



Statistische Geheimhaltung in Publikationen und bei Weitergabe von Daten

Rechtliche Rahmenbedingungen und methodische Richtlinien von Statistik Austria

Die Bundesanstalt Statistik Österreich (im folgenden kurz „Statistik Austria“ genannt) ist gesetzlich verpflichtet, vertrauliche Informationen über natürliche Personen sowie juristische Personen und Personengemeinschaften zu schützen. Um diese Vorgaben erfüllen zu können, müssen verschiedene Geheimhaltungsmaßnahmen angewendet werden, die eine unrechtmäßige Identifikation von einzelnen natürlichen Personen, juristischen Personen sowie Personengemeinschaften durch Informationen aus veröffentlichten Daten verhindern.

Grundsätzliches Ziel aller Verfahren ist es, einerseits die Geheimhaltungsinteressen der Respondent:innen zu gewährleisten und andererseits möglichst wenig an Information zu unterdrücken, um den Wert sowie die Verwertbarkeit der Daten für den:die Nutzer:in zu erhalten. Obwohl durch die Unterdrückung vorhandener Information zum Zwecke der Geheimhaltung der Informationsgehalt der Daten negativ beeinflusst wird, liefert die statistische Geheimhaltung schlussendlich jedoch einen wesentlichen Beitrag, das Vertrauen der Öffentlichkeit und insbesondere der Respondent:innen in Statistik Austria zu erhalten.

Im Folgenden werden zunächst die rechtlichen Rahmenbedingungen aufgezeigt. Daran schließt sich die Darstellung der methodischen Richtlinien und der hausinternen Verhaltensgrundsätze. Abschließend wird die zum Einsatz gelangende Software erläutert.

1) Rechtliche Rahmenbedingungen

1.1) Begriffsdefinitionen

Für die folgenden Ausführungen werden die Begriffsdefinitionen des Bundesstatistikgesetzes 2000, BGBl. I Nr. 163/1999 idgF, (im Folgenden kurz „BStatG“) und der Verordnung (EG) Nr. 223/2009 über Europäische Statistiken, ABl. Nr. L 87 vom 31.03.2009 S. 164, idgF, (im Folgenden kurz „Europäische StatistikVO“) sowie der Datenschutz-Grundverordnung, ABl. Nr. L 119 vom 27.04.2016 S. 1, idgF, (im Folgenden kurz „DSGVO“) verwendet.

„**Betroffene**“ sind „natürliche und juristische Personen, deren Merkmale erhoben werden“ (§ 3 Z 14 BStatG).

„**Vertrauliche Daten**“ sind „Daten, die eine direkte oder indirekte Identifizierung statistischer Einheiten möglich machen und dadurch Einzelinformationen offenlegen. Bei der Entscheidung, ob eine statistische Einheit identifizierbar ist, sind alle Mittel zu berücksichtigen, die nach vernünftigem Ermessen von einem Dritten angewendet werden könnten, um die statistische Einheit zu identifizieren“ (§ 3 Z 15 BStatG iVm Art. 3 Z 7 Europäische StatistikVO).

„**Statistische Einheit**“ ist eine „Grundbeobachtungseinheit, das heißt eine natürliche Person, ein Haushalt, ein Wirtschaftsteilnehmer oder eine sonstige Unternehmung, auf die sich die Daten beziehen“ (§ 3 Z 3 iVm Art. 3 Z 6 Europäische StatistikVO).

„**Erstellung von Statistiken**“ ist die „Gesamtheit der zur Erhebung, Speicherung, Verarbeitung, Aufbereitung, Analyse, Verbreitung und Bereithaltung der statistischen Informationen erforderlichen Tätigkeiten“ (§ 3 Z 3 BStatG).

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

„**Einwilligung der betroffenen Person**“ ist „jede freiwillig für den bestimmten Fall, in informierter Weise und unmissverständlich abgegebene Willensbekundung in Form einer Erklärung oder einer sonstigen eindeutigen bestätigenden

Handlung, mit der die betroffene Person zu verstehen gibt, dass sie mit der Verarbeitung der sie betreffenden personenbezogenen Daten einverstanden ist“ (Art. 4 Z 11 DSGVO).

1.2) Rechtsgrundlagen

Die rechtlichen Rahmenbedingungen für die gegenständliche Themenstellung definieren sich durch die DSGVO und das DSG und die Bestimmungen des BStatG und der Europäischen StatistikVO.

Datenschutz-Grundverordnung (DSGVO) und Datenschutzgesetz (DSG)

Das Grundrecht auf Datenschutz natürlicher Personen ist in Art. 8 der Europäischen Menschenrechtskonvention, Art. 7 und 8 der Charta der Grundrechte der Europäischen Union, Art. 16 des Vertrages über die Arbeitsweise der Europäischen Union, Art. 1 DSGVO und § 1 DSG verankert. Die Bestimmungen der DSGVO und des DSG führen das Grundrecht auf Datenschutz personenbezogene Daten von natürlichen Personen näher aus. Der gegenständliche Sachverhalt unterliegt hinsichtlich natürlicher Personen als Betroffene dem Anwendungsbereich der DSGVO und des DSG.

Das Grundrecht auf Datenschutz bezieht sich gemäß Art. 1 iVm Art. 4 Z 1 DSGVO nur auf personenbezogene Daten. Personenbezogene Daten sind gemäß § 4 Z 1 DSGVO alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person beziehen. Anonymisierte Daten sind daher vom Datenschutz nicht umfasst.

Bundesstatistikgesetz 2000 (BStatG)

In § 1 BStatG ist die Bundesstatistik als nicht personenbezogenes Informationssystem des Bundes definiert, das Daten über die wirtschaftlichen, demographischen, sozialen, ökologischen und kulturellen Gegebenheiten in Österreich den Bundesorganen zur Planung, Entscheidungsvorbereitung und Kontrolle von Maßnahmen sowie der Wissenschaft, der Wirtschaft und der Öffentlichkeit bereitstellt.

Die Statistiken werden gemäß § 4 Abs. 1 BStatG erstellt aufgrund (i) innerstaatlich unmittelbar wirksamer internationaler Rechtsakte (insbesondere EU-Vorgaben), (ii) Bundesgesetze oder (iii) Verordnungen.

Die Daten werden – je nach Anordnung – durch Beschaffung von Daten aus öffentlichen Registern, aus Verwaltungsregistern, Verwaltungsdaten und Statistikdaten, aufgrund eigener Ermittlungen, durch Beschaffung von Daten aus computergestützten Warenwirtschaftssystemen, aus computergestützten Verkehrsüberwachungssystemen und Transportsystemen, von Satelliten

tendaten, von computergestützten Nutzerdaten von Internet, Telekommunikation und Energie oder durch Befragung der Auskunftspflichtigen erhoben (§ 6 Abs. 1 BStatG).

Statistiken und Supportleistungen dürfen unter Anwendung der DSGVO und des DSG von Statistik Austria gemäß § 23 Abs. 2 BStatG auch auf Grund vertraglicher Vereinbarung für Einrichtungen des Bundes und Dritte erstellt und erbracht werden.

Nach den §§ 19 und 30 BStatG ist Statistik Austria zur Veröffentlichung der erstellten Statistiken verpflichtet.

Europäische und Nationale Sonderbestimmungen („Leges speciales“)

In europäischen und nationalen gesetzlichen Anordnungen können sich auch spezielle Veröffentlichungsbestimmungen finden.

