



STIFTERVERBAND

METHODENBERICHT

Forschung und Entwicklung im Wirtschaftssektor: Bereitstellung statistischer Daten und Analysen für das Berichtsjahr 2021

Ein Bericht der Wissenschaftsstatistik gGmbH im Stifterverband für die deutsche Wissenschaft e.V. für das Bundesministerium für Bildung und Forschung

BEAUFTRAGT VOM



Bundesministerium
für Bildung
und Forschung



FORSCHUNG
UND
ENTWICKLUNG®

INHALT

Projektbeschreibung.....	4
Befragungsmenge.....	4
Adresspflege und Neuaufnahmen.....	4
Gegenmenge.....	5
Befragungsgruppen.....	6
Prioritätsklassen.....	6
Fragebogen.....	8
Konzeption und Inhalt.....	8
Versand und Erinnerungswellen.....	8
Rücklauf.....	9
Nichtteilnehmenden-Befragung.....	11
Datenerfassung und -bearbeitung.....	12
Umgang mit fehlenden Werten.....	12
Plausibilitätsprüfungen.....	12
Geheimhaltung.....	13
Regionalisierung.....	13
Auswertung.....	14
Datenlieferung und Publikationen.....	14
Projektergebnisse und Nutzen.....	14
Einheitsleistungen.....	15
Projektverantwortliche und Ansprechpartner.....	15
Literatur.....	15
Anlagen.....	15

TABELLENVERZEICHNIS

Tabelle 1: Gruppenzuordnung in der Vollerhebung 2021.....	6
Tabelle 2: Besetzung der Prioritätsklassen in der Vollerhebung 2021	7
Tabelle 3: Urversand und Erinnerungswellen nach Befragungsgruppen	8
Tabelle 4: Rücklauf nach Prioritätsklassen.....	9
Tabelle 5: Rücklauf nach Wirtschaftszweigen und Beschäftigtengrößenklasse inkl. Gegenmenge	10
Tabelle 6: Item-Nonresponse bei Kernindikatoren	12

Projektbeschreibung

Im Rahmen der Verordnung (EU) 2019/2152 des Europäischen Parlaments und des Rates vom 27. November 2019 ist Deutschland verpflichtet, über die Forschungs- und Entwicklungsaktivitäten (FuE) der Unternehmen zu berichten. Damit lassen sich die Innovationsleistungen und -potentiale verschiedener Nationen bestimmen und miteinander vergleichen. Die EU-Verordnung gibt dabei u.a. die zu erhebenden Variablen sowie das methodische Vorgehen vor. Als zentrale Orientierung dient das gemeinschaftlich von den OECD-Mitgliedsstaaten entwickelte Frascati-Handbuch (OECD, 2018).

Im Rahmen des Projekts „Erhebung statistischer Daten und Durchführung von Analysen zu Forschung und Entwicklung im Wirtschaftssektor in Deutschland“ ist die SV Wissenschaftsstatistik gGmbH vom Bundesministerium für Bildung und Forschung (BMBF) beauftragt, die relevanten statistischen Indikatoren zu FuE in Deutschland zu erheben (Förderkennzeichen 16P4283).

Die FuE-Kernindikatoren wurden für das Berichtsjahr 2021 im Jahr 2022 als Vollerhebung erhoben und im Jahr 2023 veröffentlicht. Bei Vollerhebungen werden alle Unternehmen in die Untersuchung einbezogen, von denen bekannt ist, dass Forschung und Entwicklung durchgeführt wird oder von denen dies mit hoher Wahrscheinlichkeit anzunehmen ist. Es handelt sich dabei um aktuell ca. 29.000 Unternehmen. Die Auswertung der Daten erfolgte in der Branchengliederung der Klassifikation der Wirtschaftszweige 2008 (WZ 2008), die der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev. 2. entspricht, sowie in Beschäftigtengrößenklassen und nach Forschungsintensitäten. Zum 30.06.2023 erfolgte die endgültige Bereitstellung der FuE-Daten, die dem BMBF, dem Statistischen Bundesamt und Eurostat zur Verfügung gestellt wurden.

Befragungsmenge

Die Zielgruppe der FuE-Erhebung sind alle FuE-aktiven Unternehmen in Deutschland mit mindestens einem Beschäftigten, unabhängig von der Branchenzugehörigkeit. Die Vollerhebung für das Berichtsjahr 2021 hat den Anspruch, die Grundgesamtheit der FuE-treibenden Unternehmen und die Institutionen der industriellen Gemeinschaftsforschung (IfG) in Deutschland zu erfassen.

Adresspflege und Neuaufnahmen

Da kein allgemeines Verzeichnis FuE-treibender Unternehmen in Deutschland zur Verfügung steht, ist ein zentraler Bestandteil der Datenerhebung ein Verzeichnis dieser Art zu schaffen und zu pflegen. Die Pflege der Unternehmensdaten schließt den jährlichen Abgleich des Adressbestandes mit der Unternehmensdatenbank MARKUS des Bureau van Dijk sowie die Berücksichtigung von Veränderungen in den Konzernverflechtungen der Unternehmen mit ein.

