



AUSSDA
AUSTRIAN
SOCIAL SCIENCE
DATA ARCHIVE

Anonymisierung und Pseudonymisierung in den Sozialwissenschaften –

Praktische Strategien für verantwortungsvolles
Forschungsdatenmanagement (am Beispiel von AUSSDA)

21.01.2026

Mag. Dr. Otto Bodi-Fernandez

AUSSDA/Center for Social Research (CSR),
Institut für Soziologie
Universität Graz

Have data? Need data? | www.ausdda.at

Überblick

- * AUSSDA – The Austrian Social Science Data Archive
- * Sozialwissenschaftliche Forschungsdaten
- * Personenbezogene Daten und Datenschutz
- * Data Sharing und Datenschutz
- * Anonymisierung
- * Alternativen zur Anonymisierung (Einwilligung, Zugriffskontrolle)
- * Praxisbeispiel: Sozialer Survey Österreich (SSÖ) 2026



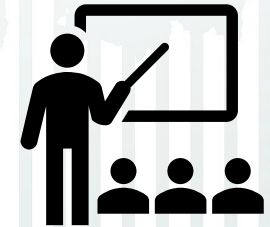
Was ist AUSSDA?

- * AUSSDA – The Austrian Social Science Data Archive
- * Nationale und zertifizierte **Infrastruktur (Core Facility)** für sozialwissenschaftliche **Forschungsdaten** in Österreich
- * Gegründet 2017
- * Seit 2020 ist AUSSDA mit dem **Core Trust Seal** als vertrauenswürdiges Repository **zertifiziert**
- * Nationaler Service Provider im **Consortium of European Social Science Data Archives** (CESSDA ERIC)



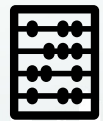
Was macht AUSSDA?

- * **Archivierung** und **Bereitstellung sozialwissenschaftlicher Forschungsdaten** für zur Nachnutzung gemäß den FAIR-Data-Prinzipien
- * **Schulung** und **Beratung** in Fragen des Forschungsdatenmanagements
 - * DMP
 - * Datenschutz
 - * Lizenzierung, etc.
- * Wir **unterstützen** Open Science, Open Access und Open Data wo immer möglich
- * Wir nehmen **Datenschutz** ernst!
- * Mission Statement:
„Wir machen sozialwissenschaftliche Daten zugänglich und nutzbar“



Sozialwissenschaftliche Forschungsdaten

- Meist **Informationen über menschliche Individuen**
- Quantitative Daten
 - Umfragen
 - Beobachtungen
- Qualitative Daten
 - Interviews
 - Gruppendiskussionen
- Sozialwissenschaftliche Daten sind häufig **personenbezogene** und **sensible Daten** → Die DSGVO gilt



Personenbezogene Daten



- **Personenbezogene Daten** → Die DSGVO gilt
- **Personenbezogene Daten:**
 - Jede Information über eine **identifizierte** oder **identifizierbare** natürliche Person ("betroffene Person")
 - Eine identifizierbare natürliche Person ist eine Person, die **direkt** oder **indirekt** identifiziert werden kann.
 - **Anonyme** oder **anonymisierte** Daten gelten nicht als personenbezogene Daten → unterliegen **nicht** mehr der **DSGVO**.
 - **Pseudonymisierte** Daten: Die Identifizierung einer Person ist nur durch Hinzufügen zusätzlicher, getrennt gespeicherter Informationen möglich. (→ weiterhin personenbezogene Daten)

Identifikatoren

- * Direkte Identifikatoren:

- * zB. Name, Email-Adresse, Telefonnummer, Sozialversicherungsnummer, Matrikelnummer, IP-Adresse

- * Indirekte Identifikatoren:

- * zB. Soziodemographische Informationen wie Geschlecht, Alter, Beruf, Wohnort
 - * Beispiel für indirekte Identifizierung durch Kombination von Merkmalen:
Interaktiver Anonymitäts-Test (Imperial College London): <https://www.ooa.world/>

Was sind pseudonymisierte Daten?

