PUBLIC

sopra steria

Privacy-Enhancing Technologies für die Verwaltung

Daten schützen und ihr Potenzial nutzen

April 2023

Inhalt

“PETs can become the foundation of a new paradigm of privacy and data protection since they provide more control to data subjects and help enhance trust in the processing of data.“

OECD (2023): Emerging Privacy Enhancing Technologies: Current Regulatory and Policy Approaches.

PUBLIC

sopra steria

Executive Summary

Von Finanzämtern, die Informationen über die finanzielle Situation der Steuerzahler*innen verarbeiten, bis zu Gesundheitsämtern, die Informationen zur Ausbreitung von Krankheiten sammeln: Organisationen im öffentlichen Sektor gehen täglich mit sensiblen Daten um.

Öffentliche Verwaltungen können viele dieser Daten nutzen, um bürgernahe Dienstleistungen zu verbessern, politische Entscheidungen auf eine datenbasierte Grundlage zu stellen oder Innovation über die Zusammenarbeit mit Wissenschaft und Privatsektor zu fördern. Dass sie Daten über Organisationen und Abteilungen hinweg kollaborativ nutzen, ist dafür von zentraler Bedeutung. Häufig steht die effektive Nutzung von Daten für Verwaltungen jedoch in einer Abwägung mit ihrem Schutz. Die Folge: Viele Datensätze werden in behördlichen Silos gehalten und kaum kollaborativ genutzt.

Die Technologiegruppe der Privacy-Enhancing Technologies (PETs) zieht die Aufmerksamkeit der internationalen Community für Dateninnovation auf sich und verspricht eine grundlegende Veränderung dessen, was in der Zusammenarbeit mit Daten möglich ist. Für die öffentliche Verwaltung können PETs ein wichtiges Mittel sein, um die Nutzung und den Schutz von Daten nicht mehr als Nullsummenspiel zu betrachten. Trotz dieser Potenziale bleibt eine systematische Auseinandersetzung mit ihnen im öffentlichen Sektor in Deutschland noch aus.

PETs basieren auf unterschiedlichen technologischen Grundlagen, können komplementär eingesetzt werden und ihre Anwendungen sind auf den Technologiemärkten in unterschiedlicher Marktreife verfügbar. Vor diesem Hintergrund kann ein strukturierter Blick auf die Potenziale und Grenzen von PETs für die öffentliche Verwaltung schwerfallen. Mit diesem Bericht möchten wir die Aufgabe angehen und ein Verständnis für PETs und ihren Wert für den öffentlichen Sektor schaffen. Dafür erläutert der Bericht die Funktionsweise und Anwendungsmöglichkeiten einzelner PETs und präsentiert Ergebnisse einer Analyse des internationalen PET-Marktes. Abschließend macht der Bericht die Möglichkeiten für den Einsatz von PETs im öffentlichen Sektor durch Fallstudien greifbar und zeigt einen von uns entwickelten Entscheidungsbaum auf, der die Auseinandersetzung mit PETs für den eigenen Anwendungsfall erleichtert.

Fünf Erkenntnisse des Berichts stechen hervor

1 PETs können für die öffentliche Verwaltung in Deutschland wichtige Bausteine sein, um den Schutz und die Nutzung von Daten nicht mehr als Nullsummenspiel zu behandeln. In der Folge können Verwaltungen mit Hilfe von PETs mit sensiblen Daten kollaborieren, um öffentliche Dienstleistungen zu verbessern, Entscheidungsgrundlagen zu festigen und Innovation zu fördern.

2 Einzelne PETs verfügen über eine unterschiedliche Marktreife. Öffentliche Verwaltungen müssen deshalb bei der Evaluierung einzelner Lösungen ein Bewusstsein dafür mitbringen, dass die Möglichkeit zur Skalierung der einen Lösung nicht mit der Möglichkeit des breitflächigen Einsatzes einer anderen Lösung einhergeht und spezifische Herausforderungen mitunter noch nicht gelöst werden können.

3 Der Einsatz von PETs im öffentlichen Sektor ist besonders dann erfolgversprechend, wenn er sich an konkreten Anwendungsfällen orientiert. Statt eine bestimmte Technologie als Ausgangspunkt für die Auseinandersetzung mit PETs zu nehmen, sind öffentliche Verwaltungen deshalb angehalten, zunächst ein präzises Verständnis für den eigenen Bedarf zu entwickeln.

4 Unsere Marktanalyse identifiziert einen geringen Anteil deutscher PET-Anbieter im internationalen Wettbewerb. Der internationale PET-Markt wird von Unternehmen aus den USA dominiert, aber auch das Vereinigte Königreich und die Schweiz treten als Hotspots für PET-Unternehmen hervor. Deutschland kann von den Praktiken und Erfahrungen anderer Länder im Aufbau eines PET-Innovationsökosystems lernen.

5 Um die Anwendung von PETs im öffentlichen Sektor zu unterstützen, können Bundes- und Landesregierungen geeignete Rahmenbedingungen schaffen und Maßnahmen ergreifen. Dazu gehört zum einen, dass die Architekturrichtlinien des Bundes PETs und ihre Grundlagen berücksichtigen. Zum anderen können PET-Leuchtturmprojekte stärker mit Forschungsgeldern gefördert werden, um in der öffentlichen Verwaltung Vertrauen in PETs aufzubauen.

Einleitung

„Daten – Gemeinsam digitale Werte schöpfen“[1]

das Motto des Digitalgipfels der Bundesregierung im Dezember 2022 war ein Fingerzeig für die Zukunft der öffentlichen Verwaltung.

Der öffentliche Sektor kann in der Tat Wertschöpfung durch Daten schaffen, sitzt er doch auf einem Schatz sensibler wie wertvoller digitaler Informationen: Jede Steuerverwaltung erhebt und speichert sensible Informationen über die finanzielle Situation von Bürger*innen, jede medizinische Behandlung in öffentlichen Krankenhäusern bringt Datenpunkte für die Akte von Patient*innen hervor.

Die öffentliche Verwaltung ist damit in einer privilegierten Situation, denn Daten haben sich zu einer der wichtigsten Ressourcen des 21. Jahrhunderts entwickelt: In der öffentlichen Verwaltung können sie Vehikel für nutzerfreundliche öffentliche Dienstleistungen, Grundlage für bessere Entscheidungen oder Gegenstand zur Kollaboration mit Forscher*innen sein.[2]

1 BMI (2022): Digital-Gipfel 2022: Daten intelligent und nachhaltig nutzen.

2 Desouza, Cevin C. et al. (2014): Big Data in the Public Sector: Lessons for Practitioners and Scholars.

Herausforderung

Die Bundesinnenministerin Nancy Faeser betonte zum Digitalgipfel: „Wir müssen weiter intensiv an einem Kulturwandel in der Verwaltung hin zu mehr datengetriebenen Handeln arbeiten.“[3] Der Kulturwandel beschreibt zwar eine der zentralen Herausforderungen für eine bessere Datennutzung in der Verwaltung. Das richtige Mindset reicht allerdings nicht aus, um in öffentlichen Verwaltungen datengetriebenes Handeln zu verankern. Neben strategischen Zielen braucht es für eine effektive, sichere und nachhaltige Nutzung von Daten vor allem Mittel und Wege in der Umsetzung. Dazu gehört zum Beispiel, dass IT-Systeme gesichert und personenbezogene Daten geschützt sind, ohne dass die Nutzung von Daten dem zum Opfer fällt.

Privacy-Enhancing Technologies

Auf den internationalen Märkten erregt eine Technologiegruppe Aufsehen, die für das von Nancy Faser beschworene Ziel des datengetriebenen Handelns großes Potenzial birgt – die öffentliche Verwaltung in Deutschland nähert sich ihr jedoch nur sehr verhalten. Sogenannte Privacy-Enhancing Technologies (PETs) setzen sich bereits in der AdTech-Branche oder der Industrie für mobile Hardware durch. Es handelt sich dabei um eine Gruppe von Grundlagentechnologien, deren Anwendung die Nutzung von Daten ermöglicht, indem sie die mit dem Datenmanagement und insbesondere der Kollaboration mit Daten verbundenen Risiken verringert.[4] PETs bieten für den öffentlichen Sektor signifikante Potenziale: Ihr Einsatz kann dazu beitragen, dass öffentliche Verwaltungen den Schutz und die Nutzung von Daten weniger als Nullsummenspiel behandeln.

PETs bieten für den öffentlichen Sektor signifikante Potenziale: Ihr Einsatz kann dazu beitragen, dass öffentliche Verwaltungen den Schutz und die Nutzung von Daten weniger als Nullsummenspiel behandeln.

3 BMI (2022): Daten intelligent und nachhaltig nutzen.

4 The Royal Society (2023): Privacy Enhancing Technologies.

Dieser Bericht

Im öffentlichen Sektor in Deutschland bleibt eine großflächige Auseinandersetzung mit PETs, die über Pilotprojekte hinausgeht, bislang aus. Dieser Bericht soll deshalb einen Überblick über die Funktionsweise von PETs vermitteln und konkrete Möglichkeiten für ihren Einsatz im öffentlichen Sektor aufzeigen.

In **Abschnitt 3** des Berichts wird zunächst aufgezeigt, welches übergeordnete Ziel mit dem Einsatz von PETs in der öffentlichen Verwaltung verfolgt werden kann. Dazu werden drei zentrale Herausforderungen für die Kollaboration mit Daten im öffentlichen Sektor abgegrenzt und eingeordnet.

In **Abschnitt 4** wird zunächst eine Begriffsbestimmung für PETs angeboten, sodann werden die am internationalen PET-Markt dominierenden Technologien und ihre grundlegenden Funktionsweisen erörtert. Das Kapitel schließt mit einem Verweis auf die Grenzen von PETs.

Abschnitt 5 beleuchtet den internationalen PET-Markt mit Blick auf seine Entstehung, globale Verteilung und Finanzierung.

In **Abschnitt 6** werden vier Projekte vorgestellt, in denen PETs im öffentlichen Sektor zum Einsatz kommen und die konkreten Anwendungsfälle für PETs greifbar machen.

Eine Konklusion wird in **Abschnitt 7** gezogen, auch werden Handlungsempfehlungen für die öffentliche Verwaltung entwickelt. Den Abschluss bildet ein Entscheidungsbaum, der als Hilfestellung zur Annäherung an PETs in der eigenen Organisation dient. Der Entscheidungsbaum hilft dabei, den eigenen Use-Case für den Einsatz von PETs zu konkretisieren und spricht auf dieser Grundlage Empfehlungen aus.

Im öffentlichen Sektor in Deutschland bleibt eine großflächige Auseinandersetzung mit PETs, die über Pilotprojekte hinausgeht, bislang aus.

Herausforderungen für datengetriebenes Handeln und Kollaboration im öffentlichen Sektor

PETs können wichtige Bausteine sein, wenn es um konkrete Wege und Werkzeuge geht, wie Organisationen im öffentlichen Sektor Daten nutzen und schützen.

Ihr technologisch innovativer Charakter macht sie zu einem wirkmächtigen Instrument für die öffentliche Verwaltung. Unter vielen besteht der größte Mehrwert von PETs für den öffentlichen Sektor womöglich darin, unterschiedlichen Akteuren die Kollaboration mit sensiblen Daten zu ermöglichen. Allerdings steht die öffentliche Verwaltung oft drei zentralen Herausforderungen gegenüber, die eine effektive Nutzung und Kollaboration mit Daten konditionieren oder gar verhindern:

Datenschutz

Informationssicherheit

Geheimschutz

Alle drei Faktoren beinhalten wichtige Kernelemente, die sie voneinander unterscheiden und die eine Abgrenzung sinnvoll machen. Gleichzeitig lassen sich die drei Herausforderungen nicht vollständig getrennt voneinander betrachten, weil sie in der konkreten Umsetzung immer wieder Überschneidungen aufweisen und auch in einem Spannungsverhältnis zueinander stehen können.

Das Ziel: Daten kollaborativ nutzen

> *„Public administrations maneuver high amounts of valuable data – that is why Privacy Enhancing Technologies have such a huge potential in the public sector."*

Maxime Agostini, Co-Founder & CEO, Sarus Tech[5]

Daten im öffentlichen Sektor können für die Verwaltung eine wertvolle Ressource sein, um die komplexen und dringenden Herausforderungen dieser Zeit zu meistern. Sie können politischen Entscheidungen eine neue Grundlage verschaffen, Services für Bürger*innen erneuern oder Nährboden für eine innovationsfördernde Zusammenarbeit zwischen Verwaltung, Unternehmen und Zivilgesellschaft sein. Damit Daten aus dem öffentlichen Sektor ihr Potenzial entfalten, sind Verwaltungen angehalten, noch stärker auf eine aktive Kollaboration mit ihnen zu setzen. Viele Behörden in Deutschland speichern ihre Datensätze in Silos, die selten mit anderen Stellen verknüpft sind und so isoliert bleiben, obwohl sie für Entscheidungen oder Prozesse einen wichtigen Mehrwert liefern könnten. Die Verknüpfung von Einwohnermeldedaten und sozioökonomischen Informationen ist dafür ein Beispiel: Wenn Stadtplaner*innen eine neue Sozialhilfeeinrichtung bauen und einen Standort festlegen, dann wären Informationen über Familien oder sozioökonomische Details wie das Einkommen von Haushalten eine wichtige Ressource für fundierte Entscheidungen. Die entsprechenden Informationen sind jedoch hochsensibel und werden zum Beispiel in Finanzämtern oder Jobcentern vorgehalten, sie sind somit für die Stadtplanung selten zugänglich.[6]

Die technische Umsetzung des effektiven Datenaustauschs wird über sogenannte *application programming interfaces* (APIs) ermöglicht, die als Schnittstellen zwischen Programmen und Organisationen funktionieren. In vielen Verwaltungen kommen APIs für den Datenaustausch innerhalb des öffentlichen Sektors und mit externen Akteuren bereits zum Einsatz. Teile der Schnittstellen beruhen jedoch auf veralteten und nicht standardisierten Protokollen. Die Folge: Wertvolle Datenquellen innerhalb der Verwaltung bleiben ungenutzt und isoliert.[7]

Eine kollaborative Datennutzung wird jedoch nicht nur durch mangelnde Schnittstellen eingeschränkt. Verwaltungen müssen vor allem gewährleisten, dass die Prozesse und Systeme des Datenmanagements sicher und die personenbezogenen sowie besonders schützenswerten und sensiblen Daten adäquat geschützt sind. Ist das nicht der Fall, dann kann der Prozess des Datenaustausches Ziel für Angriffe von außen sein oder die informationelle Selbstbestimmung von Bürger*innen gefährden. Die nachfolgend skizzierten Herausforderungen des (1) Datenschutzes, der (2) Informationssicherheit und des (3) Geheimschutzes können daher als Stellschrauben verstanden werden. Der Einsatz von PETs kann der öffentlichen Verwaltung bei der Bewältigung dieser Herausforderungen helfen und sie damit dem Ziel der effektiven Kollaboration mit Daten näherbringen.

5 Agostini, Maxime & PUBLIC (2022): Interview über PETs.

6 Klother, Marvin & PUBLIC (2023): Interview über das ATLAS-Projekt.

7 PUBLIC & Sopra Steria (2022): Öffentliche APIs und GovTech: Mit Interoperabilität Innovation fördern.

Datenschutz

Herausforderung

Von Steuerdaten im Finanzamt bis zur Akte über Sozialhilfeleistungen in der Arbeitsagentur: Der öffentliche Sektor sammelt, verwaltet, verwertet und speichert täglich hochsensible Daten, die Rückschlüsse auf Einzelpersonen und ihr Leben zulassen. Diese Daten sind für das Funktionieren der Kernverwaltung essentiell und auch für die Qualität ihrer Dienstleistungen von zentraler Bedeutung. Dass diese Daten effektiv geschützt werden müssen, ist nicht nur eine rechtliche Pflicht, sondern auch Kernargument für ein funktionierendes Vertrauensverhältnis zwischen Bürger*innen und dem Staat.[8]

Abgrenzung

Datenschutz beschreibt Maßnahmen zum Schutz von personenbezogenen Daten vor dem Missbrauch durch Dritte. Daten sind dann personenbezogen, wenn sie Informationen enthalten, die sich auf „identifizierte oder identifizierbare natürliche Personen“[9], d. h. im Fall der öffentlichen Verwaltung auf einzelne Bürger*innen, beziehen.

Anonymisierung vs. Pseudonymisierung

Anonymisierung

Auch wenn es in den Rechtswissenschaften keine einheitliche Begriffsdefinition gibt, werden darunter meist irreversible Maßnahmen verstanden, mit denen keine Rückschlüsse auf individuelle Personen möglich sind. Laut einem Urteil des Europäischen Gerichtshofs (EuGH) sind Daten dann ausreichend anonymisiert, wenn es einen „unverhältnismäßigen Aufwand an Zeit, Kosten und Arbeitskräften“[10] erfordert, Personen zu re-identifizieren.

Pseudonymisierung

Laut der Datenschutz-Grundverordnung (DSGVO) gelten personenbezogene Daten dann als pseudonymisiert, wenn sie nur durch das Hinzuziehen zusätzlicher kontextueller Informationen auf eine Person zurückführbar sind. Voraussetzung hierfür ist, dass diese kontextuellen Informationen gesondert aufbewahrt werden und durch entsprechende technische und organisatorische Maßnahmen der jeweiligen Person nicht zugewiesen werden können. Im Gegensatz zur Anonymisierung handelt es sich hierbei also um eine reversible Maßnahme.

8 OECD (2017): Embracing Innovation in Government.

9 Europäisches Parlament (2018): DSGVO Art. 4.

10 EuGH (2016): Urteil Breyer gegen Bundesrepublik Deutschland.

Rechtlicher Rahmen

Datenschutz basiert in Deutschland rechtlich auf zwei zentralen Säulen: Zunächst haben Einzelpersonen nach dem Grundgesetz das Recht auf informationelle Selbstbestimmung. Zudem regelt die DSGVO seit 2018 maßgeblich, wie personenbezogene Daten verarbeitet werden dürfen. Der Bundesbeauftragte für den Datenschutz und die Informationsfreiheit (BfDI) informiert außerdem über die Schwerpunkte der DSGVO und ihre Bedeutung für den Datenschutz in öffentlichen Verwaltungen.[11] Für die DSGVO sind acht Grundsätze entscheidend:[12]

1. **Rechtmäßigkeit, Verarbeitung nach Treu und Glaube und Transparenz:** Die Verarbeitung personenbezogener Daten muss auf rechtmäßige Weise erfolgen und für die betroffene Person transparent sein.
2. **Zweckbindung:** Die Verarbeitung personenbezogener Daten darf nur für den ausdrücklich festgelegten und legitimen Zweck erfolgen.
3. **Erforderlichkeit:** Die Verarbeitung personenbezogener Daten muss für einen bestimmten Zweck unbedingt notwendig sein, wobei keine andere Möglichkeit besteht, diesen Zweck zu erreichen.
4. **Datenminimierung:** Die verarbeiteten personenbezogenen Daten müssen auf das notwendige Minimum beschränkt sein.
5. **Richtigkeit:** Die personenbezogenen Daten müssen korrekt und auf dem neuesten Stand gehalten werden.
6. **Speicherbegrenzung:** Die personenbezogenen Daten dürfen nur für die notwendige Dauer gespeichert werden.
7. **Integrität und Vertraulichkeit:** Die personenbezogenen Daten müssen angemessen gesichert und vor unbefugtem Zugriff geschützt werden.
8. **Rechenschaftspflicht:** Die verantwortliche Stelle muss nachweisen können, dass sie die Grundsätze des Datenschutzes einhält.

11 BfDI (2020): Die DSGVO in der Bundesverwaltung (Info 6).

12 Europäisches Parlament (2018): DSGVO Art. 5.

Informationssicherheit

Herausforderung

Mehr als 1.000 Angriffe durch Schadsoftware gingen pro Tag im Erhebungszeitraum des BSI-Berichts zur Lage der IT-Sicherheit 2022 in Deutschland alleine in der Bundesverwaltung per E-Mail ein.[13] Die Zahl ist nur ein Beispiel zu den Indikatoren, die verdeutlichen, welchem Ausmaß an externen Gefahren sich der öffentliche Sektor heute gegenübersieht. Hinzu kommen Sicherheitsrisiken, die mit dem internen Umgang mit Informationen zusammenhängen. Die Kontrolle von Zugriffsrechten oder die ausreichende Qualifizierung von Beschäftigten im Umgang mit Informationen sind dafür Beispiele. Die Informationssicherheit der Systeme von Bund, Ländern und Kommunen zu gewährleisten, ist in dieser Hinsicht ein Fundament für digitale Souveränität und eine zentrale Herausforderung für öffentliche Verwaltungen.

Mehr als 1.000 Angriffe durch Schadsoftware gingen pro Tag im Erhebungszeitraum des BSI-Berichts zur Lage der IT-Sicherheit 2022 in Deutschland alleine in der Bundesverwaltung per E-Mail ein.

Abgrenzung

Informationssicherheit ist ein Oberbegriff und bezieht sich auf technische genauso wie auf nicht-technische Systeme. Das bedeutet, mit Informationssicherheit sind grundsätzlich genauso digitale Daten gemeint, wie Informationen, die analog festgehalten sind.[14] Diese weite Definition unterscheidet Informationssicherheit von der IT-Sicherheit, die explizit auf den Schutz technischer Systeme begrenzt ist.

Für die Sicherheit von Informationen lassen sich drei Grundwerte ausmachen: Vertraulichkeit, Integrität und Verfügbarkeit von Informationen. Integrität bedeutet, dass Informationen nicht unerkannt manipuliert oder verändert werden dürfen. Vertraulichkeit ist dann gewahrt, wenn unautorisierte Personen keinen Zugang zu den Informationen erhalten. Verfügbar sind Informationen dann, wenn Personen Zugang zu ihnen haben, die dazu autorisiert sind.[15] Informationssicherheit bildet damit auch den Grundstein für den Schutz von Geheimnissen und den Schutz personenbezogener Daten. Für die konkrete Umsetzung von Informationssicherheit in Unternehmen und der öffentlichen Verwaltung hat das BSI den IT-Grundschutz entworfen und darin Standards sowie Schritt-für-Schritt-Guidelines entwickelt, mit denen Organisationen ein Managementsystem für Informationssicherheit (ISMS) aufbauen können.[16]

13 BSI (2022): Die Lage der IT-Sicherheit in Deutschland 2022.

14 Eckert, Claudia (2009): IT-Sicherheit: Konzepte – Verfahren – Protokolle.

15 BSI: Online-Kurs IT-Grundschutz, Lektion 4: Schutzbedarfsfeststellung, 4.1 Grundlegende Definitionen.

16 BSI (2023): IT-Grundschutz-Kompendium – Werkzeug für Informationssicherheit.

Rechtlicher Rahmen

Die rechtliche Grundlage für die Umsetzung von Informationssicherheit in der öffentlichen Verwaltung ist in Deutschland komplex – es lassen sich jedoch einige Dokumente und Vorschriften hervorheben. So gilt der Umsetzungsplan Bund als zentrale verbindliche Leitlinie für alle Ressorts, Bundesbehörden und, wo spezifiziert, Geschäftsbereiche der Ressorts. Der Umsetzungsplan regelt die Rahmenbedingungen für den Schutz der in der Bundesverwaltung verarbeiteten Informationen und der dabei genutzten IT-Systeme, Dienste und Kommunikationsnetzinfrastrukturen. Ergänzend hat der Beauftragte der Bundesregierung für Informationstechnik (CIO Bund) eine Architekturrichtlinie für die IT des Bundes erarbeitet, die zur Weiterentwicklung der Informationstechnik auf Bundesebene dient und unter anderem die Sicherheit von IT-Systemen fördern soll. Diese wird von den technischen Spezifikationen als Anhang komplementiert. Die Spezifikationen definieren, welche Formen von Serversystemen zulässig sind (Technik und Infrastruktur), welche Arten von Protokollen den Austausch von Daten über Netzwerke ermöglichen (Netze) sowie welche Informationen wann bereitgestellt werden können (Informationen und Daten).

Des Weiteren formuliert der IT-Planungsrat als zentrales politisches Steuerungsgremium von Bund, Ländern und Kommunen für Informationstechnik sowohl Leitlinien als auch einen Umsetzungsplan. Während die Leitlinien die Informationssicherheit als ein gemeinsames verpflichtendes Ziel für die IT-gestützte ebenenübergreifende Zusammenarbeit vorgeben und dafür inhaltliche und organisatorische Rahmenbedingungen formulieren, definiert der Umsetzungsplan konkrete Maßnahmen und Messgrößen für ihren Umsetzungsstatus. Dieser wird durch die Arbeitsgruppe Informationssicherheit im IT-Planungsrat jährlich evaluiert.

Geheimschutz

Herausforderung

Der rechtskonforme Umgang mit Verschlusssachen gehört für viele Verwaltungen zum Alltag. Sie unterliegen dem Geheimschutz und sind deshalb besonders schützenswert, weil unbefugter Zugriff auf ihren Inhalt eine Gefahr für das öffentliche Interesse darstellt.[17] Die Digitalisierung der öffentlichen Verwaltung birgt gleichwohl Chancen und Risiken für den Umgang mit Verschlusssachen: Auf der einen Seite können mittels digitaler Verfahren befugte Personen dezentral auf Verschlusssachen zugreifen und Arbeitsabläufe damit deutlich vereinfachen. Auf der anderen Seite entstehen mit technologischem Fortschritt auch neue Methoden im Bereich der Spionage und gezielte Angriffe auf Geheimnisse. Der öffentliche Sektor ist deshalb angewiesen, technologische Mechanismen zu etablieren, die Geheimschutz im öffentlichen Interesse auch vor dem Hintergrund wachsender Gefahrenpotenziale sicherstellen.

