

# WISTA

## Wirtschaft und Statistik

Florian Dumpert | Jannik Reichel |  
Elisa Oertel | Holger Leerhoff |  
Christian Salwiczek

Sebastian Metzl | Igor Franjić |  
Thomas Kolvenbach

Florian Peter

Daniel Nietzer | Karola Henn |  
Chris-Gabriel Islam | Keno Krewer |  
Carsten Monzert

Maria Bolz

Senait Andom | Johanna Bingel |  
Adrian Urban

Kathrin Bank | Alexander Irmer |  
Ute Losch | Anna Sandig |  
Max Schöne

Kai Engemann | Irina Meinke  
Moritz Wittmaack

**Ethische Fragen beim Einsatz von KI/ML in der Produktion  
amtlicher Statistiken –**

**Teil 1: Identifikation**

**Teil 2: Auseinandersetzung**

**Die neue Intrahandelsstatistik ab Januar 2025**

**Die Indizes der Außenhandels- und Großhandels-  
verkaufspreise auf Basis 2021**

**Hierarchisches Klassifizieren von Scannerdaten:  
ein Methodenvergleich mit Anwendung in der Verbraucher-  
preisstatistik**

**Der neue Produktionsindex für die Gesamtwirtschaft**

**Methodische Weiterentwicklungen für Unternehmens-  
gruppen im statistischen Unternehmensregister**

**Die Zeitverwendungserhebung 2022 – ein Überblick**

**Methodik und Potenziale des Einrichtungsregisters**

**Mehrfachfallprüfung im Zensus 2022 – die neue Strategie  
zur automatisierten Identifikation und Bewertung von  
Mehrfachfällen**

# 1 | 2025

#### ABKÜRZUNGEN

D	Durchschnitt (bei nicht addierfähigen Größen)
Vj	Vierteljahr
Hj	Halbjahr
a. n. g.	anderweitig nicht genannt
o. a. S.	ohne ausgeprägten Schwerpunkt
Mill.	Million
Mrd.	Milliarde

#### ZEICHENERKLÄRUNG

–	nichts vorhanden
0	weniger als die Hälfte von 1 in der letzten besetzten Stelle, jedoch mehr als nichts
.	Zahlenwert unbekannt oder geheim zu halten
. . .	Angabe fällt später an
X	Tabellenfach gesperrt, weil Aussage nicht sinnvoll
I oder —	grundsätzliche Änderung innerhalb einer Reihe, die den zeitlichen Vergleich beeinträchtigt
/	keine Angaben, da Zahlenwert nicht sicher genug
( )	Aussagewert eingeschränkt, da der Zahlenwert statistisch relativ unsicher ist
	Abweichungen in den Summen ergeben sich durch Runden der Zahlen.
	Tiefer gehende Internet-Verlinkungen sind hinterlegt.

# INHALT

3	Editorial
4	Kennzahlen und Indikatoren
8	Aktuelle Informationsangebote
10	Kurznachrichten
15	Florian Dumpert, Jannik Reichel, Elisa Oertel, Holger Leerhoff, Christian Salwiczek <b>Ethische Fragen beim Einsatz von KI/ML in der Produktion amtlicher Statistiken – Teil 1: Identifikation</b> <i>Ethical issues concerning the use of AI/ML in the production of official statistics – part 1: identification</i>
25	Florian Dumpert, Jannik Reichel, Elisa Oertel, Holger Leerhoff, Christian Salwiczek <b>Ethische Fragen beim Einsatz von KI/ML in der Produktion amtlicher Statistiken – Teil 2: Auseinandersetzung</b> <i>Ethical issues concerning the use of AI/ML in the production of official statistics – part 2: evaluation</i>
37	Sebastian Metzl, Igor Franjić, Thomas Kolvenbach <b>Die neue Intrahandelsstatistik ab Januar 2025</b> <i>New intra-EU trade statistics as of January 2025</i>
48	Florian Peter <b>Die Indizes der Außenhandels- und Großhandelsverkaufspreise auf Basis 2021</b> <i>Index of foreign trade prices and index of wholesale selling prices, base 2021</i>

## INHALT

67	Daniel Nietzer, Karola Henn, Chris-Gabriel Islam, Keno Krewer, Carsten Monzert <b>Hierarchisches Klassifizieren von Scannerdaten: ein Methodenvergleich mit Anwendung in der Verbraucherpreisstatistik</b> <i>Hierarchical classification of scanner data: a comparison of methods and their use in consumer price statistics</i>
82	Maria Bolz <b>Der neue Produktionsindex für die Gesamtwirtschaft</b> <i>The new total market production index</i>
91	Senait Andom, Johanna Bingel, Adrian Urban <b>Methodische Weiterentwicklungen für Unternehmensgruppen im statistischen Unternehmensregister</b> <i>Further methodological developments relating to enterprise groups in the statistical business register</i>
106	Kathrin Bank, Alexander Irmer, Ute Losch, Anna Sandig, Max Schöne <b>Die Zeitverwendungserhebung 2022 – ein Überblick</b> <i>The time use survey 2022 – an overview</i>
116	Kai Engemann, Irina Meinke <b>Methodik und Potenziale des Einrichtungsregisters</b> <i>The register of institutions and collective living quarters – methodology and future potential</i>
127	Moritz Wittmaack <b>Mehrfachfallprüfung im Zensus 2022 – die neue Strategie zur automatisierten Identifikation und Bewertung von Mehrfachfällen</b> <i>Checking for multiple entries in the 2022 Census – the new strategy for the automated identification and evaluation of multiple entries</i>

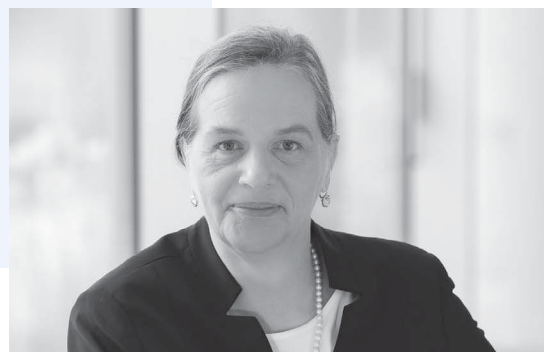
---

# EDITORIAL

---

Dr. Ruth Brand

---



## LIEBE LESERIN, LIEBER LESER,

in dieser Ausgabe unseres Wissenschaftsmagazins WISTA setzen wir die Dokumentation der Methoden des Zensus 2022 fort: Ein Beitrag erläutert das neu entwickelte Verfahren zur Identifikation und Bewertung von Mehrfachfällen. Das vollständig maschinelle Verfahren der Mehrfachfallprüfung ist ein Prozess zur Qualitätssicherung des Personenbestands des Zensus 2022, auf dessen Grundlage die Einwohnerzahl Deutschlands abgeleitet wurde.

Ab dem Berichtsmonat Januar 2025 werden in der Intrahandelsstatistik neue Schätzverfahren für fehlende Daten die Genauigkeit im Hinblick auf die Detaillerggebnisse der gehandelten Waren verbessern. Zudem ist vorgesehen, die Anmeldeschwellen zum Intrahandel in Deutschland zu erhöhen und damit Unternehmen von Statistikpflichten zu entlasten. Wie die damit verbundenen Neuerungen in der Intrahandelsstatistik die Datenqualität gewährleisten, ist einem weiteren Artikel in dieser Ausgabe zu entnehmen.

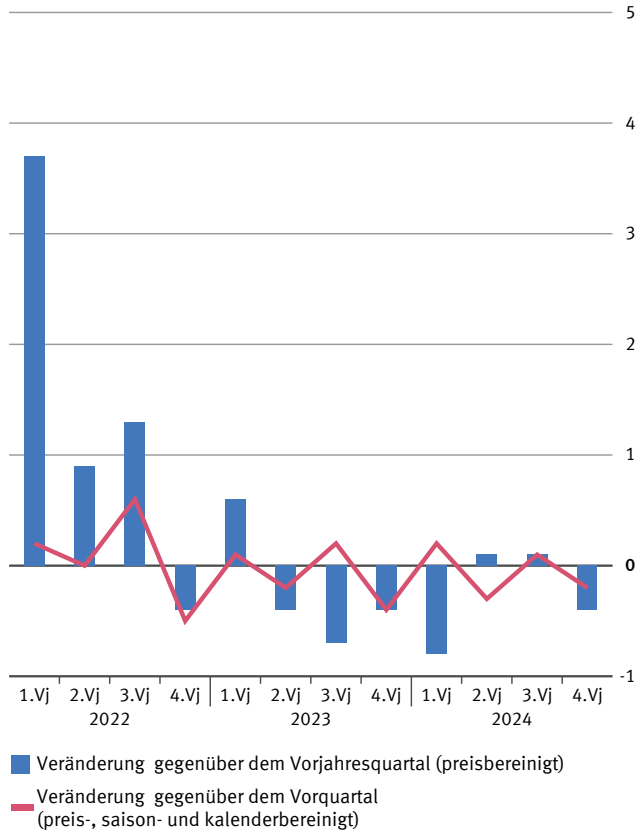
Außerdem möchte ich Ihnen den Aufsatz zum neuen Produktionsindex für die Gesamtwirtschaft empfehlen: Dieser kombiniert die vier zentralen Produktionsindizes für Industrie (in der europäischen Definition: Wirtschaftsbereiche B, C und D), Bau, Gewerbe, Handel und Dienstleistungen und bildet konjunkturelle Entwicklungen sowie strukturelle Veränderungen in der Wirtschaft ab. Er unterscheidet sich hinsichtlich der Detailtiefe, der Veröffentlichungsfrequenz und konzeptionell vom Bruttoinlandsprodukt; damit steht ein zusätzlicher Indikator für die Entwicklung der ökonomischen Aktivität in Deutschland zur Verfügung.

Wie immer bieten auch die anderen Artikel dieser Ausgabe interessanten Lesestoff aus der Vielfalt der amtlichen Statistik – ich wünsche Ihnen viel Freude bei der Lektüre.

*Ruth Brand*

Präsidentin des Statistischen Bundesamtes

**Bruttoinlandsprodukt**  
in %

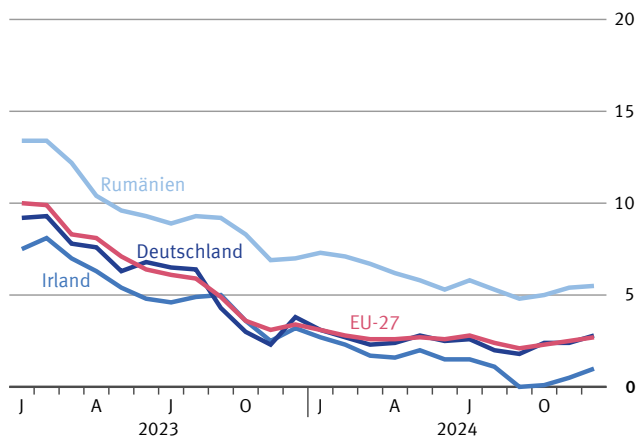


**Verbraucherpreisindex**  
2020 = 100

2024		2025	
Januar	117,6	Januar	120,3
Februar	118,1		
März	118,6		
April	119,2		
Mai	119,3		
Juni	119,4		
Juli	119,8		
August	119,7		
September	119,7		
Oktober	120,2		
November	119,9		
Dezember	120,5		

**120,3** **2,3 %**  
Veränderung zum Vorjahresmonat

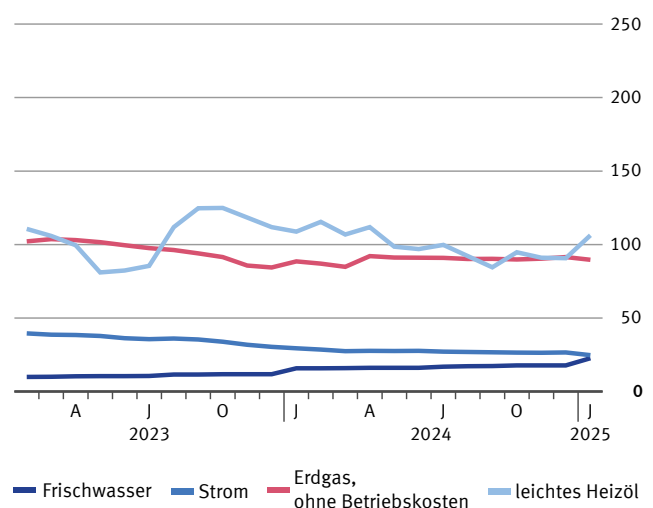
**Harmonisierter Verbraucherpreisindex insgesamt**  
Veränderung gegenüber dem Vorjahresmonat in %



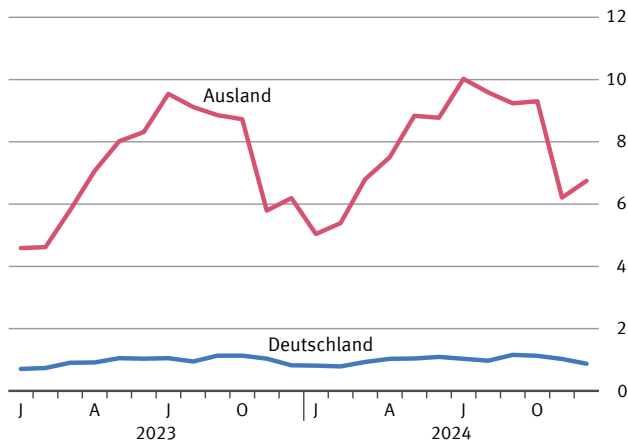
Dargestellt sind neben Deutschland und der Europäischen Union insgesamt (EU-27) die Länder mit der höchsten und der niedrigsten Veränderungsrate innerhalb der EU.

Stand: 13.02.2025

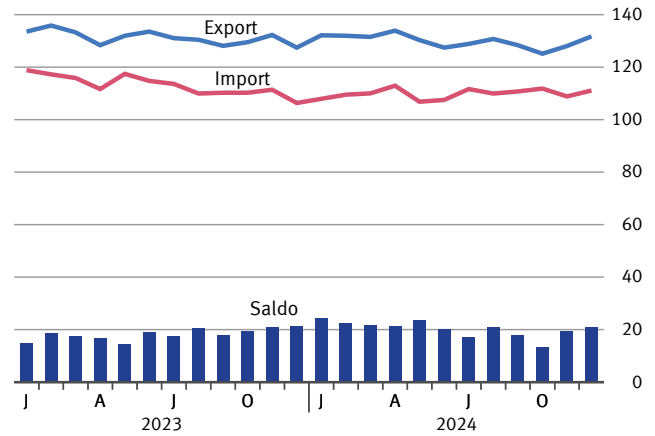
**Entwicklung der Verbraucherpreise für Energie und Wasser**  
Preisabstand in % gegenüber dem Jahr 2020



**Einsteigende Flugpassagiere auf deutschen Flughäfen nach dem Streckenziel in Mill.**

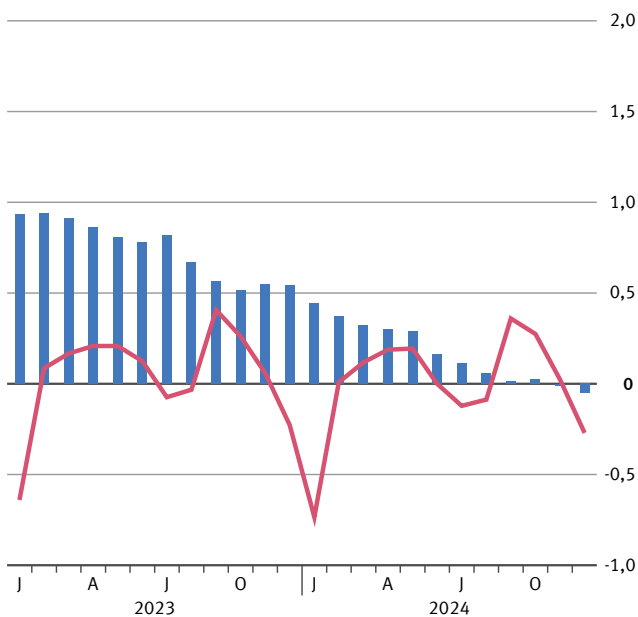


**Außenhandel in Mrd. EUR**



Kalender- und saisonbereinigte Werte nach dem Verfahren X13 JDemetra+. – Vorläufiges Ergebnis.

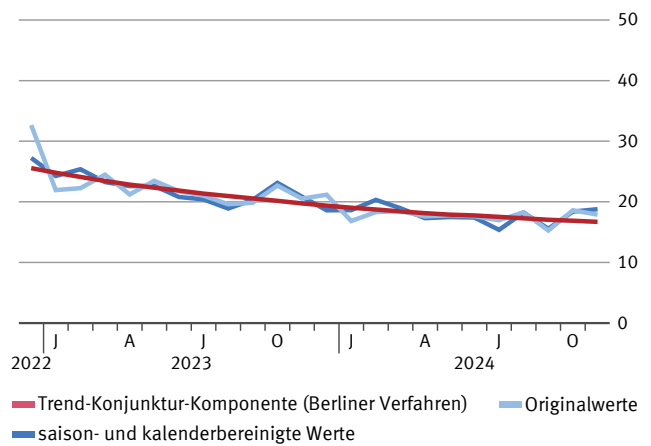
**Erwerbstätige mit Wohnort in Deutschland Veränderungen in %**



■ Veränderung gegenüber dem Vorjahresmonat  
 ■ Veränderung gegenüber dem Vormonat

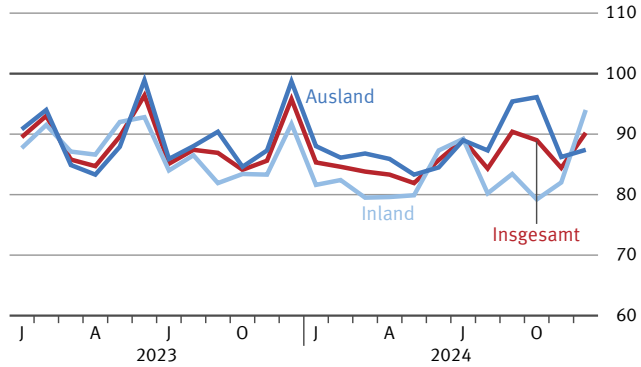
Stand: 13.02.2025

**Genehmigte Wohnungen Wohn- und Nichtwohngebäude, in 1 000**



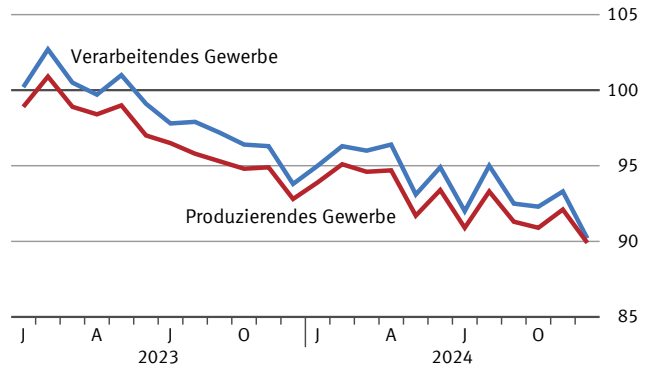
## Kennzahlen und Indikatoren

**Auftragseingang im Verarbeitenden Gewerbe**  
Volumenindex 2021 = 100



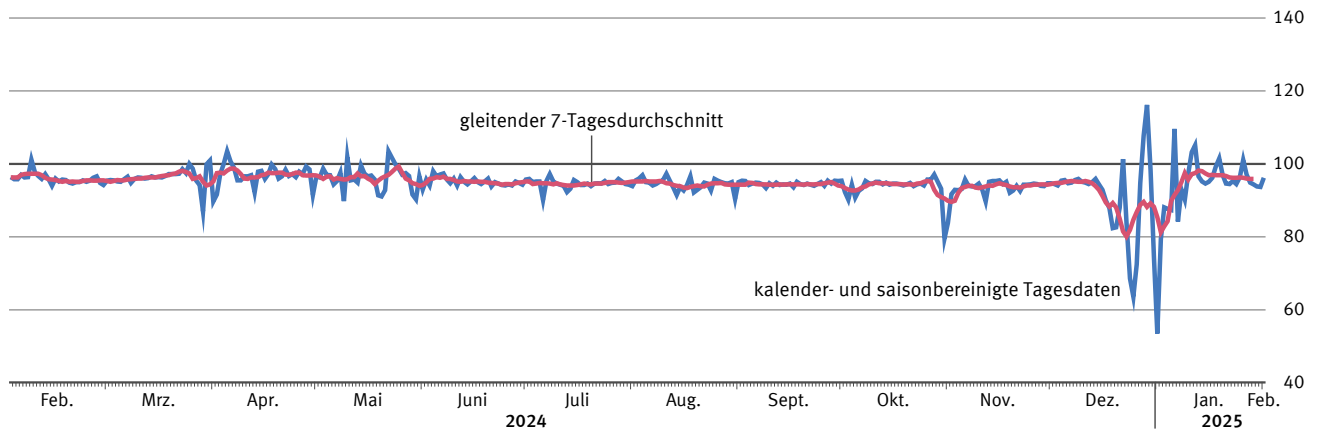
Kalender- und saisonbereinigter Wert nach dem Verfahren X13 JDemetra+. – Vorläufiges Ergebnis.

**Produktion im Produzierenden und Verarbeitenden Gewerbe**  
Index 2021 = 100



Kalender- und saisonbereinigte Werte nach dem Verfahren X13 JDemetra+. – Vorläufiges Ergebnis.

**Lkw-Maut-Fahrleistungsindex**  
2021 = 100



Quellen: Bundesamt für Logistik und Mobilität, Deutsche Bundesbank, Statistisches Bundesamt

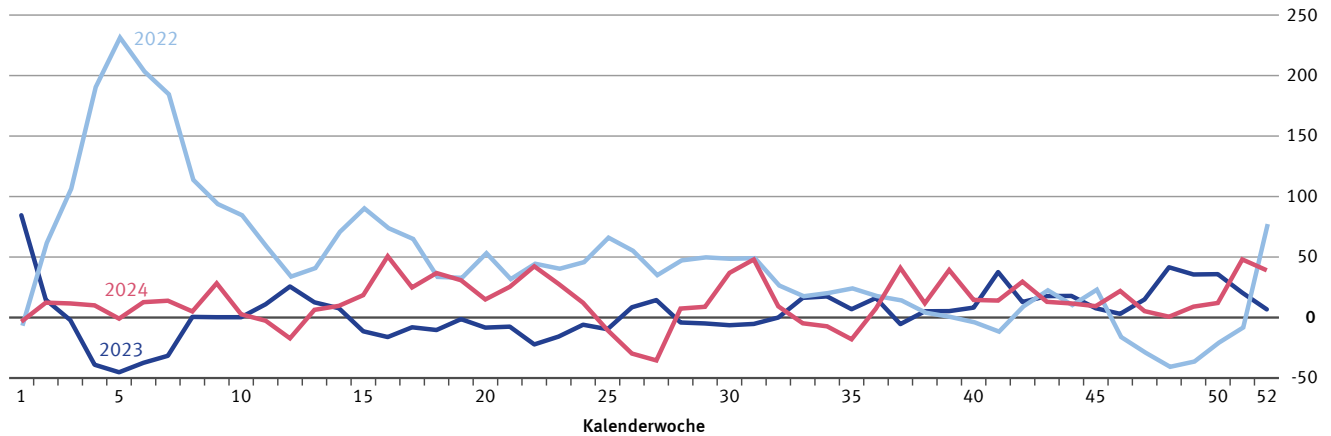
Stand: 13.02.2025



## Kennzahlen und Indikatoren

### Neue Kreditverträge nach Kalenderwochen

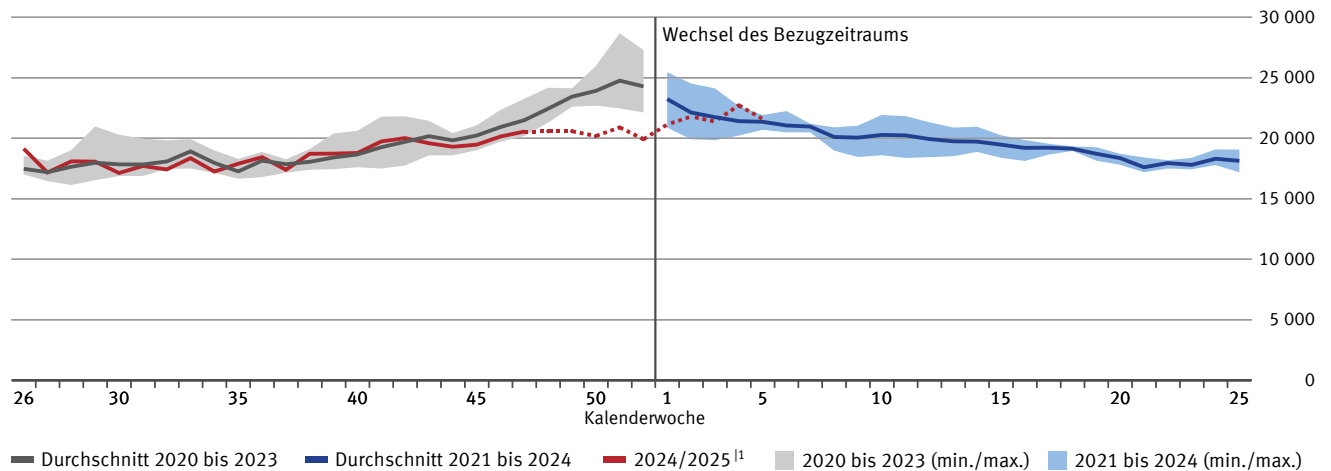
Veränderung gegenüber der entsprechenden Vorjahreswoche in %



Anfang 2022 zeigte sich mit Veränderungsraten von 100 % und mehr ein starker Anstieg im Vergleich zum Vorjahr; dabei handelt es sich um Sondereffekte, die seitens des Datenlieferanten nicht bereinigt werden konnten.

Quelle: SCHUFA Holding AG; Berechnung: Statistisches Bundesamt

### Wöchentliche Sterbefallzahlen in Deutschland



Gestrichelte Werte enthalten Schätzanteil.

1 Sonderauswertung der vorläufigen Sterbefallzahlen.

Stand: 13.02.2025



### Ukraine

Der Angriff Russlands auf die Ukraine und die damit verbundenen Sanktionen haben starke Auswirkungen auf Wirtschaft und Bevölkerung sowie den Energie-sektor. Auf einer Sonderseite zum Thema stellt das Statistische Bundesamt relevante Daten zur Verfügung. Über die Seite gelangt man auch zum [zentralen Hilfs-portal](#) der Bundesregierung für Geflüchtete aus der Ukraine.

➤ [www.destatis.de/Im-Fokus/Ukraine](http://www.destatis.de/Im-Fokus/Ukraine)



### Dashboard Deutschland

Das 2020 vom Statistischen Bundesamt entwickelte Datenportal wurde im Dezember 2024 erweitert: Es besteht nun aus dem Dashboard Integration, das über 60 Indikatoren zum Integrationsgeschehen in Deutschland umfasst, und dem Dashboard Konjunktur, das hochaktuelle und hochfrequente Zahlen zur Wirtschaftslage in Deutschland bietet. Der integrierte Pulsmesser Wirtschaft ermöglicht intuitives und einfaches Vergleichen von Daten sowie das Erkennen von konjunkturellen Entwicklungen und Zusammenhängen mithilfe täglicher, wöchentlicher, monatlicher und vierteljährlicher Indikatoren. Das Dashboard Deutschland trägt damit zu einem faktenbasierten demokratischen Diskurs der Öffentlichkeit und zur evidenz-basierten Entscheidungsfindung durch Politik und Verwaltung bei.

➤ [www.dashboard-deutschland.de](http://www.dashboard-deutschland.de)



### EXSTAT – Experimentelle Statistiken

In der Rubrik „EXSTAT – Experimentelle Statistiken“ veröffentlicht das Statistische Bundesamt regelmäßig neue, innovative Projektergebnisse. Sie entstehen auf der Grundlage neuer Datenquellen und Methoden. Im Reifegrad und in der Qualität unterscheiden sie sich von amtlichen Statistiken, insbesondere in Bezug auf Harmonisierung, Erfassungsbereich und Methodik. Dennoch sind es Ergebnisse der Statistischen Ämter des Bundes und der Länder, die interessante, neue Perspektiven auf verschiedene Themenfelder der Statistik bieten.

➤ [www.destatis.de/exstat](http://www.destatis.de/exstat)



### Ergebnisse des Zensus 2022

Wie viele Menschen leben in Deutschland? Wie wohnen und arbeiten sie? Die Ergebnisse des Zensus 2022 liefern aktuelle Antworten auf Fragen zur Bevölkerung und zu den Wohnverhältnissen in Deutschland. Alle Daten sind kostenlos und frei zugänglich in der [Zensusdatenbank](https://www.zensus2022.de/zensusdatenbank) verfügbar. Der [Zensusatlas](https://www.zensus2022.de/zensusatlas) bietet eine anschauliche Darstellung.

➤ [www.zensus2022.de](https://www.zensus2022.de)



### Klima

Der Klimawandel ist eine der größten Herausforderungen der heutigen Zeit, alle Bereiche der Gesellschaft sind betroffen. Wie beeinflusst unsere Lebens- und Wirtschaftsweise das Klima? Wie wirkt sich die Umstellung hin zu mehr Klimaschutz gesamtgesellschaftlich aus? Was bedeutet sie für unseren Alltag – vom Weg zur Arbeit bis zum aktuellen Strompreis? Wo zeigen sich die Folgen des Klimawandels? Daten und Fakten zum Thema Klima, Klimawandel und Klimaschutz sind gebündelt unter

➤ [www.destatis.de/klima](https://www.destatis.de/klima)



### Fachkräfte

Fachkräftemangel und Arbeitskräftebedarf sind zunehmend wichtige Faktoren für die wirtschaftliche Entwicklung in Deutschland. Daten und Fakten dazu bündelt das Statistische Bundesamt auf einer eigenen Sonderseite. Das Angebot umfasst die Bereiche Demografie, Erwerbstätigkeit, Bildung und Zuwanderung – und wird sukzessive erweitert.

➤ [www.destatis.de/fachkraefte](https://www.destatis.de/fachkraefte)

# KURZNACHRICHTEN

## AUS EUROPA

### Neufassung der EU-Statistikverordnung

Die EU-Statistikverordnung gibt den rechtlichen Rahmen für die Erstellung und Verbreitung europäischer Statistiken vor. Daneben regelt sie auch die Bereiche Statistische Governance und Statistische Geheimhaltung. Die Neufassung der EU-Statistikverordnung wurde am 6. Dezember 2024 als [Verordnung \(EU\) 2024/3018](#) des Europäischen Parlaments und des Rates zur Änderung der Verordnung (EG) Nr. 223/2009 über europäische Statistiken im Amtsblatt der EU veröffentlicht. Das Ziel der Überarbeitung war, das Europäische Statistische System (ESS) „fit für die Zukunft“ zu machen. Die neu gefasste Verordnung ermöglicht dem ESS, das Potenzial digitaler Datenquellen und Technologien auszuschöpfen, um die steigende Nachfrage nach neuen, detaillierteren und aktuelleren Statistiken zu decken und reaktionsfähiger zu sein.

Die wichtigsten Neuerungen sind:

- › Die Novelle sieht einen Mechanismus für die statistische Reaktion auf Krisensituationen vor (Erlass von dringenden statistischen Maßnahmen), der im Bedarfsfall eine zeitnahe und wirksame Datenerhebung gewährleistet.
- › Der Datenaustausch und die Wiederverwendung von Daten werden gefördert und vorangetrieben. Dies beinhaltet auch einen nachhaltigen Zugang zu privat gehaltenen Daten sowie Daten von Unternehmen und öffentlichen Verwaltungen für statistische Zwecke.
- › Ziel der überarbeiteten Verordnung ist es zudem, Innovationen und neue Technologien im ESS gemein-

sam zu nutzen und sie in den regulären Statistikproduktionsprozess der EU-Mitgliedstaaten zu integrieren. Experimentelle Daten erhalten in der amtlichen Statistik auf diese Weise eine neue Bedeutung.

- › Gleichzeitig stärkt die neue EU-Statistikverordnung die Rolle der nationalen statistischen Ämter in den Datenökosystemen, um so in der europäischen amtlichen Statistik auf eine Integration und Interoperabilität der Daten hinzuwirken sowie neue Standards zu setzen. Die nationalen statistischen Ämter können damit aktiv zu den Gemeinsamen Europäischen Datenräumen (Data Spaces) und zur Umsetzung des Data Governance Acts beitragen.

## AUS DEM INLAND

### Klassifikation der Wirtschaftszweige, Ausgabe 2025

Unter intensiver Beteiligung von Datennutzerinnen und Datennutzern aus Wirtschaft, Verwaltung, Wissenschaft und Gesellschaft wurde die Gliederung der Klassifikation der Wirtschaftszweige, Ausgabe 2025 (WZ 2025), erarbeitet und zum Jahresende 2024 veröffentlicht. Sie dient dazu, die wirtschaftlichen Tätigkeiten statistischer Einheiten in allen amtlichen Statistiken einheitlich zu erfassen.

Die WZ 2025 berücksichtigt die Vorgaben der europäischen statistischen Systematik der Wirtschaftszweige (NACE Revision 2.1), ist allerdings bei manchen Positionen tiefer gegliedert, um dem über die NACE hinausgehenden nationalen Informationsbedarf Rechnung zu tragen. Die WZ 2025 gilt für die Datenübermittlungen an das Statistische Amt der Europäischen Union (Eurostat)

für Bezugszeiträume ab dem 1. Januar 2025; Abweichungen hiervon regelt die Delegierte Verordnung (EU) 2023/137 der Kommission vom 10. Oktober 2022.

Neben der neuen Struktur sind auch erste Fassungen der Gegenüberstellung von WZ 2008 und WZ 2025 sowie eines Stichwortverzeichnisses der WZ 2025 verfügbar. Vor dem Hintergrund noch möglicher Änderungen an der Gegenüberstellung von NACE Revision 2 und NACE Revision 2.1 sowie im Hinblick auf die weitergehende inhaltliche Präzisierung der Abgrenzungen der nationalen Unterklassen der WZ 2025 werden die Gegenüberstellung und das Stichwortverzeichnis voraussichtlich nochmals aktualisiert. Des Weiteren ist eine Veröffentlichung der Erläuterungen zur WZ 2025 in diesem Jahr geplant.

Die WZ 2025 steht auf der Website des Statistischen Bundesamtes sowie auf dem Klassifikationsserver der Statistischen Ämter des Bundes und der Länder zur Verfügung:

➤ [www.destatis.de](http://www.destatis.de)

➤ [www.klassifikationsserver.de](http://www.klassifikationsserver.de)

## Neue Methodik zur Schätzung des Geburtsstaates in der Bevölkerungsfortschreibung

Aufgrund der Verordnung (EG) Nr. 862/2007 liefert das Statistische Bundesamt seit 2008 jährlich Daten zum Geburtsstaat der in Deutschland lebenden Bevölkerung an Eurostat. Die Daten werden für vordefinierte Geburtsstaatsgruppen untergliedert nach Alter und Geschlecht angefordert. Da das Merkmal Geburtsstaat im Bestand der Bevölkerungsfortschreibung nicht geführt beziehungsweise fortgeschrieben wird, wurde ein Verfahren zur Ermittlung des Merkmals entwickelt. Dabei wurden die Ergebnisse des Mikrozensus, der in seinen Befragungsdaten Angaben zum Geburtsstaat enthält, genutzt. Bedingt durch die Stichprobengröße lieferte dieses einfache Verfahren allerdings nicht ausreichend belastbare Ergebnisse für Untergliederungen mit kleinen Fallzahlen, wie Personen im jungen oder hohen Alter für kleine Staatengruppen.

Mit der Umstellung der Bevölkerungsfortschreibung auf den Zensus 2022 wurde deshalb ein neues Modell unter

Verwendung der Zensusdaten zum Stichtag 15. Mai 2022 entwickelt: Anhand der Zensusdaten wurde eine prozentuale Verteilung der Geburtsstaaten nach Geschlecht, Altersgruppen und Staatsangehörigkeit ermittelt. Mit dieser Verteilung können die Ergebnisse der Fortschreibung nach Staatsangehörigkeit, Geschlecht und Altersgruppe zum Stichtag 31. Dezember des Berichtsjahres auf die Geburtsstaaten aufgeteilt werden.

Vorteil der Nutzung der Zensusdaten zum 15. Mai 2022 ist, dass das vorliegende Merkmal zum Geburtsstaat den Anforderungen der internationalen Lieferverpflichtungen an Eurostat entspricht. Weiterhin können mit diesem Schätzverfahren stichprobenbedingte Einschränkungen der Belastbarkeit der Ergebnisse reduziert werden.

Dennoch ist das Modell als Übergangslösung ausgelegt. Denn im Laufe der Zeit werden sich die Verteilungen, die zum Zensusstichtag ermittelt wurden, verschieben und die Belastbarkeit des Schätzmodells einschränken. Deshalb ist die Entwicklung weiterer Schätzmodelle, zum Beispiel als ergänzende Bevölkerungsdaten nach § 5 Absatz 2 Registererprobungsgesetz, für künftige Lieferungen vorgesehen.

## VERANSTALTUNGEN

---

### Nutzerkonferenz Finanz- und Personalstatistiken

Die Nutzerkonferenz der Finanz- und Personalstatistiken fand am 29. Oktober 2024 im Statistischen Bundesamt in Wiesbaden mit zahlreichen interessierten Vertreterinnen und Vertretern von Bundesministerien und -behörden, von Landesministerien, der Zentralen Datenstelle der Landesfinanzminister und der Statistischen Ämter der Länder statt. Vorträge befassten sich mit den Themen „Die neue Statistik über die Jahresabschlüsse der doppisch buchenden Gemeinden“, „Kommunaler Schuldenvergleich“, „Umweltökonomische Gesamtrechnungen“, „Die neue Vergabestatistik“ sowie „Neuerungen in der Personalstatistik/der Gleichstellungsstatistik des Bundes“.

Die Tagesordnung sowie die Präsentationen der Vorträge sind verfügbar:

➤ [www.destatis.de](http://www.destatis.de)

### Fachausschuss „Verkehrsstatistiken“

Am 29. Januar 2025 tagte der Fachausschuss „Verkehrsstatistiken“ mit mehr als 50 Teilnehmenden aus Politik, Wissenschaft, Wirtschaft und Statistik.

Vortragende aus dem Bundesministerium für Digitales und Verkehr, dem Kraftfahrt-Bundesamt, der Bundesanstalt für Straßenwesen sowie dem Bundesamt für Logistik und Mobilität berichteten zunächst über die neuesten verkehrsstatistischen Projekte und Daten ihrer Institutionen. Themen waren unter anderem der Stand bei der Mobilitätsenerhebung in Deutschland 2023 sowie die Arbeiten an der Statistik über die Struktur der Unternehmen des gewerblichen Güterkraft- und des Werkverkehrs 2025.

Daneben lagen die inhaltlichen Schwerpunkte auf der Weiterentwicklung amtlicher Verkehrsstatistiken, deren Optimierung sowie auf einer kritischen Untersuchung des Programms der amtlichen Verkehrsstatistiken. Die Spannweite der Themen reichte dabei von der Statistik über den Öffentlichen Personennahverkehr über eine zweite Pilotstudie zur Personenbeförderung in der Binnenschifffahrt bis hin zur Schienengüterverkehrsstatistik. Das Statistische Bundesamt erläuterte zudem seine Ideen, auf welche Statistiken mit geringer Nachfrage, für die es alternative Datenquellen in guter Qualität gibt, gegebenenfalls künftig verzichtet werden könnte.

### STATISTIK VISUALISIERT

---

#### Dashboard Integration – Informationen zum Integrations- geschehen auf einen Blick

Ein gutes Viertel der Bevölkerung im Jahr 2023 war entweder selbst nach Deutschland eingewandert oder hatte zwei eingewanderte Elternteile. Die Integration dieser Menschen mit Einwanderungsgeschichte betrifft alle gesellschaftlichen Bereiche und ist Grundvoraussetzung für Teilhabe und Zusammenhalt. Mit dem Dashboard Integration bietet das Statistische Bundesamt eine Datenplattform mit umfassenden Informationen zu Stand und Entwicklung der Integration in Deutschland. Das Dashboard informiert und ermöglicht Analysen zum Integrationsgeschehen anhand von rund 500 interakti-

ven Grafiken. Es zeigt die 60 Indikatoren des 14. Integrationsberichts der Beauftragten der Bundesregierung für Migration, Flüchtlinge und Integration. Das Dashboard Integration ermöglicht Vergleiche von verschiedenen Bereichen wie Bildung, Arbeit, Wohnen, Gesundheit oder Partizipation zwischen Eingewanderten (erste Generation) und ihren Nachkommen (zweite Generation). Es enthält dabei neben Ergebnissen der amtlichen Statistik auch Daten von anderen Behörden und der wissenschaftlichen Forschung und tritt unter dem Dach des Dashboard Deutschland als weiteres Informationsangebot neben das Dashboard Konjunktur.



**Dashboard Integration**  
ein Angebot des Statistischen Bundesamtes

➤ [www.dashboard-integration.de](http://www.dashboard-integration.de)

### NEUERSCHEINUNGEN

---

#### Bundesbericht Wissenschaftlerinnen und Wissenschaftler in einer frühen Karrierephase 2025

Der „Bundesbericht Wissenschaftlerinnen und Wissenschaftler in einer frühen Karrierephase 2025“ präsentiert statistische Daten und Forschungsbefunde zu Promovierenden und Promovierten in Deutschland. Schwerpunktthema der Ausgabe 2025 ist die Etablierung der Tenure-Track-Professur, weitere Themen sind Arbeits- und Beschäftigungsbedingungen, Übergänge in die Qualifizierung und Qualifizierungsbedingungen in der Promotionsphase, Karrierewege und berufliche Perspektiven nach der Promotion sowie die Bedeutung der Befristung von Arbeitsverträgen. Die Vereinbarkeit von Familie und akademischer Karriere, Mobilitätsverhalten und die Effekte der Coronapandemie auf Wissenschaftlerinnen und Wissenschaftler in einer frühen Karrierephase werden ebenfalls behandelt. Basis des Berichtes sind Daten aus amtlichen Statistiken sowie aus regelmäßig durchgeführten Befragungen und Erhebungen. Vergleichbarkeit und Einordnung der Befunde stehen bei der Aufbereitung der Daten im Mittelpunkt.

➤ [www.bmbf.de](http://www.bmbf.de)



### Interaktive Eurostat-Veröffentlichung „Migration and asylum in Europe – 2024 edition“

Die Ausgabe 2024 der interaktiven Veröffentlichung „Migration und Asyl in Europa“ des Statistischen Amtes der Europäischen Union (Eurostat) enthält wichtige Daten und Trends zu Migration und Asyl. Sie befasst sich mit Themen wie Menschen, die in die EU, innerhalb der EU und aus der EU migrieren, mit dem internationalen Schutz von Nicht-EU-Bürgerinnen und -Bürgern sowie mit irregulärer Migration.

Die Veröffentlichung ermöglicht es, Indikatoren zu Migration und Asyl auf EU- und Länderebene zu untersuchen, zu visualisieren und zu vergleichen. Sie bietet teilweise auch Daten für die Schweiz, Norwegen, Liechtenstein und Island.

➤ [ec.europa.eu](https://ec.europa.eu)

### Neue Veröffentlichungen der OECD

#### Survey of Adult Skills 2023

Die Lesekompetenz und alltagsmathematische Kompetenz Erwachsener sinkt oder stagniert in den meisten Ländern der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD). In Deutschland haben sich die Durchschnittsergebnisse in den Bereichen Lesekompetenz und alltagsmathematische Kompetenz 2022–2023 gegenüber denen von 2011–2012 kaum verändert. Sowohl in Lesekompetenz als auch in alltagsmathematischer Kompetenz verstärkte sich allerdings das Gefälle zwischen den leistungsstärksten und den leistungsschwächsten Erwachsenen.

Die Internationale Studie über Lese-, Rechen- und Analysefähigkeiten Erwachsener „Survey of Adult Skills 2023“ soll Erkenntnisse dazu liefern, wie Kompetenzentwicklung und -nutzung die Beschäftigungsaussichten und die Lebensqualität verbessern und das Wirtschaftswachstum ankurbeln. Sie untersucht auch, wie Lese-, Mathematik- und Problemlösungskompetenzen am Arbeitsplatz eingesetzt werden. In der Erhebung wurden die Kompetenzen von rund 160 000 Menschen im Alter von 16 bis 65 Jahren in 31 Ländern gemessen.

➤ [www.oecd.org](https://www.oecd.org)

Aus dieser Studie ist auch eine Ländernotiz zu Deutschland verfügbar:

➤ [www.oecd.org](https://www.oecd.org)

#### Regions and Cities at a Glance 2024

Preissteigerungen, demografische Veränderungen, Klimawandel – die jüngsten Krisen treffen Haushalte in Städten und Regionen auf vielfältige Weise. Der Abstand zwischen den Regionen mit dem höchsten und dem niedrigsten Wachstum der Haushaltseinkommen vergrößerte sich zwischen 2014 und 2022 von 2,6 auf 3,6 Prozentpunkte. Das inflationsbereinigte Wachstum der Haushaltseinkommen verlangsamte sich von 2 % auf 0,85 % pro Jahr. Darüber hinaus werden im nächsten Jahrzehnt 54 % der Metropolregionen, 74 % der Regionen in der Nähe von mittelgroßen oder großen funktionalen städtischen Gebieten und 64 % der Regionen, die weit von einem mittelgroßen oder großen funktionalen städtischen Gebiet entfernt sind, eine schrumpfende Erwerbsbevölkerung aufweisen. In polaren und kalten Regionen stiegen 2023 die Temperaturen mehr als doppelt so stark wie in den trockenen OECD-Regionen (2°C gegenüber 0,8°C).

Der OECD-Bericht „Regions and Cities at a Glance 2024“ bietet international vergleichbare Daten, um Orte zu identifizieren, an denen die wirtschaftlichen oder sozialen Entwicklungen und die Faktoren, die diese beeinflusst haben, stärker oder schwächer waren.

➤ [www.oecd.org](https://www.oecd.org)

#### Pensions Outlook 2024

Das Pensionsvermögen ist in den entwickelten Industrieländern auf durchschnittlich 55 % des Bruttoinlandsprodukts gestiegen. In einigen Schwellenländern ist der Anteil ähnlich hoch. Eine Ausnahme bildet Deutschland, wo nur ein geringer Anteil des Rentensystems aktienfinanziert ist.

Laut „Pensions Outlook 2024“ können aktienbezogene Renten eine wichtige Finanzierungssäule der Altersvorsorge sein, wenn sie inklusiv und wirkungsorientiert ausgestaltet sind. Der OECD-Bericht zeigt auf, vor welchen Herausforderungen die Rentensysteme stehen und wie diesen begegnet werden kann.

➤ [www.oecd.org](https://www.oecd.org)





# ETHISCHE FRAGEN BEIM EINSATZ VON KI/ML IN DER PRODUKTION AMTLICHER STATISTIKEN – TEIL 1: IDENTIFIKATION

Florian Dumpert, Jannik Reichel, Elisa Oertel, Holger Leerhoff,  
Christian Salwiczek

➤ **Schlüsselwörter:** Künstliche Intelligenz – Maschinelles Lernen – Ethik – Recht – Qualität

## ZUSAMMENFASSUNG

Künstliche Intelligenz (KI) hat mit ihrem Teilgebiet Maschinelles Lernen (ML) Einzug gehalten in die Verwaltung im Allgemeinen sowie in die amtliche Statistik in Deutschland im Speziellen. Welche ethischen Fragen sind jedoch beim Einsatz von KI/ML in der Produktion amtlicher Statistiken zu stellen? Der vorliegende erste Teil des Artikels identifiziert ethische Fragestellungen beim Einsatz von KI/ML in der amtlichen Statistik anhand bereits bekannter Vorarbeiten anderer Staaten und zwischenstaatlicher Organisationen. Ein separater zweiter Teil des Artikels untersucht, ob die nationalen Vorgaben der bestehenden und meist nicht KI/ML-spezifischen Referenzrahmen wie Recht und Qualität bereits ausreichen, um den hier identifizierten, auf Risikoszenarien basierenden Fragestellungen hinreichend zu begegnen.

➤ **Keywords:** artificial intelligence – machine learning – ethics – legislation – quality

## ABSTRACT

*Artificial intelligence (AI) and its subset of machine learning (ML) have found their way into public administration in general and into official statistics in particular in Germany. What ethical issues need to be considered, however, when using AI/ML in the production of official statistics? This first part of the article identifies ethical issues concerning the use of AI/ML for official statistics on the basis of previous work conducted and published in other countries and by inter-governmental organisations. A separate second part of the article looks at the national requirements of existing reference frameworks, such as legislation and quality, which, in most cases, do not relate specifically to AI/ML. The second part examines whether these requirements are already enough to adequately address the issues which have been identified in this first part of the article on the basis of risk scenarios.*

### Dr. Florian Dumpert

ist Diplom-Mathematiker und leitet das Referat „Künstliche Intelligenz, Big Data“ des Statistischen Bundesamtes. Im Statistischen Verbund der Statistischen Ämter des Bundes und der Länder leitet er im Co-Vorsitz-Modell den Arbeitskreis Maschinelles Lernen.

### Jannik Reichel

hat einen Master in Business Intelligence and Smart Services absolviert und arbeitete zum Erstellungszeitpunkt der Aufsätze im Referat „Künstliche Intelligenz, Big Data“ des Statistischen Bundesamtes als Data Scientist an einem Projekt zur Frühschätzung der Einkommenssteuerstatistiken mittels Machine Learning und KI.

### Elisa Oertel

ist Bioinformatikerin und leitet ein Data-Science-Team im Landesamt für Statistik Niedersachsen. Auch als Vorstandsmitglied in einem deutschlandweiten Philosophieverein widmet sie sich der Verankerung ethischer Prinzipien in der technologischen Entwicklung.

### PD Dr. Holger Leerhoff

leitet im Amt für Statistik Berlin-Brandenburg das Referat „Querschnittsanalysen, EXSTAT und Geoservice“. Der Privatdozent für Philosophie an der Universität Oldenburg ist Mitglied verschiedener Arbeitskreise des Statistischen Verbundes, etwa zum Maschinellen Lernen und zur Nutzung neuer digitaler Daten.

### Christian Salwiczek

leitet das Referat „Unternehmensregister/Profiling/LCU“ beim Statistischem Amt Nord und hat die Co-Leitung des Analyse-Netzwerks des Statistischen Verbundes inne. In verschiedenen Arbeitskreisen arbeitet er an der Modernisierung der amtlichen Statistik mit und befasst sich in unterschiedlichen Gremien auch mit ethischen Fragestellungen.

# 1

## Einleitung

Bei der Einführung neuer Technologien spielen ethische Fragestellungen eine wichtige Rolle. So ist zunächst nicht klar, an welchen Normen und Zielen, Zwecken und Werten die Akteure ihr Handeln orientieren sollen, wenn über Künstliche Intelligenz mit ihrem Teilgebiet Maschinelles Lernen (KI/ML)<sup>1</sup> gesprochen wird. Dies zeigt sich in der aktuell wiedergeführten breiten gesellschaftlichen Debatte<sup>2</sup> ebenso wie in Publikationen und Fachtagungen<sup>3</sup>. Ein menschenzentrierter, dem Wohl der Allgemeinheit dienender Einsatz von KI/ML wird gefordert, gleichzeitig sollen Daten und sogenannte Reallabore bereitgestellt werden, um rechtlich noch nicht zugelassene KI/ML-Anwendungen entwickeln zu können. Häufig ist dabei zunächst unklar, was sich genau hinter dem Begriff der „ethischen Fragestellung“ verbirgt. Es gibt verschiedene Ansätze, den Begriff zu definieren und zu den so herausgearbeiteten Fragestellungen Position zu beziehen, beispielsweise die „Ethik-Leitlinien für eine vertrauenswürdige KI“ der von der Europäischen Kommission eingesetzten Hochrangigen Expertengruppe für Künstliche Intelligenz (Renda, 2020)<sup>4</sup>, die „OECD AI Principles“ der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung – OECD (OECD, 2024)<sup>5</sup>,

die Stellungnahme des Deutschen Ethikrats (Deutscher Ethikrat, 2023)<sup>6</sup> oder der „Verhaltenskodex des Bundes für menschenzentrierte und vertrauenswürdige Datenwissenschaft“ (Schweizer Bundesrat, 2023)<sup>7</sup>. Diese Darstellungen gehen das Thema sehr allgemein an. Ziel dieses Beitrags ist jedoch, Ethik mit einem Fokus auf KI/ML in der amtlichen Statistik zu beleuchten. Die deutsche amtliche Statistik ist dabei keineswegs Vorreiter: Statistische Ämter anderer Staaten und weitere zwischenstaatliche Institutionen haben sich bereits vor einiger Zeit mit Fragen von Ethik und ethischer Verantwortung auseinandergesetzt. Eine Einordnung für Deutschland stand bislang aus<sup>8</sup>; der vorliegende Aufsatz soll diese Lücke schließen. Ziel des Aufsatzes ist nicht, die Ethik der amtlichen Statistik als solche und übergreifend zu behandeln. Die [Fundamental Principles of Official Statistics](#) der United Nations Statistics Division sehen in ihrem Grundsatz 2 neben beruflichen Standards und wissenschaftlichen Grundsätzen auch die Orientierung an einer Berufsethik vor. Die Wirtschaftskommission für Europa der Vereinten Nationen (UNECE) widmete sich ethischen Fragen im Kontext der amtlichen Statistik im Allgemeinen in einem [Workshop zur Ethik in modernen statistischen Organisationen](#).<sup>9</sup> Dieser Aufsatz unter-

- 1 Zum Begriff und zur Geschichte von KI/ML siehe Seising (2024).
- 2 Die Diskussion wird mit unterschiedlichen Schwerpunkten seit Langem und immer wieder geführt, siehe beispielsweise Weizenbaum (1977).
- 3 Exemplarisch für verschiedene Formate siehe Ferdinand/Kind (2018), Nezik (2023) sowie Novotny/von Lindern (2023).
- 4 Renda (2020) bezieht sich auf die englische Fassung (High-Level Expert Group on Artificial Intelligence: [Ethics Guidelines for Trustworthy AI](#)) und beschreibt den Prozess hin zu einer umfassenden europäischen Strategie unter Berücksichtigung der ethischen Auswirkungen der Künstlichen Intelligenz und der Auswirkungen auf den Menschen (ohne speziellen Bezug zur amtlichen Statistik). Die endgültige Fassung der „Ethics Guidelines for Trustworthy AI“ schlägt schließlich drei Elemente vor: die Einhaltung rechtlicher Vorschriften, die Ausrichtung an genau spezifizierten ethischen Grundsätzen und eine sogenannte „soziotechnische“ Robustheit (Renda, 2020, hier: Seite 655). Dabei gilt, dass ethische Grundsätze (genannt werden die Achtung der menschlichen Autonomie, Schadensverhütung, Fairness und Erklärbarkeit) zu einem gewissen Grad unbestimmt sind und daher einer Konkretisierung durch Anforderungen (requirements) bedürfen. Ein Assessment Framework aus 131 Fragen leistet dies (Renda, 2020, hier: Seite 661).
- 5 Die OECD nennt in ihren AI Principles integratives Wachstum, nachhaltige Entwicklung und Wohlbefinden; Menschenrechte und demokratische Werte einschließlich Fairness und Privatsphäre; Transparenz und Erklärbarkeit; Robustheit, Sicherheit und Schutz; Verantwortlichkeit.

- 6 Kapitel 8.2 dieser Stellungnahme behandelt ethische Fragen algorithmischer Automatisierung im Verwaltungshandeln. Diskutiert werden hier jedoch automatisierte Entscheidungssysteme mit direkter Wirkung auf Personen, ein Anwendungsfall, der für die deutsche amtliche Statistik nicht einschlägig ist, wie in diesem Aufsatz später noch herausgearbeitet wird. Konkrete Empfehlungen für den Einsatz von KI in der Verwaltung finden sich auf Seite 330 ff., darunter insbesondere die Empfehlung, dass für Softwaresysteme in der öffentlichen Verwaltung Qualitätskriterien verbindlich und transparent festgelegt werden müssen (zum Beispiel in Bezug auf Genauigkeit, Fehlervermeidung und Unverzerrtheit). Der Leitgedanke der Stellungnahme des Deutschen Ethikrats ist, dass der Einsatz von KI/ML die Handlungsspielräume von Menschen generell erweitern sollte und keinesfalls verkleinern darf. Dieses Prinzip erscheint aus Sicht der Autorin und der Autoren für die amtliche Statistik sehr gut anwendbar und steht nicht im Widerspruch zu bisherigen Einsatzfeldern, wie sie in Kapitel 3 dargestellt werden.
- 7 Der Verhaltenskodex bezieht sich auf Datenwissenschaften allgemein und ist nicht spezifisch auf KI/ML-Methoden ausgerichtet. Der ethische Umgang mit Daten und Ergebnissen ist definiert als „Verwendung von Daten und Ergebnissen in einer ethischen, nachhaltigen und verantwortungsvollen Weise“ (Schweizer Bundesrat, 2023, hier: Seite 3). Er ist neben Datensicherheit, Erklärbarkeit, Reproduzierbarkeit und Transparenz eine von zwölf aufgelisteten Dimensionen, anhand derer sich Datenwissenschaften in der Verwaltung orientieren sollen.
- 8 Das gilt nach Kenntnis der Autorin und der Autoren wiederum sowohl für den Einsatz von KI/ML in der amtlichen Statistik als auch in Behörden im Allgemeinen.
- 9 Die UNECE stellt in diesem Zusammenhang fest: „Developing a cohesive definition of ethics in official statistics, along with its constituent components, is imperative“ (United Nations, 2024, hier: Seite 3).

stellt hingegen bereits die ethische Vertretbarkeit der amtlichen Statistik als solcher. Sein expliziter Fokus liegt auf dem Einsatz von KI/ML in der Produktion der deutschen amtlichen Statistik, das heißt innerhalb der Phasen und Teilprozesse des Geschäftsprozessmodells Amtliche Statistik (GMAS; siehe Blumöhr und andere, 2017, hier: Kapitel 3). Nationale und internationale Vorarbeiten fließen in die Ausführungen ein und bilden Anknüpfungspunkte für die folgenden Überlegungen. Der Beitrag erhebt dabei nicht den Anspruch, ein detaillierter Leitfaden für den ethisch vertretbaren Einsatz von KI/ML in der amtlichen Statistik zu sein; vielmehr soll er die bereits heute umgesetzten praktischen Anwendungsfälle in der amtlichen Statistik mit den damit verbundenen ethischen Fragestellungen verknüpfen.

Dass eine Auseinandersetzung mit diesem Thema prinzipiell erforderlich ist, zeigen Negativbeispiele etwa aus Österreich (Lopez, 2019; Allhutter und andere, 2020) und den Niederlanden (Fenger/Simonse, 2024; hier: Kapitel 7). Bei diesen hatte KI/ML-gestütztes Verwaltungshandeln – und auch die Produktion amtlicher Statistiken ist Verwaltungshandeln, wenngleich etwas anders gelagert – unmittelbar negativen Einfluss auf Menschen. Auch die Kommission Zukunft Statistik<sup>10</sup> fordert in ihrem Abschlussbericht mit Blick auf eine Sicherstellung der gesellschaftlichen Akzeptanz des Einsatzes von KI/ML, dass „ethische Anforderungen zu Erklärbarkeit und Fairness berücksichtigt und geprüft werden“ sollen (Kommission Zukunft Statistik, 2024; hier: Seite 43 f.).

Kapitel 2 dieses Aufsatzes stellt in einem Überblick die Vorarbeiten anderer Staaten und zwischenstaatlicher Organisationen vor, die sich bereits mit ethischen Fragestellungen im Kontext des Einsatzes von KI/ML in der amtlichen Statistik befassen haben. Anwendungsfälle von KI/ML in der deutschen amtlichen Statistik werden in Kapitel 3 skizziert, bevor in Kapitel 4 die Identifikation von ethischen Fragestellungen erfolgt. Das Zwischenfazit in Kapitel 5 schließt die Identifikation von ethischen Fragen beim Einsatz von KI/ML in der Produktion amtlicher Statistiken ab und stellt die Verbindung zum

zweiten Teil des Artikels her, der auf die Auseinandersetzung mit den identifizierten ethischen Fragestellungen fokussiert (Dumpert und andere, 2025).

## 2

### Vorarbeiten mit Bezug zur amtlichen Statistik

---

Das „Framework for Responsible Machine Learning Processes“ von Statistics Canada (Statistics Canada, 2021), welches als komplementär zum dortigen Qualitätsrahmenwerk gesehen wird, umfasst vier Hauptthemen: Respekt für die Menschen, Respekt für die Daten, fundierte Methoden und solide Anwendungen. Näher ausgeführt umfassen diese Themen wiederum Unterpunkte wie Fairness, Verantwortlichkeit, Datenschutz, Informationssicherheit, Vertraulichkeit, Transparenz, Reproduzierbarkeit, Qualität der Trainingsdaten und Erklärbarkeit. ➤ [Übersicht 1](#) geht detailliert auf diese Aspekte ein. Zu den einzelnen Unterpunkten stellt das Rahmenwerk auch Operationalisierungen in Form von Richtlinien bereit.

Die UK Statistics Authority veröffentlichte 2021 „Ethical considerations in the use of Machine Learning for research and statistics“ (UK Statistics Authority, 2021), einschließlich einer „Ethics checklist“. Sie geht eher auf traditionelle (ethische) Fragestellungen wie Transparenz und Datenschutz ein, benennt aber ergänzend Verzerrungen und Fehler in den Modellen, die dann zu schlechten Entscheidungen führen, als ethisch problematisch. ➤ [Übersicht 2](#)

Die UNECE schließlich hat 2023 ein „International Framework on Responsible AI for Official Statistics“ erarbeitet.<sup>11</sup> Dieses fordert beispielsweise, dass statistische Ämter sicherstellen müssen, dass KI/ML-Algorithmen transparent sind und dass Nutzende verstehen, welchen Einfluss KI/ML-Modelle auf die Datenqualität haben. Das Rahmenwerk nennt Kernprinzipien wie Verantwortlichkeit (Accountability), Transparenz, Erklärbarkeit, Fairness, Nichtdiskriminierung und Validität. Details hierzu enthält ➤ [Übersicht 3](#).

---

10 Die [Kommission Zukunft Statistik](#) (KomZS) wurde vom Statistischen Bundesamt eingerichtet und mit der Erarbeitung von Empfehlungen für eine vorausschauende Programmplanung und eines Zielbilds der amtlichen Statistik für das Jahr 2030 beauftragt. Der Abschlussbericht der Kommission (Kommission Zukunft Statistik, 2024) wurde am 16. Januar 2024 der Leitung des Statistischen Bundesamtes überreicht.

---

11 Veröffentlichung in Vorbereitung.

## Übersicht 1

### Kanadische Dimensionen: Framework for Responsible Machine Learning Processes

Dimension	Kurzbeschreibung
Respect for People	
Value to Canadians	Der Einsatz von ML-Algorithmen schafft einen Mehrwert für die kanadische Bevölkerung, sei es bei den Produkten selbst oder durch eine höhere Effizienz im Produktionsprozess.
Prevention of harm	Durch den Einsatz von ML könnten schutzbedürftige Bevölkerungsgruppen zu Schaden kommen, wenn sensible Informationen über sie an die Öffentlichkeit gelangen.
Fairness	Der Grundsatz der Verhältnismäßigkeit zwischen Mitteln und Zielen ist zu beachten und ein Gleichgewicht zwischen konkurrierenden Interessen und Zielen herzustellen. Fairness stellt sicher, dass Einzelpersonen und Gruppen frei von ungerechten Vorurteilen, Diskriminierung und Stigmatisierung sind.
Accountability	Verantwortlichkeit ist die rechtliche und ethische Verpflichtung einer Person oder einer Organisation, für ihre Aktivitäten verantwortlich zu sein und die Ergebnisse in transparenter Weise offenzulegen. Algorithmen sind nicht verantwortlich; jemand ist für die Algorithmen verantwortlich.
Respect for Data	
Privacy	Bei der Beschaffung sensibler Informationen haben die Regierungen Verpflichtungen in Bezug auf die Erhebung, Verwendung, Weitergabe und Aufbewahrung persönlicher Daten. Der Schutz der Privatsphäre bezieht sich im Allgemeinen auf Informationen über einzelne Personen.
Security	Unter Sicherheit versteht man die Vorkehrungen, die Organisationen treffen, um zu verhindern, dass vertrauliche Informationen in unangemessener Weise erlangt oder offengelegt werden.
Confidentiality	Vertraulichkeit bezieht sich auf den Schutz, keine identifizierbaren Informationen über ein Individuum (zum Beispiel eine Person, ein Unternehmen oder eine Organisation) zu veröffentlichen. Sie impliziert ein Vertrauensverhältnis zwischen dem Auskunftgebenden und der Organisation, die die Daten sammelt.
Sound Application	
Transparency	Transparenz bedeutet, dass klar begründet werden muss, warum ein bestimmter Algorithmus und bestimmte Lerndaten für die untersuchte Anwendung am besten geeignet sind. Um Transparenz zu erreichen, sollten ML-Entwickler eine umfassende Dokumentation erstellen und auch den Code für andere zugänglich machen, ohne dabei die Vertraulichkeit oder den Datenschutz zu gefährden.
Reproducibility of process and results	Reproduzierbarkeit des Prozesses bedeutet, dass es eine ausreichende Dokumentation und gemeinsame Nutzung des Codes gibt, sodass der ML-Prozess von Grund auf neu erstellt werden könnte. Reproduzierbarkeit der Ergebnisse bedeutet, dass dieselben Ergebnisse zuverlässig erzielt werden, wenn alle Betriebsbedingungen gleich sind. Es gibt keine Ad-hoc-Schritte oder menschliche Eingriffe, die die Ergebnisse verändern könnten.
Sound Methods	
Quality learning data	Die Repräsentativität in Bezug auf die Verteilung der Eingabe- oder Merkmalsvariablen ist gegeben.
Valid inference	Im KI/ML-Kontext bedeutet das, dass Vorhersagen, die anhand des trainierten Modells für noch nie zuvor gesehene Daten gemacht wurden, zu einem hohen Prozentsatz recht nahe an ihren jeweiligen wahren Werten liegen oder, im Fall von kategorialen Daten, dass Vorhersagen zu einem hohen Prozentsatz korrekt sind.
Rigorous modeling	Präzise Modellierung bedeutet, dass die Algorithmen verifiziert und validiert werden müssen. Dies ermöglicht es den Nutzenden und Entscheidungsträgern, dem Algorithmus in Bezug auf seine Einsatzfähigkeit, Zuverlässigkeit und Robustheit zu vertrauen.
Explainability	Ein erklärbares Modell ist ein Modell, das ausreichend dokumentiert ist, sodass klar ist, wie die Ergebnisse verwendet werden sollten und welche Art von Schlussfolgerungen oder weiteren Untersuchungen unterstützt werden können. Dies ähnelt der Transparenz mit dem Unterschied, dass diese eine Dokumentation erfordert, die erklärt, warum und wie der maschinelle Lernprozess verwendet wurde.

Quelle: [Statistics Canada, 2021](#); Übersetzung zum Teil mittels deepL.com

# Ethische Fragen beim Einsatz von KI/ML in der Produktion amtlicher Statistiken – Teil 1: Identifikation

## Übersicht 2

### Britische Dimensionen: Ethical considerations in the use of Machine Learning for research and statistics

Dimension	Kurzbeschreibung
The potential for bias	In vielen Fällen ist die Verzerrung beim Maschinellen Lernen das Ergebnis kognitiver menschlicher Fehler, das heißt Probleme werden von denjenigen verursacht, die die Systeme entwickeln oder trainieren. Zum Beispiel können die Trainingsdaten unvollständig oder nicht repräsentativ für die richtige Population sein. Verzerrungen sind jedoch nicht immer auf menschliches Versagen zurückzuführen, sondern können auch als Folge früherer gesellschaftlicher Normen in Datensätze oder Modelle eingebracht werden.
Transparency and explainability	Transparenz bezieht sich auf die Verpflichtung der Beteiligten, dafür zu sorgen, dass die Entscheidungen, die sie in Bezug auf ihre Daten, Analysen und Methoden treffen, offen und ehrlich dokumentiert werden. Außerdem sollen diese in einer Weise kommuniziert werden, die es anderen ermöglicht, sie zu bewerten. Erklärbarkeit kann dann verwendet werden, um Techniken des Maschinellen Lernens, die wir als Menschen angemessen verstehen, denen wir vertrauen und mit denen wir umgehen können, zu definieren. Obwohl sie nicht gleichbedeutend sind, sind Fragen der Transparenz und der Erklärbarkeit untrennbar miteinander verbunden.
Accountability	Jede Person, die an der Entwicklung eines maschinellen Lernsystems beteiligt ist, trägt ein gewisses Maß an Verantwortung für die Berücksichtigung der Auswirkungen des Systems, ob gut oder schlecht. Es gibt keine Verantwortung der Maschine oder des Systems. Verantwortlichkeit ist die Bereitschaft, die Verantwortung für einen Prozess zu übernehmen und damit auch für alle Fragen und Probleme, die sich daraus ergeben könnten, geradezustehen.
Confidentiality	Beim Maschinellen Lernen – insbesondere bei der Verwendung von prädiktiven Analysen – werden häufig Daten aus verschiedenen Quellen miteinander verknüpft. Neue Datenquellen (wie Daten aus sozialen Medien oder biometrische Daten) werden zunehmend von Analystinnen und Analysten genutzt, um soziale Trends zu erkennen. Die Nutzung dieser Quellen wirft aber auch ethische Fragen zur Vertraulichkeit auf. Vereinfacht ausgedrückt, bezieht sich Vertraulichkeit auf die Maßnahmen, die ergriffen werden, um sicherzustellen, dass Daten und betroffene Personen vor einer Identifizierung geschützt werden.

Quelle: [UK Statistics Authority, 2021](#); Übersetzung zum Teil mittels deepL.com

## Übersicht 3

### Dimensionen der UNECE: International Framework on Responsible AI for Official Statistics

Dimension	Kurzbeschreibung
Ethical Purpose, Public Good and Environmental Sustainability	KI/ML-Systeme sollten für Zwecke entwickelt und eingesetzt werden, die mit ethischen Standards übereinstimmen und dem öffentlichen Wohl dienen.
Accountability	Organisationen und Einzelpersonen, die für die Entwicklung und den Einsatz von KI/ML-Systemen verantwortlich sind, sollten für ihre Handlungen verantwortlich sein (also nicht das KI/ML-System) und alle unbeabsichtigten Folgen des Einsatzes von KI/ML-Systemen sollten korrigiert werden.
Transparency and Explainability	KI/ML-Systeme sollten in ihrer Konzeption, Entwicklung und Nutzung transparent sein. Sie sollten erklärbar sein, damit Einzelne verstehen können, wie KI/ML-Systeme Entscheidungen treffen und Maßnahmen ergreifen.
Fairness, Non-Discrimination, and Inclusiveness	KI/ML-Systeme sollten so konzipiert sein, dass sie Fairness gewährleisten und die Diskriminierung von Einzelpersonen oder Gruppen aufgrund von persönlichen Merkmalen wie Rasse, Geschlecht, Alter oder sozioökonomischem Status vermeiden. KI/ML-Systeme sollten aktiv vermeiden, bestehende Vorurteile zu verstärken.
Privacy and Security	KI/ML-Systeme sollten die Privatsphäre der einzelnen Personen respektieren und ihre persönlichen Daten schützen. KI/ML-Systeme sollten vor unbefugtem Zugriff und Cyberangriffen geschützt sein.
Validity and Robustness	KI/ML-Systeme sollten so konzipiert sein, dass sie valide und verlässliche Ergebnisse liefern, und sie sollten robust gegenüber Angriffen und Fehlern von außen sein.

Quelle: UNECE (Wirtschaftskommission für Europa der Vereinten Nationen), 2023 (Veröffentlichung in Vorbereitung)



### 3

## KI/ML in der deutschen amtlichen Statistik

Der Auftrag der amtlichen Statistik in Deutschland ist gesetzlich definiert, exemplarisch sei hier aus § 1 des Gesetzes über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) zitiert: „Die Statistik für Bundeszwecke (Bundesstatistik) hat im föderativ gegliederten Gesamtsystem der amtlichen Statistik die Aufgabe, laufend Daten über Massenerscheinungen zu erheben, zu sammeln, aufzubereiten, darzustellen und zu analysieren. [...] Durch die Ergebnisse der Bundesstatistik werden gesellschaftliche, wirtschaftliche und ökologische Zusammenhänge für Bund, Länder einschließlich Gemeinden und Gemeindeverbände, Gesellschaft, Wirtschaft, Wissenschaft und Forschung aufgeschlüsselt.“ Der Einsatz von KI/ML ist hierfür nicht zwingend erforderlich, kann jedoch in einigen typischen Anwendungsfällen zu einer effizienteren Statistikproduktion beitragen und die Qualität von Ergebnissen verbessern (Dumpert, 2024). Andere statistische Verfahren, deterministische Ansätze oder manuelle Tätigkeiten können ebenfalls geeignet sein, die Aufgabe des jeweiligen Teilprozesses zu bearbeiten; auch Kombinationen aus KI/ML und anderen Herangehensweisen sind denkbar. Im Folgenden werden einige typische Anwendungsfälle skizziert, die Liste ist jedoch nicht abschließend zu verstehen.

