, inkonsistente biografische oder anderen Zeitangaben aufwiesen und in der Einzelfallprüfung ein inkonsistentes Gesamtbild aufwiesen.

Schließlich konnte eine Nettoanalysestichprobe von 8.822 Wissenschaftlerinnen und Wissenschaftlern realisiert werden. Damit konnte auch die Ausschöpfungsquote mit 14,7 % (RR1, bzw. 16,7% RR3) gegenüber der letzten Befragung 2016 deutlich gesteigert werden. Die Nettoanalysestichprobe setzt sich aus 1.721 von Professor*innen, sowie 4.216 promovierten und 2.885 nicht promovierten Wissenschaftler*innen des Mittelbaus zusammen.

Tabelle 3. Rücklauf nach Statusgruppe

Responsestatus	Statusgruppe			Total
	Prof	Postdoc	Prädoc	
eingeladen	13.927	22.950	23.125	60.002
beendet	1.721	4.216	2.885	8.822
Quote	12,4 %	18,4 %	12,5 %	14,7 %

Die mittlere Bearbeitungsdauer betrug 22,4 Minuten (Median⁶). Da einige Befragte Bearbeitungszeiten von über einer Stunde aufweisen, weicht das arithmetische Mittel deutlich vom Median ab. Solche vereinzelt sehr großen Bearbeitungsdauern sind vermutlich auf Unterbrechungen zurückzuführen, bei denen die Befragung technisch nicht abgebrochen bzw. unterbrochen wurde. Angekündigt waren 25 Minuten Bearbeitungszeit. Da die Statusgruppen aufgrund teilweise differenzierter Filterführung höchst unterschiedliche Bearbeitungsdauern aufweisen, war diese Ankündigung an der Gruppe der Professor*innen mit den höchsten Bearbeitungsdauern ausgerichtet (vgl. Tabelle 4).

Tabelle 4. Bearbeitungsdauern nach Statusgruppe und Befragungsmodul

	Median	mean	sd	N
Statusgruppe				
Prof	25.8	30.8	18.6	1.721
Postdoc	22.7	28.3	19.7	4.216
Prädoc	19.9	25.5	18.4	2.885
Total	22.4	27.9	19.2	8.822
Modulsplit				
A-Data Sharing	22.1	27.6	19.4	2.209
B-Forschungsinformationen	21.7	27.3	19.4	2.170
C-Wissenstransfer	21.6	26.9	18.6	2.300
D-Autorschaft	24.6	29.8	19.1	2.143
Total	22.4	27.9	19.2	8.822

⁶ Der Median ist hier als mittlerer Lageparameter aussagekräftiger als das arithmetische Mittel, da er im Gegensatz zum arithmetischen Mittel robust gegen Ausreißer ist. Da bei Onlinesurveys stets Ausreißer entstehen, z.B. durch Unterbrechungen ohne die Befragung zu schließen, gibt das arithmetische Mittel stets ein nach oben verzerrtes Bild der „mittleren“ Bearbeitungsdauer an.

2.3 Samplebeschreibung

Die befragten Personen verteilen sich relativ gleichmäßig über die Fächergruppen. 16,4 % kommen aus den Geisteswissenschaften, 22,3 % aus den Sozialwissenschaften und 19,2 % aus den Lebenswissenschaften, 23,5 % vertreten die Naturwissenschaften und 16,0 % die Ingenieurwissenschaften (vgl. Tabelle 5). Die Klassifikation der im Fragebogen auswählbaren 48 Fachgebiete erfolgte anhand der DFG Fachsystematik der Amtsperiode 2016-2019. Dargestellt sind in Tabelle 5 die DFG-Wissenschaftsbereiche mit einer zusätzlichen Unterteilung nach Geistes- und Sozialwissenschaften.

Tabelle 5. Samplestatistik

	Anzahl	Prozent	Cum.
Fachgebiet nach DFG			
Geisteswissenschaften	1.444	16,37	16,37
Sozial- und Verhaltenswissenschaften	1.968	22,31	38,68
Lebenswissenschaften	1.689	19,15	57,82
Naturwissenschaften	2.075	23,52	81,34
Ingenieurwissenschaften	1.409	15,97	97,31
Ohne Zuordnung	237	2,69	100,00
Total	8.822	100,00	
Position			
Professor*in	1.555	17,63	17,63
Juniorprofessor*in	166	1,88	19,51
Akademische*r Rat/Rätin	614	6,96	26,47
Wissenschaftliche*r Mitarbeiter*in	6.280	71,19	97,65
- davon promoviert	3.476	55,35 %	
- davon nicht promoviert	2.804	44,65 %	
Sonstige wissenschaftliche Position	207	2,35	100,00
Total	8.822	100,00	
Geschlecht			
Männlich	5.120	58,38	58,38
Weiblich	3.597	41,01	99,40
Divers	53	0,60	100,00
Total	8.822	100,00	

17,6 % der Teilnehmenden im Sample sind Professor*innen, 1,9 % Juniorprofs und 71,2 % Wissenschaftliche Mitarbeitende, von denen 55,4 % promoviert und 44,7 % nicht promoviert sind. 58,4 % der Wissenschaftler*innen sind männlich, 41,0 % weiblich und 0,6 % der Befragten ordnen sich der Kategorie divers zu.

Die realisierte Stichprobe lässt sich auch mit der Hochschulpersonalstatistik vergleichen, die vom statistischen Bundesamt herausgegeben wird. Für den Vergleich ziehen wir die Personalstatistik von 2018 heran, weil die Daten von 2019 zum Redaktionsschluss noch nicht vorlagen (Tabelle 6). Der Vergleich zeigt, dass die verschiedenen Fachgebiete und Fächergruppen in der Breite sehr gut abgebildet sind. Gegenüber der Statistik des statistischen Bundesamtes sind die Geistes- und Sozialwissenschaften, sowie die Naturwissenschaften im Sample etwas überrepräsentiert, während Ingenieur*innen etwas unterrepräsentiert sind.

Tabelle 6. Fächerverteilung im Vergleich mit der Personalstatistik des StaBu 2018

	StaBu Personalstatistik 2018						WiBef 2019					
	Prof		WiMi		Total		Prof		WiMi		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
Geisteswissenschaften	7.187	29,5	15.345	12,4	22.532	15,2	370	21,9	1.074	15,6	1.444	16,8
Sozial- und Verhaltenswissenschaften	5.827	23,9	22.505	18,2	28.332	19,1	460	27,2	1.508	21,9	1.968	22,9
Medizin (ohne Kliniken)	1.639	6,7	14.822	12,0	16.461	11,1	134	7,9	525	7,6	659	7,7
Agrar-, Forstwiss. u. Veterinärmedizin	523	2,1	3.842	3,1	4.365	2,9	30	1,8	132	1,9	162	1,9
Naturwissenschaften	5.586	22,9	36.112	29,2	41.698	28,2	481	28,4	2.462	35,7	2.943	34,3
Ingenieurwissenschaften	3.641	14,9	31.050	25,1	34.691	23,4	216	12,8	1.193	17,3	1.409	16,4
Total	24.403	100	123.676	100	148.079	100	1.691	100	6.894	100	8.585	100
Nicht zugeordnet							30	1,7	207	2,9	237	2,7

2.4 Gewichtung

Eine Gewichtung dient dazu, Ungenauigkeiten von Schätzern zu verhindern, die durch bekannte Verzerrungen der Stichprobe hervorgerufen werden können. Es gibt zwei Quellen solcher Verzerrungen der Stichprobe: Selektive Auswahl der Befragten durch das Erhebungsdesign und Selektivität durch bewusste Teilnahme bzw. Nicht-Teilnahme (Non-Response) seitens der Befragten (vgl. Kroh et al. 2015). Wir gehen auf beide Formen jeweils mit einer entsprechenden Gewichtung ein.

