

Claußen, Sven

Article

Metadatenmodell der amtlichen Statistik

WISTA - Wirtschaft und Statistik

Provided in Cooperation with:

Statistisches Bundesamt (Destatis), Wiesbaden

Suggested Citation: Claußen, Sven (2026) : Metadatenmodell der amtlichen Statistik, WISTA - Wirtschaft und Statistik, ISSN 1619-2907, Statistisches Bundesamt (Destatis), Wiesbaden, Vol. 78, Iss. 1, pp. 113-126

This Version is available at:

<https://hdl.handle.net/10419/337207>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

METADATENMODELL DER AMTLICHEN STATISTIK

Sven Claußen

➤ **Schlüsselwörter:** Metadaten – Metadatenmanagement – Standardisierung – Digitalisierung – Automatisierung

ZUSAMMENFASSUNG

Für die Statistikproduktion ist ein gut funktionierendes, übergreifendes Management von Metadaten von großer Bedeutung, um die notwendigen Geschäftsprozesse qualitativ hochwertig und effizient umzusetzen. Mithilfe einer zentralen Systemkomponente werden dazu verschiedene Teilsysteme in einer Gesamtarchitektur für Metadaten miteinander kombiniert. Diese dient in den Produktionsprozessen sowohl als zentrale Bezugsquelle als auch zur Speicherung von Metadaten. Konzeptionelle Grundlage ist hierbei das Metadatenmodell der amtlichen Statistik (MedaS), das auf einem internationalen Standard basiert und 2025 in einer neuen Version im Metadatenmanagementsystem implementiert wurde. MedaS bietet darüber hinaus die Grundlage für eine verbesserte Kommunikation zwischen Fachbereichen, Informationstechnik und Managementebene.

➤ **Keywords:** metadata – metadata management – standardisation – digitalisation – automation

ABSTRACT

In statistics production, a well-functioning and comprehensive metadata management system is crucial to ensure that the necessary business processes are implemented efficiently and to a high quality standard. A central system component is used to combine various individual solutions into an overall metadata architecture to this end. In the production processes, it is used both as a central source of reference and for storing metadata. The conceptual framework for this is the "Metadata Model of Official Statistics" (MedaS), which is based on an international standard. A new version of MedaS was implemented in the metadata management system in 2025. MedaS also facilitates communication among specialised statistics units, IT experts and management.



Dr. Sven Claußen

ist Diplom-Informatiker und Referent im Referat „Metadaten, Qualität und Datenmanagement“ des Statistischen Bundesamtes. Zu seinen Arbeitsschwerpunkten zählen die Konzeption und Qualitätssicherung von Methoden und Verfahren zur Verwaltung von Metadaten.

1

Einleitung

Die Statistikproduktion wird durch eine Vielzahl von IT-Werkzeugen unterstützt, die untereinander Daten und Metadaten austauschen. Wichtige Basis für eine erfolgreiche Interoperabilität ist ein standardisiertes Informationsmodell, das ein gemeinsames Verständnis für die notwendigen Informationsobjekte sowie deren Eigenschaften und Beziehungen zueinander vermittelt. Im Statistischen Verbund¹ wird hierzu [das Metadatenmodell der amtlichen Statistik \(MedaS\)](#) eingesetzt, das auf einem internationalen Standard der Wirtschaftskommission für Europa der Vereinten Nationen (UNECE) aufsetzt. Im Zusammenhang mit dem [Geschäftsprozessmodell Amtliche Statistik \(GMAS\)](#) wird mithilfe von MedaS definiert, welche Objekte in jedem Teilprozess benötigt und produziert werden. MedaS hat auch außerhalb der Informationstechnik (IT) eine besondere Bedeutung, da sich durch einen einheitlichen Sprachgebrauch die Kommunikation auf verschiedenen Ebenen verbessert.

Der Beitrag stellt das Metadatenmodell der amtlichen Statistik (MedaS) vor. Adressaten sind beispielsweise diejenigen, die in statistischen Organisationen arbeiten. Er skizziert die Vorteile und den Nutzen des Modells für die Statistischen Ämter des Bundes und der Länder, wenn es sowohl fachlich als auch technisch flächendeckend angewendet wird.

Nach einem Überblick über den Modernisierungsbedarf, den Stand der Standardisierung und das Metadatenmanagement des Statistischen Verbundes in Kapitel 2 erläutert Kapitel 3 ausführlich das Metadatenmodell der amtlichen Statistik. Wie konkret die Anwendung von MedaS als Basis für die Architektur des Metadatenmanagements im Statistischen Verbund erfolgt, wird in Kapitel 4 thematisiert. Ein Fazit in Kapitel 5 rundet den Artikel ab.

1 Den Statistischen Verbund bilden die Statistischen Ämter des Bundes und der Länder.

2

Sachstand und Weiterentwicklung

2.1 Modernisierungsbedarf

Die Statistikproduktion steht zunehmend vor der Herausforderung, bei schwindenden Ressourcen weiterhin eine gute Datenqualität zu gewährleisten. Höhere Ansprüche, insbesondere an die Aktualität, sind unter diesen Rahmenbedingungen nicht mehr ohne Weiteres zu erfüllen. Erst durch Investitionen in die Modernisierung der Produktionsprozesse lassen sich langfristig und nachhaltig in größerem Umfang Kosten einsparen, ohne die bestehenden Qualitätsansprüche zu vernachlässigen.

Die Modernisierung umfasst eine übergreifende Standardisierung und Kopplung von Standardwerkzeugen, um das vorhandene Automatisierungspotenzial möglichst effizient zu nutzen (Statistisches Bundesamt, 2022). Anstelle von manuellen Export- und Importvorgängen werden benötigte Daten und Metadaten ohne Zusatzaufwand direkt zwischen den Komponenten ausgetauscht. Dies setzt eine IT-Architektur voraus, die verteilte Systeme auf Basis von Standards miteinander koppelt (Lindenstruth/Claußen, 2017; Blumöhr und andere, 2017). Die inhaltlichen sowie fachlichen Standards sind Voraussetzung für Kompatibilität und Austausch der Informationsobjekte in passenden Formaten, um diese automatisiert verarbeiten zu können.