**Datenerhebung:** Ein Einsatzfeld für KI/ML stellt der Bereich der Bildverarbeitung dar, häufig für Fernerkundungsdaten (zum Beispiel Satelliten- oder Luftbilder). Solche Daten können – entsprechend aufbereitet – genutzt werden, um beispielsweise die Korrektheit eines Registers für Gebäude zu prüfen, das Vorhandensein von Swimmingpools, Steingärten, Solar- und Photovoltaikanlagen oder Baustellentätigkeiten zu ermitteln oder um aus dem beobachteten Vegetationsstand den Ertrag der Ernte auf einem Feld zu prognostizieren. Die aus Überflügen gelieferten Bilder müssen dazu hinsichtlich ihres Inhalts erfasst werden („Was ist auf diesem Bild zu sehen?“). Gerade dieses Einsatzfeld ist mit Methoden, die nicht auf KI/ML beruhen, häufig nicht effizient zu bearbeiten.

**Datenintegration:** Ein weiteres Beispiel für den Einsatz von KI/ML findet sich in der Zusammenführung von

Informationen aus verschiedenen Quellen. Dies kann beispielsweise durch Record Linkage oder Statistical Matching erreicht werden. KI/ML wird hier alleine oder ergänzend zu anderen statistischen Verfahren eingesetzt und steigert die Ergebnisqualität.

**Klassifikation und Codierung:** Die Klassifikation beziehungsweise Codierung von statistischen Einheiten anhand beschreibender Daten, die häufig als Text vorliegen, stellt ein weiteres Anwendungsfeld für KI/ML in der amtlichen Statistik dar. Für Unternehmen und Betriebe kann dies etwa die wirtschaftliche Tätigkeit betreffen, für Personen beispielsweise ihre berufliche Tätigkeit, für Artikel in einem Supermarkt etwa die Zuweisung eines Konsumgüterschlüssels. Basierend auf dieser Zuordnung werden später, also in einer späteren GMAS-Phase, separate Auswertungen zum Beispiel des mittleren Umsatzes, des Medianverdienstes oder von Preissteigerungen vorgenommen. Die Zuordnung zu einem Wirtschaftszweig kann darüber hinaus relevant sein, wenn es darum geht, die Auswahlgrundgesamtheit für eine Stichprobenziehung festzulegen.

**Validierung, Plausibilisierung und Imputation:** Die Prüfung auf Korrektheit der vorliegenden Daten (oder bereits gebildeter Aggregate) sowie die gegebenenfalls erforderliche Korrektur von als fehlerhaft eingestuften Einträgen eines Datensatzes sind ein weiteres Einsatzgebiet. Gleiches gilt prinzipiell für die Imputation fehlender Werte.

Anwendungsfälle von KI/ML sind in aller Regel den Phasen 4 (Daten gewinnen), 5 (Daten aufbereiten) und 6 (Ergebnisse analysieren) des GMAS (mit einem Schwerpunkt auf Phase 5) zuzuordnen. Es ist zu erkennen, dass der Einsatz von KI/ML in der Statistikproduktion meist – Ausnahmen werden noch genannt werden – keinen direkten Einfluss auf Menschen, Unternehmen oder Umwelt nimmt<sup>12</sup>; Wirkungen des Einsatzes entfalten sich allenfalls mittelbar. Beispiele für eine solche mittelbare Wirkung könnten die Förderfähigkeit oder -bedürftigkeit einer Branche (nicht eines Unternehmens!) für Subventionen oder mögliche Umlagezahlungen von oder an Körperschaften sein. Diese Wirkungen ergeben sich

12 Das dürfte die amtliche Statistik auch nicht unmittelbar. Sie dient nur dazu, Massenerscheinungen abzubilden, ihre Ergebnisse können für Planungs- und politische Steuerungsprozesse verwendet werden. Sie dient nicht dazu, Einzelfälle zu regeln oder zu beurteilen (Kühling, 2023, hier insbesondere Randnummer 11 der Kommentierung von § 1 Bundesstatistikgesetz auf Seite 39).

dann auf Basis eines im Vorfeld politisch beschlossenen Automatismus oder auf Basis eines in Interpretation und Würdigung der Zahlen erfolgten Entschlusses einer gesetzgebenden, ausführenden oder rechtsprechenden Institution. Die Heranziehung zu einer Statistik, gegebenenfalls unter Auskunftspflicht, kann – als Ausnahme vom zuvor Genannten – im Zuge einer Festlegung der Auswahlgrundlage für eine Stichprobenziehung mithilfe von KI/ML als unmittelbare Wirkung gegenüber einer Person, einem Haushalt oder einem Unternehmen oder Betrieb resultieren.

Diese Beispiele zielen auf die Verarbeitung der Daten in den Statistischen Ämtern des Bundes und der Länder ab. Analoges gilt prinzipiell für die Nutzung von KI/ML außerhalb der unmittelbaren Statistikproduktion in den statistischen Ämtern (einschließlich – je nach Abgrenzung – der Analyse). Beispiele hierfür könnten das Anfertigen von Übersetzungen, das Zusammenfassen von (insbesondere auch fremdsprachigen) Dokumenten, das Erstellen von Texten oder Designs anhand konkreter Vorgaben, ein Einsatz in der Personalverwaltung, ein KI/ML-gestützter First-Level-Support für die Befragten in Form eines Chatbots, die Generierung von Code in der Softwareentwicklung und vieles mehr sein. Diese Anwendungen stammen zum Teil aus dem Bereich der sogenannten generativen KI. Sie unterscheiden sich von den zuvor genannten Beispielen dadurch, dass zwar gegebenenfalls **mit**, jedoch nicht **an** den statistischen (Einzel-)Daten gearbeitet wird, und werden im Folgenden nicht betrachtet.<sup>13</sup>

## 4

### Identifikation ethischer Risiken

---

Die Auswertung der in den Kapiteln 1 und 2 vorgestellten Vorarbeiten ohne beziehungsweise mit Bezug zur amtlichen Statistik zeigen, dass die ethischen Werte Fairness,

Verantwortung, Transparenz, Genauigkeit, Schutz der Privatsphäre und Erklärbarkeit besonders wichtig und relevant in ihrer Ausgestaltung sind. Übertragen auf die Situation der deutschen amtlichen Statistik und unter Berücksichtigung der in Kapitel 3 genannten Beispiele und Anwendungsfälle von KI/ML konnten die folgenden zu vermeidenden Szenarien identifiziert werden, die ethische Risiken darstellen:

- a) **Produktion falscher Statistiken:** Wegen des Einsatzes von KI/ML werden Statistiken (beispielsweise Indikatoren oder Tabellen) fehlerhaft produziert. Dadurch entstehen – mittelbar – Fehlallokationen oder Fehlentscheidungen durch diejenigen, die amtliche Statistiken interpretieren oder für ihre Entscheidungsfindung heranziehen (beispielsweise durch die Politik).
- b) **Verstärkung unerwünschter Verzerrungen:** Durch den Einsatz von KI/ML werden unerwünschte Verzerrungen (etwa in den Trainingsdaten der Algorithmen) verstärkt. Dadurch wird der Auftrag, Massenerscheinungen zu beschreiben, durch die amtliche Statistik nicht mehr korrekt ausgeführt, was wiederum das Risiko der Produktion fehlerhafter Statistiken nach sich ziehen könnte.
- c) **Schaffung des „gläsernen Bürgers“:** Durch den Einsatz von KI/ML werden Informationen über eine statistische Einheit (etwa eine natürliche Person) in einer Weise verknüpft, die die statistische Einheit „gläsern“ macht und eine Profilbildung ermöglicht.
- d) **Übertragung der Verantwortung auf einen Algorithmus und fehlende Transparenz:** Die eingesetzten KI/ML-Verfahren werden für Entscheidungen verantwortlich gemacht, die im Laufe der Statistikproduktion zu treffen sind, das heißt keine Person trägt die Verantwortung für eine Entscheidung. Darüber hinaus können wesentliche Schritte weder nachvollzogen noch dokumentiert werden.
- e) **Offenlegen schützenswerter Informationen:** Schutzbedürftige Daten (unabhängig davon, ob schutzbedürftig aus Sicht des Statistikrechts, des Datenschutzes, der Verschwiegenheitspflicht der Beschäftigten oder einer anderen Norm) verlassen den geschützten Bereich der amtlichen Statistik. Diese Daten, die im Vertrauen auf eine Nutzung ausschließlich zu statistischen Zwecken, gegebenenfalls unter Auskunftspflicht, erteilt wurden, stehen dann Unbefugten zur Verfügung.

<sup>13</sup> Dargestellt werden im Folgenden nur die identifizierten Risiken, das heißt die Negativseiten. Die diesen entgegenstehenden Vorteile des Einsatzes von KI/ML für die Statistikproduktion wurden in Kapitel 3 skizziert. Auch darüber hinaus kann der Einsatz von KI/ML Vorteile bringen, indem er beispielsweise den nicht gestatteten oder nicht intendierten Abfluss von schutzbedürftigen Daten verhindert oder die Transparenz von Bearbeitungsschritten gerade durch den Einsatz eines leicht verständlichen Algorithmus erhöht. Die Vorteile des Einsatzes von KI/ML außerhalb der unmittelbaren Statistikproduktion werden in diesem Aufsatz nicht näher ausgeführt, da sie nicht spezifisch für die amtliche Statistik sind.


- f) **Fehlende ökologische Nachhaltigkeit:** Der Einsatz von KI/ML, insbesondere von sogenanntem Deep Learning und den zu dieser Klasse gehörenden Large Language Models, ist häufig rechenintensiv. Dadurch verursacht er einen extrem hohen Verbrauch an Energie und Wasser, insbesondere beim erstmaligen und beim gegebenenfalls erneut erforderlichen Training, jedoch auch in geringerem Umfang bei der Anwendung. Dies ist besonders dann problematisch, wenn der Einsatz dieser ökologisch unter Umständen nicht nachhaltigen Klasse von Verfahren gar nicht erforderlich wäre (beispielsweise, weil nachhaltigere Verfahren annähernd ebenso gute Ergebnisse liefern).

## 5

---

### Zwischenfazit

---

Im vorliegenden ersten Teil der zweiteiligen Betrachtung ethischer Fragen beim Einsatz von KI/ML in der Produktion amtlicher Statistiken wurden anhand von Einsatzbereichen von KI/ML in Deutschland und von internationalen Vorarbeiten ethische Risiken identifiziert. Während einige andere Staaten und zwischenstaatliche Organisationen bereits separate ethische Richtlinien zum Einsatz von KI/ML in der amtlichen Statistik veröffentlicht haben, steht eine Einordnung für die deutsche amtliche Statistik noch aus. Der zweite Teilaufsatz („Auseinandersetzung“) greift die Fragestellungen aus Kapitel 4 einzeln auf und führt die geltenden Referenzrahmen Recht und Qualität ein, um den einzelnen Punkten zu begegnen (Dumpert und andere, 2025). 



## LITERATURVERZEICHNIS

---

Allhutter, Doris/Cech, Florian/Fischer, Fabian/Grill, Gabriel/Mager, Astrid. *Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective*. In: Frontiers in Big Data. Jahrgang 3. 2020. Seite 1 ff. DOI: [10.3389/fdata.2020.00005](https://doi.org/10.3389/fdata.2020.00005)

Blumöhr, Torsten/Teichmann, Corina/Noack, Anke. [\*Standardisierung der Prozesse: 14 Jahre AG SteP\*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 5/2017, Seite 58 ff.

Deutscher Ethikrat. *Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz*. 2023. [Zugriff am 6. Dezember 2024]. Verfügbar unter: [www.ethikrat.org](http://www.ethikrat.org)

Dumpert, Florian. [\*Maschinelles Lernen im Statistischen Bundesamt. Ein Überblick über die Historie seit 2015 und aktuelle Entwicklungen\*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 4/2024, Seite 17 ff.

Dumpert, Florian/Reichel, Jannik/Oertel, Elisa/Leerhoff, Holger/Salwiczek, Christian. [\*Ethische Fragen beim Einsatz von KI/ML in der Produktion amtlicher Statistiken – Teil 2: Auseinandersetzung\*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 1/2025, Seite 25 ff.

Fenger, Menno/Simonse, Robin. *The implosion of the Dutch surveillance welfare state*. In: Social Policy & Administration. Jahrgang 58. Ausgabe 2/2024, Seite 264 ff. DOI: [10.1111/spol.12998](https://doi.org/10.1111/spol.12998)

Ferdinand, Jan-Peter/Kind, Sonja. *Big Social Data – die gesellschaftspolitische Dimension von Prognose- und Ratingalgorithmen*. In: Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag – Themenkurzprofil. 2018. DOI: [10.5445/IR/1000133904](https://doi.org/10.5445/IR/1000133904)

Kommission Zukunft Statistik (KomZS). *Bericht und Empfehlungen der Kommission Zukunft Statistik. Version 1.0*. 2024. [Zugriff am 6. Dezember 2024]. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

Kühling, Jürgen. *Bundesstatistikgesetz: BStatG*. Kommentar. München 2023.

Lopez, Paola. *Reinforcing Intersectional Inequality via the AMS Algorithm in Austria*. In: Proceedings of the STS Conference. Graz 2019. Seite 289 ff. DOI: [10.3217/978-3-85125-668-0-16](https://doi.org/10.3217/978-3-85125-668-0-16)

Nezik, Ann-Kathrin. *Hast du ein Bewusstsein?* In: Die ZEIT. Ausgabe 3 vom 12. Januar 2023 (aktualisiert am 15. Januar 2023).

Novotny, Rudi/von Lindern, Jakob. *Alles unter Kontrolle ...* In: Die ZEIT. Ausgabe 16 vom 13. April 2023 (aktualisiert am 16. April 2023).

OECD (Organisation für wirtschaftliche Zusammenarbeit und Entwicklung). *OECD AI Principles overview*. 2024. [Zugriff am 6. Dezember 2024]. Verfügbar unter: [oecd.ai](https://oecd.ai)

## LITERATURVERZEICHNIS

---

Renda, Andrea. *Europe: Toward a Policy Framework for Trustworthy AI*. In: Dubber, Markus D./Pasquale, Frank/Das, Sunit (Herausgeber). *The Oxford Handbook of Ethics of AI*. 2020. Oxford University Press. Seite 650 ff.

DOI: [10.1093/oxfordhb/9780190067397.013.41](https://doi.org/10.1093/oxfordhb/9780190067397.013.41)

Schweizer Bundesrat. *Verhaltenskodex des Bundes für menschenzentrierte und vertrauenswürdige Datenwissenschaft*. 2023. [Zugriff am 6. Dezember 2024]. Verfügbar unter: [www.bfs.admin.ch](https://www.bfs.admin.ch)

Seising, Rudolf. *Von der Mannigfaltigkeit der KI-Geschichte*. In: zur debatte. Jahrgang 54. Ausgabe 2/2024, Seite 46 ff.

Statistics Canada. *Framework for Responsible Machine Learning Processes at Statistics Canada*. 2021. [Zugriff am 8. Januar 2025].

Verfügbar unter: [www150.statcan.gc.ca](https://www150.statcan.gc.ca)

UK Statistics Authority. *Ethical considerations in the use of Machine Learning for research and statistics*. 2021. [Zugriff am 6. Dezember 2024].

Verfügbar unter: [uksa.statisticsauthority.gov.uk](https://uksa.statisticsauthority.gov.uk)

UNECE (Wirtschaftskommission für Europa der Vereinten Nationen). *International Framework on Responsible AI for Official Statistics*. 2023 (Veröffentlichung in Vorbereitung).

United Nations (Vereinte Nationen). *Follow-up on the Conference of European Statisticians seminar “Data ethics – a key enabler of social acceptability”*. 2024. [Zugriff am 6. Dezember 2024]. Verfügbar unter: [unece.org/](https://unece.org/)

Weizenbaum, Joseph. *Die Macht der Computer und die Ohnmacht der Vernunft*. Frankfurt am Main 1977.

## RECHTSGRUNDLAGEN

---

Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) in der Fassung der Bekanntmachung vom 20. Oktober 2016 (BGBl. I Seite 2394), das zuletzt durch Artikel 14 des Gesetzes vom 8. Mai 2024 (BGBl. I Nr. 152) geändert worden ist.

# ETHISCHE FRAGEN BEIM EINSATZ VON KI/ML IN DER PRODUKTION AMTLICHER STATISTIKEN – TEIL 2: AUSEINANDERSETZUNG

Florian Dumpert, Jannik Reichel, Elisa Oertel, Holger Leerhoff,  
Christian Salwiczek

➤ **Schlüsselwörter:** Künstliche Intelligenz – Maschinelles Lernen – Ethik –  
Recht – Qualität

## ZUSAMMENFASSUNG

Auch die amtliche Statistik nutzt Methoden der Künstlichen Intelligenz (KI) und ihres Teilbereichs Maschinelles Lernen (ML). Dies wirft verschiedene ethische Fragestellungen auf, die im ersten Teil des Artikels („Identifikation“) mithilfe von Vorarbeiten aus anderen Staaten und von zwischenstaatlichen Organisationen ermittelt wurden. Der vorliegende zweite Teil des Artikels greift diese Fragestellungen einzeln auf und beleuchtet sie unter den geltenden Referenzrahmen Recht und Qualität. Dieses Vorgehen dient der Prüfung, ob es eines separaten ethischen Rahmenwerkes für den Einsatz von KI/ML in der amtlichen Statistik bedarf, oder ob geltende, zum Teil nicht KI/ML-spezifische Richtlinien ausreichen, um den Fragestellungen hinreichend zu begegnen.

➤ **Keywords:** artificial intelligence – machine learning – ethics – legislation –  
quality

## ABSTRACT

*Methods of artificial intelligence (AI) and of machine learning (ML), as one of its subsets, are also used in official statistics. This presents a variety of ethical issues, which have been identified in the first part of the article (“Identification”) using previous work carried out in other countries and by inter-governmental organisations. This second part of the article looks at these issues individually and examines them in the light of the legislation and quality reference frameworks that currently apply. The objective is to explore whether there is a need for a separate ethical reference framework for the use of AI/ML in official statistics or whether the existing guidelines, some of which do not relate specifically to AI/ML, are sufficient to adequately address the issues identified.*

### Dr. Florian Dumpert

ist Diplom-Mathematiker und leitet das Referat „Künstliche Intelligenz, Big Data“ des Statistischen Bundesamtes. Im Statistischen Verbund der Statistischen Ämter des Bundes und der Länder leitet er im Co-Vorsitz-Modell den Arbeitskreis Maschinelles Lernen.

### Jannik Reichel

hat einen Master in Business Intelligence and Smart Services absolviert und arbeitete zum Erstellungszeitpunkt der Aufsätze im Referat „Künstliche Intelligenz, Big Data“ des Statistischen Bundesamtes als Data Scientist an einem Projekt zur Frühschätzung der Einkommenssteuerstatistiken mittels Machine Learning und KI.

### Elisa Oertel

ist Bioinformatikerin und leitet ein Data-Science-Team im Landesamt für Statistik Niedersachsen. Auch als Vorstandsmitglied in einem deutschlandweiten Philosophieverein widmet sie sich der Verankerung ethischer Prinzipien in der technologischen Entwicklung.

### PD Dr. Holger Leerhoff

leitet im Amt für Statistik Berlin-Brandenburg das Referat „Querschnittsanalysen, EXSTAT und Geoservice“. Der Privatdozent für Philosophie an der Universität Oldenburg ist Mitglied verschiedener Arbeitskreise des Statistischen Verbundes, etwa zum Maschinellen Lernen und zur Nutzung neuer digitaler Daten.

### Christian Salwiczek

leitet das Referat „Unternehmensregister/Profiling/LCU“ beim Statistischem Amt Nord und hat die Co-Leitung des Analyse-Netzwerks des Statistischen Verbundes inne. In verschiedenen Arbeitskreisen arbeitet er an der Modernisierung der amtlichen Statistik mit und befasst sich in unterschiedlichen Gremien auch mit ethischen Fragestellungen.

## 1

### Einleitung

Die amtliche Statistik ist durch viele Gesetze und durch Qualitätsanforderungen geprägt. Gleichzeitig spielen Künstliche Intelligenz (KI) und deren Teilbereich Maschinelles Lernen (ML) auch dort eine wichtige Rolle. Einige Anwendungsfälle stellt der erste Teil des Artikels vor (Dumpert und andere, 2025). Ebenso wurden dort anhand von Vorarbeiten internationaler statistischer Institutionen sechs ethische Risiken identifiziert, die beim Einsatz von KI/ML in der Produktion amtlicher Statistiken eintreten können: a) Produktion falscher Statistiken, b) Verstärkung unerwünschter Verzerrungen, c) Schaffung des „gläsernen Bürgers“, d) Übertragung der Verantwortung auf einen Algorithmus und fehlende Transparenz, e) Offenlegen schützenswerter Informationen und f) Fehlende ökologische Nachhaltigkeit. Diese Risiken greift der vorliegende zweite Teil des Artikels auf. Er erörtert die Frage, ob die deutsche amtliche Statistik ebenfalls ein ethisches Rahmenwerk benötigt, wie es beispielsweise in Kanada oder im Vereinigten Königreich existiert (Statistics Canada, 2021; UK Statistics Authority, 2021). Kapitel 2 dieses Aufsatzes führt hierzu die Referenzrahmen, die für die amtliche Statistik in Deutschland gelten, ein, die im dritten Kapitel auf die oben genannten, bereits identifizierten ethischen Risiken angewendet werden. Offene Fragestellungen behandelt Kapitel 4. Der Abgleich mit den im ersten Teil des Artikels genannten Rahmenwerken internationaler statistischer Institutionen findet in Kapitel 5 statt. Das Fazit wird schließlich im Kapitel 6 gezogen.

## 2

### Referenzrahmen

#### 2.1 Einordnung

Ein Referenzrahmen gibt vor, wie Fragestellungen aus bestimmten Bereichen zu begegnen ist – im Kontext dieses Artikels sind dies ethische Fragestellungen. Dabei sind verschiedene Referenzrahmen denkbar. Beispiele sind das Recht (Verfassung, Gesetze, Verordnungen und

so weiter), informelle Normen (Traditionen, religiöse und weltanschauliche Normen und so weiter) und Selbstbindungen, die offen oder verdeckt bestehen können. Für Deutschland gilt das Recht als ein für alle verbindlicher Referenzrahmen. Privatpersonen und Unternehmen steht es immer frei, den Maßstab für sich strenger anzulegen und auf bestimmte Handlungen zu verzichten, die nach dem Recht zulässig wären. Behörden sind jedoch durch den Grundsatz der Gesetzmäßigkeit der Verwaltung (Artikel 20 Absatz 3 Grundgesetz) eingeschränkt, das heißt für jegliches Verwaltungshandeln ist eine Rechtsgrundlage erforderlich. Bestehen gesetzliche Verpflichtungen, so ist diesbezüglich ein Unterlassen seitens einer Behörde nicht zulässig. Bei Ermessensnormen ist sie nicht zum Handeln verpflichtet, muss aber bei Ausübung des Ermessens rechtliche Grundregeln beachten.

Bindend für die amtliche Statistik sind zwei Referenzrahmen. Zum einen – vorrangig – das Recht mit seinen für die amtliche Statistik einschlägigen Normen und bindenden Gerichtsbeschlüssen. Zum anderen (in obigem Sinne rechtlich zulässige) Selbstverpflichtungen, die durch die amtliche Statistik selbst eingegangen wurden. Diese finden sich wieder in den Qualitätsrahmenwerken der amtlichen Statistik in Deutschland und Europa.

#### 2.2 Recht

Prominente Beispiele für einschlägige Normen im Referenzrahmen Recht sind das Bundesstatistikgesetz, die Fachstatistikgesetze (mit einer zum Teil sehr detaillierten Festlegung, welche Merkmale in welchen Ausprägungen überhaupt durch die amtliche Statistik erhoben werden dürfen) und das Urteil des Bundesverfassungsgerichts zur Volkszählung aus dem Jahr 1983<sup>1</sup>. Der mit wenigen Ausnahmen gültige Grundsatz „Keine Statistik ohne Gesetz“ stellt sicher, dass es das Parlament, also die Legislative, ist, das im Rahmen der Gesetzgebung die Abwägung auch ethischer Aspekte vornimmt. Die Gesellschaft verhandelt hier, was sie zulassen möchte und was nicht. Für den Fall, dass Höherrangiges dabei außer Acht bleibt oder bewusst ausgeblendet wird, steht bezüglich der Norm an sich der Weg zum Bundesverfassungsgericht offen. Eine Darstellung und Einord-

---

<sup>1</sup> BVerfG, Urteil vom 15. Dezember 1983 - 1 BvR 209/83, 1 BvR 484/83, 1 BvR 440/83, 1 BvR 420/83, 1 BvR 362/83, 1 BvR 269/83 - Rn. (1 - 215).

nung der Entwicklung des Statistikrechts in Deutschland findet sich in Leischner/Kolbe (2024), eine aktuelle Kommentierung des Bundesstatistikgesetzes in Kühling (2023). Darüber hinaus enthält die europäische KI-Verordnung (2024/1689) – insoweit einschlägig für die Anwendungsfälle der amtlichen Statistik – Vorgaben zu erlaubten und untersagten Einsatzzwecken sowie zu Transparenz und Qualitätssicherung von KI/ML. Soweit personenbezogene Daten verarbeitet werden, ist die europäische Datenschutz-Grundverordnung (2016/679) maßgeblich zu beachten.

### 2.3 Qualität

---

Mit den verschiedenen Rahmenwerken zur Qualität, die als Selbstverpflichtung, aber hinsichtlich einiger Qualitätskriterien auch über den Artikel 12 der europäischen Statistikverordnung (223/2009) verbindlich für die deutsche amtliche Statistik sind, ist ein weiterer Referenzrahmen geschaffen. Die Rahmenwerke (Quality Assurance Framework, Code of Practice, Qualitätshandbuch) umfassen Vorgaben zur Institution (das statistische Amt), zu deren Prozessen und zu deren Produkten. Grundsätze sind hier beispielsweise fachliche Unabhängigkeit, solide Methodik, Genauigkeit und Zuverlässigkeit, Aktualität und Pünktlichkeit, oder Zugänglichkeit und Klarheit (Statistische Ämter des Bundes und der Länder, 2021, hier: Kapitel 3). Eine Ableitung von Qualitätsdimensionen für KI/ML in der amtlichen Statistik aus den allgemeinen Qualitätsrahmenwerken sowie ein Vorschlag für eine Operationalisierung derselben in Form von Qualitätsrichtlinien finden sich in Saidani und andere (2023). Die dort genannten Qualitätsdimensionen hat das zuständige Gremium (das Abteilungsleitungsgremium Fachstatistik – ALG FS) mit Beschluss vom 10. Oktober 2023 für die Statistikproduktion der Statistischen Ämter des Bundes und der Länder übernommen, die Qualitätsrichtlinien anschließend in leicht abgewandelter Form am 8. Oktober 2024.

Der oben genannte Vorrang des Rechts bezieht sich zum einen auf den zweiten Referenzrahmen (Qualität) und würde zum anderen auch für ein hypothetisches ethisches Rahmenwerk gelten. Dass ein von einer Behörde eingesetztes ethisches Rahmenwerk nicht höherrangiger sein kann als vom Parlament geschaffene Gesetze, ergibt sich aus Artikel 20 Absatz 3 Grundgesetz. Es müsste außerdem im Einklang mit den Werten der Ver-

fassung, insbesondere den Grundrechten, einschließlich der EU-Grundrechtecharta, und den Staatszielen stehen.

## 3

---

### Begegnung der ethischen Fragestellungen

---

#### 3.1 Herangehensweise

---

Die Autorin und die Autoren erkennen an, dass eigene Rahmenwerke für ethische Fragestellungen auch in der amtlichen Statistik ihre Berechtigung haben können. Diese sind dann komplementär zu anderen Rahmenwerken, beispielsweise zur Qualität, zu sehen und könnten dazu dienen, aus vielen Referenzrahmen stammende Zielvorgaben gemeinsam zu betrachten. Außerdem könnten sie Hilfestellung bieten im Umgang mit Zielkonflikten (zum Beispiel Schnelligkeit der Bereitstellung oder Genauigkeit von amtlichen Statistiken gegenüber ökologischer Nachhaltigkeit der Produktion derselben). Ebenfalls wird anerkannt, dass das Recht im Allgemeinen nicht jede Fragestellung antizipieren und im Detail regeln kann, etwa weil bei Schaffung der Rechtsnorm ein Umstand noch gar nicht absehbar war, da sich die Gesellschaft oder die Technologie weiterentwickeln. Hier kann ein parallel existierendes ethisches Rahmenwerk hilfreich sein (Floridi, 2018; Renda, 2020).

Für die deutsche amtliche Statistik gehen die Autorin und die Autoren einen anderen Weg. Statt ein weiteres ethisches Rahmenwerk zu entwickeln und einzuführen, wurden – in Anerkennung obiger Tatsachen – zunächst ethisch kritische Situationen in Dumpert und andere (2025, hier: Kapitel 4) identifiziert, die in der deutschen amtlichen Statistik auftreten können. In den folgenden Abschnitten dieses Kapitels wird geprüft, ob die Umsetzung<sup>2</sup> bestehender Vorgaben aus den in Kapitel 2 genannten Referenzrahmen bereits ausreicht, um diese ethisch kritischen Situationen zu vermeiden.

---

2 Die Betonung der Umsetzung ist erforderlich, denn das Bestehen zum Beispiel rechtlicher Vorgaben bedeutet nicht, dass diese auch eingehalten werden: Ihre Beachtung kann aus Unkenntnis oder aufgrund von Sachzwängen unterbleiben.



### 3.2 Produktion falscher Statistiken

---

**Vorgaben:** Aus den Grundsätzen, dass Methodik und eingesetzte Verfahren dem aktuellen Stand der wissenschaftlichen Forschung entsprechen und die Ergebnisse einer Statistik genau sein sollen<sup>3</sup>, folgt die Verpflichtung, diese Genauigkeit herbeizuführen und auch zu überprüfen. Es liegt dabei jedoch im Wesen einer Statistik und auch aller statistischen Verfahren (einschließlich KI/ML), dass die möglicherweise auf einen Einzelfall angewandten Schätzungen und Prädiktionen häufig nicht exakt zutreffen. Dies ist der Preis dafür, dass nicht alle benötigten Informationen messfehlerfrei und ohne Abgrenzungs- und Interpretationsschwierigkeiten unmittelbar erhoben werden können, sondern zumindest teilweise geschätzt werden. Auf den konkreten Anwendungsfall KI/ML heruntergebrochen – und analog auch für den Einsatz anderer Verfahren gültig – mag die Zuordnung eines bestimmten Betriebs zu einem Wirtschaftszweig nicht korrekt sein. Solange jedoch sichergestellt ist, dass die Statistik an sich den Genauigkeitsansprüchen genügt, ist dies unkritisch. Das ethische Risiko einer fehlerhaften Statistik wird also bereits durch den Referenzrahmen Qualität adressiert.

**Umsetzung:** Die Prüfung der Genauigkeit obliegt den für eine Statistik zuständigen Fachbereichen, das heißt die Bewertung, (i) ob fortlaufend eine hinreichende Genauigkeit erreicht wird, und die darauf aufbauende Entscheidung, (ii) ob und für welche Beobachtungen ein KI/ML-Verfahren eingesetzt wird.<sup>4</sup> Sollte darüber hinaus einem Einzelfall, im obigen Beispiel einem Betrieb, ein unmittelbarer rechtlicher Nachteil aus einer falschen Zuordnung oder einer grob falschen Schätzung entstehen (etwa weil er nun irrtümlich für eine Statistik auskunftspflichtig wird, die eine andere Branche betrifft), hat der Betrieb die Möglichkeit, die KI/ML-basierte oder -unterstützte Zuordnung infrage zu stellen. Es erfolgt dann gegebenenfalls eine manuelle Korrektur des Eintrags.

---

3 Grundsätze G07, G08 und G12 des Qualitätshandbuchs (Statistische Ämter des Bundes und der Länder, 2021).

4 Dies setzen die Fachbereiche beispielsweise durch stichprobenartige Prüfungen von Ergebnissen aus einem KI/ML-Verfahren um. Darüber hinaus besteht die Möglichkeit, für die Statistik besonders einflussreiche Beobachtungen (zum Beispiel aufgrund ihres im Vergleich besonders großen Umsatzes) stets manuell zu plausibilisieren, gegebenenfalls mittels direkter Rücksprache bei der oder dem Meldenden oder mittels Bestätigung der Schätzung durch die Meldende oder den Meldenden.

### 3.3 Verstärkung unerwünschter Verzerrungen

---

**Vorgaben:** Saidani und andere (2023) betrachtet das Konzept Fairness<sup>5</sup> als einen übergreifenden Aspekt und fasst dieses als Umstand, dass statistische Aggregate für bestimmte Subgruppen nicht systematisch über- oder unterschätzt werden (Saidani und andere, 2023, hier: Seite 292 f.; zu diesem Fairnessbegriff siehe auch Kim und andere, 2019). Die amtliche Statistik muss die Verhältnisse so darstellen, wie sie vorherrschen. Wie darauf reagiert wird, ist eine politische, ökonomische oder persönliche Entscheidung – mithin keine, die die amtliche Statistik trifft. Für die korrekte, nicht verzerrende Darstellung ist entscheidend, dass alle relevanten Subgruppen mit gleicher Güte durch die statistischen Verfahren im Produktionsprozess behandelt werden. Das ist durch die amtliche Statistik bereits durch die Vorgaben zur Qualität sicherzustellen, weswegen die Verstärkung unerwünschter Verzerrungen kein offenes, KI/ML-spezifisches ethisches Risiko darstellt.

**Umsetzung:** Sofern möglich, wird der oben genannte Fairnessaspekt bereits während der methodischen Untersuchung geprüft, ob ein und gegebenenfalls welches KI/ML-Verfahren in einem Teilschritt der Statistikproduktion eingesetzt werden kann. Hierzu ist es erforderlich, dass von Seiten der Fachstatistik die relevanten Subgruppen benannt werden. Im Produktionsprozess nachgelagerte Prüfungen finden beispielsweise im GMAS<sup>6</sup>-Teilprozess 6.2 „Ergebnisse validieren“ statt (etwa durch den Abgleich der Statistik mit anderen relevanten Daten, die Untersuchung von Unstimmigkeiten in der Statistik und die Validierung der Statistik bezüglich Erwartungen und bereichsbezogenem Fachwissen).

### 3.4 Schaffung des „gläsernen Bürgers“

---

**Vorgaben:** Das ethische Risiko, einen „gläsernen Bürger“ zu schaffen, kann in der deutschen amtlichen Statistik nicht eintreten, da es bereits durch den rechtlichen Referenzrahmen ausgeschlossen wird: „[E]ine umfas-

---

5 Bothmann und andere (2024) diskutiert den Fairnessbegriff breiter, Schenk/Kern (2024) liefert eine ausführliche Diskussion des Zusammenhangs von Fairness und Qualität in der amtlichen Statistik und Bothmann/Peters (2024) betrachtet Fairness als Qualitätskriterium im Maschinellen Lernen am Beispiel von Mietspiegeln.

6 Geschäftsprozessmodell Amtliche Statistik; siehe Blumöhr und andere, 2017, hier: Kapitel 3.

sende Registrierung und Katalogisierung der Persönlichkeit durch die Zusammenführung einzelner Lebens- und Personaldaten zur Erstellung von Persönlichkeitsprofilen der Bürger ist auch in der Anonymität statistischer Erhebungen unzulässig“<sup>17</sup>. Verschiedene Regelungen in statistischen Gesetzen verhindern eine Profilbildung. So ist unter anderem die Wiederherstellung eines Personenbezugs nach Löschung von identifizierenden Hilfsmerkmalen (wie Name oder Anschrift) oder nach Aggregation von Einzeldaten nach § 21 Bundesstatistikgesetz verboten. Auch die Möglichkeit, statistische Daten zusammenzuführen, ist in § 13a Bundesstatistikgesetz und in anderen fachstatistischen Gesetzen abschließend geregelt. Für das Verbot der Profilbildung ist unerheblich, wie und mit welchen Mitteln diese erfolgt. Ob KI/ML erweiterte Möglichkeiten zur Profilbildung bietet als bisher verfügbare Techniken, spielt somit für dieses ethische Risiko keine Rolle.

**Umsetzung:** Die Beschäftigten in den Statistischen Ämtern des Bundes und der Länder sind bezüglich der rechtlichen Vorgaben besonders sensibilisiert. Organisatorische und technische Maßnahmen wirken außerdem einem Datenzugriff einzelner Beschäftigter über das für die Produktion der jeweiligen Fachstatistik notwendige Maß hinaus entgegen.

### 3.5 Übertragung der Verantwortung auf einen Algorithmus und fehlende Transparenz

---

**Vorgaben:** Aus juristischer Sicht ist das handelnde Rechtssubjekt die öffentlich-rechtliche juristische Person, vertreten durch die Behörde oder Körperschaft selbst, hier also das statistische Amt, und nicht eine Maschine oder ein Algorithmus. Selbst bei tatsächlich vollständig automatisiert ablaufenden Vorgängen wird dies der jeweiligen Behörde zugerechnet, da diese handelt und sich hierfür der Maschine lediglich bedient. Artikel 22 Absatz 1 Datenschutz-Grundverordnung räumt darüber hinaus Menschen das Recht ein, nicht zum Gegenstand automatisierter Entscheidungsfindung gemacht zu werden. Dies kann im Grundsatz auch auf Prozesse der amtlichen Statistik zutreffen. Die Durchfüh-

rung von Prozessschritten durch einen Algorithmus darf außerdem nicht dazu führen, dass wesentliche Schritte der behördlichen Tätigkeit weder nachvollzogen noch dokumentiert werden (Verletzung von Transparenz- und Rechenschaftspflichten). Denn dies könnte dazu führen, dass eine behördliche Tätigkeit nicht mehr gerichtlich kontrolliert werden kann (Verletzung der Rechtsweggarantie, Artikel 19 Absatz 4 Grundgesetz) und dass Kontrollinstanzen (Aufsichtsbehörden, Parlamente) das Verwaltungshandeln nicht mehr effektiv kontrollieren können (Verletzung des Grundsatzes der Rechtmäßigkeit der Verwaltung, Artikel 20 Absatz 3 Grundgesetz). Das Risiko, dass die Verantwortung alleine einem Algorithmus zugewiesen wird, wird darüber hinaus durch die allgemeine Qualitätssicherung für amtliche Statistiken adressiert, die bei den Fachreferentinnen und Fachreferenten der Statistischen Ämter des Bundes und der Länder liegt (Saidani und andere, 2023, hier: Seite 22). Darüber hinaus besagt G01.4 des Qualitätshandbuchs (Statistische Ämter des Bundes und der Länder, 2021): „Die Leiterinnen und Leiter der nationalen statistischen Ämter und von Eurostat und gegebenenfalls die Leiterinnen und Leiter anderer statistischer Stellen tragen die alleinige Verantwortung für die Festlegung der statistischen Methoden, Standards und Verfahren sowie des Inhalts und des Zeitplans der statistischen Veröffentlichungen.“ Vorgaben zur Transparenz aus Qualitätssicht ergeben sich durch die Grundsätze G06.4 (öffentliche Zugänglichkeit der verwendeten Methoden und Verfahren) und G15.6 (Information der Nutzerinnen und Nutzer über die Methodik der statistischen Prozesse) des Qualitätshandbuchs sowie durch die Qualitätsdimension „Erklärbarkeit“ in Saidani und andere (2023, hier: Seite 279 ff.).

**Umsetzung:** Während die formelle Lage klar ist, muss bei der Umsetzung immer wieder aufs Neue darauf geachtet werden, dass die Verantwortung auch tatsächlich bewusst wahrgenommen wird. Dazu gehört – wie bei jeder Automatisierung – zum einen, dass diese in der Konzeption kritisch geprüft und getestet wird, und zum anderen, dass im laufenden Betrieb immer wieder stichprobenartige Prüfungen der Ergebnisse vorgenommen werden. Auch können besonders einflussreiche Fälle stets vom Menschen bearbeitet oder bei automatisierter Bearbeitung deren Ergebnisse zumindest explizit und häufig manuell geprüft werden. Darüber hinaus sind alle

---

<sup>17</sup> BVerfG, Urteil vom 15. Dezember 1983 - 1 BvR 209/83, 1 BvR 484/83, 1 BvR 440/83, 1 BvR 420/83, 1 BvR 362/83, 1 BvR 269/83 - Rn. (169).

Beteiligten für den sogenannten Automation Bias<sup>8</sup> zu sensibilisieren; sofern möglich, ist dieser durch entsprechende Gestaltung der Prozesse abzumildern. Bezüglich der Transparenz gilt, dass Codes in der Regel zwar nicht öffentlich zugänglich gemacht werden, jedoch die zugrunde liegenden KI/ML-Algorithmen frei verfügbare R-Packages oder Python-Module und somit vollkommen transparent sind.

### 3.6 Offenlegen schützenswerter Informationen

---

**Vorgaben:** Die jeweils einschlägigen Rechtsnormen untersagen schädliches Handeln und gemäß den Rechtsnormen nicht untersagtes Handeln ist aus Sicht des Gesetzgebers akzeptabel. Neben der Verpflichtung aus § 16 Bundesstatistikgesetz zur Geheimhaltung statistischer Einzeldaten über persönliche und sachliche Verhältnisse haben nach Artikel 20 Absatz 4 der europäischen Statistikverordnung das Statistische Bundesamt „und andere einzelstaatliche Stellen und die Kommission (Eurostat) innerhalb ihrer jeweiligen Zuständigkeitsbereiche alle erforderlichen rechtlichen, administrativen, technischen und organisatorischen Maßnahmen [zu ergreifen], um den physischen und logischen Schutz vertraulicher Daten zu gewährleisten“. Die allgemeine Verpflichtung zur Informationssicherheit (insbesondere mit dem Schutzziel Vertraulichkeit) folgt aus der europäischen NIS-2-Richtlinie und – zumindest für die Bundesverwaltung – aus § 8 Absatz 1 BSI-Gesetz und dem Umsetzungsplan Bund 2017 (Bundesministerium des Innern, 2017). Wenn personenbezogene Daten verarbeitet werden, darf dies nur auf Grundlage einer bestimmten und geeigneten Rechtsgrundlage erfolgen. Eine Datenübermittlung an externe Dienstleister macht regelmäßig einen Auftragsverarbeitungsvertrag erforderlich. Die Einhaltung der datenschutzrechtlichen Vorgaben ist gerade auch bei Auftragsverarbeitern zu dokumentieren und zu kontrollieren.

---

8 „Automatisierungsbias [...] beschreibt die menschliche Neigung, von automatisierten Systemen erzeugte Empfehlungen oder Entscheidungen übermäßig stark zu gewichten; und dies selbst in Situationen, in denen die von diesen Systemen bereitgestellten Informationen falsch sind oder in denen menschliches Urteilsvermögen genauer wäre. Dieses kognitive Phänomen verdeutlicht die Neigung von Menschen, Maschinen übermäßiges Vertrauen entgegenzubringen, obwohl sie über besseres Wissen oder Verständnis verfügen.“ (Übersetzung aus Ruschmeier/Hondrich [2024], Seite 1).

**Umsetzung:** Unbestritten besteht ein praktisches Risiko für den Abfluss von Daten, beispielsweise aus Versehen bei der Eingabe von Daten in ein Online-Übersetzungstool oder bei der Weitergabe von Modellen an Adressaten außerhalb der amtlichen Statistik<sup>9</sup>. Dieses Risiko wird durch den Einsatz von kommerziellen (generativen) KI/ML-Systemen größer, ist jedoch nicht KI/ML-spezifisch. Auch dies ist kein eigens anzusprechendes ethisches Risiko, sondern vielmehr eine Frage der – bereits existierenden – Sensibilisierung der Beschäftigten oder von technischen Lösungen. Diese Einschätzung gilt unabhängig von der im konkreten Fall eingesetzten Infrastruktur (zum Beispiel On-Premise oder Cloud).

## 4

---

### Offene Fragestellung: Ökologische Nachhaltigkeit

---

**Vorgaben:** Im Jahr 2020 wurde ein Rahmenwerk vorgeschlagen, das den Energieverbrauch und CO<sub>2</sub>-Ausstoß bei der Nutzung von Maschinellern greifbar machen sollte (Henderson und andere, 2020). Bereits zuvor wurden verschiedene Ansätze entwickelt, den Energieverbrauch von Maschinellern zu erfassen (García-Martín und andere, 2019). Samsi und andere (2023) führt eine solche Untersuchung für Large Language Models durch. Jüngst wurde dieser Punkt zudem in einer öffentlichen Anhörung zum Thema „Künstliche Intelligenz als Schlüsseltechnologie“ aufgegriffen (Deutscher Bundestag, 2024). Dass ethische Grundsätze nicht nur menschen-, sondern ebenso „planetenzentriert“ sein sollen, schlägt auch Renda (2020, hier: Seite 660) vor. Weiters erkennt Deutschland an, „dass die Umweltauswirkungen der KI- und IKT-Infrastruktur, einschließlich der direkten Auswirkungen wie Energie- und Wasserverbrauch sowie der indirekten ermöglichen und systemischen Auswirkungen der Anwendungen, systematisch und standardisiert über die Sektoren hinweg gemessen werden müssen“ (OECD, 2024, hier:

---

9 So enthalten beispielsweise Support Vector Machines per Konstruktion statistische Einzeldaten im gelernten Modell, nämlich die namensgebenden Support-Vektoren. Diese sind erforderlich, um eine Support Vektor Machine zur Klassifikation oder Regression verwenden zu können. Auch k-Nearest-Neighbour-Modelle enthalten per definitionem Einzeldaten (im ungünstigsten Fall das gesamte Datenmaterial).



Seite 158 f.). Nach Kenntnis der Autorin und der Autoren wird der Aspekt der (ökologischen) Nachhaltigkeit derzeit jedoch nicht umfänglich durch die Referenzrahmen angesprochen. Artikel 20a Grundgesetz ist hierfür nur abstrakt heranzuziehen, wenngleich eine unmittelbare Wirkung des Artikels für die Verwaltung nicht ausgeschlossen wird (Calliess, 2022, hier: Rn. 206 f.). Ausflüsse des Energieeffizienzgesetzes (hier insbesondere die Paragraphen 6 sowie 11 bis 15) könnten diese Lücke schließen; dies bleibt hinsichtlich seiner Wirkung auf die amtliche Statistik jedoch abzuwarten. Gleiches gilt beispielsweise für Maßnahmenprogramme, wie jenes der Bundesregierung zur [Nachhaltigkeit](#). Die Qualitätsaspekte liefern dennoch erste Anhaltspunkte, wenngleich lediglich indirekt: Saidani und andere (2023, hier: Seite 283) fordert, dass bei „(annähernd) gleicher Ausprägung der anderen Qualitätsdimensionen [...] das erklärbarere Modell verwendet“ werden soll. Häufig sind besser erklärbare Modelle auch die ressourcenschonenderen. Da Ressourcen außerdem wirtschaftliche Aufwände produzieren, korreliert auch die Qualitätsdimension der Wirtschaftlichkeit (Saidani und andere, 2023, hier: Seite 290 f.) mit dieser ethischen Fragestellung.

Aus dem Referenzrahmen Qualität ließe sich somit ableiten, dass bei KI/ML-Anwendungen in der amtlichen Statistik zumindest geprüft oder begründet werden muss, dass die Ergebnisse nicht in ähnlicher Güte durch ressourcenschonendere Verfahren erzeugt werden können.

**Ausblick:** Nichtsdestotrotz ist in der Zukunft ein besonderes Augenmerk auf diese ethische Fragestellung zu legen, die anderenfalls – mangels bindender Vorgaben aus einem Referenzrahmen – dem Vergessen oder einer dem ethischen Risiko nicht angemessenen Behandlung anheimfallen könnte.

## 5

### Referenzrahmen und ethische Rahmenwerke im Abgleich

Während in Kapitel 3 den im ersten Teil des Artikels herausgearbeiteten ethischen Risiken im Detail begegnet wurde und Kapitel 4 offene Punkte identifizierte,

#### Übersicht 1

Vergleich der kanadischen Dimensionen mit den Referenzrahmen

Dimension	Entsprechungen in den Referenzrahmen
Respect for People	
Value to Canadians	Dies lässt sich unmittelbar übersetzen in die Forderung nach einer höheren Qualität der Prozesse oder der Produkte gemäß Qualitätshandbuch (Qualität als „fitness for use“).
Prevention of harm	Der Abwägungsprozess, was erhoben und veröffentlicht wird, erfolgt – unabhängig von KI/ML – im Parlament bei der Erstellung der Fachstatistikgesetze. Zusätzlich greift auf individueller Ebene (Geheimhaltung).
Fairness	Fairness bei {S2023-ALG FS}.
Accountability	{Verantwortlichkeit}.
Respect for Data	
Privacy	Datenschutzrecht, Trennungs- und Abschottungsgebot. Dass die amtliche Statistik Daten erheben darf, folgt aus dem Bundesstatistikgesetz, aus Fachstatistikgesetzen und Verfassungsgerichtsurteilen.
Security	{Informationssicherheit}.
Confidentiality	{Geheimhaltung}.
Sound Application	
Transparency	{Transparenz}; der Code wird aktuell in der Regel nicht zugänglich gemacht, jedoch sind die zugrunde liegenden Algorithmen frei verfügbare R-Packages oder Python-Module und somit vollkommen transparent.
Reproducibility of process and results	Reproduzierbarkeit bei {S2023-ALG FS}.
Sound Methods	
Quality of learning data	Genauigkeit bei {S2023-ALG FS}.
Valid inference	Genauigkeit bei {S2023-ALG FS}.
Rigorous modeling	Genauigkeit und Robustheit bei {S2023-ALG FS}.
Explainability	Reproduzierbarkeit und Erklärbarkeit bei {S2023-ALG FS}.

Quelle der kanadischen Dimensionen: [Statistics Canada, 2021](#)  
Erläuterung der { }-Entsprechungen siehe Kapitel 5 im Text.

## Übersicht 2

### Vergleich der britischen Dimensionen mit den Referenzrahmen

Dimension	Entsprechungen in den Referenzrahmen
The potential for bias	Fairness bei {S2023-ALG FS}.
Transparency and explainability	{Transparenz}, Erklärbarkeit bei {S2023-ALG FS}.
Accountability	{Verantwortlichkeit}.
Confidentiality	{Geheimhaltung}.

Quelle der britischen Dimensionen: [UK Statistics Authority, 2021](#)  
Erläuterung der {}-Entsprechungen siehe Kapitel 5 im Text.

vergleichen die [Übersichten 1 bis 3](#) die in Dumpert und andere (2025) beschriebenen, aus den kanadischen, britischen und UNECE-Vorarbeiten resultierenden Dimensionen mit den in diesem Kapitel definierten Referenzrahmen Recht und Qualität. Dabei gelten folgende Festlegungen:

- › {S2023-ALG FS}: Wird dies als Entsprechung in den Referenzrahmen genannt, so ist darunter zu verstehen, dass die Dimension beziehungsweise der Aspekt für die deutsche amtliche Statistik in Saidani und andere (2023) hergeleitet wird und vom zuständigen Gremium (dem ALG FS) mit Beschluss vom 10. Oktober 2023 bestätigt wurde.
- › Die Kürzel Gxx.x beziehen sich auf die Qualitätsgrundsätze gemäß Qualitätshandbuch (Statistische Ämter des Bundes und der Länder, 2021).
- › {Geheimhaltung}: Dies fasst § 16 Bundesstatistikgesetz und G05 des Qualitätshandbuchs zusammen.

- › {Verantwortlichkeit}: Dies fasst „Verwaltungshandeln wird stets der Behörde zugerechnet“ (siehe Abschnitt 3.5), G01.4 und G06.3 zusammen.
- › {Transparenz}: Dies fasst G06.4, G15.6 und {S2023-ALG FS} zusammen.
- › {Informationssicherheit}: Dies fasst Artikel 20 Absatz 4 der europäischen Statistikverordnung, die europäische NIS-2-Richtlinie, § 8 Absatz 1 BSI-Gesetz, den Umsetzungsplan Bund 2017 und G05.5 zusammen.

Im Ergebnis zeigt der Abgleich, dass tatsächlich allen in den ethischen Rahmenwerken anderer Staaten und zwischenstaatlicher Organisationen genannten Dimensionen, auch den nicht bereits in Kapitel 3 detailliert bearbeiteten, mit Ausnahme der ökologischen Nachhaltigkeit durch die Referenzrahmen Recht und Qualität angemessen begegnet wird.

## 6

### Fazit

Als Ergebnis wird festgehalten, dass die Umsetzung der Vorgaben der bestehenden und meist nicht KI/ML-spezifischen Referenzrahmen wie Recht und Qualität bereits ausreicht, um den auf Risikoszenarien basierenden ethischen Fragestellungen ausreichend zu begegnen. Eines besonderen ethischen Rahmenwerks bedarf es zurzeit für die deutsche amtliche Statistik nicht. Solange die Referenzrahmen Recht und Qualität in der aufgezeigten Strenge bestehen und umgesetzt werden,


## Übersicht 3

### Vergleich der UNECE-Dimensionen mit den Referenzrahmen

Dimension	Entsprechungen in den Referenzrahmen
Ethical Purpose, Public Good and Environmental Sustainability	§ 1 Bundesstatistikgesetz für den ethischen Zweck und das Gemeinwohl. Die ökologische Nachhaltigkeit findet aktuell keine Entsprechung in den Referenzrahmen.
Accountability	{Verantwortlichkeit}.
Transparency and Explainability	{Transparenz}, Erklärbarkeit bei {S2023-ALG FS}.
Fairness, Non-Discrimination, and Inclusiveness	Fairness bei {S2023-ALG FS}.
Privacy and Security	Dies folgt aus Datenschutzrecht und aus dem Trennungs- und Abschottungsgebot sowie {Informationssicherheit}.
Validity and Robustness	Genauigkeit bei {S2023-ALG FS}; Robustheit folgt aus {Informationssicherheit}.

Quelle: UNECE (Wirtschaftskommission für Europa der Vereinten Nationen), 2023 (Veröffentlichung in Vorbereitung)  
Erläuterung der {}-Entsprechungen siehe Kapitel 5 im Text.

besteht – außer möglicherweise bei der ökologischen Nachhaltigkeit – kein gesonderter Handlungsbedarf aus ethischer Sicht beim Einsatz von KI/ML in der Statistikproduktion. Ein Abgleich mit den zwischenstaatlichen Ansätzen und den Ansätzen anderer Staaten zeigt außerdem, dass auch diese nicht über rechtliche Vorgaben und Qualitätsaspekte hinausgehen (siehe die Übersichten 1 bis 3 in Kapitel 5). Der Abgleich weist zudem darauf hin, dass es derzeit keine weiteren, zwar bereits propagierten, aber für die deutsche amtliche Statistik nicht berücksichtigten ethischen Risiken gibt.

Recht und Qualitätsvorgaben unterliegen dem Wandel;<sup>10</sup> Gleiches gilt für KI/ML selbst. Möglicherweise sind die vorgenommenen Bewertungen bei Änderung eines Bestandteils eines Referenzrahmens oder bei einer unerwarteten Ausweitung der Fähigkeiten von KI/ML-Verfahren zu revidieren.<sup>11</sup> Diese Einsicht ist gleichzeitig Ermunterung und Auftrag, die Umstände stets im Blick zu haben und regelmäßig zu prüfen, ob die in Kapitel 2 herangezogenen Referenzrahmen weiterhin alle ethischen Fragestellungen für die amtliche Statistik beantworten. 

---

10 Beispielsweise empfiehlt die Kommission Zukunft Statistik mit Blick auf Gesundheitsdaten, „sich am Aufbau von Forschungsnetzwerken zu beteiligen und proaktiv digitale Schnittstellen zur Zusammenführung von Datensätzen zu realisieren. Hierfür sind die erforderlichen Rechtsgrundlagen zu schaffen.“ (Kommission Zukunft Statistik, 2024, hier: Seite 4)

11 Renda (2020, hier: Seite 655) schreibt – solche Situationen wahrnehmend oder vorausahnend – hierzu: „Besides the need to ensure a comprehensive approach to trust in AI systems is the belief that both law and ethics are needed and that in some cases the two may even clash (for example, when existing legislation does not reflect technological developments and ends up forcing market players to engage in unethical behavior); whereas in most cases they will be complementary (i.e., ethics can help in interpreting the law, or can recommend behavior that is not directly required or mandated by law).“

## LITERATURVERZEICHNIS

---

Blumöhr, Torsten/Teichmann, Corina/Noack, Anke. [\*Standardisierung der Prozesse: 14 Jahre AG SteP\*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 5/2017, Seite 58 ff.

Bothmann, Ludwig/Peters, Kristina. *Fairness als Qualitätskriterium im Maschinellen Lernen – Rekonstruktion des philosophischen Konzepts und Implikationen für die Nutzung außergesetzlicher Merkmale bei qualifizierten Mietspiegeln*. In: AStA Wirtschafts- und Sozialstatistisches Archiv. Jahrgang 18. Ausgabe 2/2024, Seite 185 ff. DOI: [10.1007/s11943-024-00346-0](#)

Bothmann, Ludwig/Peters, Kristina/Bischl, Bernd. *What is fairness? On the role of protected attributes and fictitious worlds*. Version 6, 2024. DOI: [10.48550/arXiv.2205.09622](#)

Bundesministerium des Innern. *Umsetzungsplan Bund 2017. Leitlinie für Informationssicherheit in der Bundesverwaltung*. Berlin 2017. [Zugriff am 2. Januar 2025]. Verfügbar unter: [www.bmi.bund.de](http://www.bmi.bund.de)

Calliess, Christian. *Artikel 20a*. In: Dürig, Günter/Herzog, Roman/Scholz, Rupert (Herausgeber). *Grundgesetz. Kommentar*. 2022.

Deutscher Bundestag. *Öffentliche Anhörung zu KI als Schlüsseltechnologie*. In: *Parlamentsnachrichten*. 2024. [Zugriff am 4. Dezember 2024]. Verfügbar unter: [www.bundestag.de](http://www.bundestag.de)

Dumpert, Florian/Reichel, Jannik/Oertel, Elisa/Leerhoff, Holger/Salwiczek, Christian. [\*Ethische Fragen beim Einsatz von KI/ML in der Produktion amtlicher Statistiken – Teil 1: Identifikation\*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 1/2025, Seite 15 ff.

Floridi, Luciano. *Soft ethics, the governance of the digital and the General Data Protection Regulation*. In: *Philosophical Transactions of the Royal Society A*. Volume 376. Ausgabe 2133. 2018. DOI: [10.1098/rsta.2018.0081](#)

García-Martín, Eva/Rodrigues, Crefeda Faviola/Riley, Graham/Grahn, Håkan. *Estimation of energy consumption in machine learning*. In: *Journal of Parallel and Distributed Computing*. Band 134. Ausgabe Dezember 2019, Seite 75 ff. DOI: [10.1016/j.jpdc.2019.07.007](#)

Henderson, Peter/Hu, Jieru/Romoff, Joshua/Brunskill, Emma/Jurafsky, Dan/Pineau, Joelle. *Towards the Systematic Reporting of the Energy and Carbon Footprints of Machine Learning*. In: *Journal of Machine Learning Research*. Jahrgang 21. Ausgabe 248/2020, Seite 1 ff. [Zugriff am 6. Dezember 2024]. Verfügbar unter: [jmlr.org](http://jmlr.org)

Kim, Michael P./Ghorbani, Amirata/Zou, James. *Multiaccuracy: Black-Box Post-Processing for Fairness in Classification*. In: *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. 2019. Seite 247 ff. DOI: [10.1145/3306618.3314287](#)

Kommission Zukunft Statistik (KomZS). *Bericht und Empfehlungen der Kommission Zukunft Statistik*. Version 1.0. 2024. [Zugriff am 6. Dezember 2024]. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

## LITERATURVERZEICHNIS

---

Kühling, Jürgen. *Bundesstatistikgesetz: BStatG*. Kommentar. München 2023.

Leischner, Sonja/Kolbe, Angela. [Zum Einfluss des Grundrechts auf informationelle Selbstbestimmung auf die Bundesstatistik](#). In: WISTA Wirtschaft und Statistik. Ausgabe 3/2024, Seite 17 ff.

OECD (Organisation für wirtschaftliche Zusammenarbeit und Entwicklung). *OECD-Bericht zu Künstlicher Intelligenz in Deutschland*. Paris 2024. DOI: [10.1787/8fd1bd9d-de](#)

Renda, Andrea. *Europe: Toward a Policy Framework for Trustworthy AI*. In: Dubber, Markus D./Pasquale, Frank/Das, Sunit (Herausgeber). *The Oxford Handbook of Ethics of AI*. Oxford 2020. Seite 650 ff. DOI: [10.1093/oxfordhb/9780190067397.013.41](#)

Ruschemeier, Hannah/Hondrich, Lukas J. *Automation bias in public administration – an interdisciplinary perspective from law and psychology*. In: *Government Information Quarterly*. Band 41, Ausgabe 3/2024, 101953. DOI: [10.1016/j.giq.2024.101953](#)

Saidani, Younes/Dumpert, Florian/Borgs, Christian/Brand, Alexander/Nickl, Andreas/Rittmann, Alexandra/Rohde, Johannes/Salwiczek, Christian/Storfinger, Nina/Straub, Selina. *Qualitätsdimensionen maschinellen Lernens in der amtlichen Statistik*. In: *ASTA Wirtschafts- und Sozialstatistisches Archiv*. Jahrgang 17. Ausgabe 3–4/2023, Seite 253 ff. DOI: [10.1007/s11943-023-00329-7](#)

Samsi, Siddharth/Zhao, Dan/ McDonald, Joseph/Li, Baolin/Michaleas, Adam/Jones, Michael/Bergeron, William/Kepner, Jeremy/Tiwari, Devesh/Gadepally, Vijay. *From Words to Watts: Benchmarking the Energy Costs of Large Language Model Inference*. 2023. DOI: [10.48550/arXiv.2310.03003](#)

Schenk, Patrick Oliver/Kern, Christoph. *Connecting Algorithmic Fairness to Quality Dimensions in Machine Learning in Official Statistics and Survey Production*. In: *ASTA Wirtschafts- und Sozialstatistisches Archiv*. Jahrgang 18. Ausgabe 2/2024, Seite 131 ff. DOI: [10.1007/s11943-024-00344-2](#)

Statistics Canada. *Framework for Responsible Machine Learning Processes at Statistics Canada*. 2021. [Zugriff am 8. Januar 2025]. Verfügbar unter: [www150.statcan.gc.ca](#)

Statistische Ämter des Bundes und der Länder. *Qualitätshandbuch der Statistischen Ämter des Bundes und der Länder (Version 1.21)*. 2021. [Zugriff am 6. Dezember 2024]. Verfügbar unter: [www.destatis.de](#)

UK Statistics Authority. *Ethical considerations in the use of Machine Learning for research and statistics*. 2021. [Zugriff am 6. Dezember 2024]. Verfügbar unter: [uksa.statisticsauthority.gov.uk](#)

## RECHTSGRUNDLAGEN

---

Gesetz über das Bundesamt für Sicherheit in der Informationstechnik (BSI-Gesetz – BSIg) vom 14. August 2009 (BGBl. I Seite 2821), das zuletzt durch Artikel 12 des Gesetzes vom 23. Juni 2021 (BGBl. I Seite 1982) geändert worden ist.

Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) in der Fassung der Bekanntmachung vom 20. Oktober 2016 (BGBl. I Seite 2394), das zuletzt durch Artikel 14 des Gesetzes vom 8. Mai 2024 (BGBl. I Nr. 152) geändert worden ist.

Gesetz zur Steigerung der Energieeffizienz in Deutschland (Energieeffizienzgesetz – EnEfG) vom 13. November 2023 (BGBl. I Nr. 309).

Grundgesetz für die Bundesrepublik Deutschland in der im Bundesgesetzblatt Teil III, Gliederungsnummer 100-1, veröffentlichten bereinigten Fassung, das zuletzt durch Artikel 1 des Gesetzes vom 19. Dezember 2022 (BGBl. I Seite 2478) geändert worden ist.

Richtlinie (EU) 2022/2555 des Europäischen Parlaments und des Rates vom 14. Dezember 2022 über Maßnahmen für ein hohes gemeinsames Cybersicherheitsniveau in der Union, zur Änderung der Verordnung (EU) Nr. 910/2014 und der Richtlinie (EU) 2018/1972 sowie zur Aufhebung der Richtlinie (EU) 2016/1148 (NIS-2-Richtlinie) (Amtsblatt der EU Nr. L 333, Seite 80).

Verordnung (EG) Nr. 223/2009 des Europäischen Parlaments und des Rates vom 11. März 2009 über europäische Statistiken und zur Aufhebung der Verordnung (EG, Euratom) Nr. 1101/2008 des Europäischen Parlaments und des Rates über die Übermittlung von unter die Geheimhaltungspflicht fallenden Informationen an das Statistische Amt der Europäischen Gemeinschaften, der Verordnung (EG) Nr. 322/97 des Rates über die Gemeinschaftsstatistiken und des Beschlusses 89/382/EWG, Euratom des Rates zur Einsetzung eines Ausschusses für das Statistische Programm der Europäischen Gemeinschaften (Amtsblatt der EU Nr. L 87, Seite 164).

Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung) (Amtsblatt der EU Nr. L 119, Seite 1).

Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Amtsblatt der EU Nr. L 144, Seite 1).

# DIE NEUE INTRAHANDELSSTATISTIK AB JANUAR 2025

Sebastian Metzl, Igor Franjić, Thomas Kolvenbach

➤ **Schlüsselwörter:** Außenhandel – Intrahandel – Schätzung – Mikrodaten-  
austausch – fehlende Daten

## ZUSAMMENFASSUNG

Im Rahmen der Außenhandelsstatistik wird der Warenverkehr zwischen Deutschland und den anderen Mitgliedstaaten der Europäischen Union durch eine Befragung von Unternehmen erfasst. Neue innovative Verfahren und technischer Fortschritt im Statistischen Bundesamt und auf europäischer Ebene haben es ermöglicht, die Schätzverfahren für fehlende Daten in dieser Erhebung auf eine verbesserte und strukturierte Grundlage zu stellen. Daher steigt ab dem Monatsanfang Januar 2025 die Genauigkeit der Statistik im Hinblick auf die Detaillierung der gehandelten Waren. Zudem legen die neuen Schätzverfahren das Fundament für eine künftige Entlastung der kleinen und mittleren Unternehmen von Bürokratiekosten, die durch Meldepflichten zur Außenhandelsstatistik entstehen.

➤ **Keywords:** foreign trade – intra-EU trade – estimation – microdata exchange –  
missing data

## ABSTRACT

*In foreign trade statistics, trade in goods between Germany and other EU Member States is recorded by surveying enterprises. New innovative methods and technical advancements within the Federal Statistical Office and at the European level have allowed estimation techniques for missing data in this survey to be placed on a stronger and more structured foundation. From reporting month January 2025 onwards, intra-EU trade statistics will therefore be more accurate in terms of the detailed results for traded products. Furthermore, the new estimation techniques lay the basis for a future reduction of the administrative burden on small and medium-sized enterprises that is associated with data reporting requirements for the purpose of foreign trade statistics.*

### Dr. Sebastian Metzl

ist Betriebswirt und als Referent im Referat „Grundsatzfragen, Qualitätssicherung, Verbreitung“ in der Gruppe „Außenhandel“ des Statistischen Bundesamtes tätig. Seine Arbeitsschwerpunkte liegen bei den Schätzverfahren für fehlende Daten und dem Mikrodaten austausch im Intrahandel.

### Igor Franjić

ist Betriebswirt und arbeitet als Wissenschaftlicher Mitarbeiter im Referat „Grundsatzfragen, Qualitätssicherung, Verbreitung“ in der Gruppe „Außenhandel“ des Statistischen Bundesamtes. In einem EU-Projekt zu globalen Wertschöpfungsketten entwickelt er neue Indikatoren für die Außenhandelsstatistik nach Unternehmensmerkmalen. Zuvor betreute er den Mikrodaten austausch im Intrahandel.

### Thomas Kolvenbach

hat Economics (M. Sc.) studiert und ist im Statistischen Bundesamt als Sachbearbeiter im Referat „Grundsatzfragen, Qualitätssicherung, Verbreitung“ in der Gruppe „Außenhandel“ beschäftigt. Der Schwerpunkt seiner Tätigkeit liegt in der Aufbereitung zur Nutzung von Steuerdaten und der Analyse zur künftigen Nutzung von Mikrodaten für die Außenhandelsstatistik.



## 1

### Einleitung

Der Außenhandel hat maßgeblichen Einfluss auf die wirtschaftliche Entwicklung in Deutschland. Gegenstand der Außenhandelsstatistik ist der Warenverkehr Deutschlands mit anderen Staaten. Der Handel mit den Partnerländern wird dabei wert- und mengenmäßig unter anderem nach detaillierten Warenkategorien (achtstellige Warennummern des Warenverzeichnisses für die Außenhandelsstatistik) aufgeteilt. Die Nutzerinnen und Nutzer können beispielsweise nachvollziehen, wie viele neue E-Autos aus China im August 2023 nach Deutschland importiert wurden oder zu welchem Wert Moselwein im März 2024 aus Deutschland nach Slowenien exportiert worden ist. Um solche Daten zu erheben nutzt das Statistische Bundesamt verschiedene Quellen.

Für den Warenverkehr mit Ländern außerhalb der Europäischen Union (EU), den sogenannten Drittländern, stehen die relevanten Angaben aus Zollanmeldungen zur Verfügung. Bis auf wenige Ausnahmen muss jede Warenbewegung mit Drittländern aus fiskalischen Gründen und zur Durchsetzung außenwirtschaftsrechtlicher Vorgaben bei den Zollbehörden angemeldet werden. Aus statistischer Sicht gleicht diese Situation einer Vollerhebung, bei der praktisch alle Warenbewegungen erfasst werden. Bis Dezember 1992 wurde so der gesamte Warenverkehr mit dem Ausland erhoben.

Mit Vollendung des europäischen Binnenmarkts im Jahr 1993 und der damit einhergehenden Abschaffung der Warenkontrollen zwischen den EU-Mitgliedstaaten entfielen die Zollanmeldungen für einen erheblichen Teil des deutschen Außenhandels. Unabhängig davon bestand der Bedarf an statistischen Daten zum inner-europäischen Warenverkehr (Intrahandel)<sup>1</sup> jedoch fort – beispielsweise für die Berechnung der nationalen Zahlungsbilanz und des deutschen Bruttoinlandsprodukts.

---

1 Ein Teil des Intrahandels wird weiterhin über Zollanmeldungen abgedeckt. So sind Zollanmeldungen für Warenverkehre mit steuerrechtlichen Sondergebieten (zum Beispiel den Kanarischen Inseln) abzugeben. Auch werden Waren, die zur Be- oder Verarbeitung aus einem Drittland in das Zollgebiet der EU eingeführt und innerhalb dieser weiterversendet werden (zollrechtliche Veredelungsverkehre), von den Zollbehörden überwacht, bis sie in den freien Verkehr übergeführt werden oder die EU wieder verlassen. Die bei der Bewegung der Waren innerhalb der EU anfallenden Zollanmeldungen gehören zum Intrahandel.

Deshalb wurde die sogenannte Intrastat-Erhebung als zweiter Weg etabliert.

Ab dem Berichtsmonat Januar 2025 werden neue Schätzverfahren in der Intrahandelsstatistik eingeführt. Dadurch wird die Detailgenauigkeit der Ergebnisse in Hinsicht auf die gehandelten Waren verbessert und die Erhebung auf eine zukünftige Entlastung der Unternehmen von Meldepflichten an die Außenhandelsstatistik vorbereitet. Kapitel 2 beschreibt zunächst, wie die Erhebung des innergemeinschaftlichen Warenverkehrs derzeit abläuft. Die Gründe für die Einführung neuer Schätzverfahren und den Übergang zu einer neuen Schätzmethode erläutert Kapitel 3. Einzelheiten dazu, wie der Gesamtwert berechnet wird und die Zuschätzfaktoren festgelegt werden, enthält Kapitel 4. Die Kapitel 5 und 6 legen dar, wie fehlende Daten meldepflichtiger Unternehmen sowie der Schätzwert für Befreiungen aufgeteilt werden. Ein Fazit sowie ein Ausblick beschließen den Beitrag.

## 2

### Die Erhebung des innergemeinschaftlichen Warenverkehrs

Die Unternehmen melden für die Intrastat-Erhebung monatlich für rein statistische Zwecke bestimmte Informationen über die innerhalb der EU exportierten beziehungsweise importierten Waren an die nationalen statistischen Ämter und nicht mehr an die Zollbehörden. Zur Entlastung kleiner und mittlerer Unternehmen wurden in diesem Zusammenhang sogenannte Anmeldeschwellen eingeführt. Demnach müssen Unternehmen erst dann Intrastat-Anmeldungen für einen bestimmten Berichtsmonat abgeben, wenn der Gesamtwert ihrer Exporte in andere EU-Mitgliedstaaten (Versendungen) beziehungsweise ihrer Importe aus anderen EU-Mitgliedstaaten (Eingänge) im laufenden Jahr oder im Vorjahr einen bestimmten Wert überschreitet. Die Schwellen sind so gewählt, dass bislang nur etwa 10 % der exportierenden beziehungsweise importierenden Unternehmen in Deutschland zur Abgabe von Intrastatmeldungen verpflichtet sind. Dabei ist die europarechtliche Vorgabe zu berücksichtigen, dass wertmäßig mindestens 95 % der Versendungen eines Mitgliedstaats durch Unternehmensmeldungen erhoben werden müssen (Abdeckungs-



grad). Bei den Eingängen aus anderen EU-Mitgliedstaaten ist der entsprechende Abdeckungsgrad bisher national im Außenhandelsstatistikgesetz auf 93 % festgelegt. Die Anmeldeschwellen liegen bisher entsprechend bei 500 000 Euro bei Versendungen in andere EU-Mitgliedstaaten und 800 000 Euro bei den Eingängen von dort.

