ANONYMISIERUNG FÜR VERNETZTE MOBILITÄTSSYSTEME

— Anonymität und Mobilität

Whitepaper zum Begriffs- und Domänenverständnis des Kompetenzcluster ANYMOS – Anonymisierung für vernetzte Mobilitätssysteme

ANYMOS-Webseite: www.anymos.de

— Autoren

Lukas Kneis[1], Maria Rill[1], Sascha Alpers[1], Michael Friedewald[2], Bernd Beckert[2], Josef Brücklmayr[1], Oliver Denninger[1], Konstantin Krauß[2], Frederik M. Metzger[2], Alexandra Wins[1]

[1] FZI Forschungszentrum Informatik, Haid- und Neu-Straße 10-14, 76131 Karlsruhe, {kneis, m.rill, alpers, brueckmayr, denninger, wins}@fzi.de, Webseite: www.fzi.de

[2] Fraunhofer-Institut für System- und Innovationsforschung, Breslauer Straße 48, 76139 Karlsruhe, {michael.friedewald, bernd.beckert, konstantin.krauss, frederik.metzger}@isi.fraunhofer.de, Webseite: www.isi.fraunhofer.de

— Publikation: 25. August 2023

— ANYMOS-Webseite: www.anymos.de

— Herzlichen Dank für die Förderung

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

Finanziert von der
Europäischen Union
NextGenerationEU

– Inhaltsverzeichnis

1 Motivation 4

2 Anonymisierung 5

2.1 Technische Perspektive 7

2.2 Rechtliche Perspektive 8

3 Mobilität und Verkehr 10

3.1 Individualverkehr 11

3.2 Öffentlicher Verkehr 13

3.3 Mobility as a Service 14

4 Ausblick 15

5 Quellen 16

6 Impressum 18

1 Motivation

In ANYMOS werden Anforderungen und Methoden für eine Anonymisierung und anschließende Auswertung von zuvor personenbezogenen Daten untersucht. Dabei wird im Kompetenzcluster die Anwendungsdomäne Mobilität betrachtet und sich auf den Personenverkehr, da durch die Mobilität von Gütern nicht immer unmittelbar personenbezogene Daten anfallen[1], fokussiert. Die Notwendigkeit des Kompetenzclusters ANYMOS ergibt sich daraus, dass im Mobilitätsbereich bei zahlreichen Anwendungen große Datenmengen anfallen und es aufgrund der zu erwartenden Entwicklungen zu einem weiteren Anstieg dieser Datenmenge kommen wird. Um diese Daten in Zukunft sinnvoll nutzen zu können, ohne dabei durch die Verwendung personenbezogener Daten Persönlichkeitsrechte und/oder rechtliche Vorgaben zu verletzen, muss zunächst erforscht werden, wann diese Daten gesammelt werden und inwieweit sie auch nach einer Anonymisierung noch über einen Nutzwert verfügen.

Im zweiten Abschnitt des Whitepapers wird daher zunächst Anonymität beschrieben und das Spannungsfeld zwischen juristischem und technischen Begriffsverständnis erörtert.

Im dritten Abschnitt erfolgt eine Strukturierung der Mobilitätsdomäne. Dadurch soll das gemeinsame Verständnis der Begrifflichkeiten und der Relevanz der verschiedenen Themenbereiche für das Kompetenzcluster ANYMOS gefördert werden.

Abschließend wird ein Ausblick – auch auf die weiteren Arbeiten in ANYMOS gegeben.

[1] Eine Ausnahme ist beispielsweise der Versand von Paketen an natürliche Personen. Das Tracking ermöglicht dann bspw. die Information, wer wann von wem (juristische oder andere natürliche Person) eine Sendung erhalten hat und ggf. (je nach Zustellart bzw. gewünschter Zustellart) Informationen über die Anwesenheit an der Zustelladresse.

2 Anonymisierung

Anonymität beschreibt einen Zustand, in dem eine Person agieren kann, ohne dass Informationen oder Handlungen auf die Person zurückgeführt werden können. Heute besteht ein Spannungsfeld zwischen Individualität, sowie der damit oft verbundenen Individualisierung und Anonymität[2]. Das nachfolgende Beispiel verdeutlicht dies: Ein Mobilitätsnachfrager möchte, ohne das Dritte dies mit seiner Person in Verbindung bringen oder bringen können, von seinem Wohnsitz in Karlsruhe zu einer Büroanschrift in Berlin reisen und nutzt dazu einen Routingdienst. Er erwartet aber gleichzeitig möglichst für ihn passende, d.h. individualisierte, Vorschläge (bspw. hinsichtlich der Reisezeitschätzung – welche auch auf seiner typischen PKW-Fahrgeschwindigkeit in verschiedenen Situationen basiert) zu erhalten.

Anonymität ist ein – weder immer mögliches noch hinreichendes – Mittel von natürlichen Personen um ihre digitale Souveränität (vgl. Beyerer, Müller-Quade, und Reussner 2018) zu wahren. Es ist nicht immer möglich, weil es Dienste gibt, die nur erbracht werden können, wenn eine Zuordnung zu einer Person möglich ist. Ein Beispiel ist die Nutzung eines Carsharing-Angebotes. Aus rechtlichen Gründen muss der Carsharing-Anbieter die Fahrerlaubnis des Fahrers prüfen und auch die Nutzungsdauer eines Fahrzeuges mit dem Fahrer in Verbindung bringen können (bspw. um im Falle eines Geschwindigkeitsverstoßes den Behörden die Identität des Fahrers nennen zu können). Es ist nicht hinreichend, weil Anonymität alleine noch keine digitale Souveränität gewährleistet. Hierzu sind (ergänzend) andere Faktoren, wie beispielsweise Transparenz bei der Verarbeitung von Nutzerdaten notwendig.

Das Grundrecht auf informationelle Selbstbestimmung in Deutschland hat, seitdem es das Bundesverfassungsgericht 1983 mit seinem sogenannten Volkszählungsurteil[3] aus Art. 1 und Art. 2 des Grundgesetztes geschöpft hat, eine besondere Bedeutung und prägt nachfolgend[4] die Ausbildung einer Eigenschaft im internationalen Wettbewerb. Diese kann von Vorteil sein, wenn die Produkte und Dienstleistungen wegen dieser Eigenschaft verstärkt gekauft werden, aber auch von Nachteil, wenn sie bestimmte Innovationen zu verhindern scheint.[5]

Die digitale Souveränität einzelner Personen steht im Spannungsfeld der digitalen Souveränität weiterer Akteure. Dabei gibt es verschiedene Stakeholdergruppen mit teils widersprüchlichen Anforderungen. Stakeholdergruppen stammen beispielsweise aus den natürlichen Personen (Bürger bzw. Gesellschaft), aus Organisationen (insbesondere juristische Personen) oder der öffentlichen Verwaltung. Auch Staaten und Staatengruppen sind Stakeholder digitaler Souveränität.