Weitere Bestimmungen

Statistik Austria verfügt über eine hausinterne Datensicherheitsvorschrift (im Folgenden kurz „DSV“ genannt). Sie stellt eine Dienstanweisung dar, deren Verletzung zu dienstrechtlichen Konsequenzen führt (Punkt 1.2 DSV). Sie gilt für die Mitarbeiter:innen von Statistik Austria, die vertrauliche und personenbezogene Daten automationsunterstützt oder in einem manuell geführten Dateisystem verarbeiten. Sämtliche Mitarbeiter:innen werden über das Grundrecht auf Datenschutz und die bestehenden Geheimhaltungspflichten (Datengeheimnis, Statistikgeheimnis, Amtsverschwiegenheit) belehrt und zur Einhaltung verpflichtet (Punkt 7 DSV).

Zudem haben sich die Dienstnehmerinnen und Dienstnehmer von Statistik Austria aufgrund einer schriftlichen Verpflichtungserklärung gegenüber derselben dazu verpflichtet, die datenschutzrechtlichen Bestimmungen einzuhalten.¹

Exkurs: EU-Verordnung über europäische Statistiken (Europäische StatistikVO)

Auf europäischer Ebene ist in Bezug auf die statistische Geheimhaltung die Verordnung (EG) Nr. 223/2009 des Europäischen Parlaments und des Rates vom 11. März 2009 über europäische Statistiken wesentlich. Die Verordnung führt den Grundsatz der Wahrung der statistischen Geheimhaltung in einzelnen Bestimmungen näher aus.

In Kapitel I, Artikel 2 wird folgende Definition für Statistische Geheimhaltung gegeben:

„Statistische Geheimhaltung bedeutet, dass direkt für statistische Zwecke oder indirekt aus administrativen oder sonstigen Quellen eingeholte vertrauliche Angaben über einzelne statistische Einheiten geschützt werden müssen, wobei die Verwendung der eingeholten Angaben für nichtstatistische Zwecke und ihre unrechtmäßige Offenlegung untersagt sind.“

In Artikel 3 Z 7 wird definiert, was unter vertraulichen Daten zu verstehen ist. Als vertraulich gelten jene Daten, die eine direkte oder indirekte Identifizierung statistischer Einheiten möglich machen und dadurch Einzelinformationen offenlegen. Bei der Entscheidung, ob eine statistische Einheit identifizierbar ist, sind alle Mittel zu berücksichtigen, die nach vernünftigem Ermessen

1 <https://www.statistik.at/ueber-uns/aufgaben-und-grundsaeetze/rechtsgrundlagen/daten-schutz>.

von einem Dritten angewendet werden könnten, um die statistische Einheit zu identifizieren;

In Artikel 20 wird der Schutz und die Verwendung vertraulicher Daten geregelt. So müssen sowohl auf Unions- als auch auf nationaler Ebene rechtliche, administrative, technische und organisatorische Maßnahmen zum physischen und logischen Schutz vertraulicher Daten getroffen werden. Weiters wird festgeschrieben, dass die einzelstaatlichen Stellen und EUROSTAT dazu verpflichtet sind, die Grundsätze und Leitlinien in diesem Bereich zu harmonisieren. Der Kommission werden diesbezüglich Durchführungsbefugnisse übertragen.

Vertrauliche Daten, die ausschließlich für die Erstellung europäischer Statistiken erhoben wurden, dürfen von den Nationalen Statistischen Ämtern (NSÄ) und anderen einzelstaatlichen Stellen und vom Statistischen Amt der Europäischen Union (EUROSTAT) ausschließlich für statistische Zwecke verwendet werden. Eine Ausnahme kann nur dann gemacht werden, wenn eine statistische Einheit in unmissverständlicher Weise ihre Zustimmung zur Verwendung der Daten zu anderen Zwecken gegeben hat. Weiters wird in Artikel 20 das Prinzip der „passiven Geheimhaltung“ vorgesehen. Demnach dürfen statistische Ergebnisse in Ausnahmefällen auch dann verbreitet werden, wenn das statistische Datensubjekt indirekt identifiziert werden könnte. Die Voraussetzungen hierfür sind jedoch sehr restriktiv und sehen unter anderem eine einschlägige Bestimmung in einem im ordentlichen Gesetzgebungsverfahren zu erlassenden Rechtsakt vor (wie er gegenwärtig im Bereich der Handelsstatistik gegeben ist) oder die eindeutige Zustimmung des Datensubjekts.

In Artikel 21 ist die Übermittlung vertraulicher Daten geregelt. So ist die Übermittlung vertraulicher Daten innerhalb des ESS (Europäisches Statistisches System) und zwischen dem ESS und dem ESZB (Europäisches System der Zentralbanken) zulässig, sofern die Übermittlung für die effiziente Entwicklung, Erstellung und Verbreitung europäischer Statistiken oder die Verbesserung der Qualität europäischer Statistiken erforderlich ist. Im Fall der Übermittlung an das ESZB muss die Notwendigkeit der Übermittlung zudem hinreichend gerechtfertigt worden sein. Jede weitere Übermittlung über die erste Übermittlung hinaus erfordert jedenfalls die ausdrückliche Einwilligung der Stelle, die die Daten erhoben hat. Bei diesen Bestimmungen handelt es sich um Ermächtigungsklauseln, eine Verpflichtung zur Übermittlung von vertraulichen Daten innerhalb des ESS und an das ESZB kann jedoch nur in einem vom Europäischen Parlament und vom Rat im ordentlichen Gesetzgebungs-

verfahren erlassenen Rechtsakt und nicht durch einen Rechtsakt der Kommission vorgesehen werden. Die übermittelten Daten dürfen ausschließlich für statistische Zwecke verwendet werden und nur Personen, die mit statistischen Angelegenheiten befasst sind, innerhalb ihres spezifischen Arbeitsbereichs zugänglich sein.

Der Zugang zu vertraulichen Daten für wissenschaftliche Zwecke ist in Artikel 23 der Verordnung geregelt. Ein Zugang zu vertraulichen Daten, die an EUROSTAT übermittelt wurden und nur eine indirekte Identifikation der statistischen Einheiten ermöglichen, kann von EUROSTAT für wissenschaftliche Zwecke gewährt werden, wenn das NSA oder eine andere einzelstaatliche Stelle, die die betreffenden Daten übermittelt hat, einer solchen Verwendung zugestimmt hat. Weiters kann auch von einem NSA oder einer anderen einzelstaatlichen Stelle, die für die Erstellung von europäischen Statistiken zuständig ist, ein Zugang zu diesen Daten für wissenschaftliche Zwecke gewährt werden.

In Artikel 25 wird normiert, dass Daten aus Quellen, die der Öffentlichkeit rechtmäßig zugänglich sind und die nach nationalem Recht öffentlich zugänglich bleiben, für den Zweck der Veröffentlichung der aus diesen Daten gewonnenen Statistiken nicht als vertraulich gelten.

Exkurs: Europäische Durchführungsverordnung über europäische Statistiken in Bezug auf den Zugang zu vertraulichen Daten für wissenschaftliche Zwecke

Gemäß der Verordnung (EU) Nr. 557/2013 der Kommission kann EUROSTAT unter bestimmten Bedingungen zur Durchführung statistischer Analysen für wissenschaftliche Zwecke Zugang zu vertraulichen Daten, die sich zur Entwicklung, Erstellung und Verbreitung europäischer Statistiken im Besitz von EUROSTAT befinden, gewähren. Es ist immer sicherzustellen, dass das für vertrauliche statistische Daten erforderliche hohe Schutzniveau gewährleistet ist. Folgender Verfahrensablauf gewährleistet dieses hohe Schutzniveau: Die Forschungseinrichtung, die Zugang beantragen will, ist zuerst von EUROSTAT als ordentliche Forschungseinrichtung anzuerkennen; vorbehaltlich der positiven Erledigung kann die anerkannte Forschungseinrichtung einen Forschungsvorschlag vorlegen, der von EUROSTAT auf seine Ordnungsmäßigkeit geprüft wird; erweist sich der Forschungsvorschlag als ordnungsgemäß, erfolgt der Zugang entweder durch EUROSTAT oder durch eine andere von der EUROSTAT akkreditierte Zugangseinrichtung.

