



Bern, 12. Februar 2025

Bericht über die Anwendung von Datenwissenschaft zur Wahrung der Privatsphäre

Technische Lösungen zur Wahrung der Geheimhaltung für eine gemeinsame Nutzung von Daten in der Bundesverwaltung

Aktenzeichen: 031.21-688/6/7/1



Inhaltsverzeichnis

1	Einleitung	4
2	Kernkonzepte	7
2.1	Grundbegriffe	7
2.2	Input und Output Privacy	9
3	Privacy-Enhancing-Technologies	14
3.1	Trusted Execution Environment	14
3.2	Homomorphe Verschlüsselung	15
3.3	Secure Multiparty Computation	16
3.4	Distributed Learning	17
3.5	Null-Wissen-Beweis	18
3.6	Differential Privacy	19
3.7	Synthetische Daten	21
4	Fallbeispiel 1: Plattform für eine vertrauliche Anwendung der Datenwissenschaft ohne Datenübermittlung	23
4.1	Hintergrund und Ziele	23
4.2	Grundsatz	25
4.3	Grenzen und Herausforderungen	27
5	Fallbeispiel 2: Sichere Datenanreicherung durch vertrauliche Verknüpfung	29
5.1	Hintergrund und Ziele	29
5.2	Grundsatz	31
5.3	Grenzen und Herausforderungen	34
6	Rechtsgrundlagen und Gouvernanz	36
6.1	Bearbeitung identifizierender Daten – Übersicht der Rechtsgrundlagen	36
6.2	Grenzen der bestehenden Rechtsgrundlagen	38
6.2.1	Techniken zur Exfiltration identifizierender Daten: Umgehung der Anonymisierung	38
6.2.2	Auswirkungen auf die Rechtsgrundlagen zur Reglementierung der von der Bundesverwaltung gehaltenen Daten	39
6.3	Die «Five Safes»: Gouvernanz-Rahmen für eine datenschutzwahrende Datenwissenschaft	40
7	Empfehlungen für die Bundesverwaltung	43

Zusammenfassung

Der Bundesrat hat die Erarbeitung eines Konzepts über die Anwendung von Datenwissenschaft zur Wahrung der Privatsphäre in der Bundesverwaltung auf Basis von Privacy-Enhancing-Technologies (PET) in Auftrag gegeben. Mithilfe dieser datenschutzfördernden Technologien kann anhand von Daten der Bundesverwaltung datenschutzkonform Wissen generiert werden. Konkret lassen sich durch den Einsatz datenwissenschaftlicher Methoden Personendaten der Bundesverwaltung nutzen, ohne dass sie sichtbar sind oder übermittelt werden. Das Offenlegungsrisiko wird dabei kontrolliert. Angesichts der steigenden Nachfrage nach Datensätzen und Datenverknüpfungen der Bundesverwaltung prüft das beim BFS angesiedelte Kompetenzzentrum für Datenwissenschaft (DSCC), wie dieser Bedarf mithilfe datenwissenschaftlicher Methoden und unter Einhaltung der höchsten datenschutzrechtlichen Standards einfacher und schneller gestillt werden kann, und vor allem, wie sich der damit verbundene administrative und technische Aufwand verringern lässt.

Dieser Bericht orientiert sich an mehreren internationalen Bemühungen wie dem Erlass des US-Präsidenten (Executive Order) zur sicheren Verwendung von KI und den von der UNO in Zusammenarbeit mit den internationalen Statistiklieferanten durchgeführten Arbeiten zu den PET. Er zeigt anhand von zwei Fallbeispielen, die die gesamte Schweizer Gesellschaft tangieren, wie sich PET auf die gemeinsame Nutzung von Daten der Bundesverwaltung auswirken. Im ersten Fallbeispiel wird die Funktionsweise einer Plattform beschrieben, die eine vertrauenswürdige und sichere Anwendung der Datenwissenschaft ermöglicht, ohne dass die Daten über einen Fernzugriff übermittelt werden müssen. Es handelt sich dabei um eine standardisierte und automatisierte technische Lösung auf Basis des «Privacy by Design»-Prinzips, das die Datenlieferung ersetzt. Das Potenzial der von der Bundesverwaltung gehaltenen Daten kann allerdings nur ausgeschöpft werden, wenn – wie im zweiten Fallbeispiel beschrieben – zur Anreicherung der Informationen mehrere Datenquellen vertraulich zusammengeführt werden. Diese Art der Datenverknüpfung ist moderner und ressourcenschonender als die aktuelle Methode. Mit dem Einsatz von PET wird sichergestellt, dass Personendaten trotz gemeinsamer Nutzung geschützt sind. Die beiden Fallbeispiele illustrieren, wie Daten von verschiedenen Bundesämtern sowie von externen Akteuren wie der wissenschaftlichen Forschung, (zwischen-)staatlichen und nichtstaatlichen Organisationen oder der Privatwirtschaft effizienter und breiter wiederverwendet werden können, dies unter Wahrung der Vertraulichkeit.

Werden die technischen Fortschritte, insbesondere im Bereich der PET, vor dem Hintergrund der aktuellen Rechtsgrundlagen betrachtet, zeigen sich jedoch Grenzen. So kann zum Beispiel technisch nicht garantiert werden, dass sich eine Anonymisierung nicht rückgängig machen lässt. Zudem ist die Wiederverwendung von Daten durch die strikte Umsetzung des Zweckbindungsgrundsatzes in der Regel nur Bundesorganen für nicht personenbezogene Zwecke wie Statistik, Planung und Forschung erlaubt. PET bieten neue Instrumente, mit denen die Privatsphäre in einem breiteren Kontext gewährleistet wird, als dies die aktuellen rechtlichen Grundlagen vorsehen. Für diesen breiteren Kontext muss zunächst ein risikobasierter Gouvernanz-Rahmen geschaffen werden, der die Einhaltung der Rechtsstaatlichkeit garantiert. Er soll sicherstellen, dass die Vorteile einer breiteren Wiederverwendung von Daten nicht zum Nachteil für die Schweizer Bevölkerung, Haushalte und Unternehmen wird, und so die politische und gesellschaftliche Akzeptanz erhöhen.

PET bieten der Bundesverwaltung neue Möglichkeiten, um an Effizienz zu gewinnen und Mehrwert für das Gemeinwohl zu schaffen. Die Bundesverwaltung muss daher dafür sorgen, dass sie über die nötige IT-Infrastruktur und das technische Fachwissen zur Nutzung dieser Technologien verfügt. Mit einem geeigneten Rechts- und Gouvernanz-Rahmen werden PET entscheidend dazu beitragen, dass der Bund seine Aufgaben zugunsten der Schweizer Bevölkerung und unter Einhaltung der Verfassungsgrundlagen mit zunehmender Effizienz erfüllen kann. PET müssen die Bundesverwaltung bei ihren Aufgaben unterstützen, ohne die bestehenden Aufgaben wie jene der Militärverwaltung oder des Nachrichtendienstes des Bundes (NDB) zu beeinträchtigen.

1 Einleitung

Der Bundesrat (BR) hat das Eidgenössische Departement des Innern (EDI) bzw. das Bundesamt für Statistik (BFS) im Rahmen der Umsetzung der Datenwissenschaftsstrategie des Bundes (DSStB) beauftragt, ein Konzept für technische Datenschutzlösungen zur gemeinsamen Datennutzung in der Bundesverwaltung zu erarbeiten und den vorliegenden Bericht über die Anwendung von Datenwissenschaft zur Wahrung der Privatsphäre zu erstellen (Ziffer 6 des Bundesratsbeschlusses vom 2. Dezember 2022; EXE 2022.2746).

Um die Herausforderungen zu bewältigen, denen sich die Schweizer Gesellschaft heute in Bereichen wie Gesundheit, Klima, Mobilität oder Umwelt gegenüberstellt, und um fundierte Entscheidungen zu treffen, greift der Bund bei der Politikgestaltung zunehmend auf Daten zurück. Damit diese Praxis systematisiert werden kann, muss das Potenzial der von den verschiedenen Verwaltungseinheiten erhobenen Daten voll ausgeschöpft werden. Aus den vorhandenen Daten kann dank ihrer Qualität, Vollständigkeit und vor allem Zuverlässigkeit ein grosser Mehrwert für das Gemeinwohl geschaffen werden. Der Nutzung sind aktuell jedoch Grenzen gesetzt. Die bereichsübergreifende bundesinterne und -externe Zusammenarbeit muss erleichtert werden. Dabei sind berechtigten Bedenken von Privatpersonen, Unternehmen und Haushalten in der Schweiz in Bezug auf den Schutz ihrer Daten ernst zu nehmen. Die Bundesverwaltung muss daher eine datenwissenschaftliche Praxis entwickeln, die bei der Nutzung der von den Bürgerinnen und Bürgern anvertrauten Daten die Wahrung der Privatsphäre und die Rechtsstaatlichkeit gewährleistet.

Sowohl die Schweizer Wirtschaft als auch die Gesellschaft würden von Rahmenbedingungen profitieren, mit denen das Potenzial der Verwaltungsdaten sowie das der Methoden aus dem Bereich der künstlichen Intelligenz (KI) voll ausgeschöpft werden kann. Die Schaffung solcher Rahmenbedingung entspricht dem Willen des Bundesrats, vertrauenswürdige Datenräume und die digitale Selbstbestimmung zu fördern, denn «die Verknüpfung von Daten bietet ein grosses Potenzial für Gesellschaft, Wirtschaft, Wissenschaft und Verwaltungen»¹. Die Motion 22.3890 der WBK-S vom 22.8.2022, in der die Schaffung eines Rahmengesetzes für die Wiederverwendung von Daten² gefordert wird, weist darauf hin, dass Lösungen gefunden werden müssen, mit denen sämtliche Daten (und nicht nur die offenen Daten) ausgewertet werden können.

Auf internationaler Ebene betont das von der Biden-Regierung erlassene US-Präsidialdekret «Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence»³ die Notwendigkeit von Rahmenbedingungen, mit denen sich die Risiken einer Offenlegung von Personendaten amerikanischer Bürgerinnen und Bürger durch KI-Systeme kontrollieren lassen. In Europa weist das «Input Privacy Preservation Techniques Project»⁴ der Wirtschaftskommission der Vereinten Nationen für Europa auf die Relevanz der Privacy-Enhancing-Technologies (PET) für statistische Organisationen

¹ <https://www.news.admin.ch/de/nsb?id=87780>

² <https://www.parlament.ch/de/ratsbetrieb/suche-curia-vista/geschaefte?AffairId=20223890>

³ <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/>

⁴ <https://unece.org/sites/default/files/2021-11/November%20workshop%202021%20HLG-MOS-PPT%20%281%29.pdf>

hin. Und das «Office for National Statistics» des Vereinigten Königreichs prüft die Möglichkeit, diese Technologien zu nutzen, um Sterblichkeitsdaten zu schützen.⁵

Der im Rahmen der DSStB erstellte Bericht «Konkrete Anwendungsfälle von Datenwissenschaft zum Gemeinwohl entlang des Prozesses der Politikgestaltung»⁶ hat aufgezeigt, wie komplex die volle Ausschöpfung des Datenpotenzials ist, und dass sie die Einbindung vieler verschiedener Stakeholder und den Aufbau zahlreicher (organisatorischer, ethischer, rechtlicher und technischer) Kompetenzen erfordert. Eine der zentralen Schwierigkeiten bei der Anwendung der Datenwissenschaft ist die Einführung eines Datengouvernanz-Rahmens, der den gesamten Lebenszyklus der Daten abdeckt (Generierung, Erhebung, Speicherung, Sicherheit, Zugang, Löschen usw.) und auf drei Kernelementen beruht:

- 1) IT-Infrastruktur: Sie wird unter anderem im Rahmen der Entwicklung einer kollaborativen Plattform für Datenwissenschaft adressiert (Ziffer 4 des Bundesratsbeschlusses vom 2. Dezember 2022; EXE 2022.2746) und ist Gegenstand eines separaten Berichts.
- 2) Verwaltung von Datensätzen: Mit dem Programm Nationale Datenbewirtschaftung (NaDB; EXE 2023.1946)⁷ fördert der Bundesrat die langfristige Mehrfachnutzung von Daten nach dem Once-Only-Prinzip. Zusätzlich laufen Projekte zur Harmonisierung und Standardisierung von Metadaten wie die vom EDI (BFS) eingerichtete I14Y-Interoperabilitätsplattform⁸ und die vom EDI (BAR) angebotenen Linked Data Services LINDAS⁹. Da diese Projekte alle von der öffentlichen Verwaltung erhobenen Daten erfassen und ihre Interoperabilität sicherstellen, bilden sie die Grundlage dafür, dass das Potenzial der bei der öffentlichen Verwaltung vorhandenen Daten voll ausgeschöpft werden kann. Metadaten grenzen nicht nur den Datenbestand der Bundesverwaltung präzise ab, sie sind auch für die Anwendung einer datenschutzwahrenden Datenwissenschaft von entscheidender Bedeutung.
- 3) Datenschutzmanagement: Um die Privatsphäre der Bürgerinnen und Bürger ihren Erwartungen entsprechend zu schützen, wird die Nutzung von Daten der Bundesverwaltung heute entweder strikt auf ihre Wiederverwendung zu Erhebungszwecken beschränkt oder es besteht eine Anonymisierungspflicht, insbesondere bei der Veröffentlichung als «Open Government Data» (OGD). Angesichts der wissenschaftlichen Fortschritte der letzten zwanzig Jahre, insbesondere solcher, die bei Attacken genutzt werden, um persönliche Informationen aus statistischen Veröffentlichungen oder KI-Modellen zu extrahieren, erscheint die Anonymisierung technisch unrealistisch. Um das Vertrauen der Bürgerinnen und Bürger in den Staat zu wahren und die Wiederverwendung der von der Bundesverwaltung gehaltenen Daten zu ermöglichen, muss der Rahmen für die Datenschutz-Gouvernanz erneuert werden. Hierzu braucht es neue Rechtsinstrumente, die auf Technologien zur Stärkung der Geheimhaltung, sogenannten Privacy-Enhancing-Technologies (PET), basieren und gewährleisten, dass die «Five Saves»¹⁰, d. h. «Safe People», «Safe Project», «Safe Setting», «Safe Data» und «Safe Output» bei der Wiederverwendung von Daten der Bundesverwaltung eingehalten werden.

⁵ <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/methodologies/applyingdifferentialprivacyprotection-toonsmortalitydatapilotstudy>

⁶ [Konkrete Anwendungsfälle von Datenwissenschaft zum Gemeinwohl entlang des Prozesses der Politikgestaltung | Publikation](#)

⁷ <https://www.bfs.admin.ch/bfs/de/home/nadb/nadb.html>

⁸ <https://www.i14y.admin.ch/>

⁹ <https://lindas.admin.ch/>

¹⁰ <https://ukdataservice.ac.uk/help/secure-lab/what-is-the-five-saves-framework/>

Damit das Potenzial der von der Bundesverwaltung gehaltenen Daten zum Wohl der gesamten Schweizer Bevölkerung voll ausgeschöpft werden kann, müssen die verschiedenen Akteure im IT- und im Gesetzgebungsbereich folglich zusammenarbeiten. Im vorliegenden Bericht werden die technischen Lösungen zur Implementierung eines Rahmens für Datenschutz-Gouvernanz beschrieben. Der Bericht enthält Beispiele für die praktische Anwendung der Lösungen, mit denen das Datenschutzmanagement im Rahmen der Datengouvernanz umgesetzt werden kann. Zudem erläutert er das Potenzial der PET zur Erschliessung neuer Möglichkeiten der Datennutzung unter Wahrung der Vertraulichkeit. PET fördern namentlich die gemeinsame Analyse von Daten, die ansonsten zu sensibel wären, als dass sie an Dritte weitergegeben oder für einen anderen Zweck als den ursprünglichen Erhebungszweck wiederverwendet werden könnten.

Auf Ebene der UNO hat der Expertenausschuss «UN Committee of Experts on Big Data and Data Science for Official Statistics» (UNCEBD) 2018 eine Arbeitsgruppe eingesetzt, die die nationalen Statistikämter bei der Nutzung von PET unterstützen soll (UNTask Team on Privacy-Enhancing-Technologies). Sie hat einen Leitfaden zum Thema veröffentlicht.¹¹ Das Datenschutzmanagement, insbesondere die Kontrolle des Offenlegungsrisikos, ist ein wichtiges Thema der Statistikämter, da sie gemäss Grundsatz 5 des Verhaltenskodex für europäische Statistiken (Code of Practice)¹² die Geheimhaltung der Daten, die ihnen anvertraut werden, gewährleisten müssen. Die OECD hat einen Bericht veröffentlicht, der den politischen Entscheidungstragenden helfen soll, besser zu verstehen, wie PET für einen einfacheren, datenschutzwahrenden Datenaustausch eingesetzt werden können.¹³ Im vorgängig erwähnten Erlass der Biden-Regierung werden PET im Übrigen klar als technisches Kernelement für eine sichere und effiziente Datennutzung in der öffentlichen Verwaltung bezeichnet.

Vor diesem Hintergrund schlägt dieser Bericht technische Konzepte mit grösstmöglichem Anwendungsbereich auf Bundes-, nationaler und internationaler Ebene vor. Diese Konzepte müssen folglich tragfähig, exportierbar, wiederverwendbar und mit Initiativen zur Standardisierung von Metadaten (I14Y-Plattform) und zur Förderung eines Gemeinschaftsvokabulars auf Basis der von LINDAS vorgeschlagenen Ontologien kompatibel sein.

Zunächst werden die Grundbegriffe und Stakeholder, die bei der gemeinsamen Nutzung identifizierender Daten involviert sind, definiert. Danach werden verschiedene PET vorgestellt und ihre jeweiligen Vor- und Nachteile für eine Verwendung in der Bundesverwaltung erläutert.

Schliesslich folgt die Präsentation zweier Fallbeispiele sowie der bestehenden Rechtsgrundlagen. Es wird aufgezeigt, wie PET zur Umsetzung der datenschutzrechtlichen Grundprinzipien beitragen können. Abschliessend werden Empfehlungen für einen Datenschutz-Gouvernanzrahmen formuliert, der die in diesem Bericht vorgestellten Technologien unter Einhaltung der datenschutzrechtlichen Bestimmungen bestmöglich nutzt. Dieser Gouvernanz-Rahmen muss sicherstellen, dass die eingesetzten Technologien die Bundesverwaltung bei ihren Aufgaben unterstützen, ohne ihre bestehenden Aufgaben wie jene der Militärverwaltung oder des Nachrichtendienstes des Bundes (NDB) zu beeinträchtigen.

¹¹ <https://unstats.un.org/bigdata/task-teams/privacy/guide/>

¹² <https://unstats.un.org/unsd/dnss/docs-ngaf/National%20Code%20of%20Good%20Practice.pdf>

¹³ https://www.oecd-ilibrary.org/science-and-technology/emerging-privacy-enhancing-technologies_bf121be4-en

2 Kernkonzepte

2.1 Grundbegriffe

Bevor die verschiedenen Technologien zur Nutzung datenwissenschaftlicher Methoden für eine gemeinsame Datennutzung in der Bundesverwaltung im Detail vorgestellt werden, müssen die in diesem Bericht verwendeten Begriffe erläutert werden. Datenschutz ist naturgemäss interdisziplinär und hängt stark vom Kontext ab, in dem die Daten verwendet werden. Aus diesem Grund ist es wichtig, die im vorliegenden Bericht verwendeten Begriffe klar abzugrenzen. Massgebend für die Definitionen sind einerseits die beschriebenen Fallbeispiele, d. h. die gemeinsame Nutzung von Daten der Bundesverwaltung, und andererseits die Kompatibilität mit den internationalen englischsprachigen Standardbegriffen. Die englische Bezeichnung der Konzepte steht jeweils in Klammern, damit der Bezug zur massgeblichen wissenschaftlichen Literatur gewährleistet ist.

