November 2026 ISSN 2628-4154



Beitragende Autor*innen

Franziska Danner
Elisabeth Soller
Wibke Michalk
Jonathan Gescher
Tim Leibert
Wladimir Sgibnev
Anne Mehl
Mariana Batista
Thuy Chinh Duong
Stephan Leppler
Svetlana Zubareva
Louis Stille-Hönig
Markus Lienkamp
Marlene Grebe
Michael Ortgiese
Pradeep Burla
Kerstin Stark
AIAMO (Greta Schmalhorst et al.)
Ann-Katrin Voit

DVWG Journal für Mobilität und Verkehr

Nr. 27 (2026)

Daten in der Mobilität

Wissenschaftliche Leitung
Prof. Dr.-Ing. Jeanette Klemmer
Dr. Antje-Mareike Dietrich

Eine Open-Access-Publikation der Deutschen Verkehrswissenschaftlichen Gesellschaft e. V.

Inhaltsverzeichnis

Editorial 1
Jeanette Klemmer, Antje-Mareike Dietrich

Nonresponse in Befragungen – Zum Umgang in der Erhebung von Mobilität 3
Franziska Danner, Elisabeth Soller, Wibke Michalk

Beifang als Best-Practice? Mit Routinedaten der Datenarmut ländlicher Räume begegnen 13
Jonathan Gescher, Tim Leibert, Wladimir Sgibnev, Anne Mehl

Repräsentativität bei Mobilitätserhebungen mit GPS-Smartphone-Tracking: Methodische Ansätze im Überblick 24
Mariana Batista, Thuy Chinh Duong, Stephan Leppler

Generieren, verarbeiten, veröffentlichen: Einheitlicher Umgang mit Mobilitätsdaten in interdisziplinären Forschungsclustern 32
Svetlana Zubareva, Louis Stille-Hönig, Markus Lienkamp

Mobilitätsdaten als Grundlage für eine ganzheitliche ÖPNV-Planung: Die PIMO-Plattform 42
Marlene Grebe, Michael Ortgiese, Pradeep Burla, Kerstin Stark

Entwicklung eines konzeptionellen Frameworks zur Erfassung, Qualitätssicherung und Bereitstellung urbaner Umweltdaten als Grundlage für ein umweltsensitives Mobilitätsmanagement im Projekt AIAMO 52
Greta Schmalhost, Timo Houben, Ema Vosgerau, Florian Berger, Michael Weger, Joost Hemmena, Alex Romanjuta, Paul Remmler, David Schäfer, Jan Bumberger, Thomas Trabert

Mobilitätsdaten zwischen Effizienz und Gerechtigkeit: Ein Governance-Framework für eine nachhaltige und inklusive Mobilitätsplanung 62
Ann-Katrin Voit

Quellennachweis Deckblatt: igorkocka-city-10230533.jpg von IgorKocka

Journal für Mobilität und Verkehr

ISSN 2628-4154

www.dvwg.de

Editorial

Prof. Dr.-Ing. Jeanette Klemmer (FH Münster)

Dr. Antje-Mareike Dietrich (WelfenAkademie e.V. – Berufsakademie)

Seit Jahren ist ein großer Wandel in der Erfassung und Auswertung von Mobilitätsdaten zu verzeichnen. Treiber dieser Entwicklung sind insbesondere die Verbesserung der Erhebungstechnik, die Erweiterung der Erreichbarkeit von Verkehrsteilnehmenden über Smartphone-Apps sowie die Nutzung von Künstlicher Intelligenz für die Datenauswertung. Wir freuen uns sehr, in der nun vorliegenden Ausgabe einen Überblick über die Breite des Forschungsfelds und aktuelle Projekte in Deutschland geben zu können.

Befragungen sind ein zentrales Instrument zur Evaluation der Wirkung von Mobilitätsmaßnahmen. Gleichzeitig sind diese Erhebungsmethoden oftmals kostenintensiv und die Bereitschaft, an Umfragen teilzunehmen schwankt (und sinkt). In ihrem Beitrag ***„Nonresponse in Befragungen – Zum Umgang in der Erhebung von Mobilität“*** gehen die Autorinnen Franziska Danner, Elisabeth Soller und Wibke Michalk (alle TH Rosenheim) der Frage nach, wie wir auch mit geringer Beteiligung zu wertvollen Erkenntnissen kommen und welche Möglichkeiten es gibt, auch aus Nonresponse in Befragungen Schlüsse ziehen zu können.

Im anschließenden Artikel ***„Beifang als Best-Practice? Mit Routinedaten der Datenarmut ländlicher Räume begegnen“*** thematisieren Jonathan Gescher et al. ebenfalls die Herausforderung der Datenarmut und zeigen die Potenziale und Herausforderungen der Nutzung von sogenannten „Beifang-Daten“ auf, die bei der Nutzung von Mobilfunk, in der Abrechnung der Krankenkassen und aus Arbeitsmarkt-Daten (z.B. Pendler*innen-Ströme) anfallen.

Einen Schwerpunkt auf mögliche Alternativen zur Datengenerierung legen Mariana Batista*, Thuy Chinh Duong und Stephan Leppler. In ihrem Beitrag ***„Repräsentativität bei Mobilitätserhebungen mit GPS-Smartphone-Tracking: Methodische Ansätze im Überblick“*** geben Sie einen Überblick über methodische Ansätze zur Teilnehmerrekrutierung, Verfahren zur statistischen Gewichtung und Kalibrierung sowie Ansätze zur kontinuierlichen Datenvalidierung.

Dazu passt auch der Beitrag von Svetlana Zubareva, Louis Stille-Hönig und Markus Lienkamp. In ***„Generieren, verarbeiten, veröffentlichen: Einheitlicher Umgang mit Mobilitätsdaten in interdisziplinären Forschungsclustern“*** betrachten sie den Umgang mit Mobilitätsdaten und beschreiben die Notwendigkeit der transparenten Dokumentation, der qualitativ hochwertigen und datenschutzkonformen Aufbereitung sowie der einheitlichen Validierung. Sie zeigen einen möglichen Workflow der Datenbereinigung und -aufbereitung am Beispiel des Münchener Forschungsclusters MCube auf.

Neben der Datengewinnung und Validierung stehen zunehmend auch die Nutzung vorhandener Daten für eine ganzheitliche Planung in der fachlichen Diskussion. Vor diesem Hintergrund braucht es Tools und Plattformen, die Daten integrieren, verarbeiten und für die Planung niedrigschwellig nutzbar machen. Marlene Grebe et al. stellen in ihrem Beitrag ***„Mobilitätsdaten als Grundlage für eine ganzheitliche ÖPNV-Planung: Die PIMO-Plattform“*** einen Ansatz dafür vor.

Ebenfalls auf dem Plattform-Gedanken basiert der Beitrag ***„Entwicklung eines konzeptionellen Frameworks zur Erfassung, Qualitätssicherung und Bereitstellung urbaner Umweltdaten als Grundlage für ein umweltsensitives Mobilitätsmanagement im Projekt AIAMO“*** von Greta Schmalhorst et al.. Die Autor*innen zeigen den konzeptionellen Weg von der reinen Datenerhebung bis zur Bereitstellung verlässlicher Entscheidungsgrundlagen für die Verkehrsplanung und umweltsensitive Verkehrssteuerung auf.

Das JMV zum Thema Mobilitätsdaten wird abgerundet von einem zusammenfassenden Überblick von Ann-Katrin Voit. Sie untersucht in ihrem Beitrag **„Mobilitätsdaten zwischen Effizienz und Gerechtigkeit: Ein Governance-Framework für eine nachhaltige und inklusive Mobilitätsplanung“,** unter welchen methodischen und institutionellen Bedingungen Mobilitätsdaten für eine effiziente und zugleich sozial inklusive Verkehrsplanung eingesetzt werden können.

Die Zusammenstellung zeigt, wie vielfältig die Fragestellungen rund um Mobilitätsdaten sind. Ziel aller Arbeiten ist es, eine möglichst belastbare Grundlage für die Verkehrsforschung und -planung zu erhalten. Dies ist im Sinne einer zielgruppengerechten Gestaltung von Verkehrsangeboten.

Wir wünschen eine anregende Lektüre!

Antje-Mareike Dietrich und Jeanette Klemmer

Journal für Mobilität und Verkehr

ISSN 2628-4154

www.dvwg.de

Nonresponse in Befragungen – Zum Umgang in der Erhebung von Mobilität

Franziska Danner*, Elisabeth Soller, Wibke Michalk

Siehe Autor*innenangaben

Abstract

Befragungen sind ein zentrales Instrument in der Evaluation der Wirkung von Mobilitätsmaßnahmen, stehen jedoch zunehmend vor geringen und schwankenden Teilnahmequoten (vgl. Koch/Blohm 2016) und der Herausforderung, einen Nonresponse Bias ausschließen zu wollen (vgl. Groves/Peytcheva 2008; Peer et al. 2022). Der Beitrag untersucht deshalb: Wie können Wirkungen betrieblicher Mobilitätsmaßnahmen auch bei geringer Befragungsbeteiligung belastbar eingeordnet werden? Der Ansatz von Koch/Blohm (2016) wird erweitert, um das Ausmaß eines Nonresponse Bias bei geringen Rücklaufquoten einzuschätzen. Exemplarisch dient hier das Projekt CaMoRoT, in dem eine Interventionsstudie durchgeführt wird, die die Wirkung einzelner Interventionen wie auch deren Kombination evaluiert (z.B. Einführung eines Bike-Sharing-Angebots). Untersucht werden die Wahrnehmung und Haltung gegenüber Angeboten und der berichtete Modal Split. Zusätzlich kann die tatsächliche Nutzung der Angebote über Nutzungsdaten ausgewertet werden. Um mögliche Verzerrungen zu identifizieren und die Unsicherheit der Ergebnisse zu reduzieren, werden diese anhand mehrerer Dimensionen geprüft: Erstens werden die berichteten Verhaltensänderungen im Mobilitätsverhalten mit objektiven Nutzungsdaten der Mobilitätsmaßnahmen abgeglichen. Zweitens wird die Selektivität anhand einer möglichen Panelattrition geprüft, indem wiederholt Teilnehmende mit später ausgeschiedenen Befragten verglichen werden. Drittens wird ein Individual-Fixed-Effects-Modell eingesetzt, um die beobachteten Veränderungen in Mobilitätsverhalten, Nachhaltigkeit und Technologieaffinität innerhalb derselben Personen unter Kontrolle zeitinvarianter individueller Heterogenität zu untersuchen. Die Ergebnisse werden anhand externer Referenzdaten, die unter Anderem der MiD entspringen, eingeordnet. Die Kombination der Validierungsschritte ermöglicht es einzuschätzen, wie stark ein Nonresponse-Bias sein kann und ob in zukünftigen Befragungen weitere Attribute erhoben werden müssen.

Schlagwörter:

Betriebliche Mobilität, Nonresponse Bias, Umfragequote

1 Einleitung

Die zunehmende Verfügbarkeit digitaler Daten eröffnet neue Möglichkeiten zur Erfassung und Analyse individuellen Mobilitätsverhaltens: Insbesondere sensorbasierte Technologien, GPS-Daten und mobile Anwendungen ermöglichen es, zurückgelegte Wege und genutzte Verkehrsmittel mit hoher räumlicher und zeitlicher Auflösung zu erfassen (vgl. Sadeghian et al. 2021). Damit stehen der Mobilitätsforschung zunehmend Datenquellen zur Verfügung, die klassische Befragungen ergänzen können. Die Beobachtung realisierten Mobilitätsverhaltens bildet jedoch nur einen Teil der relevanten Informationen ab: Einstellungen und Präferenzen, die Wahrnehmung und Akzeptanz bestehender Mobilitätsangebote sowie subjektiv empfundene Barrieren oder Motive der Verkehrsmittelwahl lassen sich aus beobachteten Bewegungsdaten nur eingeschränkt ableiten. Befragungen bleiben daher insbesondere für die Erklärung von Mobilitätsentscheidungen ein wesentliches Erhebungsinstrument. Gleichzeitig wird die Gewinnung belastbarer Befragungsdaten zunehmend herausfordernd: Sinkende Teilnahmequoten

*korrespondierende Autorin

stellen Forschende der empirischen Sozial- und Mobilitätsforschung vor eine Herausforderung (vgl. Brick/Wiliams 2013; Wittwer et al. 2024).

Eine geringe Beteiligung ist nicht allein aufgrund einer reduzierten Stichprobengröße problematisch, da diese keine hinreichende Beurteilung eines möglichen Nonresponse Bias erlaubt (vgl. Groves 2006). Für die Aussagekraft der Ergebnisse ist vielmehr entscheidend, ob sich Teilnehmende und Nichtteilnehmende systematisch hinsichtlich der untersuchten Merkmale unterscheiden.

Vor diesem Hintergrund gewinnt die Validierung von Befragungsergebnissen anhand ergänzender und von der Befragung unabhängiger Datenquellen an Bedeutung. Insbesondere im Kontext der betrieblichen Mobilität können bereits vorhandene Daten dazu beitragen, die Zusammensetzung der Befragungsstichprobe und die Belastbarkeit der daraus abgeleiteten Ergebnisse zu beurteilen (vgl. Su et al. 2022). Ein solcher Ansatz verschiebt den Fokus von der alleinigen Maximierung der Teilnahmequote hin zu einer empirischen Prüfung möglicher Verzerrungen. Der vorliegende Beitrag untersucht daher, inwieweit die Ergebnisse einer Befragung zur betrieblichen Mobilität trotz begrenzter Teilnahme als belastbar angesehen werden können. Er ist wie folgt aufgebaut: Zuerst wird ein Überblick über die aktuelle Forschung zum Thema Nonresponse Bias gegeben. Nachfolgend wird das Vorgehen einer Nonresponse Analyse im Kontext betrieblicher Mobilität an einer urbanen Hochschule beschreiben, sowie dann die Ergebnisse hinsichtlich ihrer Belastbarkeit eingeordnet.

Als empirischer Anwendungsfall dient das Projekt Campus Mobilität Rosenheim Traunstein (CaMoRoT) an der Technischen Hochschule Rosenheim. Dort werden verschiedene Maßnahmen des betrieblichen Mobilitätsmanagements implementiert, unter anderem ein Angebot für Bike Sharing, und mithilfe einer Ex-ante und einer zweiten Befragung und Nutzungsdaten evaluiert. Der Anwendungsfall ermöglicht damit zu untersuchen, welchen Beitrag die Kombination verschiedener Datenquellen zur belastbareren Einordnung von Befragungsergebnissen leisten kann.

2 Verwandte Arbeiten

2.1 Nonresponse (Bias)

Nonresponse bezeichnet den Ausfall von für eine Erhebung vorgesehenen Beobachtungen (vgl. Brick/Wiliams 2013). In den vergangenen Jahrzehnten wurden für zahlreiche Befragungsformen sinkende Teilnahmequoten dokumentiert (vgl. Daikeler et al. 2020; König et al. 2021) zeigen anhand des deutschen IAB-Betriebspanels für den Zeitraum 2001–2017 Veränderungen der Nonresponse-Raten und der Zusammensetzung der Nichtteilnehmenden. Parallel verändert die zunehmende Digitalisierung die Erhebungspraxis. Obwohl Online-Befragungen den Zugang zur Befragung erheblich vereinfachen und den zeitlichen sowie organisatorischen Aufwand der Teilnahme reduzieren, lässt sich daraus nicht zwangsläufig eine hohe Teilnahmebereitschaft ableiten (vgl. Daikeler et al. 2020). Als Ursachen von Nonresponse diskutiert die Literatur verschiedene Faktoren: Groves und Corning (2000) erklären Teilnahmeentscheidungen als Ergebnis einer individuellen Bewertung unterschiedlicher Eigenschaften einer Befragung. Andere Studien bestätigen dieses Zusammenspiel aus Erreichbarkeit, individuellen Merkmalen und Eigenschaften der Befragung: unter anderem spielen Alter und sozioökonomische Situation, aber auch Zeitaufwand, Befragungsmodus, Thema, Incentives und Kontaktstrategie insbesondere aber auch das Bildungsniveau eine Rolle (vgl. Distler et al. 2026; Zegras et al. 2018).