Dem Vorteil der Entlastung von kleinen und mittleren Unternehmen durch die Anmeldeschwellen steht der Nachteil gegenüber, dass durch die Intrastatmeldungen nicht der gesamte innereuropäische Warenverkehr erhoben wird. Unternehmen, deren Versendungen beziehungsweise Eingänge unter der jeweiligen Anmeldeschwelle liegen (below-threshold trade), müssen keine Intrastatmeldungen über ihre Warenverkehre abgeben. Darüber hinaus gibt es aber auch Unternehmen oberhalb der Anmeldeschwelle, die ihrer bestehenden Meldepflicht für die Exporte beziehungsweise Importe eines Monats nicht fristgerecht nachkommen. In diesem Zusammenhang spricht man von sogenannten Antwortausfällen (non-response). Für die so entstehenden Differenzen zum gesamten EU-Warenverkehr müssen Schätzungen vorgenommen werden. Die Schätzungen teilen sich entsprechend auf in sogenannte Zuschätzungen für Befreiungen für die Warenverkehre von Unternehmen unterhalb der Anmeldeschwellen einerseits und Zuschätzungen für Antwortausfälle von Unternehmen oberhalb der Anmeldeschwellen, die ihrer Meldepflicht nicht oder nicht vollständig nachgekommen sind, andererseits. Die Zuschätzungen lassen sich grundsätzlich mithilfe der Daten aus den monatlichen Umsatzsteuer-Voranmeldungen berechnen, da im Wesentlichen alle Unternehmen zu dieser Meldung an die Finanzverwaltung verpflichtet sind. Die relevanten Angaben aus den Umsatzsteuer-Voranmeldungen liegen im Statistischen Bundesamt für statistische Zwecke vor. Mit diesen Angaben lassen sich die Gesamthöhe des Intrahandels und der Anteil der befreiten Unternehmen an allen innergemeinschaftlichen Lieferungen und Erwerben im jeweiligen Monat berechnen. Jedoch können die Umsatzsteuer-Voranmeldungen die Intrastatmeldung nicht ersetzen, da neben konzeptionellen Unterschieden in den Umsatzsteuer-Voranmeldungen nur ein monatlicher Gesamtwert je Unternehmen ohne Differenzierung nach Waren und Partnerländern vorliegt. Deshalb werden Schätzverfahren benötigt, die die statistisch relevanten Merkmale ergänzen.

### ➤ Exkurs: Intrastat und Umsatzsteuer

Während die Außenhandelsstatistik den physischen, grenzüberschreitenden Verkehr mit Waren erfasst, dient die Umsatzsteuer-Voranmeldung dazu, steuerliche Tatbestände für fiskalische Zwecke zu erfassen. Diese unterschiedliche Zielsetzung führt zu bedeutenden konzeptionellen Unterschieden zwischen Intrastat und Umsatzsteuer, die über die Tatsache hinausgehen, dass die Daten der Umsatzsteuer-Voranmeldungen wesentlich weniger Detailinformationen über die innereuropäischen Warenverkehre enthalten als Intrastatmeldungen. So werden für die Umsatzsteuer bestimmte Warenverkehre, die für die Außenhandelsstatistik relevant sind, undifferenziert gemeinsam mit Dienstleistungen und anderen Tatbeständen erfasst, sodass die Steuerdaten in diesem Fall für die Außenhandelsstatistik nicht nutzbar sind. Auch ist der steuerrelevante Zeitraum nicht identisch mit dem Berichtszeitraum der Außenhandelsstatistik. Zudem ist die umsatzsteuerrechtliche Bemessungsgrundlage nicht gleich dem statistischen Wert einer Ware. Aus diesen und weiteren Gründen können die Umsatzsteuerdaten die gesonderte Erhebung des Intrahandels nicht ersetzen und auch für sich genommen die Höhe des Außenhandels nicht hinreichend genau quantifizieren.

## 3

---

### Einführung neuer Schätzverfahren

---

Die bisherigen Schätzverfahren im Intrahandel erzeugen zuverlässig Schätzungen für Antwortausfälle und den Warenverkehr von Unternehmen unterhalb der Anmeldeschwelle. Diese weisen jedoch nur eine vergleichsweise geringe Detailtiefe hinsichtlich der Art der Ware auf, und zwar für die zweistelligen Kapitel (WA-2-Steller) des Warenverzeichnisses für die Außenhandelsstatistik. Weitere Merkmale, die zugeschätzt werden, sind die Partnerländer (Bestimmungsländer, Ursprungs- und Versandungsländer) sowie die deutschen Bundesländer.

Bei den derzeit geringen Zuschätzungen für Befreiungen von 1,6 % export- und 4,0 % importseitig ist eine solche Aufteilung der Zuschätzungen ausreichend. Für die Antwortausfälle werden bislang zum Zeitpunkt der Erstveröffentlichung eines Monats durchschnittlich 6,5 % auf Versendungs- und 8,8 % auf Eingangsseite geschätzt. Bis zur Veröffentlichung der endgültigen

Zahlen für das Berichtsjahr 2023 konnten nachträglich eingehende Meldungen diese Schätzungen auf 1,1 % export- und 2,2 % importseitig reduzieren, sodass das bisherige Schätzverfahren zu keinen größeren Qualitätsverlusten führte. Eine Aufteilung nach achtstelligen Warennummern war bisher – ohne weitere Zusatzinformationen – qualitativ nicht sinnvoll.

Die Nutzung einer neuen Datenquelle in Form des europäischen Mikrodatenaustauschs (Franjić und andere, 2024) ermöglicht es dabei, künftig importseitig für die Befreiungen auch auf Ebene der achtstelligen Warennummern zuzuschätzen. Zudem konnte eine neue Schätzmethode für Antwortausfälle meldepflichtiger Einheiten entwickelt werden, die auf tief gegliederten historischen Zeitreihen und Steuerdaten basiert. Im Zuge dieser Neukonzeption wurde auch das Verfahren zur Berechnung des Gesamtwertes der Warenverkehre mit anderen EU-Mitgliedstaaten überarbeitet.

Mit der Novelle von Außenhandelsstatistikgesetz und Außenhandelsstatistik-Durchführungsverordnung zum 1. Januar 2025 werden die Anmeldeschwellen in Deutschland erhöht. Sie steigen bei Versendungen in andere EU-Mitgliedstaaten von 500 000 Euro auf 1 Million Euro und bei Eingängen aus anderen Mitgliedstaaten von 800 000 Euro auf 3 Millionen Euro. Damit entlastet das Statistische Bundesamt einen wesentlichen Teil der bisher meldepflichtigen Unternehmen von Statistikpflichten. Infolgedessen erhöht sich der Anteil des Außenhandels, der geschätzt werden muss, deutlich. Dies ist insbesondere bei den Importen der Fall. Ohne eine Zuschätzung der Importe von Unternehmen unterhalb der Anmeldeschwelle auf Ebene der achtstelligen Warennummern wäre der Informationsverlust auf der Detailebene durch die Meldeschwellenanhebung erheblich.

Durch die neuen Schätzmethoden bleibt die Qualität der Außenhandelsstatistik auf dem aktuellen Niveau oder verbessert sich leicht, obwohl der Anteil der Zuschätzungen steigt. Die Umsetzung der neuen Schätzverfahren erfolgt im Kalendermonat März 2025 mit der erstmaligen Veröffentlichung der Außenhandelsstatistik für den Berichtsmonat Januar 2025. Die Details der neuen Methoden werden in den nächsten Kapiteln erläutert.

## 4

### Berechnung des Gesamtwertes und Festlegung der Zuschätzfaktoren

Eine zentrale Herausforderung bei der Erhebung des Intrahandels besteht darin, den Gesamtwert der importierten beziehungsweise exportierten Waren eines Berichtsmonats festzulegen. Hierfür wird auf Basis der Umsatzsteuer-Voranmeldungen anhand komplexer Berechnungen ein fiktiver Wert ermittelt. In einem ersten Schritt werden die angegebenen Werte für die innergemeinschaftlichen Lieferungen beziehungsweise Erwerbe aus den Umsatzsteuer-Voranmeldungen mit den von der Außenhandelsstatistik erhobenen statistischen Werten vergleichbar gemacht. Außerdem werden Warenbewegungen, die in der Umsatzsteuer-Voranmeldung nicht als innergemeinschaftliche Lieferungen beziehungsweise Erwerbe anzumelden sind (beispielsweise der Import von Erdgas oder Veredelungsgeschäfte), mit eingerechnet. Danach werden die Steuerwerte ausländischer Unternehmen, die in Deutschland Waren veredeln (be- oder verarbeiten), abgezogen – diese Warenverkehre wären ansonsten sowohl in den Steuerwerten als auch in den Zuschlägen für Veredelungsverkehre enthalten. Schließlich wird noch ein Korrekturfaktor für die Steuerwerte berechnet, der den typischen Umfang der Fehler und Korrekturen der Daten eines Berichtsmonats aus den Umsatzsteuer-Voranmeldungen widerspiegelt.

Der Abgleich dieses fiktiven Gesamtwertes mit dem Wert der tatsächlich erhobenen Intrastatmeldungen ergibt die Zuschätzungen. Für Befreiungen wird anhand der Steuerwerte des Vorjahres von Unternehmen unterhalb der Anmeldeschwelle ein prozentualer Faktor ermittelt. Zu diesem hinzugerechnet wird ein Faktor für befreite Unternehmen, die Veredelung durchführen, die nicht in den Steuerwerten enthalten sind. Von diesem Maximalwert der befreiten Warenverkehre wird der Wert der Intrastatmeldungen der sogenannten freiwilligen Melder, also der Unternehmen, die unterhalb der Meldeschwelle liegen, aber trotzdem Meldungen zum Intrahandel abgeben, abgezogen.

Unter Berücksichtigung der befreiten Unternehmen und der freiwilligen Melder wird anhand des fiktiven Gesamtwertes der Faktor für die Zuschätzungen für Antwortausfälle ermittelt. Aus der Differenz zwischen fik-

tivem Gesamtwert und erhobenen Werten kann damit grundsätzlich der Faktor für die Zuschätzungen für Antwortausfälle errechnet werden. Der prozentuale Faktor für die Zuschätzung wird für die Erstveröffentlichung des jeweiligen Monats festgelegt. Mit jeder Revision der Monatsergebnisse verringern sich die Zuschätzungen für Antwortausfälle, da das Statistische Bundesamt laufend säumige Unternehmen anmahnt, ihre fehlenden Meldungen nachzureichen.

Aufgrund rechtlicher Vorgaben und den damit verbundenen Lieferverpflichtungen gegenüber dem Statistischen Amt der Europäischen Union (Eurostat) liegt zwischen der Pressemitteilung mit ersten Außenhandelsergebnissen und den tatsächlichen Veröffentlichungszeitpunkten aller Detailergebnisse rund eine Woche. In dieser Woche gehen Nachmeldungen für den aktuellen Berichtsmonat ein. Zurzeit wird die Differenz des Gesamtwertes zwischen Pressemitteilung und Veröffentlichung der Detailergebnisse durch einen Faktor von 1,8 % bei den Exporten und 2,1 % bei den Importen als Zuschlag für diese fehlenden Meldungen in der Pressemitteilung berücksichtigt.

Die Zuschätzfaktoren werden regelmäßig auf erforderliche Anpassungen überprüft. Dabei können sich die Außenhandelsergebnisse durch die geänderten Zuschätzfaktoren für das laufende Jahr oder das Vorjahr rückwirkend nochmals ändern.

Die Gesamtwerte für Befreiungen und Antwortausfälle ergeben sich durch die Zuschätzfaktoren. Diese sind jedoch nach den relevantesten Merkmalen für die Außenhandelsstatistik aufzuteilen. Die unterschiedlichen Verfahren dafür beschreibt das folgende Kapitel 5.

## 5

### Aufteilung der Antwortausfälle

---

Bei den Meldeausfällen von Unternehmen werden zwei Kategorien unterschieden:

- › teilweise Meldeausfälle,
- › komplette Meldeausfälle.

Bei **teilweisen Meldeausfällen** hat ein Unternehmen zwar Intrastatmeldungen für einen bestimmten Berichtsmonat abgegeben. Deren Wert deckt aber nicht die übli-

che Höhe der Meldungen verglichen mit den Vormonaten beziehungsweise den der Finanzverwaltung gemeldeten Werten der innergemeinschaftlichen Lieferungen beziehungsweise Erwerbe des Unternehmens ab. Bei **kompletten Meldeausfällen** gibt ein Unternehmen für den entsprechenden Monat keine Meldung ab, obwohl alle Vormonatswerte beziehungsweise die Steuerwerte eine Meldung erwarten lassen.

Teilweise Meldeausfälle lassen sich auf Unternehmensebene schwer identifizieren, weil der vergleichsweise geringe Wert der Intrastat-Anmeldungen für den aktuellen Berichtsmonat auch normale monatliche Schwankungen in den Auslandsumsätzen eines Unternehmens widerspiegeln kann. Zudem liegen die Vergleichsdaten aus den Umsatzsteuer-Voranmeldungen zum Zeitpunkt der Erstveröffentlichung noch nicht vor. Entsprechend wäre der gesamte Wert der Schätzfehler einer Schätzung auf Unternehmensebene für teilweise Meldeausfälle sehr hoch. Dagegen lassen sich komplette Meldeausfälle mit relativ hoher Sicherheit feststellen, wodurch die Schätzfehler niedriger sind als bei den teilweisen Ausfällen. Aus diesem Grund werden die teilweisen Antwortausfälle auch weiterhin nach dem bisherigen System aufgeteilt, während die kompletten Meldeausfälle nach einem neuen, unternehmensbezogenen Bottom-up-Modell aufgeteilt werden.

### 5.1 Bottom-up-Schätzung

---

Ab Berichtsmonat Januar 2025 werden die Meldeausfälle von Unternehmen, die keine Intrastatmeldungen für einen Berichtsmonat abgegeben haben (komplette Meldeausfälle), grundsätzlich durch ein neu entwickeltes Bottom-up-Verfahren geschätzt. Hierfür sind verschiedene Verfahrensschritte erforderlich, die im Folgenden erläutert werden.

Ob ein Unternehmen, dessen Meldungen zum Zeitpunkt der Erstveröffentlichung eines Berichtsmonats fehlen, zu einem späteren Zeitpunkt nachmeldet, wird mittels einer logistischen Regression auf Basis eines vergleichbaren Monats geschätzt. Das logistische Modell berechnet anhand verschiedener Einflussgrößen eine Wahrscheinlichkeit, dass ein aktuell fehlendes Unternehmen später nachmeldet. Als Einflussgrößen haben sich dabei folgende Kriterien herauskristallisiert:

- › Je länger ein aktuell fehlendes Unternehmen schon zur Intrahandelsstatistik in einer Verkehrsrichtung meldet, desto wahrscheinlicher ist eine Nachmeldung.
- › Je häufiger ein fehlendes Unternehmen in der Vergangenheit einen Wert nachgemeldet hat, desto wahrscheinlicher wird eine Nachlieferung im aktuellen Monat.
- › Je mehr korrekte monatliche Nichtmeldungen die Unternehmenszeitreihe aufweist, desto unwahrscheinlicher ist eine Nachmeldung für den aktuellen Monat.

Ein Unternehmen wird für die weiteren Schritte nur berücksichtigt, wenn die Wahrscheinlichkeit für eine Nachmeldung bei mehr als 50 % liegt. Zusätzlich werden Unternehmen herausgenommen, von denen bekannt ist, dass sie aktuell keinen Intrahandel betreiben. Das ist insbesondere dann der Fall, wenn Unternehmen dem Statistischen Bundesamt mitteilen, dass sie keine Warenverkehre mit anderen EU-Mitgliedstaaten für den aktuellen Monat oder auf längere Sicht mehr durchführen.

Im nächsten Schritt werden die zu schätzenden Warennummern der im ersten Schritt ausgewählten Unternehmen festgelegt. Dies geschieht wieder mithilfe einer Vergleichsperiode aus der Vergangenheit und einer logistischen Regression. Dafür gelten folgende Kriterien:

- › Je länger ein Unternehmen eine aktuell fehlende Warennummer meldet, desto eher wird diese Ware auch im aktuellen Monat nachgeliefert.
- › Je häufiger eine fehlende Warennummer in der Vergangenheit vom Unternehmen nachgemeldet wurde, desto wahrscheinlicher wird eine Nachlieferung im aktuellen Monat.
- › Je mehr Lücken in der Zeitreihe einer bestimmten Warennummer eines Unternehmens existieren, desto unwahrscheinlicher ist eine Nachmeldung im aktuellen Monat.

Auch hier gilt, dass eine Warennummer eines Unternehmens im weiteren Verlauf nur berücksichtigt wird, wenn die Nachmeldewahrscheinlichkeit über 50 % liegt.

Um den aktuellen Wert einer zu schätzenden Ware eines fehlenden Unternehmens zu bestimmen, wird zunächst der zuletzt vom Unternehmen gemeldete statistische Wert der Exporte beziehungsweise Importe dieser Ware

herangezogen. Dieser Wert wird dann anhand der Marktentwicklung fortgeschrieben. Um die Marktentwicklung zu bestimmen, werden die aktuellen Meldungen anderer Unternehmen genutzt, die den Export beziehungsweise Import dieser oder gegebenenfalls vergleichbarer Waren an das Statistische Bundesamt melden.

Anschließend werden die geschätzten statistischen Werte je Warennummer anhand der historischen Verteilung der Anmeldungen des jeweiligen Unternehmens auf Partnerländer, Bundesländer und Arten des Geschäfts aufgeteilt. Das Merkmal Art des Geschäfts unterscheidet dabei zwischen bestimmten Klauseln im Geschäftsvertrag wie Kauf/Verkauf, Lohnveredelung oder Finanzierungsleasing. Dies ist insbesondere für die Deutsche Bundesbank von Bedeutung.

Meldet ein Unternehmen nach oder stellt sich aufgrund anderer dem Statistischen Bundesamt vorliegender Informationen heraus, dass eine Schätzung falsch war, wird der entsprechende Schätzwert für das Unternehmen gelöscht. Mit jeder Revision der Monatsergebnisse verringern sich damit die Bottom-up-Schätzungen und werden durch echte Meldungen der Unternehmen ersetzt.

Das hier beschriebene Verfahren ist sehr rechen- und zeitintensiv. Daher werden die Bottom-up-Schätzungen nur zur Erstveröffentlichung der Detailergebnisse berechnet. Zum Zeitpunkt der monatlichen Pressemitteilung erfolgt die Verteilung der Zuschätzungen für Antwortausfälle weiterhin nach dem bisherigen Top-down-Ansatz.

Das neue Schätzverfahren ermöglicht, schneller als bisher monatsgerechte Außenhandelsergebnisse zu veröffentlichen und dadurch auch die Nachmeldungen von Unternehmen zu einem früheren Zeitpunkt zu berücksichtigen. Damit kann ab Berichtsmonat Januar 2025 die erste Revision der Ergebnisse eines Berichtsmonats bereits mit der Erstveröffentlichung des darauffolgenden Berichtsmonats erfolgen. Bisher fand die erste Revision eines Berichtsmonats zusammen mit der Veröffentlichung des zweiten Folgemonats statt. Insgesamt wurden für jeden Berichtsmonat sechs aufeinanderfolgende monatliche Revisionen durchgeführt. Ab 2025 wird es nun sieben Revisionen geben. Die abschließende Revision der Jahresergebnisse erfolgt unverändert im Oktober des Folgejahres. Ab diesem Zeitpunkt werden die Ergebnisse als endgültig betrachtet.



## 5.2 Top-down-Schätzung

---

Bei den teilweisen Meldeausfällen ergibt eine Zuschätzung auf Ebene der achtstelligen Warennummern keine qualitativ hochwertige Aussage. Hier ist der Schätzfehler zu groß, der bei einer Abkehr vom bisherigen Modell entstünde. Für den Teil der Antwortausfälle, den die oben vorgestellte Bottom-up-Schätzung nicht abdeckt, erfolgt die Verteilung deshalb weiterhin nur auf Ebene der Kapitel des Warenverzeichnisses (WA-2-Steller), der Partnerländer und der Bundesländer. Neu ab Januar 2025 ist, dass auch die Top-down-Schätzung für Zwecke der Zahlungsbilanz und des Bruttoinlandsprodukts nach Arten des Geschäfts aufgeteilt wird. Wie bereits oben angeführt, werden zum Zeitpunkt der Bekanntgabe der ersten monatlichen Außenhandelsergebnisse die Zuschätzungen für Antwortausfälle ausschließlich nach dem Top-down-Ansatz verteilt. Erst bei der Veröffentlichung der detaillierten Ergebnisse für den aktuellen Monat rund eine Woche nach der Pressemitteilung erfolgen Bottom-up-Schätzungen.

Für die Verteilung der Zuschätzungen wird angenommen, dass bei Unternehmen mit Eingängen beziehungsweise Versendungen über 10 Millionen Euro grundsätzlich keine Meldeausfälle vorliegen. Aufgrund einer stringenten Eingangskontrolle für größere Unternehmen kann davon ausgegangen werden, dass sich Antwortausfälle vorwiegend auf kleine und mittlere Unternehmen verteilen. Daher wird berechnet, wie sich die Meldewerte von Einheiten, die einen Steuerwert unter 10 Millionen Euro aufweisen, auf Warenkapitel, Partnerländer, Bundesländer und Arten des Geschäfts verteilen. Diese Struktur der Unternehmen wird dann für die Verteilung der Zuschätzungen verwendet. Da dieses Verfahren einen Gesamtwert mithilfe von Quoten in einzelne Unterpositionen aufteilt, wird es als Top-down-Schätzung bezeichnet.

## 6

---

### Aufteilung des Schätzwertes für Befreiungen

---

Wie die Schätzung für die Antwortausfälle von meldepflichtigen Einheiten sind auch die Zuschätzungen für die Unternehmen unterhalb der Meldeschwelle auf weitere Merkmale der Außenhandelsstatistik zu verteilen. Dabei werden ab Berichtsmonat Januar 2025 unterschiedliche Verfahren für Eingänge und Versendungen angewendet, weil

- › die Lieferverpflichtungen gegenüber Eurostat für Eingänge und Versendungen,
- › die Datenlage für Eingänge und Versendungen und
- › die Meldeschwellen für die jeweiligen Verkehrsrichtungen

unterschiedlich sind.

Das Statistische Bundesamt muss wertmäßig mindestens 95 % der Versendungen durch Unternehmensmeldungen direkt erheben. Da europarechtlich kein Abdeckungsgrad bei den Eingängen mehr vorgeschrieben ist, ergibt sich hier die Möglichkeit, auf nationaler Ebene die Anmeldeschwelle deutlich stärker anzuheben als bei der Versendung. Entsprechend erhöht sich der Umfang der Warenverkehre von Unternehmen unter der Meldeschwelle, die beim Eingang zu schätzen sind, und ist deutlich größer als bei der Versendung. Um die Datenqualität auch beim Eingang – trotz Schwellenanhebungen – auf hohem Niveau zu gewährleisten, sind seit dem Jahr 2022 die nationalen statistischen Ämter der EU-Mitgliedstaaten verpflichtet, tiefgegliederte Mikrodaten über Warenversendungen untereinander auszutauschen. Dieser Mikrodatabaustausch (microdata exchange – MDE) erfolgt über einen zentralen europäischen Server, der die Daten überprüft, nach Bestimmungsland der Waren aufteilt und an den jeweiligen empfangenden Mitgliedstaat sendet. Dadurch erhält jeder Mitgliedstaat nur die Daten, in denen er selbst als Bestimmungsland aufgeführt ist. Diese Daten dürfen dann von den Mitgliedstaaten ausschließlich für statistische Zwecke genutzt werden. Basierend auf diesen ausgetauschten Mikrodaten war es möglich, eingangsseitig ein qualitativ hochwertiges Verfahren zur Aufteilung der Zuschätzungen für befreite Unternehmen zu entwickeln.

Versendungsseitig wird das bisherige Verfahren beibehalten aufgrund des hohen Abdeckungsgrades von 95 % und weil Zusatzinformationen fehlen.

## 6.1 Versendung

Die Aufteilung des Schätzwertes für die deutschen Versendungen auf einzelne Merkmale erfolgt nach einem von Eurostat favorisierten System. Dazu werden die Unternehmen, die knapp über der Meldeschwelle liegen, untersucht. Die Aufteilung des statistischen Wertes der deutschen Versendungen auf die einzelnen Partnerländer, Warenkapitel, Bundesländer und Arten des Geschäfts dieser Einheiten werden auf den Gesamtwert der Schätzungen für Unternehmen unterhalb der Anmeldeschwelle übertragen. Die entsprechende Annahme ist, dass die Exporte beziehungsweise Importe von Unternehmen knapp über der Anmeldeschwelle den Exporten beziehungsweise Importen von Unternehmen knapp unter der Anmeldeschwelle sehr ähnlich sind. Dieses Verfahren wird international als JATT-Methode (Just-Above-the-Threshold-Trade) bezeichnet. Da im Gegensatz zu den Eingangsdaten kein Mikrodaten austausch bei den Versendungsdaten erfolgt, wird diese Methode im Statistischen Bundesamt als das am besten geeignete Schätzverfahren angesehen.

## 6.2 Eingang

Anders als bei der Versendung liegen beim Eingang die zuvor erwähnten spiegelbildlichen Versendungsdaten der anderen Mitgliedstaaten (im Folgenden: MDE-Daten) vor.

Charakteristisch für die Intrahandelsstatistik ist eine sehr asymmetrische Verteilung der gemeldeten Daten: Im Eingang wie in der Versendung entfallen etwa zwei Drittel des Gesamtwertes der Warenverkehre auf die sehr geringe Zahl von etwa 3 % der meldepflichtigen Unternehmen. Gleichzeitig werden bestimmte Waren, zum Beispiel Erdöl, Erdgas oder neue Personenkraftwagen, hauptsächlich von großen Unternehmen importiert. Dadurch weisen diese Warenbereiche auf Detailebene der achtstelligen Warennummern einen hohen Abdeckungsgrad an erhobenen Werten auf. Dies gilt aber nicht für alle Waren in gleichem Maße. Waren, die vorwiegend von kleinen Unternehmen importiert

werden, zum Beispiel Freizeitartikel, Motorräder und Mopeds oder Türen und Fenster, weisen eine Untererfassung auf, weil die Zuschätzungen für die Warenverkehre dieser von der Meldepflicht befreiten Unternehmen derzeit nur auf Kapitelebene erfolgen. Solche Produkte stammen aber zumeist wiederum von großen – im anderen EU-Mitgliedstaat meldepflichtigen – Unternehmen, die diese Waren aus dem EU-Ausland nach Deutschland versenden. Die entsprechenden Warenverkehre können also durch die MDE-Daten abgebildet werden. Die MDE-Daten, die mindestens 95 % der Versendungen des jeweiligen Mitgliedstaates abdecken, lassen sich deshalb für die Verbesserung der Zuschätzungen für die Eingänge der Unternehmen verwenden, die in Deutschland von der Intrastat-Meldepflicht befreit sind.

Eine direkte Übernahme der Versendungsdaten der anderen EU-Mitgliedstaaten als deutsche Eingangsdaten ist derzeit nicht vertretbar. Detaillierte Analysen haben gezeigt, dass die Daten, die das Statistische Bundesamt durch den Mikrodaten austausch von anderen Mitgliedstaaten laufend erhält, nicht ohne Weiteres verwendbar sind. Sie erfüllen die an die Qualität und Nutzbarkeit für die deutsche Importstatistik gesetzten Kriterien – zumindest als hauptsächliche Datenquelle – derzeit nicht. Dazu sind die Probleme zu schwerwiegend, die mit der Verfügbarkeit, Übereinstimmung und Geheimhaltung der Daten einhergehen, und die Vollständigkeit der Datenlieferungen ist unzureichend (Franjić und andere, 2024). Insbesondere ist die Umsetzung eines Einstromverfahrens verbunden mit einem vollständigen Wegfall der Intrastat-Importerhebung derzeit nicht realistisch.

Stattdessen werden die MDE-Daten genutzt, um ein realistisches Bild der Eingänge von Unternehmen, die unter der Meldeschwelle liegen und deshalb nicht zur Meldung ihrer innereuropäischen Wareneingänge verpflichtet sind, zu erhalten. Dazu wird aus den MDE-Daten, die diesen Unternehmen zugeordnet werden können, die Verteilung des statistischen Wertes auf die verschiedenen Merkmale der Außenhandelsstatistik berechnet. Diese Verteilung wird dann auf die Gesamtschätzung für Einheiten unter der Meldeschwelle angewendet.

Der erste Schritt dabei ist, MDE-Daten den Unternehmen unterhalb der Meldeschwelle zuzuordnen. Zu diesem Zweck wird die in den MDE-Daten angegebene Umsatzsteuer-Identifikationsnummer der Handelspartner ver-

wendet. Die Umsatzsteuer-Identifikationsnummer ist ein eindeutiger Identifikator für Einheiten, die am umsatzsteuerlichen Waren- oder Dienstleistungsverkehr in der EU teilnehmen. Der Handelspartner in Deutschland ist in der Regel der deutsche Kunde, der die Waren vom Lieferanten im EU-Ausland erhält. Somit ist es möglich, die im Statistischen Bundesamt vorliegenden deutschen Eingangsdaten und die Daten der Umsatzsteuer-Voranmeldungen einerseits mit den empfangenen MDE-Daten andererseits über die Umsatzsteuer-Identifikationsnummer des betreffenden Unternehmens zu verknüpfen. Die MDE-Daten aus den anderen Mitgliedstaaten können damit sowohl den in Deutschland zu Intrastat meldepflichtigen Unternehmen als auch den Unternehmen unterhalb der Anmeldeschwelle zugeordnet werden.

Um den Schätzwert für Eingänge unterhalb der Anmeldeschwelle aufzuteilen, werden nach der Verknüpfung sämtliche Datensätze aus dem Pool der MDE-Daten herausgenommen, die

- › meldepflichtigen Einheiten beziehungsweise
- › ausländischen Einheiten

zugeordnet sind oder bei denen

- › der Handelspartner als unbekannt angegeben ist und gleichzeitig der Statistische Wert des einzelnen MDE-Datensatzes über der Anmeldeschwelle liegt beziehungsweise
- › der Statistische Wert der Meldung, geteilt durch die Besondere Maßeinheit, über der Anmeldeschwelle liegt.

Diese sorgfältige Abschichtung der MDE-Daten führt zur Annahme, dass die übrigbleibenden Daten Unternehmen unterhalb der Meldeschwelle zugeordnet werden können. Anschließend wird die Verteilung dieser Daten über Partnerländer, achtstellige Warennummern, Bestimmungsbundesländer und Arten des Geschäftes errechnet. Zuletzt wird dann der in Kapitel 4 berechnete Gesamtwert der Zuschätzung anhand dieser Verteilung aufgeteilt und es werden die dazugehörigen Mengenangaben (Eigenmasse und Besondere Maßeinheit) berechnet.

Eine Besonderheit liegt beim Merkmal Bestimmungsbundesland vor. Dieses Merkmal ist nicht direkt in den MDE-Daten enthalten, da es sich um Versendungsdaten der anderen EU-Mitgliedstaaten handelt und eine Erhe-

bung der ausländischen Bestimmungsregion in den Versendungsmeldungen nicht vorgesehen ist. Stattdessen wird das Sitzbundesland des Handelspartners, welcher im Statistischen Bundesamt über seine in den MDE-Daten enthaltene Umsatzsteuer-Identifikationsnummer identifiziert werden kann, verwendet. Da es sich bei den Einheiten unter der Meldeschwelle um kleine Unternehmen handelt, ist die damit verbundene Annahme, dass die Ware in das gleiche Bundesland geliefert wird, in dem sich auch der Unternehmenssitz befindet, vertretbar.

## 7

---

### Fazit und Ausblick

---

Die hier beschriebene indirekte Nutzung der MDE-Daten in der deutschen Importstatistik ermöglicht es, die Anmeldeschwelle im Eingang deutlich anzuheben und somit eine große Zahl von Unternehmen von der Meldepflicht zu befreien – und das, ohne Qualitätseinbußen im Hinblick auf die Detailtiefe hinnehmen zu müssen.


Erstmals ist es durch die neuen Schätzverfahren möglich, für die Eingänge die Zuschätzungen für Befreiungen und teilweise die Zuschätzungen für Antwortausfälle auf Ebene der achtstelligen Warennummern darzustellen. Zusätzlich erfolgen die Zuschätzungen aufgeteilt nach den Arten des Geschäfts, was vor allem für die Berechnung der Leistungsbilanz der Deutschen Bundesbank von Bedeutung ist. Damit verbessert sich die Qualität der Datenzulieferungen für die Leistungsbilanz und die Volkswirtschaftlichen Gesamtrechnungen deutlich.

Die verbesserten Schätzverfahren ermöglichen eine frühere Berücksichtigung der Nachmeldungen von Unternehmen in den veröffentlichten Ergebnissen. Entsprechend wird ab der ersten Veröffentlichung der Ergebnisse für Berichtsmonat Januar 2025 die erste Revision eines Berichtsmonats vier Wochen früher als bisher erfolgen. Infolgedessen wird es für jeden Berichtsmonat sieben statt wie bisher sechs aufeinanderfolgende monatliche Revisionen sowie die Revision der Jahresergebnisse im Oktober des Folgejahres geben.

Aufgrund der modernisierten Schätzverfahren werden auch die Außenhandelsergebnisse der Vorjahre 2022 bis 2024 nochmals leicht angepasst. Diese Korrekturen finden zeitnah zum Veröffentlichungstermin des Januar-



ergebnisses in den Kalendermonaten März und April 2025 statt.

Das Statistische Bundesamt wird auch in Zukunft mithilfe neuer Methoden und der Erschließung neuer Datenquellen Qualitätsverlusten, die durch den Wegfall von Direkterhebungen (beispielsweise durch die Erhöhung der Anmeldeschwellen) drohen, begegnen. Eine Entlastung der Unternehmen durch den Wegfall von statistischen Berichtspflichten bedeutet immer, dass statistische Informationen wegfallen. Die neuen Verfahren können diesen Informationsverlust so gut wie möglich ausgleichen. 

### LITERATURVERZEICHNIS

---

Eurostat (Statistisches Amt der Europäischen Union). *European business statistics compilers' manual for international trade in goods statistics – 2021 edition*. [Zugriff am 14. Januar 2025]. Verfügbar unter: [ec.europa.eu](https://ec.europa.eu)

Eurostat (Statistisches Amt der Europäischen Union). *European business statistics compilers' manual for international trade in goods statistics – detailed data 2024 edition*. [Zugriff am 14. Januar 2025]. Verfügbar unter: [ec.europa.eu](https://ec.europa.eu)

Franjić, Igor/Kolvenbach, Thomas/Tong, Mingyong. *Der europäische Mikrodaten-austausch – neue Datenquelle für die Außenhandelsstatistik*. In: WISTA Wirtschaft und Statistik. Ausgabe 3/2024, Seite 55 ff.

Hastie, Trevor/Tibshirani, Robert/Friedman, Jerome. *The Elements of Statistical Learning*. Ausgabe 2. New York 2009.

### RECHTSGRUNDLAGEN

---

Gesetz über die Statistik des Warenverkehrs mit dem Ausland (Außenhandelsstatistikgesetz – AHStatG) vom 14. Juni 2021 (BGBl. I Seite 1751). [Novelle verabschiedet, aber bis zum Redaktionsschluss noch nicht im Bundesgesetzblatt veröffentlicht.]

Verordnung zur Durchführung des Gesetzes über die Statistik des Warenverkehrs mit dem Ausland (Außenhandelsstatistik-Durchführungsverordnung – AHStatDV) vom 7. Juli 2021 (BGBl. I Seite 2580). [Novelle verabschiedet, aber bis zum Redaktionsschluss noch nicht im Bundesgesetzblatt veröffentlicht.]



Florian Peter

ist Volkswirt und leitet das Referat „Erzeugerpreise, Außenhandelspreise, Großhandelsverkaufspreise“ des Statistischen Bundesamtes. Seine Arbeitsschwerpunkte sind Energiepreise und die Weiterentwicklung der Erzeugerpreise gewerblicher Produkte, Außenhandelspreise und Großhandelsverkaufspreise.

# DIE INDIZES DER AUSSENHANDELS- UND GROSSHANDELSVERKAUFS- PREISE AUF BASIS 2021

Florian Peter

📌 **Schlüsselwörter:** Indexrevision – Wägungsschema – Warenkorb – Berichtsstellen – Preisrepräsentanten – Stichprobe

## ZUSAMMENFASSUNG

In der Regel im Rhythmus von fünf Jahren werden die Indizes der Außenhandelspreise, also der Einfuhr- und Ausfuhrpreise, sowie die Großhandelsverkaufspreise auf eine neue Basis umgestellt und grundlegend überarbeitet. Dabei werden das jeweilige Wägungsschema, der Warenkorb, die Berichtsstellen- und die Preisrepräsentantenstichprobe für das neue Basisjahr ermittelt und erforderliche methodische Änderungen vorgenommen. Die revidierten Ergebnisse für die neue Basis 2021 = 100 hat das Statistische Bundesamt für die Außenhandelspreise im April 2024 und für die Großhandelsverkaufspreise im August 2024 veröffentlicht. Sie ersetzen die bisherigen Ergebnisse auf Basis 2015. Der Beitrag beschreibt die Arbeiten zur Basisumstellung, die wichtigsten Neuerungen und Veränderungen seit der letzten Basisjahrumstellung und analysiert die Unterschiede zwischen den Ergebnissen auf neuer und alter Basis.

📌 **Keywords:** index revision – weighting scheme – basket of goods – reporting units – price representatives – sample

## ABSTRACT

*In general, the index of foreign trade prices, that is export and import prices, and the index of wholesale selling prices are changed to a new base year and fundamentally reviewed every five years. This involves updating the weighting scheme, the basket of goods and the samples of reporting units and price representatives on the basis of the new base year and also implementing any necessary methodological changes. The Federal Statistical Office published the revised results for the new base year (2021 = 100) in April 2024 for foreign trade prices and in August 2024 for wholesale selling prices. They replace the results on base 2015. This article describes the work required to change the base year and the most important new developments and changes since the last rebasing; it also analyses the differences between the results calculated for the new base year and those on the old base.*

## 1

### Einleitung

Die Indizes der Außenhandelspreise messen auf repräsentativer Grundlage die Entwicklung der Preise aller Waren, die zwischen Deutschland und dem Ausland gehandelt werden (Einfuhren und Ausfuhren). Die Indizes der Großhandelsverkaufspreise messen die Entwicklung der Preise aller Waren, die von Großhandelsunternehmen innerhalb Deutschlands verkauft werden. Die Indexberechnung folgt in allen drei Fällen dem Laspeyres-Konzept, das heißt Warenkorb und Wägungsschema werden für eine Basisperiode festgelegt und bleiben bis zur nächsten Basisjahrumstellung unverändert. Der Index zeigt somit die Preisveränderung dieses gewichteten Warenkorbs zwischen der jeweils aktuellen Berichtsperiode und dem Basisjahr. Das Basisjahr wurde 2024 auf das Jahr 2021 umgestellt. In der Regel erfolgen solche Revisionen im Abstand von fünf Jahren, um methodische Weiterentwicklungen, Veränderungen der Handelsstruktur und die Einführung neuer Produkte zu erfassen. Die Indexrevisionen auf das Basisjahr 2021 weichen von diesem normalen Fünf-Jahresrhythmus ab. Hintergrund dafür ist bei den Außenhandelspreisen die Einführung der European-Business-Statistics-Verordnung (EBS-Verordnung)<sup>1</sup>, die den europäischen Unternehmensstatistiken ein einheitliches Regelwerk vorgibt und als nächste Basis das Jahr 2021 nennt. Die Großhandelsverkaufspreise werden nicht über die EBS-Verordnung geregelt. Da die Großhandelsverkaufspreise allerdings für die Deflationierung der Großhandelsstatistik genutzt werden, die wiederum Teil der EBS-Verordnung ist, wurde auch für die Statistik der Großhandelsverkaufspreise 2021 als neues Basisjahr gewählt. Mit der Indexrevision wurden die Warenkörbe und Wägungsschemata, das heißt die Gewichte der Warenkorbpositionen zur Indexberechnung, an die Strukturen der Handelsmärkte im neuen Basisjahr angepasst.

Mit dem Wechsel der Basisperiode wurden nicht nur Gewichte und Warenkörbe an die Verhältnisse des Jahres 2021 angepasst, sondern auch die Berichtsstellen- und Preisrepräsentantenstichproben überarbeitet. Mit

diesen neuen Berechnungsgrundlagen wurden die jeweiligen Gesamtindizes und alle Subindizes von Beginn des neuen Basiszeitraums, Januar 2021, neu berechnet. Die bisher berechneten Indizes auf der Basis 2015 = 100 für den Zeitraum von Januar 2021 bis zum Umstieg auf das neue Basisjahr, also Dezember 2023 für die Außenhandelspreise und März 2024 für die Großhandelsverkaufspreise, wurden durch die neuen Ergebnisse auf Basis 2021 ersetzt.

Kapitel 2 stellt die Umstellungsarbeiten und die Analyse der Ergebnisse für die Außenhandelspreise dar, Kapitel 3 geht auf die Arbeiten für den Großhandelsverkaufspreisindex ein.

## 2

### Außenhandelspreise

#### 2.1 Warenkorb und Wägungsschema

##### Datenquellen

Für die Umstellung der Außenhandelspreise auf das neue Basisjahr 2021 wurden sowohl die Warenkörbe (die Auswahl der Güter für die Preisbeobachtung) als auch die Wägungsschemata (die Gewichtung, mit der diese Güter in die Indexberechnung einfließen) neu ermittelt. Der Gesamtindex der Einfuhr- und Ausfuhrpreise wird als Laspeyres-Index berechnet, ebenso die weiteren Indizes für Aggregate unterhalb des Gesamtindex, also für die Indizes der verschiedenen Güterabteilungen bis hin zu den Güterunterkategorien. Bis zu den Güterunterkategorien liegen Gewichtungsinformationen zu Umsätzen aus der Außenhandelsstatistik vor. Für die Berechnung der Indizes der einzelnen Güterarten aus den Preisreihen der Warenkorbpositionen, also der niedrigsten Hierarchiestufe der Preisindizes, liegen hingegen keine tieferen Gewichtungsinformationen vor. Daher werden auf dieser Stufe ungewichtete Indizes als Mittelwert aus den einzelnen Preismeldungen zur jeweiligen Güterart berechnet, sogenannte Elementarindizes. In Deutschland wird für den Einfuhr- und Ausfuhrpreisindex auf der untersten Ebene der Carli-Index<sup>2</sup> verwendet.

<sup>1</sup> Die Verordnung trat im November 2019 in Kraft. Um sie in allen Mitgliedstaaten frühestmöglich umsetzen zu können, wurde das nächste Basisjahr der betroffenen Unternehmensstatistiken von dem regulären Jahr 2020 auf 2021 verschoben.

<sup>2</sup> Zum Carli-Index siehe Peter, 2024, Seite 43 f.

Die Wägungsschemata auf der neuen Indexbasis 2021 beruhen auf den Einfuhr- und Ausfuhrwerten im Jahr 2021, die von der Außenhandelsstatistik für rund 10 000 Warenarten in der Gliederung des Warenverzeichnisses für die Außenhandelsstatistik (WA), jeweils weiter unterteilt nach Ursprungs- und Bestimmungsländern, erhoben wurden (Statistisches Bundesamt, 2022a).

Zunächst wurden die Daten der Außenhandelsstatistik um bestimmte Warenarten bereinigt, für die entweder eine Preiserfassung vergleichbarer Produkte im Zeitablauf nicht oder nur schwer möglich<sup>3</sup> oder aufgrund der Durchführungsverordnung (EU) 2020/1197<sup>4</sup> nicht verpflichtend ist. Diese Waren machten 2021 rund 2,5 % des Einfuhr- und 3,1 % des Ausfuhrwertes aus.

Die Wägungsschemata wurden gebildet, indem die bereinigten Außenhandelsdaten zu Warenkorbpositionen zusammengefasst und ihnen entsprechend ihrer Außenhandelswerte Gewichte zugewiesen wurden. Das Wägungsschema des Einfuhrpreisindex auf der neuen Basis 2021 umfasst 1 331 Positionen, das des Ausfuhrpreisindex 1 329 Positionen. Zusätzlich wurden die Warenkorbpositionen nach ihrer Herkunft beziehungsweise Bestimmung unterteilt in die Ländergruppen „Län-

der der Eurozone“ beziehungsweise „Länder außerhalb der Eurozone“. Diese Unterscheidung wird getroffen, um zu beurteilen, inwieweit geänderte Wechselkurse die Preisentwicklung beeinflussen.

## Ergebnisse und Analyse

Die Zusammenfassungen zu den Indexexpositionen der Wägungsschemata erfolgten auf der Basis des Güterverzeichnisses für Produktionsstatistiken, Ausgabe 2019 (GP 2019). Auch die Indexergebnisse werden in dieser Gliederung veröffentlicht. Dabei mussten die Daten der Außenhandelsstatistik sehr aufwendig umgeschlüsselt werden, insbesondere vom Warenverzeichnis für die Außenhandelsstatistik in das GP 2019. Zusätzlich werden die Einfuhr- und Ausfuhrpreisindizes auch in der Gliederung der Warengruppen der Ernährungswirtschaft und der gewerblichen Wirtschaft (EGW<sup>5</sup>) dargestellt.

Die Darstellung der Wägungsschemata nach der Gliederung der EGW ermöglicht einen schnellen Überblick über die Zusammensetzung der Ein- und Ausfuhr bezüglich der Verarbeitungsstufe der importierten und exportierten Güter. Ein Vergleich der Grobstrukturen der Wägungsschemata des Einfuhr- und Ausfuhrpreisindex auf neuer Indexbasis nach dieser Gliederung zeigt deutliche Unter-

3 Hierzu zählen beispielsweise Luft-, Wasser- und Raumfahrzeuge sowie Antiquitäten und Kunstgegenstände.

4 Beispielsweise der Handel von Kernbrennstoffen und der Handel mit Waffen und Munition, siehe Durchführungsverordnung (EU) 2020/1197, hier: Anhang Teil B Tabelle 4 (Einfuhrpreise) und Tabelle 5 (Ausfuhrpreise).

5 Das Wägungsschema des Ausfuhrpreisindex enthält auch die Güterabteilung „Reparatur, Instandhaltung und Installation von Maschinen und Ausrüstungen (einschließlich Wartung)“, die in der Gliederung nach EGW keiner Warengruppe zugeordnet werden kann und daher bei diesem Vergleich ausgeschlossen wurde.

**Tabelle 1**

**Wägungsschema der Außenhandelspreise auf Basis 2021**

	Index der Einfuhrpreise			Index der Ausfuhrpreise <sup>1</sup>		
	insgesamt	aus Ländern der Eurozone	aus Ländern außerhalb der Eurozone	insgesamt	in Länder der Eurozone	in Länder außerhalb der Eurozone
Promille						
Gesamtindex	1 000	367,06	632,94	988,34	379,41	608,93
Güter der Ernährungswirtschaft	90,47	50,24	40,23	61,14	33,02	28,12
Güter der Gewerblichen Wirtschaft	909,53	316,82	592,71	927,2	346,39	580,81
Rohstoffe und Halbwaren	170,41	61,76	108,65	70,74	41,01	29,73
Rohstoffe	89,33	23,23	66,1	15,11	11,2	3,91
Halbwaren	81,08	38,53	42,55	55,63	29,81	25,82
Fertigwaren	739,12	255,06	484,06	856,46	305,38	551,08
Vorerzeugnisse	107,51	63,5	44,01	114,23	48,54	65,69
Enderzeugnisse	631,61	191,56	440,05	742,23	256,84	485,39

1 Ohne Güterabteilung 33 „Reparatur, Instandhaltung und Installation von Maschinen und Ausrüstungen (einschließlich Wartung)“; Güterverzeichnis für Produktionsstatistiken, Ausgabe 2019 (GP 2019).

## Die Indizes der Außenhandels- und Großhandelsverkaufspreise auf Basis 2021

schiede. Der Ausführpreisindex wird wesentlich stärker von Fertigwaren dominiert als der Einfuhrpreisindex, Rohstoffe und Halbwaren haben im Einfuhrpreisindex ein höheres Gewicht. Rohstoffe werden in größerem Umfang aus Ländern außerhalb der Eurozone eingeführt als aus Ländern der Eurozone. [↗ Tabelle 1](#)

Weitere Analysen lassen sich durch einen Vergleich der Wägungsschemata auf alter und neuer Basis erstellen. [↗ Tabelle 2](#) stellt die Wägungsanteile des Einfuhrpreisindex nach Güterabteilungen für die beiden Basisjahre

einander gegenüber. [↗ Tabelle 3](#) enthält die gleiche Gegenüberstellung für die Ausführpreise. Bei der Bewertung der Strukturverschiebungen muss unterschieden werden zwischen Veränderungen in den Wägungsanteilen, die auf Preiseffekte und solchen, die auf Mengeneffekte zurückzuführen sind. Hat sich das Gewicht einer Warenkorposition ausschließlich aufgrund eines Preisanstiegs erhöht, führt das unter sonst gleichbleibenden Bedingungen zu einem Anstieg des Gesamtindex, obwohl die reale Handelsstruktur unverändert blieb. Bei einem reinen Mengeneffekt hängt der Einfluss auf

**Tabelle 2**

Einfuhrpreisindex – Wägungsschemata auf Basis 2015 und auf Basis 2021 im Vergleich

	2015	2021	2015 in Preisen von 2021	2015 zu 2021	2015 in Preisen von 2021 zu 2021
	Wägungsanteile in Promille			Veränderung in %	
01 Erzeugnisse der Landwirtschaft und Jagd	36,03	30,06	36,04	– 16,57	– 16,59
02 Forstwirtschaftliche Erzeugnisse	1,03	0,64	0,92	– 37,86	– 30,43
03 Fische und Fischereierzeugnisse	0,84	0,65	0,82	– 22,62	– 20,73
05 Kohle	4,68	4,17	7,16	– 10,90	– 41,76
06 Erdöl und Erdgas	73,74	68,88	94,72	– 6,59	– 27,28
07 Erze	6,88	10,25	12,36	+ 48,98	– 17,07
08 Steine und Erden, sonstige Bergbauerzeugnisse	1,87	1,57	1,78	– 16,04	– 11,80
10 Nahrungsmittel und Futtermittel	52,8	50,56	51,75	– 4,24	– 2,30
11 Getränke	6,91	6,55	6,66	– 5,21	– 1,65
12 Tabakerzeugnisse	1,02	2,7	1,01	+ 164,71	+ 167,33
13 Textilien	12	11,36	11,31	– 5,33	+ 0,44
14 Bekleidung	34,16	31,57	31,78	– 7,58	– 0,66
15 Leder und Lederwaren	15,14	13,2	14,11	– 12,81	– 6,45
16 Holz und Holz-, Kork-, Korb-, Flechtwaren ohne Möbel	7,97	9,00	9,54	+ 12,92	– 5,66
17 Papier, Pappe und Waren daraus	17,79	13,61	17,27	– 23,50	– 21,19
19 Kokereierzeugnisse und Mineralölerzeugnisse	24,69	20,63	25,99	– 16,44	– 20,62
20 Chemische Erzeugnisse	90,92	83,85	97,1	– 7,78	– 13,65
21 Pharmazeutische und ähnliche Erzeugnisse	41,09	60,27	38,23	+ 46,68	+ 57,65
22 Gummi- und Kunststoffwaren	33,67	33,89	32,39	+ 0,65	+ 4,63
23 Glas und -waren, Keramik, Steine und Erden	12,36	11,43	11,97	– 7,52	– 4,51
24 Metalle	59,46	66,04	75,95	+ 11,07	– 13,05
25 Metallerzeugnisse	31,51	31,57	30,92	+ 0,19	+ 2,10
26 Datenverarbeitungsgeräte, elektronische und optische Erzeugnisse	119,54	116,06	90,96	– 2,91	+ 27,59
27 Elektrische Ausrüstungen	60,24	73,83	56,37	+ 22,56	+ 30,97
28 Maschinen	87,02	83,05	81,43	– 4,56	+ 1,99
29 Kraftwagen und Kraftwagenteile	110,03	100,08	103,56	– 9,04	– 3,36
30 Sonstige Fahrzeuge	6,66	8,68	6,08	+ 30,33	+ 42,76
31 Möbel	14,48	14,19	13,68	– 2,00	+ 3,73
32 Waren anderweitig nicht genannt	25,65	26,34	23,33	+ 2,69	+ 12,90
35 Energieversorgung	0,71	3,84	1,96	+ 440,85	+ 95,92
38 Dienstleistungen zu Abfällen und Wertstoffen	9,11	11,48	12,86	+ 26,02	– 10,73

Güterabteilungen des Güterverzeichnis für Produktionsstatistiken, Ausgabe 2019 (GP 2019).

**Tabelle 3**

**Ausfuhrpreisindex – Wägungsschemata auf Basis 2015 und auf Basis 2021 im Vergleich**

	2015	2021	2015 in Preisen von 2021	2015 zu 2021	2015 in Preisen von 2021 zu 2021
	Wägungsanteile in Promille			Veränderung in %	
01 Erzeugnisse der Landwirtschaft und Jagd	9,63	7,47	10,38	– 22,43	– 28,03
02 Forstwirtschaftliche Erzeugnisse	0,39	0,87	0,42	+ 123,08	+ 107,14
03 Fische und Fischereierzeugnisse	.	0,12	.	.	.
05 Kohle	.	0,15	.	.	.
06 Erdöl und Erdgas	8,36	10,02	10,99	+ 19,86	– 8,83
08 Steine und Erden, sonstige Bergbauerzeugnisse	1,4	1,26	1,41	– 10,00	– 10,64
10 Nahrungsmittel und Futtermittel	46,9	47,12	46,65	+ 0,47	+ 1,01
11 Getränke	5,08	4,72	4,96	– 7,09	– 4,84
12 Tabakerzeugnisse	3,47	1,64	3,57	– 52,74	– 54,06
13 Textilien	8,98	8,54	8,59	– 4,90	– 0,58
14 Bekleidung	13,82	17,26	13,4	+ 24,89	+ 28,81
15 Leder und Lederwaren	6,00	8,04	5,76	+ 34,00	+ 39,58
16 Holz und Holz-, Kork-, Korb-, Flechtwaren ohne Möbel	6,41	8,16	7,64	+ 27,30	+ 6,81
17 Papier, Pappe und Waren daraus	18,2	16,28	18,42	– 10,55	– 11,62
19 Kokereierzeugnisse und Mineralölerzeugnisse	14,18	12,95	15,86	– 8,67	– 18,35
20 Chemische Erzeugnisse	97,9	104,65	101,66	+ 6,89	+ 2,94
21 Pharmazeutische und ähnliche Erzeugnisse	52,2	74,07	48,22	+ 41,90	+ 53,61
22 Gummi- und Kunststoffwaren	38,79	40,43	38,43	+ 4,23	+ 5,20
23 Glas und -waren, Keramik, Steine und Erden	13,56	13,74	13,77	+ 1,33	– 0,22
24 Metalle	41,17	44,2	50,34	+ 7,36	– 12,20
25 Metallerzeugnisse	37,44	36,01	38,42	– 3,82	– 6,27
26 Datenverarbeitungsgeräte, elektronische und optische Erzeugnisse	87,4	90,88	75,21	+ 3,98	+ 20,83
27 Elektrische Ausrüstungen	65,92	75,28	65,93	+ 14,20	+ 14,18
28 Maschinen	159,48	152,95	159,51	– 4,09	– 4,11
29 Kraftwagen und Kraftwagenteile	202,19	157,29	192,62	– 22,21	– 18,34
30 Sonstige Fahrzeuge	5,69	6,56	5,46	+ 15,29	+ 20,15
31 Möbel	9,1	9,11	9,41	+ 0,11	– 3,19
32 Waren anderweitig nicht genannt	23,15	25,11	22,81	+ 8,47	+ 10,08
33 Reparatur, Instandhaltung und Installation von Maschinen und Ausrüstungen (einschließlich Wartung)	13,31	11,66	14,5	– 12,40	– 19,59
35 Energieversorgung	1,94	3,94	5,5	+ 103,09	– 28,36
38 Dienstleistungen zu Abfällen und Wertstoffen		9,52	.	.	.

Güterabteilungen des Güterverzeichnis für Produktionsstatistiken, Ausgabe 2019 (GP 2019).

den Gesamtindex vom Verlauf der jeweiligen Preisreihe ab und kann dementsprechend steigend oder dämpfend wirken. Dargestellt sind in den Tabellen auch die Gewichte der Basis 2015 inflationiert zu Preisen von 2021. Der Vergleich der inflationierten Gewichte mit den tatsächlichen Gewichten zeigt, ob die Änderungen des Wägungsschemas in den verschiedenen Güterabteilungen mehr auf Preis- oder mehr auf Mengenänderungen zurückzuführen sind.

Bei der Analyse der Veränderung der Wägungsanteile ist zu beachten, dass ein Rückgang des Gewichts einer Warenkorbposition nicht bedeuten muss, dass auch der Außenhandelsumsatz dieser Warenkorbposition gesunken ist. Die Gewichte geben den Anteil einer Warenkorbposition am gesamten Außenhandel an. Daher kann ein Rückgang eines einzelnen Gewichts daraus resultieren, dass der Umsatz weniger stark gestiegen ist als der Gesamtumsatz aller Warenkorbpositionen. Zuletzt kam



es noch zu geringen Verschiebungen der Gewichte durch den Wechsel der Klassifikation von GP 2009 zu GP 2019, dies hatte allerdings nur sehr geringe Auswirkungen auf die Gewichtung innerhalb einzelner Güterabteilungen.

Tabelle 2 zeigt, dass es beim Wägungsschema der Einfuhrpreise zum Teil zu größeren Verschiebungen der Gewichte einzelner Güterabteilungen gekommen ist. Vor allem die Güterabteilung 21 „Pharmazeutische und ähnliche Erzeugnisse“ hat infolge der Corona-Pandemie sowohl absolut als auch prozentual deutlich an Bedeutung gewonnen. Den prozentualen Anstieg dieser Güterabteilung übertraf noch der Anstieg der Güterabteilung 07 „Erze“ im Wägungsschema der Einfuhrpreise, hier wurden sowohl Eisen- als auch Nichteisenerze in größerem Umfang als 2015 eingeführt. Der höhere Einfuhrwert von Kraftträdern und Fahrrädern sorgte für den Anstieg der Güterabteilung 30 „Sonstige Fahrzeuge“. Bei elektrischem Strom führten die Preissteigerungen im Jahr 2021 zu einem deutlichen prozentualen Anstieg bei dieser Güterabteilung.

Die gesunkenen Einfuhrwerte von Kohle sowie Erdöl und Erdgas, und das bei deutlich gestiegenen Preisen, führten zu den niedrigeren Gewichten dieser Güterabteilungen. Trotz gestiegener Einfuhrwerte haben die Güterabteilungen 01 „Erzeugnisse der Landwirtschaft und Jagd“, 29 „Kraftwagen und Kraftwagenteile“ sowie 17 „Papier, Pappe und Waren daraus“ im Wägungsschema auf Basis 2021 niedrigere Gewichte als auf der alten Basis.

Auch bei der Gewichtung der Ausfuhrpreise ergaben sich auf Basis 2021 gegenüber der Basis 2015 größere Verschiebungen zwischen den Güterabteilungen. Wie bei den Einfuhrpreisen stieg das Gewicht der Güterabteilung 21 „Pharmazeutische und ähnliche Erzeugnisse“ sehr stark an. Gleiches trifft auf die Gewichte der Güterabteilungen 14 „Bekleidung“ sowie 27 „Elektrische Ausrüstungen“ zu. Auch die Güterabteilung 16 „Holz und Korkwaren (ohne Möbel); Flecht- und Korbmacherwaren“ hat aufgrund der Preissteigerungen bis 2021 ein höheres Gewicht am Warenkorb der Ausfuhrpreise als zuvor. Der Anstieg der Ausfuhrwerte dieser Abteilungen führte dazu, dass die beiden größten Güterabteilungen 29 „Kraftwagen und Kraftwagenteile“ sowie 28 „Maschinen“ trotz gestiegener Ausfuhrwerte von 2015 bis 2021 deutlich an Wägungsanteilen am Warenkorb 2021 verloren haben.

## 2.2 Berichtsstellen- und Preisrepräsentantenstichprobe

Auf Grundlage der neuen Wägungsschemata und Warenkörbe wurden die Stichproben für die Berichtsstellen und Preisrepräsentanten vollständig überarbeitet. Zunächst wurden alle bisherigen Preisrepräsentanten dahingehend geprüft, ob und welchen neuen Warenkorbpositionen sie durch den Wechsel der Güterklassifikation zugeordnet werden können. Danach wurde festgelegt, für welche Warenkorbpositionen neue Preisrepräsentanten benötigt werden. Für einige Warenkorbpositionen, besonders bei solchen mit gestiegenen Wägungsanteilen, war die Anzahl der bestehenden Preisreihen zu gering. Für diese Positionen waren deshalb neue Preisreihen aufzubauen, wofür in der Regel neue Berichtsstellen herangezogen und neue Preisrepräsentanten definiert wurden. Schließlich wurden diejenigen Preisrepräsentanten und Berichtsstellen identifiziert, die aus der Berichtspflicht entlassen werden konnten.

Die Stichprobe der Berichtsstellen wird in der Außenhandelspreisstatistik durch eine gezielte Auswahl von Firmen mit hohen Import- beziehungsweise Exportwerten für die betroffene Warenkorbposition (und Ländergruppe) gebildet. Der Berichtskreis der Einfuhr- und Ausfuhrpreisstatistik ist dabei nicht institutionell, sondern funktional abgegrenzt.

### ▢ Abgrenzung des Berichtskreises – funktional oder institutionell?

Die Unterscheidung zwischen funktionaler und institutioneller Abgrenzung des Berichtskreises ist für den Umfang des potenziellen Berichtskreises der Außenhandelspreise entscheidend. Eine institutionelle Abgrenzung bedeutet, dass dem potenziellen Berichtskreis für eine bestimmte Warenkorbposition im Ausfuhrpreisindex nur Unternehmen angehören, welche die Güter, die zu der Warenkorbposition gehören, produzieren und exportieren. Dem Berichtskreis des Einfuhrpreisindex würden nur Unternehmen angehören, die Güter einer Warenkorbposition importieren, die sie selber verbrauchen oder in ihrer Produktion einsetzen. Die funktionale Abgrenzung erweitert den Berichtskreis auf alle Unternehmen, die die Güter einer Warenkorbposition exportieren oder importieren. Dadurch gehören etwa auch Groß- und Zwischenhändler, die selbst nicht produzieren, zum Berichtskreis.

Insgesamt werden sowohl für die Einfuhrpreisstatistik als auch für die Ausfuhrpreisstatistik auf Basis 2021 jeweils etwa 4 800 Berichtsstellen befragt.

## 2.3 Ergebnisse der Indexneuberechnung

Mit den neuen Berechnungsgrundlagen (Warenkorb, Wägungsschema, Berichtsstellen- und Preisrepräsentantenstichprobe) wurden alle Indexergebnisse ab Januar 2021 neu berechnet.

Im Allgemeinen wird erwartet, dass für einen parallel berechneten Zeitraum ein Preisindex auf einer neuen Basis weniger stark ansteigt (beziehungsweise noch stärker sinkt) als der Index auf alter Basis (Neubauer, 1996, hier: Seite 57). Dieser sogenannte Laspeyres-Effekt wird durch mehrere Substitutionseffekte erklärt. Es wird davon ausgegangen, dass die Preiselastizitäten der Güter des Warenkorbs negativ sind und somit Güter, deren Preise steigen, durch Güter mit niedrigeren Preisen oder mit weniger stark steigenden Preisen ersetzt werden. Das Gewicht der Positionen mit einer günstigeren Preisentwicklung im Warenkorb sollte steigen, wenn deren Handelsumsätze hierdurch stärker steigen als der Gesamtumsatz; außerdem sollten die Unternehmen mit einer günstigeren Preisentwicklung mit einer größeren Wahrscheinlichkeit in die Befragung aufgenommen werden, da ihre Umsatzwerte steigen. Für den Zeitraum Januar 2021 bis Dezember 2023 wurden für die Einfuhr- und Ausfuhrpreise jeweils Indizes auf unterschiedlicher Basis berechnet. Sowohl bei den Einfuhrpreisen als

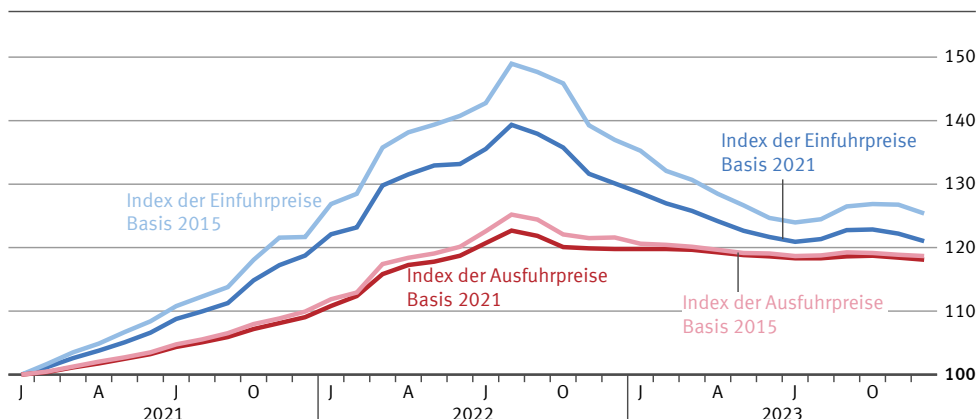
auch bei den Ausfuhrpreisen lag die Veränderungsrate der Gesamtindizes für diesen Zeitraum auf neuer Basis unter der auf alter Basis. So betrug die Veränderungsrate Dezember 2023 gegenüber Januar 2021 für den Einfuhrpreisindex auf alter Basis 25,4 %, auf neuer Basis 21,0 %, bei den Ausfuhrpreisen lagen die Raten bei 18,7 % auf alter Basis und 18,1 % auf neuer Basis.

➤ **Grafik 1** bildet die Verläufe der Gesamtindizes der Einfuhr- und Ausfuhrpreise auf Basis 2015 und Basis 2021 für den vergleichbaren Zeitraum ab. Die Indizes sind für diesen Vergleich jeweils auf Januar 2021 = 100 normiert. Es zeigt sich, dass über den gesamten Vergleichszeitraum sowohl der Index der Einfuhrpreise als auch der Index der Ausfuhrpreise auf neuer Basis 2021 unter den Vergleichsindizes auf alter Basis 2015 liegen. Die größte Differenz zwischen den Indizes auf alter und neuer Basis erreichten die Einfuhrpreise im Oktober 2022 mit 10,1 Prozentpunkten, die größte Differenz bei den Ausfuhrpreisindizes wurde im September 2022 mit 2,6 Prozentpunkten erreicht.

Eine Differenzierung der Abweichungen nach möglichen Ursachen ist nur eingeschränkt möglich. Zumeist treten verschiedene Ursachen gemeinsam auf, die zu den unterschiedlichen Verläufen auf neuer und alter Basis führen. Mit dem Austausch von Warenkorpositionen, der neuen Gewichtung der Warenkorpositionen und der Aktualisierung von Berichtsstellen liegen mehrere Gründe für einen veränderten Verlauf vor, deren Einflüsse nur sehr schwer voneinander zu trennen sind.

### Grafik 1

Index der Einfuhrpreise und Index der Ausfuhrpreise auf alter und neuer Basis im Vergleich  
Januar 2021 = 100



## Die Indizes der Außenhandels- und Großhandelsverkaufspreise auf Basis 2021

Neben dem Gesamtindex zeigt ein Vergleich der Veränderungsraten der einzelnen Güterabteilungen zwischen Januar 2021 und Dezember 2023 ein differenzierteres Bild. [↗ Tabelle 4](#)

Von 31 zu vergleichenden Güterabteilungen bei den Einfuhrpreisen lagen bei 17 die Veränderungsraten auf Basis 2021 unter denen auf Basis 2015. Die größte negative Abweichung verzeichnet die Güterabteilung

**Tabelle 4**

### Veränderungsraten der Außenhandelspreise im Vergleich

	Einfuhrpreisindex			Ausfuhrpreisindex		
	Basis 2021	Basis 2015	Differenz in Prozentpunkten	Basis 2021	Basis 2015	Differenz in Prozentpunkten
	Veränderung Januar 2021 zu Dezember 2023 in %			Veränderung Januar 2021 zu Dezember 2023 in %		
Gesamtindex <sup>1</sup>	+ 21,0	+ 25,4	– 4,4	+ 18,1	+ 18,7	– 0,6
01 Erzeugnisse der Landwirtschaft und Jagd	+ 25,6	+ 27,7	– 2,1	+ 29,9	+ 23,5	+ 6,4
02 Forstwirtschaftliche Erzeugnisse	+ 44,7	+ 56,3	– 11,6	+ 30,5	+ 32,5	– 2,0
03 Fische und Fischereierzeugnisse	+ 56,7	+ 56,4	+ 0,3	+ 69,9	+ 50,8	+ 19,1
05 Kohle	+ 148,9	+ 136,3	+ 12,6	+ 70,6	+ 79,7	– 9,1
06 Erdöl und Erdgas	+ 83,2	+ 101,6	– 18,4	+ 138,8	+ 76,6	+ 62,2
07 Erze	+ 11,7	+ 13,9	– 2,2	–	–	–
08 Steine und Erden, sonstige Bergbauerzeugnisse	+ 30,3	+ 28,7	+ 1,6	+ 27,6	+ 25,00	+ 2,6
10 Nahrungsmittel und Futtermittel	+ 32,0	+ 33,1	– 1,1	+ 32,0	+ 33,0	– 1,0
11 Getränke	+ 19,3	+ 17,8	+ 1,5	+ 19,2	+ 15,3	+ 3,9
12 Tabakerzeugnisse	+ 23,1	+ 18,3	+ 4,8	+ 23,2	+ 22,4	+ 0,8
13 Textilien	+ 16,9	+ 18,3	– 1,4	+ 13,9	+ 15,9	– 2,0
14 Bekleidung	+ 12,6	+ 12,6	+ 0,0	+ 10,5	+ 9,7	+ 0,8
15 Leder und Lederwaren	+ 15,6	+ 16,0	– 0,4	+ 6,3	+ 7,7	– 1,4
16 Holz und Holz-, Kork-, Korb-, Flechtwaren ohne Möbel	+ 30,4	+ 31,7	– 1,3	+ 22,8	+ 26,3	– 3,5
17 Papier, Pappe und Waren daraus	+ 30,3	+ 31,8	– 1,5	+ 26,2	+ 28,1	– 1,9
19 Kokereierzeugnisse und Mineralölerzeugnisse	+ 74,5	+ 77,4	– 2,9	+ 69,7	+ 69,8	– 0,1
20 Chemische Erzeugnisse	+ 21,8	+ 23,1	– 1,3	+ 22,8	+ 23,9	– 1,1
21 Pharmazeutische und ähnliche Erzeugnisse	+ 10,4	+ 9,4	+ 1,0	+ 7,1	+ 9,1	– 2,0
22 Gummi- und Kunststoffwaren	+ 13,8	+ 12,6	+ 1,2	+ 21,0	+ 18,1	+ 2,9
23 Glas und -waren, Keramik, Steine und Erden	+ 20,1	+ 21,5	– 1,4	+ 27,0	+ 27,8	– 0,8
24 Metalle	+ 19,0	+ 23,1	– 4,1	+ 26,1	+ 26,6	– 0,5
25 Metallerzeugnisse	+ 20,2	+ 19,5	+ 0,7	+ 20,8	+ 21,9	– 1,1
26 Datenverarbeitungsgeräte, elektronische und optische Erzeugnisse	+ 7,5	+ 8,7	– 1,2	+ 8,7	+ 9,3	– 0,6
27 Elektrische Ausrüstungen	+ 16,8	+ 16,5	+ 0,3	+ 16,9	+ 18,6	– 1,7
28 Maschinen	+ 14,8	+ 14,6	+ 0,2	+ 16,6	+ 16,9	– 0,3
29 Kraftwagen und Kraftwagenteile	+ 14,5	+ 14,7	– 0,2	+ 13,3	+ 13,6	– 0,3
30 Sonstige Fahrzeuge	+ 21,5	+ 20,4	+ 1,1	+ 13,6	+ 10,2	+ 3,4
31 Möbel	+ 17,4	+ 18,1	– 0,7	+ 20,1	+ 21,3	– 1,2
32 Waren anderweitig nicht genannt	+ 12,0	+ 11,1	+ 0,9	+ 11,4	+ 11,7	– 0,3
33 Reparatur, Instandhaltung und Installation von Maschinen und Ausrüstungen (einschließlich Wartung)	–	–	–	+ 6,8	+ 6,2	+ 0,6
35 Energieversorgung	+ 27,3	+ 27,3	+ 0,0	+ 27,3	+ 27,3	+ 0,0
38 Dienstleistungen zu Abfällen und Wertstoffen	+ 13,6	+ 14,3	– 0,7	+ 13,6	+ 11,5	+ 2,1

<sup>1</sup> Nach dem Güterverzeichnis für Produktionsstatistiken, Ausgabe 2019 (GP 2019).