In diesem Bericht bezeichnet eine **identifizierbare Einheit** (Identifiable Unit) bzw. eine **Datenschutz-einheit** (Privacy Unit) alle natürlichen oder juristischen Personen oder deren Gruppierungen (z. B. Haushalte), die nichtöffentliche und schützenswerte Informationen generieren. Bei natürlichen Personen wird eine identifizierbare Einheit auch betroffene Person oder Datensubjekt (Data Subject) genannt. Die Begriffe in diesem Bericht sind allgemein formuliert, da der Schutz einer Personengruppe, beispielsweise eines Haushalts, umfangreichere Massnahmen erfordert als der ihrer einzelnen Mitglieder.

Identifizierende Daten umfassen alle erhobenen Angaben zu einer identifizierbaren Person oder Personengruppe. Da die Identifizierbarkeit von Daten nicht durch ihre Art bestimmt wird, gelten alle Daten in Verbindung mit einer identifizierbaren Einheit als identifizierend. Die modernen Verfahren zur Reidentifizierung wie Angriffe durch Verknüpfung (siehe Kapitel 6.2 für weiterführende Informationen) verwenden nicht den Wert einzelner Daten, sondern eine Kombination mehrerer Daten.

Handelt es sich bei der identifizierbaren Einheit um eine natürliche Person im Sinne der Datenschutzgesetze, z. B. dem Bundesgesetz über den Datenschutz ([Art. 5 Bst. a DSG](#); [SR 235.1](#)), wird von **Personnendaten** (Personal Data) gesprochen. Bei juristischen Personen wird in der Gesetzgebung der Begriff **Daten juristischer Personen** verwendet (LOGA, [SR 172.10](#)). Die Anwendung einer datenschutzwohnenden Datenwissenschaft geht über diese Gesetze hinaus. Sie richtet sich beispielsweise auch nach der Verordnung über die Durchführung von statistischen Erhebungen des Bundes ([SR 431.012.1](#)), die Folgendes festhält: «Die Ergebnisse der Erhebungen werden in einer Form veröffentlicht oder zugänglich gemacht, die jede Identifizierung der befragten Personen, Haushalte, Unternehmungen oder Betriebe ausschliesst.» Dieses Beispiel, das sowohl den Schutz von natürlichen als auch von juristischen Personen und deren Gruppen einschliesst, begründet auch, warum in diesem Bericht zwischen Personendaten und identifizierenden Daten unterschieden wird.

Als **Datenprodukt** (Data Product) werden alle Produkte bezeichnet, die von identifizierenden Daten abgeleitet und mithilfe datenwissenschaftlicher Methoden erstellt wurden. Sie werden in der Regel in drei Hauptkategorien gegliedert:

- **Mikrodaten:** Datensätze mit Angaben zu einer Datenschutzeinheit, also zu einer oder mehreren Personen
- **Ergebnisse von Datenanalysen:** Zu dieser Kategorie gehören neben der Statistikproduktion alle objektiven Informationen zu einer Gruppe von Datenschutzeinheiten. Diese Daten sind für die Politikgestaltung besonders nützlich.
- **Machine-Learning- und KI-Modelle:** Sie können mithilfe von identifizierenden Daten trainiert werden. Dies birgt das Risiko, dass diese Daten beim Trainingsprozess oder bei der Verwendung des Modells offengelegt werden.

Die Anwendung einer datenschutzwahrenden Datenwissenschaft stellt sicher, dass die zur Produktstellung verwendeten identifizierenden Daten geschützt bleiben. In diesem Zusammenhang müssen zwei Ziele unterschieden werden:

- **Der Schutz der Privatsphäre** (Privacy): Er soll sicherstellen, dass die Datenprodukte eine Reihe von Standards erfüllen. Letztere werden nicht nur von den Gesetzen, sondern auch von der aktuellen gesellschaftlichen und kulturellen Realität beeinflusst und definieren, was zum Zeitpunkt der Erstellung des Datenprodukts akzeptabel bzw. nicht akzeptabel ist. Die datenschutzrechtliche Sensibilität der Daten ist kontextabhängig und kann daher nicht im Voraus bestimmt werden. Ein gutes Beispiel zur Veranschaulichung dieser Kontextabhängigkeit ist das Geschlecht einer natürlichen Person. Dieses gilt oft als nicht sensibel bzw. nicht schützenswert, da allgemein bekannt. Allerdings wird die Bekanntgabe, dass sich die Geschlechterzusammensetzung eines Haushalts von einem Jahr zum anderen verändert hat, insofern zur sensiblen Information, als es sich um eine Geschlechtsänderung handeln muss. Solche Fälle können bei der Erstellung eines Datenprodukts nicht vorausgesehen werden, zumal die Sensibilität einer Information auf einem gesellschaftlich beeinflussten Werturteil der jeweiligen Epoche beruht.
- Die **Vertraulichkeit** (Confidentiality): Sie umfasst alle Massnahmen zur Geheimhaltung von identifizierenden Daten. Vertraulichkeit ist somit objektiv, beständig, deutlich stärker und technisch schwieriger umzusetzen als Massnahmen zum Schutz der Privatsphäre.

Die im vorliegenden Bericht erläuterte datenschutzwahrende Datenwissenschaft soll nicht nur prüfen, ob die Produkte den aktuellen Standards entsprechen, sondern auch die Vertraulichkeit der identifizierenden Daten sämtlicher Datenschutzeinheiten sicherstellen. Hierzu wird ein formeller, von der Entwicklung der Normen unabhängiger Rahmen mit einem objektiv beurteilbaren Zweck erstellt. Als Vorbild dient das «United States Census Bureau», das rechtlich verpflichtet ist, in sämtlichen von ihm bereitgestellten Datenprodukten die Vertraulichkeit der Personendaten zu gewährleisten.¹⁴

Die Vertraulichkeit identifizierender Daten muss sowohl während der Erstellung des Datenprodukts als auch bei der Bereitstellung und Nutzung sichergestellt sein. Folglich muss zunächst die **Datensicherheit** gewährleistet werden, damit niemand unbefugt auf identifizierende Daten zugreifen kann. Die wichtigsten Werkzeuge zur Gewährleistung der Datensicherheit verwenden kryptografische Verfahren, Authentisierungsverfahren und Zugriffsberechtigungen. Die Vertraulichkeit wiederum muss sicherstellen, dass bei der Bereitstellung des Produkts keine identifizierbaren Daten **offengelegt** werden, d. h., dass bei der willentlichen Veröffentlichung eines Ergebnisses keine identifizierbaren Daten an die Öffentlichkeit gelangen. Diese beiden Aspekte der Vertraulichkeit ergänzen sich und definieren zwei Kategorien von Technologien, die im nächsten Kapitel beschrieben werden.

Um die Ziele einer datenschutzwahrenden Datenwissenschaft in das DSG (SR 235.1)¹⁵ und die Datenschutz-Grundverordnung (DSGVO)¹⁶ einzuordnen, werden folgende Definitionen in Erinnerung gerufen:

- **Anonymisierung**: Verfahren, mit dem identifizierende Daten unumkehrbar so verändert werden, dass ohne unverhältnismässigen Aufwand kein direkter oder indirekter Rückschluss auf eine identifizierbare Person möglich ist.

¹⁴ <https://www.govinfo.gov/content/pkg/USCODE-2007-title13/pdf/USCODE-2007-title13.pdf>

¹⁵ <https://www.fedlex.admin.ch/eli/cc/2022/491/de>

¹⁶ <https://eur-lex.europa.eu/eli/reg/2016/679/oj>

- **Pseudonymisierung:** Verfahren, bei dem die Personendaten so abgeändert werden, dass ohne zusätzliche Informationen kein Rückschluss auf die betreffende Person möglich ist. Voraussetzung ist, dass diese zusätzlichen Informationen gesondert aufbewahrt werden und durch technische und organisatorische Massnahmen so geschützt sind, dass die Daten nicht einer identifizierbaren Person zugeordnet sind.

Aus diesen Definitionen geht klar hervor, dass die Gewährleistung der Vertraulichkeit identifizierender Daten dem Konzept der Anonymisierung entspricht und pseudonymisierende Massnahmen nicht ausreichen, um die Vertraulichkeit ohne zusätzliche organisatorische und/oder vertragliche Massnahme zu garantieren.

Die jüngsten wissenschaftlichen Fortschritte haben gezeigt, dass die Anonymisierung rein technisch nicht garantiert werden kann und sich nie ganz ausschliessen lässt, dass identifizierende Daten offengelegt werden. Bei der Veröffentlichung eines Datenprodukts besteht, unabhängig von seiner Art, immer ein Restrisiko, das identifizierende Daten offengelegt werden. Mithilfe technischer Lösungen lässt sich dieses Risiko jedoch kontrollieren. Sie ermöglichen es, den Kompromiss zwischen Vertraulichkeit und Nutzen, den sogenannten **Privacy-Utility Trade-off**, zu steuern. Wie der Begriff nahelegt, ist es nicht möglich ist, ein Datenprodukt zu veröffentlichen, das die identifizierenden Daten exakt widerspiegelt und gleichzeitig ein hohes Mass an Vertraulichkeit gewährleistet. Je stärker das Offenlegungsrisiko eingeschränkt wird, desto geringer ist der Nutzen des Produkts und umgekehrt. Mit den nachfolgend erläuterten PET lässt sich dieser Kompromiss quantifizieren und bewusst steuern. Die Festlegung eines akzeptablen Offenlegungsrisikos und/oder eines Mindestnutzens ist ein politischer und gesellschaftlicher Drahtseilakt und lässt sich technologisch nicht lösen.

2.2 Input und Output Privacy

Die Erstellung und die Veröffentlichung eines Datenprodukts aus identifizierenden Daten sind Teil einer Wertschöpfungskette, an der folgende Stakeholder beteiligt sind:

- die identifizierbare Person (Data Subject): Datenschutzeinheit, die die identifizierenden Daten generiert.
- der Datenanbieter (Input Party): Organisation, die die identifizierenden Daten erhebt und ihre Verwertung durch die Erstellung eines Datenprodukts in Auftrag gibt. Das kann zum Beispiel ein nationales Statistikamt sein, das eine Statistik wie die Armutsquote produziert, oder ein Unternehmen, das Daten anhand von vernetzten Objekten sammelt.
- der Rechenanbieter (Computing Party): Organisation, die die nötige Rechenleistung für die Erstellung des Datenprodukts bereitstellt. Die Berechnungen erfolgen in der Regel auf Remote-Servern, d. h. über eine Cloud-Infrastruktur. Diese kann entweder von einer vertrauenswürdigen Organisation, z. B. einem Bundesamt, betrieben werden, oder von einem Privatunternehmen, das sämtlichen Kunden den Zugang zu gemeinsam genutzten und somit öffentlichen Servern ermöglicht.
- der Datenkonsument (Result Party): Personenkreis, an den das Datenprodukt verteilt wird. In der öffentlichen Verwaltung wird z. B. jedes OGD-Produkt uneingeschränkt bereitgestellt. Da die Öffentlichkeit den grössten Verteilerkreis bildet und das Risiko, dass identifizierende Daten extrahiert werden, hier am grössten ist, bedarf es einer strikten Offenlegungskontrolle. Ein Produkt kann auch einem kleineren Kreis zugänglich gemacht werden, z. B. den Mitarbeitenden der Bundesverwaltung. Da diese dem Amtsgeheimnis unterstehen, ist das Extraktionsrisiko kleiner als bei einer Bereitstellung für die breite Öffentlichkeit.

Privacy-Enhancing-Technologies (PET; datenschutzfördernde Technologien) sind technische Lösungen, mit denen das Risiko, dass die Vertraulichkeit identifizierender Daten verloren geht, entlang der gesamten Wertschöpfungskette abhängig vom gegenseitigen Vertrauen der Stakeholder gesteuert werden kann. «Vertrauen» ist hier im weitesten Sinn zu verstehen, d. h. als beliebiger Grund, der gegen die

Weitergabe identifizierender Daten an einen anderen Stakeholder spricht. Das kann z. B. ein Gesetz sein, das den Austausch identifizierender Daten zwischen den Stakeholdern verbietet, oder das fehlende Vertrauen der betroffenen Personen in den Datenanbieter oder fehlende Rechtsgrundlagen bzw. fehlende organisatorische Massnahmen zur Verhinderung von Datenextraktionsversuchen. Das fehlende Vertrauen zwischen zwei oder mehreren Stakeholdern kann mit PET kompensiert werden. Diese lassen sich in zwei Hauptkategorien unterteilen:

- Technologien zur Sicherstellung der **Vertraulichkeit der Inputs** (Input Privacy): Sie sorgen dafür, dass zwei oder mehrere der oben beschriebenen Stakeholder Daten für eine Berechnung übermitteln können und jeder Stakeholder dabei nur auf die eigenen identifizierenden Daten zugreifen kann. In den meisten Fällen basieren diese Technologien auf Verschlüsselungstechniken und gehören damit zu den Lösungen, mit denen die Datensicherheit gewährleistet wird.
- Technologien zur Sicherstellung der **Vertraulichkeit der Outputs** (Output Privacy): Sie zielen darauf ab, die Offenlegung identifizierender Daten aus dem Datenprodukt zu verhindern. Das Produkt selbst wird dabei willentlich veröffentlicht. Dieses Verfahren ist auch als (statistische) Offenlegungskontrolle bekannt und wird von den nationalen Statistikämtern bereits seit mehreren Jahren weiterentwickelt.

Je nach Vertrauen zwischen den verschiedenen Stakeholdern der Wertschöpfungskette identifizierender Daten kombiniert die für die Vertraulichkeit verantwortliche Organisation – in der Regel der Datenanbieter – mehrere Technologien, um die Input und Output Privacy sicherzustellen. Die Beispiele der Abbildungen 1, 2 und 3 zeigen, wo entlang der Wertschöpfungsketten der identifizierenden Daten Input- und Output-PET ein Vertrauensdefizit zwischen den Parteien schliessen.

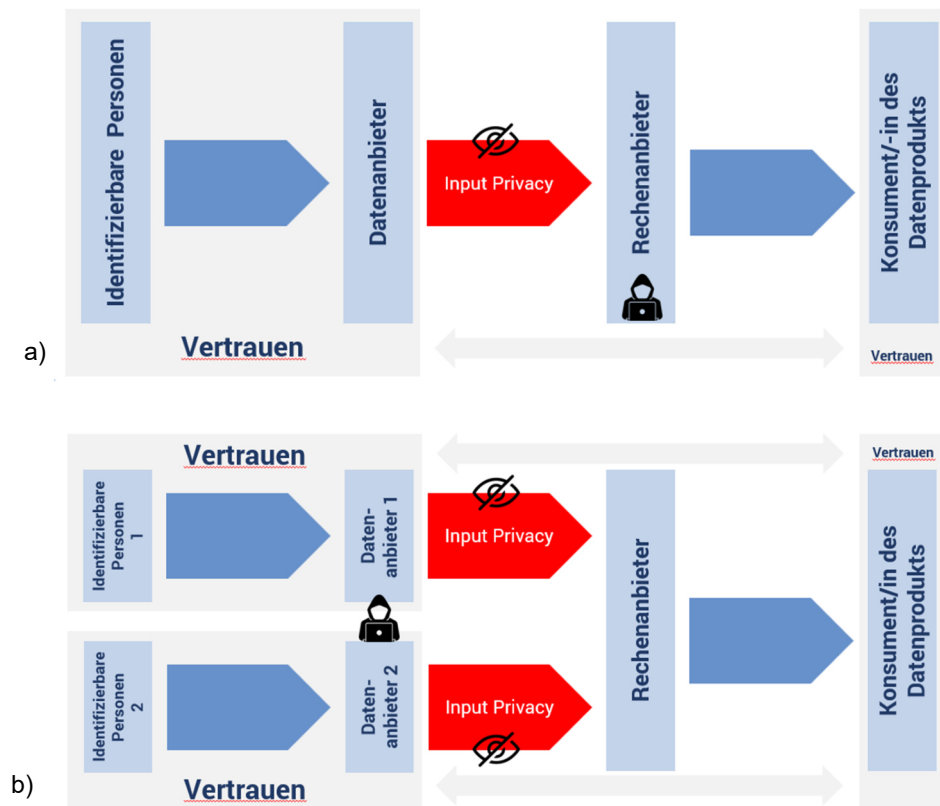


Abbildung 1: Beispiel eines Vertrauensschemas mit Verwendung von Input-PET mit a) einem einzigen Datenanbieter und b) zwei Datenanbietern. In beiden Fällen verbindet den Datenanbieter und die Datennutzenden ein Vertrauensverhältnis.

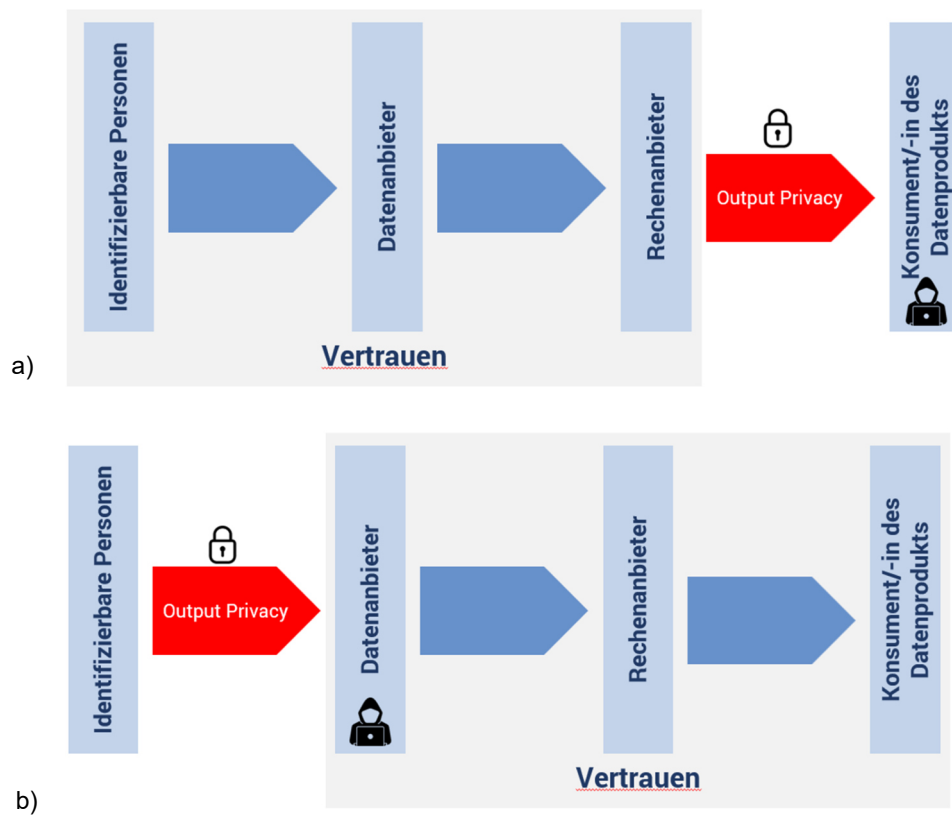


Abbildung 2: Beispiel eines Vertrauensschemas mit Verwendung von Output-PET mit a) dem Referenzschema für «Open Government Data» (OGD) und b) dem Schema bei fehlendem Vertrauen der identifizierenden Personen (Einheiten) in den Datenanbieter.