Designgewichtung

Die Designgewichtung gleicht die durch das Erhebungsdesign induzierten unterschiedlichen Auswahlwahrscheinlichkeiten aus. Im Rahmen der Wissenschaftsbefragung werden Designgewichte berechnet, um die designbedingte Überrepräsentation von Professor*innen und Postdocs und die daraus folgende Unterrepräsentanz von Prädocs auszugleichen.

Die Wissenschaftsbefragung 2019 wurde als disproportional geschichtete (stratifizierte) Zufallsstichprobe mit größeren Ziehungswahrscheinlichkeiten für Professor*innen und Postdocs durchgeführt, um in diesen Gruppen eine hinreichende Fallzahl für detaillierte Analysen zu sichern. Weil die Ziehungswahrscheinlichkeiten von Professor*innen und Postdocs erhöht wurden (Oversampling), sind Prädocs unterrepräsentiert. Diese designbedingte Unterrepräsentanz von Prädocs wird mit einem Designgewicht ausgeglichen, welches wir im Scientific Use File (SUF) anbieten. Dabei ergibt sich das Designgewicht w_D unmittelbar aus dem disproportionalen Schichtungsansatz als Inverse der Ziehungswahrscheinlichkeit (vgl. Tabelle 7).

Non-Response-Gewichtung

Um mögliche Verzerrungen in der Stichprobe auszugleichen, die sich durch ein selektives Antwortverhalten (Non-Response) der Befragten ergeben, greifen wir auf die Informationen aus der Bruttoadressstichprobe zurück. Hier ziehen wir wiederum die Statuspositionen heran, die aus der Bruttostichprobe bekannt sind. Die Non-Responsegewichtung entspricht einer Anpassung an die entsprechende Verteilung in der *Bruttostichprobe* und wird als umgekehrte Teilnahmewahrscheinlichkeit der jeweiligen Statusgruppen berechnet.

Kombiniertes Design- und Non-Response-Gewicht

In einem dritten Schritt kombinieren wir die Designgewichtung und die Non-Response-Gewichtung in einem Gewicht. Dieses kombinierte Wahrscheinlichkeitsgewicht passt die Nettostichprobe somit direkt an die *Auswahlgesamtheit* an und berücksichtigt sowohl das stratifizierte Design als auch den Gruppenspezifischen Non-Response. Das entspricht faktisch einer Kombination des Designgewichts mit dem Non-Response Gewicht und macht somit die einzelne Verwendung dieser beiden Gewichte überflüssig (vgl. Tabelle 7).

Die Gewichtung sollte verwendet werden, wenn durchschnittliche, über alle Statusgruppen hinweggehende, Größen – für alle Wissenschaftler*innen – geschätzt werden. Bei statusgruppendifferenzierenden Analysen wird dieses Gewicht nicht benötigt. Ebenso wenig wie in multivariaten Analysen, wenn auf die Statusgruppen kontrolliert wird. Da sich die Angaben der Statusgruppen bei den allermeisten gemessenen Merkmalen (wissenschaftlichen Praktiken, als auch bei Einstellungen) deutlich unterscheiden, empfiehlt es sich jedoch ohnehin stets diese Gruppenunterschiede mit zu berücksichtigen und auszuweisen. Werden die Gewichte verwendet, so gelten die Daten als repräsentativ für die als Auswahlrahmen definierte Menge an Wissenschaftlerinnen und Wissenschaftlern.

Tabelle 7. Design- und Non-Response- Gewichte

	Profes	Postdocs	Prädocs	Gesamt
Auswahlgesamtheit	27.900	46.232	93.250	167.382
Disproportional geschichtete Stichprobe	13.927	22.950	23.125	60.002
Ziehungswahrscheinlichkeit	0,499	0,496	0,248	0,358
Designgewicht: w_D	2,003	2,014	4,032	
Nettostichprobe	1.721	4.216	2.885	8.822
Teilnahmewahrscheinlichkeit	0,124	0,184	0,125	0,147
Non Responsegewicht: w_A	8,092	5,444	8,016	
Kombiniertes Gewicht: w_K	16,212	10,966	32,322	18,973
Kombiniertes Wahrscheinlichkeitsgewicht, normiert	0,854	0,578	1,704	Mean=1

Randanpassung (Kalibrierung)

Je nach Fragestellung kann auch eine Anpassung (Poststratifizierung) an bekannte Randverteilungen der Grundgesamtheit vorgenommen werden. Hier bieten sich Daten des statistischen Bundesamtes an. Diese weisen jährlich die Zahlen für das hauptberufliche, wissenschaftlich-künstlerische Personal aus (vgl. Tabelle 6).

Die Zahlen für 2019 lagen zum Redaktionsschluss dieses Kapitels noch nicht vor. Falls Nutzer*innen des Scientific Use Files diese Randanpassung vornehmen möchten, empfehlen wir die Angaben zum „höchsten Hochschulabschluss“ (siehe Statistisches Bundesamt, Fachserie 11, Reihe 4.4, 2018: S. 88) in Verbindung mit der „Personalgruppe“ zu verwenden. Ob eine Randanpassung sinnvoll ist, hängt von der konkreten Fragestellung ab und ist im Einzelfall zu prüfen (vgl. Hdioud und Tirari 2019).

3 Zum Fragebogen

3.1 Themenübersicht

Die Wissenschaftsbefragung ist eine breit aufgestellte Mehrthemenbefragung zu den Arbeits- und Forschungsbedingungen in der Wissenschaft. Über die Kernfragen hinaus, die in jeder Welle laufen, sind Schwerpunktthemen in den Fragebogen integriert. Zu den Kernfragen gehören beispielsweise Fragen zu den Forschungsbedingungen, zur Forschungspraxis, zum wissenschaftlichen Nachwuchs, zu Wissenstransfer und Wissenschaftspolitik. Alle Themenblöcke sowie die Schwerpunkte der diesjährigen Wissenschaftsbefragung können der Grafik in Abbildung 2 entnommen werden. Schwerpunktthemen waren in diesem Jahr zum einen Fachkulturen (unter der Fragestellung, inwiefern sich die Forschungspraktiken zwischen den Fächern unterscheiden), Open Data (In welchem Maß werden Forschungsdaten von Wissenschaftler*innen geteilt?) und Forschungsinformationen (Über welche Praktiken und Verfahren werden Forschungsinformationen an Hochschulen erfasst?). Zum anderen lagen die Schwerpunkte auf Autorschaft (Welche Normen existieren für die Verteilung von Autorschaften?), Wissenstransfer (Inwiefern kooperieren Wissenschaftler*innen mit Akteur*innen innerhalb und außerhalb der Wissenschaft?) und aktuellen wissenschaftspolitischen Themen.

