Interuniversitäre Forschungsplattform
Zukunft – Technik – Gesellschaft (Z-T-G)

Privacy enhancing AI: Teilen ohne weitergeben

Endbericht des Forschungsprojekts Z-T-G 004

Projekt:	Z-T-G 004 (AI Privacy)
Datum:	25. August 2025
Ansprechperson:	Roman Lukas Prunč, roman.prunc@tugraz.at Christian Dayé, christian.daye@tugraz.at
Mitwirkende:	Forscher:innenteam des Z-T-G (in alphabetischer Reihenfolge): Christian Dayé[1], Lea Demelius[2], Andreas Trügler[2, 3], Roman Lukas Prunč[1], Bernhard Wieser[1] [1] Technische Universität Graz, STS Unit [2] Know Center GmbH, Graz [3] Universität Graz, Sermilik Research Station

Inhaltsverzeichnis

Zusammenfassung

Im Zuge des Projekts Privacy Enhancing AI: Teilen ohne Weitergeben (Z-T-G 04) wollten wir die Möglichkeiten erfassen, welche sich durch neue kryptografische Verfahren für politische Entscheidungsfindungen ergeben und welche Rolle Vertrauen hierbei spielt, beziehungsweise wie dieses beeinflusst wird. Konkret lauteten die beiden Forschungsfragen, welche unsere Untersuchung leiteten:

Forschungsfrage 1: Wie schätzen relevante Stakeholder die Möglichkeiten ein, durch die Kombination von Kryptographie und KI eine breitere Datengrundlage für Entscheidungsprozesse bereitzustellen?

Forschungsfrage 2: Welche Faktoren fördern das Vertrauen der Öffentlichkeit in solche kryptographischen KI-Technologien und welche sind diesem abträglich?

Zur Beantwortung dieser Fragen haben wir zwei zentrale Zugänge gewählt. Für die Erfassung der Perspektiven relevanter Expert:innen haben wir eine Reihe an leitfadengestützten Interviews durchgeführt. Zudem haben wir die Perspektive von interessierten Bürger:innen durch die Ausarbeitung von Futures Wheels im Rahmen eines Workshops und einer universitären Lehrveranstaltung eingeholt. Als Referenzprojekt diente uns die sogenannte Covid Heatmap, die es durch ihren Bezug zu einem hochpolarisierten Thema ermöglichte, die uns interessierenden Fragen in einem speziellen Kontext zu verorten, dessen Problematik allen bekannt ist.

Digitalisierung bedeutet neue Herausforderungen für die Wahrung des Grundrechts auf Privatheit und es gibt eine öffentliche Sensibilisierung für das in diesem Zusammenhang bedeutende Thema Datenschutz. Insbesondere bei sensiblen Daten gibt es in der Bevölkerung das Bedürfnis nach besonderen Schutzvorkehrungen. PETs können solche Mechanismen darstellen, mit denen neue Erkenntnisse gewonnen werden können, während zugleich der Datenschutz gewahrt wird.

Von politischer Seite muss dabei jedoch im Sinne demokratischer Prinzipien klar auf Transparenz und Freiwilligkeit gesetzt werden, da ansonsten Vorwürfe von Kontrolle und Überregulierung angebracht werden können.

Zudem ist es nicht nur die implementierte Technologie an sich, welche die Perspektive von Bürger:innen formt und auch deren Umsetzung ist nur zweitrangig, wenn auch nicht unwesentlich. Im Vordergrund steht stets der Zweck, zu welchem ein System eingeführt wird. Ist dieser belastet, kann dies selbst das datenschutzfreundlichste System nicht ausgleichen. Vielmehr noch werden die Zweifel, die dem Gesamtthema eigen sind, auch auf die Technologie übertragen.

Einleitung

Problemstellung

Digitale Datensätze ermöglichen schnelle und zielgerichtete computergestützte Auswertungen und Analysen, sowie auch eine Erweiterung der Perspektive durch die erleichterte Einbeziehung und Kombination unterschiedlicher Quellen. Doch diesen technischen Möglichkeiten gegenüber steht die Wahrung der Grundrechte und daraus folgende gesetzliche Bestimmungen, die mache Implementierungen einschränken. Dies ist etwa der Fall, wenn das Recht auf Privatheit durch die Verarbeitung von Daten mit Personenbezug betroffen ist. Der Schutz derartiger Informationen wird EU-weit durch die Datenschutz-Grundverordnung (DS-GVO) geregelt. Insbesondere wenn solche Daten von öffentlichem Interesse sind, ist es dementsprechend erforderlich, Lösungen zu finden, welche vorhandene Möglichkeiten nutzen, ohne die betreffenden Normen zu verletzen. Das betrifft zum Beispiel Fälle, in denen dadurch eine relevante Wissensgrundlage für politische Entscheidungen geschaffen werden kann. Exemplarisch für solche Problemstellungen steht die Frage nach Infektionszahlen, -herden und -clustern sowie Verbreitungsmustern während der Covid-19 Pandemie. In diesem Kontext bekam die Frage, wie sensible Daten als Entscheidungsgrundlage herangezogen werden können, während gleichzeitig ein Höchstmaß an Privatheit gewährleistet bleibt, besondere politische Bedeutung. Einen Lösungsansatz hierfür stellen neue Methoden der Kryptographie dar, die eine Verarbeitung und analytische Auswertung von sensiblen Datensätzen durch machine learning Algorithmen ermöglichen, ohne dabei Informationen über Einzelpersonen preiszugeben.

Davon ausgehend war es das Ziel dieses Projekts, die Potentiale von „privacy enhancing technologies“ (PETs) zu erörtern und ihre Nutzung im Kontext politisch-administrativer Entscheidungen zu bewerten. Dabei wurde erfasst, wie relevante Expert:innen sowie Bürger:innen den Mehrwert einschätzen, der sich durch kryptographisch geschützte Datenanalyse, auch in Kombination mit machine learning gestützter Verarbeitung, ergibt und wo sie dabei besondere Herausforderungen sehen.

Das öffentliche Interesse, zum Beispiel Infektions-Hotspots zu kennen – und sie somit bewusst vermeiden zu können –, liegt auf der Hand. Dementsprechend groß ist daher auch das Potential der Möglichkeit zur Analyse von Daten ohne deren Weitergabe („privacy preserving data analytics“). Doch wenn der Schutz sensibler Daten durch den Einsatz einer bestimmten Technologie übernommen wird, wirft das in profunder Weise Fragen nach dem gesellschaftlichen Vertrauen in diese Technologie und letztendlich auch der Akzeptanz ihrer Implementierung auf. Somit ist es nicht die technische Möglichkeit allein, die den Erfolg solcher KI-Anwendungen bedingt, sondern auch der Grad, mit dem die Öffentlichkeit diesen KI-Methoden vertraut und damit „trustworthy data sharing“ ermöglicht wird. Aus diesem Grund stellt sich die Frage, was in dieser Hinsicht vertrauensbildend wirkt bzw. Anlass zu Zweifel und Misstrauen gibt.

Vor diesem Hintergrund wurde im Rahmen dieses Projekts ausdrücklich nach den Faktoren gesucht, die zur Ausbildung eines solchen Vertrauens beitragen oder dieses gegebenenfalls unterminieren. Bewusst wurden dabei nicht nur technische, sondern auch soziale Aspekte bewertet, und zwar sowohl aus Expert:innen-Perspektive als auch aus der Sicht einer informierten Öffentlichkeit. Als veranschaulichendes Beispiel und grundlegender Anknüpfungspunkt diente dabei die im Zuge des DDAI COMET Moduls entstandene Covid-19 Heatmap, bei der Daten des Gesundheitsministeriums mit Telekom-Anbietern abgeglichen werden, um die Covid-19 Ansteckungsherde der vergangenen Tage zu identifizieren.

Konkret ausformuliert lauteten die Forschungsfragen, die dem Projekt zugrunde lagen, wie folgt:

Forschungsfrage 1: Wie schätzen relevante Stakeholder die Möglichkeiten ein, durch die Kombination von Kryptographie und KI eine breitere Datengrundlage für Entscheidungsprozesse bereitzustellen?

Forschungsfrage 2: Welche Faktoren fördern das Vertrauen der Öffentlichkeit in solche kryptographischen KI-Technologien und welche sind diesem abträglich?

Zur Beantwortung dieser Fragen wurden unterschiedliche Herangehensweisen gewählt. Zum einen wurden durch leitfadengestützte Expert:inneninterviews sowohl die Potenziale von PETs zur Schaffung einer breiteren Wissensbasis als auch die Rolle des Vertrauens im Zuge darauf basierender öffentlicher Entscheidungen erhoben. Komplementär dazu wurde die Beantwortung der zweiten Forschungsfrage zudem durch die Erfassung von Bürger:innenperspektiven mittels der Ausarbeitung von Futures Wheels sowie Gesprächen im Zuge einer Informationsveranstaltung vorangetrieben.

Das Projekt zielt damit letztlich darauf ab, gesellschaftspolitische Entscheidungen nachvollziehbarer zu machen und die Wissensbasis, auf deren Grundlage sie getroffen werden, einer breiten Öffentlichkeit zur Verfügung zu stellen. Auch können dadurch Hinweise abgeleitet werden, mit welchen Maßnahmen die Anwendung kryptographischer KI-Methoden begleitet werden soll. Auf diese Weise befasst sich das Projekt mit wesentlichen Prinzipien des Open Government, nämlich mit Transparenz von Entscheidungsfindungen und Verantwortlichkeit, und klärt, wie „privacy enhancing technologies" dazu einen vertrauenswürdigen Beitrag leisten können.

Öffentliches Interesse

Die Bereitstellung von wissenschaftlichem Wissen ist als Grundlage gesellschaftlicher Entscheidungsprozesse von großer Wichtigkeit. Dies ist jedoch nicht nur für die Lenkung der Entscheidungsprozesse selbst von großem Wert, sondern auch für die Möglichkeit, die Grundlagen dieser Entscheidungen für die breite Öffentlichkeit nachvollziehbar zu machen. Die Herausforderung dabei ist es, Daten und sich daraus ergebende Fakten zu drängenden gesellschaftlichen Problemstellungen für Bürger:innen und Politiker:innen verfügbar zu machen, dabei jedoch keinerlei Privatheitsgrundsätze zu verletzen. Eine solche Möglichkeit kann sich durch den Einsatz moderner kryptografischer Methoden wie PETs eröffnen. Dadurch können sowohl separate Datensätze unterschiedlicher Eigner für eine gemeinsame Auswertung eingesetzt werden, ohne die in ihnen enthaltenen Informationen mit den jeweils anderen zu teilen, als auch einzelne Datensätze bereitgestellt werden, ohne dabei Rückschlüsse auf Individuen zu erlauben. Insbesondere kann damit dem Bestreben von Open Government Initiativen entsprochen werden, durch das Verfügbarmachen der Wissensgrundlagen, auf deren Basis politische Entscheidungen getroffen werden, Transparenz und Accountability dieser Prozesse zu gewährleisten.

Die Möglichkeit, separate Datensätze miteinander zu verbinden sowie die Auswertung ohne Personenbezug, versprechen enorme Erkenntnisfortschritte, indem zuvor unerfassbare Zusammenhänge sichtbar gemacht werden. Das beginnt bei gesundheitsbezogenen und epidemiologischen Daten, wie es etwa während der Pandemie deutlich wurde (Bampoulidis et al., 2022), betrifft aber ebenso wirtschafts- und sozialpolitische Aufgaben und Entscheidungen. Jedoch sind die Chancen, welche sich durch solche Auswertungen bieten auch begleitet von einer Reihe technischer, juristischer und politischer Fragen, welche es zu beantworten gilt.

Datenschutz

Datenschutz ist eines der wichtigsten Mittel, um die Wahrung des Grundrechts auf Privatheit zu gewährleisten. Insbesondere im Sinne einer informationellen Privatheit, also dem Schutz vor unerwünschtem Zugriff auf persönliche Informationen (siehe Rössler, 2002) Die Verarbeitung und Bereitstellung von Daten und die Kombination distinkter Datenquellen zu analytischen Zwecken unterliegt dementsprechend vor allem dann einem besonderen Schutz, wenn es sich um personenbezogene Daten handelt. Die maßgeblichen Bestimmungen dazu sind in den Ländern der Europäischen Union und dementsprechend auch in Österreich durch die seit 2018 geltende Datenschutz-Grundverordnung (DS-GVO) festgelegt.

Personenbezogene Daten im Sinne der DS-GVO sind in Art. 4 Z 1 definiert als alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person beziehen. Dabei wird in ErwGR 26 angeführt, dass für die Feststellung der Identifizierbarkeit alle Mittel berücksichtigt werden, die von dem Verantwortlichen oder einer anderen Person

nach allgemeinem Ermessen wahrscheinlich genutzt werden, um die natürliche Person direkt oder indirekt zu identifizieren. Nicht in den Geltungsbereich der DS-GVO fallen demzufolge Daten, bei denen kein Personenbezug hergestellt werden kann. Dies sind neben anonymisierten Daten auch verschlüsselte Daten in Bezug auf alle, die nicht in Besitz des entsprechenden Schlüssels sind. In diesem Sinne kann der Einsatz von PETs dazu dienen Datenschutz und den von der DS-GVO ebenso als Schutzziel definierten freien Datenverkehr zugleich zu verwirklichen (Bierbauer & Helminger, 2023).

Dies entspricht auch dem langfristigen Zugang der Europäischen Kommission, die sich bereits 2007 dafür eingesetzt hat, sowohl die Entwicklung als auch den Einsatz von PETs zu fördern (KOM(2007) 228), da sie sich von diesen einen verbesserten Schutz der Privatsphäre versprach (Mitteilung der Kommission an das Europäische Parlament und den Rat über die Verbesserung des Datenschutzes durch Technologien zum Schutz der Privatsphäre, 2007).

Im Sinne der entsprechenden Formulierungen der DS-GVO sowie der politischen Willensbekundung von Seiten der Kommission stellt sich der Einsatz von PETs als Schutzmaßnahme bei der Verarbeitung von Datensätzen, welche personenbezogene Daten enthalten, nicht als Umgehungsmaßnahme dar, sondern als ein erwünschtes Mittel zur Eröffnung neuer analytischer Zugänge.

Künstliche Intelligenz

Der Ausgangspunkt der Künstlichen Intelligenz in Wissenschaft und Forschung wird zumeist mit dem 1956 abgehaltenen Dartmouth Summer Research Project on Artificial Intelligence festgesetzt. Die dem Projekt zugrundeliegende Annahme war, dass es möglich sei, menschliche Denkprozesse maschinell nachzubilden. Dementsprechend erwartete man sich „that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer." (McCarthy et al., 2006, S. 12)

Nahezu 70 Jahre später ist der Begriff Künstliche Intelligenz fest im alltäglichen Wortschatz verankert. Jedoch umfasst er eine Reihe unterschiedlicher Systeme, weswegen eine Vielzahl unterschiedlicher Definitionen existiert, die jeweils unterschiedliche Aspekte in den Vordergrund stellen. In diesem Sinne unterscheiden Russel und Norvig (2012) vier verschiedene Ausprägungen künstlicher Intelligenz: Menschliches Denken, Menschliches Handeln, Rationales Denken und Rationales Handeln (S.23). Eine andere Unterscheidung von KI-Systemen nach ihrer generativen oder prädiktiven Funktionsweise (*Generative AI vs. Predictive AI*, 2024) hat insbesondere in den letzten Jahren mit der Veröffentlichung populärer Large Language Models (LLM) zunehmend an Bedeutung gewonnen.