2.2 Standardisierung

Auf der ganzen Welt erledigen statistische Organisationen ähnliche Aufgaben. Die dabei eingesetzten Prozesse unterscheiden sich zwar im Detail, jede dieser Aktivitäten verwendet und produziert aber ähnliche Informationen (zum Beispiel definieren alle Organisationen Grundgesamtheiten für ihre statistischen Beobachtungen, verwenden statistische Klassifikationen, erstellen Datensätze und verbreiten Informationen). Grundlage der Statistikproduktion sind insbesondere die international abgestimmten

Standards [Generic Statistical Business Process Model \(GSBPM\)](#) und [Generic Statistical Information Model \(GSIM\)](#), die in Arbeitsgruppen der UNECE mit Beteiligung des Autors in den Jahren 2007 bis 2019 entwickelt wurden (UNECE, 2025). Das GSBPM definiert den gesamten Geschäftsprozess der Erstellung einer Statistik von der Planung bis zur Auswertung anhand von acht Phasen und 44 Teilprozessen (Gehle/Lüüs, 2017). Das GSIM standardisiert alle Informationsobjekte, die bei der Gestaltung und Erstellung von Statistiken notwendig sind, durch eine konsistente Beschreibung und die Beziehungen untereinander. Vor einer nationalen Adaptierung und technischen Umsetzung sind aufgrund bewusst vorhandener Freiheitsgrade insbesondere die Objektattribute umfassender zu spezifizieren.

Der Statistische Verbund hat GSBPM und GSIM in eigenen Standards adaptiert. GMAS und MedaS sind nicht nur die deutschen Übersetzungen der internationalen Modelle, sondern passen sie durch Präzisierungen und Ergänzungen auch an die Bedürfnisse und Gepflogenheiten der amtlichen Statistik in Deutschland an. Die Verwendung standardisierter Begrifflichkeiten und zusätzliche Beispiele erleichtern das Verständnis von GMAS und MedaS. Gleichwohl bleiben Bedeutungen der adaptierten Geschäftsprozesse und Metadatenobjekte sowie ihre Beziehungen untereinander bestehen, um auch langfristig die Kompatibilität zu GSBPM und GSIM zu sichern.

2.3 Metadatenmanagement der Statistischen Ämter des Bundes und der Länder

Das Metadatenmanagement des Statistischen Verbundes basiert auf GMAS und MedaS, die sich sinnvoll ergänzen. Der Fokus dieses Artikels liegt auf der einführenden Beschreibung von MedaS, das 2025 in der Version 2.0 erstmals umfassend alle Informationsobjekte der Statistikproduktion enthält. Von diesem Standard profitieren nicht nur IT-Verantwortliche, sondern alle am Produktionsprozess Beteiligten, da MedaS (in Kombination mit GMAS) statistik- und organisationsübergreifend die gemeinsame Sprache für die Beschreibung der Statistikproduktion im Statistischen Verbund bildet. Trotz eingedeutsch-

ter Begrifflichkeiten sind die integrierten Verweise auf GSIM auch bei der internationalen Zusammenarbeit, zum Beispiel bei Studienbesuchen, Kooperationen sowie Sitzungen des Statistischen Amtes der Europäischen Union (Eurostat) und der UNECE, von großem Nutzen.

Die Aufgabe, geeignete technische Infrastrukturkomponenten für das Management von Metadaten bereitzustellen, obliegt dem Statistischen Bundesamt. Eine der wichtigsten Komponenten ist hierbei das Metadatenmanagementsystem, das als zentrale Drehscheibe die Speicherung, Verteilung und Standardisierung aller notwendigen Metadatenobjekte übernimmt. Langjährig bewährte metadatenhaltende Systeme wie die Statistikdatenbank, die Erhebungsdatenbank oder der Klassifikationsserver werden dabei nicht ersetzt, sondern deren Metadaten (zum Beispiel Klassifikationen) mithilfe sogenannter Adapter übernommen und in das standardisierte MedaS transformiert. Die Zwischenspeicherung und Bereitstellung über zentrale Standardschnittstellen gewährleisten einen einfachen, dauerhaften Zugriff und eine hohe Verfügbarkeit notwendiger Metadaten. Für den weiteren Ausbau – um letztlich alle MedaS-Objekte über moderne Bedienoberflächen verwalten zu können – kann das Metadatenmanagementsystem sukzessive um sogenannte Anwendungsmodule erweitert werden. Diese lassen sich insofern schlank und ressourcengünstig entwickeln, als die Speicherung der Inhalte direkt im Metadatenmanagementsystem erfolgen kann. Eine eigene Datenbank in den Anwendungsmodulen sowie aufwendige Modelltransformationen sind nicht notwendig.

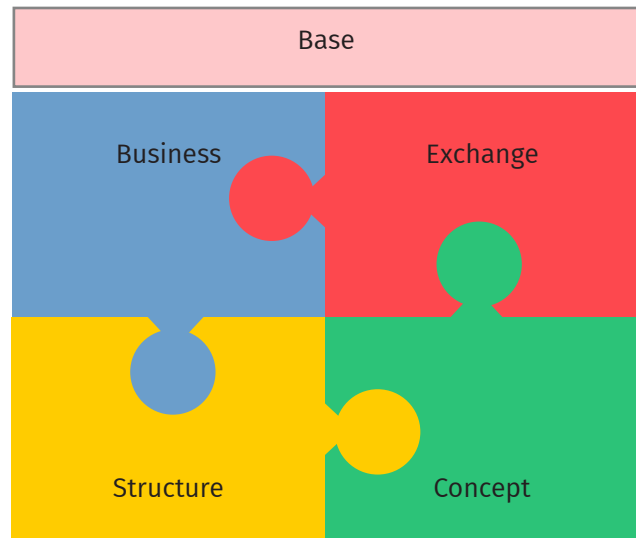
3

Das Metadatenmodell der amtlichen Statistik

Ein Informationsmodell beschreibt alle Objekte, die Informationen über die reale Welt spezifizieren – sogenannte Informationsobjekte – und deren Beziehungen zueinander. Für spezielle Bedürfnisse wie die Statistikproduktion kann es spezialisierte Informationsmodelle geben, wie GSIM oder MedaS, die auf die Verarbeitung und Verwaltung von Daten und Metadaten in statistischen Organisationen fokussieren. MedaS als deutsche Adaption des internationalen Standards GSIM wurde im Wesentlichen an den deutschen Sprachgebrauch angepasst und um einige wenige Objekte sowie zahlreiche Objektattribute ergänzt. Die Einteilung aller Informationsobjekte in