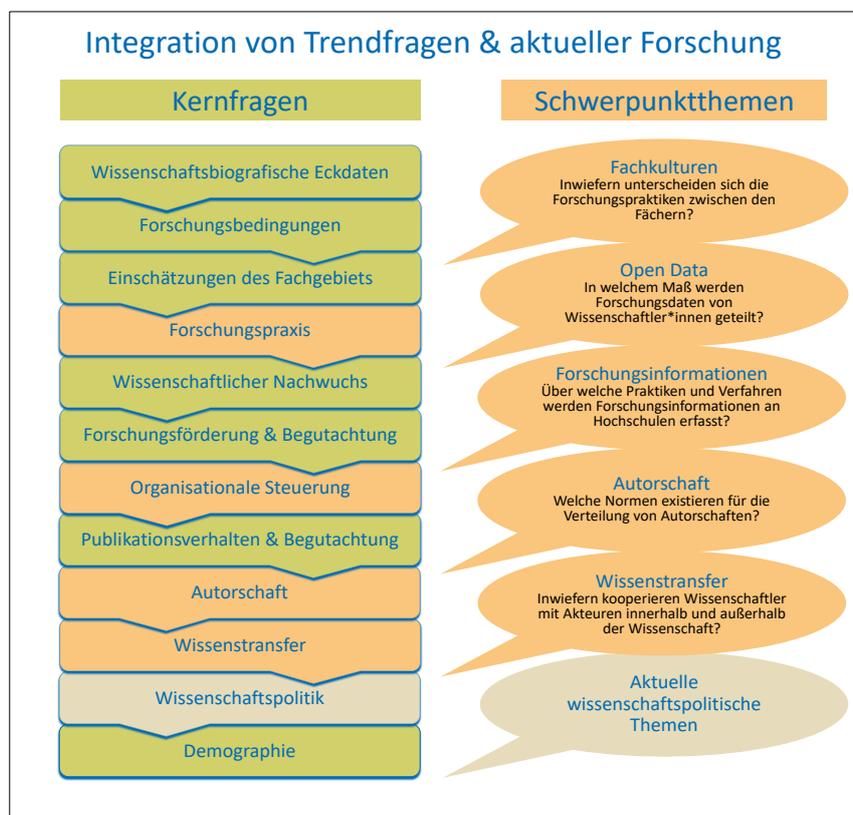


Abbildung 2. Themenblöcke der Wissenschaftsbefragung

Ein Teil der Fragen aus den vorangegangenen Wissenschaftsbefragungen der Jahre 2009 und 2016 wurde übernommen, um zeitliche Veränderungen gezielt untersuchen zu können. Eine Übersicht über diese Trendfragen, die in mehreren Befragungswellen erhoben wurden, ist in Anhang B - Übersicht der Trendfragen gegeben. Im Block zu wissenschaftspolitischen Themen wurden einige Fragen (z.B. zur Exzellenzstrategie und zum Promotionsrecht an Fachhochschulen) aktualisiert und weitere Fragen, z.B. zum Zustand des Wissenschaftssystems und zu Reformen, ergänzt. Einige Items der Frage ff7 (Ansichten über die Förderung durch die DFG) wurden aus früheren Wissenschaftsbefragungen übernommen, die ursprünglich aus den Allensbacher Hochschullehrerbefragungen 1976/77 und 1983/84 stammen (vgl. dazu Böhmer et al. 2011: 173). Die Frage de3 (Diskriminierung) wurde aus Gründen der Vergleichbarkeit mit der Allgemeinbevölkerung aus Beigang et al. 2017 übernommen.

3.2 Modularisierung

Die wichtigste Einflussgröße für eine hohe Responserate und eine gute Datenqualität ist ein ansprechender, in sich konsistenter und nicht zu umfangreicher Fragebogen. Damit der Fragebogen nicht zu lang und die Belastung für die einzelnen Befragten nicht zu groß wird, wurde der Fragebogen systematisch modularisiert und einzelne Module randomisiert. So bekommen zwar alle Befragten die Fragen aus dem Kerninstrument, sowie einige zentrale Fragen zur Einschätzung des Wissenschaftssystems, jedoch jeweils nur eines von vier Vertiefungsmodulen. Auf diese Weise ist es möglich, das eigentlich sehr umfangreiche Befragungsprogramm von 162 Fragen und 420 Items für die Befragten auf 25 Minuten zu reduzieren (vgl. Tabelle 4).

Die Befragten wurden dafür zufällig auf vier verschiedene Gruppen aufgeteilt, denen jeweils ein Schwerpunktthema (Modul) zugeordnet war. Die Zusammenstellung der Module/Fragen orientierte sich an der thematischen Gliederung des Instruments. Tabelle 8 zeigt den Modulplan, dem man die entsprechenden Modulsplits *Msplit/A* bis *Msplit/D* entnehmen kann, wobei jeder Buchstabe für ein Modulthema steht.

Modul A: Open data / data sharing

Modul B: Forschungsinformationen

Modul C: Wissenstransfer

Modul D: Autorschaft

Darüber hinaus wurden in den Themenblöcken wissenschaftlicher Nachwuchs und Forschungsförderung/Begutachtung jeweils zwei lokale Randomsplits (*Rsplit_1* und *Rsplit_2*) vorgenommen, bei denen jeweils eine Hälfte der Befragten andere Fragen zugeteilt bekam als die andere Hälfte. Die beiden Randomsplits waren jeweils voneinander sowie von den vier Modulsplits unabhängig. Im Modulplan (Tabelle 8) ist zudem dargestellt, welche Fragen für Professor*innen gefiltert waren.