Von Nonresponse zu unterscheiden ist Nonresponse Bias. Dieser entsteht nicht bereits dadurch, dass Erhebungen fehlen, sondern wenn sich Respondierende und Nicht-respondierende hinsichtlich für die Untersuchung relevanter Merkmale systematisch unterscheiden (vgl. Groves/Petytcheva 2008). Damit ist eine geringe Response Rate zunächst lediglich ein Indikator für ein mögliches Risiko (vgl. Tversky/Kahneman 1971). Die grundlegende Meta-Analyse von Groves und Peytcheva (2008) zeigt, dass der Zusammenhang zwischen Nonresponse Rate und Nonresponse Bias stark zwischen Untersuchungen und Zielvariablen variiert: Eine niedrige Rücklaufquote muss nicht zwangsläufig zu verzerrten Ergebnissen führen, umgekehrt garantiert eine hohe Rücklaufquote keine unverzerrten Schätzungen. Auch Fuchs et al. (2013) zeigen anhand des Euro-

pean Social Survey, dass zusätzliche Kontaktversuche zwar die Response Rate erhöhen können, dies jedoch nicht automatisch mit einer entsprechenden Verbesserung der Datenqualität einhergeht. Entscheidend ist also vielmehr, ob die Teilnahmebereitschaft mit der jeweiligen Zielvariable zusammenhängt. Brick und Tourangeau (2017) betonen daher, dass Feldmaßnahmen nicht ausschließlich auf eine maximale Response Rate, sondern gezielt auf die Verringerung der Unterrepräsentation relevanter Gruppen ausgerichtet werden sollten. Um einen eventuellen Nonresponse Bias einzuschätzen, gibt es verschiedene Verfahren. Liegen Informationen über die gesamte Ausgangsstichprobe oder Population vor, können Respondierende und Nonrespondierende hinsichtlich bekannter Merkmale verglichen werden: Schouten et al. (2009) entwickelten hierfür sogenannte R-Indikatoren, welche die Repräsentativität der Response anhand von Response-Propensities quantifizieren. Daneben können Register- und Sampling-Frame-Daten sowie Paradaten, beispielsweise die Anzahl notwendiger Kontaktversuche, zur Identifikation selektiver Teilnahme genutzt werden (vgl. Olsen 2013). Eine weitere verbreitete Strategie ist der Vergleich früher und später Respondierender, wobei angenommen wird, dass schwer erreichbare bzw. spät antwortende Personen den Nonrespondierenden ähnlicher sind, diese Continuum of Resistance-Annahme ist allerdings nicht in jedem Kontext empirisch gesichert (vgl. Clarsen et al. 2021).

2.2 Evaluation im Mobilitätskontext

Im Kontext betrieblicher Mobilität werden häufig Mitarbeiterbefragungen eingesetzt, in denen Verkehrsmittelwahl, Pendelverhalten, Einstellungen oder Veränderungen zur Evaluation einer Intervention erfragt werden (vgl. Cairns et al. 2010; Van Malderen et al. 2012; Petrunoff et al. 2015). Methodisch stellt sich hier ein ähnliches Problem: Werden Wirkungen anhand freiwilliger Mitarbeiterbefragungen beurteilt, können Beschäftigte mit z.B. besonderem Interesse an nachhaltiger Mobilität eine höhere Teilnahmebereitschaft besitzen, was die berichtete Wirkung einer Mobilitätsmaßnahme verzerren könnte: Wittwer et al. (2024) identifizierten in einer Nachbefragung von Nichtteilnehmenden der Haushaltserhebung SrV (Mobilität in Städten) 2018 systematische Unterschiede in der Teilnahmewahrscheinlichkeit, insbesondere hinsichtlich des Bildungsniveaus. Trotzdem wurden die zentralen Ergebnisse zum Mobilitätsverhalten durch diese Überrepräsentation kaum verzerrt. Auch Zegras et al (2018) finden, dass eine systematische Nicht-Teilnahme an einer Haushalts-Mobilitätsbefragung vor allem vom Bildungsniveau abhängt.

2.3 Standardvorgehen zur Identifikation des Bias

Empirische Untersuchungen aus der Mobilitätsforschung zeigen jedoch, dass eine selektive Stichprobenzusammensetzung nicht zwangsläufig mit einer relevanten Verzerrung zentraler Mobilitätskennzahlen einhergeht. Um dies einzuordnen, nutzen Hubrich et al. (2025) ein mehrstufiges Vorgehen zur Analyse von Nonresponse und Stichprobenselektivität: Im ersten Schritt wird untersucht, ob bestimmte Bevölkerungsgruppen in der realisierten Stichprobe über- oder unterrepräsentiert sind. Hierzu werden die Stichprobenstrukturen anhand von Merkmalen wie Alter, Geschlecht oder Staatsangehörigkeit verglichen und Unterschiede in den Teilnahmewahrscheinlichkeiten untersucht. In der SrV (Mobilität in Städten) 2023 werden zudem designbedingte Unterschiede berücksichtigt, beispielsweise die höhere Auswahlwahrscheinlichkeit von Personen aus größeren Haushalten. Im zweiten Schritt wird geprüft, ob diese Selektivität auch für die untersuchten Mobilitätskennwerte relevant ist. Entscheidend ist, ob Nichtantwortende innerhalb vergleichbarer Gruppen beispielsweise andere Verkehrsmittel nutzen oder ein anderes Mobilitätsverhalten aufweisen als Teilnehmende. Frühere Nonresponse-Befragungen der SrV zeigten, dass solche zusätzlichen Unterschiede im Mobilitätsverhalten weitgehend vernachlässigt werden konnten und die strukturelle Selektivität durch Gewichtung ausgeglichen werden konnte. Das Vorgehen unterscheidet damit zwischen der Frage, wer an einer Befragung teilnimmt, und der Frage, ob Unterschiede in der Teilnahme tatsächlich die interessierenden Mobilitätsergebnisse beeinflussen. Die Belastbarkeit der Ergebnisse wird somit nicht allein anhand der Ausschöpfungsquote, sondern anhand der tatsächlichen Selektivität und ihrer Ergebnisrelevanz beurteilt.

Dieses Vorgehen dient im Folgenden als methodischer Referenzrahmen. Seine Übertragung auf die Evaluation betrieblicher Mobilitätsmaßnahmen ist jedoch nur eingeschränkt möglich. Für die Analyse

der strukturellen Selektivität werden zunächst Informationen über die Grundgesamtheit benötigt, um die Zusammensetzung der realisierten Stichprobe einordnen zu können. Hubrich et al. (2025) können hierfür auf Informationen aus dem Einwohnermelderegister zurückgreifen und Merkmale der gezogenen Personen auch unabhängig von ihrer späteren Teilnahme betrachten. Für die Prüfung, ob die festgestellte Selektivität auch für das Mobilitätsverhalten gegeben ist, wären Mobilitätsinformationen über die Nichtteilnehmenden erforderlich. In früheren SrV-Erhebungen wurden diese über gesonderte Nonresponse-Befragungen erhoben; für die SrV 2023 konnte auf die daraus gewonnenen Erkenntnisse zurückgegriffen werden. Bei betrieblichen Mobilitätsbefragungen stehen solche Informationen jedoch häufig nicht zur Verfügung. Damit stellt sich die methodische Herausforderung, wie beide Prüfschritte mit den verfügbaren Populations-, Befragungs- und ergänzenden Mobilitätsdaten möglichst belastbar abgebildet bzw. durch alternative Evidenz ergänzt werden können.

Die allgemeine Surveyforschung stellt eine Vielzahl diagnostischer Verfahren zur Untersuchung des Nonresponse Bias zur Verfügung und auch in der Mobilitätsforschung gibt es, insbesondere anhand von Haushaltserhebungen, entsprechende Ansätze. Jedoch wird die mögliche Verzerrung durch selektive Teilnahme bei der Evaluation betrieblicher Mobilitätsmaßnahmen an Hochschulen bislang deutlich weniger systematisch adressiert. Daraus ergibt sich die Forschungsfrage: Wie können Wirkungen betrieblicher Mobilitätsmaßnahmen an Hochschulen auch bei geringer Befragungsbeteiligung belastbar eingeordnet werden?

3 Methode: mehrstufige Validierung der Übertragbarkeit

Im Falle der Evaluation der betrieblichen Mobilitätsmaßnahmen, die an der TH Rosenheim seit 2025 umgesetzt worden sind, liegen keine Mobilitätsinformationen für Personen vor, die bereits an der ersten Befragung (Ex-ante) nicht teilgenommen haben. Ebenso existiert keine längere Reihe methodisch identischer Erhebungen, anhand derer die Stabilität der Selektivität geprüft werden könnte. Es liegen Befragungsdaten der Hochschulangehörigen ex ante vor, nach der Einführung der ersten Maßnahmen sowie Buchungsdaten des Bike-Sharing Angebots vor. Der Standardansatz wird dementsprechend folgendermaßen angepasst, wie in Tabelle 1 nachzulesen.

Tabelle 1: Methodenübersicht

Stufe	Prüfung	Einordnung
1. Stichprobenselektivität	Vergleich Befragungsstichprobe mit der gesamten TH Rosenheim	Standardvorgehen
2. Mobilitätsbezogene Selektivität	Vergleich Mobilitätsverhalten mit Verhaltensgruppe und MiD-Referenzdaten	Standardvorgehen mit Ergänzung
3. Panelattrition	Vergleich von Personen mit Teilnahme an beiden Wellen mit Ausfällen nach Ex-ante-Umfrage	Ergänzung
4. Objektive Validierung	Abgleich der Befragungsergebnisse mit objektiven Nutzungsdaten	Ergänzung

Quelle: Eigene Darstellung

3.1 Analyse der Stichprobenselektivität

Im ersten Schritt wird untersucht, ob die realisierte Befragungsstichprobe strukturell von der Referenzpopulation (alle Hochschulangehörige der TH Rosenheim) abweicht. Dabei werden die Teilnehmenden der Ex-ante Umfrage (n=260), die Teilnehmenden der zweiten Umfrage (n=125) sowie die Personen, die an beiden Befragungen (Panel) teilgenommen haben (n=15), getrennt anhand der verfügbaren Strukturinformationen verglichen. Anhand der verfügbaren soziodemografischen und hochschulbezogenen Merkmale wird geprüft, ob einzelne Gruppen gegenüber der Grundgesamtheit über- oder unterrepräsentiert sind. Da keine Individualdaten für die gesamte Grundgesamtheit vorliegen, dient der Vergleich mit den verfügbaren aggregierten Strukturinformationen der Hochschulangehörigen als pragmatische Annäherung, um mögliche strukturelle Selektivitäten der Befragungsstichprobe dennoch sichtbar zu machen.

3.2 Prüfung der mobilitätsbezogenen Selektivität

Der zweite Schritt betrifft die Frage, ob festgestellte strukturelle Selektivitäten auch für die untersuchten Mobilitätskennwerte relevant sind. Hubrich

et al. (2025) nutzen hierfür eine direkte Gegenüberstellung des Mobilitätsverhaltens von Teilnehmenden und Nichtteilnehmenden. Da für Hochschulangehörige ohne Teilnahme an den Befragungen keine Mobilitätsinformationen vorliegen, wird dieser Teil des Standardvorgehens durch einen externen Plausibilitätsvergleich ergänzt: Das Mobilitätsverhalten der Studierenden in der Ex-ante Befragungsstichprobe wird dazu mit einer vergleichbaren Gruppe verglichen. Außerdem werden weitere Mobilitätscharakteristika der gesamten Stichprobe, anhand von Altersklassen, vergleichbaren Referenzgruppen aus der Studie Mobilität in Deutschland 2023 (nachfolgend: MiD) gegenübergestellt (vgl. infas et al. 2025). Ziel ist zu prüfen, ob das in der Befragungsstichprobe beobachtete Mobilitätsverhalten gegenüber geeigneten externen Referenzen auffällige Abweichungen aufweist. Damit wird geprüft, ob die eventuell festgestellten Unterschiede für die Mobilitätsergebnisse überhaupt relevant sind.

3.3 Analyse der Panelattrition

Im nächsten Schritt wird die Selektivität zwischen den beiden Befragungswellen genauer analysiert. Hierfür werden Personen, die an beiden Umfragen teilgenommen haben, mit Personen verglichen, die an der Ex-ante, jedoch nicht an der zweiten Umfrage teilgenommen haben. Im Gegensatz zum initialen Nonresponse liegen für diese späteren Ausfälle bereits Informationen aus der ersten Befragungswelle vor. Dadurch können nicht nur soziodemografische Merkmale, sondern auch das bereits vor dem Ausfall beobachtete Mobilitätsverhalten verglichen werden. Damit wird geprüft, ob die Ergebnisse dadurch verzerrt sein könnten, dass bestimmte Personengruppen nicht mehr teilgenommen haben.

3.4 Abgleich mit objektiven Nutzungsdaten

Als weitere Evidenzquelle werden objektive Nutzungsdaten der Fahrräder aus dem Bike Sharing herangezogen. Dieser Schritt ermöglicht einen von der Befragung unabhängigen Abgleich des beobachteten Mobilitätsverhaltens. Dazu werden die in der Befragung ermittelten Nutzungsmuster und Veränderungen den tatsächlich registrierten Buchungsdaten gegenübergestellt und gruppenspezifische (statusbezogene) Nutzungsmuster betrachtet. Geprüft wird insbesondere, ob Befragungs- und Nutzungsdaten hinsichtlich Richtung und Struktur der beobachteten Veränderungen zu vergleichbaren Befunden führen. Damit wird geprüft ob sich die berichteten Nutzungsmuster aus der Befragung auch in der tatsächlichen Nutzung wiederfinden.

3.5 Individuelle Fixed-Effects

Zusätzlich wird die Panelstruktur der CaMoRoT-Umfragen genutzt, um die Wirkung der Mobilitätsmaßnahmen mittels Individual Fixed Effects zu untersuchen. Dabei werden Veränderungen innerhalb derselben Personen zwischen den beiden Befragungswellen betrachtet. Zeitinvariante individuelle Unterschiede, beispielsweise stabile Mobilitätspräferenzen oder andere über den Untersuchungszeitraum konstante persönliche Eigenschaften, werden dadurch kontrolliert (vgl. Angrist/Pischke 2009). Diese Analyse prüft, ob die beobachteten Veränderungen auch bei einer ausschließlich intraindividuellen Betrachtung bestehen bleiben, also ausschließlich durch die Einführung der Mobilitätsmaßnahmen herbeigeführt worden sind.

4 Ergebnisse

Die Belastbarkeit der ermittelten Wirkungen wird im Folgenden nicht anhand eines einzelnen Prüfkriteriums beurteilt, sondern aus der Gesamtschau der einzelnen Analyseschritte eingeordnet.

4.1 Selektivität der Stichprobe

Zuerst wird die generelle Selektivität der Beantwortenden beurteilt. Tabelle 2 berichtet die soziodemographische Zusammensetzung der Stichprobe, wobei Zahlen <10 aus datenschutzrechtlichen Gründen nicht berichtet werden, dies gilt insbesondere für Aussagen über das Panel. Die Zusammensetzung unterscheidet sich insbesondere hinsichtlich Alter und Status zwischen den Stichproben. Während in der Ex-ante Umfrage 52,3 % der Befragten der Altersgruppe 18–24 Jahre angehörten, lag dieser Anteil in Umfrage 2 bei 19,2 %. Auch hinsichtlich des Status zeigen sich Unterschiede: Der Anteil der Studierenden betrug in der Ex-ante Umfrage 61,5 % und in Umfrage 2 32,0 %. Die Ergebnisse zeigen also strukturelle Unterschiede zwischen beiden Befragungswellen. Eine Aussage über die Vergleichbarkeit mit der Referenzpopulation ist zum aktuellen Zeitpunkt aufgrund fehlender Vergleichsdaten hinsichtlich der betrachteten Merkmale nicht möglich.

Tabelle 2: Soziodemographische Beschreibung der Stichprobe

Merkmal	Kategorie	Ex-ante Umfrage	Umfrage 2	Beide
Stichprobengröße	n	260	125	15
Alter	18 - 24	126 (52,3 %)	24 (19,2 %)	0 (0,0 %)
	25 - 34	46 (19,1 %)	32 (25,6 %)	<10
	35 - 44	28 (11,6 %)	24 (19,2 %)	<10
	45 - 54	19 (7,9 %)	29 (23,2 %)	<10
	55 - 64	20 (8,3 %)	16 (12,8 %)	<10
	> 64	<10	0 (0,0 %)	0 (0,0 %)
	keine Angabe	<10	0 (0,0 %)	<10
Geschlecht	Weiblich	118 (49,0 %)	62 (49,6 %)	<10
	Männlich	119 (49,4 %)	61 (48,8 %)	<10
	Divers	<10	<10	<10
	Sonstige / keine Angabe	<10	<10	<10
Status	Studierende	160 (61,5 %)	40 (32,0 %)	0 (0,0 %)
	Professorinnen und Professoren	11 (4,2 %)	12 (9,6 %)	<10
	Lehrbeauftragte	<10	0 (0,0 %)	<10
	Wissenschaftliche Mitarbeitende	17 (6,5 %)	21 (16,8 %)	<10
	Nichtwissenschaftliche Mitarbeitende	50 (19,2 %)	50 (40.0 %)	<10
	Keine Angabe	20 (7.7 %)	<10	<10
International	International	19 (7,3 %)	<10	0 (0,0 %)
	Nicht international	220 (84,6 %)	114 (91,2 %)	13 (86,7 %)
	Keine Angabe	21 (8,1 %)	<10	<10
Hauptcampus	Rosenheim	224 (92,9 %)	113 (90,4 %)	11 (78,6 %)
	Mühldorf	<10	<10	0 (0,0 %)
	Traunstein	<10	<10	<10
	Burghausen	<10	<10	<10
	keine Angabe	<10	<10	0 (0,0 %)

Quelle**:** Eigene Darstellung

4.2 Ergebnisse mobilitätsbezogene Selektivität

Neben der generellen Selektivität wird untersucht, ob diese Selektivität sich auch auf die für diese Studie interessanten, also mobilitätsbezogenen Merkmale bezieht. Der Vergleich des Modal Splits der Studierenden aus der Ex-ante Umfrage mit einer Vergleichsgruppe die aus der MiD (gefiltert auf Studierende auf deren Weg zur Hochschule) hervorgeht zeigt deutliche Unterschiede (siehe Abbildung 1). Der Individualverkehr liegt in der vorliegenden Gruppe mit 38 % deutlich höher als der Referenzwert von 16 %, gleiches zeigt sich im Fahrradverkehr, mit 32 % gegenüber 21% in der Referenzpopulation. Der ÖPNV-Anteil ist mit 24 % deutlich niedriger als der Referenzwert von 45 %. Der Fußverkehr macht mit 6% den niedrigsten Anteil aus, während dieser in der Referenzgruppe bei 18 % liegt.

