Dauerhafte Etablierung eines einheitlichen Prozesses für die Sicherstellung der Geheimhaltung im Statistischen Verbund

Ein Überblick für Bayern

Dipl.Math.oec.Univ. Andreas Nickl, Dipl.Soz.Wiss.Univ. Patrick Rothe

Der Schutz vertraulicher Angaben in Form der statistischen Geheimhaltung spielt in der amtlichen Statistik traditionell eine zentrale Rolle. In diesem Themenfeld geht der nachfolgende Beitrag auf die aktuellen Entwicklungen im Bereich der Standardisierung und Harmonisierung ein und beschreibt den Weg hin zur Etablierung eines innerhalb des Statistischen Verbunds einheitlichen Prozesses zur Sicherstellung der statistischen Geheimhaltung. Neben der Darstellung des prototypischen Ablaufs wird auch ein Blick auf die Beweggründe hierfür, auf den erwarteten Nutzen dieses Vorgehens sowie den aktuellen Stand der Umsetzung geworfen.

Ausgangslage

Die statistische Geheimhaltung gehört seit jeher zu einer der zentralen Grundlagen der amtlichen Statistik. Gesetzlich geregelt ist diese – auch als sogenanntes Statistikgeheimnis bezeichnet – in § 16 Bundesstatistikgesetz (BStatG). Ihr Ziel ist die Sicherstellung der Wahrung des Grundrechts auf informationelle Selbstbestimmung, wie es sich seit dem sogenannten „Volkszählungsurteil" von 1983 als prägend für den Schutz individueller Daten darstellt. So sind Einzelangaben durch die amtliche Statistik grundsätzlich geheim zu halten, es sei denn, eine rechtliche Ausnahmeregelung ermöglicht zu einem explizit geregelten Zweck hiervon abzuweichen. Für einzelne festgelegte Nutzergruppen existieren privilegierte Formen des Datenzugangs, beispielsweise für die empirisch forschende unabhängige Wissenschaft im Rahmen des sogenannten Wissenschaftsprivilegs. In der überwiegenden Vielzahl der Anwendungsfälle jedoch gilt die Vorgabe absoluter Anonymität für die Ergebnisse, die von den Statistischen Ämtern veröffentlicht werden.

Um diesen rechtlichen Anforderungen gerecht zu werden, verfügt die amtliche Statistik über ein breites Angebot an Instrumenten aus dem methodischen Werkzeugkasten (Rothe 2015a; Rothe 2015b). Handelte es sich bei den hierzu eingesetzten Verfahren in der Vergangenheit primär um informationsreduzierende Methoden – wie beispielsweise das Zellsperrverfahren, in der Regel in Kombination mit einer Mindestfallzahlregel –, so kamen in jüngerer Zeit neue innovative Vorgehensweisen aus dem Bereich der datenverändernden Verfahren – beispielsweise das beim Zensus 2011 eingesetzte Verfahren SAFE (Gießing et al. 2014) oder aber die post-tabulare stochastische Überlagerung (Cell-Key-Methode) (Höhne/Höninger 2019) – als neue Möglichkeiten hinzu. Bereits hieran ist zu erkennen, dass es nicht nur eine einzige Lösung gibt, mit der das Ziel der statistischen Geheimhaltung erreicht werden kann, sondern dass hierfür mehrere Wege und Ansätze zur Verfügung stehen. Dies gilt analog auch innerhalb eines gewählten Verfahrens hinsichtlich dessen konkreter Anwendung und Parametrisierung.

Wie kam es zur Etablierung eines einheitlichen Prozesses für die Sicherstellung der statistischen Geheimhaltung?

So vielfältig die zur Verfügung stehenden Möglichkeiten sind, so unterschiedlich kann zumindest theoretisch dieses Ziel erreicht werden. Auch