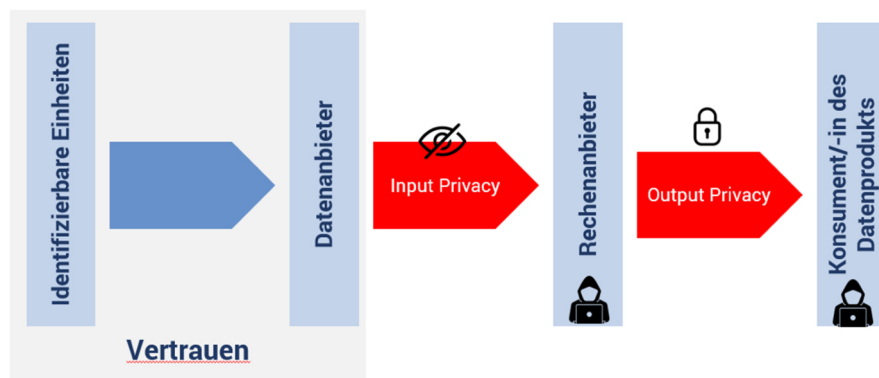


Abbildung 3: Beispiel eines Vertrauensschemas mit einer Kombination von Input- und Output-PET

3 Privacy-Enhancing-Technologies

3.1 Trusted Execution Environment

Prinzip

Eine sogenannte Trusted Execution Environment (vertrauenswürdige Ausführungsumgebung, TEE) auch Confidential Computing, d. h. vertrauliche Datenverarbeitung genannt, dient dazu, bei der Ausführung eines Informatikprogramms die Codes und Daten vor Software- und physischen Angriffen zu schützen, um ihre Sicherheit und Integrität zu gewährleisten. Dank dieser Technologie kann ein Informatikprogramm auf einer Remote-Infrastruktur privat ausgeführt werden, auch wenn zwischen dem Datenanbieter und den Datenkonsumierenden kein Vertrauensverhältnis besteht. TEE garantieren die Integrität des Codes und schützen die Daten vor Manipulationen innerhalb der Umgebung, indem sie unbefugte Zugriffe von aussen verhindern. Die Integrität des ausgeführten Codes wird von einer Ausführungsbestätigung bescheinigt, die vom Datenanbieter mithilfe eines kryptografischen Echtheitsnachweises generiert wird. Für die Einrichtung dieser TEE werden Hardware-Komponenten bestimmter Hersteller oder IT-Dienstleister benötigt. Sie sind in der Regel in spezifischen Prozessorenmodellen verbaut.

Grenzen und Risiken

Welche TEE implementiert werden können, hängt vom Prozessorenmodell ab. Einige lassen sich zudem nur über spezifische Cloud-Dienste implementieren. Dadurch entsteht eine Abhängigkeit vom jeweiligen Anbieter. Bei einigen TEE-Modellen ist zudem der verfügbare Arbeitsspeicher beschränkt. Diese Beschränkung kann zwar mit spezifischen Lösungen umgangen werden, von einem übermässigen Gebrauch dieser Möglichkeit wird aber abgeraten.

Da die Technologie von der Hardware abhängt, weist sie technische Schwachstellen auf. Diese können in Zusammenhang mit dem Zugriff auf das Gerät stehen oder mit nachträglich erkannten Sicherheitslücken, durch die sich das System nicht mehr updaten lässt und somit veraltet ist. Zudem unterscheiden sich die Systeme je nach Anbieter und verwendeter Technologie erheblich, was die Einrichtung einer TEE zusätzlich erschwert.

Der Code ist in einer TEE nicht zwingend geheim, er kann der Systemadministratorin bzw. dem Systemadministrator bekannt sein. Obwohl die Umgebung die Ausführung des Codes schützt, ist der Code selbst möglicherweise nicht ausreichend vor Angriffen von aussen, d. h. über die von ihm bereitgestellten Schnittstellen, abgesichert. Dadurch ist das System potenziell anfällig für Seitenkanalattacken (Side-Channel-Attacken).

Stärken und Ausgereiftheit

Mehrere Branchen verwenden bereits TEE, insbesondere, um die Ausführung sensibler Codes auf Smartphones, die als nicht vertrauenswürdige Geräte gelten, zu schützen. Auch Anwendungen von Finanzdienstleistern verwenden TEE häufig für den sicheren Zahlungsverkehr und zur Identifizierung. Ein grosser Vorteil dieser Technologie ist, dass sie sich einfach in andere PET integrieren lässt, ohne dass die Ausführung des Codes merklich verlangsamt wird.

Nicht zuletzt bieten viele Anbieter von Cloud-Diensten (Azure, AWS usw.) TEE auf ihren Plattformen an, was ihre Integration in den Organisationen vereinfacht. Aktuell wird vor allem daran gearbeitet, die Risiken möglicher Angriffe auf TEE zu identifizieren und eine Norm festzulegen. Einige Akteure wie Globalplatform und Enarx sind hier besonders aktiv.

Einige TEE haben gegenüber anderen PET den Vorteil, dass bestehende Algorithmen (z. B. zur Datenanalyse) flexibel in das sichere Umfeld übertragen werden können, ohne dass der Quellcode erheblich geändert werden muss. Allerdings hängt diese Möglichkeit stark davon ab, in welchem Rahmen die TEE implementiert werden. Die entsprechende Forschung steht erst am Anfang.

In der Datenwissenschaft lassen sich TEE in verschiedenen Bereichen einsetzen. Beispielsweise kann ein Anbieter (Eigentümer eines Lernmodells) sein Modell auf einer TEE bereitstellen, damit ein Kunde (Dateninhaber) es mit der Gewissheit nutzen kann, dass seine Daten sicher übermittelt werden. Der Kunde selbst kann nicht direkt auf das Modell und der Anbieter nicht direkt auf die Daten zugreifen.

3.2 Homomorphe Verschlüsselung

Prinzip

Die homomorphe Verschlüsselung ist eine kryptografische Technik, die es ermöglicht, dank der mathematischen Eigenschaften der Daten Berechnungen anhand von verschlüsselten Daten durchzuführen. Die Daten müssen folglich für statistische Analysen, Hypothesentests oder Interferenzen mit KI-Modellen nicht entschlüsselt werden.

Mit dieser Technik kann zum Beispiel ein Anbieter sensibler Daten («A») die Bearbeitung und Speicherung identifizierender Daten an einen Rechenanbieter («B») in einer Cloud-Infrastruktur auslagern. Angenommen, «A» ist ein Mitarbeiter eines Bundesamts, der das Machine-Learning-Programm von «B» verwenden möchte. Das Modell birgt das Risiko, dass identifizierende Daten aus Steuererklärungen zu Betrugszwecken genutzt werden. «A» verschlüsselt die Daten homomorph und übermittelt sie an «B». «B» lässt die verschlüsselten Daten durch das Modell laufen und erhält ein verschlüsseltes Ergebnis, das er an «A» schickt. «A» entschlüsselt das Ergebnis, um das Betrugsrisiko in Erfahrung zu bringen. «B» hingegen erhält keine Informationen zu den Daten.

Zum Zeitpunkt der Berichterstellung ist die homomorphe Verschlüsselung die einzige Technologie, bei der die Daten während des gesamten Prozesses verschlüsselt bleiben. Der Datenanbieter muss dem Rechenanbieter somit nicht vertrauen, da dieser zu keinem Zeitpunkt auf die verschlüsselten Daten zugreifen kann.

Die «Fully Homomorphic Encryption» (FHE) ist die fortschrittlichste Form der homomorphen Verschlüsselung. Mit ihr lassen sich unbegrenzt viele Berechnungen anhand verschlüsselter Daten durchführen, sodass auch komplexere Anwendungen wie das Training tiefer neuronaler Netze möglich sind.

Grenzen und Risiken

Die homomorphe Verschlüsselung birgt ähnliche Risiken wie die anderen Verschlüsselungstechniken. Sie erfordert insbesondere eine sichere Speicherung der Verschlüsselungscodes sowie einen Schutz gegen «Brute Force»-Attacken, Kryptanalysen, Seitenkanalattacken und Datenabfang.

Die grösste Einschränkung der homomorphen Verschlüsselung ist der grosse Bedarf an IT-Ressourcen: Verschlüsselte Daten sind signifikant voluminöser als unverschlüsselte Daten und benötigen daher mehr Speicherplatz. Ausserdem sind Berechnungen anhand von verschlüsselten Daten erheblich langsamer. Die homomorphe Verschlüsselung eignet sich daher nicht für Echtzeitanwendungen.

Die homomorphe Verschlüsselung ist eine Input-PET. Das heisst: Die Daten bleiben beim Rechenanbieter zwar verschlüsselt, es gibt aber keine Garantie, dass sie nach der Entschlüsselung des Datenprodukts durch die Konsumentin oder den Konsumenten nicht offengelegt werden.

Schliesslich lassen sich mit der homomorphen Verschlüsselung nicht alle Analysen durchführen: Nur lineare Operationen sind exakt berechenbar, bei komplexeren Operationen muss auf Näherungswerte zurückgegriffen werden. Die Verschlüsselungsgarantie beeinträchtigt somit die Genauigkeit.

Stärken und Ausgereiftheit

Die homomorphe Verschlüsselung wird in der Praxis erfolgreich verwendet und es gibt Open-Source-Bibliotheken. Zu den gängigsten Anwendungen zählt die private Mengenüberschneidung (Private Set Intersection). Sie stellt fest, ob es zwischen zwei Datensätzen unterschiedlicher Datenanbieter, die ihre identifizierenden Daten nicht miteinander teilen möchten, Übereinstimmungen gibt. Eine weitere häufige Anwendung ist die Bereitstellung von Machine-Learning-Modellen wie im genannten Beispiel des Betrugsrisikos.

Auf dem Gebiet der homomorphen Verschlüsselung wird aktiv geforscht und die Technologie hat in den letzten fünf Jahren grosse Fortschritte gemacht. Insbesondere wurden spezialisierte Hardware-Lösungen entwickelt, die die Rechenzeit reduzieren und die Technologie zukunftsfähig machen. Sollten die Fortschritte bei der homomorphen Verschlüsselung weiter so stark voranschreiten, könnte sie in vielen Anwendungen eingesetzt werden und würde in verschiedenen Bereichen ein riesiges Potenzial bieten.

3.3 Secure Multiparty Computation

Prinzip

Mit der sicheren Mehrparteienberechnung (Secure Multiparty Computation, MPC oder SMPC) können mehrere Teilnehmende gleichzeitig Berechnungen durchführen, ohne dass die privaten Inputs (Daten) der Teilnehmenden offengelegt werden. MPC ist ein Teilbereich der Kryptografie. Im Gegensatz zur herkömmlichen Kryptografie, die vor systemexternen Angriffen schützt, dient MPS dazu, die Datensicherheit zwischen den verschiedenen Teilnehmenden zu garantieren. Diese einigen sich auf den anzuwendenden Algorithmus und verwenden dann ein MPC-Protokoll, um den Output dieses Algorithmus aus den geheimen Inputs zu berechnen, ohne die Inputs preiszugeben. Die Teilnehmenden erfahren jeweils nur das, was sich aus dem Output ableiten lässt. MPC stellt sicher, dass die Inputs der einzelnen Parteien geheim bleiben (Input Privacy) und garantiert die Integrität des Ergebnisses.

Es wird zwischen zwei Kategorien von MPC-Algorithmen unterschieden: Die erste Kategorie umfasst alle Arten von Algorithmen, d. h. auch solche, die zur Erstellung eines Datenprodukts verwendet werden. Sie nutzt häufig das Konzept der Geheimnisteilung (Secret Sharing), bei der die Inputs in mehrere Teile aufgeteilt und auf die einzelnen Parteien verteilt werden. Anschliessend führt der MPC-Algorithmus die Berechnungen auf den Teilmengen durch. Sie werden erst wieder zusammengeführt, wenn die Berechnungen abgeschlossen sind, sodass die Inputs der einzelnen Parteien während der gesamten Durchführung des Algorithmus geheim bleiben. Die zweite Kategorie umfasst nur die Algorithmen für die Produktion hochspezifischer Datenprodukte (z. B. Verknüpfung privater Datensätze) und sind gezielt für diese Produkte optimiert.

Grenzen und Risiken

MPC-Protokolle bieten ein hohes Mass an Sicherheit mit einer formalen Garantie, setzen aber generell kryptografische Verfahren voraus, die hohe Berechnungs- und Kommunikationskosten verursachen. Sie eignen sich somit nur für bestimmte Aufgaben.

Die Risikomodelle für MPC sind unterschiedlich eng gefasst und berücksichtigen auch die Möglichkeit, dass Parteien korrupt sein könnten oder eine Kollusionsgefahr besteht. Einige Algorithmen sind selbst dann sicher, wenn die Mehrheit der Parteien korrupt ist. Wichtig ist, dass vor der Verwendung von MPC ein Gefahrenmodell festgelegt und die geeigneten Algorithmen und Bibliotheken mit Bedacht gewählt werden.

Damit MPC überhaupt verwendet werden kann, muss zur Ausführung von Algorithmen häufig der Code neu geschrieben werden. Bibliotheken für MPC-geeignete Open-Source-Software bieten ausserdem nicht die gleiche Bandbreite an datenwissenschaftlichen Funktionen, was bei der Anpassung eines Codes für MPC zusätzlichen Aufwand bedeutet.

Nicht zuletzt schränkt MPC zwar den Zugriff der Parteien auf die Input-Daten der anderen Teilnehmenden ein, garantiert aber nicht die Vertraulichkeit des berechneten Ergebnisses. Tatsächlich lassen sich aus dem Ergebnis der Berechnung manchmal sensible Informationen ableiten. Vor der Verwendung von MPC ist daher gründlich zu überlegen, welche Berechnungen durchgeführt werden sollen. Oft muss MPC mit Techniken zum Schutz der Output Privacy wie Differential Privacy kombiniert werden.

Stärken und Ausgereiftheit

MPC kommt bei Anwendungen mit kritischer Input Privacy zum Einsatz, etwa bei gemeinsamen Datenanalysen im Gesundheits- und Finanzwesen sowie in anderen sensiblen Bereichen. In Estland zum Beispiel wurden MPC-Techniken im Rahmen einer Studie über den Zusammenhang zwischen studienbegleitender Erwerbstätigkeit und Studienabbruch verwendet. Für die Studie wurden Daten von zwei voneinander unabhängigen staatlichen Stellen benötigt, die nicht über die nötigen Rechtsgrundlagen für eine gemeinsame Datennutzung verfügen. Ohne MPC wäre die Studie daher nicht möglich gewesen.

Die grosse Mehrheit der MPC-basierten Projekte greifen auf proprietäre Software-Bibliotheken zurück. Es gibt noch keine Open-Source-Bibliotheken für MPC, die reif genug für eine Produktivsetzung sind. Aktuell liegt der Forschungsfokus auf der Optimierung der MPC-Produkte zur Entlastung der Computer- und Kommunikationssysteme. Die Effizienz von MPC sollte sich mit den Fortschritten der kryptografischen Techniken verbessern, sodass die Technik für eine grössere Bandbreite an Anwendungen infrage käme. Die Kombination von MPC und Differential Privacy (siehe unten) zur Geheimhaltung der Outputs birgt ebenfalls grosses Potenzial, allerdings steht die Forschung dort erst am Anfang.

3.4 Distributed Learning

Prinzip

Beim Distributed Learning (verteilttes Lernen) können mehrere Parteien einen Machine-Learning-Algorithmus, insbesondere ein neuronales Netz, so trainieren, dass die Sicherheit der Trainingsdaten bei beiden Parteien (Input Privacy) garantiert ist. Die Daten bleiben auf den Geräten des Dateninhabers und werden nie zwischen den Parteien ausgetauscht. Einzig auf die Modellparameter, d. h. das Lernergebnis, haben alle Parteien Zugriff. Beim Distributed Learning muss eine Drittpartei, d. h. der Rechenanbieter, den Prozess koordinieren. Er hat zu keinem Zeitpunkt Zugang zu den Inputdaten, spielt aber eine zentrale Rolle bei der Aggregation und Verteilung der Trainingsparameter an die Parteien.

Es wird zwischen zwei Arten von Distributed Learning unterschieden: dem Federated Learning (föderiertes Lernen) und dem Split Learning (fraktioniertes Lernen). Beim Federated Learning verfügen alle Teilnehmenden über ihre eigene Kopie des Modells, das sie individuell mit ihren eigenen Daten trainieren. Zur Aggregation senden sie die Parameter ihres Modells an einen zentralen Server. Daraus wird ein globales Modell erstellt und an die einzelnen Teilnehmenden für eine neue Trainingsrunde verteilt. Beim Split Learning wird das Modell so aufgeteilt, dass alle Teilnehmenden auf ihrer jeweiligen Infrastruktur einen anderen Teil des Modells mit ihren eigenen Daten trainieren. Alle diese Teile produzieren ein Zwischen-Output. Während des Trainingsprozesses trainiert auch der zentrale Server ein Modell, das die Zwischen-Outputs miteinander kombiniert. Im Unterschied zum Federated Learning muss das Split Learning zwingend sequenziell erfolgen. Beim Split Learning sind die Input-Daten jedoch besser geschützt als beim Federated Learning, da den verschiedenen Parteien mehr Informationen zu den Parametern und der Architektur des Modells verborgen bleiben und sie so weniger Möglichkeiten haben, in das Modell hineinzusehen.

Grenzen und Risiken

Distributed Learning verursacht mehr Kommunikationskosten für die Datenübermittlung zwischen den Parteien und dem zentralen Server als die klassische, zentralisierte Trainingsmethode der Machine-

Learning-Modelle. Bei Techniken, die bei jedem Schritt des Modelloptimierungsalgorithmus eine Synchronisierung zwischen den Parteien erfordern, nimmt der Lernprozess zudem mehr Zeit in Anspruch. Verteilt trainierte KI-Modelle sind generell weniger leistungsfähig als zentralisiert trainierte Modelle.

Distributed Learning dient dazu, die Vertraulichkeit der privaten Input-Daten zu gewährleisten. Da nicht ausgeschlossen werden kann, dass die während der Trainingsphase übermittelten Modellparameter und das daraus abgeleitete Modell Informationen zu den Daten der Teilnehmenden preisgeben, braucht es zusätzlich eine Technologie, mit der die Vertraulichkeit der Outputs sichergestellt werden kann, damit sich das Offenlegungsrisiko kontrollieren lässt.

Distributed Learning ist zudem anfällig für Korruption, sowohl auf Ebene der Teilnehmenden als auch des Aggregationsservers. Teilnehmende können zum Beispiel falsche Trainingsdaten eingeben, um das Verhalten des Modells gezielt für eigene Interesse zu verändern.

Stärken und Ausgereiftheit

Distributed Learning fördert die Zusammenarbeit zwischen Organisationen: Sie können ohne Datenzentralisierung gleichzeitig ein Modell trainieren, wodurch sich bestimmte Input-Privacy-Probleme gar nicht erst stellen. Darüber hinaus fördert die Vielfalt der beim Distributed Learning verwendeten Datenquellen die Entwicklung robusterer und leistungsfähigerer Modelle. Gesundheitsdaten stehen zum Beispiel nicht immer in ausreichender Menge zur Verfügung, um ein leistungsstarkes KI-Modell für medizinische Diagnosen zu trainieren. Beim Distributed Learning kann mit Daten aus mehreren Spitälern ein robustes KI-Modell trainiert und gleichzeitig die Input Privacy verbessert werden.

Es gibt mehrere Open-Source-Bibliotheken, die das Distributed Learning von KI-Algorithmen ermöglichen. Sie lassen sich gut in klassische Machine-Learning-Bibliotheken integrieren und machen diese Technologie somit nutzbar.

3.5 Null-Wissen-Beweis

Prinzip

Ein Null-Wissen-Beweis (Zero Knowledge Proof, ZKP) ist ein kryptografisches Protokoll, bei dem ein Beweiser einem Verifizierer beweist, dass eine Aussage wahr ist, ohne die zugrunde liegenden Informationen preiszugeben. Mit dem Null-Wissen-Beweis lässt sich zum Beispiel folgende Aussage verifizieren: Auf die gesamte Finanzgeschichte des Unternehmens X wurde ein Machine-Learning-Modell zur Beurteilung der Kreditwürdigkeit angewendet, ohne Details zum Modell oder zu den Finanzdaten des Unternehmens offenzulegen. Null-Wissen-Beweise sind probabilistisch, d. h., es bleibt immer ein Restrisiko, dass die Aussage als falsch eingestuft wird, obwohl sie stimmt. Dieses Risiko lässt sich allerdings durch den Einsatz zusätzlicher Rechenleistung willkürlich reduzieren.