Vergleicht man die gesamte Menge der Ex-ante Beantwortenden (also auch die Mitarbeitenden) mit den Ergebnissen der MiD, so findet sich folgendes Bild: Beim Modal Split liegen die Ex-ante Befragten beim MIV bei 37,4 % (MiD: 50 %), beim ÖPNV bei 24,8 % (MiD: 8 %), beim Fahrrad bei 31,7 % (MiD: 13 %) und beim Fußverkehr bei 6,2 % (MiD: 28 %).

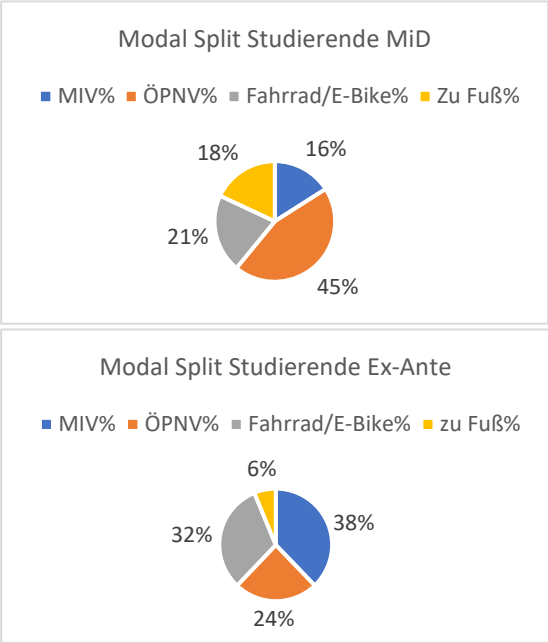

Abbildung 1: Modal Split Vergleich (Eigene Abbildung)

Die Pkw-Verfügbarkeit liegt in der Ex-ante Umfrage bei 71,7 % und in Umfrage 2 bei 83,2 %, gegenüber 88 % in der MiD. Der Führerscheinbesitz beträgt bei der Ex-ante Umfrage 95 % gegenüber 85 % in der MiD. Die Ex-ante Stichprobe unterscheidet sich also hinsichtlich mehrerer mobilitätsbezogener Merkmale, was bedeutet, dass die festgestellte strukturelle Selektivität somit mit Abweichungen bei relevanten Mobilitätsmerkmalen einhergeht.

4.3 Ergebnisse Panelattrition

Während die gesamte Stichprobe bereits auf Selektivität eingeschätzt worden ist, wird nun das Panel, also die Beantwortenden beider Umfragen, auf Selektivität geprüft. Studierende machten 61,5 % der ausschließlich an der Ex-ante Umfrage Teilnehmenden aus, waren im Panel jedoch nicht vertreten. Ein interessanter Punkt kann hier herausgegriffen werden: in der Ex-ante Umfrage und in der Umfrage 2 nehmen sowohl bei den Professorinnen und Professoren als auch bei den nichtwissenschaftlichen Mitarbeitenden ungefähr gleich viele Personen teil, die Zahl der Panelpersonen ist jedoch deutlich kleiner. Dies weist auf eine beständige, aber nicht wiederholte Teilnahme hin.

Bei den mobilitätsbezogenen Merkmalen fallen die Unterschiede zwischen der Ex-ante-Umfrage und dem Panel insgesamt weniger einheitlich aus. Die Anteile der Autofahrenden waren mit 2,7 pp Unterschied vergleichbar. Im Panel zeigte sich jedoch ein um 5,4 pp geringerer durchschnittlicher MIV-Anteil und ein um 9,1 pp höherer ÖPNV-Anteil. Die durchschnittliche Entfernung zur Hochschule sowie der Fahrrad-/E-Bike-Anteil unterschieden sich nur gering. Weitere Aussagen über das Panel können aufgrund einer zu geringen Fallzahl nicht dargestellt werden. Insgesamt bestehen damit Hinweise auf selektive Panelattrition, deren mögliche Auswirkungen auf die gemessenen Veränderungen aufgrund der kleinen Panelgröße jedoch vorsichtig einzuordnen sind.

4.4 Objektives Nutzungsverhalten

Um zu untersuchen, wie plausibel es ist, dass die Ergebnisse der Studie belastbar sein könnten, wird zusätzlich eingeordnet, ob das berichtete Mobilitätsverhalten auch dem Tatsächlichen entspricht. Insgesamt zeigen die Buchungsdaten, dass Bike Sharing nach den ersten drei Monaten bereits von 3,1 % der Mitarbeitenden und 2,7 % der Studierenden, das entspricht 3 % aller Hochschulangehörigen, genutzt wird. Gleichzeitig bedeutet das einen strukturellen Unterschied zu den Befragungsteilnehmenden, bei denen der Anteil der Nutzenden bereits bei 13 % liegt: Umfrage 2 zeigt, dass 9,4 % der teilnehmenden Mitarbeitenden und 12,5 % der Studierenden berichten, das Bike Sharing gelegentlich oder selten zu nutzen.

4.5 Individuelle Fixed-Effects

Die geplante Fixed-Effects-Analyse konnte aufgrund der geringen Fallzahl von lediglich n = 15 Personen mit vollständigen Beobachtungen in beiden Befragungswellen nicht belastbar durchgeführt werden. Eine Schätzung auf dieser Grundlage wäre mit hoher statistischer Unsicherheit verbunden und würde keine verlässlichen Aussagen über intraindividuelle Veränderungen ermöglichen.

5 Zusammenfassung und Diskussion

Diese Studie setzt sich anhand verschiedener Perspektiven damit auseinander, ob betriebliche Mobilitätsmaßnahmen belastbar evaluiert werden können, wenn es an Beantwortungen fehlt. Da die Ergebnisrelevanz des initialen Nonresponse aufgrund fehlender Informationen über Nichtteilnehmende nicht direkt geprüft werden kann, wird diese Analyse durch externe Mobilitätsreferenzen, die Untersuchung der Panelattrition und objektive Nutzungsdaten ergänzt. Eine geplante Fixed-Effects-

Analyse, die eine zusätzliche Robustheitsprüfung der beobachteten intraindividuellen Veränderungen ermöglichen sollte, konnte aufgrund geringer Fallzahl nicht durchgeführt werden.

Die Studie findet mithilfe mehrerer Dimensionen das differenzierte Bild der Selektivität der vorliegenden Stichprobe: Erstens zeigt die generelle Selektivitätsanalyse eine soziodemografische Selektivität hinsichtlich Alter und Status. Zweitens zeigt die mobilitätsbezogene Analyse Selektivität hinsichtlich dieser Faktoren. Drittens finden sich Hinweise auf eine strukturelle Panelattrition und viertens stützt der objektive Abgleich des Nutzungsverhaltens die Befragungsergebnisse in ihrer grundlegenden Tendenz, nicht jedoch in ihrer exakten Größenordnung. Insgesamt häufen sich also die Hinweise auf eine mehrdimensionale Selektivität. Zur Einordnung der Studienergebnisse sind also weitere Maßnahmen nötig, zum Beispiel ein Vergleich der Wirkungseinschätzungen mit und ohne entsprechender Gewichtung der Personengruppen entsprechend der Ergebnisse der Selektivitätsanalyse.

Eine wichtige Limitation dieser Einordnung ist die geringe Stichprobe des Panels. Trotzdem zeigt die Methodik, dass sich eine potenzielle Selektivität auf einzelne Charakteristika eingrenzen lässt. Künftige Untersuchungen könnten sich weiter auf die Isolierung dieser Merkmale konzentrieren oder sie entsprechend antizipieren, z.B. durch eine entsprechende sachliche Gewichtungslösung der Stichprobe. Alternativ braucht es ein differenzierteres Bild der Zusammensetzung der Hochschule oder eine geeignete Nonresponse-Befragung, um die Charakteristika, nach denen sich die Stichprobe selektiert, einordnen zu können. Darauf aufbauend könnte zukünftig außerdem untersucht werden, ob alternative Evaluationsmethoden in diesem Kontext entwickelt werden können, die ein geeigneteres Bild der tatsächlichen Wirkung betrieblicher Mobilitätsmaßnahmen im Hochschul-Kontext abgeben.

Literaturverzeichnis

Angrist, Joshua D./Jörn-Steffen Pischke (2009): Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press. https://doi.org/10.1515/9781400829828

Brick, J. Michael/Douglas Williams (2013): Explaining rising nonresponse rates in cross-sectional surveys, in: The ANNALS of the American Academy of Political and Social Science, Bd. 645, Nr. 1, S. 36–59, [online] https://journals.sagepub.com/doi/full/10.1177/0002716212456834 .

Brick, J. Michael/Roger Tourangeau (2017): Responsive Survey Designs for Reducing Nonresponse Bias, in: Journal of Official Statistics, Bd. 33, Nr. 3, S. 735–752, [online] https://doi.org/10.1515/jos-2017-0034 .

Cairns, Sally/Carey Newson/Adrian Davis (2010): Understanding successful workplace travel initiatives in the UK, in: Transportation Research Part A: Policy and Practice, Bd. 44, Nr. 7, S. 473–494, [online] https://doi.org/10.1016/j.tra.2010.03.010

Clarsen, Benjamin/Jens Christoffer Skogen/Thomas Sevenius Nilsen/Leif Edvard Aarø (2021): Revisiting the continuum of resistance model in the digital age: A comparison of early and delayed respondents to the Norwegian counties public health survey, in: BMC Public Health, Bd. 21, Nr. 1, Artikel 730, [online] https://doi.org/10.1186/s12889-021-10764-2

Daikeler, Jessica/Michael Bošnjak/Katja Lozar Manfreda (2020): Web versus other survey modes: An updated and extended meta-analysis comparing response rates, in: Journal of Survey Statistics and Methodology, Bd. 8, Nr. 3, S. 513–539, [online] https://doi.org/10.1093/jssam/smz008.

Distler, Christoph/Mustafa Coban/Mark Trappmann (2026): Nonresponse in web surveys: A deeper look into education bias, in: Journal of Survey Statistics and Methodology, smag031, [online] https://doi.org/10.1093/jssam/smag031.

Fuchs, Marek/Daniela Bossert/Sabrina Stukowski (2013): Response rate and nonresponse bias – Impact of the number of contact attempts on data quality in the European Social Survey, in: Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique, Bd. 117, Nr. 1, S. 26–45, [online] https://doi.org/10.1177/0759106312465547 .

Groves, Robert M. (2006): Nonresponse rates and nonresponse bias in household surveys, in: Public Opinion Quarterly, Bd. 70, Nr. 5, S. 646–675, [online] https://doi.org/10.1093/poq/nfl033 .

Groves, Robert M./Emilia Peytcheva (2008): The impact of nonresponse rates on nonresponse bias: A meta-analysis, in: Public Opinion Quarterly, Bd. 72, Nr. 2, S. 167–189, [online] https://doi.org/10.1093/poq/nfn011.

Groves, Robert M./Eleanor Singer/Amy Corning (2000): Leverage-saliency theory of survey participation: Description and an illustration, in: Public Opinion Quarterly, Bd. 64, Nr. 3, S. 299–308, [online] https://doi.org/10.1086/317990

Hubrich, Stefan/Rico Wittwer/Frank Ließke/Sebastian Wittig/Fabian Schönherr/Regine Gerike (2025): Methodenbericht zum Forschungsprojekt „Mobilität in Städten – SrV 2023", Dresden: Technische Universität Dresden.

infas, DLR, IVT und infas 360 (2025): Mobilität in Deutschland 2023 (im Auftrag des BMV) [Datensatz]. Bereitgestellt durch den BASt MobilityData-Campus. Kontakt: mobilitydatacampus@bast.de. https://www.bast.de/mdc [abgerufen am 01.09.2026].

Koch, Achim/Michael Blohm (2016): Nonresponse Bias, in: GESIS Survey Guidelines, Mannheim: GESIS – Leibniz-Institut für Sozialwissenschaften, [online] https://doi.org/10.15465/gesis-sg_en_004

König, Corinna/Joseph W. Sakshaug/Jens Stegmaier/Susanne Kohaut (2021): Trends in establishment survey nonresponse rates and nonresponse bias: Evidence from the 2001–2017 IAB Establishment Panel, in: Journal of Official Statistics, Bd. 37, Nr. 4, S. 931–953, [online] https://doi.org/10.2478/jos-2021-0040 .

Olson, Kristen (2013): Paradata for nonresponse adjustment, in: The ANNALS of the American Academy of Political and Social Science, Bd. 645, Nr. 1, S. 142–170, [online] https://doi.org/10.1177/0002716212459475 .

Peer, Eyal/David Rothschild/Andrew Gordon/Zak Evernden/Ekaterina Damer (2022): Data quality of platforms and panels for online behavioral research,

in: Behavior Research Methods, Bd. 54, Nr. 4, S. 1643–1662, [online] https://doi.org/10.3758/s13428-021-01694-3.

Petrunoff, Nick/Chris Rissel/Li Ming Wen/Julie Martin (2015): Carrots and sticks vs carrots: Comparing approaches to workplace travel plans using disincentives for driving and incentives for active travel, in: Journal of Transport & Health, Bd. 2, Nr. 4, S. 563–567, [online] https://doi.org/10.1016/j.jth.2015.06.007

Sadeghian, Peyman/Johan Håkansson/Xiaoyun Zhao (2021): Review and evaluation of methods in transport mode detection based on GPS tracking data, in: Journal of Traffic and Transportation Engineering (English Edition), Bd. 8, Nr. 4, S. 467–482, [online] https://doi.org/10.1016/j.jtte.2021.04.004

Schouten, Barry/Fannie Cobben/Jelke Bethlehem (2009): Indicators for the representativeness of survey response, in: Survey Methodology, Bd. 35, Nr. 1, S. 5–18, [online] https://www.researchgate.net/profile/Jelke-Bethlehem/publication/267836796_Indicators_for_the_Representativeness_of_Survey_Response/links/547458fc0cf245eb436dd8ae/Indicators-for-the-Representativeness-of-Survey-Response.pdf?_sg%5B0%5D=started_experiment_milestone&origin=journalDetail&_rtd=e30%3D .

Su, Tao/Michael E. Renda/Jinhua Zhao (2022): Examining the discrepancies between self-reported and actual commuting behavior at the individual level, in: Transportation Research Record, Bd. 2676, Nr. 2, S. 118–131, [online] https://doi.org/10.1177/03611981211037240 .

Tversky, Amos/Daniel Kahneman (1971): Belief in the law of small numbers, in: Psychological Bulletin, Bd. 76, Nr. 2, S. 105–110, [online] https://doi.org/10.1037/h0031322.

Van Malderen, Laurent/Bart Jourquin/Isabelle Thomas (2012): Employer transport plan: Do they change the commuting behaviour of employees?, in: 52nd European Congress of the European Regional Science Association (ERSA).

Autor*innenangaben

Franziska Danner
Technische Hochschule Rosenheim am Campus Chiemgau
Stadtplatz 32
83278 Traunstein

E-Mail: franziska.danner@th-rosenheim.de

Elisabeth Soller
Technische Hochschule Rosenheim am Campus Chiemgau
Stadtplatz 32
83278 Traunstein

E-Mail: elisabeth.soller@th-rosenheim.de

Wibke Michalk
Technische Hochschule Rosenheim am Campus Chiemgau
Stadtplatz 32
83278 Traunstein

E-Mail: wibke.michalk@th-rosenheim.de

Journal für Mobilität und Verkehr

ISSN 2628-4154

www.dvwg.de

Beifang als Best-Practice? Mit Routinedaten der Datenarmut ländlicher Räume begegnen

Jonathan Gescher, Tim Leibert, Wladimir Sgibnev, Anne Mehl

Siehe Autor*innenangaben

Abstract

Analysen der Mobilitätsmuster in ländlichen Räumen stellen angesichts ländlicher Datenarmut eine Herausforderung für die Forschung dar. Die Nutzung von Routinedaten kann hier Abhilfe schaffen und einen Schritt in eine stärker evidenzbasierte Debatte darstellen. In diesem Artikel stellen wir exemplarisch Arbeitsmarkts-, Mobilfunkverbindungs- und Krankenkassenabrechnungsdaten vor und diskutieren Einschränkungen und Potenziale der Nutzung solcher Routinedaten für die Mobilitätsforschung sowie praktische Herangehensweisen, um solche Datensätze nutzbar zu machen.