Dieses Spannungsfeld ist auch im Kontext der Mobilität entscheidend. Eine Kommune und mobilitätsdienstleistende Unternehmen können innerhalb ihres Gestaltungsspielraums bessere (für die Allgemeinheit oder für einzelne – je nach Zielfunktion) Entscheidungen treffen, wenn sie das Mobilitätsverhalten der Nutzer verstehen. Hierzu sind aussagekräftige Daten hilfreich. Anonyme Daten –

[2] Vergleiche Abschnitt „Individualität und Anonymität“ in (Zarnow 2017)

[3] BVerfGE 65, 1 – 71, https://www.bundesverfassungsgericht.de/e/rs19831215_1bvr020983.html

[4] Zur Wirkungsgeschichte des Volkszählungsurteils vergleiche (Landesbeauftragter für den Datenschutz Baden-Württemberg 2014)

[5] Die Gefährdung des Grundrechtes auf informationelle Selbstbestimmung durch Mobilität, beziehungsweise durch die Art und Weise wie Mobilitätsangebote realisiert werden beschreibt (Wagner 2021, Seite 149) fest.

das heißt Daten die keiner natürlichen Person zugeordnet sind oder zugeordnet werden können – können in diesem Spannungsfeld helfen sofern ihr Datennutzwert für die geplante Verwendung ausreicht. Inzwischen ist das Datenschutzrecht – welches dieses Grundrecht gewährleisten soll – in Europa großteils in der Datenschutz-Grundverordnung (DSGVO) durch die Europäische Union kodifiziert und harmonisiert.

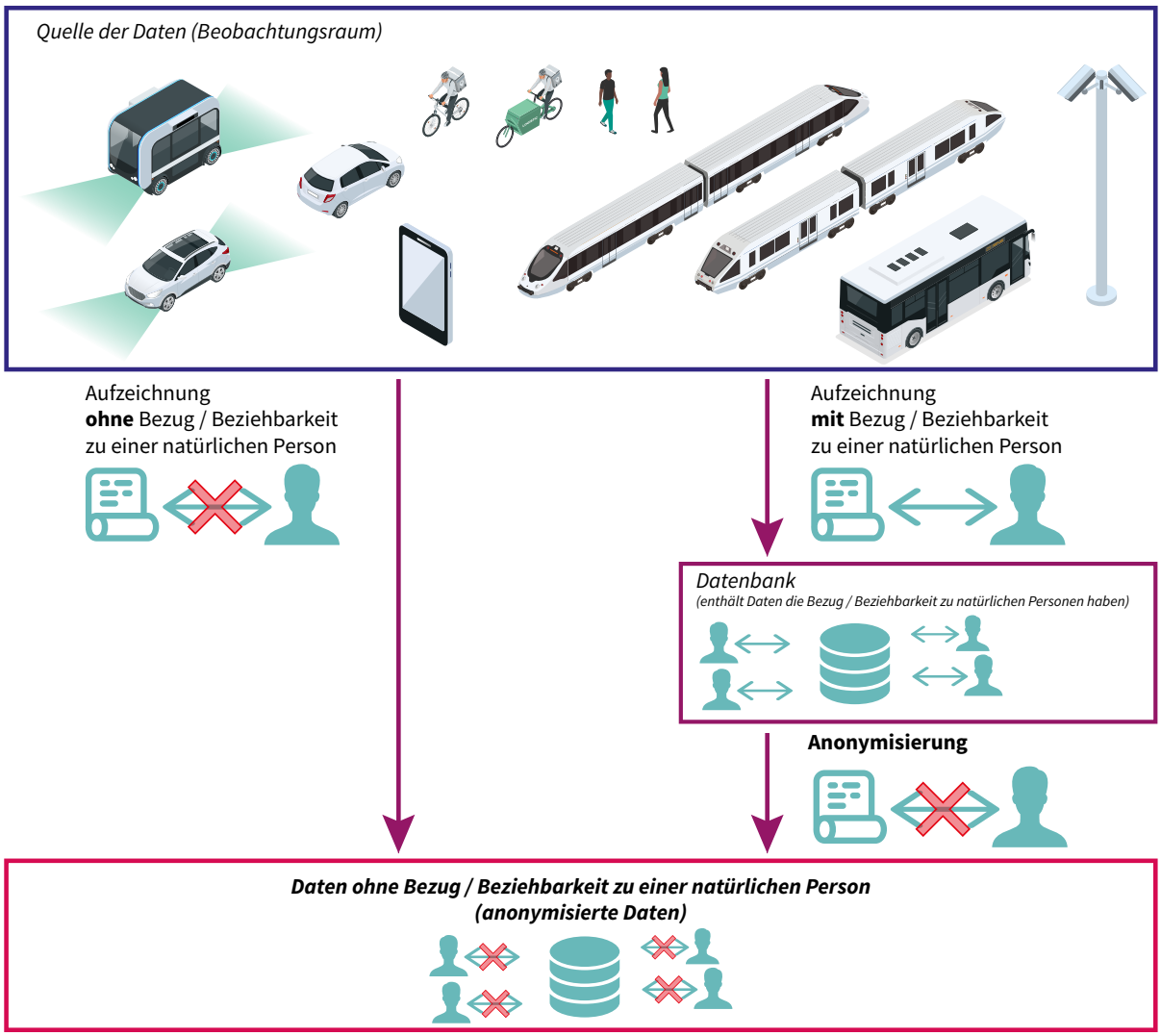

Abbildung 1: Übersicht über verschiedene Wege zu anonymisierten Daten

Dabei gibt es zwei Möglichkeiten anonyme Daten zu erhalten (siehe Abbildung 1). Zum einen können diese wie Abbildung 1 zeigt direkt so erhoben werden, dass es keine Möglichkeit gibt Daten einer konkreten natürlichen Person zuzuordnen. Zum anderen können Daten personenbezogen oder personenbeziehbar aufgezeichnet und nachträglich anonymisiert werden. Es verbleibt die Herausforderung festzustellen, wann Daten keine Möglichkeit zu einem Bezug zu einer natürlichen Person haben. Dies wird nachfolgend in zwei Perspektiven erörtert. Dabei ist zu beachten, dass die Einschätzung ob Daten keinen Bezug und keine Beziehbarkeit zu einer natürlichen Person aufweisen jeweils für einen bestimmten Zeitpunkt erfolgt und zu späteren Zeitpunkten überprüft werden muss. Hin-

tergrund ist auch, dass sich die Möglichkeiten einen Bezug zu einer Person herzustellen ändern können – beispielsweise, weil es zu einem späteren Zeitpunkt „neue“ weitere Datenquellen gibt mit denen eine Kombination möglich ist.