In der EU ist die Anwendung von KI durch die Verordnung (EU) 2024/1689 über künstliche Intelligenz (AI Act) geregelt. Angesichts der Breite an Bedeutungen, die der Begriff der Künstlichen Intelligenz umfassen kann, kommt der Definition innerhalb des Gesetzestextes eine besondere Bedeutung zu. Art 3 Z 1 bestimmt als „KI-System" ein maschinengestütztes System, das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können. Etwas konkreter wird zudem ErwGR 12 in dem es heißt:

> Ein wesentliches Merkmal von KI-Systemen ist ihre Fähigkeit, abzuleiten. Diese Fähigkeit bezieht sich auf den Prozess der Erzeugung von Ausgaben, wie Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen, die physische und digitale Umgebungen beeinflussen können, sowie auf die Fähigkeit von KI-Systemen, Modelle oder Algorithmen oder beides aus Eingaben oder Daten abzuleiten. Zu den Techniken, die während der Gestaltung eines KI-Systems das Ableiten ermöglichen, gehören Ansätze für maschinelles Lernen, wobei aus Daten gelernt wird, wie bestimmte Ziele erreicht werden können, sowie logik- und wissensgestützte Konzepte, wobei aus kodierten Informationen oder symbolischen Darstellungen der zu lösenden Aufgabe abgeleitet wird. Die Fähigkeit eines KI-Systems, abzuleiten, geht über die einfache Datenverarbeitung hinaus, indem Lern-, Schlussfolgerungs- und Modellierungsprozesse ermöglicht werden. Die Bezeichnung „maschinenbasiert" bezieht sich auf die Tatsache, dass KI-Systeme von Maschinen betrieben werden. […] KI-Systeme sind mit verschiedenen Graden der Autonomie ausgestattet, was bedeutet, dass sie bis zu einem gewissen Grad unabhängig von menschlichem Zutun agieren und in der Lage sind, ohne menschliches Eingreifen zu arbeiten. Die Anpassungsfähigkeit, die ein KI-System nach Inbetriebnahme aufweisen könnte, bezieht sich auf seine Lernfähigkeit, durch sie [*sic*] es sich während seiner Verwendung verändern kann.

Insbesondere aufgrund des selbstlernenden Faktors sind die Ergebnisse, welche von KI-Systemen produziert werden, nicht immer für menschliche Betrachter nachvollziehbar beziehungsweise erklärbar. Diese Unzugänglichkeit des zugrundeliegenden Algorithmus wird zumeist mit dem Begriff „black box" beschrieben und verdeutlicht eines der großen Probleme, die sich beim Einsatz künstlicher Intelligenz für Entscheidungsfindungsprozesse ergeben. Dementsprechend gibt es unter dem Titel „Explainable Artificial Intelligence" (XAI) Bestrebungen, insbesondere für den politisch-administrativen Bereich, erklärbare und transparente KI-Anwendungen zu gestalten (European Data Protection Supervisor, 2023).

Privacy Enhancing Technologies (PETs)

Die Benennung Privacy Enhancing Technologies umfasst eine Reihe an Methoden „zum Schutz der Privatsphäre durch Eliminierung oder Verminderung personenbezogener Daten

oder durch Vermeidung einer unnötigen und/oder unerwünschten Verarbeitung von personenbezogenen Daten ohne Verlust der Funktionalität des betreffenden Informationssystems." (Mitteilung der Kommission an das Europäische Parlament und den Rat über die Verbesserung des Datenschutzes durch Technologien zum Schutz der Privatsphäre, 2007) Dabei unterscheiden sich PETs teilweise sehr stark in ihrem grundlegenden Zugang und ihrem Einfluss auf die Daten sowie etwaiger unter ihrem Einsatz durchgeführter Verarbeitungen. Zu diesen Methoden zählen zum Beispiel die Generierung eines auf realen Informationen aufbauenden synthetischen Datensatzes für Analysezwecke (Baumgartner et al., 2023), die Secure Multiparty Computation (SMPC), mittels welcher mehrere Parteien Daten gemeinsam analysieren können, ohne diese miteinander zu teilen (Bierbauer & Helminger, 2023), oder das sogenannte Federated Learning bei dem mehrere Parteien gemeinsam ein Modell trainieren (Alkaeed et al., 2024; Torkzadehmahani et al., 2022).

Im Rahmen dieses Projekts standen zwei weitere spezifische PET-Ansätze besonders im Fokus: Differential Privacy und Homomorphe Verschlüsselung

Differential Privacy (DP)

Die Grundidee hinter Differential Privacy ist es, einen Datensatz durch die Einbeziehung zufälliger Variablenausprägungen beziehungsweise Rauschen so abzuändern, dass ein tatsächlicher Rückschluss auf individuelle Parameter nicht mehr möglich ist. Da diese „Störungen" kontrolliert zugefügt werden und auf statistisch festgelegten Verteilungen basieren, können sie auf der Aggregatebene herausgerechnet werden. Somit entsprechen Ergebnisse von Berechnungen, die an mit DP geschützten Daten durchgeführt wurden annähernd jedoch nicht exakt jenen, welche durch den originalen Datensatz erreicht worden wären. Dadurch können Gesamtergebnisse erzielt werden, ohne den individuellen Schutz der jeweiligen Datensubjekte zu verletzen und Rückschlüsse auf diese zuzulassen (Alkaeed et al., 2024; Torkzadehmahani et al., 2022).

Homomorphe Verschlüsselung – Homomorphic Enryption (HE)

Homomorphe Verschlüsselung bedeutet ein mathematisches Verändern der Daten, wobei dabei jedoch deren grundlegende Struktur beziehungsweise die in ihnen vorhandenen Relationen erhalten bleiben. Die Besonderheit, die sich durch den Einsatz von HE ergibt, ist, dass es diese Methode erlaubt, Berechnungen direkt an verschlüsselten Daten durchzuführen. Auch die Ergebnisse dieser Operationen verbleiben verschlüsselt und können lediglich mit dem richtigen Schlüssel korrekt ausgelesen werden. Durch dieses Verfahren ist es zu keinem Zeitpunkt der Berechnung notwendig, die Daten zu entschlüsseln, wodurch sich eine Vielzahl an neuen Möglichkeiten gemeinsamer Datennutzung ergibt, auch im Sinne der vereinigten Auswertung mehrere distinkter Datensätze (Alkaeed et al., 2024). Der große Vorteil dieser Methode ist, dass die Ergebnisse der Berechnungen schließlich exakt jenen entsprechen, welche mit unverschlüsselten Daten gewonnen werden.

PETs und Künstliche Intelligenz

Obwohl PETs, wie auch andere kryptographische Methoden, nur bedingt auf dem Einsatz von KI-Systemen beruhen, stellen sie dennoch eine Schlüsseltechnologie dar, wenn es darum geht, die Potentiale von KI für gesellschaftlich relevante Fragestellungen zu nutzen. Insbesondere die beiden für dieses Projekt zentralen PETs, HE und DP, beruhen auf mathematischen Operationen, welche nicht auf den Einsatz von KI angewiesen sind. Dadurch sind sie in ihrer Funktionsweise transparent und beweisbar und ihr Einsatz muss nicht auf dem Vertrauen in die Berechnungen eines black box Algorithmus beruhen. Dennoch können hier machine learning Modelle eingesetzt werden, um die einzelnen Parameter der Verschlüsselung zu definieren (Alkaeed et al., 2024). Auch bei der Generierung synthetischer Datensätze können KI-Systeme zum Tragen kommen, um diese entsprechend einer realen Vorlage zu erstellen.

KI kommt im Kontext von PETs insbesondere dann zum Einsatz, wenn es um die Überprüfung der Sicherheit einer Verschlüsselung geht. Einerseits durch simulierte Angriffe (Jovic et al., 2022; Krček et al., 2022), aber auch durch die Erkennung von Anomalien in der Funktion eines bestehenden Systems, wodurch mögliche reale Angriffe entdeckt werden können.

Eine besondere Rolle kommt jedoch den PETs in umgekehrter Richtung zu, nämlich als Schutzmechanismus bei der Anwendung von KI-Systemen. Durch ihren Einsatz können machine learning Algorithmen trainiert oder für die Auswertung von Datensätzen genutzt werden, ohne dabei personenbezogene Informationen preiszugeben. Jedoch bringen spezifische PETs jeweils auch spezifische Vor- und Nachteile in diesem Kontext mit sich. Während DP hier eine schnelle Verarbeitung erlaubt, sind die dadurch generierten Ergebnisse nicht präzise. HE im Gegenzug erlaubt exakte Berechnungen sowie die Verschlüsselung von Daten und Model, verlangt jedoch auch eine höhere Rechenleistung, wodurch der Einsatz für große Berechnungen ineffizient wird (Alkaeed et al., 2024; Baumgartner et al., 2023; Sousa & Kern, 2023; Torkzadehmahani et al., 2022). Somit unterscheiden sich die beiden Ansätze auch grundlegend hinsichtlich ihres Schutzziels. DP verhindert Rückschlüsse von der Auswertung auf Individuen in den Eingangsdaten (output privacy), während HE die Daten selbst im Zuge der Auswertung schützt, aber keine Rückschlüsse auf Einzelne aus den Ergebnissen verhindert (input privacy).

Corona Heatmap

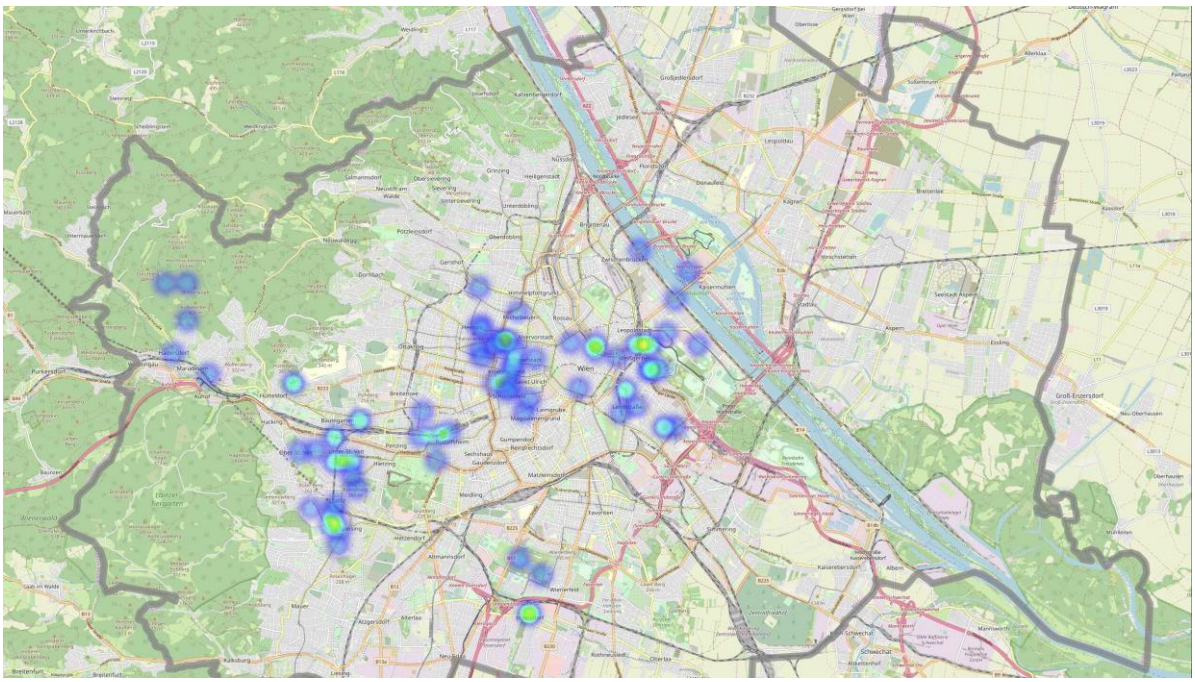
Als Referenzbeispiel für die Anwendung von PETs diente die im Zuge des DDAI COMET Moduls[1] entstandene Corona Heatmap, welche vom Know Center gemeinsam mit der TU Graz entwickelt wurde. Dabei handelt es sich um eine technische Lösung, bei der zwei

[1] https://www.know-center.at/research/comet-modul/ddai-data-driven-artificial-intelligence/

distinkte Datensätze von unterschiedlichen Quellen miteinander genutzt und ausgewertet werden können (*COVID Heatmap – COVID Heatmap*, n.d.).

Konkret präsentierte das Projekt eine Lösung, mittels der eine Nachverfolgung von möglichen Covid-19 Infektionsherden möglich ist. Dafür werden Infektionsdaten der Gesundheitsbehörde mit den Bewegungsdaten von Mobilfunkbetreibern abgeglichen, wodurch Bereiche identifiziert werden können, in denen sich positiv getestete Personen vermehrt aufgehalten haben. Es muss an dieser Stelle betont werden, dass dieses Projekt niemals mit realen Daten umgesetzt, sondern lediglich mit simulierten Daten veranschaulicht wurde (Bampoulidis et al., 2022; *COVID Heatmap – COVID Heatmap*, n.d.).

Abbildung 1: Corona Heatmap[2]



Damit weder die Mobilfunkanbieter erfahren, welche Personen infiziert sind, noch die Gesundheitsbehörde Einzelne überwachen kann, wurden in die Implementierung Schutzmechanismen in Form von PETs integriert. Die mittels HE verschlüsselten Telefonnummern infizierter Personen werden an den Mobilfunkbetreiber gesendet, welcher darauf aufbauend wiederum aggregierte Bewegungsprofile berechnet, ohne diese selbst entschlüsseln zu können. Diese werden danach an die Behörde zurückgesendet, welche den notwendigen Schlüssel besitzt und somit als einzige das Ergebnis auslesen kann. Um zusätzlichen individuellen Schutz zu gewährleisten, wird neben HE auch DP

[2] https://covid-heatmap.isec.tugraz.at/wp-content/uploads/2020/04/blue_heat.jpg

eingesetzt, damit die Spuren Einzelner nicht erkennbar sind (Bampoulidis et al., 2022; *COVID Heatmap – COVID Heatmap*, n.d.).

Das Projekt Corona Heatmap verdeutlicht somit klar, wie durch die Kombination zweier Datensätze ein gesellschaftlicher Mehrwert geschaffen werden kann, ohne dabei den Schutz personenbezogener Daten zu verletzen.

Vertrauen

Im Rahmen dieses Projekts wird eine Reihe von zentralen Herausforderungen thematisiert, die sich im Zusammenhang mit der Auswertung von personenbezogenen beziehungsweise sensiblen Daten ergeben. Damit rücken nicht nur technische Möglichkeiten zum Schutz von Privatheit und die Erhaltung fundamentaler Freiheitsrechte in den Fokus, sondern auch die Frage, wie solche Instrumente Akzeptanz, Vertrauens- und Glaubwürdigkeit gewinnen können und wo die Herausforderungen dabei liegen (Taddeo, 2009).