Schlagwörter:

Routinedaten, Ländliche Räume, Mobilfunkdaten, Arbeitsmarktdaten, Krankenkassenabrechnungsdaten

1 Ländliche Mobilitätsdaten-Armut

Menschen im „ländlichen Raum“ fahren Auto, sind infrastrukturell unterversorgt und müssen weit pendeln. Wer nach Klischees zu ländlicher Mobilität sucht, wird schnell fündig – oft pointiert und häufig gestützt auf anekdotische Evidenz oder einzelne Fallstudien. Dass die kritische Einordnung dabei teils auf der Strecke bleibt, liegt nicht nur an der hohen Diversität ländlicher Räume, sondern auch an einer übergreifenden Gemeinsamkeit – es mangelt schlicht an Daten.

Mobilitätserhebungen sind komplex und zwangsläufig lückenhaft. Während einige dieser Lücken methodisch bedingt sind, gibt es zusätzlich eine systematische Verzerrung des (Daten-)Wissens zu Alltagsmobilität – es konzentriert sich auf die Zentren. Größere Städte können etwa eigene Mobilitätserhebungen oder Sonderauswertungen übergreifender Studien durchführen. Hier können Daten auf Kreisebene verwendet werden und der Datenschutz, etwa bei Zensusdaten, schränkt die Verfügbarkeit von Gemeindedaten weniger stark ein.Auch Erhebungsstellen für Flussdaten, beispielsweise Fahrradzählstellen, finden sich überwiegend in urbanen Zentren. Ver-stärkt wird die Diskrepanz durch ein Ost-West-Gefälle bei der Datenversorgung: nur drei ostdeutsche Bundesländer (Brandenburg, Thüringen, Mecklenburg-Vorpommern) bestellen regionale Vertiefungsmodule bei der Erhebung zur „Mobilität in Deutschland“. Darüber hinaus ist kein einziges Vertiefungsmodul auf Kreis-, Kommunal- oder Verkehrsverbundsebene (von insgesamt etwa 60!) in Ostdeutschland angesiedelt (MID 2023). Die SrV-Umfragen der TU Dresden decken die neuen Bundesländer zwar deutlich besser ab, lassen allerdings ländliche Räume außen vor (SRV 2023).

Angelehnt an das Konzept der Digitalen Spaltung (Kersting 2025) können wir im Falle von Mobilitäts-daten von einer regionalisierten Daten-Spaltung sprechen. Die resultierende ländliche Mobilitätsdaten-Armut hat, unserer Einschätzung nach, realweltliche Auswirkungen, etwa angesichts metropolenzentrierter Debatten zur Mobilitätswende oder der Qualität politischer und planerischer Entscheidungen zur Alltagsmobilität in ländlichen Räumen.

Alltagsmobilität bezeichnet dabei die räumliche Bewegung von Individuen, die, im Gegensatz zur dauerhaften Ortsveränderung (wie einem Umzug), meist regelmäßig stattfindet. Der Begriff ‚Mobilität' lässt sich dabei als multidimensionaler Begriff auffassen, der sowohl die zur Verfügung stehenden Mobilitäts-optionen (wie die Verfügbarkeit von Verkehrsmitteln) als auch tatsächlich realisiertes Mobilitätsverhalten (wie Wegelängen) beinhaltet (Forschungsgesellschaft für Straßen- und Verkehrswesen 2020, S. 25). Während andere Autor:innen unter Mobilität teilweise nur die Handlungsmöglichkeit zur Bewegung verstehen (z. B. Herget 2016), nutzen wir somit ein weiteres Verständnis der Begrifflichkeit. Darüber hinaus begreifen wir Mobilität als sozial und kulturell konstituierte Lebenspraxis (Wilde 2014, S. 174), deren Erforschung somit neben der Datenperspektive immer auch die Nutzendensicht beinhalten sollte.

Im Gegensatz zu Flussdaten lassen sich Verflechtungsdaten, also Daten mit einem Start- und Zielort, hingegen nicht einfach „zählen". Von Studien mit Standorttrackern abgesehen (e.g. van Dülmen 2022), sind Forschende wie Planende hier oftmals auf Sekundär- oder Routinedaten angewiesen, also Daten, die routinemäßig für andere Zwecke erfasst werden. Eine bevorzugte Nutzung solcher Routinedaten – anstatt einer Vermehrung grundständiger Umfragen – entspricht auch dem Grundsatz der Datenminimierung der EU-Datenschutz-Grundverordnung. In diesem Artikel widmen wir uns daher auf den ersten Blick mobilitätsunabhängigen Datenquellen, aus denen wir dennoch wertvolle Schlüsse für die Alltags-mobilität in ländlichen Räumen ziehen können: Arbeitsmarkts-, Mobilfunkverbindungs- und Krankenkassenabrechnungsdaten. Die Ergebnisse stammen aus einem vom Freistaat Sachsen geförderten Pilot-projekt zur Nachhaltigen Mobilität in periphrasierten Regionen (Peri-Mobil[1]). Anhand exemplarischer Anwendungen diskutieren wir im Folgenden die Einschränkungen und Potenziale der Nutzung solcher Routinedaten für die Mobilitätsforschung sowie praktische Herangehensweisen, um solche Datensätze nutzbar zu machen.

2 Pendeldaten

Pendeldaten sind die am breitesten verfügbare, flächendeckende Mobilitätsdatenquelle in Deutschland. Sie fließen als Verflechtungsdaten in zahlreiche Analysen des Verkehrs und Mobilitätsgeschehens ein (z.B. Krebs und Pflüger 2023; Oberst 2020). Methodisch bedingt haben diese Daten jedoch zahlreiche Einschränkungen.

Pendeldaten sind über die Bundesagentur für Arbeit verfügbar (Bundesagentur für Arbeit 2025) und bieten auf Gemeindeebene jährlich aktualisierte Informationen. Sie basieren auf einer Vollerhebung aller sozialversicherungspflichtig Beschäftigten und sind eine Sekundärstatistik, d.h. die Daten werden nicht primär zur Erfassung von Pendler-Mobilität erhoben. Stattdessen gilt als „Pendeln", wenn der gemeldete Wohnort und der gemeldete Arbeitsort in unterschiedlichen Gemeinden liegen (Hadam 2021). Die Verflechtungsmatrix wird differenziert nach Geschlecht bereitgestellt, was detailliertere Auswertungen des Pendelverhaltens ermöglicht. Die Daten sind damit besonders nützlich, um Trends in der Arbeits-mobilität zu erkennen und regionale Unterschiede zu analysieren.

Ein Nachteil der Pendlerdaten liegt darin, dass sie aufgrund der Erfassungslogik potenzielle statt tatsächlicher Mobilität erfassen (Information und Technik Nordrhein-Westfalen (IT.NRW) 2023). Dabei bleibt außen vor, ob, wie oft und von wo die Beschäftigten den Arbeitsort aufsuchen (Homeoffice, Teilzeit, Zweitwohnungen, Beschäftigung in einer Niederlassung) (Fuchs et al. 2025). Zusätzlich sind Selbstständige, Beamtinnen und Beamte, Studierende und andere nicht sozialversicherungspflichtig Beschäftigte in dieser Statistik nicht enthalten. Das spiegelt sich in der Diskrepanz zwischen durchschnittlichen Pendel-wegen in den Pendlerdaten und der MID-Befragung wider (Nitschke et al. 2022). Fehlende Routen und Zeiten der Pendelbewegungen

[1] https://leibniz-ifl.de/forschung/projekt/nachhaltige-mobilitaet-in-peripherisierten-regionen-perimobil

schränken die Nutz-barkeit für Mobilitätsanalysen zusätzlich ein.

In Ermangelung von Alternativen werden Pendlerdaten dennoch zur Analyse von Alltagsmobilitätsmustern eingesetzt. Eine Überschneidung mit Mustern der Bildungs-, Gesundheits- und Freizeitmobilität ist dabei nur bedingt gegeben.

3 Mobilfunkdaten

Eine mögliche Ergänzung bieten Mobilfunkdaten. Diese haben als Mobilitätsdatenquelle insbesondere während der Covid-19 Pandemie an Aufmerksamkeit gewonnen (Szocska et al. 2021; Alarcon et al. 2021). In der Pandemie lag der Fokus besonders auf der Anzahl der Bewegungen; Mobilfunkdaten sind jedoch auch eine vielversprechende Datenquelle für die Erforschung von Mobilitätsmustern (Schmidt und Männel 2017). Für unsere Auswertungen haben wir eine Lizenz der Firma Teralytics für Sachsen, Sachsen-Anhalt und Thüringen erworben, da Mobilfunkdaten in Deutschland nicht frei verfügbar sind. Die Daten basieren auf anonymisierten Mobilfunkverbindungen von Telefonica-Kund:innen, welche ca. ein Drittel des deutschen Mobilfunkmarktes ausmachen. Die bereitgestellten Daten werden von Teralytics auf die Gesamtbevölkerung hochgerechnet. Dabei berücksichtigt das Unternehmen regionale Marktanteile des Anbieters, um eine Annäherung an tatsächliche Bewegungen zu erreichen (Statistisches Bundesamt (Destatis) 2023; Teralytics GmbH 2022).

Bewegungen basieren auf sogenannten Call Detail Records (CDR). Dabei wird registriert, wenn sich eine SIM-Karte zwischen Funktürmen bewegt und mit ihnen in Interaktion tritt (mobile Daten, Anrufe, SMS). Hält sich die SIM-Karte mindestens 30 Minuten in einer Funkzelle auf, gilt die Bewegung als abgeschlossen. Diese Bewegungen werden in Gemeindegrenzen und kleinere Gebietseinheiten umgerechnet, wobei es aufgrund variierender Größen der Funkzellen zwangsläufig zu kleinräumigen Ungenauigkeiten kommt (Schmidt und Männel, 2017). Die Daten enthalten keine Strecken oder sozio-demographische Merkmale der Personen. Eine Einsicht in die Rohdaten oder den genauen Prozess der Datenaufbereitung ist nicht möglich. Dies ist zu Forschungszwecken aus mehreren Gründen problematisch, unter anderem aufgrund der eingeschränkten Vergleichbarkeit über mehrere Jahre. Wechsel in der Methodik können genauso wenig nachvollzogen werden wie die Qualität der Hochrechnungen. Einen Überblick über diese und weitere Einschränkungen geben Bock und Schönduwe (2021). Seit Ende 2023 stellt Teralytics regelmäßige Regionalreports zur Datenqualität bereit (Teralytics Gmbh 2023), die eine unabhängige Validierung der Daten jedoch nicht ersetzen können.

Wir haben eine einfache Plausibilisierung der Daten anhand der Pendeldaten durchgeführt. Dafür wurde ein Korrelationsmodell zwischen Teralytics-Bewegungen für Arbeitstage im Juni 2023 und offiziellen Pendlerdaten für 2022 berechnet (Tabelle 1). Die Korrelationen sind insgesamt hoch, sodass Mobilfunkverflechtungen sowohl für periphere als auch zentrale Regionen plausibel erscheinen. Da Pendlerdaten zahlreiche Einschränkungen haben und keine Daten für den identischen Zeitraum verfügbar waren, können die bestehenden Unterschiede durch beide Datenquellen erklärt werden.

Tabelle 1: Korrelationsergebnisse zwischen Teralytics-Daten an Werktagen im Juni 2023 und Pendlerdaten 2022 für Gemeinden in Sachsen und Sachsen-Anhalt, eigene Berechnungen basierend auf Teralytics 2023, Bundesagentur für Arbeit 2025

Pearson-Korrelationskoeffizient für:		
Gebietstyp	Eingehende Mobilität	Ausgehende Mobilität
Sehr zentral	0,94	0,82
Zentral	0,86	0,91
Peripher	0,78	0,79
Sehr peripher	0,81	0,86

Ein großer Vorteil von Mobilfunkdaten ist der hohe Detaillierungsgrad. Um sie für die Identifikation regionaler Mobilitätsmuster oder die Kommunikation mit Bewohner:innen einzusetzen, muss jedoch zuerst die Komplexität reduziert werden. Dies ist mittels eines von Teralytics bereitgestellten Dashboards möglich, welches den Nutzenden erlaubt Daten abzubilden und nach Zeit und Tag zu filtern. Dies kommt jedoch mit mehreren Einschränkungen: Neben kleineren Problemen mit der Geschwindigkeit und der Auswahl der Daten, ist die Bewegungsvisualisierung (Stand Anfang 2024) ein Hindernis. Bewegungen werden nur als Pfeildarstellungen abgebildet, deren Lesbarkeit eingeschränkt ist und deren Skalierung sich nicht über mehrere Zeiträume fixieren lässt, d.h. die Pfeildicke ist für den Maximalwert an Bewegungen in einem Zeitraum immer identisch, unabhängig davon, ob es sich um 10 oder um 1.000

handelt. Diese Form der Darstellung verhindert eine schnelle visuelle Erfassung von Mustern über den Zeitverlauf. Wir haben uns daher für eine eigene Visualisierung der innertäglichen Bewegungsmuster anhand zweier Fallstudiengemeinden entschieden. Um Unterschiede in Bewegungsumfang- und zielen sichtbar zu machen, erstellen wir dazu zwei- und dreidimensionale Karten. Diese bilden für jede Stunde eines durchschnittlichen Monats an Arbeitstagen sowie Wochenenden die Bewegungen ab.

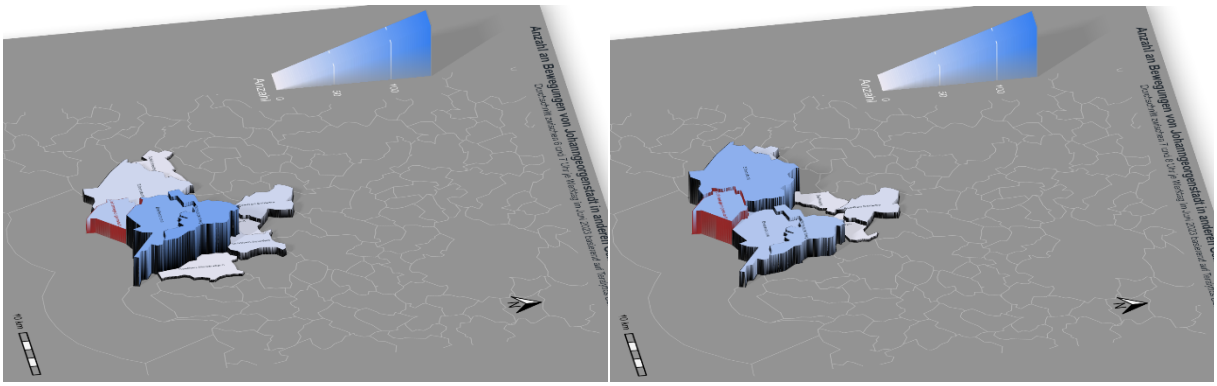

Abbildung 1: Karte der Mobilfunkbewegungen innerhalb Johanngeorgenstadts und aus Johanngeorgenstadt in andere Gemeinden, startend zwischen 6:00 und 7:00 Uhr (oben) und 7:00 und 8:00 Uhr (unten) an Werktagen im Juni 2023, eigene Darstel-lung basierend auf Teralytics 2023

Abbildung 1 zeigt exemplarisch am Beispiel der Stadt Johanngeorgenstadt im Erzgebirgskreis – eine der beiden Fallstudien im Projekt PeriMobil – das Potenzial von Mobilfunkdaten besonders in zeitlich hoch aufgelösten Anwendungsfällen. Dargestellt sind Bewegungen innerhalb der Stadt und aus Johanngeorgenstadt hinaus, die zwischen 6 Uhr und 7 Uhr sowie 7 Uhr und 8 Uhr an Werktagen starten. Dabei fällt auf, dass ab 6 Uhr die östlichen Nachbargemeinden Breitenbrunn und Schwarzenberg die Hauptzielorte sind, während ab 7 Uhr das westlich gelegene Eibenstock eine wichtigere Rolle einnimmt.

Eine Schwäche der Mobilfunkdaten (wie vieler weiterer Datenquellen) ist, dass sie keine Informationen über den Zweck einer Bewegung oder weitere Indikatoren (Alter, Beruf, etc.) enthalten. Muster können so zwar identifiziert aber nur schwer erklärt werden. Die Verknüpfung mit qualitativen Bausteinen kann hier einen großen Erkenntnisgewinn bringen. So können beispielsweise Kindergarten- und Schulschlusszeiten eindeutig in den Bewegungsmustern erkannt und von Bewohner:innen kontextualisiert werden.

Die Auffälligkeiten in den morgendlichen Bewegungen erklären sich demnach durch die Infrastruktur in- und zwischen den Gemeinden. In Breitenbrunn liegt die Klinik Erlabrunn, die im Schichtbetrieb arbeitet - einer der größten Arbeitgeber in der Region. Schwarzenberg ist nicht nur ein Ziel für Arbeitnehmer:innen, sondern auch einer der nächsten weiterführenden Schulstandorte. Aufgrund der Anbindung müssen Schüler:innen aus Johanngeorgenstadt schon vor 7 Uhr mit öffentlichen Verkehrsmitteln aufbrechen, während der Bus zur Oberschule im östlich angrenzenden Eibenstock Johanngeorgenstadt nach 7 Uhr verlässt.