Tabelle 8. Modulplan

WiBef 2019: Modulplan		Modultypen (wer erhält was)							
		A/Prof	A/NI/M	B/Prof	B/NI/M	C/Prof	C/NI/M	D/Prof	D/NI/M
0=neu x=trend Tr.									
1. Wissenschaftliche und Biographische Eckdaten									
Art der Einrichtung	x k	x	x	x	x	x	x	x	x
Position, Besoldungsstufe, promoviert, habilitiert	x k	x	x	x	x	x	x	x	x
Akademischer Lebenslauf (HSA, Promotion, Erstberufung)	x k	x	x	x	x	x	x	x	x
Fachgebiet	x k	x	x	x	x	x	x	x	x
Karriereziel	x k	x	x	x	x	x	x	x	x
Funktionsrollen im Wissenschaftssystem	o k	x	x	x	x	x	x	x	x
2. Forschungs- und Arbeitsbedingungen									
2.1. Forschungsbedingungen									
Stellenfinanzierung	x k	x	x	x	x	x	x	x	x
Arbeitsvertrag (Befristung, Tenure Track, Laufzeit, Befristungsgrund, Anzahl)	x k	x	x	x	x	x	x	x	x
Vertragliche und tatsächliche Arbeitszeit	x k	(x)	x	(x)	x	(x)	x	(x)	x
Zeitbudgets für Tätigkeiten/ Lehrverpflichtung.	x k	x	x	x	x	x	x	x	x
Zufriedenheit Beruf	x k	x	x	x	x	x	x	x	x
Zufriedenheit Ausstattung Professur (Bayern)	o k	x		x		x		x	
2.2 Einschätzungen der Forschungssituation im Fachgebiet									
Allgemeine evaluative Einschätzungen ("In meinem Fachgebiet")	(x) k	x	x	x	x	x	x	x	x
2.3 Forschungspraxis									
Forschungspraxis (semantisches Differenzial)	o k	x	x	x	x	x	x	x	x
Nutzung von Daten (open data)	o Msplit/A	x	x						
Erzeugung von Daten (open data)	o Msplit/A	x	x						
Gründe gegen Open Data	o Msplit/A	x	x						
2.4 Wissenschaftlicher Nachwuchs									
Betreute Promovierende (Anzahl) /Erstbetreuer	x k	x		x		x		x	
Qualität von Qualifizierungsarbeiten	o k	x		x		x		x	
Plagiate (Umgang mit, Häufigkeit)	o Rsplit_1/1	x		x		x		x	
Einschätzungen zur Situation des Nachwuchses	x Rsplit_1/2	x		x		x		x	
3. Forschungsförderung/Drittmittelsituation									
3.1. Forschungsförderung & Begutachtung									
Anz. Drittmittelanträge	x k	x	x	x	x	x	x	x	x
Einschätzung von Forschungsförderern (DFG, BMBF, ERC): Aufwand, Qualität,	o Rsplit_2/1	x	x	x	x	x	x	x	x
Qualität von Begutachtung bei Förderanträgen (Entwicklung letzte 5 Jahre)	x k	x	x	x	x	x	x	x	x
Allensbach (DFG) Items	x k	x	x	x	x	x	x	x	x
Ausgestaltung von Begutachtungsprozessen (Fairness, Vorleistungen, Losverfahren)	o Rsplit_2/2	x	x	x	x	x	x	x	x
3.2. Organisationale Steuerung									
Grundfinanzierung (Anteile, Leistungsabhängigkeit)	x k	x		x		x		x	
Erhebung FI (Umfang & Art der Erhebung)	o Msplit/B			x	x				
Institutionelle Erhebung (Aufwand, Zwecke, Einschätzung)	o Msplit/B			x	x				
Kernsatzsatz Forschung (Kenntnis)	o Msplit/B			x	x				
4. Publikationen und Wissenstransfer									
4.1. Publikationsverhalten und Begutachtung									
Publikationsoutput/-formate (letzte 5J.)	o k	x	x	x	x	x	x	x	x
Kriterien Fachzeitschrift	(x) k	x	x	x	x	x	x	x	x
Open Access (Anteil Publikationen, Einstellung zu)	o k	x	x	x	x	x	x	x	x
Publikationsformate (Einschätzung)	o Msplit/C					x	x		
Qualität der Manuskriptbegutachtung (letzte 5J.)	x k	x	x	x	x	x	x	x	x
Eigene Gutachtertigkeit	k	x	x	x	x	x	x	x	x
4.2. Autorschaft									
Predatory Journals (E-Mail-Anfragen, Umgang, Erkennung)	o Msplit/D							x	x
Predatory Journals (Einschätzung)	o Msplit/D							x	x
Autorschaftsansprüche (Tätigkeiten)	(x) Msplit/D							x	x
Autorschaftsansprüche (weitere Aspekte)	o Msplit/D							x	x
Ablehnung Autorschaft									
Anonym Publizieren	o Msplit/D							x	x
Verfassen wiss. Publikationen (Typologie, Unterstützung)	o Msplit/D							x	x
Bewertung Zusammenarbeit mit Scientific Writers	o Msplit/D							x	x
4.3. Wissenstransfer									
Gruppen außerhalb Wiss. (Zusammenarbeit, Austausch)	o k	x	x	x	x	x	x	x	x
Erfahrungen (Disziplinenwechsel, Tätigkeit außerhalb Wiss.)	o k	x	x	x	x	x	x	x	x
Eig. Forschung für Gruppen (Relevanz, Kenntnis)	o Msplit/C					x	x		
Nichtwiss. Publikationen (Transfer-/Praktikerformate)	Msplit/C					x	x		
Digitale Plattformen zur Wissenschaftskommunikation	o k	x	x	x	x	x	x	x	x
Mendeleynutzung	o gefiltert	x	x	x	x	x	x	x	x
5. Wissenschaftspolitische Einstellungen									
Evaluation Wiss-system (Materiell, Ideell)	x k	x	x	x	x	x	x	x	x
Konkrete Reformvorschläge	o k	x	x	x	x	x	x	x	x
Promotionsrecht an Fachhochschulen	(x) k	x	x	x	x	x	x	x	x
Bewertung d. Exzellenzstrategie (ExIn)	x k	x	x	x	x	x	x	x	x
Open Data Diskurs (Kenntnis, Einschätzung)	o Msplit/A	x	x						
6. Demographie									
Geschlecht, Geburtsjahr	x k	x	x	x	x	x	x	x	x
Gleichstellung AGG - Diskriminierungserfahrungen	o k	x	x	x	x	x	x	x	x
Gleiche Behandlung Männer Frauen (Bayern)	o k	x	x	x	x	x	x	x	x
Staatsbürgerschaft	x k	x	x	x	x	x	x	x	x
Partnerschaft (zusammen/ getrennt lebend)	x k	x	x	x	x	x	x	x	x
Doppelter Haushalt/ Fernpendeln	o k	x	x	x	x	x	x	x	x
Kinder (Anzahl, Alter)	x k	x	x	x	x	x	x	x	x
Work-to-Family Conflict	o k	x	x	x	x	x	x	x	x

4 Hinweise zum Scientific Use File

4.1 Allgemeine Datennutzungshinweise

[Voraussetzungen der Datennutzung] Die Daten der Wissenschaftsbefragung 2019 werden durch das FDZ des DZHW entsprechend der europäischen Datenschutzgrundverordnung (EU-DSGVO) anonymisiert bereitgestellt und ausschließlich zur wissenschaftlichen Nutzung freigegeben.⁷ Das FDZ bietet ein *Scientific Use File* (SUF) für die wissenschaftliche Sekundärnutzung an.

Das SUF kann über die Website des FDZ (<https://fdz.dzhw.eu>) beantragt werden. Voraussetzung ist die Anstellung der Datennutzerin/des Datennutzers an einer wissenschaftlichen Einrichtung. Studierende oder Promovierende ohne eine Anstellung an einer wissenschaftlichen Einrichtung müssen gemeinsam mit einer betreuenden Mitarbeiterin bzw. einem betreuenden Mitarbeiter einen Datennutzungsantrag stellen. Für die Nutzung des SUF wird ein Datennutzungsvertrag abgeschlossen.

[Datenzugang] Über den *Digital Object Identifier* (DOI) [10.21249/DZHW:scs2019:2.0.0](https://doi.org/10.21249/DZHW:scs2019:2.0.0) ist eine Website mit zentralen Informationen zur Studie sowie weiteren Dokumentationsmaterialien erreichbar.

Das SUF wird über zwei Zugangswege angeboten, die hinsichtlich des Speicherortes, der Möglichkeit der eigenständigen Verknüpfung mit externen Daten und der Kontrollmöglichkeiten des FDZ unterschiedlich restriktiv sind.

- **Remote-Desktop:** Die Daten werden auf einem Terminal-Server des FDZ bereitgestellt. Über eine besonders gesicherte Verbindung zwischen dem lokalen Computer der nutzenden Person und dem Terminal-Server des FDZ können die Daten mit der auf dem Terminal-Server vorhandenen Software analysiert werden. Das Übertragen der Daten auf den lokalen Computer ist nicht möglich. Analyseergebnisse werden erst nach einer Prüfung auf datenschutzrechtliche Unbedenklichkeit durch das FDZ freigegeben und zur Verfügung gestellt.
- **On-Site:** Die Daten werden in den Räumlichkeiten des FDZ in einer kontrollierten Umgebung an einem speziell gesicherten Computer zur Analyse bereitgestellt. Wie beim Remote-Desktop-Zugang werden Analyseergebnisse erst nach einer Prüfung auf datenschutzrechtliche Unbedenklichkeit durch das FDZ freigegeben und zur Verfügung gestellt.

⁷ Das Datenschutzkonzept des FDZ ist angelehnt an den Portfolio-Ansatz von Lane, Heus und Mulcahy (2008, S. 6 ff.), an dem sich bereits das Leibniz-Institut für Bildungsverläufe (LifBi) (vgl. Koberg, 2016, 699ff.) und das FDZ der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung (vgl. Hochfellner, Müller, Schmucker und Roß, 2012, S. 9 f.) orientieren. Das FDZ des DZHW hat diesen Ansatz an die Anforderungen der eigenen Datenbestände angepasst und nutzt vier Kategorien von Maßnahmen zur Sicherstellung des Datenschutzes, die in unterschiedlicher Weise kombiniert werden können: Rechtlich-institutionelle Maßnahmen, informationelle Maßnahmen, technische Maßnahmen und statistische Maßnahmen.