Es wird unterschieden zwischen Unternehmen, die eine FuE-Aktivität in vorherigen Erhebungen gemeldet haben, und solchen, bei denen ein entsprechender Verdacht besteht. Die zweite Gruppe der Unternehmen, bei denen Forschungs- und Entwicklungsaktivitäten anzunehmen sind, werden durch Schlüsselindikatoren aus verschiedenen Informationsquellen identifiziert. Diese sind vor allem:

- Förderdaten des Bundes (ZIM, Förderkatalog)
- veröffentlichte Informationen der Unternehmen aus Jahresabschlüssen und Geschäftsberichten
- Informationen aus Medien und kommerziellen Unternehmensdatenbanken (v.a. MARKUS)
- Mitgliedslisten von Wirtschaftsverbänden mit innovativem Tätigkeitsschwerpunkt

Im Gesamtadressbestand befinden sich zurzeit ca. 125.000 Adressen von Unternehmen, von denen etwa 29.000 Unternehmen zum aktiven Bestand gehören. D. h. diese Unternehmen sind wirtschaftlich aktiv und weisen mindestens diskontinuierlich FuE-Aktivitäten auf. Alle Unternehmen, für die ein Datensatz vorliegt, aber die nicht mehr existieren (zum Beispiel aufgrund einer Insolvenz oder nicht mehr FuE-aktiv sind) gehören zum passiven Bestand. Unternehmen, die innerhalb von sechs Jahren (drei Vollerhebungen) keine Rückmeldung zu FuE gegeben haben und für die gleichzeitig keine FuE-relevanten Sekundärdaten vorliegen, werden aus dem aktiven Bestand entfernt. Das Gleiche gilt für nicht mehr existierende Unternehmen ohne Daten.

Die Befragungsmenge der Vollerhebung 2021 ist die Menge aller FuE-aktiven Unternehmen und setzt sich aus Unternehmen des Bestandes und Neuaufnahmen zusammen. Eine Neuaufnahme von Unternehmen in den Adressbestand findet bei begründetem Verdacht statt, dass FuE-Aktivitäten stattfinden. Dazu werden Datenbestände aus Förderprogrammen und Verbänden sowie weitere öffentlich zugängliche Informationen recherchiert.

Außerdem werden seit 2013 in ausgewählten forschungsintensiven Branchen alle Unternehmen bestimmter Beschäftigtengrößenklassen befragt. Grundlage für die Auswahl der Adressen für die Sonderbefragung bildet hier der Adressbestand der Unternehmensdatenbank MARKUS des Bureau van Dijk. 2021 wurden alle Unternehmen der Wirtschaftszweige C 20 sowie C 21 (Pharma- und Chemiebranche) ab einer Beschäftigtenzahl von 50 befragt. Bereits im Adressbestand enthaltene Unternehmen wurden weiterhin auch mit geringerer Beschäftigtenanzahl abgefragt. Zusätzlich wurden alle Unternehmen des Verbands Forschender Arzneimittelhersteller (VfA) kontaktiert. Damit sollte sichergestellt werden, dass sämtliche FuE-aktiven Unternehmen in der Erhebung Berücksichtigung finden.

Gegenmenge

Um den Adressbestand der FuE-Erhebung kontinuierlich zu aktualisieren und bisher nicht erfasste Unternehmen zu identifizieren, die nicht Teil der üblichen Informationsquellen sind, wird bei jeder Vollerhebung eine Gegenmengenanalyse durchgeführt.

Grundlage für die Auswahl der Adressen für die Gegenmengenanalyse bildet der Adressbestand der MARKUS-Datenbank. Insgesamt wurden in der Erhebung 2021 1.997 Unternehmen für die Gegenmengenanalyse herangezogen, die aus den Kernbranchen Kraftfahrzeugbau, Maschinenbau, Elektro oder IuK stammen und zwischen 10 und 250 Beschäftigte haben. Um die Kosten für die Befragung möglichst gering zu halten, wurden die Unternehmen für die Gegenmengenanalyse nur im Urversand befragt.

Von den 1.997 befragten Unternehmen haben 112 Unternehmen einen Fragebogen zurückgeschickt (5,6 Prozent). Davon berichteten 56 Unternehmen eine FuE-Aktivität, die folglich in die Auswertung der Erhebung 2021 miteinbezogen wurde. Außerdem werden die neu identifizierten Unternehmen für künftige Erhebungen in den Adressbestand aufgenommen. Die FuE-Gesamtaufwendungen der 56 identifizierten FuE-treibenden Unternehmen belaufen sich auf etwa 12 Millionen Euro. Dies macht einen Anteil von etwa 0,012 Prozent an der FuE-Gesamtsumme aller befragten Unternehmen aus.

Da bei den Unternehmen in der Gegenmengenanalyse grundsätzlich kein Verdacht auf FuE-Aktivitäten besteht, zeigen die Ergebnisse, dass der bisher etablierte Adressbestand bereits eine sehr gute

Abdeckung erzielt und von hoher Qualität ist. Auf dieser Grundlage ist anzunehmen, dass keine systematische Untererfassung der FuE-Aktivitäten vorliegt.