Die Verknüpfung der Mobilfunkmuster mit qualitativen Methoden war darüber hinaus besonders hilfreich für die Erklärung nichtrealisierter Bewegungen, die andernfalls in den Gesprächen nicht aufgekommen wären. In Mobilfunkdaten konnten deutlich asymmetrische Bewegungsmuster beobachtet werden, welche scheinbar mit der Lage unserer Untersuchungsregionen an (innerdeutschen) Grenzen zusammenhängen. Obwohl keine physischen Barrieren durch die Grenzen bestehen, zeigt sich in Arzberg, der zweiten Fallstudie an der brandenburgischen Grenze, ein deutlich Sachsen-bezogenes, in Johanngeorgenstadt ein fast ausschließlich Sachsen/Deutschland-bezogenes Bewegungsprofil. In Diskussionsrunden wurde deutlich, dass zum Grenzübertritt meist ein spezifischer Anlass erforderlich

ist, welcher nicht im eigenen (Bundes-)Land erfüllt werden kann.

Distanzanalyse

Neben der Analyse konkreter Verflechtungen können mittels Mobilfunkdaten auch übergeordnete Bewegungsmuster identifiziert werden. Mobilfunkdaten enthalten die Luftliniendistanz zwischen den kleinräumigen Gebietseinheiten. Dies erlaubt ein Bild davon, wie viele Bewegungen im Nahraum, über mittlere Distanzen und weite Strecken zurückgelegt wer-den. Abbildung 2 zeigt die kumulierten Prozentwerte nach Distanz und Raumtyp für Sachsen. In den Mobilfunkdaten ist dabei ein deutlicher Rückgang der zu-rückgelegten Entfernungen mit zunehmender Zentralität der Gemeinden sichtbar. Die Ergebnisse sind dabei jedoch besonders auf sehr kurze Distanzen mit Vorsicht zu interpretieren, da die Zahl der Bewegungen stark von der Anzahl und Größe der Funkzellen in einem Gebiet abhängt. So können Bewegungen in einem Gebiet innerhalb einer Funkzelle erfolgen, während dieselbe Distanz in einem anderen Gebiet mit einem Funkturmwechsel verbunden ist (Modifiable Areal Unit Problem, siehe z.B. Williams et al. 2015). Unabhängig von der genauen Zahl, gibt die Distanzauswertung dennoch einen plausiblen Eindruck der Bewegungsmuster in verschiedenen Raumtypen wieder, deren signifikante Unterschiede über längere Distanzen sich nicht allein methodisch erklären lassen.

Während in sehr zentralen Kommunen 77% aller Bewegungen innerhalb einer Luftliniendistanz von unter 5 km erfolgen, sind es in sehr peripheren Gebieten nur 42%. Sachsenweit entfallen 90% aller Mobilfunkbewegungen auf Strecken bis 15 km, in sehr zentralen Gebieten sind es 10 km, in sehr peripheren 22 km.

Die Aufzeichnung von Mobilfunkdaten erfolgt nicht, um daraus Bewegungen von Personen abzuleiten. Dass es dennoch möglich ist, macht diese Routinedaten zu einer spannenden Datenquelle nicht zuletzt für ansonsten datenarme Regionen.

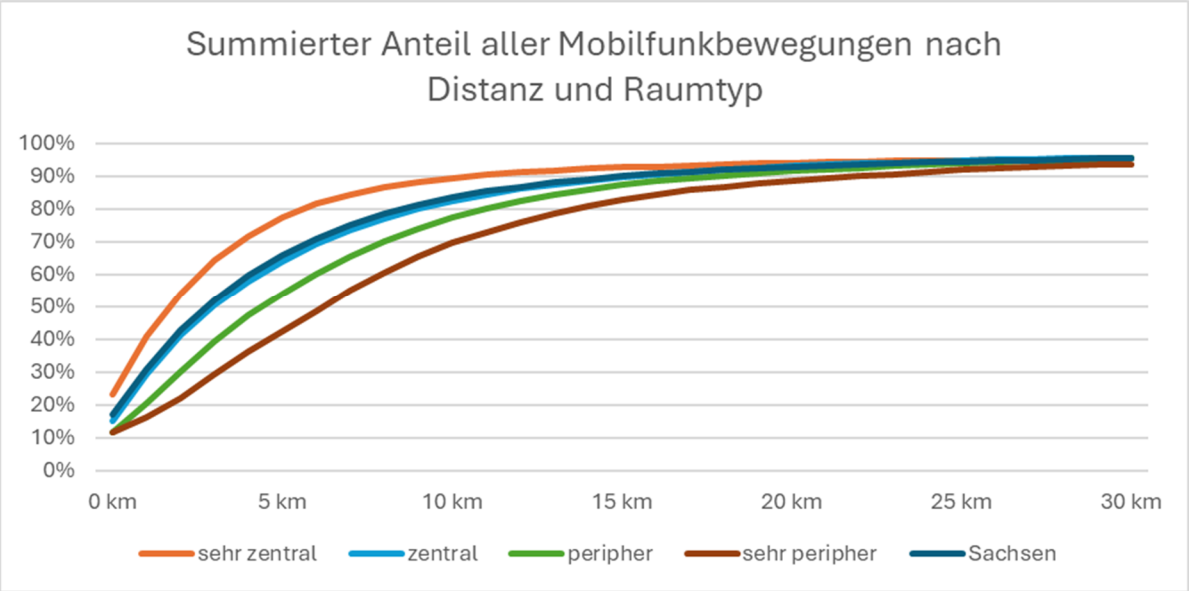

Abbildung 2: Summierter Anteil aller Mobilfunkbewegungen nach Distanz und Raumtyp im Juni 2023, eigene Berechnungen basierend auf Teralytics 2023

Insbesondere die hohe zeitliche Auflösung bietet Potenziale für die Mobilitätsforschung. Abgesehen von hohen Kosten bleibt die fehlende Verfügbarkeit von Rohdaten und die intransparente Aufbereitung für die wissenschaftliche Nutzung problematisch und verhindern weitere Analysen zu Geschwindigkeit, Routen, etc. (Bock und Schönduwe 2021; Statistisches Bundesamt (Destatis) 2023). Die aufbereiteten Daten haben dennoch einen hohen Wert zur Identifikation lokaler Bewegungsmuster. Für den größtmöglichen Nutzen ist eine Verbindung der Datenanalyse mit qualitativen Bestandteilen entscheidend, um den Kontext der Bewegungen zu verstehen. Dabei kann eine visuelle Aufbereitung unterstützend wirken. Hier bietet sich neben der codebasierten Darstellung für Einzelgemeinden das vom Leibniz-Institut für Länderkunde entwickelte

hin&weg-Tool (siehe nächster Abschnitt) für Analysen an, die über die Möglichkeiten des Teralytics-Dashboards hinausgehen. Aufgrund des Potentials von Mobilfunkdaten für die Mobilitätsforschung, wäre eine geregelte Bereitstellung, wie sie auch von Seiten des Statistischen Bundesamts diskutiert wird (Saidani et al. 2022), begrüßenswert.

4 Krankenkassen-Abrechnungsdaten

Während Mobilfunkdaten tatsächliche Bewegungen ohne Kontext der Mobilitätsursachen bereitstellen, ermöglichen Krankenkassenabrechnungsdaten das Gegenteil. Die Routinedaten enthalten sowohl die Gemeinde, in der die Ärztin bzw. der Arzt ihre/seine Praxis hat, als auch die Wohnsitzgemeinde der Patient:innen. Ob, wann und wie häufig diese dazwischen mobil waren, verraten die Daten nicht. Dennoch bieten sie eine einzigartige Gelegenheit für das Verständnis eines zentralen Elements der Daseinsvorsorge.

Mobilität zur Inanspruchnahme von Gesundheitsleistungen stellt eine kleine Untergruppe der Alltagsmobilität dar. Dabei kommt der wohnortnahen Versorgung mit ärztlichen Leistungen gerade in dünner besiedelten Gebieten eine besondere politische und soziale Bedeutung zu, wie nicht zuletzt in Interviews im Rahmen des PeriMobil-Projektes deutlich wurde. Krankenkassenabrechnungsdaten liefern hier eine wertvolle Ergänzung zu bestehenden Mobilitätsdatenquellen um den speziellen Bereich der Gesundheitsmobilität abzubilden und ein Grundanfordernis an Verkehrsverbindungen aufzuzeigen.

Daten

74,6 Millionen Menschen in Deutschland sind gesetzlich versichert. Das entspricht ca. 90% der Bevölkerung (Verband der Ersatzkassen (vdek) 2025). Begeben sich diese Menschen in ärztliche Versorgung, wird dies von den Krankenkassen zu Abrechnungszwecken erfasst und es entstehen Routinedaten. Diese enthalten Informationen zu Diagnosen und Krankheitshäufigkeiten, für die Mobilitätsforschung entscheidend sind jedoch die Praxis- und Versichertenadressen (Ohlmeier et al. 2014). Aus diesen lassen sich Muster der Inanspruchnahme medizinischer Leistungen flächendeckend abbilden. Die Daten geben an, wie viele Personen in einem Quartal einen Arzt oder eine Ärztin einer bestimmten Fachgruppe aufgesucht haben – jedoch nicht, wie oft (Sundmacher et al. 2018).

Die im PeriMobil-Projekt genutzten Daten wurden uns für die Fallstudiengemeinden Johanngeorgenstadt und Arzberg von der Kassenärztlichen Vereinigung Sachsen (im Folgenden – KV) für das dritte Quartal 2022 bis zum zweiten Quartal 2023 zur Verfügung gestellt. Die bereitgestellten Daten enthalten nur Informationen zu Ärzt:innen und Patient:innen in Sachsen, über alle Krankenversicherungsträger hinweg. Inanspruchnahmen in oder aus anderen Bundesländern können hingegen nicht abgebildet werden. Dies wäre bei einem Datenbezug von einer einzelnen Krankenversicherung (etwa AOK Plus oder Barmer) möglich, allerdings wäre der Datensatz auf die jeweiligen Versicherten beschränkt und mit entsprechenden Datenschutz-Problemen in kleinen Gemeinden verbunden.

Auswertung

Die bereitgestellten Abrechnungsdaten enthalten komplexe Zeitreihen-Verflechtungen, deren Auswertung mit klassischen Statistikprogrammen nicht oder nur sehr eingeschränkt möglich ist. Aus diesem Grund wurde die IfL-eigene hin&weg-Software zur Analyse und Visualisierung verwendet. Die Anwendung wurde speziell für die Auswertung komplexer Verflechtungsdaten entwickelt und steht Open-Source und kostenfrei zur Verfügung (hin-und-weg.online).

Eine explorative Auswertung der Abrechnungsdaten für die Gemeinde Arzberg (Gescher 2024) folgt im nächsten Abschnitt. Die Analyse zeigt die Auswirkungen einer Praxisschließung auf die Muster der Inanspruchnahme medizinischer Leistungen. Die KV-Daten ermöglichen es, Ausweichbewegungen auf weiter entfernte Praxisstandorte nachzuvollziehen und so wichtige Änderungen der Alltagsmobilität zu identifizieren.

Die Aufschlüsselung der Daten nach Fachgebieten hat das Potenzial unterschiedliche Versorgungsstrukturen aufzuzeigen und macht deutlich, welche Strecken die Bevölkerung in peripheren Gebieten vor allem zu Fachärzt:innen zurücklegen muss.

Besonders bei Fachärzt:innen gilt es die Einschränkungen der Datenquelle immer mit zu beden-

ken. Da es sich um Abrechnungsdaten handelt, müssen nicht zwangsläufig auch Personen zwischen Wohn- und Praxisstandort mobil gewesen sein. Bei Laborärzt:innen liegt z.B. nur ein Versand der Proben nahe. Zusätzlich können keine Aussagen über die Anzahl an Besuchen und damit den Umfang der Mobilität zwischen Wohn- und Praxisgemeinden getroffen werden. Wie auch bei Pendeldaten stellen Zweitwohnsitze und im Falle der Gesundheitsdaten zusätzlich private wie berufliche Aufenthalte an anderen Orten zusätzliche Einschränkungen der Interpretierbarkeit dar. So kann nicht unterschieden werden, ob (weit entfernte) Praxen aufgrund ihrer Spezialisierung und/oder eines Versorgungsmangels im Nahraum oder aufgrund anderweitiger Aufenthalte am Praxisort aufgesucht wurden.

Berücksichtigt man die methodischen Einschränkungen, bieten diese dennoch eine aufschlussreiche Datenquelle für die Abbildung von Gesundheitsmobilität. Bewegungen, die teils entgegen der Pendel-ströme verlaufen und die in einer mobilfunkbasierten Betrachtung von Alltagsmobilität untergehen, können hierüber herausgearbeitet werden. So kann ein Teil der Mindestmobilitätsanforderungen zur Sicherstellung der Daseinsvorsorge in Forschung und Planung berücksichtigt werden. Dies wird im Folgenden exemplarisch für die Gemeinde Arzberg demonstriert.

Ergebnisse

Arzberg in Sachsen, gelegen im ostelbischen Flach-land, hat etwa 1.800 Einwohner:innen, 19 Ortsteile, einen Einkaufsmarkt, eine Grundschule und – bis Februar 2023 – eine Hausarztpraxis. Anhand der KV-Abrechnungsdaten kann konkret nachvollzogen werden, was ein Wegfall der lokalen hausärztlichen Versorgung in einer peripheren Region bedeutet. Im Folgenden konzentrieren wir uns auf Daten für Hausarztbesuche mit Bezug zur Gemeinde Arzberg, die sowohl die Patient:innen enthalten, die die Hausarztpraxis in Arzberg aufsuchen, als auch diejenigen, die in Arzberg wohnen und medizinische Versorgung in anderen Gemeinden in Anspruch nehmen. Die Daten liegen nur für Sachsen vor, Arztbesuche im nahegelegenen Brandenburg oder in anderen Bundesländern werden ebenso wie Besuche von dort nicht erfasst.

Im dritten Quartal 2022 wurden insgesamt 902 Patient:innen aus Arzberg in den Abrechnungsdaten erfasst, davon suchten 256 einen Hausarzt im benachbarten Beilrode auf, 236 gingen nach Torgau, 124 nach Belgern-Schildau. Mit 244 Patientinnen und Patienten versorgte die Hausarztpraxis im Arzberger Ortsteil Triestewitz in diesem Quartal die zweitmeisten Arzberger Patient:innen (27,5 %). Dazu kamen Patient:innen aus Torgau (94), Beilrode (67) und anderen Gemeinden (27), welche zu Praxisbesuchen nach Arzberg kamen (siehe Abbildungen 3 und 4).

Im zweiten Quartal 2023 war die Hausarztpraxis in Arzberg bereits geschlossen, daher wurden auch keine einpendelnden Patient:innen mehr registriert. Dafür stieg die Zahl der Arzberger:innen, die zu einem Hausarzt nach Beilrode fuhren, um 50 % an auf 384. Auch nach Torgau (+5 %) und Belgern-Schildau (+14 %) reisten mehr Patient:innen. Dazu kommt eine weitere auffällige Entwicklung: Während im dritten Quartal 2022 902 Patient:innen einen Arzt aufsuchten, waren es im zweiten Quartal 2023 etwa 10 % weniger (818).

Als Ursache kommt eine Erkältungswelle im Spätsommer 2022 (TK-Gesundheitsreport 2024) infrage, die jedoch deutlich schwächer ausfiel als im ersten Quartal 2023. Auch könnten Patient:innen neben der Ärztin in Arzberg zusätzlich Ärzte in anderen Gemeinden aufgesucht haben, daher bis zur Schließung doppelt gezählt worden sein. Dass im ersten Quartal 2023, in dem der Krankenstand deutschlandweit sehr hoch lag (TK Gesundheitsreport 2024), dennoch etwa 4 % weniger einen Hausarzt aufgesucht haben als im vierten Quartal 2022, legt jedoch nahe, dass es aufgrund der Praxisschließung zu (vorübergehenden) Aufschiebungen von Arztterminen kam oder Personen ganz auf einen Hausarztbesuch verzichtet haben.