Die wichtigsten Anwendungsbereiche von Null-Wissen-Beweisen sind Authentifizierung, Identifizierung und Kryptowährungen. In der Datenwissenschaft können Null-Wissen-Beweise den Austausch und die Zusammenarbeit zwischen Stakeholdern in Szenarien, in denen identifizierende Informationen geschützt werden müssen, vereinfachen. Damit lässt sich zum Beispiel die Art der identifizierenden Daten bestimmen, ohne sie preiszugeben, oder die Identität einer Person überprüfen, ohne persönliche Informationen zu verraten. So ist zum einen der Datenfluss gegen Lecks gesichert, zum anderen teilen die Stakeholder Daten nicht mit einem allfälligen Rechenanbieter.

Grenzen und Risiken

Die Überprüfung von Null-Wissen-Beweisen ist sehr ressourcenintensiv. Sie erfordert entweder eine umfangreiche Kommunikation zwischen dem Verifizierer und dem Beweiser oder eine hohe Rechenkapazität. Aus diesem Grund werden Null-Wissen-Beweise derzeit in der Praxis nur bei sehr einfachen Problemen eingesetzt, die entweder umfangreiche Datensätze oder Echtzeit-Daten verwenden.

Die Erstellung neuer ZKP-Protokolle ist insofern technisch schwierig, als sich menschliche Probleme nicht so einfach in eine mathematische Aussage umwandeln lassen. Für ihre Implementierung werden daher nicht nur IT-Entwicklerinnen und -Entwickler benötigt, sondern auch Personen mit guten mathematischen und kryptografischen Kenntnissen. Mit dem Aufkommen von Quanten-Rechenleistungen werden viele ZKP-Protokolle nicht mehr sicher sein.

Neben ihrer Eigenschaft als Technologie der Vertraulichkeit der Inputs setzen Null-Wissen-Beweise oft ein gewisses Mass an Vertrauen zwischen den Datenanbietern und den Rechenanbietern voraus und lassen sich daher nur beschränkt auf spezifische Vertrauensschemas anwenden.

Stärken und Ausgereiftheit

Die Technologie ist unterschiedlich ausgereift. Einige ZKP-Protokolle werden bereits gut beherrscht und angewendet, andere hingegen befinden sich noch in der Forschungsphase. Da Null-Wissen-Beweise für Blockchains, bei denen die Transaktionen ohne den Austausch identifizierender Daten überprüft werden können, unverzichtbar sind, werden grossen Anstrengungen unternommen, um die bestehenden Grenzen zu überwinden. Ausserdem wird an einer Taxonomie und an einer Standardisierung gearbeitet.

Ein typisches Beispiel für eine datenwissenschaftliche Anwendung von Null-Wissen-Beweisen besteht darin, die erfolgreiche Ausführung eines Algorithmus zu bestätigen, ohne die verwendeten Daten offenzulegen. Im Rahmen von Machine Learning und KI kann der Eigentümer eines Modells das Ergebnis des Modells bestätigen, ohne das Modell zu teilen. Null-Wissen-Beweise sind zudem ein wichtiger Bestandteil bei der Implementierung von E-Voting-Systemen, da sie die Gültigkeit der Wahlen und Abstimmungen und die Anonymität der Wählenden bzw. Abstimmenden garantieren. Im Februar 2023 hat die EU eine auf Null-Wissen-Beweisen basierende «digitale Identität» verabschiedet, mit der sich alle Personen, die in der EU wohnhaft sind oder einen EU-Pass besitzen, authentifizieren und persönliche Informationen zur Nutzung von Clouds für das Ausleihen von Velos oder die Bezahlung der Steuern beglaubigen können.

3.6 Differential Privacy

Prinzip

Differential Privacy steht nicht für ein Tool oder einen Algorithmus, sondern für einen methodischen Rahmen. Dieser definiert die mathematischen Eigenschaften, die ein für statistische Analysen verwendeter Algorithmus erfüllen muss, um das Offenlegungsrisiko bei der Weitergabe des zugehörigen Datenprodukts mathematisch zu quantifizieren. Differential Privacy ist folglich eine Technologie zur Kontrolle der Output Privacy.

Die Technologie beruht auf der Randomisierung des Outputs eines Algorithmus, bei dem Extraktionen von identifizierenden Daten aus dem Datenprodukt unzuverlässig gemacht werden. Ein Algorithmus, der die Voraussetzungen der Differential Privacy erfüllt, garantiert, dass das Risiko, identifizierende Daten offenzulegen – unabhängig von den Zusatzinformationen und den technischen Mitteln, die dem Angreifer zur Verfügung stehen – kontrolliert wird. Dieses Risiko wird mit einem Privacy-Loss-Budget quantifiziert. Welches Budget verwendet wird, hängt von der Sensibilität der Daten und dem Verwendungszweck des geschützten Datenprodukts ab.

Differential Privacy kann auch direkt auf dem ursprünglichen Datensatz angewendet werden. In diesem Fall spricht man von Local Differential Privacy. Dabei werden Daten unzuverlässig gemacht, indem sie willkürlich gestört werden, sodass sie nicht gegen die identifizierbare Einheit verwendet werden können. Diese Methode, auch Randomized-Response-Technik genannt, wird seit mehreren Jahren von nationalen Statistikämtern, Befragungsinstituten und seit Kurzem auch in der Privatwirtschaft verwendet, um die Ehrlichkeit einer Antwort abzuschätzen oder sich zu vergewissern, dass die Datenerhebung zu sensiblen Themen wie ausserehelichen Beziehungen oder Telemetrie (Lokalisierung, installierte Apps, Tastaturerfassung usw.) der PCs akzeptiert wird.

Grenzen und Risiken

Zur Implementierung von Differential-Privacy-Mechanismen müssen willkürliche Störfaktoren (sogenanntes Rauschen) in den Algorithmus eingebettet werden. Dazu muss evaluiert und klar kommuniziert werden, in welchem Rahmen und mit welchem Vertrauensgrad die gestörten Datenprodukte oder Mikrodaten verwendet werden dürfen. Die Evaluation wird von Fall zu Fall vorgenommen und erfordert ein hohes Mass an Fachwissen in den Bereichen Datenvertraulichkeit und mathematische Statistik.

Derzeit wird in zahlreichen Projekten daran gearbeitet, Standards oder Richtlinien festzulegen, die abhängig vom Kontext und der Sensibilität des Datenprodukts das Ausmass der Störung kalibrieren.¹⁷ Die Wahl des Privacy-Loss-Budgets und damit die Justierung des Privacy-Utility Trade-off (Kompromiss zwischen Nutzen und Vertraulichkeit) bleibt aber vor allem eine politische Herausforderung, die Diskussionen zwischen Fachpersonen der Differential Privacy, Regulationsbehörden und politischen Entscheidungstragenden erfordert. Es gibt keine Universallösung, da die Wahl des Privacy-Loss-Budgets vom Datenschutz-Gouvernanz-Rahmen und somit von den «Five Safes»¹⁸ abhängt. Besonders wichtig ist dabei, dass die Erwartungen der Empfangenden, d. h. der «Safe People», berücksichtigt werden, da die Zuverlässigkeit des veröffentlichten Produkts nur für bestimmte Verwendungszwecke gewährleistet werden kann.

Zudem setzt Differential Privacy fundierte Fachkenntnisse voraus, die über mathematisches und technisches Wissen hinausgehen. Die Implementierung ist in Bezug auf die IT-Infrastruktur nicht trivial, da sowohl das Privacy-Loss-Budget als auch die für die Analysen verwendeten Datenversionen streng überwacht werden müssen.

Stärken und Ausgereiftheit

An der 2006 formell eingeführten Differential Privacy¹⁹ wird in akademischen Kreisen intensiv geforscht und experimentiert. Die entsprechenden Forschungen knüpfen an die Arbeiten zur Kontrolle des Offenlegungsrisikos statistischer Daten an und sind eine Reaktion auf die zunehmenden Kenntnisse über Angriffe zur Offenlegung identifizierender Daten wie Reconstruction, Linking und Membership Inference Attacks. Heute wird Differential Privacy sowohl in der Privatwirtschaft (Apple, Google, Microsoft) als auch von staatlichen Einrichtungen (United States Census Bureau) verwendet, sodass die verfügbaren Open-Source-Bibliotheken ausreichend ausgereift und zuverlässig sind. Darüber hinaus ist ein Differential-Privacy-Algorithmus nur geringfügig komplexer als ein klassischer Algorithmus, und nicht mit nennenswerten IT-Zusatzkosten verbunden.

Differential Privacy als Technologie zur Gewährleistung der Output Privacy ist aus folgenden Gründen besonders vielversprechend:

- Das Risiko, dass identifizierende Daten aus einem Datenprodukt extrahiert werden, kann mathematisch zuverlässig eingeschränkt werden, unabhängig von den Zusatzinformationen und technischen Mitteln, die dem Angreifer zur Verfügung stehen. Diese Garantie stellt sicher, dass Extraktionsversuche zu keinen zuverlässigen Ergebnissen führen, und definiert das höchste Zuverlässigkeitsniveau, das ein Angriff erreichen kann.

¹⁷ <https://www.nist.gov/news-events/news/2023/12/nist-offers-draft-guidance-evaluating-privacy-protection-technique-ai-era>

¹⁸ <https://ukdataservice.ac.uk/help/secure-lab/what-is-the-five-safes-framework/>

¹⁹ Dwork, C. (2006). Differential Privacy. In: Bugliesi, M., Preneel, B., Sassone, V., Wegener, I. (Hrsg.) Automata, Languages and Programming. ICALP 2006. Lecture Notes in Computer Science, vol 4052. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11787006_1

- Das kumulative Offenlegungsrisiko kann auch bei mehreren aufeinanderfolgenden, voneinander unabhängigen Veröffentlichungen von Datenprodukten unterschiedlicher Art quantifiziert werden.
- Über das Hinzufügen von Rauschen kann transparent informiert werden, ohne die Vertraulichkeit zu gefährden, was wiederum eine öffentliche Debatte zum Kompromiss zwischen Vertraulichkeit und Nutzen ermöglicht.
- Die von Differential Privacy garantierte Offenlegungskontrolle kann für die Wiederverwendung von Daten der Bundesverwaltung in der Bundesverwaltung selbst oder in der Schweizer Gesellschaft (Forschung, Privatwirtschaft, Zivilgesellschaft) genutzt werden, ohne dass diese Daten ausgetauscht werden müssen. Dabei wird das Risiko, dem sich die Personen durch die Angabe ihrer Daten aussetzen, kontrolliert, und sie haben die Gewissheit, dass keine identifizierbaren Daten zuverlässig extrahiert werden können. Kapitel 2 erläutert dieses Anwendungsbeispiel sowie sein grosses gesellschaftliches Potenzial im Detail.

3.7 Synthetische Daten

Prinzip

Bei der Generierung synthetischer Daten wird ein alternativer Datensatz erzeugt, der einen Detaillierungsgrad auf Stufe einer identifizierbaren Person aufweist, aber die identifizierenden Daten des Originaldatensatzes nicht eindeutig offenlegt. Es wird zwischen zwei Arten von synthetischen Datensätzen unterschieden:

- Vollständig simulierte Daten oder künstliche Daten, die nur die Merkmale der Originaldaten nachbilden. Solche synthetischen Datensätze haben die gleichen Metadaten (Variablennamen, Variablenarten, mögliche Werte usw.) wie die identifizierenden Daten, enthalten aber ansonsten ausschliesslich zufällig generierte Informationen, die hauptsächlich zum Testen von Datenanalyse-Pipelines verwendet werden.
- Die Daten werden so erzeugt, dass sie bestimmte statistische Eigenschaften des Original-Datensatzes nachbilden. Somit werden die Daten trotz der teilweise stochastischen Methode nicht rein zufällig erzeugt. Da in diesem Fall Informationen zum Datensatz geteilt werden, ist das Risiko, dass identifizierende Daten offengelegt werden, nicht null und muss kontrolliert werden.

Die Erzeugung synthetischer Daten der ersten Kategorie basiert in der Regel auf einfachen Zufallszahlengeneratoren. Bei der zweiten Kategorie kommen komplexere Algorithmen aus der angewandten Mathematik sowie Machine-Learning- und KI-Algorithmen zum Einsatz (Random Forest, neuronale Netze, Generative Adversarial Networks usw.).

Grenzen und Risiken

Jeder synthetische Datensatz reproduziert bestimmte Eigenschaften des Original-Datensatzes. Die Datenkonsumierenden müssen daher transparent über Art und Umfang dieser Eigenschaften informiert werden, damit sie verstehen, in welchem Kontext und unter welchen Voraussetzungen der Datensatz zuverlässig verwendet werden kann. Im Rahmen eines Forschungsprojekts zum Beispiel muss ein synthetischer Datensatz garantieren, dass die Erkenntnis nicht auf den zufällig generierten Daten, sondern auf einem Signal in den identifizierenden Originaldaten beruht.

Die meisten Generatoren synthetischer Daten bieten keine formelle Garantie, dass diese nicht offengelegt werden. Die jüngsten Entwicklungen in diesem Bereich haben sogar gezeigt, dass solche Generatoren das gleiche Offenlegungsrisiko aufweisen wie jedes andere Datenprodukt. Dieses Problem kann überwunden werden, indem diese Generatoren mit einer Technologie zur Output-Privacy-Kontrolle wie Differential Privacy kombiniert werden.

Versprechen, dass ein Datensatz sämtliche Eigenschaften eines Originaldatensatzes reproduziert, ohne identifizierende Daten offenzulegen, beruhen auf einem Trugschluss und sind technisch nicht realisierbar.

Stärken und Ausgereiftheit

Synthetische Datensätze bieten die Möglichkeit, Mikrodaten bereitzustellen, insbesondere zuhanden der Öffentlichkeit, und dabei gleichzeitig das Risiko zu begrenzen, dass identifizierende Daten offengelegt werden. Diese Lösung lässt sich flexibel anwenden, einfach umsetzen und bietet zahlreiche Möglichkeiten, identifizierende Daten einer Verwaltung zu nutzen. Allerdings müssen die Generatoren synthetischer Daten dazu mit einer Technologie der Output-Privacy-Kontrolle kombiniert werden. Die Erzeugung synthetischer Daten ist ausgereift und es gibt zahlreiche Open-Source-Implementierungen, die als R- oder Python-Paket verfügbar sind. Simulierte Daten können in einer Verwaltung problemlos produktiv eingesetzt werden. Bei der zweiten Art synthetischer Daten, die bestimmte statistische Eigenschaften der identifizierenden Daten reproduzieren, muss die wissenschaftliche Forderung hingegen noch klären, wie zuverlässig die generierten Daten wirklich sind. Durch diese Unsicherheit ist ihre Verwendung in einer Verwaltung momentan noch nicht spruchreif.

4 Fallbeispiel 1: Plattform für eine vertrauliche Anwendung der Datenwissenschaft ohne Datenübermittlung

4.1 Hintergrund und Ziele

Die verschiedenen Bundesorgane verfügen über eine grosse Menge an Daten, die zur Ausübung ihrer Aufgaben für das öffentliche Interesse unverzichtbar sind. Die Bundesverwaltung investiert viel Zeit und Ressourcen in die Erhebung, Auswahl und Konsolidierung dieser Datensätze. Sie sind eine zuverlässige und einzigartige Informationsquelle, deren gemeinwohlorientiertes Potenzial jedoch noch nicht voll ausgeschöpft wird. Daten von natürlichen oder juristischen Personen dürfen gemäss den Rechtsgrundlagen im Bereich Datenschutz und Statistikgeheimnis in der Regel nur für den Zweck verwendet werden, für den sie erhoben wurden. Ihre Wiederverwendung für nicht personenbezogene Zwecke ist hingegen möglich, namentlich für Forschung, Planung oder Statistik ([Art. 39 Abs. 1 DSG, SR 235.1](#)). Solche Anfragen werden genau geprüft, um ihre Legitimität und das Offenlegungsrisiko abzuklären. Entsprechend lang und komplex ist der administrative Prozess für die Erteilung der Datenzugriffsrechte.

Das Potenzial der von der Bundesverwaltung erhobenen Daten kann folglich nur ausgeschöpft werden, wenn diese in verschiedenen Bereichen wiederverwendet werden.

- Erstens könnten andere Bundesämter und Verwaltungseinheiten die Daten nutzen, um die öffentliche Politik gestützt auf objektive Indikatoren effizient mitzugestalten. Das BFS zum Beispiel könnte Einzeldaten zur Demografie und zu Todesfällen liefern, damit das Bundesamt für Gesundheit (BAG) seine Entscheidungen proaktiv und möglichst realitätsnah anpassen kann. Im Bildungsbereich könnten die Daten des BFS zu den Bildungsverläufen der Studierenden für die Kantone von Interesse sein, um den gesamtschweizerischen Bildungsbedarf möglichst genau einzuschätzen.
- Zweitens könnten Forschungslabore die Daten verwenden, um ihren Wissensstand zu erweitern und die Wettbewerbsfähigkeit der Schweiz zu stärken. Hochschulforschende könnten die Daten der Bundesämter zur Erarbeitung innovativer Lösungen gegen Armut und Schulabbrüche analysieren.
- Drittens könnte die Privatwirtschaft die Daten nutzen, um die Wirtschaftskraft der Schweiz zu fördern. Schweizer Unternehmen könnten beispielsweise die Daten der Bundesämter zur Ergänzung ihrer eigenen Daten verwenden, um strategische Entscheide für Investitionen und Produktentwicklungen zu treffen.
- Viertens könnten die Daten bei globalen Krisen wie der Covid-19-Pandemie auch gemeinsam mit anderen Staaten genutzt werden. Zur Koordination einer effizienten globalen Reaktion auf eine Pandemie könnte die Schweiz beispielsweise ihre epidemiologischen Daten anderen Ländern zur Verfügung stellen.

Diese Wiederverwendung der Daten muss jedoch zwingend politisch und gesellschaftlich akzeptiert sein, d. h. es muss sichergestellt werden, dass ihre Vorteile nicht auf Kosten der Schweizer Bevölkerung, Haushalte und Unternehmen gehen. Ansonsten besteht die Gefahr, dass das Vertrauen in die Bundesstatistik oder die Bereitschaft, an Erhebungen teilzunehmen, deutlich sinken. Um das Vertrauen in den Staat zu wahren, muss das Offenlegungsrisiko bei der Verwendung von identifizierenden Daten innerhalb und ausserhalb der Bundesverwaltung zwingend eingeschränkt werden.

Das hier präsentierte Fallbeispiel untersucht die mögliche Wiederverwendung von Daten bei gleichzeitiger Kontrolle des Offenlegungsrisikos. Wichtig ist dabei, dass die Daten von einem vertrauenswürdigen zentralen Dienst erhoben werden. Mit dem Einsatz von PET wird dafür gesorgt, dass das Vertrauen der Personen, die dem Dienst ihre Daten bereitstellen, bei einer möglichen Wiederverwendung der Daten gewahrt oder sogar gestärkt wird. Dieses Vertrauen in den zentralen Dienst sicherzustellen ist Aufgabe der interdisziplinären Gouvernanz. Dazu braucht es neben technischen auch rechtliche und organisatorische Massnahmen. Auf die Koordination dieser Massnahmen wird in Kapitel 6.3 eingegangen.