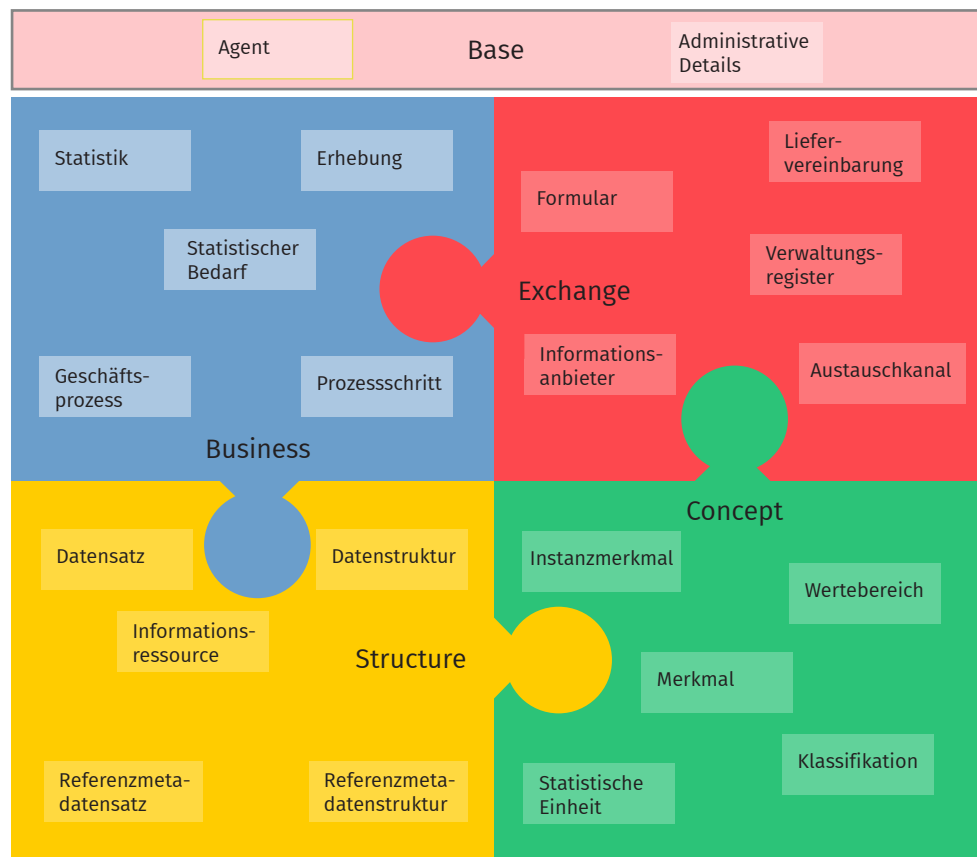
Grafik 1

Die fünf Hauptgruppen von Informationsobjekten im Metadatenmodell der amtlichen Statistik



Grafik 2

Auswahl einiger Informationsobjekte mit Zuordnung zu den fünf Hauptgruppen im Metadatenmodell der amtlichen Statistik



fünf Hauptgruppen wurde einschließlich der Originalbezeichnungen und farblichen Kennzeichnungen beibehalten (UNECE, 2025; Lindenstruth/Claußen, 2017). [↗ Grafik 1](#)

Die aktuelle Version des Metadatenmodells der amtlichen Statistik MedaS 2.0 enthält 131 Informationsobjekte, die den fünf Hauptgruppen zugeordnet sind. [↗ Grafik 2](#)

In einem Informationsmodell stehen die Informationsobjekte miteinander in Beziehung.

[↗ Grafik 3](#) zeigt die folgenden Zusammenhänge zwischen den Informationsobjekten:

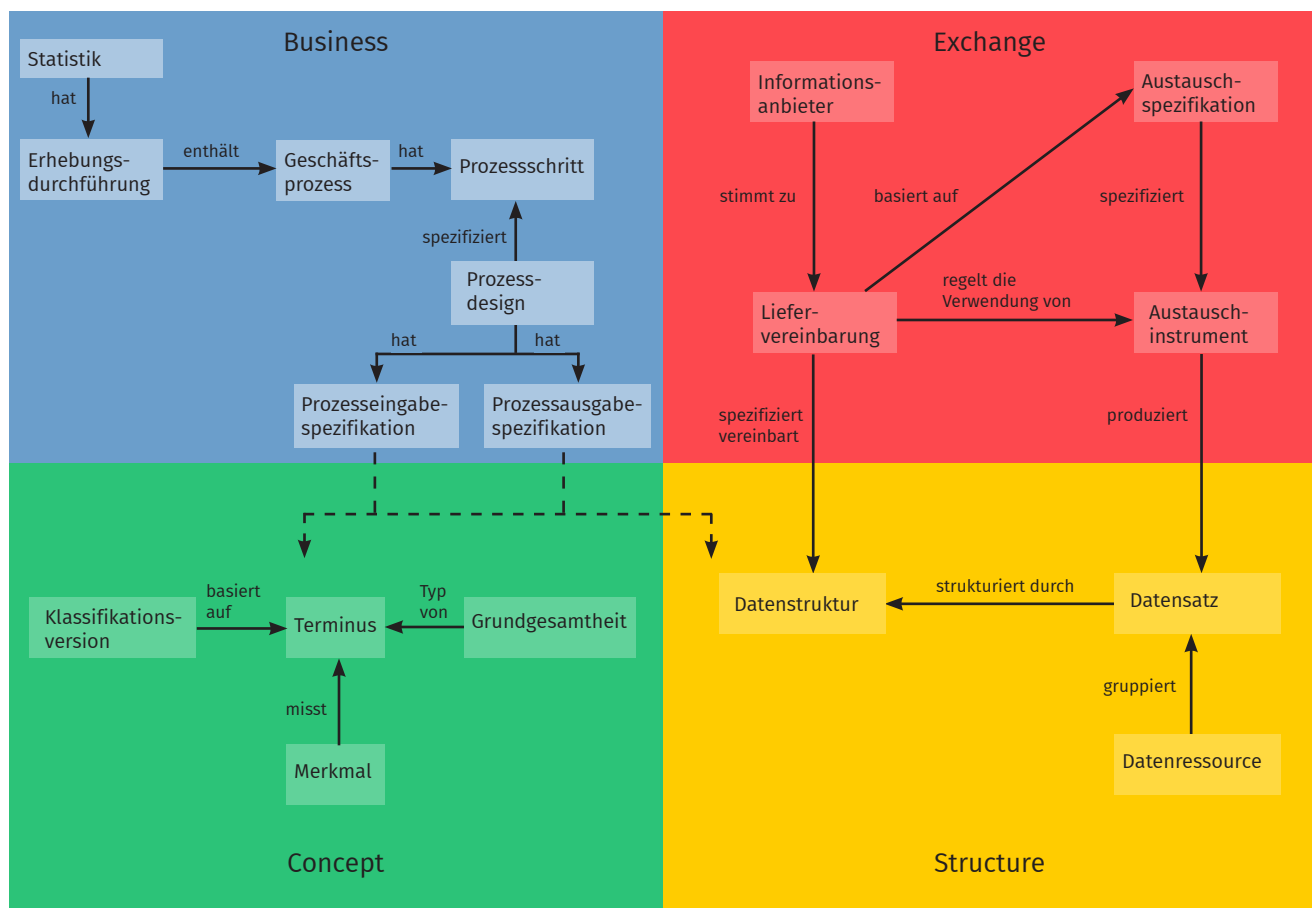
Eine statistische Organisation initiiert eine Statistik. Die Statistik entspricht einer fortlaufenden Aktivität wie einer Erhebung oder einer Ergebnisreihe und

hat eine Erhebungsdurchführung, beispielsweise vierteljährlich oder jährlich. Die Erhebungsdurchführung umfasst eine Reihe von Geschäftsprozessen. Die Geschäftsprozesse bestehen aus mehreren Prozessschritten, die durch ein Prozessdesign spezifiziert werden. Diese Prozessdesigns haben Prozesseingangsspezifikationen und Prozessaussgangsspezifikationen. Bei den Spezifikationen handelt es sich häufig um Informationen, die sich auf Konzepte und Strukturen beziehen (zum Beispiel Klassifikation, Merkmal, Grundgesamtheit, Datenstruktur und Datensatz).