Die betreffende einzelstaatliche statistische Stelle, die die Daten an EUROSTAT übermittelt hat, hat ihre Genehmigung für den Zugang zu erteilen. Der Zugang zu den Daten ist von EUROSTAT zu verwehren, wenn die einzelstaatliche statistische Stelle, die diese Daten geliefert hat, diese Genehmigung nicht erteilt.

EUROSTAT hat auf seiner Homepage folgende Informationen abrufbar zu halten: die „Leitlinien für die Beurteilung von Forschungseinrichtungen, Forschungsvorschlägen und Zugangseinrichtungen“, die Liste der anerkannten Forschungseinrichtungen, die Liste der akkreditierten Zugangseinrichtungen und die Liste der Datensätze für Forschungszwecke mit der entsprechenden Dokumentation und den zur Verfügung stehenden Zugangsarten.

1.3) Die in Bezug auf die statistische Geheimhaltung maßgeblichen Bestimmungen im Einzelnen

1.3.1) Gesetzliche Vorgaben

Informationen, die Statistik Austria aufgrund einzelner Meldungen zur Kenntnis gelangen, sind streng vertraulich zu behandeln und dürfen ausschließlich für Zwecke der „amtlichen Statistik“ und der Einräumung des Zugangs der Wissenschaft zu Statistikdaten Verwendung finden. Dies bedeutet, dass eine personen- und unternehmensbezogene Identifizierung von Meldeeinheiten einem:einer Sachbearbeiter:in nur mehr zu folgenden Zwecken ermöglicht wird (§ 5 Abs. 2; § 25a Abs. 3; §§ 31, 31a und 31b BStatG):

- Überprüfung der Erfüllung der Auskunftspflicht,
- Plausibilisierung, Berichtigung oder Vervollständigung von Auskünften,
- Zusammenführung von Daten über dieselbe statistische Einheit bei einer Erhebung, die aus unterschiedlichen Datenquellen erfolgt,
- Erstellung, Ergänzung und Berichtigung der statistischen Register über juristische Personen, Einrichtungen, Unternehmen und ihrer Betriebe und Arbeitsstätten sowie über Arbeitsgemeinschaften und Forschungsstätten,
- Erstellung von durch Bundesgesetz bzw. innerstaatlich unmittelbar wirksamen internationalen Rechtsakt angeordneten Verlaufsstatistiken,
- Sicherstellung der Prüftätigkeit internationaler Organisationen auf Grund eines völkerrechtlich verbindlichen Vertrages,
- Entlastung der Respondent:innen bei wiederholten zeitnahen Erhebungen in der Art der Befragung über die gleichen Erhebungsmerkmale,

- Erstellung der Volkswirtschaftlichen Gesamtrechnungen
- Einräumung des Zugangs der Wissenschaft zu Statistikdaten

Zudem enthält das BStatG weitere Maßnahmen zum Schutz von Geheimhaltungsinteressen von Personen bzw. Unternehmen sowie zur Sicherung der Vertraulichkeit von Einzeldaten. Dazu gehört etwa, dass die Identitätsdaten von natürlichen Personen zum frühest möglichen Zeitpunkt beseitigt und durch ein Pseudonym ersetzt werden müssen (§ 15 Abs. 1 BStatG) oder dass die Identität von Personen und Unternehmen verschlüsselt werden muss, falls die Beibehaltung des Personen- und Unternehmensbezuges notwendig ist (§ 15 Abs. 2 BStatG). Weiters wird im BStatG geregelt, dass alle Personen, die mit Aufgaben der Bundesstatistik betraut sind, über alle personen- und unternehmensbezogenen Daten zur Verschwiegenheit verpflichtet sind (§ 17 Abs. 3 BStatG) und dass Verstöße gegen die Verschwiegenheitspflicht als Verletzung des Amtsgeheimnisses nach § 310 StGB strafrechtlich verfolgt werden. Nach § 17 Abs. 2 BStatG dürfen personen- und unternehmensbezogene Daten an Dritte nur dann übermittelt werden, wenn ein innerstaatlich unmittelbar wirksamer internationaler Rechtsakt oder bundesgesetzliche Bestimmungen dies vorsehen.

Wissenschaftliche Nutzung von Statistikdaten (§ 31 BStatG):

Hierbei ist zwischen dem Zugang gemäß § 31 Abs. 1 und 2 BStatG und dem Zugang gemäß § 31 Abs. 3 ff BStatG zu unterscheiden.

Entsprechend § 31 Abs. 1 und 2 BStatG hat Statistik Austria fachlich geeigneten Personen und wissenschaftlichen Einrichtungen für wissenschaftliche Zwecke Zugang zu nicht personen- und unternehmensbezogenen statistischen Daten einzuräumen. Dabei müssen Datensicherheitsmaßnahmen angewendet werden, die das Ermitteln mit Mitteln, die vernünftiger Weise angewendet werden können, und die Abspeicherung von personenbezogenen und unternehmensbezogenen statistischen Daten auf externe Datenträger unmöglich machen. Anwendungsfälle hierfür sind Standardisierte Anonymisierte Einzeldatensätze, die als Scientific Use Files außer Haus gegeben werden können, sowie Datensätze im Safe Center oder für das kontrollierte Fernrechnen.

Entsprechend § 31 Abs. 3 ff BStatG hat Statistik Austria ab 1. Juli 2022 im Austrian Micro Data Center Zugang zu statistischen Daten, bei denen keine Identifizierung der Betroffenen durch Namen, Anschrift oder anhand einer öffentlich zugänglichen Identifikationsnummer möglich ist, via Remote Access (gesicherter Online-Zugang durch Bereitstellen einer virtuellen Desktop-Infrastruktur (VDI) mit einer Zwei-Faktor-Authentifizierung und Zugriffs-Protokollierung) zu gewähren. Zugang ist zu gewähren zu Daten des Unternehmensregisters für die Verwaltung

(§ 25 BStatG), des Registers für statistische Einheiten (§ 25a BStatG), der fachstatistischen Register gemäß § 26 BStatG, des Bildungsstandregisters gemäß § 19 des Bildungsdokumentationsgesetzes 2020 sowie des Gebäude- und Wohnungsregisters gemäß § 1 des GWR-Gesetzes. Diese Daten können mit weiteren eigenen Statistikdaten anhand des bPK-AS beziehungsweise der verschlüsselten Unternehmenskennzahl und mit externen Datenbeständen verknüpft werden, die die wissenschaftliche Einrichtung zur Verfügung stellt. Zugang darf nur akkreditierten wissenschaftlichen Einrichtungen zu ausführlich beschriebenen Forschungsvorhaben (bei geplanter Ablehnung Konsultation des Statistikrates) gewährt werden.

Für die wissenschaftliche Nutzung von Statistikdaten gilt der Datenminimierungsgrundsatz und es bedarf einer abschließenden Prüfung der Ergebnisse des Forschungsvorhabens, damit ein Rückschluss auf statistische Einheiten ausgeschlossen werden kann. Die wissenschaftlichen Einrichtungen, die wissenschaftliche Leitung und alle Personen, die am Forschungsvorhaben mitwirken, unterliegen dem Statistikgeheimnis. Die Abspeicherung von vertraulichen Daten auf externe Datenträger, das Fotografieren, das Abschreiben oder die Anfertigung einer Bildschirmkopie vertraulicher Daten oder die Verwendung vertraulicher Daten für andere als wissenschaftliche Zwecke stellt jedenfalls eine Verletzung des Statistikgeheimnisses dar.