06 „Erdöl und Erdgas“. Zum Teil ist die unterschiedliche Entwicklung hier auf die veränderte Gewichtung innerhalb der Güterabteilung zurückzuführen. Außerdem wurde Flüssigerdgas neu als eigene Warenkorbposition aufgenommen. Allerdings hat auch die Berechnungsmethode der Indizes auf unterster Warenkorbebene einen Einfluss auf die Indexverläufe vor und nach der Indexrevision. Die Berechnung der Indizes auf unterster Hierarchiestufe mit dem Carli-Index kann bei besonders hohen und inhomogenen Preisveränderungen, wie sie in der Energiekrise nach 2021 bei Erdöl und Erdgas vorlagen, innerhalb einer Warenkorbposition dazu führen, dass es durch die rechnerische Umbasierung der Preisindizes zu unterschiedlichen Verläufen auf neuer und alter Basis kommt.<sup>6</sup> Die Auswirkungen der verwendeten Formel für die Berechnung der Elementarindizes bei der Umbasierung 2021 werden noch weiter untersucht. Dabei werden auch Vergleichsrechnungen mit alternativen Indexformeln Teil der Analyse sein, um zu entscheiden, ob bei stärker schwankenden Preisen die Verwendung des Carli-Index für die Berechnung der Elementarindizes der Außenhandelspreise in Deutschland weiterhin die beste Option darstellt.

Die negativen Abweichungen der Güterabteilungen forstwirtschaftliche Erzeugnisse und Metalle lassen sich zum größten Teil durch die veränderte Berichtsstellenstichprobe und durch Änderungen im Warenkorb und der Gewichtung erklären.

Eine auffallend große positive Abweichung, also eine höher liegende Veränderungsrate des Index auf neuer Basis gegenüber dem Index auf alter Basis, gab es in der Güterabteilung Kohle. Auch hier führte eine Mischung aus geänderter Berichtsstellenstichprobe (vor allem bei der Warenkorbposition Koks-kohle) und veränderter Gewichtung zu der Abweichung.

Bei den Ausführpreisen lag bei 19 der 31 zu vergleichenden Güterabteilungen die Veränderungsrate zwischen Januar 2021 und Dezember 2023 auf Basis 2021 unter der auf Basis 2015. Auch hier gab es die größte negative Abweichung in der Güterabteilung Erdöl und Erdgas. Der Einfluss der Berechnungsmethode ist hier deutlich größer als bei den Einfuhrpreisen. Die größte positive Abweichung verzeichnete die Güterabteilung Fische und Fischereierzeugnisse. Dabei führten auch Veränderun-

gen in der Berichtsstellen- und Produktstichprobe zu dem abweichenden Verlauf, vor allem in der zweiten Jahreshälfte 2023.

Im Zuge einer Basisumstellung müssen auch die bisherigen Indexergebnisse vor Januar 2021 auf das neue Basisjahr 2021 umgerechnet werden. Diese langen Indexreihen, mit Monats- und Jahreswerten teilweise ab 1962, werden mittels multiplikativer Verknüpfung der Werte auf alter Basis mit den neuen Indexergebnissen gebildet. Hierfür werden die Indexwerte auf Basis 2015 vor 2021 mit den Indexwerten auf Basis 2021 ab 2021 mithilfe des Quotienten der Januarwerte 2021 auf neuer und alter Basis verkettet.

## 3

### Großhandelsverkaufspreisindex

#### 3.1 Warenkorb und Wägungsschema

##### Datenquelle

Im Gegensatz zu den Außenhandelspreisindizes, die die Preisentwicklung der gehandelten Güter auf Ebene der Güterabteilungen widerspiegeln, stellt der Großhandelsverkaufspreisindex die Preisentwicklung von Verkäufen der Unternehmen des Wirtschaftszweiges Großhandel und der darin enthaltenen Wirtschaftszweigklassen dar. Das bedeutet, dass zum Beispiel der Preisindex für den Großhandel mit lebenden Tieren nicht nur die Preisentwicklung beim Handel mit lebenden Tieren enthält, sondern die Preisentwicklung von allen beobachteten Gütern berücksichtigt, die von Unternehmen mit diesem wirtschaftlichen Schwerpunkt gehandelt werden. Der Großhandelsverkaufspreisindex wird daher in zwei Stufen berechnet. Zunächst wird die Preisentwicklung von den im Großhandel gehandelten Gütern berechnet, unabhängig vom wirtschaftlichen Schwerpunkt der Unternehmen, die die entsprechenden Güter handeln. Aus diesen Preisentwicklungen werden danach die Indizes für die Wirtschaftszweige, gewichtet nach den Umsätzen der gehandelten Güter des Wirtschaftszweiges, berechnet. Dementsprechend erfolgt auch die Festlegung des Wägungsschemas des Großhandelsverkaufspreisindex in einem mehrstufigen Verfahren.

<sup>6</sup> Für weiterführende Informationen zur Indexrevision 2021 der Erzeugerpreise gewerblicher Produkte siehe Peter (2024).

Die Datengrundlage für das Wägungsschema bildet die Jahrerhebung im Großhandel, zu deren Erhebungsprogramm der Jahresumsatz, Investitionen, Warenbezüge und Warenbestände am Anfang eines Jahres<sup>17</sup> gehören. Der Gesamtumsatz wird nach ausgeübter wirtschaftlicher Tätigkeit und nach Produktgruppen gemäß dem Abschnitt G „Großhandelsleistungen (ohne Handelsleistungen mit Kraftfahrzeugen)“ der Classification of Products by Activity (CPA 2.2; Statistische Güterklassifikation in Verbindung mit den Wirtschaftszweigen) weiter aufgegliedert erfasst. Für den Großhandel umfasst diese Gliederung 60 Produktgruppen (Statistisches Bundesamt, 2022b).

In einem ersten Schritt werden die Gewichte der Wirtschaftszweigklassen und -unterklassen des Großhandels anhand der durch die Jahrerhebung im Großhandel ermittelten Umsätze gebildet. Die Umsätze auf dieser ersten Stufe sind die Umsätze der Großhandelsunternehmen nach ihrem wirtschaftlichen Schwerpunkt. Großhandelsunternehmen handeln in der Regel aber nicht nur mit Produkten, die ihrem wirtschaftlichen Schwerpunkt entsprechen, sondern häufig mit Waren unterschiedlicher Wirtschaftszweige. So handeln Unternehmen, die nach ihrem wirtschaftlichen Schwerpunkt dem Wirtschaftszweig Großhandel mit Bekleidung zugeordnet werden, häufig auch mit Produkten, die dem Wirtschaftszweig Großhandel mit Schuhen zugeordnet werden. Um dies im Wägungsschema abzubilden, werden in einem zweiten Schritt die einzelnen Wirtschaftszweigklassen nach ihren Handelstätigkeiten mit verschiedenen Produktgruppen gewichtet. Die Gewichte für die ersten beiden Wägungsschritte lassen sich somit aus den regulären Ergebnissen der Jahrerhebung im Großhandel ableiten.

Das aus den ersten beiden Schritten gewonnene Wägungsschema und die Produktgliederung sind für die Zwecke der Preiserhebung noch sehr grob. Die tiefste Gliederung der CPA 2.2 besteht im Abschnitt G zum Beispiel aus der Produktgruppe „Getreide, Rohtabak, Saatgut und Futtermittel“. Die jeweiligen Umsätze für die einzelnen Produkte, also für verschiedene Getreidesorten, für Rohtabak, für Saatgut oder für Futtermittel sind allerdings nicht bekannt. Diese in der Regel sehr heterogenen Produktgruppen umfassen unterschiedlichste

Produkte mit genauso unterschiedlichen Preisentwicklungen. Daher wird eine weitere Gliederung und Gewichtung der Produktgruppen benötigt, um die Entwicklung der Großhandelsverkaufspreise zu ermitteln. In einem dritten Schritt werden deshalb die einzelnen gehandelten Produktgruppen nochmals detaillierter nach einzelnen Produkten gegliedert und gewichtet. Hierfür wurde die Jahrerhebung im Großhandel für das Jahr 2021 wie schon für 2015 (Peter, 2019) um eine tiefere Produktgliederung erweitert.

Der Index der Großhandelsverkaufspreise gibt die Preisentwicklung der vom Großhandel im Inland verkauften Güter an. Die zur Gewichtung verwendete Jahrerhebung im Großhandel erfasst allerdings den Großhandel insgesamt, also auch den Exporthandel der inländischen Großhändler. Für das Basisjahr 2021 wurde nun erstmals eine Bereinigung der Großhandelsumsätze um den Außenhandelsanteil vorgenommen, um die Handelsstrukturen für den Großhandelsverkaufspreisindex besser abbilden zu können.

Die Bereinigung der Großhandelsumsätze erfolgt mit den Umsatzdaten der Außenhandelsstatistik. Eine Herausforderung war hierbei, die Großhandelsstatistik mit der Außenhandelsstatistik zu verknüpfen. In einem ersten Schritt waren die Produktklassifikationen der beiden Statistiken zusammenzuführen. Beide Statistiken nutzen unterschiedliche Klassifikationssysteme zur Produktklassifikation: die Außenhandelsstatistik das Warenverzeichnis für die Außenhandelsstatistik (WA), also eine Klassifikation nach Warennummern, die Großhandelsstatistik die Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008), beziehungsweise eine für die Erhebung des Basisjahres des Großhandelsverkaufspreisindex eigens erstellte Produktklassifikation. Es existierte jedoch keine Zuordnung von der WA zur Produktklassifikation im Großhandel. Eine solche wurde erstellt, um die beiden Statistiken auf Produktgruppen-Ebene zu verknüpfen.

In einem zweiten Schritt wurden die beiden Statistiken auf Unternehmensebene verknüpft um festzustellen, welches Großhandelsunternehmen auch Außenhandel betreibt. Dazu mussten die von der Großhandelsstatistik genutzten Identifikatoren des statistischen Unternehmensregisters und das von der Außenhandelsstatistik verwendete eigene Register zusammengeführt und miteinander verknüpft werden. Um diese Verknüpfung zu

<sup>17</sup> Des Weiteren erfasst werden die Anzahl der Beschäftigten, die Bruttoentgelte, die Sozialabgaben und Subventionen.



**Tabelle 5**

**Außenhandelsanteil am Großhandelsumsatz nach Wirtschaftszweigunterklassen**

	Außenhandelsanteil am Großhandelsumsatz
	%
46.21.0 Großhandel mit Getreide, Rohtabak, Saatgut und Futtermitteln	10,21
46.22.0 Großhandel mit Blumen und Pflanzen	9,6
46.23.0 Großhandel mit lebenden Tieren	2,69
46.24.0 Großhandel mit Häuten, Fellen und Leder	73,73
46.31.0 Großhandel mit Obst, Gemüse und Kartoffeln	2,56
46.32.0 Großhandel mit Fleisch und Fleischwaren	9,71
46.33.0 Großhandel mit Milch, Milcherzeugnissen, Eiern, Speiseölen und Nahrungsfetten	8,96
46.34.0 Großhandel mit Getränken	4,55
46.35.0 Großhandel mit Tabakwaren	3,62
46.36.0 Großhandel mit Zucker, Süßwaren und Backwaren	7,1
46.37.0 Großhandel mit Kaffee, Tee, Kakao und Gewürzen	15,63
46.38.1 Großhandel mit Fisch und Fischerzeugnissen	17,09
46.38.2 Großhandel mit Mehl und Getreideprodukten	16,54
46.38.9 Großhandel mit Nahrungs- und Genussmitteln a. n. g.	10,36
46.41.0 Großhandel mit Textilien	30,8
46.42.1 Großhandel mit Bekleidung	25,2
46.42.2 Großhandel mit Schuhen	22,78
46.43.1 Großhandel mit Foto- und optischen Erzeugnissen	25,27
46.43.2 Großhandel mit elektrischen Haushaltsgeräten	8,91
46.43.3 Großhandel mit Geräten der Unterhaltungselektronik	11,22
46.44.1 Großhandel mit keramischen Erzeugnissen und Glaswaren	21,71
46.44.2 Großhandel mit Wasch-, Putz- und Reinigungsmitteln	16,62
46.45.0 Großhandel mit kosmetischen Erzeugnissen und Körperpflegemitteln	15,11
46.46.1 Großhandel mit pharmazeutischen Erzeugnissen	22,19
46.46.2 Großhandel mit medizinischen und orthopädischen Artikeln, Dental- und Laborbedarf	13,28
46.47.0 Großhandel mit Möbeln, Teppichen, Lampen und Leuchten	19,79
46.48.0 Großhandel mit Uhren und Schmuck	27,09
46.49.1 Großhandel mit Spielwaren und Musikinstrumenten	18,43
46.49.2 Großhandel mit Fahrrädern, Fahrradteilen und -zubehör, Sport- und Campingartikeln (ohne Campingmöbel)	12,22
46.49.3 Großhandel mit Lederwaren, Reisegepäck, Geschenk- und Werbeartikeln	27,92
46.49.4 Großhandel mit Karton, Papier, Pappe, Schreibwaren, Bürobedarf, Büchern, Zeitschriften und Zeitungen	12,6
46.49.5 Großhandel mit nicht elektrischen Haushaltsgeräten, Haushaltswaren aus Metall sowie sonstigen Gebrauchs- und Verbrauchsgütern a. n. g.	21,07
46.51.0 Großhandel mit Datenverarbeitungsgeräten, peripheren Geräten und Software	12,51
46.52.0 Großhandel mit elektronischen Bauteilen und Telekommunikationsgeräten	28,86
46.61.0 Großhandel mit landwirtschaftlichen Maschinen und Geräten	25,95
46.62.0 Großhandel mit Werkzeugmaschinen	21,18
46.63.0 Großhandel mit Bergwerks-, Bau- und Baustoffmaschinen	8,44
46.64.0 Großhandel mit Textil-, Näh- und Strickmaschinen	14,68
46.65.0 Großhandel mit Büromöbeln	2,91
46.66.0 Großhandel mit sonstigen Büromaschinen und -einrichtungen	62,96
46.69.1 Großhandel mit Flurförderzeugen und Fahrzeugen a. n. g.	9,25
46.69.2 Großhandel mit sonstigen Maschinen	25,57

Klassen der Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008).

## noch Tabelle 5

Außenhandelsanteil am Großhandelsumsatz nach Wirtschaftszweigunterklassen

	Außenhandelsanteil am Großhandelsumsatz
	%
46.69.3 Großhandel mit sonstigen Ausrüstungen und Zubehör für Maschinen sowie mit technischem Bedarf	50,92
46.71.1 Großhandel mit festen Brennstoffen	7,41
46.71.2 Großhandel mit Mineralölerzeugnissen	3,87
46.72.1 Großhandel mit Eisenerzen, Eisen, Stahl, Eisen- und Stahlhalbzeug	16,53
46.72.2 Großhandel mit NE-Erzen, NE-Metallen und NE-Metallhalbzeug	13,29
46.73.2 Großhandel mit Roh- und Schnittholz	17,41
46.73.3 Großhandel mit sonstigen Holzhalbwaren sowie Bauelementen aus Holz	11,8
46.73.4 Großhandel mit Baustoffen und Bauelementen aus mineralischen Stoffen	3,8
46.73.5 Großhandel mit Flachglas	13,83
46.73.6 Großhandel mit Anstrichmitteln	11,99
46.73.7 Großhandel mit Sanitärkeramik	0,67
46.73.8 Großhandel mit Tapeten und Bodenbelägen (ohne Teppiche)	7,97
46.74.1 Großhandel mit Werkzeugen und Kleisenwaren	23,15
46.74.2 Großhandel mit Installationsbedarf für Gas, Wasser und Heizung	6,66
46.74.3 Großhandel mit Metall- und Kunststoffwaren für Bauzwecke	17,01
46.75.0 Großhandel mit chemischen Erzeugnissen	25,61
46.76.0 Großhandel mit sonstigen Halbwaren	40,67
46.77.0 Großhandel mit Altmaterialien und Reststoffen	21,77

Klassen der Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008).

ermöglichen, wurde auf die Ergebnisse eines Projekts zur Mikrodatenverknüpfung der Außenhandelsstatistik und der Unternehmensstatistiken (Kruse und andere, 2021) zurückgegriffen. Zum Zeitpunkt der Erstellung des Wägungsschemas des Großhandelsverkaufspreisindex standen allerdings nur Ergebnisse aus dem Projekt für das Jahr 2020 zur Verfügung; diese wurden mithilfe der Umsätze aus der Außenhandelsstatistik für 2021 fortgeschrieben.

Mit den verknüpften Produktklassifikationen und Einzeldaten der Großhandels- und Außenhandelsstatistik sowie der Information, welche Großhandelsunternehmen auch Außenhandel betreiben, war es möglich, die Großhandelsumsätze auf Ebene der tieferen Produktgliederung auf inländische Umsätze zu bereinigen. In [Tabelle 5](#) sind als Ergebnis des Bereinigungsverfahrens die Außenhandelsanteile am Großhandel auf Ebene der Wirtschaftszweigunterklassen<sup>8</sup> abgetragen.

<sup>8</sup> Bei der Analyse wurden die Wirtschaftszweigunterklassen ohne ausgeprägten Schwerpunkt nicht berücksichtigt. Die Großhandelsunternehmen werden für den Berichtskreis des Großhandelsverkaufspreisindex funktional abgegrenzt, daher sind die Unternehmen beziehungsweise ihre Außenhandelsumsätze, die mit wirtschaftlichem Schwerpunkt den Wirtschaftszweigunterklassen ohne ausgeprägten Schwerpunkt angehören, in den anderen Wirtschaftszweigunterklassen berücksichtigt.

Die Anteile unterscheiden sich deutlich zwischen den Wirtschaftszweigunterklassen und liegen zwischen 0,7 % bei Sanitärkeramik und 63 % bei sonstigen Büromaschinen und -einrichtungen.

Ein wichtiges Ergebnis dieser Bereinigung ist, dass die Außenhandelsstrukturen im Großhandel nicht homogen über den gesamten Großhandel sind und eine Bereinigung um den Außenhandelsanteil zu einer besseren Abbildung der Preisentwicklung des inländischen Großhandels führt. Dies hat allerdings auch Auswirkungen auf die Vergleichbarkeit der Wägungsanteile im Wägungsschema auf neuer Basis 2021 und alter Basis 2015.

## Ergebnisse und Analyse

Mit den Umsätzen aus der um die tiefere Produktgliederung erweiterten und um die Außenhandelsumsätze bereinigten Jahreserhebung im Großhandel wurden der Warenkorb und das Wägungsschema für den Großhandelsverkaufspreisindex gebildet. [Tabelle 6](#) stellt das Wägungsschema auf Ebene der Wirtschaftszweigunterklassen der WZ 2008 dar. Bei der Analyse der Veränderungen der Wägungsanteile gelten dieselben



**Tabelle 6**

**Index der Großhandelsverkaufspreise – Wägungsschemata auf Basis 2015 und Basis 2021 im Vergleich**

	2015	2021	2015 in Preisen von 2021	2015 zu 2021	2015 in Preisen von 2021 zu 2021
	Wägungsanteile in Promille			Veränderung in %	
46.21.0 Großhandel mit Getreide, Rohtabak, Saatgut und Futtermitteln	33,07	27,01	38,03	– 18,32	– 28,98
46.22.0 Großhandel mit Blumen und Pflanzen	3,41	5,14	3,55	+ 50,73	+ 44,79
46.23.0 Großhandel mit lebenden Tieren	14,08	11,26	12,64	– 20,03	– 10,92
46.31.0 Großhandel mit Obst, Gemüse und Kartoffeln	27,22	28,63	26,89	+ 5,18	+ 6,47
46.32.0 Großhandel mit Fleisch und Fleischwaren	15,19	13,81	14,52	– 9,08	– 4,89
46.33.0 Großhandel mit Milch, Milcherzeugnissen, Eiern, Speiseölen und Nahrungsfetten	11,93	11,28	12,47	– 5,45	– 9,54
46.34.0 Großhandel mit Getränken	20,38	28,16	19,77	+ 38,17	+ 42,44
46.35.0 Großhandel mit Tabakwaren	19,38	17,59	20,6	– 9,24	– 14,61
46.36.0 Großhandel mit Zucker, Süßwaren und Backwaren	7,34	10,65	6,88	+ 45,10	+ 54,80
46.37.0 Großhandel mit Kaffee, Tee, Kakao und Gewürzen	5,88	3,83	5,3	– 34,86	– 27,74
46.38.1 Großhandel mit Fisch und Fischerzeugnissen	3,81	2,91	3,65	– 23,62	– 20,27
46.38.2 Großhandel mit Mehl und Getreideprodukten	1,33	1,2	1,28	– 9,77	– 6,25
46.38.9 Großhandel mit Nahrungs- und Genussmitteln a. n. g.	33,59	36,25	31,86	+ 7,92	+ 13,78
46.39.1 Großhandel mit tiefgefrorenen Nahrungsmitteln, ohne ausgeprägten Schwerpunkt	1,38	2,04	1,32	+ 47,83	+ 54,55
46.39.9 Großhandel mit sonstigen Nahrungs- und Genussmitteln, Getränken und Tabakwaren, ohne ausgeprägten Schwerpunkt	22,45	43,2	21,84	+ 92,43	+ 97,80
46.41.0 Großhandel mit Textilien	4,62	3,05	4,47	– 33,98	– 31,77
46.42.1 Großhandel mit Bekleidung	17,31	10,84	16,74	– 37,38	– 35,24
46.42.2 Großhandel mit Schuhen	4,53	4,24	4,34	– 6,40	– 2,30
46.43.1 Großhandel mit Foto- und optischen Erzeugnissen	6,29	5,2	5,92	– 17,33	– 12,16
46.43.2 Großhandel mit elektrischen Haushaltsgeräten	8,80	9,86	8,24	+ 12,05	+ 19,66
46.43.3 Großhandel mit Geräten der Unterhaltungselektronik	17,75	12,65	15,26	– 28,73	– 17,10
46.44.1 Großhandel mit keramischen Erzeugnissen und Glaswaren	2,08	1,98	2,01	– 4,81	– 1,49
46.44.2 Großhandel mit Wasch-, Putz- und Reinigungsmitteln	4,55	5,26	4,37	+ 15,60	+ 20,37
46.45.0 Großhandel mit kosmetischen Erzeugnissen und Körperpflegemitteln	10,06	8,49	9,37	– 15,61	– 9,39
46.46.1 Großhandel mit pharmazeutischen Erzeugnissen	64,55	62,47	57,85	– 3,22	+ 7,99
46.46.2 Großhandel mit medizinischen und orthopädischen Artikeln, Dental- und Laborbedarf	26,49	37,29	25,1	+ 40,77	+ 48,57
46.47.0 Großhandel mit Möbeln, Teppichen, Lampen und Leuchten	7,77	8,07	6,97	+ 3,86	+ 15,78
46.48.0 Großhandel mit Uhren und Schmuck	3,47	2,23	3,55	– 35,73	– 37,18
46.49.1 Großhandel mit Spielwaren und Musikinstrumenten	3,60	4,48	3,54	+ 24,44	+ 26,55
46.49.2 Großhandel mit Fahrrädern, Fahrradteilen und -zubehör, Sport- und Campingartikeln (ohne Campingmöbel)	6,15	9,51	5,69	+ 54,63	+ 67,14
46.49.3 Großhandel mit Lederwaren, Reisegepäck, Geschenk- und Werbeartikeln	4,44	2,55	4,2	– 42,57	– 39,29
46.49.4 Großhandel mit Karton, Papier, Pappe, Schreibwaren, Bürobedarf, Büchern, Zeitschriften und Zeitungen	18,55	15,42	18,39	– 16,87	– 16,15
46.49.5 Großhandel mit nicht elektrischen Haushaltsgeräten, Haushaltswaren aus Metall sowie sonstigen Gebrauchs- und Verbrauchsgütern a. n. g.	6,27	4,58	6,1	– 26,95	– 24,92
46.51.0 Großhandel mit Datenverarbeitungsgeräten, peripheren Geräten und Software	51,12	57,29	38,23	+ 12,07	+ 49,86
46.61.0 Großhandel mit elektronischen Bauteilen und Telekommunikationsgeräten	33,08	32,7	26,91	– 1,15	+ 21,52
46.52.0 Großhandel mit landwirtschaftlichen Maschinen und Geräten	13,10	14,46	13,23	+ 10,38	+ 9,30
46.62.0 Großhandel mit Werkzeugmaschinen	6,19	7,59	5,99	+ 22,62	+ 26,71

Klassen der Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008).

# Die Indizes der Außenhandels- und Großhandelsverkaufspreise auf Basis 2021

## noch Tabelle 6

### Index der Großhandelsverkaufspreise – Wägungsschemata auf Basis 2015 und Basis 2021 im Vergleich

	2015	2021	2015 in Preisen von 2021	2015 zu 2021	2015 in Preisen von 2021 zu 2021
	Wägungsanteile in Promille			Veränderung in %	
46.63.0 Großhandel mit Bergwerks-, Bau- und Baustoffmaschinen	5,70	8,93	5,61	+ 56,67	+ 59,18
46.65.0 Großhandel mit Büromöbeln	2,19	2,15	2,21	– 1,83	– 2,71
46.66.0 Großhandel mit sonstigen Büromaschinen und -einrichtungen	2,36	0,96	2,11	– 59,32	– 54,50
46.69.1 Großhandel mit Flurförderzeugen und Fahrzeugen a. n. g.	4,34	6,25	4,37	+ 44,01	+ 43,02
46.69.2 Großhandel mit sonstigen Maschinen	24,66	22,22	23,91	– 9,89	– 7,07
46.69.3 Großhandel mit sonstigen Ausrüstungen und Zubehör für Maschinen sowie mit technischem Bedarf	38,43	28,75	35,67	– 25,19	– 19,40
46.71.1 Großhandel mit festen Brennstoffen	1,67	2,15	1,55	+ 28,74	+ 38,71
46.71.2 Großhandel mit Mineralölserzeugnissen	107,97	107,2	113,91	– 0,71	– 5,89
46.72.1 Großhandel mit Eisenerzen, Eisen, Stahl, Eisen- und Stahlhalbzeug	33,22	33,9	49,46	+ 2,05	– 31,46
46.72.2 Großhandel mit NE-Erzen, NE-Metallen und NE-Metallhalbzeug	20,38	19,66	25,37	– 3,53	– 22,51
46.73.1 Großhandel mit Holz, Baustoffen, Anstrichmitteln und Sanitärkeramik, ohne ausgeprägten Schwerpunkt	8,57	6,86	8,9	– 19,95	– 22,92
46.73.2 Großhandel mit Roh- und Schnittholz	4,22	4,75	5,0	+ 12,56	– 5,00
46.73.3 Großhandel mit sonstigen Holzhalbwaren sowie Bauelementen aus Holz	4,60	8,26	4,96	+ 79,57	+ 66,53
46.73.4 Großhandel mit Baustoffen und Bauelementen aus mineralischen Stoffen	23,93	30,25	25,18	+ 26,41	+ 20,14
46.73.5 Großhandel mit Flachglas	1,04	0,83	0,87	– 20,19	– 4,60
46.73.6 Großhandel mit Anstrichmitteln	3,28	3,12	3,5	– 4,88	– 10,86
46.73.7 Großhandel mit Sanitärkeramik	2,05	2,62	2,08	+ 27,80	+ 25,96
46.73.8 Großhandel mit Tapeten und Bodenbelägen (ohne Teppiche)	1,35	2,24	1,35	+ 65,93	+ 65,93
46.74.1 Großhandel mit Werkzeugen und Kleisenwaren	12,83	11,93	12,32	– 7,01	– 3,17
46.74.2 Großhandel mit Installationsbedarf für Gas, Wasser und Heizung	21,38	24,36	23,19	+ 13,94	+ 5,05
46.74.3 Großhandel mit Metall- und Kunststoffwaren für Bauzwecke	14,92	16,26	16,42	+ 8,98	– 0,97
46.75.0 Großhandel mit chemischen Erzeugnissen	31,34	26,7	32,45	– 14,81	– 17,72
46.76.0 Großhandel mit sonstigen Halbwaren	9,88	4,34	10,11	– 56,07	– 57,07
46.77.0 Großhandel mit Altmaterialien und Reststoffen	12,51	15,62	17,1	+ 24,86	– 8,65
46.90.1 Großhandel mit Rohstoffen, Halb- und Fertigwaren, ohne ausgeprägten Schwerpunkt	7,89	2,58	8,34	– 67,30	– 69,06
46.90.2 Großhandel mit Rohstoffen und Halbwaren, ohne ausgeprägten Schwerpunkt	5,75	4,34	6,39	– 24,52	– 32,08
46.90.3 Großhandel mit Fertigwaren, ohne ausgeprägten Schwerpunkt	52,33	39,6	50,61	– 24,33	– 21,75

Klassen der Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008).

Vorüberlegungen wie für die Außenhandelspreise. Zu unterscheiden sind reale Veränderungen und solche Veränderungen, die nur durch die Preisentwicklung verursacht wurden, sowie relative und absolute Veränderungen der Wägungsanteile. Die Bereinigung der Großhandelsumsätze um die Außenhandelsumsätze wurde erstmals für das Wägungsschema 2021 durchgeführt, sie führt zu einer stärkeren Aussagekraft der Großhandelspreisstatistik. Daher haben sich für viele der Positionen der WZ 2008 die Wägungsanteile 2021 gegen-

über 2015 erheblich geändert und die Strukturen der Wägungsschemata auf Basis 2015 und Basis 2021 sind nur eingeschränkt vergleichbar. So hat die WZ 46.69.3 „Großhandel mit sonstigen Ausrüstungen und Zubehör für Maschinen sowie mit technischem Bedarf“ einen Außenhandelsanteil von über 50%; die Bereinigung um die Außenhandelsumsätze in der Wirtschaftszweigunterklasse führt mit dazu, dass der Wägungsanteil im neuen Wägungsschema um 25 % gesunken ist. Dagegen sind in der WZ 46.73.4 „Großhandel mit Baustoffen und

Bauelementen aus mineralischen Stoffen“ nur 3,8% Außenhandelsumsatz enthalten. Verbunden mit großen Preissteigerungen zwischen 2015 und 2021 stieg der Wägungsanteil dieser Wirtschaftszweigunterklasse um 26% im neuen Wägungsschema.

Annähernd verdoppelt hat sich das Gewicht der WZ 46.39.9 „Großhandel mit sonstigen Nahrungs- und Genussmitteln, Getränken und Tabakwaren, ohne ausgeprägten Schwerpunkt“. Ebenso sind die Gewichte der WZ 46.73.3 „Großhandel mit sonstigen Holzhalbwaren sowie Bauelementen aus Holz“, zum Teil auch bedingt durch starke Preisanstiege zwischen 2015 und 2021, und der WZ 46.63.0 „Großhandel mit Bergwerks-, Bau- und Baustoffmaschinen“ sehr stark angestiegen. Sehr stark an Gewicht am Großhandel insgesamt verloren haben die WZ 46.90.1 „Großhandel mit Rohstoffen, Halb- und Fertigwaren, ohne ausgeprägten Schwerpunkt“ und die WZ 46.49.3 „Großhandel mit Lederwaren, Reisegepäck, Geschenk- und Werbeartikeln“.

### 3.2 Berichtsstellen- und Preisrepräsentantenstichprobe

---

Im letzten Schritt vor der Indexneuberechnung wurden die Stichproben der Berichtsstellen und der Preisrepräsentanten an das neue Wägungsschema angepasst. Bei der Auswahl der Berichtsstellen und der Preisrepräsentanten können, wie bei den Außenhandelspreisen, zwei unterschiedliche Konzepte verfolgt werden. Die Unternehmen können institutionell abgegrenzt werden und nur nach Preisen für Waren befragt werden, die auch zu ihrem wirtschaftlichen Schwerpunkt gehören, oder funktional abgegrenzt werden und nach den Preisen aller gehandelten Waren unabhängig vom wirtschaftlichen Schwerpunkt befragt werden. Der Vorteil der institutionellen Abgrenzung wäre, dass in der Erhebung überwiegend die Großhändler erfasst werden, die sich auf den Handel mit bestimmten Waren spezialisiert haben und entsprechende Bedeutung in diesem Markt besitzen. Von Nachteil wäre, dass Großhändler, die ein breites Produktportfolio anbieten und möglicherweise auch eine bedeutende Rolle im Großhandel mit Waren eines anderen Wirtschaftszweiges spielen, nicht in der Stichprobe für diese zweite Handelstätigkeit berücksichtigt würden. Um auch diese Großhändler bei der Stichprobe der Berichtsstellen adäquat zu berücksichtigen und die

Preisentwicklung möglichst vollständig für die relevanten Produktgruppen abzubilden, wurde bei der Auswahl der Berichtsstellen nicht der wirtschaftliche Schwerpunkt des Unternehmens als Auswahlkriterium herangezogen, sondern die tatsächlichen Handelstätigkeiten der Unternehmen.

Die Auswahl der Berichtsstellen, also der Firmen, die künftig ihre Verkaufspreise melden sollen, erfolgt durch eine gezielte Stichprobe. In einem ersten Arbeitsschritt wird die Zielstichprobe definiert und dabei die Anzahl der notwendigen Preisbeobachtungen für jede Handelstätigkeit und Produktgruppe festgelegt. Die Anzahl der notwendigen Preisbeobachtungen bestimmt sich nach dem Wägungsanteil und der Streuung der Veränderungsraten der bisherigen Preisentwicklungen der entsprechenden Positionen. Ein höherer Wägungsanteil wird grundsätzlich mit einer höheren Anzahl von Preisreihen hinterlegt, um der größeren Bedeutung der Handelstätigkeit im Gesamtindex gerecht zu werden. Allerdings kann eine relativ homogene Preisentwicklung auch mit weniger Preisreihen valide abgebildet werden. Daher wurde für Produkte mit bisher beobachteter homogener Preisentwicklung die Anzahl der Preisreihen reduziert.

Danach erfolgte der Abgleich der so gewonnenen Zielstichprobe mit der Stichprobe der Indexbasis 2015. Als Ergebnis wurde bestimmt, welche bisherigen Preiserhebungen weitergeführt, welche Neuaufnahmen gewonnen werden mussten und auf welche bisherigen Preiserhebungen verzichtet werden kann. Auf dieser Basis wurden die neu einzubeziehenden Unternehmen herangezogen und bisher berichtspflichtige Unternehmen gegebenenfalls aus der Berichtspflicht entlassen.

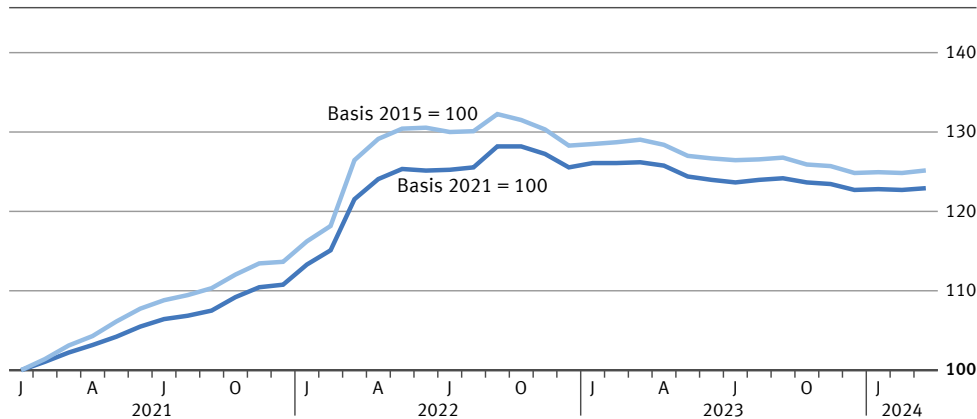
### 3.3 Ergebnisse der Indexneuberechnung

---

Mit den neuen Berechnungsgrundlagen (Warenkorb, Wägungsschema, Berichtsstellen- und Preisrepräsentantenstichprobe) erfolgte eine Neuberechnung aller Indexergebnisse ab Januar 2021. Wie bei den Außenhandelspreisen können unterschiedliche Gründe für Abweichungen der Indexverläufe auf neuer und alter Basis vorliegen. Eine genaue Identifikation nur einer Ursache für abweichende Entwicklungen ist kaum möglich, da in der Regel verschiedene Ursachen gleichzeitig auftreten. Bei den Großhandelsverkaufspreisen wird der Vergleichszeitraum Januar 2021 bis März 2024 betrach-

**Grafik 2**

Index der Großhandelsverkaufspreise auf Basis 2015 und Basis 2021 im Vergleich  
Januar 2021 = 100



tet. Die Veränderungsrate für diesen Zeitraum betrug für den Index auf alter Basis 25,2%, für den auf neuer Basis 22,9%. Für den grafischen Vergleich wurden in [Grafik 2](#) die Indizes jeweils auf Januar 2021 = 100 normiert. Es zeigt sich, dass die Indizes sehr ähnliche Verläufe haben, aber der Index auf neuer Basis über den gesamten Zeitraum unter dem auf alter Basis liegt und insgesamt weniger stark ansteigt. Der größte Abstand zwischen den Indizes wird im Juni 2022 mit 5,4 Prozentpunkten erreicht. Danach nähern sich die Indizes wieder leicht an, aber auch im März 2024 betrug der Abstand noch 2,2 Prozentpunkte. Neben den veränderten Produkt- und Berichtsstellenstichproben sind die Veränderungen des Wägungsschemas der Hauptgrund für den Unterschied im Verlauf. So wirken sich der Rückgang der Wägungsanteile insbesondere der WZ 46.72.2 „Großhandel mit NE-Erzen, NE-Metallen und NE-Metallhalbzug“ und der WZ 46.21.0 „Großhandel mit Getreide, Rohtabak, Saatgut und Futtermitteln“ deutlich auf den neuen Indexverlauf aus.

Werden die einzelnen Wirtschaftszweigklassen in [Tabelle 7](#) analysiert, zeigt sich ein uneinheitliches Bild der Abweichungen. Die größte positive Abweichung (der Index auf Basis 2021 liegt über dem Index auf Basis 2015) weist die WZ 46.37.0 „Großhandel mit Kaffee, Tee, Kakao und Gewürzen“ mit 12,6 Prozentpunkten auf. Die größten negativen Abweichungen gab es bei der WZ 46.23.0 „Großhandel mit lebenden Tieren“ mit – 16,1 Prozentpunkten und bei der WZ 46.31.0 „Großhandel mit Obst, Gemüse und Kartoffeln“ mit – 11,8 Prozentpunkten. Auch für die Großhandelsverkaufspreise gilt, wie bei den Außenhandelspreisen, dass einzelne Gründe für die Abweichungen nur schwer identifiziert werden können. Unterschiedliche Ursachen, der Austausch von Warenkorbpositionen, die neue Gewichtung und der Tausch von Berichtsstellen, die zu den unterschiedlichen Verläufen auf neuer und alter Basis führen, treten meist in Kombination auf.

**Tabelle 7**

**Veränderungsraten der Großhandelsverkaufspreise im Vergleich**

	Veränderungsrate Januar 2015 zu Juli 2018 in %		Differenz in Prozent- punkten
	Basis 2015	Basis 2021	
Gesamtindex	+ 25,2	+ 22,9	– 2,3
46.21 Großhandel mit Getreide, Rohtabak, Saatgut und Futtermitteln	+ 4,7	+ 6,9	+ 2,2
46.22 Großhandel mit Blumen und Pflanzen	+ 16,4	+ 11,7	– 4,7
46.23 Großhandel mit lebenden Tieren	+ 75,9	+ 59,8	– 16,1
46.31 Großhandel mit Obst, Gemüse und Kartoffeln	+ 40,3	+ 28,5	– 11,8
46.32 Großhandel mit Fleisch und Fleischwaren	+ 32,1	+ 33,8	+ 1,7
46.33 Großhandel mit Milch, Milcherzeugnissen, Eiern, Speiseölen und Nahrungsfetten	+ 40,1	+ 33,4	– 6,7
46.34 Großhandel mit Getränken	+ 17,6	+ 18,0	+ 0,4
46.35 Großhandel mit Tabakwaren	+ 17,6	+ 14,4	– 3,2
46.36 Großhandel mit Zucker, Süßwaren und Backwaren	+ 34,1	+ 26,0	– 8,1
46.37 Großhandel mit Kaffee, Tee, Kakao und Gewürzen	+ 29,7	+ 42,3	12,6
46.38 Großhandel mit sonstigen Nahrungs- und Genussmitteln	+ 29,2	+ 28,4	– 0,8
46.39 Großhandel mit Nahrungs- und Genussmitteln, Getränken und Tabakwaren, ohne ausgeprägten Schwerpunkt	+ 30,2	+ 26,9	– 3,3
46.41 Großhandel mit Textilien	+ 20,1	+ 14,7	– 5,4
46.42 Großhandel mit Bekleidung und Schuhen	+ 10,2	+ 10,6	+ 0,4
46.43 Großhandel mit Foto- und optischen Erzeugnissen, elektrischen Haushaltsgeräten und Geräten der Unterhaltungselektronik	+ 12,1	+ 11,7	– 0,4
46.44 Großhandel mit keramischen Erzeugnissen, Glaswaren und Reinigungsmitteln	+ 25,2	+ 19,0	– 6,2
46.45 Großhandel mit kosmetischen Erzeugnissen und Körperpflegemitteln	+ 13,2	+ 11,8	– 1,4
46.46 Großhandel mit pharmazeutischen, medizinischen und orthopädischen Erzeugnissen	+ 9,9	+ 7,9	– 2,0
46.47 Großhandel mit Möbeln, Teppichen, Lampen und Leuchten	+ 27,7	+ 26,7	– 1,0
46.48 Großhandel mit Uhren und Schmuck	+ 19,3	+ 25,6	+ 6,3
46.49 Großhandel mit sonstigen Gebrauchs- und Verbrauchsgütern	+ 24,1	+ 24,7	+ 0,6
46.51 Großhandel mit Datenverarbeitungsgeräten, peripheren Geräten und Software	+ 3,3	+ 0,8	– 2,5
46.52 Großhandel mit elektronischen Bauteilen und Telekommunikationsgeräten	+ 12,3	+ 3,1	– 9,2
46.61 Großhandel mit landwirtschaftlichen Maschinen und Geräten	+ 23,2	+ 19,9	– 3,3
46.62 Großhandel mit Werkzeugmaschinen	+ 17,0	+ 17,1	+ 0,1
46.63 Großhandel mit Bergwerks-, Bau- und Baustoffmaschinen	+ 21,8	+ 20,7	– 1,1
46.65 Großhandel mit Büromöbeln	+ 24,5	+ 21,3	– 3,2
46.66 Großhandel mit sonstigen Büromaschinen und -einrichtungen	+ 10,6	+ 14,3	+ 3,7
46.69 Großhandel mit sonstigen Maschinen und Ausrüstungen	+ 21,9	+ 21,1	– 0,8
46.71 Großhandel mit festen Brennstoffen und Mineralöl-erzeugnissen	+ 49,8	+ 49,6	– 0,2
46.72 Großhandel mit Erzen, Metallen und Metallhalbzeug	+ 27,8	+ 26,4	– 1,4
46.73 Großhandel mit Holz, Baustoffen, Anstrichmitteln und Sanitärkeramik	+ 32,0	+ 32,7	+ 0,7
46.74 Großhandel mit Metall- und Kunststoffwaren für Bauzwecke sowie Installationsbedarf für Gas, Wasser und Heizung	+ 33,0	+ 28,6	– 4,4
46.75 Großhandel mit chemischen Erzeugnissen	+ 31,2	+ 33,5	+ 2,3
46.76 Großhandel mit sonstigen Halbwaren	+ 28,1	+ 22,2	– 5,9
46.77 Großhandel mit Altmaterialien und Reststoffen	+ 10,2	+ 9,7	– 0,5


Klassen der Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008).

### 4

---

#### Ausblick

---

Derzeit wird daran gearbeitet, sowohl die europäischen als auch die nationalen Wirtschaftszweig- und Güterklassifikationen zu überarbeiten. Diese sollen schrittweise ab 2025 in den Unternehmensstatistiken umgesetzt werden. Sowohl für den Index der Großhandelsverkaufspreise als auch für die Indizes der Außenhandelspreise ist die Implementierung der neuen Klassifikationen für die nächste Indexrevision geplant. Die EBS-Verordnung sieht für die Außenhandelspreise als nächstes Basisjahr 2025 vor, die Revision auf dieses Basisjahr wird vermutlich Anfang 2028 durchgeführt werden. Auch der Großhandelsverkaufspreisindex wird als nächstes Basisjahr 2025 ausweisen. Dies verkürzt die jetzige Basisperiode auf vier Jahre. Nach 2025 ist vorgesehen, wieder den bisherigen Fünf-Jahresrhythmus aufzunehmen. 

## LITERATURVERZEICHNIS

---

Kruse, Hendrik W./Meyerhoff, Annette/Erbe, Anette. [Neue Methoden zur Mikrodaten-verknüpfung von Außenhandels- und Unternehmensstatistiken](#). In: WISTA Wirtschaft und Statistik. Ausgabe 5/2021, Seite 53 ff.

Statistisches Bundesamt. *Statistischer Bericht Außenhandel Dezember 2021*. 2022a. [Zugriff am 17. Januar 2025]. Verfügbar unter: [www.statistischebibliothek.de](http://www.statistischebibliothek.de)

Statistisches Bundesamt. *Qualitätsbericht – Strukturstatistik im Handels- und Dienstleistungsbereich*. 2022b. [Zugriff am 20. Januar 2025]. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

Peter, Florian. [Index der Erzeugerpreise gewerblicher Produkte \(Inlandsabsatz\) auf Basis 2021](#). In: WISTA Wirtschaft und Statistik. Ausgabe 5/2024, Seite 37 ff.

Peter, Florian. [Index der Erzeugerpreise gewerblicher Produkte \(Inlandsabsatz\) auf Basis 2015](#). In: WISTA Wirtschaft und Statistik. Ausgabe 2/2019, Seite 46 ff.

Neubauer, Werner. *Preisstatistik*. München 1996.

## RECHTSGRUNDLAGEN

---

Durchführungsverordnung (EU) 2020/1197 der Kommission vom 30. Juli 2020 zur Festlegung technischer Spezifikationen und Einzelheiten nach der Verordnung (EU) 2019/2152 des Europäischen Parlaments und des Rates über europäische Unternehmensstatistiken, zur Aufhebung von zehn Rechtsakten im Bereich Unternehmensstatistiken (Amtsblatt der EU Nr. L 271, Seite 1).

Verordnung (EU) 2019/2152 des Europäischen Parlaments und des Rates vom 27. November 2019 über europäische Unternehmensstatistiken, zur Aufhebung von zehn Rechtsakten im Bereich Unternehmensstatistiken (Amtsblatt der EU Nr. L 327, Seite 1).



# HIERARCHISCHES KLASSIFIZIEREN VON SCANNERDATEN: EIN METHODEN- VERGLEICH MIT ANWENDUNG IN DER VERBRAUCHERPREISSTATISTIK

Daniel Nietzer, Karola Henn, Chris-Gabriel Islam, Keno Krewer,  
Carsten Monzert

➤ **Schlüsselwörter:** Maschinelles Lernen – Hierarchisches Lernen –  
Verbraucherpreisindex – Klassifizierung – COICOP

## ZUSAMMENFASSUNG

Die Nutzung von Scannerdaten in der Verbraucherpreisstatistik hat viele Vorteile, bringt allerdings auch einige Herausforderungen mit sich. Eine der Herausforderungen ist die Klassifizierung der Artikel nach dem vom Verbraucherpreisindex verwendeten Klassifikationssystem COICOP. Aufgrund der großen Datenmenge bei Scannerdaten ist eine automatisierte Klassifizierung mittels maschinellen Lernens zwingend, um die Daten effizient in der amtlichen Statistik verwenden zu können. Dieser Beitrag analysiert einen Ansatz aus der aktuellen Forschung, der die zugrunde liegende hierarchische Struktur der COICOP-Klassifikation bei der automatisierten Klassifizierung berücksichtigt. Außerdem werden die Auswirkungen von Fehlklassifizierungen auf den Verbraucherpreisindex untersucht.

➤ **Keywords:** machine learning – hierarchical learning – consumer price index – classification – COICOP

## ABSTRACT

*The use of scanner data in consumer price statistics offers many benefits but also presents certain challenges. One such challenge is the classification of articles according to the COICOP system, which is used for the consumer price index (CPI). Due to the sheer volume of scanner data, automated classification using machine learning is essential to be able to use the data efficiently in official statistics. This article examines a current research approach that incorporates the underlying hierarchical structure of the COICOP classification during the automated classification of articles. The impact of misclassification on the CPI is also analysed.*

### Daniel Nietzer

hat Volkswirtschaftslehre an den Universitäten Göttingen und Mannheim studiert. Er schrieb seine diesem Beitrag zugrunde liegende Masterarbeit im Referat „Methoden und Kommunikation in der Preisstatistik“ des Statistischen Bundesamtes.

### Karola Henn

ist Diplom-Finanz- und Wirtschaftsmathematikerin und Referentin im Referat „Methoden und Kommunikation in der Preisstatistik“ des Statistischen Bundesamtes. Hauptsächlich beschäftigt sie sich mit der Einführung von Scannerdaten in den Verbraucherpreisindex, ein Schwerpunkt ist dabei die Klassifizierung von Scannerdaten.

### Dr. Chris-Gabriel Islam

hat Wirtschaftsmathematik und Lehramt in Mathematik und Geschichte in Darmstadt und Mexiko-Stadt studiert sowie in Göttingen und Hasselt (Belgien) promoviert. Er hat bis 2024 das Projekt „Scannerdaten in der Verbraucherpreisstatistik“ geleitet.

### Dr. Keno Krewer

hat Physik in Mainz studiert und am Max-Planck-Institut für Polymerforschung promoviert. Er verantwortete die Methodik des Projekts „Scannerdaten in der Verbraucherpreisstatistik“ im Statistischen Bundesamt.

### Carsten Monzert

hat Volkswirtschaftslehre in Frankfurt am Main und Mainz studiert. Er ist Data Engineer im Referat „Datenlabor“ des Statistischen Bundesamtes.

### Danksagung

Die diesem Artikel zugrunde liegende Masterarbeit ist am Lehrstuhl für Ökonometrie der Georg-August-Universität Göttingen entstanden. In diesem Zusammenhang bedanken sich Autorin und Autoren herzlich bei Prof. Dr. Helmut Herwartz für die Kooperation.

## 1

### Einleitung

Um den Verbraucherpreisindex (VPI) in Deutschland zu ermitteln, erheben vorwiegend die Statistischen Ämter der Länder dezentral die Preise vor Ort. Neue Datenquellen und technische Weiterentwicklungen ermöglichen es, in Digitalisierungsprojekten die Datenerhebung angesichts dynamischer und individualisierter Preissetzungen weiterzuentwickeln. So werden beispielsweise seit mehreren Jahren Preise automatisiert per Web Scraping im Internet erhoben (Blaudow/Ostermann, 2020). Auch Scannerdaten als digital vorliegende Datenquelle bieten sich an, um die Datenerhebung weiterzuentwickeln. Scannerdaten können die Preismessung zeitlich und räumlich sowie in Bezug auf den Sortimentsumfang ausweiten. Ziel ist es, die Genauigkeit und Qualität des Verbraucherpreisindex zu verbessern, indem die Preiserhebung erweitert wird und sie damit die Preisentwicklung im Einzelhandel repräsentativer darstellt (Bieg, 2019).

Bei Scannerdaten handelt es sich um Scannerkassendaten, die zu jedem verkauften Artikel Absatz, Umsatz, zugehörige Artikelinformationen und Regionalangaben für einen bestimmten Zeitraum enthalten. Um sie in der laufenden Statistikproduktion zu nutzen, fordert das Statistische Bundesamt Datenlieferungen von Einzelhandelsunternehmen an, die dann größtenteils von einer zentralen Stelle der Handelskette geliefert werden. Für die Anforderung dieser Transaktionsdaten wurde zunächst das Preisstatistikgesetz geändert. Dabei wurde insbesondere § 7b Absatz 3 Preisstatistikgesetz, der die Anforderungen von Transaktionsdaten regelt, aufgenommen. Außerdem wurde über § 3 Absatz 1a Preisstatistikgesetz ausdrücklich die Übermittlung von Umsatz- und Absatzinformationen berücksichtigt. Eine Besonderheit unter anderem bei den Scannerdaten ist auch die Regelung in § 7c Preisstatistikgesetz, die eine statistische Auswertung auch zu anderen Zwecken als der Preisstatistik ermöglicht. So werden Scannerdaten bereits für [experimentelle Konjunkturstatistiken](#) verwendet.

Scannerdaten werden in einer eigens entwickelten Fachanwendung aufbereitet, dabei werden Angaben plausibilisiert, imputiert und klassifiziert, um anschließend Preisindizes für die entsprechenden Güter zu berech-

nen. Da der Verbraucherpreisindex nach der COICOP<sup>1</sup> klassifiziert wird, müssen auch Scannerdaten nach der COICOP gegliedert ausgewertet werden. Dieser Schritt der Klassifizierung erfolgt aktuell in einem mehrstufigen Verfahren, das unter anderem Maschinelles Lernen beinhaltet. Auf diesen Schritt baut die in diesem Beitrag betrachtete Machbarkeitsstudie auf und versucht, das aktuelle Klassifizierungsverfahren durch Verfahren des hierarchischen Klassifizierens zu verbessern. Auch das hierarchische Klassifizieren hat zum Ziel, die Genauigkeit und Qualität des Verbraucherpreisindex zu verbessern, indem die Klassifizierung weiter verbessert wird.

Die Integration von Scannerdaten in den Verbraucherpreisindex führt dazu, dass deutlich mehr Artikel in den Index einbezogen werden können, da Scannerdaten das komplette Sortiment der verkauften Güter abbilden. Auch dehnen Scannerdatenlieferungen den möglichen Erfassungszeitraum erheblich aus und bilden damit die Preisänderungen umfassender und somit realistischer ab. Während die bisherigen Vor-Ort-Erhebungen die Preise in der Regel nur an einem Tag im Monat erfassen, ist bei Scannerdaten geplant, für die Indexberechnung die Daten der ersten beiden vollen Wochen eines Monats zu nutzen. Damit können beispielsweise Sonderangebote, die nur in einer Woche oder nur an bestimmten Wochentagen gelten, besser berücksichtigt werden.<sup>2</sup>

Um Scannerdaten im Verbraucherpreisindex produktiv nutzen zu können, müssen die Daten entsprechend dem Klassifikationssystem der Verbraucherpreisstatistik klassifiziert werden. Eine korrekte Klassifizierung ist notwendig, um den Verbraucherpreisindex möglichst verzerrungsfrei abzubilden. Das aktuelle Vorgehen bei den Scannerdaten sieht vor, dass für die Klassifizierung der Scannerdaten ein flaches maschinelles Lernverfahren angewendet wird, das die Hierarchie der COICOP nicht nutzt. In der im Folgenden beschriebenen Machbarkeitsstudie wurde versucht, die Hierarchie der Klassifikation

1 Die Klassifikation der Verwendungszwecke des Individualverbrauchs (Classification of Individual Consumption by Purpose – COICOP) ist eine Klassifikation der Statistikabteilung der Vereinten Nationen zur Klassifizierung und Analyse der Konsumausgaben der privaten Haushalte, privaten Organisationen ohne Erwerbszweck und des Staates nach dem Verwendungszweck, siehe [ec.europa.eu](http://ec.europa.eu)

2 Auch bei den Scannerdaten kann es vorkommen, dass Sonderangebote, die nur für einen kurzen Zeitraum gelten, in die Berechnung für diesen Zeitraum einfließen. Der Effekt ist aber immer geringer als bei einer Vor-Ort-Erhebung, bei der im Zweifelsfall ein Sonderangebot, das nur einen Tag gilt, den kompletten Erfassungszeitraum repräsentiert. Des Weiteren wäre es denkbar, in Zukunft die ersten drei oder vier vollen Wochen in den Preisindex einfließen zu lassen.

zu nutzen, um bei maschinellen Klassifizierungsschritten bessere Ergebnisse zu erzielen. Kapitel 2 stellt die Theorie und Methodik eines hierarchischen Klassifizierungsansatzes sowie die Ergebnisse der Machbarkeitsstudie vor. Kapitel 3 untersucht Abweichungen, die sich bei Anwendung verschiedener Klassifizierungsansätze im Vergleich zu einem Referenzindex ergeben. Abschließend folgen ein Fazit sowie eine Empfehlung für mögliche weitere Entwicklungsarbeiten.

## 2

### Hierarchisches Klassifizieren als neuer Ansatz

#### 2.1 Einführung in Maschinelles Lernen und Klassifizieren

Die COICOP-Klassifikation ist ein hierarchisches Klassifikationssystem, das bis zur vierten Ebene (COICOP-5-Steller) europäisch harmonisiert ist. Für den nationalen Verbraucherpreisindex werden die einzelnen Positionen tiefer gegliedert und erhalten einen zehnstelligen Code (10-Steller). Um Preisindizes aus Scannerdaten zu erstellen, wird jedem gelieferten Artikel<sup>3</sup> eine entsprechende Klasse auf dieser Ebene zugeordnet. Klassifizierung von Daten ist ein typischer Anwendungsfall von Maschinellem Lernen, der auch bei Scannerdaten effiziente Ergebnisse liefert.<sup>4</sup>

„Machine Learning oder maschinelles Lernen ist ein Teilbereich der künstlichen Intelligenz. Der Schwerpunkt liegt dabei auf dem Trainieren von Computern, um aus Daten und Erfahrungen zu lernen und sich stets zu verbessern – anstatt explizit dafür programmiert zu werden.“ (SAP Deutschland, 2025) Zu unterscheiden sind zwei Teilbereiche des maschinellen Lernens: das unüberwachte Lernen und das überwachte Lernen.

Beim **unüberwachten Lernen** versucht die Maschine, in den Daten selbstständig Muster und Zusammenhänge zu erkennen. Hierbei gibt es keine Zielvorgabe, es handelt sich also um einen explorativen Ansatz.

Klassifizierungsaufgaben hingegen sind ein Beispiel für **überwachtes Lernen**, bei dem ein Algorithmus auf Grundlage bereits manuell klassifizierter Daten (Trainingsdaten) Muster erkennt, um neue Artikel in die richtige Kategorie beziehungsweise Klasse einzusortieren.

Bei dieser Art des Lernens werden Kenntnisse aus einem bekannten Datensatz genutzt, um für einen anderen unbekannten Datensatz Vorhersagen zu treffen. Die Maschine lernt Zusammenhänge und Muster der Daten und wendet diese anschließend auf den unbekannten Datensatz an. Für das Trainieren eines solchen Algorithmus wird ein Datensatz von Klassifizierungsbeispielen benötigt, welcher neben den Variablen, die für das Klassifizieren genutzt werden, die korrekten Klassen enthält. Dieser Datensatz wird in zwei Teile unterteilt: einen Trainings- und einen Testdatensatz. Die Teilung wird vorgenommen um zu verhindern, dass der Klassifikator im Training den Datensatz auswendig lernt (overfitting), aber im Echtbetrieb bei der Klassifizierung von neuen Daten schlechtere Ergebnisse liefert. Deswegen werden Klassifikatoren im Bereich des überwachten Lernens mit Trainingsdaten trainiert und anschließend mithilfe der Testdaten evaluiert. Dieses trainierte Modell wird als Klassifikator bezeichnet. Beispiele für Methoden des überwachten Lernens sind neuronale Netze, Support Vector Machines und das in dieser Untersuchung angewendete Verfahren Random Forest (James und andere, 2013).

Das hierarchische Klassifizieren ist eine Form des überwachten Lernens und hat in den letzten Jahrzehnten an Bedeutung gewonnen. Beim hierarchischen Klassifizieren wird die vorhandene Hierarchie des Klassifikationssystems genutzt, das heißt der Klassifikator klassifiziert sukzessive auf jeder Ebene beginnend von der obersten.

Bisher sind die meisten der heutzutage verwendeten Klassifikatoren flache Klassifikatoren, die die hierarchische Struktur des Klassifikationssystems nicht berücksichtigen und direkt Vorhersagen auf der untersten Ebene treffen. Vorbereitend auf die Einführung in den Verbraucherpreisindex werden flache Klassifikatoren genutzt, hierbei werden die beobachteten Artikel direkt einem 10-Steller zugeordnet, während beim hierar-

<sup>3</sup> Ein Datensatz enthält alle Verkäufe eines Artikels einer Filiale in einer Kalenderwoche. Das heißt, Artikel sind mehrfach in Datenlieferungen vorhanden, da die Artikel in unterschiedlichen Filialen verkauft werden. Außerdem sind die Artikel auch in Datenlieferungen weiterer Wochen vorhanden. Jeder Artikel muss allerdings nur einmal klassifiziert werden.

<sup>4</sup> Zu weiteren Aspekten von Maschinellem Lernen in der amtlichen Statistik siehe Dumpert (2024).

chischen Klassifizieren die Artikel schrittweise immer feineren Klassen zugeordnet werden. Die Zuordnung auf der 10-Steller-Ebene erfolgt erst im letzten Schritt. In der COICOP-Klassifikation bedeutet dies, dass beim hierarchischen Klassifizieren ein Granny-Smith-Äpfel nicht direkt dem 10-Steller 0116130200 „Äpfel“ zugeordnet wird, sondern zuerst dem COICOP-2-Steller 01 „Nahrungsmittel und alkoholfreie Getränke“, dann dem COICOP-3-Steller 011 „Nahrungsmittel“. Die Zuordnung wird über die 4-, 5- und 7-Steller fortgesetzt, bis der Granny-Smith-Äpfel im letzten Schritt dem 10-Steller 0116130200 „Äpfel“ zugeordnet wird. [↗ Grafik 1](#)

Hierarchische Klassifikatoren besitzen keinen grundsätzlichen Vorteil gegenüber flachen Klassifikatoren. Welches Verfahren besser geeignet ist, hängt von dem genutzten Klassifikator und Klassifikationssystem ab und ist für jeden Anwendungsfall zu untersuchen. Grundsätzlich kann der Ansatz des hierarchischen Klassifizierens eine Klassifizierung in zwei Aspekten verbessern: Zum einen könnten mehr Artikel zur korrekten Klasse zugeordnet werden. Zum anderen führen selbst falsche Zuordnungen zu geringeren Fehlern auf oberen Ebenen,

falls zumindest eine Oberklasse<sup>5</sup> korrekt zugeordnet ist. Bei der Preisstatistik entspricht die oberste Ebene dem Verbraucherpreisindex. In dem in Grafik 2 dargestellten Fall könnte zwar der Teilindex der untersten Klasse durch eine Fehlklassifizierung verzerrt sein, die Verzerrung des Teilindex der gemeinsamen Oberklasse könnte aber reduziert werden. Von einer Fehlklassifizierung eines Artikels sind die in [↗ Übersicht 1](#) genannten Indizes betroffen.

### Übersicht 1

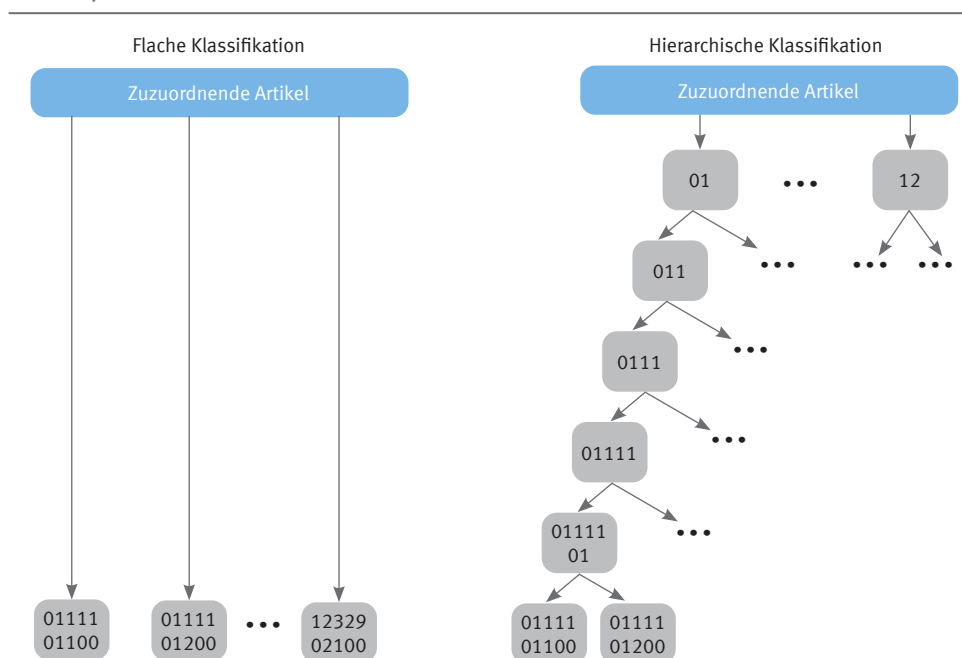
#### Relevante Indizes bei einer Fehlklassifizierung

Nummer	Bezeichnung	Erklärung
(1)	Spender-Index	Der Index der Klasse (10-Steller), in die der Artikel tatsächlich gehört
(2)	Empfänger-Index	Der Index der Klasse (10-Steller), in die der Artikel zugeordnet wurde
(3)	Spender-Aggregate	Indizes von Oberklassen, zu denen nur der Spender-Index beiträgt
(4)	Empfänger-Aggregate	Indizes von Oberklassen, zu denen nur der Empfänger-Index beiträgt
(5)	Gemeinsame Aggregate	Indizes von Oberklassen, die sowohl Spender als auch Empfänger enthalten

- 5 Im hierarchischen Klassifizieren bezeichnet der Begriff Oberklasse eine übergeordnete Klasse und der Begriff Unterklasse eine untergeordnete Klasse: Äpfel ist eine Unterklasse der Oberklasse Früchte.

### Grafik 1

#### Vergleichende Darstellung einer flachen und einer hierarchischen Klassifikation am Beispiel der COICOP<sup>1</sup>



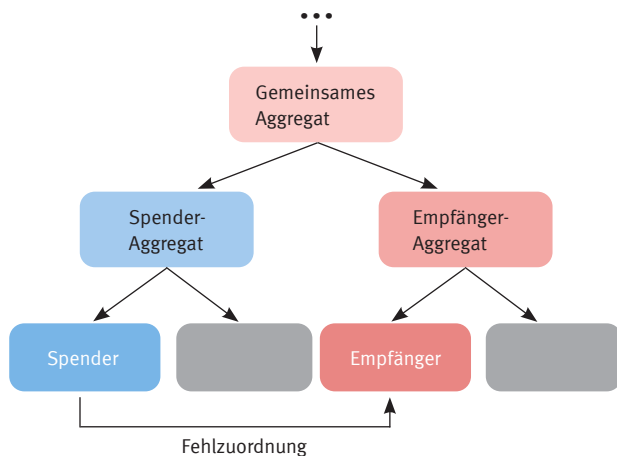
1 Classification of Individual Consumption by Purpose – Klassifikation der Verwendungszwecke des Individualverbrauchs.

Für die Indizes (1) und (2) der einzelnen Klassen gilt, dass nah verwandte Klassen häufig ähnliche Preisentwicklungen haben. Wenn sich die Preisreihen aus der Spender-Klasse ähnlich wie die der Empfänger-Klasse verhalten, ist die Auswirkung der Fehlzurordnung gering. Dies ist der erste Grund, warum Fehlzurordnungen mit geringen Abständen zwischen dem korrekten 10-Steller und dem zugeordneten 10-Steller nur wenig Einfluss auf die Qualität der Indizes haben.

Der zweite Grund liegt in der Aggregation: Je höher aggregiert wird, desto geringer sind die Auswirkungen der Fehlzurordnung einzelner Artikel, sofern die Zurordnung der Oberklassen korrekt erfolgt ist. Die Fehler im Spender- beziehungsweise im Empfänger-Aggregat erben ihr Verhalten von den Einzelindizes, im gemeinsamen Aggregat heben sich die Auswirkungen der Fehlzurordnung teilweise auf. ➤ Grafik 2

**Grafik 2**

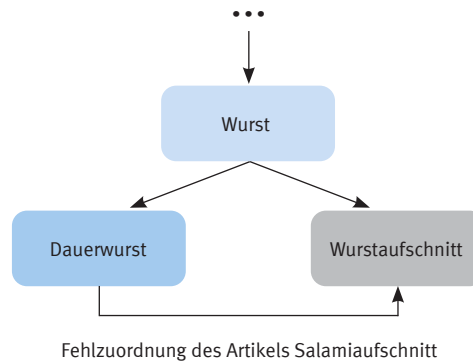
Auswirkung einer Fehlzurordnung eines Artikels aus dem Spender- zum Empfänger-10-Steller



➤ Grafik 3 verdeutlicht dies mit einem Beispiel: Der Artikel „Salamiaufschnitt“ ist eigentlich eine Dauerwurst, wird aber fälschlicherweise dem Wurstaufschnitt zugeordnet. Eine Preisänderung des „Salamiaufschnitts“ wird jetzt dem Wurstaufschnitt zugerechnet, im Teilindex für Dauerwurst fehlt sie. Im gemeinsamen Aggregat „Wurst“ ist die Änderung wieder enthalten. Lediglich die Gewichtung aus dem neuen Teilindex, mit der die Preisänderung des Salamiaufschnitts eingeht, könnte verfälscht sein. Der Salamiaufschnitt geht fälschlicherweise mit dem Gewicht der Klasse Wurstaufschnitt multipliziert

**Grafik 3**

Beispielhafte Fehlzurordnung des Salamiaufschnitts



mit dem Anteil des Salamiaufschnitts an der Klasse des Wurstaufschnitts ein. Die Gewichte der Klassen sowohl in der Oberklasse als auch in der Unterklasse basieren innerhalb der Scannerdatenindizes auf Umsatzanteilen von Dauerwurst beziehungsweise Wurstaufschnitt. Ein Artikel mit gegebenem Umsatz wird also einen kleineren Umsatzanteil (kleines Gewicht) an einer umsatzstarken (schwergewichtigen) Klasse oder einen größeren Anteil (großes Gewicht) an einer umsatzschwachen (leichtgewichtigen) Klasse aufweisen. Daher ist das Gewicht des Artikels im gemeinsamen Aggregat auch bei Fehlzurordnung ähnlich. Im gemeinsamen Aggregat ist also der Effekt der Fehlzurordnung wieder teilweise korrigiert. Je näher die falsche Klasse mit der korrekten Klasse verwandt ist, desto früher wird ein gemeinsames Aggregat gebildet und werden die Auswirkungen der Fehlzurordnung reduziert. Daher führt eine nähere Verwandtschaft zwischen korrekten und falsch zugeordneten Klassen zu einer geringeren Verzerrung in den Indizes höherer Aggregate.

## 2.2 Typisierung hierarchischer Klassifikatoren

Bei hierarchischen Klassifikatoren werden mit lokalen und globalen Klassifikatoren zwei Ansätze unterschieden, hierarchische Strukturen zu berücksichtigen. Beim globalen Klassifikator (auch Big-Bang-Klassifikator genannt) wird ein einziger Klassifikator gebaut und trainiert, welcher die hierarchische Struktur der Daten berücksichtigt (Silla/Freitas, 2011). Bei lokalen Klassifikatoren werden auf unterschiedlichen Ebenen der Hierarchie Klassifikatoren trainiert, die jeweils den Output



aus den vorherigen Klassifikatoren übernehmen, um Unterklassen vorherzusagen. Es gibt drei verschiedene Arten von lokalen Klassifikatoren:

- › lokale Klassifikatoren je Klasse (LCN),
- › lokale Klassifikatoren je Oberklasse (LCPN) und
- › lokale Klassifikatoren je Ebene (LCL).

Bei einem LCN-Ansatz wird je Klasse ein binärer Klassifikator trainiert; in Teil (b) von [Grafik 4](#) sind jeweils die Einheiten markiert, für die eigene Klassifikatoren trainiert werden. Bei einem LCPN-Ansatz wird je Oberklasse ein Klassifikator gebildet, der zwischen den zugehörigen Unterklassen entscheidet, das ist in Teil (c) von Grafik 4 dargestellt. Dass für den LCL-Ansatz je Ebene ein Klassifikator trainiert wird, der zwischen allen Unterklassen entscheidet, zeigt Teil (d) von Grafik 4. Um den Gegensatz zu flachen Klassifikatoren zu verdeutlichen, ist dieser in Teil (a) von Grafik 4 abgebildet.