Befragungsgruppen

Die Befragungsmenge wird in Gruppen eingeteilt, die sich aus bisherigem Meldeverhalten und der Zugehörigkeit in der FuE-Erhebung begründen. Es werden folgende Gruppen unterschieden:

- Gruppe A: Kerngruppe (mindestens eine valide Meldung 2017 – 2020 oder ein Förderhinweis)
- Gruppe B: kürzlich erfolgte Neuaufnahmen und diskontinuierlich forschende Unternehmen
- Gruppe C: aktuelle Neuaufnahmen der Erhebung 2021
- Gruppe D: Prüfgruppe
- Gruppe E: Institutionen für Gemeinschaftsforschung (IfG)
- Gruppe F: Gegenmenge (Unternehmen ohne FuE-Verdacht)

Tabelle 1: Gruppenzuordnung in der Vollerhebung 2021

Gruppe	Anzahl der Unternehmen	Anteil an der Befragungsmenge
A (Kerngruppe)	14.732	47,45 %
Sondergruppe	132	0,43 %
Aktuelle FuE-Meldung	5.699	18,36 %
Förderhinweis 5013263	8.901	28,67 %
B (kürzliche Neuaufnahme/diskontinuierlich forschend)	8.284	26,68 %
Neuaufnahmen 2017-2020	6.012	19,37 %
Diskontinuierlich forschende Unternehmen	2.272	7,32 %
C (aktuelle Neuaufnahmen)	3.536	11,39 %
Neuaufnahme Verbände	486	1,57 %
Neuaufnahme Förderdaten	3.050	9,82 %
D (Prüfgruppe)	2.223	7,16 %
Reaktivierte Unternehmen	482	1,55 %
Unternehmen der WZ Vollerhebung (C20, C21)	1.741	5,61 %
E (Institutionen für Gemeinschaftsforschung)	273	0,88 %
F (Gegenmenge)	1.997	6,43 %
Gesamt	31.045	100,00 %

Prioritätsklassen

Die Prioritätsklassen wurden 2013 erstmalig gebildet, um den Unternehmen Prioritäten für die Datenbearbeitung zuzuweisen. Es werden folgende hierarchische Prioritätsklassen unterschieden:

- Sehr hohe Priorität (SHP)
Die Klasse „sehr hohe Priorität“ umfasst die 100 FuE-stärksten Unternehmen in Deutschland auf Basis des Berichtsjahres 2019.
- Hohe Priorität (HP+)
In dieser Klasse sind Unternehmen, die 5 Millionen Euro oder mehr an FuE aufwenden und Teil der Versandgruppe A sind.
- Hohe Priorität (HP-)
In dieser Klasse sind Unternehmen, die weniger als 5 Millionen Euro an FuE aufwenden und Teil der Befragungsgruppe A sind. Des Weiteren sind Unternehmen aus den Befragungsgruppen B und C in dieser Klasse; es besteht ein FuE-Verdacht, der aber nicht abgesichert ist. Außerdem befinden sich in dieser Klasse die Institutionen für industrielle Gemeinschaftsforschung (Gruppe E).
- Mittlere Priorität (MP)
In dieser Klasse sind Unternehmen aus der Prüfgruppe (Gruppe D) und aus der Gegenmenge (Gruppe F). Es ist unsicher, ob die Unternehmen (noch) FuE betreiben.

Tabelle 2 zeigt die Besetzungen der Prioritätsklassen und die Anteile an der Befragungsmenge. Je nach Prioritätsklasse werden unterschiedliche Formen der Datenbearbeitung bei fehlenden Werten (Item-Nonresponse) und Beobachtungen (Unit-Nonresponse) angewendet. Je nach Art des Datenausfalls wurde entweder eine automatische Ergänzung oder eine manuelle Plausibilisierung und Ergänzung durchgeführt. Eine ausführliche Beschreibung der Vorgehensweise findet sich im Abschnitt zu Datenerfassung und -bearbeitung dieses Berichts.

Tabelle 2: Besetzung der Prioritätsklassen in der Vollerhebung 2021

Prioritätsklasse	Anzahl der Unternehmen	Anteil an der Befragungsmenge
Sehr hohe Priorität (SHP)	100	0,32 %
Hohe Priorität (HP+)	1.015	3,27 %
Hohe Priorität (HP-)	25.949	83,59 %
Mittlere Priorität (MP)	3.981	12,82 %
Gesamt	31.045	100,00 %

Fragebogen

Konzeption und Inhalt

Gemäß EU-Verordnung besteht eine Meldepflicht des vollständigen Variablensatzes mit allen Untergliederungen für die ungeraden Berichtsjahre. Daher erfasst die Vollerhebung 2021 den vollständigen Variablensatz.