wenn sich die Zahl tatsächlicher Umsetzungen für gewöhnlich auf eine überschaubare Anzahl in Frage kommender Varianten – beispielsweise hinsichtlich angewandter Regeln oder der Parametrisierung – konzentrieren, so ergibt sich doch ein Nebeneinander von angewandten Regelungen und Herangehensweisen. Dies erwies sich im Rahmen der Zusammenarbeit innerhalb des Statistischen Verbunds[1] zunehmend als kontraproduktiv, beispielsweise dann, wenn Teilergebnisse unterschiedlicher Stellen, bei denen die Geheimhaltung nach unterschiedlichen Regeln durchgeführt wurde, im Rahmen einer koordinierten Auswertung für einen externen Kunden zusammengeführt wurden. Dieser musste damit umgehen, dass sich hier unterschiedliche Herangehensweisen bei der Umsetzung der statistischen Geheimhaltung durch Unterschiede in den jeweils zur Verfügung gestellten Ergebnissen bemerkbar machen. Dies ist nur ein Beispiel, das illustriert, weshalb eine abgestimmte und einheitliche Geheimhaltung von den Statistischen Ämtern als ein wichtiges Ziel eingestuft wurde und letztlich in die Einführung eines mehrstufigen, standardisierten Prozesses zur Gewährleistung einer einheitlichen Geheimhaltung mündete. Neben der methodischen Komponente ist die rechtliche Dimension nicht zu vernachlässigen, da unterschiedliche Interpretationen der gegebenen Rechtsgrundlage ebenfalls in voneinander abweichenden Vorgehensweisen in der alltäglichen Praxis führen können.

Ganz wesentlich sind neben der Erfassung des Status Quo („Wo stehen wir?“) und der Klärung des konkreten Geheimhaltungsbedarfs („Was müssen wir tun – und was nicht?“) die daran anschließende, fachlich fundierte Entscheidung für die geeignete Geheimhaltungsmethode (Rohde et al. 2018) sowie die Ausarbeitung eines abgestimmten Geheimhaltungskonzepts. Dabei sollten auch moderne Entwicklungen sowohl methodischer als auch technologischer Art mitgedacht werden. So wird insbesondere eine teilweise oder vollständige Automatisierung der im Geheimhaltungsprozess anfallenden Arbeitsgänge unter den Gesichtspunkten von Sicherheit und Effizienz als erstrebenswert angesehen. Diese Punkte stellen wichtige Ziele dar bei der Aufgabe, innerhalb einer spezifischen Fachstatistik mit ihren individuellen Eigenschaften eine konsistente und einheitliche Geheimhaltung zu verwirklichen und sind somit handlungsleitend für den im Folgenden detailliert vorgestellten idealtypischen Prozess.

Vorgehensweise eines Implementierungsprozesses für Geheimhaltungsverfahren

Wie eingangs beschrieben, ist der Schutz von Einzelangaben bei der Verarbeitung und Veröffentlichung statistischer Ergebnisse in der amtlichen Statistik von zentraler Bedeutung. Neben der nationalen Regelung der Einhaltung der Geheimhaltung im § 16 Abs. 1 BStatG hat dies auch Eingang in den Verhaltenskodex für europäische Statistiken (Amt für Veröffentlichungen der Europäischen Union 2018) gefunden:

> **Grundsatz 5 – Statistische Geheimhaltung und Datenschutz**
> Die Anonymität der Datenlieferanten, die Geheimhaltung ihrer Angaben, deren ausschließliche Verwendung für statistische Zwecke und die Sicherheit der Daten sind unter allen Umständen gewährleistet.

Die konkrete Ausgestaltung der Geheimhaltung erfolgt für jede einzelne Fachstatistik. Im „Handbuch zur statistischen Geheimhaltung“ (Gießing et al. 2018) sind die Methoden der statistischen Geheimhaltung ausführlich dargelegt.

Wenn eine Statistik erstmals erhoben oder das Auswertungskonzept überarbeitet wird, muss hierzu auch ein adäquates Geheimhaltungsverfahren bestimmt und eingeführt werden. Die Sicherstellung der Geheimhaltung sollte ein einheitlicher Prozess im Statistischen Verbund sein beziehungsweise die Geheimhaltung muss zwischen den Anwendungsbereichen abgestimmt werden.

Zunächst wird ein Prozess dargestellt, der Statistiken schrittweise zur Entwicklung eines Konzeptes für ein einheitliches Geheimhaltungsverfahren mit möglichst automatisierter Umsetzung führt.

1 Im Statistischen Verbund haben sich das Statistische Bundesamt sowie alle 14 Statistischen Ämter der Länder zusammengeschlossen.

Diese gemeinsame Entwicklung einer Geheimhaltungsmethode für eine Statistik im Statistischen Verbund kann als ein Standardprozess aufgefasst werden, der sich in vier aufeinander aufbauende Stufen gliedert:

I. Geheimhaltungsbedarf klären,
II. Geheimhaltungsmethode wählen,
III. Geheimhaltungskonzept erstellen,
IV. Umsetzung automatisieren.

Während die ersten drei Stufen (Bedarfsprüfung, Methodenwahl, Konzeptentwicklung) die eigentliche Planungsphase umfassen, verfolgt die vierte Stufe (Digitalisierung) das Ziel der praktischen Umsetzung und eventuell der Automatisierung des entwickelten Geheimhaltungsprozesses. Mit Digitalisierung ist dabei eine möglichst weitgehend automatisierte Umsetzung gemeint. Theoretisch durchläuft jede Statistik alle vier Stufen und tritt jeweils nach Abschluss einer Stufe in die darauf folgende ein.