17 Bundesministerium der Justiz (2021): Sicherheitsüberprüfungsgesetz (SÜG) § 1.

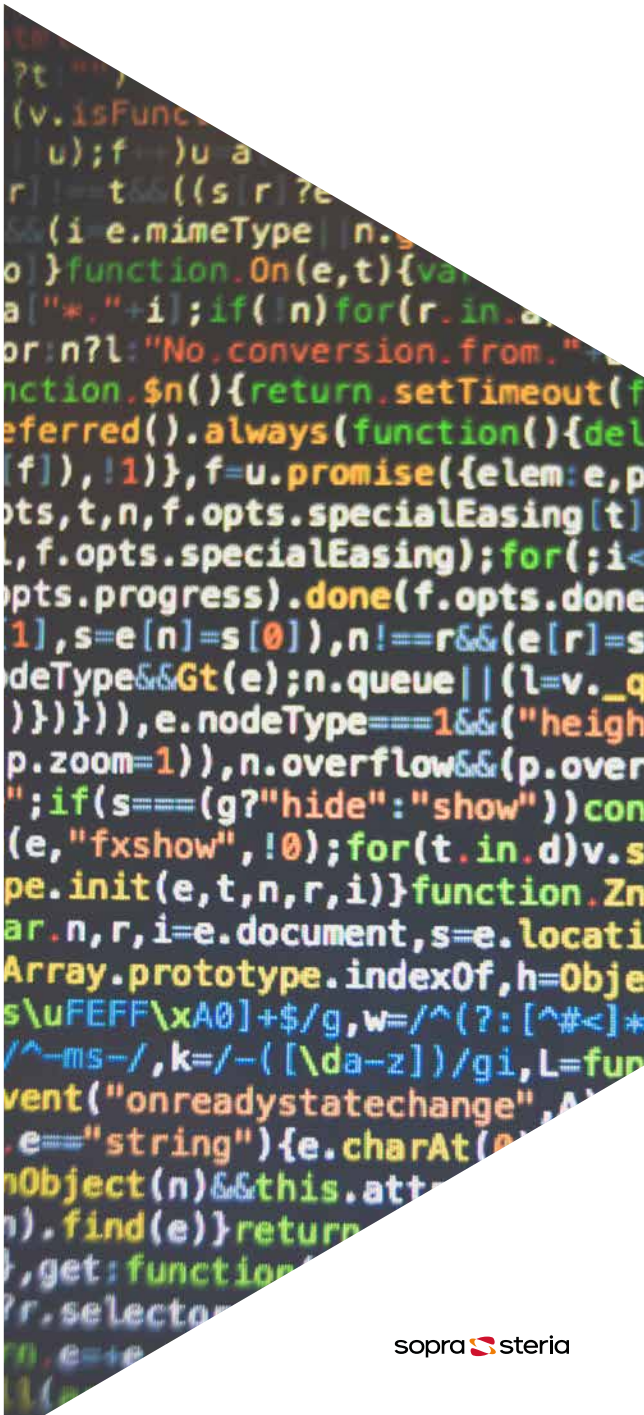
Abgrenzung

Der Geheimschutz bezieht sich auf als Verschlusssache eingestufte Informationen und soll unbefugten Zugriff, Offenlegung oder Missbrauch verhindern. Informationen werden dann als Verschlusssache eingestuft, wenn ihre Offenlegung der nationalen Sicherheit, der öffentlichen Ordnung oder anderen wichtigen Interessen Deutschlands schaden könnte.[18] Das Format der Informationen spielt dafür zunächst keine Rolle: Gesprochenes Wort kann genauso unter den Begriff fallen wie elektronisch gespeicherte Daten.

Dem Einsatz von technischen Systemen zur Verarbeitung und Speicherung von Verschlusssachen in der Verwaltung kommt in der Verschlusssachenanweisung (VSA) eine wichtige Rolle zu. Die Verschlusssachen-IT (VS-IT) unterliegt entsprechenden Sicherheitsanforderungen, wie etwa an die Zulassung der verwendeten IT-Produkte oder ein angemessenes Risikomanagementsystem.[19] Die Anforderungen selbst sind oft als Verschlusssache eingestuft und werden somit nicht öffentlich transparent gemacht. Für ihre Überprüfung ist das BSI zuständig.

Rechtlicher Rahmen

In Deutschland bildet das Gesetz über die Sicherheitsüberprüfung von Personen (Sicherheitsüberprüfungsgesetz – SÜG) den gesetzlichen Rahmen für Geheimschutz. Auf Grundlage dieses Gesetzes regelt die VSA den Umgang mit Verschlusssachen in der öffentlichen Verwaltung. Der Geheimschutz fällt hauptsächlich in den Verantwortungsbereich des Bundesnachrichtendienstes, des Bundesamts für Verfassungsschutz, des Bundeskriminalamts und des Bundesamts für Sicherheit in der Informationstechnik.



18 Bundesministerium der Justiz (2023): Verschlusssachenanweisung § 2.

19 Bundesministerium der Justiz (2023): Verschlusssachenanweisung § 50 Abs. 3 i. V. m. Abs. 8a.

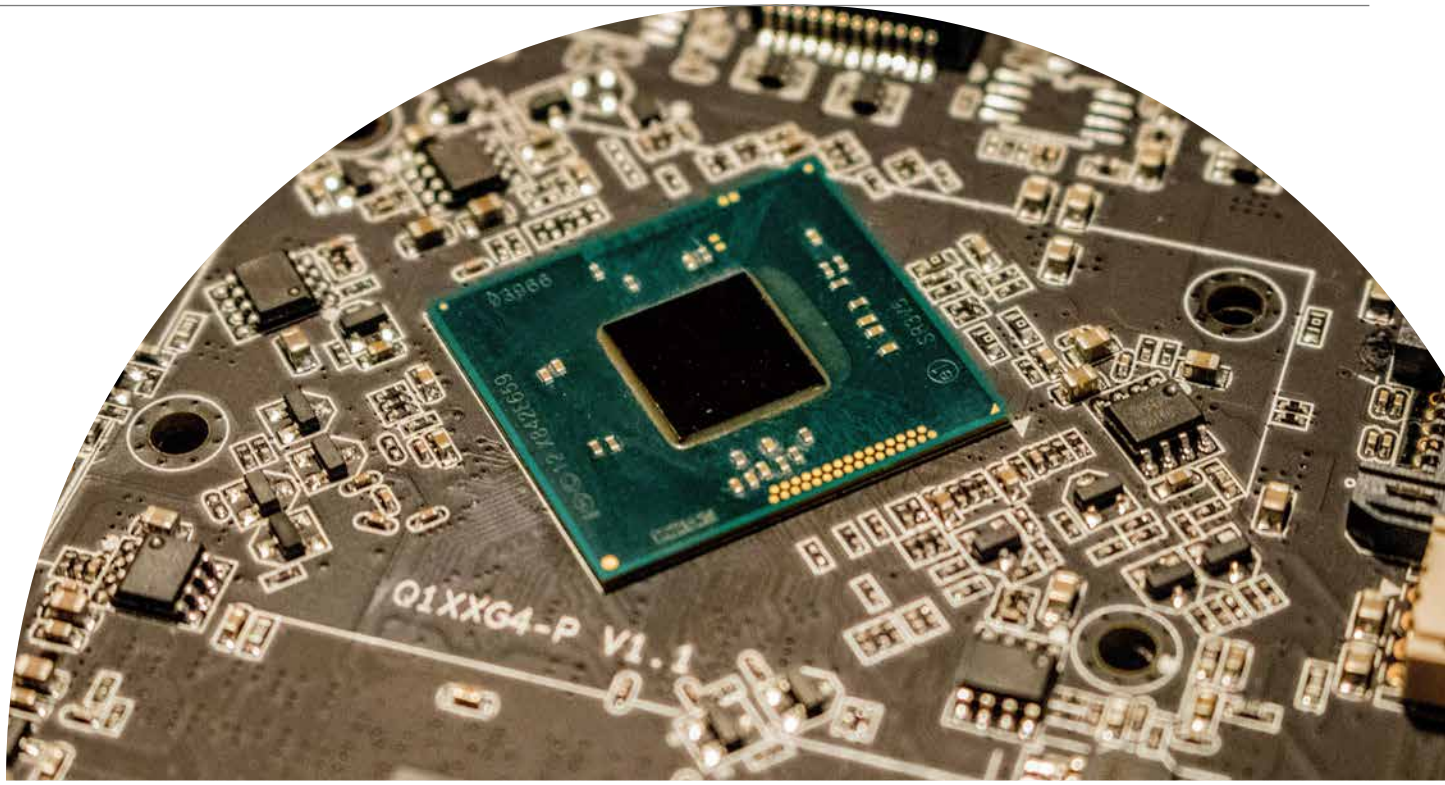
Privacy-Enhancing Technologies erklärt

1 Was sind Privacy-Enhancing Technologies?

Der Begriff der Privacy-Enhancing Technologies wurde erstmals Ende der 1990er Jahre eingeführt. Die wachsende Besorgnis über die Auswirkungen neuer Technologien auf die Privatsphäre veranlasste Forscher*innen und politische Entscheidungsträger*innen, nach Wegen zum Schutz persönlicher Daten zu suchen. Viele Privacy-Enhancing Technologies wurden mit der Überzeugung entwickelt, dass Informationen besser geschützt werden können, wenn Technologien selbst die Erfassung und Verarbeitung von Daten hinsichtlich der Erfüllung von Mindeststandards kontrollieren, anstatt sich auf rechtliche und politische Maßnahmen zu verlassen.[20]

Viele Privacy-Enhancing Technologies wurden mit der Überzeugung entwickelt, dass Informationen besser geschützt werden können, wenn Technologien selbst die Erfassung und Verarbeitung von Daten hinsichtlich der Erfüllung von Mindeststandards kontrollieren, anstatt sich auf rechtliche und politische Maßnahmen zu verlassen.

20 Goldberg, Ian; Wegner, David; Brewer, Eric (1997): Privacy-Enhancing Technologies for the Internet.



Die Agentur der Europäischen Union für Cybersicherheit (ENISA) definiert PETs als eine Technologiegruppe, die Funktionen zum Schutz der Privatsphäre oder des Datenschutzes unterstützen[21] und damit auf Grundsätze wie beispielsweise Datensparsamkeit abzielen. Häufig verfolgen diese Techniken, Werkzeuge und Systeme das Ziel, die Verwendung von Daten möglich zu machen, ohne vollen Zugang zu den Rohdaten zu erfordern.

Mit Blick auf die drei Herausforderungen, die dieser Bericht hervorhebt, wird deutlich: Privacy-Enhancing Technologies leisten insbesondere einen Beitrag zum Datenschutz. Wie die folgenden Kapitel aufzeigen, generieren Teile der Technologien jedoch auch in Bereichen der Informationssicherheit und des Geheimschutzes Mehrwerte für die öffentliche Verwaltung. In der Folge können sie Organisationen dazu befähigen, sicher mit Daten zu kollaborieren, indem sie Risiken einschränken, die gewöhnlich mit der Kollaboration an Daten über Organisationen und ihre Einheiten hinweg einhergehen.[22]

Dass die Definition von PETs nicht konkreter ist, liegt daran, dass die unterschiedlichen Werkzeuge auf verschiedene Technologien zurückgreifen – etwa Verschlüsselung, anonyme Kommunikation oder künstlich produzierte Daten. So unterschiedlich wie der technologische Unterbau sind auch die Anwendungsmöglichkeiten für PETs.

Die Agentur der Europäischen Union für Cybersicherheit (ENISA) definiert PETs als eine Technologiegruppe, die Funktionen zum Schutz der Privatsphäre oder des Datenschutzes unterstützen19 und damit auf Grundsätze wie beispielsweise Datensparsamkeit abzielen.

21 ENISA (2016): Readiness Analysis for the Adoption and Evolution of Privacy Enhancing Technologies.

22 The Royal Society (2023): Privacy Enhancing Technologies.

2 Privacy-Enhancing Technologies und ihre Grundlagen

PETs beruhen auf unterschiedlichen technologischen Grundlagen und befinden sich an verschiedenen Punkten der Marktreife. Der internationale Markt für PETs wird durch technologische Innovationen aus Forschung und Entwicklung sowie den steigenden Bedarf an Informationssicherheit und Datenschutz weiter wachsen.[23] Es ist vor diesem Hintergrund nicht einfach, ein strukturiertes Bild vom aktuellen Stand des PET-Angebots zu erhalten und Lösungen zu identifizieren, die den Anwendungsfällen und Bedarfen des öffentlichen Sektor entsprechen. Dieses Kapitel soll zunächst einen Blick auf die Grundlagen von PETs werfen, die auf dem internationalen PET-Markt besonders präsent sind und die gleichzeitig Mehrwerte für den Einsatz im öffentlichen Sektor haben können.

Drei Parameter sind beim Blick auf PETs entscheidend, die alle Lösungen und ihre Einsatzmöglichkeiten konditionieren, unabhängig vom Kontext des konkreten Anwendungsfalles:

PETs sind je nach zu lösender Herausforderung kombinierbar und lassen sich als komplementäre Werkzeuge im Baukasten der sicheren Datenverarbeitung von Organisationen verstehen.

1. PETs sind kombinierbar

Jede Technologie innerhalb der PETs hat ihre Stärken und ihre Grenzen. Die Vielschichtigkeit der Lösungen und ihrer Anwendungsfälle ist der Grund dafür, dass sie nicht im direkten Wettbewerb gegeneinander stehen.[24] PETs sind je nach zu lösender Herausforderung kombinierbar und lassen sich als komplementäre Werkzeuge im Baukasten der sicheren Datenverarbeitung von Organisationen verstehen. Die Anbieter auf dem PET-Markt greifen die Möglichkeit zur Kombination von PETs bereits vielfach auf und bieten mitunter ein flexibles Portfolio an technologischen Komponenten, das je nach Herausforderung in der Anwendung erweitert werden kann. Die Lösungen, die konkret zum Einsatz kommen, werden dann beispielsweise als Plattform aufbereitet, um Kund*innen eine sichere Datenverarbeitung oder -auswertung zu ermöglichen.

23 The Royal Society (2023): Privacy Enhancing Technologies.

24 Williams, Ellison Anne (2022): Privacy-enhancing technologies – myths and misconceptions.

2. PETs haben eine unterschiedliche Marktreife

PETs haben in den letzten Jahren nicht zuletzt aufgrund des wachsenden Bewusstseins für Datenschutz und Informationssicherheit mehr Aufmerksamkeit aus der Privatwirtschaft und dem öffentlichen Sektor erhalten. Die Lösungen auf dem PET-Markt befinden sich aktuell jedoch auf unterschiedlichen Stufen der technologischen Entwicklung und Marktreife. Das geht unter anderem darauf zurück, dass die Angebote auf unterschiedlichen technologischen Grundlagen beruhen, wie das folgende Kapitel aufzeigt. Nutzer*innen müssen deshalb bei der Evaluierung einzelner Lösungen ein Bewusstsein dafür mitbringen, dass die Möglichkeit zur Skalierung der einen Lösung nicht mit der Möglichkeit des breitflächigen Einsatzes einer anderen Lösung einhergeht und spezifische Herausforderungen mitunter noch nicht gelöst werden können. Gleichzeitig müssen Nutzer*innen sorgfältig abwägen, ob die ausgewählten Lösungen gut in ihre bestehende IT-Infrastruktur integriert werden können.[25] Das ist besonders für die Integration im öffentlichen Sektor relevant, wo maßgeschneiderte Lösungen häufig erwartet werden.[26]

3. PET-Lösungen richten sich an unterschiedliche Nutzer*innen

Die Angebote auf dem PET-Markt befinden sich nicht nur auf unterschiedlichen Stufen der Marktreife – die Technologien können in der Anwendung auch diverse Formen annehmen und richten sich folglich an unterschiedliche Nutzer*innen. Die meisten PET-Lösungen werden in der Anwendung heute von ausgebildeten Informatiker*innen oder ähnlichen Profilen betreut, die eine hohe technische Expertise besitzen. Nicht nur die Einbindung in die IT-Infrastruktur, sondern auch die Betreuung und Anwendung dieser Lösung erfordert erweiterte Qualifikationen im Umgang mit technischen Systemen und Daten. No-Code- oder Low-Code-Anwendungen sind auf dem PET-Markt vorhanden, jedoch noch nicht weit verbreitet. Diese Art von Lösungen geht für gewöhnlich mit einer intuitiven Oberfläche einher und erlaubt auch technisch weniger Qualifizierten die Betreuung der Anwendung.

25 Information Commissioner's Office (2022): Anonymisation, pseudonymisation and privacy enhancing technologies guidance, Chapter 5.

26 The Royal Society (2023): Privacy Enhancing Technologies.

3 Die wichtigsten PETs im Überblick

Dieser Abschnitt stellt die wichtigsten PETs vor, indem grundlegende Funktionsweisen erklärt und anhand von Beispielen verdeutlicht werden. Die vorgestellten PETs haben sich in der Analyse als besonders relevant für den Einsatz im öffentlichen Sektor erwiesen. Die Reihenfolge der Darstellung ergibt sich entlang der Frage, bei welcher der drei Herausforderungen die jeweiligen PETs im öffentlichen Sektor Mehrwerte bieten können.

A Differential Privacy

Bei welcher Herausforderung kann Differential Privacy der öffentlichen Verwaltung helfen?

- **Datenschutz**
- Informationssicherheit
- Geheimschutz

Funktionsweise

Technisch gesehen ist Differential Privacy ein mathematisches Kriterium.[27] Bei hohem Erfüllungsgrad dieses Kriteriums können Analyst*innen Statistiken und Auswertungen also aus einem Datensatz mit personenbezogenen Daten generieren, ohne dass Rückschlüsse auf Einzelpersonen möglich sind.[28] Anwendungen, die auf Differential Privacy zurückgreifen, reichern für die Analyse Statistiken und Auswertungen mit *noise* (Rauschen) an, also einer künstlichen Veränderung gegenüber den Rohdaten. Die künstliche Veränderung der Daten für die Analyse hat das Ziel, dass die statistischen Eigenschaften der Daten erhalten bleiben, spezifische Rückschlüsse auf eine Person aus der Datenbank jedoch erschwert werden. Grundlage für die Darstellung des Mechanismus und die Erfüllung des Kriteriums ist der Parameter Epsilon (ε).

27 Nissim, Kobbi; Steinke, Thomas (2018): Differential Privacy: A Primer for a Non-technical Audience.

28 Dwork, Cynthia; Roth, Aaron (2014): The Algorithmic Foundations of Differential Privacy.

Dieser *Differential-Privacy-Parameter* beschreibt das Maß an Rauschen, das bei der Analyse hinzugefügt wird. Je höher der Parameter, desto mehr Rauschen wird hinzugefügt und desto schwieriger wird es, spezifische Informationen aus der Datenbank abzuleiten.[29]

Mit einem höheren Grad an hinzugefügtem Rauschen geht jedoch ein gradueller Präzisionsverlust in den Auswertungen einher.

Beispiel

Schulbehörden führen in vielen Fällen eine Schulstatistik mit Daten, die zum Beispiel Aufschluss über schulische Leistungen geben. Die Behörde möchte Statistiken und Analysen aus diesen Daten erstellen und veröffentlichen, um über die Wirkung ihrer Schulpolitik zu informieren. Dabei muss sie sicherstellen, dass die Veröffentlichung keine Rückschlüsse auf Einzelpersonen zulässt. Um dies zu verhindern, kann mittels Differential-Privacy-Anwendungen ein Rauschen für die Analyse der einzelnen Schüler*innen hinzugefügt werden. Dies würde es erschweren, spezifische Informationen über eine Person aus der Datenbank abzuleiten, während gleichzeitig allgemeine statistische Auswertungen und Analysen auf der Datenbank durchgeführt werden können – wie etwa die Zeugnisdurchschnitte in verschiedenen Stadtteilen. Der Differential-Privacy-Parameter lässt sich dabei wie ein Drehknopf verstehen, der die Balance zwischen Datenschutz und Genauigkeit der Analyse angibt. Wird der Parameter auf stärkeren Datenschutz ausgerichtet, dann wird der Statistik mehr Rauschen hinzugefügt, die Re-Identifizierung wird schwieriger – die Daten verlieren jedoch auch an Genauigkeit. Wird der Parameter stärker auf Genauigkeit ausgerichtet, passiert das Gegenteil.

> *„The key feature of Differential Privacy is that it provides the ability to extract information from highly sensitive data without violating privacy.“*[30]

Marco Gaboardi, Co-Founder & Chief Scientist, DPella

29 Dwork, Cynthia; Roth, Aaron (2014): The Algorithmic Foundations of Differential Privacy.

30 Gaboardi, Marco; Russo, Alejandro & PUBLIC (2023): Interview über PETs.

Verbreitung

Die Informatik entwickelt bereits seit etwa zwanzig Jahren robuste Theorien für Differential Privacy und konnte damit den Grundstein für die praktische Anwendung legen. Ein prominentes Beispiel für den Einsatz von Differential Privacy ist die Veröffentlichung der Daten des US-amerikanischen Zensus. Seit 2008 greift das Volkszählungsamt auf Differential-Privacy-Anwendungen zurück, um Rückschlüsse auf Einzelpersonen aus seinen veröffentlichten Statistiken zuzulassen. Es reagiert damit auf die wachsende Anzahl und Qualität von Versuchen, den Zugang zu großen Datenbasen und Statistiken zu Missbrauchszwecken und zur Re-Identifizierung von Individuen zu nutzen. Das US-amerikanische Census-Bureau ist damit die erste Organisation für den breitflächigen Einsatz von Differential Privacy.[31]

Übersicht

Vorteile

Ermöglicht hohen Grad an Datenschutz und liefert ein Bemessungsinstrument für die Balance zwischen Datenschutz und Datengenauigkeit.

Nachteile

Hoher Schutz der Privatsphäre geht mit Verlust von Präzision in den Analysen einher.

Marktreife

Gering - (wenngleich die mathematischen Prinzipien von Differential Privacy schon länger diskutiert werden), Differential-Privacy-Lösungen befinden sich noch im Entwicklungsstadium und ihre Anwendung und Betreuung erfordert in der Regel ein hohes technisches Verständnis.

Komplementarität

Wird häufig mit anderen PETs kombiniert.

B Synthetische Daten

Bei welcher Herausforderung können synthetische Daten der öffentlichen Verwaltung helfen?

- **Datenschutz**
- Informationssicherheit
- Geheimschutz

Funktionsweise

Synthetische Daten sind künstlich generierte Daten. Sie basieren auf echten Daten, wurden jedoch so verändert, dass Rückschlüsse auf Einzelpersonen in den Daten erschwert werden. Der Prozess der Erstellung synthetischer Daten kann auf verschiedene Arten erfolgen. Eine Möglichkeit besteht darin, Machine-Learning-Algorithmen wie Generative-Adversarial-Networks (GANs) oder Variational Autoencoders (VAEs) zu verwenden. Diese Algorithmen lernen aus echten Daten und generieren dann neue Daten, die den echten ähnlich sind, aber keine identifizierbaren persönlichen Informationen enthalten sollen.

Ein weiterer Ansatz besteht darin, die Daten so zu verändern, dass lediglich identifizierbare persönliche Informationen wie Namen oder Adressen entfernt oder verschlüsselt werden. In einigen Fällen können auch bestimmte Merkmale der Daten, wie beispielsweise Altersgruppen oder Postleitzahlen, kategorisiert, aggregiert, summiert oder geclustert werden, um die Privatsphäre der betroffenen Personen zu schützen.[32]

31 Abowd, John M. (2018): Protecting the Confidentiality of Americas Statistic's: Adopting Modern Disclosure Avoidance Methods at the Census Bureau.

32 Wagner, Paul (2021): Privacy Enhancing Technologies and Synthetic Data.

„In der Zukunft werden wir zurückblicken und denken: So wie heute mit echten Daten zu arbeiten, ist so absurd, wie damals im Flugzeug zu rauchen."[32]

Andreas Ponikiewicz, *VP Global Sales, Mostly AI*

Mit der Veränderung der Daten kann je nach Qualität der Anwendung ein Präzisionsverlust einhergehen – ähnlich wie bei Differential Privacy ist auch beim Einsatz synthetischer Daten die Balance zwischen Datenschutz und Nutzbarkeit der Daten ein wichtiger Parameter.

Beispiel

Die öffentliche Verwaltung kann synthetische Daten bereitstellen, um Forschenden eine Vorabprüfung eines potentiell interessanten Datensatzes zu ermöglichen. Das Projekt für synthetische Datengenerierung für britische Langzeitstudien (SYLLS) stellt synthetische Daten auf Basis einer Langzeitstudie für Forschende bereit. Die Datensätze des Projektes verlinken Zensusdaten mit administrativen Informationen wie Familienstand, sozialem Status und Religionszugehörigkeit in England, Wales, Schottland und Nordirland. Die originalen Datensätze sind aufgrund ihrer sensiblen Informationen nur für einen kleinen Kreis an Forschenden zugänglich und damit nicht für eine freie wissenschaftliche Analyse verfügbar. Im Rahmen des Projektes wurden Teile der Daten synthetisiert und einem größeren Kreis an Forschenden zur Verfügung gestellt, damit diese sich mit den Variablen der generellen Charakteristika für ihre Forschungsarbeiten vertraut machen können.[34]

Verbreitung

Synthetische Daten nehmen auf dem internationalen PET-Markt eine große Rolle ein. Die Gesundheitsforschung und die Finanzindustrie sind zwei dominierende Felder für ihren Einsatz. Die Projekte zur Implementierung befinden sich jedoch auch hier in der Regel noch im Status der Pilotierung. Im Gesundheitswesen werden synthetische Daten beispielsweise verwendet, um klinische Studien durchzuführen oder neue Therapien zu entwickeln, ohne dabei die Privatsphäre der Patienten zu gefährden. In der Finanzindustrie werden synthetische Daten verwendet, um Risikobewertungen durchzuführen oder Betrug zu erkennen, ohne dabei vertrauliche Informationen offenzulegen.[35]

Übersicht

Vorteile

Ermöglicht einen hohen Grad an Datenschutz und bietet vielfältige Anwendungsmöglichkeiten, wie zum Beispiel das Trainieren von KI-Modellen.

Nachteile

Ein höherer Grad an Manipulation kann mit dem Verlust von Präzision in den Analysen einhergehen und die Re-identifizierung wird erschwert – damit sind sie nicht für alle Arten der Analyse geeignet. Die Qualität synthetischer Daten kann je nach Anwendung unterschiedlich ausfallen.

Marktreife

Mittel – synthetische Daten spielen für das Trainieren von KI-Modellen bereits eine Rolle. Konkrete Anwendungsfälle mit Transferwert für den öffentlichen Sektor sind darüber hinaus jedoch noch selten oder befinden sich in der Pilotierung.