Nachdem in den Vollerhebungen der Vorjahre ein Schwerpunkt auf FuE im Ausland gelegt worden ist, wurden in die Erhebung 2021 Fragen zu folgenden Schwerpunkten integriert:

- Bekanntheit der Forschungszulage
- Auswirkungen der Covid-19-Pandemie
- Stellenwert von Klimaschutz

Auf Fragen zur Altersstruktur, zum fachlichen Hintergrund und der Herkunft des FuE-Personals wurde zur Entlastung der befragten Unternehmen verzichtet, da sich eine Variation dieser Kennzahlen in einer kurzfristigen Betrachtung erfahrungsgemäß nicht erkennen lässt.

Versand und Erinnerungswellen

Die Unternehmen wurden in der Erhebung 2021 schriftlich kontaktiert. Das Anschreiben enthielt eine kurze Erläuterung der FuE-Erhebung sowie die Bitte um Mitwirkung an der Erhebung. Gleichzeitig beinhaltete das Anschreiben einen Online-Code, mit dem die Unternehmen einerseits ihren individuellen Fragebogen online ausfüllen oder andererseits den Fragebogen ausdrucken und papierbasiert ausfüllen konnten. Die Programmierung und das Hosting des Online-Fragebogens wurden von einem externen Dienstleister (Meap GmbH, Witten) durchgeführt.

Im Rahmen des Urversands wurden im April 2022 alle Unternehmen aus der Befragungsmenge kontaktiert. Zusätzlich gab es je nach Gruppenzugehörigkeit der Unternehmen bis zu zwei Erinnerungswellen (Tab. 3). Zwischen dem Urversand und den einzelnen Erinnerungen erfolgte in der FuE-Erhebung 2021 regelmäßig eine Meldeanalyse, die die zum Stichtag fehlenden Meldungen feststellte. Diese Menge wurde als Grundgesamtheit für die nächste Erinnerungswelle festgelegt. Darüber hinaus wurden im Rahmen einer telefonischen Nachfassaktion solche Unternehmen kontaktiert, die noch keine Meldung abgegeben haben und der höchsten Prioritätsklasse (SHP) zugeordnet werden, da diese einen signifikanten Einfluss auf die Gesamtergebnisse haben können.

Tabelle 3: Urversand und Erinnerungswellen nach Befragungsgruppen

Gruppe	Urversand	1. Erinnerung	2. Erinnerung
A	04.04.2022	13.05.2022	30.06.2022
B	04.04.2022	13.05.2022	30.06.2022
C	04.04.2022	13.05.2022	—
D	04.04.2022	13.05.2022	—
E	04.04.2022	13.05.2022	30.06.2022
F	04.04.2022	—	—

Rücklauf

Die bereinigte Rücklaufquote wird errechnet als Verhältnis der antwortenden Unternehmen zu der Netto-Grundgesamtheit (Anzahl der insgesamt befragten Unternehmen abzüglich 883 technischer Ausfälle und ohne die Gegenmenge¹). Dadurch ergibt sich für das Berichtsjahr 2021 eine bereinigte Rücklaufquote von 24 Prozent und liegt damit zwischen den Werten der beiden letzten Vollerhebungen (2019: 29 Prozent, 2017: 22 Prozent). Aufbauend auf den Rückmeldungen der Unternehmen konnten Informationen zu weiteren 677 Unternehmen aus Geschäftsberichten oder Jahresabschlüssen recherchiert werden. Außerdem sind Informationen Dritter wie Presseberichte, Verbandsberichte etc. ebenfalls wertvolle Informationen, die zur Durchführung einer Schätzung herangezogen werden.

Als technischer Ausfall sind Unternehmen zu werten, die

- bei der jeweiligen Muttergesellschaft erfasst wurden (Anzahl: 480),
- Konkurs oder Liquidation meldeten (Anzahl: 350),
- unbekannt verzogen sind (Anzahl: 14),
- mit einem anderen Unternehmen fusionierten (Anzahl: 22) oder
- sonstige technische Ausfälle sind (Anzahl: 17).

In der folgenden Tabelle 4 findet sich der Rücklauf aufgeschlüsselt nach Prioritätsklassen. In den höheren Prioritätsklassen ist der Rücklauf anteilig höher und die Quote der Ausfälle geringer.

Tabelle 4: Rücklauf nach Prioritätsklassen

Prioritätsklasse	Rücksendung	Absage	Technischer Ausfall	Keine Rückmeldung	Gesamt
Sehr hohe Priorität (SHP)	51	1	1	47	100
Hohe Priorität (HP+)	256	25	24	710	1.015
Hohe Priorität (HP-)	6.347	171	679	18.752	25.949
Mittlere Priorität (MP)	283	8	179	3511	3.981
Gesamt	6.937	205	883	23.020	31.045

Weiterhin findet sich in Tabelle 5 der Rücklauf aufgeschlüsselt nach Unternehmensgröße und Wirtschaftszweigklassifikation.

¹100 der 883 technischen Ausfällen entfallen dabei auf die Gegenmenge.