- * Direkte Identifikatoren (zB Namen) werden durch einen **Code ersetzt**
- * **Zuordnung** der Daten zu einer konkreten Person ist nur über diesen Code möglich
- * Zuordnungstabelle mit Namen und Code wird **getrennt** und zugriffsgeschützt aufbewahrt.
- * **Re-Identifizierung** ist über Daten und Zuordnungstabelle prinzipiell möglich → Daten gelten weiterhin als **personenbezogen** → **DSGVO gilt!**
- * Anwendung: zB bei Längsschnittstudien

Was sind anonyme Daten?

- * DSGVO enthält keine Begriffsbestimmung
- * Daten ohne Personenbezug
 - * Daten die **a priori** keinen Personenbezug enthalten (Wetterdaten, Temperatur,...)
 - * Daten bei denen Personenbezüge **entfernt** wurden (→ Anonymisierung)
- * DSGVO kommt nicht zur Anwendung
- * Entscheidung der österr. Datenschutzbehörde:
 - * Von anonymen Daten kann ausgegangen werden, wenn die **Gruppe**, auf die sich **identifizierte Angaben** beziehen, **mehr als 5 Personen** umfasst

“Sensible Daten” (Art. 9 DSGVO)

- * „**Besondere Kategorien** personenbezogener Daten“ → Daten, die besonders sensibel sind und daher einen stärkeren Schutz erfordern.
 - * rassische und ethnische Herkunft
 - * politische Meinungen
 - * religiöse oder weltanschauliche Überzeugungen
 - * Gewerkschaftszugehörigkeit
 - * genetischen Daten
 - * biometrischen Daten (zur eindeutigen Identifizierung)
 - * Gesundheitsdaten
 - * Daten zu Sexualleben oder der sexuellen Orientierung

Datenschutz

- Datenschutzrechtliche Vorgaben sind zu beachten, wenn **personenbezogene** Daten **lebender natürlicher** Personen im Spiel sind.
- Rechtlicher Rahmen
 - Datenschutzgrundverordnung (DSGVO)
 - Datenschutzgesetz (DSG)
 - Forschungsorganisationsgesetz (FOG)
- „Verbot mit Erlaubnisvorbehalt“: Verarbeitung personenbezogener Daten ist grundsätzlich verboten und nur unter bestimmten Bedingungen erlaubt:
 - wenn dies gesetzlich oder durch andere Rechtsvorschriften erlaubt oder vorgeschrieben ist
 - oder die jeweils betroffene Person, um deren Daten es geht, hierin eingewilligt hat.

Datenschutz

- Geeignete Rechtfertigungsgründe für wissenschaftliche Forschung (u. A.):
 - Einwilligung (bei „sensiblen Daten“ ausdrückliche Einwilligung)
 - Wissenschaftliche Forschung im öffentlichen Interesse

 - Art. 89 DSGVO: **Garantien und Ausnahmen** in Bezug auf die Verarbeitung zu
 - *im öffentlichen Interesse liegenden Archivzwecken,*
 - *wissenschaftlichen oder historischen Forschungszwecken*
 - *statistischen Zwecken*

 - Geeignete Garantien zur Datenminimierung:
 Pseudonymisierung/Anonymisierung , Nutzungsverträge, Zugangsbeschränkungen, etc.
- Das Datenschutzrecht erlaubt die Verarbeitung personenbezogener Daten zu Forschungszwecken auch ohne Einwilligung, sofern Schutzmaßnahmen zur Datenminimierung getroffen werden.