Komplementarität

Wird häufig mit anderen PETs kombiniert.

33 Klisch, Sabine; Ponikiewicz, Andreas & PUBLIC (2022): Interview über PETs.

34 Calcraft, Paul et al. (2021): Accelerating public policy research with synthetic data.

35 Langevin, Alex (2021): Synthetic Data Augmentation of Imbalanced Data Sets With Generative Adversarial Networks Under Varying Distributional Assumptions: A Case Study in Credit Card Fraud Detection.

36 Je nach Use-Case und der Notwendigkeit zur Re-Identifizierung kann das ein Vorteil oder Nachteil sein.

C Federated Learning

Bei welcher Herausforderung kann Federated Learning der öffentlichen Verwaltung helfen?

> *„Federated Learning turns around a present problem in training machine learning models: Now the analysis comes to the data, not the data to the analysis."*
>
> *Vertreter*in eines PET-Unternehmens*

Funktionsweise

Der Begriff des Federated Learnings wurde 2016 eingeführt und hat seither in der Forschung und Entwicklung große Aufmerksamkeit erhalten.[37] Federated Learning ermöglicht mehreren Parteien mit dezentralen Daten, gemeinsam ein Machine-Learning (ML)-Modell zu trainieren, ohne die lokalen Daten auszutauschen. Die Technologie greift dafür auf Entwicklungen aus verschiedenen Bereichen wie maschinellem Lernen und der Datenschutzforschung zurück.[38] Die Technologie reagiert auf eines der zentralen Probleme bei der Entwicklung von ML-Modellen: ML-Modelle sind auf eine große Menge von qualitativ hochwertigen Daten angewiesen, um präzise zu funktionieren. Diese Daten sind jedoch in vielen Anwendungsbereichen nicht zentral verfügbar: Sie werden – zum Beispiel durch Smartphone-Nutzer*innen oder in verschiedenen Krankenhäusern – dezentral generiert und gespeichert. Neben der physischen Trennung ergeben sich außerdem in vielen Fällen datenschutzrechtliche Grenzen, die eine Zusammenführung der Daten einschränken.

Beim zentralisierten Federated Learning betreibt eine zentrale Instanz das ML-Modell, das es zu trainieren gilt. Dieses Modell führt in einzelnen Schritten immer wieder Trainings mit den lokalen Daten der dezentralen Parteien durch. Die dezentralen Parteien generieren durch ihre Trainings jeweils ein Update des ML-Modells, verschlüsseln es und geben es zurück an die zentrale Instanz. Aus den lokalen Trainings entsteht in der zentralen Instanz ein verbessertes ML-Modell, das im nächsten Schritt wieder zur Analyse an die dezentralen Parteien gesendet wird.[39] In diesem iterativen Lernprozess können isolierte Datensilos verbunden und die Qualität der ML-Modelle verbessert werden. Die dezentralisierte Form des Federated Learnings funktioniert ebenfalls über iteratives Lernen auf verstreuten Datenquellen, diese kommunizieren jedoch direkt miteinander und umgehen damit die Abhängigkeit von einer zentralen Instanz.

Federated Learning kann dezentrale Datenquellen für das Trainieren von ML-Modellen nutzbar machen – die Technologie selbst beinhaltet jedoch keine eigenen Sicherheitsmechanismen. Wenn lokale Updates aus den dezentralen Parteien nicht ausreichend verschlüsselt und geschützt sind, dann können Angriffe Rückschlüsse auf die personenbezogenen Daten der einzelnen Parteien ermöglichen.[40]

37 McMahan, Brendan et al. (2017): Communication-Efficient Learning of Deep Networks from Decentralized Data.

38 Li, Qinbin et al. (2021): A Survey on Federated Learning Systems: Vision, Hype and Reality for Data Privacy and Protection.

39 Kairouz et al. (2021): Advances and Open Problems in Federated Learning.

40 Kairouz et al. (2021): Advances and Open Problems in Federated Learning.

D Homomorphe Verschlüsselung

Bei welcher Herausforderung kann homomorphe Verschlüsselung der öffentlichen Verwaltung helfen?

- Datenschutz
- Informationssicherheit
- Geheimschutz

Beispiel

Die Verbindung von Datensilos in Polizeibehörden ist ein möglicher Anwendungsfall für Federated Learning im öffentlichen Sektor: Lokale Polizeibehörden erheben und speichern von Fall zu Fall sensible Daten wie Recherche- oder Überwachungsergebnisse. Angenommen, eine übergeordnete Polizeibehörde wollte ein System zur Früherkennung von Straftaten etablieren, so ist das Modell im Lernprozess auf eine breite Datenbasis angewiesen. Die einzelnen lokalen Datensilos beinhalten gegebenenfalls nicht genug Daten, um ein qualitativ hochwertiges Modell zu trainieren, das verlässliche Früherkennung betreibt. Die Polizeibehörden könnten deshalb davon profitieren, ein ML-Modell auf der gemeinsamen Grundlage ihrer Daten zu trainieren. Die Daten können jedoch aufgrund von Datenschutzrichtlinien nicht einfach zwischen den Behörden ausgetauscht werden. Ein Federated-Learning-Ansatz würde es einem zentralen Konsortium erlauben, auf die dezentralen Datensätze zuzugreifen, ohne dass die lokalen Polizeibehörden ihre Rohdaten bewegen oder freigeben müssen.[41]

Überblick

Vorteile

Kann dezentrale Datenquellen für das Trainieren von ML-Modellen erschließen.

Nachteile

Bietet keine originären Funktionen zum Schutz von Daten und ist deshalb anfällig für Cyberangriffe, wenn Analysen nicht zusätzlich geschützt werden.

Marktreife

Mittel – trotz hohen Interesses von großen Techunternehmen sind die Standardisierung und Skalierung noch nicht erreicht.

Komplementarität

Wird häufig mit anderen PETs kombiniert, um Rückschlüsse auf personenbezogene Daten zu verhindern.

Funktionsweise

Homomorphe Verschlüsselung wird auch als „heiliger Gral der Verschlüsselung"[42] bezeichnet. Die Technologie basiert auf kryptographischen Grundlagen und bietet die Möglichkeit, Daten im Ruhezustand, während des Datentransfers und während der Verarbeitung verschlüsselt zu halten. So können Nutzende Analysen auf Datensätzen vollziehen und das Ergebnis zu ihrer Anfrage erhalten, haben jedoch zu keinem Zeitpunkt Zugriff auf die Rohdaten. Technisch betrachtet nutzt die homomorphe Verschlüsselung wie andere Verschlüsselungsverfahren auch einen *Public Key* zur Verschlüsselung von Daten, die nur mit dem passenden *Private Key* entschlüsselt werden können. Der Unterschied besteht darin, dass bei der homomorphen Verschlüsselung Manipulationen der verschlüsselten Daten durchgeführt werden können, während die Vertraulichkeit der Daten erhalten bleibt.[43]

41 Li, Qinbin et al. (2021): A Survey on Federated Learning Systems: Vision, Hype and Reality for Data Privacy and Protection.

42 Tourky, Dalia; ElKawkagy, Mohamed; Keshk, Arabi (2016): Homomorphic encryption the „Holy Grail" of cryptography.

43 Hackermoon (2020): Homomorphic Encryption: Introduction And Use Cases.

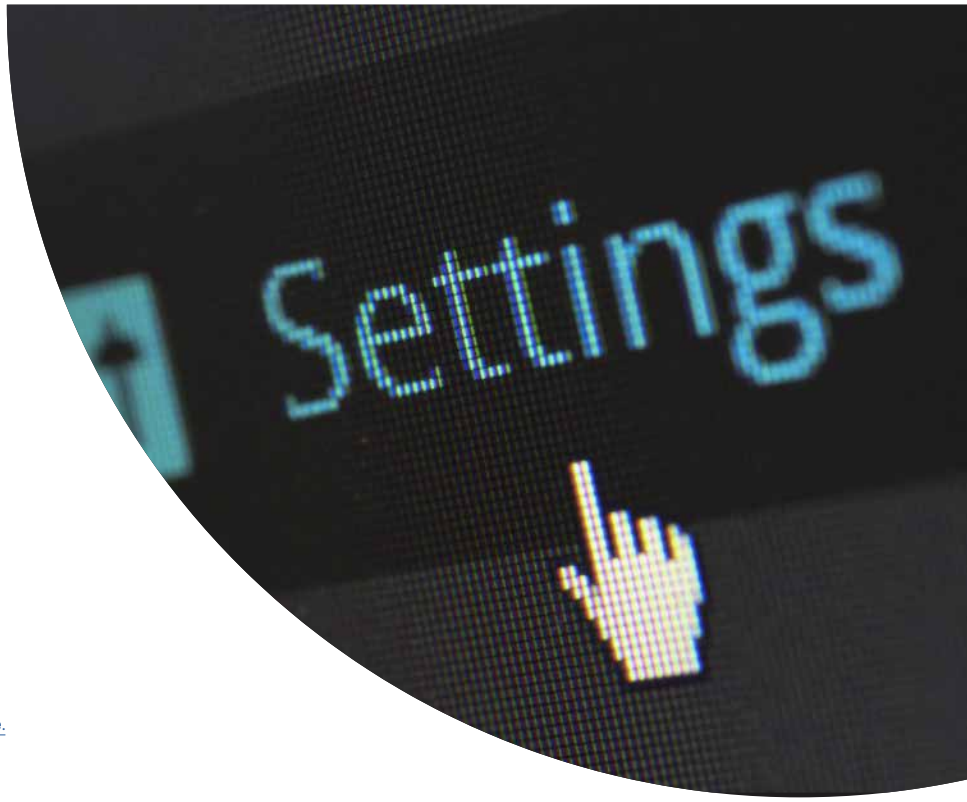
2009 legte Craig Gentry mit seiner Dissertation an der Stanford University erstmals ein Konzept voll-homomorpher Verschlüsselung vor und beschreibt die Funktion der Technologie als Blackbox: Verschiedene Nutzer*innen können mit ihren Händen auf das Innere der Box zugreifen, es verändern und Schlüsse aus dem Inhalt ziehen, ihn jedoch nicht entnehmen oder einsehen. Einzig der Besitzer des privaten Schlüssels ist in der Lage, auf den ursprünglichen Inhalt zuzugreifen.[44] [45] Generell kann zwischen voll-homomorpher und partiell-homomorpher Verschlüsselung unterschieden werden.[46] Voll-homomorphe Verschlüsselung erlaubt den Nutzenden die Berechnung beliebiger Funktionen ohne Einschränkungen hinsichtlich der Art der unterstützten Operationen oder ihrer Komplexität. Die partiell-homomorphe Verschlüsselung hingegen unterstützt lediglich die Durchführung von Additionen und Multiplikationen mit den geschützten Daten. In beiden Fällen erfordern komplexe Analysen und größere Datenmengen hohe Rechenkapazitäten und schränken die Anwendung homomorpher Verschlüsselung in der Praxis ein.[47]

Beispiel

Das PET-Unternehmen Duality hat 2020 in einem Bericht die theoretische Möglichkeit aufgezeigt, homomorphe Verschlüsselung für die Analyse von Daten in groß angelegten Studien zur Genomforschung zu nutzen.[49] Für die Studie wurde ein datenschützender Rahmen vorgestellt, der die Auswertung von 25.000 Individualdaten unter hohen Datenschutzanforderungen ermöglicht und die Rechenzeit gegenüber der vorherigen Auswertungsmethode um das 30-fache verkürzt.[50]

„There have been many misconceptions about homomorphic encryption – however, the technology is no magic."[48]

Kurt Rohloff, CTO & Co-Founder, Duality Technologies



44 Gentry, Craig (2009): A Fully Homomorphic Encryption Scheme.

45 Office of the Director of National Intelligence (2017): How would you explain homomorphic encryption?

46 Information Commissioner's Office (2022): Anonymisation, pseudonymisation and privacy enhancing technologies guidance, Chapter 5.

47 Information Commissioner's Office (2022): Anonymisation, pseudonymisation and privacy enhancing technologies guidance, Chapter 5.

48 Cohen, Ronen; Rohloff, Kurt & PUBLIC (2022): Interview über PETs

49 Eurekalert (2020): Duality Technologies researchers accelerate privacy-enhanced collaboration on genomic data.

50 Blatt, Marcel; Gusev, Alexander; Polyakov, Yuriy; Goldwasser, Shafi (2020): Secure large-scale genome-wide association studies using homomorphic encryption.

Verbreitung

Der Einsatz von homomorpher Verschlüsselung ist weder im privaten noch im öffentlichen Sektor weit vorangeschritten. Das liegt in erster Linie an der fehlenden Effizienz aktueller Modelle voll-homomorpher Verschlüsselung gegenüber der Analyse von unverschlüsselten Daten. Große Techunternehmen und kleinere Innovatoren übertragen derzeit die mathematischen Grundsätze auf skalierbare Lösungen und versuchen, die Effizienz zu erhöhen.[51] Insbesondere im Bereich des Cloud Computings birgt die Technologie jedoch großes Potenzial, zukünftig Kollaboration in Umfeldern zu ermöglichen, in denen hohe Datenschutzanforderungen und geringes Vertrauen zwischen den Akteuren vorherrschen.

E Secure Multiparty Computation

Bei welcher Herausforderung kann Secure Multiparty Computation der öffentlichen Verwaltung helfen?

- **Datenschutz**
- Informationssicherheit
- **Geheimschutz**

Überblick

Vorteile

Ermöglicht die Nutzung und den Schutz sensibler Daten. Stellt eine neue Lösung dar für das Outsourcing und die Freigabe von Daten in Fällen, in denen der Datenschutz bisher ein Hemmnis war.

Nachteile

Die komplexen Rechenvorgänge erfordern ein erhebliches Maß an Rechenleistung, das die zurzeit verfügbare Hardware an ihre Grenzen bringt. Dies hat hohe Kosten und lange Rechenzeiten zur Folge.

Marktreife

Gering – technische Limitationen schränken den Einsatz von homomorpher Verschlüsselung aktuell stark ein.[52]

Komplementarität

Homomorphe Verschlüsselung bietet bereits umfassende Möglichkeiten zum Schutz von Daten, kann jedoch beispielsweise mit SSI-Technologien kombiniert werden, um den Zugang zu Datensätzen zu regeln.

Funktionsweise

Secure Multiparty Computation (SMPC) ist ein kryptografisches Protokoll, das mehreren Parteien die Zusammenarbeit an gemeinsamen Daten erlaubt, ohne dass die einzelnen Parteien ihren eigenen Daten-Input offenlegen. Anwendungen, die auf diese Technologie zurückgreifen, erlauben somit einer oder mehreren Parteien, Analysen und Ergebnisse aus gemeinsamen Daten zu ziehen, und halten die einzelnen Daten doch geheim.[53]

SMPC-Anwendungen nutzen das Prinzip des „secret-sharing“, um Outputs sichtbar und Inputs geheim zu halten. Beim *secret-sharing* wird ein geheimer Wert in mehrere Anteile aufgeteilt und jeder Anteil wird an einen anderen Teilnehmer verteilt. Die jeweiligen Teilnehmenden kennen nur ihren eigenen Anteil des *secret* und haben keine Informationen über die anderen Anteile oder den ursprünglichen Geheimwert. Um das Geheimnis zu rekonstruieren, muss eine Mindestanzahl von Teilnehmern ihre Anteile nach einer vorgegebenen Regel kombinieren.[54]

51 Microsoft bietet mit Microsoft SEAL derzeit Open-Source-Bibliotheken für den Aufbau von Ende-zu-Ende-verschlüsselten Datenspeicherungs- und -berechnungsdiensten.

52 Homomorphic Encryption Standardisation (2023): Participants.

53 IEEE (2021): 2842-2021 - IEEE Recommended Practice for Secure Multi-Party Computation.

54 Information Commissioner's Office (2022): Anonymisation, pseudonymisation and privacy enhancing technologies guidance, Chapter 5.

Beispiel 1

Ein imaginäres Beispiel verdeutlicht die Funktionsweise: Eine Gewerkschaft möchte die durchschnittlichen krankheitsbedingten Fehltage von drei Mitgliedern errechnen. Die einzelnen Daten sind jedoch hochsensibel und die drei Mitglieder (A, B und C) möchten die Daten nicht ohne Weiteres transparent machen. Deshalb greifen sie auf SMPC zurück. Die SMPC-Anwendung teilt die Informationen jeder Partei in drei zufällig erzeugte „geheime Anteile" auf: Der Input von Person A – ihre Krankheitstage im vergangenen Jahr – beträgt 8 Tage. Dieser Betrag wird in geheime Anteile von 2, 5 und 1 aufgeteilt. Der Input von Person B beträgt 12 Tage, der von Person C 7 Tage. Person A behält einen seiner Anteile, den zweiten verteilt es an Person B und den dritten an Person C. Die Personen B und C verfahren mit ihren Eingangsdaten genauso. Wenn dieser Vorgang abgeschlossen ist, hat jede Person drei geheime Anteile. Person A hat zum Beispiel den geheimen Anteil, den es von seiner eigenen Eingabe behalten hat, zusammen mit einem geheimen Anteil von Person B und einem weiteren von Person C. Die geheimen Anteile können nicht verraten, wie hoch der Input der einzelnen Personen war – das bedeutet, Person A erfährt nicht die Anzahl der Fehlzeiten von Person B oder C. Jede Partei addiert anschließend ihre geheimen Anteile. So wird ein Teilergebnis berechnet – für jede Person und für die Gesamteingaben aller drei Personen. Die SMPC-Anwendung teilt dann die Gesamtsumme durch die Anzahl der Parteien – in diesem Fall drei –, das ergibt die durchschnittlichen Fehlzeiten der drei Personen: 9 Tage.

Beispiel 2

Der Boston Women's Workforce Council (BWWC) hatte es sich 2015 zur Aufgabe gemacht, die Unterschiede in der durchschnittlichen Bezahlung zwischen Frauen und Männern in verschiedenen Industrien möglichst wahrheitsgetreu zu erfassen, um auf dieser Grundlage besser informierte politische Entscheidungen zu treffen. Vor der Zusammenarbeit mit der Boston University war der BWWC dafür auf freiwillige Datenspenden von Unternehmen aus Boston angewiesen. Für die datenspendenden Organisationen bestand das Risiko, dass ihre Daten im Falle eines erfolgreichen Angriffes auf die Daten offengelegt würden, auch wenn sie dem BWWC in seinem Versprechen zur Diskretion vertrauten.[55] Die Softwarelösung der Boston University vermochte 2015 dieses Dilemma zu lösen. Mit der Plattform konnten deutlich mehr Unternehmen dafür gewonnen werden, ihre Daten für die Analyse zur Verfügung zu stellen, weil technologisch sichergestellt werden konnte, dass keine Partei Zugriff auf die einzelnen Dateninputs hat. Der BWWC verwendet die Lösung noch heute, um Daten über die Bezahlung marginalisierter Gruppen effektiv und sicher zu erfassen und veröffentlicht auf dieser Grundlage jährliche Reports.[56]

> *„The use of Secure Multiparty Computation in the public sector is being seriously discussed, but it has not yet been widely deployed."*[57]

Andrei Lapets, VP, Engineering & Applied Cryptography, Magnite; außerordentlicher Professor für die Praxis der Informatik, Boston University

55 Boston University (2015): Computational Thinking Breaks a Logjam.

56 BWWC (2022): Gender and Racial Wage Gaps in Boston by the Numbers.

57 Lapets, Andrei & PUBLIC (2023): Interview zu PETs.

Verbreitung

SMPC-Anwendungen können verschieden ausgestaltet werden und unterschiedliche Rollen vorsehen: In der Forschung und Entwicklung findet insbesondere die Peer-to-Peer-Variante Anwendung. Hier erhalten alle Beteiligten die gleichen Rollen und Rechte für den Input der Daten und die Berechnung des Outputs. Im Client-Server-Modell hingegen wird zwischen dem Serviceanbieter, der die Anwendung stellt, und externen Akteuren, die lediglich Dateninput liefern, unterschieden.[58] Beide Modelle befinden sich jedoch nach wie vor in einem frühen Entwicklungsstadium und werden weder in der Industrie noch im öffentlichen Sektor breitflächig angewendet. Zwei Hürden stellen sich für die Skalierung von SMPC-Anwendungen: Zum einen erlauben aktuelle SMPC-Protokolle nur einfache Rechnungen und keine komplexen Analysen. SMPC-Protokolle erfordern in der Regel mehrere Kommunikations- und Berechnungsrunden zwischen den Parteien, was rechenintensiv und zeitaufwändig sein kann, insbesondere bei der Arbeit mit großen Datensätzen. Zum anderen sind die Sicherheitsmechanismen von SMPC-Modellen durch böswillige partizipierende Parteien angreifbar: Die Parteien müssen sich auf das zu verwendende Protokoll einigen, Schlüssel und kryptografisches Material austauschen und sicherstellen, dass keine Partei vom Protokoll abweicht. Jeder Fehler oder jede böswillige Abweichung vom Protokoll können die Sicherheit und den Datenschutz der Berechnung gefährden.

Überblick

Vorteile

Ermöglicht die Zusammenarbeit an gemeinsamen Daten, ohne dass die einzelnen Parteien ihre Daten offenlegen und bietet damit hohen Schutz gegen externe Angriffe.

Nachteile

Erlaubt in der aktuellen Marktreife lediglich simple Rechnungen. Kann durch gezielte Manipulation der Beteiligten an Sicherheit einbüßen.

Marktreife

Gering - befindet sich noch im Entwicklungsprozess, Fortschritte lassen jedoch erste Anwendungsfälle im öffentlichen Sektor zu.

Komplementarität

Da der Output von SMPC-Modellen unverschlüsselt ausgegeben wird, wird die Technologie häufig mit Differential Privacy kombiniert.

58 Lapets, Andrei (2018): Accessible Privacy-Preserving Web-Based Data Analysis for Assessing and Addressing Economic Inequalities.

F Trusted-Execution-Environments (TEE)

Bei welcher Herausforderung können Trusted-Execution-Environments der öffentlichen Verwaltung helfen?

- Datenschutz
- Informationssicherheit
- Geheimschutz

Funktionsweise

TEEs sind isolierte Teile eines Prozessors, die es den Nutzer*innen erlauben, Teile des Codes oder Daten vom Rest des Betriebssystems zu isolieren und zu schützen. Der im TEE ausgeführte Code kann nicht eingesehen oder verändert werden, selbst wenn ein Angreifer in der Lage ist, bösartigen Code mit vollen Rechten auf demselben Prozessor auszuführen.[59] Die Sicherheitsmechanismen von TEEs bestehen gewöhnlich aus zwei Komponenten: Eine Hardwarekomponente sichert den Teil des Prozessors, der als TEE funktionieren soll. Softwarekomponenten steuern die Funktion des Codes oder die Verarbeitung der Daten innerhalb des isolierten Bereichs und organisieren den Austausch mit dem Rest des Betriebssystems. Trusted-Execution-Environments lassen sich abstrakt als sicherer Raum innerhalb eines Gebäudes verstehen, den nur befugtes Personal betreten kann um darin sensible Tätigkeiten auszuführen. Die Hardwarekomponenten des TEEs funktionieren als Tür, die lediglich autorisierten Personen Zugang gewährt. Die Softwarekomponenten des Trusted-Execution-Environments funktionieren als Manager, der die Aufsicht für alle Abläufe innerhalb des Raumes hat und darauf achtet, dass befugte Personen keine sicherheitsgefährdenden Gegenstände in den Raum bringen. Die Kombination aus Software- und Hardware-Komponente unterscheidet TEEs vom Sandboxing (softwarebasiert) und Hardware-Secure-Models (hardwarebasiert).

> *Trusted-Execution-Environments lassen sich abstrakt als sicherer Raum innerhalb eines Gebäudes verstehen, den nur befugtes Personal betreten kann um darin sensible Tätigkeiten auszuführen.*

TEEs sind daher besonders für die Anwendung in Umgebungen geeignet, in denen das Vertrauen zwischen zwei oder mehr Parteien gering ist und hohe Sicherheitsstandards für sensible Codes oder Daten erhalten werden sollen, ohne Einschränkungen in der Ausübung von Anwendungen hinzunehmen. Eine weitere, besondere Bedeutung, die auch für die öffentliche Verwaltung eine relevante Frage darstellt, wird ihnen für das Outsourcen von Anwendungen und Daten auf externen Server oder Clouds zugesprochen. Mit diesem Schritt

59 Seker, Ensar (2021): Trusted Execution Environment (TEE), Implementations, Drawbacks.

Verbreitung

Trusted-Execution-Environments sind bereits in vielen elektronischen Geräten des Alltags verbaut – wie etwa in Smartphones oder Smart-TVs. ARM und Intel haben mit „TrustZone“ und „Software Guard Extensions“ (SGX) die am weitesten verbreiteten TEE-Produkte entwickelt, die unter anderem auf Android-Geräten zur Anwendung kommen.[61, 62] Einige PET-Anbieter, die sogenannte Data-Clean-Rooms anbieten, nutzen TEEs als Sicherheitskomponente ihrer Lösung und kombinieren sie mit weiteren PETs, um den Datenschutz und die Nutzbarkeit der ausgehenden Daten zu erhöhen.

sind in vielen Fällen Sicherheitsbedenken gegenüber den Cloud- und Server-Anbietern und ihrer Infrastruktur verbunden. Denn selbst wenn Daten in der Cloud im Ruhezustand verschlüsselt sind, müssen sie für die Bearbeitung in der Regel entschlüsselt werden und sind damit anfälliger für externe Zugriffe. Da die Bearbeitung bei TEEs jedoch innerhalb des isolierten Bereichs stattfindet, bedeutet ein erfolgreicher Angriff auf die Cloud nicht die Preisgabe der eigenen Daten oder Codes. Weil die Daten innerhalb des isolierten Bereichs unverschlüsselt sind, entsteht außerdem kein zusätzlicher Rechenaufwand.

Das größte Risiko entsteht für TEEs durch sogenannte Side-Channel-Attacks, die Informationen aus der Art und Weise extrahieren, wie das TEE mit dem Rest des Systems kommuniziert.[60]

Überblick

Vorteile

Bietet hohe Sicherheitsstandards für sensible Codes oder Daten durch Isolation vom Hauptprozessor. Kann technologische Basis für das Outsourcen auf Clouds oder externe Server sein. Dank Verwendung unverschlüsselter Daten wird keine erhöhte Rechenkapazität erfordert.