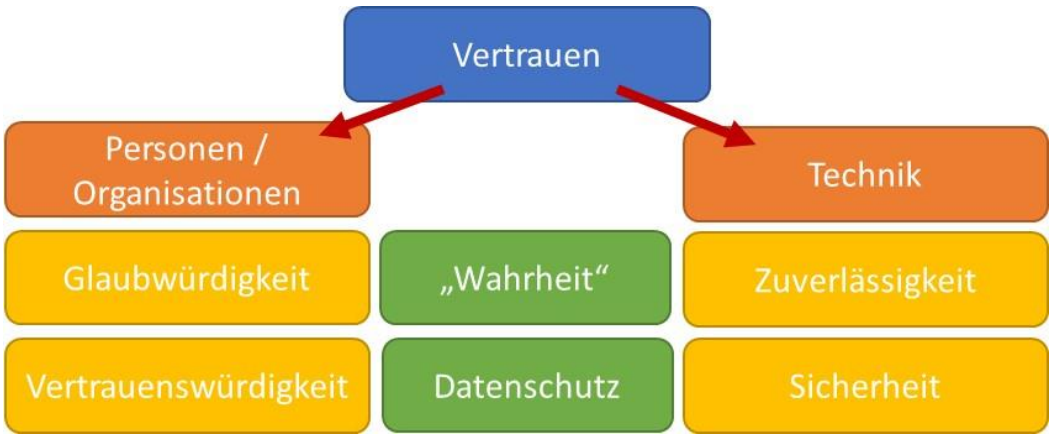
Vertrauen ist ein zentrales Element in der Beziehung zwischen Menschen und Technologie. In Zeiten zunehmender Digitalisierung, in denen sensible Daten alltäglich über digitale Infrastrukturen übertragen werden, gewinnt das Vertrauen, das Menschen in die den digitalen Raum regelnden Schutzmechanismen legen, sowohl in gesellschaftlicher Hinsicht wie auch aus Sicht einer funktionierenden demokratisch-staatlichen Ordnung immens an Bedeutung.

Während technische Sicherheit oft durch mathematische Modelle und Zertifizierungsverfahren nachgewiesen werden kann, bedeutet dies nicht zwangsläufig, dass Nutzer:innen solchen Technologien auch vertrauen (Beck, 1986; Pavlou, 2003). Dieses Spannungsverhältnis zwischen technischer Evidenz und subjektivem Vertrauen bildet das Ausgangsinteresse, mit dem wir uns der Erhebung der Sicht von Bürger:innen auf Privacy-Enhancing Technologies (PETs) zuwandten.

In den Sozialwissenschaften gilt Vertrauen als ein vielschichtiges Konzept, das nicht allein rationalen Abwägungen folgt, sondern auch von sozialen, kulturellen und emotionalen Faktoren geprägt ist (Giddens, 1996; Misztal, 1996). Frühere Studien zeigen, dass technisches Verständnis, mediale Berichterstattung (Rogers, 1995), institutionelle Glaubwürdigkeit (Siegrist et al., 2000) sowie persönliche Erfahrungen mit digitalen Technologien (Gefen et al., 2003) zentrale Einflussgrößen für das Vertrauen in technische Systeme darstellen können. Studien neueren Datums haben demgegenüber betont, dass Vertrauen in Technik, in Personen und in soziale Institutionen voneinander zu unterscheidende psychologische Phänomene sind und es daher ratsam ist, sich derartigen Fragen mit einem offenen, multifaktoriellen Denkrahmen anzunähern (Uttenthal, 2024).

Eines der Erkenntnisziele im Rahmen des Projekts war es daher, diese verschiedenen Arten von Vertrauen und ihre Bedeutung für den Erfolg einer Technologie wie PETs zu erfassen und zu verstehen, welche Faktoren Vertrauenswürdigkeit fördern oder zunichtemachen können. Als wesentliche Orientierung diente dabei die in Abbildung 2 dargestellte Untergliederung der Vertrauensarten.

Abbildung 2: Zwei Arten von Vertrauen



Forschungsdesign und -methoden

Erhebung der Expert:innensicht

Den Kern der Erfassung von Expert:innenperspektiven bildeten qualitative, leitfadengeleitete Interviews. Für die Erstellung und Verfeinerung des entsprechenden Leitfadens wurde zunächst, neben der Befassung mit entsprechender Literatur auch der Rahmen eines Panels der STS Konferenz Graz 2024 genutzt. Im Zuge des Vortrages „Privacy enhancing Technologies: socio-technical reconfigurations of trust" befragte Bernhard Wieser das Publikum per Mentimeter Umfrage[3] zu Vertrauen und Anwendung von PETs in verschiedenen Kontexten und suchte danach die Diskussionen mit den Anwesenden. Diese ersten Ergebnisse lieferten Denkanstöße, um gezielte Aspekte in den nachfolgenden Gesprächen ins Auge zu fassen. Insbesondere waren dies eine detailliertere Betrachtung der epistemischen Gemeinschaften, denen die Interviewpartner:innen angehören, sowie die Feststellung, dass Vertrauen in eine Technologie und deren Unterstützung nicht unbedingt Hand in Hand gehen, da für letzteres auch der Anwendungskontext von Bedeutung ist.

Auf diesem Wissen aufbauend wurde ein Interviewleitfaden entworfen (Siehe Abbildung 3) und verfeinert, der leicht und flexibel an die jeweiligen Gesprächspartner:innen angepasst werden kann. Im Wesentlichen gliedert er sich in drei große Blöcke, die jedoch teilweise ineinandergreifen. Diese sind:

- Gegenwart: Fragen nach dem Ist-Zustand
- Vertrauen: Fragen zur Einstellung gegenüber der Technologie und deren Einfluss auf gesellschaftliche Akzeptanz, also auf die Bereitschaft, die Implementierung eines Systems hinzunehmen
- Zukunft: Fragen nach Potenzialen und möglichen Entwicklungen

Als Vorbereitung auf jedes der Interviews wurde stets eine extensive Literaturrecherche vorgenommen, um nicht nur einen tiefergehenden Einblick in das jeweilige Fachgebiet des Gegenübers zu gewinnen, sondern auch um mit etwaigen Publikationen der Interviewpartner:innen im Vorhinein bekannt zu sein und das Gespräch entsprechend gestalten zu können, ohne unnötige Redundanzen aufkommen zu lassen. Insgesamt wurden 6 Interviews mit Gesprächspartner:innen aus den Bereichen Informatik, Rechtswissenschaft, Datenschutz und Technikfolgenabschätzung durchgeführt. Jedes Interview dauerte zwischen ein- und eineinhalb Stunden. Wenn möglich, wurden die Interviews im persönlichen Kontakt durchgeführt, doch bei zu großer räumlicher Distanz wurde die online Kommunikationsplattform Webex[4] genutzt.

[3] https://www.mentimeter.com/
[4] https://www.webex.com/

Die Auswertung der Interviews erfolgte im Sinne einer strukturierenden qualitativen Inhaltsanalyse (Mayring, 2015) welche durch die vorhergehend festgelegten Kategorien und theoretischen Hintergründe geleitet wurde.

Abbildung 3: Interviewleitfaden

Thema	Fragen	Aspekte
Die Gegenwart		
Privacy Enhancing Technologies (PETs)	Können Sie uns kurz beschreiben, in welcher Rolle bzw. aus welcher Perspektive Sie sich mit Privacy beschäftigten?	Was ist der persönliche Zugang/Background? Wie verwenden Sie es? Gibt es konkrete Anwendungsfälle? Was für Verschlüsselungsmethoden werden eingesetzt? Gäbe es Alternativen? Spielen AI-gestützten Datenverschlüsselungsverfahren eine Rolle?
Aktuelle Einsatzgebiete	Können Sie uns über ihren direkten Bereich hinaus konkrete Beispiele nennen, wo Privacy besonders relevant ist? Wo die Verfahren, mit denen Sie sich befassen zum Einsatz kommen oder kommen werden, könnten, sollten? (Wo auch AI-gestützte Datenverschlüsselungsverfahren aktuell bereits zum Einsatz kommen?) Und unterscheiden sich diese von Ihrem Einsatzgebiet bzw. von den von Ihnen eingesetzten Methoden?	- Industrie? - Business to business? - Business to customer? - Öffentlichkeit? Kriminalitätsbekämpfung?
	Was ist aus Ihrer Sicht das Besondere an diesen Verfahren?	Was sind die spezifischen Vorteile? Wie sicher ist es wirklich? Wie sieht die rechtliche Situation aus? Bedeutet das eine Neuerung? GDPR? Ist es theoretisch denkbar, es zu „knacken"? Welches Grundwissen setzt die Anwendung voraus? Wer entwickelt und implementiert die Verfahren? Open Source? Wie überprüft man, ob es funktioniert?

Thema	Fragen	Aspekte
Entwicklungsschritte	Was sind aus Ihrer Sicht aktuell die größten Herausforderungen für die Durchsetzung von AI-gestützten Datenverschlüsselungsverfahren?	- Technisch Hardware - Technisch Software - Rechtlich - Soziale Akzeptanz - Vertrauen? -
Vertrauen		
Vertrauen als Faktor sozialer Akzeptanz	Stichwort soziale Akzeptanz: Wie nehmen Sie aktuell die öffentliche Meinung bezüglich ihres Einsatzgebietes wahr (Sammeln der Daten und ihr Zweck)? Welche Rolle spielt hier das Vertrauen in die Methode der Verschlüsselung bzw in Privacy Technologien?	- In Österreich - In der EU - Anderswo - Gibt es das Thema in der Öffentlichkeit? Also wird es in seiner Spezifität wahrgenommen? - Was wird erfasst?
	Ist Vertrauen ein zentraler Faktor für die Akzeptanz seitens der Bürger:innen oder Kund:innen bei der Verarbeitung von personenbezogenen Daten? Welche Faktoren spielen ihrer Ansicht nach eine Rolle für Vertrauen?	- Technisch/Informatisches Wissen? Technologieskepsis? - Vertrauen in Institutionen die Technologien anwenden - Thema im Rahmen dessen die Technologie zum Einsatz kommt - Ist die notwendige Ungenauigkeit förderlich oder hinderlich bei der Schaffung von Vertrauen? - Vertrauen in Methode vs. Vertrauen auf Ergebnisse - Wie relevant sind unscharfe Ergebnisse für Politik(beratung)
	Wie könnte Vertrauen gestärkt werden?	- Transparenz? - Technisches Wissen? - Wissenschaft? - Kommunikation?

Thema	Fragen	Aspekte
Die Zukunft		
	Was glauben Sie, wie werden die weiteren Entwicklungen im AI gestützten Privacy-Bereich in der näheren Zukunft aussehen?	-
Potentiale	Wenn wir an den Bereich der öffentlichen Verwaltung denken (Beispiel Covid Heatmap), glauben sie, dass AI gestützte Verschlüsselungsverfahren Möglichkeiten schaffen können, die Entscheidungsprozesse beeinflussen und leiten können?	- Mögliche Einsatzgebiete - Politischer Wille/Bereitschaft - unterscheiden sich Bereiche (z.B. Gesundheitswesen, Steuerwesen usw.)
Risiken	Sehen Sie Faktoren, die dazu führen könnten, dass die spezifische Privacy stärkende Verfahren/Technologien sich nicht durchsetzen können?	- Rechtlich - Technisch - Gesellschaftlich - Einstellung zur Verwendung der Daten: Kontrolle vs. Exploitation Wo könnte man hier ansetzen? Wissen vermitteln? Vertrauen stärken?

Schlussendlich wurden durch den Austausch mit anderen Forscher:innen, die mit ähnlichen Themen befasst sind, im Zuge zweier Fachpanels bei Konferenzen (EASST-4S 2024, STS Graz 2025) weitere Einblicke und neue sozialwissenschaftliche Perspektiven zum Einsatz von PETS gewonnen.

Erhebung der Bürger:innensicht

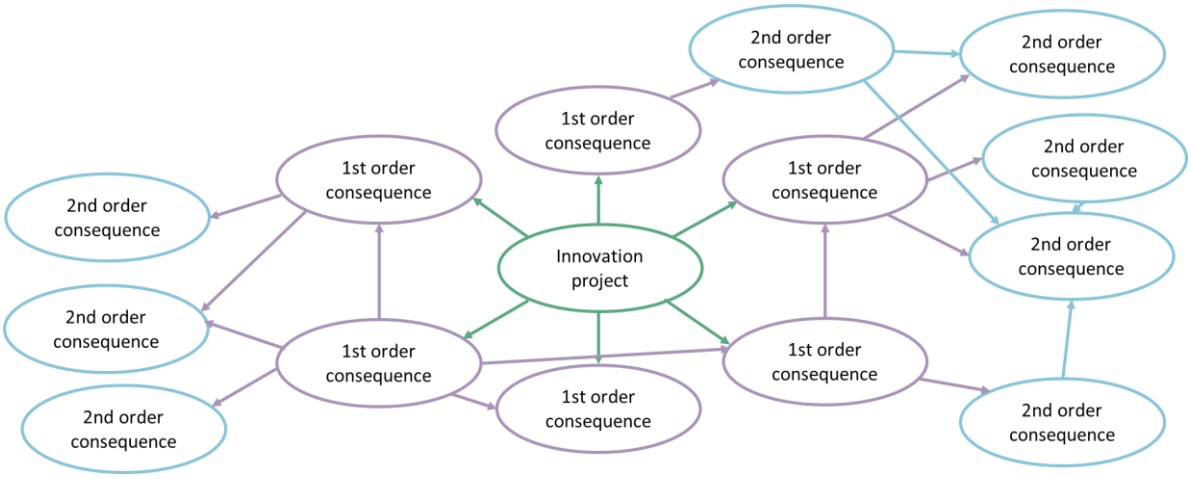
Für diesen Erhebungsteil des Projekts Z-T-G 004 (AI Privacy) wurde ein partizipatives Forschungsdesign gewählt. Aufbauend auf ersten Beobachtungen im Rahmen eines Tags der offenen Tür im April 2024, bei dem wir jungen Menschen neuere PET-Ansätze beispielhaft vorstellten und anschließend mit ihnen über ihre Sicht auf Aspekte des Vertrauens in Verschlüsselungen sprachen, gestalteten wir ein Workshopformat für Studierende, in dessen Rahmen diese in Gruppen Futures Wheels gestalteten und entlang dieser über mögliche Konsequenzen der Einführung der Corona Heatmap diskutierten. Diese offene Herangehensweise ermöglicht es, nicht nur individuelle Argumentationsmuster zu erfassen, sondern im Prinzip auch gesellschaftliche Deutungsrahmen zu identifizieren, die das Vertrauen in technische Innovationen strukturieren.

Ausgehend von einer kurzen Schilderung der Futures Wheels und des konkreteren methodischen Settings, in dem diese im Rahmen des Z-T-G 004 (AI Privacy) Projekts eingesetzt wurden, wird in diesem Berichtsabschnitt explorativ beschrieben, welche assoziativen Überlegungen Bürger:innen haben, wenn ihnen Einsatzfelder von PETs vorgestellt werden. Anschließend werden die Ergebnisse rückgebunden an das Kernthema des Projektteils und somit an die Frage, welche Faktoren Bürger:innen mit dem Einsatz von PETs im öffentlichen Interesse assoziativ in Verbindung bringen und wie sich diese Faktoren mit einem etwaigen Vertrauen in diese Technologie in Beziehung setzen lassen.

Die Methode des *Futures Wheel*, vom kanadischen Zukunftsforscher Jerome C. Glenn in den frühen 1970er Jahren erstmals vorgestellt, stellt ein heuristisches Instrument der explorativen Zukunftsforschung dar, das darauf abzielt, systematisch die direkten und indirekten Konsequenzen eines möglichen Ereignisses, einer technologischen Innovation oder einer sozialen Entwicklung zu antizipieren (Glenn, 2009). Zentral ist dabei der visuelle und iterative Aufbau eines „Rades“: Ausgehend von einem zentralen Impuls – etwa der Einführung einer neuen Verschlüsselungstechnologie – werden im ersten Schritt die unmittelbaren Primärfolgen identifiziert, die sich aus dieser Veränderung ergeben könnten. In weiteren konzentrischen Kreisen werden sodann die Sekundär- und Tertiäreffekte dieser ersten Wirkungen erfasst und verknüpft. Diese Struktur erlaubt es, sowohl intendierte als auch unbeabsichtigte Konsequenzen in ihrer wechselseitigen Verschränkung sichtbar zu machen (s. Abbildung). Methodisch eignet sich das *Futures Wheel* insbesondere für partizipative Settings, in denen heterogene Akteur:innen mit unterschiedlichen Wissensbeständen gemeinsam über Zukünfte reflektieren. Es fördert dabei nicht nur systemisches Denken, sondern macht auch normativ unterlegte Erwartungen und

Befürchtungen transparent, was es zu einem wertvollen Instrument für reflexive Technikfolgenabschätzung und Vertrauensforschung im digitalen Kontext macht.