Auf der Website sind Informationen u.a. über akkreditierte wissenschaftliche Einrichtungen und Forschungsvorhaben zu veröffentlichen.

Nach § 19 Abs. 2 und 3 BStatG sind die Statistiken grundsätzlich „in solcher Weise zu veröffentlichen, dass ein Rückschluss auf Angaben über bestimmte oder bestimmbare Betroffene ausgeschlossen werden kann, es sei denn, dass der Betroffene an der Geheimhaltung der Angaben kein schutzwürdiges Interesse hat.“ Ansonsten ist die Veröffentlichung nur mit Zustimmung² der:des Betroffenen zulässig. Dabei sind gemäß § 19 Abs. 3 BStatG insbesondere konkrete Hinweise der Betroffenen über die Möglichkeit von Rückschlüssen auf Angaben, an deren Geheimhaltung ein schutzwürdiges Interesse der:des Betroffenen besteht, zu berücksichtigen.

Der in § 19 BStatG verwendete Begriff „schutzwürdiges Geheimhaltungsinteresse“ ist im BStatG nicht definiert. Um die in § 19 Abs. 2 BStatG festgelegten

2 Zustimmung ist in Analogie zum Begriff „Einwilligung der betroffenen Person“ des Art. 4 Z 11 DSGVO die freiwillig für den bestimmten Fall, in informierter Weise und unmissverständlich abgegebene Willensbekundung in Form einer Erklärung oder einer sonstigen eindeutigen bestätigenden Handlung, mit der der Betroffene zu verstehen gibt, dass er mit der Verwendung der ihn betreffenden personen- und unternehmensbezogenen Daten einverstanden ist.

Grenzen genau zu erfassen, ist daher die Definition der Begriffe „vertrauliche Daten“ und „schutzwürdiges Geheimhaltungsinteresse“ näher zu erörtern.

1.3.1.1) Vertrauliche Daten

Ein Personen- oder Unternehmensbezug liegt vor, wenn die Identität der:des Betroffenen bestimmt oder bestimmbar ist. Bestimmt bedeutet, dass die Daten dem Betroffenen eindeutig zugeordnet sind, etwa durch Benennung des Namens oder anhand einer öffentlich zugänglichen Identifikationsnummer in Zusammenhang mit der Anführung der Daten. Bestimmbar ist die Identität, wenn die Identifizierung einer statistischen Einheit durch andere Mittel als die direkte Identifizierung möglich ist. Bei der Entscheidung, ob eine statistische Einheit identifizierbar ist, sind alle Mittel zu berücksichtigen, die nach vernünftigem Ermessen von einem Dritten angewendet werden könnten, um die statistische Einheit zu identifizieren.

Hinsichtlich der Identifizierbarkeit ist jeweils auf den durchschnittlich zu erwartenden Informationsstand des angesprochenen Empfängerkreises abzustellen und Zusatzkenntnisse, welche bei den Nutzer:innen der Statistik üblicherweise vorhanden sind, zu berücksichtigen. Die Veröffentlichung solcher statistischer Daten, die mit bei den Benutzer:innen der Statistik üblicherweise vorhandenen Zusatzkenntnissen auf den:die einzelne:n Betroffene:n rückführbar sind, ist untersagt, wenn diese Veröffentlichung nicht in einer Rechtsvorschrift vorgeschrieben ist.

Daten aus Quellen, die der Öffentlichkeit rechtmäßig zugänglich sind und die nach nationalem Recht öffentlich zugänglich bleiben, gelten für den Zweck der Verbreitung der aus diesen Daten gewonnenen Statistiken nicht als vertraulich (Art. 25 Europäische StatistikVO).

1.3.1.2) Schutzwürdiges Geheimhaltungsinteresse

Das Vorliegen des schutzwürdigen Geheimhaltungsinteresses hängt vom konkreten Fall ab und ist für jeden Einzelfall zu prüfen. Das Vorliegen eines schutzwürdigen Geheimhaltungsinteresses ist jedenfalls dann ausgeschlossen, wenn die Daten

- allgemein verfügbar oder
- auf die:den Betroffene:n nicht rückführbar sind.

1.3.1.2.1) Allgemeine Verfügbarkeit von Daten

Allgemein verfügbar sind Daten, die für jedermann zugänglich sind. Hierzu gehören Daten, die aufgrund der Rechtsordnung öffentlich zugänglich sind (z. B. Da-

ten im Grundbuch, Firmenbuch, Unternehmensdaten, die einer Veröffentlichungspflicht unterliegen, etc) sowie Daten, die auf sonstige (zulässige) Weise z. B. durch Herausgabe eines Telefonbuchs, eines Einwohnerverzeichnisses oder durch Medienberichte öffentlich wurden. Die allgemeine Verfügbarkeit bedarf einer genauen Prüfung insbesondere dahingehend, ob die allgemeine Zugänglichkeit im Zeitpunkt der beabsichtigten Verwendung tatsächlich noch besteht.

Festzuhalten ist, dass die Veröffentlichung der Daten eine zulässige sein muss. Öffentlich bedeutet, dass jedermann, ohne Ansehen der Person und des Zweckes der Einsicht, die Einsicht freisteht. Die Ärztekammer darf z. B. Daten aus dem öffentlichen Teil der Ärzteliste an jedermann übermitteln.

Der allgemeinen Verfügbarkeit entzogen sind somit nur Daten, die geheim gehalten werden können. Daten, welche für jedermann offensichtlich sind, sind daher ebenfalls allgemein verfügbar; an solchen Daten besteht grundsätzlich kein schutzwürdiges Geheimhaltungsinteresse. Wird hingegen eine offensichtliche Information mit einer zusätzlichen personen- bzw. unternehmensbezogenen, nicht allgemein verfügbaren Information verknüpft, so stellt dies einen Eingriff in schutzwürdige Geheimhaltungsinteressen dar.

1.3.1.2.2) Vorgangsweise bei Rückführbarkeit

Ist ein Rückschluss auf Betroffene nicht vermeidbar, darf eine Veröffentlichung gemäß § 19 Abs. 2 BStatG nur nach vorheriger ausdrücklicher schriftlicher Zustimmung der:des Betroffenen vorgenommen werden.

Erfolgt die Veröffentlichung der Statistik in einem gesetzlich vorgegebenen Detaillierungsgrad aufgrund einer ausdrücklichen und verfassungskonformen gesetzlichen Ermächtigung bzw. Verpflichtung, so ist eine Beeinträchtigung des schutzwürdigen Geheimhaltungsinteresses ausgeschlossen.

1.3.1.2.3) Weitere nationale Rechtsgrundlagen – gesetzliche Ermächtigungen

Es gibt Statistiken, die aufgrund von speziellen europäischen bzw. bundesgesetzlichen Bestimmungen Sonderregelungen der statistischen Geheimhaltung unterliegen. Im Folgenden wird auf diese abweichenden Regelungen eingegangen.