Das Projekt besteht in einer Plattform, auf der sich Personen mit den entsprechenden Rechten einloggen können, um unter Wahrung der Offenlegungskontrolle Datenprodukte zu erzeugen, ohne dass sie zu irgendeinem Zeitpunkt Einsicht in die Originaldaten erhalten. Zwischen Dateneinhabern und Datennutzenden werden somit keine identifizierenden Informationen ausgetauscht. Die Plattform verfolgt zwei übergeordnete Ziele: Zum einen soll sie die Möglichkeit bieten, identifizierende Daten auch dann zu nutzen, wenn die Wiederverwendung als zu riskant eingestuft oder die Weitergabe gesetzlich verboten ist, zum anderen dafür sorgen, dass rasch Algorithmen zur Datenanalyse entwickelt werden können, auch wenn noch kein formeller Datenaustauschprozess festgelegt wurde. Mit PET wie Differential Privacy lässt sich das Offenlegungsrisiko unabhängig von der Verwendung der identifizierenden Daten feinmaschig kontrollieren. Folglich sorgt die Plattform dafür:

- das Vertrauen der Bevölkerung in den Staat aufrechtzuerhalten, indem sichergestellt wird, dass das Risiko negativer Auswirkungen für Einzelpersonen bei der Wiederverwendung identifizierender Daten begrenzt ist;
- einen schnelleren Zugriff auf geschützte Daten für andere Zwecke als den ursprünglichen Erhebungszweck zu gewährleisten. In diesem Fall kann die Plattform so lange als Übergangslösung dienen, bis das Forschungslabor Zugang zu den Originaldaten erhält. Zur Wahrung der Vertraulichkeit wird den Ergebnissen der Algorithmen mit Differential-Privacy-Methoden kontrolliertes Rauschen hinzugefügt. Bei einer hohen Menge solcher Störungen ist das Offenlegungsrisiko sehr gering und der Zugriff kann mit einer minimalen Rechtsgrundlage erlaubt werden. Die Ergebnisse sind hingegen mit einem hohen Mass an Unsicherheit behaftet, was ihren Nutzen entsprechend einschränkt. Den Forschenden bietet diese Lösung jedoch die wertvolle Möglichkeit, den Code und das Testprotokoll zu testen und so früher mit der Arbeit zu beginnen.

Die Plattform funktioniert ähnlich wie ein Remote Access Service, unterscheidet sich jedoch von herkömmlichen Lösungen vor allem dadurch, dass die identifizierenden Daten von den Nutzenden garantiert nie eingesehen werden können und dass die von ihnen ausgeführten Algorithmen nicht von einer Person überprüft werden müssen. Bei Remote-Access-Plattformen müssen Fachpersonen im Bereich der Offenlegungskontrolle die Ergebnisse systematisch kontrollieren, was mit einem hohen Kosten- und Zeitaufwand verbunden ist. Mit Differential Privacy kann dank einer zeit- und technologiebeständigen mathematischen Quantifizierungsmethode der gesamte Prozess unter Wahrung der Offenlegungskontrolle automatisiert werden.

Um das Potenzial dieses Fallbeispiels in der Bundesverwaltung zu veranschaulichen, werden nachfolgend einige konkrete Anwendungen beim Eidgenössischen Departement des Innern beschrieben.

- Sterblichkeitsstatistik: Die vom BFS für die Sterblichkeitsstatistik erhobenen Daten sind auch für andere Zwecke nützlich. Auf ihrer Grundlage lässt sich z. B. die Auswirkung bestimmter Faktoren wie der Covid-19-Pandemie auf die Sterblichkeit schätzen. Die Universität Mailand hat mit diesen schützenswerten Einzeldaten eine Studie über die Schweizer Bevölkerung durchgeführt. Dazu wurde wie vorgeschrieben ein Datenlieferungsvertrag aufgesetzt und unterzeichnet, mit dem die Verantwortung für den Datenschutz der antragstellenden Universität übertragen wurde. Das Reputationsrisiko für die Bundesverwaltung im Falle eines Datenlecks auf Seite der Universität bleibt aber bestehen. Mit der in diesem Fallbeispiel beschriebenen Plattform könnte dieses Risiko kontrolliert werden. Gleichzeitig würde sie den Verwaltungsaufwand für die Forschung reduzieren und damit die Durchführung solcher Studien deutlich vereinfachen.
- Schweizerische Lohnstrukturerhebung: Wie die umfassende Berichterstattung in den Medien zeigt, hat die Öffentlichkeit grosses Interesse an den Daten dieser Erhebung. Das BFS hat daher neben der eigentlichen Statistikproduktion den Lohnrechner Salarium entwickelt. Mit dieser Webanwendung kann nach bestimmten Kriterien wie Region, Geschlecht, Wirtschaftszweig und Wochenstunden der Monatslohn berechnet werden. Die Ergebnisse dieser im übrigen meistbesuchten BFS-Seite werden beispielsweise für individuelle Lohnverhandlungen, Lohnverhandlungen zwischen Sozialpartnern und zur Überprüfung der branchenüblichen Löhne verwendet. Sie

sind von unschätzbarem Wert für die Schweizer Gesellschaft, die direkt von ihrer Wiederverwendung, z. B. in der Forschung, profitiert. Auf ihrer Grundlage könnten sozial- und wirtschaftspolitische Massnahmen beurteilt und effizient und gerecht angepasst werden. Das hier präsentierte Fallbeispiel würde die Ausschöpfung dieses Potenzials unter Wahrung der Privatsphäre der Bürgerinnen und Bürger vereinfachen und beschleunigen und so die Akzeptanz für die Wiederverwendung der Daten erhöhen.

- Programm DigiSanté: Mit diesem Programm wird die digitale Transformation im Gesundheitswesen gefördert. Das vierte Massnahmenpaket verfolgt zwei Ziele: Es soll einerseits für eine effizientere Nutzung der von der öffentlichen Verwaltung gehaltenen Daten sorgen, damit diese zur besseren Planung und Steuerung des Gesundheitswesens durch die Behörden und die beteiligten Akteurinnen und Akteure eingesetzt werden können, und andererseits den Zugang zu diesen Daten für Forschende aus dem akademischen und privatwirtschaftlichen Bereich vereinfachen. Mit der in diesem Kapitel präsentierten Plattform lassen sich diese Herausforderungen technisch umsetzen. Sie garantiert die Vertraulichkeit der Gesundheitsdaten bei ihrer Wiederverwendung und schafft damit einen «Datenraum für gesundheitsbezogene Forschung».

Mit anderen Worten: Die Plattform erhöht automatisch das Nutzungspotenzial von Daten der Bundesverwaltung und garantiert die Vertraulichkeit der identifizierenden Daten. Damit begrenzt sie das Reputationsrisiko für die Bundesverwaltung und negative Auswirkungen für die Bürgerinnen und Bürger.

4.2 Grundsatz

Parteien und Vertrauensschema

Die Architektur der Plattform ist so gestaltet, dass die spezifischen Merkmale des Vertrauensverhältnisses innerhalb der öffentlichen Verwaltung bestmöglich genutzt werden können.

- Die identifizierbaren Einheiten übermitteln ihre Daten im Rahmen eines gesetzlich verankerten Vertrauensverhältnisses an den Datenanbieter, in der Regel ein Bundesamt.
- Da jedes Bundesamt einen vertrauenswürdigen Rechenanbieter wie das Bundesamt für Informatik und Telekommunikation beauftragen kann, erübrigt sich der Einsatz von Technologien zur Sicherstellung der Input Privacy. Der Rechenanbieter ist für die Datensicherheit verantwortlich und muss somit die nötigen Massnahmen ergreifen, um unberechtigte Zugriffe auf die Daten und die Plattform zu verhindern. Zudem ist er für die Identifizierung und Legitimierung der Plattformnutzenden zuständig.
- Datenkonsumierende können Bundesangestellte anderer Bundesämter oder die breite Öffentlichkeit sein. Es wird in jedem Fall davon ausgegangen, dass zwischen dem datenanbietenden Amt und den Datenkonsumierenden kein Vertrauensverhältnis besteht. Folglich muss zwingend eine Technologie zur Output-Privacy-Kontrolle (insbesondere Differential Privacy) verwendet werden, um das Offenlegungsrisiko zu verringern.

Dieses Vertrauensmodell wird in der öffentlichen Verwaltung, die über direkt vom Bund verwaltete Rechenkapazitäten verfügt, häufig verwendet.

Beschreibung des Dienstes

Abbildung 4 veranschaulicht die Funktionsweise des Dienstes. Sie zeigt, wie die Daten einer autorisierten Nutzerin oder einem autorisierten Nutzer zur Verfügung gestellt werden. Aus Sicht der Nutzerin oder des Nutzers funktioniert der Dienst wie folgt:

- 1) Die Nutzerin oder der Nutzer verbindet sich aus der Ferne mit der Plattform und erhält so Zugriff auf eine virtuelle Umgebung, auf der die Programmiersprachen Python und R ausgeführt werden. Da die meisten Datenanalysen im privatwirtschaftlichen und akademischen Bereich auf diesen Sprachen beruhen, ist ihre Beherrschung unverzichtbar für die Nutzung der Plattform.
- 2) Die Nutzerin oder der Nutzer kann in einer Bibliothek einen umfassenden Katalog der verfügbaren Datensätze sowie deren Art und Darstellung (Metadaten) einsehen. Dort ebenfalls aufgeführt ist das jeweils zugewiesene Privacy-Loss-Budget.
- 3) Die Nutzerin oder der Nutzer wählt den gewünschten Datensatz aus und lädt einen vollständig simulierten künstlichen Datensatz in die eigene Rechenumgebung hoch, der die Metadaten des Original-Datensatzes getreu wiedergibt. Mithilfe dieses künstlichen Datensatzes können Algorithmen zur Datenanalyse entwickelt und getestet werden.
- 4) Anschliessend kann die Nutzerin oder der Nutzer den fertigen Algorithmus an einen Remote-Server der Bundesverwaltung übermitteln. Bevor der Algorithmus ausgeführt wird, wird geprüft, ob er einen Mechanismus zur Offenlegungskontrolle enthält. Fehlt ein solcher Mechanismus, wird die Ausführung des Algorithmus verweigert. In einem nächsten Schritt überprüft der Server die Zugriffsrechte der Nutzerin oder des Nutzers für den Datensatz und die Angemessenheit des zugewiesenen Privacy-Loss-Budgets.
- 5) Sind die Voraussetzungen in Bezug auf Datenzugriff und Vertraulichkeit erfüllt, lädt der sichere Server die identifizierenden Daten von einem sicheren Speicherplatz herunter.
- 6) Anschliessend wird der Algorithmus selbstständig auf den Original-Datensatz angewendet und das Ergebnis mithilfe von Differential Privacy unter Wahrung der Datenvertraulichkeit berechnet.
- 7) Das geschützte Ergebnis wird gespeichert und an die Nutzerin bzw. den Nutzer übermittelt. Durch die Speicherung kann die Nutzerin oder der Nutzer auch auf frühere Ergebnisse zugreifen, ohne dass das Privacy-Loss-Budget für die gleiche Analyse belastet wird. Die Behörden haben durch die Archivierung der Algorithmen und der entsprechenden Ergebnisse die Möglichkeit, die Aktivität der Nutzenden und somit die Verwendung der Datensätze zu überwachen.

Mit dieser Methode können Nutzende Algorithmen ohne direkten Zugriff auf identifizierende Daten ausführen und dem datenliefernden Amt garantieren, dass das Offenlegungsrisiko begrenzt ist.

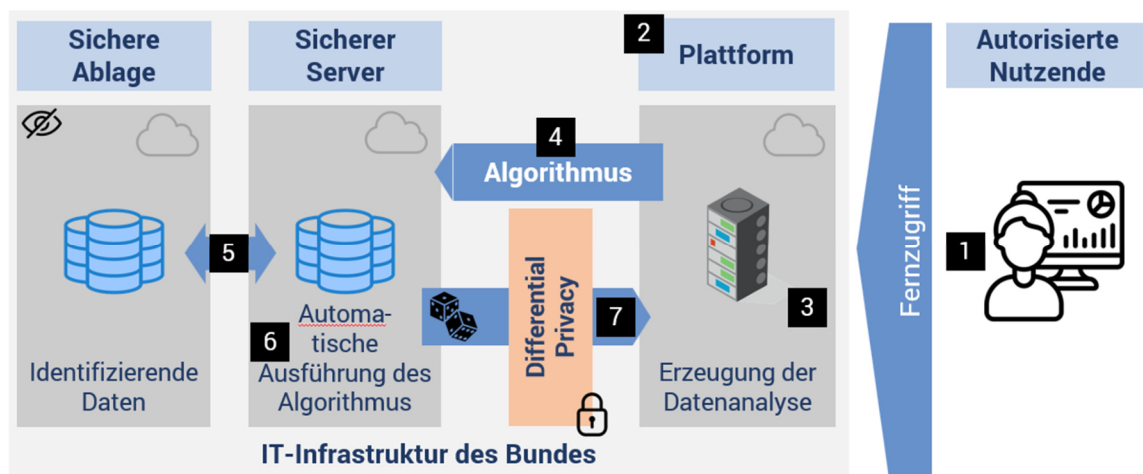


Abbildung 4: Funktionsschema einer Plattform für eine vertrauliche Anwendung der Datenwissenschaft ohne Datentransfer

Damit eine Datenanalyse zulässig ist, muss zwingend ein Mechanismus zur Offenlegungskontrolle wie ein auf Differential Privacy basierender Randomisierungsalgorithmus integriert sein. Die Plattform entwickelt diese Algorithmen nicht intern, sondern zieht dazu anerkannte Open-Source-Bibliotheken wie «OpenDP», «SmartnoiseSQL» zur Datenmanipulation und -transformation sowie «DiffPrivLib» für Machine Learning und künstliche Intelligenz hinzu. Dies sind nur einige Beispiele.

Der sichere Server ist direkt mit dem Speicherplatz der Datensätze verbunden. Im Zuge der zunehmenden Verbreitung des Once-Only-Prinzips entstehen immer neue Kataloge, in denen die Datensätze der Bundesverwaltung und vor allem die Metadaten enthalten sind. Ein Beispiel dafür ist die I14Y-Norm: Sie vereinfacht die Erzeugung künstlicher Datensätze, mit denen die von den Nutzenden für die Datenanalyse erstellten Algorithmen überprüft werden. Dank der direkten Integration mehrerer solcher Kataloge in die Plattform, wie sie u. a. im Rahmen des «fso-metadata»-Pakets des BFS möglich ist, entstehen Synergieeffekte, sodass sich das Potenzial der Daten der Bundesverwaltung voll ausschöpfen lässt.

Um die Machbarkeit einer solchen Plattform aufzuzeigen, hat das Kompetenzzentrum für Datenwissenschaft des Bundes (DSCC) eine **Open-Source**-Implementierung entwickelt, die auf seinem GitHub²⁰ verfügbar ist. **Bei der Entwicklung dieser Implementierung wurde darauf geachtet, dass sie sich auf eine Standard-Informatikstruktur übertragen lässt. Sie steht somit allen Einheiten der Bundesverwaltung zur Verfügung, die mit solchen Diensten experimentieren oder sie in Produktion setzen möchten.** Die technischen Details dazu sind einem Bericht enthalten (<https://arxiv.org/abs/2406.17087>).

4.3 Grenzen und Herausforderungen

Beim vorliegenden Fallbeispiel werden die besonderen Merkmale des Vertrauensverhältnisses innerhalb der Bundesverwaltung zum Vorteil genutzt. Da mit einem vertrauenswürdigen Rechenanbieter zusammengearbeitet wird, kann der Fokus einzig und allein auf die Vertraulichkeit der Ergebnisse gelegt werden. Diese Vertraulichkeit wird durch den Einsatz von Differential Privacy sichergestellt, deren grösste Herausforderung in der Bestimmung des Privacy-Loss-Budgets besteht. Wie erläutert geht es beim Privacy-Loss-Budget darum, den Kompromiss zwischen dem Nutzen der Daten und deren Offenlegungsrisiko präzise festzulegen. Dazu wird für jede Nutzerin und jeden Nutzer ein maximal zulässiger Vertraulichkeitsverlust definiert. Wie hoch bzw. tief die zulässige Schwelle angesetzt wird, ist nicht zuletzt eine politische Frage. Diese lässt sich wie folgt formulieren: «Wie hoch darf das Risiko für die Datenschutzeinheiten des Datensatzes maximal sein, damit es noch akzeptabel ist?» oder «Wie gering darf der Nutzen sein, damit die Zuverlässigkeit des Datensatzes gewährleistet ist?». Diese beiden Fragestellungen entsprechen den Grundsätzen «Confidentiality first» und «Utility first». Sie ermöglichen es, Prioritäten zu setzen und das Privacy-Loss-Budget bewusst und nachvollziehbar festzulegen.

Ein sprechendes Beispiel ist die Datennutzung in der Forschung: Der aktuelle Prozess für den Zugriff auf identifizierende Daten ist langsam und schwerfällig. Er setzt einen formellen Rechtsrahmen voraus, der die Verantwortung für den Datenschutz auf das Forschungslabor überträgt. Dieser Prozess ist mit einem hohen Verwaltungsaufwand verbunden und verlangsamt die Forschung erheblich, was in gewissen Situationen wie der Covid-19-Pandemie problematisch sein kann. Die beiden folgenden Anwendungen zeigen exemplarisch, wie die Plattform in solchen Fällen effizient genutzt werden kann:

- Vereinfachung und/oder Vorarbeiten: Eine Forscherin möchte so schnell wie möglich auf die Daten zugreifen, damit sie mit den Forschungsarbeiten beginnen kann. Die Plattform bietet ihr die Möglichkeit, die Machbarkeit der Analysen zu überprüfen und, bei schlüssigem Ergebnis, die

²⁰ <https://github.com/dsc-admin-ch/lomas>

Algorithmen vorzubereiten und zu testen, bis der Verwaltungsprozess für den Zugriff auf die Originaldaten abgeschlossen ist. Die Vorteile liegen auf der Hand.

- Für die Forscherin: Das Forschungsprojekt schreitet schneller voran.
- Für die Bundesverwaltung: Wenn die Abklärungen ergeben, dass die Analysen nicht durchführbar sind, werden Ressourcen gespart.
- Ermöglichen: Es werden Daten bereitgestellt, die zuvor als zu sensibel eingestuft wurden, aber jetzt genutzt werden können, da sie nie in ihrer ursprünglichen Form offengelegt werden und das Offenlegungsrisiko unter Kontrolle ist. Dadurch ergeben sich folgende Vorteile:
 - Für die Forscherin: Sie kann Daten analysieren, die sonst nicht verfügbar wären.
 - Für die Bundesverwaltung: Das Potenzial der erhobenen Daten kann voll ausgeschöpft werden. Da sich der finanzielle Aufwand für die Erhebung und Konsolidierung der Datensätze nicht nur für das zuständige Organ lohnt, resultiert zudem ein Reputationsgewinn.

Bei diesen beiden Anwendungsbeispielen braucht es nur ein geringes Privacy-Loss-Budget, um das Risiko auszugleichen, sei es wegen der Datensensibilität, oder weil sie nur für kurze Zeit verwendet werden und umfassende Kontrollen der möglichen Nutzenden bzw. die Unterzeichnung formeller Verträge unter diesen Umständen schwierig wären.

Obwohl das Offenlegungsrisiko für jede Nutzerin und jeden Nutzer mithilfe des Privacy-Loss-Budgets bewertet und kontrolliert wird, gibt es keine technische Massnahme, die verhindern kann, dass mehrere Nutzende ihre Budgets kombinieren und so identifizierende Informationen aus der Plattform extrahieren. Eine solche Kollusion muss durch rechtsverbindliche Regeln zur Nutzung der Plattform ausdrücklich untersagt werden. Es kann also nicht komplett auf einen rechtlichen Rahmen für die Nutzung der Plattform verzichtet werden. Solange das Privacy-Loss-Budget für die Nutzenden begrenzt bleibt, ist der Rechtsrahmen vermutlich weniger restriktiv als die Vereinbarungen zur Weitergabe identifizierender Daten. Dadurch werden Ressourcen eingespart und den Nutzenden wird der Zugriff auf die Daten vereinfacht. In der Bundesverwaltung können die Identitätskontrolle und die Rückverfolgbarkeit der Nutzenden mithilfe von Authentifizierungsprotokollen wie eIAM sichergestellt werden.