Abbildung 4: Futures Wheel mit komplexem Relationsgeflecht



Erarbeitet wurden die Futures Wheels in zwei verschiedenen Settings der universitären Lehre. Zum einen wurde die Methode in der von Christian Dayé geleiteten Lehrveranstaltung „Futurology" im Wintersemester 2024/25 an der TU Graz eingesetzt, zum anderen im Rahmen der Research Week der europaweiten Unite!-Allianz technischer Hochschulen (https://www.unite-university.eu/, letzter Zugriff 17. Juni 2025), die von 14. bis 18. Oktober 2024 in Autrans bei Grenoble, Frankreich, stattfand. An der Lehrveranstaltung an der TU Graz nahmen 30 Studierende teil. Im Rahmen der Research Week in Grenoble besuchten 127 Teilnehmer:innen das Modul zu Transferable Skills, das das von Filip Rozborski, Paula de Pablo Sanz und Christian Dayé konzipierte und moderierte Seminar „Empowering futuristic minds for responsible research and innovation" enthielt.

In diesen Settings wurden jeweils Gruppen von vier bis acht Studierenden gebildet, die gemeinsam in ca. 60 Minuten ein Futures Wheels entwickeln sollten. Der Ausgangsimpuls, also das Ereignis, zu dem das Futures Wheel gestaltet werden sollte, war in unserem Fall ein in einem Szenario beschriebenes Innovationsprojekt, das die Leiter:innen in die Lehrveranstaltung mitbrachten. Diese Szenarien bezogen sich auf viele unterschiedliche Technologien, darunter auch Genomeditierung oder Roboterhunde. Zwei dieser Szenarien allerdings bezogen sich auf den Einsatz von PETs im öffentlichen Interesse.

Box 1 und Box 2 geben die beiden Szenarien in der Lehrsprache Englisch wieder. Sie gleichen sich, wie gesagt, darin, dass nicht näher beschriebene PETs genutzt werden, um Datensätze ohne Kenntlichmachung persönlicher Informationen zu kombinieren. Szenario #1 bezieht sich auf eine Corona Heatmap und insofern direkt auf das Referenzprojekt unserer Technologiefolgenabschätzung. Allerdings ist das Szenario nicht in Österreich, sondern in Frankreich verortet. Grund für diese Entscheidung war einerseits, dass somit

verhindert werden sollte, dass die Studierenden in Österreich zu stark an ihre eigenen Erfahrungen dachten; andererseits sollten, nachdem die Teilnehmer:innen der Unite! Research Week von verschiedenen europäischen Technischen Hochschulen kamen, auch die Szenarien quer durch Europa gestreut sein.

In Szenario #2 hingegen werden PETs genutzt, um private und nicht ordnungsgemäß den griechischen Finanzbehörden gemeldete Swimming-Pools automationsgestützt in Satellitenbildern aufzuspüren. Die Überlegung hinter diesem zweiten Szenario war, dadurch eine Vergleichsbasis zu schaffen für einen Einsatz von PETs, der nicht mit dem „heißen" Thema Corona zusammenhängt.

Die meisten Aspekte der Szenarien sind technisch bereits realisiert oder zumindest möglich – sie sind also, um Bertrand de Jouvenel (1967) zu zitieren, „futuribles." Was hingegen Fiktion ist, sind die Kollaborationen, die in den Szenarien beschrieben werden – also jene zwischen dem Gesundheitsministerium und den Mobilfunkbetreibern in Szenario #1 bzw. zwischen den Steuerbehörden und der erfundenen bulgarischen Firma in Szenario #2.

Box 1: Szenario #1: Corona24 HeatMap in France

The scenario

On June 1, 2025, in the midst of a new pandemics caused by a mutation of the COVID-19 virus, the French Ministry of Health publishes a web-based app, the Corona24 HeatMap. In collaboration with the five large mobile phone companies, the Ministry combines its own data on diagnosed diseases with the positioning data recorded by these companies to create a nation-wide map that shows where the diseased persons had stayed during the infection period. The reason, a spokesperson of the Ministry explains, is that such a heatmap can help to identify infection hotspots.

The spokesperson further states that through the use of a Privacy Enhancing Technology that follows the principles of homomorphic encryption, the Corona24 HeatMap poses no risk to privacy. All personal data remain with the state authorities. As the requests for positioning data come in an encrypted form for which only the Ministry holds the key, the mobile phone companies have no means to identify which datasets have been accessed.

Box 2: Szenario #2 – Private swimming pools in Greece

The scenario

On September 1, 2025, after another summer of wildfires and water shortage, the Greek government announces that it started to explore satellite data for private swimming pools across its territory. While this had been done a decade ago for a specified neighborhood of Athens, it was only with the help of AI that the authorities could extend this analysis of satellite images for the whole country. Two reasons motivated this step. First, the authorities wanted to be able to assess the available water capacities in case of further wildfires. Second, they combined these data with data by the Ministry of Finance to assess whether or not the owners of these private swimming pools have the correct license and had paid the required "tax on luxurious living" (article 44 of Law 4111/2013).

The artificial intelligence system that analysed the satellite data has been provided by a Bulgarian company called YAILLOWBIRD. The government's spokesperson guaranteed that due to the use of a Privacy Enhancing Technology that follows the principles of homomorphic encryption, YAILLOWBIRD has not been given access to the personal data held by the Ministry of Finance.

Die Sicht von Expert:innen

Die Meinungen der Expert:innen zusammengefasst, ergibt sich ein größtenteils klares und zumeist einheitliches Bild, wie der Einsatz von PETs und kryptografischen Verfahren für den Bereich öffentlicher Entscheidungsfindungen einzuordnen ist. Datenschutz ist ein heikles öffentliches Thema, bei dem PETs einen wertvollen Beitrag leisten können, um mögliche Problemfelder zu vermeiden.

Das öffentliche, wie auch das akademische Interesse an Datenschutz ist in den letzten Jahren stark gestiegen. Dieser Trend zeichnet sich besonders seit 2013 ab, seit der NSA Skandal öffentlich geworden ist. Auch das Inkrafttreten der DS-GVO im Jahre 2016 hat einen großen Beitrag zu dieser Zunahme und zur Bewusstseinsbildung geleistet, und zwar auch abseits von öffentlich debattierten Zwischenfällen, wie zum Beispiel des GIS-Datenlecks des Jahres 2020. Insbesondere die in der Verordnung und damit in juristischer Form festgehaltene „privacy by design" Zugangsweise zur Gestaltung technologischer Systeme trägt zu einem nachhaltigen Umdenken bei.

Datenschutz, so verglich es ein bei einer NGO tätiger Interviewpartner, ist wie Brandschutz. Zumeist braucht man ihn nicht, aber im Ernstfall ist er essenziell. In diesem Sinne muss Datenschutz auch gut vorbereitet sein, vorrausschauend gestaltet werden und alle Eventualitäten abdecken. Es gibt in diesem Fall kein Trial and Error, denn eine ausgenutzte Lücke im System hat nachhaltige Folgen und kann nicht wieder korrigiert werden. Nichteinhaltungen, so zeigen Beispiele der Vergangenheit, sind zumeist nicht in schlechten Absichten begründet, sondern in Unkenntnis.

Datenschutz bedeutet dementsprechend mehr als die bloße Einhaltung der DS-GVO, auch wenn diese natürlich maßgeblich ist. Er bedeutet Schutz vor gegenwärtigem und zukünftigem Missbrauch persönlicher Informationen. Sei dies durch Datenlecks aber auch durch potenzielle Kompetenzüberschreitungen staatlicher Akteure und Institutionen und Machtmissbrauch. Dies bedeutet neue Herausforderungen, bedingt durch die fortschreitende Digitalisierung von Identitätssystemen und des datengenerierenden Alltags, wodurch Personenbezug leicht herstellbar ist. Daten(sammlungen) sind dadurch heute nicht mehr quasi-statische Dokumente, sondern können sich stetig verändern, der gegenwärtigen Situation angeglichen oder durch zusätzliche Informationen ergänzt werden.

Von staatlicher Seite wird in Österreich der Datenschutz im Großen und Ganzen gut eingehalten, jedoch gibt es etwas Zurückhaltung bei der Anwendung neuer Technologien. Die Frage, wie PETs hier eingesetzt werden könnten, um Systeme effizienter zu gestalten und neue Möglichkeiten der Entscheidungsfindung zu schaffen, ist daher am Puls der Zeit und als hochrelevant einzustufen.

PETs und neue Möglichkeiten

Die Bewertungen des Einsatzes von PETs als geeignetes Mittel zur Schaffung neuer Grundlagen für politische Entscheidungen und Verwaltungsprozesse fielen durchwegs positiv aus. Auch im konkreten Fall des Referenzprojekts der Covid Heatmap kann die konzipierte Lösung mit dem Einsatz homomorpher Verschlüsselung als Schutzmechanismus bei der gemeinsamen Verwendung des Datensatzes ohne Preisgabe der darin enthaltenen Informationen als hochtaugliches Mittel zum Zweck verstanden werden. PETs sind „state of the art" Sicherheitsstandards, deren Funktionalität von Expert:innen nicht in Zweifel gezogen wird. Dennoch gibt es einige Punkte, denen Beachtung geschenkt werden sollte, insbesondere die Frage, wie Datenschutz durch andere Möglichkeiten erreicht werden kann, ohne Verschlüsselung notwendig zu machen und auch die Frage eines langfristig gegebenen Schutzes.

Der Einsatz von PETs eröffnet neu Potenziale, Daten, insbesondere personenbezogene, nicht nur auszuwerten, sondern auch Datensätze gemeinsam für Analysen zu verwenden, ohne dabei die Schutzprinzipien der DS-GVO zu verletzen. Die richtige Anwendung der Technologien vorausgesetzt, gelingt es dadurch, dass der Personenbezug der Daten nicht hergestellt werden kann. Diese Möglichkeit, Zusammenarbeit ohne Teilen umzusetzen, bedeutet auch, dass Daten über den Zweck hinaus verwendet werden können, für den sie ursprünglich gesammelt wurden und sich somit eventuell völlig neue Perspektiven eröffnen. Abseits der DS-GVO sollte jedoch auch mitbedacht werden, ob dadurch in andere Rechte eingegriffen wird. Im Falle der Covid Heatmap stellt sich eventuell die Frage, welche politischen Schlüsse und Konsequenzen aus dem Wissen um Infektionsherde gezogen werden und ob daraus Eingriffe in die persönliche Freiheit gerechtfertigt werden.

Im Sinne der DS-GVO erlauben es PETs zudem zwei der durch die Verordnung festgelegten Schutzziele zugleich zu verwirklichen, nämlich Datenschutz und freien Datenverkehr. Beide waren bereits in der Datenschutzrichtlinie aus 1995 festgeschrieben und können dementsprechend durchaus als Ideale auf europäischer Ebene eingeordnet werden. Diese Gleichzeitigkeit der beiden Ziele, die sich auf den ersten Blick als vermeintliche Gegenpole darstellen, ist einer der besonders attraktiven Aspekte, welche PETs in die Datenverarbeitung einbringen können. Vor allem, da dies ohne Funktionseinbußen möglich ist, stellen solche Systeme somit eine praktikable Version des Datenschutzes dar.

Ferner kann durch solche Technologien ein weiterer Spagat zwischen zwei scheinbaren Antagonisten geschafft werden. Versteht man solche Systeme wie die Covid Heatmap als Möglichkeiten, Sicherheit zu schaffen beziehungsweise zu erhöhen, dann gelingt es durch die Entfernung des Personenbezugs, den mutmaßlichen Widerspruch zwischen Privatheit und Sicherheit aufzubrechen. Durch den Einsatz geeigneter PETs können beide Faktoren bedient werden und keiner muss zu Gunsten des anderen geopfert werden.

Wenn die anvisierte Verarbeitung von Daten durch eine staatliche Institution erfolgen soll, braucht es erhöhte Voraussetzungen, insbesondere eine gesetzliche Grundlage. „Der Staat"

kann also Gesetze erlassen, die ihm die notwendigen Schritte ermöglichen, wenn die entsprechenden politischen Mehrheiten gefunden werden. Die Grenze ist solchen Beschlüssen lediglich durch Verfassungsrechte oder Unionsrecht (Grundrechte, Menschenrechte) – so wie eben durch die DS-GVO gesetzt.

Bei den angewandten Methoden, die unter den Begriff PETs fallen, wurden sowohl homomorphe Verschlüsselung als auch Differential Privacy als besonders gute Mittel zur Wahrung von Datenschutzaufgaben eingeordnet. Homomorphe Verschlüsselung wurde dabei als „Goldstandard" verstanden, da es das Auslesen von Informationen jedweder Art verunmöglicht und somit unbedingten Schutz bei der Verarbeitung bietet. Die Entwicklung der Technologie stellt einen großen Durchbruch im Datenschutz dar, da sie ein kollaboratives Element ermöglicht. Gerade aufgrund der zunehmenden Bedeutung von machine learning steigt die Wichtigkeit dieses Ansatzes, auch wenn er für große Modelle, aufgrund der Rechenintensität bisher noch unbrauchbar ist. Dabei muss auch beachtet werden, dass die Methode kontinuierlich verbessert und weiterentwickelt wird. Die Vorteile von Differential Privacy hingegen liegen in anderen Bereichen, nämlich dort wo Exaktheit nicht unbedingt erwünscht ist, wie bei Standortdaten, die generell schwer anonymisierbar sind, insbesondere auch, wenn es um Ergebnisdarstellung geht. Hier muss von Seiten der Anwendenden, wie auch bei aggregierten Daten im Allgemeinen, stets dir Frage gestellt werden, ab wann eine Diffusion ausreichend ist, um keinerlei Rückschlüsse zuzulassen und somit keinen Eingriff in die Privatsphäre darzustellen. Denn ein herstellbarer Personenbezug ist stets problematisch. Aus gesetzlicher Sicht ist es hierbei maßgeblich, ob dieser mit vertretbarem Aufwand gewonnen werden kann. Das bedeutet, zu fragen, ob es realistisch machbar ist. Im Sinne eines guten Datenschutzes, wäre es jedoch auch wünschenswert, hypothetische Szenarien mitzudenken und insbesondere zu fragen, ob Probleme, die sich zukünftig mittels neuer Methoden stellen könnten, auch vermieden werden können.