Datenschutz und Ethik

- * Datenschutz ist nicht nur eine **rechtliche Pflicht**, sondern ein zentrales **ethisches Prinzip** wissenschaftlicher Forschung.
- * **Ethische Prinzipien**
(Ethikkodex der ÖGS - Österreichischen Gesellschaft für Soziologie)
 - * Vermeidung von Schaden
 - * Informierte Einwilligung und Freiwilligkeit
 - * Vertraulichkeit
 - * Archivieren und Teilen von Forschungsdaten



AUSSDA
AUSTRIAN
SOCIAL SCIENCE
DATA ARCHIVE

Data Sharing und Datenschutz

Besondere Herausforderungen in den Sozialwissenschaften

Have data? Need data? | www.ausoda.at

Data Sharing und Datenschutz

- **Forschungsfördereinrichtungen und Journals**
→ Forderung: Daten möglichst offen zugänglich machen (Open Data)
- **Forschungsethik und Datenschutzrecht**
→ Forderung: Daten möglichst schützen

Wie können diese Anforderungen vereinbart werden?



Source: Stock Adobe

Warum Anonymisierung?

- Einhaltung datenschutzrechtlicher Vorgaben und forschungsethischer Prinzipien
- Einhaltung von Förderrichtlinien und guter wissenschaftlicher Praxis
- Schafft die Grundlage für rechtssicheres und verantwortungsvolles Archivieren und Teilen von Forschungsdaten
- Herausforderung
 - Verlust von Information
 - Aufwand

Wann anonymisieren?

- * Planung der Anonymisierung von Beginn an
- * Schritte in jeder Phase des Datenlebenszyklus
 - * Identifikatoren bei Erhebung vermeiden, wenn nicht benötigt
 - * Löschen, sobald nicht mehr nötig
 - * Ggf. Einwilligung zur Erhebung direkter Identifikatoren



Anonymisierung

- * Anonymisierte Daten unterliegen nicht der DSGVO.
 - * Einwilligung oder Rechtsgrundlage für die Datenverarbeitung nicht erforderlich.
 - * Keine rechtliche Einschränkung hinsichtlich Nachnutzungszwecke (Open Data).
 - * Achtung: Ethische Einschränkungen sind weiterhin möglich (!)
- * Grenzen der Anonymisierung
 - Verarbeite ich personenbezogene oder anonymisierte Daten? → oft unklar!
 - Umfragedaten können indirekte Identifikatoren enthalten
 - Absolute Anonymisierung ist oft nicht möglich
 - Wie viel Informationsverlust ist akzeptabel?
 - Konzept der **faktischen Anonymität**

Anonymisierung

- * **Absolute Anonymisierung** → Identifizierung ist ausgeschlossen
- * **Faktische Anonymisierung** → Identifizierung wäre nur mit unverhältnismäßigem Aufwand möglich (→ Erwägungsgrund 26 DSGVO)
- * **Pseudonymisierung** → Daten werden so bearbeitet, dass ohne Hinzunahme zusätzlicher Informationen keine Identifizierung möglich ist.
- * Pseudonymisierte Daten sind als *personenbezogene Daten* anzusehen.

Zu berücksichtigen:

- Risiko/Wahrscheinlichkeit der Identifikation?
- Welcher Schaden ist zu erwarten?
- Wie viel Informationsverlust ist akzeptabel?

direkt identifizierbar → indirekt identifizierbar → pseudonymisiert → faktisch anonym → absolut Anonym



Source: Stock Adobe

Anonymisierung: Wie viel Informationsverlust ist akzeptabel?

Wie anonymisieren?

- * Beste Lösung: Keine identifizierenden Informationen erheben
- * Zweitbeste Lösung: Informationen so bearbeiten, dass Identifizierbarkeit ausgeschlossen wird → Anonymisierung
 - * Direkte Identifikatoren entfernen z. B. Name, Telefonnummer, Sozialversicherungsnummer
 - * Indirekte Identifikatoren entfernen, ersetzen oder aggregieren z. B. Kombinationen aus Beruf, Einkommen, Alter, Geschlecht und Wohnort
- * **Wenn personenbezogene Daten nicht benötigt werden → Löschen!**
Wenn sie behalten werden → braucht es einen Grund!