Bezieht sich der Geschäftsprozess beispielsweise auf die Erhebung von Daten, gibt es einen Informationsanbieter, der auf Basis einer Liefervereinbarung der statistischen Organisation Daten zur Verfügung stellt. Diese Liefervereinbarung legt eine vereinbarte

Grafik 3

Beziehungen zwischen den Informationsobjekten im Metadatenmodell der amtlichen Statistik



Datenstruktur fest und regelt den Einsatz eines Austauschinstrument, das für die eingehenden Informationen verwendet wird. Das Austauschinstrument könnte in diesem Fall ein Fragebogen oder ein Verwaltungsregister sein, das durch die Austauschspezifikation definiert wird.

Der vom Austauschinstrument produzierte Datensatz wird in einer Datenressource (gegebenenfalls mit anderen Datensätzen) gruppiert und durch eine Datenstruktur strukturiert.

Dieses Beispiel zeigt, dass MedaS keine statistik- oder erhebungsspezifischen Bezeichnungen verwendet; dadurch ist es übergreifend einsetzbar. Es abstrahiert von Besonderheiten einer Erhebungsdurchführung und deckt dennoch die Vielfalt der Statistikproduktionen ab. So verbergen sich hinter einem Informationsanbieter alle Auskunftseinheiten, ob mit oder ohne Auskunftspflicht, ob in einer Stichprobe oder einer Vollerhebung, ob über eine Online-Befragung, ein Verwaltungsregister oder über einen klassischen Papierfragebogen. Erst durch die Befüllung der abstrakten Objekte konkretisiert sich eine Statistikproduktion, beispielsweise durch Festlegung eines Merkmalsträgertyps mit „Person“, der Definition der Grundgesamtheit mit „Personen über 18 Jahre mit Wohnsitz in Deutschland“ und der Auflistung der zugehörigen (etwa 70 Millionen) statistischen Einheiten. Als Austauschinstrument ist gegebenenfalls ein zentrales Melderegister vorgesehen. Liegen die notwendigen technischen und juristischen Voraussetzungen vor, können Werkzeuge der Statistikproduktion auf Basis der hinterlegten Datenstruktur entsprechende Datensätze aus diesem Melderegister automatisiert entnehmen und auf Basis der weiteren Metadaten weiterverarbeiten (einschließlich Geheimhaltung, Imputation und so weiter).

An dieser Stelle sei ausdrücklich darauf hingewiesen, dass MedaS kein fachliches Modell der zu verarbeitenden Daten ist. So hat beispielsweise eine Person ein Geschlecht, ein Geburtsdatum und einen Geburtsort. Ein Unternehmen ist beispielsweise durch einen Unternehmenstyp, eine Geschäftsadresse und eine Anzahl Beschäftigter (die wiederum Personen sind) gekennzeichnet. Diese Zusammenhänge obliegen einem semantischen Daten-

modell über die in der Statistik zu betrachtenden Merkmalsträgertypen, das gesondert zu definieren und innerhalb des Statistischen Verbundes aufgrund der vorhandenen Verzahnungen statistikübergreifend abzustimmen wäre.

Zwar ist in diesem Aufsatz keine detaillierte Definition aller MedaS-Objekte möglich, nachfolgend soll allerdings eine vollständige Auflistung mit einer leichtverständlichen Kurzbeschreibung das Verständnis fördern:

3.1 Base

Die Gruppe „Base“ stellt Informationen bereit, die von anderen Informationsobjekten wiederverwendet werden können, um Funktionen wie Identifikation und Versionierung zu unterstützen:

- > **Administrative Details:** speichert insbesondere technische Angaben (Status, Zeitpunkt der Erstellung und so weiter)
- > **Adresse:** umfasst räumliche Informationen, um einen Akteur lokalisieren zu können
- > **Agent:** stellt einen Akteur dar und wird durch folgende Untertypen konkretisiert:
 - > **Organisation:** eine Ansammlung von Menschen, die gemeinsam organisiert sind
 - > **Person:** Einzelpersonen, die Metadaten bearbeiten und nutzen
 - > **Software:** Akteur, insbesondere für automatisierte Prozessabläufe
- > **Agent in einer Rolle:** Verknüpfung eines Akteurs mit einer bestimmten Rolle bezüglich eines konkreten Informationsobjekts
- > **Änderungsereignis:** bestimmtes Ereignis, das einen konkreten Einfluss auf Informationsobjekte hat
- > **Artefakt:** Basisobjekt für (nahezu) alle Informationsobjekte, um grundlegende Objekteigenschaften zu bündeln
- > **Attribut:** Eigenschaft eines Objekts
- > **i18n-Text:** dient der Verwaltung mehrsprachiger Texte

- > **Informationsobjekt:** dient der Modellierung des eigenen Informationsmodells
- > **Rolle:** Akteure können je nach Kontext verschiedene Rollen einnehmen:
 - > **Besitzer:** verantwortlich für die Definition, Spezifizierung, Erstellung und Pflege eines Objekts
 - > **Betreuer:** kümmert sich um administrative und betriebliche Angelegenheiten, aber ohne Entscheidungsbefugnis
 - > **Kontakt:** verantwortlich für die Bereitstellung zusätzlicher Informationen über ein Objekt

3.2 Business

Die Gruppe „Business“ wird verwendet, um die Entwürfe und Pläne von Statistiken sowie die Prozesse zu erfassen, die zur Bereitstellung statistischer Ergebnisse durchgeführt werden. Dies umfasst auch die Identifizierung eines statistischen Bedarfs, die Geschäftsprozesse einer Statistik sowie deren Evaluierung.