5 Fallbeispiel 2: Sichere Datenanreicherung durch vertrauliche Verknüpfung

5.1 Hintergrund und Ziele

Systematische Erhebung identifizierender Daten und ihre Grenzen

Heute werden in allen gesellschaftlichen Bereichen systematisch Daten erhoben. Die Art und Weise, wie Informationen gesammelt und verwendet werden, hat sich stark verändert. In der Privatwirtschaft ist die systematische Datenerhebung insbesondere bei Banken stark verbreitet. Sie erfassen minutiös Daten zu den Finanztransaktionen und den Konsumgewohnheiten der Kundinnen und Kunden. Auch die Anbieter von «Smart Devices» erheben intensiv Daten, insbesondere in Bereichen wie personalisierter Gesundheit, in der vernetzte Geräte in Echtzeit die physiologischen Eigenschaften der Personen überwachen und aufzeichnen. Ebenfalls von diesem Trend betroffen sind Sektoren wie die Energieeffizienz, in der intelligente Sensoren den Energieverbrauch in Wohn- und Geschäftsgebäuden messen und analysieren, oder die Mobiltelefonie, in der die Betreiber Daten zur Personenmobilität erfassen.

Im öffentlichen Sektor und insbesondere im medizinischen Bereich, werden ebenfalls systematisch Daten erhoben, allerdings erfolgt dort die Datenerhebung in einem strengen Rechtsrahmen und setzt oft die ausdrückliche Einwilligung der betroffenen Personen voraus (Art. 7 Abs. 1 Humanforschungsgesetz, HFG). Hospitalisierte Patientinnen und Patienten werden beispielsweise immer häufiger gefragt, ob sie der Verwendung ihrer Daten für klinische oder epidemiologische Studien zustimmen.

Nichtregierungsorganisationen (NGO) wiederum erheben in grossem Ausmass Daten zu humanitären und Entwicklungszwecken. Sie erfassen Informationen zu Migrationsströmen, um die Bevölkerungsbewegungen besser zu verstehen und die Hilfs- und Schutzbedürfnisse der Migrantinnen und Migranten zu ermitteln. Zudem messen sie die wirtschaftliche und gesellschaftliche Entwicklung in Entwicklungsländern, um die Wirkung der Hilfsprogramme zu beurteilen und die Entwicklungspolitik zu steuern. Sie beteiligen sich auch an der Bewältigung internationaler Krisen wie Epidemien und Naturkatastrophen, indem sie anhand der erhobenen Daten den humanitären Bedarf ermitteln und die Nothilfemassnahmen koordinieren.

Hierbei ist allerdings zu berücksichtigen, dass solche Datenerhebungen unweigerlich mit Verzerrungen, Doppelspurigkeiten und allgemein mit Ungenauigkeiten behaftet sind. Aus diesem Grund ist die Verwendung dieser Daten in sensiblen Bereichen, d. h. zu anderen Zwecken als Werbeabsichten, heikel und erfordert eine gründliche Analyse und genaue Modellierung sowie Datensicherheits- und Datenschutzmassnahmen.

Auf der anderen Seite investiert die Bundesverwaltung viel Zeit und Ressourcen in die Erhebung, Auswahl und Konsolidierung zuverlässiger und vertrauenswürdiger Datensätze. Sie sind eine einzigartige Informationsquelle, die nur mit bestimmten klinischen Studien im Gesundheitsbereich vergleichbar sind. Dennoch sind einige für die Funktionsweise der Bundesverwaltung zentrale Informationen nicht verfügbar oder werden erst gar nicht erhoben, da die gesetzliche Grundlage dafür fehlt oder ihre Erhebung von der Gesellschaft aufgrund des zu hoch eingestuften Risikos für Einzelpersonen nicht akzeptiert wird. Nur wenn die Daten der verschiedenen Bundesorgane, aber auch der Forschung, der NGO und der Privatwirtschaft mit jenen des öffentlichen Sektors angereichert werden, lässt sich ihr Potenzial voll ausschöpfen. Dazu müssen die Daten verknüpft, d. h. zwei oder mehr Datenquellen miteinander kombiniert und Informationen zur gleichen identifizierbaren Person zusammengeführt werden. Die Ergebnisse dieser Datenverknüpfung können entweder analysiert oder für das Training von Machine-Learning- oder KI-Modellen verwendet werden. Wichtig ist dabei, dass die Vertraulichkeit der identifizierbaren Einheiten, aus denen die Daten stammen, gewährleistet ist.

Nutzungsmöglichkeiten der Datenverknüpfung im öffentlichen Sektor

Die Datenverknüpfung bietet dem öffentlichen Sektor erhebliche Vorteile. Sie ermöglicht die Anreicherung mit Daten anderer Verwaltungen und vereinfacht so die Verknüpfung zwischen Verwaltungseinheiten wie Bundesämtern oder mit anderen Verwaltungsebenen wie Kantonen und Gemeinden. Bundesämter wie das BAG oder das Bundesamt für Umwelt (BAFU) könnten zum Beispiel anhand der vom BFS erhobenen Daten die Wirkung ihrer Gesundheits- oder Umweltpolitik auf die Sterblichkeit untersuchen. Auch der Bundesverwaltung eröffnet die Anreicherung mit externen Daten neue Möglichkeiten. Mithilfe von Daten externer Akteure wie Banken könnten beispielsweise Wirtschafts- und Konsumindikatoren präziser geschätzt werden. Ebenso könnten dank der von den Energieversorgern erhobenen Daten Energieeffizienzprogramme verbessert oder wirksame Regeln bei drohender Energieknappheit erstellt werden. Die Mobilitätsdaten der Mobiltelefonanbieter könnten der Optimierung des öffentlichen Verkehrs und der entsprechenden Infrastruktur dienen und bei nationalen Krisen wie einer Epidemie oder einer Naturkatastrophe wichtige Entscheidungsgrundlagen liefern.

Auch für nichtöffentliche Akteure bietet die Datenverknüpfung grosses Potenzial. In der Forschung zum Beispiel, insbesondere im Gesundheitsbereich werden von den verschiedenen Akteuren grosse Mengen an Daten gehalten, die infolge einer mangelhaften Erhebung verzerrt sein können. Durch die Anreicherung dieser Datensätze mit zuverlässigen Daten wie jenen des BFS können solche Mängel korrigiert werden, sodass die wissenschaftliche Glaubwürdigkeit der Ergebnisse aus den entsprechenden Studien wiederhergestellt wird. Mit dieser Methode lässt sich das Problem der Reproduzierbarkeit, mit der sich die Wissenschaft heute konfrontiert sieht, teilweise lösen. Solche Verknüpfungen sind gemäss Art. 14a des Bundesstatistikgesetzes (BStatG; SR 431.01) für Planung, Statistik und Forschung zulässig, aber administrativ und technisch sehr kostspielig. Mithilfe von PET können Verknüpfungen in grossem Umfang durchgeführt werden, ohne, dass die Bundesverwaltung dadurch zu stark belastet wird.

Internationale NGO und Institutionen der zwischenstaatlichen Zusammenarbeit bemühen sich, das Gemeinwohl über Landesgrenzen hinaus zu fördern. Dem Datenaustausch auf internationaler Ebene oder mit bestimmten zwischenstaatlichen Organisationen sind aufgrund strenger Regelungen aber oft Grenzen gesetzt oder er ist zumindest sehr kompliziert. Für den internationalen Handel oder die Migration wären identifizierende Daten anderer Länder von grossem Nutzen, um auf globale Herausforderungen wie den Klimawandel zu reagieren.

In der Privatwirtschaft würden alle Unternehmen, die systematisch identifizierende Daten erheben, von Verknüpfungen mit den Daten der Bundesverwaltung profitieren, um ihre eigenen Daten anzureichern und zuverlässig zu machen. Dafür bräuchte es allerdings klare Rechtsgrundlagen. Gleichzeitig würde durch die Nutzung der erstellten Datenprodukte die Innovation gefördert und somit die Schweizer Wirtschaft gestärkt.

Verknüpfungen von vornehmlich identifizierenden, schützenswerten Daten müssen politisch und gesellschaftlich akzeptiert sein und es muss sichergestellt werden, dass die Vorteile der Verknüpfung nicht auf Kosten der Schweizer Bevölkerung, Haushalte und Unternehmen gehen. Um das Vertrauen in den Staat zu wahren, muss zum einen sichergestellt werden, dass das Offenlegungsrisiko bei der Verknüpfung von identifizierenden Daten innerhalb und ausserhalb der Bundesverwaltung begrenzt ist, zum anderen, dass keine Rückschlüsse auf die Datenschutzeinheit möglich sind.

Sichere Datenanreicherung durch vertrauliche Verknüpfung

Dieses Fallbeispiel zeigt, wie Daten von Parteien, die nicht berechtigt oder gewillt sind, identifizierende Daten zu teilen, aber trotzdem ein Produkt durch die Kombination von Datensätzen erstellen wollen, ihre Daten miteinander verknüpfen können. Die Verknüpfung erfolgt über eine Plattform, auf die autorisierte Parteien zugreifen können, um ihre Daten bereitzustellen, per Verknüpfung anzureichern und Datenprodukte mit kontrolliertem Offenlegungsrisiko zu erstellen. Die Beteiligten haben während des gesamten Prozesses keinen Zugriff auf die identifizierenden Original-Dateien der anderen Parteien. Mithilfe dieser Plattform soll die Bundesverwaltung die Möglichkeit erhalten, ihre eigenen Datensätze mit Daten aus

externen Quellen anzureichern oder die Datensätze der Bundesverwaltung bereitzustellen, um Daten aus externen Quellen zuverlässiger zu machen.

Sie ermöglicht somit die Zusammenarbeit von Datenanbietern, die sich gegenseitig nicht vertrauen, aber ihre Daten gemeinsam nutzen und ein Datenprodukt aus beiden Quellen erstellen möchten. Das fehlende Vertrauensverhältnis zwischen den Parteien sorgt dafür, dass, bedingt durch die technische Lösung, die verschiedenen Akteure nicht direkt auf die Ergebnisse der Verknüpfung zugreifen können. Die identifizierenden Daten werden nie in ihrer Originalform bereitgestellt. Die konsumierende Partei erhält, nachdem das Offenlegungsrisiko überprüft wurde, lediglich das Datenprodukt, das anhand der Verknüpfungsergebnisse erzeugt wurde. Dazu muss zwingend eine Output-Privacy-Enhancing-Technologie (O-PET), insbesondere Differential Privacy, in Kombination mit einer Input-Privacy-Enhancing-Technologie (I-PET) verwendet werden.

Input Privacy garantiert, dass die beteiligten Parteien nie direkt auf die zur Verknüpfung verwendeten identifizierbaren Originaldateien zugreifen können und ausschliesslich der zwischen den beiden Datenanbietern vereinbarte Code ausgeführt wird. Gleichzeitig lässt sich mithilfe von Output Privacy das Offenlegungsrisiko, das bei der Erstellung des Datenprodukts anhand der angereicherten Daten entsteht, präzise kontrollieren. Output Privacy stellt sicher, dass aus dem Ergebnis der Verknüpfung keine Rückschlüsse auf Einzelpersonen gezogen werden können. Es sind einzig Rückschlüsse auf Personengruppen möglich. Abgeklärt werden muss hingegen, von wem und zu welchem Zweck das Produkt verwendet wird und welchen Nutzen das Gemeinwesen daraus zieht. Die zur Output-Kontrolle eingesetzte Technologie macht jeden Profiling- und Extraktionsversuch für die Gewinnung von Einzeldaten unzuverlässig. Mit der im folgenden Kapitel beschriebenen Architektur können die beiden Technologien (I-PET und O-PET) optimal miteinander kombiniert werden.

Durch ihre Rolle und ihre Architektur erfüllt die Plattform gleich mehrere wichtige Funktionen:

- Sie trägt dazu bei, das Vertrauen der Bevölkerung in den Staat zu wahren, indem sie verhindert, dass die Nutzenden auf die Originaldaten zugreifen können und so das Risiko begrenzt, dass Einzelpersonen durch die Verwendung der identifizierenden Daten zu Schaden kommen.
- Der Staat kann seine Aufgaben durch den Zugriff auf neue Datenquellen effizienter erfüllen.
- Von der Bereitstellung zuverlässiger Daten aus der Bundesverwaltung profitiert die ganze Gesellschaft.

Die Plattform vereinfacht somit die gemeinsame Nutzung der bei den verschiedenen Akteuren der Schweizer Gesellschaft vorhandenen Daten, sodass alle davon profitieren, und garantiert dabei, dass die identifizierenden Einzeldaten vertraulich bleiben und das Risiko negativer Auswirkungen für die Bürgerinnen und Bürger genau kontrolliert wird.

5.2 Grundsatz

Parteien und Vertrauensschema

Entscheidendes Merkmal der Systemarchitektur ist das fehlende Vertrauensverhältnis zwischen den Datenanbietern und Datennutzenden (siehe Abbildung 5).

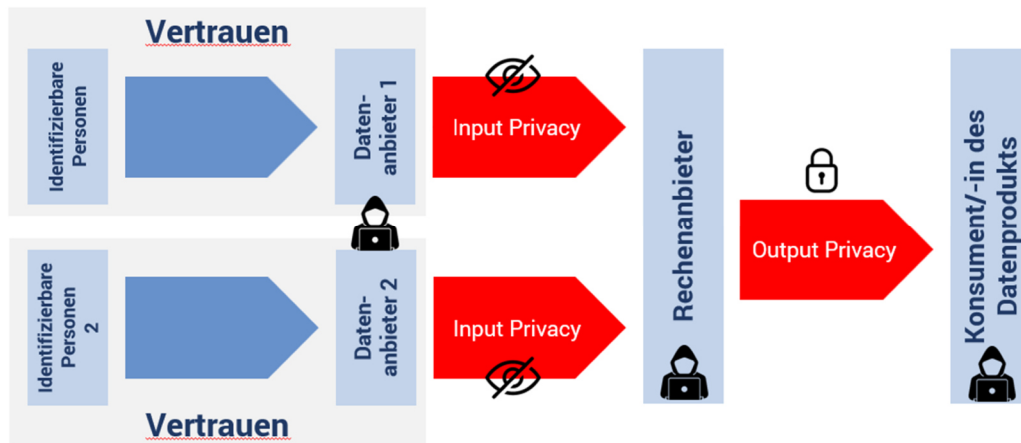


Abbildung 5: Vertrauensschema für das Fallbeispiel «Sichere Datenanreicherung durch vertrauliche Verknüpfung»

Folgende Parteien sind beteiligt:

- 1) Zwei oder mehr Gruppen aus identifizierbaren Personen, deren Daten von mindestens zwei voneinander unabhängigen Datenanbietern erhoben werden: Die Datenerhebungen erfolgen separat, decken nicht die gleichen Informationen ab und umfassen lediglich eine Teilgruppe der gemeinsamen identifizierbaren Personen. Ein Datenanbieter möchte die Informationen aus den beiden Datensätzen kombinieren, um daraus ein mit beiden Quellen angereichertes Datenprodukt zu erstellen.
- 2) Datenanbieter: Bei einem der beiden Datenanbieter handelt es sich um ein Bundesorgan, bei der anderen um einen Akteur aus der Schweizer oder internationalen Gesellschaft. Um die Vertraulichkeit der von den Parteien gehaltenen identifizierenden Daten sicherzustellen, darf keiner von beiden ohne Offenlegungskontrolle auf die identifizierenden Daten des anderen zugreifen können.
- 3) Rechenanbieter: Er darf von keiner der beiden Datenanbietern als vertrauenswürdig eingestuft werden. Angesichts der gesetzlichen Vorschriften und der IT-Sicherheitsmassnahmen gilt in der öffentlichen Verwaltung das Bundesamt für Informatik (BIT) oder jeder andere IT-Dienstleister des Bundes als vertrauenswürdig. Sind hingegen mehrere Parteien beteiligt, steht mindestens einer der beiden Datenanbieter nicht in einem Vertrauensverhältnis zur Bundesverwaltung, d. h., der Rechenanbieter darf technisch nicht in der Lage sein, auf die Daten der Datenanbieter zuzugreifen. Das fehlende Vertrauensverhältnis bedeutet aber nicht zwingend, dass auf die Recheninfrastruktur des Bundes zurückgegriffen werden muss. Die Rechenleistung kann auch an öffentliche Cloud-Anbieter ausgelagert werden. Folglich muss die Input-Privacy-Lösung für dieses Szenario ausreichende Sicherheits- und Vertraulichkeitsgarantien bieten.
- 4) Datenkonsument/-in: In diesem Fall ist mindestens einer der Datenanbieter auch der Konsument des Datenprodukts. Da zwischen den beiden Parteien kein Vertrauensverhältnis besteht, muss das Produkt zum Schutz der **Input Privacy** mit einer PET geschützt werden.

Wahl der PET

Da in diesem Fallbeispiel mehr als zwei Parteien beteiligt sind, kommen nur drei Arten von Input-Privacy-Technologien infrage: Trusted Execution Environment, Secure Multiparty Computation und homomorphe Verschlüsselung. Bei der Wahl der Technologie sind mehrere Kriterien zu beachten, insbesondere die Komptabilität mit der gewählten O-PET (Differential Privacy), die Verfügbarkeit von Open-Source-Lösungen zur Senkung des «Vendor Lock-in»-Risikos sowie die Einfachheit der Implementierung und Bereitstellung des Dienstes. Die Vor- und Nachteile der einzelnen Technologien sind in Kapitel 3 erläutert. Nachfolgend wird die Wahl der Technologien für dieses Fallbeispiel kurz begründet.

- Es gibt noch keine Open-Source-Implementierungen von Secure Multiparty Computation, die reif genug wären für eine solche Anwendung. Ausserdem birgt die Kombination von Differential Privacy mit Secure Multiparty Computation noch einige Schwierigkeiten, insbesondere in Bezug auf die ausgelagerte Multiparty-Generierung von Zufallszahlen. Die wenigen aktuell verfügbaren Techniken befinden sich noch im Forschungsstadium.
- Für die homomorphe Verschlüsselung sind zwar Open-Source-Multiparty-Lösungen vorhanden, damit sie trotz fehlender vertrauenswürdiger Drittpartei mit Differential Privacy kombiniert werden können, bedarf es aber eines Randomisierungsmechanismus, der direkt auf die verschlüsselten Daten angewendet wird, um das Datenprodukt zu entziffern und Differential Privacy anzuwenden. Diese Aufgabe ist jedoch komplex und momentan fehlt es noch an einer ausgereiften Lösung.
- Im Rahmen von TEE könnte der im Fallbeispiel 1 beschriebene Dienst bereits heute umgesetzt werden, ohne dass allzu viele Änderungen nötig wären. Er liesse sich direkt in eine Enklave integrieren und wäre eine einfache Lösung, die Input- und Output-Privacy-Technologien miteinander verbindet. Öffentlich-private Initiativen wie Gramine²¹ bieten zudem Open-Source-Lösungen, die mit verschiedenen Cloud-Anbietern genutzt werden können. In Verbindung mit einem Integritätsnachweis für den ausgeführten Code sind TEE die vollständigste Lösung für die Implementierung dieses Fallbeispiels. Im Gegensatz zu den beiden anderen Technologien beinhalten TEE jedoch keine mathematische Sicherheitsgarantie und setzen voraus, dass dem Hardwarehersteller wie Intel oder AMD vertraut wird.