Die bereitgestellten Daten weisen im Falle der Wissenschaftsbefragung 2019 für beide Zugangswege einen identischen Informationsgehalt auf. Die technischen und organisatorischen Kontrollmaßnahmen über den Zugangsweg On-Site ermöglichen es jedoch, auch Mitarbeiter*innen der Deutschen Forschungsgemeinschaft (DFG) einen Datenzugang zu gewähren (siehe Beschränkung von Nutzergruppen). Mit diesem Vorgehen wird ein Höchstmaß an Nutzbarkeit und gleichzeitig ein bestmöglicher Schutz der bereitgestellten Daten sichergestellt.

[Beschränkung von Nutzergruppen] Personen, die zum Zeitpunkt des Abschlusses des Datennutzungsvertrages bei der DFG beschäftigt sind oder im Verlauf der Datennutzung ein Beschäftigungsverhältnis bei der DFG eingehen, dürfen die bereitgestellten Daten nur über den Zugangsweg On-Site nutzen. Dies liegt darin begründet, dass die Mitarbeiter*innen der DFG über Datenbankzugänge verfügen, die, in Verbindung mit den Individualdaten der Wissenschaftsbefragung 2019, ein erhöhtes De-Anonymisierungspotenzial darstellen. Eine Bereitstellung über Remote-Desktop ist im Sinne der Wahrung der Anonymität der Befragungsteilnehmer*innen in diesem Fall zu verwehren. Während der Laufzeit des Datennutzungsvertrages neu entstandene Beschäftigungsverhältnisse der Datennutzer*innen bei der DFG sind dem FDZ unverzüglich mitzuteilen. Die Nutzung des Datenzuganges Remote-Desktop ist ab dem ersten Tag eines Beschäftigungsverhältnisses bei der DFG nicht mehr zulässig.

[Gebühren der Datenbereitstellung] SUF werden derzeit (Stand: Februar 2022) kostenfrei zur Verfügung gestellt. Änderungen bzw. die aktuelle Gebührenordnung können auf der Website des FDZ (<https://fdz.dzhw.eu>) eingesehen werden.

[Pflichten der Datennutzer*innen] Die Datennutzer*innen sind verpflichtet, folgende Regeln⁸ einzuhalten:

- **Wissenschaftliche Nutzung:** Die Daten dürfen ausschließlich für wissenschaftliche Zwecke verwendet werden. Eine kommerzielle Nutzung ist untersagt.
- **De-Anonymisierungsverbot:** Jeder Versuch der Re-Identifikation von Analyseeinheiten (z. B. Personen, Haushalten, Institutionen) ist verboten.
- **Gebot zur Mitteilung von Sicherheitslücken:** Falls Datennutzer*innen Kenntnis von Sicherheitslücken hinsichtlich Datenschutz bzw. Datensicherheit erlangen, müssen diese dem FDZ unverzüglich angezeigt werden.
- **Keine Weitergabe der Daten:** Die Daten dürfen nur durch die Personen genutzt werden, die einen Datennutzungsvertrag abgeschlossen haben.
- **Bereitstellung/Meldung von Publikationen:** Jede Art von wissenschaftlichen Arbeiten (z.B. Publikationen, Abschlussarbeiten, Dissertationen), die aus der Arbeit mit Daten des FDZ hervorgehen, ist dem FDZ anzuzeigen. Dabei ist dem FDZ eine elektronische Version der Druckfassung zur Verfügung zu stellen.
- **Zitationspflicht:** Die verwendeten Daten müssen in Veröffentlichungen, anderen Arbeiten (z.B. Abschlussarbeiten) und Vorträgen laut der Vorgaben des FDZ zitiert werden.

4.2 Anonymisierung

[Datenschutzrechtlicher Rahmen] Für personenbezogene Daten⁹, die in freiwilligen Befragungen durch das DZHW erhoben werden, gelten die EU-Datenschutz-Grundverordnung (EU-DSGVO) und

⁸ Der Datennutzungsvertrag regelt die Nutzungsbedingungen im Detail.

das Bundesdatenschutzgesetz in seiner Neufassung vom 30. Juni 2017.¹⁰ Danach sind personenbezogene Daten für die Weitergabe zur wissenschaftlichen Sekundärnutzung (ohne Vorliegen einer Einverständniserklärung zur Sekundärnutzung der personenbezogenen Daten) in der Regel derart aufzubereiten, dass „die personenbezogenen Daten ohne Hinzuziehung zusätzlicher Informationen nicht mehr einer spezifischen betroffenen Person zugeordnet werden können, sofern diese zusätzlichen Informationen gesondert aufbewahrt werden und technischen und organisatorischen Maßnahmen unterliegen, die gewährleisten, dass die personenbezogenen Daten nicht einer identifizierten oder identifizierbaren natürlichen Person zugewiesen werden können“ (Art. 4 Abs. 5 DSGVO; s. auch Art. 89 DSGVO sowie Erwägungsgrund 26 DSGVO). Das heißt, für die Weitergabe von Daten aus wissenschaftlichen Forschungsprojekten an Dritte sind die Daten derart zu anonymisieren, dass kein Bezug zur Person mehr hergestellt werden kann.

[Statistische Anonymisierungsmaßnahmen] Die Anonymität der Befragten wird über eine Kombination aus technischen Zugriffsbeschränkungen (vgl. Kapitel 4.1) und statistischen Anonymisierungsmaßnahmen und sichergestellt.

Im Rahmen der statistischen Anonymisierungsmaßnahmen wurden zunächst alle Informationen darauf geprüft, ob sich über diese Personen oder Institutionen direkt identifizieren lassen. Diese sogenannten *direkten Identifikatoren* wurden in der Befragung nicht erhoben, entsprechend sind sie auch nicht im SUF der Wissenschaftsbefragung 2019 enthalten.

Anschließend wurden die *Quasi-Identifikatoren* bestimmt, also Informationen, die in Kombination oder durch die Anspielung externer Informationen geeignet sind, eine Person indirekt zu identifizieren. Um eine eindeutige Zuordnung der Daten zu unterbinden, wurden diese Schlüsselmerkmale aggregiert oder gar nicht freigegeben. Zentrale Quasi-Identifikatoren in der Wissenschaftsbefragung 2019 sind beispielsweise vertiefende Angaben rund um die wissenschaftliche Position.

Zuletzt wurde geprüft, ob in den Daten *sensible Informationen*, z. B. zur Gesundheit, sexuellen Orientierung und zu politischen Einstellungen, enthalten waren. Diese eignen sich zwar nicht notwendig zur Re-Identifikation von Individuen oder Institutionen, jedoch können die Informationen im Falle einer De-Anonymisierung nutzbringend sein (vgl. Koberg, 2016, S. 694) und sind daher besonders schützenswert (Art. 9 DSGVO, Erwägungsgrund 51 DSGVO). In der Wissenschaftsbefragung 2019 wurden Informationen zu persönlichen Diskriminierungserfahrungen erhoben, für die bei den Befragten kein zusätzliches Einverständnis für die Sekundärnutzung eingeholt wurde. Daher wurden diese Antworten gelöscht.

⁹ „Personenbezogene Daten (sind) alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person (im Folgenden „betroffene Person“) beziehen; als identifizierbar wird eine natürliche Person angesehen, die direkt oder indirekt, insbesondere mittels Zuordnung zu einer Kennung wie einem Namen, zu einer Kennnummer, zu Standortdaten, zu einer Online-Kennung oder zu einem oder mehreren besonderen Merkmalen identifiziert werden kann, die Ausdruck der physischen, physiologischen, genetischen, psychischen, wirtschaftlichen, kulturellen oder sozialen Identität dieser natürlichen Person sind“ (Art. 4 DSGVO, S. 1).