- > **Änderungsdefinition:** Ergebnis konzeptioneller Arbeiten aufgrund eines statistischen Bedarfs
- > **Bewertung:** Evaluation zur Verbesserung von Produktionsprozessen (und Produkten) im Hinblick auf den statistischen Bedarf
- > **Erhebung:** spezifiziert Prozesse und Objekte für konkrete Erhebungsdurchführungen unter Berücksichtigung der spezifischen Anforderungen
- > **Erhebungsdurchführung:** beschreibt die konkrete Ausführung einer Iteration einer Erhebung für einen bestimmten Berichtszeitraum
- > **Geschäftsdienst:** ermöglicht den Zugriff auf eine bestimmte Geschäftsfunktion
- > **Geschäftsfall:** begründet die Durchführung eines Statistikunterstützungsprogramms, um eine neue Erhebung für eine bestehende oder neue Statistik zu initiieren
- > **Geschäftsfunktion:** definiert den Geschäftszweck, dem ein Geschäftsdienst oder ein Prozessschritt dient
- > **Geschäftsprozess:** strukturierte und wiederholbare Aktivität, die eine oder mehrere Geschäftsfunktionen erfüllt
- > **Prozessausgabe:** neue oder veränderte Objekte, die bei der Ausführung eines Prozessschritts entstehen (nutzbar als Prozesseingabe für andere Prozessschritte):
 - > **Kernausgabe:** Objekt, für das der Prozessschritt maßgeblich dient (typischerweise die Kerneingabe für einen nachfolgenden Prozessschritt oder das Endprodukt der Statistikproduktion)
- > **Prozessausführungsprotokoll:** zeichnet technische Informationen über die Ausführung eines Prozessschritts auf, um Prozessabläufe und etwaige Störungen zu dokumentieren
- > **Prozessmetrik:** zeichnet qualitative und inhaltliche Informationen über die Ausführung eines Prozessschritts auf, um das Qualitätsmaß der Kernausgabe beurteilen zu können
- > **Prozessausgabespezifikation:** definiert die Prozessausgaben für die Ausführung eines Prozessdesigns
- > **Prozessdesign:** definiert die Ausführung eines Prozessschritts als Teil eines bestimmten Geschäftsdienstes
- > **Prozesseingabe:** Objekte, die zur Ausführung eines Prozessschritts benötigt werden:
 - > **Kerneingabe:** Objekt, für das der Prozessschritt maßgeblich dient (das beispielsweise transformiert oder als Grundlage für die Erzeugung neuer Objekte verwendet wird)
 - > **Parametereingabe:** dient der Konfiguration generischer Prozessschritte, um sich an spezifische Bedürfnisse anzupassen
 - > **Prozessunterstützungseingabe:** Objekt, das für den Zweck des Prozessschritts hilfsweise benötigt wird
- > **Prozesseingabespezifikation:** definiert die notwendigen Prozesseingaben für die Ausführung eines Prozessdesigns
- > **Prozessmethode:** beschreibt unabhängig von Technologien und Werkzeugen eine bestimmte

Methodik zum Ausführen eines Prozessschritts (gegebenenfalls Anwendung bestimmter Regeln)

- > **Prozessmuster:** Definition einer Vorlage zur Beschleunigung von Designprozessen und zur gemeinsamen Nutzung und Wiederverwendung von Designmustern, die sich als effektiv erwiesen haben
- > **Prozessschritt:** implementiert das definierte Prozessdesign, um die gewünschten Ergebnisse zu erzeugen
- > **Prozessschrittinstanz:** führt einen Prozessschritt in einem bestimmten Kontext mit tatsächlichen Instanzen der Eingaben (zum Beispiel bestimmte Datensätze) aus, sodass einzigartige Ergebnisse erzielt werden, auch wenn der Prozessschritt konstant bleibt
- > **Prozesssteuerung:** steuert die Ausführung der Prozessschritte gemäß Prozesssteuerungsdesign
- > **Prozesssteuerungsdesign:** definiert den Fluss zwischen den Prozessschritten auf Basis von Entscheidungspunkten
- > **Referenzdokument:** dokumentiert – meist unstrukturiert – Methodik, Rahmenbedingungen, Grundlagen, Best Practices und so weiter in physischen Publikationen (zum Beispiel Büchern) oder elektronischen Dateien (zum Beispiel Qualitätsberichten)
- > **Regel:** ermöglicht die Prüfung oder Transformation beliebiger Objekte durch Anwendung einer hinterlegten Anweisung
- > **Statistik:** sorgt für organisatorischen Umgebungskontext, in dem Aktivitäten zur Erstellung von Statistiken innerhalb einer statistischen Organisation durchgeführt werden
- > **Statistikunterstützungsaktivität:** umfasst Funktionen wie Metadatenmanagement, Datenmanagement, Qualitätsmanagement, methodische Forschung und Designfunktionen
- > **Statistischer Bedarf:** dokumentiert eine vorgeschlagene oder auferlegte Anforderung, Anfrage oder sonstige Mitteilung, die bei einer statistischen Organisation eingegangen ist:

> **Informationsanfrage:** dokumentiert eine Anfrage nach neuen Informationen für einen bestimmten Zweck

> **Umgebungsänderung:** dokumentiert eine Änderung im Tätigkeitsbereich einer statistischen Organisation

3.3 Concept

Die Gruppe „Concept“ wird verwendet, um die Bedeutung von Daten zu definieren und ein Verständnis dafür zu vermitteln, was die Daten messen:

- > **Ausprägung:** spezifiziert übergreifend eine bestimmte Eigenschaft, die in einem beliebigen Kontext verwendet werden kann
- > **Bezeichnung:** benennt ein Objekt zur (eindeutigen) Identifizierung
- > **Datum:** enthält den tatsächlichen Wert, der einen Datenpunkt füllt, und verknüpft nützliche Informationen, beispielsweise zur Herkunft oder Qualität
- > **Ebene:** verknüpft eine Ebene einer (hierarchischen) Knotenmenge mit einem unverwechselbaren Terminus
- > **Gegenüberstellung:** dokumentiert die Zuordnung von Knotenmengen
- > **Grundgesamtheit:** spezifiziert die Menge aller statistischen Einheiten mit übereinstimmenden Identifikationskriterien (sachlich, räumlich und zeitlich)
- > **Instanzmerkmal:** beschreibt Instanzen von Daten, die gesammelt wurden (erst die Kombination von Standardmerkmal, Kontextmerkmal und Instanzmerkmal definiert das konkrete Merkmal für eine Erhebungsdurchführung)
- > **Klassifikationsfamilie:** fasst Klassifikationsreihen mit einem gemeinsamen Terminus (zum Beispiel die Wirtschaftstätigkeit) zusammen
- > **Klassifikationsreihe:** bündelt Klassifikationsversionen
- > **Knoten:** definiert ein einzelnes Element einer Knotenmenge, speziell je nach Einsatzzweck:

- > **Ausprägungselement:** ordnet eine Ausprägung einer Ausprägungsliste zu
- > **Klassifikationselement:** definiert Inhalt und Grenzen der verknüpften Ausprägung für eine Klassifikationsversion
- > **Verschlüsselungszuordnung:** verknüpft eine Ausprägung mit Schlüsselwert für eine Schlüssel-liste
- > **Knotenmenge:** bündelt Knoten basierend auf einer Menge von Termini
 - > **Ausprägungsliste:** enthält Knoten ohne Schlüsselwerte
 - > **Klassifikationsversion:** definiert Ausprägungen mit standardisierten Schlüsselwerten unter Bezugnahme auf ein- und ausschließende Charakteristiken
 - > **Schlüsselliste:** verwaltet mögliche Schlüsselwerte für aufgezählte Wertebereiche
- > **Kontextmerkmal:** verwendet ein Merkmal in einem konkreten Kontext, das heißt einer konkreten Erhebung
- > **Konzeptioneller Wertebereich:** definiert den Wertebereich des Merkmals auf Ebene des Standardmerkmals „konzeptionell“ im Sinne eines Wertebereichstyps vor, speziell:
 - > **Aufgezählter konzeptioneller Wertebereich:** konzeptioniert einen Wertebereich auf Basis einer Ausprägungsliste
 - > **Beschriebener konzeptioneller Wertebereich:** konzeptioniert einen Wertebereich durch eine Beschreibung oder Spezifikation des Wertebereichs, zum Beispiel eine Regel, eine Prozedur oder ein Intervall, in dem die Werte liegen
- > **Maßeinheit:** definiert eine Größeneinheit, in der ein Wert gemessen wird
- > **Maßeinheitentyp:** gruppiert alle Maßeinheiten, die ineinander umgerechnet werden können
- > **Merkmalsträgertyp:** gruppiert Statistische Einheiten auf Basis einer einzelnen sachlichen Eigenschaft (ohne zeitliche und geografische Angabe)
- > **Sachgebiet:** bündelt gleichartige Statistiken, die auf den gleichen Terminologien basieren
- > **Schlüsselwert:** kodiert Ausprägungen
- > **Standardmerkmal:** kombiniert die Bedeutung eines Terminus mit einem Merkmalsträgertyp und dem Konzeptionellen Wertebereich, um die zu messende Eigenschaft zu definieren
- > **Statistische Einheit:** legt den konkreten Träger der Informationen für die statistische Untersuchung fest
- > **Stichwort:** definiert die Charakteristiken eines Knotens, jedes Stichwort bezieht sich normalerweise eindeutig auf einen Knoten
- > **Stichwortverzeichnis:** bündelt Stichworte und verknüpft damit Text aus statistischen Datenquellen mit Knotenmengen
- > **Terminologie:** bündelt Termini in Abgrenzung zu anderen Terminologien
- > **Terminus:** definiert eine Begrifflichkeit, die als Basis für viele andere Objekte (Grundgesamtheit, Universum, Ausprägung, Merkmal, Merkmals-trägertyp) dient
- > **Universum:** definiert einen spezialisierten Merkmalsträgertyp ohne zeitliche und geografische Angabe
- > **Wertebereich:** definiert eine abstrakte Oberklasse für mögliche Eigenschaften, die durch die Objekte „Beschriebener Wertebereich“ und „Aufgezählter Wertebereich“ hinreichend konkretisiert werden muss:
 - > **Aufgezählter Wertebereich:** verknüpft eine Knotenmenge, die zulässige Ausprägungen und Schlüsselwerte enthält
 - > **Beschriebener Wertebereich:** konkretisiert die erlaubten Werte durch einen maschinell auswertbaren Ausdruck
 - > **Fehlwertebereich:** verknüpft ein Instanzmerkmal mit einem (aufgezählten oder beschriebenen) Wertebereich, der Aussagen zu Fehlwerten dokumentiert
 - > **Inhaltlicher Wertebereich:** verknüpft ein Kontextmerkmal mit einem (aufgezählten oder beschriebenen) Wertebereich, der die spezifischen gültigen Werte dokumentiert

- > **Zuordnung:** ordnet (im Rahmen einer Gegenüberstellung) Knoten einer Quellknotenmenge und einer Zielknotenmenge einander zu

3.4 Exchange

Die Gruppe „Exchange“ dient zum Katalogisieren der Informationen, die innerhalb einer statistischen Organisation und zwischen statistischen Organisationen über Austauschinstrumente ausgetauscht werden. Sie umfasst Informationsobjekte, die die Sammlung und Verbreitung von Informationen beschreiben:

- > **Ausgabespezifikation:** spezifiziert Produkte und verwendet dazu die Präsentationen, die sie enthalten
- > **Aussage:** enthält Hinweise in Formularen für den Interviewer oder Befragten
- > **Austauschinstrument:** spezifiziert Werkzeuge zum Empfangen oder Senden von Informationen:
 - > **Formular:** Instrument für Datenerhebungen (zum Beispiel persönlich, computergestütztes Interview, Online-Fragebogen)
 - > **Harvesting:** Sammeln von Daten zu einem bestimmten Sachthema über netzbasierte Dienste (zum Beispiel Verwaltungsdaten über Webservices auslesen oder Preise von Webseiten extrahieren)
 - > **Verbreitungsinstrument:** Werkzeug, um Daten oder Referenzmetadaten zur weiteren Nutzung anzubieten
- > **Austauschspezifikation:** legt das Design der Austauschinstrumente fest (unter anderem Informationsstruktur), speziell:
 - > **Formularspezifikation:** enthält das vollständige Formulardesign, mit Beziehung zur Formularkomponente der obersten Ebene
 - > **Formularkomponente:** definiert die Struktur der Formularspezifikation als eine Kombination aus Fragen, Formularblöcken und Aussagen, speziell:
 - > **Formularblock:** ermöglicht Wiederverwendung in mehreren Formularen
 - > **Formularlogik:** steuert Darstellung der Fragen, Aussagen, Formularkomponenten und so weiter auf Basis hinterlegter Regeln
 - > **Frage:** spezifiziert ein Antwortfeld als Teil eines Formulars, üblicherweise in Verbindung mit einem Merkmal und dessen Wertebereich
 - > **Informationsanbieter:** definiert einen Lieferanten für Informationen mit einer bestimmten Informationsstruktur
 - > **Informationsverbraucher:** definiert einen Konsumenten für Informationen mit einer bestimmten Informationsstruktur
 - > **Instanzaussage:** spezifiziert eine Aussage in einer konkreten Formularkomponente
 - > **Instanzformularblock:** spezifiziert einen Formularblock in einer bestimmten Formularkomponente
 - > **Instanzfrage:** definiert die Verwendung einer Frage in einer bestimmten Formularkomponente
 - > **Liefervereinbarung:** regelt die Nutzung von Austauschinstrumenten zwischen Informationsanbietern und -verbrauchern bezüglich Informationsstruktur (und damit die Inhalte) der auszutauschenden Informationen
 - > **Präsentation:** definiert die Darstellung von Daten und Referenzmetadaten aus Informationssätzen, zum Beispiel Tabellen, Grafiken, strukturierte Datendateien
 - > **Produkt:** verpackt Informationssätze für einen Informationsverbraucher
 - > **Register:** speichert Einzeldaten zum regelmäßigen Abruf

3.5 Structure

Die Gruppe „Structure“ wird verwendet, um Informationen im gesamten statistischen Geschäftsprozess zu strukturieren:

- > **Datenpunkt:** definiert die Zelle als Platzhalter für einen Wert, den die Zelle enthalten kann
- > **Datenstrukturkomponente:** definiert einen Bestandteil einer Datenstruktur, speziell:

- > **Identifikatorkomponente:** identifiziert die statistische Einheit
- > **Messkomponente:** dient einem beobachteten/abgeleiteten Wert
- > **Attributkomponente:** hält eine andere Information als zur Identifikation oder Messung bereit
- > **Datentupel:** enthält bestimmte Eigenschaften zu einer Statistischen Einheit in einer festen Reihenfolge (bezogen auf eine Tabelle entspricht diese einer Tabellenzeile mit einer bestimmten Satzart)
- > **Informationsressource:** bündelt Informationssätze, speziell:
 - > **Datenressource:** bündelt Datensätze
 - > **Referenzmetadatenressource:** bündelt Referenzmetadaten
- > **Informationssatz:** enthält Informationen, die während der Statistikproduktion erhoben, verarbeitet, analysiert und verbreitet werden, speziell:
 - > **Datensatz:** enthält Daten (unter anderem aus Register, Zeitreihen, Längsschnittdaten, Umfragedaten, Ereignisverlaufsdaten, Datentabellen, Datenwürfel, Data Warehouses und so weiter)
 - > **Referenzmetadatenatz:** enthält Referenzmetadaten mit einem bestimmten Referenzmetadatenbezug
- > **Informationsstruktur:** beschreibt das Aussehen einer Informationsstruktur, speziell:
 - > **Datenstruktur:** beschreibt den Aufbau eines Datensatzes durch Kombination entsprechender Datenstrukturkomponenten
 - > **Makrodatenstruktur:** beschreibt speziell die Struktur eines Makrodatensatzes
 - > **Mikrodatenstruktur:** beschreibt speziell die Struktur eines Mikrodatensatzes
 - > **Referenzmetadatenstruktur:** beschreibt den Aufbau eines Referenzmetadatenatzes durch Kombination entsprechender Referenzmetadatenattribute
- > **Referenzmetadatenattribut:** definiert den Bestandteil einer Referenzmetadatenstruktur
- > **Referenzmetadatenbezug:** identifiziert das Thema, für das Referenzmetadaten mit einer bestimmten Referenzmetadatenstruktur grundsätzlich gemeldet werden können (zum Beispiel Datensätze oder Statistiken)
- > **Referenzmetadatengegenstand:** konkretisiert ein bestimmtes Objekt, für den ein Referenzmetadatenatz angelegt wird (zum Beispiel ein bestimmter Datensatz oder eine bestimmte Statistik)
- > **Referenzmetadateninhalt:** enthält den eigentlichen Inhalt für eine bestimmte Eigenschaft eines Referenzmetadaten-themas
- > **Tupelbeziehung:** dokumentiert Abhängigkeiten zwischen Tupelstrukturen
- > **Tupelstruktur:** beschreibt den Aufbau eines Datentupels und kann darüber zusätzliche Beziehungen definieren (zum Beispiel eine Person ist Teil eines Haushalts)

4

Anwendung

MedaS ist die Basis für die Architektur des Metadatenmanagements im Statistischen Verbund. Mit dieser ist es möglich, bestehende metadatenhaltende Systeme und Werkzeuge mithilfe eines zentralen Metadatenmanagementsystems miteinander zu koppeln. Die technische Anbindung der externen Systeme, die Übernahme ausgewählter Inhalte und die gegebenenfalls notwendigen Objekttransformationen übernimmt das Metadatenmanagementsystem. Dieses verbindet somit alle Systeme mit einer zentralen Metadaten-speicherung und stellt über standardisierte Schnittstellen übergreifend notwendige Metadatenobjekte im standardisierten Informationsmodell MedaS für eine konsistente Nachnutzung durch Standardwerkzeuge (zum Beispiel das Statistikportal) und Fachanwendungen (zum Beispiel zur Übertragung von Konfigurationsparametern und Qualitätsinformationen) zur Verfügung. Entsteht neuer Bedarf für die Verwaltung weiterer Metadaten, ermöglicht das Metadatenmanagementsystem die kostengünstige Ergänzung schlanker Anwendungs-module ohne eigene Datenbank, die zur persistenten Speicherung

von Inhalten die bereits existierende Datenbank des Metadatenmanagementsystems nutzen.

Gegenüber dieser statischen Perspektive aus Sicht der Systemarchitektur offenbart ein Blick auf die Geschäftsprozesse der Statistikproduktion eine eher dynamische Sichtweise: Entlang der Prozesskette zur Statistikproduktion werden in bestimmten Teilprozessen (gemäß GMAS) Metadatenobjekte (gemäß MedaS) produziert und in anderen Teilprozessen als Eingabeobjekte benötigt. [↗ Grafik 4](#)

Wichtige Aspekte sind hierbei die konsistente Wiederverwendung bereits vorhandener Metadaten, die Reduzierung manueller Aufwände und die lückenlose Dokumentation der Produktionsprozesse, um die Qualität der Produkte und Prozesse jederzeit bewerten und stetig verbessern zu können (Europäische Kommission, 2019, hier: Seite 19, 28 f., 32, 35, 38, 40, 44, 54 ff.; Statistisches Bundesamt, 2017, hier: Seite 11, 15 f., 18 f.; Statistisches Bundesamt, 2020 hier: Seite 22, 24 f., 27). Hinsichtlich des Ziels einer digitalen, standardisierten und weitgehend automatisierten Statistikproduktion sind die oben aufgeführten MedaS-Objekte unterschiedlich relevant. Der Ausbau und die weitere IT-Unterstützung durch das Metadatenmanagementsystem erfolgen daher sukzessive nach Bedarf unter Berücksichtigung vorhandener Ressourcen, die in den weiteren Ausbau der Infrastruktur des Metadatenmanagements investiert werden können. MedaS dient dabei als Orientierung für das weitere Vorgehen, um perspektivisch eine vollständig digitale und medienbruchfreie Verwaltung aller Metadatenobjekte nutzen zu können. Durch die Kompatibilität mit dem internationalen Standard GSIM können auch Austausch und Nutzung technischer Komponenten anderer statistischer Organisationen gelingen.