Für dieses Fallbeispiel bieten sich TEE geradezu an. Sie weisen grosse Synergien mit dem Fallbeispiel im vorangehenden Kapitel auf, dessen Bestandteile nach der Verknüpfung direkt für die Erstellung von Datenprodukten wiederverwendet werden können

Beschreibung des Dienstes

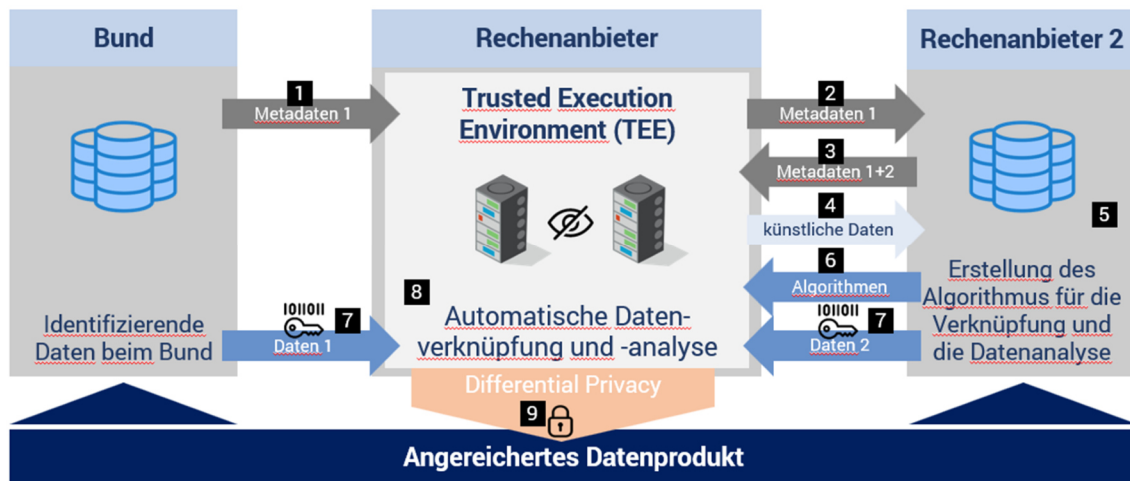


Abbildung 6: Funktionsschema für die Erstellung von Datenprodukten aus Daten, die durch eine vertrauliche Verknüpfung mit Offenlegungskontrolle angereichert wurden

²¹ <https://gramineproject.io/>

Abbildung 6 zeigt die Funktionsweise des Dienstes. Grundsätzlich wird von Folgendem ausgegangen: Anbieter 1 (in Abbildung 6 der Bund) stellt seine Datensätze zur Verfügung, während Anbieter 2 die Erstellung eines Datenprodukts verlangt und dafür auch seine eigenen Daten einbringt. Anbieter 1 und 2 haben vorgängig ein für die Erstellung des Datenprodukts und die Zusammenführung ihrer Datenbanken angemessenes Privacy-Loss-Budget vereinbart. Der Betrieb des Dienstes ist Aufgabe der Bundesverwaltung. Sie ist folglich auch für die erforderlichen Genehmigungen, die Definition des Privacy-Loss-Budgets und die Einrichtung der Enklave zuständig.

Aus Sicht von Anbieter 2 funktioniert der Dienst wie folgt:

- 1) Anbieter 1 lädt seine Metadaten in die TEE hoch.
- 2) Anbieter 2 ruft die Metadaten des von Anbieter 1 bereitgestellten Datensatzes ab.
- 3) Anbieter 2 ergänzt den Metadatenkatalog mit seinen eigenen Metadaten und standardisiert die Kodierung seiner Daten, indem er sie der Kodierung von Anbieter 1 angleicht. Anschliessend lädt Anbieter 2 den konsolidierten Metadatenkatalog in die TEE hoch.
- 4) Anhand des konsolidierten Metadatenkatalogs wird ein künstlicher Datensatz erzeugt, der die Merkmale der Datensätze 1 und 2 nachbildet. Dieser künstliche Datensatz wird an Anbieter 2 übermittelt, damit dieser den Algorithmus für die Datenverknüpfung und die Produkterstellung entwickeln und testen kann.
- 5) Die Regeln für den Vergleich zwischen Datensatz 1 und 2 sowie der Algorithmus für die Datenprodukterstellung anhand des künstlichen Datensatzes werden erstellt.
- 6) Die Verknüpfungs- und Analysealgorithmen werden in die TEE hochgeladen. Die Algorithmen werden in der TEE am künstlichen Datensatz getestet.
- 7) Beide Anbieter laden ihre identifizierenden Daten in die TEE hoch. Dieser Datentransfer ist zum Schutz vor unbefugtem Zugriff verschlüsselt.
- 8) Es wird überprüft, ob für die Ausführung des Algorithmus zur Datenprodukterstellung ein Differential-Privacy-Mechanismus und ein Privacy-Loss-Budget mit ausreichendem Schutz vorhanden sind.
- 9) Die Verknüpfung und die Ausführung des Algorithmus zur Erstellung des Datenprodukts erfolgen automatisch. Anschliessend wird das Produkt verschlüsselt an die Datenanbieter übermittelt.

Schritte 7, 8 und 9 sind ähnlich wie im Fallbeispiel 1 und basieren auf dessen Umsetzung in einer TEE. Mit diesem Vorgehen kann Anbieter 2 anhand des Verknüpfungsergebnisses ein Datenprodukt erstellen, ohne je auf die identifizierenden Daten zuzugreifen, und gleichzeitig Anbieter 1 garantieren, dass das erstellte Datenprodukt nur ein begrenztes Offenlegungsrisiko birgt.

5.3 Grenzen und Herausforderungen

Das hier präsentierte Fallbeispiel beruht weitgehend auf der Harmonisierung und Standardisierung der Metadaten der beiden Datenanbieter. Es wird davon ausgegangen, dass sie von den beiden Datenanbietern gemeinsam genutzt werden und daher die Spaltennamen, die Beschreibungen der Spalteninhalte sowie sämtliche möglichen Werte für den anderen Anbieter ersichtlich sind. Wichtig: Auf dieser Vorstufe ist es nicht möglich, Personendaten zu verknüpfen oder automatisch zu verwenden. Nur, wenn diese Harmonisierung durchgeführt wird, kann das Potenzial der in der Bundesverwaltung vorhandenen Daten voll ausgeschöpft werden. In diesem Zusammenhang kann nicht genug betont werden, wie wichtig Standardisierungsinitiativen wie I14Y sind. PET bieten sich als technische Lösung an, um Daten der Bundesverwaltung nach abgeschlossener Standardisierung unter Wahrung der Privatsphäre wiederzuverwenden. Sowohl PET als auch die Harmonisierung der Metadaten sind für die moderne Datenwissenschaft von grundlegender Bedeutung und somit untrennbar miteinander verbunden. Sie müssen gleichzeitig umgesetzt werden, damit die Daten der Schweizer Gesellschaft vollumfänglich genutzt werden können.

Dieses zweite Fallbeispiel basiert auf den im ersten Fallbeispiel beschriebenen Instrumenten. Es stellen sich somit die gleichen Schwierigkeiten und Herausforderungen, insbesondere in Bezug auf das Privacy-Loss-Budget und die Schaffung entsprechender Rechtsgrundlagen. Neben den rechtlichen und methodischen Voraussetzungen braucht es für die Umsetzung dieses Fallbeispiels spezialisierte Hardware,

d. h. Rechenchips, mit denen die Nutzenden strikt voneinander getrennt werden können (TEE / Confidential Computing). Solche Hardware gehört heute bei den meisten öffentlichen Cloud-Anbietern wie Azure, AWS oder Google Cloud zum Standardangebot. Die IT-Dienste des Bundes bieten sie hingegen noch nicht an. Ihre Einführung wäre mit Investitionen in Infrastruktur und Personalressourcen verbunden. Da die Cloud-Dienste der öffentlichen Anbieter sicher sind, wäre es durchaus denkbar, diese zu nutzen und auf interne Investitionen zu verzichten. Die Entscheidung, ob interne Kapazitäten aufgebaut oder Leistungen öffentlicher Cloud-Dienste in Anspruch genommen werden sollen, ist eine politische Frage, die in Zusammenhang mit der Datengouvernanz und der Datennutzung steht. Um das Potenzial der von der Bundesverwaltung gehaltenen Daten voll auszuschöpfen, muss diese entweder in eine Infrastruktur investieren, die es ihr ermöglicht, die Daten mithilfe der beschriebenen Technologien bereitzustellen, oder akzeptieren, dass sie die Gouvernanz bis zu einem gewissen Grad verliert, und die nötigen Rechtsgrundlagen für die Nutzung öffentlicher Cloud-Dienste wie TEE schaffen. Auch für den Betrieb eines Dienstes wie in diesem Abschnitt beschrieben braucht es neue Rechtsgrundlagen.

6 Rechtsgrundlagen und Gouvernanz

In diesem Kapitel wird aufgezeigt, wie PET im aktuell geltenden Recht geregelt sind und inwiefern sie Denkanstösse für eine einfachere, datenschutzwahrende Wiederverwendung identifizierender Daten liefern können. Die folgende Übersicht der Rechtsgrundlagen erhebt keinerlei Anspruch auf Vollständigkeit. Im Falle einer Umsetzung der in diesem Bericht beschriebenen Anwendungsfälle sind sämtliche Datenschutzbestimmungen anwendbar, einschliesslich der Datenschutzverordnung (DSV, SR 235.11) und der Geheimhaltungsvorschriften, die in den internationalen Instrumenten für Amtshilfe in Steuersachen enthalten sind.

6.1 Bearbeitung identifizierender Daten – Übersicht der Rechtsgrundlagen

Die bestehenden Rechtsgrundlagen für die Bearbeitung identifizierender Daten (siehe Kapitel 2.1) unterscheiden zwischen Daten natürlicher Personen (Personendaten) und Daten juristischer Personen. Zusammenführungen dieser beiden Datenarten werden nur indirekt thematisiert. Die Gesetzgebung stellt den Schutz von Gruppen sicher, indem sie die einzelnen Personen der Gruppe schützt. Technisch braucht es für den Schutz einer Gruppe jedoch spezifische und strengere Massnahmen als für den Schutz der einzelnen Personen in der Gruppe. Aus diesem Grund werden identifizierende Daten in diesem Bericht allgemein gefasst, d. h. als Informationen zu einer beliebigen identifizierbaren Einheit (für eine genaue Definition siehe Kapitel 2.1).

Zu beachten ist auch, dass der Begriff **Bearbeitung** im rechtlichen Sinn (Art. 5 Bst. d DSG) den gesamten Lebenszyklus der Daten und alle Verfahren der datenschutzwahrenden Datenwissenschaft²² umfasst, d. h. das Beschaffen, Speichern, Aufbewahren, Verwenden, Verändern, Bekanntgeben, Archivieren, Löschen und Vernichten von Daten. Dieser Umgang mit Daten in der Bundesverwaltung stützt sich insbesondere auf zwei Grundprinzipien:

- **Anonymisierung identifizierender Daten:** In der geltenden Gesetzgebung wird die Vertraulichkeit identifizierender Daten durch ihre Anonymisierung und Löschung gewährleistet. Anonymisierte Daten fallen nicht in den Geltungsbereich des Datenschutzgesetzes.
- **Primäre Datennutzung:** Der bestehende Rechtsrahmen beschränkt die Verwendung der von der Bundesverwaltung erhobenen Daten auf die gesetzlich vorgesehenen Zwecke. Für jede Datenbearbeitung, deren Zweck klar definiert sein muss, wird eine gesetzliche Grundlage vorausgesetzt. Die Weiterverwendung für andere Zwecke ist grundsätzlich verboten, ausser für Statistik-, Forschungs- und Planungszwecke oder falls eine Rechtsgrundlage geschaffen wird, die die Bearbeitung der zu einem anderen Zweck erhobenen Daten erlaubt.

Die Anonymisierung und die eingeschränkte Wiederverwendung der Daten entsprechen den Grundsätzen der **Rechtsstaatlichkeit** (Art. 5 der Bundesverfassung [BV; SR 101]), nach denen rechtsstaatliches Handeln eine rechtliche Grundlage erfordert, im öffentlichen Interesse liegen und verhältnismässig sein muss. Sie gewährleisten zudem den Schutz der **Privatsphäre**, indem sie jeder Person Anspruch auf Schutz vor Missbrauch ihrer persönlichen Daten zusichern (Art. 13 BV), sowie die Verwirklichung der **Grundrechte** durch den Staat (Art. 35 BV). Folglich müssen sich die Rechtsgrundlagen für die Datenbearbeitung nach den Grundsätzen **Treu und Glauben, Rechtmässigkeit, Verhältnismässigkeit, Zweckbindung und Rechtsgleichheit** richten.

Für die Bearbeitung identifizierender Daten natürlicher Personen, d. h. **Personendaten**, gilt sowohl für private Personen als auch für Bundesorgane das Bundesgesetz über den Datenschutz (DSG, SR 235.1) Das DSG führt die Grundsätze der Anonymisierung und der Zweckbindung wie folgt ein:

²² <https://www.bfs.admin.ch/bfs/de/home/dscc/dscc.html>

- Das DSG regelt ausschliesslich die Bearbeitung von **nicht anonymisierten** Personendaten, d. h. von Angaben, die sich auf eine bestimmte oder bestimmbare Person beziehen (Art. 5 DSG). Anonymisierte Daten fallen demnach nicht in den Geltungsbereich des DSG.
- Das DSG verlangt ausdrücklich, dass für die Bearbeitung von Personendaten durch ein Bundesorgan eine **Rechtsgrundlage** bestehen muss (Art. 34 DSG). Ob es eine solche Rechtsgrundlage im formellen oder materiellen Sinn braucht, hängt namentlich von der Sensibilität der Daten und dem Bearbeitungszweck ab. Führt die Bearbeitung zu einem schwerwiegenden Eingriff in die Grundrechte einer betroffenen Person, braucht es ein formelles Gesetz. Ist sie hingegen für die Aufgabe eines Bundesorgans unentbehrlich, das bereits in einem formellen Gesetz definiert ist, reicht ein Gesetz im materiellen Sinn. Auch die gemeinsame Bearbeitung von Personendaten durch ein Bundesorgan zusammen mit anderen Bundesorganen, kantonalen oder privaten Organen ist Gegenstand des DSG. In diesem Fall regelt der Bundesrat die Modalitäten (Kontrollverfahren und Verantwortung; Art. 33 DSG).

Bei der Bearbeitung von **Daten juristischer Personen** durch ein Bundesorgan kommt eine Übergangsbestimmung des neuen DSG (Art. 71 DSG) zur Anwendung. Grundsätzlich gilt aber das Regierungs- und Verwaltungsorganisationsgesetz (RVOG, SR. 172.010). Diese Gesetze führen die Grundsätze der Anonymisierung und der Primärnutzung wie folgt ein:

- Analog zu Artikel 5 DSG fallen Daten juristischer Personen nicht in den Geltungsbereich des RVOG.
- Gemäss Artikel 57r RVOG muss die Bearbeitung von Daten juristischer Personen, einschliesslich besonders schützenswerter Daten, durch Bundesorgane zweckgebunden sein, d. h., sie müssen der Erfüllung ihrer in einem Gesetz im formellen Sinn umschriebenen Aufgaben dienen.

Ein im Rahmen dieses Berichts besonders relevantes Beispiel für formelle, Bundesorgan-spezifische Rechtsgrundlagen ist die **Bundestatistik**.

- Artikel 18 des BStatG besagt, dass die Datenprodukte **anonymisiert** werden müssen und keine identifizierenden Merkmale zu den befragten Personen, Haushalten, Unternehmen oder Betrieben enthalten dürfen. Bei Verletzungen dieser Grundsätze haftet gemäss Bundesgesetz über die Verantwortlichkeit des Bundes sowie seiner Behördenmitglieder und Beamten (VG, SR 170.32) potenziell der Bund und könnte, wenn die gesetzlichen Voraussetzungen erfüllt sind, strafrechtlich verfolgt werden. Zudem würden solche Verstösse der Reputation und dem Image der Bundesverwaltung schaden.
- Die Bundesstatistik untersteht dem **BStatG** und der Statistikerhebungsverordnung (SR 431.012.1). Die für die Bundesstatistik erhobenen Daten werden in erster Linie für statistische Zwecke verwendet. Es gibt aber eine rechtliche Grundlage, die ihre Wiederverwendung in Form einer Datenlieferung oder einer Verknüpfung für Forschungs-, Planungs- und statistische Zwecke zulässt. Die Anwendung dieser Rechtsgrundlage setzt die Unterzeichnung eines Vertrags voraus, der in der Regel die datenschutzrechtliche Verantwortung auf den Datenempfänger überträgt. Die am 30. April 2025 in Kraft getretene Bundesstatistikverordnung fördert die Mehrfachnutzung der Daten durch die Anwendung des Once-Only-Prinzips, beschränkt sie aber auf Forschungs-, Planungs- und statistische Zwecke.

Im Gesundheitsbereich schützt das **Humanforschungsgesetz** (HFG, SR 810.30) nicht anonymisierte Daten durch ähnliche Anonymisierungs- und Primärnutzungsgrundsätze.

6.2 Grenzen der bestehenden Rechtsgrundlagen

6.2.1 Techniken zur Exfiltration identifizierender Daten: Umgehung der Anonymisierung

Die Zweckmässigkeit der aktuellen Rechtsgrundlagen muss unter Berücksichtigung der jüngsten Fortschritte im Bereich der Techniken zur Extraktion von identifizierenden Daten aus einem Datenprodukt beurteilt werden. Was früher als rein theoretisches Risiko galt, d. h. die Verletzung der Privatsphäre durch die Offenlegung von Attributen wie der Zugehörigkeit zu einer schützenswerten Bevölkerungsgruppe oder die direkte Reidentifizierung einer identifizierbaren Person, ist heute Realität. Die technische Umsetzung solcher Angriffe ist nicht mehr reine Theorie, sondern Realität und sogar in grossem Umfang möglich. Es werden drei Kategorien solcher Angriffe auf die Vertraulichkeit unterschieden.

Die erste Kategorie umfasst die Angriffe durch Verknüpfung, sogenannte Linkage Attacks. Dabei wird ein Satz von Mikrodaten mit einer oder mehreren Quellen von Zusatzinformationen kombiniert, um Informationen über eine identifizierbare Einheit zusammenzuführen und so ihre Reidentifizierung zu ermöglichen. Diese Mikrodaten können entweder illegal durch die Ausnutzung einer Sicherheitslücke beschafft werden oder durch legale Mittel wie einer Lieferung von pseudonymisierten Daten, einer Rekonstruktion oder einer Extraktion von Daten aus einer Publikation. Von solchen Angriffen sind hauptsächlich Mikrodaten betroffen, d. h. direkt mit identifizierbaren Personen verbundene Daten. Angesichts der zunehmenden systematischen Erhebung solcher Daten ist davon auszugehen, dass Schutzmassnahmen, bei denen Art und Ausmass der verfügbaren Zusatzinformationen eingeschätzt werden müssen, früher oder später umgangen werden. Bei einem datenwissenschaftlichen Wettbewerb haben Forschende der Universität Texas in Austin²³ am Beispiel der Reidentifikation des Filmverlaufs auf Netflix gezeigt, wie einfach solche Angriffe sind und dass es unmöglich ist vorauszusehen, wie die Angreifenden die verfügbaren Zusatzinformationen nutzen.

Die Angriffe durch Verknüpfung betreffen fast ausschliesslich Mikrodaten. Datenprodukte, in denen die Daten aggregiert oder verallgemeinert wurden, sind anderen Angriffen ausgesetzt, sogenannten Reconstruction Attacks. Dabei werden Mikrodaten, die zur Erstellung eines Datenprodukts verwendet wurden, extrahiert. Solche Angriffe wurden früher lediglich als potenzielles Risiko in Erwägung gezogen. Heute weiss man, dass sie eine echte Schwachstelle darstellen. Das hat auch die Erfahrung des US Census Bureau²⁴ gezeigt. Dieser Behörde ist es trotz Einsatz herkömmlicher Techniken zur Offenlegungskontrolle wie der Aggregation und Festlegung von Mindestgrössen für die Anzahl Personen einer Gruppe gelungen, anhand von Volkszählungstabellen aus dem Jahr 2010 einen erheblichen Teil der US-amerikanischen Bevölkerung zu rekonstruieren und zu reidentifizieren.

Bei der dritten Art dieser Angriffe handelt es sich um sogenannte Attribute Inference Attacks. Dabei wird nicht die Person reidentifiziert, sondern lediglich die Zugehörigkeit dieser Person zu einer potenzial schützenswerten Gruppe innerhalb eines Datensatzes ermittelt. Attribute Inference Attacks sind besonders attraktiv, da die Erfolgsaussichten deutlich höher sind als bei den beiden anderen Angriffskategorien. Im Gesundheitsbereich ist das Risiko solcher Angriffe überdurchschnittlich hoch, da die Zugehörigkeit zu einer Gruppe von Personen, die an einer seltenen Krankheit leiden, besonders schützenswerte Informationen preisgeben kann.