Abbildung 1 veranschaulicht den Prozess beispielhaft.

Um eine Entwicklungsstufe abzuschließen und in die darauf folgende Entwicklungsstufe aufzurücken, müssen jeweils bestimmte Stufenziele erreicht werden.[2]

Auch wenn eine Statistik bereits bearbeitet wird, können Stufenziele (z. B. Stufe I oder Stufe II) schon erreicht sein. Geheimhaltungskonzepte sind darüber hinaus regelmäßig oder im Bedarfsfall (z. B. bei der Neukonzeption der fachstatistischen Geheimhaltung) zu überprüfen.

2 Eine detaillierte Beschreibung der einzelnen Stufen und deren Ziele ist zu finden in Gießing et al. 2020.

Wie ist der aktuelle Stand der Umsetzung der einheitlichen Geheimhaltung in allen dezentralen Statistiken mit Geheimhaltungsbedarf?

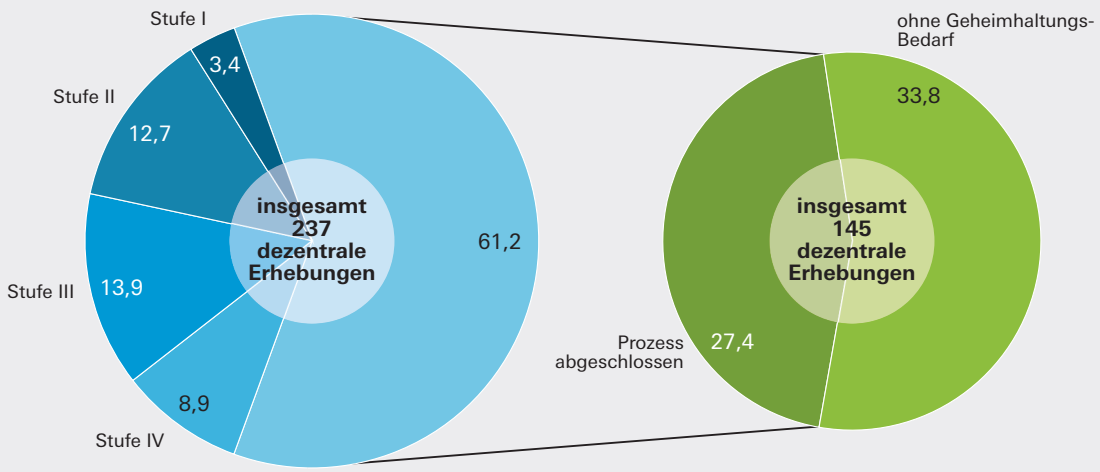
Erste Auswertungsergebnisse nach Rückmeldungen der Fachbereiche zeigen zum Stand Februar 2022 folgende Ergebnisse:
Insgesamt wurden 363 EVAS-Nummern in die Erhebung einbezogen, wobei davon 237 dezentrale Erhebungen darstellten.

EVAS ist die Abkürzung für „Einheitliches Verzeichnis aller Statistiken des Bundes und der Länder“. Es enthält alle bestehenden, historischen und geplanten Statistiken (Erhebungen, Rechenwerke und Register), die vom Statistischen Bundesamt und von den Statistischen Ämtern der Länder betreut werden.

Eine Bedarfsprüfung steht somit lediglich bei knapp 3% aller betrachteten dezentralen Erhebungen noch aus. Zu klären ist die Frage, ob für die Ergebnisse der jeweiligen Statistik grundsätzlich Geheimhaltungsbedarf besteht oder nicht. Bei annähernd 13% der Statistiken werden derzeit die in Frage kommenden Geheimhaltungsverfahren auf ihre Tauglichkeit hin untersucht. Die Stufe III „Konzeptentwicklung“ durchlaufen aktuell etwa 14% der Fälle. Die zentrale Aufgabe ist die Erarbeitung eines Fachkonzepts für eine im Verbund einheitliche und abgestimmte Geheimhaltung. In circa 9% der betrachteten dezentralen Statistiken erfolgen Aufbau und Optimierung

Abb. 1
Standardprozess zur Entwicklung einer Geheimhaltungsmethode für eine Statistik im Statistischen Verbund

Abb. 2
Anteil dezentraler Statistiken nach Einstufung des Handlungsbedarfs
Anteile in Prozent



eines automatisierten und medienbruchfreien Ablaufs der Geheimhaltung. Demzufolge befindet sich etwas mehr als ein Drittel der Erhebungen in den Stufen I bis IV. Der Großteil untergliedert sich somit in Statistiken, welche den Prozess bereits abgeschlossen haben (27 %), und in Statistiken ohne bestehenden Geheimhaltungsbedarf (34 %).