Globale und lokale Ansätze haben jeweils verschiedene Vor- und Nachteile. Die lokalen Klassifikatoren sind leicht umsetzbar, da sie auf bekannten flachen Klassifikatoren basieren. Ein Nachteil ist, dass Fehler vererbt werden können. Falls ein Artikel wie „Salami-aufschnitt“ auf der ersten Ebene der Hierarchie statt in den richtigen COICOP-2-Steller 01 „Nahrungsmittel und alkoholfreie Getränke“ in den COICOP-2-Steller 12 „Andere Waren und Dienstleistungen“ eingeordnet

wird, kann der Artikel nicht mehr dem korrekten 10-Steller zugeordnet werden.<sup>16</sup> Für den globalen Klassifikator spricht, dass nur ein einziges Modell gebildet und trainiert wird. Dieses Modell ist allerdings relativ komplex und anwendungsspezifisch. Da die globalen Klassifikatoren so komplex sind, können traditionelle Evaluationsmetriken nicht genutzt werden, was einen Vergleich mit flachen Klassifikatoren erschwert. Im Weiteren liegt der Fokus auf den lokalen Klassifikatoren, insbesondere auf LCPN und LCN [Teile (c) und (b) der Grafik 4], die auch in der Literatur am häufigsten verwendet werden (Silla/Freitas, 2011).

## 2.3 Aktuelles Vorgehen

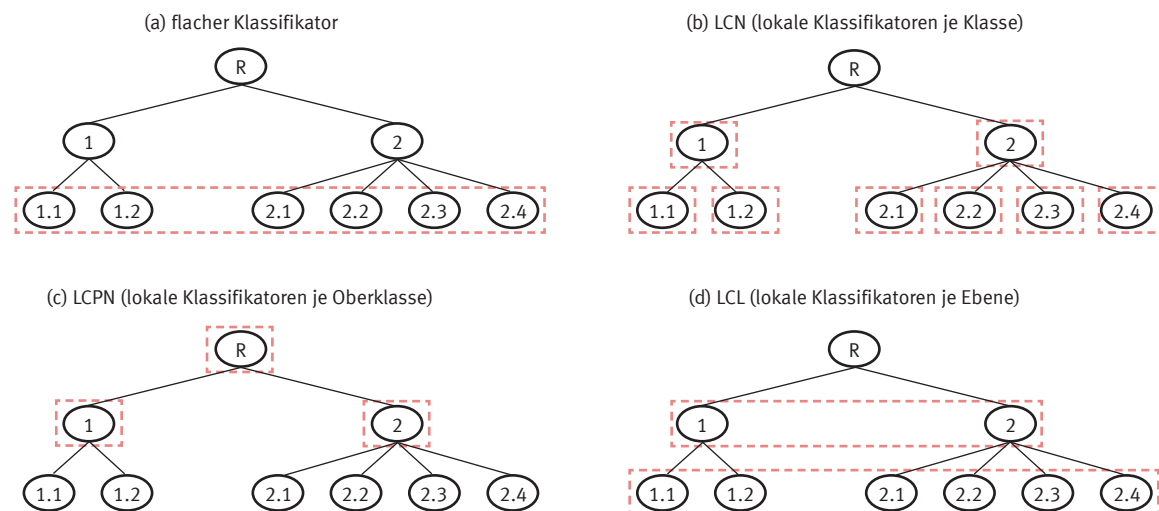
Das Projekt „Scannerdaten in der Verbraucherpreisstatistik“ bereitet die Einführung von Scannerdaten in den Verbraucherpreisindex zur nächsten Indexrevision vor. Hierfür ist geplant, die Scannerdaten der 23 umsatzstärksten Handelsketten<sup>17</sup> aus dem Bereich Lebensmitteleinzelhandel und Drogeriemärkte zu nutzen. Um Verzerrungen bei der Berechnung von Scannerdatenindizes zu vermeiden, werden speziell für die Nutzung von Massendaten entwickelte Indexformeln verwendet. Um Scannerdaten in den Verbraucherpreisindex zu inte-

6 Grund hierfür ist, dass ein Top-down-Prädiktionsansatz verwendet wurde.

7 Es werden Handelsketten einbezogen, die kumuliert mindestens 80% des Umsatzes aller Handelsketten abdecken.

**Grafik 4**

Verschiedene Arten von Klassifikatoren





grieren, wird der Index auf Basis des GEKS-Törnqvist-Preisindex<sup>8</sup> berechnet.

Um Indizes mittels Scannerdaten zu berechnen, werden nach Dateneingang alle Artikel in einem mehrstufigen Verfahren klassifiziert, wobei auch Maschinelles Lernen eingesetzt wird. In der ersten Stufe werden Lookup-Tabellen genutzt, in denen bereits bekannte Identifikatoren aus den Scannerdaten COICOP-Klassen zugeordnet sind. So kann eine Zuordnung entweder direkt über händlerinterne Warengruppen oder über den Barcode<sup>9</sup> erfolgen. Die verbleibenden Artikel werden durch ein maschinelles Lernverfahren zugeordnet. Verwendet wird hierbei das Verfahren Random Forest, das zu den flachen Klassifikatoren gehört. Zur Qualitätssicherung werden Artikel, bei denen das Lernverfahren eine geringe Vertrauenswahrscheinlichkeit<sup>10</sup> aufweist, manuell nachklassifiziert.

▼ **Tabelle 1** zeigt die Struktur des Trainingsdatensatzes, der für jede Handelskette erstellt wird, um die Klassifikatoren zu trainieren. Anschließend erfolgt ein Rebalancing des Trainingsdatensatzes, um die Güte des Klassifikators zu verbessern.<sup>11</sup> Die Spalte „Rebalanced“

in Tabelle 1 gibt an, ob der Datensatz auch nach dem Rebalancing im Trainingsdatensatz verbleibt. Beträgt der Wert FALSE, so bleibt der Wert nach dem Rebalancing erhalten, ansonsten wird er entfernt. Die Klassifikatoren werden auf Basis der Test- und Trainingsdaten trainiert und evaluiert.

## 2.4 Vergleich und Zwischenfazit

Um die Güte der angewandten flachen und hierarchischen Klassifikatoren vergleichen zu können, werden Evaluationsmetriken<sup>12</sup> verwendet. Eine höhere Güte bedeutet, dass ein Klassifikator die Artikel häufiger den korrekten Klassen zuordnet. Die Ergebnisse sind auf 0 bis 1 normiert, wobei ein höherer Wert einem besseren Ergebnis entspricht. Die Evaluation von hierarchischen Klassifikatoren birgt zusätzliche Herausforderungen, da die Klassifizierung auf verschiedenen Ebenen stattfindet und traditionelle Metriken beziehungsweise Gütemaße dies nur unzureichend messen können. Zwei Metriken, die sich für die Bewertung von hierarchischen Klassifikatoren eignen und auf die weitere Metriken aufbauen, sind „Präzision“ und „Recall“. **Präzision** ist definiert als das Verhältnis zwischen der Anzahl der korrekt zu einer Klasse zugeordneten Artikel und der Anzahl der zu dieser Klasse zugeordneten Artikel. **Recall** bezeichnet das Verhältnis zwischen der Anzahl der korrekt zu einer Klasse zugeordneten Artikel und der Gesamtzahl der tatsächlichen Artikel dieser Klasse. Um die Leistung auf der 10-Steller-Ebene zu messen, wird als traditionelle Metrik die Metrik „F1“ genutzt, welche Präzision und Recall kombiniert. Der F1-Wert kann entweder als mikro- oder makro-Version berechnet werden. In der makro-Version wird der Durchschnitt über alle F1-Werte der einzelnen Klassen gebildet. Der F1-Wert der Klassen ist das

- 8 Der GEKS-Törnqvist-Index ist ein multilateraler Preisindex, der vom Statistischen Bundesamt und anderen nationalen Statistikämtern wie dem Office for National Statistics (ONS) eingesetzt wird.
- 9 Mit Barcode ist die Global Trade Item Number (GTIN) gemeint, die jeden Artikel international identifiziert.
- 10 Die Vertrauenswahrscheinlichkeit gibt an, wie wahrscheinlich es ist, dass für eine Beobachtung eine richtige Vorhersage getroffen wird. Beim verwendeten Random-Forest-Verfahren erhält man für jeden Artikel und jede COICOP eine Wahrscheinlichkeit, ob die zugeordnete Klasse korrekt ist.
- 11 Wenn die Klassen ungleich verteilt sind, verschlechtert sich die Güte des Klassifikators, vor allem für kleine Klassen (Fernández und andere, 2018). Beim Rebalancing wird durch die Techniken des Over- und Undersamplings die Ungleichverteilung der Klassen ausgeglichen, indem aus Mehrheitsklassen Datensätze entfernt (Undersampling) und Minderheitsklassen mit aus Datensätzen anderer Handelsketten synthetisierten Daten angereichert werden (Oversampling).

- 12 Es ist wichtig, mehrere Metriken zu verwenden, da diese die Güte der Klassifikatoren nach verschiedenen Kriterien bewerten.

**Tabelle 1**

**Aufbau der Trainingsdatensätze (fiktive Daten)**

Artikel-ID	Artikelname	Händlerkategorie	COICOP-10-Steller	Mehrwertsteuersatz	Rebalanced
54957608	BASMATI REIS	Speisekammer_Verarbeitete_Lebensmittel_Reis	0111101100	7	TRUE
30529904	Zitrone Bio	Lebensmittel_Frucht_Frisch	0116117100	7	FALSE
635343270	Kaffee Shampoo	Drogerie_Dusche_Shampoo	1213211100	19	FALSE
49967686	Apfel Shampoo	Drogerie_Dusche_Shampoo	1213211100	19	TRUE

Die Spalte „Rebalanced“ gibt an, ob der Datensatz nach dem Rebalancing im Trainingsdatensatz verbleibt (TRUE) oder nicht (FALSE).

harmonische Mittel aus Präzision und Recall. Der mikro F1-Wert wird direkt ohne Berücksichtigung der Klassen berechnet. Für eine Bewertung der hierarchischen Klassifizierung ist die makro-Version besser geeignet, da Falschzuordnungen in allen Klassen gleich gewichtet werden und deshalb eine spezifische Falschzuordnung in einer großen Klasse nicht die schlechte Zuordnung in vielen kleinen Klassen überdecken kann. Als weitere Evaluationsmetrik wird zusätzlich der „hierarchische F1“ (Kiritchenko und andere, 2006), eine hierarchische Version des mikro F1 genutzt, um die Klassifizierungsgüte auf allen Ebenen der COICOP-Klassifikation zu messen.

↗ **Tabelle 2** vergleicht den aktuell genutzten flachen Klassifikator (Random Forest) und zwei Ansätze des hierarchischen Klassifizierens (LCN und LCPN)<sup>13</sup> auf sämtlichen Ebenen der COICOP-Klassifikation. Datengrundlage sind Trainingsdatensätze, welche jeweils ungefähr 13 000 Produkte enthalten. Insgesamt zeigt sich bei den verwendeten Evaluationsmetriken kein einheitliches Bild, wobei die Unterschiede zwischen den Klassifikatoren meist nur sehr gering sind. Während der flache Klassifikator im Vergleich minimal besser die hierarchische Struktur abbildet, ist der hierarchische LCN-Klassifikator minimal besser, um die unterste Klasse (10-Steller) vorherzusagen.

Neben makro F1 und hierarchischer F1 führt Tabelle 2 auch die beiden Metriken Recall und Präzision auf. Um ein vollständiges Bild zu erhalten, wird die Metrik makro F1 zusätzlich auf verschiedenen Ebenen ausgewertet,

angefangen von der 2-Steller-Ebene bis zur 7-Steller-Ebene. Damit lässt sich feststellen, auf welcher Ebene sich die Klassifizierungen unterscheiden. Da der flache Klassifikator grundsätzlich nur auf der untersten Ebene die Klassen vorhersagen kann, wurde für die Auswertung angenommen, dass die Vorhersagen des flachen Klassifikators konsistent auf höheren Ebenen sind. Wenn also ein Artikel dem 10-Steller 0112710200 „Dauerwurst“ zugeordnet wird, erhält dieser den COICOP-2-Steller 01 „Nahrungsmittel und alkoholfreie Getränke“, den COICOP-3-Steller 011 „Nahrungsmittel“ und auf den folgenden Ebenen die entsprechenden Klassen.

Die F1-Werte für die verschiedenen Ebenen zeigen, dass der LCN-Ansatz für alle Ebenen den flachen Klassifikator und auch den LCPN-Klassifikator übertrifft, der LCPN-Klassifikator liefert vergleichbare Ergebnisse zum flachen Klassifikator. Besonders deutliche Differenzen zeigen sich auf der 5-Steller-Ebene, was insbesondere für die Vergleichbarkeit innerhalb der Europäischen Union interessant sein könnte. Der dort verwendete Harmonisierte Verbraucherpreisindex (HVPI) verwendet als tiefste Gliederung diese 5-Steller-Ebene.

Für die Klassifizierung im Rahmen des Verbraucherpreisindex ist die Metrik Präzision ausschlaggebend. Während sich bei konstanter Präzision durch einen niedrigen Recall die Varianz im Index erhöht, führt eine niedrige Präzision bei konstantem Recall zu einer systematischen Verzerrung im Index (Spackman und andere, 2023). Bezogen auf die Präzision übertreffen die hierarchischen Klassifikatoren den flachen Klassifikator. Für den Recall schneidet der LCN besser und der LCPN schlechter als der flache Klassifikator ab. Zwischen LCN- und LCPN-

<sup>13</sup> Für den LCN und den LCPN wurde für jeden lokalen Klassifikator ein Random Forest genutzt.

**Tabelle 2**

Durchschnittliche Ergebnisse von Klassifikationsverfahren nach verschiedenen Evaluationsmetriken

Klassifikator Evaluationsmetrik	Flacher Klassifikator	LCN (lokale Klassifikatoren je Klasse)	LCPN (lokale Klassifikatoren je Oberklasse)
makro F1 (10-Steller)	0,861	0,864	0,860
hF (10-Steller)	0,951	0,950	0,946
Recall (10-Steller)	0,880	0,881	0,876
Präzision (10-Steller)	0,868	0,872	0,872
makro F1 (2-Steller)	0,938	0,942	0,937
makro F1 (3-Steller)	0,944	0,947	0,942
makro F1 (4-Steller)	0,944	0,947	0,940
makro F1 (5-Steller)	0,893	0,902	0,895
makro F1 (7-Steller)	0,869	0,874	0,869

Ansatz ist die Präzision gleich, der LCN-Ansatz erreicht jedoch einen höheren Recall. Der Unterschied zwischen den beiden hierarchischen Klassifikatoren beim F1-Wert wird deswegen vom Recall getrieben.

Zusammenfassend sind die Unterschiede zwischen dem flachen Klassifikator und den hierarchischen Klassifikatoren gering. Allerdings führt der LCN-Klassifikator zu leicht besseren Ergebnissen für fast alle genutzten Metriken, auch wenn die Verbesserung nur geringfügig ist.

Über die Machbarkeitsstudie hinaus besteht weiterer Optimierungsspielraum durch ein Rebalancing auf höheren Ebenen und eine individuelle Parameteroptimierung für jeden einzelnen Klassifikator innerhalb eines lokalen Ansatzes. Diese beiden Maßnahmen könnten die Güte der hierarchischen Klassifikatoren, aber auch – in Bezug auf das Rebalancing der Klassen – des flachen Klassifikators weiter verbessern.

## 3

### Auswirkungen von Fehlklassifizierung auf die Scannerdatenindizes

#### 3.1 Vorgehen

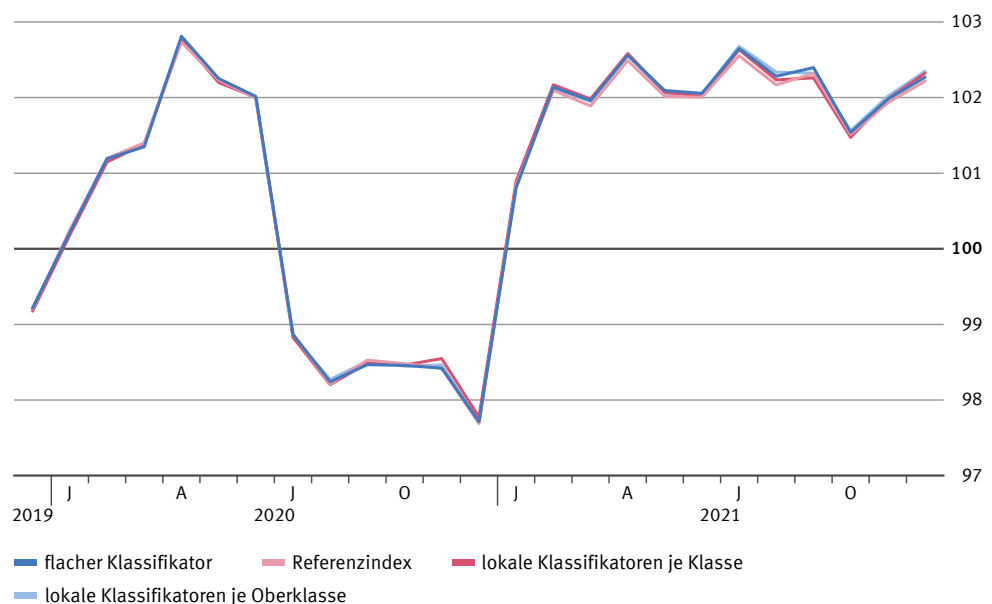
Das Ziel der Klassifizierung von Scannerdaten ist es, den Verbraucherpreisindex möglichst verzerrungsfrei abzubilden, indem die Artikel in die korrekten 10-Steller der COICOP zugeordnet werden. Eine höhere Qualität in der Klassifizierung führt zu einer geringeren Verzerrung<sup>14</sup>. Im Folgenden wird aufgezeigt, wie sich die verschiedenen Klassifizierungsansätze auf einen Index basierend auf Scannerdaten auswirken. Dazu werden Scannerdatenindizes über verschiedene Arten der Klassifizierung erstellt, um die Auswirkung der Fehlklassifizierung zu isolieren. Würde man den Verbraucherpreisindex als Referenzindex nutzen, wäre es nicht möglich, den Einfluss der Fehlklassifizierung zu bestimmen, da der Verbraucherpreisindex auf anderen Daten aus der

14 Und mathematisch gesehen auch zu einer geringeren Varianz.

Grafik 5

#### Auswirkungen unterschiedlicher Klassifikatoren auf Scannerdatenindizes

2020 = 100



traditionellen Erhebung basiert.<sup>15</sup> Daher werden die Indizes mit einem als unverfälscht angenommenen Scannerdatenindex verglichen, dessen Klassifizierung manuell erfolgt ist (Referenzindex). Der Beobachtungszeitraum umfasst 25 Monate, von Dezember 2019 bis einschließlich Dezember 2021.<sup>16</sup> Es werden lediglich Scannerdaten aus dem Geschäftstyp 04 „Discounter und Drogeriemärkte“ genutzt. Diese enthalten wöchentliche Verkaufszahlen (Absatz) und den Umsatz jedes Artikels, der in einer Filiale verkauft wurde.

Für den Vergleich werden Indizes auf Basis der in Tabelle 2 genannten Klassifikatoren berechnet und mit dem Referenzindex verglichen. Der Referenzindex basiert, wie die anderen Indizes, ausschließlich auf Artikeln, die im manuell erstellten Trainingsdatensatz enthalten sind. Während für die zu vergleichenden Indizes jeweils die Zuordnung des jeweiligen Klassifikators

genutzt wird, wird für den Referenzindex der manuell zugeordnete Trainingsdatensatz<sup>17</sup> verwendet und als korrekt angenommen. Dies gewährleistet, dass mögliche Abweichungen zwischen dem Referenzindex und den Indizes auf Basis der Klassifikatoren ausschließlich auf die Fehlklassifizierung zurückzuführen sind. In diesem Zuge werden Indizes auf 10-Steller-Ebene berechnet und durch ein gewichtetes arithmetisches Mittel auf höhere Ebenen aggregiert. Die Ergebnisse der verschiedenen Klassifikatoren zeigt [Grafik 5](#).

Um die Unterschiede zwischen den Scannerdatenindizes basierend auf den verschiedenen Klassifikatoren zu verdeutlichen, zeigt [Grafik 6](#) die absolute Abweichung jeweils zum Referenzindex für die verschiedenen Klassifikatoren.

Um das Ausmaß der Fehlklassifizierung auf den Scannerdatenindex zu bewerten, dienen die Grenzen, die das Statistische Amt der Europäischen Union (Eurostat) zur Definition eines erheblichen Methodenwechsels nutzt,<sup>18</sup> als Orientierung. Laut der EU-Verordnung han-

15 Eine Abweichung findet nur statt, wenn der falsch zugeordnete Artikel eine andere Preisentwicklung hat als die Klasse, welcher er zugeordnet wurde.

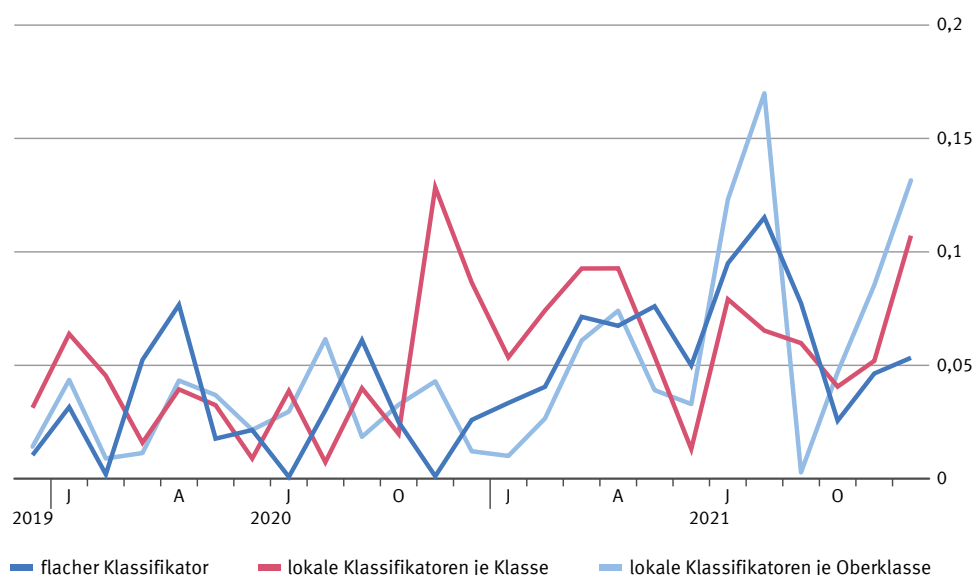
16 Der GEKS-Törnqvist-Index benötigt einen längeren Zeitraum zur Berechnung der Indizes, gängig ist dabei ein 25-Monats-Fenster. In der vorliegenden Machbarkeitsstudie können für alle betrachteten Monate die gleichen Daten aus einem beliebigen Zeitfenster verwendet werden. Im Produktionsbetrieb können nur Daten der Vergangenheit verwendet werden, sodass jeweils ein 25-Monats-Fenster in der Vergangenheit verwendet wird.

17 Es wird nicht der gesamte Trainingsdatensatz, sondern nur ein Test-Split genutzt. Dies sind 25 % des gesamten Trainingsdatensatzes. Durch die geringe Anzahl der berücksichtigten Artikel ist der Index nicht mehr repräsentativ.

18 Artikel 2 Nummer 21 der Verordnung (EU) 2016/792.

#### Grafik 6

#### Absolute Abweichung der Scannerdatenindizes vom Referenzindex in Indexpunkten



**Tabelle 3**

Angesetzte Grenzwerte für die Indexabweichung

	Verbraucherpreisindex	2-Steller	3-Steller	4-Steller	5-Steller	10-Steller
	Prozentpunkte					
Bund	<b>0,1</b>	<b>0,3</b>	<b>0,4</b>	<b>0,5</b>	<b>0,6</b>	1,0
Länder	0,2	0,6	0,8	1,0	1,2	2,0

Fett hervorgehobene Werte: bedeutender Methodenwechsel laut Verordnung (EU) 2016/792; übrige Werte in Anlehnung an die Verordnung.

delt es sich um einen bedeutenden Methodenwechsel, wenn die Veränderungsrate zum Vorjahresmonat eines Index oder eines Teilindex in einem beliebigen Zeitraum die fett hervorgehobenen Werte in [Tabelle 3](#) überschreitet. Die vorgegebenen Grenzen gelten bis zur 5-Steller-Ebene und sind auf nationale Ergebnisse ausgerichtet. Um auch für die 10-Steller und für die Bundesländerebene Grenzwerte zu erhalten, wird für den 10-Steller der Grenzwert 1,0 gesetzt, für die Bundesländerebene wird eine Verdopplung der Grenzwerte angesetzt. Die Grenzwerte auf Bundesländerebene und auf niedrigeren Ebenen der Güterklassifikation sind höher angesetzt, da weniger Artikel in den einzelnen Klassen enthalten sind: Dies kann zu einer höheren Varianz bei gleichbleibender Güte führen, da der Einfluss einzelner Artikel auf das Aggregat höher ist.

## 3.2 Ergebnisse

Wie Grafik 5 zeigt, sind die Unterschiede bei den Scannerdatenindizes, die auf Grundlage unterschiedlicher Klassifizierungsverfahren berechnet wurden, auf Bundesebene in der grafischen Darstellung kaum zu erkennen. Eine Auswertung der durchschnittlichen Abweichung vom Referenzindex unter Berücksichtigung der in Tabelle 3 aufgeführten Grenzwerte zeichnet jedoch ein differenzierteres Bild. Auf der 10-Steller-Ebene wird im betrachteten Zeitraum der Grenzwert für die Indizes auf Bundesebene auf Basis des flachen und des LCPN-Klassifikators bei 50 % und für den Index auf Basis des LCN-Klassifikators bei 44,87 % der Klassen in mindestens einem Monat überschritten. Alle drei Indizes überschreiten den Grenzwert auf der Scannerdaten-Gesamtindex-Ebene in mindestens einem Monat. Der Anteil der Klassen, die mindestens in einem Monat den Grenzwert überschreiten, ist bei den hierarchischen Klassifikatoren auf allen Ebenen niedriger oder gleich groß wie beim flachen Klassifikator, mit Ausnahme der 3-Steller-Ebene.

[Tabelle 4](#)

Ähnliche Beobachtungen wie auf der Gesamtindexebene lassen sich auch bei den Scannerdatenindizes auf Bundesländerebene beobachten. Die Indizes auf

**Tabelle 4**

Scannerdatenindizes auf Basis unterschiedlicher Klassifikationsverfahren mit mindestens einem Monatsergebnis im Zeitraum Dezember 2019 bis Dezember 2021 außerhalb der Grenzwerte

Ebene	Klassen	Flacher Klassifikator		LCN (lokale Klassifikatoren je Klasse)		LCPN (lokale Klassifikatoren je Oberklasse)	
	Anzahl		%	Anzahl	%	Anzahl	%
Scannerdaten-Gesamtindex	1	1	100	1	100	1	100
2-Steller	7	3	42,68	2	28,57	2	28,57
3-Steller	10	2	20	3	30	3	30
4-Steller	24	8	33,33	6	25	6	25
5-Steller	81	46	56,79	43	53,09	45	55,56
10-Steller	234	117	50	105	44,87	117	50

**Tabelle 5****Durchschnittliche Abweichung aufgrund von Fehlklassifizierung**

Ebene	Insgesamt			Oberhalb des Grenzwertes			Unterhalb des Grenzwertes		
	Flach	LCN	LCPN	Flach	LCN	LCPN	Flach	LCN	LCPN
Scannerdaten-Gesamtindex	0,044	0,054	0,047	0,115	0,118	0,142	0,041	0,048	0,034
2-Steller	0,126	0,090	0,091	0,852	0,454	0,448	0,067	0,063	0,065
3-Steller	0,122	0,099	0,102	0,864	0,605	0,604	0,065	0,071	0,075
4-Steller	0,173	0,146	0,154	1,047	1,057	1,024	0,107	0,100	0,108
5-Steller	0,492	0,407	0,433	2,231	2,124	1,935	0,142	0,147	0,147
10-Steller	0,623	0,555	0,599	2,627	2,666	2,573	0,224	0,217	0,230

Erläuterungen: Flach = Flacher Klassifikator; LCN = lokale Klassifikatoren je Klasse; LCPN = lokale Klassifikatoren je Oberklasse

Basis der hierarchischen Klassifikatoren haben einen kleineren Anteil an Klassen mit mindestens einem Monat oberhalb des Grenzwertes auf fast allen Ebenen. Ausnahmen auf Länderebene bilden lediglich die Scannerdaten-Gesamtindex-Ebene, bei der der LCPN-Klassifikator einen gleich hohen Anteil an Überschreitungen hat wie der flache Klassifikator, sowie die 2-Steller-Ebene, bei der dasselbe für den LCN-Klassifikator gilt.<sup>19</sup>

Weiterhin ist ebenfalls die durchschnittliche absolute Abweichung in Prozentpunkten vom Referenzindex relevant. [Tabelle 5](#) verdeutlicht, dass der LCN-Klassifikator auf allen Ebenen außer der Scannerdaten-Gesamtindex-Ebene zur geringsten Abweichung führt. Tabelle 5 zeigt außerdem, welche durchschnittlichen Abweichungen sich ergeben, wenn für die Ermittlung einerseits ausschließlich die Ebenen berücksichtigt werden, bei denen mindestens ein Monat eine Indexabweichung über dem Grenzwert aufweist (oberhalb Grenzwert), sowie andererseits Ebenen, bei denen die Indexabweichung in keinem Monat den Grenzwert überschreitet (unterhalb Grenzwert). Die Qualitätsunterschiede zwischen den Klassifikatoren sind minimal.

## 4

### Fazit und Ausblick


Die Auswertungen der Tests zeigen kein einheitliches Bild, sodass der Qualitätsvorteil eines bestimmten Klassifikators nicht valide erkennbar ist. Zusammengefasst sind die aktuell verwendete Methode des flachen Klassifizierens und die dargestellte Methode des hierarchischen Klassifizierens als qualitativ gleichwertig einzustufen. Stellt man den aktuellen Klassifizierungsprozess um, wird daher keine bedeutsame Verbesserung bei der Klassifizierung erwartet. Dennoch liefert die vorliegende Untersuchung wichtige Erkenntnisse für Projekte der amtlichen Statistik, bei denen flache Klassifikatoren bei hierarchischen Klassifikationsstrukturen genutzt werden, zum Beispiel für die Klassifikation nach der „Systematik der Einnahmen und Ausgaben der privaten Haushalte“ bei der Einkommens- und Verbrauchsstichprobe.<sup>20</sup> Der vorliegende Methodenvergleich zeigt Ansatzpunkte, hierarchische Klassifizierungsmethoden zu erproben, um die Qualität der Klassifizierung zu verbessern. Ebenso sollten die hier untersuchten Ansatzpunkte auch in der Preisstatistik weiterverfolgt werden, insbesondere sobald die Scannerdaten weiterer Einzelhandelsketten aus dem Lebensmittelbereich eingebunden sind. Da die Ergebnisse stark von der Qualität der durch die Einzelhandelsketten gepflegten Artikelbeschreibungen und Warengruppen abhängen, könnte das hierarchische Klassifizieren bei Handelsketten aus anderen Bereichen zu besseren Ergebnissen führen.

19 Der Fokus des Artikels liegt in der Interpretation der Bundesergebnisse, die Ergebnisse auf Ebene der Bundesländer werden daher nicht im Detail aufgeführt.

20 Die bisher noch offene Fragestellung, ob es sinnvoll ist, hierarchische Klassifikationsverfahren zu nutzen, nennt auch Dumpert (2021).



Aktuell befindet sich das Projekt zur Integration von Scannerdaten in den Verbraucherpreisindex in der Entwicklungs- und Aufbauphase. Es ist geplant, Scannerdaten mit der nächsten Indexrevision des Verbraucherpreisindex in den Produktionsbetrieb zu überführen. Damit kann die Preismessung zeitlich und räumlich sowie in Bezug auf den Sortimentsumfang ausgeweitet und die Preisentwicklung besser abgebildet werden. In der Folge werden Vor-Ort-Erhebungen im Lebensmittel- und Drogeriebereich größtenteils ersetzt.

Währenddessen wird weiterhin versucht, den aktuellen Klassifizierungsprozess zu verbessern. Hierzu wird die Qualität der manuell erstellten Trainingsdaten durch die Konzeption und Bereitstellung weiterer Werkzeuge erhöht sowie das bisherige Klassifikationssystem überarbeitet<sup>21</sup>. Von einem verbesserten Klassifizierungsprozess bei den Scannerdaten profitiert letztlich nicht nur die Verbraucherpreisstatistik, sondern auch weitere amtliche Statistiken, in denen die Verwendung von Scannerdaten geprüft wird. Dazu zählen aktuell der Einzelhandel (Koch/Erdemsiz, 2020), die Haushaltserhebungen und die Kaufkraftparitäten. 

---

21 Dies umfasst die Umstellung von der COICOP 1998 auf die COICOP 2018 und Umsetzung auf die nationale 10-Steller-Ebene.

## LITERATURVERZEICHNIS

---

Blaudow, Christian/Ostermann, Holger. *Entwicklung eines generischen Programms für die Nutzung von Web Scraping in der Verbraucherpreisstatistik*. In: WISTA Wirtschaft und Statistik. Ausgabe 5/2020, Seite 103 ff.

Bieg, Matthias. *Nutzung von Scannerdaten in der Preisstatistik – eine Untersuchung anhand von Marktforschungsdaten*. In: WISTA Wirtschaft und Statistik. Ausgabe 2/2019, Seite 25 ff.

Dumpert, Florian. *Machine Learning in der amtlichen Statistik – Ergebnisse und Bewertung eines internationalen Projekts*. In: WISTA Wirtschaft und Statistik. Ausgabe 4/2021, Seite 53 ff.

Dumpert, Florian. *Maschinelles Lernen im Statistischen Bundesamt*. In: WISTA Wirtschaft und Statistik. Ausgabe 4/2024, Seite 17 ff.

Fernández, Alberto/García, Salvador/Galar, Mikel/Prati, Ronaldo C./Krawczyk, Bartosz/Herrera, Francisco. *Learning from Imbalanced Data Sets*. 2018. DOI: [10.1007/978-3-319-98074-4](https://doi.org/10.1007/978-3-319-98074-4)

James, Gareth/Witten, Daniela/Hastie, Trevor/Tibshirani, Robert. *An Introduction to Statistical Learning – with Applications in R*. 2013. DOI: [10.1007/978-1-0716-1418-1](https://doi.org/10.1007/978-1-0716-1418-1)

Kiritchenko, Svetlana/Matwin, Stan/Nock, Richard/Famili, A. Fazel. *Learning and Evaluation in the Presence of Class Hierarchies: Application to Text Categorization*. In: Lamontagne, Luc/Marchand, Mario. *Advances in Artificial Intelligence*. Canadian AI 2006. DOI: [10.1007/11766247\\_34](https://doi.org/10.1007/11766247_34)

Koch, Julia/Erdemsiz, Baran. *Einsatz von Scannerdaten während der COVID-19-Pandemie*. In: WISTA Wirtschaft und Statistik. Ausgabe 4/2020, Seite 96 ff.

SAP Deutschland. *Was ist Machine Learning? Definition von Machine Learning im Detail*. [Zugriff am 14. Januar 2025]. Verfügbar unter: [www.sap.com](https://www.sap.com)

Silla, Carlos N./Freitas, Alex A. *A survey of hierarchical classification across different application domains*. In: *Data Mining and Knowledge Discovery*. Ausgabe 22/2011, Seite 31 ff. DOI: [10.1007/s10618-010-0175-9](https://doi.org/10.1007/s10618-010-0175-9)

Spackman, William/DeVilliers, Greg/Ritter, Christian/Goussev, Serge. *Identifying and mitigating misclassification: A case study of the Machine Learning lifecycle in price indices with web-scraped clothing data*. 2023. [Zugriff am 9. Januar 2025]. Verfügbar unter: [unece.org](https://unece.org)

## RECHTSGRUNDLAGEN

---

Gesetz über die Preisstatistik (Preisstatistikgesetz – PreisStatG) in der im Bundesgesetzblatt Teil III, Gliederungsnummer 720-9, veröffentlichten bereinigten Fassung, das zuletzt durch Artikel 3 des Gesetzes vom 17. Januar 2024 (BGBl. I Nr. 13) geändert worden ist.

Verordnung (EU) 2016/792 des Europäischen Parlaments und des Rates vom 11. Mai 2016 über harmonisierte Verbraucherpreisindizes und den Häuserpreisindex sowie zur Aufhebung der Verordnung (EG) Nr. 2494/95 des Rates (Amtsblatt der EU Nr. L 135, Seite 11).



**Maria Bolz**

hat einen Masterabschluss in Economics von der Universität Hohenheim und ist als Referentin im Referat „Konjunkturindizes, Saisonbereinigung“ des Statistischen Bundesamtes für die Methodik der Indexberechnung zuständig.

# DER NEUE PRODUKTIONSINDEX FÜR DIE GESAMTWIRTSCHAFT

Maria Bolz

➤ **Schlüsselwörter:** Konjunkturindikator – Indexberechnung – Indexgewichte – gewerbliche Wirtschaft – Bruttoinlandsprodukt

## ZUSAMMENFASSUNG

Der Produktionsindex für die Gesamtwirtschaft ist ein neuer zusammengesetzter Konjunkturindikator. Er kombiniert vier Konjunkturindizes, die den größten Teil der Gewerblichen Wirtschaft abdecken: die Produktionsindizes für die Industrie, das Baugewerbe, den Dienstleistungssektor sowie den deflationierten Umsatzindex für den Handel. Aufgrund des relativ breiten Erfassungsbereichs liefert der Index ein umfassendes Bild davon, wie sich die ökonomische Aktivität in Deutschland monatlich entwickelt. Der Beitrag beschreibt die methodischen Hintergründe zur Berechnung des Produktionsindex für die Gesamtwirtschaft.

➤ **Keywords:** short-term economic indicator – index calculation – index weights – non-financial business economy – gross domestic product

## ABSTRACT

*The total market production index is a new composite economic indicator. It combines four short-term indices covering most of the non-financial business economy, namely the indices of production in industry, construction and the service sector and the deflated turnover index for trade. Due to the relatively broad coverage of the index, it provides a comprehensive picture of the monthly development of economic activity in Germany. The article describes the methodological background to the calculation of the total market production index.*

## 1

### Einleitung

Der Produktionsindex für die Gesamtwirtschaft ist ein Indikator für die Entwicklung der ökonomischen Aktivität in Deutschland. Er kombiniert die vier zentralen Produktionsindizes für Industrie (in der europäischen Definition: Wirtschaftsbereiche B, C und D), Baugewerbe, Handel und Dienstleistungen und bildet konjunkturelle Entwicklungen und strukturelle Veränderungen in der Wirtschaft ab.

Kapitel 2 erläutert zunächst, welche Wirtschaftsbereiche im Produktionsindex enthalten und welche ausgeschlossen sind. Des Weiteren sind der Bezugszeitraum, das Timelag und die Veröffentlichung des Index beschrieben. Darüber hinaus werden die Indexberechnung einschließlich der zugrunde liegenden Gewichtung, die Methoden zur Berechnung der einzelnen Teilindizes und die Preis- und Saisonbereinigung thematisiert. Kapitel 3 vergleicht den deutschen mit dem europäischen Produktionsindex und stellt die Unterschiede zum Bruttoinlandsprodukt dar. Der Beitrag schließt mit einem Fazit und einem Ausblick zu möglichen Weiterentwicklungsfeldern.

## 2

### Methodik

#### 2.1 Erfassungsbereich

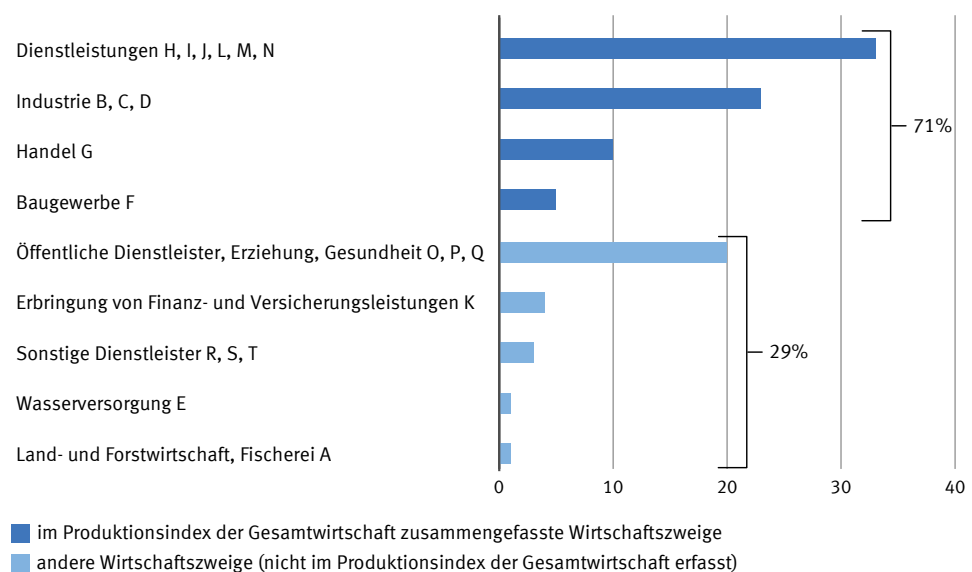
Der Produktionsindex für die Gesamtwirtschaft deckt die Industrie, das Baugewerbe, den Handel und den Dienstleistungssektor ab. [➤ Grafik 1](#) veranschaulicht den Anteil der im Produktionsindex der Gesamtwirtschaft zusammengefassten Wirtschaftszweige an der Gesamtwertschöpfung. Im Jahr 2023 entfielen rund 71 % der gesamten Bruttowertschöpfung Deutschlands auf die durch den Produktionsindex für die Gesamtwirtschaft abgedeckten Bereiche.<sup>1</sup>

Damit umfasst der Produktionsindex für die Gesamtwirtschaft alle Bereiche, die aufgrund der Verordnung über die europäische Unternehmensstatistik (EBS-Verord-

- 1 Der Produktionsindex für die Gesamtwirtschaft umfasst in den genannten Bereichen ausschließlich die wirtschaftlichen Aktivitäten von Unternehmen. Aktivitäten des Staates oder gemeinnütziger Organisationen bleiben dabei unberücksichtigt. Diese werden durch die Volkswirtschaftlichen Gesamtrechnungen erfasst.

**Grafik 1**

**Bruttowertschöpfung 2023 nach Wirtschaftsbereichen**  
in %



Wirtschaftsabschnitte der Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008).

nung) aktuell durch Konjunkturindizes abgedeckt werden und für die monatlich Daten vorliegen.<sup>12</sup>

Welche Wirtschaftsabschnitte der Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008), Teil des Produktionsindex der Gesamtwirtschaft sind, zeigt [Übersicht 1](#).

### Übersicht 1

Wirtschaftsabschnitte der Klassifikation der Wirtschaftszweige, Ausgabe 2008

Kennbuchstabe	Bezeichnung des Wirtschaftsabschnitts
A	Land- und Forstwirtschaft, Fischerei
B	Bergbau und Gewinnung von Steinen und Erden
C	Verarbeitendes Gewerbe
D	Energieversorgung
E	Wasserversorgung; Abwasser- und Abfallentsorgung und Beseitigung von Umweltverschmutzungen
F	Baugewerbe
G	Handel; Instandhaltung und Reparatur von Kraftfahrzeugen
H	Verkehr und Lagerei
I	Gastgewerbe
J	Information und Kommunikation
K	Erbringung von Finanz- und Versicherungsdienstleistungen
L	Grundstücks- und Wohnungswesen
M	Erbringung von freiberuflichen, wissenschaftlichen und technischen Dienstleistungen
N	Erbringung von sonstigen wirtschaftlichen Dienstleistungen
O	Öffentliche Verwaltung, Verteidigung; Sozialversicherung
P	Erziehung und Unterricht
Q	Gesundheits- und Sozialwesen
R	Kunst, Unterhaltung und Erholung
S	Erbringung von sonstigen Dienstleistungen
T	Private Haushalte mit Hauspersonal; Herstellung von Waren und Erbringung von Dienstleistungen durch Private Haushalte für den Eigenbedarf ohne ausgeprägten Schwerpunkt
U	Exterritoriale Organisationen und Körperschaften

Blaue Schrift: Wirtschaftsabschnitt ist Teil des Produktionsindex für die Gesamtwirtschaft.

Einige Teilbereiche von Wirtschaftsabschnitten, die Teil des Produktionsindex für die Gesamtwirtschaft sind, sind jedoch nicht in der Konjunkturstatistik und somit

auch nicht im Produktionsindex für die Gesamtwirtschaft enthalten. Dazu gehören unter anderem:

- › Verwaltung und Führung von Unternehmen und Betrieben (M70.1)
- › Forschung und Entwicklung (M72)
- › Veterinärwesen (M75)

Die EBS-Verordnung hat diese Teilbereiche vom Berichtsbereich der Konjunkturstatistiken im Europäischen Statistiksistem ausgenommen, etwa aufgrund von Schwierigkeiten der Datenerhebung oder wegen geringer ökonomischer Bedeutung der betroffenen Bereiche für die konjunkturelle Entwicklung.

Die Abgrenzung des Produktionsindex der Gesamtwirtschaft überschneidet sich mit der von Eurostat genutzten Definition der gewerblichen Wirtschaft. Diese „non-financial business economy“ umfasst die Abschnitte B bis J sowie L bis N und S95 der WZ 2008. Es handelt sich hierbei jedoch nicht um eine gesetzlich festgelegte Statistikklassifikation.

Insgesamt lässt sich der Produktionsindex der Gesamtwirtschaft als ein Indikator betrachten, der den größten Teil der gewerblichen deutschen Wirtschaft erfasst.

## 2.2 Verfügbarkeit

Der Produktionsindex für die Gesamtwirtschaft ist ab dem Monatsmonat November 2024 abrufbar, wird monatlich erstellt und steht etwa 65 Tage nach Ende des jeweiligen Bezugszeitraums zur Verfügung. Das Timelag (die Zeitspanne zwischen dem Ende des Bezugszeitraums und der Veröffentlichung der Ergebnisse) ergibt sich aus dem frühest möglichen Zeitpunkt, zu dem alle Teilindizes – also zuletzt auch der Dienstleistungsproduktionsindex – vorliegen. [Tabelle 1](#)

**Tabelle 1**

Timelags und Veröffentlichungstabellen des Produktionsindex für die Gesamtwirtschaft und seiner Teilindizes

Index	Timelag	Tabelle <sup>1</sup>
Produktionsindex für die Gesamtwirtschaft	t+65	48111-0001
Dienstleistungsproduktionsindex	t+65	47414-0006
Umsatzindex für den Handel	t+60	45212-0005
Produktionsindex für die Industrie	t+38	42153-0001
Produktionsindex für das Baugewerbe	t+38	42153-0001

1 Tabellennummer in [GENESIS-Online](#), der Datenbank des Statistischen Bundesamtes.

<sup>12</sup> Die Verpflichtung zur Lieferung von Konjunkturindizes an das Statistische Amt der Europäischen Union (Eurostat) ergibt sich aus der EBS-Verordnung. Die technischen Spezifikationen und Einzelheiten zu den Konjunkturindizes legt die Durchführungsverordnung (EU) 2020/1197 fest. Die relevanten Anforderungen finden sich insbesondere in Anhang I dieser Durchführungsverordnung.



Der Produktionsindex für die Gesamtwirtschaft wird sowohl unbereinigt als auch kalender- und saisonbereinigt veröffentlicht. Die Daten können in der Datenbank [GENESIS-Online](#), der zentralen Veröffentlichungsdatenbank des Statistischen Bundesamtes, über die Tabelle 48111-0001 abgerufen werden.

### 2.3 Berechnung

Der Produktionsindex für die Gesamtwirtschaft wird als gewichteter Mittelwert der vier Teilindizes berechnet. Die Gewichte spiegeln die Anteile der Teilbereiche an der gesamten Bruttowertschöpfung aller einbezogenen Wirtschaftszweig-Abschnitte in Deutschland im Basisjahr 2021 wider. Die Daten werden aus den jährlichen Kostenstruktur- und Strukturhebungen sowie für den Handelsbereich aus der Unternehmensstatistik des Statistischen Bundesamtes gewonnen.<sup>13</sup> ➤ **Tabelle 2** gibt die Gewichte an, mit denen die Teilindizes in den Gesamtindex eingehen.

Die vier Konjunkturindizes, die zum Produktionsindex für die Gesamtwirtschaft kombiniert werden, verwenden unterschiedliche Größen, um die ökonomische Aktivität zu messen. Für die Analyse stellt dies jedoch kein Problem dar, da bei Konjunkturstatistiken in der Regel von Interesse ist, wie sich die Ergebnisse im Zeitverlauf entwickeln, und nicht die absoluten Werte einzelner Berichtsmonate.

Um den **Produktionsindex für die Industrie** zu berechnen, wird als Messgröße in der Regel der preisbereinigte Produktionswert verwendet. Produktionswerte werden

durch die monatlichen und vierteljährlichen Produktionserhebungen ermittelt, bei der Industriebetriebe mit mehr als 50 Beschäftigten (monatlich) beziehungsweise mindestens 20 Beschäftigten (vierteljährlich) befragt werden (Statistisches Bundesamt, 2023).

Im **Baugewerbe** gibt es keine Produktionserhebung, sodass auf Arbeitsstunden und ergänzend auf Umsätze zurückgegriffen wird. Werden Arbeitsstunden herangezogen, fließt die Entwicklung der Arbeitsproduktivität mithilfe von Korrekturfaktoren grob in die Berechnungen ein. Werden Umsätze verwendet, werden diese mit Preisindizes des Baugewerbes deflationiert, um Veränderungen aufgrund von Preisänderungen auszuschließen (Linz und andere, 2024).

Beim **Dienstleistungsproduktionsindex** ist die Messgröße der preisbereinigte Nettoumsatz, der mit der Bruttowertschöpfung der Unternehmen aus der strukturellen Unternehmensstatistik gewichtet wird (Statistisches Bundesamt, 2023).

Für den **Umsatzindex im Einzelhandel** wird ebenso der Nettoumsatz als Messgröße verwendet.

Für den Produktionsindex für die Gesamtwirtschaft werden mithilfe der aufgeführten Messgrößen im nächsten Schritt Messzahlen berechnet, indem der aktuelle Wert der Messgröße eines Monats durch den Produktionswert im Basisjahr geteilt wird. Der Produktionsindex für die Gesamtwirtschaft wird als gewichteter Mittelwert der Messzahlen berechnet und gibt das Verhältnis der aktuellen Werte zu den entsprechenden Werten im Basisjahr an. Ein Indexwert von 110 bedeutet zum Beispiel, dass das monatliche Produktionsvolumen um 10% höher liegt als im Durchschnitt des Basisjahres. Das Basisjahr wird in der Regel alle fünf Jahre angepasst.

➤ **Grafik 2** stellt die vier Teilindizes des Produktionsindex für die Gesamtwirtschaft dar. Die Zeitreihe zum Gesamtindex liegt seit Januar 2015 vor, da der Dienst-

13 Kostenstrukturhebung im Bereich Verarbeitendes Gewerbe, Bergbau und Gewinnung von Steinen und Erden (42251), Strukturhebung für kleine Unternehmen im Bereich Verarbeitendes Gewerbe, Bergbau und Gewinnung von Steinen und Erden (42252), Kostenstrukturhebung im Bauhauptgewerbe (44253), Kostenstrukturhebung im Ausbaugewerbe (44254), Strukturhebung für kleine Unternehmen im Baugewerbe (44252), Strukturhebung im Dienstleistungsbereich (47415), Bereichsübergreifende Unternehmensstatistik (48112).

**Tabelle 2**

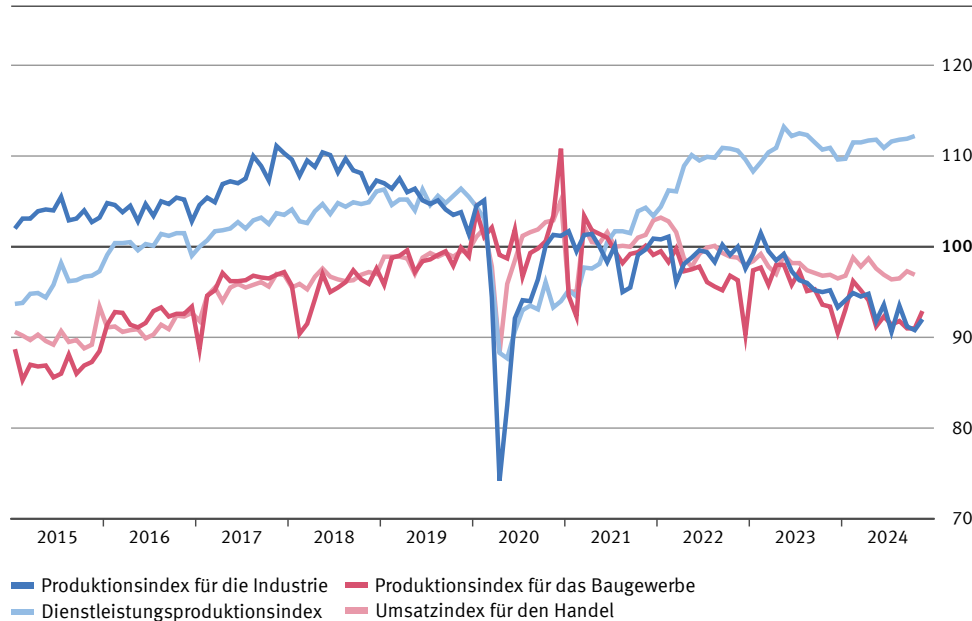
**Gewichtung des Produktionsindex für die Gesamtwirtschaft**

EVAS-Nr. <sup>1</sup>	Statistik	Wirtschaftszweig-Abschnitte	Gewichtung in %
47414	Dienstleistungsproduktionsindex	H, I, J, L, M, N	37,5
42153	Produktionsindex für die Industrie	B, C, D	34,3
45212	Umsatzindex für den Handel	G	21,1
42153	Produktionsindex für das Baugewerbe	F	7,1

1 Aus dem einheitlichen Verzeichnis aller Statistiken des Bundes und der Länder (EVAS).

**Grafik 2**

**Teilindizes des Produktionsindex für die Gesamtwirtschaft**  
2021 = 100, kalender- und saisonbereinigte Werte



leistungsproduktionsindex erst ab diesem Zeitpunkt verfügbar ist.

Hinsichtlich der geografischen Abgrenzung bezieht sich der Produktionsindex für die Gesamtwirtschaft auf Deutschland insgesamt, eine Berechnung für die Bundesländer oder tiefere regionale Abgrenzungen sind nicht verfügbar.

Der Produktionsindex für die Gesamtwirtschaft steht als Volumenindex<sup>14</sup> zur Verfügung, alle zugrundeliegenden Messgrößen, mit Ausnahme der Arbeitsstunden, sind preisbereinigt.

Für den Produktionsindex für die Gesamtwirtschaft werden sowohl (nicht saisonbereinigte) Originalwerte als auch kalender- und saisonbereinigte Werte veröffentlicht. Dabei wird eine indirekte Saisonbereinigung verwendet, bei der die kalender- und saisonbereinigten Werte der einzelnen Teilindizes aggregiert werden.<sup>15</sup>

➔ Grafik 3

4 Der Produktionsindex als Volumenindex stellt die Entwicklung des preisbereinigten Produktionswerts dar.

5 Der Unterschied zwischen direkter und indirekter Saisonbereinigung ist ausführlich erläutert in Linz und andere (2018).

## 3

### Vergleiche mit anderen Indizes und dem Bruttoinlandsprodukt

#### 3.1 Europäischer Produktionsindex für die Gesamtwirtschaft

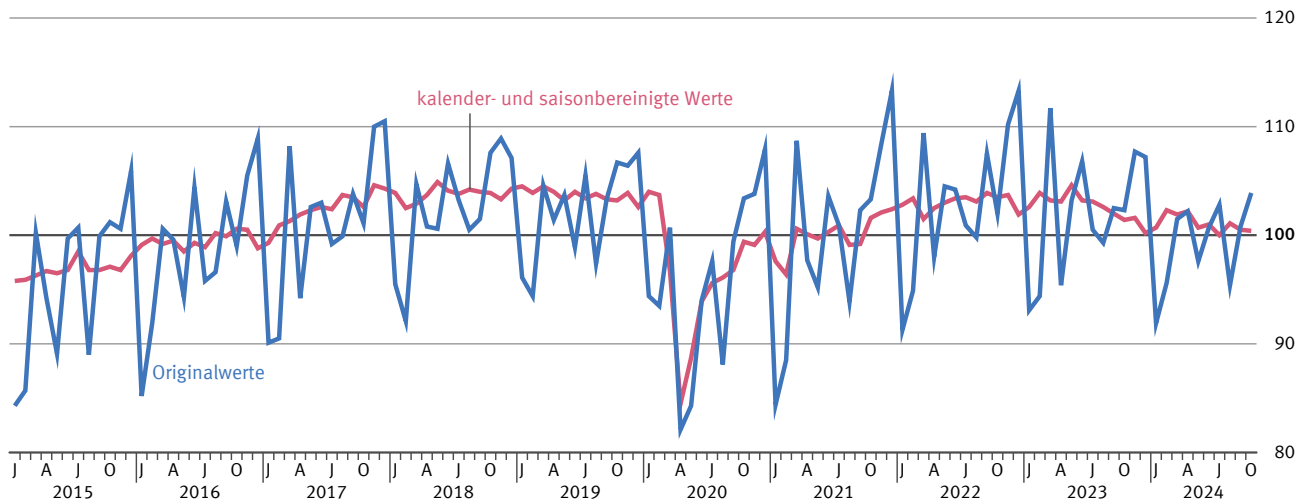
Auch Eurostat berechnet einen Produktionsindex für die Gesamtwirtschaft, der unter der Bezeichnung Total Market Production Index geführt wird. Eurostat erstellt ihn nicht nur für die Europäische Union und die Eurozone, sondern veröffentlicht ihn auch für einige einzelne Länder.<sup>16</sup> Er gibt jeden Monat einen ersten Eindruck von der Entwicklung der europäischen Wirtschaft. Da die Methodik von Eurostat<sup>17</sup> und die für den deutschen Produktionsindex für die Gesamtwirtschaft identisch sind, ist er mit den Indizes anderer Länder direkt vergleichbar.

6 Eurostat veröffentlicht den Total Market Production Index aktuell für folgende Länder: Belgien, Bulgarien, Deutschland, Estland, Finnland, Griechenland, Italien, Kroatien, Lettland, Litauen, Malta, Österreich, Rumänien, Slowakei und Zypern.

7 Eine detaillierte Methodenbeschreibung zur Berechnung des europäischen Total Market Production Index siehe unter: [ec.europa.eu](https://ec.europa.eu)

**Grafik 3**

**Produktionsindex für die Gesamtwirtschaft**  
2021 = 100



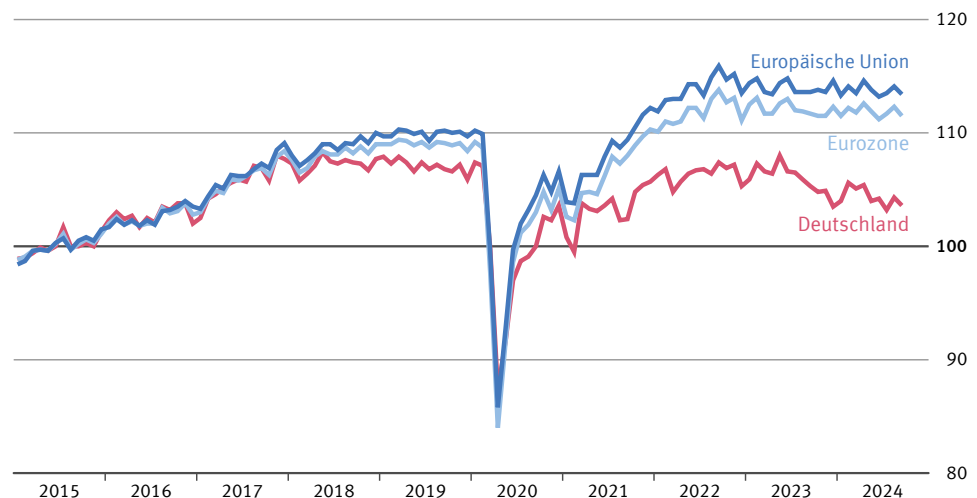
In der europäischen Datenbank können die Ergebnisse des Total Market Production Index monatlich mit einem Timelag von 65 Tagen abgerufen werden.<sup>8</sup>

<sup>8</sup> Die Tabelle ist unter folgendem Link erreichbar: [ec.europa.eu](https://ec.europa.eu). Die Eurostat-Datenbank bietet umfassende Informationen zu Konjunkturstatistiken, einschließlich spezifischer Tabellen und Zeitreihen. Sie unterstützt den Zugang zu harmonisierten Daten der EU-Mitgliedstaaten.

➤ Grafik 4 vergleicht den deutschen Produktionsindex für die Gesamtwirtschaft mit den Produktionsindizes der Europäischen Union (EU) und der Eurozone. Um die längerfristige Entwicklung seit 2015 zu betrachten, wurden die Daten nachträglich auf das Jahr 2015 normiert. Langfristig betrachtet haben sich die Produktionsindizes der EU und der Eurozone dynamischer entwickelt als der für Deutschland. Im Vergleich zum Durchschnitt des Jah-

**Grafik 4**

**Produktionsindex für die Gesamtwirtschaft**  
2015 = 100; kalender- und saisonbereinigte Werte



res 2015 stieg der Produktionsindex der EU bis 2023 um 14 %, der der Eurozone um 12 %, während der deutsche Produktionsindex nur um 6 % zulegte.

### 3.2 Bruttoinlandsprodukt

---

Im Rahmen der Volkswirtschaftlichen Gesamtrechnungen wird das vierteljährliche Bruttoinlandsprodukt berechnet – damit ist ein etablierter Konjunkturindikator für die Gesamtwirtschaft in Deutschland bereits verfügbar. Daher stellt sich die Frage, warum der Produktionsindex für die Gesamtwirtschaft als zusätzlicher Indikator sinnvoll ist.

Zunächst liefert der Produktionsindex für die Gesamtwirtschaft sehr detaillierte Ergebnisse für einzelne Wirtschaftszweige, während das Bruttoinlandsprodukt eine gröbere Einteilung nach Wirtschaftsbereichen vornimmt. Weiterhin wird der Produktionsindex monatlich veröffentlicht, sodass auch die Entwicklung innerhalb des Quartals nachvollzogen werden kann; die Ergebnisse für das Bruttoinlandsprodukt stehen hingegen nur vierteljährlich und jährlich zur Verfügung.

Die monatliche Veröffentlichungsfrequenz führt allerdings nur in den ersten beiden Monaten eines Quartals dazu, dass Informationen über die konjunkturelle Entwicklung in Deutschland früher verfügbar sind. Zum Beispiel wird das Ergebnis zum Bruttoinlandsprodukt für das erste Quartal 2025 am 23. Mai veröffentlicht. Der gesamtwirtschaftliche Produktionsindex für den Berichtsmonat Januar 2025 erscheint am 7. April und für Februar 2025 am 5. Mai. Für den Berichtsmonat März 2025 wird der Produktionsindex hingegen erst am 5. Juni veröffentlicht und somit später als die Quartalsmeldung zum Bruttoinlandsprodukt. Eine Schnellmeldung zum Bruttoinlandsprodukt für das erste Quartal 2025 ist zudem bereits am 30. April 2025 verfügbar, diese enthält allerdings keine Untergliederung nach Wirtschaftsbereichen und bezieht sich nur auf das Gesamtergebnis.

Außer der Detailtiefe und der Veröffentlichungsfrequenz gibt es konzeptionelle Unterschiede zwischen dem Produktionsindex für die Gesamtwirtschaft und dem Bruttoinlandsprodukt. Diese erlauben unterschiedliche Interpretationen der Ergebnisse und je nach Untersuchungsziel kann es sinnvoll sein, den einen oder anderen Konjunkturindikator zu verwenden.

So gibt der Produktionsindex für die Gesamtwirtschaft die Entwicklung der Produktionsaktivität an. Das Ausmaß der Produktionsaktivität in einer Volkswirtschaft ist zum Beispiel relevant für die Entwicklung von Umsätzen, Importen und Exporten, geleisteten Arbeitsstunden, Beschäftigung, Materialverbrauch, Energieverbrauch oder Verkehrsaktivität in Deutschland. Das Bruttoinlandsprodukt ist hingegen ein Maß für die Entwicklung der erwirtschafteten Leistung. Es gibt an, wieviel wirtschaftlicher Mehrwert in einer Volkswirtschaft geschaffen wurde, der zum Beispiel in Form von Löhnen und Gehältern oder Unternehmensgewinnen verteilt werden kann. So werden in der Entstehungsrechnung<sup>9</sup> der Volkswirtschaftlichen Gesamtrechnungen die im Produktionsprozess eingesetzten Vorleistungen an Waren und Dienstleistungen vom Produktionswert, also dem Wert aller produzierten Waren und Dienstleistungen, abgezogen. Ein solcher Abzug findet bei der Berechnung des Produktionsindex für die Gesamtwirtschaft nicht statt – er zeichnet die Entwicklung des Produktionsvolumens ohne Abrechnung von Vorleistungen nach. Das Bruttoinlandsprodukt und der Produktionsindex für die Gesamtwirtschaft entwickeln sich nur dann ähnlich, wenn der Anteil der Vorleistungen am Produktionswert, also die Vorleistungsquote, konstant bleibt.

Ein weiterer Unterschied ist besonders für den Bereich des Verarbeitenden Gewerbes wichtig: Das Bruttoinlandsprodukt definiert den zugrunde liegenden Produktionswert sehr weit, indem es auch den Wert von Dienstleistungen (einschließlich Handelsdienstleistungen), die durch die Unternehmen des Verarbeitenden Gewerbes erbracht werden, umfasst. Der Produktionsindex bezieht hingegen im Verarbeitenden Gewerbe im Wesentlichen den (preisbereinigten) Wert der physisch produzierten Waren ein. Der Wert von Dienstleistungen, die von den Industriebetrieben des Verarbeitenden Gewerbes erbracht werden, geht nur im Fall von Montage-, Instandhaltungs- oder Reparaturdienstleistungen ein.

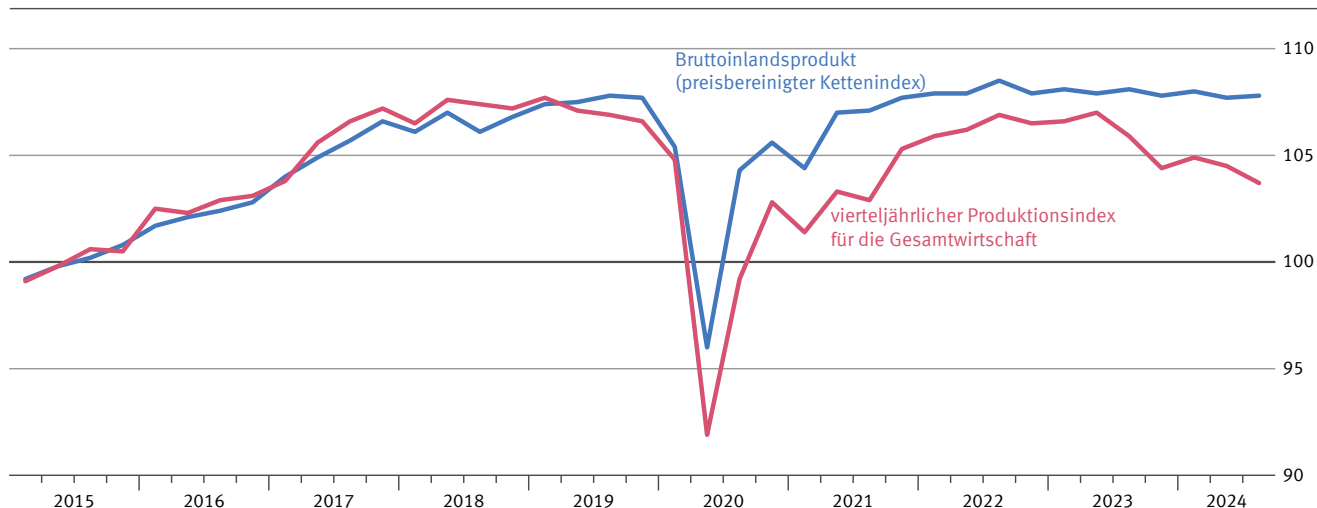
Zusätzlich zu den beiden zuletzt genannten Unterschieden umfasst der Produktionsindex für die Gesamtwirtschaft nicht alle Wirtschaftsabschnitte. Insbesondere

---

9 In den deutschen Volkswirtschaftlichen Gesamtrechnungen wird das Bruttoinlandsprodukt originär sowohl von der Entstehungs- als auch von der Verwendungsseite her berechnet. Zwischen den beiden daraus resultierenden Rechenergebnissen findet anschließend eine Abstimmung statt, die zum Veröffentlichungsergebnis des Bruttoinlandsprodukts und seiner Aggregate führt. Für weitere Erläuterungen siehe [www.destatis.de](http://www.destatis.de)

### Grafik 5

Vergleich des vierteljährlichen Produktionsindex für die Gesamtwirtschaft mit dem Bruttoinlandsprodukt  
2015 = 100; kalender- und saisonbereinigte Werte



die Abschnitte O, P und Q (öffentliche Verwaltung, Erziehung und Unterricht, Gesundheits- und Sozialwesen) gehen nicht in den Produktionsindex für die Gesamtwirtschaft ein, während das Bruttoinlandsprodukt auch diese Wirtschaftsbereiche erfasst. Darüber hinaus gibt es eine Reihe weiterer konzeptioneller Unterschiede, etwa bei der Erfassung von Steuern, Subventionen oder der Schattenwirtschaft.

➤ Grafik 5 zeigt einen Vergleich zwischen der Entwicklung des vierteljährlichen Produktionsindex für die Gesamtwirtschaft und des Bruttoinlandsprodukts für den Zeitraum von 2015 bis 2024. Die Ergebnisse wurden nachträglich auf das Jahr 2015 normiert. Für die unterschiedliche Entwicklung der beiden Indikatoren seit etwa 2018 spielen die genannten konzeptionellen Unterschiede eine wichtige Rolle, insbesondere die unterschiedliche Berücksichtigung der Vorleistungen und die breitere Abgrenzung des Produktionswertes in den Volkswirtschaftlichen Gesamtrechnungen.

### 4

## Fazit und Ausblick

Der Produktionsindex für die Gesamtwirtschaft ist ein zentraler Indikator, um die konjunkturellen Entwicklungen in Deutschland zu beobachten. Er repräsentiert den Großteil der gewerblichen Wirtschaft, indem er die Bereiche Industrie, Baugewerbe, Handel und Dienstleistungen abdeckt. Mit einem Anteil von rund 71 % dieser Wirtschaftszweige an der gesamten Bruttowertschöpfung in Deutschland bietet der Index einen Überblick über die wirtschaftliche Dynamik in Deutschland.

Das aktuelle Timelag von 65 Tagen bei der Veröffentlichung erschwert eine zeitnahe wirtschaftliche Analyse. Perspektivisch wird angestrebt, insbesondere den Dienstleistungsproduktionsindex schneller zu veröffentlichen, um die Verzögerungsdauer der Veröffentlichung des Produktionsindex für die Gesamtwirtschaft zu verkürzen.

## LITERATURVERZEICHNIS

---

Linz, Stefan/Fries, Claudia/Völker, Julia. *Saisonbereinigung der Konjunkturstatistiken mit X-12-ARIMA und mit X13 in JDemetra+*. In: WISTA Wirtschaft und Statistik. Ausgabe 4/2018, Seite 59 ff.

Linz, Stefan/Flores, Luis Federico/Bolz, Maria/Schächer, Jennifer/Eid, Nicole. *Umstellung des Produktionsindex im Produzierenden Gewerbe auf das Basisjahr 2021*. In: WISTA Wirtschaft und Statistik. Ausgabe 2/2024, Seite 55 ff.

Statistisches Bundesamt. *Klassifikation der Wirtschaftszweige 2008*. 2008. [Zugriff am 21. Januar 2025]. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

Statistisches Bundesamt. *Qualitätsbericht Produktionserhebungen*. 2024. [Zugriff am 23. Januar 2025]. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

Statistisches Bundesamt. *Qualitätsbericht Konjunkturstatistik im Dienstleistungsbereich*. 2023. [Zugriff am 23. Januar 2025]. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

## RECHTSGRUNDLAGEN

---

Durchführungsverordnung (EU) 2020/1197 der Kommission vom 30. Juli 2020 zur Festlegung technischer Spezifikationen und Einzelheiten nach der Verordnung (EU) 2019/2152 des Europäischen Parlaments und des Rates über europäische Unternehmensstatistiken, zur Aufhebung von zehn Rechtsakten im Bereich Unternehmensstatistiken (Amtsblatt der EU Nr. L 271, Seite 1).

Verordnung (EU) 2019/2152 des Europäischen Parlaments und des Rates vom 27. November 2019 über europäische Unternehmensstatistiken, zur Aufhebung von zehn Rechtsakten im Bereich Unternehmensstatistiken (Amtsblatt der EU Nr. L 327, Seite 1).