- So gilt etwa im Bereich der Außenhandelsstatistik gemäß Artikel 19 der Verordnung (EU) 2019/2152 das Prinzip der „passiven Geheimhaltung“. Dieses Prinzip besagt, dass auch aggregierte Daten, aus denen Respondent:innen indirekt identifizierbar sind, veröffentlicht werden dürfen, wenn die betroffene Meldeeinheit keinen Antrag auf Geheimhaltung stellt. Solch ein Antrag ist seitens der Statistikbehörde auf seine Berechtigung hin zu

prüfen und gegebenenfalls demselben stattzugeben oder ihn abzulehnen. Weiters ist im Bereich der Außenhandelsstatistik auch eine Weitergabe von Individualdaten beziehungsweise ein Informationsaustausch mit Behörden möglich, wenn dies zur Einleitung oder Durchführung eines Strafverfahrens nötig ist (§ 3 Handelsstatistisches Gesetz 1995, BGBl. I Nr. 173/95 idgF) oder wenn zwischen den handelsstatistischen Anmeldungen und den von den Steuerbehörden ermittelten Finanzdaten erhebliche Differenzen festgestellt werden (§ 12 Abs. 5 Handelsstatistisches Gesetz 1995, BGBl. I Nr. 173/95 idgF).

- Die Ergebnisse der Bundesstatistik zum Bildungswesen in regionaler Gliederung dürfen entsprechend § 18 Abs. 1 Bildungsdokumentationsgesetz 2020, BGBl. I Nr. 20/2021 idgF, für Zwecke der Qualitätssicherung, Bildungsplanung und Raumordnung unter Angabe von Name oder Bezeichnung und Adresse der Bildungseinrichtung veröffentlicht werden.
- Ferner gilt auch für die Arbeitsstättenzählungen eine abweichende Regelung. Gemäß § 8 Registerzählungsgesetz, BGBl. I Nr. 33/2006 idgF, kann die Veröffentlichung von anonymisierten Einzelergebnissen in Tabellenform uneingeschränkt erfolgen. Für Zwecke der Raumordnung und der Wirtschaftspolitik dürfen Einzeldaten betreffend Art der ausgeübten Tätigkeiten, Rechtsform, gesetzliche berufliche Interessenvertretung oder Rechtsträger, Größengruppe der unselbständig erwerbstätigen Personen und Unternehmenszugehörigkeit mit Namen und Anschrift der Arbeitsstätte veröffentlicht werden.
- Darüber hinaus werden – geregelt durch § 71 Abs. 1 des Wirtschaftskammergesetzes 1998, BGBl. I Nr. 103/1998 idgF – Angaben, die im Zuge statistischer Erhebungen nach dem erhoben werden, an die Wirtschaftskammer Österreich für statistische Zwecke übermittelt. Die Übermittlung ist nur gestattet für Angaben von Kammermitgliedern sowie für Kammermitglieder betreffende Daten, die nach den Bestimmungen des BStatG der Statistik Austria zugänglich werden, wenn dies der Wahrnehmung der den Wirtschaftskammern gesetzlich übertragenen Aufgaben dient. Nach § 71 Abs. 3 leg. cit. gelten auch für die Wirtschaftskammern bei Veröffentlichung dieser Daten die Bestimmungen des BStatG.
- Auch an die Oesterreichische Nationalbank werden vertrauliche Daten und gegebenenfalls Einzeldaten übermittelt, die für die Erstellung der Zahlungsbilanzstatistiken erforderlich sind. Die rechtliche Grundlage hierfür bietet § 6 Devisengesetz 2004, BGBl. I Nr. 123/2003 idgF.
- Gemäß § 25 Abs. 6 Arbeitsmarktservicegesetz, BGBl. Nr. 313/1994 idgF, darf Statistik Austria von ihr verarbeitete Stammdaten der Arbeitgeber und

Daten über Ausbildungen an das Arbeitsmarktservice und an das Bundesministerium für Arbeit, Soziales, Gesundheit und Konsumentenschutz übermitteln, soweit diese Daten für Zwecke der in deren gesetzlichen Aufgabenbereich fallender wissenschaftlicher oder arbeitsmarktstatistischer Untersuchungen, die keine personenbezogenen Ergebnisse zum Ziel haben, eine wesentliche Voraussetzung bilden.

- Das LFBIS-Gesetz, BGBl. Nr. 448/1980 idgF, regelt die Übermittlungen von land- und forstwirtschaftlichen Daten, die im Rahmen des land- und forstwirtschaftlichen Betriebsinformationssystems zwischen Statistik Austria und dem für Landwirtschaft zuständigen Bundesministerium stattfinden dürfen.
- Weiters ist gemäß § 7 Umweltkontrollgesetz, BGBl. I Nr. 152/1998 idgF, zur Ausführung der dem Umweltbundesamt und Statistik Austria gesetzlich übertragenen Aufgaben ein Austausch von Einzeldaten zwischen dem Umweltbundesamt und Statistik Austria zulässig.
- Das Bundesgesetz über die Gesundheit Österreich GmbH (GÖGG), BGBl. I Nr. 132/2006 idgF, enthält in § 15a Abs. 6 die folgende Sonderbestimmung: "Die Gesundheit Österreich GmbH ist berechtigt, bei der Bundesanstalt Statistik Österreich Informationen zum Todeszeitpunkt und zur Todesursache von Personen anzufordern, deren Daten in einem Register verarbeitet sind". Dies ist eine Ausnahme von der Wahrung des Statistikgeheimnisses im Bereich der Todesursachenstatistik.
- Gem. § 3 Abs. 1 der Art. 15a B-VG Vereinbarung zwischen dem Bund und den Ländern über die Zusammenarbeit im Bereich der Statistik, BGBl. 408/1985, können auf Ersuchen anonymisierte Einzeldaten für statistische Zwecke übermittelt werden und nach § 3 Abs. 4 Namen und Adressen, die insoweit mit weiteren Merkmalen verknüpft sind, als dies zur Durchführung solcher statistischer Erhebungen und Arbeiten unerlässlich ist, die für den Empfänger zur Wahrnehmung der ihm gesetzlich übertragenen Aufgaben eine wesentliche Voraussetzung bilden.
- Auch § 4 Abs. 2 Straßen- und Schienenverkehrsstatistikgesetz, BGBl. Nr. 142/1983, sieht eine Sonderregelung vor. Wenn für konkrete Untersuchungen oder Maßnahmen, insbesondere verkehrspolitischer Art, regionale Verflechtungsdaten der Güterströme im Inland bis zur Ebene der Bezirksverwaltungsbehörden benötigt werden, so kann jeder Bundesminister und jeder Landeshauptmann die Durchführung der hierfür erforderlichen Sonderauswertungen veranlassen. Der eine solche Sonderauswertung veranlassende Bundesminister oder Landeshauptmann hat jedoch dafür zu sorgen, dass durch die Verwendung der Sonderauswertung keine Rückschlüsse auf konkrete Geschäftsbeziehungen ermöglicht werden.

Rechtliche Schlussfolgerungen

Bei der Veröffentlichung von Statistiken sind die im BStatG enthaltenen Vorgaben zu berücksichtigen, wobei sich diese Einschränkungen ausschließlich auf personen- und unternehmensbezogene Daten – also Daten über Betroffene, deren Identität bestimmt oder bestimmbar ist – beziehen. Da veröffentlichte Statistiken meist keine personen- bzw. unternehmensbezogenen Daten enthalten, entfällt schon deshalb häufig eine datenschutzrechtliche Relevanz.

Liegen solche Daten jedoch vor, so ist darauf zu achten, dass die schutzwürdigen Geheimhaltungsinteressen der Betroffenen nicht beeinträchtigt werden. Eine Beeinträchtigung ist jedenfalls auszuschließen, wenn die Daten allgemein verfügbar sind oder die Datenveröffentlichung auf einer ausdrücklichen gesetzlichen Bestimmung beruht.

Es ist daher im Einzelfall zunächst zu prüfen, ob ein Personen- bzw. Unternehmensbezug hergestellt werden kann, ob also die Daten dem angesprochenen Empfängerkreis einen Rückschluss auf die Betroffenen ermöglichen.