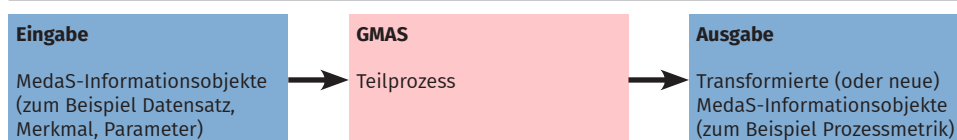
Bei der Modernisierung des Werkzeugportfolios für die Statistikproduktion wird bevorzugt auf Fertigprodukte gesetzt, die in die bestehende System-

landschaft zu integrieren sind. Zwar wird der Austausch von Inhalten über technische Schnittstellen der Komponenten im Prozess grundsätzlich unterstützt, bei voneinander abweichenden Informationsmodellen können aber größere Aufwände für Anbindung und Transformationen entstehen. Muss jedes metadatenhaltende System seine Inhalte mit allen anderen Systemen abgleichen, um einen konsistenten Stand der Informationsobjekte entlang der Produktionskette zu gewährleisten, kann die Integration zu einem exponentiell steigenden Programmieraufwand für entsprechend leistungsfähige Adapter führen. Durch den Einsatz einer zentralen Komponente wie dem Metadatenmanagementsystem, das einen abgestimmten Standard wie MedaS implementiert, reduziert sich der Aufwand spürbar. Für jedes neue System ist dann lediglich ein Adapter zur Anbindung der Webservices und Transformation auszutauschender Informationsobjekte notwendig.

Schließlich übernimmt MedaS nicht nur eine wichtige Rolle als technischer Standard, sondern definiert darüber hinaus auch eine gemeinsame Fachsprache, um ein einheitliches Verständnis für Metadaten zu erreichen. Dies fördert sowohl die organisationsinterne Kommunikation als auch den Austausch auf nationaler und internationaler Ebene über Organisationsgrenzen hinweg. Das spiegelt sich insbesondere in der Spezifikation von Geschäftsprozessen (zum Beispiel im GSBPM beziehungsweise GMAS) oder auch der methodischen Abstimmung einer Statistikproduktion wider.

Grafik 4

Definition von Informationsobjekten entlang der Prozesskette zur Statistikproduktion



5

Fazit

International und national sind standardisierte Geschäftsprozessmodelle und Informationsmodelle wichtige Grundlagen für die Modernisierung der amtlichen Statistik. MedaS und GMAS sind komplementäre Modelle, um statistische Informationen zu erstellen und zu verwalten. GMAS bietet Bausteine und eine gemeinsame Kommunikationssprache für die Modellierung des statistischen Geschäftsprozesses und identifiziert die Aktivitäten von Statistik erstellenden Organisationen, die zu Informationsausgaben führen. Diese Aktivitäten sind in Teilprozesse wie „Produkte konzipieren“ und „Aggregate berechnen“ unterteilt. MedaS hilft bei der Beschreibung von GMAS-Teilprozessen, indem es die Informationsobjekte im Prozessfluss von Statistiken definiert.

Ein gutes Metadatenmanagement ist für die effiziente Implementierung statistischer Geschäftsprozesse von wesentlicher Bedeutung. Metadaten sind in jeder Phase von GMAS vorhanden, entweder erstellt, aktualisiert oder unverändert aus einer vorherigen Phase übernommen. Im GMAS liegt der Schwerpunkt des übergreifenden Prozesses des Metadatenmanagements auf dem Erstellen, Aktualisieren, Verwenden und Wiederverwenden von Metadaten. Metadatenmanagement-Strategien und -Systeme sind daher für den Betrieb von GMAS von entscheidender Bedeutung und werden durch MedaS erleichtert.

Eine Referenzimplementierung von MedaS findet sich im Metadatenmanagementsystem des Statistischen Bundes. Dieses dient der zentralen Ablage von standardisierten Metadaten entlang der Prozesskette der Statistikproduktion, um deren Wiederverwendbarkeit zu erleichtern und Inkonsistenzen zu vermeiden. Hierzu stellt das Metadatenmanagementsystem eine Reihe von Schnittstellen zur Verfügung, über die die Standardwerkzeuge und Fachanwendungen auf Metadaten zugreifen können. Zur initialen Eingabe und Verwaltung von Metadaten können über browserbasierte Benutzeroberflächen diverse Anwendungsmodule zum Beispiel zur Dokumentation der Produkt- und Prozessqualität genutzt

werden. So ist es möglich, die Implementierungslücken ergänzend zu den existierenden metadatenhaltenden Produktionswerkzeugen sukzessive zu schließen.

Das Metadatenmodell der amtlichen Statistik, MedaS, ist die nationale Adaption des internationalen Standards GSIM, der die konzeptionelle Grundlage für die Modernisierung der amtlichen Statistik weltweit bildet. In seine Entwicklung sind im Zeitverlauf Wissen und Erfahrung aus verschiedenen statistischen Organisationen eingeflossen. Da immer mehr statistische Organisationen GSIM als Teil ihres Modernisierungsprogramms verwenden oder dies planen, wächst die Benutzerbasis des Modells stetig. Der Austausch von Wissen und gewonnenen Erkenntnissen kommt neuen und erfahrenen Nutzenden gleichermaßen zugute und das Feedback aus diversen Anwendungen ist für die Überarbeitung des Modells von wesentlicher Bedeutung. Somit kann MedaS auch künftig von der Weiterentwicklung von GSIM profitieren. 

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.

Europäische Kommission. *Qualitätssicherungsrahmen des Europäischen Statistischen Systems*. Version 2.0. 2019. [Zugriff am 13. Januar 2026]. Verfügbar unter: ec.europa.eu

Gehle, Christian/Lüüs, Hans-Peter. [*Prozessmanagement im Statistischen Bundesamt*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 5/2017, Seite 46 ff.

Lindenstruth, Thomas/Claußen, Sven. [*Metadatenmanagement als neue Integrationsarchitektur*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 5/2017, Seite 76 ff.

Statistisches Bundesamt. *Die Digitalstrategie 2025 des Statistischen Bundesamtes*. Wiesbaden 2022. (Internes Dokument)

Statistisches Bundesamt. *Das System der amtlichen Statistik: Organisation und Zusammenarbeit im nationalen, europäischen und internationalen Kontext*. Wiesbaden 2020. [Zugriff am 13. Januar 2026]. Verfügbar unter: www.statistischebibliothek.de

Statistisches Bundesamt. *Verhaltenskodex für europäische Statistiken*. Version November 2017. [Zugriff am 13. Januar 2026]. Verfügbar unter: www.destatis.de

UNECE. *Generic Statistical Business Process Model*. Version 5.2. 2025. [Zugriff am 13. Januar 2026]. Verfügbar unter: statswiki.unece.org

UNECE. *Generic Statistical Information Model*. Version 2.0. 2025. [Zugriff am 13. Januar 2026]. Verfügbar unter: statswiki.unece.org

Herausgeber

Statistisches Bundesamt (Destatis), Wiesbaden

Schriftleitung

Dr. Daniel Vorgrimler

Redaktion: Ellen Römer

Ihr Kontakt zu uns

www.destatis.de/kontakt

Erscheinungsfolge

zweimonatlich, erschienen im Februar 2026

Ältere Ausgaben finden Sie unter www.destatis.de sowie in der [Statistischen Bibliothek](#).

Artikelnummer: 1010200-26004-4, ISSN 1619-2907

© Statistisches Bundesamt (Destatis), 2026

Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.