# METHODISCHE WEITERENTWICKLUNGEN FÜR UNTERNEHMENSGRUPPEN IM STATISTISCHEN UNTERNEHMENSREGISTER

Senait Andom, Johanna Bingel, Adrian Urban

➤ **Schlüsselwörter:** Globalisierung – Wirtschaftsstatistik – Kontrolle – Qualitätssicherung – Kontrollbeziehungen

## ZUSAMMENFASSUNG

Die Einführung der Unternehmensgruppe als eigene Einheit im statistischen Unternehmensregister ermöglicht es, Unternehmensgruppen präziser darzustellen. Um dabei die Qualität zu sichern, wurde die „Manuelle Unternehmensgruppenbearbeitung“ eingeführt, die die Strukturen dieser Gruppen überprüft und bei Bedarf korrigiert. Für nicht manuell bearbeitete Gruppen wurde das Verfahren der „Automatisierten Aktualisierung“ entwickelt, das mithilfe des Ähnlichkeitsmaßes RUMS die Konsistenz der Gruppenstrukturen zwischen Berichtsjahren sicherstellt. Der Beitrag erläutert die methodischen Ansätze dieser Weiterentwicklungen, die seit 2022 schrittweise umgesetzt werden, und beschreibt deren Bedeutung für die Datenbasis im statistischen Unternehmensregister.

➤ **Keywords:** globalisation – economic statistics – control – quality assurance – links of control

## ABSTRACT

*The introduction of the enterprise group as a separate entity in the German statistical business register enables a more precise representation of enterprise groups. For the purpose of quality assurance, manual enterprise group processing was introduced to review and, if necessary, correct the structures of these enterprise groups. For groups that are not manually processed, an automated updating process was developed that uses the RUMS similarity metric to ensure the consistency of enterprise group structures across reference years. This article explains the methodological approaches of these enhancements, which have been gradually implemented since 2022, and outlines their importance for the data basis of the statistical business register.*

### Senait Andom

ist Volkswirtin (M. Sc.) und als Referentin im Referat „Unternehmensregister, Gewerbeanzeigen, Verkehr“ des Statistischen Landesamtes Baden-Württemberg tätig. Ihr Aufgabenbereich umfasst das Profiling von Unternehmen und die Koordinierung der damit verbundenen Aufgaben.

### Johanna Bingel

ist Sozialwissenschaftlerin (B. A.) und als Sachbearbeiterin im Referat „Unternehmensregister, -gruppen, Erhebungsbetreuung, Profiling“ des Statistischen Bundesamtes für die Verarbeitung der Kaufdaten zu Unternehmensgruppen zuständig. Sie wirkt bei der „Manuellen Unternehmensgruppenbearbeitung“ und der Bearbeitung des EuroGruppen-Registers beim Statistischen Amt der Europäischen Union mit.

### Adrian Urban

ist Volkswirt (M. Sc.) und Referent für das Sachgebiet Unternehmensgruppen im Referat „Unternehmensregister, -gruppen, Erhebungsbetreuung, Profiling“ des Statistischen Bundesamtes. Er koordiniert die „Manuelle Unternehmensgruppenbearbeitung“ im Statistischen Verbund und entwickelt Konzepte zur Verbesserung der Datenqualität und Analyse von Unternehmensgruppen.

## 1

### Einleitung

Unternehmensgruppen (auch Konzerne genannt) haben großen Einfluss auf das Wirtschaftsgeschehen und stehen häufig im Mittelpunkt wirtschaftspolitischer Diskussionen. Im Jahr 2022 waren rund 292 000 Unternehmensgruppen in Deutschland tätig (knapp 15% der Gesellschaften) mit 68% der Beschäftigten in Deutschland und 83% der bundesweiten Umsätze. Unternehmensgruppen bündeln ihre wirtschaftlichen Aktivitäten über einzelne Gesellschaften hinweg auf einer höheren organisatorischen Ebene. Gesellschaften (Rechtliche Einheiten), die durch Kontrollbeziehungen verbunden sind, bilden eine Unternehmensgruppe. Diese wirtschaftlichen Verflechtungen zwischen Rechtlichen Einheiten werden zunehmend komplexer und erstrecken sich auch über Staatsgrenzen hinweg. Für die Wirtschaftsstatistik ist es wichtig, die Unternehmensgruppenstrukturen genau zu erfassen, denn sie spiegeln die tatsächliche wirtschaftliche Bedeutung und den Einfluss der Marktakteure besser wider als die Betrachtung einzelner Gesellschaften. Der Bedarf an hochwertigen Informationen hierzu ist in den letzten Jahren stark gestiegen.

Wichtige Fragen, die sich im Zusammenhang mit der wirtschaftlichen Bedeutung von Unternehmensgruppen stellen, sind beispielsweise: „Welche wirtschaftliche Macht haben die 100 bedeutsamsten Unternehmensgruppen?“, „Welche Wirtschaftsbereiche werden am stärksten von Unternehmensgruppen bestimmt?“ oder „Wie hoch ist der Einfluss von Unternehmensgruppen in Deutschland, die aus der Russischen Föderation, China oder dem Mittleren Osten kontrolliert werden?“. Um solche Fragen beantworten zu können, ist eine qualitativ hochwertige Abbildung von Unternehmensgruppen im statistischen Unternehmensregister wichtig. Seit 2005 werden Informationen zu Unternehmensgruppen in der Datenbank des Unternehmensregisters geführt. Grundlage dafür ist die EU-Verordnung 2019/2152.<sup>1</sup> Aller-

dings waren die bisherigen Möglichkeiten, die Unternehmensgruppen in der Datenbank abzubilden, sehr rudimentär und von Qualitätsmängeln geprägt. Diese Mängel sowie das Fehlen der Unternehmensgruppe als Einheitentyp<sup>2</sup> erlaubten Auswertungen zu Unternehmensgruppen nur unter erschwerten Bedingungen (Kleber und andere, 2010). Mit Einführung der Unternehmensgruppe als eigenem Einheitentyp in der Datenbank des statistischen Unternehmensregisters wurde im Jahr 2022 erstmals die Möglichkeit geschaffen, dort umfangreiche Gruppeninformationen abzubilden. Zusätzlich haben einige Statistischen Ämter der Länder im Jahr 2023 die „Manuelle Unternehmensgruppenbearbeitung“ eingeführt, die wesentlich dazu beitragen soll, qualitätsgesicherte Daten zu Unternehmensgruppen zur Verfügung zu stellen.

Dieser Aufsatz erklärt grundlegende Konzepte zu Unternehmensgruppen, beschreibt Prozesse sowie die Abbildung von Unternehmensgruppen in der Datenbank des Unternehmensregisters und erläutert die Aufgaben der „Manuellen Unternehmensgruppenbearbeitung“ näher.

## 2

### Definition der Unternehmensgruppe

Um Unternehmensgruppen als statistische Einheiten zu betrachten, ist zunächst eine eindeutige und in der Praxis umsetzbare Definition der Gruppe als Zusammenschluss von Rechtlichen Einheiten (RE) erforderlich. Im Fachbereich Unternehmensregister wurde das statistische Kontrollkonzept entwickelt, das angelehnt ist an die Bestimmungen zur Unternehmensgruppe aus der EU-Verordnung 2019/2152. Darin wird die Unternehmensgruppe definiert als ein Zusammenschluss von Rechtlichen Einheiten, die über eine rechtliche Kontrollbeziehung miteinander verbunden sind. Hierzu wird weiter im statistischen Kontrollkonzept präzisiert:

„Eine Unternehmensgruppe ist ein Zusammenschluss von Rechtlichen Einheiten, wobei mindestens zwei juristische Personen i.w.S. über Kontrollbeziehungen verbunden sind. Das Gruppenoberhaupt der Unterneh-

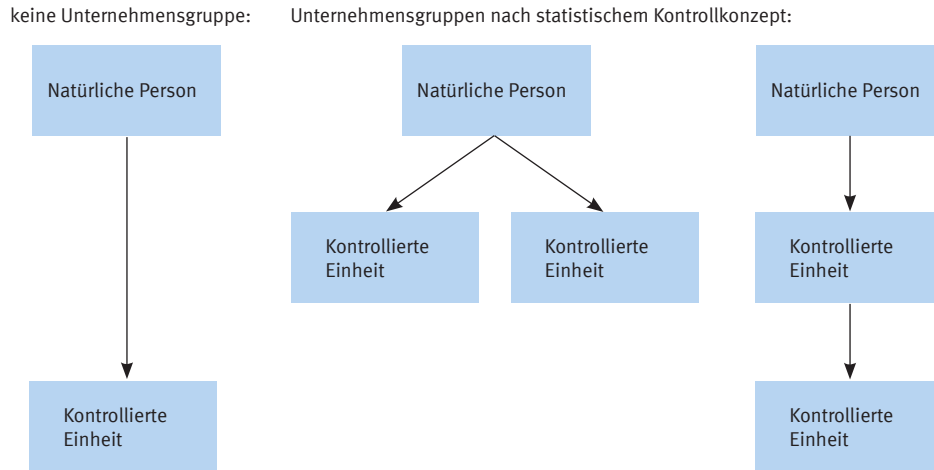
---

1 Die Europäische Union (EU) verpflichtete mit der überarbeiteten Registerverordnung (Verordnung [EU] Nr. 177/2008), später abgelöst durch die EU-Verordnung 2019/2152, die Mitgliedstaaten zur Erfassung von Unternehmensgruppen in ihren Statistikregistern. In Deutschland bestand diese Verpflichtung bereits durch § 47 des Gesetzes gegen Wettbewerbsbeschränkungen für die Arbeit der Monopolkommission. Daher beschaffte die deutsche amtliche Statistik bereits 2005 erstmals Daten zu Unternehmensgruppen und verarbeitete diese zunächst testweise, später regelmäßig.

2 Die Datenbank des statistischen Unternehmensregisters führt neben dem neuen Einheitentyp Unternehmensgruppe auch die Einheitentypen Unternehmen, Rechtliche Einheit und Niederlassung.

## Grafik 1

### Unternehmensgruppen mit natürlicher Person als Gruppenoberhaupt

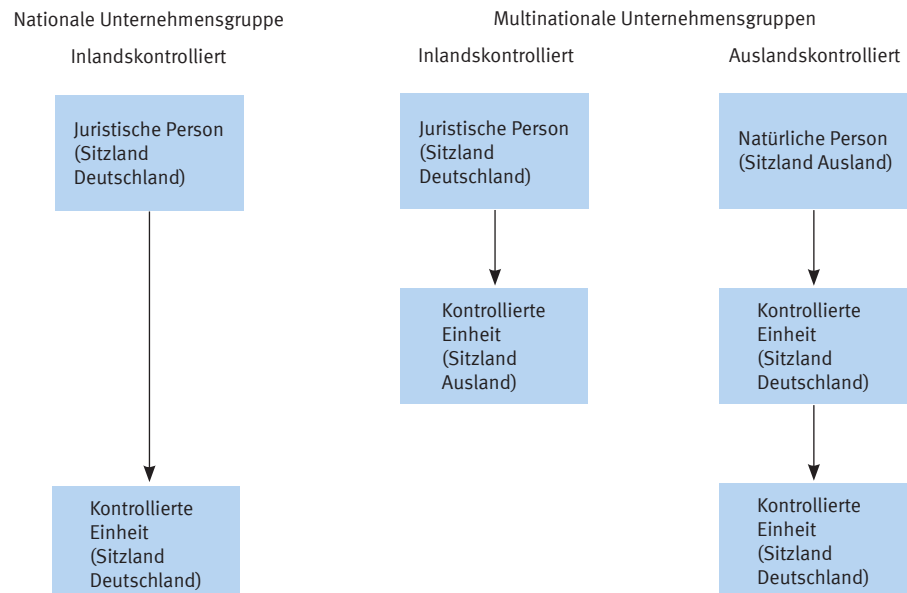


Unternehmensgruppe fungiert als das höchste Kontrollorgan in der Gruppe und wird von keiner anderen Einheit kontrolliert. Das Gruppenoberhaupt einer Unternehmensgruppe kann auch eine natürliche Person sein; in diesem Fall müssen mindestens zwei juristische Personen i. w. S. von dieser natürlichen Person kontrolliert werden. Für die globale Betrachtung einer Unternehmensgruppe ist es unerheblich, in welchem Land die Rechtlichen Einheiten ihren Sitz haben.“ [↪ Grafik 1](#)

Relevant für die deutsche amtliche Statistik beziehungsweise für das deutsche statistische Unternehmensregister sind alle Unternehmensgruppen, die mindestens eine deutsche Rechtliche Einheit enthalten. Unternehmensgruppen können in nationale und multinationale Unternehmensgruppen unterteilt werden. Eine multinationale Unternehmensgruppe enthält Rechtliche Einheiten, die mindestens in zwei verschiedenen Staaten ansässig sind. Ist dies nicht der Fall, so handelt es sich um eine nationale Unternehmensgruppe. [↪ Grafik 2](#)

## Grafik 2

### Nationale und multinationale Unternehmensgruppen



In Unternehmensgruppen gibt es wichtige Rollen/Funktionen, die einzelne Rechtliche Einheiten der Unternehmensgruppe ausüben: das Gruppenoberhaupt und die „Deutsche Entscheidungseinheit“.

Für die Außendarstellung und gegebenenfalls Befragung der Unternehmensgruppe ist es notwendig, eine repräsentative Einheit festzulegen. Diese muss ihren Sitz in Deutschland haben und sollte als strategische und operative Steuerungseinheit der Unternehmensgruppe fungieren. Es soll immer eine wirtschaftlich aktive Einheit und keine natürliche Person sein. Im Unternehmensregister wird diese Einheit als die „Deutsche Entscheidungseinheit“<sup>3</sup> gekennzeichnet. Diese Einheit kann mit dem Gruppenoberhaupt übereinstimmen, muss es jedoch nicht. Dies gilt insbesondere für Gruppen mit natürlichen Personen als höchstem Kontrollorgan sowie für Gruppen mit Gruppenoberhaupt im Ausland.

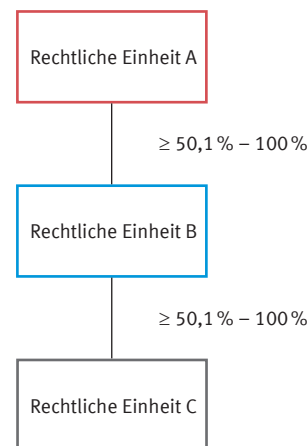
Um die Zusammensetzung einer Unternehmensgruppe zu identifizieren oder zu bestimmen, ist der Nachweis von Kontrolle zwischen den Rechtlichen Einheiten essenziell. Im statistischen Kontrollkonzept wird hierzu definiert: „Kontrolle liegt genau dann vor, wenn eine Rechtliche Einheit (Muttereinheit) die Macht hat, die finanziellen und betrieblichen Entscheidungsprozesse einer anderen Rechtlichen Einheit (Tochtereinheit) so zu beherrschen, dass sie Nutzen aus der Geschäftstätigkeit der beherrschten Einheit ziehen kann.“

Die sogenannte Muttereinheit besitzt hierbei einen bestimmenden Einfluss über die mittel- und langfristigen

3 Analog dazu gibt es in der „Europaweiten Manuellen Unternehmensgruppenbearbeitung“ die Festlegung auf eine globale Entscheidungseinheit (global decision center) in der multinationalen Unternehmensgruppe.

**Grafik 3**

Kontrolle durch Mehrheit der Kapital- beziehungsweise Stimmrechtsanteile



Strategien einer oder mehrerer untergeordneter Rechtlicher Einheiten (Tochtereinheiten). Der bestimmende Einfluss zeigt sich in der Regel in einer Anteils- beziehungsweise Stimmrechtsmehrheit. Im Grundsatz wird davon ausgegangen, dass einer Kapitalmehrheit auch eine Stimmenmehrheit in den Haupt- beziehungsweise Gesellschafterversammlungen entspricht, die sich letztlich in den Leitungs- und Kontrollorganen der betreffenden Einheit durchsetzt. ➤ Grafik 3

Kontrolle kann aber auch faktisch (wirtschaftlich und organisatorisch) vorliegen, zum Beispiel infolge einer Konsolidierung einer anderen Einheit in das eigene Rechnungswesen (Befreiungsvermerk nach § 264 Absatz 3 Handelsgesetzbuch), durch bestehende Beherrschungs- und Gewinnabführungsverträge oder bei einer identischen Geschäftsführung. ➤ Grafik 4 und ➤ Grafik 5

**Grafik 4**

Faktische Kontrolle durch ein konsolidiertes Rechnungswesen<sup>1</sup>

## 2. KONSOLIDIERUNGSKREIS

Der Konsolidierungskreis umfasst neben dem Mutterunternehmen Duravit AG, Hornberg, folgende Tochterunternehmen:

	Nennkapital EUR	Beteiligungsquote %	Eigenkapital zum 31.12.2021 EUR	Jahresergebnis 2021 EUR
Duravit S.A.S., Bischwiller/ Frankreich	4.824.016,00	100,00	9.257.688,49	281.030,00
Duravit Sanitärporzellan Meißen GmbH, Meißen <sup>1</sup>	256.000,00	100,00	256.000,00	Ergebnisabführungsvertrag
DURALOG Duravit Logistik GmbH, Hornberg <sup>1</sup>	256.000,00	100,00	256.000,00	Ergebnisabführungsvertrag
Duravit USA Inc., Atlanta/USA	88.339,00	100,00	3.236.233,70	514.226,62
DURAVIT BeLux S.P.R.L., Overijse/Belgien	19.000,00	100,00	-1.445.232,31	306.529,53
Duravit España S. L., Barcelona/Spanien	51.850,00	100,00	802.839,45	90.262,16
Duravit Egypt Sanitaryware Manufacturing Company (S.A.E.), Kairo/Ägypten	6.633.790,48	100,00	42.832.064,53	12.776.550,73

1 Öffentlich zugänglicher Konzernabschluss aus dem Bundesanzeiger.

**Grafik 5**

Faktische Kontrolle durch Beherrschungs- und Gewinnabführungsverträge<sup>1</sup>

Handelsregister B des Amtsgerichts Stuttgart	Abteilung B Wiedergabe des aktuellen Registerinhalts Abruf vom 10.06.2021 14:43
<p>Mit der "Ernst Klett Aktiengesellschaft", Stuttgart (Amtsgericht Stuttgart HRB 17521) wurde am 30.09.2014 ein Beherrschungs- und Gewinnabführungsvertrag abgeschlossen, dem die Gesellschafterversammlung am 14.10.2014 zugestimmt hat.</p>	
1. Anzahl der bisherigen Eintragungen:	
43	
2. a) Firma:	
Ernst Klett Verlag GmbH	

<sup>1</sup> Öffentlich zugängliche Information aus dem Handelsregister.

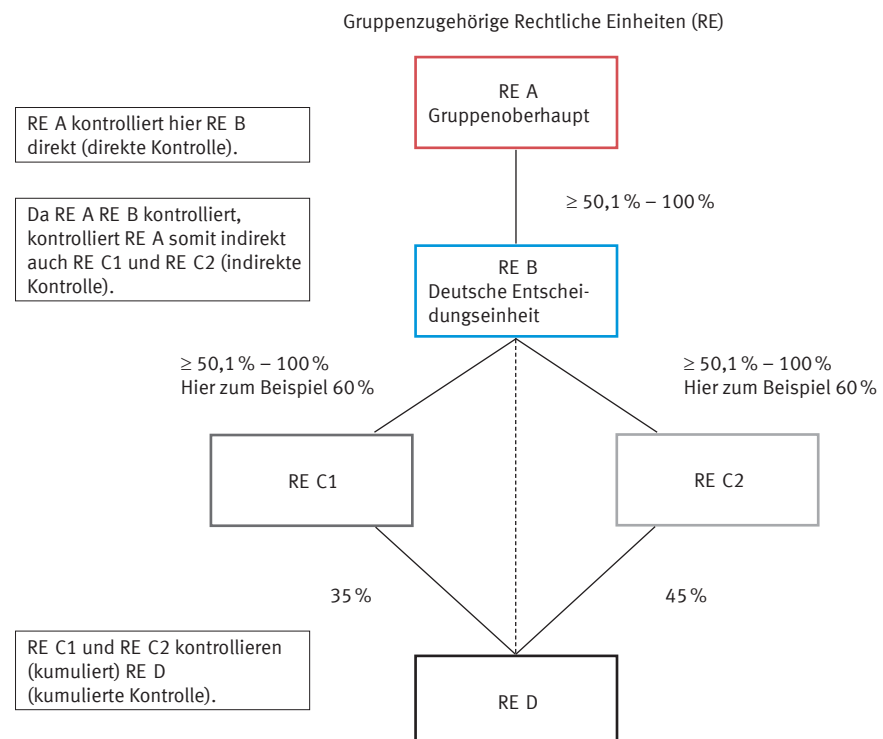
Die Kontrolle kann sowohl direkt zwischen den Einheiten vorliegen als auch indirekt, indem die Muttereinheit mittelbar über weitere Einheiten, die sie direkt kontrolliert, einen beherrschenden Einfluss auf andere Rechtliche Einheiten ausübt. Auch wenn die weiteren direkt kontrollierten Rechtlichen Einheiten individuell keine Kapital- oder Stimmrechtsmehrheit an einer anderen Rechtlichen Einheit besitzen, ist durch ihre gemeinsame

Mehrheit eine indirekt kumulierte Kontrolle gegeben. Innerhalb der Unternehmensgruppe entsteht durch direkte und indirekte Kontrollverhältnisse eine hierarchische Struktur beziehungsweise Baumstruktur. An der Spitze dieser Hierarchie steht das Gruppenoberhaupt.

➤ Grafik 6

**Grafik 6**

Baumstruktur einer Unternehmensgruppe



Zwar kann jede Einheit Kontrolle ausüben, aber nicht jede Einheit kann kontrolliert werden. Nur juristische Personen können kontrolliert werden. Natürliche Personen, Genossenschaften oder Erbgemeinschaften können per definitionem nicht kontrolliert werden.

### 3

## Nutzzende von Informationen zu Unternehmensgruppen

Der Bedarf an qualitativ hochwertigen Informationen zu Unternehmensgruppen ist aufgrund der Globalisierung und der steigenden Anforderungen bei der Bewertung multinationaler Zusammenhänge stark gestiegen. Zur Erfüllung der Nutzendenbedarfe sehen auf europäischer Ebene die EU-Verordnung 177/2008 und die diese

ablösende EU-Verordnung 2019/2152 die Führung der Unternehmensgruppen im statistischen Unternehmensregister vor.

➤ **Übersicht 1** zeigt die wichtigsten Verwendungen der Unternehmensgruppendaten des statistischen Unternehmensregisters.

### 4

## Die Verarbeitung der Unternehmensgruppenlieferung

In Deutschland gibt es keine zentrale und vollumfängliche administrative Quelle zur Erfassung oder Verwaltung von Unternehmensgruppen, die die Statistischen Ämter des Bundes und der Länder (diese bilden den Sta-

### Übersicht 1

#### Verwendungen der Unternehmensgruppendaten des statistischen Unternehmensregisters

Art/Nutzzende der Verwendung	Einsatz von Daten zu Unternehmensgruppen	Verwendung der Ergebnisse	Gesetzliche Grundlage
Profiling von Unternehmen	Profiling ist eine Methode zur Ermittlung von Unternehmen nach EU-Definition. Unternehmensgruppen bilden den Ausgangspunkt für das Profiling. Für das Profiling in Deutschland werden nur die nationalen Teile der Unternehmensgruppen betrachtet.	Die Ergebnisse des Profiling bilden die Grundlage für die Ergebnisdarstellung der Unternehmensstrukturstatistik.	EU-Verordnung 2019/2152
EuroGruppenRegister (EGR)	Nationale Teile multinationaler Unternehmensgruppen sowie deren Beziehungen ins Ausland werden jährlich von EU/EFTA-Mitgliedstaaten an Eurostat für die Befüllung des EGR übermittelt.	Zusammenführung der Daten, um ein Gesamtbild multinationaler Unternehmensgruppen zur Ermöglichung von Globalisierungsstatistiken zu erstellen	EU-Verordnung 2019/2152
Europaweite manuelle Unternehmensgruppenbearbeitung (EMU)	Manuelle Überprüfung der Gruppenstrukturen bedeutender multinationaler Unternehmensgruppen im EGR durch EU/EFTA-Mitgliedstaaten	Qualitätssicherung der Strukturen der bedeutendsten multinationalen Unternehmensgruppen im EGR	
Inward-Foreign Affiliates Statistics (IFATS)	Die IFATS nutzen die Unternehmensgruppendaten für Informationen zur Kontrolle bei Rechtlichen Einheiten und Unternehmen.	Statistik über die Unternehmen in Deutschland, die aus dem Ausland kontrolliert werden	Verordnung (EG) Nr. 716/2007
Monopolkommission	Bereitstellung von Informationen zu Unternehmensgruppen für die Bewertung des Konzentrationsgrades Rechtlicher Einheiten	Beitrag zur Erstellung der Hauptgutachten zum Stand und zur Entwicklung der Unternehmenskonzentration in Deutschland	§ 44 GWB
Large Cases Unit (LCU)	Aufdeckung von Dateninkonsistenzen bei bedeutenden multinationalen Unternehmensgruppen basierend auf verschiedenen Erhebungen	Qualitätssicherung der Volkswirtschaftlichen Gesamtrechnungen und der Wirtschaftsstatistiken	§ 1 QVWSG
Trade by Enterprise Characteristics (TEC)	Unternehmensgruppendaten werden genutzt, um zu bestimmen, von wo Außenhandel betreibende Unternehmen kontrolliert werden.	Verknüpfung von Außenhandelsdaten mit Unternehmensregisterdaten; Darstellung von Außenhandelsergebnissen nach Unternehmensmerkmalen in der TEC-Statistik (Kruse und andere, 2023)	EU-Verordnung 2020/1197
Landwirtschaftsstatistik	Analyse der Unternehmensgruppenzugehörigkeit Rechtlicher Einheiten zur Ermittlung von Eigentumsverhältnissen bei Agrarflächen	Erstellung von Statistiken zu Unternehmensgruppen in der Landwirtschaft (Statistisches Bundesamt, 2024)	EU-Verordnung 2018/1091, Durchführungsverordnung 2018/1874, § 27 Agrarstatistikgesetz

GWB: Gesetz gegen Wettbewerbsbeschränkungen; QVWSG: Gesetz zur Prüfung von Daten multinationaler Unternehmensgruppen zur Sicherung der Qualität der Volkswirtschaftlichen Gesamtrechnungen und der Wirtschaftsstatistiken



tistischen Verbund) für die Führung der Unternehmensgruppen im Unternehmensregister nutzen könnten. Deshalb werden Daten von kommerziellen Datenanbietern verwendet. Diese Anbieter führen zwar die Einheit Unternehmensgruppe nicht direkt in ihrem Datenbestand, können jedoch im großen Umfang die relevanten direkten und indirekten Kontrollbeziehungen zwischen Rechtlichen Einheiten ermitteln.

Die Datenbeschaffung erfolgt über eine öffentliche Ausschreibung zur jährlichen Lieferung von Stamm- und Beziehungsinformationen auf Ebene der Rechtlichen Einheiten. Der Stand der Kontrollbeziehungen wird zum 31. Dezember eines Berichtsjahres ermittelt, die Datenlieferung erfolgt zum 31. März des Folgejahres. Die jährliche Lieferung enthält Informationen zu Unternehmensgruppen mit mindestens einer deutschen gruppenzugehörigen Rechtlichen Einheit. Neben Stammdateninformationen und Beziehungsinformationen zu deutschen Rechtlichen Einheiten umfasst die Datenlieferung noch zusätzlich Informationen über ausländische erste Muttergesellschaften und Tochtergesellschaften sowie ausländische Gruppenoberhäupter. Die Verarbeitung dieser Unternehmensgruppenlieferung (UGL) im statistischen Unternehmensregister nennt sich **primäre UGL-Verarbeitung**.

Neben dem kommerziellen Datenanbieter gibt es weitere ergänzende sekundäre Quellen vom Bundeszentralamt für Steuern über umsatzsteuerliche Organschaften und vom Berichtskreismanagement der öffentlichen Finanzstatistiken über öffentliche Fonds, Einrichtungen und Unternehmen. Diese Quellen enthalten jedoch lediglich Informationen zu Teilmengen der Unternehmensgruppen. Die Verarbeitung dieser Informationen im Unternehmensregister heißt **sekundäre UGL-Verarbeitung**.

### 4.1 Die primäre UGL-Verarbeitung

---

Bei der primären UGL-Verarbeitung überführen die Statistischen Ämter des Bundes und der Länder die gekauften Daten in die Datenbank des Unternehmensregisters im Zuge einer ersten, weitestgehend maschinell unterstützten Verarbeitung. Diese fokussiert auf den Abgleich der Stammdateninformationen zu den Einheiten. Beziehungsdaten werden bis zu diesem Zeitpunkt nur auf technische Aspekte über formale Plausibilitätschecks geprüft und überarbeitet. Eine inhaltliche Überprüfung

erfolgt nicht. Die manuelle Überprüfung der Beziehungsinformationen auf ihre Richtigkeit findet später für ausgewählte Unternehmensgruppen statt (siehe Kapitel 6).

Bei der sogenannten Zuordnungsfindung werden die Stammdateninformationen zu deutschen Rechtlichen Einheiten (UGL-Einheiten) mit den entsprechenden Rechtlichen Einheiten in der Datenbank des Unternehmensregisters abgeglichen und verknüpft. Zudem werden alle neuen, also vor dem Berichtsjahr noch nicht existierenden und nicht von anderen Quellen des Unternehmensregisters bereits gelieferten deutschen juristischen Personen mit einem Handelsregistereintrag erstmalig in der Datenbank aufgenommen.

Darüber hinaus findet in der primären UGL-Verarbeitung auch eine Prüfung von bedeutsamen Merkmalsänderungen bei den Stammdateninformationen der UGL-Einheiten statt. Hierbei wird überprüft, ob eine Veränderung bei der UGL-Einheit korrekt in die dazugehörige Einheit in der Datenbank übernommen wurde.

### 4.2 Die sekundäre UGL-Verarbeitung

---

Im Unternehmensregister werden weitere, regelmäßig zu verarbeitende Informationen als ergänzende Pflege von Unternehmensgruppen verwendet. Dies erfolgt im Zuge der sekundären UGL-Verarbeitung. Eine ergänzende Quelle für die aus der primären UGL-Verarbeitung gewonnenen Erkenntnisse sind die steuerlichen Organkreise (STO). Steuerliche Organkreise sind für steuerliche Zwecke erfolgende Zusammenfassungen mehrerer Rechtlicher Einheiten. Sie bestehen aus einem Organträger und mehreren Organgesellschaften. Die Informationen hierzu erhält das Unternehmensregister vom Bundeszentralamt für Steuern, die Pflege von Organkreisstrukturen in der Datenbank übernehmen die Statistischen Ämter der Länder. Steuerliche Organkreise stellen konzeptionell (deutsche) Teile von Unternehmensgruppen dar, das heißt sie können eine Unternehmensgruppe in der Regel nicht vollständig beschreiben, unter anderem auch wegen Einschränkungen auf bestimmte Rechtsformen. Organkreise können jedoch eine ergänzende Quelle sein, um fehlende gruppenzugehörige Rechtliche Einheiten zu identifizieren. In den überwiegenden Fällen wird mit den Organkreisbeziehungen die Beherrschungsaussage aus den Daten des kommerziellen Anbieters bestätigt. Die Bearbeitung der Organkreise geschieht im

jährlichen Turnus. Die Verknüpfung der Gruppeninformationen aus den steuerlichen Organkreisen mit den deutschen Rechtlichen Einheiten für das jeweilige Berichtsjahr erfolgt zentral und weitgehend automatisiert durch das Statistische Bundesamt. Wenn der kommerzielle Datenanbieter eine andere Beherrschungsaussage für eine bestimmte Rechtliche Einheit liefert als die steuerlichen Organkreise, dann wird die Gruppeninformation aus den steuerlichen Organkreisen nicht mit der deutschen Rechtlichen Einheit in der Datenbank verknüpft.

Eine weitere ergänzende Quelle sind die Eignerstrukturen aus dem Berichtskreismanagement für öffentliche Einheiten (BKM). Das Berichtskreismanagement der öffentlichen Finanzstatistiken führt eine Datenbank über Eignerstrukturen (Kontrollbeziehungen) und stellt deren Jahresauszüge seit dem Berichtsjahr 2018 dem Unternehmensregister zur Verfügung (Statistisches Bundesamt, 2024a). Im Gegensatz zu steuerlichen Organkreisen sind die Gruppenstrukturen hier zumindest für den nationalen Kontext vollständig<sup>4</sup>, das heißt die BKM-Eignerdaten liefern stets eine Aussage zum höchsten Kontrollorgan innerhalb einer öffentlich kontrollierten Unternehmensgruppe. Hierdurch können in Deutschland öffentlich kontrollierte Unternehmensgruppen nahezu vollständig erfasst werden. Die Verknüpfung der Gruppeninformationen aus der Quelle BKM mit den deutschen Rechtlichen Einheiten für das jeweilige Berichtsjahr erfolgt ebenso zentral durch das Statistische Bundesamt.

Nach Abschluss der Arbeiten der primären sowie der sekundären UGL-Verarbeitung sind alle Rechtlichen Einheiten, die zu einer Unternehmensgruppe gehören, in der Datenbank des statistischen Unternehmensregisters gekennzeichnet.

---

<sup>4</sup> Ausländische Töchter in öffentlicher deutscher Kontrolle sind nicht im Berichtskreismanagement für öffentliche Einheiten enthalten.

## 5

### Einführung der Einheit Unternehmensgruppe in der Datenbank des Unternehmensregisters

---

Bei der UGL-Verarbeitung werden, wie bereits in Kapitel 4 beschrieben, lediglich Informationen über die Zugehörigkeit zu Unternehmensgruppen bei den gruppenzugehörigen Rechtlichen Einheiten in der Datenbank des Unternehmensregisters abgebildet. Diese Informationen umfassen zum Beispiel die Beziehung zur direkten Mutter und die Beziehung zum Gruppenoberhaupt der Unternehmensgruppe. Mit der primären und sekundären UGL-Verarbeitung werden noch keine Informationen für die Einheit Unternehmensgruppe in der Datenbank bereitgestellt. In der Vergangenheit waren gezielte Analysen außerhalb der Datenbank des Unternehmensregisters, zum Beispiel Auswertungen mit dem Statistikprogramm SAS, erforderlich, um Unternehmensgruppen zu identifizieren.

Um die Verarbeitung in externen Prozessen zu reduzieren, wurde im Jahr 2022 für das Berichtsjahr 2020 erstmals die Unternehmensgruppe als eigene Einheit in der Datenbank des statistischen Unternehmensregisters eingeführt. Im Nachfolgenden bezeichnet die Abkürzung „UG“ (Plural: „UGn“) die technische Einheit Unternehmensgruppe in der Datenbank. Alle deutschen Rechtlichen Einheiten, die nach den abgebildeten Beziehungsinformationen in der Datenbank auf dasselbe Gruppenoberhaupt verweisen, werden derselben UG zugeordnet. Im Gegensatz zur UGL-Verarbeitung ist die UG in der Datenbank somit nicht nur eine Sammlung von Beziehungsinformationen bei den Rechtlichen Einheiten, sondern bildet die Unternehmensgruppe als Ganzes ab, wie es die EU-Verordnung 2019/2152 verlangt.

Die UG wird mit eigenen Merkmalen in der Datenbank des statistischen Unternehmensregisters geführt. Dazu zählen eine UG-ID, ein UG-Typ (national oder multinational), ein Name sowie wirtschaftliche Kennzahlen wie Beschäftigte und Umsätze, die sich auf den deutschen Teil der Unternehmensgruppe beziehen. Zusätzlich existiert, wie in Kapitel 2 beschrieben, für jede Unternehmensgruppe eine „Deutsche Entscheidungseinheit“, die für jede UG in der Datenbank bestimmt und gekenn-

zeichnet wird. Durch die UG werden somit sowohl die Stammdaten einer Unternehmensgruppe selbst als auch alle Beziehungen zwischen den gruppenzugehörigen Rechtlichen Einheiten und der Unternehmensgruppe an einer Stelle in der Datenbank des Unternehmensregisters jeweils für ein Berichtsjahr abgebildet.

## 6

### Manuelle Unternehmensgruppenbearbeitung

Bei der UGL-Verarbeitung (siehe Kapitel 4) erfolgt grundsätzlich bezüglich der Gruppenzusammenhänge keine Überprüfung der Qualität. Für die zahlreichen Nutzenden von Unternehmensgruppendaten sind jedoch qualitätsgesicherte Informationen über Gruppenstrukturen von großer Bedeutung, da fehlerhafte Informationen gravierende Auswirkungen auf die Nutzenden haben können (siehe Kapitel 3).

So wurde mit Einführung der neuen Aufgabe Profiling von Unternehmen im Jahr 2016 erkannt, dass für bestimmte Unternehmensgruppen Informationen zu den gruppenzugehörigen Rechtlichen Einheiten unvollständig oder fehlerhaft waren. Dies hatte zur Folge, dass vor Durchführung des Profiling umfangreiche Korrekturen und Nachqualifizierungen der Unternehmensgruppen erforderlich wurden. Neben dem Profiling zeigten sich bei immer mehr Nutzenden gestiegene Qualitätsanforderungen bei den Strukturdaten zu Unternehmensgruppen. Aufgrund des Umfangs und der hohen fachlichen Anforderungen wurde im statistischen Unternehmensregister 2023 ein neuer Arbeitsprozess eingeführt – die „Manuelle Unternehmensgruppenbearbeitung“.

Für die „Manuelle Unternehmensgruppenbearbeitung“ überprüft qualifiziertes Sachbearbeitungspersonal die

Gruppenzusammensetzung in der Datenbank des Unternehmensregisters. Die Statistischen Ämter der Länder führen die „Manuelle Unternehmensgruppenbearbeitung“ immer für ein bestimmtes Berichtsjahr  $t$  durch. Der Bearbeitungszeitraum erstreckt sich von Mitte Juni  $t+1$  bis Ende März  $t+2$ . Das Statistische Bundesamt gibt die Methodik vor, erstellt den Arbeits- und Zeitplan und koordiniert die Aktivitäten.

Über das sogenannte UG-Service-Desk bietet das Statistische Bundesamt neben regelmäßigen Schulungen und Austauschformaten zur Methodik auch bilaterale Unterstützung bei methodischen und technischen Fragen an.

Im Jahr 2022 gab es rund 292 000 in Deutschland tätige Unternehmensgruppen. Da der Zeitaufwand für die Bearbeitung einer Gruppe einschließlich der Korrektur in der Datenbank sehr hoch sein kann, ist die manuelle Bearbeitung nur für die wichtigsten Gruppen möglich. Hierfür erstellt das Statistische Bundesamt eine „UG-Select-Liste“, die – basierend auf bestimmten Auswahlkriterien – die bedeutendsten Unternehmensgruppen enthält. Diese für die „Manuelle Unternehmensgruppenbearbeitung“ vorgesehenen Top-Gruppen (rund 5 000 Gruppen beziehungsweise 1,7 % aller Gruppen) decken 60 % der Beschäftigten und 77 % der Umsätze aller Unternehmensgruppen ab. [Tabelle 1](#)

Zur Erzeugung der UG-Select-Liste werden verschiedene Kriterien und Nutzergruppen berücksichtigt. [Tabelle 2](#) Bei den Mengengerüsten werden keine Doppelzählungen vorgenommen, sondern eine relevante Gruppe wird stets beim erstgenannten Kriterium gezählt.

Eine Unternehmensgruppe wird immer in ihrer Gesamtheit von einem der Statistischen Ämter der Länder betreut, wobei sich die Zuständigkeit für die Unternehmensgruppe meist anhand des Sitz-Bundeslandes der „Deutschen Entscheidungseinheit“ ergibt.

**Tabelle 1**

Ökonomische Relevanz der Gruppen der UG-Select-Liste für das Berichtsjahr 2022

	Gesamtheit der Unternehmensgruppen			Umsatz			Beschäftigte		
	nationale Gruppen	multinational inländisch kontrolliert	multinational ausländisch kontrolliert	nationale Gruppen	multinational inländisch kontrolliert	multinational ausländisch kontrolliert	nationale Gruppen	multinational inländisch kontrolliert	multinational ausländisch kontrolliert
	Anteile in %								
Gesamtheit der Gruppen	80,0	5,0	15,0	21,0	54,0	25,0	44,0	38,0	18,0
Gruppen der UG-Select-Liste	0,6	0,6	0,5	7,0	49,3	20,3	14,7	32,2	13,0

Tabelle 2

## Auswahlkriterien zur Erstellung der UG-Select-Liste für das Berichtsjahr 2022

	Konkrete Nutzende/Kriterien	Anzahl der Unternehmensgruppen für Berichtsjahr 2022
nach Nutzendengruppen		
1. Fallgruppe	Gruppen relevant für das nationale manuelle Profiling	2 963
2. Fallgruppe	Fälle der Large Cases Unit (LCU)	166
3. Fallgruppe	Fälle der Europaweiten Manuellen Unternehmensgruppenbearbeitung (EMU)	27
4. Fallgruppe	Sonstige für das EuroGruppenRegister (EGR) relevante Gruppen	247
nach weiteren Kriterien		
5. Fallgruppe	Multinationale Gruppe mit mindestens 8 Rechtlichen Einheiten, mindestens 50 Millionen Euro Umsatz sowie mindestens 250 Beschäftigten	698
6. Fallgruppe	Nationale Unternehmensgruppen mit mindestens einem Konzernabschluss, mindestens 8 Rechtlichen Einheiten, mindestens 50 Millionen Euro Umsatz sowie mindestens 250 Beschäftigten	267
7. Fallgruppe	Nationale Unternehmensgruppen ohne Konzernabschluss, mindestens 8 Rechtlichen Einheiten, mindestens 50 Millionen Euro Umsatz sowie mindestens 250 Beschäftigten	400
8. Fallgruppe	Besonders volatile Gruppen im Zeitvergleich	0 <sup>1</sup>
	Insgesamt	4 768

1 Die Anzahl der Gruppen bei diesem Kriterium beträgt 0 aufgrund eines Methodikwechsels im Berichtsjahr 2022.

Zu Beginn der manuellen Bearbeitung einer Unternehmensgruppe ist die Identifizierung und Überprüfung des Gruppenoberhauptes und der „Deutschen Entscheidungseinheit“ erforderlich, da beide Einheiten für die Abgrenzung von Unternehmensgruppen eine wesentliche Rolle spielen (siehe Kapitel 2). Zur Verifizierung des Gruppenoberhauptes können verschiedene Quellen wie Konzernabschlüsse oder Handelsregistereinträge überprüft werden. Die „Deutsche Entscheidungseinheit“ kann durch die Außenpräsenz der Unternehmensgruppe, etwa durch das Impressum der Webseite oder als die höchste konsolidierende Einheit der Gruppe, die den Konzernabschluss erstellt, identifiziert werden (Redecker/Sturm, 2017; Andom, 2024).

Nach der korrekten Identifikation beider Einheiten erfolgt eine umfassende Analyse der Unternehmensgruppe. Dabei werden relevante Quellen herangezogen und die enthaltenen Informationen kritisch bewertet, um ein fundiertes Verständnis der Gruppe zu gewinnen. Zu den wesentlichen externen Informationsquellen zählen dabei Konzernabschlüsse, Webseiten der Unternehmensgruppen, Jahresabschlüsse, Handelsregistereinträge, Pressemitteilungen oder externe Datenbanken. Folgende Fragen sind zu untersuchen: Wie ist das Geschäftsmodell der Unternehmensgruppe in Deutschland sowie global aufgestellt? Verfügt die Unternehmensgruppe über einen Konzernabschluss? Liegt eine Beteiligungsliste vor? Welche Webseiten sind vorhanden?

Anhand dieser Informationen werden die Struktur beziehungsweise die Zusammensetzung der Unterneh-

mensgruppe analysiert. Ausgehend von der Struktur der Einheit Unternehmensgruppe in der Datenbank des Unternehmensregisters (siehe Kapitel 5) erfolgt ein technisch unterstützter Abgleich der gruppenzugehörigen deutschen Rechtlichen Einheiten aus der Datenbank mit den externen vorliegenden Informationen aus verschiedenen Quellen. Im Idealfall sind dies die Beteiligungslisten aus den Konzernabschlüssen (Rommelspacher, 2021; Redecker und andere, 2021).

Ziel des Abgleiches ist, alle deutschen gruppenzugehörigen Rechtlichen Einheiten einer Unternehmensgruppe zweifelsfrei zu identifizieren und diese der Unternehmensgruppe zuzuordnen. Dabei sind häufig komplexe und teilweise intransparente Konstellationen der Gruppenstrukturen festzustellen. Insbesondere die Abgrenzung zwischen Kapital- und Stimmrechtsanteilen gestaltet sich schwierig, da diese nicht identisch sein müssen. Auch das in der Praxis häufig vertretene Rechtskonstrukt von zusammengesetzten Personenhandelsgesellschaften (beispielsweise GmbH & Co. KG) führt bei einer Abweichung von Gesellschafter und Komplementär zu erschwerten Prüfkongstellationen ebenso wie vertragliche Sondervereinbarungen (Joint Ventures), die sich nicht in den Besitzanteilen widerspiegeln. Darüber hinaus liefern unterschiedliche Quellen teils widersprüchliche Informationen, was eine sorgfältige Abwägung und Bewertung erfordert.

Sofern eine Differenz zwischen dem Unternehmensregister und externen Quellen besteht, sind gegebenen-

falls identifizierte Fehler in den Unternehmensgruppenstrukturen in der Datenbank zu korrigieren. Wie bereits beschrieben, ist der Zeitaufwand für die manuelle Bearbeitung einer Gruppe einschließlich der Korrektur in der Datenbank relativ hoch. Daher kann nur die Bearbeitung der wichtigsten Gruppen erfolgen. Die Strukturen aller weiteren nicht manuell bearbeiteten Unternehmensgruppen können nicht flächendeckend überprüft werden. Trotzdem sind für diese Gruppen jährlich ihre jeweiligen UGn in der Datenbank des Unternehmensregisters auf das aktuelle Berichtsjahr zu aktualisieren. Hierfür wurde der Prozess der „Automatisierten Aktualisierung“ der UG entwickelt.

## 7

### Automatisierte Aktualisierung der UG

Die bei der jährlichen primären und sekundären UGL-Verarbeitung gewonnenen Informationen werden nicht automatisch in einem Schritt auf die Einheit UG in der Datenbank des Unternehmensregisters übertragen. Für alle im Zuge der „Manuellen Unternehmensgruppenbearbeitung“ überprüften Gruppen wird im Laufe des Bearbeitungszyklus die jeweilige Einheit UG in der Datenbank auf das aktuelle Berichtsjahr angepasst. Dies führt zu zeitlichen Mischzuständen in der Daten-

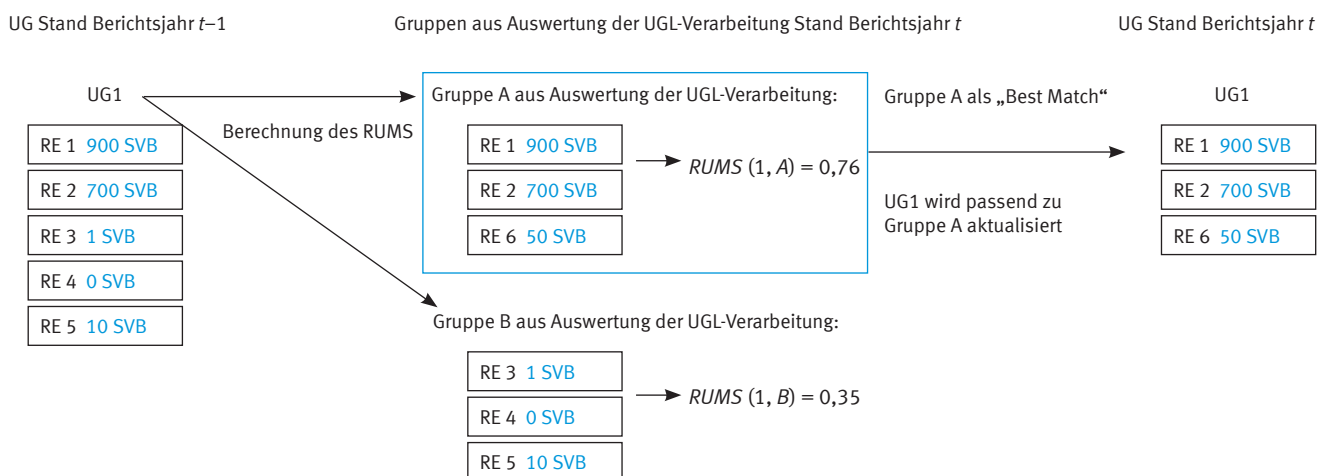
bank, da die UGn aller nicht überprüften Unternehmensgruppen sich noch auf dem Zeitstand des alten Berichtsjahres befinden. Um bis zum Ende eines jeweiligen Bearbeitungszyklus bei allen UGn in der Datenbank den gleichen Zeitstand zu erreichen, war eine Lösung für die Aktualisierung der nicht manuell geprüften Gruppen zu finden. Hierfür wurde das Verfahren der „Automatisierten Aktualisierung“ der UG entwickelt.

Jede Unternehmensgruppe, die über mehrere Berichtsjahre hinweg existiert, muss in der Datenbank als UG mit derselben UG-ID fortgeführt werden. Gleichzeitig müssen neue Gruppen als UGn hinzugefügt und die UGn nicht mehr existierender Gruppen entfernt werden. Der Abgleich der Unternehmensgruppenstrukturen zwischen den Berichtsjahren (UG  $t-1$  gegenüber UGL-Verarbeitung  $t$ ) ist methodisch anspruchsvoll: Welche Gruppe gilt als unverändert (identisch) und kann in dieser Form weitergeführt werden, welche Gruppe kann weitergeführt werden, muss aber verändert werden (ähnlich), welche Gruppe wird als neu angesehen und welche existiert nicht mehr? Aufgrund welcher Kriterien wird über die Kontinuität einer Unternehmensgruppe entschieden?

Um die Kontinuität zwischen den Unternehmensgruppen der verschiedenen Berichtsjahre zu ermitteln, hat das Statistische Bundesamt das Ähnlichkeitsmaß RUMS (Rommelspacher-Urban Metric for the Similarity of Enterprise Groups) entwickelt. Dieses Maß vergleicht

#### Grafik 7

#### Anwendung des RUMS zur Aktualisierung einer UG



RUMS: Rommelspacher-Urban Metric for the Similarity of Enterprise Groups ; UG: Unternehmensgruppe (eigene Einheit in der Datenbank des Unternehmensregisters); UGL: Unternehmensgruppenlieferung;  
RE: Rechtliche Einheit; SVB: Sozialversicherungspflichtig Beschäftigte



die Zusammensetzung der Unternehmensgruppen aus den verschiedenen Berichtsjahren anhand überschneidender Rechtlicher Einheiten sowie deren Beschäftigten und Umsätzen, um für jede UG in der Datenbank aus dem alten Berichtsjahr eine passende Gruppe des neuen Berichtsjahrs zu finden (Gruppe aus Auswertung der UGL-Verarbeitung). Das RUMS kann Werte zwischen 0 und 1 annehmen, wobei 1 eine vollständige Übereinstimmung darstellt und 0 keinerlei Ähnlichkeit bedeutet (Rommelspacher/Urban, 2023). [↗ Grafik 7](#)

Der höchste RUMS-Wert dient als Schlüssel, um die höchste Übereinstimmung für die Aktualisierung zu bestimmen. Das bedeutet, dass für jede UG in der Datenbank aus dem alten Berichtsjahr die Gruppe aus dem neuen Berichtsjahr (aus der Auswertung der UGL-Verarbeitung) mit dem höchsten RUMS-Wert als das beste Match ausgewählt wird. Sollten mehrere Gruppen im neuen Berichtsjahr ähnlich hohe RUMS-Werte aufweisen, wird die Gruppe mit dem höchsten Wert als das beste Match zur Fortführung der UG in der Datenbank genutzt (siehe Grafik 7). „Identische“ UGn mit einem RUMS-Wert von 1 werden unverändert fortgeführt, während UGn mit einem Wert von 0 stillgelegt werden (nicht mehr existierende Unternehmensgruppen). „Ähnliche“ UGn mit RUMS-Werten zwischen 0 und 1 werden in der Regel entsprechend der höchsten Übereinstimmung

in Bezug auf ihre Gruppenzusammensetzung aus Rechtlichen Einheiten aktualisiert. Zusätzlich müssen neue Unternehmensgruppen im aktuellen Berichtsjahr als neue UGn in der Datenbank angelegt werden.

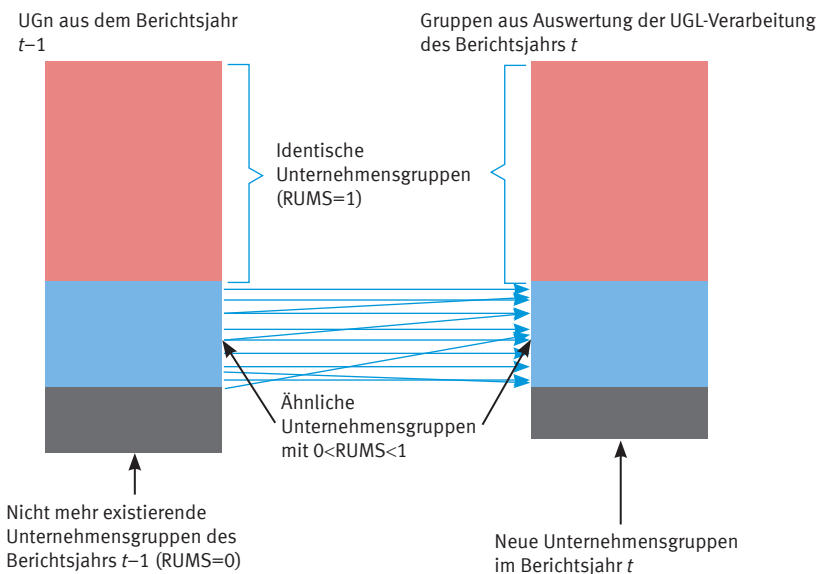
[↗ Grafik 8](#)

Nach erfolgter Anpassung aller Gruppenzusammensetzungen der UGn in der Datenbank auf das neue Berichtsjahr werden deren „Deutsche Entscheidungseinheiten“ automatisiert bestimmt. Zusätzlich werden Merkmale der Unternehmensgruppen für die UGn in der Datenbank maschinell festgesetzt, beispielsweise der Name der Unternehmensgruppe und deren Wirtschaftszweig.

Nach Abschluss der Aktualisierung sind alle UGn in der Datenbank im April  $t+2$  auf den Stand des Berichtsjahres  $t$  gebracht. [↗ Grafik 9](#)

#### Grafik 8

##### Aktualisierung der Unternehmensgruppen mit RUMS

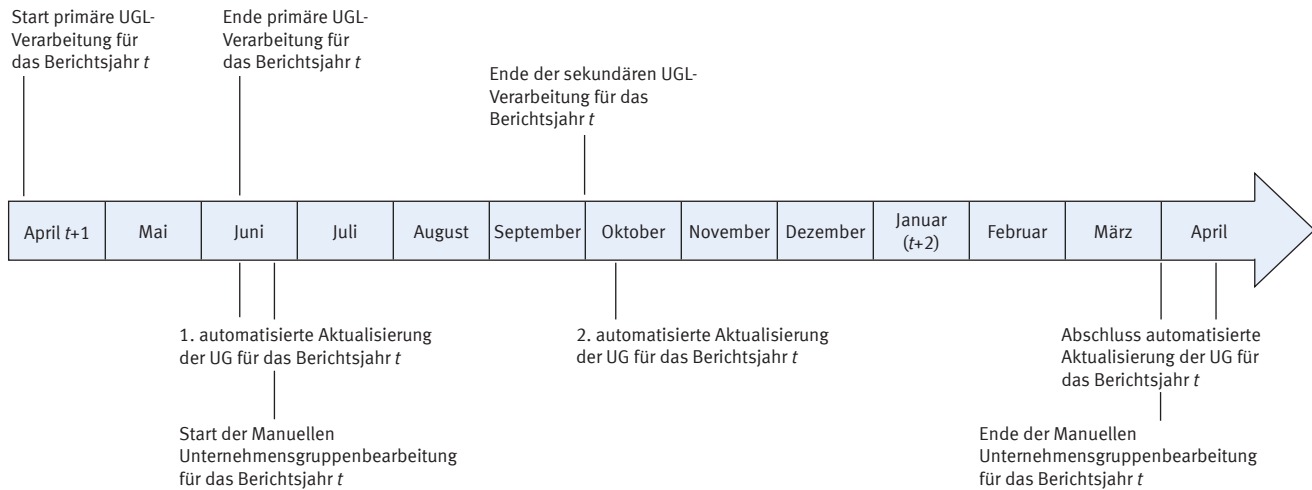


RUMS: Rommelspacher-Urban Metric for the Similarity of Enterprise Groups



**Grafik 9**

Zeitstrahl der Arbeiten zu Unternehmensgruppen im statistischen Unternehmensregister



## 8

### Fazit und Ausblick

Unternehmensgruppen haben großen Einfluss auf das Wirtschaftsgeschehen. Sie sind charakterisiert durch eine hohe Dynamik und starke Veränderungen im Zeitablauf. So sind Übernahmen, Verschmelzungen, Verschiebungen von Eigentum, Fusionen und Verkäufe an der Tagesordnung. Diese Veränderungen von Unternehmensgruppen nehmen Einfluss auf die Kennzahlen der Wirtschaft. Daher ist die Bereitstellung qualitativ hochwertiger Informationen über ihre Strukturen eine wichtige Aufgabe der Statistik.

Das statistische Unternehmensregister als Basis-Infrastruktur für die amtliche Wirtschaftsstatistik ist von ständigen Anpassungen und Weiterentwicklungen geprägt. Eine dieser Weiterentwicklungen stellt die Abbildung der Unternehmensgruppe als eigener Einheitentyp und dessen regelmäßige Aktualisierung in der Datenbank des Unternehmensregisters dar. Um die Strukturen der wichtigsten Unternehmensgruppen zu überprüfen, wurde im Jahr 2023 die „Manuelle Unternehmensgruppenbearbeitung“ eingeführt. Dadurch wurden wichtige Qualitätssteigerungen in der Bereitstellung von Unternehmensgruppendaten im Unternehmensregister erzielt, von denen die zahlreichen Nutzenden von Unternehmensgruppendaten profitieren. Die „Manuelle Unter-

nehmensgruppenbearbeitung“ befindet sich derzeit erst im zweiten Bearbeitungsjahr (Berichtsjahr 2023) und somit noch in der Aufbauphase. Perspektivisch soll die „Manuelle Unternehmensgruppenbearbeitung“ jährlich für die rund 5 000 größten Unternehmensgruppen durchgeführt werden. Zur Umsetzung dieser Aufgabe ist jedoch eine Ausweitung der derzeitigen Ressourcenausstattung nötig.

Durch die erzielten Weiterentwicklungen ist es nun für den Fachbereich Unternehmensregister möglich, ein Auswertungs- und Veröffentlichungskonzept zu entwickeln, um Informationen über Unternehmensgruppen bereitzustellen. [\[1\]](#)

## LITERATURVERZEICHNIS

---

Andom, Senait. *Profiling von Unternehmen zur Umsetzung des EU-Unternehmensbegriffs*. In: Statistisches Landesamt Baden-Württemberg. Monatshefte. Ausgabe 07/2020, Seite 18 ff. [Zugriff am 21. Januar 2025]. Verfügbar unter: [www.statistik-bw.de](http://www.statistik-bw.de).

Kleber, Birgit/Sturm, Roland/Tümmler, Thorsten. *Ergebnisse zu Unternehmensgruppen aus dem Unternehmensregister*. In: Wirtschaft und Statistik. Ausgabe 6/2010, Seite 527 ff.

Kruse, Hendrik W./Hieber, Florian/Limberg, Heiko/Zapf, Benedikt/Boddin, Dominik. *Außenhandelsaktive Unternehmen: neue Analysemöglichkeiten durch Mikrodatenverknüpfung*. In: WISTA Wirtschaft und Statistik. Ausgabe 5/2023, Seite 13 ff.

Redecker, Matthias/Rommelspacher, Simon/Sturm, Roland. *Profiling von Unternehmen im Echtbetrieb*. In: WISTA Wirtschaft und Statistik. Ausgabe 5/2021, Seite 99 ff.

Redecker, Matthias/Sturm, Roland. *Profiling von Unternehmen*. In: WISTA Wirtschaft und Statistik. Ausgabe 6/2017, Seite 9 ff.

Rommelspacher, Simon. *iProfAnT – Integrated Profiling Analysis Tool*. Vortrag und Aufsatz für die Veranstaltung „27th Meeting of the Wiesbaden Group on Business Registers“. 2021. [Zugriff am 21. Januar 2025]. Verfügbar unter: [inegi.org.mx](http://inegi.org.mx)

Rommelspacher, Simon/Urban, Adrian. *Similarity Metric for Comparison of Enterprise Groups*. In: 28th Meeting of the Wiesbaden Group. Session 3: Globalisation and Large Case Units. 2023. [Zugriff am 21. Januar 2025]. Verfügbar unter: [www.aanmelder.nl](http://www.aanmelder.nl)

Statistisches Bundesamt. *Qualitätsbericht: Finanzvermögen des öffentlichen Gesamthaushalts*. 2024a. [Zugriff am 21. Januar 2025]. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

Statistisches Bundesamt. *4 800 landwirtschaftliche Betriebe waren 2022 Teil einer Unternehmensgruppe*. Pressemitteilung Nr. 415 vom 6. November 2024. 2024b. [Zugriff am 21. Januar 2025]. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

## RECHTSGRUNDLAGEN

---

Durchführungsverordnung (EU) 2018/1874 der Kommission vom 29. November 2018 zu den für 2020 gemäß der Verordnung (EU) 2018/1091 des Parlaments und des Rates über integrierte Statistiken zu landwirtschaftlichen Betrieben und zur Aufhebung der Verordnungen (EG) Nr. 1166/2008 und (EU) Nr. 1337/2011 hinsichtlich der Liste der Variablen und ihrer Beschreibung zu liefernden Daten (Amtsblatt der EU Nr. L 306, Seite 14).

Durchführungsverordnung (EU) 2020/1197 der Kommission vom 30. Juli 2020 zur Festlegung technischer Spezifikationen und Einzelheiten nach der Verordnung (EU) 2019/2152 des Europäischen Parlaments und des Rates über europäische Unternehmensstatistiken, zur Aufhebung von zehn Rechtsakten im Bereich Unternehmensstatistiken (Amtsblatt der EU Nr. L 271, Seite 1).

Gesetz gegen Wettbewerbsbeschränkungen (GWB) in der Fassung der Bekanntmachung vom 26. Juni 2013 (BGBl. I Seite 1750, 3245), das zuletzt durch Artikel 6 des Gesetzes vom 5. Dezember 2024 (BGBl. I Nr. 400) geändert worden ist.

Gesetz über Agrarstatistiken (Agrarstatistikgesetz – AgrStatG) in der Fassung der Bekanntmachung vom 17. Dezember 2009 (BGBl. I Seite 3886), das zuletzt durch Artikel 1 des Gesetzes vom 14. November 2022 (BGBl. I Seite 2030) geändert worden ist.

Gesetz zur Prüfung von Daten multinationaler Unternehmensgruppen zur Sicherung der Qualität der Volkswirtschaftlichen Gesamtrechnungen und der Wirtschaftsstatistiken (Qualität-VGR und WS-Gesetz – QVWSG) vom 14. Juni 2021 (BGBl. I Seite 1751, 1757).

Verordnung (EU) 2018/1091 des Europäischen Parlaments und des Rates vom 18. Juli 2018 über integrierte Statistiken zu landwirtschaftlichen Betrieben und zur Aufhebung der Verordnung (EG) Nr. 1166/2008 und (EU) Nr. 1337/2011 (Amtsblatt der EU Nr. L 200, Seite 1).

Verordnung (EU) 2019/2152 des Europäischen Parlaments und des Rates vom 27. November 2019 über europäische Unternehmensstatistiken, zur Aufhebung von zehn Rechtsakten im Bereich Unternehmensstatistiken (Amtsblatt der EU Nr. L 327, Seite 1).

Verordnung (EG) Nr. 716/2007 des Europäischen Parlaments und des Rates vom 20. Juni 2007 zu gemeinschaftlichen Statistiken über die Struktur und Tätigkeit von Auslandsunternehmenseinheiten (Amtsblatt der EU Nr. L 171, Seite 17).

Kathrin Bank, Alexander Irmer,  
Ute Losch, Anna Sandig und  
Max Schöne

begleiteten in unterschiedlichen  
Funktionen im Referat „Freiwillige  
Haushaltserhebungen – Aufberei-  
tung, Auswertung und Analyse“  
des Statistischen Bundesamtes die  
Zeitverwendungserhebung 2022  
von der Konzeption bis zur  
Veröffentlichung.

# DIE ZEITVERWENDUNGSERHEBUNG 2022 – EIN ÜBERBLICK

Kathrin Bank, Alexander Irmer, Ute Losch, Anna Sandig,  
Max Schöne

🔗 **Schlüsselwörter:** Einsamkeit – Gender Care Gap – Zeitverwendung –  
unbezahlte Arbeit – HETUS

## ZUSAMMENFASSUNG

Die Zeitverwendungserhebung ist eine freiwillige Haushaltsbefragung, die alle zehn Jahre in Deutschland durchgeführt wird. Sie gibt Aufschluss darüber, für welche Aktivitäten verschiedene Bevölkerungsgruppen in unterschiedlichen Lebenslagen täglich ihre Zeit aufwenden. Darüber hinaus stellt die Zeitverwendungserhebung Fragen zum subjektiven Wohlbefinden und trägt dadurch zu einer umfassenden Messung des Wohlstands in Deutschland bei. Der Aufsatz beschreibt, wie die Datenerhebung der Zeitverwendungserhebung 2022 konzipiert und durchgeführt wurde sowie die anschließende Datenaufbereitung. Darüber hinaus wird das angewandte Verfahren zur Hochrechnung skizziert und beschrieben, in welcher Form die Daten bereitgestellt werden.

🔗 **Keywords:** loneliness – gender care gap – time use – unpaid work – HETUS

## ABSTRACT

*The time use survey is a voluntary household survey that is conducted in Germany every ten years. It provides an insight into how much time different population groups, experiencing different situations in life, spend on various activities throughout the day. Additionally, the time use survey asks respondents about their subjective well-being, and thereby helps to provide a comprehensive measurement of the well-being of the population. This article describes how data collection for the time use survey 2022 was designed and conducted, and discusses the subsequent processing of the data. Furthermore, it outlines the extrapolation method used and explains how the data are made available.*

## 1

### Einleitung

In der Zeitverwendungserhebung (ZVE) werden seit 1991/1992 etwa alle zehn Jahre private Haushalte in Deutschland befragt, wieviel Zeit die Personen eines Haushalts für ihre Aktivitäten aufwenden und wann sie im Tagesverlauf diesen Tätigkeiten nachgehen. Ziel der Erhebung ist, eine Datenbasis über die Zeitverwendung verschiedener Bevölkerungsgruppen in unterschiedlichen Lebenslagen zu liefern. Genutzt werden diese Daten insbesondere, um gesellschaftspolitische Maßnahmen zu evaluieren, und für Vergleiche zwischen den Mitgliedstaaten der Europäischen Union im Rahmen der harmonisierten europäischen Zeitverwendungserhebung (Harmonised European Time Use Survey – HETUS). Die Statistischen Ämter der Länder übernehmen in Deutschland die Datenerhebung.

Nachdem in den Jahren 1991/1992, 2001/2002 und 2012/2013 zunächst jeweils etwa 5 000 Haushalte befragt wurden, verdoppelte sich der Stichprobenumfang für die Zeitverwendungserhebung 2022 auf rund 10 000 Haushalte. Ausgewählt wurden diese Haushalte auf Basis einer Quotenstichprobe (siehe Kapitel 4). Die Befragung selbst bestand aus drei Teilen: dem Haushaltsfragebogen, der von einem erwachsenen Haushaltsmitglied ausgefüllt wurde, sowie dem Personenfragebogen und dem Tagebuch, welche jeweils von allen Haushaltsmitgliedern ab 10 Jahren ausgefüllt wurden. Im Tagebuch wurden jeweils an drei Tagen die Tagesabläufe über 24 Stunden dokumentiert. Die Zeitverwendungserhebung bildet somit vollständige Tagesabläufe der beteiligten Haushalte ab und ermöglicht dadurch Aussagen zu treffen zu deren Erwerbsarbeit, Bildung, Haushaltsführung, Kinderbetreuung, Freizeitaktivitäten und freiwilliges Engagement sowie die dazugehörigen Wegezeiten und das dafür genutzte Fortbewegungsmittel. Darüber hinaus erfasste die Zeitverwendungserhebung 2022 soziodemografische Merkmale und stellte Fragen zum subjektiven Wohlbefinden (zum Beispiel Zeitstress, Zeitwünsche, Einsamkeit). Besonderes Augenmerk lag auf der unbezahlten Arbeit, wie Kinderbetreuung, Hausarbeit oder Ehrenämter, da auf Basis der Zeitver-

wendungserhebung der sogenannte Gender Care Gap<sup>1</sup> berechnet wird, der in den Gleichstellungsbericht der Bundesregierung einfließt.

Neben der Möglichkeit, die drei Erhebungsteile auf Papier auszufüllen, kam bei der Zeitverwendungserhebung 2022 erstmals auch ein digitales Erhebungsinstrument zum Einsatz. So konnten Befragte über einen Browser mittels Web-App oder auf mobilen Endgeräten mithilfe einer mobilen App ihre Angaben eintragen (Bitzer und andere, 2022).

Vorherige Zeitverwendungserhebungen waren als Erhebungen für besondere Zwecke gemäß § 7 Bundesstatistikgesetz durchgeführt worden. Am 1. Juli 2021 ist das Gesetz über die statistische Erhebung der Zeitverwendung (Zeitverwendungserhebungsgesetz) als rechtliche Grundlage in Kraft getreten, das die künftige Durchführung der Zeitverwendungserhebung und die damit verbundene Möglichkeit zeitlicher Vergleiche sicherstellt.

Der vorliegende Aufsatz liefert einen allgemeinen Überblick über die Zeitverwendungserhebung 2022. In Kapitel 2 werden die unterschiedlichen Befragungsmodi und Erhebungsinstrumente vorgestellt. Kapitel 3 geht auf die in der Erhebung erfassten Aktivitäten ein. Kapitel 4 befasst sich mit der Stichprobenplanung und der Feldphase, während in Kapitel 5 die Erfassung, Plausibilisierung und Aufbereitung nach der Feldphase beschrieben werden. Kapitel 6 erklärt das angewandte Verfahren zur Hochrechnung und Kapitel 7 fasst die wesentlichen Ergebnisse zusammen. Kapitel 8 beschreibt die Bereitstellung der Daten, Kapitel 9 gibt einen Ausblick auf künftige Zeitverwendungserhebungen.

1 Der Gender Care Gap misst die Ungleichheit in der unbezahlten Arbeit zwischen Männern und Frauen, indem er die Differenz des Zeitaufwands für unbezahlte Arbeit von Frauen und Männern ins Verhältnis zum Zeitaufwand für unbezahlte Arbeit der Männer setzt (BMFSFJ, 2018).

## 2

### Befragungsmodi und Erhebungsinstrumente

Die Zeitverwendungserhebung 2022 wurde als Mixed-Mode-Erhebung konzipiert und durchgeführt. Dabei wurde ein Online-First-Ansatz verfolgt (Bitzer und andere, 2022): Die Teilnahme über den digitalen Meldeweg – das heißt mittels mobiler App oder Web-App – war die Regel, für die alternative Teilnahme mithilfe von Papierunterlagen mussten sich Teilnehmende aktiv entscheiden (opt out). Der digitale Meldeweg bietet den Vorteil, dass Fragebogen und Tagebuch flexibel über Smartphone, Tablet, Notebook, Laptop oder PC ausgefüllt werden können. Automatische Plausibilitätsprüfungen während des Ausfüllens reduzieren die Anzahl unvollständiger oder fehlerhafter Antworten. Außerdem entfällt die manuelle Erfassung von Erhebungsdaten, und damit auch potenzielle Fehler in der Erfassung. Durch die Möglichkeit der Teilnahme mithilfe von Papierunterlagen konnten auch Personen ohne Zugang zum Internet an der Erhebung teilnehmen. Die Durchführung der Erhebung im Mixed-Mode-Design erwies sich als vorteilhaft, da verschiedene Zielgruppen unterschiedlich stark durch die verschiedenen Erhebungsmodi angesprochen wurden, was sich positiv auf die Teilnahmebereitschaft auswirkte. Detaillierte Ausfüllhinweise unterstützten die Teilnehmenden, bei individuellen Fragen konnten sie mit dem zuständigen Statistischen Landesamt Kontakt aufnehmen.

Um eine möglichst hohe Datenqualität zu erreichen, wurden die Erhebungsinstrumente vor Beginn der Feldphase in Form von qualitativen Pretests erprobt. Testpersonen notierten dabei ihre Erfahrungen und mögliche Schwierigkeiten bei der Handhabung der Erhebungsinstrumente. Anschließend wurden individuelle Leitfadeninterviews mit den Testpersonen geführt. Die hieraus gewonnenen Informationen dienten dazu, die Erhebungsinstrumente gezielt zu optimieren.