Liegt Personen- bzw. Unternehmensbezug vor, so ist als nächstes zu klären, ob die personen- bzw. unternehmensbezogenen Daten allgemein zugänglich sind. Ist dies nicht der Fall, so liegen schutzwürdige Geheimhaltungsinteressen jedenfalls vor.

Sind personen- bzw. unternehmensbezogene Daten allgemein zugänglich, ist allenfalls noch zu prüfen, ob eine Verknüpfung mit Daten, welche nicht allgemein zugänglich sind, vorliegt. Ist dies auch nicht der Fall, so liegen keine schutzwürdigen Geheimhaltungsinteressen vor.

2) Methodische Richtlinien

Im folgenden Abschnitt werden die in Statistik Austria gültigen methodischen Richtlinien für die Geheimhaltung in Publikationen und bei extern zugänglichen Daten definiert. Dabei ist zu beachten, dass diese Richtlinien nicht stereotyp für das Anonymisieren von Daten verwendet werden können. Vielmehr hängt die Wahl geeigneter statistischer Methoden, Gliederungsstufen oder Kodierungen wesentlich vom konkreten Datenmaterial, dessen Einheiten es zu schützen gilt, ab. Oft können sich die Datenschutzmaßnahmen auf unterschiedliche Darstellungsmerkmale beziehen, sodass beispielsweise die Optionen bestehen, entweder die regionale oder die branchenmäßige Gliederungstiefe stärker auszuschöpfen. Auch hier kann keine allgemein gültige Strategie vorgegeben werden, sondern die zu wählende Vorgangsweise hängt vom konkreten Datenmaterial ab.

Die anzuwendenden Datenschutzmaßnahmen hängen jedenfalls vom Inhalt der Publikation ab; es ist zu unterscheiden, ob statistische **Aggregate, etwa in Form von Tabellen (Häufigkeitstabellen, Wertetabellen)** veröffentlicht oder **Einzel-daten (Mikrodaten)** zugänglich gemacht werden.

Häufigkeitstabellen geben für jede Zelle einer mehrdimensionalen Tabelle die Anzahl an Einheiten an, die in diese Merkmalskombination fallen. Wertetabellen enthalten numerische (meist nicht-negative) Größen wie Mittelwerte oder Merkmalssummen. Mikrodaten können nur als entsprechend anonymisierte Datensätze für alle oder einen Teil der Einheiten der Grundgesamtheit weitergegeben werden.

Besondere Beachtung sollte die Tatsache finden, dass eine Veröffentlichung als Karte oder als Grafik einer Veröffentlichung als Häufigkeits- bzw. Wertetabelle gleichzusetzen ist. Daher müssen die dargestellten Daten, z.B. Häufigkeitstabellen auf Rasterzellen, den Geheimhaltungsrichtlinien entsprechen. Es stellt keinen zusätzlichen Schutzmechanismus dar, dass die Hintergrunddaten von Karten bzw. Grafiken nicht direkt zum Download angeboten werden.

Es ist daher notwendig, unterschiedliche Geheimhaltungsverfahren auf die verschiedenen Datenformen anzuwenden. Einen Überblick über statistische Methoden zur Geheimhaltung von Mikrodaten, Häufigkeits- sowie Wertetabellen bietet das [Handbuch](#) des [Center of Excellence on Statistical Disclosure Control](#).

2.1) Regeln für die Abgabe aggregierter Daten

2.1.1) Wertetabellen

- Die Zeilen und Spalten einer Wertetabelle bestehen aus Gliederungsmerkmalen wie zB Region, ÖNACE, Staatsangehörigkeit oder Größenklassen. Die Ausprägungen dieser Gliederungsmerkmale können zur Identifikation von Einheiten beitragen. Im Bereich der Wirtschaftsstatistik dürfen nur Zellen publiziert werden, zu denen wenigstens x Einheiten beitragen. Alle Zellen, zu deren Wert weniger als x Meldeeinheiten beitragen, werden unterdrückt (**x-Regel**). Diese Vorgehensweise wird als Primärunterdrückung bezeichnet. In der Statistik Austria wurde bis zum Jahr 2007 $x=4$ gewählt, ab dem Jahr 2008 reichen $x=3$ beitragende Einheiten zur Publikation des Werts einer Tabellenzelle aus. Gilt in einer Zelle $x < 3$, so spricht man von einer „**schutzwürdigen Zelle**“.
- Um einen Antrag von betroffenen Unternehmen auf Schutzwürdigkeit eines Tabellenfeldes zu überprüfen, kann in der Wirtschaftsstatistik die (**n, k**)-**Dominanz- oder Konzentrationsregel** verwendet werden. Diese besagt, dass ein Zellenwert dann zu unterdrücken ist, wenn auf die n größten Einheiten dieser Zelle mindestens $k\%$ der Zellensumme entfallen. Die Frage, ob die Werte n und k publiziert werden sollen, wird in verschiedenen Ländern unterschiedlich gehandhabt. Die Dominanzregel ist im anglo-amerikanischen Raum weit verbreitet, in Österreich wird sie im Bereich der Wirtschaftsstatistik nur implizit eingesetzt und zwar dann, wenn ein Unternehmen ein schutzwürdiges Interesse zur Unterdrückung eines Zellwertes vorbringt und eine Dominanz behauptet wird. Dem schutzwürdigen Interesse wird stattgegeben, wenn nach entsprechender Prüfung eine Dominanz vorliegt.
- In schutzwürdigen Zellen werden Werte nicht publiziert, sondern die Zelle enthält anstelle einer Zahl ein „G“ (für Geheim). Um aus publizierten Randverteilungen nicht die unterdrückten Werte von schutzwürdigen Zellen rückrechnen zu können, müssen zusätzliche Zellen unterdrückt werden.

Mit anderen Worten: Zusätzlich zur Primärunterdrückung werden auch Aggregate mit mehr als x Meldeeinheiten unterdrückt, um sicherzustellen, dass durch Differenzbildung oder andere mathematische Operationen nicht auf Werte geschlossen werden kann, die bereits durch primäre Geheimhaltung unterdrückt wurden. Diese Vorgehensweise wird als „**sekundäre Geheimhaltung**“ bezeichnet.

2.1.2) Häufigkeitstabellen

- Die Methoden zur Geheimhaltung schutzwürdiger Zellen in Wertetabellen können im Prinzip auch für Häufigkeitstabellen angewendet werden. Allerdings hat sich in der Statistik Austria die Konvention durchgesetzt, in diesen Tabellen keine Wertunterdrückung anzuwenden. Die Existenz detaillierter Ausprägungen (zB regionaler Gliederungen) erleichtert Nutzer:innen von Daten die Identifikation statistischer Einheiten. Daher muss bei der Auswahl der zu verkreuzenden Merkmale und ihrer Ausprägungen entsprechend vorsichtig vorgegangen werden.
- Seit der Registerzählung 2011 werden aus Datenschutzgründen bereits die Mikrodaten durch die Methode „Target Record Swapping“ geschützt. Daher konnte ab diesem Berichtsjahr die Datenschuttschwelle für die Datenabgabe gesenkt werden. Kleinräumige Statistikprodukte werden zusätzlich durch Unterdrückung unterhalb einer Datenschuttschwelle geschützt (z.B. Personendaten ab 4 Personen mit Hauptwohnsitz).
- Andere Methoden zur Vermeidung von Tabellenfeldern mit zu geringer Besetzung sind die Änderung der Merkmalsdimensionen oder die Vergrößerung von Ausprägungen. Eine Alternativvariante besteht in der Erzeugung von sicheren (anonymisierten) Mikrodaten und der anschließenden Tabellierung ausgehend von den sicheren Mikrodaten.