Trotz der Möglichkeit, Daten durch den Einsatz verschiedener PETs zu schützen, so mahnte einer der befragten Technikfolgenexperten, soll dies keine Art Freifahrtschein bei der Konzeptualisierung einer Technologie darstellen. Es bleibt essenziell, Datenschutz von Anfang an mitzudenken, ihn als Teil des Systems zu integrieren und ihn nicht erst nach der Entwicklung anzuflicken. Damit gemeint ist auch, sich zu überlegen, welche Daten in einem System benötigt werden, und diese auf ein quantitatives und identifizierendes Minimum zu reduzieren. Das Prinzip der Datenminimierung, das eine wesentliche Säule von gutem Datenschutz darstellt, sollte stets im Fokus stehen. Und genau zu einer solchen Reduktion von gespeicherten personenbezogenen Daten können PETs beitragen. Das Covid Heatmap Beispiel verdeutlicht dies abermals gut. Durch die Nutzung der Infektions- und Bewegungsdatensätze mittel zwischengeschalteter homomorpher Verschlüsselung, ist es nicht mehr notwendig, Daten zu kopieren, einen gemeinsamen Datensatz, der alle diese Informationen enthält, zu erstellen und diesen dann auf einem Server zu speichern. Statt Daten zu speichern, die lediglich einer singulären Auswertung dienen, werden lediglich die

spezifischen Ergebnisse für die jeweils benötigten Fälle erstellt. Die miteinander kombinierten Informationen sind nicht abermals abrufbar, sind nicht der Gefahr eines Angriffs ausgesetzt, da Systeme mit wenigen Daten in diesem Sinne unattraktiv sind, und können nicht zu einem späteren Zeitpunkt missbräuchlich für andere Zwecke verwendet werden, wenn sich etwa die politische Ausrichtung der Regierung ändert. Damit wird auch eine Zukunftsperspektive in die Datenschutzüberlegungen miteinbezogen. Ein Positivbeispiel, wie Datenschutz auf staatlicher Ebene gut gelöst ist, stellen bereichsspezifische Personenkennzeichen (bPK) dar, die es seit nunmehr über 20 Jahren in Österreich gibt und durch die eine zentrale staatliche Datensammlung, die alle Informationen an einem Ort enthält, verhindert wird. Dennoch sind alle relevanten Informationen für die jeweiligen Behörden zugänglich.

Auch in die Kerbe der zukünftigen Verwertbarkeit schlägt die Frage nach dem Bestand einer Verschlüsselung. Unter der Phrase „Save now, decrypt later" ist die Idee zusammengefasst, Daten zu speichern, selbst wenn diese verschlüsselt sind, und zu warten, bis eine Methode entwickelt wurde, um diese auszulesen. Im Falle der homomorphen Verschlüsselung ist die gegenwärtig herrschende Annahme, dass diese noch in naher Zukunft nicht aufgebrochen werden kann, doch ist dies, auch hinsichtlich der Möglichkeiten von Angriffen durch künstliche Intelligenz-Systeme, nicht gesichert. Die Fehleinschätzung, dass etwas unknackbar sei, hat sich im Laufe der Geschichte mehrfach zugetragen. Damit verbunden ist auch die Frage, wie lange Daten relevant oder nützlich verbleiben, also wie lange ihr Schutz gewährleistet sein muss. Als Lösungen, die hierbei einen dauerhaften Schutz versprechen, wurden, neben der zuvor genannten Datenminimierung, einerseits selbstlöschende Daten aufgebracht, also Daten, die nur eine begrenzte Zeit überhaupt Informationen enthalten und sich danach unwiderruflich zerstören, sowie auch synthetische Daten, die keinen direkten Personenbezug enthalten.

Vertrauen in kryptographische Methoden

Wenn Vertrauen für den Bestand eines Verhältnises grundlegend wichtig ist, bedeutet dies auch, dass dabei eine Form von Risiko eingegangen wird. Der Bruch des Vertrauens bedeutet in solchen Situationen die Verursachung von Schaden. Dementsprechend ist das Gegenteil von Vertrauen, nämlich Misstrauen, kryptografischen Verfahren in gewisser Weise inhärent, und zwar im Sinne eines Unwillens, Informationen gegenüber einer anderen Partei preiszugeben und der Idee, dies durch eine Technologie zu verhindern. Aus diesem Grunde ist es notwendig, Informationen zu verschlüsseln, wenn zum Beispiel zwei Parteien über unterschiedliche Daten verfügen, diese nicht miteinander teilen wollen, jedoch zu einem bestimmten Zweck kooperieren möchten oder gar müssen. Es ist in solchen und ähnlichen Fällen daher eine Bedingung, dass alle Beteiligten sich der Sicherheit der angewandten Schutzmethode bewusst sind, um die Kollaboration einzugehen und eingehen zu wollen. Sich auf die Technologie verlassen zu können

bedeutet ferner, dass keine Notwendigkeit des Vertrauens gegenüber Kooperationspartnern besteht und ersetzt für eine so geartete Situation den Bedarf nach einer dritten, vertrauenswürdigen Partei, zu welcher Operationen ausgelagert werden. Es ist hierbei also eine Art Vertrauen in die kryptografische Technologie, welche Vertrauen in Personen ersetzt.

Doch worauf gründet sich das Vertrauen in die Methode im Falle von PETs? Zu einem wichtigen Teil aus der Nachweisbarkeit ihrer Funktionsweise und der Überprüfbarkeit. Das neuartige Element das modernen Kryptografiemethoden, wie homomorpher Verschlüsselung, eigen ist, ist die Transparenz der Ansätze. Während Verschlüsselungen ursprünglich aufgrund der Unkenntnis über die ihnen zugrundeliegenden Systeme erfolgreich waren, hat sich zu modernen Methoden hin eine Art Paradigmenwechsel vollzogen. Gegenwärtige kryptografische Verfahren werden öffentlich, teilweise im Rahmen von Wettbewerben (NIST) entwickelt und weiterentwickelt. Es gibt kollektive Bestrebungen aus der entsprechenden Community, Schwachstellen aufzudecken und zu verbessern. So lange bis keine mehr gefunden werden. Entschlüsselung kann nunmehr lediglich durch den Besitz des notwendigen Schlüssels erfolgen. Das Wissen über die Methode ist für diese Zwecke irrelevant. Die Sicherheit der Systeme wird sozusagen bewiesen und nicht verschwiegen. Diese Situation weist ein paradox anmutendes Element auf, wonach Methoden umso stärker und dementsprechend sicherer sind, je transparenter sie sind und je mehr Leute sich an ihrer Entwicklung beteiligen. Diese gemeinschaftliche Unbrechbarkeit der Verschlüsselung gemeinsam mit dem dieser zugrundeliegenden mathematischen Nachweis ihrer Funktion fungiert auch als vertrauensbildende Maßnahme, insbesondere bei Personen, welche mit der Thematik vertraut sind.

Der Methode zu vertrauen, bedeutet auch ein temporär begrenztes Vertrauen, im Sinne, dass diese als im Moment der Anwendung sicher verstanden wird, mit dem Bewusstsein, dass sich dies zukünftig ändern kann. Deswegen trägt auch der kontinuierliche Verbesserungsdrang aus der Community vertrauensbildend bei, indem er das dynamische Element, welches digitalem Datenschutz innewohnt, erfasst.

Einer der interviewten Experten für Technikfolgenabschätzung warf in diesem Kontext die Frage auf, ob es denn überhaupt angebracht ist, von Vertrauen in eine Technologie zu sprechen. Vielmehr gelte dieses Vertrauen den Menschen, welche für die Technologie verantwortlich zeichnen. In einem Fall wie bei homomorpher Verschlüsselung ist das mathematische Verfahren, auf welchem die Methode gründet für Menschen aus dem Feld und mit Spezialwissen aus verwandten Bereichen nachvollziehbar, doch darüber hinaus wird es schwer fassbar. Dementsprechend gilt das Vertrauen in die Technologie eben den Fachleuten, welche zusichern, dass das System funktioniert, sowie der damit befassten wissenschaftlichen Gemeinschaft. Das Wissen über die Offenheit der Methode und der kollektiven Weiterentwicklung kann hier einen wertvollen Beitrag leisten.

Während PETs in diesem Sinne offen sind, verhält es sich bei künstlichen Intelligenzsystemen zumeist anders. Insbesondere gilt das bei selbstlernenden und „selbstverbessernden" beziehungsweise sich adaptierendenden Systemen. Wenn deren Entscheidungsgrundlagen nicht nachvollziehbar sind, wie durch den Begriff der black box verdeutlicht, sie dabei nicht auf menschlichem Ermessen gründen und sich die Outputs zudem stetig verändern könnten, gelten sie nicht als zuverlässig. Dementsprechend wird auch einer Institution nicht vertraut, welche ihre Entscheidung auf solchen Systemen aufbaut, die sie nicht oder nur bedingt kontrollieren kann. In der Kombination von PETs mit KI-Systemen ist es daher essenziell, dass beide Teile vertrauenswürdig gestaltet sind, denn ansonsten entbehrt die gesamte Anwendung der Vertrauensgrundlage.

Wenn nun ein System, wie die Covid Heatmap implementiert werden soll, entstehen Vertrauensfragen, die über die rein technische Seite hinausgehen. Es geht dabei um Fragen der Politik und Demokratie. Insbesondere da das Thema Covid medial omnipräsent und stark polarisierend war.

Bezüglich des Vertrauens in PETs anwendende Institutionen, wie der staatlichen Verwaltung, welche dem Prinzip der Rechtsstaatlichkeit unterliegt, stellen sich daher andere Fragen. Zum einen jene bezüglich des Vertrauens in die Fähigkeit, die Methoden richtig und sauber anzuwenden. Auf der Kehrseite schwingt hier das Element eines Inzweifelziehens notwendiger Kompetenzen mit. Auf der anderen Seite steht Vertrauen in den Willen einer korrekten Anwendung. Die negative Seite hierzu bedeutet die Annahme einer bewusst falschen Anwendung und die Unterstellung eines entsprechenden Motivs, aus dem heraus sich die missbräuchliche Verwendung von Daten begründet.

Das generelle, bereits bestehende Vertrauen in den Staat und dessen Lenkungsmechanismen sowie seine rechtsstaatliche Durchsetzungskraft bestimmt auch die Einstellung zur Anwendung von PETs im Rahmen von Entscheidungsfindungsprozessen. Insbesondere dann, wenn dieses Vertrauen beschädigt ist, wird dies auf die Anwendung der Technologie übertragen. In diesen Fällen muss an einer ganz anderen Ebene entgegengewirkt werden.

Neben der allgemeinen Einstellung zu Staat und Regierung ist es für das Vertrauen in ein PETs gestütztes Entscheidungssystem, so wie das Referenzprojekt Covid Heatmap, in dem personenbezogene Daten eine Grundlage bilden, maßgebend, wie Bürger:innen eingebunden werden. Die erste Frage, die sich in diesem Zusammenhang stellt, ist die Frage, inwiefern Zwang ausgeübt wird. Ist die Teilnahme beziehungsweise die Bereitstellung der eigenen Daten verpflichtend oder beruht sie auf Freiwilligkeit? Die Möglichkeit, zu entscheiden, was mit persönlichen Informationen geschieht ist wesentlich für eine breite Zustimmung. Dabei reicht es für diese Zwecke eine einfach zugängliche Opt-out Option zu bieten, mittels welcher Bürger:innen ihre Zustimmung verweigern können. Fehlt eine solche Möglichkeit der autonomen Entscheidung, schafft dies Unmut und Misstrauen bezüglich des Zwecks und der Absichten, die hinter dem System stehen.

Des Weiteren ist es wichtig, dass die eingesetzten Methoden nachvollziehbar und transparent kommuniziert werden. Das bedeutet nicht, dass etwa die mathematischen Grundlagen einer homomorphen Verschlüsselung allen verständlich sein müssen, aber doch, dass sie prinzipiell offen sind und die Informationen leicht zugänglich und klar sind. Diese Punkte sind umso wichtiger, je sensibler die eingesetzten Daten und der zugehörige Themenbereich sind, so wie zum Beispiel Gesundheitsdaten, und/oder wenn es Teile der Bevölkerung gibt, die den Zweck des Systems ablehnen (wie dies zum Beispiel im Fall der Covid Tracing Maßnahmen geschehen ist – dazu weiter unten etwas mehr). Fehlen Freiwilligkeit und Transparenz, schafft dies eine Basis für Ablehnung aufgrund mangelnden Vertrauens. Zudem ist ein solcher Zugang demokratiepolitisch zu hinterfragen.

Es bedarf also eines gewissen Aufklärungs- und Informationsaufwands, wenn Systeme personenbezogene Daten verarbeiten. Selbst wenn diese geschützt sind und niemals offengelegt werden. Da, wie erwähnt, methodische Fragen spezialisiertes Wissen voraussetzen und dadurch eben nicht allgemein verständlich sind, braucht es bei der Kommunikation der Maßnahmen nicht nur Personen und Institutionen, die für die Sicherheit der Systeme bürgen, sondern dazu auch Erklärungen, die breit nachvollziehbar sind. In diesem Kontext spielen auch NGOs eine bedeutende Rolle als Kontrollinstanzen und damit als vertrauensbildende Institutionen sowie auch als komplexitätsreduzierende Kommunikatoren an die Öffentlichkeit. Durch Personen und Organisationen außerhalb des staatlichen Gefüges kann somit ein entscheidender Beitrag zur Akzeptanz eines Systems beigetragen werden.

Eine weitere vertrauensbildende Maßnahme ist die Bindung staatlicher Maßnahmen an Gesetze. Einerseits die Notwendigkeit, Verfassungsgesetzen zu entsprechen, was bedeutet, dass gewisse Schutzgarantien stetig bestehen bleiben und Grundrechte nicht verletzt werden dürfen. Andererseits die Notwendigkeit, auch jede einzelne Maßnahme auf einem entsprechenden Gesetz zu begründen oder dieses zu beschließen. Das heißt, dass politische Mehrheiten dafür gefunden werden müssen. Auch die Gesetzmäßigkeit von Maßnahmen bedarf einer klaren Kommunikation, um Vertrauen zu schaffen und zu erhalten. Es muss deutlich sein, dass alles grundrechtskonform gelöst wurde und keine Ambiguitäten enthalten sind. Sobald hier etwas unklar erscheint oder Anlass für Debatten und Auslegungsfragen bietet, schafft dies Zweifel.

Das Beispiel der Covid Pandemie hat deutlich gezeigt, wie wichtig die oben genannten Faktoren und auch die entsprechende Kommunikation sind beziehungsweise, welche Konsequenzen es nach sich ziehen kann, wenn dies vernachlässigt wird. Eine Interviewpartnerin sah hier einen Mitgrund für den drastischen Anstieg der Maßnahmengegner zu jener Zeit. Die stetige anlassbezogene Gesetzgebung und Nachbesserung der Rechtsnormen schaffte Unsicherheit und bei manchen auch Bedenken über die demokratische Qualität der Prozesse.

Datenschutzrecht ist in erster Linie als Abwehrrecht gegen den Staat und dessen Eingriffe in die Privatsphäre gedacht worden. Wenn sich der Staat dies jedoch anlassbezogen nach und nach „zurechtzimmert", ist der Glaube daran erschüttert. Selbst wenn der Grund für die Verarbeitung der Daten nachvollziehbar ist. Die pandemische Situation stellte eine unvorhergesehene Herausforderung dar und zu Beginn herrschte im Wesentlichen politischer und Konsens über die Notwendigkeit der Maßnahmen. Aus der restlichen Entwicklung können entsprechende Lehren gezogen werden. Das Beispiel der elektronischen Gesundheitsakte hat auch gezeigt, wie Misstrauen gegenüber einem System geschürt wurde, gewachsen ist und wie die Opt-out Option dazu beigetragen hat, die Debatte nicht eskalieren zu lassen.

Die Einhaltung aller zuvor genannten vertrauensfördernden Maßnahmen ist für jedes System von Anfang an notwendig. Denn hat sich Misstrauen einmal breitgemacht, ist es nahezu unmöglich dieses wieder loszuwerden. Das bedeutet schließlich auch, dass das System gesellschaftlich zum Scheitern verurteilt ist. Und dabei ist es dann vollkommen irrelevant, wie es umgesetzt wurde. Ein Negativbeispiel in diesem Zusammenhang stellte die Stopp Corona App dar. Aufgrund kleinerer Probleme zu Beginn, der schlechten Kommunikation und des schlussendlich polarisierenden Themas geriet die App in Verruf und wurde dieses schlechte Image nicht mehr los. Und das, obwohl sie letztlich als System auch von unabhängigen Stellen als datenschutzfreundlich eingestuft wurde.