Die Zeitverwendungserhebung besteht aus drei Erhebungsteilen:

- › Der [Haushaltsfragebogen](#) lieferte grundlegende Angaben zum Haushalt und wurde von nur einer Person im Haushalt ausgefüllt. Konkret wurden Informationen

über die Anzahl der im Haushalt lebenden Personen, deren Verwandtschaftsbeziehungen zueinander und soziodemografische Angaben wie Alter, Geschlecht oder Staatsangehörigkeit erfragt. Darüber hinaus waren Fragen zur Wohnsituation, zum Haushaltsnettoeinkommen und zu durch den Haushalt wahrgenommene Unterstützungs- sowie Pflegeleistungen enthalten. Ebenso gab es Fragen zur Zeit, die Kinder unter 10 Jahren in der Schule, in Betreuung oder bei außerschulischen Aktivitäten verbrachten.

- › Alle Personen des Haushalts ab 10 Jahren füllten den [Personenfragebogen](#) aus. Dort machten sie Angaben zu ihrem Familienstand, ihrer Schul- und Berufsbildung, Erwerbstätigkeit, ihrem Bruttoerwerbseinkommen und dem Aufwand für ehrenamtliche Tätigkeiten und Unterstützungsleistungen. Weitere Fragen betrafen den Gesundheitszustand, die Lebenszufriedenheit und das subjektive Zeitempfinden der teilnehmenden Personen.
- › Das [Tagebuch](#) stellte das Kernstück der Zeitverwendungserhebung dar und lieferte detaillierte Informationen über die Zeitverwendung der einzelnen Haushaltsmitglieder ab 10 Jahren. In der Papierform hatte es eine tabellarische Struktur, die den Befragten das Ausfüllen erleichterte. Das Tagebuch wurde an vorab festgelegten Tagen geführt – zwei zusammenhängende Wochentage und ein Tag des diesen folgenden Wochenendes. Ein sogenannter Anschreibetag begann dabei um 4:00 Uhr morgens und erstreckte sich über 24 Stunden. Teilnehmende notierten im 10-Minuten-Takt, welche Haupt- und Nebenaktivitäten sie ausführten, und gaben an, welche weiteren Personen gegebenenfalls anwesend waren. Personen, die auf dem digitalen Meldeweg an der Erhebung teilnahmen, konnten dafür mithilfe eines Suchalgorithmus oder einer direkten Auswahl aus 174 vordefinierten Aktivitäten wählen oder alternativ Klartextangaben machen. Personen, die mit einem Papierfragebogen teilnahmen, schrieben ihre Aktivitäten in eigenen Worten auf. Zusätzlich konnten Teilnehmende je 10-Minuten-Takt ankreuzen, ob sie für die Ausübung der Aktivität Informations- und Kommunikationstechnologien wie Smartphones oder Laptops genutzt haben. Der Ort, an dem sich die teilnehmende Person zum Zeitpunkt der Aktivität aufhielt, wurde ebenfalls notiert. Bewegte sich eine Person zwischen zwei Orten fort, war das dafür genutzte Verkehrsmittel zu notieren. Um zu veranschaulichen, in welchem Detailgrad das Tage-



## Grafik 1

### Musterseite des Tagebuchs der Zeitverwendungserhebung 2022

#### Musterseite von Vater Peter

Uhrzeit	Was haben Sie gemacht? (Hauptaktivität)	Was haben Sie gleichzeitig gemacht? (Nebenaktivität)	Dabei Computer/ Tablet/Smartphone genutzt?	Ort oder Verkehrsmittel	Wer war dabei?					
					Haushaltsmitglieder					
					Alleine	Partner/ Partnerin	Kinder unter 10 Jahren	Eltern (-teil)	Andere Haushalts- mitglieder	Andere bekannte Personen
16.00 – 16.10	im Büro gearbeitet		<input checked="" type="checkbox"/>	im Büro	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
16.10 – 16.20	zum Kindergarten gefahren	Radio gehört	<input type="checkbox"/>	Auto	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16.20 – 16.30			<input type="checkbox"/>		<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16.30 – 16.40	Sohn abgeholt	mit Erzieherin unterhalten	<input type="checkbox"/>	Kindergarten	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

buch auszufüllen war, enthielt es eine ausführliche Ausfüllanweisung sowie mehrere ausgefüllte Musterseiten. [➤ Grafik 1](#) Dies sollte es den Teilnehmenden ermöglichen, alle relevanten Komponenten ihres Tagesablaufes zu erfassen. Jeder der drei Anschreibetage endete mit einer subjektiven Einschätzung zum Tagesablauf. Sie lieferte Informationen darüber, welche Aktivitäten die größte Freude und welche gar keine Freude bereitet haben und wofür sich die Teilnehmenden mehr Zeit gewünscht hätten. Damit ergänzten sie die im Personenfragebogen gestellten Fragen zum subjektiven Zeitempfinden und ermöglichten es, Muster der persönlichen Zeitverwendung mit Blick auf das subjektive Wohlergehen einer Person auszuwerten.

Aufgrund der hierarchischen Struktur der Aktivitätenliste auf drei Ebenen sind Auswertungen entweder bezogen auf allgemeine Lebensbereiche oder für spezifischere Aktivitäten innerhalb dieses Bereichs möglich. Zum Beispiel könnte eine Auswertung die durchschnittliche Zeitverwendung für 3 Qualifikation/Bildung, für 31 Unterricht in der Schule oder für einzelne Schulfächer, zum Beispiel 314 Kunst/Musik enthalten. Darüber hinaus sind Zusammenfassungen von Aktivitäten auch aus unterschiedlichen Aktivitätsbereichen möglich. Auf der obersten Ebene in der Aktivitätenliste der Zeitverwendungserhebung 2022 existieren 9 einstellige Codes, auf den Ebenen darunter 48 zweistellige Codes und 174 dreistellige Codes.

## 3

### Klassifikation der Aktivitäten

Um die frei angeschriebenen Aktivitäten im Tagebuch statistisch auszuwerten, ist es erforderlich, diese zu systematisieren. Das zentrale Instrument dafür ist eine Klassifikation der Aktivitäten (Aktivitätenliste), welche auf der Activity Coding List des HETUS (Eurostat, 2020) beruht und um nationale Besonderheiten ergänzt wurde. Bei einer Teilnahme mit Papierunterlagen wurden die frei formulierten Aktivitäten aus dem Tagebuch während der Datenaufbereitung im Statistischen Bundesamt codiert (siehe Kapitel 5). Im Gegensatz dazu wählten die Online-Teilnehmenden die zutreffende Aktivität und den dazugehörigen Code weitgehend selbst aus der Aktivitätenliste aus.<sup>12</sup>

Mit der Zeitverwendungserhebung 2022 fanden im Vergleich zur Zeitverwendungserhebung 2012/2013 auch neue Codes Eingang in die Aktivitätenliste. Beispielsweise wurde der Bereich 52 Unterstützung für andere Haushalte in sechs dreistellige Codes unterteilt (Codes 521 bis 525 und 529), sodass die Art der Unterstützung nun differenzierter erfasst werden konnte. Dadurch war es nun möglich, etwa danach zu unterscheiden, ob andere Haushalte bei der Pflege von Erwachsenen unterstützt wurden oder bei der Kinderbetreuung. Die einzelne Auflistung von Arbeitsgemeinschaften in der Schule hingegen entfiel aufgrund geringer Fallzahlen und gehört in der Zeitverwendungserhebung 2022 zu Schulaktivitäten in der unterrichtsfreien Zeit (Code 330). Den Aktivitäten, die zu spezifisch oder zu allgemein sind, um zugeordnet werden zu können, standen für die jeweiligen Lebensbereiche die dreistelligen Aktivitäts-codes mit der Bezeichnung sonstige oder unbestimmte Aktivitäten zur Verfügung. Insgesamt hatte die im Vergleich zur Zeitverwendungserhebung 2012/2013 noch stärkere Angleichung an die Activity Coding List des HETUS den größten Einfluss auf Änderungen in der Aktivitätenliste.

<sup>12</sup> Eine Ausnahme stellte die Codierung von Wegezeiten dar, die auch bei einer Teilnahme über die mobile App beziehungsweise die Web-App erst im Statistischen Bundesamt erfolgte. Dieses Vorgehen sollte die Befragten entlasten.

Während einige Aktivitätscodes bereits als Online-/Offline-Aktivität oder als Aktivitäten alleine/in Gesellschaft definiert sind (zum Beispiel im Bereich 76 Spiele), konnten andere Aktivitäten durch die zusätzlichen Spalten im Aufbau des Tagebuchs (siehe Kapitel 2) weiter spezifiziert werden. Diese Angaben bieten Zusatzinformationen darüber, ob während der Aktivitäten ein Computer, Tablet oder Smartphone genutzt wurde, welches Verkehrsmittel bei den Wegezeiten genutzt wurde und ob weitere Personen während der Aktivität anwesend waren. Eine Neuerung der Zeitverwendungserhebung 2022 war, dass auch Eltern(teile) angegeben werden konnten.

## 4

### Stichprobenplanung

Für die Zeitverwendungserhebung 2022 wurden private Haushalte über unterschiedliche Wege, wie Aufrufe in Social Media, direkte Anschreiben der Statistischen Landesämter oder Flyer und Plakate, angeworben – Haushalte konnten sich freiwillig für die Teilnahme anmelden. Aus dem so gebildeten Pool teilnahmebereiter Haushalte wurde dann unter Beachtung eines Quotenplans zufällig eine Stichprobe gezogen. Gemeinschaftsunterkünfte, wie Seniorenwohnheime und die dort lebenden Personen, wurden nicht berücksichtigt.

Gemäß den HETUS-Richtlinien wurden als Stichprobeneinheiten nicht Personen, sondern Haushalte gezogen. Dadurch wurden auch Daten zu Beziehungen der Personen eines Haushalts untereinander gewonnen, was wiederum Analysen zur Aufteilung hauswirtschaftlicher Aufgaben unter den Haushaltsmitgliedern ermöglichte.

Um eine Nettostichprobe von etwa 10 000 Haushalten zu erreichen, wurde mit einer Bruttostichprobe von knapp 12 000 Haushalten geplant. Dies sollte den erwartbaren Ausfall von Haushalten im Verlauf der Befragung kompensieren. Die Auswahl der Haushalte erfolgte durch ein Quotenverfahren, welches sich an der Verteilung der Quotenmerkmale in der Bevölkerung gemäß Mikrozensus orientiert. Als Quotenmerkmale wurden dabei das Bundesland, der Haushaltstyp und die soziale Stellung der Haupteinkommensperson gewählt. Diese Merkmale sind besonders eng mit der Zeitverwendung der Bevölkerung verbunden. Haushalte mit Selbstständigen, Be-

amten/Beamtinnen oder sonstigen Nichterwerbstätigen als Haupteinkommensperson sowie Alleinerziehende, Paare mit Kind(ern) und sonstige Haushalte wurden überproportional in die Stichprobe einbezogen. Diese Disproportionalität gewährleistete, dass eine ausreichende Anzahl dieser kleineren oder besonders interessierenden Bevölkerungsgruppen teilnahm.

Für jedes Bundesland wurden zusätzlich Richtwerte für die Zahl der Haushalte je Gemeindegrößenklasse vorgegeben, um eine zum Mikrozensus möglichst ähnlich relative Verteilung der Haushalte nach Gemeindegrößenklasse zu erzielen. Dieses Vorgehen stellte sicher, dass Menschen sowohl aus ländlichen als auch aus dicht besiedelten Regionen in der Stichprobe vertreten waren.

Die hohe Belastung für teilnehmende Haushalte bei Haushaltserhebungen führt – ohne Teilnahmepflicht – erfahrungsgemäß zu geringen Teilnahme- und hohen Ausfallraten. Dies begründet für die (freiwillige) Zeitverwendungserhebung die Wahl eines Quotenverfahrens anstelle einer (einfachen) Zufallsstichprobe: Ein Haushalt, der sich mit Abgabe seiner Teilnahmeerklärung grundsätzlich bereit erklärt hat, an einer bestimmten Art von Erhebung teilzunehmen, verweigert im konkreten Fall mit geringerer Wahrscheinlichkeit die Teilnahme. Auch bricht er nach Beginn der Befragung diese seltener ab als ein zufällig ausgewählter Haushalt. Durch die Wahl eines Quotenverfahrens waren daher vergleichsweise weniger Haushalte anzuwerben, um eine Nettostichprobe einer gegebenen Größe zu erzielen. Auch die oben beschriebene Disproportionalität war einfacher zu erzielen.

Nachteilig ist allerdings, dass mit einem Quotenverfahren keine Wahrscheinlichkeit angegeben werden kann, mit der ein Haushalt in die Stichprobe gezogen wurde. Ein solches nicht probabilistisches Stichprobenverfahren eignet sich im engeren Sinne nicht für eine Extrapolation auf die Grundgesamtheit. Jedoch kann auch eine Quotenstichprobe, deren Quotierungsmerkmale in einem engen Zusammenhang zur Zeitverwendung stehen, im Rahmen der Hochrechnung an detaillierte Ergebnisse des Mikrozensus zur Gesamtbevölkerung angepasst werden (siehe Kapitel 6).

Da Jahreszeit, Wochentage und Feiertage die persönliche Zeitverwendung beeinflussen, wurden Daten kontinuierlich während des gesamten Jahres 2022 erhoben.

Die Stichprobenhaushalte wurden für diesen Zweck gleichmäßig auf die 365 Tage des Jahres verteilt und jedem Haushalt dabei drei Tage vorgegeben, an denen alle Personen ab 10 Jahren ein Tagebuch führen sollten. Dies gewährleistete, dass auch nicht alltägliche Ereignisse und ungewöhnliche Tagesabläufe erfasst wurden, wie etwa jene an Feiertagen, im Urlaub, bei Familienfesten und bei Krankheit.<sup>3</sup>

## 5

### Erfassung, Plausibilisierung und Aufbereitung

Die Anwerbung der Haushalte<sup>4</sup> sowie der Versand von Zugangskennungen<sup>5</sup> für die mobile App beziehungsweise die Web-App einerseits und von Erhebungsunterlagen auf Papier andererseits lag ebenso in der Verantwortung der Statistischen Ämter der Länder wie die anschließende Bearbeitung und die Unterstützung der teilnehmenden Haushalte bei Rückfragen. Teilnehmende Personen übermittelten die ausgefüllten Fragebogen an das für sie zuständige Statistische Landesamt. Dort wurden die Unterlagen mittels einer dialoggestützten Anwendung auf Vollständigkeit geprüft und in eine Datenbank überführt.<sup>6</sup> Unklarheiten wurden im direkten Kontakt mit den befragten Personen behoben. Anschließend ging die Verantwortung für diese Daten innerhalb der Anwendung an das Statistische Bundesamt über, das die Erhebungsdaten aufbereitete, das heißt plausibilisierte und codierte. Die beschriebenen Arbeiten liefen kontinuierlich während des gesamten Erhebungsjahres.

Zusätzliche Filterführungen in der Anwendung, die die Eingabe einiger unplausibler Datenkonstellationen verhinderten, unterstützten die Erfassung sowie die Plausibilisierung der Daten. So waren etwa Eingaben zu Kindern unter 10 Jahren nicht möglich, wenn zuvor eingegeben wurde, dass im Haushalt kein Kind unter 10 Jahren lebt. Darüber hinaus gab es Fehlermeldungen und Warnhinweise, die auf (etwaige) unplausible Datenkonstellationen hinwiesen, wie die Angabe im Haushaltsfragebogen, eine Person sei in Deutschland geboren, bei gleichzeitiger Nennung ihres Zuzugsjahres nach Deutschland.

Die aufwendige Codierung der frei beschriebenen Aktivitäten im Papier-Tagebuch<sup>7</sup> wurde in der Anwendung zwar technisch unterstützt, trotzdem war für eine korrekte Codierung ein Interpretations- und Abstraktionsprozess notwendig, da sich der zu vergebende Code je nach Lebenssituation der tagebuchführenden Person unterschied. Dies lässt sich am Beispiel der Eintragung „mit dem Kind gespielt“ illustrieren: Handelte es sich bei der ausfüllenden Person um ein Elternteil oder ein deutlich älteres Geschwisterkind, war die Angabe unter „473 Spielen und Sport mit Kindern des Haushalts“ aus dem Bereich „Kinderbetreuung“ zu erfassen, während der Code „761 Gesellschaftsspiele und Spiele miteinander“ zu vergeben war, wenn ein Geschwisterkind im ähnlichen Alter diese Eintragung vorgenommen hatte. Nahm hingegen etwa eine Tagesmutter, welche das Kind betreute, die Eintragung vor, war die Codierung „210 Haupterwerbstätigkeit“ richtig. Bei der Codierung der Tagebucheinträge waren also Informationen aus allen Tagebuchspalten und gegebenenfalls den anderen Erhebungsteilen mit einzubeziehen.

Neben der Erfassung, Plausibilisierung und Aufbereitung mithilfe der Anwendung erfolgten maschinelle Korrekturen im Statistikprogramm SAS. Ziel der maschinellen Korrekturen war es, weitere inhaltliche Inkonsistenzen zu beseitigen und internationale (HETUS-)Anforderungen umzusetzen. Dies umfasste unter anderem den Tausch von Haupt- und Nebenaktivitäten in bestimmten

---

3 Zusätzlich konnten Befragte für jeden Tag selbst einschätzen, ob es sich um einen gewöhnlichen oder ungewöhnlichen Tag handelte.

4 Das Statistische Bundesamt unterstützte mit bundesweiter Werbung mittels Influencern über Social Media. Hierdurch sollten besonders schwer erreichbare Zielgruppen, wie etwa Selbstständige, Alleinerziehende, Großfamilien, Haushalte mit Migrationshintergrund und auch einkommensschwache Bevölkerungsgruppen erreicht werden.

5 In der Zeitverwendungserhebung 2022 war es erstmals möglich, über den Browser oder das mobile Endgerät teilzunehmen (siehe Kapitel 2).

6 Bei Teilnahme über die mobile App beziehungsweise die Web-App wurden Erhebungsauskünfte dem Erfassungsprogramm automatisiert über eine Schnittstelle zugeführt. Diese Daten durchliefen durch die Art der Nutzungsführung und durch Nutzungshinweise direkt bei Eingabe gewisse Plausibilitätsprüfungen und es konnten bereits zu diesem Zeitpunkt einige Inkonsistenzen verhindert werden (siehe Kapitel 2).

---

7 Auch in der mobilen App beziehungsweise in der Web-App war es möglich, Freitexte einzutragen, allerdings stellte dies eher die Ausnahme dar und die Teilnehmenden wählten zumeist vordefinierte Aktivitäten aus (siehe Kapitel 3).

Fällen<sup>18</sup> und das Löschen von unplausiblen Nebenaktivitäten. Darüber hinaus wurden Haushalts- und Personentypisierungen vorgenommen, die sowohl für die Berechnung der Hochrechnungsfaktoren als auch für spätere Auswertungen relevant waren.

Nach Abschluss der Aufbereitung enthielten die Daten der Zeitverwendungserhebung 2022 Informationen zu 9774 Haushalten mit 19526 Personen ab 10 Jahren. Erfasst und codiert waren die Aktivitäten von insgesamt 57 769 Tagebuchtagen.

## 6

### Hochrechnung

Die Stichprobe der Zeitverwendungserhebung 2022 gestaltete sich in mehrfacher Hinsicht nicht proportional zur Bevölkerung in Deutschland (siehe Kapitel 4). Dies wurde zum einen durch eine disproportionalen Stichprobenstruktur verursacht. Beispielsweise waren größere Haushalte überproportional in der Zeitverwendungserhebung vertreten. Zum anderen gibt es weitere Merkmale, die zwar bei der Stichprobenplanung nicht einbezogen werden konnten, aber mit großer Wahrscheinlichkeit das Zeitverhaltensverhalten der Befragten beeinflussten. Dazu zählen unter anderem das Geschlecht und das Alter der Haushaltsmitglieder. Ungewichtete Auswertungen hätten daher zu verzerrten Ergebnissen geführt. So lag die ungewichtete Haushaltsgröße in der Zeitverwendungserhebung bei 2,5 Personen, gewichtet bei 2,0 Personen; ohne die Gewichtung wäre sie also überschätzt worden. Durch die Verwendung von Hochrechnungsfaktoren wurde die Stichprobe an die Grundgesamtheit angepasst und die Disproportionalität korrigiert.

Für die Zeitverwendungserhebung 2022 wurden Hochrechnungsfaktoren für Haushalte, Personen und Tagebuchtage erstellt, um die Stichprobe an den Mikrozensus 2022 anzupassen. Die Hochrechnung wurde wie in der Zeitverwendungserhebung 2012/2013 mithilfe von SAS nach dem Verfahren der verallgemeinerten Regres-

sionsschätzung (Generalized REGression, kurz GREG) durchgeführt. Dieses Verfahren besteht aus zwei Schritten. Im ersten Schritt wurden für jede Quotierungszelle, das heißt für jede sinnvolle Kombination von Quotenmerkmalen, freie Hochrechnungsfaktoren als Quotient aus hochgerechneten Mikrozensushaushalten und an der Zeitverwendungserhebung teilnehmenden Haushalten gebildet. Im zweiten Schritt folgte die Anpassung an weitere relevante Merkmale mittels einer Regressions-schätzung. Diese Schätzung gewährleistet, dass die endgültigen Hochrechnungsfaktoren die Verteilung der Anpassungsmerkmale wiedergeben und gleichzeitig so wenig wie möglich von den freien Faktoren des ersten Schritts abweichen.

Da bei der Zeitverwendungserhebung 2022 einige Quotierungszellen nur sehr gering besetzt sind, wurden für die freie Hochrechnung ausschließlich die Quotenmerkmale „Haushaltstyp“ und „Soziale Stellung der Haupteinkommensperson“ einbezogen. Zusätzlich zu den Quotenmerkmalen wurden einige Merkmale bei der Anpassung verwendet, die einen großen Einfluss auf die Zeitverwendung haben und deren Verteilung in der Stichprobe sich von der Verteilung im Mikrozensus unterscheidet. Das waren unter anderem das Merkmal „früheres Bundesgebiet/neue Länder“ sowie Gemeindegrößenklassen, Geschlecht, Alter, Wochentage und Erhebungsquartale.

## 7

### Ausgewählte Ergebnisse

Nach den Ergebnissen der Zeitverwendungserhebung 2022 liegt die Hauptlast der unbezahlten Arbeit in Deutschland noch immer vor allem bei Frauen. Unbezahlte Arbeit setzt sich dabei aus Haushaltsführung, Kinderbetreuung, der Pflege von Angehörigen, freiwilligem und ehrenamtlichem Engagement sowie der Unterstützung haushaltsfremder Personen zusammen. Mit durchschnittlich knapp 30 Stunden leisteten 2022 Frauen ab 18 Jahren in allen Altersgruppen rund 9 Stunden mehr unbezahlte Arbeit je Woche als Männer derselben Altersgruppe. Der daraus berechnete Gender Care Gap liegt damit bei 44,3%. In der vorangegangenen Zeitverwendungserhebung 2012/2013 lag dieser noch bei 52,4%. Somit ist die Lücke beim Zeitaufwand für

---

<sup>18</sup> Die HETUS 2018 Guidelines geben für bestimmte Aktivitätskombinationen vor, welche Aktivität als Haupt- und welche als Nebenaktivität codiert werden muss: Passive Aktivitäten (zum Beispiel „zu Besuch sein/Besuch haben“) sollen stets Nebenaktivitäten sein und eine zugleich genannte andere (aktive) Aktivität (zum Beispiel Essen und Trinken) die Hauptaktivität.



unbezahlte Arbeit zwischen Frauen und Männern zwar kleiner geworden, jedoch weiterhin beträchtlich. Werden bezahlte und unbezahlte Arbeit zusammen betrachtet, arbeiteten Frauen im Jahr 2022 mit durchschnittlich rund 46 Stunden knapp 1,5 Stunden mehr je Woche als Männer. 2012/2013 hatte dieser Unterschied nur etwa 1 Stunde betragen. Der Umfang an insgesamt geleisteter Arbeit<sup>9</sup> von Erwachsenen im Erwerbsalter von 18 bis 64 Jahren hängt stark davon ab, ob sie in einem Haushalt mit Kind(ern) oder ohne Kind leben. Alleinerziehende und Elternteile in Paarhaushalten mit Kind(ern) arbeiteten durchschnittlich gut 57,5 Stunden je Woche und damit 11 Stunden je Woche mehr als Personen derselben Altersgruppe, die in Haushalten ohne Kind leben. Jede vierte erwerbstätige Mutter empfand dabei ihre Zeit für Erwerbsarbeit als zu knapp bemessen, während jeder vierte erwerbstätige Vater fand, dass er zu viel Zeit im Job verbringt.

In der Zeitverwendungserhebung 2022 wurde erstmals das Thema „Einsamkeit“ abgefragt. Dabei stellte sich heraus, dass sich jede sechste Person ab 10 Jahren (16 %) in Deutschland oft einsam fühlt. Junge Erwachsene von 18 bis 29 Jahren sind dabei am stärksten von Einsamkeit betroffen (24 %). Am wenigsten ausgeprägt ist das Gefühl der Einsamkeit bei Personen ab 65 Jahren (10 %)<sup>10</sup>. Weitere Ergebnisse stehen im Internetangebot des Statistischen Bundesamtes auf der Themenseite [Zeitverwendung](#) zur Verfügung.

## 8

### Bereitstellung der Daten

Für Institutionen unabhängiger wissenschaftlicher Forschung im Inland stellen die Forschungsdatenzentren des Bundes und der Länder anonymisierte Mikrodaten aus der Zeitverwendungserhebung 2022 für die kontrollierte Datenfernverarbeitung sowie für Gastwissenschaftsarbeitsplätze an ihren Standorten bereit. Weiterhin steht für jene Institutionen ein Scientific-Use-File (faktisch anonymisierte Mikrodaten) bereit. Darüber hinaus wird für Personen außerhalb inländischer wissen-

schaftlicher Einrichtungen ein Public-Use-File (absolut anonymisierte Mikrodaten) angeboten.

Scientific-Use-Files und Public-Use-Files haben stets zwei sehr gegensätzliche Anforderungen zu erfüllen. Zum einen sollen sie ein hohes Analysepotenzial bieten, zum anderen müssen sie den strengen Vorgaben der faktischen beziehungsweise absoluten Anonymität genügen. Scientific-Use-File und Public-Use-File der Zeitverwendungserhebung 2022 liefern daher fachlich tief gegliederte Auswertungsmöglichkeiten zulasten der Regionalinformationen. Zudem wird der Schutz der Daten dadurch gewährleistet, dass Substichproben aus den originalen Datensätzen der Zeitverwendungserhebung 2022 gezogen wurden. Der Scientific-Use-File besteht aus einer 95-prozentigen Zufallsstichprobe, während für den Public-Use-File eine 80-prozentige Zufallsstichprobe gezogen wurde. Die Datensätze und weitere Informationen zur Anonymisierung sind auf der [Website der Forschungsdatenzentren zur Zeitverwendungserhebung](#) zu finden.

## 9

### Ausblick

Die Einzeldaten der Zeitverwendungserhebung 2022 gehen in den HETUS der Datenrunde 2020<sup>11</sup> ein, dadurch lassen sich mit den ZVE-Daten auch Vergleiche auf europäischer Ebene durchführen. Hierfür wird Eurostat anonymisierte Mikrodaten für die Wissenschaft bereitstellen.

Aufgrund wachsender Datenbedarfe etwa im Hinblick auf eine bessere Messung des Wohlbefindens der Bevölkerung oder im Bereich der geschlechtsspezifischen Aufgabenteilung in Familie und Gesellschaft gewinnen Zeitverwendungserhebungen nicht nur in Deutschland zunehmend an Bedeutung. Daher wurden künftige europäische Time-Use-Surveys auf eine gesetzliche Grundlage gestellt: Die Verordnung (EU) 2019/1700 des


---

9 Die Summe aus bezahlter und unbezahlter Arbeit.

10 Dabei ist jedoch zu beachten, dass Hochbetagte in der Stichprobe unterrepräsentiert sind und Personen in Alten- und Pflegeheimen nicht befragt wurden.

---

11 Bis zur Datenrunde 2020 erfolgten beziehungsweise erfolgen HETUS-Datenlieferungen an Eurostat aus nationalen Zeitverwendungserhebungen auf Basis eines Gentlemen's Agreement. Die HETUS-Datenerhebungsjahre stellten nur Empfehlungen dar. Die in den HETUS 2020 eingehenden Zeitverwendungserhebungen europäischer Staaten wurden wegen der Corona-Pandemie meistens in den Jahren 2022 oder 2023 durchgeführt.

Europäischen Parlamentes und des Rates (kurz: IESS-Verordnung) regelt auch das Themengebiet Zeitverwendung. Nach Artikel 3 Absatz 2 dieser Verordnung ist die Datenerhebung für den Bereich Zeitverwendung jedoch fakultativ. Ein Datenerhebungsjahr für den ersten europäischen Time-Use-Survey unter der IESS-Verordnung ist noch nicht gesetzlich festgelegt<sup>12</sup>. Das Zeitverwendungserhebungsgesetz sieht aufgrund der in § 5 geregelten 10-jährlichen Erhebungsperiodizität vor, dass die nächste Zeitverwendungserhebung in Deutschland im Jahr 2032 stattfinden wird. 

---

12 Nach dem Entwurf der IESS-Verordnung ist der nächste europäische Time-Use-Survey für das Jahr 2032 vorgesehen.



### LITERATURVERZEICHNIS

---

Bitzer, Eric/van Hoorde, Joël/Schöne, Max. *Erstmaliger App-Einsatz für die Zeitverwendungserhebung 2022*. In: WISTA Wirtschaft und Statistik. Ausgabe 4/2022, Seite 63 ff.

Bundesministerium für Familie, Senioren, Frauen und Jugend (BMFSFJ). *Zweiter Gleichstellungsbericht der Bundesregierung*. 2. Auflage. Berlin 2018. [Zugriff am 8. Januar 2025]. Verfügbar unter: [www.bmfsfj.de](http://www.bmfsfj.de)

Statistisches Amt der Europäischen Union (Eurostat). *Harmonised European Time Use Surveys (HETUS) 2018 Guidelines. Re-edition*. Luxemburg 2020. [Zugriff am 8. Januar 2025]. Verfügbar unter: [ec.europa.eu](http://ec.europa.eu)

Statistisches Amt der Europäischen Union (Eurostat). *Harmonisierte Europäische Zeitverwendungserhebungen (HETUS) – Informationen zu den Daten*. 2024. [Zugriff am 20. September 2024]. Verfügbar unter: [ec.europa.eu](http://ec.europa.eu)

### RECHTSGRUNDLAGEN

---

Gesetz über die statistische Erhebung der Zeitverwendung (Zeitverwendungserhebungsgesetz – ZVEG) vom 2. Juni 2021 (BGBl. I Seite 1293).

Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) in der Fassung der Bekanntmachung vom 20. Oktober 2016 (BGBl. I Seite 2394), das zuletzt durch Artikel 14 des Gesetzes vom 8. Mai 2024 (BGBl. 2024 I Nr. 152) geändert worden ist.

Verordnung (EU) 2019/1700 des Europäischen Parlamentes und des Rates zur Schaffung eines gemeinsamen Rahmens für europäische Statistiken über Personen und Haushalte auf der Grundlage von Einzeldaten aus Stichprobenerhebungen, zur Änderung der Verordnungen (EG) Nr. 808/2004, (EG) Nr. 452/2008 und (EG) Nr. 1338/2008 des Europäischen Parlaments und des Rates sowie zur Aufhebung der Verordnung (EG) Nr. 1177/2003 des Europäischen Parlaments und des Rates und der Verordnung (EG) Nr. 577/98 des Rates (Amtsblatt der EU Nr. L 261, Seite I/1).



**Kai Engemann**

studierte Empirische Demokratieforschung (M. A.) und ist seit 2022 wissenschaftlicher Mitarbeiter im Referat „Gebäude und Wohnungen“ des Statistischen Bundesamtes. Er ist für die Eignungsprüfungen von Datenquellen, einschließlich der Daten der öffentlichen Verwaltung, zuständig und erarbeitet Nutzungspotenziale des Einrichtungsregisters.



**Irina Meinke**

studierte Politikwissenschaft mit Schwerpunkt Empirie sowie Technologie- und Innovationsmanagement und ist seit 2013 im Statistischen Bundesamt tätig. Sie hat das Qualitätsmanagementsystem für den Zensus 2022 als stellvertretende Projektleitung mitverantwortet. Derzeit leitet sie das Teilprojekt „Einrichtungsregister“, das Daten zu allen Gemeinschaftsunterkünften und Wohnheimen in Deutschland führt.

# METHODIK UND POTENZIALE DES EINRICHTUNGSREGISTERS

Kai Engemann, Irina Meinke

➤ **Schlüsselwörter:** Statistikregister – Einrichtungen – Gemeinschaftsunterkünfte – Wohnheime – Verwaltungsdatennutzung

## ZUSAMMENFASSUNG

Ein zentraler Datenbestand zu allen Einrichtungen in Deutschland, wie er in Form des Einrichtungsregisters aufgebaut werden soll, existiert bisher nicht. Ziel ist, diesen Datenbestand nicht nur für den Zensus und darüber hinaus für Bundes- und Landesstatistiken nutzbar zu machen, sondern auch als relevante Datengrundlage für die Öffentlichkeit, Wissenschaft, Politik und Verwaltung zu etablieren. Das Einrichtungsregister kann einen wertvollen Beitrag leisten, um übergreifende sozialpolitische oder planerische Entscheidungen in Bund, Ländern und Gemeinden zu unterstützen: Wie viele Einrichtungen einer bestimmten Bereichsart gibt es in Deutschland, in einem Bundesland oder in einer Gemeinde? Wo kann der Bau von bestimmten Einrichtungen geplant werden?

➤ **Keywords:** statistical register – institutions – collective living quarters – residential homes – use of administrative data

## ABSTRACT

*A centralised database of all institutions and facilities in Germany, as planned in the form of the “register of institutions and collective living quarters”, does not currently exist. The aim is not only to utilise this collection of data for the census and for official federal and Land statistics but also to establish it as a valuable data resource for the public, the scientific community, policy-makers and public administration. The register can make an important contribution to supporting overarching social policy or planning decisions at the federal, Land and municipality levels. For example, the register data can be analysed to identify how many facilities of a certain type there are in Germany, or in a particular Land or municipality, or where the construction of certain facilities can be planned.*

### 1

## Einleitung

---

Derzeit konzipieren die Statistischen Ämter des Bundes und der Länder ein Einrichtungsregister, das in den nächsten Jahren aufgebaut werden soll. Es wird Informationen zu allen Einrichtungen in Deutschland umfassen. Eine Einrichtung ist dadurch gekennzeichnet, dass in ihr bestimmte Personengruppen mit ähnlicher Lebenslage wohnen. Dies können beispielsweise Pflegebedürftige, Ältere, Geflüchtete oder Personen in Ausbildung oder im Studium sein. Im Einrichtungsregister werden nur Merkmale zu den Einrichtungen selbst und keine Informationen zu den in den Einrichtungen lebenden Personen erfasst (Kapitel 2 und 3). Aufbau und Pflege des Einrichtungsregisters regelt derzeit § 10 Registerzensuserprobungsgesetz. Zusätzlich benötigte Regelungen sind in Folgegesetzen auszuarbeiten.

In den Zensusrunden 2011 und 2022 wurde bereits ein Datenbestand zu Einrichtungen mithilfe von umfassenden manuellen Recherchen der Statistischen Ämter der Länder und durch Befragungen von Trägern und Leitungen der Einrichtungen für den jeweiligen Zensusstichtag aufgebaut. Dieser sogenannte Sonderanschriftenbestand diente als Grundlage für die Erhebung an Gemeinschaftsunterkünften und Wohnheimen sowie für die Haushaltegenerierung im Zensus 2011 und 2022. Im Zuge der Weiterentwicklung der Zensusmethodik wird mit dem Einrichtungsregister ein ähnlicher Datenbestand zu allen Einrichtungen aufgebaut. Dabei wird die Art und Weise, wie die Pflege des Datenbestands erfolgt, weiterentwickelt. Darüber hinaus werden die Auswertungs- und Nutzungspotenziale ausgeweitet.

Ziel des Einrichtungsregisters ist es, einen dauerhaften Datenbestand über Einrichtungen als Statistikregister aufzubauen, der in regelmäßigen Erhebungszyklen aktualisiert, qualitätsgesichert und historisiert wird. Dies ist in Deutschland einmalig. Für die regelmäßige Pflege des Einrichtungsregisters werden vorrangig bereits bestehende Datenquellen verwendet (Kapitel 4), die systematisch im Hinblick auf ihre Qualität evaluiert werden (Kapitel 5).

Eine weitere Zielsetzung des Einrichtungsregisters ist, dass es nicht ausschließlich dem Zensus, sondern darüber hinaus auch anderen Bundes- und Landesstatistiken

zur Verfügung stehen soll. Das Einrichtungsregister hat vielseitige Auswertungspotenziale, die sich unmittelbar aus seinem Datenbestand oder durch Verknüpfungen mit anderen Datenbeständen ergeben (Kapitel 6).

### 2

## Bereichsarten

---

Das Einrichtungsregister wird sämtliche Einrichtungen in Deutschland umfassen, beispielsweise solche für pflegebedürftige, ältere, geflüchtete oder junge Menschen. Unter dem Begriff „Einrichtungen“ werden in diesem Zusammenhang alle Wohnheime sowie Gemeinschaftsunterkünfte verstanden.

**Wohnheime** erfüllen die Wohnbedürfnisse bestimmter Bevölkerungskreise. Ein Beispiel für Wohnheime, die im Einrichtungsregister geführt werden, sind Studierendenwohnheime. **Gemeinschaftsunterkünfte** erfüllen ebenso die Wohnbedürfnisse bestimmter Bevölkerungskreise, gewährleisten zusätzlich aber auch ein gewisses Maß an Betreuung und/oder Versorgung. Beispiele für Gemeinschaftsunterkünfte sind Einrichtungen für ältere oder pflegebedürftige Menschen, Einrichtungen für Kinder und Jugendliche oder Flüchtlingsunterkünfte. Zu den Gemeinschaftsunterkünften gehören auch Justizvollzugsanstalten und Notunterkünfte, beispielsweise für Wohnungslose.

Unter einem anderen Begriffsverständnis können auch Kindertagesstätten, Schulen, Sportplätze oder Parks als (öffentliche) Einrichtungen bezeichnet werden. Diese Einheiten werden im Einrichtungsregister jedoch nicht erfasst und gehören nicht zur Grundgesamtheit, da hier grundsätzlich keine Personen wohnhaft gemeldet sein können. Für die im Einrichtungsregister erfassten Einrichtungen hingegen gilt, dass die Personen, die in diesen Einrichtungen leben, nach dem Bundesmeldegesetz in der Einrichtung mit Haupt- oder Nebenwohnsitz meldepflichtig wohnhaft gemeldet sein können.

### ➤ Beispiele für reguläre und besondere Meldepflicht

Studierende, die in einem Studierendenwohnheim wohnen, oder Geflüchtete, die in einer Flüchtlingsunterkunft wohnen, sind laut § 17 Bundesmeldegesetz dazu aufgefordert, sich zwei Wochen nach Einzug beim örtlichen Meldeamt an der Anschrift der Unterkunft wohnhaft zu melden. Hier greift die [reguläre Meldepflicht](#).

In Einrichtungen für Menschen mit Behinderung oder in Einrichtungen für ältere oder pflegebedürftige Menschen gilt eine [besondere Meldepflicht](#). Laut § 32 Bundesmeldegesetz haben sich die dort untergebrachten Personen beim örtlichen Meldeamt in der Einrichtung wohnhaft zu melden, wenn diese Personen keine andere Wohnung mehr haben und länger als drei Monate in der Einrichtung leben. Sollten sich die Personen nicht selbst anmelden können, so erledigt das die Einrichtungsleitung für sie.

Insgesamt werden 15 Bereichsarten im Einrichtungsregister erfasst, davon werden drei als Wohnheime und zwölf als Gemeinschaftsunterkünfte gelistet.

### ➤ Übersicht 1

#### Übersicht 1

##### Bereichsarten im Einrichtungsregister

<a href="#">Wohnheime im Einrichtungsregister</a>
Studierendenwohnheim
Arbeitendenwohnheim
Sonstiges Wohnheim
<a href="#">Gemeinschaftsunterkünfte im Einrichtungsregister</a>
Internat
Kloster
Mutter/Vater-Kind-Einrichtung
Gemeinschaftsunterkunft von Flüchtlingen
(Not-)Unterkunft für Wohnungslose
Sonstige sozialtherapeutische Einrichtungen
Einrichtung für ältere und/oder pflegebedürftige Menschen
Einrichtung für Menschen mit Behinderung
Einrichtung für Kinder und Jugendliche
Krankenhaus, vor allem Palliativstation, Psychiatrische Klinik und Maßregelvollzugseinrichtung
Justizvollzugsanstalt
Kaserne der Bundeswehr und Polizei

Zudem werden im Einrichtungsregister auch Daten zu sogenannten Sonderfällen geführt. Sonderfälle sind beispielsweise ausländische diplomatische Vertretungen in Deutschland oder Unterkünfte von in Deutschland stati-

onierten ausländischen Streitkräften. Diese Sonderfälle sind im Zensus methodisch besonders zu berücksichtigen: Sie werden bei der Ermittlung der Bevölkerungszahl teilweise ausgeschlossen und im Rahmen von Erhebungen nicht befragt.

## 3

### Zentrale Merkmale

Im Einrichtungsregister werden Informationen zu den Einrichtungen selbst, nicht jedoch zu den darin lebenden Personen geführt.

Das Einrichtungsregister wird das Merkmal [Bereichsart](#) erfassen. Diese Einteilung ist wichtig, um später zu den erfassten Bereichsarten passgenaue Auswertungen für unterschiedliche Interessengruppen durchführen zu können. So könnten für verschiedene Ministerien auf Bundes- und Landesebene jeweils unterschiedliche Bereichsarten von Relevanz sein: Für den regelmäßig vom Bundesministerium für Gesundheit erstellten [Pflegebericht](#) könnten insbesondere die Daten zu Einrichtungen für ältere und/oder pflegebedürftige Menschen bedeutsam sein.

Außerdem werden die Merkmale [Name](#) sowie [Anschrift der Einrichtung](#) enthalten sein. Die Anschrift wird im Einrichtungsregister gemäß EU-Richtlinie der europäischen Geodateninfrastruktur (Infrastructure for Spatial Information in Europe – INSPIRE) georeferenziert. Mit INSPIRE werden Geoinformationen in Europa über Verwaltungsgrenzen hinweg nutzbar gemacht. Auch kleinräumige kartografische Darstellungen und Analysen werden ermöglicht, zum Beispiel über die Kartierung nach Gitterzellen (siehe auch Kapitel 6).<sup>1</sup>

Weiterhin wird das Einrichtungsregister das Merkmal [Anzahl der Einrichtungsplätze](#) für jede Einrichtung führen, sofern diese Information verfügbar ist. Nicht bei allen Bereichsarten kann dieses Merkmal ausgewiesen werden. Das ist vor allem abhängig davon, ob es in der Datenquelle vorhanden ist, die zur Pflege dieser

<sup>1</sup> Ein strategisches Ziel der Statistischen Ämter des Bundes und der Länder ist es, Geokoordinaten als Attribut zu etablieren und als weiteren Standard in den statistischen Produktionsprozess aufzunehmen. Dies soll sicherstellen, dass statistische Informationen adäquat, zeitgemäß und entsprechend der internationalen Standards bereitgestellt und verbreitet werden.

Bereichsart verwendet wird (siehe auch Kapitel 4). Die Anzahl der Einrichtungsplätze wird für Auswertungszwecke erfasst. So können grundsätzlich verfügbare Kapazitäten je Bereichsart ausgewertet werden. Weiterhin dient das Merkmal dazu, Plausibilitätsprüfungen im Einrichtungsregister durchzuführen. Diese ermöglichen es, unplausible Eintragungen oder fehlerhafte Datensätze zu ermitteln.

Außerdem wird das Einrichtungsregister die Information dazu enthalten, ob es möglich ist, in einer Einrichtung einen **eigenen Haushalt** zu führen. In diesem Sinne liegt eine eigene Haushaltsführung vor, wenn die Art des Wohnens oder der Unterbringung ein selbstständiges Wirtschaften in einer eigenen Wohneinheit ohne die Betreuung oder Versorgung durch Dritte ermöglicht. Nach dieser Definition lassen Wohnheime eine eigene Haushaltsführung zu, während dies für Gemeinschaftsunterkünfte nicht gilt. Diese Einteilung ist insbesondere für statistische Zwecke relevant, speziell für die im Zensus erforderliche Haushaltegenerierung (siehe auch Kapitel 6).

Zudem werden **Informationen zu den Trägern der Einrichtungen** geführt, wie der Name und die Anschrift des Trägers, des Verwalters oder des Eigentümers. Basierend auf diesen Informationen sind Aussagen zur Trägerlandschaft von Einrichtungen möglich. Denkbar wäre zum Beispiel, dass Einrichtungen nach der Art des Trägers (beispielsweise privater Träger, öffentlicher Träger oder Träger der freien Wohlfahrtspflege) ausgewertet werden. Ebenso wird das Einrichtungsregister **Kontaktinformationen** für eine mögliche Kontaktaufnahme mit den Trägern oder Einrichtungen enthalten, um relevante Informationen zu ihren Einrichtungen einzuholen, falls notwendig.

Die zentralen Merkmale im Einrichtungsregister werden sein:

- › Bereichsart (Art der Einrichtung)
- › Name der Einrichtung
- › Anschrift der Einrichtung
- › Anzahl der Einrichtungsplätze
- › Möglichkeit zur eigenen Haushaltsführung
- › Name und Anschrift des Trägers der Einrichtung
- › Kontaktinformationen des Trägers oder der Leitung der Einrichtung

## 4

### Datenquellen zur Pflege des Einrichtungsregisters

Der Datenbestand des Einrichtungsregisters wird regelmäßig zu pflegen und zu aktualisieren sein, da im Laufe der Zeit Einrichtungen eröffnen, ihre Bestimmung ändern oder schließen.

Für die regelmäßige Pflege des Einrichtungsregisters werden bereits vorhandene Daten insbesondere aus der öffentlichen Verwaltung, zum Beispiel von Ministerien und Ämtern, zu verwenden sein, um die Daten belastungsarm zu gewinnen. Sowohl auf nationaler als auch auf europäischer Ebene ist die Verwendung von Verwaltungsdaten in den vergangenen Jahren rechtlich gestärkt worden.<sup>12</sup>

Zum Aufbau und zur Pflege des Einrichtungsregisters gibt es keine übergreifende, alle Bereichsarten abdeckende Datenquelle, die regelmäßig über Daten zu allen Wohnheimen und Gemeinschaftsunterkünften in Deutschland verfügt. Aufgrund der Heterogenität der geführten Bereichsarten werden unterschiedliche Quellen die Daten liefern. Neben den Verwaltungsdaten werden auch Statistikdaten sowie öffentlich verfügbare Verzeichnisse verwendet. Auch Daten von Verbänden und privaten Anbietern werden auf ihre Eignung geprüft.

Über die wiederkehrenden Datenlieferungen lässt sich der Datenbestand des Einrichtungsregisters regelmäßig aktualisieren. Vereinfacht gelten folgende Regeln: Liefern die Datenquellen zu einer bisher bekannten Einrichtung wiederholt keine Daten mehr, gilt diese als „geschlossen“. Eine bisher bekannte Einrichtung, zu der eine Datenquelle erneut Informationen liefert, gilt weiterhin als „aktuell“. Über die Datenlieferungen könnten auch Informationen zu Änderungen einzelner Merkmale ins Einrichtungsregister gelangen, zum Beispiel, dass eine bestimmte Einrichtung nunmehr keine Einrichtung für Kinder und Jugendliche, sondern eine Mutter/Vater-Kind-Einrichtung ist.

2 Siehe Artikel 17a „Zugang zu Verwaltungsunterlagen sowie deren Verwendung und Integration“ der EU-Statistikverordnung Nr. 223/2009 und § 5a „Nutzung von Verwaltungsdaten“ des Bundesstatistikgesetzes sowie Engelter/Sommer (2017).



## 4.1 Daten der amtlichen Statistik

---

Besonders relevant für die regelmäßige Aktualisierung des Einrichtungsregisters sind Daten aus der amtlichen Statistik, insbesondere aus thematisch verwandten Statistiken. Dazu zählen die Statistik der stationären Pflegeeinrichtungen, die Grunddaten der Krankenhäuser, die Grunddaten der Vorsorge- oder Rehabilitationseinrichtungen, die Statistik untergebrachter wohnungsloser Personen oder auch das statistische Unternehmensregister. Diese Datenquellen wurden bereits entsprechend den Standards der amtlichen Statistik intensiv geprüft, aufbereitet und qualitätsgesichert.

Einmalig soll für den initialen Aufbau des Einrichtungsregisters der Sonderanschriftenbestand des Zensus 2022 verwendet werden. Diese Daten spiegeln den Stand zum Zensusstichtag 15. Mai 2022 wider. Da seitdem Einrichtungen neu hinzugekommen sind, ihre Bestimmung geändert haben oder geschlossen worden sein können, ist es erforderlich, diesen initialen Datenbestand zu aktualisieren (Boragk und andere, 2024).

## 4.2 Daten der öffentlichen Verwaltung

---

Nicht für alle im Einrichtungsregister geführten Bereichsarten liegen Daten der amtlichen Statistik vor. Daher wird die Nutzung von [Datenquellen, die der öffentlichen Verwaltung zuzuordnen sind](#), zentral für die Pflege des Einrichtungsregisters sein. Eine Vielzahl von relevanten und potenziell geeigneten Verwaltungsdatenquellen konnte bereits für die Pflege des Einrichtungsregisters identifiziert werden. Dazu zählen Daten des Bundesamtes für Migration und Flüchtlinge über Gemeinschaftsunterkünfte für Flüchtlinge (Erstaufnahmeeinrichtungen) sowie Daten der Bildungsministerien über Internate oder Daten der Jugendämter der Länder über Einrichtungen für Kinder und Jugendliche.

Daten der öffentlichen Verwaltung haben den Vorteil, dass die zuständigen (Verwaltungs-)Stellen häufig auch zentrale administrative Tätigkeiten der Einrichtungen, zum Teil auch die Trägerschaft selbst, wahrnehmen. Allerdings gibt es auch verschiedene Herausforderungen, die bei der weiteren Konzeption des Einrichtungsregisters zu berücksichtigen sind. So weisen die Datenlieferanten eine gewisse Heterogenität auf. Dies betrifft zum Teil die Qualität der Daten, den Abdeckungsgrad

und die Zugehörigkeit der Dateneigner zu unterschiedlichen Verwaltungsebenen (vor allem Bund oder Land). Bei der Planung der Übermittlungswege für den Dateneintrag und bei der Aufbereitung des Datenbestands ist diese Heterogenität zu berücksichtigen.

## 4.3 Daten von Verbänden

---

Zudem sollen Daten von Verbänden (zum Beispiel Verbände der freien Wohlfahrtspflege) für das Einrichtungsregister genutzt werden. In einigen Bereichsarten, wie den Einrichtungen für ältere und/oder pflegebedürftige Menschen und den Einrichtungen für Kinder und Jugendliche, führen die Verbände der freien Wohlfahrtspflege einen erheblichen Anteil der Einrichtungen. Die Daten stellen daher für bestimmte Bereichsarten eine wichtige Ergänzung zu den Daten der öffentlichen Verwaltung dar.

## 4.4 Daten privater Anbieter und allgemein zugängliche Daten

---

Zusätzlich können Verzeichnisse privater Dateneigner genutzt werden, um das künftige Einrichtungsregister zu pflegen. Bei bestimmten Bereichsarten wie Studierendenwohnheimen führen private Anbieter einen gewissen Anteil der Einrichtungen, deren Daten sind in besonderem Maße in Verzeichnissen privater Dateneigner enthalten. Im Rahmen der Novellierung der EU-Statistikverordnung 223/2009 durch die EU-Verordnung 2024/3018 wurde § 17 ausgeweitet: Private Datenanbieter können dazu verpflichtet werden, den nationalen Statistikämtern oder dem Statistischen Amt der Europäischen Union (Eurostat) auf Verlangen und unter bestimmten Bedingungen (unter anderem im Einklang mit den Grundsätzen der statistischen Geheimhaltung und der Kostenwirksamkeit) Daten für die Erstellung europäischer Statistiken zur Verfügung zu stellen. Zudem sollen Daten, die allgemein zugänglich sind, für den Aufbau und die Pflege des Einrichtungsregisters verwendet werden (beispielsweise Verzeichnisse im Internet, unter anderem von privaten Datenanbietern, wie der Wohnheimfinder für Studierende des Deutschen Akademischen Austauschdienstes).



### 4.5 Manuelle Recherchen von Einrichtungen

Nur in den Fällen, in denen keine bekannten Datenquellen vorhanden sind, soll für die betroffenen Bereichsarten auf manuelle Recherchen von Einrichtungen durch die Mitarbeitenden der statistischen Ämter zurückgegriffen werden. Auf Befragungen von Trägern zu ihren Einrichtungen soll möglichst vollständig verzichtet werden, da sie eine höhere Belastung der Befragten nach sich ziehen würden.

#### ➤ Ausblick: Innovationspotenziale bei manuellen Recherchen von Einrichtungen

Ergänzende manuelle Recherchen von Einrichtungen durch die statistischen Ämter können nicht ganz vermieden werden. Auch hier besteht belastungsreduzierendes Innovationspotenzial: Wo möglich, soll eine automatisierte Unterstützung im Datenbearbeitungsprozess, insbesondere bei der Datengewinnung, umgesetzt werden, um manuelle Arbeiten zu minimieren. So könnten beispielsweise die automatisierte Suche und Ausgabe von relevanten Datenpunkten in Open-Source-Software wie OpenStreetMap die manuellen Recherchen von Einrichtungen unterstützen. Das Verfahren wird zurzeit für einzelne Bereichsarten, bei denen keine geeignete Datenquellen identifiziert werden konnten, erprobt.

## 5

### Qualitätssicherung

Der § 5a Bundesstatistikgesetz ermöglicht dem Statistischen Bundesamt, infrage kommende Datenquellen von Stellen der öffentlichen Verwaltung auf ihre Eignung hin zu untersuchen. Der Prozess der Eignungsprüfungen ist in zwei Prüfschritte aufgeteilt: Im ersten Schritt werden die Metadaten (beschreibende Informationen über vorhandene Datenbestände, die von den Verwaltungsstellen zu übermitteln sind) geprüft. Im zweiten Schritt werden die eigentlichen Daten, das heißt ihre Inhalte, geprüft. Hierfür sind dem Statistischen Bundesamt auf Anforderung formal anonymisierte Einzelangaben für weitere Eignungsprüfungen von der Verwaltungsstelle zu übermitteln. Für Eignungsprüfungen von Datenquellen, die zur Pflege des Einrichtungsregisters herangezogen werden sollen, sind die folgenden Prüfkriterien zentral:

- › Führt die Datenquelle die für die jeweilige Bereichsart relevanten Einheiten und Merkmale? Hierfür werden die in den Datenquellen verwendeten Definitionen der vorhandenen Einheiten und Merkmale mit den im Einrichtungsregister geltenden Definitionen verglichen.
- › Deckt die Datenquelle die zu erfassenden Einrichtungen (Einheiten) der jeweiligen Bereichsart ausreichend ab? Hierbei ist zu prüfen, inwiefern und gegebenenfalls in welchem Ausmaß eine Unter- oder Übererfassung von Einheiten in der Datenquelle zu erwarten ist.
- › Kann die Datenquelle alle zentralen Merkmale, die das Einrichtungsregister führt, liefern? Kann auf fehlende Merkmale verzichtet werden?
- › Wie häufig und zu welchem Zeitpunkt wird die Datenquelle aktualisiert?
- › Welche Datenübermittlungswege kommen infrage? Kann die Datenquelle die von der Statistik vorgegebene Datenstruktur liefern?
- › Wie ist insgesamt die Qualität der Datenquelle zu bewerten, beispielsweise mit Blick auf mögliche enthaltene Fehler? Inwiefern sind qualitätssichernde Maßnahmen zu ergreifen?

Stellt sich bei den Eignungsprüfungen heraus, dass die Daten der öffentlichen Verwaltung geeignet sind, ist die regelmäßige Übermittlung der Einzeldaten für die konkrete statistische Verwendung in den kommenden gesetzlichen Folgeregelnungen des Registerzensuserprobungsgesetzes festzulegen.

Im Rahmen einer regelmäßigen Evaluierung ist insbesondere für die Schritte Datengewinnung und Datenaufbereitung systematisch die Qualität der eingegangenen Datenquellen zu analysieren und sind gegebenenfalls qualitätssichernde Maßnahmen umzusetzen. Hierfür werden aussagekräftige Qualitätskennzahlen – in Anlehnung an die Qualitätsdatenblätter des Statistischen Verbunds<sup>3</sup> – ausgearbeitet (Meinke/Hentschke, 2022; Tümmeler/Meinke, 2019).

<sup>3</sup> Den Statistischen Verbund bilden die Statistischen Ämter des Bundes und der Länder.

Eine mögliche Auswahl der Qualitätskennzahlen sind

- › Lieferversäumnisse bei Datenlieferanten am Dateneingang,
- › fälschlicherweise erfasste Einheiten (aufgrund zum Beispiel von Dubletten),
- › Antwortausfälle bei relevanten Merkmalen und die
- › Quote der festgestellten sowie bearbeiteten Fehler nach verschiedenen Fehlerschlüsseln.

Basierend auf den Qualitätskennzahlen ist unter anderem zu erkennen, ob sich die Qualität einer bestimmten Datenquelle über die Jahre verändert hat. Beispielsweise kann eine verschlechterte Datenqualität dazu führen, dass eine Datenquelle nicht mehr zur Aktualisierung des Einrichtungsregisters herangezogen wird. Turnusmäßig wiederkehrende Eignungsprüfungen identifizieren neue Datenquellen oder bereits bekannte, die zwischenzeitlich eine verbesserte Qualität erreicht haben. So kann künftig eine kontinuierliche Evaluierung und damit auch Verbesserung der Pflege des Einrichtungsregisters erfolgen.

## 6

### Nutzungs- und Auswertungspotenziale

---

Das Einrichtungsregister wird im Kontext des Zensus entwickelt und soll darüber hinaus auch von anderen Bundes- und Landesstatistiken genutzt werden. Zudem können mithilfe des Einrichtungsregisters EU-Lieferverpflichtungen erfüllt und weiterführende Auswertungen auf nationaler Ebene durchgeführt werden, die politische Entscheidungsträger in Bund, Ländern und Gemeinden unterstützen.

#### 6.1 Nutzungspotenziale innerhalb des Zensus

---

Das Einrichtungsregister wird zunächst im Kontext der Weiterentwicklung des Zensus aufgebaut, bei dem die [Ermittlung der Bevölkerungszahl](#) in Deutschland im Fokus steht (Körner und andere, 2019; Söllner/Körner, 2022). Dabei ist das Einrichtungsregister ein wichtiger Baustein, denn zur Bevölkerung Deutschlands gehören natürlich auch Personen, die in Einrichtungen leben. Im

Einrichtungsregister werden zwar keine Informationen zu den dort lebenden Personen erhoben oder geführt, sondern nur zu den Einrichtungen selbst. Abhängig von der methodischen Ausgestaltung des künftigen Zensus liefert das Einrichtungsregister relevante Informationen: entweder, um Personen mit Einrichtungen zu verknüpfen, oder als Grundlage, um den Berichtskreis zu ermitteln für eine möglicherweise durchzuführende Erhebung von Personen, die in Einrichtungen leben.

Im Zensus wird außerdem eine [Haushaltgenerierung](#) durchgeführt, bei der das Einrichtungsregister unterstützt. Die Haushaltgenerierung ordnet die an einer Anschrift gemeldeten Personen einem privaten Haushalt zu. Personen, die in Gemeinschaftsunterkünften betreut oder versorgt werden, werden von der Haushaltgenerierung ausgeschlossen, da per definitionem keine eigene Haushaltsführung vorliegt und sie damit keinen privaten Haushalt darstellen.

Im Zensus findet außerdem eine [Gebäude- und Wohnungszählung](#) statt, bei der das Einrichtungsregister den Erhebungsprozess und die Bestimmung des Berichtskreises unterstützt. In der Gebäude- und Wohnungszählung werden Wohnheime befragt, Gemeinschaftsunterkünfte sind jedoch nicht relevant und von der Befragung ausgeschlossen. Das Einrichtungsregister liefert hierfür die Daten zu Wohnheimeigentümern, die als Auskunftspflichtige herangezogen werden.

#### 6.2 Nutzungspotenziale außerhalb des Zensus

---

Das Einrichtungsregister kann als Statistikregister von der amtlichen Statistik auch außerhalb des Zensus als relevante Datengrundlage genutzt werden. Einerseits liegen Vorteile für andere Statistiken in der [gemeinsamen Berichtskreispflege](#) bei sich thematisch überschneidenden Bereichsarten<sup>4</sup>. So könnten im Einrichtungsregister erfasste Einheiten, die den anderen thematisch verwandten Bundes- und Landesstatistiken aufgrund unterschiedlicher Erhebungsmethodik nicht bekannt wären, übermittelt werden. Andererseits könnten die in den jeweils thematisch verwandten Statistiken

---

<sup>4</sup> Beispielsweise mit der Statistik der stationären Pflegeeinrichtungen, mit den Grunddaten der Krankenhäuser, mit den Grunddaten der Vorsorge- oder Rehabilitationseinrichtungen, mit der Statistik untergebrachter wohnungsloser Personen.

erfassten und relevanten Einheiten für die Pflege des Einrichtungsregisters genutzt werden, sodass eine solche gemeinsame Berichtsreisepflege zu einem effizienteren Statistikerstellungsprozess führt. Darüber hinaus kann ein Zusammenwirken des Einrichtungsregisters mit thematisch verwandten Statistiken auch vielseitige Auswertungspotenziale schaffen.

### 6.3 Auswertungspotenziale

Das Einrichtungsregister sollen die Statistischen Ämter des Bundes und der Länder nutzen können, um in ihrem Zuständigkeitsbereich Auswertungen durchzuführen. Solche Auswertungen sind abhängig von den europäischen und nationalen Anforderungen und können somit erst mit Verabschiedung der jeweiligen nationalen und europäischen Regelungsgrundlagen final festgelegt werden. Dabei sind auch datenschutzrechtliche Überlegungen zu berücksichtigen. Der rechtliche Rahmen ist in den Folgeregelungen des Registerzensuserprobungsgesetzes weiter auszuarbeiten.

Mithilfe des Einrichtungsregisters werden [Lieferverpflichtungen an Eurostat](#) aus der zu erwartenden, zukünftigen EU-Rahmenverordnung über Bevölkerungsstatistiken (ESOP<sup>5</sup>) erfüllt. Dabei können beispielsweise folgende Fragestellungen unter anderem aus den Daten des Einrichtungsregisters, gegebenenfalls verknüpft mit anderen Zensusmerkmalen, beantwortet werden: Wie viele Gemeinschaftsunterkünfte gibt es in Deutschland? Wie viele Menschen in Deutschland leben in Gemeinschaftsunterkünften? Wie viele Menschen in Deutschland leben in Einrichtungen?

Denkbar sind außerdem auch Auswertungen, die [unmittelbar aus dem Datenbestand des Einrichtungsregisters](#) erfolgen (also auch unabhängig vom Zensus). Durch die verschiedenen, im Einrichtungsregister geführten Arten von Wohnheimen und Gemeinschaftsunterkünften ergibt sich ein vielfältiges Auswertungspotenzial. Dieses könnten Datennutzende aus den unterschiedlichsten thematischen Tätigkeitsbereichen sowohl bei Ämtern und Ministerien des Bundes und der Länder sowie kommunalen Stellen als auch auf europäischer

Ebene nutzen: So veröffentlicht das Bundesministerium für Familie, Senioren, Frauen und Jugend beispielsweise regelmäßig einen Familienbericht, einen Kinderbericht, einen Jugendbericht sowie einen Altersbericht.<sup>16</sup> Diese Berichte weisen thematische Überschneidungen zu verschiedenen im Einrichtungsregister geführten Bereichsarten auf. Anhand der Daten des Einrichtungsregisters könnten in Zukunft Analysen zu den verschiedenen Bereichsarten, die in den unterschiedlichen Berichten thematisiert sind, durchgeführt werden.

Durch eine mögliche [Verknüpfung des Datenbestands](#) des Einrichtungsregisters mit den Datenbeständen anderer amtlichen Statistiken ergeben sich vielseitige Auswertungsmöglichkeiten, ohne zusätzliche Belastungen durch Befragungen hervorzurufen. So könnte eine Verknüpfung der Daten der Statistik der Sterbefälle mit dem Einrichtungsregister zeigen, inwiefern sich der Sterbeort in bestimmten Einrichtungen über die Zeit entwickelt. Solche Auswertungen könnten für das Robert Koch-Institut sowie als politische Entscheidungsgrundlage von hoher Relevanz sein, insbesondere nach besonderen Ereignissen wie einer Pandemie oder Extremwetter.

Da das Einrichtungsregister perspektivisch georeferenziert werden soll, werden zudem [kleinräumige kartografische Darstellungen und Analysen](#) ermöglicht, ähnlich wie im [Krankenhausatlas](#), im [Zensusatlas](#), im [Agraratlas](#) und im [Unfallatlas](#). Solche interaktiven Karten verdeutlichen die kleinräumigen Verteilungen von bestimmten Einheiten, beispielsweise von Einrichtungen für ältere, geflüchtete oder behinderte Menschen. Auch können solche Karten für ganz Deutschland die Erreichbarkeit der nächsten Pflegeeinrichtung von verschiedenen Standorten aus zeigen. Denkbar wäre zudem, weitere Mobilitätsoptionen wie den öffentlichen Personennahverkehr einzubinden. Dabei könnten folgende Fragestellungen im Fokus stehen: Wie stellt sich die durchschnittliche Wegstrecke zur nächsten Einrichtung dar? Sind die Einrichtungen gut mit öffentlichen Verkehrsmitteln erreichbar? Darüber hinaus wären [Simulationsstudien](#) denkbar, die demografische Veränderungsprozesse wie eine älter werdende Bevölkerung in bestimmten Regionen in die Zukunft projizieren und dabei die gegenwärtige örtliche Verteilung der Einrichtungen für ältere und/

<sup>5</sup> ESOP steht für European statistics on population. In der ESOP sind auch für das Einrichtungsregister relevante Merkmale enthalten: Living quarters (LVQ), Housing arrangements (HAR) und Household status (HST).

<sup>6</sup> Eine Übersicht aller Berichte des Bundesministeriums für Familie, Senioren, Frauen und Jugend findet sich auf dessen Webseite: [www.bmfsfj.de](http://www.bmfsfj.de)

oder pflegebedürftige Menschen analysieren könnten. Daraus könnten Entscheidungsträger dann Handlungsmaßnahmen ableiten, beispielsweise ob entsprechende Einrichtungen neu zu bauen wären.

## 7


---

### Fazit

---

Mit dem Einrichtungsregister soll ein regelmäßig aktualisierter, qualitätsgesicherter und historisierter, dauerhafter Datenbestand zu allen Wohnheimen und Gemeinschaftsunterkünften in Deutschland aufgebaut werden. Ein solcher Datenbestand ist in Deutschland einmalig.

Für die regelmäßige Pflege des Einrichtungsregisters sollen verschiedene Daten vorrangig aus Statistik und Verwaltung verwendet werden. Es sollen auch Daten von Verbänden und privaten Anbietern herangezogen werden, sofern sie die Qualitätsstandards der amtlichen Statistik erfüllen. Um die herangezogenen Datenquellen systematisch evaluieren zu können und eine angemessene Qualität sicherzustellen, sollen im Einrichtungsregister verschiedene Qualitätskennzahlen etabliert werden.

Das Einrichtungsregister wird im Kontext der Weiterentwicklung der Zensusmethodik aufgebaut, wo es für verschiedene Module und Verfahren zur Ermittlung der Bevölkerungszahl sowie weiterer Zensus-Merkmale benötigt wird. Es hat vielseitige Verwendungsmöglichkeiten, die auch von anderen relevanten Bundes- und Landesstatistiken genutzt werden können. Die geführten Informationen dienen dazu, EU-Lieferverpflichtungen im Rahmen der zukünftigen EU-Rahmenverordnung über Bevölkerungsstatistiken (ESOP) zu erfüllen. Ebenso können die Daten sozialpolitische oder planerische Entscheidungen auf nationaler Ebene in Bund, Ländern und Kommunen unterstützen. Das Einrichtungsregister birgt vielseitige kleinräumige Auswertungspotenziale, die sich unmittelbar aus dem eigenen Datenbestand oder anhand von Verknüpfungen mit anderen Datenbeständen ergeben. Die Auswertungsmöglichkeiten erstrecken sich von Häufigkeitsauswertungen über kartografische Darstellungen und Erreichbarkeitsanalysen bis hin zu komplexen Simulationsstudien. 

### LITERATURVERZEICHNIS

---

Boragk, Lisa/Gaedke, Annika/Gemmeke, Charlotte/Meyn, Christoph. [Die Ermittlung des Berichtskreises am Beispiel der Sonderbereiche im Zensus 2022](#). In: WISTA Wirtschaft und Statistik. Ausgabe 6/2024, Seite 41 ff.

Engelter, Marion/Sommer, Kay. [Die Novellierung des Bundesstatistikgesetzes 2016](#). In: WISTA Wirtschaft und Statistik. Ausgabe 6/2016, Seite 11 ff.

Körner, Thomas/Krause, Anja/Ramsauer, Kathrin. [Anforderungen und Perspektiven auf dem Weg zu einem künftigen Registerzensus](#). In: WISTA Wirtschaft und Statistik. Sonderheft Zensus 2021. 2019, Seite 74 ff.

Meinke, Irina/Hentschke, Janine. [Kern-Qualitätskennzahlen im Zensus 2022](#). In: WISTA Wirtschaft und Statistik. Ausgabe 3/2022, Seite 25 ff.

Söllner, René/Körner, Thomas. [Der Registerzensus: Ziele, Anforderungen und Umsetzungsansätze](#). In: WISTA Wirtschaft und Statistik. Ausgabe 4/2022, Seite 13 ff.

Statistisches Bundesamt. *Die Datenstrategie des Statistischen Bundesamtes. Version 4.1*. 2021. [Zugriff am 16. Januar 2025]. Verfügbar unter: [www.statistischebibliothek.de](http://www.statistischebibliothek.de)

Tümmler, Thorsten/Meinke, Irina. [Aufbau des Qualitätsmanagements für den Zensus 2021](#). In: WISTA Wirtschaft und Statistik. Sonderheft Zensus 2021. 2019, Seite 59 ff.

## RECHTSGRUNDLAGEN

---

Bundesmeldegesetz (BMG) vom 3. Mai 2013 (BGBl. I Seite 1084), das zuletzt durch Artikel 6 des Gesetzes vom 23. Oktober 2024 (BGBl. I Nr. 323) geändert worden ist.

Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) in der Fassung der Bekanntmachung vom 20. Oktober 2016 (BGBl. I Seite 2394), das zuletzt durch Artikel 14 des Gesetzes vom 8. Mai 2024 (BGBl. I Nr. 152) geändert worden ist.

Gesetz zur Erprobung von Verfahren eines Registerzensus (Registerzensuserprobungsgesetz – RegZensErpG) vom 9. Juni 2021 (BGBl. I Seite 1649).

Richtlinie 2007/2/EG des europäischen Parlaments und des Rates vom 14. März 2007 zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft (INSPIRE) (Amtsblatt der EU Nr. L 108, Seite 1).

Verordnung (EU) 2015/759 des Europäischen Parlaments und des Rates vom 29. April 2015 zur Änderung der Verordnung (EG) Nr. 223/2009 über europäische Statistiken (Amtsblatt der EU Nr. L 123, Seite 90).

Verordnung (EU) 2024/3018 des Europäischen Parlaments und des Rates vom 27. November 2024 zur Änderung der Verordnung (EG) Nr. 223/2009 über europäische Statistiken.

Verordnung (EG) Nr. 223/2009 des europäischen Parlaments und des Rates vom 11. März 2009 über europäische Statistiken und zur Aufhebung der Verordnung (EG, Euratom) Nr. 1101/2008 des Europäischen Parlaments und des Rates über die Übermittlung von unter die Geheimhaltungspflicht fallenden Informationen an das Statistische Amt der Europäischen Gemeinschaften, der Verordnung (EG) Nr. 322/97 des Rates über die Gemeinschaftsstatistiken und des Beschlusses 89/382/EWG, Euratom des Rates zur Einsetzung eines Ausschusses für das Statistische Programm der Europäischen Gemeinschaften (Amtsblatt der EU Nr. L 87, Seite 164).