Dies legen auch Monsees und Westhal (2024) in einer aktuellen deutschlandweiten Studie zu Effekten von Hausarztschließungen nahe. Zu der Herausforderung, eine neue Praxis zu finden, kommt in ländlichen Räumen die Erreichbarkeit als wichtiges Thema hinzu. Zwar sind die (Auto-)Fahrzeiten nach Beilrode mit zehn Minuten und nach Torgau mit 15 Minuten überschaubar, immerhin 17 % der Arzberger Hausarztbesucher:innen müssen jedoch den knapp halbstündigen Weg per Fähre nach Belgern-

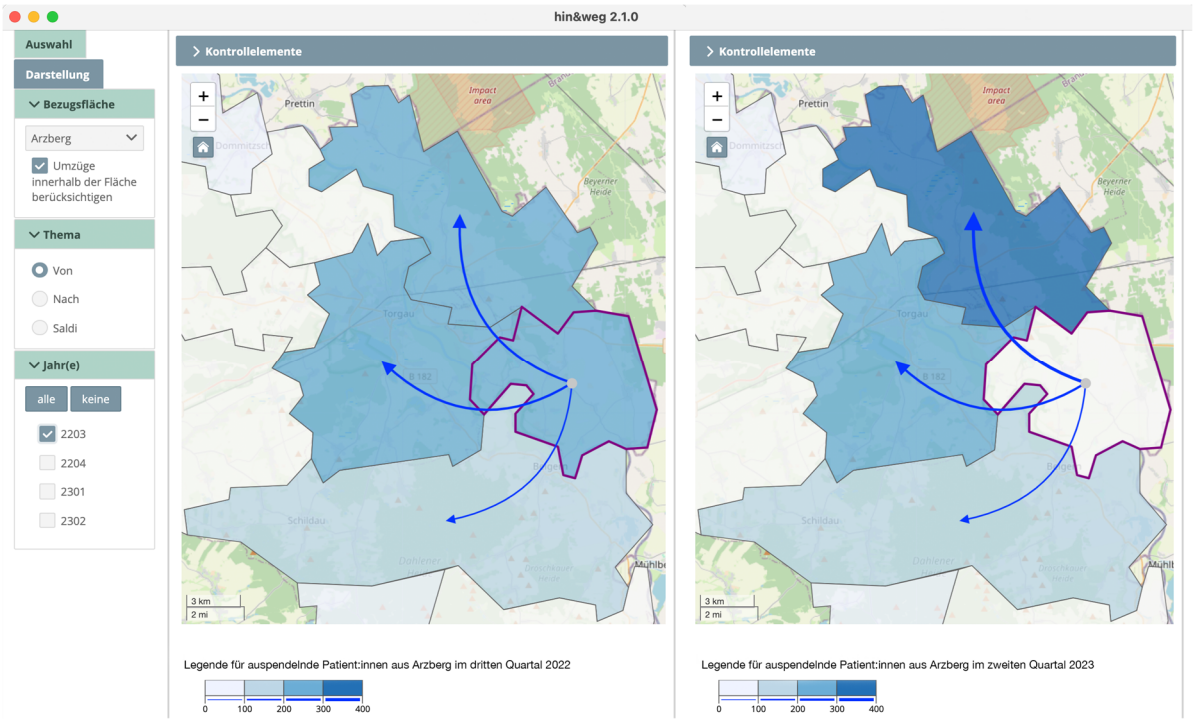

Abbildung 3: Vergleich der Bewegungsmuster zur Inanspruchnahme hausärztlicher Leistungen: Vor (links) und nach (rechts) der Schließung der Hausarztpraxis in Arzberg (rot umrandet). Patient:innen müssen in die Nachbargemeinden ausweichen, es sind keine Hausarztbesuche im Gemeindegebiet mehr möglich; Bearbeiteter Screenshot der hin&weg-Anwendung bar KV Sachsen

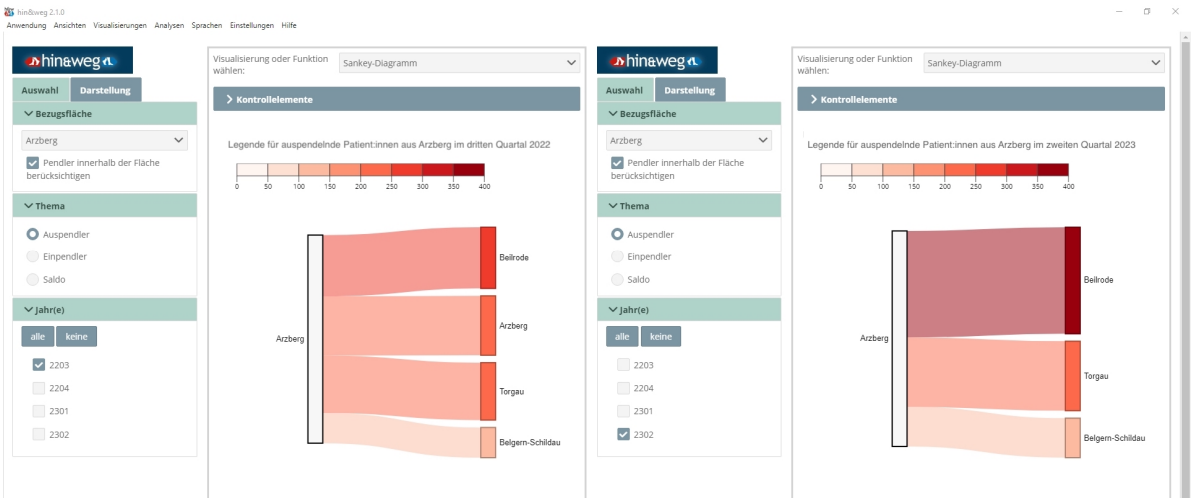

Abbildung 4: Vergleich der Praxisstandorte Arzberger Patient:innen im dritten Quartal 2022 vor der Schließung der Hausarztpraxis in der Gemeinde, und im zweiten Quartal 2023. Bearbeiteter Screenshot der hin&weg-Software basierend auf Daten der KV Sachsen

Schildau auf sich nehmen. Dazu kommt, dass Arzberg eine überdurchschnittlich alte Bevölkerung aufweist. Gerade (Hoch-)Betagte sind für Arzttermine oftmals auf Fahrdienste anderer, in der Regel Familienmitglieder oder – im Fall von Arzberg auch des ehrenamtlichen Bürgerbusses – angewiesen. Die Fahrzeiten verdoppeln sich dadurch oftmals oder es müssen Wartezeiten durch die Fahrdienstleistenden in Kauf genommen werden.

Daten-Visualisierung mit hin&weg

Arzberg ist kein Einzelfall. In vielen ländlichen Räumen stellt der Mangel an Hausärzt:innen als erste Anlaufstelle bei gesundheitlichen Fragen und Beschwerden ein gravierendes Problem dar. Ein genaues Verständnis der gesundheitsbezogenen Mobilitätsmuster kann Krankenkassen und Politik unterstützen, Versorgungslücken zu identifizieren und Ausweichbewegungen nach Praxisschließungen nachzuvollziehen. Die Open-Source-Software hin&weg ermöglicht solche Analysen. Die Anwendung wurde speziell für die Auswertung komplexer Verflechtungsdaten entwickelt. Basierend auf den Verflechtungsmatrizen und einer Shape-Datei ermöglicht es hin&weg Sankey-, Chord-, oder Balkendiagramme, Karten und Zeitreihen zu erstellen und dabei interaktiv zwischen betrachteten Gebieten und Jahren zu wechseln. Die Anwendung wurde als abgeschottetes Stand-Alone-Programm entwickelt und funktioniert daher auch komplett ohne Internetzugriff und transferiert keine Daten nach außen. Dies ist für viele Datenhalter entscheidend im Umgang mit sensiblen Gesundheits- aber auch weiteren Verflechtungsdaten. Die Software wurde gezielt so lizensiert, dass sie frei verändert werden kann, um sie noch besser auf spezielle Anwendungsfälle, wie etwa Abrechnungsdaten, zuzuschneiden. Komplexe Mobilitätserfordernisse in der Gesundheitsvorsorge können so räumlich und statistisch abgebildet werden; die in der Anwendung implementierte Zeitauswahl erlaubt intertemporale Vergleiche und Zeitreihen, um Muster zu erkennen und zu validieren.

Krankenkassenabrechnungsdaten erlauben neben Analysen zu Hausärztinnen und Hausärzten auch die Inanspruchnahme fachärztlicher Leistungen zu untersuchen. Hier zeigt sich, dass Menschen in peripheren Gebieten teils sehr weite Wege in die Zentren (Leipzig, Dresden) in Kauf nehmen (müssen), um eine adäquate Versorgung zu erhalten.

5 Fazit

Tiefgehende Analysen der Mobilitätsmuster in ländlichen Räumen stellen eine Herausforderung für die Forschung dar. Angesichts ländlicher Datenarmut fällt es schwer, Annahmen und Erfahrungswissen zu kontextualisieren, zu bestätigen oder, wo nötig, zu widerlegen. Die Nutzung von Routinedaten kommt mit ihren eigenen Herausforderungen und kann dennoch einen Schritt in eine evidenzbasiertere Debatte darstellen. Mobilfunkdaten können vor allem durch ihre breite Verfügbarkeit und hohe zeitlichen Auflösung einen Beitrag leisten. Krankenkassenabrechnungsdaten ermöglichen es, ein Mindestmaß an notwendiger Mobilitätsinfrastruktur abzubilden, um einen der zentralen Bausteine der Daseinsvorsorge zu erreichen. Beide Datenquellen sind dabei mit Einschränkungen verbunden. Mobilfunkdaten sind proprietär, in der Aufbereitung bedingt nachvollziehbar und liefern keine Informationen über Mobilitätsgründe. Während diese bei den Krankenkassendaten naheliegend sind, bilden diese keine erfasste, sondern potenzielle Mobilität ab, deren Ausmaß und Häufigkeit nicht aus den Daten ersichtlich ist. Beide Quellen stützen sich aufgrund ihrer speziellen Herkunft auf sehr breite Stichproben, die weit über übliche Befragungsdaten hinausgehen und somit flächendeckende Mobilitätsanalysen besonders in ländlichen Räumen ermöglichen. Dabei stellt das Modifiable Areal Unit Problem (MAUP), wie bei allen gebietsbezogenen Datenquellen, eine zusätzliche Herausforderung dar, beispielsweise indem weite Pendelstrecken oder Inanspruchnahmen innerhalb einer Gemeinde nicht abgebildet werden während kürzere Strecken über Gemeindegrenzen enthalten sind.

Es wird noch ein weiter Weg sein, bis diese und weitere Routinedaten (z.B. Schuleinzugsbereiche) über Modellprojekte wie PeriMobil hinaus für Planung und Forschung genutzt werden können. Neben Fragen des Datenschutzes spielen hier die Fragmentierung der Institutionen, die die Daten erheben und verwalten, eingeschränkte (personelle) Kapazitäten zur Bereitstellung und teilweise fehlende Bereitschaft zur Herausgabe der Daten eine Rolle. Nichtsdestotrotz bieten Routinedaten ein großes Potential, Wegebeziehungen und zeitliche Muster der Alltagsmobilität in ländlichen Räumen besser zu verstehen.

Literaturverzeichnis

Alarcon, Pedro A. de, Salevsky, Alejandro, Gheti-Kao, Daniel, Rosalen, Willian, Duarte, Marby C., Cuervo, Carlos et al. (2021): The contribution of telco data to fight the COVID-19 pandemic: Experience of Telefonica throughout its footprint. In: Data & Policy 3. DOI: 10.1017/dap.2021.6.

Bäumer, M., Dubernet, I., Follmer, R. & Wawrzyniak, B. (2025): MiD 2023. Regionale Ergebnisse – ein Überblick. Studie von infas, DLR, IVT und infas 360 im Auftrag des Bundesministeriums für Verkehr und digitale Infrastruktur, Berlin.

Bundesagentur für Arbeit (2025): Pendleratlas. https://statistik.arbeitsagentur.de/DE/Navigation/Statistiken/Interaktive-Statistiken/Pendleratlas/Pendleratlas-Nav.html

Bock, B. & Schönduwe, R. (2021): Black-Box Mobilität. Mobilitätsanalysen mit Nachfragedaten. Wirkungen technischer und erhebungsbedingter Eigenschaften auf die Aussagekraft von Mobilitätsdaten. In: Discussion Paper SP III (601).

Forschungsgesellschaft für Straßen- und Verkehrswesen (2020): Begriffsbestimmungen für das Straßen- und Verkehrswesen, Köln.

Fuchs, M., Fritzsche, B. & Sujata, U. (2025): Entwicklung und Struktur der Pendlerverflechtungen Sachsen-Anhalts. In: IAB-Regional. Berichte und Analysen aus dem Regionalen Forschungsnetz. IAB Sachsen-Anhalt-Thüringen 1. Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg. https://doi.org/10.48720/IAB.RESAT.2501

Hadam, S. (2021): Pendler Mobil: Die Verwendung von Mobilfunkdaten zur Unterstützung der amtlichen Pendlerstatistik: Können Mobilfunkdaten die amtliche Pendlerrechnung unterstützen? In: AStA Wirtschafts- und Sozialstatistisches Archiv 15 (3-4), S. 197–235. DOI: 10.1007/s11943-021-00294-z.

Herget, M. (2016): Mobilität von Familien im ländlichen Raum. Arbeitsteilung, Routinen und typische Bewältigungsstrategien. Springer VS, Wiesbaden. https://doi.org/10.1007/978-3-658-11068-0

Hubrich, S. (2024): Teilnehmende Städte und Gemeinden an Mobilität in Städten – SrV 2023. Technische Universität Dresden, Dresden.

Information und Technik Nordrhein-Westfalen (IT.NRW) (2023): Qualitätsbericht Pendlerrechnung der Länder.

Kersting, N. (2025): Digitale Ungleichheiten und digitale Spaltung. In: Klenk, T., Nullmeier, F. & Wewer, G. (eds) Handbuch Digitalisierung in Staat und Verwaltung. Springer VS, Wiesbaden. https://doi.org/10.1007/978-3-658-37373-3_19

Krebs, O. & Pflüger, M. (2023): On the road (again): Commuting and local employment elasticities in Germany. Regional Science and Urban Economics 99. https://doi.org/10.1016/j.regsciurbeco.2023.103874

Monsees, D. & Westphal, M. (2024): Disruptions in Primary Care: Can Resigning GPs Cause Persistently Negative Health Effects?, CESA Working Paper 4. Fern Universität, Hagen https://www.fernuni-hagen.de/igas/docs/cesa_wp_04.pdf

Nitschke, L., Quentin, P., Kanisius, F., Schluckebier, K., Burlon, N. S. & Buscher, J. et al. (2022): Pendeln verstehen: status quo, Forschungsstand und Perspektiven: ISOE-Institut für sozial-ökologische Forschung, Frankfurt am Main.

Oberst, C. A. (2020): Mobilität, Wohnen und der Arbeitsmarkt. In: Depenheuer, Hertzsch et al. (Hg.) 2020 – Wohneigentum für breite Schichten 18. Springer VS, Berlin, Köln. https://doi.org/10.1007/978-3-662-61287-3_6

Ohlmeier, C., Frick, J., Prütz, F., Lampert, T., Ziese, T., Mikolajczyk, R. & Garbe, E. (2014): Nutzungsmöglichkeiten von Routinedaten der Gesetzlichen Krankenversicherung in der Gesundheitsberichterstattung des Bundes. In: Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz 57 (4), S. 464–472. DOI: 10.1007/s00103-013-1912-1.

Saidani, Y., Bohnensteffen, S. & Hadam, S. (2022): Qualität von Mobillfunkdaten. Projekterfahrungen und Anwendungsfälle aus der amtlichen Statistik. Hg. v. Statistisches Bundesamt (Destatis) (WISTA, 5).

Schmidt, A. & Männel, T. (2017): Potenzialanalyse zur Mobilfunkdatennutzung in der Verkehrsplanung, Stuttgart.

Statistisches Bundesamt (Destatis) (2023): Mobilitätsindikatoren auf Basis von Mobilfunkdaten.

https://www.destatis.de/DE/Service/EXSTAT/Datensaetze/mobilitaetsindikatoren-mobilfunkdaten.html, zuletzt geprüft am 22.05.2024.

Sundmacher, L.; Schang, L.; Schuettig, W.; Flemming, R.; Frank-Tewaag, J. & Geiger, I. et al. (2018): Gutachten zur Weiterentwicklung der Bedarfsplanung i.S.d. §§ 99 ff. SGB V zur Sicherung der vertragsärztlichen Versorgung, München, Leipzig, Bonn, Greifswald, Köln.

Szocska, M., Pollner, P., Schiszler, I., Joo, T., Palicz, T. & McKee, M. et al. (2021): Countrywide population movement monitoring using mobile devices generated (big) data during the COVID-19 crisis. In: Scientific Reports 11 (1). DOI: 10.1038/s41598-021-81873-6.

Techniker Krankenkasse (2024): Gesundheitsreport 2024. Arbeitsunfähigkeiten. https://www.tk.de/resource/blob/2168508/ee48ec9ef5943d2d40dc10a76bedf290/gesundheitsreport-au-2024-data.pdf, zuletzt geprüft am 14.08.2025.

Teralytics GmbH (2022): Teralytics Produktvorstellung.

Teralytics GmbH (2023): Matrix v5 Kurzdistanz Bericht_Kurzdistanz_Qualitätsbericht (1). Thüringen, Sachsen, Sachsen-Anhalt.

Van Dülmen, C. (2022): Alles anders und doch gleich – Fragile Alltagsmobilität sozial benachteiligter BewohnerInnen ländlicher Peripherien vor und während der Covid-Pandemie. Journal für Mobilität und Verkehr 14. https://doi.org/10.34647/jmv.nr14.id91

Verband der Ersatzkassen (vdek) (2025): Daten zum Gesundheitswesen: Versicherte. https://www.vdek.com/presse/daten/b_versicherte.html, zuletzt geprüft am 15.08.2025.