Um für Häufigkeitstabellen zu verhindern, dass auf Einzelpersonen bzw. Unternehmen rückgeschlossen werden kann, werden – abgestimmt auf die betreffende Publikation – unterschiedliche Verfahren angewandt. Diese Verfahren reichen von Nicht-Veröffentlichung von Daten für tiefgehende Gliederungsstufen (z. B. Spitalsentlassungsstatistik) bis hin zur Aggregation beziehungsweise Umkodierung von Klassifizierungsmerkmalen (z. B. Top- bzw. Bottom-Kodierung der Variable Alter bei EU-SILC).

Bei der Publikation von Häufigkeitstabellen ist es wichtig darauf zu achten, dass die verkreuzten Merkmale ihrer Anzahl und Gliederungstiefe nach so gewählt werden, dass die datenschutzrechtlichen Belange der beitragenden Einheiten sichergestellt sind. Man könnte argumentieren, dass geringe Zellbesetzungen und damit eine mögliche Identifizierung einer Einheit in einer Häufigkeitstabelle keine zusätzliche Information über die betreffende Einheit freisetzt. Dennoch zeigt ein sehr einfaches Beispiel die dahinterstehende Problematik. So könnte die numerische Variable Umsatz in mehrere disjunkte Umsatzklassen umkodiert werden und in einer Häufigkeitstabelle die Anzahl von Unternehmen nach Bundesland und den definierten Umsatzklassen ausgegeben werden. Es ist offensichtlich, dass bei eindeutiger Identifikation eines Unternehmens in einer Zelle schützenswerte Information einer beitragenden Einheit (über den Umsatz) aufgedeckt wird.

2.2) Regeln für die Abgabe von Einzeldaten

- Für die sichere Abgabe von Einzeldaten muss zunächst überlegt werden, welche Merkmale (Schlüsselmerkmale) ein eventuell an Datenmissbrauch Interessierter zur Verfügung haben könnte, die ihm die Identifizierung einzelner Personen/Einheiten ermöglicht. Merkmale in diesem Sinne können beispielsweise aus öffentlich abrufbaren Registern stammen oder branchenspezifisches Insiderwissen sein. Wenn die Kombination dieser Merkmalsausprägungen zu Zellen mit geringer Besetzung führt, dann ist das Risiko einer Identifizierung von Einheiten groß.
- Um das Risiko der Identifizierung von Einzeldaten so gering wie möglich zu halten, gibt es zwei Arten von Maßnahmen: Datenreduktion (Subsampling, Vergrößerung von Merkmalsausprägungen, Löschung von Variablen oder Datensätzen) und Datenverschmutzung (Verschmutzung der Daten, data swapping, Mikroaggregation, Löschung und Imputation). Bei Anwendung aller dieser Maßnahmen muss jedoch immer darauf geachtet werden, dass durch die Datensicherungsmaßnahmen die Datenqualität nicht zu stark leidet.

Anonymisierte Mikrodatensätze können durch das (kombinierte) Anwenden unterschiedlicher Verfahren erstellt werden, auf die nun kurz eingegangen wird:

- **Umkodierung:**

Ein probates Mittel stellt die Umkodierung von Ausprägungen in den Schlüsselvariablen dar. Dazu zählt unter anderem Top- bzw. Bottom-Coding beziehungsweise Aggregation. Als Top- bzw. Bottom-Coding wird das Einziehen einer oberen/unteren Grenze bei einer quantitativen Variablen verstanden. Originalwerte, die höher/niedriger als die jeweilige Grenze sind, werden durch die obere/untere Grenze ersetzt.

Unter Aggregation versteht man das Zusammenfassen mehrerer Ausprägungskombinationen einer (Schlüssel)variablen zu einer neuen Ausprägungskombination.

- **bewusstes Ersetzen:**

Bei dieser Methode werden kritische Werte (die etwa aus der Erfahrung bekannt sind) durch fehlende Werte ersetzt.

- **Mikroaggregation:**

Mikroaggregation bezeichnet ein Verfahren für numerische Variablen. Statistische Einheiten werden in Gruppen zu jeweils mindestens k Einheiten unterteilt. Schließlich wird ein Aggregat (typischerweise der Gruppen-Mittelwert)

anstelle der Echtwerte für ausgewählte oder alle Variablen einer Gruppe veröffentlicht. Dadurch wird sichergestellt, dass jede Ausprägung dieser Variablen zumindest k-mal im veröffentlichten Datensatz aufscheint.

- **Verschmutzung:**

Auch das „Verschmutzen“ von numerischen Variablen durch das Überlagern mit (korrelierten oder unkorrelierten) Zufallszahlen aus einer gewählten Verteilung, die typischerweise Mittelwert 0 besitzt („adding noise“) stellt ein probates Mittel dar, das Risiko zur korrekten Identifikation einer statistischen Einheit zu vermindern, da exakte Verknüpfungen mit externen Datenquellen nicht mehr möglich sind.

- **Swapping:**

Ein weiteres Verfahren zur Anonymisierung sowohl kategorieller als auch numerischer Variablen ist das sogenannte „Swapping“. Beim „Rank Swapping“ werden die Werte nach einer Variablen gereiht („ranking“). Anschließend werden – für ausgewählte Variablen - Werte einer statistischen Einheit mit zufällig ausgewählten Werten anderer Einheiten innerhalb einer gewissen Grenze (der Rang zweier vertauschter Werte darf sich etwa nicht um mehr als p Prozent gemessen an der Anzahl der Beobachtungen unterscheiden) vertauscht („swapping“).

Beim „Target Record Swapping“ werden zuerst sogenannte „risky Records“ gesucht - das sind Personen bzw. Haushalte, die aufgrund ihrer Merkmalskombinationen leicht zu identifizieren wären. Einzelne Merkmale dieser Personen/Haushalte werden dann mit den Merkmalen anderer, ähnlicher Personen/Haushalte getauscht. Dabei wird darauf geachtet, dass die wichtigsten Eckzahlen nicht verzerrt werden. Der Prozentsatz an vertauschten Datensätzen darf zur Wahrung der Wirksamkeit des Verfahrens nicht publiziert werden.

- **Subsampling:**

Zudem ist es möglich, das Risiko für eine Identifizierung zu senken, wenn nur ein zufällig gezogenes Subsample der Daten veröffentlicht wird. In Verbindung mit weiteren Anonymisierungsmaßnahmen wird es Angreifern zusätzlich erschwert, Verknüpfungen zwischen externen Datenquellen und den veröffentlichten Mikrodaten vorzunehmen, da selbst eine eindeutige Verknüpfung nicht gleichbedeutend mit einer positiven Identifizierung einer Meldeeinheit ist. So wurden etwa 10% Haushaltsstichproben aus den Volkszählungen 1971 bis einschließlich 2001 für Forschungszwecke an die University of Minnesota geliefert.

Im Umgang mit Mikrodaten werden die gesetzlichen Bestimmungen des BStatG umgesetzt, indem die Identitätsdaten von natürlichen Personen zum frühest möglichen Zeitpunkt beseitigt und durch ein nicht rückführbares Pseudonym ersetzt werden und, falls die Beibehaltung des Personen- bzw. Unternehmensbezuges notwendig ist, die Identität von Personen und Unternehmen verschlüsselt wird. Die Schlüssel werden getrennt von den Daten selbst aufbewahrt.

Werden Mikrodaten veröffentlicht, so wird eine Gruppe von meist kategoriellen Variablen als sogenannte „**Schlüsselvariablen**“ definiert. Es wird dabei angenommen, dass ein:e Angreifer:in Informationen über eine Person oder ein Unternehmen für diese Variablen besitzen kann. Meldeeinheiten mit Ausprägungskombinationen in den Schlüsselvariablen, die in einem Datensatz sehr selten auftreten, weisen ein hohes Risiko für eine korrekte Identifizierung auf.