# MEHRFACHFALLPRÜFUNG IM ZENSUS 2022 – DIE NEUE STRATEGIE ZUR AUTOMATISIERTEN IDENTIFIKATION UND BEWERTUNG VON MEHRFACHFÄLLEN

Moritz Wittmaack

➤ **Schlüsselwörter:** Record Linkage – Datendeduplikation – Melderegister – amtliche Einwohnerzahl – Personenbestand

## ZUSAMMENFASSUNG

Die kombinierte Mehrfachfallprüfung der Melderegisterdaten und der Fehlbestände der primärstatistischen Erhebungen ist ein vom Statistischen Bundesamt entwickeltes vollständig maschinelles Verfahren, um die Qualität der Zensusergebnisse zu sichern: Es bringt potenzielle Mehrfachfälle – mehrere dieselbe Person repräsentierende Datensätze – aus dem Personenbestand des Zensus 2022 zusammen, prüft diese anhand fester Regeln auf Zusammengehörigkeit und stellt die melderechtliche Zulässigkeit der vorliegenden Mehrfachmeldungen sicher. Besonders herausfordernd war hierbei das Fehlen eines registerübergreifenden, persistenten Personenidentifikators. Der Artikel erläutert die Hintergründe, die zur neuen Mehrfachfallprüfung geführt haben, und stellt den Ablauf des Verfahrens vor.

➤ **Keywords:** record linkage – data deduplication – population register – official population – population stock

## ABSTRACT

*To guarantee the quality of the census results, the Federal Statistical Office developed a fully automated method to check for multiple entries in a data pool that combines population register data and missing data entries identified in primary data surveys. The method gathers potential multiple entries – i.e. two or more data records representing the same person – from the population stock of the 2022 Census, checks whether the entries refer to the same entity using a fixed set of rules, and ensures that the multiple entries identified are permitted under Germany's registration law. The lack of a persistent personal identifier across all registers proved to be a particular challenge in this context. This article explains the background to the introduction of the new multiple entry check and describes the individual steps of this new process.*

### Moritz Wittmaack

studierte Soziologie und empirische Sozialforschung in Bielefeld und Köln. Von Dezember 2019 bis Mai 2023 war er als Wissenschaftlicher Mitarbeiter im Statistischen Bundesamt mit der Konzeption sowie der Umsetzung der Mehrfachfallprüfung des Zensus 2022 betraut. Seit Juni 2023 ist er Referent im Referat „Lohn- und Einkommensteuer“.

Maßgeblich an der Entwicklung der Mehrfachfallprüfung als vollständig maschinelles Verfahren beteiligt waren Bernd Michel (†) und Paul Konstantin Schmidtke. Philipp Hadeball und Thomas Franke unterstützten die Umsetzung des entwickelten Verfahrens mittels SAS Version 9.4.

## 1

## Einleitung

Die Mehrfachfallprüfung ist ein Prozess zur Qualitätssicherung des Personenbestands des Zensus 2022, auf dessen Grundlage die Einwohnerzahl Deutschlands abgeleitet wurde (Bretschki und andere, 2024). Für die Integration der verschiedenen Datenquellen des Personenbestands stand kein registerübergreifender, persistenter Personenidentifikator zur Verfügung. Daher konnte aufgrund der dezentralen Melderegisterführung nicht davon ausgegangen werden, dass jede zum Zensusstichtag in Deutschland lebende Person mit genau einem einwohnerzahlrelevanten Melderegisterdatensatz im Personenbestand enthalten sein würde.

Um die Qualität der Zensusergebnisse zu sichern, sah § 21 Absatz 1 Zensusgesetz 2022 deshalb vor, dass das Statistische Bundesamt anhand der von den Meldebehörden gelieferten Melderegisterdaten prüft, ob Personen mit mehr als einer alleinigen Wohnung oder Hauptwohnung oder nur mit einer oder mehreren Nebenwohnungen gemeldet sind. Bei nicht melderechtskonformen mehrfachen Meldungen einer Person war zu entscheiden, welcher Melderegisterdatensatz relevant für die Ermittlung der Einwohnerzahl ist (§ 21 Absatz 2 Zensusgesetz 2022). Außerdem war für Personen, die an Sonderbereichsanschriften wohnten, dort aber nicht gemeldet waren, der Wohnungsstatus festzustellen (§ 21 Absatz 3 Zensusgesetz 2022). Ob diese Personen zusätzlich unter gegebenenfalls abweichenden Angaben an derselben oder einer anderen Anschrift in Deutschland gemeldet waren, klärte die Mehrfachfallprüfung.

Das dazu entwickelte fehlertolerante Verfahren bringt alle dieselbe Person repräsentierenden Datensätze zusammen, prüft auf Zusammengehörigkeit und bestimmt letztlich, welcher der vorhandenen Datensätze zur Ermittlung der Einwohnerzahl herangezogen wird. Anstatt wie im Zensus 2011 strittige Fälle händisch zu beurteilen und den Wohnsitz mithilfe einer postalischen Befragung festzustellen (Diehl, 2012), wurde im Zensus 2022 ein vollständig maschinelles Verfahren angewandt.

Kapitel 2 beschreibt den Aufbau des Personenbestands, Kapitel 3 die sich aus der Art der Datenintegration ergebenden Datenfehler. Hiernach stellt Kapitel 4 das entwickelte Verfahren entlang seiner Anwendung im Zensus

2022 vor, die Kapitel 5 und 6 befassen sich mit der Suche nach prozessproduzierten beziehungsweise dauerhaften Mehrfachfällen. Anschließend untersucht Kapitel 7 die Wirkung der Mehrfachfallprüfung auf den Personenbestand. Ein kurzes Fazit beschließt den Artikel.

## 2

## Generierung des stichtagsrelevanten Personenbestands

Entsprechend internationaler Vorgaben erfolgt die Feststellung der amtlichen Einwohnerzahl in Deutschland alle zehn Jahre durch einen Zensus. Wie der vorangegangene Zensus 2011 wurde der Zensus 2022 registergestützt durchgeführt (Dittrich und andere, 2022). Ein Personenbestand in Form eines zentralen Registers aller in Deutschland wohnhaften Personen existiert nicht und war für den Zensus 2022 zu erstellen. Datenquelle hierfür waren die Melderegisterdaten der etwa 11 000 Gemeinden Deutschlands, welche in rund 5 500 Melderegister führende Stellen organisiert sind.

Die Daten des Melderegisterdatenabzugs zum Zensusstichtag (15. Mai 2022; MRZ1) stellten die Grundlage des Personenbestands dar. Drei Monate nach dem Stichtag erfolgte ein weiterer Melderegisterdatenabzug (MRZ2), der die Datengrundlage erweiterte. Dies diente dazu, Änderungen in den Melderegistern nachzuziehen, die sich vor oder am Zensusstichtag ereignet hatten, aber noch nicht im Melderegisterdatenbestand am Zensusstichtag enthalten waren. Beispiele dafür sind Veränderungen des Personenstands, Geburten, Todesfälle oder die Meldung eines Umzugs. Die Daten des MRZ2 wurden anschriftenbasiert in den mit MRZ1 befüllten Personenbestand integriert. Hierfür wurde auf personenidentifizierende Merkmale zurückgegriffen, weil kein registerübergreifender und persistenter Personenidentifikator zur Verfügung stand. Bei erfolgreicher Anbindung an einen vorhandenen Datensatz wurde dieser mit den MRZ2-Informationen aktualisiert. Konnte der MRZ2-Personendatensatz nicht angebunden werden, wurde er neu aufgenommen.

Neben den Melderegistern wurden auch die Daten der beiden primärstatistischen Erhebungen des Zensus 2022 anschriftenbasiert integriert. Hierbei handelt es

sich einmal um die Sonderbereichserhebung.<sup>1</sup> Die Vollerhebung der Sonderbereichsanschriften dient dazu, die Melderegister um Unter- und Übererfassungen zu korrigieren (Boragk und andere, 2024). Eine Untererfassung bedeutet, dass eine Person an einer Anschrift wohnt, aber dort nicht gemeldet ist (Fehlbestand). Eine Übererfassung liegt vor, wenn eine Person an einer Anschrift gemeldet, dort aber nicht wohnhaft ist (Karteileiche).

Die zweite primärstatistische Erhebung ist die Haushaltsstichprobe. Auch diese Haushaltebefragung bei einer Stichprobe der sonstigen Wohnanschriften bereinigt zum einen die Melderegister um Unter- und Übererfassungen. Zum anderen wurden mit einer Unterstichprobe der ausgewählten Wohnanschriften sowie bei einem Teil der Wohnheime zusätzliche zensusrelevante Informationen erhoben, welche nicht in den Melderegistern enthalten waren (Klink/Lorentz, 2022).

Bei erfolgreicher Anbindung der primärstatistischen Personendaten an einen bestehenden Melderegisterdatensatz wurde dieser um die primärstatistischen Merkmale ergänzt. Hierbei konnte die Existenz der durch den Melderegisterdatensatz repräsentierten Person bestätigt werden, wenn die Person angetroffen wurde und die Anbindung erfolgreich war („gemeldet und wohnhaft“). Wurde die Person bei erfolgreicher Anbindung nicht angetroffen, konnte deren Existenz nicht bestätigt werden (Karteileiche). War kein Melderegisterdatensatz zu finden, um die erhobenen primärstatistischen Merkmale an der Anschrift anzubinden, wurde ein neuer Personendatenatz angelegt (Fehlbestand).

Eine weitere Herausforderung bei der Generierung des stichtagsrelevanten Personenbestands war, dass die Melderegister aufgrund der dezentralen Pflege melde-rechtswidrige Konstellationen von Meldungen derselben Person enthalten. Die Meldedaten werden zwar seit 2007 gemäß dem XMeld-Schema gehalten und mit einem Protokollstandard anlassbezogen ausgetauscht.<sup>2</sup> Allerdings können etwa fehlerhafte Eingaben, Fehlnutzungen von Feldern oder Datenverluste bei Datenbankmigra-

tionen zu fehlerhaften Einträgen führen. Auch ist nicht auszuschließen, dass Karteileichen in MRZ1 und MRZ2 enthalten waren, welche auf die Zeit vor dem standardisierten Datenaustausch im Meldewesen zurückzuführen sind. Daher waren alle Melderegisterdatensätze zu einer Person zusammenzubringen und es war zu entscheiden, welcher von ihnen zur Ermittlung der Einwohnerzahl herangezogen werden sollte.

### 3

## Arten von Mehrfachfällen

Durch die sukzessive Datenintegration entstanden Mehrfachfälle, beispielsweise weil sich Namensbestandteile zwischen den Datenlieferungen verändert haben. Der Personenbestand war um diese prozessproduzierten Mehrfachfälle zu bereinigen, bevor nach den Mehrfachfällen innerhalb der Melderegister gesucht werden konnte. Bei Mehrfachfällen an einer Anschrift (technische Mehrfachfälle) wurde die Integration der Datensätze zu derselben Person nachgeholt. Bei Mehrfachfällen an unterschiedlichen Anschriften, welche auf Umzüge im Zeitraum zwischen MRZ1 und MRZ2 zurückzuführen sind (temporäre Mehrfachfälle), wurde der jeweils veraltete Melderegisterdatensatz inaktiv gesetzt. ➤ Grafik 1 auf Seite 130

Allerdings sind Mehrfachfälle im Personenbestand nicht zwangsläufig auf Fehler innerhalb der Melderegister oder auf Fehler bei der Erstellung des Personenbestands zurückzuführen. In Abgrenzung zu den im Datenintegrationsprozess entstehenden Mehrfachfällen werden diese im Folgenden als dauerhafte Mehrfachfälle bezeichnet. Sie lassen sich noch weiter differenzieren:

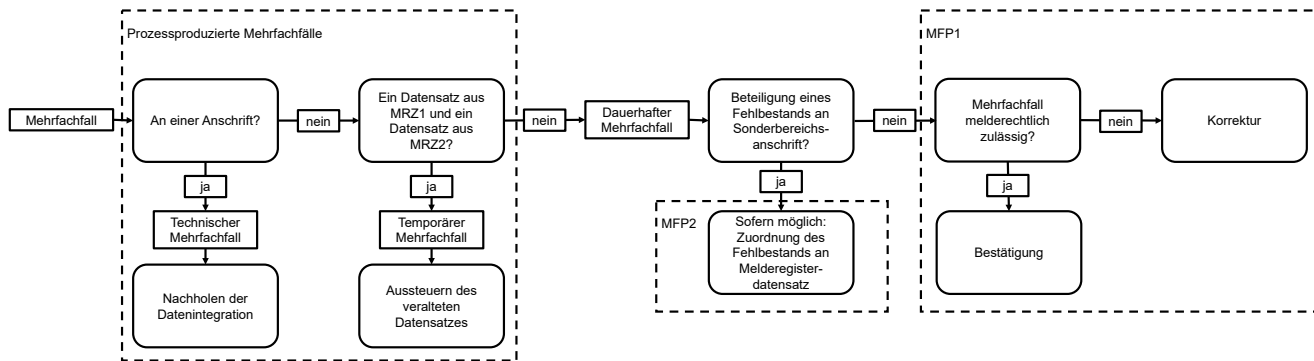
- › Mehrfache Meldungen des Wohnsitzes einer Person in den Melderegistern einer oder mehrerer Gemeinden sind gemäß § 21 Bundesmeldegesetz legitim, sofern zwischen Haupt- und Nebenwohnung unterschieden wird. Hier galt es, den Wohnungsstatus melderechtlich zulässiger Mehrfachfälle durch die Mehrfachfallprüfung zu bestätigen.
- › Unzulässige dauerhafte Mehrfachfälle liegen bei mehreren einwohnerzahlrelevanten Melderegisterdatensätzen (alleinige Wohnung und Hauptwohnung) einer Person vor. Hier war genau ein Melderegisterdatensatz zu bestimmen, welcher die Person repräsentiert.

1 Sonderbereiche bezeichnen Anschriften, die sich aufgrund des Meldeverhaltens ihrer Bewohnenden von gewöhnlichen Anschriften unterscheiden, etwa aufgrund einer besonders hohen Fluktuation (zum Beispiel Studierendenunterkünfte) oder abweichenden Meldeverpflichtungen (zum Beispiel Kasernen).

2 Zur Einführung des standardisierten Datenaustauschs im Meldewesen siehe Tramer (2007) und für nähere Erläuterungen zu XMeld siehe KoSIT (2021a; 2021b).

Grafik 1

Arten der Mehrfachfälle im Zensus 2022 – Entstehung und Handlungsbedarfe



MRZ: Melderegisterauszug; MFP: Mehrfachfallprüfung

Die weiteren einwohnerzahlrelevanten Melderegisterdatensätze dieser Person wurden genau wie eine Nebenwohnung ohne Anbindung an einen einwohnerzahlrelevanten Datensatz als nicht existent ausgezeichnet. Diese Suche und Verarbeitung entsprechender dauerhafter Mehrfachfälle heißt nachfolgend Mehrfachfallprüfung 1 (MFP1).

- › Für bestimmte Sonderbereichsanschriften besteht nach dem Bundesmeldegesetz eine Ausnahme von der Meldepflicht. Hier wurde geprüft, ob die Fehlbestände der Sonderbereichserhebung tatsächlich Untererfassungen der Melderegister aufdeckten oder ob die durch den Fehlbestand der Sonderbereichserhebung repräsentierte Person auch in den Melderegistern enthalten war. In einem an die Mehrfachfallprüfung anschließenden Prozess wurde dann entschieden, wo diese Person gezählt wurde (Bretsch und andere, 2024). Diese Suche wird nachfolgend MFP2 genannt.

## 4

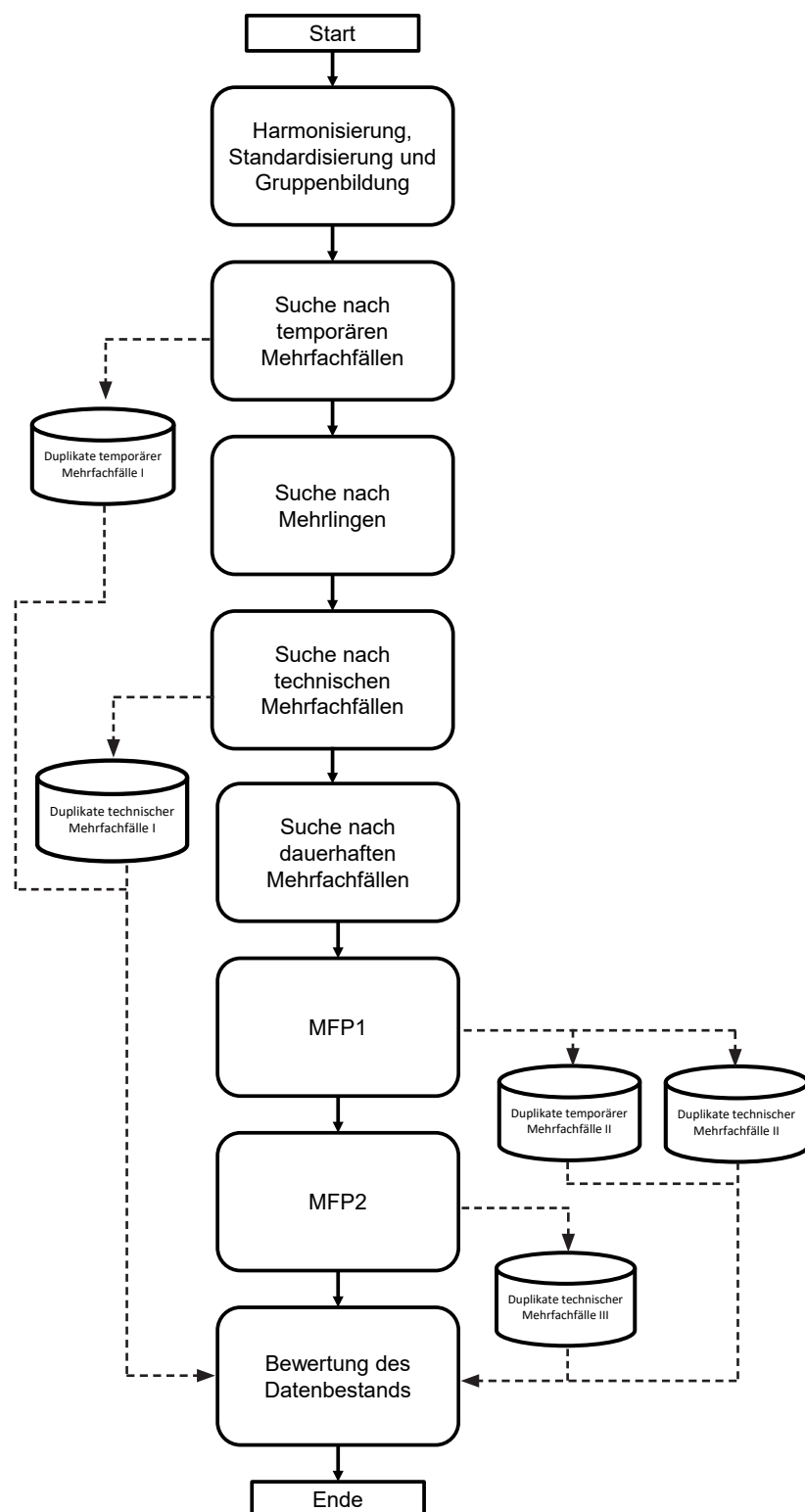
### Ablauf der Mehrfachfallprüfung

Der Suche nach Mehrfachfällen lagen die Annahmen zugrunde, dass sich bestimmte Merkmale einer Person im Lebensverlauf nicht oder nur selten ändern und dass diese Änderungen mit den Informationen aus den Melderegistern oder den primärstatistischen Erhebungen nachvollzogen werden können. Hierzu wurden die Merkmale Geschlecht, Geburtsdatum, Geburtsort und die vorhandenen Namensinformationen als „Quasi-Identifikatoren“ (Schnell, 2020, hier: Seite 151 f.) genutzt: Über gleiche oder zulässig ähnliche Merkmalsausprägungen sollten Datensätze identifiziert werden, welche mit an Sicherheit grenzender Wahrscheinlichkeit dieselbe Person im Personenbestand repräsentierten.

Die Mehrfachfallprüfung folgte dem in der Fachliteratur üblichen Vorgehen aus Harmonisierung und Standardisierung, Blocken, Vergleichen und Entscheidungsfindung (beispielsweise Christen, 2012, hier: Seite 23 f.; Schnell, 2016, hier: Seite 665 f.). Die manuelle Bewertung strittiger Mehrfachfälle im Zensus 2011 wurde gegen ein vollständig maschinelles Verfahren ausgetauscht. Hierbei ersetzte die weiter unten beschriebene Analyseeinheit „Kette“ die bis dato zentrale Analyseeinheit „Dublette“ für das Verarbeiten und Bewerten der Mehrfachfälle. ➡ Grafik 2

Grafik 2

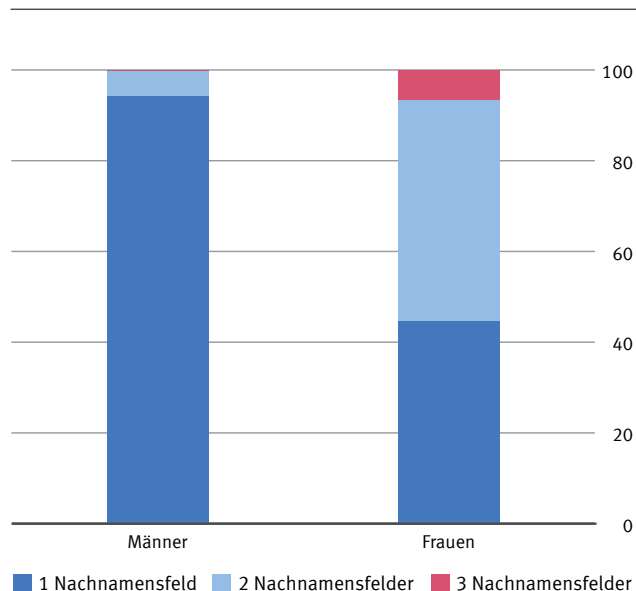
Schematische Darstellung der Mehrfachfallprüfung des Zensus 2022



## 4.1 Harmonisierung und Standardisierung

Da Nachnamen sich insbesondere mit der Eheschließung ändern können, wurden neben dem Vornamen und dem aktuellen Familiennamen sofern vorhanden auch der Geburtsname sowie der Familienname vor erster Änderung auf Ähnlichkeit abgeglichen. Dass in Deutschland deutliche Unterschiede zwischen Männern und Frauen in Hinblick auf Nachnamenswechsel im Lebensverlauf existieren, zeigt [Grafik 3](#).

**Grafik 3**  
Befüllte Nachnamensfelder vor Standardisierung in %



Die Angaben zum Geschlecht und zum Geburtsdatum wurden so übernommen, wie sie in den jeweiligen Ausgangsfeldern enthalten waren.<sup>13</sup> Für die fehleranfälligen Freitextfelder zu Namen und Geburtsort wurden folgende Standardisierungsschritte vorgenommen:

1. Alle Buchstaben kleinschreiben
2. Diakritika (Akzentzeichen), Umlaute sowie ß in ihre Stammbuchstaben überführen
3. Bindestriche durch Leerzeichen ersetzen

<sup>3</sup> Die Angaben zum Geschlecht enthalten die Ausprägungen „männlich“, „weiblich“, „divers“ und „ohne Angabe“. Aufgrund der geringen Besetzungen von „divers“ und „ohne Angabe“ werden die nachfolgenden Grafiken nur nach Männern und Frauen aufgeschlüsselt.

4. Sonderzeichen, Hochkommata und mehrfache sowie führende Leerzeichen entfernen

5. Klammern und deren Inhalte entfernen

Da gleiche Ausprägungen innerhalb der drei Nachnamensvarianten keinen informativen Mehrwert bringen, wurde der standardisierte Geburtsname nur behalten, sofern er vom aktuellen Familiennamen abwich. Der Familienname vor erster Änderung wurde nur dann behalten, wenn er wiederum weder dem standardisierten aktuellen Familiennamen noch dem standardisierten Geburtsnamen glich. Die Standardisierung des Geburtsortes wurde noch ausgeweitet: Lag im Ausgangsfeld eine Befüllung vor, welche auf Unkenntnis des Geburtsortsnamens hindeutete, beispielsweise „nicht bekannt“, dann blieb das standardisierte Geburtsortsfeld leer. Die sich aus der Standardisierung ergebenden Felder werden im Folgenden mit dem tiefgestellten Zusatz „standardisiert“ gekennzeichnet.

Zudem wurde für jedes der vier Namensfelder jeweils ein weiteres Feld angelegt, in welchen die in den jeweiligen standardisierten Namensfeldern enthaltenen Namen alphabetisch sortiert wurden.<sup>14</sup> Dies reduziert das Risiko, dass Personen mit Doppelnamen aufgrund unterschiedlicher Reihenfolgen der Namenseinträge nicht gefunden werden. Wie bei den standardisierten Nachnamensfeldern wurden die sortierten Geburts- und Familiennamensfelder vor erster Änderung nur bei den zuvor beschriebenen Abweichungen behalten. Die sich aus der Sortierung ergebenden Felder werden im Folgenden mit dem tiefgestellten Zusatz „sortiert“ gekennzeichnet.

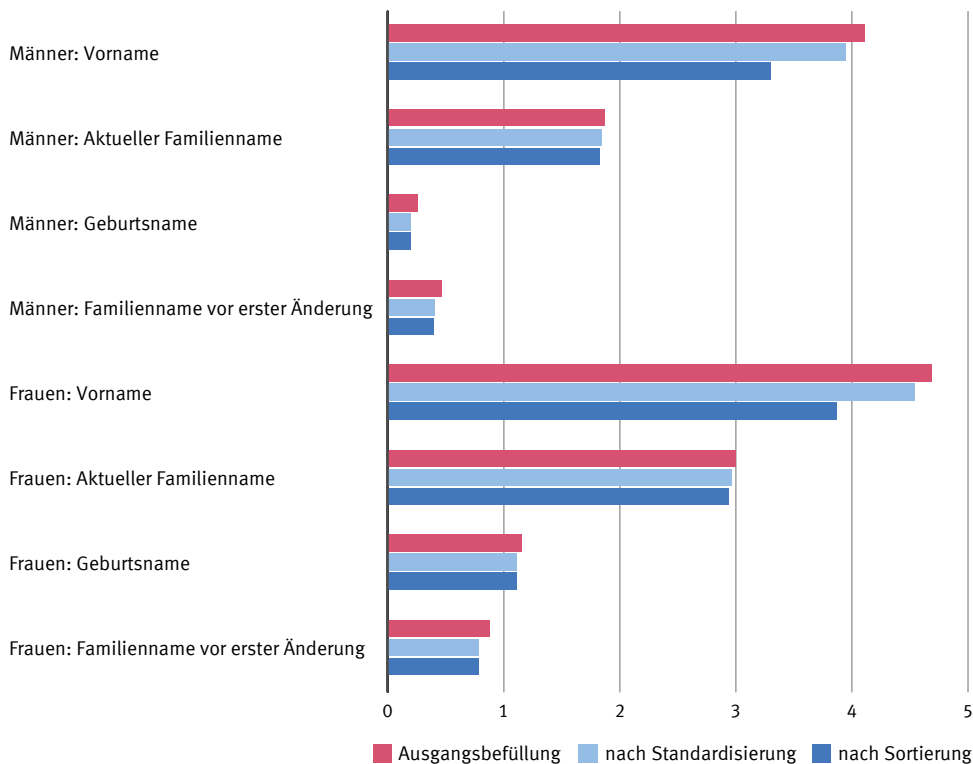
Zuletzt wurde das standardisierte Vornamensfeld jeweils mit den drei Nachnamensvarianten kombiniert und alphabetisch sortiert, sodass drei weitere, als Vollnamen bezeichnete Felder entstanden. Auch hier wurden die Vollnamensfelder nur dann behalten, wenn sie sich voneinander unterschieden. Dies kompensiert unter anderem Unterschiede in der Übertragung von einzelnen Namen in die jeweiligen Felder für Personen aus Kulturkreisen, in welchen nicht in Vor- und Familiennamen aufgeteilt wird, oder einfache Vertauschungen in den Namenseingabefeldern. Die Wirkung der beschriebenen Aufbereitungsschritte stellt [Grafik 4](#) dar.

<sup>4</sup> Die Standardisierung und Sortierung wurde auch für die Vornamen der ersten zehn Kinder und die Vor- und aktuellen Familiennamen der Ehe-/Lebenspartner sowie der gesetzlichen Vertreter angewandt.



## Grafik 4

Unterschiedliche Ausprägungen der Namensfelder  
Mill.



Auch die Felder zur aktuellen Anschrift und zur Zuzugsanschrift, welche nur für Melderegisterdatensätze zur Verfügung stand, wurden vergleichbar zu den Namensfeldern standardisiert. Die unterschiedlichen Einträge im Straßensfeld der aktuellen und der Zuzugsanschrift wurden jeweils auf Basis des im Zensus 2011 entwickelten Straßenthesaurus (Schöneich, 2012) standardisiert. Des Weiteren wurde für alle Melderegisterdatensätze ein Datumsfeld regelbasiert aus den verschiedenen Meldungsdaten zur Wohnung generiert. Dieses diente in der Mehrfachfallprüfung als Proxy für den tatsächlichen Wohnbeginn an einer Anschrift mit dem aktuellen Wohnstatus (Universaldatum).

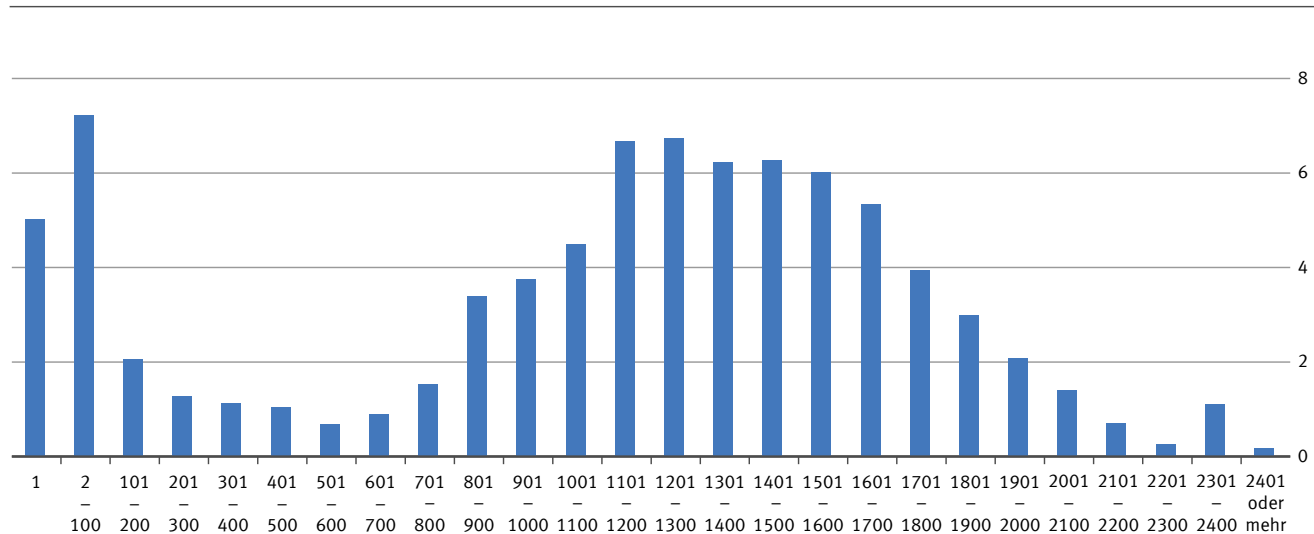
## 4.2 Gruppenbildung

Die Felder zum Geschlecht und zum Geburtsdatum grenzten distinkte Suchräume für die paarweisen Abgleiche ab (Blocking). Datensätze, welche in diesen beiden Feldern identische Ausprägungen aufwiesen, gehörten zu derselben Obergruppe, innerhalb der die Namensabgleiche stattfanden. Insgesamt gab es 81 566 Obergruppen, sie enthielten im Durchschnitt 1 094 Datensätze. [↗ Grafik 5](#)

Um die Anzahl der paarweisen Abgleiche innerhalb der Obergruppen noch weiter zu reduzieren, wurde der Geburtsort<sub>standardisiert</sub> als weitere, gleitende Einschränkung des Suchraums herangezogen. Innerhalb jeder Obergruppe musste sich der erste Buchstabe des Geburtsort<sub>standardisiert</sub> gleichen oder mindestens einer der abzugleichenden Datensätze wies einen leeren Geburtsort<sub>standardisiert</sub> auf. Dies reduzierte die durchgeführten Abgleiche um 93%: Während ohne Berücksichtigung des ersten Buchstabens des Geburtsort<sub>standardisiert</sub> rund 65,3 Milliarden paarweise Abglei-

**Grafik 5**

Obergruppen nach gruppierter Anzahl enthaltener Datensätze in 1 000



che durchzuführen gewesen wären, waren es so lediglich rund 4,61 Milliarden.

### 4.3 Vergleichen

Um die Ähnlichkeit der Namensinformationen zu bestimmen, wurde wie im Zensus 2011 auf die Jaro-Winkler-Stringmetrik ( $JW$ ) zurückgegriffen. Hierbei handelt es sich um eine Metrik, bei welcher die Ähnlichkeit zweier Strings ( $a$  und  $b$ ) auf der Basis der Reihenfolge identischer Zeichen innerhalb eines von der Länge der zu vergleichenden Strings abhängigen zulässigen Suchradius durch einen Wert im Wertebereich null bis eins ausgedrückt wird (Winkler, 1990). Für die Anwendung in der Mehrfachfallprüfung wurden die von Winkler (1990) vor-

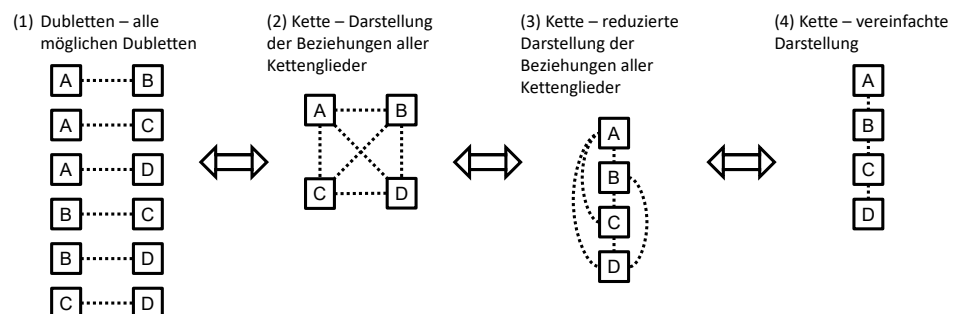
geschlagenen Grundeinstellungen genutzt. Allerdings wurden auch namentrennende Leerzeichen innerhalb des zulässigen Positionsradius als gemeinsame Zeichen gewertet.

### 4.4 Dubletten und Ketten

Wenn die miteinander verglichenen Datensätze (Elemente) einen bestimmten Schwellenwert erreichten oder überschritten ( $JW(a,b) \geq a$ ), wurden sie als initial paarig betrachtet und als sogenannte Dubletten weiterverarbeitet. Dubletten dienen als Zwischenergebnis der Identifizierung von Mehrfachfällen und als Voraussetzung für die Verkettung sowie für die anschließende detaillierte Prüfung der als potenzielle Mehrfachfälle

**Grafik 6**

Zusammenhang der Konzepte Dublette und Kette



derselben Person ausgemachten Datensätze. Ketten entstehen aus einer oder mehreren Dubletten mit mindestens einem gemeinsamen Element. Aus den beiden Dubletten  $D(A,B)$  und  $D(B,C)$  kann die dreigliedrige Kette  $K(A,B,C)$  abgeleitet werden. In sprachlicher Abgrenzung zur Dublette werden die distinkten Elemente einer Kette als Kettenglieder bezeichnet. Ketten können auch in die maximal mögliche Anzahl sie konstituierender Dubletten zurückgewandelt werden. [↗ Grafik 6](#)

Ein wesentlicher Vorteil der Kette im Vergleich zur Dublette besteht für die Datenverarbeitung darin, dass alle derselben Kette angehörigen Datensätze in einem Long-Format abgelegt und mit einer gemeinsamen Identifikationsnummer für die jeweilige Kette versehen werden können.

## 4.5 Maschinelle Klärung strittiger Fälle: Kettentransitivität und Repräsentanz

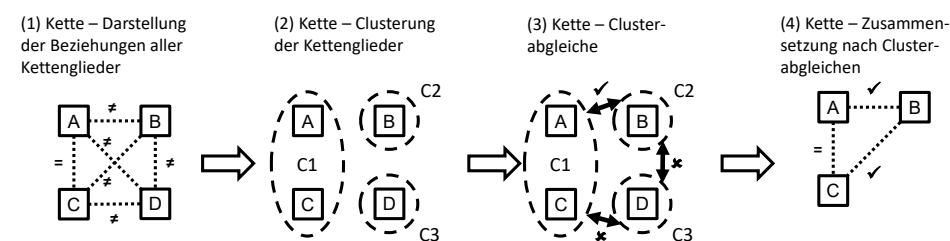
Um sicherzustellen, dass alle Kettenglieder einer Kette dieselbe Person repräsentieren, muss jede Kette vollständig transitiv sein. Aufgrund der Verwendung des Schwellenwerts  $\alpha$  und dem Vorgehen bezüglich des Geburtsorts konnten zwei Fälle vorkommen: Datensätze  $(A,C)$  wurden aufgrund der initialen Paarigkeitsbestimmung mit einem weiteren Datensatz  $(B)$  zu Kettengliedern derselben Kette verknüpft, welche im direkten Abgleich nicht paarig bewertet wurden ( $\alpha \leq JW(a_A, b_B) < 1, \alpha \leq JW(a_B, b_C) < 1$  und  $JW(a_A, b_C) < \alpha$ ). Oder es wurden mitunter auch Datensätze zu Dubletten, deren Geburtsorte<sub>standardisiert</sub> sich ab dem zweiten Buchstaben unterschieden. Derart „strittige“ Ketten wurden näher untersucht. Hierzu wurden die Kettenglieder strittiger Ketten auf Basis der Ausprägungen der standardisier-

ten Namensfelder sowie dem Geburtsort<sub>standardisiert</sub> in distinkte Cluster unterteilt. Im Anschluss wurden die Kettenglieder verschiedener Cluster derselben Kette mithilfe weiterer Informationen aus den Melderegisterfeldern auf Gemeinsamkeiten abgeglichen (Clusterabgleich). Der Vorteil dieser nachgelagerten finalen Paarigkeitsprüfung bestand darin, dass die Anzahl der aufwendigen Abgleiche auf die initial verketteten Datensätze, welche Namens- beziehungsweise Geburtsortsdiskrepanzen aufwiesen, begrenzt wurde. Lagen keine Informationen vor, welche die Diskrepanz zwischen den Clustern innerhalb einer Kette überbrückten, so wurde die strittige Kette entlang der untereinander unpaarigen Cluster aufgetrennt. [↗ Grafik 7](#)

Jede Person darf nur durch einen einwohnerzahlrelevanten Datensatz im Personenbestand repräsentiert werden. Um dies zu erreichen, wurde für jede Kette ein führendes Kettenglied bestimmt. Alle weiteren Kettenglieder wurden als Duplikate bezeichnet und Datensätze, die nicht mit mindestens einem weiteren Datensatz verkettet werden konnten, als Unikate. Das führende Kettenglied wurde auf Basis fester Regeln bestimmt, welche sich darin unterschieden, ob es sich um prozessproduzierte Mehrfachfälle, Mehrfachfälle der Melderegister oder Mehrfachfälle mit Fehlbeständen der Sonderbereichserhebung handelte. Das führende Kettenglied ermöglichte, jede Kette in die Anzahl der zur Kettenkonstitution notwendigen Dubletten zu überführen, indem jedes Duplikat einzig auf das führende Kettenglied der Kette verweist. Dies war wichtig, damit die an die Mehrfachfallprüfung anschließenden Prozesse, welche aufgrund ihrer IT-Infrastruktur teilweise keine zeilenübergreifenden Operationen zuließen, die Ergebnisse der Mehrfachfallprüfung verarbeiten konnten. [↗ Grafik 8](#)

### Grafik 7

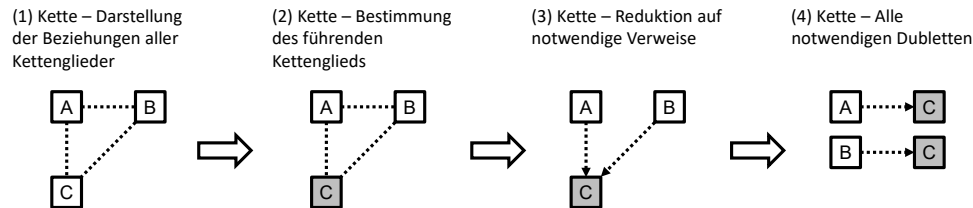
Schematische Darstellung zum Vorgehen des Clusterabgleichs zur Sicherstellung von Kettentransitivität



Erläuterungen: = identische Ausprägung; ≠ unterschiedliche Ausprägung; ✓ Gemeinsamkeiten vorhanden; \* keine Gemeinsamkeiten vorhanden.

Grafik 8

## Schematische Darstellung zum Vorgehen der Bestimmung des führenden Kettenglieds



Anmerkung: Der grau hinterlegte Datensatz wurde als führendes Kettenglied bestimmt.

## 5

## Suche nach prozessproduzierten Mehrfachfällen

Zunächst wurde der Personenbestand um die prozessproduzierten Mehrfachfälle bereinigt. Für die Suche nach temporären Mehrfachfällen wurden die ausschließlich in MRZ1 gelieferten mit den ausschließlich in MRZ2 gelieferten Datensätzen innerhalb der Untergruppen abgeglichen. Diese Suche zielte darauf ab, triviale temporäre Mehrfachfälle zu finden. Daher wurde die Gleichheit der standardisierten Vornamen sowie mindestens einer der standardisierten Nachnamensvarianten zur Paarigkeitsbestimmung genutzt. Die nicht trivialen temporären Mehrfachfälle wurden bei der späteren Suche nach dauerhaften Mehrfachfällen gefunden. Im Anschluss wurden die Dubletten entlang übereinstimmender Elemente verknüpft. Sofern mehr als zwei Datensätze aus MRZ2 in der entstandenen Kette enthalten waren, wurde der gemäß Universaldatum jüngste Datensatz als führendes Kettenglied ausgewählt und alle anderen MRZ2-Datensätze aus der Kette ausgesteuert. Alle Duplikate wurden mit einem Verweis zum entsprechenden führenden Kettenglied versehen, in eine Tabelle geschrieben und aus den nachfolgenden Prozessen der Mehrfachfallprüfung ausgesteuert.

Für technische Mehrfachfälle sollte bei der Integration der als Duplikate markierten Datensätze in die führenden Kettenglieder verhindert werden, dass im Elternhaus zusammenlebende gleichgeschlechtliche Mehrlinge mit ähnlich klingenden Vornamen oder mit gleichen Vornamen in unterschiedlicher Reihenfolge versehentlich zusammengefasst wurden. Daher wurden zum einen unter allen an einer Anschrift befindlichen Datensätzen derselben Untergruppe nach Dubletten mit identischen standardisierten Nachnamensvarian-

ten, aber unterschiedlichen standardisierten Vornamen gesucht, sofern sie zum Stichtag das 27. Lebensjahr noch nicht vollendet hatten.<sup>15</sup> Zum anderen wurden gleichgeschlechtliche Mehrlinge in den Melderegisterdatensätzen der Eltern identifiziert. Sofern diese Mehrlinge im Personenbestand gefunden werden konnten, wurden auch sie zu Mehrlingsketten verknüpft in einer Tabelle gesammelt.

Im Anschluss wurde an allen Anschriften mit mindestens zwei Datensätzen von alleinigen Wohnungen und Hauptwohnungen nach technischen Mehrfachfällen gesucht. Aufgrund des beschränkten Suchraums konnte auch eine Abweichung von einem Jahr im Geburtsdatum als weitere gleitende Einschränkung des Suchraums genutzt werden. Da in erster Linie triviale technische Mehrfachfälle gefunden werden sollten, wurde Übereinstimmung in mindestens einer Vollnamensvariante als Paarigkeitskriterium gewählt. Die paarigen Datensätze wurden verkettet und auf enthaltene Mehrlinge gemäß der zuvor erstellten Mehrlingsketten kontrolliert. Innerhalb der mehrlingsfreien Ketten wurden anschließend die führenden Kettenglieder bestimmt. An Anschriften der Sonderbereichserhebung sowie der Haushaltsstichprobe wurde dafür der Existenzbefund gemäß Erhebung und an den sonstigen Anschriften das jüngste Universaldatum als Kriterium herangezogen. Anschließend wurde die Integration der Duplikate in die führenden Kettenglieder nachgeholt. Die Duplikate wurden dann in eine Tabelle geschrieben und aus den nachfolgenden Prozessen der Mehrfachfallprüfung ausgesteuert.

<sup>15</sup> Die Anzahl der im Elternhaus lebenden Kinder nimmt mit dem Alter ab: Während etwa drei Viertel der 25-Jährigen in Deutschland im Jahr 2022 das Elternhaus bereits verlassen hatten, waren es bei den 30-Jährigen bereits über 90 % (Statistisches Bundesamt, 2023). Um auch möglichen systematischen, kulturell bedingten Verzögerungen der Haushaltsgründung von Menschen mit Migrationshistorie gerecht zu werden, wurden auf das relativ zeitkonstante Durchschnittsalter von etwa 24 Jahren für das Verlassen des Elternhauses in Deutschland noch drei Jahre aufgeschlagen.

## 6

### Suche nach dauerhaften Mehrfachfällen

Nach Aussonderung der trivialen Duplikate temporärer und technischer Mehrfachfälle wurde innerhalb der beschriebenen Untergruppen<sup>6</sup> für alle abzugleichenden Datensätze der maximale  $JW(a,b)$  über die neun Vergleichskombinationen der drei Vornamen von Datensatz 1 ( $v_i$ ) und Datensatz 2 ( $v_j$ ) berechnet ( $\beta$ ), wobei die Vornamen sowohl von links nach rechts ( $JW(\vec{v}_i, \vec{v}_j)$ ) als auch von rechts nach links ( $JW(\vec{v}_j, \vec{v}_i)$ ) abgeglichen wurden:

$$\beta = \text{MAX} \left( JW(\vec{v}_i, \vec{v}_j), JW(\vec{v}_j, \vec{v}_i) \right)$$

wobei

$v_i$  = Vorname i von Datensatz 1 mit  $i=1,2,3$

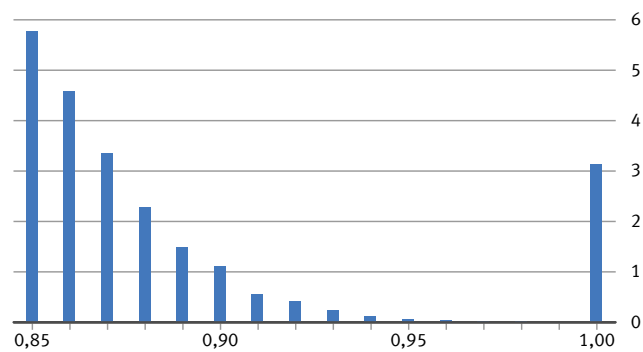
$v_j$  = Vorname j von Datensatz 2 mit  $j=1,2,3$ .

Als Grenzwert für die initiale Bewertung der Paarigkeit der verglichenen Datensätze wurde 0,85 gewählt. Dieser Grenzwert ergab sich aus der Abwägung, einerseits möglichst wenige Datensätze als fälschlicherweise nicht paarig zu klassifizieren. Hierauf zielten auch die Abgleiche in und gegen Leserichtung ab, um nicht erfasste Zweitnamen zu kompensieren, welche durch die Sortierung der Namen innerhalb der Vornamen einen Einfluss auf den JW-Wert haben konnten.<sup>7</sup> Auch die Wahl des maximalen JW-Werts aller neun Vornamensabgleiche beider Vergleichsrichtungen sollte die Gefahr minimieren, vorzeitig Datensätze als unpaarig zu klassifizieren, welche sich mithilfe weiterer Informationen noch als Mehrfachfall derselben Person entpuppen konnten. Andererseits führte jede Absenkung des Schwellenwerts zu einem starken Anstieg der entstehenden Dubletten und damit zu einem Anstieg der notwendigen Rechen-

operationen der folgenden Schritte. Als Zwischenergebnis der Suche nach dauerhaften Mehrfachfällen ergab sich das in [Grafik 9](#) dargestellte maximale Dubletten-set mit 23,216 Millionen Dubletten.

**Grafik 9**

Dauerhafte Dubletten nach maximaler Jaro-Winkler-Distanz über alle Abgleichskombinationen der Vornamen Mill.



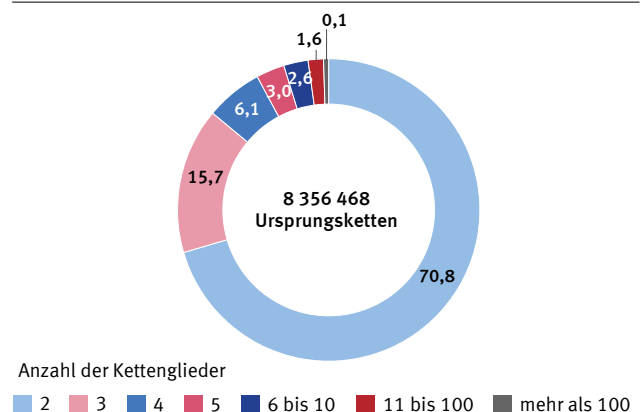
Die Werte für die kaum sichtbaren Dubletten zwischen den Jaro-Winkler-Werten zwischen 0,90 und 1,00 sind: 562 000 (0,91); 420 000 (0,92); 233 000 (0,93); 116 000 (0,94); 59 000 (0,95); 30 000 (0,96); 19 000 (0,97); 22 000 (0,98); 6 000 (0,99)

Aus der Verkettung der dauerhaften Dubletten entstanden 8,357 Millionen sogenannte Ursprungsketten mit einer durchschnittlichen Länge von drei Kettengliedern, welche die Basis der nachfolgenden Schritte bildeten.

[Grafik 10](#)

**Grafik 10**

Ursprungsketten nach Anzahl der Kettenglieder in %



6 Für Fehlbestände der Haushaltsstichprobe bestand noch die Zusatzeinschränkung, dass sie nur innerhalb der Anschrift mit Melderegisterdaten abgeglichen wurden. Hintergrund dafür war, tatsächliche Aufdeckungen von Untererfassungen innerhalb der Melderegister von prozessproduzierten Fehlbeständen der Haushaltsstichprobe abzugrenzen, welche das Ergebnis einer nicht erfolgten Zusammenführung mit dem Melderegisterpendant der Person waren.

7 Zum Beispiel bei  $v_i$  „judith-weiss“ und  $v_j$  „anita-judith-weiss“ ergibt  $JW(\vec{v}_i, \vec{v}_j) = 0,74$  und  $JW(\vec{v}_j, \vec{v}_i) = 0,93$ .

## 6.1 Suche nach Mehrfachfällen in den Melderegistern (MFP1)

Aus den Ursprungsketten wurden die miteinander verketteten Melderegisterdatensätze herausgezogen und als initiale MFP1-Ketten abgelegt. Innerhalb jeder als „strittig“ markierten MFP1-Kette wurde jedes Kettenmitglied eines Clusters mit allen Kettengliedern der anderen Cluster paarweise abgeglichen. Sofern eine der folgenden Bedingungen für mindestens einen der entstehenden paarweisen Abgleiche zutraf, galt die Diskrepanz zwischen den Clustern als überbrückt. Alle in diesen Clustern enthaltenen Datensätze galten dann als dieselbe Person repräsentierend.

- 1) Dieselbe aktuelle oder letzte Lebens- beziehungsweise Ehepartnerschaft konnte auf Basis der Lebenspartnerschafts- beziehungsweise Eheschließungsdaten oder der -auflösungsdaten in Kombination mit vier von fünf hinterlegten Angaben zum oder zur aktuellen oder letzten Lebens- oder Ehepartner/-partnerin (Vorname<sub>sortiert</sub>, Familienname<sub>sortiert</sub>, Geburtsname<sub>sortiert</sub>, Geschlecht, Geburtsdatum) ausgemacht werden.
- 2) Mindestens eines der hinterlegten Kinder stimmte in Bezug auf zwei von drei Merkmalen überein (Vorname<sub>sortiert</sub>, Geschlecht, Geburtsdatum).
- 3) Mindestens einer der hinterlegten gesetzlichen Vertreter stimmte in Bezug auf drei von vier Merkmalen überein (Vorname<sub>sortiert</sub>, Nachname<sub>sortiert</sub>, Geschlecht, Geburtsdatum).
- 4) Mindestens eine aktuelle Anschrift stimmte überein.
- 5) Mindestens eine aktuelle Anschrift stimmte mit mindestens einer Zuzugsanschrift überein.
- 6) Mindestens eine Zuzugsanschrift stimmte überein.

Die Bedingungen 3) bis 6) wurden allerdings nur als die Diskrepanz überbrückend gewertet, wenn einer der folgenden detaillierten Namensabgleiche zutraf:

- a) Der Vorname<sub>sortiert</sub> und mindestens einer der drei sortierten Nachnamensvarianten (Aktueller Familienname<sub>sortiert</sub>, Geburtsname<sub>sortiert</sub> und Familienname vor erster Änderung<sub>sortiert</sub>) stimmten überein.
- b) Mindestens einer der Vollnamen stimmte überein.

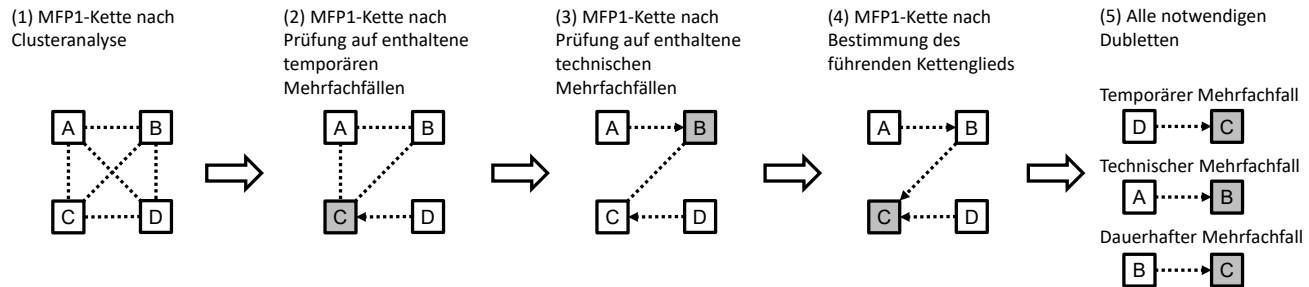
- c) Bei gleichem Vornamen<sub>sortiert</sub> betrug der maximale JW-Wert mindestens eines Abgleichs der sortierten Nachnamensvarianten mindestens 0,9.
- d) Bei einer unterschiedlichen Anzahl von Namen im Vornamen<sub>sortiert</sub> stimmte mindestens einer der drei sortierten Nachnamensvarianten überein und alle in dem weniger Worte beinhaltenden Vornamen<sub>sortiert</sub> enthaltenen Namen wurde mit einem JW-Wert von mindestens 0,95 in dem Vornamen<sub>sortiert</sub> des Vergleichsdatensatzes gefunden.
- e) Bei mindestens einer gleichen sortierten Nachnamensvariante betrug der JW-Wert Vorname<sub>sortiert</sub> mindestens 0,9.
- f) Bei einer unterschiedlichen Anzahl von Namen in den sortierten Nachnamensvarianten stimmte der Vorname<sub>sortiert</sub> überein und alle in der weniger Worte beinhaltenden sortierten Nachnamensvariante enthaltenen Namen wurden mit einem JW-Wert von mindestens 0,95 in einer der sortierten Nachnamensvarianten des Vergleichsdatensatzes gefunden.
- g) Bei einer unterschiedlichen Anzahl von Namen in den Vollnamensfeldern wurden alle in dem weniger Worte beinhaltenden Vollnamensfeld enthaltenen Namen mit einem JW-Wert von mindestens 0,95 in einem Vollnamensfeld des Vergleichsdatensatzes gefunden.

Wenn keine Diskrepanz überbrückenden Übereinstimmungen zwischen Clustern gefunden wurden, dann wurden die Ketten entlang der nicht übereinstimmenden Cluster aufgetrennt. So entstanden aus der initialen MFP1-Kette teilweise neue Ketten und teilweise Unikate. Alle nach den Clusterabgleichen bestehenden MFP1-Ketten wurden auf verkettete Mehrlinge geprüft und sofern notwendig entlang dieser Mehrlinge aufgetrennt. Innerhalb der bestehenden Ketten wurden erneut prozessproduzierte Mehrfachfälle gesucht und die zuvor beschriebenen Schritte zum Umgang mit prozessproduzierten Mehrfachfällen vollzogen. Als letztes wurde innerhalb jeder MFP1-Kette das einwohnerzahlrelevante Kettenglied mit dem jüngsten Universaldatum zum führenden Kettenglied bestimmt. Bei reinen Nebenwohnungsketten wurde ersatzweise die Nebenwohnung mit dem jüngsten Universaldatum ausgewählt. Alle Duplikate erhielten wieder den Verweis zum führenden Kettenglied. [➤ Grafik 11](#) stellt das beschriebene Vorgehen schematisch dar.



**Grafik 11**

Schematische Darstellung der sukzessiven Kontrolle der MFP1-Ketten auf enthaltene temporäre und technische Mehrfachfälle und der Bestimmung des Führenden Kettenglieds der MFP1-Kette



Anmerkung: Der grau hinterlegte Datensatz wurde in dem jeweiligen Abschnitt als führendes Kettenglied bestimmt.  
MFP1: Mehrfachfälle in den Melderegistern.

## 6.2 Suche nach Mehrfachfällen mit Fehlbeständen der Sonderbereichserhebung (MFP2)

Für die MFP2 wurden alle nach MFP1 nicht als temporäre oder technische Duplikate ausgesteuerten Kettenglieder der Ursprungsketten, welche mindestens einen Fehlbestand der Sonderbereichserhebung enthielten, als initiale MFP2-Ketten abgelegt. Innerhalb dieser MFP2-Ketten wurden die Fehlbestände der Sonderbereichserhebung mit allen als führende Kettenglieder hervorgegangenen Datensätzen sowie den Unikaten<sup>18</sup> paarweise abgeglichen. Hierbei wurde ein Fehlbestand der Sonderbereichserhebung nur dann einem Melde-registerdatensatz zugeordnet, um eine finale MFP2-Kette zu bilden, wenn die in MFP1 angewandten Bedingungen der Namensinformationsabgleiche zutrafen. Sofern aus diesen Namensinformationsabgleichen mehrere Melde-registerdatensätze hervorgingen, welchen derselbe Fehlbestand der Sonderbereichserhebung zugeordnet werden konnte, wurden zusätzlich die Informationen zum Familienstand, zur Staatsangehörigkeit und zum Geburtsstaat zur finalen Zuordnung herangezogen. Innerhalb der finalen MFP2-Ketten wurde nochmals nach technischen Mehrfachfällen gesucht und, sofern vorhanden, das Prozedere zum Umgang mit technischen Mehrfachfällen wiederholt. Als führendes Kettenglied wurde der jeweils enthaltene Fehlbestand der Sonder-

bereichserhebung ausgewählt und die Duplikate der finalen MFP2-Kette erhielten den Verweis zu diesem.

## 7

### Wirkung der Mehrfachfallprüfung auf den stichtagsrelevanten Personenbestand

Nach dem Durchlauf der MFP2 wurden die prozessproduzierten Mehrfachfälle konsolidiert: Hatte sich das führende Kettenglied eines temporären Mehrfachfalls im Zuge der sukzessiven Abfolge der Schritte der Mehrfachfallprüfung als Duplikat eines technischen Mehrfachfalls entpuppt, dann wurden die Verweise der zugehörigen Duplikate auf das entsprechende führende Kettenglied dieses technischen Mehrfachfalls umgewidmet. Hier-nach wurde der Personenbestand konsolidiert, wobei alle Duplikate der prozessproduzierten Mehrfachfälle als nicht existent gemäß der Melderegisterkonsolidierung ausgezeichnet wurden ( $\text{Existenz}_{\text{MRKONS}}$ ). Die übrigen Melderegisterdatensätze erhielten abhängig vom initialen Wohnungsstatus aus dem Melderegister ( $\text{Wohnungsstatus}_{\text{MR}}$ ) sowie der Anzahl der in der finalen MFP1-Kette enthaltenen alleinigen Wohnungen, Hauptwohnungen und Nebenwohnungen eine Existenzsetzung nach MFP1 ( $\text{Existenz}_{\text{MFP1}}$ ) und einen Wohnungsstatus nach MFP1 ( $\text{Wohnungsstatus}_{\text{MFP1}}$ ). Die Bewertung der Datensätze gemäß Mehrfachfallprüfung zeigt [Übersicht 1](#).

<sup>18</sup> Dies umfasst sowohl die im Lauf der MFP1 aus der initialen MFP1-Kette herausgelösten Datensätze als auch jene, welche die einzigen Melderegisterdatensätze innerhalb der Ursprungskette ausmachten und daher nicht in die MFP1 eingingen.

## Übersicht 1

### Vergabe der Existenz und des Wohnungsstatus gemäß MFP1

Art des Mehrfachfalls	Wohnungsstatus <sub>MR</sub>	Existenz <sub>MRKONS</sub>	Anzahl weiterer MFP1-Kettenglieder mit Wohnungsstatus <sub>MR</sub> AW/HW	Anzahl weiterer MFP1-Kettenglieder mit Wohnungsstatus <sub>MR</sub> NW	Existenz <sub>MFP1</sub>	Wohnungsstatus <sub>MFP1</sub>
Duplikat eines temporären Mehrfachfalls	AW/HW	n				
Duplikat eines technischen Mehrfachfalls	AW/HW	n				
Unikat nach MFP1	AW	e	0	0	e	AW
	HW	e	0	0	e	AW
	NW	e	0	0	n	
Duplikat einer finalen MFP1-Kette	AW/HW	e	≤ 1	≤ 0	n	
	NW	e	0	≤ 1	n	
	NW	e	≤ 1	≤ 0	e	NW
Führendes Kettenglied einer finalen MFP1-Kette	AW	e	≤ 1	0	e	AW
	HW	e	≤ 1	0	e	AW
	AW	e	≤ 0	≤ 1	e	HW
	HW	e	≤ 0	≤ 1	e	HW
	NW	e	0	≤ 1	n	

MFP1: Mehrfachfälle in den Melderegistern; MR: Melderegister; AW: Alleinige Wohnung; HW: Hauptwohnung; NW: Nebenwohnung; n: nicht existent; e: existent

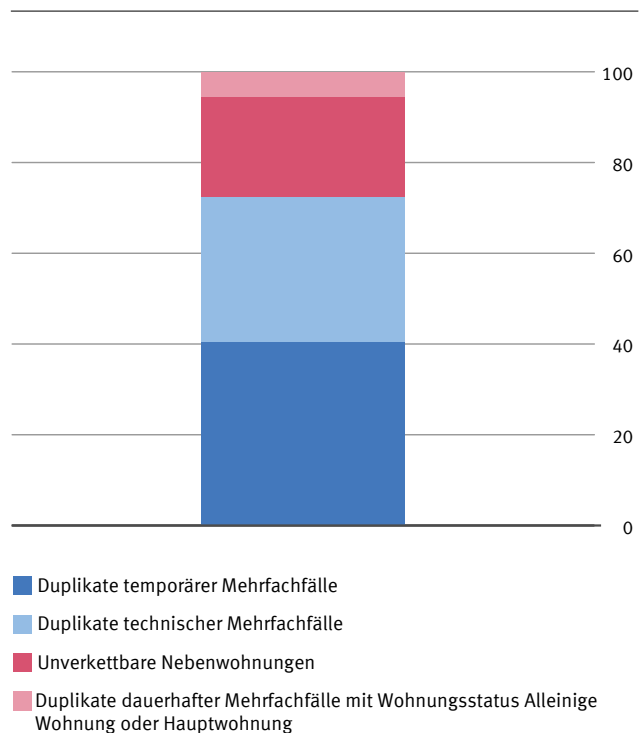
Die Ergebnisse dieser (Nicht-)Existenzsetzung und Wohnungsstatusvergabe gemäß Mehrfachfallprüfung zeigen die Grafiken 12 und 13: Den Großteil der rund 692 700 als nicht existent markierten Datensätze machten die prozessproduzierten Mehrfachfälle an verschiedenen Anschriften mit rund 279 900 (40,4 %) sowie an einer Anschrift mit rund 222 500 (32,1 %) aus. Weitere 152 700 (22 %) Nicht-Existenzsetzungen entfielen auf Nebenwohnungen, welche keinem einwohnerzahlrelevanten Melderegisterdatensatz zugeordnet werden konnten. Die überzähligen einwohnerzahlrelevanten Melderegisterdatensätze beliefen sich auf rund 37 600 (5,4 %). Dass trotz der im Vergleich zum Zensus 2011 erweiterten Methode verhältnismäßig wenige melderegisterimmanente, melderechtlich unzulässige einwohnerzahlrelevante Mehrfachfälle gefunden wurden, ist mutmaßlich auf den automatisierten Datenaustausch der melderegisterführenden Stellen zurückzuführen.

#### ↳ Grafik 12

Innerhalb des Personenbestands reduzierte die Mehrfachfallprüfung die Anzahl der alleinigen Wohnungen um 0,7 %. Allerdings war diese Abnahme nicht nur auf Nichtexistenzsetzungen zurückzuführen: In einigen Fällen konnten eine oder mehrere Nebenwohnungen einer originären alleinigen Wohnung zugeordnet werden, sodass der Wohnungsstatus auf Hauptwohnung umge-

#### Grafik 12

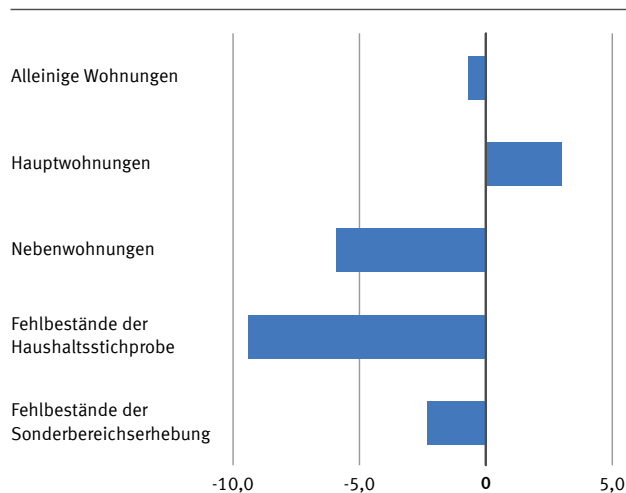
Grund für die Nicht-Existenzsetzung gemäß Mehrfachfallprüfung in %



ändert wurde. Hierdurch ist die Zunahme der Zahl der Hauptwohnungen nach Mehrfachfallprüfung um 3,0 % im Prä-Post-Vergleich zum initialen Personenbestand zu erklären. Die Abnahme der Zahl der Nebenwohnungen um 5,9 % ist vorwiegend auf die Nicht-Existenzsetzung jener Nebenwohnungen zurückzuführen, für welche keine einwohnerzahlrelevanten Datensätze gefunden wurden. Hierbei handelte es sich vermutlich um inaktive Nebenwohnsitze, wie ein Vergleich der durchschnittlichen Universaldaten zeigte: Während das Universaldatum der durch die Mehrfachfallprüfung in ihrer Existenz bestätigten Nebenwohnungen im Durchschnitt auf März 2011 datierte, betrug das durchschnittliche Universaldatum der nicht existent gesetzten Nebenwohnungen April 1992. Mithilfe der erweiterten Methode der Mehrfachfallprüfung konnten 9,4 % der initialen Fehlbestände der Haushaltsstichprobe nachträglich mit einer originären Karteileiche aus den Melderegistern an derselben Anschrift zusammengeführt werden. Dies gelang auch für 2,3 % der initialen Fehlbestände der Sonderbereichserhebung. ➤ Grafik 13


### Grafik 13

Veränderung der Datensatzart im Personenbestand nach gegenüber vor der Mehrfachfallprüfung in %



## 8

### Fazit

Die Mehrfachfallprüfung ist eine wichtige qualitätssichernde Maßnahme im Personenbestand des Zensus 2022. Der Perspektivwechsel von der Dublette zur Kette als zentrale Analyseeinheit und der Einsatz der Clusteranalyse strittiger Ketten ermöglichte die Entwicklung eines vollständig maschinellen Verfahrens, um Mehrfachfälle zu suchen und zu bewerten. Im Ergebnis wurden vor allem prozessproduzierte Mehrfachfälle aufgedeckt. 

## LITERATURVERZEICHNIS

---

Bretsch, Corinna/Seibel, Steffen/Vorndran, Ingeborg/Meyn, Christoph. [\*Ermittlung der Einwohnerzahl im Zensus 2022\*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 6/2024, Seite 17 ff.

Christen, Peter. *Data Matching. Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection*. Heidelberg 2012. DOI: [10.1007/978-3-642-31164-2](#)

Diehl, Eva-Maria. [\*Methoden der Mehrfachfallprüfung im Zensus 2011\*](#). In: Wirtschaft und Statistik. Ausgabe 6/2012, Seite 473 ff.

Dittrich, Stefan/Bretsch, Corinna/Stepien, Halina Danuta/Vorndran, Ingeborg/Michel, Bernd/Kleber, Birgit/Timm, Ulrike/Pfahl, Miriam. [\*Der Zensus 2022 – mit Online First an der Schwelle zu einem Registerzensus\*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 6/2022, Seite 90 ff.

Boragk, Lisa/Gaedke, Annika/Gemmeke, Charlotte/Meyn, Christoph. [\*Die Ermittlung des Berichtskreises am Beispiel der Sonderbereiche im Zensus 2022\*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 6/2024, Seite 41 ff.

Dreschmitt, Kai/Pfahl, Miriam. [\*Die Personenerhebung im Zensus 2022\*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 6/2024, Seite 29 ff.

Klink, Steffen/Lorentz, Kai. [\*Auswahlplan und Stichprobenhauptziehung für den Zensus 2022\*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 1/2022, Seite 13 ff.

KoSIT (Koordinierungsstelle für IT-Standards). *XMeld*. 2021a [Zugriff am 7. Januar 2025]. Verfügbar unter: [www1.osci.de](http://www1.osci.de)

KoSIT (Koordinierungsstelle für IT-Standards). *Datensatz für das Meldewesen. Einheitlicher Bundes-/Länderteil (DSMeld)*. Bearbeitungsstand: 28. Oktober 2021 (12. Änderung, wirksam ab 1. November 2021). 2021b. [Zugriff am 7. Januar 2025]. Verfügbar unter: [www1.osci.de](http://www1.osci.de)

Schnell, Rainer. *Record Linkage*. In: Wolf, Christof/Joye, Dominique/Smith, Tom W./Fu, Yang-Chih (Herausgeber). *The SAGE Handbook of Survey Methodology*. London 2016. Seite 662 ff.

Schnell, Rainer. *Record Linkage als zentraler Baustein der Forschung mit Registern und Big Data-Nutzungen*. In: Klumpe, Bettina/Schröder, Jette/Zwick, Markus (Herausgeber). *Qualität bei zusammengeführten Daten. Befragungsdaten, administrative Daten, neue digitale Daten: miteinander besser?* Wiesbaden 2020. Seite 151 ff. DOI: [10.1007/978-3-658-31009-7\\_11](#)

Schöneich, Cordula. [\*Der Straßenthesaurus im Zensus 2011\*](#). In: Wirtschaft und Statistik. Ausgabe 11/2012, Seite 957 ff.

## LITERATURVERZEICHNIS

---

Statistisches Bundesamt. *Junge Menschen verlassen ihr Elternhaus im Schnitt im Alter von 23,8 Jahren*. Zahl der Woche Nr. 36 vom 5. September 2023. [Zugriff am 7. Januar 2025]. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

Tramer, Karl. *Elektronischer Datenaustausch im Meldewesen. Modell für überörtlichen Datenaustausch in Deutschland?* In: VM Verwaltung & Management. Jahrgang 13. Heft 6/2007, Seite 300 ff. DOI: [10.5771/0947-9856-2007-6](https://doi.org/10.5771/0947-9856-2007-6)

Winkler, William E. *String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage*. 1990. [Zugriff am 7. Januar 2025]. Verfügbar unter: [eric.ed.gov](http://eric.ed.gov)

## RECHTSGRUNDLAGEN

---

Bundesmeldesetz vom 3. Mai 2013 (BGBl. I Seite 1084), das zuletzt durch Artikel 6 des Gesetzes vom 23. Oktober 2024 (BGBl. I Nr. 323) geändert worden ist.

Gesetz zur Durchführung des Zensus im Jahr 2022 (Zensusgesetz 2022) vom 26. November 2019 (BGBl. I Seite 1851), das durch Artikel 2 des Gesetzes vom 3. Dezember 2020 (BGBl. I Seite 2675) geändert worden ist.

**Herausgeber**  
Statistisches Bundesamt (Destatis), Wiesbaden

---

**Schriftleitung**  
Dr. Daniel Vorgrimler  
Redaktion: Ellen Römer

---

**Ihr Kontakt zu uns**  
[www.destatis.de/kontakt](http://www.destatis.de/kontakt)

---

**Erscheinungsfolge**  
zweimonatlich, erschienen im Februar 2025  
Ältere Ausgaben finden Sie unter [www.destatis.de](http://www.destatis.de) sowie in der [Statistischen Bibliothek](#).

---

Artikelnummer: 1010200-25001-4, ISSN 1619-2907

---

© Statistisches Bundesamt (Destatis), 2025  
Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.