2.1 Technische Perspektive

Dieser Abschnitt soll den Begriff der Anonymisierung aus technischer Sicht beleuchten. Zahlreiche Anwendungsfälle erfordern die Verarbeitung von Daten, die Informationen über einzelne Personen enthalten und es erlauben, diese Personen eindeutig zu identifizieren – sogenannte personenbezogene Daten (Stummer 2023). Häufig werden Bestände solcher Daten für Analysezwecke an Dritte zur Verfügung gestellt, beispielsweise für statistische Auswertungen. Um die Anonymität der damit betroffenen Personen zu wahren, ist es erforderlich, vor der Weitergabe der Daten jegliche Informationen daraus zu entfernen, die eine Zuordnung der in einem Datensatz enthaltenen Information zu einer einzelnen Person ermöglichen. Man spricht hier vom Prozess der Anonymisierung (Buchmann 2015; Stummer 2023).

Das Ziel von Anonymisierung ist es, die Gefahr eines Rückschlusses von Informationen aus den anonymisierten Daten auf einzelne Personen – eine sogenannte Deanonymisierung oder Reidentifizierung – zu minimieren. Dass generell ein Restrisiko dafür besteht, ist allgemein akzeptiert (Stummer 2023). Allerdings ist dieses Restrisiko größer, als man intuitiv vermuten würde. Es ist zur Anonymisierung von Datensätzen nicht ausreichend, lediglich Merkmale zu entfernen, die explizit zur Identifizierung vorgesehen sind, wie etwa Namen oder Identifikationsnummern. In Praxisbeispielen konnte mehrfach gezeigt werden, dass allgemeinere Merkmale für sich jeweils keine Reidentifizierung erlauben, allerdings in Kombination miteinander schnell dazu dienen können, einzelne Personen eindeutig oder mit hoher Wahrscheinlichkeit zu identifizieren (Ohm 2009). Beispielsweise beschreibt (Golle 2006), dass ein Großteil der Bevölkerung der Vereinigten Staaten von Amerika lediglich anhand der Kombination aus ihrem Geschlecht, ihrem Geburtsdatum und ihrer Postleitzahl eindeutig bestimmbar sind. Auch solche quasi-identifizierenden Merkmale müssen bei der Anonymisierung berücksichtigt werden. Sie können beispielsweise entfernt oder verallgemeinert werden (aus dem Geburtsdatum könnte der Geburtsmonat oder das Geburtsjahr werden).

Wird der Grad an Anonymisierung eines Datenbestands bewertet, dann wird dabei immer gemessen, wie hoch die Wahrscheinlichkeit einer Reidentifizierung im schlimmsten bzw. im wahrscheinlichsten Fall ist (Buchmann 2015). Ein verbreitetes Maß ist hier das Modell der k-Anonymität (Sweeney 2002). Es besagt, dass bei einem Datenbestand, in dem jede Person einem Datensatz zugeordnet werden kann, nach der Anonymisierung jede Person nur noch mit einer Gruppe von gleichartigen Datensätzen in Verbindung gebracht werden kann. Die Menge aller Datensätze im Bestand wird durch die Anonymisierung in Äquivalenzklassen eingeteilt, sodass nur noch diese Klassen unterscheidbar sind, aber nicht mehr die Datensätze innerhalb jeder Klasse. Die Größe k der kleinsten Äquivalenzklasse gibt die Sicherheit der Anonymisierung an. Die Wahrscheinlichkeit, eine Person aus einer Äquivalenzklasse eindeutig zu identifizieren, beträgt 1/k. Je mehr Datensätze in der Klasse vorhanden sind, desto unwahrscheinlicher ist die Gefahr einer Reidentifizierung (Buchmann 2015).

Bei der Bewertung der k-Anonymität gibt es einige Aspekte zu beachten. Zunächst müssen die Äquivalenzklassen groß genug sein, um Anonymität ausreichend zu gewährleisten. Zweitens müssen die Datensätze innerhalb einer Äquivalenzklasse hinsichtlich ihrer sensiblen Informationen ausreichend

heterogen sein. Wenn, im Extremfall, alle Datensätze innerhalb einer Klasse identische Werte bei sensibler Information enthalten, so ist es gar nicht mehr nötig, zur Reidentifikation eine Person einem Datensatz zuzuordnen, da die Information über sie ohnehin aus dem anonymisierten Datensatz ausgelesen werden kann. Entsprechend hat auch die Verteilung, wie häufig welche Information in einer Äquivalenzklasse vorhanden ist, einen Einfluss auf den Grad der Anonymisierung. Demnach muss beim Prozess der Anonymisierung beachtet werden, dass in den entstehenden Äquivalenzklassen die Werte sensibler Information adäquat verteilt sind (Li, Li, und Venkatasubramanian 2007).

2.2 Rechtliche Perspektive

Im folgenden Abschnitt möchten wir die Begriffsdefinition zu „Anonymisierung“ aus rechtlicher Sicht umreißen.

Daten, die einen Personenbezug im Sinne des Art. 4 Nr.1 DSGVO aufweisen, unterliegen datenschutzrechtlichen Regelungen. Die Verarbeitung solcher Daten ist dementsprechend zunächst verboten, außer es liegt eine Erlaubnis vor. Diese Erlaubnis kann nach Art. 6 Abs. 1 DSGVO beispielsweise durch die Einwilligung (Art. 6 Abs. 1 lit a DSGVO) des Betroffen oder aufgrund einer rechtlichen Verpflichtung (Art. 6 Abs. 1 lit c DSGVO) oder zur Erfüllung eines Vertrages (Art. 6 ABs. 1 lit b DSGVO) gegeben sein.[6] Sollen personenbezogene Daten verarbeitet werden, so müssen neben der Erlaubnis weitere Aspekte erfüllt sein. Beispielsweise ist ein Zweck für die Verarbeitung festzulegen und es muss der Grundsatz der Datensparsamkeit beachtet werden.

Sollen Daten mit Dritten geteilt werden, ohne dass dies unter den ursprünglich festgelegten Verarbeitungszweck fällt oder diese Datenweitergabe nicht durch die ursprüngliche Erlaubnis abgedeckt ist, so muss der Personenbezug dieser Daten entfernt werden, sprich die Daten werden anonymisiert. Die Anonymisierung von Daten stellt jedoch einen eigenen Verarbeitungsvorgang dar, der ebenfalls einer Erlaubnis bedarf, beispielsweise in Form einer Einwilligung oder rechtlichen Erlaubnis. Für die anonymisierten Daten sind dann im Weiteren die Regelungen der DSGVO nicht mehr relevant.[7]

„Anonyme Daten“ ist nicht in den Begriffsbestimmungen des Art. 4 DSGVO definiert. In Erwägungsgrund 26 wird jedoch ausgeführt, dass es sich hierbei um solchen Daten handelt, die entweder von vornherein keinerlei Personenbezug (Abbildung 1, links) aufweisen oder deren Personenbezug so entfernt wurde (Abbildung 1, rechts), dass der Aufwand an Kosten, Arbeit und Zeit die Daten auf eine Person beziehen zu können unverhältnismäßig groß ist. Der Begriff „anonyme Daten“ wird somit in Abgrenzung zum Begriff der „personenbezogenen Daten“ erläutert. Daher ist, bei der Beurteilung der Frage, ob ein Datum ein anonymes Datum ist, relevant, ob der Personenbezug (noch) vorliegt oder nicht. Bei der Frage der Personenbeziehbarkeit von Daten wurden in der Literatur zwei Meinungen diskutiert: der relative und der absolute Ansatz bezüglich der Personenbeziehbarkeit. Streitgegenstand ist hierbei, welches Zusatzwissen hinzuzurechnen ist bei der Beurteilung, ob eine Person identifizierbar ist und somit ein Personenbezug vorliegt. Der absolute Ansatz zieht das sogenannte

[6] Weitere Legitimationsgrundlagen finden sich in den Art.6 Abs. 1 lit d – f DSGVO.