¹⁰ Die DSGVO gilt grundsätzlich innerhalb der EU und somit ebenfalls für das DZHW. Das BDSG in seiner Neufassung vom 30. Juni 2017 (Gesetz zur Anpassung des Datenschutzrechts an die Verordnung (EU) 2016/679 und zur Umsetzung der Richtlinie (EU) 2016/680 (Datenschutz-Anpassungs- und Umsetzungsgesetz EU DSAnpUG-EU)) kommt teils zusätzlich zur Anwendung, da die DZHW GmbH juristisch als öffentliche Stelle des Bundes betrachtet wird (§ 2 Abs. 3 BDSG). Der Bund hält die absolute Mehrheit der Anteile der DZHW GmbH und das Institut erfüllt Aufgaben der öffentlichen Verwaltung des Bundes im weitesten Sinn.

Darüber hinaus empfehlen Ebel und Meyermann, offene Angaben zu löschen „selbst wenn die jeweiligen Fragestellungen an sich unproblematisch sind. Denn es besteht die Gefahr, dass Studienteilnehmer*innen bei eigentlich unbedenklichen Fragen mit offener Antwortmöglichkeit kritische Informationen preisgegeben haben, die zu einer Identifikation führen könnten“ (Ebel & Meyermann, 2015, S. 5). Nicht codierte offene Angaben wurden daher gelöscht.

Die nachfolgende Tabelle 9 stellt in Kurzform die durchgeführten statistischen Anonymisierungsmaßnahmen dar. Genauere Informationen zur Generierung der einzelnen Variablen sind im Datensatzreport zu finden.

Tabelle 9. Maßnahmen der statistischen Anonymisierung nach Zugangsweg

Merkmal	Variablen	Anonymisierungsmaßnahme im SUF
WISSENSCHAFTLICHE BIOGRAPHIE		
Andere wissenschaftliche Einrichtung (offene Nennung)	wb1do	Löschung
Beschäftigungsposition/ Promotion	wb2 wb2o wb5	Löschung, aber: Aggregation zu den Kategorien: - Professor*in - Postdoc - Prädoc (neue Variable: stgroup)
Art der Professur	wb3	Löschung
Andere Besoldungsstufe der Professur (offene Nennung)	wb4o	Löschung
Jahr des Absolvierens einer Etappe des akademischen Lebenslaufs	wb6a1 wb6b1 wb6c1	Löschung, aber: Aggregation zum „akademischen Alter bezogen auf den Hochschulabschluss (nur Nicht-Promovierte)“ (= Jahre seit dem Hochschulabschluss) - bis zu 1 Jahr - 2 bis 3 Jahre - 4 bis 5 Jahre - 6 bis 7 Jahre - 8 bis 9 Jahre - 10 Jahre und mehr (neue Variable: aage_hsa) Aggregation zum „akademischen Alter bezogen auf die Promotion“ (= Jahre seit der Promotion) - bis 2 Jahre - 3 bis 5 Jahre - 6 bis 10 Jahre - 11 bis 15 Jahre - 16 bis 20 Jahre - 21 bis 30 Jahre

Merkmal	Variablen	Anonymisierungsmaßnahme im SUF
		- 31 Jahre und mehr (neue Variable: aage_prom)
		Aggregation zu den Kategorien: - 1991 und früher - 2019/2020 (Rest einzeln ausgewiesen) (neue Variable: wb6c1_r)
Land des Absolvierens einer Etappe des akademischen Lebenslaufs	wb6a2 wb6b2 wb6c2	Löschung, aber: Aggregation zu den Kategorien: - Deutschland - EU - Europa ohne EU - Sonstiges (neue Variablen: wb6a2_r, wb6b2_r, wb6c2_r)
Fachgebiet	wb8 wb8o	Löschung, aber: Aggregation zu DFG-Fachgebieten (5 bzw. 10 Kategorien) sowie DFG- Wissenschaftsbereiche (4 Kategorien) (neue Variablen: dfg4, dfg5, dfg10)
Andere Position als langfristiges Karriereziel (offene Nennung)	wb9o	Löschung
Funktionsrolle bei Organen der DFG	wb10c	Löschung
FORSCHUNGSBEDINGUNGEN		
Gesamtlaufzeit befristeter Vertrag	fb4	Löschung, aber: Aggregation zu den Kategorien: - mehr als 4 bis unter 5 Jahre - 5 bis unter 6 Jahre - 6 und mehr Jahre (Rest einzeln ausgewiesen) (neue Variable: fb4_r)
Sonstiger Befristungsgrund (offene Nennung)	fb5o	Löschung
Anzahl befristeter Arbeitsver- träge an wissenschaftlichen Einrichtungen	fb6	Löschung, aber: Top-Codierung für Anzahl>20 (neue Variable: fb6_r)
Vertraglich vereinbarte Wochenarbeitszeit	fb7	Löschung, aber: Rundung zu ganzen Stunden (neue Variable: fb7_r)
Tatsächliche durchschnittli- che Wochenarbeitszeit	fb8	Löschung, aber: Rundung zu ganzen Stunden (neue Variable: fb8_r)
Sonstige Tätigkeit mit Arbeitszeitanteil (offene Nennung)	fb10ho	Löschung

Merkmal	Variablen	Anonymisierungsmaßnahme im SUF
FACHGEBIETSEINSCHÄTZUNGEN		
-	-	-
FORSCHUNGSPRAXIS		
Genutzte Typen von Daten, die von Dritten bereitgestellt wurden	fp3a1_ma fp3a2_ma	Löschung
Selbst erhobene/produzierte Typen von Daten	fp6a1_ma fp6a2_ma	Löschung
WISSENSCHAFTLICHER NACHWUCHS		
Anzahl betreuter Promovierender (allgemein und als Erstbetreuer*in)	wn1 wn2	Löschung, aber: Top-Codierung für Anzahl>10 (neue Variablen: wn1_r, wn2_r)
FORSCHUNGSFÖRDERUNG		
Anzahl eingereichter Anträge	ff1a ff1b ff1c ff1d ff1e ff1f ff1g	Löschung, aber: Aggregation bei „DFG“ und „BMBF“ zu den Kategorien: - 4 bis 10 - mehr als 10 (Rest einzeln ausgewiesen) (neue Variable: ff1a_r, ff1c_r) bzw. Aggregation bei „Europ. Union“, „Nat. Stiftungen“, „Weitere in Industrie/Wirtschaft“ und „Sonstige“ zu den Kategorien: - 2 bis 10 - mehr als 10 (Rest einzeln ausgewiesen) (neue Variable: ff1d_r, ff1e_r, ff1f_r, ff1g_r) bzw. Top-Codierung für Anzahl>2 bei „Exzellenzstrategie“ (neue Variable: ff1b_r)
ORGANISATIONALE STEUERUNG		
Andere Leistungskriterien, von denen Grundmittel für Forschung abhängig sind (offene Nennung)	os2ho	Löschung
Kenntnis Kerndatensatz Forschung	os10_mb	Löschung, aber: Aggregation zu den Kategorien: - Ja, aber ich kenne nur den Begriff - Ja, und ich habe ein grobes Ver-