Eine Besonderheit bei der Veröffentlichung von Mikrodaten stellen die „Mikrodatsätze für Forschung und Lehre“ dar. Obwohl alle Standardisierten bzw. Aufgabenspezifischen Datensätze anonymisiert veröffentlicht werden, soll der Zugang zu den Daten für Forschende und Benutzer:innen erst nach einem Registrierungsverfahren sowie, als zusätzlichen Datensicherheitsaspekt, nach Anerkennung der Nutzungsbestimmungen gewährt werden.

2.3) Weitergabe von Daten an EUROSTAT

Es gelten die jeweiligen EU-Rechtsgrundlagen.

3) Allgemeine hausinterne „Verhaltens“-Grundsätze

Vor dem Hintergrund dieser rechtlichen und methodischen Rahmenbedingungen bzw. Richtlinien, sind zudem folgende hausinterne „Verhaltens“-Grundsätze zu beachten:

- Die Entscheidung, ob Einzeldaten bzw. aggregierte Daten ausreichend geschützt sind und daher zur Veröffentlichung freigegeben werden können, liegt im Aufgabenbereich des:der für die jeweilige Statistik verantwortlichen Direktor:in. Er:Sie kann sich dabei des methodischen Rats des Methodikbereichs und des juristischen Rats des:der Datenschutzbeauftragten bzw. des Rechtsbereiches bedienen. In Fällen mangelnder Einigung liegt die Entscheidung bei dem:der fachstatistischen Generaldirektor:in.

- Angewendete Verfahren zur Wahrung der Geheimhaltung sollen in der jeweiligen Standarddokumentation unter dem Punkt „Behandlung vertraulicher Daten“ dokumentiert werden.
- Die Verfahren müssen international gültigen Standards im Bereich statistischer Geheimhaltung entsprechen. Ein Überblick über anerkannte Methoden wird im [Handbuch für Statistische Geheimhaltung](#) des CENEX SDC gegeben. Die Verfahren zur Veröffentlichung vertraulicher Daten sind Teil des Schulungsprogramms von Statistik Austria.

Sind national erhobene, jedoch vertrauliche Daten an die EU (EUROSTAT) zu übermitteln, gilt die Verordnung (EG) Nr. 223/2009 des Europäischen Parlaments und des Rates über die europäischen Statistiken sowie die Einzelbestimmungen in den Rechtsakten der jeweiligen Fachbereiche. Die Daten sind in den im speziellen Rechtsakt vorgesehenen Format und Detaillierungsgrad an EUROSTAT zur Erstellung von Europäischen Statistiken zu übermitteln. Vertrauliche Daten werden so gekennzeichnet, dass sie auch von EUROSTAT nicht publiziert werden dürfen, aber sehr wohl in gesamteuropäische Aggregate eingehen können.

- Die Geheimhaltungsmethoden müssen kontinuierlich verbessert und angepasst werden, um den laufend verbesserten technischen Möglichkeiten von Datenangreifern gerecht zu werden. Dem entspricht es, dass sich Statistik Austria an europäischen und internationalen Forschungsprojekten zum Thema Geheimhaltung beteiligt.
- Der Umgang der Mitarbeiter:innen mit personenbezogenen und vertraulichen Daten ist, wie erwähnt, durch eine hausinterne Datensicherheitsvorschrift geregelt. Die Mitarbeiter:innen müssen sich schriftlich zur Einhaltung des Datengeheimnisses verpflichten. Aufgrund § 17 Abs. 4 BStatG werden Verstöße gegen die statistische Geheimhaltung als Verletzung des Amtsgeheimnisses gemäß § 310 StGB strafrechtlich verfolgt und mit Freiheitsstrafen bis zu drei Jahren geahndet.
- In der Statistik Austria ist ein:e Mitarbeiter:in als Datenschutzbeauftragte:r bestellt. Diese:r soll die Mitarbeiter:innen in der Umsetzung der datenschutzrechtlichen Bestimmungen unterstützen, Informationsdienste leisten und Hilfestellungen bei datenschutzrechtlichen Durchführungen bieten.

4) Zum Einsatz gelangende Software

Zur praktischen Umsetzung der gesetzlichen Vorschriften hinsichtlich Geheimhaltung für spezifische Daten ist die Anwendung unterschiedlicher Softwarepakete notwendig. Im Folgenden werden kurz die hausintern verwendeten Programme beschrieben.

4.1) Software für Geheimhaltung von aggregierten Daten

4.1.1) sdcTable

In der Statistik Austria wurde ein freies R-Paket – [sdcTable](#) - zur Geheimhaltung von Tabellendaten entwickelt, mit dem "sichere" (anonymisierte) Tabellen erzeugt werden können. Die Software kann im Batch-Modus verwendet werden und bietet die Möglichkeit als Eingangsparameter Mikrodaten oder bereits voraggregierte Tabellen zu verarbeiten. Nachdem alle, die Tabelle definierenden, Klassifikationsvariablen spezifiziert wurden, können unterschiedliche Varianten zur Primärunterdrückung (Fallzahlregel, Dominanzregeln, manuelles Sperren spezifischer Tabellenzellen) angewendet werden.

Zur Absicherung der Primärsperren kann aus unterschiedlichen Algorithmen gewählt werden. Außerdem können in sdcTable zwei Tabellen, die gemeinsame Zellen aufweisen (also verlinkt sind) in einem Schritt anonymisiert werden. Dabei wird sichergestellt, dass die gemeinsamen Zellen einen konsistenten Status in beiden Tabellen aufweisen.

4.1.2) STATcube

In der Statistik Austria wird von der Direktion IT die Software Suite SuperSTAR unter dem Namen STATcube zur Verfügung gestellt. Diese Software eignet sich sowohl für die Tabellierung von aggregierten Daten als auch von Mikrodaten. STATcube unterstützt standardmäßig folgende primäre Geheimhaltungsregeln:

- x-Regel,
- (n, k)-Dominanz oder Konzentrationsregel,
- Rundungsregel (random rounding).

Die sekundäre Geheimhaltung wird durch Verwendung vorgerechneter Geheimhaltungsmatrizen unterstützt (z. B. erstellt mit dem R-Paket sdcTable aus Punkt 4.1.1.).

Die SuperSTAR Software implementiert obige Regeln mithilfe von Softwaremodulen. Weitere Module können mittels der offengelegten Schnittstellendefinition erstellt werden.

4.2) Software zur Anonymisierung von Mikrodaten

4.2.1) sdcMicro

Hausintern wurde ein flexibles R-Paket - [sdcMicro](#) - entwickelt, in dem Standardmethoden zur Anonymisierung von Einzeldaten implementiert sind. So besteht etwa die Möglichkeit Variablen umzukodieren, lokale Unterdrückung und Mikroaggregation durchzuführen, Swapping anzuwenden oder stochastischen Noise auf Variablen aufzutragen. Zusätzlich können wichtige Kennzahlen wie die Verteilung der Schlüsselvariablen in der Stichprobe bzw. der Population berechnet werden und basierend auf diesen Daten globale sowie individuelle Risikomaßzahlen berechnet werden. Auch Methoden zur Beurteilung der Datenqualität nach Anwendung unterschiedlicher Anonymisierungsmethoden sind in sdcMicro vorhanden.

Ebenfalls in dem R-Paket sdcMicro enthalten ist eine grafische Oberfläche, mit der viele Methoden von sdcMicro unkompliziert angewendet werden können.

4.2.2) STATcube

Hiezu gilt das unter Punkt 4.1.1) Gesagte.