[7] Siehe Erwägungsgrund 26 zur DSGVO. Erwägungsgründe haben keinen normativen Charakter, sondern sind erläuternder Natur. Sie legen dar, was der Gesetzgeber sich bei bestimmten Themen gedacht hat und können zur Auslegung von gesetzlichen Regelungen verwendet werden.

„Weltwissen“ heran (also jegliches irgendwo auf der Welt verfügbare Wissen). Der relative Ansatz betrachtet lediglich das Wissen, beziehungsweise die Mittel und Fähigkeiten desjenigen, der zu den Daten Zugang hat. Laut EWG 26 tendiert die DSGVO zum relativen Ansatz. Auch das Gericht der Europäischen Union (EuG) folgt der Theorie des relativen Personenbezugs in seiner aktuellen Rechtsprechung.

Am 23. April 2023 entschied das EuG in einem Urteil (EuG, Urteil vom 23.04.2023 - T-557/20)[8], dass pseudonymisierte Daten, die einem Dritten übermittelt werden, genau dann keinen Personenbezug mehr aufweisen, wenn es für den Dritten keine Möglichkeit gibt, die Pseudonymisierung rückgängig zu machen. Hier zeigt sich, dass es bei der Frage, ob es sich um anonyme oder personenbezogene Daten handelt, auf die Mittel und Fähigkeiten desjenigen ankommt, der auf die Daten zugreifen kann und somit universelle/allgemeingültige Aussagen zur Einordnung eines Datensatzes nicht zwangsläufig gemacht werden können. Die ehemalige Art. 29 Datenschutzgruppe hat bereits 2014 eine Stellungnahme zu Anonymisierungstechniken und der rechtlichen Einordnung dieser veröffentlicht.[9] Diese wird nach wie vor vom European Data Privacy Board (EDPB) für Handlungsempfehlungen herangezogen.[10]

Mobilitätsdaten können einen Personenbezug aufweisen. Dementsprechend sind auch bei der Weitergabe solcher Daten die datenschutzrechtlichen Vorgaben einzuhalten. Gleichzeitig ist sowohl national als auch international der Datenaustausch erwünscht und wird grundsätzlich durch entsprechende Gesetzgebung im Rahmen der Digitalstrategie wie dem Data Act, Data Governance Act, sowie im Speziellen durch die Intelligente Verkehrssysteme Richtlinie sowie dem geplanten Mobilitätsdatengesetz forciert.[11]

Konsequenz für ANYMOS:

Die Daten, die auf der Plattform zur Verfügung gestellt werden, könnten einerseits für die Nutzenden als anonyme Daten nutzbar sein. Zu berücksichtigen ist hierbei jedoch, welche Mittel und Fähigkeiten die Nutzenden zur Reidentifizierung der Daten haben bzw. ihnen zugerechnet werden müssen. Andererseits könnten diese Daten für die bereitstellenden Akteure einen Personenbezug aufweisen, weswegen bei diesen die Pflichten der DSGVO zu beachten wären und somit unter anderem eine Legitimationsgrundlage für die weitere Übermittlung der Daten vorliegen muss.

[8] https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX%3A62020TJ0557

[9] https://cnpd.public.lu/content/dam/cnpd/fr/publications/groupe-art29/wp216_en.pdf

[10] https://edpb.europa.eu/sites/default/files/files/file1/edpb_replyec_questionnaireresearch_final.pdf S. 11 Rn.46.

[11] https://digital-strategy.ec.europa.eu/de/policies/mobility-data

3 Mobilität und Verkehr

Mobilität wird im Rahmen des Kompetenzclusters ANYMOS als **räumliche Mobilität** aufgefasst. Räumliche Mobilität ist hierbei als Ortsveränderung beziehungsweise Wechsel zwischen mindestens zwei Standorten definiert. Metriken für die Mobilität sind die **Mobilitätsrate**, das **Mobilitätsstreckenbudget** und das **Mobilitätszeitbudget**. Die Mobilitätsrate gibt hierbei die Anzahl an Ortsveränderungen je Zeiteinheit an, das Mobilitätsstreckenbudget die durchschnittliche Wegelänge pro Person innerhalb eines gewählten Zeitraums. Als Mobilitätszeitbudget werden die durchschnittlichen Wegestunden bezeichnet, die eine Person innerhalb des gewählten Zeitraums aufwendet (Stock und Bernecker 2014).

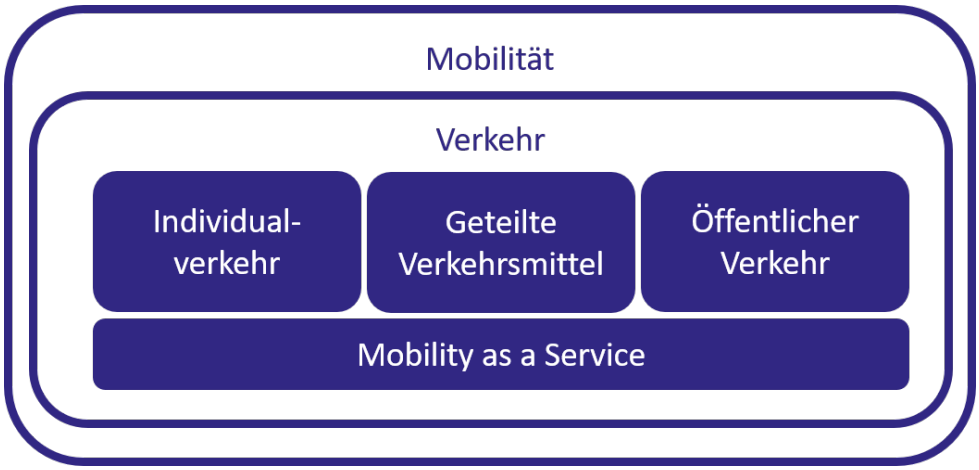

Abbildung 2: Strukturierung der Mobilitätsdomäne

Der **Verkehr** ist eine Teilmenge der Mobilität und bezeichnet eine Raumüberwindung unter Verwendung technischer und organisatorischer Einrichtungen. Metriken für den Verkehr sind die Anzahl der beförderten Personen in Form des **Verkehrsaufkommens** und die **Verkehrsleistung** als Produkt der Beförderungsfälle und jeweiligen Entfernung. Die Verkehrsleistung bildet auch die Grundlage für den Modal Split (Kagerbauer 2021). Die Anzahl an Fahrzeugen, beziehungsweise die Auslastung der Verkehrsinfrastruktur entspricht somit entgegen dem allgemeinen Sprachgebrauch nicht direkt dem Verkehrsaufkommen.