Indem sie die Offenlegungskontrollen umgehen, offenbaren die drei Angriffskategorien die Verwundbarkeit sämtlicher Datenprodukte. Es gibt keine technische Lösung, die zu 100 Prozent verhindern kann, dass solche Angriffe erfolgreich sind. Aus diesem Grund ist es nicht möglich, die im DSG und in der DSGVO beschriebene Anonymisierung technisch zu garantieren. Sie würde eine unwiderrufliche Verän-

²³ https://www.cs.utexas.edu/~shmat/shmat_oak08netflix.pdf

²⁴ www2.census.gov/library/working-papers/2023/adrm/ces/CES-WP-23-63.pdf

derung der identifizierenden Daten erfordern, was technisch nicht machbar ist. Rechtlich kann die Anonymisierung hingegen garantiert werden, sofern das Erfolgsrisiko solcher Angriffe im Kontext der Datenbearbeitung ausreichend gering eingeschätzt wird.

Bei der Bereitstellung von Datenprodukten, unabhängig davon, ob diese Mikrodaten oder andere aggregierte Daten enthalten, besteht eine erhebliche Gefahr, dass identifizierende Daten offengelegt werden. Dieses Restrisiko kann nur vollständig beseitigt werden, wenn ganz auf die Erstellung und Weiterleitung von Produkten verzichtet wird. Das Offenlegungsrisiko ist somit der Preis, den man bezahlen muss, wenn man die Vorteile von Datenprodukten nutzen will. Es handelt sich um einen Kompromiss zwischen Vertraulichkeit und Nutzen, bei dem das eine nur auf Kosten des anderen möglich ist. Um diesen Kompromiss formell zu quantifizieren, müssen bei der Erstellung eines Datenprodukts das Offenlegungsrisiko und die Vorteile (d. h. der Nutzen) mithilfe von Output-PET wie Differential Privacy gegeneinander abgewogen werden.

6.2.2 Auswirkungen auf die Rechtsgrundlagen zur Reglementierung der von der Bundesverwaltung gehaltenen Daten

Dass die unwiderrufliche Anonymisierung technisch nicht garantiert werden kann, zeigt, dass das Prinzip bereits überholt ist. Folglich ist auch der aktuelle Rechtsrahmen zu eng gefasst, um das Potenzial der bei der Bundesverwaltung vorhandenen Daten vollumfänglich nutzen zu können. Es braucht differenziertere Rechtsgrundlagen, die das verbleibende Offenlegungsrisiko mithilfe von PET wie Differential Privacy ausdrücklich quantifizieren.

Darüber hinaus wird die Wertschöpfung aus den Daten der Bundesverwaltung auch durch die Zweckbindung stark eingeschränkt. Die aktuellen Rechtsgrundlagen bieten nur wenig Handlungsspielraum für die Wiederverwendung dieser Daten mithilfe einer Datenwissenschaft, die den Verhaltenskodex des Bundes für menschenzentrierte und vertrauenswürdige Datenwissenschaft erfüllt (Punkt 3 des Bundesratsbeschlusses vom 2. Dezember 2022; EXE 2022.2746²⁵). Auch bei der gemeinsamen Datenbearbeitung wie im Fallbeispiel 2 dieses Berichts oder im Kapitel «Ganzheitliche Lagebilder in Krisensituationen» auf Seite 17 des Berichts «Konkrete Anwendungsfälle von Datenwissenschaft zum Gemeinwohl entlang des Prozesses der Politikgestaltung» beschrieben (Punkt 2 des Bundesratsbeschlusses vom 2. Dezember 2022; .EXE: 2022.2746²⁶) ist der Handlungsspielraum gering.

Eine Erweiterung des bestehenden Rechtsrahmens in Bezug auf die Zweckgebundenheit wäre dank PET jedoch denkbar. Auf technischer Ebene vereinfachen PET die gemeinsame Nutzung von Daten, ohne dass diese ausgetauscht werden müssen, und gewährleisten die formelle Kontrolle des Offenlegungsrisikos bei der Wiederverwendung der Daten. Sie reduzieren das Risiko, dass identifizierenden Einheiten durch die Verletzung der Privatsphäre bei der Wiederverwendung ihrer Daten Schaden entsteht. PET ermöglichen somit eine datenschutzwahrende Wiederverwendung von Daten für nicht personenbezogene Zwecke durch Bundesorgane im Rahmen von Artikel 39 DSGVO.

PET allein können aber nicht sicherstellen, dass die Daten nicht missbräuchlich verwendet werden, d. h., dass bei ihrer Bearbeitung sämtliche Grundrechte eingehalten werden. Sie bieten auch keine Gewähr, dass ein Projekt, bei dem Daten wiederverwendet werden, ethisch ist und dem Gemeinwohl zugutekommt. Die Erstellung solcher Datenprodukte muss daher in einem breiter gefassten Rechtsrahmen geregelt werden, der nicht nur den Schutz der Privatsphäre, sondern auch die Art der Publikation, die

²⁵ <https://www.bfs.admin.ch/bfs/de/home/statistiken/kataloge-datenbanken.assetdetail.29325685.html>

²⁶ <https://dam-api.bfs.admin.ch/hub/api/dam/assets/29325687/master>

benötigten Daten und den Empfänger berücksichtigt. Wenn die rechtsstaatliche Wiederverwendung dieser Daten gewährleistet werden soll, braucht es einen Gouvernanz-Rahmen, der alle diese Aspekte und die damit verbundenen Risiken abdeckt.

6.3 Die «Five Safes»: Gouvernanz-Rahmen für eine datenschutzwahrende Datenwissenschaft

Da die Anonymisierung technisch nicht garantiert werden kann und die Wertschöpfung aus Daten durch die Zweckbindung eingeschränkt ist, bedarf es eines differenzierteren Gouvernanz-Rahmens, der das gesamte Spektrum an möglichen Eingriffen in die Grundrechte und die Privatsphäre der Personen berücksichtigt. Die Binarität dieser auf Konformität basierenden Grundsätze kann die Komplexität des Offenlegungsrisikos wie wir es heute kennen nicht erfassen und bietet nicht die nötigen Voraussetzungen dafür, dass die Bundesverwaltung ihre Daten zur Bewältigung der aktuellen Herausforderungen auswerten kann. Daher braucht es einen Rahmen, der das gesamte **Spektrum** der Risiken und der zu ihrer Kontrolle eingeführten technischen, organisatorischen und rechtlichen Massnahmen berücksichtigt. Zudem bedarf es Instrumenten, die Projekte zur Wiederverwendung von Daten **interdisziplinär und umfassend** begreifen und bei der Abwägung zwischen Risiko und Nutzen, insbesondere in Bezug auf das Offenlegungsrisiko von identifizierenden Daten, eine fundierte Entscheidung ermöglichen.

Das vor über zwanzig Jahren formulierte «Five-Safes»-Modell kommt in vielen nationalen Statistikämtern und Datenarchivierungsdiensten für die Forschung zum Einsatz. Damit wird die Wiederverwendung von Daten in spezifischen Projekten anhand von fünf Dimensionen geprüft: Art des Projekts, verwendete Daten, Empfänger, Umfeld und schliesslich Ergebnisse. Das Modell deckt sämtliche Risiken ab, die bei der Produktion und Weitergabe von Datenprodukten aus identifizierenden Daten auftreten können. Es wird unter anderem von den nationalen Statistikämtern des Vereinigten Königreichs (ONS UK), Australiens (Australian Bureau of Statistics) und Neuseelands (Stats New Zealand), aber auch vom UK Data Service²⁷, von Research Data Scotland²⁸ und vom Data Innovation Program der kanadischen Provinz British Columbia²⁹ verwendet.

Die «Five Safes» haben sich als Rahmenwerk für die Bearbeitung von schützenswerten identifizierenden Daten bewährt. Sie wurden vor dem Aufkommen von PET formuliert, dank ihrer Vollständigkeit und Flexibilität decken sie aber auch Projekte ab, die PET verwenden. Abbildung 7 zeigt im Detail, wie das Modell in eine menschenzentrierte und vertrauenswürdige Datenwissenschaft eingebunden werden könnte.

²⁷ <https://ukdataservice.ac.uk/help/secure-lab/what-is-the-five-safes-framework/>

²⁸ <https://www.researchdata.scot/our-work/data-explainers/what-is-the-five-safes-framework/>

²⁹ <https://www2.gov.bc.ca/gov/content/data/about-data-management/data-innovation-program/privacy-security>

Five Safes: Gouvernanz-Rahmen für die sichere und ethische Nutzung des Potenzials öffentlicher Daten

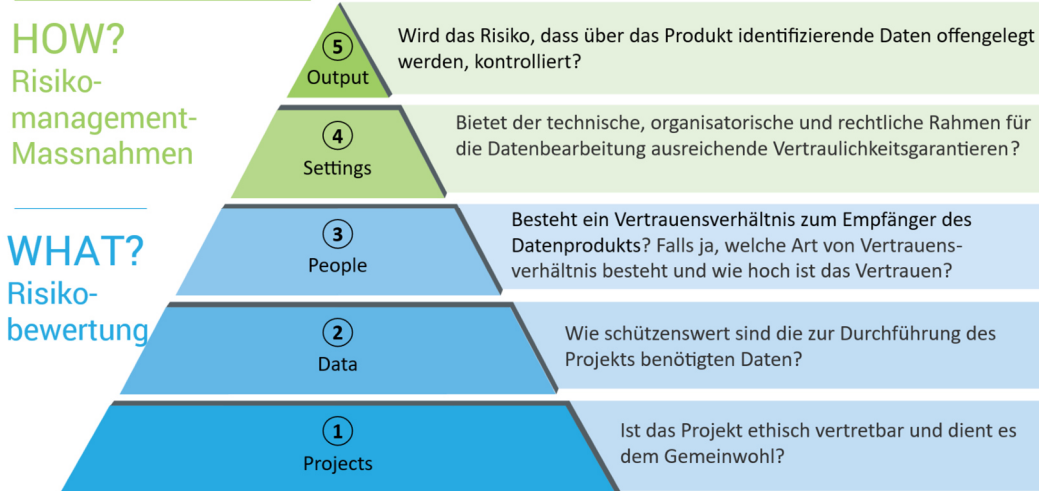


Abbildung 7: Die «Five Saves» – ein Gouvernanz-Rahmen für eine die Grundrechte und die Privatsphäre wahrende Wiederverwendung von Daten der Bundesverwaltung

Auf Basis des «Five Saves»-Modells kann die Erstellung eines Datenprodukts anhand der folgenden Dimensionen kontextabhängig bewertet werden:

- 1) **«Safe Project»:** Ist das Projekt ethisch vertretbar und dient es dem Gemeinwohl? Entspricht es einer menschenzentrierten und vertrauenswürdigen Datenwissenschaft, die das staatliche Handeln unterstützt?³⁰ Zur Beantwortung dieser Fragen müssen die Datenprodukte anhand ihrer Sensibilität und ihrer Nutzen bzw. Risiken für die Gesellschaft klassiert werden. Eine solche Klassifikation wurde für KI-Anwendungen in der EU im Rahmen der KI-Gesetzgebung erstellt.³¹ Sie bietet sich als Rahmenwerk für die Wiederverwendung von Daten in der Bundesverwaltung an und ist mit der Risikoanalyse bei der Datenschutz-Folgenabschätzung (DSFA, Art. 22 DSGVO) vergleichbar.
- 2) **«Safe Data»:** Wie schützenswert sind die zur Durchführung des Projekts benötigten Daten? Massgebend ist die **Sensibilität** der zur Erstellung des Datenprodukts verwendeten Originaldaten und nicht des Datenprodukts an sich. Dessen Sensibilität wird mit der Dimension «Safe Project» beurteilt. Die Sensibilität einer Information muss nach den Kriterien von Artikel 5 DSGVO³² bewertet und eingestuft werden.
- 3) **«Safe People»:** Besteht ein Vertrauensverhältnis zum Empfänger des Datenprodukts? Falls ja, welche Art von Vertrauensverhältnis besteht und wie hoch ist das Vertrauen? Das Vertrauen wird hier als Wahrscheinlichkeit verstanden, mit der die Empfängerin oder der Empfänger des Datenprodukts versucht, identifizierende Daten zu extrahieren. Das Vertrauensverhältnis kann unterschiedliche Formen annehmen, z. B. Berufsgeheimnis, Vertrag, gesetzliche Ermächtigung oder Akkreditierung. Die Einheiten der Bundesverwaltung unterstehen dem Amtsgeheimnis,

³⁰ <https://www.bfs.admin.ch/bfs/de/home/dscc/dscc.assetdetail.29325686.html>

³¹ <https://digital-strategy.ec.europa.eu/fr/policies/regulatory-framework-ai>

³² https://www.fedlex.admin.ch/eli/cc/2022/491/de#art_5

wodurch das Offenlegungsrisiko begrenzt ist. Mit bundesexternen Mitarbeitenden kann das Vertrauen bis zu einem gewissen Grad mit vertraglichen Massnahmen sichergestellt werden. Umgekehrt gibt es bei OGD-Datenprodukten keine Einschränkungen für die Empfangenden, die somit nicht durch ein Vertrauensverhältnis gebunden sind.

- 4) **«Safe Settings»:** Bietet der technische, organisatorische und rechtliche Rahmen für die Datenbearbeitung ausreichende Vertraulichkeitsgarantieren? Mit der Dimension «Safe People» kann ein Vertrauensschema festgelegt werden, das den Einsatz von PET formell regelt, sodass die unter Punkt 1 bis 3 aufgeführten Risiken, die zu einem möglichen Vertrauensdefizit zwischen den an der Datenprodukterstellung beteiligten Parteien führen können, kompensiert werden.
- 5) **«Safe Outputs»:** Wird das Risiko, dass über das Produkt identifizierende Daten offengelegt werden, kontrolliert? Mithilfe von PET kann die Garantie der Offenlegungskontrolle an die Sensibilität der Daten, das Vertrauen in die Empfängerin oder den Empfänger und die Art des Projekts angepasst werden.

Diese fünf Dimensionen lassen sich in zwei Kategorien zusammenfassen (siehe Abbildung 7):

- Die Dimensionen 1 bis 3, die sich auf die Bewertung der Projektrisiken beziehen und die Frage «what?» beantworten, entsprechen der «Beschreibung der geplanten Bearbeitung» und der «Bewertung der Risiken» einer DSFA (Art. 22 DSGVO). Es handelt sich dabei insofern in erster Linie um eine **subjektive** Bewertung, als die soziale, kulturelle und politische Akzeptanz des Projekts überprüft wird. Ein als zu hoch eingestuftes Risiko in einer dieser Dimensionen kann ein Hindernis sein und ein Verbot des Projekts rechtfertigen.
- Die Risikomanagement-Massnahmen beantworten die Frage «how?» und umfassen die Dimensionen «Settings» und «Output». In dieser Kategorie, die dem letzten Teil einer DSFA entsprechen, kann das Risiko von unbefugten Zugriffen, eingeschränkter Nutzung und Offenlegung **objektiv** bewertet und mit organisatorischen, rechtlichen und/oder technischen Massnahmen abgeschwächt werden. Hier bieten PET eine **technische Lösung**, um einem in der ersten Kategorie als potenziell hoch eingestuftes Risiko entgegenzuwirken.

Dieser **risikobasierte Ansatz** ist mit Reglementen und Vorschriften vergleichbar, die im Rahmen bestehender Rechtsgrundlagen gelten, wie beispielsweise die vom DSGVO vorgeschriebene Datenschutz-Folgenabschätzung³³ oder das Formular des BFS für Datenverknüpfungsanfragen³⁴. Aus diesem Grund wird in diesem Kapitel aufgezeigt, wie die vorgängig beschriebenen PET die Umsetzung eines solchen Gouvernanz-Rahmens vereinfachen könnten. Mit PET kann das Offenlegungsrisiko von identifizierenden Daten unabhängig von ihrer Art (Daten von natürlichen oder juristischen Personen, Unternehmen oder Personengruppen) und vom Kontext ihrer Wiederverwendung kontrolliert werden. Sie sind somit ein wesentliches Element für die Anwendung einer datenschutzwahrenden, menschenzentrierten und vertrauenswürdigen Datenwissenschaft.

³³ <https://www.bj.admin.ch/bj/de/home/staat/datenschutz/info-bundesbehoerden.html>

³⁴ <https://www.bfs.admin.ch/bfs/de/home/dienstleistungen/datenverknuepfungen/fuer-dritte.html>

7 Empfehlungen für die Bundesverwaltung

Um die Verfügbarkeit der Daten sicherzustellen, müssen die Katalogisierung und vor allem die Harmonisierung der Daten weiter vorangetrieben werden. Damit werden zum einen die erhobenen Daten standardisiert, zum anderen verbessert sich die Interoperabilität zwischen den Systemen und Datenbanken, was nicht nur den Zugriff auf die Daten, sondern auch ihre Wiederverwendung durch die verschiedenen Stellen und Departemente sowie die Anwendung von Techniken zur Offenlegungskontrolle vereinfacht.

Entscheidend ist auch, dass bundesintern die nötigen datenwissenschaftlichen und PET-Kenntnisse vorhanden sind. Datenwissenschaftlerinnen und -wissenschaftler sind unverzichtbar, um die riesigen Datenmengen zu analysieren und zu interpretieren. Ebenso entscheidend sind Fachpersonen für PET, die sicherstellen, dass sensible Daten vor unbefugtem Zugriff und Verletzungen der Vertraulichkeit geschützt sind. Zudem braucht es für das Prototyping und die Einrichtung der Technologien eine geeignete IT-Infrastruktur. Hierzu gehören leistungsstarke Server, Rechenchips zur Einrichtung von TEE sowie Software-Tools für den Test und die Implementierung von Diensten zur Umsetzung einer datenschutz-wahrenden Datenwissenschaft.

Die in den beiden Fallbeispielen beschriebenen Dienste zur Umsetzung einer vertraulichen Datenwissenschaft, ist eine direkte Antwort auf eine der aktuellen Herausforderungen der Bundesverwaltung: PET können dank ihres Anwendungspotenzials helfen, dem steigenden und immer breiteren Bedarf an diesen Diensten (insbesondere im Rahmen des Projekts DigiSanté) zu begegnen, sie einfacher und schneller bereitzustellen, die dafür nötigen Ressourcen zu reduzieren und dabei gleichzeitig die höchsten datenschutzrechtlichen Standards zu gewährleisten. Diese Dienste sind nicht nur für das EDI relevant, sie betreffen die ganze Bundesverwaltung und sollten daher allen Bundesorganen zur Verfügung stehen.

Um die Wiederverwendung von Daten der Bundesverwaltung zu regeln, braucht es zudem einen Gouvernanz-Rahmen wie das «Five Safes»-Modell. Er soll eine strukturierte Grundlage bilden, um das Risiko für die Verletzung von Grundrechten, insbesondere der Privatsphäre bzw. der Datenvertraulichkeit, zu bewerten, und die eingeführten organisatorischen, rechtlichen und technischen Massnahmen zur Begrenzung dieser Risiken beschreiben. Dazu müssen Rechtsgrundlagen für die Wiederverwendung der Daten festgelegt werden, die in diesen Gouvernanz-Rahmen fallen. Die Gesetzgebung muss die Verantwortlichkeiten und Pflichten der Stakeholder klar und präzise regeln und einen soliden Rechtsrahmen für die Erfüllung der Aufgaben der Bundesverwaltung bieten.

Schliesslich müssen die technischen und organisatorischen Massnahmen zur Verwendung von Daten der Bundesverwaltung in datenwissenschaftlichen Projekten transparent kommuniziert werden. Nur so können das Vertrauen und die Unterstützung der gesamten Schweizer Gesellschaft sichergestellt werden. Im Rahmen dieser Transparenz sind alle Massnahmen zur Umsetzung des «Privacy by Design»-Prinzips detailliert zu beschreiben, um aufzuzeigen, was die Bundesverwaltung für den Schutz der Privatsphäre unternimmt.