Marktreife

Weit – viele Produkte haben bereits Marktreife und werden in der Breite verbaut.

Komplementarität

Wird häufig mit anderen PETs kombiniert, um Datenoutputs zu verschlüsseln oder zu manipulieren.

Nachteile

Die verarbeitbare Datenmenge in TEEs ist begrenzt. TEEs sind anfällig für Side-Channel-Attacks.

60 Information Commissioner's Office (2022): Anonymisation, pseudonymisation and privacy enhancing technologies guidance, Chapter 5.

61 Buchner, Nicolas; Kinkelin, Holger; Rezabek, Filip (2022): Survey on Trusted Execution Environments.

62 Android (n. a.): Vertrauensvolles TEE.

G Zero-Knowledge Proof

Bei welcher Herausforderung können Zero-Knowledge Proofs der öffentlichen Verwaltung helfen?

- Datenschutz
- Informationssicherheit
- Geheimschutz

> *"Mit Technologien wie dem Zero-Knowledge Proof wollen wir den Nutzenden die Kontrolle über die eigenen Daten zurückgeben."*
>
> **Helge Michael,** Head of Lissi Identity, Neosfern.

Funktionsweise

Der Zero-Knowledge Proof ist ein kryptografisches Protokoll, das die Verifizierung einer Information zwischen einer beweisenden Partei (Prover) und einer verifizierenden Partei (Verifier) sicherstellt - ohne, dass das Beweismittel selbst offengelegt wird. Zero-Knowledge Proofs schaffen damit vertrauenswürdige Transaktionen, bei denen mathematische Verfahren zum Schutz der Privatsphäre der Nutzenden eingesetzt werden. Sie stellen eine Möglichkeit dar, Self-Sovereign Identity (SSI) Systeme technologisch so auszugestalten, dass Verifizierungen datensparsam abgewickelt werden können.[63,64] Damit schaffen sie mögliche Abhilfe für einen Balanceakt, der bei digitalen Verifizierungen häufig zwischen dem Anspruch des Beweisenden (Prover) auf Datenschutz und dem berechtigten Interesse zur Prüfung von Informationen durch den Verifizierenden (Verifier) besteht.[65] Mit Hilfe dieser Technologie könnte zum Beispiel ein Gast in einer Bar (Prover) dem Türsteher (Verifier) beweisen, dass er oder sie volljährig ist, ohne dass er oder sie das eigene Alter preisgeben muss.

Beispiel

Die grundsätzliche Funktion der Technologie lässt sich am besten an einem Beispiel beschreiben, dessen Grundlage der Kryptograf Jean-Jacques Quisquater 1990 gelegt hat:[66] Das Beispiel dreht sich um eine Höhle. Im Eingang der Höhle eröffnen sich zwei Gänge, einer zur Rechten und einer zur Linken. Beide Gänge sind im hinteren Teil der Höhle durch eine Tür miteinander verbunden, die sich nur mit einem Zauberwort, dem Geheimnis, öffnen lässt. Wer also den rechten Gang geht und das Geheimnis nicht kennt, kann auch nur zur Rechten wieder hinauskommen. Wer hingegen das Geheimnis kennt, kann auch aus dem linken Gang zum Ausgang der Höhle zurückkehren. Die Protagonist*innen des Beispiels sind Vincent, der als Verifier (Verifizierender) auftritt, und Paula, die als Prover (Beweisende) auftritt. Paula kennt das geheime Wort, möchte es aber nicht preisgeben. Vincent weiß um den Aufbau der Höhle und die Funktion der Tür und möchte wissen, ob Paula das Geheimnis wirklich kennt. Um beiden Wünschen gerecht zu werden, testet Vincent Paulas Wissen. Vincent stellt sich dafür zunächst außerhalb der Höhle auf und lässt Paula einen Eingang ihrer Wahl gehen. Sobald Paula im Inneren der Höhle ist, stellt sich Vincent an den Eingang der Höhle und ruft Paula zu, auf welcher Seite sie aus der Höhle herauskommen soll. Dieses Spiel wiederholen beide 25 mal hintereinander. In allen 25 Gelegenheiten kommt Paula zu der Seite aus der Höhle, die Vincent ihr ansagt. Angenommen, Paula kennt das Geheimnis nicht: In diesem Fall hat sie mit jedem Versuch eine Chance von 50%, dass Vincent

63 van Rijmenam, Mark (2019): How Zero-Knowledge Proof Increases Your Privacy While Enabling Trustless Transactions.

64 SSIs werden im folgenden Kapitel grundsätzlich behandelt.

65 Fraunhofer FIT (2021): Self-Sovereign Identity Foundations, Applications, and Potentials of Portable Digital Identities.

66 Quisquater, Jean-Jacques (1990): How to Explain zero-Knowledge Proofs to your Children.

zufällig den Weg als Ausgang fordert, zu dem sie auch hineingelangt ist. Das wäre bei wenigen Wiederholungen möglich - wird bei jedem weiteren Versuch jedoch unwahrscheinlicher. Nach dem 25ten erfolgreichen Versuch kann Vincent zu einer Wahrscheinlichkeit von etwa 33 Millionen (2^{25}) zu eins davon ausgehen, dass Paula das Geheimnis kennt, ohne dass sie das Geheimnis preisgeben muss.

Auch für die öffentliche Verwaltung bietet die Technologie als konkrete Ausgestaltung von Verifizierungen in SSI-Systemen relevante Anwendungsmöglichkeiten: Wenn beispielsweise ein Bürger bescheinigen möchte, dass er für ein Sozialhilfeprogramm leistungsberechtigt ist, dann muss er in der Regel Informationen vorlegen, die seine Berechtigung beweisen. In herkömmlichen Verfahren muss er dafür ganze Dokumente über seine finanzielle berufliche Situation offenlegen. Wenn das Verifizierungssystem für diesen Prozess auf Zero-Knowledge Proofs zurückgreift, dann kann der Bürger seine Berechtigung beweisen, ohne dass er die dafür benötigte - in vielen Fällen hochsensible - Information offenlegt.

Anwendungen, die sich der Technologie dieses Protokolls zu Nutzen machen, simulieren die Interaktion beider Parteien in kürzester Zeit und erlauben so datensparsame und sichere Verifizierung. Zero-Knowledge Proofs werden deshalb als wichtiger technologischer Baustein von SSI-Systemen anerkannt.[67]

Verbreitung

Zero-Knowledge Proofs werden mathematisch schon seit über 30 Jahren erforscht und entwickelt. Die technische Umsetzung in Industrie und Verwaltung lässt jedoch noch auf sich warten. Mit der gesteigerten Aufmerksamkeit für Distributed-Ledger Technologien wie Blockchain gewinnt jedoch auch der Zero-Knowledge Proof an Relevanz, weil er Blockchain-basierte Prozesse datensparsamer und recheneffizienter ausgestalten kann.[68] Für den verbreiteten Einsatz im öffentlichen Sektor bedarf es insbesondere der Standardisierung und Zulassung der Protokolle durch das BSI, auf denen Zero-Knowledge Proofs basieren.

Überblick

Vorteile

Kann Verifizierungen durchführen, ohne Beweismittel freizugeben.

Nachteile

Die technische Infrastruktur muss für die Anwendung von Zero-Knowledge Proofs ausgelegt sein.

Marktreife

Gering - Trotz steigendem Interesse aus dem öffentlichem Sektor und der Industrie sind Standardisierungen und Produkte noch in der Entwicklungsphase.

Komplementarität

Wird aufgrund der klar abgegrenzten Anwendung selten mit anderen PETs kombiniert.

67 Fraunhofer FIT (2021): Self-Sovereign Identity Foundations, Applications, and Potentials of Portable Digital Identities.

68 Eberhardt, Jacob; Tai, Stefan (2021): ZoKrates - Scalable Privacy-Preserving Off-Chain Computations.

4 Exkurs: Self-Sovereign Identity Systeme

Self-Sovereign Identities (SSIs) gelten als neues Paradigma für das Management digitaler Identitäten.[69] In Verbindung mit anderen Technologien bilden SSIs die Grundlage, Informationen zwischen unterschiedlichen Instanzen digital auszutauschen und zu verifizieren. SSI-Systeme greifen dafür auf Grundsätze der Kryptographie und Distributed Ledger Technologien (DLS) zurück und gewinnen mit der Entwicklung beider Felder an Bedeutung - auch für die Verwaltung.[70] SSIs sind dabei nicht originär *Privacy-Enhancing*: Sie können beispielsweise technologisch so ausgestaltet werden, dass sie auf radikaler Transparenz beruhen und Grundsätzen des Datenschutzes wie der Datenminimierung nicht zwingend entsprechen. SSI-Systeme können jedoch auch so ausgestaltet werden, dass Verifizierungen nicht nur sicher, sondern auch datensparsam ablaufen. Entscheidend sind dafür die spezifischen technologischen Komponenten, mit denen SSI-Systeme gebaut werden. Der Schutz sensibler Daten kann bei Identifikations- und Authentifizierungsverfahren beispielsweise durch Zero-Knowledge Proofs oder Selective Disclosure gestärkt werden. Gleichzeitig bieten SSIs das Potenzial, technische Systeme durch Dezentralisierung sicherer zu machen.

Identität und Verifizierung in der digitalen Welt stellen Nutzer*innen und Anbieter bisher vor eine Abwägung zwischen Datenschutz, Sicherheit und Benutzerfreundlichkeit. In herkömmlichen digitalen Verfahren können nicht alle drei Eigenschaften gleichzeitig gewährleistet werden. Eine hohe Sicherheit bedeutet häufig aufwändige Verfahren für Nutzer*innen. Komfortable Lösungen gehen auf Kosten des persönlichen Datenschutzes oder sind weniger sicher. Ein Übermaß an Komfortabilität im Sinne der Benutzerfreundlichkeit hingegen wird schnell zum Risiko für die Sicherheit.

Beispiele für herkömmliche Identifikations- und Authentifizierungsverfahren, bei denen Abwägungen getroffen werden

Sicherheit zu Lasten von Benutzerfreundlichkeit

Per Video oder sogar durch einen Besuch der Poststation durchgeführte Identifizierungsverfahren

Benutzerfreundlichkeit zu Lasten von Datenschutz

Sogenannte Single Sign-on-Dienste, wie “mit Google einloggen” oder vergleichbare Angebote von Facebook und LinkedIn, die Nutzerdaten zentral verwalten

Benutzerfreundlichkeit zu Lasten von Sicherheit

Die Verwendung von unsicheren Passwörtern oder Pin-Codes, die einfach zu merken aber auch nur geringen Schutz bieten

69 Fraunhofer FIT (2021): Self-Sovereign Identity Foundations, Applications, and Potentials of Portable Digital Identities.

70 Bayerisches Staatsministerium für Digitales (n.A.): Digitale Identitäten - SSI.

Während Nutzer*innen im privaten Bereich oft bereit sind, Kompromisse einzugehen, bei denen sie bewusst oder unbewusst den Schutz ihrer persönlichen Daten oder die Sicherheit vor unbefugten Zugang zu ihren Online-Konten aufgeben, steht in Öffentlichen Verwaltungen diese Abwägung nicht zur Disposition. Der bisherige Umgang mit digitalen Identitäten und der Verifizierung von Informationen ist für die Verwaltung daher ein großes Hindernis, digitale Dienstleistungen zugänglich zu machen und stellt zum Beispiel eine Blockade für eine nutzerfreundliche Umsetzung des Onlinezugangsgesetzes (OZG) dar.

Von auf SSI-Grundsätzen beruhenden Verfahren wird erwartet, die notwendige Abwägung zwischen Datenschutz, Sicherheit und Benutzerfreundlichkeit abzulösen. Der Anspruch ist es, Identitätsnachweise und Verifizierungen benutzerfreundlich, datensparsam und sicher zu ermöglichen. Für den öffentlichen Sektor sind SSIs daher ein potentieller Game Changer: Sie können im eGovernment den Zugang zu Dienstleistungen erheblich vereinfachen, aber auch Zugriffsrechte auf Informationen steuern und damit interne Prozesse im Bereich des Geheimschutzes erneuern.

Funktionsweise

Die essentiellen Bausteine von SSI-Lösungen sind *Verifiable Credentials*, *Verifiable Presentations* und die drei Rollen von *Issuer*, *Holder* und *Verifier* sowie *Decentral Identifiers* und *Digital Wallets*.

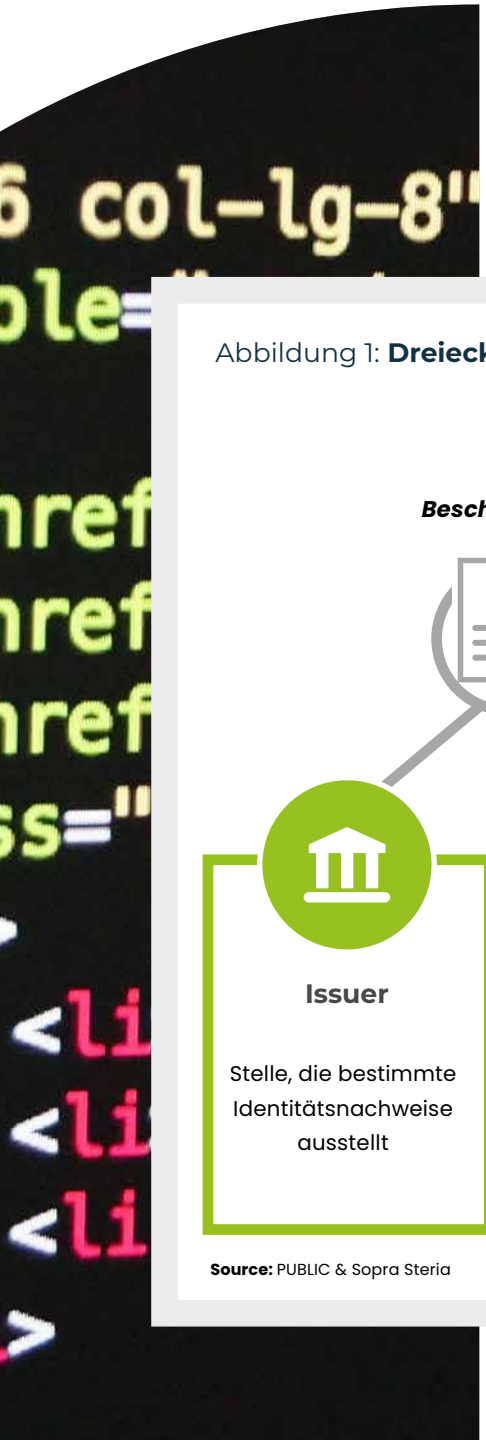
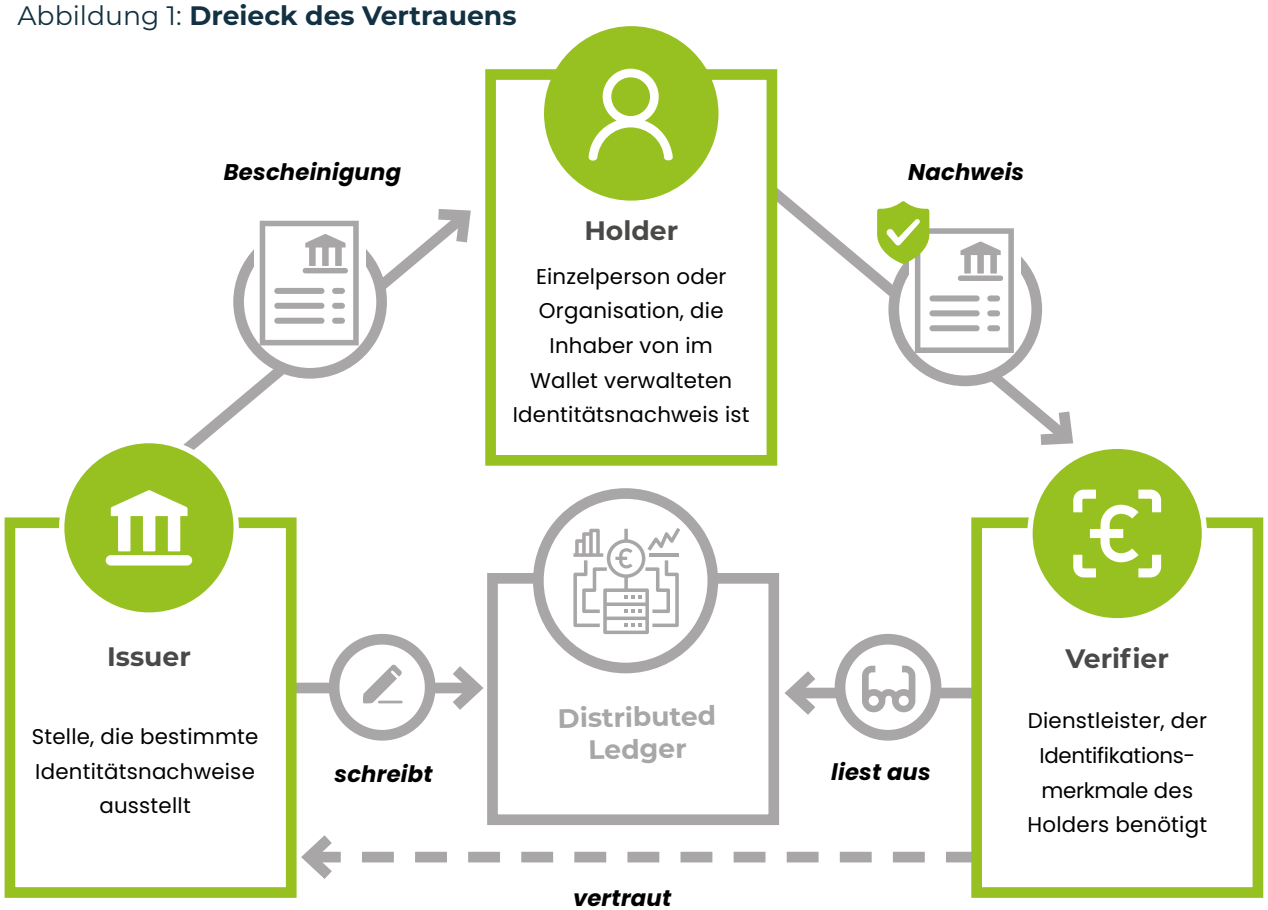


Abbildung 1: **Dreieck des Vertrauens**



Source: PUBLIC & Sopra Steria

Das Verifiable Credential ist zentraler Bestandteil, das die Form eines Zertifikats einnimmt. Der Prozess von SSIs baut auf dem Dreieck des Vertrauens der drei Rollen auf, in dem zwei Akteure bilateral und verschlüsselt kommunizieren: Ein durch den *Issuer* ausgestelltes *Verifiable Credential* gehört dem *Holder*. Der *Holder* kann dem *Verifier* eine *Verifiable Presentation* einzelner Attribute des *Verifiable Credentials* präsentieren, der dies mit Hilfe von *Decentral Identifiers* validiert und so die Echtheit der *Verifiable Presentation* bestätigt. Dieser Prozess der Verifizierung basiert auf Technologien der Blockchain, über die der *Verifier* Zugriff auf Distributed Ledgers hat, die die Echtheit von Nachweisen bestätigen, aber keine Informationen über den Holder selbst enthalten. [71]

Das Dreieck des Vertrauens kann man sich beispielsweise für den Nachweis über den Besitz der Fahrerlaubnis vorstellen: Das Kraftfahrt-Bundesamt agiert als Issuer der Fahrerlaubnis eines *Holders* (*Verifiable Credential*), der seine Berechtigung zum Führen eines Fahrzeugs nachweisen muss. *Verifier* sind in diesem Fall vielfältig, es können Autovermietungen, Versicherungen oder Verkehrsbehörden sein. Der Fahrzeugführer präsentiert der Versicherung den Nachweis über seine Fahrerlaubnis (*Verifiable Presentation*). Der *Verifier* bekommt die Echtheit über den Distributed Ledger beispielsweise mit dem Scan eines QR-Codes bestätigt. Weitere mögliche Inhalte von Wallets sind vielfältig. Der Personalausweis wird als Basis-ID gesehen. Ein Studierendenausweis im Wallet kann als Berechtigungsnachweis für Ermäßigungen in Museen verwendet werden. Bildungsabschlüsse können für Bewerbungsverfahren abgelegt und vom potentiell zukünftigen Arbeitgeber verifiziert werden.

Komponenten für Datenschutz und Sicherheit

In der spezifischen Architektur von SSIs bieten sich verschiedene Prinzipien und Technologien an, um Verifizierungen so durchzuführen, dass sie Privatsphäre wahren und Daten schützen: Neben der bereits beschriebenen Technologie des Zero-Knowledge Proofs ermöglicht etwa der Prozess des *Selective Disclosure,* dass ausschließlich die Informationen übermittelt werden, die für eine Verifizierung auch relevant sind. Dieser Prozess ist vergleichbar mit der Idee, für den Nachweis des Alters nicht den gesamten Personalausweis vorzeigen zu müssen, sondern lediglich das Geburtsdatum. Weiter lässt sich in SSI-Systemen ein *Zero Trust Model* aufbauen, das wichtige Sicherheitskomponenten gegenüber traditionellen Systemen mitbringt. In SSI-Systemen bedeutet Zero Trust, dass jede Authentifizierungs- und Autorisierungsanfrage geprüft wird, unabhängig von der Identität des*der Benutzer*in oder des verwendeten Geräts. Zero Trust trägt dazu bei, unbefugten Zugriff zu verhindern und das Risiko von Sicherheitsverletzungen zu verringern, indem es davon ausgeht, dass jede Anfrage potenziell bösartig ist und eine Überprüfung erfordert. Es ist damit ein proaktiver und granularer Sicherheitsansatz, der zum Schutz sensibler Daten und Ressourcen beitragen kann - selbst in komplexen und dynamischen Umgebungen.[72] Mit dieser Voraussetzung bieten SSI-Systeme gegenüber zentralisierten Systemen die Möglichkeit, solche Angriffe zu erkennen, zu isolieren und zu blockieren, die über den erfolgreichen Angriff auf einen Bestandteil das gesamte System infiltrieren möchten.

71 Strüker, Jens et al. (2021): Self-Sovereign Identity Foundations, Applications, and Potentials of Portable Digital Identities.

72 Trust over IP Foundation (2022): No, I don't trust you - implementing zero-trust architecture in the world of self-sovereign identity (SSI).

Ausblick

Damit SSI-Systeme ihre Potenziale im öffentlichen Sektor entfalten, müssen Verwaltungen verschiedene Hürden nehmen: Ihr Aufbau ist hochkomplex und erfordert hohe technische Expertise. Teile der kryptographischen Verfahren, auf denen SSI-Systeme beruhen, sind außerdem noch nicht ausreichend durch das BSI anerkannt. Hinzu kommt eine Einschränkung, die sich wie bei PETs auch in SSI-Systemen wiederfindet: Erhöhte Sicherheitsanforderungen können auch in SSI-Systemen mit hohen Anforderungen an die Rechenleistung und damit zu langsameren Abläufen führen - etwa wenn aufwändigere Prüfungen von Behauptungen eingefordert werden.

Dennoch können öffentliche Verwaltungen mit SSIs eine Architektur aufbauen, die Verifizierungen ermöglicht, dabei sensible Daten schützt und Sicherheitsvorteile gegenüber traditionellen, zentralisierten IT-Systemen bietet. Damit eröffnen sich dem öffentlichen Sektor eine Vielzahl an Einsatzmöglichkeiten: Mit ihrer Hilfe können Zugriffsrechte für sensible Informationen gesteuert oder auch digitale Transaktionen im Bereich von digitalen Verwaltungsdienstleistungen gestaltet und gesichert werden. Die Möglichkeit zum Einsatz von Zero-Knowledge Proofs und Selective Disclosure machen SSIs potenziell zu Daten-schützenden Systemen. Ihr dezentraler Charakter bietet außerdem Schutz vor Angriffen, die das schwächste Glied als Einfallstor für die Kompromittierung ganzer Systeme wählen. Deshalb wird von SSIs erwartet, die Grundlage für Schutzmechanismen in IT-Systemen zu bilden, die eine umfassende Resilienz gegenüber Angriffen bieten.

Deshalb wird von SSIs erwartet, die Grundlage für Schutzmechanismen in IT-Systemen zu bilden, die eine umfassende Resilienz gegenüber Angriffen bieten.

5 Was sind Privacy-Enhancing Technologies nicht?

Es gehört zur Dynamik wachsender Technologiemärkte, dass sich überhöhte Erwartungen aufbauen und ein klarer Fokus dafür erschwert wird, was bestimmte Lösungen tatsächlich leisten können. Gleiches gilt auch für die Entwicklung von PETs. Anbieter stehen vor der Herausforderung, den Use-Case ihres Produktes deutlich zu machen, ohne potenziellen Kunden Mehrwerte zu versprechen, die sie nicht leisten können. Organisationen hingegen suchen in vielen Fällen nach ganzheitlichen Lösungen, insbesondere wenn es um den DSGVO-konformen Umgang mit Daten geht. Für einen mündigen Blick auf die Entwicklung von PETs lohnt sich deshalb die Vergegenwärtigung ihrer Grenzen und dessen, was PETs nicht leisten können. Diese Grenzen liegen häufig darin begründet, dass einzelne PET-Lösungen nur einen Baustein für die größere Aufgabe des Datenmanagements erfüllen. Trotz der folgenden Einschränkungen sind PETs in ihrer Kombination und Umsetzung in konkrete, technologiebasierte Produkte eine wertvolle Komponente für Gesamtlösungen.