Der **Modal Split** gibt die Verteilung der Verkehrsmittel gemessen an ihrem Anteil an der Gesamtverkehrsleistung an. Die Verkehrsmittel nutzen jeweils einen Verkehrsträger. Dieser Verkehrsträger stellt das Medium, bzw. die technische Infrastruktur, dar, die das Verkehrsmittel benötigt. Als Verkehrsträger ergeben sich somit Straße, Schiene, Wasser und Luft. Im Personenverkehr spielen im Nahverkehr vor allem Straße und Schiene, sowie im Fernverkehr Straße, Schiene und Luft eine Rolle. **Nahverkehr** bezeichnet hierbei Strecken von unter 50 Kilometern und unter einer Stunde Reisedauer. Dementsprechend wird bei Distanzen über 50 Kilometern und mehr als einer Stunde Reisedauer von **Fernverkehr** gesprochen (Stock und Bernecker 2014).

Eine wachsende Bedeutung im Verkehr spielen **Multimodalität** und **Intermodalität**. Multimodalität bezeichnet die Verwendung verschiedener Verkehrsmittel im Verlauf eines Zeitraums. Wird also beispielsweise innerhalb einer Woche neben Autofahrten auch das Fahrrad genutzt, so spricht man von Multimodalität. Bei der Intermodalität handelt es sich um einen Spezialfall der Multimodalität bei dem auch innerhalb eines einzelnen Weges verschiedene Verkehrsmittel genutzt werden. Ein Weg ist hierbei definiert als Ortsveränderung einer Person von Start bis Ziel. Wird innerhalb eines Zeitraums nur ein Verkehrsmittel genutzt, spricht man von **Monomodalität**. Für die Monomodalität wichtig ist der Umgang mit Fußwegen. Werden Fußwege berücksichtigt, sinkt der Anteil von monomodalem Verhalten. Bei der Verwendung von Metriken zur Modalität sollte daher klar angegeben werden, ob Fußwege grundsätzlich als eigener Modus gezählt werden (Kagerbauer 2021). Wenn dies der Fall ist, können zusätzliche Voraussetzungen bestimmt werden, die ein Fußweg erfüllen muss, um als eigene Etappe zu zählen. Abhängig vom jeweiligen Umfeld müssen hier individuell sinnvolle Kriterien gewählt werden. Die so gewählten Kriterien sollten dann klar kommuniziert werden.

Die zunehmende Verhaltensänderung hin zu einem multi- bzw. intermodalen Verhalten wird durch eine Reihe verschiedener Faktoren getrieben. So bleibt das Auto zwar häufig das Hauptverkehrsmittel, aufgrund ökologischer Gesichtspunkte findet allerdings eine Orientierung hin zu Alternativen statt. Dies ermöglicht gerade in Ballungsräumen eine höhere Auswahl an Mobilitätsmöglichkeiten. Dank der Verbreitung von Smartphones und neuer Produkte im Mobility as a Service-Umfeld können sich Nutzer auch unterwegs über diese Alternativen informieren und so verschiedene Verkehrsmittel nutzen (Kagerbauer, 2021). Die hierbei zusätzlich anfallenden Nutzerdaten machen diese Entwicklung auch für das Kompetenzcluster ANYMOS relevant.

3.1 Individualverkehr

Der **Individualverkehr** ist gekennzeichnet durch die freie Bestimmung über Fahrzeit, Fahrroute und Geschwindigkeit. Er wird unterteilt in den **Nichtmotorisierten Verkehr (NMV)**, bestehend unter anderem aus „zu Fuß gehen“ und Fahrradfahren, sowie dem **Motorisierten Individualverkehr (MIV)**. Verkehrsmittel des MIV sind Personenkraftwagen (PKW), Motorräder oder auch eigene Boote und eigene Flugzeuge (Dziekan und Zistel 2018; Kagerbauer 2021).

In Bezug auf das Verkehrsmittel PKW sind die **Innenraumüberwachung** und das **autonome Fahren** von Interesse. Bei der Innenraumüberwachung wird mit Hilfe von Sensoren das Fahrzeuginnere ausgewertet. Aktuell wird diese Technologie primär zur Müdigkeitserkennung eingesetzt, für das autonome Fahren sind jedoch weitere Techniken vorgeschrieben, um ein sicheres Hand-Over und Take-Over zu garantieren. Daneben sind weitere Einsatzmöglichkeiten, wie beispielsweise eine Erhöhung der Sicherheit beim Einsatz von Airbags oder automatisierte Notrufe bei gesundheitlichen Problemen der Insassen, denkbar. Für die Innenraumüberwachung werden unter anderem Kameras, akustische Sensoren oder Sensoren innerhalb des Lenkrades eingesetzt. Besonders bei der Messung von Vitaldaten ist eine Anonymisierung wichtig, da es sich hierbei um sensible Gesundheitsdaten handelt (Laufer 2022; Russ u. a. 2021).

Das autonome Fahren ist von der SAE International in 6 Level unterteilt. Level 0 entspricht hierbei keiner dauerhaften Unterstützung, bei Level 1 und Level 2 findet jeweils eine Unterstützung durch Assistenzsysteme statt. Erst ab Level 3 spricht man von autonomem Fahren. Mit Level 3 kann das Fahrzeug in bestimmten Situationen selbstständig fahren, in allen anderen Situationen muss der

Fahrer das Steuer übernehmen. Während das Auto gemäß den Bestimmungen von Level 3 autonom agiert, geht die Haftung bei Unfällen vom Fahrer auf den Hersteller des Fahrzeuges über. Um dies im Zweifelsfall nachvollziehen zu können, muss das Auto allerdings während des Betriebs Daten sammeln und über eine Black Box verfügen. Hierbei wird durch die Innenraumüberwachung sichergestellt, dass der Fahrer weiterhin dazu in der Lage ist, innerhalb eines vorgegebenen Zeitraums das Steuer zu übernehmen und sich nicht unsachgemäß verhält (Huber 2020). Bei einer Automatisierung von Level 4 ist eine Übergabe der Kontrolle an einen Fahrer nicht mehr zwingend vorgesehen, das Fahrzeug kann allerdings trotzdem nicht jederzeit an einem beliebigen Ort eingesetzt werden, sondern ist nach wie vor einigen Einschränkungen unterworfen. Erst mit Level 5 ist vorgesehen, dass das Fahrzeug immer und überall autonom agieren kann (SAE International 2021).