Anonymisierungsmethoden (quantitative Daten)

- * **Variablen entfernen**
 - * *zB. Namen, Adressen, sonstige Identifikatoren*
- * **Reduzieren der Genauigkeit (Kategorien zusammenfassen)**
 - * *zB deataillierte geographische Angaben, detaillierte Berufsangaben*
- * **Bedeutungsinhalte bei Textvariablen verallgemeinern**
 - * *zB offene Antworten kategorisieren*
- * **Beschränken der oberen und unteren Bereiche kontinuierlicher Variablen**
 - * *zB: Alter, Haushaltsgröße, Kinderzahl*

AUSSA Data Deposit Guideline (Beispiele)

Direkte Identifikatoren entfernen:

Name, Email, Matrikelnummer, etc. → Löschen

Räumliche Angaben:

Postleitzahl/Gemeinde → Bundesland

Zeitliche Angaben:

Genaues Geburtsdatum → Geburtsjahr (oder Kategorien)

Berufliche Angaben (zB ISCO-08 Code):

Soziologie (2632) → Sozialwissenschaftler (263)

[Link zu: AUSSDA Data Deposit Guideline](#)

AUSSA Data Deposit Guideline (Beispiele)

Bildungsabschluss, Nationalität, Konfession, Beschäftigungsstatus etc:

kleine Fallzahlen ($n < 20$) → Zusammenfassen mit anderen Kategorien

zB Uni, FH, Päd. Hochschule → Hochschulabschluss

Kontinuierliche Merkmale mit einzelnen Fällen an den Rändern:

Haushaltsgröße: → 1, 2, 3, 4, 5 Personen und mehr

Alter: → in Jahren (bis 79), 80 Jahre und älter

Offene Antworten → Löschen oder Kategorisieren

genau prüfen: Offene Antworten können auf personenbezogene Angaben zu Dritten Personen enthalten!

[Link zu: AUSSDA Data Deposit Guideline](#)

Anonymous or not?

Example A

- 🌐 name: -
- 🌐 age: 45-50
- 🌐 gender: male
- 🌐 occupation: scientist
- 🌐 employer: University of Graz
- 🌐 discipline: sociology

Example B

- 🌐 name: -
- 🌐 age: 45-50
- 🌐 gender: male
- 🌐 occupation: scientist
- 🌐 employer: University of Graz
- 🌐 discipline: social sciences

Qualitative Daten

- * Anonymisierung von Beginn planen (wann und von wem wird anonymisiert)
 - * Interviewtranskripte gründlich lesen
 - * Definieren von Identifikatoren zB.: Name, Firma, Wohnort
 - * Erstellen eines Anonymisierungsplans (bei AUSSDA erforderlich!)
 - * Max Meier → [Name]
 - * McDonalds → [Betrieb]
 - * Sinabelkirchen → [Gemeinde]
-
- * For example, see [Kalova & Füst \(2023\)](#)

Siehe: AUSSDA
Data Deposit Guideline for Qualitative Data

Anonymisierungsmethoden (qualitative Daten)

* Erstellen eines Anonymisierungsplans

Example Identifier	Type of identifier	Change?	Example replacements	Justification
Researcher names	direct	yes	[P1], [P2], etc.	Names are direct identifiers and are therefore removed and replaced with P1, P2, etc.
Nuclear Physics	indirect	no	[Discipline]	We do not categorise or change the discipline given that this is important information for data reuse and will not allow identification of participants together with other available information.
University of Vienna	indirect	yes	[Researcher's University]	The university used in combination with the discipline could lead to identification. We changed the name of the researchers' universities to <i>[Name of university]</i> .
Austria	indirect	no	[Country in Central Europe]	No change since the country will provide important contextual information and cannot be used to identify interviewees.

Siehe: AUSSDA
Data Deposit Guideline for Qualitative Data



AUSSDA
AUSTRIAN
SOCIAL SCIENCE
DATA ARCHIVE

Alternativen zu Anonymisierung

Was, wenn vollständige Anonymisierung nicht möglich ist?