Geheimhaltungs-Informationsplattform

Als zentrales Werkzeug für die interne Berichterstattung der statistikspezifischen Geheimhaltung der Fachbereiche der Statistischen Ämter des Bundes und der Länder dient in absehbarer Zeit eine Geheimhaltungs-Informationsplattform (GHIP). Diese bietet den Fachbereichen die Möglichkeit, durch eine nutzerfreundliche Web-Oberfläche eine einfache und zeitnahe Dokumentation über Stand und Praxis ihrer statistikspezifischen Geheimhaltung zu erstellen. Anhand einer zentralen Bereitstellung relevanter Daten wird das kontinuierliche Monitoring der Umstellung der Geheimhaltung auf standardisierte und automatisierte Verfahren durch die GHIP vereinfacht. Ebenso erleichtert dies die Berichterstattung an Leitungsebenen und Gremien des Statistischen Verbunds. Die GHIP wird als neues Modul des Metadatenmanagementsystems[3] des Statistischen Verbunds umgesetzt.

Durch die Bereitstellung einer zentralen, metadatengetriebenen Informationsplattform werden die Dokumentationsaufwände der Fachbereiche reduziert, zudem können diese sich leichter an der Geheimhaltungspraxis vergleichbarer Fachstatistiken orientieren. Ebenso ermöglicht dies eine Beschleunigung bei der Umstellung der zentralen und dezentralen Statistiken auf eine standardisierte und automatisierte Geheimhaltung, wie es in den strategischen Maßnahmen der amtlichen Statistik gefordert ist.

Der Aufwand der Fachbereiche für regelmäßige Berichterstattungen wird reduziert, da diese mit der GHIP selbst angeben, welcher Stand erreicht ist: Die Abfragen zum Monitoring des Umsetzungsstands bei der statistikspezifischen Geheimhaltung können entfallen. Ebenso wird es der Leitungsebene erleichtert, sich schnell einen Überblick über den Umsetzungsstand hinsichtlich der automatisierten Geheimhaltung zu verschaffen, da alle relevanten Informationen an einem Ort vorgehalten werden, auf deren Grundlage sich

3 Nähere Informationen zum Metadatenmanagement sind zu finden in Lindenstruth/Claußen 2017.

automatisiert oder on-demand Standardauswertungen ausgeben lassen, die wiederum in den Entscheidungsprozess einfließen können.

Welchen Nutzen bringt die Etablierung eines vereinheitlichten Geheimhaltungsprozesses mit sich?

Es sprechen mehrere Gründe dafür, einen umfassenden, aber somit auch arbeitsaufwendigen Prozess zur Standardisierung und Neugestaltung der statistischen Geheimhaltung in allen Bereichen der amtlichen Statistik zu etablieren. Der zu erzielende mittel- und langfristige Nutzen übersteigt die durchaus hohen Aufwände, die damit einhergehen, deutlich: So erleichtern identische Regelungen zum Umgang mit der statistischen Geheimhaltung sowohl in den einzelnen betroffenen Ämtern als auch in der Zusammenarbeit zwischen diesen deren praktische Durchführung. In den einzelnen Fachabteilungen, Referaten, Sachgebieten oder Teams tätige und mit Fragen der statistischen Geheimhaltung betraute Personen erhalten durch einheitliche und gut dokumentierte Vorgehensweisen Handlungssicherheit, die zu einer nachvollziehbaren und vergleichbaren Durchführung der statistischen Geheimhaltung mit sicheren Ergebnissen führt. Aus der Perspektive der Datennutzenden ergibt sich insbesondere durch die Sicherstellung einer abgestimmten Geheimhaltung eine Vergleichbarkeit der Ergebnisse unabhängig davon, in welchem Statistischen Amt diese generiert wurden. Darüber hinaus resultieren aus Nutzersicht durch die Neugestaltung der statistischen Geheimhaltung in vielen Bereichen neue Möglichkeiten, insbesondere was die technische Ausgestaltung sowohl des – idealerweise automatisierten – Geheimhaltungsprozesses als auch die zeitgemäße Darbietung der Daten für die Öffentlichkeit, beispielsweise mittels flexibler Auswertungsdatenbanken angeht. Nicht zuletzt dienen die im Rahmen der Prozessetablierung erhobenen Angaben zum konkreten Vorgehen in den einzelnen Fachstatistiken als fundierte Datenbasis für Entscheidungen sowohl innerhalb der einzelnen Statistischen Ämter als auch des gesamten Statistischen Verbunds. Durch eine kontinuierliche Pflege und Aktualisierung dieser Informationen mithilfe der geplanten Informationsplattform soll dies dauerhaft gewährleistet werden. Bei all dem stellt ein standardisierter Prozessablauf ein wichtiges Hilfsmittel dar, das den Betroffenen Orientierung geben soll bei der Konzeption und Umsetzung der notwendigen Maßnahmen. Damit fügt sich dies auch als ein Instrument zur Qualitätssicherung in entsprechende Anstrengungen der Statistischen Ämter in Deutschland sowie des Europäischen Statistischen Systems ein.