Autonome Fahrzeuge nutzen hierbei eine Vielzahl verschiedener Sensoren wie zum Beispiel Radar, Lidar und Kameras. Aus Sicht der Anonymisierung sind hierbei vor allem die Kameradaten relevant, da hierbei Personen im öffentlichen Raum aufgezeichnet werden. Diese haben einer Aufzeichnung nicht zugestimmt, eventuell sind sie sich dieser Aufzeichnung nicht einmal bewusst. Hierbei spielt auch eine Rolle, ob Daten nur live im Auto verarbeitet werden, oder ob eine Weitergabe an externe Server und eine dauerhafte Speicherung erfolgt (Bloom und Emery 2022).

Wird ein Individualverkehrsmittel entweder gemeinsam oder zeitlich versetzt von verschiedenen Personen genutzt, welche häufig weder Eigentümer noch Halter des Fahrzeugs sind, kann von einem **geteilten Verkehrsmittel** gesprochen werden. Hierunter fallen unter anderem Taxis, Bike- und Carsharing, Fahrgemeinschaften oder Ride-Hailing-Dienste (Kagerbauer 2021). Eine zunehmende Nutzung von geteilten Verkehrsmitteln ist Teil des Trends „Nutzen statt Besitzen“ und wird sich voraussichtlich auch in den nächsten Jahren fortsetzen (Scholl, Schulz, und Süßbauer 2010). Da bei vielen dieser Dienste die genauen Standortdaten von Start und Ziel bekannt sind, besteht ein Bedarf für die Anonymisierung dieser Daten, um keine genauen Routen nachvollziehen zu können und gleichzeitig den Nutzwert zu erhalten.

Für das Kompetenzcluster ANYMOS relevant sind hierbei anfallende Daten von Kameras und Akustiksensoren sowohl im Bereich der Innenraumüberwachung als auch aus dem Außenbereich des Fahrzeugs. Ebenfalls relevant sind Standortdaten, welche sowohl im Individualverkehr als auch bei geteilten Verkehrsmitteln anfallen. Hierbei muss zwischen Daten unterschieden werden, die nur während des Betriebs genutzt und somit nicht persistiert werden, und solchen Daten, die dauerhaft gespeichert werden. Sowohl Standortdaten als auch weitere im Verlauf des Kompetenzclusters ANYMOS identifizierte Daten können leichter einer Person, beziehungsweise einer kleinen Personengruppe zugeordnet werden. Zusätzlich findet durch die Außensensoren eine Überwachung des öffentlichen Raumes statt, wobei ebenfalls personenbezogene Daten entstehen. Durch eine Anonymisierung sollen sowohl die Rechte der Insassen als auch der Außenstehenden gewahrt werden und gleichzeitig der Nutzwert erhalten bleiben.

3.2 Öffentlicher Verkehr

Der **öffentliche Verkehr** ist dadurch gekennzeichnet, dass er im Rahmen von Beförderungsbedingungen für jeden zugänglich ist. Der öffentliche Verkehr ist typischerweise fahrplangebunden, es kann also nicht frei über Fahrzeit, Fahrroute und Geschwindigkeit entschieden werden. Der öffentliche Verkehr unterteilt sich in den **öffentlichen Personennahverkehr (ÖPNV)** und den **öffentlichen Personenfernverkehr (ÖPFV)** (Dziekan und Zistel 2018; Kagerbauer 2021).

Der ÖPNV tritt typischerweise als Linienverkehr auf und wird im alltäglichen Leben genutzt, also für Aktivitäten wie dem Weg zur Arbeit oder zum Einkaufen. Typische Verkehrsmittel sind Busse, U-Bahnen, Straßenbahnen und Eisenbahnen im Nahverkehr. Die Verkehrsträger Luft und Wasser spielen im ÖPNV derzeit nur eine untergeordnete Rolle. Typische Verkehrsmittel im ÖPFV sind der Fernzug, der Reisebus oder das Flugzeug. Auch hier spielt der Verkehrsträger Wasser nur eine untergeordnete Rolle (Dziekan und Zistel 2018; Kagerbauer 2021).

Das **Ticketing** im öffentlichen Verkehr ist in den letzten Jahren zunehmend accountbasiert. Hierdurch können zunehmend personenbezogene Daten gesammelt werden, wodurch beispielsweise eine Nachverfolgung der Routen einer Person möglich wird. Ebenfalls datenschutzrelevant ist die **automatisierte Fahrgastzählung**. Hierbei soll automatisiert die aktuelle Auslastung eines öffentlichen Verkehrsmittels erfasst werden. Systeme zur automatisierten Fahrgastzählung lassen sich nach ihrem Interaktionsgrad mit den Fahrgästen charakterisieren. Ist für ein solches System eine direkte Interaktion des Fahrgastes nötig spricht man von einem **integrierten System**. Hierzu zählen beispielsweise SmartCards, welche durch An- und Abmeldung in der Lage sind Start und Ende einer Fahrt zu erfassen. Ein solches System wird beispielsweise in London eingesetzt (Bagchi und White 2005). In Deutschland wird stattdessen bei integrierten Systemen nur der Zustieg, nicht der Ausstieg erfasst.

Müssen die Fahrgäste nicht direkt mit dem System interagieren, so spricht man von einem **unabhängigen System**. Hierfür werden verschiedene Techniken zur Fahrgastzählung eingesetzt. Unter anderem Drucksensoren, Kameras oder Wi-Fi-(Grgurević, Juršić, und Rajič 2022). Je nach verwendetem System ist zu ermitteln, inwiefern die hierbei ermittelten Daten überhaupt über einen Personenbezug verfügen.

Durch Auswertung der durch Ticketing und Fahrgastzählung gesammelten Daten kann ein genauerer Blick auf Nutzerverhalten und Nutzerbedürfnisse erhalten werden, jedoch müssen die Rechte der Fahrgäste gewahrt werden. Für das Kompetenzcluster ANYMOS interessant ist somit die Anonymisierung dieser Daten bei gleichzeitiger Erhaltung des Nutzwertes. So sollen neben optimierter Verkehrsplanung auch verbesserte Ticketingoptionen und eine verbesserte Routenplanung ermöglicht werden. Wie auch im Individualverkehr, fallen zusätzlich Kameradaten an, die jedoch nicht zwingend aus technischen Gründen, sondern auch aus Gründen der Überwachung und eventuellen Strafverfolgung aufgenommen werden. Dies bringt andere Anforderungen mit sich als beispielsweise Daten die nur zur Live-Auswertung benötigt werden. Wann gesammelte Daten eine Identifikation ermöglichen, muss im Rahmen von ANYMOS und auch projektspezifisch in weiteren Vorhaben ermittelt werden.