Merkmal	Variablen	Anonymisierungsmaßnahme im SUF
		<p>ständnis von seinen Inhalten bzw. kenne mich gut mit seinen Inhalten aus</p> <ul style="list-style-type: none"> - Nein, ist mir nicht bekannt <p>(neue Variable: os10_mb_r)</p>
Quelle Kenntnis Kerndatensatz Forschung	os11a_mb os11b_mb os11c_mb os11d_mb os11e_mb os11f_mb os11g_mb os11h_mb os11ho_mb	Löschung
PUBLIKATIONSVERHALTEN		
Anzahl Publikationen	pu1a pu1b pu1c pu1d pu1e pu1f pu1g pu1h	<p>Löschung, aber:</p> <p>Aggregation bei „Originalartikel“ zu den Kategorien:</p> <ul style="list-style-type: none"> - 5 bis 19 - 20 bis 34 - 35 bis 49 - 50 bis 64 - 65 bis 79 - mehr als 79 <p>(Rest einzeln ausgewiesen)</p> <p>(neue Variable: pu1a_r)</p> <p>bzw.</p> <p>Aggregation bei „Überblicksartikel“ zu den Kategorien:</p> <ul style="list-style-type: none"> - 4 bis 10 - mehr als 10 <p>(Rest einzeln ausgewiesen)</p> <p>(neue Variable: pu1b_r)</p> <p>bzw.</p> <p>Top-Codierung für Anzahl>1 bei „Monografien“</p> <p>(neue Variablen: pu1c_r)</p> <p>bzw.</p> <p>Aggregation bei „Buchbeiträge“ zu den Kategorien:</p> <ul style="list-style-type: none"> - 3 bis 5 - 6 bis 10

Merkmal	Variablen	Anonymisierungsmaßnahme im SUF
		<p>- mehr als 10 (Rest einzeln ausgewiesen) (neue Variable: pu1d_r)</p> <p>bzw.</p> <p>Top-Codierung für Anzahl>2 bei „Special Issues“, „Kommentare“, „Editorials“ und „Handbücher“ (neue Variablen: pu1e_r, pu1f_r, pu1g_r, pu1h_r)</p>
Anzahl Gutachtertätigkeit	pu8a pu8b pu8c pu8d pu8e pu8f	<p>Löschung, aber:</p> <p>Aggregation bei „Zeitschriftenartikel“ zu den Kategorien:</p> <ul style="list-style-type: none"> - 11 bis 15 - 16 bis 20 - 21 bis 30 - mehr als 30 <p>(Rest einzeln ausgewiesen) (neue Variable: pu8a_r)</p> <p>bzw.</p> <p>Aggregation bei „Förderanträge“ zu den Kategorien:</p> <ul style="list-style-type: none"> - 3 bis 10 - 11 bis 19 - mehr als 19 <p>(Rest einzeln ausgewiesen) (neue Variable: pu8b_r)</p> <p>bzw.</p> <p>Aggregation bei „Mitwirkung an Evaluationsverfahren“ zu den Kategorien:</p> <ul style="list-style-type: none"> - 2 bis 4 - mehr als 4 <p>(Rest einzeln ausgewiesen) (neue Variable: pu8c_r)</p> <p>bzw.</p> <p>Top-Codierung für Anzahl>2 bei „Mitwirkung an Akkreditierungsverfahren“ (neue Variablen: pu8d_r)</p> <p>bzw.</p>

Merkmal	Variablen	Anonymisierungsmaßnahme im SUF
		Top-Codierung für Anzahl>3 bei „Berufungsverfahren“ (neue Variablen: pu8e_r)
		bzw.
		Aggregation bei „Promotionen“ zu den Kategorien: - 3 bis 10 - mehr als 10 (Rest einzeln ausgewiesen) (neue Variable: pu8f_r)
AUTORSCHAFT		
Recherchequelle bzgl. Identifikation von „Predatory-Journal“/„Fake Journal“ (offene Nennung)	au3do_md	Löschung
Gegenstand des Autorschaftskonflikts (offene Nennung)	au8_md	Löschung
Gründe für Ablehnung einer Autorschaft (offene Nennung)	au10_md	Löschung
Aufgaben beim Verfassen von Publikationen, bei denen professionelle Unterstützung in Anspruch genommen wird	au12a_md au12b_md au12c_md au12d_md au12e_md au12f_md	Löschung
Erfahrung mit Scientific Writern	au13_md	Löschung
WISSENSTRANSFER		
Sonstige Personen/Gruppen außerhalb der Wissenschaft für Austausch (offene Nennung)	wt2go	Löschung
Sonstiger Bereich außerhalb der Wissenschaft, für den eigene Forschung relevant ist (offene Nennung)	wt5ho_mc	Löschung
Sonstige gesellschaftliche Gruppen, die Forschungsergebnisse nutzen bzw. zur Kenntnis nehmen (offene Nennung)	wt6io_mc	Löschung
Sonstige Plattformen hinsichtlich Nutzung für wissenschaftliche Zwecke	wt8eo	Löschung

Merkmal	Variablen	Anonymisierungsmaßnahme im SUF
(offene Nennung)		
Nutzungszweck Mendeley (offene Nennung)	wt9do	Löschung
Grund Artikelaufnahme in Mendeley Library	wt10a wt10b wt10c wt10d wt10e wt10eo	Löschung
WISSENSCHAFTSPOLITIK		
-	-	-
DEMOGRAFIE		
Geschlecht	de1	Löschung, aber: Recodierung der Kategorie „divers“ zu Missing (neue Variablen: de1_r)
Geburtsjahr	de2	Löschung
Persönliche Diskriminierungserfahrung	de3a de3b de3c de3d de3e de3f de3g de3h	Löschung
Staatsbürgerschaft	de5a de5b	(Erste) Staatsbürgerschaft: Löschung, aber: Aggregation zu den Kategorien: <ul style="list-style-type: none"> - Deutschland - EU - Europa ohne EU - Sonstiges (neue Variable: de5a_r) Zweite Staatsbürgerschaft: Löschung
Doppelter Haushalt	de7a de7b de7c	Löschung, aber: Aggregation zu den Kategorien: <ul style="list-style-type: none"> - ja (doppelter Haushalt bei befragter Person und/oder Partner*in vorhanden) - nein (neue Variable: de7_r)
Pendeldistanz	de9	Löschung, aber: Aggregation zu den Kategorien: <ul style="list-style-type: none"> - kein Pendeln - Nahpendler bis 50km - Nahpendler 51-150km

Merkmal	Variablen	Anonymisierungsmaßnahme im SUF
		<ul style="list-style-type: none"> - Fernpendler 151-250km - Fernpendler ab 251km (neue Variable: de9_r)
Anzahl Kinder	de10	Löschung, aber: Aggregation bei Anzahl>2 (neue Variable: de10_r)
Alter Kinder	de11a	Löschung, aber: Aggregation zu den Kategorien: <ul style="list-style-type: none"> - keine Kinder oder keine Altersangabe - mind. ein Kind 12 oder jünger - kein Kind 12 oder jünger (neue Variable: de11_r)
	de11b	
	de11c	
	de11d	
	de11e	
	de11f	