Schließlich zeigt das Covid Beispiel auch noch auf, dass es Fälle und Themen gibt, die hochpolarisierend sind und in denen die Gründe für die Ablehnung eines Systems nicht in diesem selbst zu suchen sind. Denn unabhängig von Umsetzung und Kommunikation kann eine Applikation auf Widerspruch stoßen, wenn es ihr Zweck ist, der abgelehnt wird. In diesem Falle handelt es sich nicht um eine Frage bloßen Vertrauens, sondern wird zu einer Frage der Akzeptanz, es geht um Fragen „warum" etwas getan wird und nicht mehr „wie" es umgesetzt wird.

Auf das Corona Heatmap Beispiel bezogen, kann davon ausgegangen werden, dass aufgrund der starken Lagerbildung während der Pandemie das System auf große Ablehnung gestoßen wäre. Egal wie sauber es ausgeführt worden wäre. Gleichzeitig hätte es aber wohl auch viel Zustimmung gegeben. Der Grund ist in der Einstellung zum Pandemiemanagement generell zu finden und ist vom System unabhängig. Es sind politische Fragen, die hier im Vordergrund stehen. Konkret in diesem Fall die Frage, ob ein staatliches Covid-Management wichtiger ist als Privatheit. Jedoch ist auch im Falle von bejahenden Stimmen zu beachten, dass die Pandemie eine Ausnahmesituation mit Notfallgesetzgebung war und als solche verstanden werden muss. Es bedeutet nicht, dass das Sammeln von Bewegungsdaten, selbst in aggregierter Form, akzeptiert wird, wenn es als dauerhafte Option gedacht wird.

Ein Interviewpartner führte an, dass so ein System auch mittels des Missbrauchs von Bewegungsdaten zur Nachverfolgung von Demonstrationszügen verwendet werden

könnte und dabei den normativen Vorgaben der DS-GVO entsprechen würde[5]. Diese Perspektive wurde gewählt, um zu veranschaulichen, dass auch ein geschütztes System nicht per se ethisch unbedenklich ist und die Frage stellt sich, ob diese Problematik nicht auch für die Heatmap gilt. Sowohl bei der Heatmap, als auch beim Demonstrationsmonitoring sind es Bewegungsdaten. Bei beiden stellt sich im Vorhinein die Frage, was das öffentliche Interesse ist, das damit bedient wird. Und was wiegt höher – Sicherheit oder individuelle Freiheit? Die reine technologische Möglichkeit sowie juristische Erlaubtheit machen den Einsatz eines Systems noch nicht unproblematisch. Insbesondere dann nicht, wenn dadurch Eingriffe in Freiheit oder Demokratie ermöglicht werden.

Es ist eine feine Grenze, wenn es darum geht, zu bewerten, ob PETs in diesem Sinne neue Möglichkeiten eröffnen oder Maßnahmen zur Umgehung gesetzlicher Vorgaben darstellen. Ihr Einsatz eröffnet auch Gefahren der Überregulierung, Kontrolle und Restriktion. Im Falle der Covid Heatmap wäre die Implementierung als informatives System zum Schutz der Bevölkerung eine völlig andere Maßnahme als ein Überwachungssystem zur Durchsetzung von Verboten. Daher bedeutet die Eliminierung des Personenbezugs von Daten nicht automatisch, dass ein Vorgehen unproblematisch ist.

All das Vorhergenannte zeigt schließlich, dass der Erfolg solcher Systeme zur öffentlichen Entscheidungsfindung ein „Alignment“ aus faktischer Notwendigkeit, technischer Umsetzung und korrekter Wahrnehmung erfordert.

[5] Inwiefern dies mit anderen Gesetzen in Konflikt stünde, sei für das bewusst plakative Beispiel dahingestellt.

Die Sicht von Bürger:innen

Die Futures Wheels

Die nachfolgenden Abbildungen stellen die Futures Wheels dar, die in den genannten Settings von europäischen Studierenden unter Anleitung erarbeitet wurden.

Abbildung 5: Szenario #1, Futures Wheel 1

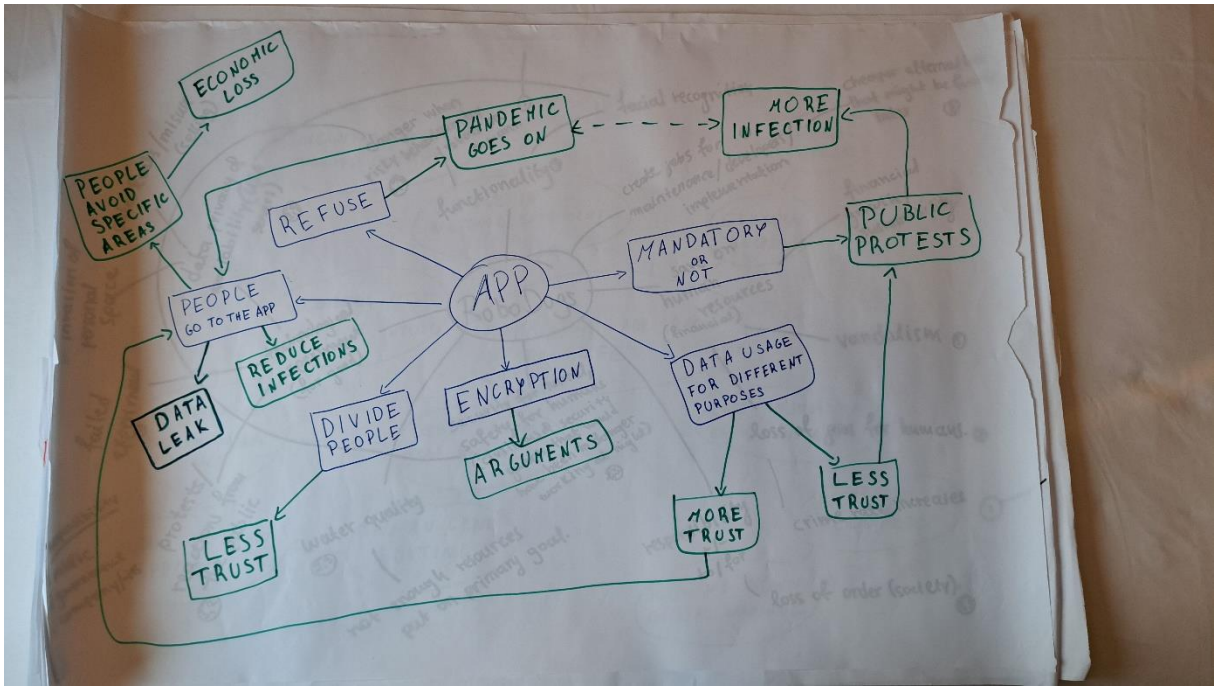


Abbildung 6: Szenario #1, Futures Wheel 2

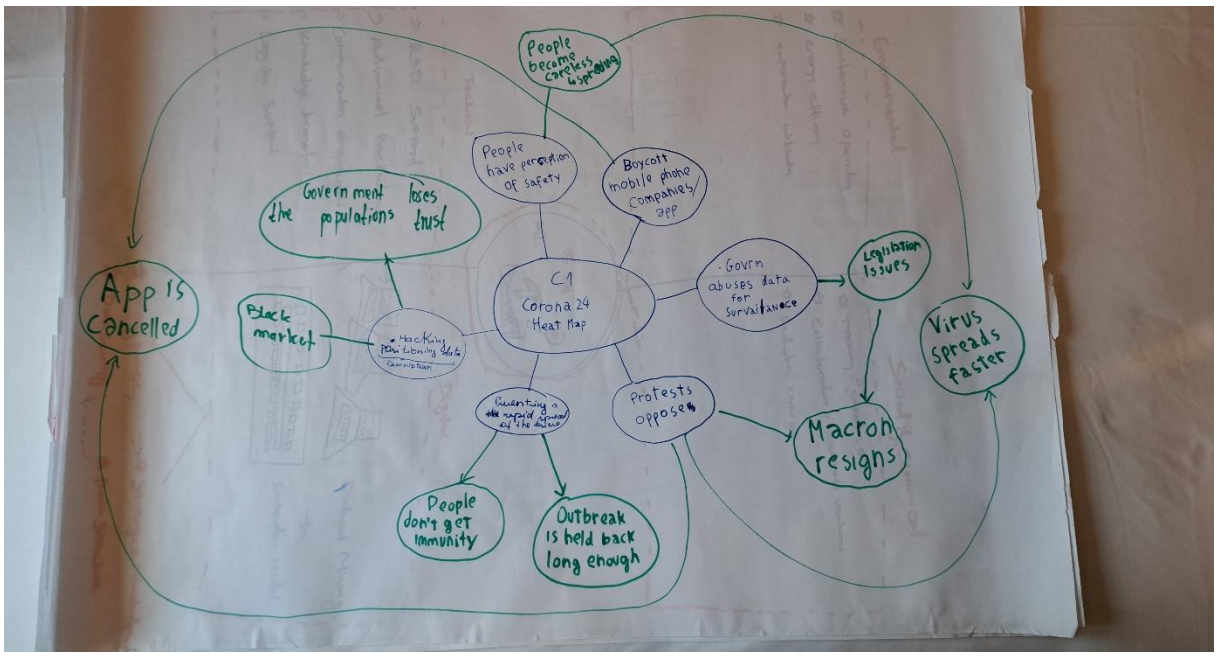


Abbildung 7: Szenario #2, Futures Wheel 1

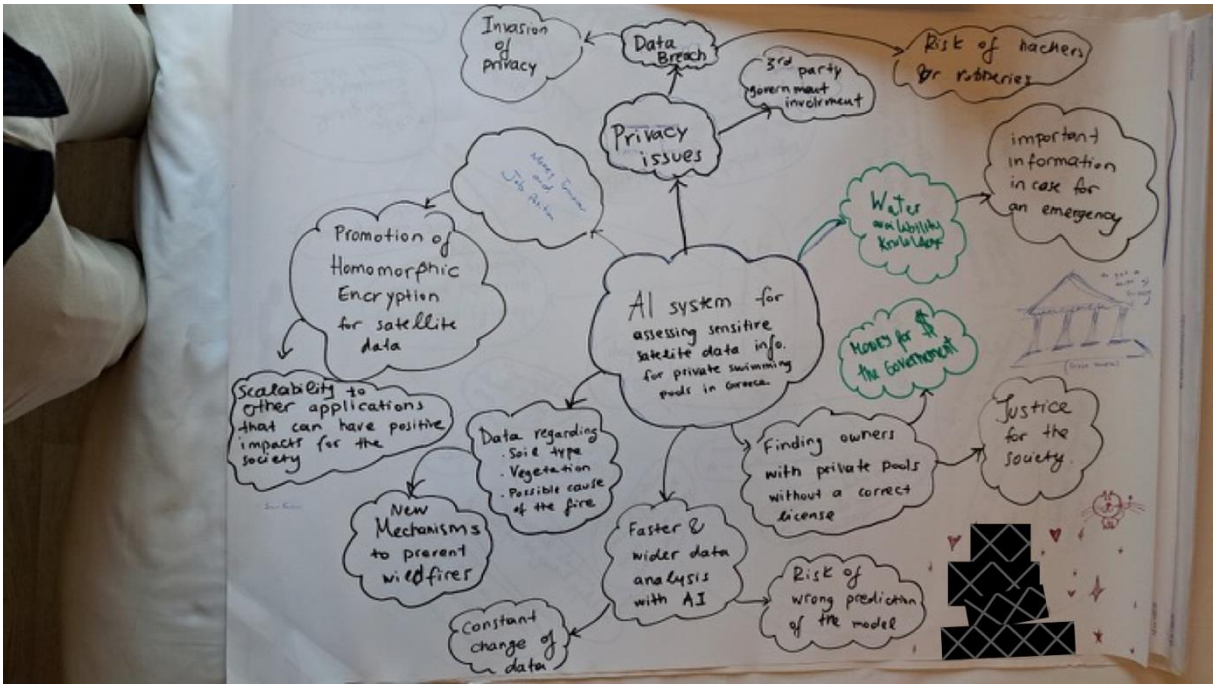


Abbildung 8: Szenario #2, Futures Wheel 2

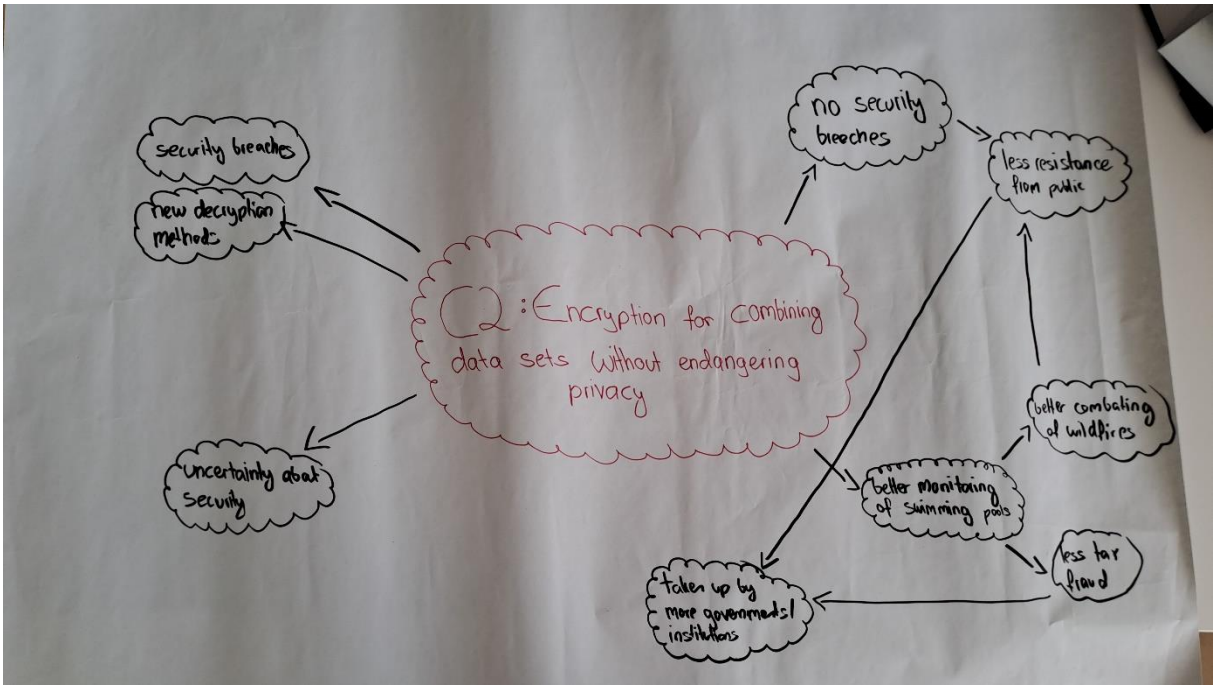
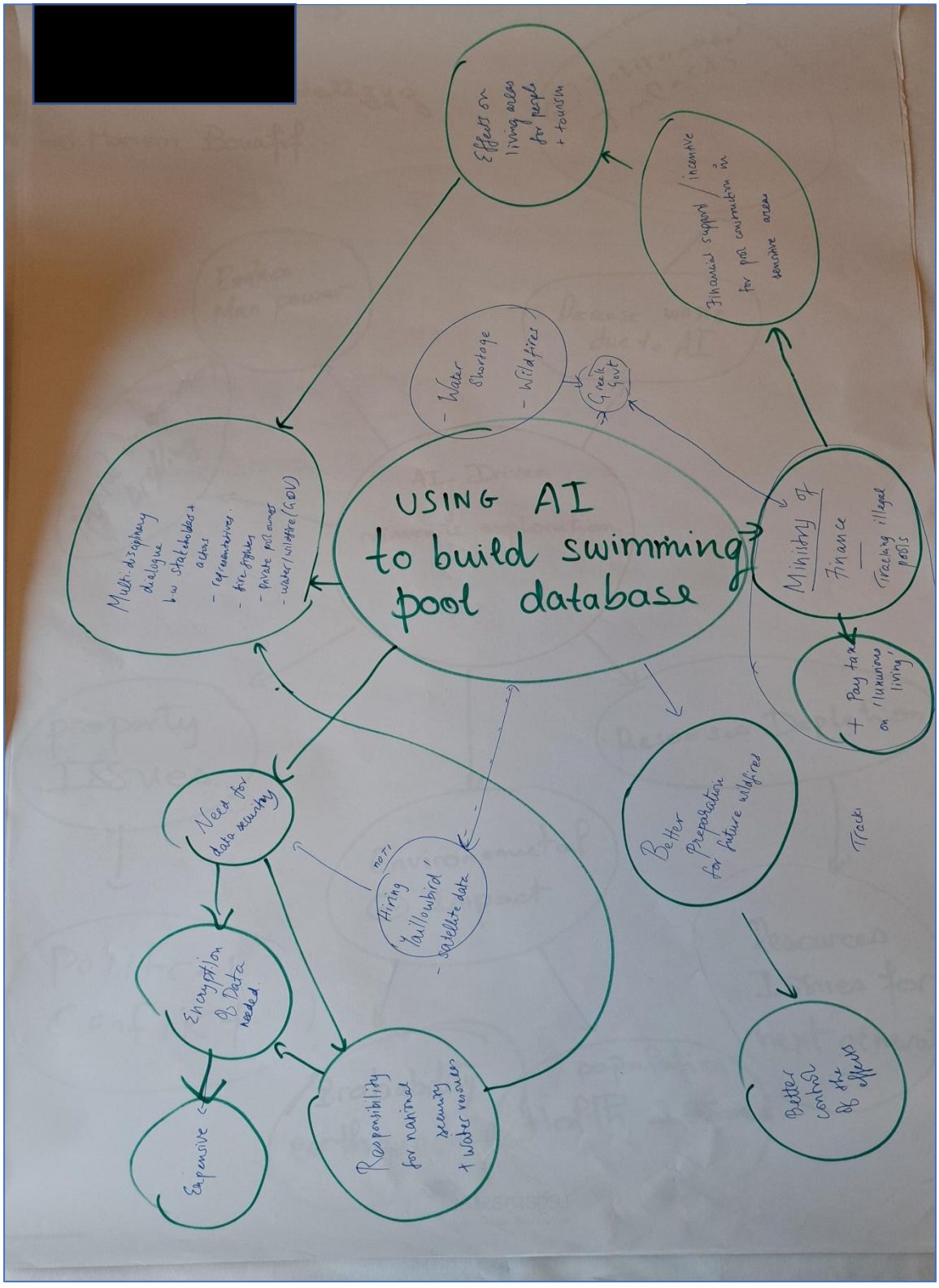