3.3 Mobility as a Service

Mobility as a Service (MaaS) bezeichnet die Integration verschiedener Mobilitätsdienste in ein einheitliches Interface. Merkmale von MaaS-Diensten sind unter anderem die Förderung der Multimodalität und Intermodalität durch Verknüpfung verschiedener Mobilitätsdienstleister, die Verfügbarkeit übergreifender Tarifoptionen, wie beispielsweise eines Abonnements, und die Personalisierung der Routenvorschläge. MaaS-Dienste befinden sich aktuell noch im Aufbau, werden in den nächsten Jahren aber zunehmend an Bedeutung gewinnen (Jittrapirom u. a. 2017).

Analog zur SAE-Klassifizierung im Bereich autonomes Fahren lassen sich auch MaaS-Dienste in verschiedene Level unterteilen. Die Kriterien sind hierbei die **informationelle Integration**, die **transaktionale Integration** und die **operative Integration**. Es sollen also Informationen über verschiedene Verkehrsmittel verknüpft werden (informationelle Integration), diese sollen gemeinsam buchbar sein (transaktionale Integration), und der Umstieg zwischen verschiedenen Verkehrsmitteln soll reibungslos funktionieren (operative Integration). Level 0 entspricht keiner Integration, ab Level 1 sind in einer Plattform Informationen über einen Teil der verfügbaren Verkehrsmittel abrufbar. Ein MaaS-Dienst auf Level 2 bietet neben einer teilweisen informationellen Integration auch eine teilweise operative oder transaktionale Integration, teilweise auch beides. Wenn einige Wege mit einer vollständigen Integration zurückgelegt werden können, ist Level 3 erreicht. Für Level 4 ist allerdings eine vollständige Integration einiger Verkehrsträger gefordert. Bei Level 5 ist eine vollständige Integration in allen Situationen erreicht, es sind also alle Verkehrsträger vollständig operativ, informationell und transaktional in die Plattform integriert (Lyons, Hammond, und Mackay 2019).

Durch die umfassende Integration verschiedener Services hat ein MaaS-Anbieter Zugriff auf eine besonders hohe Datenmenge. Diese Daten bieten Möglichkeiten zur Verbesserung des Dienstes, zum Beispiel durch bessere Anpassung an Nutzerbedürfnisse oder die aktuelle Verkehrssituation. Es gilt somit, die hierfür wichtigen Daten nutzen zu können und gleichzeitig die Rechte der Fahrgäste zu wahren. In Verbindung mit der zunehmenden Entwicklung von MaaS-Diensten zeigt sich hierdurch die Relevanz für ANYMOS.

4 Ausblick

Mit einem wachsenden Grad an Automatisierung geht eine Erleichterung der Mobilität einher. Sei es durch eine Planung intermodaler Routen oder der Automatisierung der alltäglichen Fahrten mit dem Auto, diese Entwicklung geht mit einem steigenden Grad an anfallenden Daten einher. Vor allem die Zunahme an MaaS-Diensten und geteilten Verkehrsmitteln, die steigende Vernetzung von (autonomen) Fahrzeugen und der Wunsch nach einer besseren Mobilitäts- und Routenplanung im öffentlichen Verkehr treiben diese Entwicklung. Die Anonymisierung dieser Nutzerdaten ist ein wichtiger Faktor, um in all diesen Bereichen in Zukunft die Spannungsfelder Datenschutz und Datenverwertung in Einklang zu bringen. ANYMOS soll hierzu einen wichtigen Beitrag liefern.

Für Mobilitätsdaten ist im kommenden Jahr 2024 das sog. Mobilitätsdatengesetz zu erwarten, welches unter anderem Regeln für die Nutzung von Mobilitätsdaten beinhalten soll (Bundesministerium für Digitales und Verkehr 2022). Die Berücksichtigung der datenschutzrechtlichen Aspekte und insbesondere deren Umsetzung wird im Rahmen des Kompetenzclusters verfolgt und fließt in die Gestaltung der Plattform ein. Hierbei wird auch die weitere Ausgestaltung des Data Acts berücksichtigt. Des Weiteren wird ein Handlungsleitfaden des European Data Privacy Board (EDPB) für das Jahr 2024 erwartet. Auch diese Ergebnisse werden im Rahmen des Kompetenzclusters einfließen.

Als Grundlage für die weitere Arbeit im Kompetenzcluster ANYMOS wird zunächst, wie in Abbildung 3 zu sehen, eine Anonymitätsdefinition erstellt werden. Diese soll die in der Mobilitätsdomäne und ihren jeweiligen Anwendungsfällen auftretenden Anforderungen an Anonymisierung abdecken. Zusätzlich zu dieser Anonymitätsdefinition wird im späteren Verlauf von ANYMOS auch eine Definition des Datennutzwertes erstellt. Hierbei sollen die Besonderheiten von anonymisierten Daten berücksichtigt werden, um den praktischen Nutzen dieser Daten bestimmen zu können. An dieser Stelle zeigt sich auch das in diesem Kompetenzcluster untersuchte Spannungsfeld zwischen Anonymität und Datennutzung. Die hierfür im Kompetenzcluster ANYMOS erarbeiteten Lösungen sollen am Ende der Laufzeit im November 2025 durch Demonstratoren veranschaulicht werden.

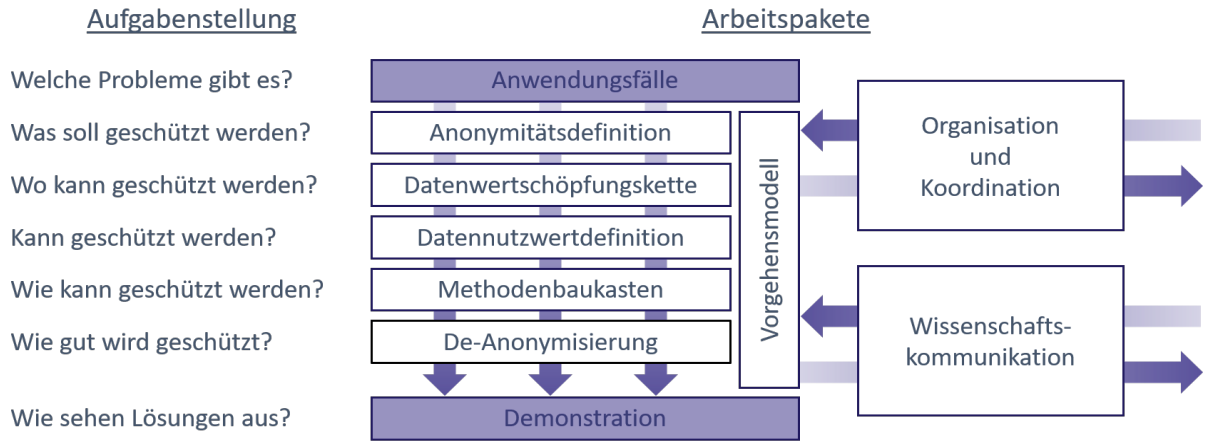

Abbildung 3: Aufgabenstellung und Struktur der Arbeitspakete im Kompetenzcluster ANYMOS

5 Quellen

Bagchi, Mousumi, und Peter White. 2005. „The Potential of Public Transport Smart Card Data". *Transport Policy* 12(5):464–74. doi: 10.1016/j.tranpol.2005.06.008.