Have data? Need data? | www.ausdda.at

DATA SHARING

Einwilligung

Anonymisierung

Zugriffs-
beschränkung

DATENSCHUTZ

Informierte Einwilligung

- * Informierte Einwilligung ist ein **ethisches Prinzip**.
- * Informierte Einwilligung ist eine (von mehreren möglichen) **Rechtsgrundlagen** zur Verarbeitung personenbezogener Daten.
 - * Anforderungen:
 - * Verständlich
 - * Informiert
 - * Freiwillig
 - * Zweckgebunden
- * → Information sollte auch Angaben über Archivieren und Teilen von Daten enthalten.

Informierte Einwilligung

* Grenzen der Einwilligung

- * Die Einwilligung kann **widerrufen** werden (in diesem Fall müssen die Daten gelöscht werden).
- * Die Einwilligung muss an bestimmte **Zwecke** der Datenverarbeitung gebunden sein. In der Forschung sind zukünftige Zwecke schwer vorhersehbar.
- * Einwilligungserklärungen sind in der Praxis oft **mangelhaft formuliert**.

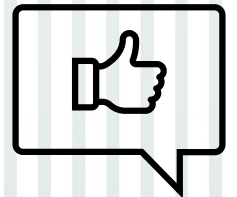
Informierte Einwilligung

Formulierungen **vermeiden**, die Data Sharing verhindern!!!

„Die erhobenen Daten werden ausschließlich im Forschungsprojekt verwendet und nicht an Dritte weitergegeben. Nach Abschluss des Projektes werden die erhobenen Daten gelöscht.“



„Ihre Antworten werden getrennt von Ihren Kontaktdaten (Name, Adresse) gespeichert, sodass keine Rückschlüsse auf Ihre Person gezogen werden können. Die Kontaktdaten werden nach Abschluss der Erhebung gelöscht.“

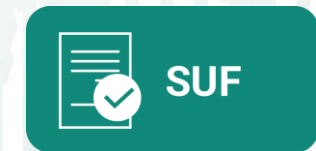


Zugangsbeschränkungen

- Wenn vollständige Anonymisierung nicht möglich ist:
- Zugriff einschränken, um Datenaustausch zu ermöglichen
 - Missbrauchsrisiko wird minimiert – Daten werden geschützt
- Lizenzen
 - Mittels Lizenz wird die Nachnutzung von Daten **auf wissenschaftliche Zwecke beschränkt**.
 - Die DSGVO erlaubt die Nachnutzung von Daten auch ohne Einwilligung zu Forschungszwecken (Forschungsprivileg Art. 89 DSGVO iVm mit nationalem Recht).
 - Es müssen **geeignete Garantien** für die **Datenminimierung** vorhanden sein (z. B. Anonymisierung, Pseudonymisierung, Zugangsbeschränkung).

Lizenzmodelle bei AUSSDA

- 
Open Access Files (OA):
sind als Open Data unter Creative Commons Attribution International 4.0 License (CC BY)
- 
Scientific Use Files (SUF):
Zugang für registrierte User - stehen nur für wissenschaftliche Zwecke zur Verfügung – zB aus Datenschutzgründen



Zugangsklassen für SUF-Files



Restricted **account-based** access:

einloggen → zustimmen → runterladen

zB. Pseudonymisierte Daten, faktisch anonymisierte Daten

Restricted **controlled** access

*einloggen → Zugriff beantragen → Formular ausfüllen (inkl. Projektbeschreibung) → Freigabe durch AUSSDA-Mitarbeiter*in für „sensible“ Daten*

Zusammenfassung

- Daten erheben nötig?
 - Wenn schon vorhanden, dann vermeiden (Sekundäranalyse)
- Personenbezogene Daten nötig?
 - wenn nicht, dann vermeiden!
 - wenn ja, Rechtgrundlage heranziehen (zb Einwilligung, öffentliches Interesse) , ggf. pseudonymisieren, getrennt aufbewahren
- Personenbezogene Daten nicht mehr nötig?
 - Löschen bzw.
 - Anonymisieren
- Wie sollen Daten archiviert werden?
 - Open Access (OA): möglich bei absolut anonymen Daten, ansonsten...
 - SUF – account based: für pseudonymisierte/‘faktisch anonymisierte‘ Daten
 - SUF – controlled access: für „sensible“ Daten mit Restrisiko der Identifizierbarkeit