> *„PET should not be regarded as ‚silver bullet' solutions. They cannot substitute legal frameworks but operate within them, so that their applications will need to be combined with legally binding and enforceable obligations to protect privacy and data protection rights."*[73]
>
> OECD (2023)

Vier Grenzen von PETs sind von besonderer Bedeutung:

1. **PETs sind kein Ersatz für Datenstrategien:** PETs sind Bausteine für die sichere Nutzung von Daten – strategische, übergeordnete Überlegungen zu ihrer Implementierung sind jedoch nach wie vor erforderlich. PETs kommen meistens dann zielführend zum Einsatz, wenn sie in eine ganzheitliche Datenstrategie eingebettet werden.
2. **PETs sind kein Ersatz für IT-Sicherheit:** Die Sicherheit von IT-Systemen muss durch Kernkomponenten und Standards umgesetzt werden, die zum Beispiel vom BSI festgehalten sind. PETs können darauf aufbauen und einen neuen Umgang mit Daten ermöglichen, bilden aber nicht den Grundstein für die Sicherheit von IT-Systemen.
3. **PETs sind keine allumfassenden Lösungen zur Einhaltung der DSGVO:** Sie garantieren nicht die Einhaltung rechtlicher Grundsätze, können aber Funktionen bereitstellen, die eine Organisation in ihrer Auslegung der Vorschriften unterstützen.
4. **PETs sind kein Ersatz für Datenethikprogramme:** PETs müssen mit Sorgfalt implementiert werden, um die Grundsätze der Datenethik zu erfüllen, sie können Datenethikprogramme aber nicht ersetzen.

73 OECD (2023): Emerging Privacy Enhancing Technologies: Current Regulatory and Policy Approaches.

Der internationale PET-Markt

Grundlagentechnologien – von homomorpher Verschlüsselung bis zu Differential Privacy – bilden Grundpfeiler, müssen jedoch für den Einsatz in der Praxis in praktikable Anwendungen gegossen werden.

Dafür nehmen innovative Unternehmen ihre Funktionen auf und entwickeln Produkte, die die Technologien mit weiteren wichtigen Komponenten anreichern und an konkreten Herausforderungen ausrichten. Als Produkte, die den Weg zu den Nutzenden finden, entstehen so etwa Softwarelösungen, die auf dem Server des Klienten installiert werden können. Diese Unternehmen konkurrieren mit ihren Lösungen auf einem Markt, über den das folgende Kapitel einen Überblick verschafft.

Das Kapitel betrachtet Zeitpunkt und Ort der Gründung, die verwendeten Grundlagentechnologien sowie Zahlen zu investiertem Risikokapital. Die Erkenntnisse deuten auf einen wachsenden Markt hin, auf dem deutsche Anbieter bisher unterrepräsentiert sind. Das Interesse bestimmter Branchen ist ein Indiz für den Mehrwert, den PET-Lösungen bieten. Marktreife Lösungen für den öffentlichen Sektor beschränken sich bisher weitestgehend auf den Gesundheitssektor. Best-Practice-Beispiele aus anderen Ländern zeigen, wie Deutschland die Attraktivität als Standort verbessern und von Interaktionen mit der Wissenschaft und PET-Anbietern profitieren kann.

1. Seit wann entwickelt sich der PET-Markt?

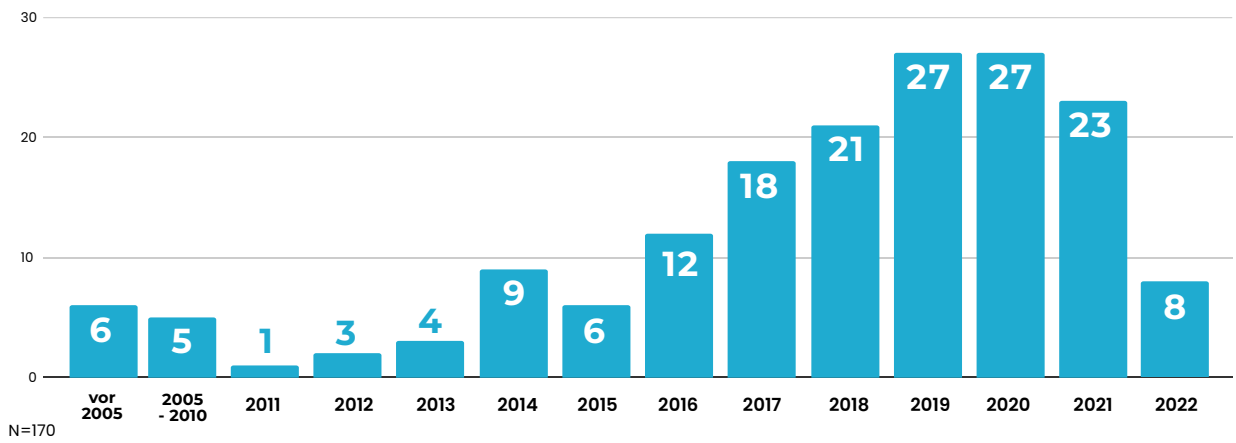
Der Blick auf die Gründungsjahre von PET-Unternehmen beantwortet zwei wichtige Fragen: Zum einen lässt er Rückschlüsse darüber zu, wie neuartig PET-Lösungen tatsächlich sind, zum anderen kann er Hinweise darauf geben, welche Umstände Impulse für die Entwicklung des PET-Markts sind. Abbildung 2 zeigt die Gründungsjahre der PET-Unternehmen unserer Analyse. Ein erster bemerkenswerter Höhepunkt bei den Neugründungen lässt sich 2014 verzeichnen, auf den ein nahezu jährlicher Anstieg an Gründungen bis 2019 folgt. Seit 2020 ist ein Abflachen der Kurve zu erkennen.

Für Neugründungen in Europa lässt sich die Verabschiedung der DSGVO als wichtiger Impuls für die Nachfrage verstehen.

Für Neugründungen in Europa lässt sich die Verabschiedung der DSGVO als wichtiger Impuls für die Nachfrage verstehen. Der erste Entwurf der Verordnung wurde 2012 von der EU-Kommission präsentiert, 2016 finalisiert und 2018 verabschiedet. Der zeitliche Zusammenhang mit Neugründungen von PET-Unternehmen ist nicht nur auf Abbildung 2 zu erkennen, sondern wurde auch von verschiedenen Interviewpartner*innen bestätigt:

> *„Data regulations tend to increase the demand for PET solutions and often act as an effective leverage on the market."*[74]
>
> ***Alejandro Russo,*** *Co-Founder & CEO von DPella*

Abbildung 2: **Gründungsjahre von PET-Unternehmen**



74 Gaboardi, Marco; Russo, Alejandro & PUBLIC (2023): Interview über PETs.

2. Wo werden PET-Lösungen entwickelt?

Der Blick auf die regionale Verteilung von PET-Unternehmen kann Erkenntnisse darüber liefern, wo sich regionale Hotspots für die Entwicklung von PET-Lösungen befinden und in welchem Verhältnis der europäische Markt gegenüber anderen internationalen Märkten steht. Abbildung 3 zeigt die Verteilung von PET-Unternehmen unserer Analyse über die Kontinente. Mit beinahe der Hälfte aller PET-Unternehmen dominiert der nordamerikanische Kontinent den internationalen PET-Markt. Wenn man jedoch den europäischen Markt inklusive dem Vereinigten Königreich und der Schweiz betrachtet, dann gleicht sich die Anzahl der PET-Unternehmen in Europa der in den USA an.[75]

PET-Hotspot Schweiz: *Mit über 5% der weltweiten PET-Unternehmen tritt die Schweiz in Relation zu ihrer Größe als PET-Hotspot auf. Das Land weist in unserer Analyse damit etwa so viele PET-Unternehmen auf wie Deutschland.*

Forschungsförderung und Innovationswettbewerbe

Der Gründungsstandort ist für PET-Unternehmen ein entscheidender Erfolgsfaktor. Ein günstiges Umfeld hängt von verschiedenen Bedingungen ab, die in einem Innovationsökosystem durch Interaktionen zwischen verschiedenen Akteuren geschaffen werden. Zielgerichtete Interaktionen von Akteuren aus Wissenschaft, Wirtschaft und öffentlicher Verwaltung tragen dazu bei, einen attraktiven Standort für die Gründung und den Unternehmenserfolg zu schaffen. Öffentliche Institutionen können durch Forschungsförderungsprogramme und Innovationswettbewerbe dazu beitragen, ein erfolgreiches Innovationsökosystem zu kreieren.

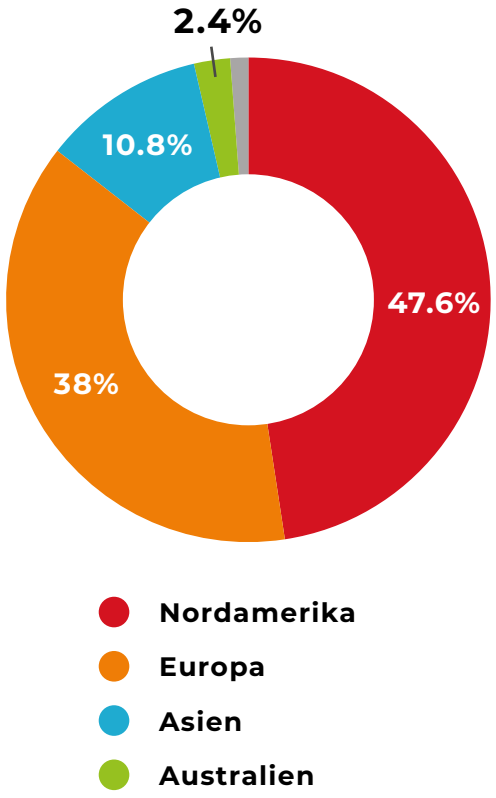
75 Europa (inkl. UK und CH)= 63; USA=68.

Eine OECD-Studie[76] hat in Kanada, der Türkei, Großbritannien, Singapur sowie den USA Programme mit Vorbildcharakter identifiziert, welche die Forschung und Entwicklung im Bereich von PETs fördern. Beispielsweise hat Singapur 2022 das „Digital Trust Centre" initiiert, das mit einer Investition von umgerechnet 35 Millionen Euro durch öffentliche Forschungseinrichtungen gefördert wird. Ziel des Digital Trust Centre an der Nanyang Technological University in Singapur ist es, Spitzenforschung im Bereich von Vertrauens- und Sicherheitstechnologien zu betreiben, die Fähigkeiten talentierter Forscher*innen auszubilden und Ideen aus der Forschung in marktreife Lösungen umzusetzen.

Innovationswettbewerbe beteiligen neben Wissenschaftler*innen auch Unternehmen. Beispiele für Innovationswettbewerbe zur Förderung von PETs finden sich unter anderem in Frankreich[77], Mexiko, Großbritannien und den USA. Großbritannien und die USA haben 2022 zusammen einen mit 1,6 Millionen USD Preisgeld dotierten Innovationswettbewerb initiiert.[78] An dem Format haben sich in der ersten Phase rund 60 Teams aus der Wissenschaft und kommerziellen PET-Unternehmen beteiligt, die Prototypen für PET-Lösungen im Bereich der Bekämpfung von Finanzkriminalität und von Pandemien entwickelten. Unter ihnen fanden sich auch führende PET-Anbieter der Marktstudie. Der britisch-amerikanische Innovationswettbewerb zeigt, wie die Interaktion von öffentlichen Institutionen mit kommerziellen PET-Anbietern in Form von Innovationswettbewerben dazu beiträgt, PET-Lösungen aus der Forschung und der Privatwirtschaft auf konkrete Anwendungsfälle in der öffentlichen Verwaltung zu übertragen.

> *„The U.S.-U.K. PETs Challenge elevates awareness about the opportunity while closing the gap between research and reality to realize the societal benefit of these technologies."*
>
> *Prämiertes Team eines PET-Anbieters*[79]

Abbildung 3: **Verteilung der PET-Unternehmen nach Kontinenten**



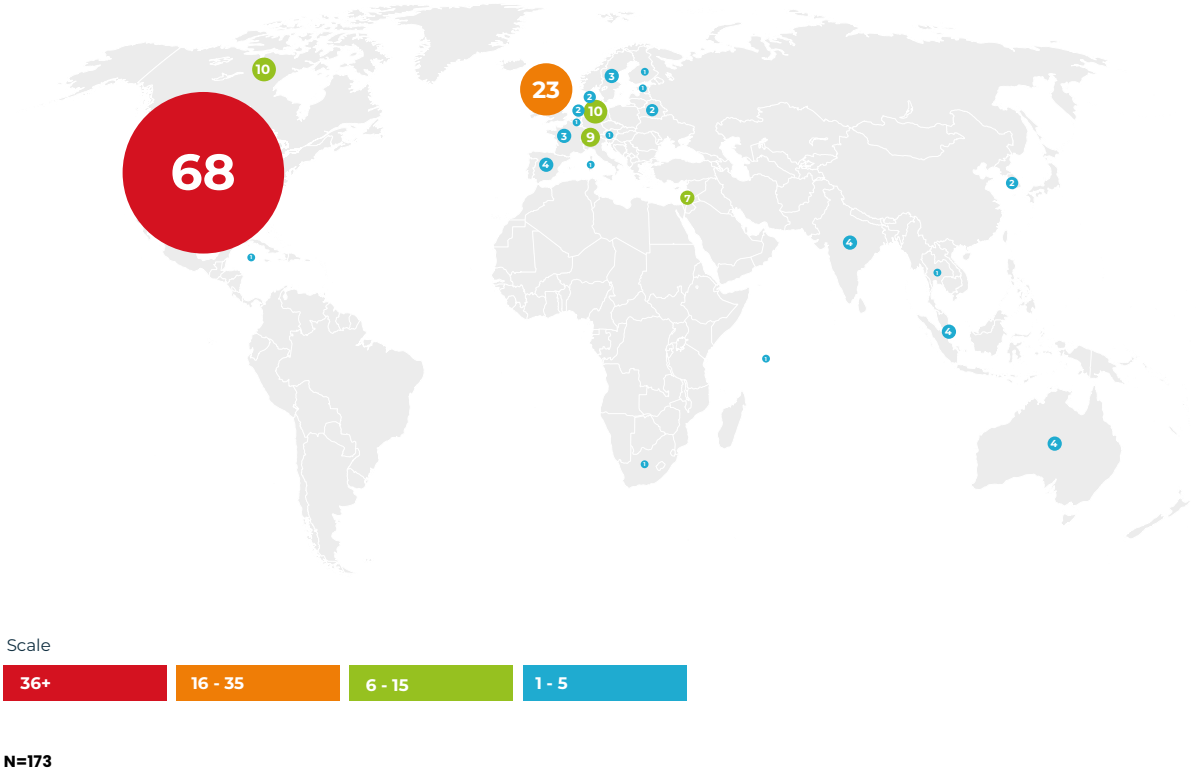
76 OECD (2023): Emerging Privacy Enhancing Technologies: Current Regulatory and Policy Approaches.

77 National Commission on Informatics and Liberty (2022): Launch of 7th edition of CNIL-Inria Privacy Award.

78 White House (2022): U.S. and U.K. Launch Innovation Prize Challenges in Privacy-Enhancing Technologies to Tackle Financial Crime and Public Health Emergencies.

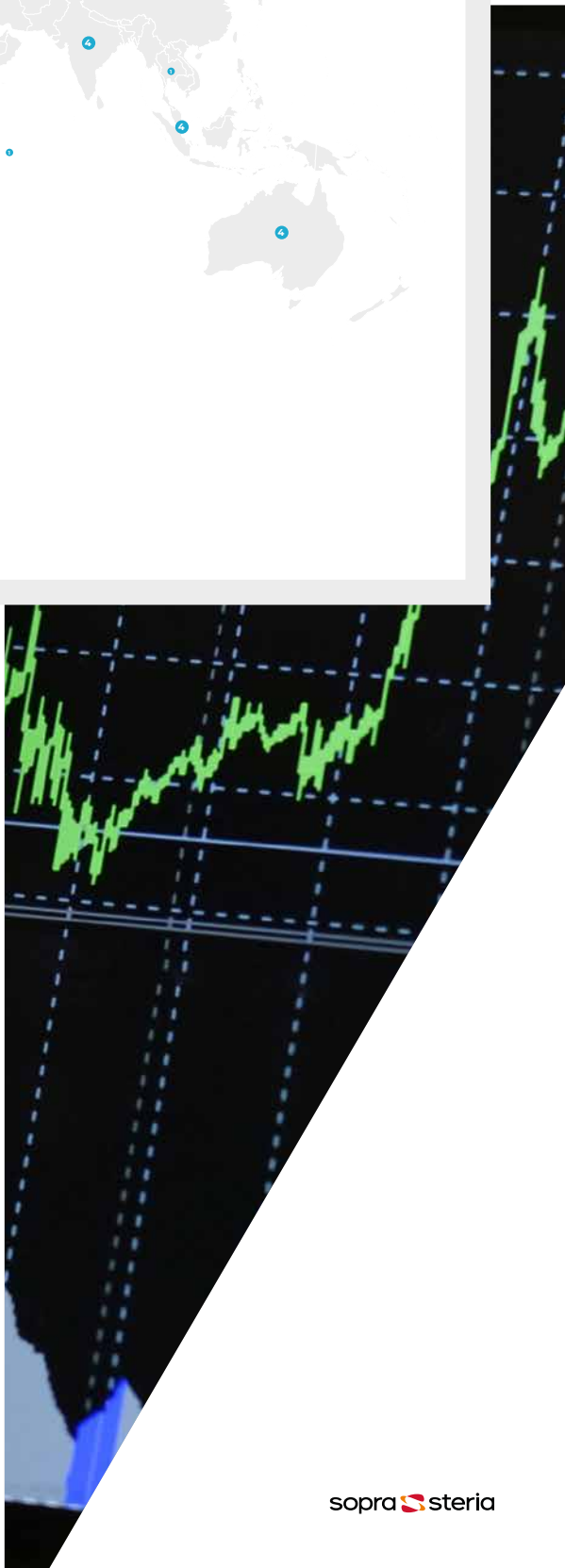
79 Chung (2022): Meet the Winners of the U.S. PETs Prize Challenge.

Abbildung 4: **Geographische Verteilung der PET-Unternehmen**



Deutschland hinkt hinterher

In Deutschland wurden insgesamt nur zehn Unternehmen identifiziert, die PETs als Kern der von ihnen angebotenen Dienstleistungen verwenden. Bis auf zwei Ausnahmen wurden alle deutschen PET-Anbieter nach 2016 gegründet. Mit einer Risikokapitalfinanzierung von mehr als 10 Millionen Euro ist Apheris das in Deutschland am besten finanzierte PET-Unternehmen. Neben Apheris konnten die deutschen Start-ups Edgeless Systems und Adaptant Solutions ein Funding von jeweils mehr als einer Million Euro einsammeln. Insgesamt haben deutsche PET-Unternehmen rund 16 Millionen Euro Funding erhalten. Obwohl die Anzahl deutscher PET-Unternehmen fast 6 % aller globalen PET-Unternehmen ausmacht, wurden nur 0,76 % des globalen Fundings in PET-Anbieter in Deutschland investiert.

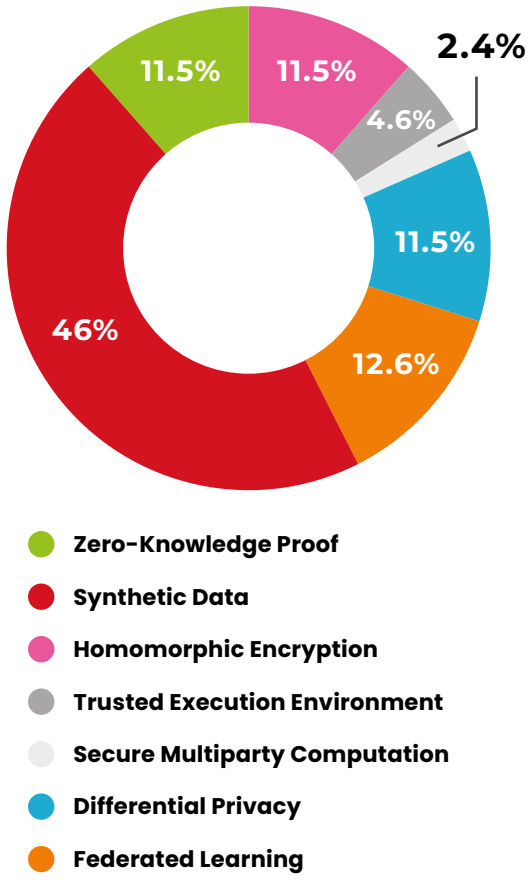


3. Welche Technologien dominieren den Markt?

PET-Unternehmen greifen zur Entwicklung ihrer Lösungen auf unterschiedliche PETs zurück. Die Lösungen stehen nicht zwingend in Konkurrenz zueinander. Die Analyse der verwendeten Technologien gibt jedoch wertvolle Hinweise auf aktuelle Verteilungen und Prioritäten auf dem internationalen Markt für PETs. Abbildung 5 zeigt, dass eine große Anzahl der Unternehmen Lösungen anbietet, die auf synthetische Daten zurückgreifen – beinahe jedes zweite Unternehmen setzt auf diese Technologie. Unternehmen, die Zero-Knowledge-Proofs, Differential Privacy oder homomorphe Verschlüsselung einbinden, folgen mit etwa 10 % aller Unternehmen etwa gleichauf. Der Anteil an Unternehmen, die SMPC oder TEE anbieten, fällt mit etwa 5 % jedoch ab.

Abbildung 5 zeigt, dass eine große Anzahl der Unternehmen Lösungen anbietet, die auf synthetische Daten zurückgreifen – beinahe jedes zweite Unternehmen setzt auf diese Technologie.

Abbildung 5: **Unternehmen pro Technologie**



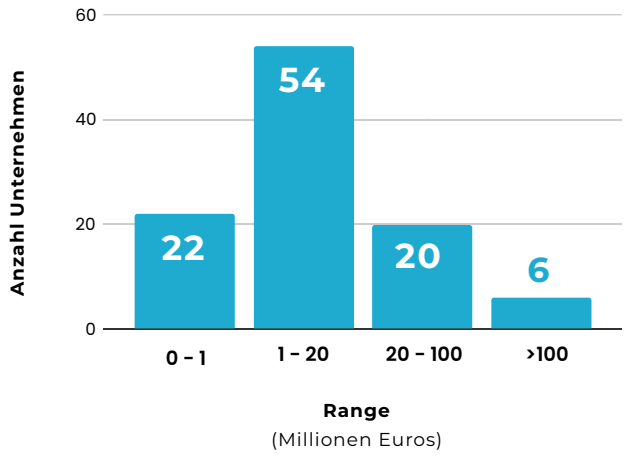
N=78 (N bedingt sich durch die konkrete Angabe einer Technologie in der Unternehmensbeschreibung auf Crunchbase)

4. Wie viel Geld fließt in den Markt?

Die Analyse von Risikokapital, das in den PET-Markt investiert wird, kann für diesen Bericht zwei Hinweise liefern: Sie erleichtert eine Aussage über den Reifegrad der Lösungen auf dem PET-Markt und verleiht ihm eine Größenordnung. Neben den Summen an investiertem Kapital drückt der Ursprung des Kapitals das Interesse strategischer Investoren aus. Dadurch lassen sich Early Adopters von skalierbaren PET-Lösungen und deren Industrien identifizieren. Von den in unserer Analyse identifizierten 195 PET-Unternehmen haben insgesamt 102 Unternehmen Risikokapital in der Gesamthöhe von 2,1 Milliarden US-Dollar erhalten. Zwei Drittel der Unternehmen befinden sich aktuell in der Seed-Phase. Die Tatsache, dass ein derart großer Teil der Unternehmen auf einer frühen Entwicklungsstufe steht, bedeutet auch, dass sich viele Lösungen für die Anwendung von PETs ebenfalls noch in einem frühen Stadium der Marktreife befinden. Das wiederum bietet zwei Lesarten an: Zum einen ist zu erwarten, dass ein Teil der PET-Lösungen ihre Use-Cases noch unter Beweis stellen muss und gegebenenfalls nicht langfristig Profit abwirft. Eine alternative Lesart lässt bei einer hohen Anzahl an Unternehmen in diesem Entwicklungsstadium ein hohes Innovationspotenzial für die kommenden Jahre erwarten.

Von den in unserer Analyse identifizierten 195 PET-Unternehmen haben insgesamt 102 Unternehmen Risikokapital in der Gesamthöhe von 2,1 Milliarden US-Dollar erhalten.

Abbildung 6: **Funding im PET-Markt**



N=102

Internationale PET-Champions

Folgende sechs Unternehmen aus unserer Analyse haben ein Funding von über 100 Millionen US-Dollar eingesammelt: Owkin, AppsFlyer, Privatar, Mina, R3 und Permutive. Vier von ihnen sitzen in den Vereinigten Staaten und zwei im Vereinigten Königreich. Zu den hervorstechenden Industrien, die von den PET-Champions targetiert werden, gehören der Finanz-, der AdTech- und der Gesundheitssektor.

Strategische Investments von Corporates

Ein Blick auf die Akteure, die in PET-Unternehmen investieren, zeigt, für welche Industrien PETs eine besondere Rolle spielen. Die summenmäßig größte Beteiligung eines Corporates, das in ein PET-Unternehmen investiert hat, ist der französische Pharmakonzern Sanofi, der 2021 180 Millionen US-Dollar in das Unternehmen Owkin investiert hat. Sanofi erhofft sich davon Erkenntnisse aus der Forschung an Krebsbehandlungen, die mittels künstlicher Intelligenz unter Wahrung der Privatsphäre von Patienten durchgeführt wird. Dabei kommen die Grundlagentechnologien des Federated Learning und der Multiparty Computation zum Einsatz. Neben dem Gesundheits- und Pharmasektor fällt der AdTech-Sektor auf. Der CRM-Anbieter Salesforce hat sich 2020 an dem Unternehmen AppsFlyer beteiligt. Während die genaue Beteiligung von Salesforce Ventures nicht bekannt ist, deutet die Größe der Investitionsrunde von 210 Millionen auf einer Bewertung von zwei Milliarden US-Dollar auf die Relevanz des Sektors hin. AppsFlyer bietet Lösungen, um Daten aus dem Marketing unter Einhaltung geltender Datenschutzbestimmungen zu analysieren. Zentral dafür sind Attributionsmodelle, die es erlauben, Kosten im Marketing zu optimieren. Auch in anderen Industrien investieren etablierte Corporates in PET-Unternehmen. Am australischen Unternehmen Data Republic haben sich 2018 Singapore Airlines und Qantas beteiligt. Das PET-Unternehmen verspricht Lösungen, um Kundendaten teilen zu können, ohne Rückschlüsse auf persönliche Daten zu erhalten. Auch Telekommunikationsunternehmen setzen auf PETs. So hat Swisscom Ventures in das PET-Unternehmen Inpher investiert, das verschiedene Lösungen anbietet, die auf den Grundlagentechnologien der Secure-Multiparty-Computation, der homomorphen Verschlüsselung und des Federated Learning basieren.