Tabelle 5: Rücklauf nach Wirtschaftszweigen und Beschäftigtengrößenklasse inkl. Gegenmenge

I. Wirtschaftsgliederung ¹⁾ II. Beschäftigten- größenklassen	Rückmeldungen der Unternehmen								
	Rücksendung		Absage		Technischer Ausfall		Keine Rückmeldung		Gesamt
	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl

I. NACH DER WIRTSCHAFTSGLIEDERUNG

A 01-03	46	33,1	4	2,9	2	1,4	87	62,6	139
B 05-09	12	19,0	0	0,0	1	1,6	50	79,4	63
C 10-33	3.682	25,5	116	0,8	384	2,7	10.276	71,1	14.452
10-12	78	24,4	3	0,9	2	0,6	237	74,1	320
13-15	102	26,2	3	0,8	7	1,8	278	71,3	390
16-18	71	22,6	2	0,6	9	2,9	232	73,9	314
19	3	23,1	0	0,0	0	0,0	10	76,9	13
20	290	21,7	10	0,7	33	2,5	1.004	75,1	1.337
21	86	22,8	3	0,8	9	2,4	279	74,0	377
22	213	25,1	6	0,7	15	1,8	615	72,4	849
23	142	28,0	0	0,0	13	2,6	352	69,4	507
24	71	19,7	1	0,3	22	6,1	267	74,0	361
25	435	24,5	13	0,7	36	2,0	1.290	72,7	1.774
26	672	28,7	19	0,8	53	2,3	1.599	68,3	2.342
27	244	26,5	7	0,8	33	3,6	637	69,2	921
28	891	27,2	31	0,9	87	2,7	2.263	69,2	3.270
29	86	20,6	5	1,2	35	8,4	295	70,6	418
30	53	25,7	1	0,5	4	1,9	148	71,8	206
30.3	27	32,1	0	0,0	3	3,6	54	64,3	84
31-33	245	23,3	12	1,1	26	2,5	770	73,1	1.053
D,E 35-39	156	26,9	7	1,2	18	3,1	398	68,7	579
F 41-43	192	26,4	3	0,4	11	1,5	521	71,7	727
J 58-63	929	22,8	17	0,4	79	1,9	3.054	74,9	4.079
K 64-66	14	11,1	4	3,2	5	4,0	103	81,7	126
M 69-75	1.372	29,1	31	0,7	113	2,4	3.200	67,9	4.716
71	653	30,3	10	0,5	43	2,0	1.446	67,2	2.152
72	496	30,2	14	0,9	44	2,7	1.090	66,3	1.644
G-I, L, N-U	474	16,8	14	0,5	90	3,2	2.241	79,5	2.819
Unbekannt	60	1,8	9	0,3	180	5,4	3.090	92,4	3.345

II. NACH BESCHÄFTIGTENGROSSENKLASSEN

unter 250 Beschäftigte	6.003	24,9	138	0,6	518	2,1	17.455	72,4	24.114
250 bis 499 Beschäftigte	405	27,4	11	0,7	57	3,9	1.007	68,0	1.480
500 und mehr Beschäftigte	471	25,8	45	2,5	111	6,1	1.200	65,7	1.827
Unbekannt	58	1,6	11	0,3	197	5,4	3.358	92,7	3.624
Insgesamt	6.937	22,3	205	0,7	883	2,8	23.020	74,2	31.045

Anmerkung:

1) Die Wirtschaftsgliederung basiert auf der Klassifikation der Wirtschaftszweige des statistischen Bundesamtes, Ausgabe 2008 (WZ 2008).

Rücksendungen enthalten auch Unternehmen, die melden, dass sie weder forschen noch entwickeln. Wenn diese Unternehmen keine Angaben zu Ihrer Wirtschaftszweigklassifikation und/oder Unternehmensgröße tätigen, werden diese Merkmale nicht nachträglich recherchiert und verbleiben damit unbekannt.

Nichtteilnehmenden-Befragung

Um die möglichen Gründe für eine Nichtteilnahme an der FuE-Erhebung zu erfahren und den Rücklauf in den kommenden Erhebungen weiter zu verbessern, ist im Herbst 2022 im Anschluss an die Feldphase eine Nichtteilnehmenden-Befragung durchgeführt worden. Dazu sind alle Unternehmen aus der dritten Prioritätsklasse, von denen bis dahin keine Rückmeldung auf den FuE-Fragebogen erfolgte und zu denen Emailkontakte vorlagen, per Mail kontaktiert und um Teilnahme an einer fünfminütigen Online-Befragung gebeten worden. Es wurden 6.163 Unternehmen im Rahmen der Nichtteilnehmenden-Befragung kontaktiert.