Fazit/Ausblick

Dieser Beitrag bietet einen Überblick über den aktuellen Stand der Arbeiten zur Etablierung eines einheitlichen Prozesses zur Sicherstellung der statistischen Geheimhaltung.

Sowohl die amtliche Statistik in Deutschland als auch die Nutzenden ihrer Veröffentlichungen und Daten werden von einer entsprechenden Prozessgestaltung mittel- bis langfristig profitieren – auch wenn im Rahmen der aktuellen Schritte diverse Anstrengungen als Vorleistungen in das Vorhaben eingebracht werden müssen. Die zu erwartenden Vorteile für die Arbeit der Statistischen Ämter dürften diese jedoch deutlich überwiegen. Nicht zuletzt werden es auch die Auskunftspflichtigen und freiwillig Befragten sein, die von einer optimierten Vorgehensweise im Bereich der statistischen Geheimhaltung profitieren werden, indem hierdurch der Schutz ihrer Einzelangaben weiterhin auf sehr hohem Niveau sichergestellt und zugleich durch konsistente Vorgehensweisen eine schnellere und nutzerfreundliche Bereitstellung von Ergebnissen für Datennutzende möglich wird.

Literatur

Amt für Veröffentlichungen der Europäischen Union (2018): Verhaltenskodex für europäische Statistiken.

Bundesstatistikgesetz in der Fassung der Bekanntmachung vom 20. Oktober 2016 (BGBl. I S. 2394), das zuletzt durch Artikel 2 des Gesetzes vom 14. Juni 2021 (BGBl. I S. 1751) geändert worden ist.

Gießing, Sarah / Habla, Heike / Höninger, Julia / Hoffmeister, Rita / Merz, Franz-Josef / Richter, Alexander / Scharnhorst, Sebastian / Schmidtke, Kerstin / Spies, Lydia / Tonte, Andreas / Uhrich, Stefanie (2018): Handbuch zur Statistischen Geheimhaltung. Stand 23. Dezember 2020. Internes Dokument des Statistischen Verbunds.

Gießing, Sarah / Heinzl, Felix / Kleber, Birgit / Wilke, Achim (2014): Geheimhaltung beim Zensus 2011. In: Bayern in Zahlen 11/2014, S. 673–681.

Gießing, Sarah / Höninger, Julia / Ramic, Arijana Amina / Rohde, Johannes (2020): Ein Implementierungsprozess für Geheimhaltungsverfahren in vier Stufen. In: Zeitschrift für amtliche Statistik Berlin-Brandenburg 02/2020, S. 6–10.

Höhne, Jörg / Höninger, Julia (2019): Die Cell-Key-Methode – ein Geheimhaltungsverfahren. In: Statistische Monatshefte Niedersachsen, S. 18–23.

Lindenstruth, Thomas / Claußen, Sven (2017): Metadatenmanagement als neue Integrationsarchitektur. In: Wirtschaft und Statistik 05/2017, S. 76–86.

Rohde, Johannes / Seifert, Christiane / Gießing, Sarah (2018): Entscheidungskriterien für die Auswahl eines Geheimhaltungsverfahrens. In: Wirtschaft und Statistik 03/2018, S. 90–104.

Rothe, Patrick (2015a): Statistische Geheimhaltung – Der Schutz vertraulicher Daten in der amtlichen Statistik. Teil 1: Rechtliche und methodische Grundlagen. In: Bayern in Zahlen 5/2015, S. 294–303.

Rothe, Patrick (2015b): Statistische Geheimhaltung – Der Schutz vertraulicher Daten in der amtlichen Statistik. Teil 2: Herausforderungen und aktuelle Entwicklungen. In: Bayern in Zahlen 08/2015, S. 482–489.