Wilde, M. (2014): Mobilität und Alltag. Einblicke in die Mobilitätspraxis älterer Menschen auf dem Land. Studien zur Mobilitäts- und Verkehrsforschung 25. Springer VS, Wiesbaden. https://doi.org/10.1007/978-3-658-03819-9

Williams, N. E., Thomas, T. A., Dunbar, M., Eagle, N., & Dobra, A. (2015). Measures of human mobility using mobile phone records enhanced with GIS data. PloS one, 10(7), e0133630. DOI: 10.1371/journal.pone.0133630

Autor*innenangaben

Jonathan Gescher
Wissenschaftlicher Mitarbeiter
Leibniz-Institut für Länderkunde, Schongauerstraße 9, 04328, Leipzig, Deutschland
E-Mail: J_Gescher@leibniz-ifl.de

Tim Leibert
Projektleiter
Leibniz-Institut für Länderkunde, Schongauerstraße 9, 04328, Leipzig, Deutschland
E-Mail: T_Leibert@leibniz-ifl.de

Wladimir Sgibnev
Projektleiter, Koordinator der Forschungsgruppe Mobilitäten und Migration
Leibniz-Institut für Länderkunde, Schongauerstraße 9, 04328, Leipzig, Deutschland
E-Mail: W_Sgibnev@leibniz-ifl.de

Anne Mehl
Fußverkehrsplanerin
Bezirksamt Friedrichshain-Kreuzberg
Yorckstraße 4-11, 10965, Berlin
E-Mail: Anne.Mehl@ba-fk.berlin.de

Journal für Mobilität und Verkehr

ISSN 2628-4154

www.dvwg.de

Repräsentativität bei Mobilitätserhebungen mit GPS-Smartphone-Tracking: Methodische Ansätze im Überblick

Dr.-Ing. Mariana Batista*, Thuy Chinh Duong, Stephan Leppler

Siehe Autor*innenangaben

Abstract

Smartphonebasiertes GPS-Tracking hat sich als leistungsfähiges Instrument zur Erhebung hochauflösender longitudinaler Mobilitätsdaten etabliert und bietet gegenüber konventionellen Verfahren signifikante Vorteile hinsichtlich zeitlicher Auflösung und Genauigkeit. Wie bei allen empirischen Erhebungsmethoden ist jedoch die Repräsentativität der Stichprobe entscheidend für die Validität und Generalisierbarkeit der Ergebnisse. Dieser Artikel gibt einen Überblick über methodische Ansätze zur Stärkung der Repräsentativität in Smartphone-Tracking-Studien. Er beschreibt praxisorientierte Strategien zur Teilnehmerrekrutierung, Verfahren zur statistischen Gewichtung und Kalibrierung sowie Ansätze zur kontinuierlichen Datenvalidierung. Durch die Verbindung etablierter Methoden der klassischen Umfrageforschung mit modernen Erkenntnissen aus dem Bereich Smartphone-Tracking bündelt der Beitrag Ansätze zur Reduktion von Selektionseffekten und zur Generierung verlässlicher, hochwertiger Mobilitätsdaten als belastbare Grundlage für Verkehrsplanung und Mobilitätsforschung.

Schlagwörter:

GPS-Tracking, Repräsentativität, Längsschnittdaten

1 Einleitung

Smartphone-basiertes GPS-Tracking hat sich als innovatives Instrument in der Mobilitätsforschung etabliert. Neben akademischen Institutionen wie der ETH Zürich, der TU München und der TU Dresden setzen auch führende Markt- und Sozialforschungsinstitute wie infas und YouGov verstärkt auf diese Technologie. Die Methode ermöglicht eine passive, hochauflösende Erfassung longitudinaler Mobilitätsdaten – von präzisen Routenverläufen und Weglängen bis hin zu dynamischen Modal-Split-Veränderungen. Gegenüber traditionellen Erhebungsformen (z. B. Wegetagebüchern oder Telefoninterviews) bietet das Tracking signifikante Vorteile: eine überlegene zeitliche Granularität, minimierte Erinnerungsfehler (recall error) und ein hohes Potenzial für Langzeitbeobachtungen. Damit liefert es eine fundierte Datengrundlage für Verkehrsplanung, Einnahmenaufteilung und strategische Mobilitätsentscheidungen.

Wie bei allen empirischen Erhebungsverfahren ist die Repräsentativität der Stichprobe eine zentrale Voraussetzung für die Validität und Generalisierbarkeit der Ergebnisse. Während traditionelle Erhebungen meist auf Zufallsstichproben beruhen, stehen Smartphone-Tracking-Projekte häufig vor der Herausforderung selektiver Teilnahmestrukturen (self-selection bias); sie erfordern dann zusätzliche methodische Absicherungen. Repräsentativität wird hier nicht als isolierter Schritt, sondern als koordinierter Gesamtprozess verstanden: Dieser umfasst eine zielgerichtete Rekrutierungsstrategie, statistische Gewichtungs- und Hochrechnungsverfahren sowie die kontinuierliche Validierung anhand externer Referenzdaten. In ihrer Gesamtheit sorgen diese

*korrespondierende Autorin

Elemente dafür, dass die Mobilitätskennziffern stabil und belastbar bleiben.

Der vorliegende Artikel versteht sich als strukturierte Zusammenstellung methodischer Ansätze zur Sicherung der Datenqualität in Smartphone-basierten Mobilitätserhebungen (Abbildung 1). Auf Basis etablierter Forschungsprinzipien und praktischer Projekterfahrung werden Strategien vorgestellt, die flexibel an unterschiedliche Forschungsfragen und operative Kontexte angepasst werden können.

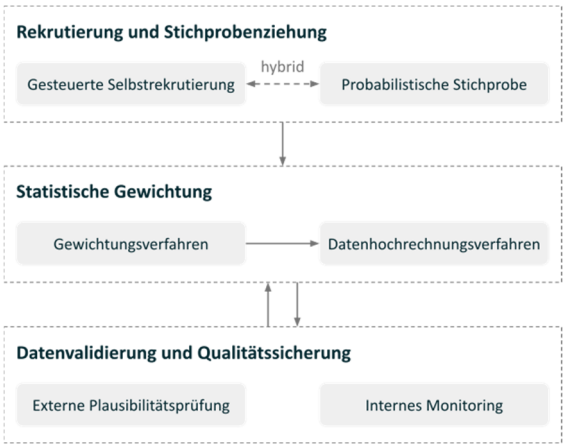

Abbildung 1: Überblick über die Ansätze (Eigene Darstellung, 2026)

2 Stichprobenziehung und Rekrutierungsstrategien zur Stärkung der Repräsentativität

Die Wahl des Sampling-Designs beeinflusst maßgeblich die Aussagekraft der Ergebnisse. Über reine Selbstrekrutierung hinaus gelten hybride Ansätze – die Verknüpfung klassischer Stichprobenverfahren mit digitalen Rekrutierungskanälen – in der Praxis als vielversprechend, um die Repräsentativität zu steigern (Heinonen et al. 2024; Lawson et al. 2023; Loder et al. 2024). Gleichzeitig prägt der gewählte Akquisekanal spezifische Verzerrungsmuster, die im nachgelagerten Gewichtungsprozess adressiert werden müssen: Während eine app-basierte Selbstrekrutierung besonders hochmobile Bevölkerungsgruppen ansprechen kann, erreichen postalische Einladungen häufiger Personen mit geringerer Mobilitätsfrequenz. Die hybride Kombination verschiedener Kanäle kann zu einer ausgeglicheneren Ausgangsstichprobe beitragen und strukturelle Verzerrungen bereits in der Akquisephase reduzieren.

2.1 Gesteuerte Selbstrekrutierung

Die Selbstrekrutierung ist aufgrund ihrer großen Reichweite und geringen Teilnahmehürden weit verbreitet. Um eine hinreichende Diversität der Stichprobe zu gewährleisten, bedarf es jedoch einer gezielten Steuerung (Meister et al. 2025; Heinonen et al. 2024; Piras et al. 2025). Die nachfolgend aufgeführten Kanäle lassen sich kontextspezifisch kombinieren, um die Repräsentativität aktiv zu fördern:

- Gezielte Online-Kampagnen: Spezifische demografische Segmente können über soziale Medien und Suchmaschinen mithilfe einer präzisen Zielgruppendefinition angesprochen werden (Micro-Targeting).
- Integration in bestehende Apps: Kooperationen mit etablierten Apps (vgl. BVG Fahrinfo-App – Meine Wege) ermöglichen den Zugang zu relevanten Nutzergruppen, wie z. B. Fahrgästen des öffentlichen Verkehrs.
- Medienbasierte Rekrutierung mit Screening: Über Print- und Rundfunkmedien lassen sich Zielgruppen ansprechen, die über digitale Kanäle möglicherweise schwerer erreichbar sind. Ein anschließendes Screening-Verfahren ermöglicht die Auswahl geeigneter Teilnehmender aus einem vorqualifizierten Interessiertenpool.
- Gezielte Ansprache bestehender Kundendatenbanken: Postalische Einladungen oder E-Mail-Kampagnen an Abonnent:innen sowie registrierte App-Nutzende ermöglichen eine zielgenaue und persönliche Akquise.
- Flyer und Vor-Ort-Präsenz: Der Einsatz von Flyern an Mobilitätsknotenpunkten, in Bürgerzentren oder bei öffentlichen Veranstaltungen bietet eine Möglichkeit, Personen direkt vor Ort anzusprechen. Kleine Werbegeschenke (z. B. Aufkleber oder Gutscheine) können das Interesse und die Anmeldequote steigern.
- Einbindung von Partnernetzwerken: Die Nutzung der Reichweite institutioneller Akteure und lokaler Multiplikatoren (z. B. E-Mail-Verteiler, Social-Media-Kanäle) kann die Rekrutierung durch persönliche oder institutionell vermittelte Empfehlungen unterstützen.
- Incentivierungssysteme: Gutscheine oder Verlosungen können als Anreize eingesetzt werden, um die Teilnahmebereitschaft über verschiedene Nutzungsgruppen hinweg zu steigern.

- Institutionelles Branding und transparente Kommunikation: Die Kooperation mit renommierten Institutionen (z. B. Hochschulen oder Verkehrsunternehmen) sowie eine transparente Kommunikation zum Datenschutz erhöhen die Glaubwürdigkeit und die Teilnahmequote.

2.2 Probabilistische Stichprobenverfahren

Zufallsstichproben finden auch in Smartphone-Tracking-Studien Anwendung. Theoretisch gelten Smartphonebesitz, digitale Kompetenzen und Installationsbereitschaft als Teilnahmevoraussetzung, wodurch methodisch eine bedingte Wahrscheinlichkeitsstichprobe entsteht. In der Praxis ist die Smartphone-Penetration jedoch so weit fortgeschritten, dass demografische Verzerrungen (z. B. nach Alter) kaum noch primär auf mangelnde Technikaffinität zurückzuführen sind. Die Wahl des passenden Verfahrens erfordert stets eine Abwägung zwischen statistischer Repräsentationsgüte und ökonomischem Aufwand, orientiert an der Struktur der Zielpopulation. Abbildung 2 veranschaulicht vier grundlegende probabilistische Stichprobenverfahren.

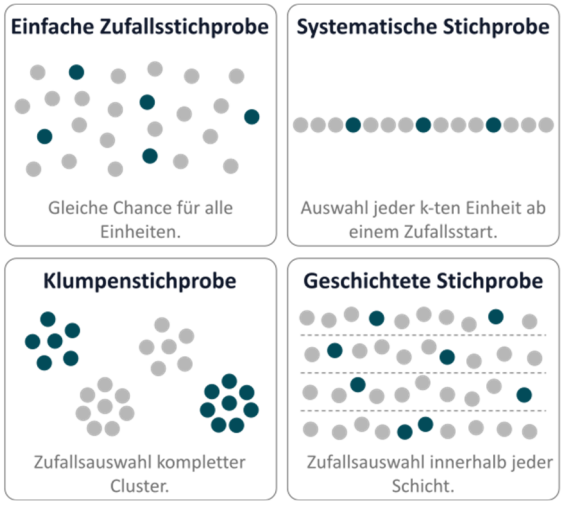

Abbildung 2: Überblick über probabilistische Stichprobenverfahren (Eigene Darstellung nach United Nations Statistics Division, 2008)

Die einfachste Form ist die einfache Zufallsstichprobe (Simple Random Sampling), bei der jede Person der Population die gleiche Auswahlwahrscheinlichkeit hat. Sie ist methodisch transparent, setzt jedoch einen vollständigen Stichprobenrahmen voraus, beispielsweise eine Kundendatenbank. Die systematische Stichprobe (Systematic Sampling) ermöglicht eine strukturierte Auswahl: Ausgehend von einem Zufallsstart wird jede k-te Person einer geordneten Liste ausgewählt. Bei großen Populationen kann dieses Verfahren die praktische Durchführung vereinfachen und eine gleichmäßige Verteilung der Stichprobe über den Auswahlrahmen begünstigen. Die beiden weiteren Verfahren berücksichtigen die Struktur der Population. Bei der Klumpenstichprobe (Cluster Sampling) werden zunächst natürliche Gruppen, beispielsweise Gemeinden, zufällig ausgewählt. Bei einer einstufigen Klumpenstichprobe werden anschließend alle Personen innerhalb der ausgewählten Gruppen einbezogen. Das kann die Erhebungskosten bei geografisch weit verteilten Populationen senken, geht jedoch häufig mit einer geringeren statistischen Präzision einher. Die geschichtete Stichprobe (Stratified Sampling) unterteilt die Population dagegen in sich gegenseitig ausschließende Schichten, beispielsweise nach Alter oder Raumtyp, aus denen jeweils Zufallsstichproben gezogen werden. Dadurch lassen sich wichtige Teilgruppen gezielt berücksichtigen und Stichprobenfehler bei geeigneter Schichtung reduzieren. Voraussetzung sind Informationen über die Merkmale, anhand derer die Schichten gebildet werden.

Nach der Auswahl müssen die Personen über verschiedene Rekrutierungskanäle bzw. Kontaktstrategien angesprochen werden. Postalische Einladungen an registerbasierte Zufallsstichproben (z. B. EWO-Daten) erreichen die Bevölkerung flächendeckend und bilden eine gute Basis für einen ausgewogenen Teilnehmendenpool. Einladungen per E-Mail setzen Adressen aus Datenbanken oder Registern voraus und lassen sich mit einer Schichtung verbinden, sodass schon die Kontaktliste die demografische Vielfalt abbildet. Gruppen mit geringer Online-Affinität, etwa ältere Menschen, lassen sich telefonisch oder persönlich vor Ort ansprechen. Das wirkt dem Non-Response-Bias entgegen, ist aber aufwendiger.

In der Praxis werden Auswahl und Ansprache häufig zu hybriden Designs kombiniert (Lawson et al. 2023; Masse et al. 2025; Bürbaumer et al. 2024). Eine Möglichkeit ist die Verknüpfung mit etablierten Mobilitätspanels. Sie stärkt die Stichprobenstruktur und erlaubt es, Ergebnisse über verschiedene Datentypen hinweg gegeneinander zu prüfen. Reichen die eigenen Kanäle nicht aus, können professionelle Access-Panels gezielt Quoten für bestimmte demogra-

fische oder verhaltensbezogene Gruppen füllen. Solche Stichproben werden häufig schon vor der finalen Gewichtung statistisch angepasst, um ihre Balance zu verbessern.

3 Statistische Gewichtung und Hochrechnung

Selbst bei optimierter Rekrutierung lassen sich systematische Stichprobenverzerrungen (Bias) kaum vollständig vermeiden. In Smartphone-Tracking-Studien – die primär auf Selbstrekrutierung oder hybriden Ansätzen basieren – erfordert die Analyse daher statistische Gewichtungsverfahren, um Abweichungen von der Zielpopulation zu korrigieren. Dieser Abschnitt beschreibt gängige Verfahren zur Gewichtung und Hochrechnung vor, die sich in der Praxis sowohl bei selbstrekrutierten als auch bei hybriden Stichproben anwenden lassen, um Indikatoren über das Mobilitätsverhalten der Gesamtbevölkerung zu gewinnen.

3.1 Gewichtungsverfahren

Vor der Gewichtung ist zu klären, in welchen Dimensionen die Stichprobe von der Zielpopulation abweicht. Je nach Studie kommen unterschiedliche Merkmale infrage. Vier Dimensionen sollten jedoch in der Regel immer berücksichtigt werden. Erstens kann die soziodemografische Struktur der Teilnehmenden von der Bevölkerung abweichen, etwa hinsichtlich Bildungsstands, Erwerbsstatus oder Haushaltszusammensetzung. Solche Unterschiede sollten durch eine demografische Gewichtung ausgeglichen werden. Zweitens kann die räumliche Verteilung verzerrt sein. Wenn die Rekrutierung nicht geografisch geschichtet erfolgt, können beispielsweise Menschen in Städten überproportional häufig in der Stichprobe vertreten sein, was das ermittelte Mobilitätsverhalten verzerren kann. Drittens beeinflusst die verhaltensbezogene Zusammensetzung die Ergebnisse. Je nach Rekrutierungskanal, beispielsweise über Apps, Gamification oder Anreize, beteiligen sich überdurchschnittlich häufig Personen, die besonders mobil oder an Mobilität interessiert sind und bestimmte Verkehrsmittel bevorzugen. Hier kann der Abgleich mit etablierten Referenzdatensätzen helfen, Wegehäufigkeiten und Modal Split besser an die Verhältnisse in der Bevölkerung anzupassen. Viertens ist die Panelmortalität (longitudinal dropout) zu berücksichtigen. Die Langzeitbeobachtung ist ein wesentlicher Vorteil des Trackings gegenüber Stichtagserhebungen. Allerdings scheiden Teilnehmende im Zeitverlauf aus. Die Gewichtung sollte deshalb nicht nur einzelne Datenlücken, sondern auch das Ausscheiden von Teilnehmenden über den Beobachtungszeitraum berücksichtigen, um die zeitliche Konsistenz der Ergebnisse zu wahren.