Insgesamt haben sich 860 Unternehmen zurückgemeldet, was einer Rücklaufquote von knapp 14 Prozent entspricht. Im Rahmen dieser Befragung hatten die Unternehmen neun mögliche Gründe für die fehlende Teilnahme zur Auswahl. Mit rund 40 Prozent ist der am häufigsten gewählte Grund für die Nichtteilnahme der Zeitaufwand der Befragung. Außerdem wird von circa 18 Prozent angegeben, dass sie keine validen Aussagen zu ihren Forschungs- und Entwicklungsaufwendungen machen könnten. 16 Prozent der Unternehmen empfinden den Fragebogen zudem als zu lang. Seltener hingegen wurde angegeben, dass der Fragebogen nicht angekommen sei (11 Prozent), an keinen Befragungen ohne Meldepflicht teilgenommen wird (ebenfalls 11 Prozent), der Online-Fragebogen zu benutzerunfreundlich ist (3 Prozent), prinzipiell nicht an Umfragen teilgenommen wird (3 Prozent) und die Gefahr des Abflusses von Geschäftsgeheimnissen droht (2 Prozent). Außerdem konnten die Befragten selbst mehrere Gründe für die Nichtteilnahme als Freitext angeben. 279 der 860 Befragten nutzten diese Option. Etwa ein Drittel (88 befragte Unternehmen) ging dabei auf fehlende personelle Kapazitäten bzw. Zeitmangel ein. Weitere selbstgenannte Gründe umfassen das Versäumen der Teilnahme (8 Prozent) und die Umstrukturierung des Unternehmens, z. B. Verkauf oder Insolvenz (4 Prozent). 3 Prozent begründen ihre Nichtteilnahme mit der großen Masse an Umfrageanfragen. Seltener wurden Probleme bezüglich der Zuständigkeiten bzw. der internen Weiterleitung (2,5 Prozent) und Probleme mit dem Zugang oder der Applikation (1 Prozent) thematisiert.

Des Weiteren wurden die Unternehmen nach der Höhe der Kernindikatoren gefragt, um mit Hilfe der Nichtteilnehmenden-Befragung auch die Qualität der Schätzungen verbessern zu können. 374 Unternehmen haben in diesem Rahmen noch Angaben zur Höhe ihrer FuE-Aufwendungen im Jahr 2021 gemacht. Mithilfe dieser Angaben wurde untersucht, inwiefern die im Rahmen der Nichtteilnehmenden-Befragung angegebenen Daten von den anderweitig geschätzten Daten abweichen. Hierfür wurden die Angaben der Unternehmen mit der Schätzung verglichen, welche verwendet worden wäre, wenn die Unternehmen keine Angaben zu FuE getätigt hätten. Für die 374 Unternehmen wäre die Summe um 46 Mio. Euro höher gewesen als im Rahmen der Nichtteilnehmenden-Befragung angegeben. Dabei wären 216 Fälle überschätzt und 158 Fälle unterschätzt worden. Die maximale Überschätzung läge bei etwa 7 Mio. Euro, die maximale Unterschätzung dagegen bei 2,8 Mio. Euro. Der Median der absoluten Abweichung liegt bei +26 Tsd. Euro. Dabei fallen die Abweichungen bei jenen Fällen höher aus, zu denen bereits bei der vorherigen Vollerhebung auch keine Meldungen zur FuE-Tätigkeit vorlagen.

Insgesamt konnte durch die Nichtteilnehmenden-Befragung die Qualität der Daten durch die hinzugekommenen Angaben zur FuE-Tätigkeit verbessert werden. Zudem zeigte sich, dass ein Teil der Fragebögen über den Postweg nicht bei der bearbeitenden Person im Unternehmen angekommen ist. Deshalb wurden bereits in der folgenden Erhebung Unternehmen zusätzlich per Mail an die Teilnahme an der FuE-Erhebung erinnert.

Datenerfassung und -bearbeitung

Umgang mit fehlenden Werten

Antwortausfälle betreffen entweder den gesamten Fragebogen (Unit-Nonresponse) oder einzelne Fragen im Fragebogen (Item-Nonresponse). Da die Angaben der antwortenden Unternehmen als Basis für die Schätzung und Imputation der fehlenden Werte herangezogen werden, ist ein ausreichendes und nicht verzerrtes Meldeverhalten bei Kernindikatoren wichtig. Tabelle 6 bietet einen Überblick über das Meldeverhalten der antwortenden Unternehmen.

Tabelle 6: Item-Nonresponse bei Kernindikatoren

Kernindikator	Beschäftigte	Umsatz	Gesamte FuE-Aufwendungen	Interne FuE-Aufwendungen	FuE-Personal
Item-Nonresponse (FuE-Aktivität = ja)	12%	13%	15%	23%	34%
Item-Nonresponse (FuE-Aktivität = nein)	5%	9%	—	—	—

Das Vorgehen zum Schätzen der Unit-Nonresponse-Fälle ist abhängig von der Prioritätsklasse. Unternehmen der höchsten Prioritätsklassen (SHP und HP+) werden bei Unit-Nonresponse durch Einzelfallprüfung bearbeitet und ggf. persönlich kontaktiert, um Angaben aus Geschäftsberichten oder Vorjahresdaten zu verifizieren und zu aktualisieren. Der Mehraufwand liegt in der Relevanz dieser Unternehmen begründet, da die höchsten Prioritätsklassen solche Unternehmen umfassen, die auch die höchsten FuE-Aufwendungen aufweisen und somit das Gesamtergebnis maßgeblich beeinflussen. Bei Fällen von Item-Nonresponse ist eine manuelle Schätzung der fehlenden Werte auf Basis der gegebenen Teilantworten sowie den Vorjahresinformationen und Branchen- und Größenklassenverteilungen vorgesehen.