AUSSDA
AUSTRIAN
SOCIAL SCIENCE
DATA ARCHIVE

Praxisbeispiel

Sozialer Survey Österreich (SSÖ) 2026

Have data? Need data? | www.ausstda.at

Sozialer Survey Österreich (SSÖ) 2026

- Repräsentative Bevölkerungsumfrage (seit 1986)
- Durchführung:
 - Lead: Universität Graz
 - Partner*innen: Universitäten Linz, Salzburg und Wien
- Themen 2026:
 - Einstellungen zur Rolle des Staates (ISSP Modul)
 - Gesellschaftliche Herausforderungen in Österreich
 - Persönliche Lebenssituation
- Zufallsstichprobe (brutto n=3800) aus dem Zentralen Melderegister (ZMR)
 - Genehmigung durch BMI
- Erhebungsmodus: Push-to-Web (Einladung zur Teilnahme an Online-FB)
- Zielpersonen werden postalisch angeschrieben
 - Einladungsschreiben
 - Informationsblatt (Informationen zur Studie und zum Datenschutz)
- Rechtsgrundlage: Wissenschaftliche Forschung im öffentlichen Interesse
- Informationsblatt informiert über
 - Datenverarbeitung in der Studie
 - Archivierung und Bereitstellung der Befragungsdaten

Pseudonymisierung während Feldphase

- Postalische Abwicklung durch **externes Umfrageinstitut**
 - Externes Institut erhält Kontaktdaten (Name, Adresse) ausschließlich zu diesem Zweck
 - **Kein Zugriff auf Befragungsdaten** durch externes Institut
 - Externes Institut **löscht Kontaktdaten** nach Abschluss der Feldphase
- Jede Zielperson erhält per Brief einen individuellen **Zugangscode** zur Online-Befragung
- Fragebogen erhebt **keine direkten Identifikatoren**
- Kontaktdaten bleiben **getrennt** von Befragungsdaten gespeichert
- Zuordnungstabelle liegt **nur an der Uni Graz**
- Zugriff: **nur Projekt- & Studienleitung**
- Zweck der Kontaktdaten: **Erneute Kontaktaufnahme** für Erinnerungs- und Dankeschreiben
- → Daten sind **pseudonymisiert** (volle Datenschutzpflicht)

Nach Abschluss der Feldphase

- Löschung der Zuordnungstabellen
 - Datenbereinigung, Prüfung auf etwaige Identifikatoren
 - Datenanalysen gemäß Forschungszweck
 - Vor Archivierung und Bereitstellung:
 - Aggregation potenziell identifizierender Variablen
 - Prüfung seltener Merkmalskombinationen
 - Entfernung feinräumlicher Regionalangaben
- Minimierung des Re-Identifizierungsrisikos
- Ziel: faktisch anonyme Forschungsdaten mit wissenschaftlichem Informationsgehalt

Archivierung bei AUSSDA

● Scientific Use File (SUF)

- Zugang nur für registrierte Forschende
- Nutzungsvereinbarung & Zweckbindung (wiss. Zwecke)
- Schutz vor Re-Identifikation bei Erhalt analytischer Qualität
- Je höher Risiko, desto strengere Zugriffskontrolle

● Open Access (OA)

- Nur bei vollständiger Anonymisierung möglich
- Reduzierte Version durch Entfernen aller soziodemographischen Angaben
- Frei zugänglich und uneingeschränkt nachnutzbar

Vielen Dank für Ihre Aufmerksamkeit!

Mag. Dr. Otto Bodi-Fernandez

AUSSDA/Center for Social Research (CSR)

Institut für Soziologie, Universität Graz

Mail: otto.bodi@uni-graz.at

AUSSDA Website: <https://www.aussda.at>

Mail: info@aussda.at

Have data? Need data? | www.aussda.at