Fallstudien: PETs in der Praxis

01 ATLAS:

Datentreuhänder für anonymisierte Analysen in kommunalen Datenräumen

Auf einen Blick

- **Projektbeteiligte:** Polyteia, Hasso-Plattner-Institut für Digital Engineering, Technische Universität Berlin (TU Berlin), SINE Foundation, KIProtect, Verband Region Rhein-Neckar
- **Art der Verwaltung:** Kommunalverwaltungen
- **PETs im Einsatz:** ScrambleDB, Differential Privacy, Secure Multiparty Computation
- **Ziel des Projekts:** Isolierte, sensible Daten in Kommunalverwaltungen für die Auswertung nutzbar machen

Kommunen erheben und speichern große Mengen sensibler Daten. Viele dieser sensiblen Daten können insbesondere dann Mehrwerte für das öffentliche Leben generieren, wenn Kommunen sie erstens miteinander verbinden und zweitens kollaborativ nutzen.

Hintergrund

Kommunen erheben und speichern große Mengen sensibler Daten. Viele dieser sensiblen Daten können insbesondere dann Mehrwerte für das öffentliche Leben generieren, wenn Kommunen sie erstens miteinander verbinden und zweitens kollaborativ nutzen. Die Sozialraumplanung ist dafür ein gutes Beispiel: Sie verfolgt den Gedanken, dass sozioökonomische Parameter bei der Stadt- und Raumplanung Gewicht finden müssen. Wenn eine Kommune also den Bau einer neuen Kita plant, dann ist für die Wahl des Standortes relevant, in welchem Bezirk besonders viele Familien mit jungen Kindern wohnen. Und wenn eine Stadt den

Bau eines neuen Jobcenters plant, dann sind Einkommensverhältnisse und Informationen zu Sozialhilfebezügen für die Wahl eines geeigneten Standortes ebenso wichtig. Den Stellen, die in diesem Beispiel die Stadtplanung betreiben, stehen die entscheidenden Daten in vielen Fällen jedoch nicht zur Verfügung. Das liegt daran, dass sensible Daten – wie etwa Informationen über Sozialhilfeleistungen – häufig in dezentralen Silos gesichert werden und für weitere Zwecke außerhalb der zuständigen Behörde nicht zur Analyse freigegeben sind. Das hat gute Gründe – denn diese Daten sind besonders schützenswert und unterliegen insbesondere dann speziellen Sicherheits- und Datenschutzanforderungen, wenn sie personenbezogen sind. Um Wege zu finden, diese Daten trotzdem nutzbar zu machen, fördert die Bundesregierung im Bundesrahmenprogramm zur IT-Sicherheit „Digital. Sicher. Souverän." unter anderem Projekte, die Technologien und Verfahren zur Anonymisierung und Depersonalisierung erforschen und erproben.[80]

80 BMBF (2021): Forschungsnetzwerk Anonymisierung für eine sichere Datennutzung.

Umsetzung

Das Konsortium um das Berliner GovTech-Start-up Polyteia führt im Rahmen des Bundesprogramms das Forschungsprojekt *Datentreuhänder für anonymisierte Analysen in kommunalen Datenräumen* durch. „Kollektive Datennutzung bei dezentraler Datensouveränität" ist das zentrale Stichwort des Projektes.[81] Gemeinsam mit dem Verband Region Rhein-Neckar erörtern die Projektpartner im ersten Schritt einen konkreten Use-Case mit Bedarf für die Verknüpfung von Datensilos in Kommunen.[82] Zur Debatte steht dafür die bereits genannte Sozialraumplanung. Konkret könnten etwa Einwohnermeldedaten mit Daten aus lokalen Jobcentern verknüpft werden, um soziale Gesichtspunkte, wie die Abhängigkeit von Sozialleistungen, stärker in die räumliche Kommunalplanung einzubeziehen.[83]

Für den ausgewählten Use-Case erproben und erforschen die Projektbeteiligten drei elementare Funktionen, die die kollaborative Nutzung sensibler Daten in Kommunen ermöglichen:

1. Die automatische Pseudonymisierung beim Vorgang der Datenspeicherung in kommunalen Behörden;
2. Die zentrale Auswertung pseudonymisierter Daten über Silo-Grenzen hinaus mittels eines Datentreuhänders;
3. Die dezentrale Auswertung von Daten aus Daten-Silos der Kommunalverwaltung.

81 Aurel Stenzel in Bielawa, Helen (2023): Neue Krypto-Verfahren für verknüpfte Datensilos.

82 Zum Redaktionsschluss für den Bericht befand sich das Projekt in diesem Stadium.

83 Das Projekt arbeitet gezielt mit sog. Optionskommunen zusammen, in denen anstelle der Bundesagentur für Arbeit kommunale Träger die Zuständigkeit für die Aufgaben des SGB II übernehmen.

Mit Abschluss des für drei Jahre geplanten Projektes möchten die Beteiligten exemplarisch zeigen, wie kommunale Behörden die drei genannten Kernfunktionen mittels PETs und eines dafür entworfenen Arbeitsprozesses durchführen können. Die technologischen Grundlagen werden als Open-Source-Lösung entwickelt.

> *„Mit dem ATLAS-Projekt möchten wir Technologien so weiterentwickeln, dass Kommunen mit ihrer Hilfe sensible Daten einfach und schnell verknüpfen und analysieren können."*[84]

Marvin Klother, *Key Account & Strategic Partnership Executive, Polyteia*

Damit der Demonstrator nicht nur theoretische Mehrwerte aufzeigt, sondern auch für den Einsatz in der Praxis geeignet ist, werden alle drei Aspekte und die dafür verwendeten Lösungen laufend unter zwei Gesichtspunkten geprüft und bewertet:

1. Zum einen bewerten die Projektbeteiligten die Praktikabilität der Anwendungen. Da die Lösungen auch in der täglichen Verwaltungspraxis zum Einsatz kommen sollen, wird ein besonderes Augenmerk auf die einfache Bedienbarkeit der Instrumente zur Datenauswertung gelegt. Relevant ist dafür auch, dass die Lösungen trotz des Umgangs mit großen Datenmengen nicht anfällig für Latenzprobleme sind.
2. Insbesondere die Verknüpfung dezentraler Datensilos stellt für kommunale Verwaltungen einen neuen Prozess dar, der mit datenschutzrechtlichen Bedenken verbunden sein kann. Zum anderen werden deshalb alle drei Projektschritte mit Blick auf die Informationssicherheit und die Einhaltung der Sicherheitsanforderungen evaluiert. Diese Prüfung soll den beteiligten Kommunen ein klares Bild davon vermitteln, inwiefern die Maßnahmen eine sichere und DSGVO-konforme Kollaboration mit sensiblen Daten ermöglichen.[85]

84 Klother, Marvin & PUBLIC (2023): Interview über das ATLAS Projekt.

85 Klother, Marvin & PUBLIC (2023): Interview über das ATLAS Projekt.

PETs und ihre Rolle

In Bezug auf die drei Funktionen des Projektes kommen unterschiedliche PETs zum Einsatz. Für die (1) Pseudonymisierung beim Vorgang des Datenspeicherns wird das Protokoll ScrambleDB weiterentwickelt. Mit dessen Hilfe werden Datensätze automatisiert in Datenbruchteile gestückelt, um die Re-Identifizierung zu erschweren.[86] Auch die (2) zentralisierte Auswertung von Daten aus dezentralen Silos soll auf Grundlage des ScrambleDB-Protokolls ermöglicht werden. Die Institution des Datentreuhänders wird jedoch durch die Anwendung von Differential Privacy unterstützt, um mittels eines messbaren Kriteriums sicherzustellen, dass ausgehende Analysen keine Rückschlüsse auf Einzelpersonen zulassen. Die Funktion der (3) dezentralen Analyse verteilter Daten möchte das Konsortium unter Anwendung von Secure-Multiparty-Computation ermöglichen.[87]

Für die Anwendung auf den ausgewählten Use-Case greifen die Projektbeteiligten also nicht auf ein am Markt existierendes Produkt zurück, sondern adaptieren die Grundlagen von ScrambleDB und Secure-Multiparty-Computation für den spezifischen Bedarf.[88]

Learnings für den Einsatz von PETs im öffentlichen Sektor

1. **Ausrichtung an Use-Cases:** Das Projekt entwickelt die Anwendung für PETs auf Grundlage eines konkreten Bedarfs. Der vom Use-Case getriebene Ansatz ist eine wichtige Voraussetzung dafür, dass die entwickelten Lösungen auch in der Verwaltungspraxis einen Mehrwert liefern. Andersherum laufen Lösungen, die auf PETs zurückgreifen, Gefahr, nicht an die Realität der Verwaltungsarbeit angepasst zu sein. In Verwaltungen, die sich mit dem Einsatz von PETs beschäftigen, sollte die Definition des eigenen Bedarfs deshalb an erster Stelle stehen.
2. **Augenmerk auf Praktikabilität und Performance:** Viele PET-basierte Lösungen erfordern im Umgang hohe technische Expertise. Das kann besonders für kommunale Verwaltungen eine Hürde darstellen, da die einzelnen Behörden nicht über entsprechend qualifiziertes Personal verfügen. Das Projekt verfolgt mit der stetigen Prüfung auf Praktikabilität und Performance deshalb eine wichtige Fragestellung für die erfolgreiche Implementierung von PETs im öffentlichen Sektor.
3. **Rechtssicherheit schafft Vertrauen:** Öffentliche Verwaltungen müssen darauf vertrauen können, dass die Prozesse und Systeme, mit denen sie Daten speichern, bewegen und auswerten, technisch und rechtlich konform sind. Der Einsatz von PETs steht im öffentlichen Sektor noch am Anfang und deshalb kann die wissenschaftliche Evaluierung von Sicherheitsaspekten wie im ATLAS-Projekt eine wichtige Voraussetzung für die Vertrauensbildung der Verwaltung in Bezug auf die Technologie sein.

86 Lehmann, Anja (2019): ScrambleDB: Oblivious (Chameleon) Pseudonymization-as-a-Service.

87 Klother, Marvin & PUBLIC (2023): Interview über das ATLAS Projekt.

88 Da sich das Projekt zu Redaktionsschluss noch in der Frühphase befand, können keine weiteren Spezifika über den Einsatz und die Entwicklung der PETs ausgeführt werden.

02
FAIR TREATMENT:

Federated analytics and AI research across TREs for adolescent mental health

Auf einen Blick

- **Projektbeteiligte:** Cambridge University, Bitfount, AIMES, Intermine, CRATE, Anna Freud National Centre for Children and Families
- **Art der Verwaltung:** Krankenhäuser und Einrichtungen für soziale Dienste
- **PETs im Einsatz:** Federated Learning/Federated Analysis; Differential Privacy
- **Ziel des Projekts:** Isolierte, sensible Daten zugänglich für medizinische Forschung machen

„Integrated health services require integrated health data. This is why we initiated the FAIR TREATMENT project."

Anna Moore, Clinical Lecturer, Cambridge University[89]

Hintergrund

Die Anzahl psychischer Erkrankungen im Erwachsenenalter nimmt international zu – und in vielen Fällen liegen die Auslöser für diese Erkrankungen bereits in den ersten 20 Jahren des Lebens. Mediziner*innen und Wissenschaftler*innen stehen deshalb vor der Aufgabe, die Indizien für Auslöser psychischer Erkrankungen frühzeitig zu erkennen und diese Erkenntnisse in präventive Eingriffe umzuwandeln. Ein zentrales Problem für das Erkennen von Indizien und die Entwicklung von präventiven Maßnahmen ist die fehlende Datenlage. Einzelne Krankenhäuser oder andere medizinische Einrichtungen erfassen und verarbeiten zwar regelmäßig sensible und medizinisch wertvolle Daten. Diese Daten werden jedoch häufig isoliert gesammelt und gespeichert – und selten standardisiert.[90] Die einzelnen Datensätze verfügen meist nicht über genug Tiefe und Details, um allein die Grundlage für neue medizinische Erkenntnisse

89 Moore, Anna & PUBLIC (2022): Interview über das FAIR TREATMENT Projekt.

90 Moore, Anna et al. (2022): FAIR TREATMENT: Federated analytics and AI Research across TREs for AdolescenT MENTal health.

zu bilden. Zudem sind für die Entwicklung präventiver Interventionen häufig auch soziodemographische Daten oder Daten aus Sozialämtern von hohem Wert, die wiederum isoliert erhoben und verarbeitet werden. Da es sich in vielen Fällen um hochsensible Daten handelt, können die einzelnen Institutionen sie auch nicht einfach untereinander austauschen.

Umsetzung

Das Projekt FAIR TREATMENT möchte genau diese Hürde überwinden: Mit Hilfe sogenannter Trusted-Research-Environments, also geschützter Datenräume, sollen medizinische Daten aus verschiedenen Krankenhäusern und Einrichtungen für soziale Dienste des Vereinigten Königreichs miteinander verbunden und für wissenschaftliche Untersuchungen zugänglich gemacht werden. In der ersten Phase des Projektes, die für diesen Bericht untersucht wurde, haben die Projektbeteiligten einen Demonstrator eingeführt, der die Grundlage für dieses Ziel bildet. Vier zentrale Herausforderungen wurden dafür identifiziert:[91]

1. Die Standardisierung und Harmonisierung der Daten aus unterschiedlichen Krankenhäusern;
2. Der dezentrale, flexible Zugriff auf die zusammengeführten Daten;
3. Die Etablierung hoher Sicherheits- und Datenschutzstandards;
4. Die aktive Einbeziehung der datengebenden Bevölkerung.

> *„The healthcare sector holds vast treasures of data. The challenge lies in the standardization and secure processing of this data."*
>
> *Anna Moore, Clinical Lecturer, Cambridge University*[92]

PETs und ihre Rolle

Die technische Bewältigung dieser Herausforderungen lässt sich in zwei Ebenen unterteilen: Auf der ersten Ebene wurden drei regionale Trusted-Research-Environments installiert, die Daten aus lokalen Krankenhäusern und Einrichtungen für soziale Dienste zusammenführen. Auf der zweiten Ebene sollen autorisierte Forscher*innen Analysen dezentral auf den gemeinsamen Daten der Trusted-Research-Environments durchführen können.

Ebene 1

1. Mit Hilfe der Open-Source-Software CRATE wurden die Daten aus den dezentralen Datenquellen zunächst pseudonymisiert, um die Re-Identifizierung im weiteren Verlauf des Projektes zu erschweren.
2. Für die Infrastruktur der Trusted-Research-Environments wurden Datenräume aufgebaut, die mit hohen Sicherheitskomponenten versehen wurden, um den Zugang zu reglementieren.
3. Die Standardisierung und Harmonisierung der Daten innerhalb der Trusted-Research-Environments erfolgte mit der Open-Source-Software Intermine.

91 Moore, Anna & PUBLIC (2022): Interview über das FAIR TREATMENT Projekt.

92 Moore, Anna & PUBLIC (2022): Interview über das FAIR TREATMENT Projekt.

Ebene 2

4. Um Forscher*innen die Analyse über die drei regionalen Trusted-Research-Environments zu ermöglichen, hat das PET-Start-up Bitfount eine Federated-Analysis-Anwendung entwickelt, die einen gleichzeitigen Zugriff auf die gesammelten, pseudonymisierten und standardisierten Daten zulässt und ähnlich wie Federated Learning funktioniert.
5. Die Datenoutputs aus den Trusted-Research-Environments wurden von Bitfount außerdem mit Privacy Checks versehen, die auf Differential Privacy zurückgreifen: Durch sie kann sichergestellt werden, dass Analysen aus den Daten keine Rückschlüsse auf Einzelpersonen zulassen.

PETs kommen in diesem Projekt insbesondere auf der Ebene zwei zum Einsatz: Ein Federated-Analysis-Ansatz ermöglicht den Zugriff auf unterschiedliche Datensammlungen, während Differential Privacy als Kontrollinstanz dafür fungiert, dass Analysen für die medizinische Forschung unter hohen Privatsphärestandards durchgeführt werden. In Kombination erhöhen sie daher gleichzeitig die Nutzbarkeit und den Schutz von Daten.[93]

Learnings für den Einsatz von PETs im öffentlichen Sektor

1, **PETs sind komplementär:** PETs haben sich im FAIR-TREATMENT-Projekt in zweierlei Hinsicht als komplementär erwiesen. Zum einen verhalten sie sich komplementär zu anderen Sicherheitsmechanismen, wie etwa beim Aufbau der Trusted-Research-Environments. Gleichzeitig können PETs selbst miteinander kombiniert werden: Der Einsatz von Federated Analysis und Differential Privacy leistet einen Beitrag dazu, die Nutzbarkeit von hochsensiblen Daten und Datenschutz zu verbinden.

2. **Start-ups als Anbieter**: Für die technische Realisierung des Projektes spielen das Start-up Bitfount und Open-Source-Software eine entscheidende Rolle. Beides sind wichtige Komponenten im Innovationsprozess von PETs und können für den öffentlichen Sektor und seinen Zugang zu PETs eine wertvolle Komponente sein.

3. **Transparenz entscheidend:** Das zentrale Augenmerk des Projektes lag neben der technischen Realisierung in der Einbeziehung der Bürger*innen, deren Daten verwendet wurden. Transparenz über die Datenverarbeitung wurde als relevantes Vehikel hervorgehoben, um Unterstützung für das Forschungsvorhaben zu schaffen.[94, 95]

> *„The people providing data involved in our project, especially the parents, were not fundamentally afraid of sharing their data. Rather, they were surprised that the data were not yet being used for research purposes and were willing to contribute.“*
>
> **Anna Moore,** Clinical Lecturer, Cambridge University

93 Kaufman, Lauren & PUBLIC (2022): Interview über das FAIR TREATMENT Projekt.

94 Moore, Anna & PUBLIC (2022): Interview über das FAIR TREATMENT Projekt.

95 Moore, Anna & PUBLIC (2022): Interview über das FAIR TREATMENT Projekt.

03
NESSI Nachweisplattform ELSTER

Self-Sovereign Identities

Auf einen Blick

- **Projektbeteiligte:** **Bayerisches Landesamt für Steuern, Sparkassen-Finanzgruppe, Fraunhofer FIT, Friedrich-Alexander-Universität Erlangen-Nürnberg, mgm technology partners, secunet security Networks AG**
- **Art der Verwaltung: Finanz- und Steuerverwaltung**
- **PETs im Einsatz: Selective Disclosure**
- **Ziel des Projekts: Einen datensparsamen, sicheren Prozess zur Verifizierung zwischen Nutzenden, Steuerbehörden und Finanzinstituten etablieren**

Hintergrund

ELSTER ist die zentrale Software zur elektronischen Abwicklung der Steuererklärung in Deutschland und vielen Bürger*innen bekannt. Der Steuerbescheid, den sie von ihren Finanzämtern über ELSTER erhalten, dient in einigen anderen Bereichen als Nachweis für die Inanspruchnahme von Dienstleistungen. Das gilt zum Beispiel für die Bescheinigung der eigenen Bonität. Insbesondere Selbstständige sind auf ihren Einkommensteuerbescheid angewiesen, um ihre Bonität zu bescheinigen, weil sie keine klassischen Gehaltsnachweise vorlegen können. Bürger*innen müssen Bonität zum Beispiel dann nachweisen können, wenn sie einen Kredit bei einer Bank aufnehmen möchten. Im Regelfall erhalten sie ihren Einkommensteuerbescheid aktuell in Papierform oder als elektronisches PDF-Dokument. Für die Bürger*innen und die Bank, bei der der Kredit beantragt werden soll, ergeben sich dadurch mehrere Hindernisse für eine digitale, effiziente Antragstellung: Die Dokumente erzeugen häufig Medienbrüche im Prozess und können mitunter nur schwer automatisiert ausgelesen werden. Außerdem sind Bescheinigungen in Papierform oder PDF-Format anfällig für Fälschungen und lassen sich schwer auf ihre Gültigkeit überprüfen.[96] Neben Hürden, die den Prozess beeinflussen, gehen mit dem Einreichen des gesamten Steuerbescheides auch Bedenken zur Datensparsamkeit einher: Denn im Einkommensteuerbescheid finden sich auch Informationen, die nicht zwingend für die Entscheidung über den Kredit relevant sind und so dennoch sichtbar werden.

96 Bayerisches Landesamt für Steuern (2022): NESSI Nachweisplattform ELSTER Self-Sovereign Identities

Das Projekt NESSI hat einen Demonstrator entwickelt, mit dem die aufgezeigten Herausforderungen und Schwachstellen gelöst werden sollen.

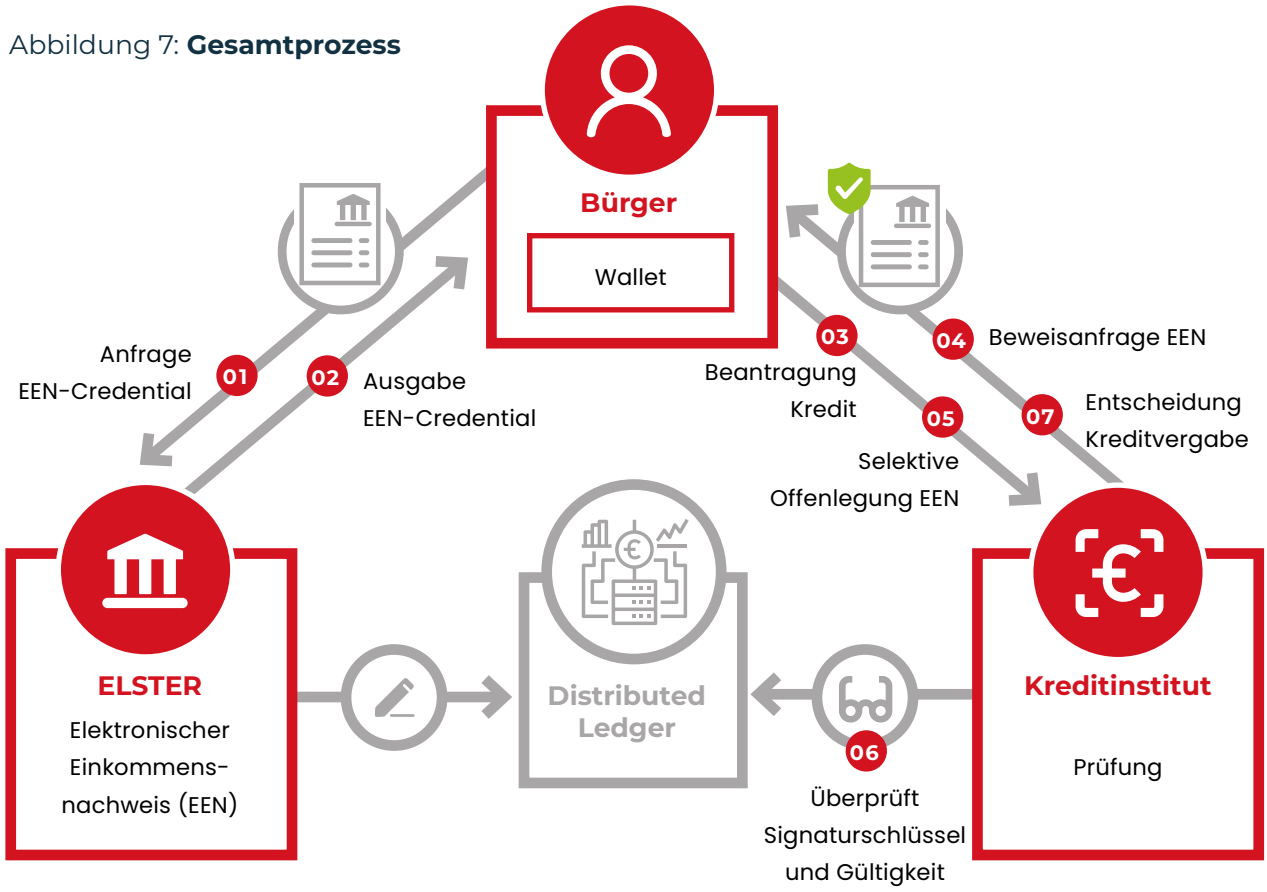
Umsetzung

Das Projekt NESSI hat einen Demonstrator entwickelt, mit dem die aufgezeigten Herausforderungen und Schwachstellen gelöst werden sollen. Wie Abbildung 7 verdeutlicht, werden die drei beteiligten Parteien über ein Blockchain-Register miteinander verbunden. Das Projekt greift dafür auf die Infrastruktur zurück, die das Consortium IDUnion für digitale Identitäten entwickelt hat.[97] Die Nutzenden benötigen für die Umsetzung des Prozesses ein digitales Wallet, das in diesem Fall von dem SSI-Startup Lissi bereitgestellt wurde. Der Prozess für den Nachweis der Bonität für einen Kreditantrag erfolgt mit der dargestellten Infrastruktur in sieben Schritten:

1. Auf dem elektronischen Steuerbescheid, der über das Portal Mein-ELSTER bereitgestellt wird, ist ein QR-Code angegeben. Dieser QR-Code wird mit dem Wallet gescannt, um die Abfrage des Einkommensnachweises und weiterer Stammdaten zu initiieren.
2. Da im einscannenden Wallet eine Basis-Identität hinterlegt ist, kann die Finanzbehörde dem Nutzer die Informationen digital und sicher auf das Wallet übertragen.
3. Anschließend beantragt der Nutzer den Kredit beim Finanzinstitut.
4. Das Kreditinstitut möchte eine Bestätigung für die Bonität des Nutzers und fordert deshalb Informationen aus dem Einkommensnachweis. Dafür fordert die Bank nur die Informationen aus dem Nachweis an, die sie tatsächlich benötigt.
5. Der Nutzer gibt die benötigten Informationen über das Wallet frei.
6. Das Kreditinstitut überprüft den Signaturschlüssel und die Gültigkeit.
7. Letztlich entscheidet die Bank auf Grundlage der Informationen über die Kreditvergabe.