Abbildung 9: Szenario #2, Futures Wheel 3



Analyse der Futures Wheels

Was verraten uns nun die fünf Futures Wheels über die Sichtweise europäischer Bürger:innen auf PETs?

Szenario #1: Corona24 HeatMap in Frankreich

Zur Erinnerung: Das Szenario beschrieb die Veröffentlichung der Corona24 HeatMap durch das französische Gesundheitsministerium in Zusammenarbeit mit Mobilfunkanbietern. Ziel ist es, Infektionsherde durch die Verknüpfung staatlicher Diagnosedaten mit Bewegungsdaten aus den Mobilfunknetzen zu visualisieren.

Trotz der Anwendung fortschrittlicher Verschlüsselungstechnologie bleiben tiefgreifende Fragen in Bezug auf Vertrauen, Kontrolle, Datenschutz und gesellschaftlicher Kohäsion offen. Die beiden vorliegenden Futures Wheels erfassen mögliche Entwicklungspfade auf strukturierte Weise und offenbaren sowohl divergente als auch sich überschneidende Dynamiken.

Das erste Wheel (Abbildung 5) ist stark auf die unmittelbaren gesellschaftlichen Reaktionen auf die App fokussiert. Der zentrale Begriff „App" wird in Relation gesetzt zu Fragen der Freiwilligkeit, Datennutzung, Verschlüsselung und gesellschaftlicher Spaltung. Die Verzweigungen zeigen eine differenzierte Auseinandersetzung mit Vertrauen, Datenschutz, politischer Legitimität und potenziellen Auswirkungen wie Protesten oder ökonomischem Verlust.

Das zweite Wheel (Abbildung 6) hat einen breiteren Fokus, der über unmittelbare App-bezogene Debatten hinausgeht und systemische, politische Risiken inklusive etlicher Eskalationen aufzeigt. Der zentrale Begriff „Corona24 Heat Map" ist Ausgangspunkt für Szenarien wie Regulierungsprobleme, Regierungsvertrauensverlust, Schwarzmarktbildung und sogar den Rücktritt des Präsidenten. Die semantische Dichte ist hier geringer, aber die narrative Reichweite deutlich größer.

Beide Wheels thematisieren das Vertrauen der Bevölkerung in Staat und Technologie als zentrale Faktoren. Im ersten Wheel wird gezeigt, wie Vertrauen durch Datenlecks, Zwangsnutzung oder fragwürdige Datennutzung sinken kann – was wiederum zu öffentlichen Protesten führen kann. Im zweiten Wheel führt der Missbrauch von Daten durch den Staat direkt zum Vertrauensverlust, was im weiteren Verlauf den Rücktritt politischer Verantwortungsträger begünstigt. Vertrauen wirkt in beiden Wheels als systemisches Kippelement. Das erste Wheel fokussiert stärker auf kurzfristige Reaktionen – öffentliche Debatten, Vertrauensverlust, Verhaltensänderungen in Mobilität. Das zweite Wheel hingegen integriert mittel- bis langfristige Konsequenzen, darunter strukturelle Probleme wie die mangelnde Immunitätsbildung in der Bevölkerung oder legislative Folgeprobleme.

Ein zentrales Unterscheidungsmerkmal ist der Fokus auf die (Un-)Freiwilligkeit der Nutzung. Das erste Wheel thematisiert explizit die Debatte um eine Pflichtnutzung („mandatory or not"), was zu gesellschaftlicher Spaltung führen kann („divide people"). Das zweite Wheel hingegen thematisiert eher implizit die Akzeptanz und Protestbereitschaft ohne formale Differenzierung zwischen freiwilliger und verpflichtender Nutzung.

Auch hinsichtlich ökonomischer und sozialer Nebeneffekte unterscheiden sich die beiden Futures Wheels. Das erste Wheel verweist auf lokale ökonomische Verluste infolge von Gebietsscheu („People avoid specific areas") und damit verbundene gesellschaftliche Konsequenzen wie weniger Vertrauen und stärkere Polarisierung. Das zweite Wheel denkt weiter: Schwarzmarktbildung, Boykottaktionen und sogar gesellschaftliche Apathie durch falsche Sicherheitserwartungen („People become careless") werden als systemische Nebenwirkungen betrachtet.

Das zweite Wheel zeichnet ein deutlich dramatischeres Bild möglicher politischer Konsequenzen. Die Eskalation über „Legislation Issues" und „Govrn abuses data for surveillance" bis zu „Macron resigns" zeigt eine Kaskade von institutionellen Vertrauenskrisen, die bis zur Destabilisierung staatlicher Führung reichen. Das erste Wheel bleibt in dieser Hinsicht technokratischer – es zeigt, wie Vertrauen durch technische Missstände erodiert, ohne jedoch explizit den Zusammenbruch staatlicher Autorität zu antizipieren.

Zusammen zeigen beide Futures Wheels ein komplementäres Zukunftsbild:

- Das erste Wheel legt den Finger auf die soziale Sensibilität gegenüber Überwachungstechnologie, die durch scheinbar „sichere" Verschlüsselung nicht vollständig aufgelöst wird.
- Das zweite Wheel offenbart, dass technologische Lösungen in pandemischen Krisen nur dann nachhaltig funktionieren können, wenn gesellschaftliche Akzeptanz, transparente Kommunikation und politische Integrität gewährleistet sind.

Insgesamt sprechen beide Wheels eine deutliche Warnung aus: Ohne reflektierte politische Rahmung und partizipative Einbindung droht selbst eine gut gemeinte Gesundheitsinnovation wie die Corona24 HeatMap zum Katalysator gesellschaftlicher und politischer Krisen zu werden.

Szenario #2: Private Swimming-Pools in Griechenland

Zur Erinnerung: Das Szenario beschreibt eine Kollaboration zwischen den griechischen Steuerbehörden und einer bulgarischen Firma, die sich auf die KI-gestützte Analyse von Satellitenbildern spezialisiert hat. Durch den Abgleich von auf den Bildern gefundenen

Swimming-Pools mit den Daten der Steuerbehören über jene privaten Pools, für die ordnungsgemäß eine Luxussteuer geleistet wird, können einerseits Fälle von Steuerhinterziehung geahndet werden. Andererseits verspricht das auch einen besseren Überblick über Wasserressourcen bei Waldbränden.

Die drei Futures Wheels zu den Swimming-Pools sind Ausdruck eines kollaborativen Zukunftsprozesses, in dem urbane Schwimmbäder als Ansatzpunkt für multiple Transformationspotenziale betrachtet werden. In ihrer Gesamtheit spiegeln die Darstellungen eine zunehmende Komplexität, Tiefe und Integration gesellschaftlicher, ökologischer und kultureller Dimensionen wider. Diese Schnittmenge legt nahe, dass die Gruppen – unabhängig voneinander – jene gesellschaftlichen Grundprobleme identifizierten, die im Szenario angelegt waren: Polarisierung, Klimawandel, Digitalisierung.

Alle Wheels zu diesem Szenario sind geprägt von einer funktionalen, output-orientierten und letztlich optimistischen Perspektive. Der Fokus liegt auf einer umfassenden Nutzung der vorhandenen Ressourcen zum Wohle der Allgemeinheit – das erscheint nicht problematisch, ebensowenig wie der Einsatz von PETs bzw. der Kollaboration mit einer nicht-griechischen Firma.

Eine Ausnahme bietet das zweite Wheel (Abbildung 8), das – etwas entgegen der Methodik des Futures Wheels - die Optionen Sicherheitslücke und keine Sicherheitslücke links und rechts des Ausgangsevents platziert. Allerdings bleiben die Folgen der Sicherheitslücke hier auch wenig ausdifferenziert: Man werde neue Verschlüsselungsmethoden entwickeln, und es könne sich Unsicherheit über die Datensicherheit breitmachen.

Die zwei anderen Wheels in Abbildung 7und Abbildung 9 sind vergleichsweise komplexer: Die radialen Achsen sind mit qualitativ anspruchsvolleren Konzepten angereichert – Reaktionskapazität im Brandfall, Steuereinnahmen, gesellschaftliche Gerechtigkeit sind untereinander verknüpft. Die Effekte sind nicht nur funktional oder utilitaristisch, sondern beinhalten auch normative Perspektiven. Die Systemtiefe in diesen beiden Wheels ist deutlich höher: anstelle linearer Kausalitäten treten komplexe Wirkungsgefüge zutage.

Zusammenfassung

Abschließend ist zunächst zu wiederholen, dass sich die Szenarien über den Einsatz von PETs, die den Ausgangspunkt der Futures Wheels bildeten, absichtlich unterschieden. Beide Szenarien thematisieren den Einsatz von PETs seitens des Staats mit einer möglichen Orientierung am Allgemeinwohl. Sie tun dies jedoch mit sehr unterschiedlichen Ausgangspunkten und normativen Vorannahmen:

- Szenario #1 geht von einer krisenhaften Ausnahmesituation (Pandemie) aus, in der technologisches Handeln unter einem hohen Legitimationsdruck steht. Datenschutz, Vertrauen und gesellschaftliche Spaltung stehen im Zentrum der Analyse.

- Szenario #2 basiert auf einem technokratischen Normalmodus: Digitalisierung wird funktional zur Effizienzsteigerung eingesetzt (Steuerkontrolle, Brandvorsorge), ohne dass größere gesellschaftliche Konflikte antizipiert oder durchgespielt werden.

Diese Logiken aus den Szenarien bildeten sich auch in den in den Futures Wheels entfalteten Zukunftsvorstellungen ab.

Im den Futures Wheels zu Szenario #1 stehen PETs explizit im Zentrum der Debatte – nicht in technischer Tiefe, aber als politisch aufgeladene Schutzmaßnahme gegen Überwachung und Machtmissbrauch. Die Wheels zeigen:

- Ambivalente Wirksamkeit von PETs: Trotz starker Verschlüsselung bleibt das gesellschaftliche Vertrauen fragil. PETs können Misstrauen nicht automatisch kompensieren.

- Politische und soziale Kippdynamiken: Datenschutzverletzungen führen nicht nur zu Misstrauen, sondern können systemische Instabilität erzeugen (Proteste, Legitimationsverlust, Rücktritt des Präsidenten).

- PETs sind hier Teil eines hochgradig konflikthaften gesellschaftlichen Aushandlungsprozesses – sie lösen keine Probleme, sondern rahmen sie anders.

In den Futures Wheels zu Szenario #2 erscheinen PETs nur am Rande und weitgehend voraussetzungslos:

- In einem Wheel wird die Möglichkeit einer Sicherheitslücke angedeutet, aber deren Folgen bleiben abstrakt und folgenlos: „Es werden neue Verschlüsselungsverfahren entwickelt."

- Die Tatsache, dass eine ausländische Firma mit sensiblen Bilddaten arbeitet, wird nicht als soziales oder politisches Problem diskutiert.

- PETs fungieren als stilles Hintergrundversprechen technologischer Reibungslosigkeit – ihre gesellschaftliche Aushandlung wird ausgeblendet.

Während PETs im Falle der Corona24 HeatMap als umstrittene, symbolisch aufgeladene Governance-Technologie auftreten, erscheinen sie im Falle der Suche nach illegalen Swimming-Pools entpolitisiert und rein ihrer Funktionalität wegen relevant. Die Frage ihrer gesellschaftlichen Akzeptabilität stellt sich in Frankreich offen und kritisch, in Griechenland bleibt sie latenter Konsens.

In beiden Szenarien stellt sich die grundlegende Frage, ob ein Eingriff in Privatheit gebilligt wird oder nicht. Insofern rückt die Frage nach dem Zweck in den Vordergrund, zu dem PETs eingesetzt werden, und zwar konkreter nach dessen antizipierter Konfliktgeladenheit:

Aspekt	***Szenario #1***	***Szenario #2***
Konfliktdichte	Hoch: Proteste, Polarisierung, politische Krise	Niedrig: Akzeptanz, Funktionalität
Vertrauensfrage	Zentral, kritisch, instabil	Implizit vorhanden, kaum problematisiert
Regierungsrolle	Ambivalent, potenziell autoritär oder technokratisch	Technisch-rational, kooperativ
Bürgerperspektive	Als Widerstandsträger und Vertrauensgeber	Als passive Nutznießer einer besseren Infrastruktur
Gesellschaftliche Tiefe	Explizit: Soziale Spaltung, Legitimitätsverlust	Implizit: Ressourcengerechtigkeit, Umweltbewusstsein

Zusammenfassend kann gesagt werden, dass die Futures Wheels wertvolle Beiträge zur partizipativen Technologiefolgenabschätzung leisten. Sie tun das, indem sie antizipativ jenen Diskursraum aufspannen, in dem sich mögliche künftige Anwendungen von PETs durch staatliche Akteure aller Voraussicht nach bewegen werden. Hervorzuheben ist hier, nach einer eingehenden Analyse der Unterschiede, abschließend auch der Umstand, dass das Potential von PETs zur Beförderung gesamtgesellschaftlicher Ziele – Gesundheit, Katastrophenschutz, Gerechtigkeit – von den Teilnehmer:innen an unseren Workshops annähernd ohne Einwände geteilt wurde.