Antwortausfälle bei Unternehmen der mittleren Prioritätsklassen (HP- und MP) werden kollektiv geschätzt, unter Verwendung von Informationen aus externen Quellen, Branchen- und Größenklassenverteilungen sowie Vorjahresdaten. Die Unternehmen werden ebenfalls manuell geprüft, sofern Widersprüche auftauchen.

Plausibilitätsprüfungen

Die veröffentlichten FuE-Daten müssen nicht nur in sich konsistent, sondern auch plausibel zu den Vorjahren sein. Gerade die Entwicklung einzelner Großunternehmen hat Einfluss auf die sektoralen und regionalen Aggregate. Daher ist die Prüfung der Validität der Rohdaten entscheidend für die Zuverlässigkeit der gebildeten Indikatoren. Die Plausibilität der FuE-Daten wird in zwei Stufen geprüft:

Offensichtliche Fehler

Dabei handelt es sich vor allem um Summen- und Übertragungsfehler. Aber auch Zahlendreher, falsche Einheiten (zum Beispiel Angaben in Millionen Euro statt in Tausend Euro) oder Verwechslung von Zeilen oder Spalten (zum Beispiel Männer und Frauen) führen zu offensichtlichen Fehlern. Diese werden in der Einzelfallprüfung sofort bei der Dateneingabe erkannt und von der zuständigen Sachbearbeitung korrigiert. Sollten außergewöhnlich umfangreiche Datenmanipulationen nötig sein, erfolgen

diese in Absprache mit den Referenten oder der Projektleitung. In Einzelfällen ist eine Nachfrage bei den betroffenen Unternehmen nötig.

Unplausible Daten

Zunächst unplausible FuE-Daten können, aber müssen nicht falsch sein. Eine Verdopplung der internen FuE-Aufwendungen eines Unternehmens kann zum Beispiel Anzeichen einer abweichenden Unternehmensabgrenzung gegenüber dem Vorjahr oder einer abweichenden Definition von FuE seitens des Unternehmens sein. Da eine Überprüfung solcher Angaben notwendig ist, wird eine fallspezifische Untersuchung durchgeführt. Dies geschieht über einen Vergleich der FuE-Daten mit den allgemeinen Entwicklungen des Unternehmens, eine Recherche von externen Informationen (z.B. Geschäftsberichte, Förderdaten) und gegebenenfalls auch über eine Kontaktaufnahme mit dem Unternehmen. Auch auffällig konstante Angaben eines Unternehmens können Grund für weitere Recherchen sein.

Geheimhaltung

Weit über den gesetzlichen Datenschutz hinaus verpflichtet sich die SV Wissenschaftsstatistik gegenüber den befragten Unternehmen zur Geheimhaltung der gemeldeten Daten. Dies betrifft einerseits die Nutzung der Mikrodaten (zum Beispiel über das Forschungsdatenzentrum) aber auch die Veröffentlichung der aggregierten FuE-Daten. Ohne einen unverhältnismäßig hohen Aufwand soll auch aus den Veröffentlichungstabellen kein Rückschluss auf einzelne Unternehmen möglich sein. Um dies zu gewährleisten, werden Maßnahmen zur primären und sekundären Geheimhaltung getroffen.

Primäre Geheimhaltung

Die primäre Geheimhaltung betrifft zum einen Aggregate, die zu gering besetzt sind. Entsprechend werden nur Aggregate veröffentlicht, die mindestens drei Unternehmen beinhalten. Zum anderen sind Aggregate betroffen, die stark von einem Unternehmen dominiert werden (mehr als 70 Prozent des Gesamtwertes). Hier wird eine Besetzung von mehr als zehn Unternehmen gefordert.

Sekundäre Geheimhaltung

Die sekundäre Geheimhaltung stellt sicher, dass auch nicht aufgrund von Rückrechnungen auf eigentlich geheim zu haltende Aggregate geschlossen werden kann. In einigen Fällen ist die Zahl der primär geheim zu haltenden Aggregate groß genug, um dies zu verhindern. In anderen Fällen ist es nötig, auch Aggregate auszublenden, die nicht der primären Geheimhaltung unterliegen, um die sekundäre Geheimhaltung sicherzustellen.

Entsprechend der Vorgehensweise von Eurostat wird die Geheimhaltung von Aggregaten bei allen Pflichtlieferungen bzw. auf darauf aufbauenden Tabellen in eigenen Veröffentlichungen erreicht, indem die betroffenen Aggregate nicht ausgewiesen werden. Bei Regionaldaten auf Kreisebene wird die Geheimhaltung erreicht, indem Aggregate zusammengefasst werden.

Regionalisierung

Die der FuE-Erhebung zugrunde liegende EU-Verordnung sieht eine regionale Darstellung der internen FuE-Aufwendungen sowie des FuE-Personals auf NUTS-1 und NUTS-2-Ebene vor. In Deutschland entspricht dies den Bundesländern und den Regierungsbezirken. Darüber hinaus werden interessierten Personen die FuE-Daten auf Kreisebene (NUTS-3) für einen geringen finanziellen Beitrag zur Verfügung gestellt.