97 IDUnion: Über das Projekt.

Abbildung 7: **Gesamtprozess**



Die Gültigkeitsprüfung entlang der sieben Schritte erfolgt über das Blockchain-basierte Register. Auf diesem Register werden in Abgrenzung zu anderen Blockchain Anwendungen nicht die einzelnen Details der Einkommensnachweise gespeichert, sondern drei Kategorien von Metadaten,[98] die Rückschlüsse auf personenbezogene Details kaum ermöglichen.[99] Die Anwendung birgt zum einen den Vorteil eines voll digitalen Prozesses ohne Medienbrüche. Zum anderen ist die ausstellende Behörde zu jeder Zeit in der Lage, das Credential zurückzuziehen und dem Nachweis damit die Gültigkeit zu entziehen - etwa wenn neue Informationen über den Antragsteller seine Bonität anzweifeln.

Für die Bewilligung eines Kredits sind in der Realität mehr Informationen für den Antragsteller notwendig, als ein Einkommensnachweis. Das Projekt NESSI hatte zunächst zum Ziel, zu zeigen, dass Verifizierungsprozesse mit hochsensiblen Daten digital und sicher über SSI abzubilden sind. Auf dieser Grundlage lassen sich theoretisch weitere Credentials abbilden und auch komplizierte Abfragen ermöglichen.[100]

Für den breitflächigen Einsatz von SSI-Lösungen in der Verwaltung identifiziert das Projekt jedoch eine zentrale Hürde: Das BSI hat bisher neuartige kryptographische Verfahren, auf denen etwa das für NESSI verwendete Register beruht, noch nicht bewertet. Die positive Bewertung der Verfahren ist jedoch Voraussetzung dafür, dass derartige Lösungen im öffentlichen Sektor skaliert werden können.[101, 102]

98 Daten über den Credential-Aussteller, Daten über die Schemata der Credentials und die verschiedenen Versionen des Revocation Registries.

99 Bayerisches Landesamt für Steuern (2022): NESSI Nachweisplattform ELSTER Self-Sovereign Identities

100 Bayerisches Landesamt für Steuern (2022): NESSI Nachweisplattform ELSTER Self-Sovereign Identities.

101 Michael, Helge; PUBLIC(2023): Interview zu PETs.

102 Bayerisches Landesamt für Steuern (2022): NESSI Nachweisplattform ELSTER Self-Sovereign Identities.

PETs und ihre Rolle

Datenschützende Elemente treten in diesem Projekt auf zwei Ebenen auf: Zunächst ist die SSI-Infrastruktur selbst so aufgebaut, den Nutzenden die Gewalt über ihre Daten in die eigene Hand zu geben. Im Kontrast zu den bekannten Single Sign-on-Diensten, bei denen Facebook, Google und andere Plattformen zentral Nutzerdaten verwalten und für Authentifizierung bei unterschiedlichen (privaten) Diensten nutzen, baut die SSI-Infrastruktur auf dezentraler Verwaltung von Identifikationsdaten auf. Der für NESSI verfolgte Ansatz unterscheidet sich darüber hinaus durch die Art der gespeicherten Daten von klassischen Blockchain-Infrastrukturen, zu denen es relevante Datenschutzbedenken gibt.[103]

Konkrete PETs kommen jedoch insbesondere im Prozess der Verifizierung zum Einsatz. Digitale Wallets - in diesem Fall von Lissi - lassen zum einen den Prozess des Selective Disclosures zu und können auch Zero-Knowledge Proofs abbilden. Würde der Nutzer der Bank seinen gesamten Einkommensnachweis zur Verfügung stellen, dann können auf diesem Wege personenbezogene Informationen eingesehen werden, die für die eigentliche Transaktion irrelevant sind. Den gleichen Prozess könnte das Wallet auch über Zero-Knowledge Proofs umsetzen. In diesem Fall würde die Bank nicht das Einkommen selbst erfahren, sondern auf Grundlage vorher abgestimmter Bedingungen lediglich erfahren, ob die Bonität des Nutzers gegeben ist.[104] Dies erfolgt beispielsweise durch die Freigabe von Range-Proofs, die nur ein Mindesteinkommen bestätigen, nicht aber das genaue Einkommen mitteilen.

Learnings für PETs im öffentlichen Sektor

1. SSI Systeme können Nutzer*innen Kontrolle über Daten zurückgeben: NESSI hat gezeigt, dass datensparsame und sichere Verifizierungen über SSI-Systeme funktionieren. Das bringt für die drei beteiligten Akteure nicht nur Komfort im Prozess, sondern kann Abhängigkeiten der Nutzer*innen von zentralen Instanzen reduzieren.
2. Um zu skalieren, brauchen neuartige kryptographische Verfahren Rückenwind vom BSI: Grundlagentechnologien wie der Zero-Knowledge Proof können Transaktionen und Verifizierungen in SSI Systemen sicher und datensparsam durchführen. Die fehlende Bewertung der zugrundeliegenden Verfahren durch das BSI blockiert jedoch maßgeblich die Skalierung von innovativen Lösungen.

103 Krishnan, Saravanan et al. (2020): Handbook of Research on Blockchain Technology.

104 Michael, Helge; PUBLIC: Interview zu PETs.

Standardisierung von Identifizierungs- und Authentifizierungsverfahren: Bestrebungen von IDUnion und Implikationen der eIDAS-Verordnung

Das im Anwendungsfall verwendete Wallet Lissi (Let's initiate Self-sovereign Identity) ist zunächst als Forschungsinitiative entstanden. Nachdem erste Prototypen positive Resonanz erfahren haben, schlossen sich große Banken wie die Commerzbank, ING-Diba und die Deutsche Bank sowie die Bundesdruckerei der Entwicklung von Lissi an. Das vom Bundesministerium für Wirtschaft und Klimaschutz unterstützte Vorhaben Lissi führt das IDUnion Projekt. Das aus mehr als 60 Konsortiumsmitgliedern bestehende IDUnion Projekt hat das Ziel, die Offenheit eines Ökosystems für die dezentrale Identitätsverwaltung zu fördern.

Regulatorisch trat bereits 2014 die erste Version der eIDAS-Verordnung (electronic IDentification, Authentic and trust Services) als rechtliche Grundlage für digitale Identitäten in Kraft. Ihr Ziel ist es, einen einheitlichen Rahmen für die Anerkennung elektronischer Identifizierungen (eIDs) und einen europäischen Binnenmarkt für Vertrauensdienste zu schaffen. Damit soll das Vertrauen bei elektronischen Transaktionen gestärkt werden. Aktuell wird über die eIDAS 2.0-Verordnung verhandelt, die ab 2024 die Standards für das Digital Identity (EUDI) Wallet setzt. Vergleichbar mit der DSGVO wird von der eIDAS-Verordnung erwartet, globale Standards zu setzen.[105, 106] Große Tech-Konzerne mit dem Interesse ihre eigenen zentralen Single Sign-On-Dienst weiter zu skalieren werden gezwungen werden, dezentrale SSI-basierende Authentifizierungsverfahren ihrer Nutzer*innen zu akzeptieren. Darin liegt die Chance, die digitale Souveränität aller Bürger*innen im digitalen Raum zu stärken.

Eine flächendeckende Nutzung digitaler Identitäten setzt voraus, dass Sicherheit über die Zulassung bestimmter Verfahren besteht. Deshalb sind auch nationale Regulierungsbehörden gefragt, Standards für die Zulassung von kryptografischen Verfahren zu setzen.

105 Lissi (2022): eIDAS and the European Digital Identity Wallet: Context, status quo and why it will change the world.

106 Johnson, Alastir (2022): EIDAS 2.0 Turns To Self-Sovereign Identification To Bring Users Ownership And Control.

04
Cyber Defence Alliance:

Der Einsatz von PETs für die effektive Bekämpfung von Cyberkriminalität in der Finanzwirtschaft

Auf einen Blick

- **Projektbeteiligte:** **Cyber Defence Alliance, Duality Technologies, vier internationale Finanzinstitute, Londoner Polizei**
- **Art der Verwaltung:** **Finanzwirtschaft, Betrugserkennung**
- **PETs im Einsatz:** **Homomorphe Verschlüsselung**
- **Ziel des Projekts:** **Sichere und datenschutzkonforme Zusammenarbeit mit sensiblen Kundendaten, um organisierte Cyberkriminalität in der Finanzwirtschaft effektiv zu bekämpfen**

Hintergrund

Die Cyber Defence Alliance (CDA) ist ein gemeinnütziges öffentlich-privates Konsortium mit Hauptsitz in London. Es arbeitet mit privaten Institutionen des Finanzsektors und Strafverfolgungsbehörden zusammen, um proaktiv Informationen zur Bekämpfung von Cyberkriminalität und Bedrohungen auszutauschen. Zum Kern der CDA gehört es, Informationen ihrer Partner zu analysieren und sie in verwertbare Informationen für die Industrie und die Strafverfolgungsbehörden umzuwandeln.[107]

Die Cyberkriminalität im Finanzsektor hat drei Eigenschaften, die das Erkennen und Bekämpfen von Betrugsfällen erheblich erschweren:

1. Internationalisierung: Regionale Grenzen stellen für das organisierte Verbrechen im Finanzsektor eine geringe Hürde dar. Wenn die Methoden des Betrugs sich internationalisieren, dann sind auch die Strafverfolgung und die Industrie darauf angewiesen, ihre Mechanismen zur Erkennung und Bekämpfung auf einen internationalen Kontext auszuweiten. [108]

107 GlobeNewswire (2020): Cyber Defence Alliance (CDA) partners with Anomali to better enable sharing of Threat Intelligence among banking members.

108 Center for Strategic & International Studies (2017): Forces Shaping the Next Generation of Cyber Threats to Financial Institutions.

2, Vielschichtigkeit: Cyberkriminalität im Finanzwesen unterliegt einer Arbeitsteilung, bei der grob zwei Komponenten unterschieden werden: Kriminelle Dienstleistungen, die die Infrastruktur für Cyberkriminalität bereitstellen oder den direkten Zugang zu sensiblen Daten ermöglichen, und kriminelle Dienstleistungen, die den nicht-autorisierten Zugang zu Daten monetarisieren. Die Strafverfolgung muss mit Interventionen deshalb diverse und komplexe Formen der Cyberkriminalität erkennen und schnell reagieren können.[109]

Die Möglichkeit zum schnellen und sicheren Informationsaustausch zwischen Finanzinstituten und Strafverfolgungsbehörden ist ein Vehikel, das das Erkennen von Cyberkriminalität im Finanzwesen und die Möglichkeiten zur Reaktion maßgeblich steigern kann. Schnell muss der Informationsaustausch erfolgen, weil das organisierte Verbrechen die zeitraubenden Prozesse ausnutzt, die Institutionen für gewöhnlich nutzen, um sensible Daten für Ermittlungszwecke auszutauschen. Sicher müssen sie besonders deshalb sein, weil die Daten von Finanzinstituten in vielen Fällen hochsensible Kundendaten mit Personenbezug oder Geschäftsgeheimnisse beinhalten.[110]

Um diese Hürden zu nehmen, hat die CDA eine Kooperation mit vier internationalen Finanzinstituten, der Londoner Polizei und dem PET-Anbieter Duality Technologies initiiert. Ziel des Projektes war es einen Rahmen aufzubauen, über den private Finanzinstitute untereinander und mit der Londoner Polizei Analysen auf Grundlage gemeinsamer, ursprünglich dezentraler Daten durchführen können – schnell, sicher und vertrauensvoll.[111]

> *„The limitations for the use of PETs in the public and private sectors are not only in the technologies themselves. It is first about understanding what the technology can and cannot do."*
>
> **Ronen Cohen,** Vice President Strategy, Duality Technologies[112]

109 Center for Strategic & International Studies (2017): Forces Shaping the Next Generation of Cyber Threats to Financial Institutions.

110 World Economic Forum (2020): Cyber Information Sharing: Building Collective Security.

111 World Economic Forum (2020): Cyber Information Sharing: Building Collective Security.

112 Cohen, Ronen; Rohloff, Kurt & PUBLIC (2022): Interview über PETs

PETs und ihre Rolle

Für die technische Umsetzung der Projektziele hat Duality Technologies eine Plattform auf Grundlage homomorpher Verschlüsselung bereitgestellt.[113] Diese Anwendung verschlüsselt die Daten im Ruhezustand und hält sie während des Transfers zwischen den Partnern und während der Analyse weiterhin verschlüsselt. Die sensiblen Daten der einzelnen Finanzinstitute müssen aus diesem Grund zu keinem Zeitpunkt den anderen Partnern des Projektes offenbart werden. Die Lösung verschlüsselt die Daten jedoch nicht nur, sie schafft über die gemeinsame Plattform auch ein System, das allen Parteien die bereitgestellten, verschlüsselten Daten zur Analyse automatisiert verfügbar macht. Im Vorhinein hatten sich die Projektpartner darüber geeinigt, welche Daten bereitgestellt werden sollten.[114]

Learnings für den Einsatz von PETs im öffentlichen Sektor

1. **Sicherheit und Vertrauen durch homomorphe Verschlüsselung**: Der Einsatz von homomorphen Verschlüsselungsmethoden kann dort eine vertrauensvolle Kollaboration an sensiblen Daten ermöglichen, wo einzelne Akteure zuvor Sicherheitsbedenken gegenüber der Zusammenarbeit priorisiert haben.
2. **Öffentlich-private Partnerschaften können PETs in den öffentlichen Sektor tragen:** Zusammenschlüsse zwischen Akteuren aus der Industrie und dem öffentlichen Sektor können geeignete Wege für die Annäherung an PETs für die öffentliche Verwaltung sein. Sie minimieren in vielen Fällen das finanzielle Risiko für die öffentliche Seite und können auf die Expertise der Industrie bauen, um Use-Cases zu identifizieren.
3. **Technologisches Verständnis ist der Grundstein für den Einsatz von PETs:** Dass der Einsatz homomorpher Verschlüsselung im öffentlichen Sektor nicht weiter fortgeschritten ist, liegt nicht zwingend am technologischen Reifegrad. Die Vorarbeit liegt in der Aufklärung von Entscheider*innen, die mit einem gesteigerten Verständnis für Technologien sowie deren Möglichkeiten und Grenzen besser Bedarfe identifizieren und so einen Impuls aus der Nachfrage senden können.[115]

113 Cohen, Ronen; Rohloff, Kurt & PUBLIC (2022): Interview über PETs

114 World Economic Forum (2020): Cyber Information Sharing: Building Collective Security.

115 Cohen, Ronen; Rohloff, Kurt & PUBLIC (2022): Interview über PETs

Ausblick und Handlungsempfehlungen

Privacy-Enhancing Technologies können der öffentlichen Verwaltung dabei helfen, Daten zu nutzen und Daten zu schützen. Eine Balance, die von vielen Organisationen des öffentlichen Sektors lange als Nullsummenspiel betrachtet wurde, kann mit Hilfe von PETs neu austariert werden.

Die Technologien sind damit nicht nur in ihrem engen Wortsinn als Tools zum Schutz von Daten zu verstehen, sondern können dem öffentlichen Sektor neue Mittel für die aktive Nutzung von Daten in die Hand geben und damit einen Beitrag für die digitale Handlungsfähigkeit der Verwaltung leisten.

Öffentliche Verwaltungen in Deutschland haben sich der Potenziale von PETs noch nicht strategisch und in der Fläche angenommen. Dieser Bericht soll einen Beitrag dazu leisten, über PETs zu informieren und ihren Use-Case für den öffentlichen Sektor greifbar zu machen. Damit sie ihre Potenziale auch in der Praxis der öffentlichen Verwaltung entfalten können, sind jedoch politischer Wille und Initiative auf verschiedenen Ebenen der Verwaltung erforderlich. Aus den Erkenntnissen, die im Rahmen dieses Berichtes gesammelt wurden, lassen sich fünf konkrete Handlungsempfehlungen hervorheben, die insbesondere die Bundes- und Landesebene adressieren. Für die konkrete Auseinandersetzung von öffentlichen Organisationen mit den Potenzialen von PETs bietet das Anwendungsframework im folgenden Kapitel eine praktische Entscheidungshilfe.

Bundesebene

PETs in die Architekturrichtlinie des Bundes aufnehmen

Die Architekturrichtlinie des Bundes und die zu ihr gehörenden technischen Spezifikationen sind die zentrale Vorgabe für die IT-Aufstellung der Bundesverwaltung. Wenn die Bundesregierung *privacy-* und *security-by-design* in der öffentlichen Verwaltung sicherstellen möchte, dann sollten PETs und ihre Grundlagen in der Fortschreibung der Architekturrichtlinie berücksichtigt und verankert werden. Damit würde eine Grundlage dafür geschaffen, dass Verwaltungen PETs für die eigenen IT-Systeme beschaffen und effektiv nutzen können.

Mit dem *Bundesprogramm Privacy-Enhancing Technologies* Leuchtturmprojekte fördern

Viele öffentliche Verwaltungen zögern beim Einsatz von PETs. Neben fehlenden Standards, die den rechtlichen Rahmen für die Verwendung bilden, fehlt es an erfolgreichen Use-Cases und der aktiven Interaktion zwischen Wissenschaft, PET-Anbietern und öffentlichen Auftraggebern. Die Bundesregierung kann diese Blockade mit einem Förderprogramm aufbrechen. Bereits vorhandene (Forschungs-)Förderungen konzentrieren sich stark auf die Nutzbarmachung von Daten durch Anonymisierung. Ein breiter angelegtes Förderprogramm, das Pilotprojekte für den Einsatz von PETs im öffentlichen Sektor unterstützt, kann Leuchtturmprojekte mit Vorbildfunktion hervorbringen. Die erfolgreiche Durchführung von staatlich geförderten Projekten, die den End-to-End-Einsatz von PETs demonstrieren, wirken als wichtiger Impuls und als Anregung für Organisationen im öffentlichen Sektor. Eine solche Initiative sollte auch die aktive Kommunikation und Aufbereitung der Use-Cases beinhalten, um ihrem Leuchtturmcharakter gerecht zu werden.

Mit Standardisierung und Bewertung Vertrauen schaffen

Institutionen auf Bundesebene sind angehalten, sich in Prozesse zur Standardisierung von Technologien zu engagieren und Bewertungen zur Einsatzmöglichkeit von PETs abzugeben. Standards und positive Bewertungen können Vorbehalte öffentlicher Verwaltungen gegenüber PETs abbauen und deren Einsatz so begünstigen. Insbesondere das BSI könnte mit der Bewertung von neuen kryptografischen Verfahren den Grundstein dafür legen, dass zum Beispiel SSI-Systeme ihr Potenzial für den öffentlichen Sektor entfalten.

Landesebene

Mit PET-Trainingsprogrammen aufklären

Eine zentrale Hürde für die Verbindung zwischen Herausforderungen in der öffentlichen Verwaltung und PET-Lösungen ist ein Informationsdefizit auf Seiten der Verwaltung. Für viele Beschäftigte öffentlicher Organisationen ist die aufkommende Technologiegruppe noch nicht greifbar. Mit geförderten Trainingsprogrammen und Weiterbildungen kann dieses Defizit in der Breite bearbeitet werden. Hier können sich IT-Referate, aber auch andere Abteilungen mit Datenbezug, mit den Möglichkeiten, Funktionsweisen und Grenzen von PETs vertraut machen. Diese Grundlage kann als Einstieg dafür dienen, die eigenen Herausforderungen im Management von Daten mit potenziellen Lösungen auf dem PET-Markt in Verbindung zu bringen.

Mit PET-Innovationswettbewerben Co-Creation ermöglichen

Für viele Organisationen im öffentlichen Sektor ist der eigene Use-Case schwer mit den Lösungen auf dem PET-Markt in Verbindung zu bringen. Darüber hinaus besteht für gewöhnlich kein enger Draht zwischen privaten Anbietern wie Start-ups und potenziellen Kunden im öffentlichen Sektor. Landesministerien können Mittel für PET-Innovationswettbewerbe zur Verfügung stellen und diese Hürde überwinden. Challenge-Programme können für gewöhnlich über Planungs- und Innovationswettbewerbe abgebildet werden und bieten die Möglichkeit, dass Verwaltungen und Anbieter von technologischen Lösungen gemeinsam an einem Produkt arbeiten, das den spezifischen Anforderungen des Auftraggebers entspricht.

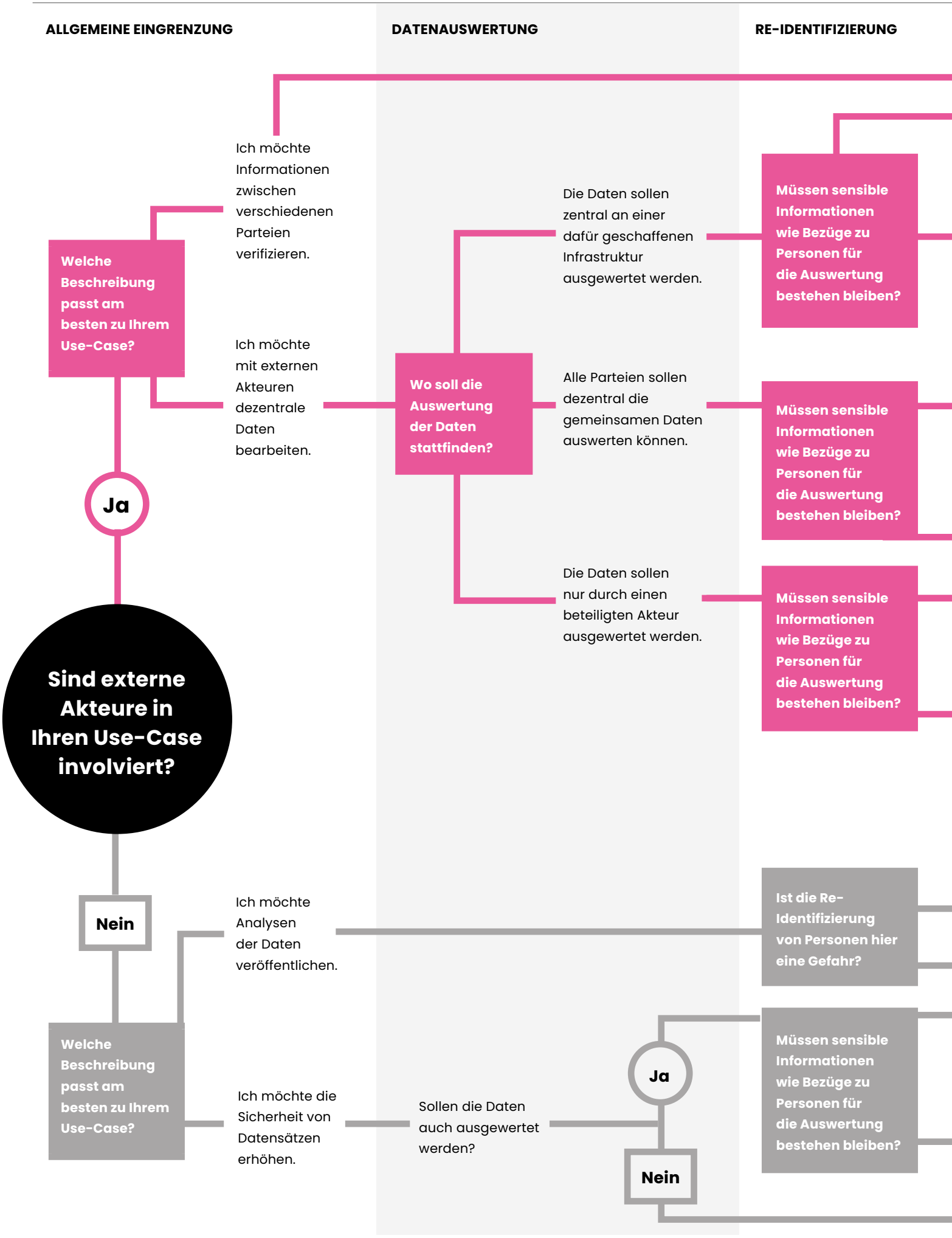
Anwendungsframework

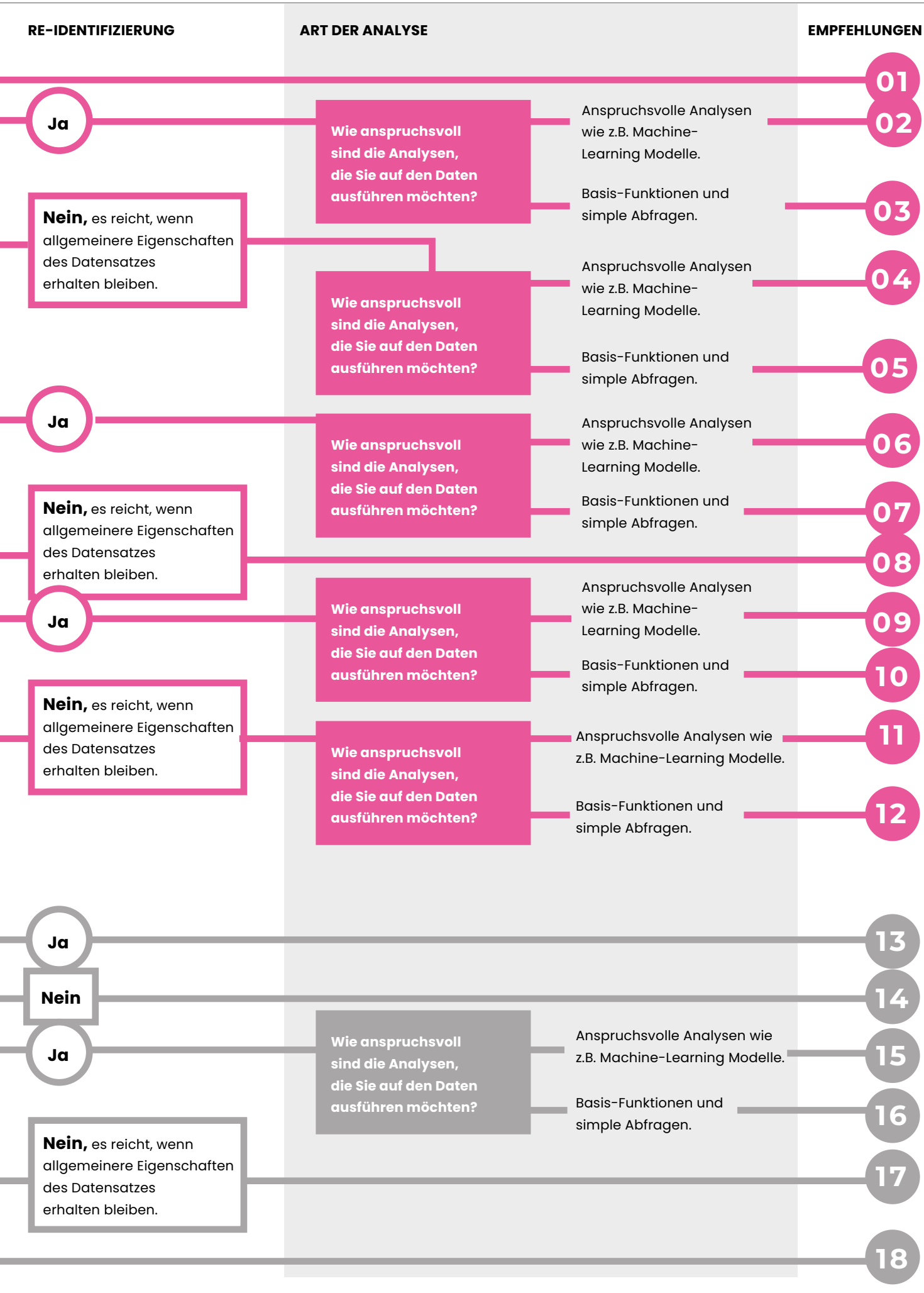
Für einzelne Organisationen, Abteilungen und Beschäftigte ist es nicht einfach zu erkennen, ob eine PET-Lösung zu den Herausforderungen der eigenen Organisation passt, und wenn ja, welche.