Schluss

Im Zuge des Projekts Privacy Enhancing AI: Teilen ohne Weitergeben (Z-T-G 04) wollten wir die Möglichkeiten erfassen, welche sich durch neue kryptografische Verfahren für politische Entscheidungsfindungen ergeben und welche Rolle Vertrauen hierbei spielt beziehungsweise wie dieses beeinflusst wird. Konkret lauteten die beiden Forschungsfragen, welche unsere Untersuchung anleiteten:

Forschungsfrage 1: Wie schätzen relevante Stakeholder die Möglichkeiten ein, durch die Kombination von Kryptographie und KI eine breitere Datengrundlage für Entscheidungsprozesse bereitzustellen?

Forschungsfrage 2: Welche Faktoren fördern das Vertrauen der Öffentlichkeit in solche kryptographischen KI-Technologien und welche sind diesem abträglich?

Zur Beantwortung dieser Fragen haben wir zwei zentrale Zugänge gewählt. Für die Erfassung der Perspektiven relevanter Expert:innen haben wir eine Reihe an leitfadengestützten Interviews durchgeführt. Zudem haben wir die Perspektive von interessierten Bürger:innen durch die Ausarbeitung von Futures Wheels im Rahmen eines Workshops und einer universitären Lehrveranstaltung eingeholt. Als Referenzprojekt diente uns die sogenannte Covid Heatmap, die es durch ihren Bezug zu einem hochpolarisierten Thema ermöglichte, die uns interessierenden Fragen in einem speziellen Kontext zu verorten, dessen Problematik allen bekannt ist.

Zentrale Ergebnisse

Datenschutz als Mittel zur Wahrung von Privatheit ist in den letzten Jahren zu einem deutlich bedeutenderen Thema geworden, das insbesondere in Kombination mit anderen gesellschaftlich debattierten Themen oder sensiblen Daten erhöhte Beachtung findet. Dies ist unter anderem auch durch die zunehmende Digitalisierung des Alltags zu erklären, in dem immer mehr persönliche Daten anfallen, die ausgewertet, aber auch missbräuchlich verwendet werden können. Dies meint einerseits Auswertungen zu Zwecken, die von den Datensubjekten unerwünscht sind durch die ursprünglich erhebende Institution, als auch die unbefugte Erlangung der Daten durch Dritte.

Hinsichtlich der Möglichkeit, durch den Einsatz von PETs neue Datengrundlagen zu schaffen und dabei datenschutzfreundlich zu verfahren, wurde von Seiten der Expert:innen eine positive Bewertung abgegeben. Die von uns betrachteten Methoden der homomorphen Verschlüsselung und der Differential Privacy bieten, korrekt eingesetzt, gegenwärtig die Möglichkeit, den Personenbezug eines Datensatzes für Dritte unkenntlich zu machen. Sie sind dadurch ein geeignetes Mittel, um neue Datengrundlagen für Entscheidungen zu

schaffen und neue Perspektiven zu eröffnen, ohne datenschutzrechtlich bedeutsame Gefährdungen zu produzieren.

Im Sinne der DS-GVO ermöglicht es der Einsatz von Technologien wie homomorpher Verschlüsselung, den zwei großen Schutzzielen der Verordnung, Datenschutz und freier Datenverkehr, zugleich zu entsprechen.

Dennoch sollte immer abgewogen werden, ob der Einsatz von PETs notwendig ist, oder ob die betreffenden Ziele auch auf anderem, von Anfang an datensparsamerem Wege, erreicht werden können. In Fällen in denen Datensätze gemeinsam verwendet werden, kann ein solcher Weg auch durch den Einsatz von PETs erreicht werden, indem die Notwendigkeit der Erstellung eines neuen, kombinierten Datensatzes entfällt. Durch diese Nichtspeicherung schaffen sie auch Schutz vor zukünftigem Missbrauch.

Vertrauen spielt eine große Rolle bei der Entwicklung von neuen datenbasierten Systemen für die Entscheidungsfindung, und zwar umso mehr, je sensibler die verwendeten Daten sind, denn dies erhöht das mit dem System verbundene Risiko.

Vertrauen äußert sich in diesem Kontext in zwei Formen: Als Vertrauen in die Methode und als Vertrauen in die Institution, welche die Methode anwendet. In beiden Fällen kann ein Fehlen von Vertrauen die Implementierung der entsprechenden Technologie nachhaltig verhindern.

Aus der Perspektive der umsetzenden Institutionen und Personen ist Vertrauen in und eine Kenntnis über die Methode eine notwendige Bedingung, denn durch ihren Einsatz wird Datenschutz und somit die Wahrung des Rechts auf Privatheit garantiert. Zugleich entfällt die Notwendigkeit, den anderen Partner:innen bei einer Kollaboration vertrauen zu müssen, da sich Sicherheit aus der eingesetzten Methode ergibt.

Das Vertrauen in die Methode wiederum gründet bei Datenverarbeitern wie -subjekten im Falle von PETs auf deren Transparenz, Nachvollziehbarkeit und Beweisbarkeit. Dies wirkt umso mehr, da es, wie im Falle der homomorphen Verschlüsselung, kollektive Bestrebungen gibt, die Technologien zu prüfen und zu verbessern und dadurch eine Gemeinschaft von Expert:innen dafür sorgt, dass diese als sicher einzustufen sind. Dementsprechend sorgt das Wissen über diese stetige fachspezifische Befassung mit der Sicherheit von PETs auch dafür, dass ein Vertrauen in die Technologie über das Vertrauen in das Urteil der Fachleute abgeleitet wird. Zu einem positiven Standpunkt gegenüber einem solchen System, das von staatlicher Seite eingesetzt wird, können bei Bürger:innen zudem Einschätzungen nichtstaatlicher Institutionen, wie Datenschutz NGOs, beitragen, ohne dass sich jede:r einzelne mit der Technologie selbst vertraut machen muss.

Vertrauen in die Institution, im Falle unseres Referenzprojektes in den Staat und die staatliche Verwaltung, ist wiederum unabhängig von der Technologie, kann diese jedoch

hemmen. Abseits des generellen staatsbezogenen Vertrauens, kann das Vertrauen in die Umsetzung sehr stark durch den demokratischen Zugang beeinflusst werden. Besonders relevant sind hier Fragen nach Transparenz und Freiwilligkeit. Deren Ermangelung schafft das Gefühl von Kontrolle und trägt zur Ausbildung von Mistrauen bei.

Ebenso vertrauensbildend ist die klare Befolgung rechtsstaatlicher Prinzipien und die klare Kommunikation darüber. Wenn der Eindruck entsteht, dass versucht wird, bestehende Normen zu umgehen oder sukzessive auszuhöhlen, wirkt dies stark vertrauenshemmend.

Hat sich einmal Misstrauen bezüglich eines Systems eingeschlichen, ist es de facto egal, wie gut es umgesetzt ist. Es wird höchstwahrscheinlich keinerlei Akzeptanz mehr erfahren.

Schließlich kann auch Vertrauen in die Methode und in die Institution die Hemmnisse nicht kompensieren, die schlagend werden, wenn der Zweck des Systems abgelehnt wird. Im Falle der Covid Heatmap könnte es so erklärt werden, dass selbst wenn Bürger:innen davon ausgehen, dass bei der Umsetzung datenschutzrechtlich alles unbedenklich umgesetzt wurde, sie die Heatmap dennoch ablehnen, weil sie solche Maßnahmen zur Pandemieregulierung ablehnen.

Die Ergebnisse der Futures Wheels schließen hier an, indem sie aufzeigen, dass es große Unterschiede, auch im Vertrauen zu den Technologien, zwischen den zwei gewählten Szenarien gab.

Das erste Beispiel, die Covid Heatmap, das den Teilnehmer:innen nahe war und ein Szenario zeichnete, welches sie selbst unmittelbar betreffen könnte, zeigte eine große Skepsis gegenüber der Technologie. Hier herrschte schon a priori Misstrauen in Bezug auf die Thematik, welches nicht durch die Technologie kompensiert werden kann.

Im zweiten Szenario, bei dem die Interessen von Poolbesitzern gegen öffentliches Interesse ausgespielt wurden, wurden selbst im Hinblick auf den Datenschutz problematische Faktoren in Kauf genommen.

Schließlich soll festgehalten werden, dass sowohl die befragten Expert:innen als auch die Bürger:innen höher gebildete gesellschaftliche Segmente abbilden. In diesem Sinne könnte es ein interessanter Ansatzpunkt für nachfolgende Untersuchungen sein, eine breiter aufgestellte empirische Erhebung zu diesem Thema durchzuführen und diese Ergebnisse zu prüfen.

Erkenntnisse

Digitalisierung bedeutet neue Herausforderungen für die Wahrung des Grundrechts auf Privatheit. Als Mittel zu diesem Zweck spielt Datenschutz eine entscheidende Rolle. Insbesondere bei sensiblen Daten gibt es in der Bevölkerung das Bedürfnis nach besonderen Vorkehrungen. PETs können solche Mechanismen darstellen, mit denen neue Erkenntnisse gewonnen werden können, während zugleich der Datenschutz gewahrt wird.

Von politischer Seite muss dabei jedoch im Sinne demokratischer Prinzipien klar auf Transparenz und Freiwilligkeit gesetzt werden, um nicht Vorwürfen von Kontrolle und Überregulierung ausgesetzt zu sein und in Konsequenz Vertrauensverluste zu erfahren.

Zudem ist es nicht nur die Technologie, die implementiert wird, welche die Perspektive von Bürger:innen formt. Vielmehr ist deren ordentliche Umsetzung nur zweitrangig, wenn auch nicht unwesentlich. Im Vordergrund steht stets der Zweck, zu welchem ein System eingeführt wird. Ist dieser belastet, kann dies selbst das datenschutzfreundlichste System nicht ausgleichen. Vielmehr noch, werden die Zweifel, die dem Gesamtthema eigen sind auch auf die Technologie übertragen.

Literatur

Alkaeed, M., Qayyum, A., & Qadir, J. (2024). Privacy preservation in Artificial Intelligence and Extended Reality (AI-XR) metaverses: A survey. *Journal of Network and Computer Applications*, *231*, 103989. https://doi.org/10.1016/j.jnca.2024.103989

Bampoulidis, A., Bruni, A., Helminger, L., Kales, D., Rechberger, C., & Walch, R. (2022). Privately Connecting Mobility to Infectious Diseases via Applied Cryptography. *Proceedings on Privacy Enhancing Technologies*, *2022*(4), 768–788. https://doi.org/10.56553/popets-2022-0132

Baumgartner, M., Veeranki, S. P. K., Hayn, D., & Schreier, G. (2023). Introduction and Comparison of Novel Decentral Learning Schemes with Multiple Data Pools for Privacy-Preserving ECG Classification. *Journal of Healthcare Informatics Research*, *7*(3), 291–312. https://doi.org/10.1007/s41666-023-00142-5

Beck, U. (1986). *Risikogesellschaft: Auf dem Weg in eine andere Moderne*. Suhrkamp.

Bierbauer, D., & Helminger, L. (2023). Offenlegung von Daten unter Wahrung der Privatsphäre mittels SMPC (Secure Multiparty Computation). *Austrian Law Journal*, *10*, 1 135. https://doi.org/10.25364/01.10:2023.1.1

COVID Heatmap – COVID Heatmap. (n.d.). Retrieved 20 June 2025, from https://covid-heatmap.isec.tugraz.at/index.html

de Jouvenel, B. (1967). *The Art of Conjecture*. Basic Books.

European Data Protection Supervisor. (2023). *TechDispatch: Explainable artificial intelligence. #2/2023*. Publications Office. https://data.europa.eu/doi/10.2804/802043

Gefen, Karahanna, & Straub. (2003). Trust and TAM in Online Shopping: An Integrated Model. *MIS Quarterly*, *27*(1), 51. https://doi.org/10.2307/30036519

Generative AI vs. predictive AI: What's the difference? | IBM. (2024, August 9). https://www.ibm.com/think/topics/generative-ai-vs-predictive-ai-whats-the-difference

Giddens, A. (1996). *Die Konsequenzen der Moderne* (J. Schulte, Trans.). Suhrkamp.

Jovic, A., Jap, D., Papachristodoulou, L., & Heuser, A. (2022). Traditional Machine Learning Methods for Side-Channel Analysis. In L. Batina, T. Bäck, I. Buhan, & S. Picek (Eds.), *Security and Artificial Intelligence* (Vol. 13049, pp. 25–47). Springer International Publishing. https://doi.org/10.1007/978-3-030-98795-4_2

Krček, M., Li, H., Paguada, S., Rioja, U., Wu, L., Perin, G., & Chmielewski, Ł. (2022). Deep Learning on Side-Channel Analysis. In L. Batina, T. Bäck, I. Buhan, & S. Picek (Eds.), *Security and Artificial Intelligence* (Vol. 13049, pp. 48–71). Springer International Publishing. https://doi.org/10.1007/978-3-030-98795-4_3

Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12., vollständig überarbeitete und aktualisierte Aufl). Beltz.

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *AI Magazine*, *27*(4), 12. https://doi.org/10.1609/aimag.v27i4.1904

Misztal, B. A. (1996). *Trust in modern societies: The search for the bases of social order*. Polity press.

Mitteilung der Kommission an das Europäische Parlament und den Rat über die Verbesserung des Datenschutzes durch Technologien zum Schutz der Privatsphäre (2007). https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:52007DC0228

Pavlou, P. A. (2003). *Consumer Acceptance of Electronic Commerce: Integrating Trust and Risk with the Technology Acceptance Model* (SSRN Scholarly Paper 2742286). Social Science Research Network. https://papers.ssrn.com/abstract=2742286

Rogers, E. M. (1995). *Diffusion of innovations* (4. ed). Free Press.

Rössler, B. (2001). *Der Wert des Privaten*. Suhrkamp.

Russell, S. J., & Norvig, P. (with Kirchner, F.). (2012). *Künstliche Intelligenz: Ein moderner Ansatz* (3., aktualisierte Auflage). Pearson.

Siegrist, M., Cvetkovich, G., & Roth, C. (2000). Salient Value Similarity, Social Trust, and Risk/Benefit Perception. *Risk Analysis*, *20*(3), 353–362. https://doi.org/10.1111/0272-4332.203034

Sousa, S., & Kern, R. (2023). How to keep text private? A systematic review of deep learning methods for privacy-preserving natural language processing. *Artificial Intelligence Review*, *56*(2), 1427–1492. https://doi.org/10.1007/s10462-022-10204-6

Taddeo, M. (2009). Defining Trust and E-Trust: From Old Theories to New Problems. *International Journal of Technology and Human Interaction*, *5*(2), 23–35. https://doi.org/10.4018/jthi.2009040102

Torkzadehmahani, R., Nasirigerdeh, R., Blumenthal, D. B., Kacprowski, T., List, M., Matschinske, J., Spaeth, J., Wenke, N. K., & Baumbach, J. (2022). Privacy-Preserving Artificial Intelligence Techniques in Biomedicine. *Methods of Information in Medicine*, *61*(S 01), e12–e27. https://doi.org/10.1055/s-0041-1740630

Uttenthal, M. (2024). A conceptual analysis of trust. *Social Science Information*, *63*(3), 392–410. https://doi.org/10.1177/05390184241270835