Die Regionalisierung der FuE-Daten erfolgt nach Forschungsstätten. Im Fragebogen wurden die Unternehmen gebeten, die Postleitzahl sowie den Ort ihres Hauptforschungsstandorts anzugeben. Aufgrund dieser Angaben konnte jeweils die aktuell gültige Gemeindegrenznummer zugeordnet werden, aus der die Zugehörigkeit zu Bundesland, Regierungsbezirk und Kreis eindeutig zu erkennen ist. Darüber hinaus wurden die Unternehmen gebeten, den Anteil der FuE-Beschäftigten pro Forschungsstätte anzugeben.

Auswertung

Datenlieferung und Publikationen

Die Liefertermine für die Ergebnisse der Erhebung 2021 waren:

- Oktober 2022: Pflichtlieferung an die OECD. Hierbei handelt es sich um die Ergebnisse, die auf den Daten der Fortschreibung beruhen. Die Pressemitteilung enthält die Bekanntgabe von Trendzahlen, die die vorläufige Summe der internen FuE-Aufwendungen sowie die vorläufige Anzahl des FuE-Personals abbilden.
- Juni 2023: Pflichtlieferung der endgültigen Ergebnisse an Eurostat. Keine Pressearbeit.

Die Veröffentlichung der endgültigen FuE-Daten 2021 erfolgte in folgenden erhebungsspezifischen Publikationen:

- Facts - Forschung und Entwicklung in der Wirtschaft 2021 (SV Wissenschaftsstatistik), erschienen im April 2023, online verfügbar unter: <https://www.stifterverband.org/fue-facts-2021>
- Policy Paper - Drei Jahre Forschungszulage: Ein erstes Fazit zur Resonanz (SV Wissenschaftsstatistik), erschienen im Mai 2023; online verfügbar unter: https://www.stifterverband.org/medien/drei_jahre_forschungszulage
- Gutachten und Studien zum deutschen Innovationssystem (Expertenkommission Forschung und Innovation)
- Bundesbericht Forschung und Innovation sowie Datenportal (BMBF)
- sonstige Publikationen des BMBF und anderer Bundesministerien
- sonstige Publikationen der EU und der OECD

Projektergebnisse und Nutzen

Die FuE-Erhebung hat das Ziel, Daten zu Forschung und Entwicklung im Wirtschaftssektor zu erheben. Die FuE-Daten werden in ihrer aggregierten Form an das BMBF, die EU sowie die OECD gemeldet und von Politik und Wirtschaft nachgefragt (bspw. Wirtschaftsverbände, Industrie- und Handelskammern). Außerdem werden die aggregierten Ergebnisse der Erhebung auch in interaktiver Form im Datenportal des Stifterverbands dem interessierten Publikum zur Verfügung gestellt (https://stifterverband.shinyapps.io/FuE_Daten/). Ziel ist es, hierdurch einen neuen und modernen Zugang zu den Daten zu ermöglichen und umfassend über die Erkenntnisse zu Forschung und Entwicklung im deutschen Wirtschaftssektor zu informieren. Die FuE-Daten liegen darüber hinaus im

Forschungsdatenzentrum auf Mikroebene vor und können zu wissenschaftlichen Forschungszwecken oder in der Lehre genutzt werden (<https://www.fdz-wissenschaftsstatistik.de>).

Einheitsleistungen

Neben Lieferung, Auswertung und Publikation der FuE-Daten gehören die Einheitsleistungen ebenfalls zu den vertraglich vereinbarten Leistungen des Projektes. Diese umfassen den 9. FuE-Workshop zum Thema: „Wissenstransfer bei technologischen und sozialen Innovationen: Alte Probleme und neue Lösungen?“ am 08. September 2022 (Einheitsleistung 1). Das Protokoll der Veranstaltung befindet sich in der Anlage 2 zu diesem Bericht.

Projektverantwortliche und Ansprechpartner

Dr. Gero Stenke
T 0201/8401-426
gero.stenke@stifterverband.de
Wissenschaftsstatistik gGmbH
im Stifterverband für die Deutsche Wissenschaft
Baedekerstraße 1
45128 Essen

Dr. Johannes Schmitt
T 0201/8401-412
johannes.schmitt@stifterverband.de
Wissenschaftsstatistik gGmbH
im Stifterverband für die Deutsche Wissenschaft
Baedekerstraße 1
45128 Essen

Literatur

OECD (2018). Frascati-Handbuch 2015: Leitlinien für die Erhebung und Meldung von Daten über Forschung und experimentelle Entwicklung. OECD Publishing, Paris.
<http://dx.doi.org/10.1787/9789264291638-de>

Anlagen

Anlage 1: Fragebogen der Erhebung zu Forschung und Entwicklung im Wirtschaftssektor 2021

Anlage 2: Protokoll der Einheitsleistung 1