Literaturverzeichnis

- The American Association for Public Opinion Research (AAPOR) (2016). Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 9th edition. [https://www.aapor.org/Standards-Ethics/Standard-Definitions-\(1\).aspx](https://www.aapor.org/Standards-Ethics/Standard-Definitions-(1).aspx), zugegriffen am 19. Mai 2020.
- Beigang, S., Fetz, K., Kalkum, D., & Otto, M. (2017): Diskriminierungserfahrungen in Deutschland. Ergebnisse einer Repräsentativ- und einer Betroffenenbefragung. Baden-Baden: Nomos.
- Böhmer, S., Neufeld, J., Hinze, S., Klode, C., & Hornbostel, S. (2011). Wissenschaftlerbefragung 2010: Forschungsbedingungen von Professorinnen und Professoren an deutschen Universitäten. IfQ Working Paper, 8.
- Ebel, T. & Meyermann, A. (2015). Hinweise zur Anonymisierung von quantitativen Daten. Forschungsdaten Bildung informiert. Bd.3. Frankfurt am Main: Deutsches Institut für Internationale Pädagogische Forschung.
- Gabler, S. und Häder, S. (2015). Stichproben in der Theorie. Mannheim, GESIS – Leibniz-Institut für Sozialwissenschaften (SDM Survey Guidelines). DOI: 10.15465/sdm-sg_008
- Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly*, 73(2), 349–360. <https://doi.org/10.1093/poq/nfp031>
- Hdioud, B. & Tirari, M.H. (2019). Criteria for choosing between calibration weighting and survey weighting. *Survey Methodology*. 44. 339-349.
- Hochfellner, D., Müller, D., Schmucker, A. & Roß, E. (2012). Datenschutz am Forschungsdatenzentrum (FDZ-Methodenreport Nr. 6). Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung (IAB).
- Koberg, T. (2016). Disclosing the National Educational Panel Study. In H.-P. Blossfeld, J. v. Maurice, M. Bayer & J. Skopek (Hrsg.), *Methodological Issues of Longitudinal Surveys. The example of the National Educational Panel Study* (S. 691–708). Wiesbaden: Springer VS. doi:10.1007/978-3-658-11994-2
- Kroh, M., Siegers, R., & Kühne, S. (2015). Gewichtung und Integration von Auffrischungstichproben am Beispiel des Sozio-oekonomischen Panels (SOEP). In Schupp&Wolf (2015). *Nonresponse Bias*: 409-444.
- Lane, J., Heus, P. & Mulcahy, T. (2008). Data access in a cyber world: Making use of cyberinfrastructure. *Transactions on Data Privacy*, 1(1), 2–16.
- Neufeld, J., Johann, D. (2018). Wissenschaftlerbefragung 2016. *Methoden- und Datenbericht*. (Version 1.0.0). Hannover: fdz.DZHW.
- Hofmann, Y. E., & Kanamüller, A. (2019). Wie zufrieden sind die Professorinnen und Professoren an den Universitäten Bayerns? Eine fächerspezifische Standortbestimmung auf Basis der landesweiten Professorenbefragung ProfQuest. *Beiträge Zur Hochschulforschung*, (2), 84–95.
- Schupp, J. & Wolf, C. (2015). *Nonresponse Bias: Qualitätssicherung sozialwissenschaftlicher Umfragen*. Springer Fachmedien Wiesbaden. 10.1007/978-3-658-10459-7.

Anhang A - Einladungsschreiben

[Betreff: DZHW-Studie: Wie steht es um die Wissenschaft in Deutschland?]

[anrede],

hiermit laden wir Sie herzlich zur DZHW-Wissenschaftsbefragung 2019 ein, dem großen deutschen Monitoring von der Wissenschaft für die Wissenschaft.

Der Wissenschaftsrat, die Politik und die Medien greifen Ergebnisse des Deutschen Zentrums für Hochschul- und Wissenschaftsforschung (DZHW) in aktuellen wissenschaftspolitischen Debatten auf. Ihre Erfahrungen und Einschätzungen, sowie die Ihrer Kolleginnen und Kollegen bilden die empirische Grundlage für eine umfangreiche Bestandsaufnahme der Arbeits- und Forschungsbedingungen in Deutschland.

Die Befragung dauert 25 Minuten und umfasst ein breites Spektrum wichtiger Themen. Zur Befragung gelangen Sie über den folgenden personalisierten Link:

[https://www.wb.dzhw.eu/current?p=\[PIN\]](https://www.wb.dzhw.eu/current?p=[PIN])

Für die Hochschul- und Wissenschaftsforschung ist Ihre Einschätzung von großem Wert. Wir würden uns daher sehr über Ihre Teilnahme freuen!

Mit herzlichem Dank für Ihre Unterstützung und freundlichen Grüßen

Dr. Jens Ambrasat
Projektleiter Wissenschaftsbefragung

Prof. Dr. Stefan Hornbostel
Leitung der Abteilung 2 – Forschungssystem und Wissenschaftsdynamik
Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW)
Schützenstraße 6a
D-10117 Berlin

Hinweise zum Datenschutz

Unsere Datenerhebung erfolgt im Einklang mit den Vorgaben der europäischen Datenschutzgrundverordnung (DSGVO). Ihre Teilnahme ist selbstverständlich freiwillig. Ihre Angaben werden ausschließlich zu Forschungszwecken genutzt und zuerst anonymisiert, bevor sie für wissenschaftliche Auswertungen zur Verfügung gestellt werden. Ausführliche Informationen zum Datenschutz und den Ihnen zustehenden Rechten finden Sie unter: https://www.wb.dzhw.eu/pdf/wiBef2019_Datenschutz.pdf
Datenschutzbeauftragter: Martin Fuchs (datenschutz@dzhw.eu)

Ergebnisse der Studie

Die zentralen Ergebnisse werden wir im kommenden Jahr als „Barometer für die Wissenschaft“ veröffentlichen. Weitere Ergebnisse unserer bisherigen Befragungswellen sind auf unserer Projekthomepage verfügbar:

<https://www.wb.dzhw.eu>

Für weitere Informationen stehen Ihnen die Mitarbeiter:
Jens Ambrasat & Christophe Heger unter wibef2019@dzhw.eu zur Verfügung

Anhang B - Übersicht der Trendfragen

V. 2019	Label 2019	V. 2016	V. 2009	Trendfile 09/16
wb2-wb2_r	Professur/Mittelbau	v1_1	position	v1_5a
wb3	Art der Professur	v1_2	professur	v1_5b
wb4	Besoldungsstufe	v1_3	besoldung	v1_6b
wb6c1	Erstberufung	v1_4	jahr_prof	v1_9
wb5	Sind Sie promoviert	v1_5	.	v1_7a
wb6b1	Jahr der ersten Promotion	v1_6	jahr_dr	v1_7b
wb6b2	Land der ersten Promotion	.	F6_1c	.
wb7	Habilitiert, ja/nein	v1_8	.	v1_8a
wb1a-wb1d	Art der Einrichtung	v1_10, v1_10_andere	a_dem_0_1_2_ inst_a_a	v1_10
wb8	Fachgebiet	v1_11	F014	v1_11a
fb2 –fb3	Stelle befristet	v1_13	.	.
wb9	Langfristiges Karriereziel	v1_25	.	.
fa1-fa11	Einschätzungen z. Fachgebiet	v4_2 bis v4_10	e_a1_lom_3_ 4_2_b	.
pu2a-pu2k	Auswahlkriterien Fachzeitschrift	v5_2	.	.
au5a_md- au6g_md	Autoren- schaft/Danksagungen	v5_5	D_Betr_3_6_2 _A	v4_1 - v4_6
de1	Geschlecht	v10_5	gender	v1_2
de2	Geburtsjahr	v10_7	gebjahr	v1_1
de5a	Staatsbürgerschaft, erste	v10_8a	staatsbürger- schaft	v1_3
de5b	Staatsbürgerschaft, zweite	v10_8a	staatsbürger- schaft	.
de6	Familienstand	v10_11	.	.
de6	Feste Partnerschaft	v10_12	.	.
de7b- de7c	Doppelte Haushaltsführung	v10_13 v10_14	.	.
de10	Kinder Ja/nein	v10_15	.	.
de11a- de11f	Anzahl Kinder	v10_16	.	.