Beyerer, Jürgen, Jörn Müller-Quade, und Ralf Reussner. 2018. „Karlsruher Thesen zur Digitalen Souveränität Europas". *Datenschutz und Datensicherheit - DuD* 42(5):277–80. doi: 10.1007/s11623-018-0940-2.

Bloom, Cara, und Josiah Emery. 2022. „Privacy Expectations for Human-Autonomous Vehicle Interactions". S. 1647–54 in *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. Napoli, Italy: IEEE. doi: 10.1109/RO-MAN53752.2022.9900615.

Buchmann, Erik. 2015. „Wie kann man Privatheit messen?: Privatheitsmaße aus der Wissenschaft". *Datenschutz und Datensicherheit - DuD* 39(8):510–14. doi: 10.1007/s11623-015-0461-1.

Bundesministerium für Digitales und Verkehr. 2022. „BMDV startet Prozess für ein Mobilitätsdatengesetz". *Bundesministerium für Digitales und Verkehr*. Abgerufen (https://bmdv.bund.de/SharedDocs/DE/Pressemitteilungen/2022/081-wissing-mobiliitaetsdatengesetz.html).

Dziekan, Katrin, und Meinhard Zistel. 2018. „Öffentlicher Verkehr". S. 347–72 in *Verkehrspolitik*, herausgegeben von O. Schwedes. Wiesbaden: Springer. doi: 10.1007/978-3-658-21601-6_16.

Golle, Philippe. 2006. „Revisiting the Uniqueness of Simple Demographics in the US Population". S. 77–80 in *Proceedings of the 5th ACM workshop on Privacy in electronic society*. Alexandria Virginia USA: ACM. doi: 10.1145/1179601.1179615.

Grgurević, Ivan, Karlo Juršić, und Vinko Rajič. 2022. „Review of Automatic Passenger Counting Systems in Public Urban Transport". S. 1–15 in *5th EAI International Conference on Management of Manufacturing Systems*, *EAI/Springer Innovations in Communication and Computing*, herausgegeben von L. Knapčíková, D. Peraković, A. Behúnová, und M. Periša. Cham: Springer. doi: 10.1007/978-3-030-67241-6_1.

Huber, Christian. 2020. „Automatisiertes und autonomes Fahren – wer haftet?" S. 697–711 in *Handbuch Industrie 4.0: Recht, Technik, Gesellschaft*, herausgegeben von W. Frenz. Berlin, Heidelberg: Springer. doi: 10.1007/978-3-662-58474-3_36.

Jittrapirom, Peraphan, Valeria Caiati, Anna-Maria Feneri, Shima Ebrahimigharehbaghi, María J. Alonso González, und Jishnu Narayan. 2017. „Mobility as a Service: A Critical Review of Definitions, Assessments of Schemes, and Key Challenges". *Urban Planning* 2(2):13–25. doi: 10.17645/up.v2i2.931.

Kagerbauer, Martin. 2021. „Multimodalität". S. 179–98 in *Stadtverkehrsplanung Band 1*, herausgegeben von D. Vallée, B. Engel, und W. Vogt. Berlin, Heidelberg: Springer. doi: 10.1007/978-3-662-59693-7_7.

Landesbeauftragter für den Datenschutz Baden-Württemberg. 2014. „LfD Baden-Württemberg: 30 Jahre Volkszählungsurteil — aktueller denn je". *Datenschutz und Datensicherheit - DuD* 38(2):133–133. doi: 10.1007/s11623-014-0055-3.

Laufer, Patrick. 2022. „Privatsphäre und Grenzen bei der Innenraumsensorik". *ATZ - Automobiltechnische Zeitschrift* 124(9):16–23. doi: 10.1007/s35148-022-0888-2.

Li, Ninghui, Tiancheng Li, und Suresh Venkatasubramanian. 2007. „t-Closeness: Privacy Beyond k-Anonymity and l-Diversity“. S. 106–15 in *2007 IEEE 23rd International Conference on Data Engineering*. Istanbul: IEEE. doi: 10.1109/ICDE.2007.367856.

Lyons, Glenn, Paul Hammond, und Kate Mackay. 2019. „The Importance of User Perspective in the Evolution of MaaS“. *Transportation Research Part A: Policy and Practice* 121:22–36. doi: 10.1016/j.tra.2018.12.010.

Ohm, Paul. 2009. „Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization“. *UCLA Law Review* 57:1701-- 1777.

Russ, Philipp, Frank Laakmann, Martin Seyffert, und Michele Vivaldelli. 2021. „Wachsamer Blick in den Innenraum“. *ATZelektronik* 16(7–8):48–51. doi: 10.1007/s35658-021-0643-6.

SAE International, SAE International. 2021. „Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles“. *SAE International*. Abgerufen 11. August 2023, https://www.sae.org/standards/content/j3016_202104/.

Scholl, Gerd, Lasse Schulz, und Elisabeth Süßbauer. 2010. „Nutzen statt Besitzen - Perspektiven für ressourceneffizienten Konsum durch innovative Dienstleistungen“. *Materialeffizienz und Ressourcenschonung* (Paper zu Arbeitspaket 12).

Stock, Wilfried, und Tobias Bernecker. 2014. „Einführung in die Thematik: Grundlegende Begriffe und ihre empirische Darstellung“. S. 1–62 in *Verkehrsökonomie*. Wiesbaden: Springer. doi: 10.1007/978-3-658-02308-9_1.

Stummer, Sarah. 2023. „Personenbezogenheit vs. Anonymität: Ein Mapping des rechtlichen und technischen Begriffsverständnisses von ‚Personenbezogenheit‘, ‚Pseudonymität‘ und ‚Anonymität‘“. *Datenschutz und Datensicherheit - DuD* 47(6):354–60. doi: 10.1007/s11623-023-1776-y.

Sweeney, Latanya. 2002. „K-Anonymity: A Model for Protecting Privacy“. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 10(05):557–70. doi: 10.1142/S0218488502001648.

Wagner, Manuela. 2021. *Das neue Mobilitätsrecht: Der Rechtsrahmen zum automatisierten und vernetzten Fahren*. Baden-Baden: Nomos. doi: 10.5771/9783748925927.

Zarnow, Christopher. 2017. „Religionsproduktive Differenzen: Bausteine Zu Einer Theologie Des Urbanen“. *Praktische Theologie* 52(4):220–26. doi: 10.14315/prth-2017-0409.

6 Impressum

— DOI: 10.5445/IR/1000161584

Herausgeber

FZI Forschungszentrum Informatik

Haid-und-Neu-Str. 10-14

76131 Karlsruhe

+49 721 9654-0

fzi@fzi.de

www.fzi.de

This work is licensed under the Creative Commons Attribution 4.0 International license (CC BY 4.0). You can read the license terms here: http://creativecommons.org/licenses/by/4.0