Das gilt selbst dann, wenn die grundlegende Funktionsweise einzelner PETs bekannt ist und der Blick auf erfolgreiche Use-Cases im öffentlichen Sektor Hinweise liefert. Das Raster, das den Kern dieses abschließenden Kapitels bildet, soll eine Entscheidungshilfe für die öffentliche Verwaltung darstellen. Es soll im ersten Schritt den Blick dafür schärfen, welche Komponenten für den eigenen Use-Case des Managements von sensiblen Daten zentral sind – von der Kollaboration mit externen Parteien bis zur Art der Analyse. Im zweiten Schritt spricht es eine Empfehlung aus, indem es den eigenen Use-Case mit PET-Technologien in Verbindung bringt. Damit soll ein strategischer Blick auf die Vielfältigkeit der PET-Lösungen für den eigenen Fachbereich erleichtert werden. Das Raster abstrahiert dabei notwendigerweise von spezifischen Anwendungsfällen. Es soll dazu dienen, eine strategische Richtung vorzugeben und Klarheit über geeignete PET-Anwendungen zu schaffen, und nicht als endgültige Antwort auf eine spezifische Herausforderung verstanden werden.

Vor dem Blick auf das Raster lohnt die Verinnerlichung vier zentraler Fragen:

1. **Was ist das gewünschte Ergebnis?** PETs unterscheiden sich darin, wie sie funktionieren und welchen Zweck sie erfüllen können. Vor der Entscheidung für eine Technologie oder ein Produkt sollten Nutzer*innen ihren Anwendungsfall so genau wie möglich definieren – einschließlich des Ergebnisses, das mit PETs angestrebt wird.
2. **Wer ist mit der Betreuung der Lösung betraut?** Bevor eine PET-Lösung eingeführt wird, sollte der Kreis der Nutzer*innen und Beteiligten berücksichtigt werden. Einige Lösungen werden als Low- oder No-Code-Anwendung angeboten, während andere tiefgreifende technische Kenntnisse erfordern.
3. **Wie sensibel sind die Daten?** Die Klassifizierung der Daten, mit denen eine öffentliche Einrichtung umgeht, ist ausschlaggebend für die Eignung und Notwendigkeit einer PET-Lösung. Je höher die Daten auf einer Sensibilitätsskala eingestuft sind, desto höher ist der Bedarf an PETs: Biometrische Daten zum Beispiel müssen mit höchsten Sicherheitsstandards behandelt werden. Geolokalisierungsdaten hingegen können anders behandelt werden.
4. **Wie sind die Daten strukturiert?** Strukturierte Daten, semi-strukturierte Daten und unstrukturierte Daten stellen an die mit ihnen betraute Lösung grundlegend unterschiedliche Herausforderungen. Bevor eine PET-Lösung in Betracht gezogen wird, sollte deshalb geprüft werden, ob sie für die Verarbeitung der Datenformate geeignet ist und ob eine Standardisierung der Daten vorausgehen muss.


ALLGEMEINE EINGRENZUNG
DATENAUSWERTUNG
RE-IDENTIFIZIERUNG
Sind externe Akteure in Ihren Use-Case involviert?
Ja
Welche Beschreibung passt am besten zu Ihrem Use-Case?
Ich möchte Informationen zwischen verschiedenen Parteien verifizieren.
Ich möchte mit externen Akteuren dezentrale Daten bearbeiten.
Wo soll die Auswertung der Daten stattfinden?
Die Daten sollen zentral an einer dafür geschaffenen Infrastruktur ausgewertet werden.
Alle Parteien sollen dezentral die gemeinsamen Daten auswerten können.
Die Daten sollen nur durch einen beteiligten Akteur ausgewertet werden.
Müssen sensible Informationen wie Bezüge zu Personen für die Auswertung bestehen bleiben?
Müssen sensible Informationen wie Bezüge zu Personen für die Auswertung bestehen bleiben?
Müssen sensible Informationen wie Bezüge zu Personen für die Auswertung bestehen bleiben?
Nein
Welche Beschreibung passt am besten zu Ihrem Use-Case?
Ich möchte Analysen der Daten veröffentlichen.
Ist die Re-Identifizierung von Personen hier eine Gefahr?
Ich möchte die Sicherheit von Datensätzen erhöhen.
Sollen die Daten auch ausgewertet werden?
Ja
Nein
Müssen sensible Informationen wie Bezüge zu Personen für die Auswertung bestehen bleiben?


RE-IDENTIFIZIERUNG
ART DER ANALYSE
EMPFEHLUNGEN
Ja
Wie anspruchsvoll sind die Analysen, die Sie auf den Daten ausführen möchten?
Anspruchsvolle Analysen wie z.B. Machine-Learning Modelle.
Basis-Funktionen und simple Abfragen.
01
02
03
Nein, es reicht, wenn allgemeinere Eigenschaften des Datensatzes erhalten bleiben.
Wie anspruchsvoll sind die Analysen, die Sie auf den Daten ausführen möchten?
Anspruchsvolle Analysen wie z.B. Machine-Learning Modelle.
Basis-Funktionen und simple Abfragen.
04
05
Ja
Wie anspruchsvoll sind die Analysen, die Sie auf den Daten ausführen möchten?
Anspruchsvolle Analysen wie z.B. Machine-Learning Modelle.
Basis-Funktionen und simple Abfragen.
06
07
Nein, es reicht, wenn allgemeinere Eigenschaften des Datensatzes erhalten bleiben.
08
Ja
Wie anspruchsvoll sind die Analysen, die Sie auf den Daten ausführen möchten?
Anspruchsvolle Analysen wie z.B. Machine-Learning Modelle.
Basis-Funktionen und simple Abfragen.
09
10
Nein, es reicht, wenn allgemeinere Eigenschaften des Datensatzes erhalten bleiben.
Wie anspruchsvoll sind die Analysen, die Sie auf den Daten ausführen möchten?
Anspruchsvolle Analysen wie z.B. Machine-Learning Modelle.
Basis-Funktionen und simple Abfragen.
11
12
Ja
13
Nein
14
Ja
Wie anspruchsvoll sind die Analysen, die Sie auf den Daten ausführen möchten?
Anspruchsvolle Analysen wie z.B. Machine-Learning Modelle.
Basis-Funktionen und simple Abfragen.
15
16
Nein, es reicht, wenn allgemeinere Eigenschaften des Datensatzes erhalten bleiben.
17
18
EMPFEHLUNGEN

Empfehlungen

01 **SSI-Lösungen** sind für diesen Use-Case relevant. Sie ermöglichen Datenschutz, Nutzerfreundlichkeit und Sicherheit für die Verifizierung von Informationen zwischen verschiedenen Akteuren.

02 Ein **Trusted Execution Environment** könnte es ermöglichen, die zentrale Analyse sicher durchzuführen. Die Übertragung der Daten wird jedoch nicht durch die vertrauenswürdige Ausführungsumgebung gesichert. Außerdem sind Trusted Execution Environments Grenzen in der Speicherung von ruhenden Daten gesetzt - das könnte eine Einschränkung sein.

03 **Homomorphe Verschlüsselung** könnte eine attraktive Lösung sein - sie ermöglicht die gemeinsame Arbeit an verschlüsselten Daten, ohne auf die Rohdaten zuzugreifen.

04 Eine Kombination aus **Differential Privacy oder synthetischen Daten und Federated Learning** könnte es Ihnen ermöglichen, die Re-Identifizierung Ihrer Daten zu verhindern und dezentrale Datensätze für eine zentrale Analyse zu verbinden. Diese Technologien bieten jedoch keine Sicherung der Daten.

05 **Homomorphe Verschlüsselung** könnte eine attraktive Lösung sein, um Daten im Ruhezustand, während der Übertragung und der Verarbeitung verschlüsselt zu halten. Verschiedene Formen der homomorphen Verschlüsselung erfüllen diese Anforderung jedoch unterschiedlich - deshalb lohnt ein genauer Blick auf die Lösung.

06 Es könnte schwierig sein, eine praktikable Lösung zu finden. Vielleicht möchten Sie jedoch direkt mit Anbietern von eingebetteten PET-Lösungen sprechen, um dies herauszufinden.

07 Je nach den Details Ihres Use-Cases könnten vertrauenswürdige Ausführungsumgebungen oder **homomorphe** Verschlüsselung eine attraktive Lösung sein.

08 Eine Kombination aus **Differential Privacy oder synthetischen Daten und Federated Learning** könnte es Ihnen ermöglichen, die Re-Identifizierung Ihrer Daten zu verhindern und dezentrale Datensätze für eine zentrale Analyse zu verbinden. Diese Technologien bieten jedoch keine Sicherung der Daten und Sicherheitsstandards sollten deshalb auf anderem Wege erfüllt werden.

09 **Federated Analysis** könnte eine Lösung sein, um auf dezentralisierte Daten zuzugreifen, ohne die Datensätze übertragen zu müssen, bietet jedoch keine zusätzliche Sicherheit für die ruhenden Daten.

10 **Secure Multiparty Computation** im client-server Modell kann eine Möglichkeit bieten, einfache Analysen mit dezentralen Daten unter hohen Datenschutzstandards durchzuführen.

11 **Federated Analysis** könnte eine Lösung sein, um auf dezentralisierte Daten zuzugreifen, ohne die Datensätze übertragen zu müssen, bietet jedoch keine zusätzliche Sicherheit für die ruhenden Daten.

12 **Secure Multiparty Computation** im client-server Modell kann eine Möglichkeit bieten, einfache Analysen mit dezentralen Daten unter hohen Datenschutz Standards durchzuführen.

Empfehlungen

13 **Differential Privacy** kann eine attraktive Lösung sein, um Analysen zu veröffentlichen und Re-Identifizierung stark zu erschweren.

14 PETs werden für diesen Use-Case voraussichtlich keinen Mehrwert leisten. Beachten Sie jedoch, dass häufig auch aus Makro-Daten Rückschlüsse auf Individuen gezogen werden können.

15 Einige **Trusted Execution Environments** können eine attraktive Lösung sein, denn sie erlauben anspruchsvolle Analysen unter hohen Sicherheitsstandards.

16 Je nach den Einzelheiten des Use-Cases könnten **Trusted Execution Environments oder homomorphe Verschlüsselung** eine attraktive Lösung sein. Wichtig ist jedoch, dass die Daten bereits an der Quelle verschlüsselt werden.

17 Dieser Use-Case kommt selten vor - gegebenenfalls könnte es **Differential Privacy** ermöglichen, die Analyse von Daten mit ähnlichen Merkmalen zu erlauben, ohne auf die Rohdaten zugreifen zu müssen. Differential Privacy bietet jedoch keine Sicherheit für die Rohdaten selbst.

18 PETs leisten für diesen Use-Case für gewöhnlich keinen Mehrwert. Konzentrieren Sie sich auf die Grundlagen, um Datenschutz- und Sicherheitsgrundsätze zu gewährleisten. Bewerten Sie außerdem regelmäßig Ihre vorhandene Infrastruktur und deren potenzielle Risiken, um Sicherheitslücken zu schließen.

4. Appendix: Methoden

In diesem Abschnitt legen wir unser methodisches Vorgehen für die systematische Marktanalyse und die Untersuchung von Fallstudien dar. Weiter verdeutlichen wir den Prozess der Auswahl und Durchführung der Expert*inneninterviews, die die initiale Literaturrecherche ergänzen und insbesondere das Anwendungsframework informieren.

Systematische Marktanalyse

Die systematische Marktanalyse dient der Erkundung der Anbieterseite von PET-Lösungen. Sie wurde durchgeführt, um den Markt auf Merkmale wie die regionale Verteilung von PET-Unternehmen oder seine zeitliche Entwicklung zu untersuchen und um herauszufinden, welche PET-Technologien von Anbietern in welchem Anteil aufgegriffen werden. Die Grundlage der Marktanalyse bildet ein Datensatz, der durch ein Suchmuster auf der Datenplattform Crunchbase erstellt wurde. Crunchbase stellt unter anderem Daten über Unternehmen und Investoren bereit und erfasst neben Variablen wie Namen, Gründungsjahr, Finanzierung, Branche und Größe des Unternehmens auch Unternehmensbeschreibungen („Description“).

Für das Suchmuster wurden auf Grundlage einer initialen Literaturrecherche Keywords und eine Bedingung definiert. Als Keywords wurden folgende Begriffe ausgewählt: „Homomorphic Encryption“, „Secure Multiparty Computation“, „Federated Learning“, „Differential Privacy“, „Synthetic Data“, „Trusted Execution Environments“, „Privacy-Enhancing Technologies“, „Data Clean Room“, „privacy-preserving“, „confidential computing“, „privacy engineering“. Als einzige weitere Bedingung wurde festgelegt, dass das Unternehmen aktiv sein muss. Die Keywords wurden auf das Textfeld der Unternehmensbeschreibungen angewendet. In der Folge ergab das Suchmuster einen Datensatz von zunächst 199 Unternehmen. Die 199 Unternehmen führten mindestens eines der ausgewählten Keywords in der Unternehmensbeschreibung und sind zum Zeitpunkt der Abfrage der Plattform zufolge aktiv gewesen. Bevor wir uns auf das beschriebene Suchmuster als Filter festgelegt haben, wurden elf weitere Suchmuster getestet. Die Suchmuster unterschieden sich in der Funktion, in der die Keywords zueinander standen („and“; „or“), in der Auswahl der Keywords selbst. Im Vergleich zu anderen Suchmustern stellt das Suchmuster eine verhältnismäßig enge Variante dar, da sie sich auf die Kerntechnologien in Kombination mit Oberbegriffen beschränkt, die sich in der Literaturrecherche als eng verwoben mit der Gruppe der Privacy-Enhancing-Technologies herauskristallisiert haben. So konnte mit größerer Wahrscheinlichkeit sichergestellt werden, dass Privacy-Enhancing Technologies zum Kerngeschäft der gefilterten Unternehmen zählen.

Im nächsten Schritt wurden die Unternehmen im Datensatz manuell daraufhin überprüft, ob Privacy-Enhancing Technologies in der Tat Teil ihrer angebotenen Lösung oder ihres Kerngeschäfts sind. In der Folge wurden vier Unternehmen ausgemustert und der finale Datensatz von 195 Unternehmen weltweit als Grundlage für die Analyse herangezogen. Die Datenbasis bietet eine geeignete Quelle, um (1) die zeitliche Entwicklung (Gründungsjahre), (2) die geographische Verteilung der Unternehmen (Unternehmenssitz), (3) Anhaltspunkte über die Finanzierungen auf dem PET-Markt (Funding) und (4) die Verwendung einzelner Technologien (Nennung der Technologie in der Unternehmensbeschreibung) zu analysieren. Unterschiede in den

Methoden

N-Werten der unterschiedlichen Analysen sind im Falle von (1), (2) und (3) durch fehlende Daten im Datensatz von Crunchbase bedingt und durch fehlende Nennungen der einzelnen Technologien im Falle von (4).

Dennoch resultiert die Auswahl der Plattform Crunchbase und die Entscheidung für ein Suchmuster notwendigerweise in einer Annäherung an den gesamten Markt. Der Datensatz erhebt deshalb nicht den Anspruch, die gesamte Anbieterseite für PET-Lösungen abzubilden.

Fallstudien

Bei der Auswahl der Fallstudien, die für diesen Bericht analysiert wurden, handelt es sich um ein informiertes Convenience Sample. Die zwei Bedingungen, die alle vier Fallstudien vereinen, sind

(1) der Einsatz von Privacy-Enhancing-Technologies für ein definiertes Ziel und

(2) die Beteiligung einer Organisation des öffentlichen Sektors.

Für die Durchführung der Fallstudien haben wir je Fallstudie mindestens ein semi-strukturiertes Interview mit Projektbeteiligten geführt. Bedingung für die Auswahl der Interviewpartner*innen war ihre Zugehörigkeit zur öffentlichen Organisation oder der Anbieterseite der PET-Lösung. Ergänzend zu den Interviews haben wir die Cases im Rahmen einer Literaturrecherche untersucht. Die Kombination aus Literaturrecherche und Interview ermöglichte, die Ergebnisse in der Regel über mehrere Quellen hinweg zu triangulieren.

Begleitende Interviews

Neben den Interviews für die Fallstudien haben wir elf semi-strukturierte Interviews mit Vertreter*innen von Start-ups mit einer Dauer von 30–60 Minuten geführt. Die ausgewählten Start-ups wurden über die Datenbasis für die Marktanalyse und Kontakte von PUBLIC ausgewählt. Die Start-ups aller interviewten Vertreter*innen bieten eine oder mehrere Lösungen an, die auf PETs zurückgreifen. Die Interviews dienten dem Zweck, die Erkenntnisse aus der initialen Literaturrecherche zu unterfüttern und als kritisches Korrektiv für die daraus gezogenen Erkenntnisse zu wirken. In den Interviews wurde insbesondere der öffentliche Sektor als Anwendungsbereich für PETs thematisiert. Vor diesem Hintergrund flossen die Erkenntnisse aus den begleitenden Interviews insbesondere in den Prozess zur Erstellung des Entscheidungsbaumes ein.

Bibliografie

A - I

Abowd, John M. (2018): Protecting the Confidentiality of Americas Statistic's: Adopting Modern Disclosure Avoidance Methods at the Census Bureau

Agostini, Maxime & PUBLIC (2022): Interview über PETs

Android (n. a.): Vertrauensvolles TEE

Bayerisches Landesamt für Steuern (2022): NESSI Nachweisplattform ELSTER Self-Sovereign Identities

Blatt, Marcel; Gusev, Alexander; Polyakov, Yuriy; Goldwasser, Shafi (2020): Secure large-scale genome-wide association studies using homomorphic encryption

BMI (2022): Daten intelligent und nachhaltig nutzen

BMI (2022): Digital-Gipfel 2022: Daten intelligent und nachhaltig nutzen

Boston University (2015): Computational Thinking Breaks a Logjam

BSI (2022): Die Lage der IT-Sicherheit in Deutschland 2022

BSI (2023): IT-Grundschutz-Kompendium – Werkzeug für Informationssicherheit

BSI (n. a.): Online-Kurs IT-Grundschutz, Lektion 4: Schutzbedarfsfeststellung, 4.1 Grundlegende Definitionen

BSI (n. a.): Geheimschutz

Buchner, Nicolas; Kinkelin, Holger; Rezabek, Filip (2022): Survey on Trusted Execution Environments

Bundesministerium der Justiz (2021): Sicherheitsüberprüfungsgesetz (SÜG) § 1

Bundesministerium der Justiz (2023): Verschlusssachenanweisung

BWWC (2022): Gender and Racial Wage Gaps in Boston by the Numbers

Center for Strategic & International Studies (2017): Forces Shaping the Next Generation of Cyber Threats to Financial Institutions

CIO Bund (2022): Architekturrichtlinie für die IT des Bundes

Desouza, Cevin C. et al. (2014): Big Data in the Public Sector: Lessons for Practitioners and Scholars

Cohen, Ronen; Rohloff, Kurt & PUBLIC (2022): Interview über PETs

Dwork, Cynthia; Roth, Aaron (2014): The Algorithmic Foundations of Differential Privacy

Eberhardt, Jacob; Tai, Stefan (2021): ZoKrates - Scalable Privacy-Preserving Off-Chain Computations

Eckert, Claudia (2009): IT-Sicherheit: Konzepte - Verfahren - Protokolle

EuGH (2016): Urteil Breyer gegen Bundesrepublik Deutschland

Eurekalert (2020): Duality Technologies researchers accelerate privacy-enhanced collaboration on genomic data

Europäisches Parlament (2018): DSGVO Art. 4

ENISA (2016): Readiness Analysis for the Adoption and Evolution of Privacy Enhancing Technologies

Gentry, Craig (2009): A Fully Homomorphic Encryption Scheme

FITKO (n.a.): Koordination IT-Planungsrat

Fraunhofer FIT (2021): Self-Sovereign Identity Foundations, Applications, and Potentials of Portable Digital Identities

GlobeNewswire (2020): Cyber Defence Alliance (CDA) partners with Anomali to better enable sharing of Threat Intelligence among banking members

Goldberg, Ian; Wegner, David; Brewer, Eric (1997): Privacy-Enhancing Technologies for the Internet

Hackermoon (2020): Homomorphic Encryption: Introduction And Use Cases

Homomorphic Encryption Standardisation (2023): Participants

IDUnion (n. a.): Über das Projekt

IEEE (2021): 2842-2021 - IEEE Recommended Practice for Secure Multi-Party Computation

Information Commissioner's Office (2022): Anonymisation, pseudonymisation and privacy enhancing technologies guidance, Chapter 5

J - W

Johnson, Alastir (2022): EIDAS 2.0 Turns To Self-Sovereign Identification To Bring Users Ownership And Control

Kairouz et al. (2021): Advances and Open Problems in Federated Learning

Klisch, Sabine; Ponikiewicz, Andreas & PUBLIC (2022): Interview über PETs.

Klother, Marvin & PUBLIC (2023): Interview über das ATLAS Projekt

Krishnan, Saravanan et al. (2020): Handbook of Research on Blockchain Technology

Lapets, Andrei (2018): Accessible Privacy-Preserving Web-Based Data Analysis for Assessing and Addressing Economic Inequalities

Lapets, Andrei & PUBLIC (2023): Interview zu PETs

Kaufman, Lauren & PUBLIC (2022): Interview über das FAIR TREATMENT Projekt

Lehmann, Anja (2019): ScrambleDB: Oblivious (Chameleon) Pseudonymization-as-a-Service

Li, Qinbin et al. (2021): A Survey on Federated Learning Systems: Vision, Hype and Reality for Data Privacy and Protection

Lissi (2022): eIDAS and the European Digital Identity Wallet: Context, status quo and why it will change the world.

McMahan, Brendan et al. (2017): Communication-Efficient Learning of Deep Networks from Decentralized Data

Michael, Helge & PUBLIC (2023): Interview zu PETs

Moore, Anna et al. (2022): FAIR TREATMENT: Federated analytics and AI Research across TREs for AdolescenT MENTal health

Moore, Anna & PUBLIC (2022): Interview über das FAIR TREATMENT Projekt

Nissim, Kobbi; Steinke, Thomas (2018): Differential Privacy: A Primer for a Non-technical Audience

OECD (2017): Embracing Innovation in Government

OECD (2023): Emerging Privacy Enhancing Technologies: Current Regulatory and Policy Approaches

Office of the Director of National Intelligence (2017): How would you explain homomorphic encryption?

PUBLIC & Sopra Steria (2022): Öffentliche APIs und GovTech: Mit Interoperabilität Innovation fördern

Quisquater, Jean-Jacques (1990): How to Explain zero-Knowledge Proofs to your Children

Gaboardi, Marco; Russo, Alejandro & PUBLIC (2023): Interview über PETs

Seker, Ensar (2021): Trusted Execution Environment (TEE), Implementations, Drawbacks

Stern, Klaus (1980): Das Staatsrecht der Bundesrepublik Deutschland: Band II

Strüker, Jens et al. (2021): Self-Sovereign Identity Foundations, Applications, and Potentials of Portable Digital Identities

The Royal Society (2023): Privacy Enhancing Technologies

Tourky, Dalia; ElKawkagy, Mohamed; Keshk, Arabi (2016): Homomorphic encryption the "Holy Grail" of cryptography

van Rijmenam, Mark (2019): How Zero-Knowledge Proof Increases Your Privacy While Enabling Trustless Transactions

Wagner, Paul (2021): Privacy Enhancing Technologies and Synthetic Data

Williams, Ellison Anne (2022): Privacy-enhancing technologies – myths and misconceptions

World Economic Forum (2020): Cyber Information Sharing: Building Collective Security

Danksagungen

Wir bedanken uns bei der Vielzahl von Menschen, die uns bei der Erarbeitung dieses Berichts unterstützt haben:

Wir danken den vielen Vertrer*innen von PET-Unternehmen, die im Rahmen der Expert*inneninterviews einen wertvollen Beitrag zu den Inhalten des Berichts geleistet haben. Ebenso danken wir den Expert*innen, welche mit ihren Erfahrungen über Anwendungsfälle von PETs im öffentlichen Sektor die Fallstudien bereichert haben.

Kontakt

Autoren

Jakob Kollotzek | PUBLIC

Research Associate
jakob.kollotzek@public.io

Robert Seethaler | Sopra Steria

Lead Expert
robert.seethaler@soprasteria.com

Leon Rückert | PUBLIC

Junior Analyst
leon.rueckert@public.io

PUBLIC

Wir beraten Regierungen, öffentliche Verwaltungen und Unternehmen mit Public-Sector-Bezug, die ihre Dienstleistungen und Systeme digital erneuern wollen. Wir vernetzen und unterstützen Gründer:innen, politische Entscheidungsträger:innen und Investor:innen, um gemeinsam die bestmögliche Zukunft für den öffentlichen Sektor zu identifizieren und zu gestalten.

Website: de.public.io
Twitter: @PUBLIC_Germany
Email: kontakt@public.io