Um diesen Verzerrungen methodisch zu begegnen, kommen unter anderem folgende statistische Gewichtungsverfahren zum Einsatz:

Poststratifizierung: Die bereits erhobene Stichprobe wird anhand zentraler soziodemografischer und räumlicher Merkmale (z. B. Alter, Geschlecht, Region, Gemeindetyp) in Strata unterteilt. Die Gewichte der einzelnen Gruppen werden so angepasst, dass ihre Anteile in der Stichprobe den bekannten Strukturen der Gesamtbevölkerung entsprechen. Dieses Verfahren kann territoriale Disparitäten (z. B. Stadt-Land-Gefälle) sowie demografische Unterrepräsentationen adressieren.

Verhaltensbasierte Kalibrierung: Hierbei werden die Trackingdaten mit etablierten Referenzdatensätzen (z. B. MiD, ÖV-Panels) abgeglichen und die Gewichte entsprechend angepasst. Ziel ist die Korrektur verhaltensbezogener Verzerrungen, die beispielsweise durch eine Über- oder Unterrepräsentation spezifischer Verkehrsmittel (ÖV, aktive Mobilität), Wegelängen (z. B. Langstrecken) oder überdurchschnittlich mobiler Personen entstehen können. Die Referenzdaten dienen dabei als Orientierung und nicht als Zielwert, den die Trackingdaten reproduzieren müssen.

Raking Verfahren (Iterative Proportionale Anpassung): Dieses Verfahren passt die Gewichte iterativ entlang mehrerer Dimensionen an. Dadurch werden sie an bekannte Bevölkerungskennwerte angeglichen, was Verzerrungen in mehreren Variablen gleichzeitig reduziert. Raking wird nach der Poststratifizierung (oder unabhängig davon) angewendet und erlaubt eine simultane Berücksichtigung soziodemografischer, territorialer und mobilitätsbezogener Faktoren.

Modellierung longitudinaler Ausfälle (Attrition): Panelmortalität und Datenlücken können durch eine explizite Modellierung des Abbruchprozesses adressiert werden. Mithilfe von Attritionsmodellen (z. B. via logistische Regression oder Überlebenszeitanalyse) wird die Abbruchwahrscheinlichkeit geschätzt. Diese dient als Basis für Invers-Wahr-

scheinlichkeitsgewichte oder kalibrierte longitudinale Gewichte, um nicht-zufällige Abbrüche auszugleichen.

Diese Verfahren entsprechen etablierten Standards großskaliger Mobilitätsstudien und lassen sich grundsätzlich auf die Smartphone-basierte Datenerhebung übertragen. Dies unterstreicht, dass Repräsentativität eine methodenübergreifende Herausforderung darstellt und kein spezifisches Defizit einzelner Erhebungsinstrumente.

3.2 Datenhochrechnungsverfahren

Analog zu traditionellen Erhebungen beruht auch Smartphone-Tracking auf Stichproben, die nicht zwangsläufig die Gesamtbevölkerung abbilden. Während die statistische Gewichtung die strukturelle Zusammensetzung der Stichprobe korrigiert, dient die Datenhochrechnung als weiterer Schritt dazu, die gewichteten Daten auf die Populationsebene zu übertragen. Durch die Skalierung auf verschiedene räumliche Einheiten – von Städten über Regionen bis hin zu nationalen Aggregaten – lassen sich umfassende Mobilitätsindikatoren wie das Gesamtverkehrsaufkommen, regionale Modal Splits oder durchschnittliche Wegelängen schätzen.

Für eine belastbare Hochrechnung kommen verschiedene, sich ergänzende Ansätze zum Einsatz. Ausgangspunkt sind gewichtete Datensätze, auf deren Grundlage unter Einbeziehung soziodemografischer, räumlicher und verhaltensbezogener Merkmale aggregierte Mobilitätsindikatoren geschätzt werden. Hierfür eignen sich inferenzstatistische und prädiktive Modelle, beispielsweise Regressionsansätze, generalisierte additive Modelle oder probabilistische Imputationsverfahren. Zur Validierung und Feinjustierung können die Modellergebnisse mit aggregierten externen Daten abgeglichen werden, etwa mit Ticketingdaten, Verkehrszählungen oder zensusbasierten Pendlerstatistiken. Diese Triangulation kann dazu beitragen, die Belastbarkeit von Modal-Split-Schätzungen einzuschätzen und das Gesamtverkehrsaufkommen besser abzubilden.

Sind einzelne Räume oder Zeiträume nur dünn oder lückenhaft erfasst, können ergänzend räumliche und zeitliche Hochrechnungsverfahren eingesetzt werden. Die zeitliche Skalierung überträgt kurze Beobachtungsfenster auf vollständige Werktage, Wochen oder Saisons. Die räumliche Hochrechnung ermöglicht es, Schätzungen auf weitere geografische Einheiten zu übertragen und räumlich konsistente Indikatoren abzuleiten. Schließlich können modellbasierte synthetische Daten dazu beitragen, Datenlücken im Längsschnitt oder unterbesetzte Stichprobenzellen auszugleichen. Hierfür kommen beispielsweise probabilistisches Resampling, prädiktive Verhaltensmodellierung oder agentenbasierte Simulationen infrage, die auf beobachteten Mobilitätsmustern aufbauen.

4 Datenvalidierung und Qualitätssicherung

Die Validität von Smartphone-Tracking-Ergebnissen korreliert unmittelbar mit der Integrität und Plausibilität der erhobenen Daten. Dieser Abschnitt beschreibt zentrale Strategien zur Sicherung der Genauigkeit und Konsistenz. Die Qualitätssicherung stützt sich auf eine Kombination aus externem Benchmarking und kontinuierlicher, nutzerbasierter Validierung. Diese Prozesse sind eng mit den zuvor beschriebenen Strategien zur Repräsentativitätssicherung verzahnt und können sowohl vor als auch nach der statistischen Gewichtung erfolgen.

4.1 Externe Plausibilitätsprüfungen und Benchmarking

Der Abgleich zentraler Mobilitätsindikatoren mit anerkannten Referenzstudien dient der Beurteilung, inwieweit beobachtete Muster mit etablierten Referenzwerten übereinstimmen. Dazu lassen sich Kenngrößen wie Modal Split, Wegelängen und Wegehäufigkeiten mit Erhebungen wie „Mobilität in Deutschland“ (MiD) oder dem „System repräsentativer Verkehrsbefragungen“ (SrV) vergleichen. Ergänzend können soziodemografische und mobilitätsbezogene Merkmale, beispielsweise Pkw-Besitz oder ÖPNV-Nutzungshäufigkeit, anhand amtlicher Statistiken überprüft werden. Deutliche Abweichungen können auf systematische Verzerrungen durch Rekrutierung, Antwortausfälle oder Klassifikationsalgorithmen hinweisen und Anlass für eine Überprüfung der Stichprobenzusammensetzung, der Gewichtung oder der Datenverarbeitung geben.

Bei der Interpretation solcher Vergleiche ist jedoch zu berücksichtigen, dass auch die Referenzerhebungen mit Unsicherheiten behaftet sind. Sie können selbst Berichtsverzerrungen und eigene Fehlerrisiken aufweisen (z. B. Untererfassung kurzer Wege). Zudem bilden sie die Mobilität zu bestimmten Erhebungszeitpunkten ab, werden meist nur in mehrjährigen Abständen durchgeführt und können

daher von der aktuellen Situation abweichen. Benchmarking dient deshalb primär der Plausibilitätsprüfung und Einordnung der Ergebnisse, nicht der uneingeschränkten Anpassung an einzelne Referenzwerte.

Ergänzend zu externen Vergleichen sind interne und räumliche Konsistenzprüfungen erforderlich. So können beobachtete Mobilitätsmuster und Aktivitätsräume mit Bevölkerungsdichten und Flächennutzungsdaten verglichen werden, um die Plausibilität ihrer räumlichen Verteilung zu beurteilen. Darüber hinaus können automatische Prüfverfahren unplausible Datenpunkte identifizieren, etwa aufgrund extremer Reisegeschwindigkeiten oder unrealistischer Aufenthaltsdauern. Solche Auffälligkeiten lassen sich je nach Prüfverfahren kennzeichnen, einer manuellen Überprüfung zuführen oder nach festgelegten Regeln korrigieren.

4.2 Internes Monitoring, Nutzervalidierung und Panel-Retention

Ein wesentlicher Baustein der Qualitätssicherung ist das fortlaufende Monitoring der erfassten Daten sowie die Validierung automatisierter Wegeerkennungen durch die Nutzenden selbst (Molloy et al. 2023). Da automatisierte Tracking- und Klassifikationsverfahren systemimmanente Unschärfen aufweisen, dient direktes Nutzerfeedback als entscheidende Referenz (ground truth), um die Erkennungsgenauigkeit zu überwachen und die zugrunde liegende Algorithmik iterativ zu optimieren. Das Monitoring trägt dazu bei, das spezifische Mehrwertpotenzial des Smartphone-Trackings – die Erhebung hochauflösender Längsschnittdaten über Monate oder Jahre – voll auszuschöpfen und Datenreihen für tiefgreifende Analysen von Verhaltensänderungen und Saisonalitätseffekten zu gewinnen. Um diese Datenkontinuität zu gewährleisten, ist eine langfristige Nutzerbindung (panel retention) entscheidend. Basierend auf empirischen Projekterfahrungen skizziert dieser Abschnitt bewährte Strategien zur Förderung des Engagements und zur Sicherstellung einer dauerhaft hohen Datenqualität.

Nutzervalidierung und Klassifikationsgenauigkeit

Damit die erfassten Daten möglichst verlässlich sind, ist eine punktuelle oder gar vollständige Validierung der getrackten Daten hilfreich. Hierbei sind die interaktiven Funktionen der App so auszugestalten, dass ein möglichst qualifiziertes und aussagekräftiges Nutzerfeedback erhalten wird. In vielen Apps können Teilnehmende erkannte Wegsegmente, Verkehrsmittel oder Ticketinformationen bestätigen oder korrigieren. Diese Rückmeldungen sind hilfreich, um die Klassifikationsgüte zu quantifizieren und zu verbessern. Die aktive Einbindung kann zudem Transparenz schaffen und das Vertrauen in die Erfassungstechnologie stärken. Ergänzend können regelmäßige Push-Benachrichtigungen die Teilnehmenden dazu anregen, ihre Mobilitätsprotokolle zeitnah zu prüfen. Das kann Datenlücken verringern, das Bewusstsein für das eigene Mobilitätsverhalten schärfen und so das langfristige Interesse an der App-Nutzung festigen.

Strategien zur langfristigen Teilnahme und Nutzerbindung

Für die langfristige Teilnahme legen Beobachtungen aus der Praxis und aus vergleichbaren Panel-Studien nahe, dass transparente Kommunikation, kontinuierliches Feedback und durchdachte Anreize zusammenwirken können (Bürbaumer et al. 2024; Loder et al. 2024). Eine offene Kommunikation über Forschungsziele und Datennutzung kann Vertrauen aufbauen und das Engagement der Teilnehmenden fördern. Die Reputation der Trägerorganisation (z. B. Universität oder Verkehrsbehörde) kann dabei als Glaubwürdigkeitsanker wirken und die Teilnahmequoten erhöhen. Nach initialer Ansprache können gezielte, wiederholte Nachfassaktionen per E-Mail oder über Partnernetzwerke wie Kommunen oder NGOs helfen, inaktive Personen zurückzugewinnen und die Aktivität insgesamt aufrechtzuerhalten.

Ein weiterer Ansatz sind Rückmeldungen und spielerische Elemente. Individuelle Auswertungen zum eigenen Mobilitätsverhalten, etwa zur CO_2-Bilanz oder zum Modal Split, bieten einen direkten Mehrwert und können die Bindung der Teilnehmenden stärken. Gamification-Ansätze mit Wettbewerbselementen wie Challenges oder Ranglisten, zum Beispiel zur Anzahl der Validierungen oder Vielfalt der genutzten Verkehrsmittel, können die intrinsische Motivation ansprechen und eine verstetigte App-Nutzung fördern. Anreize lassen sich zudem dynamisch anpassen, etwa wenn sie auf das individuelle Teilnahmeverhalten reagieren oder gezielt unterrepräsentierte Gruppen adressieren. Das kann

zur Stabilisierung der Stichprobe beitragen. Eine längerfristige Anreizplanung, etwa mit Boni bei Meilensteinen oder regelmäßigen Verlosungen, kann die Motivation über ausgedehnte Erhebungszeiträume stützen. Wie stark diese Maßnahmen wirken, hängt vom Kontext der Studie ab und ist im Einzelfall zu prüfen.

5 Fazit

Smartphone-basiertes GPS-Tracking bietet erhebliche Potenziale für die Mobilitätsforschung, insbesondere durch die hochauflösende Erfassung räumlicher und zeitlicher Bewegungsmuster. Die Aussagekraft der gewonnenen Daten hängt jedoch wesentlich davon ab, inwieweit die untersuchten Personen die Zielpopulation abbilden und sich die beobachteten Mobilitätsmuster auf diese übertragen lassen. Die Annäherung an repräsentative Ergebnisse erfordert daher eine sorgfältige methodische Planung über den gesamten Forschungsprozess hinweg.

Die vorgestellten Ansätze verdeutlichen, dass die Verbesserung der Repräsentativität nicht auf eine einzelne Maßnahme reduziert werden kann, sondern ein Zusammenspiel verschiedener methodischer Schritte erfordert. Bereits bei der Rekrutierungsstrategie können kombinierte digitale, analoge und partnerbasierte Kanäle sowie gegebenenfalls probabilistische Auswahlverfahren dazu beitragen, Selektionsverzerrungen zu verringern. Gewichtung und Hochrechnung setzen anschließend an verbleibenden Abweichungen an, indem sie die Stichprobenzusammensetzung an relevanten soziodemografischen, räumlichen und verhaltensbezogenen Merkmalen ausrichten und die Übertragung der Ergebnisse auf die Zielpopulation unterstützen. Externe Referenzdaten und Datenvalidierungen durch Teilnehmende ergänzen diese Maßnahmen, indem sie die Plausibilitätsprüfung der Ergebnisse ermöglichen und gleichzeitig zur kontinuierlichen Optimierung der algorithmischen Klassifikation beitragen.

Zusammenfassend lässt sich festhalten, dass Smartphone-Tracking ein hochflexibles und erkenntnisreiches Instrument darstellt, das konventionelle Erhebungsmethoden nicht nur ergänzt, sondern in einigen Aspekten auch übertrifft. Seine Stärke liegt insbesondere in der detaillierten Erfassung tatsächlicher Bewegungsmuster über längere Zeiträume. Die Übertragbarkeit der Ergebnisse auf die Gesamtbevölkerung bleibt dabei eine eigenständige methodische Herausforderung, die sich nicht allein durch statistische Anpassungsverfahren lösen lässt. Eine abgestimmte Kombination aus geeigneten Rekrutierungsstrategien, statistischer Gewichtung, Hochrechnung und systematischer Qualitätssicherung kann dazu beitragen, die Aussagekraft der Daten zu verbessern und ihr Potenzial für unterschiedliche Anwendungskontexte besser auszuschöpfen. Auf dieser Grundlage lassen sich Studien konzipieren, die methodische Anforderungen und projektspezifische Rahmenbedingungen berücksichtigen und fundierte Beiträge zur Verkehrsplanung und Mobilitätsforschung leisten können.

Literaturverzeichnis

Berliner Verkehrsbetriebe (BVG): Fahrinfo-App – Meine Wege, [online] https://www.bvg.de/de/abos-und-tickets/alle-apps/fahrinfo-app/meine-wege [abgerufen am 16.09.2026].

Bürbaumer, M., Kirchberger, C., Berger, M. & Dorner, J. (2024): Recruitment, participant motivation and response rates in a smartphone-based travel survey: Mobility Panel in aspern Seestadt, in: Transportation Research Procedia, Bd. 76, S. 283–295, DOI: 10.1016/j.trpro.2023.12.055.

Heinonen, S., Meyer de Freitas, L., Meister, A., Schwab, L., Roth, J., Hintermann, B., Götschi, T. & Axhausen, K. W. (2024): The e-biking in Switzerland (EBIS) study: methods and dataset, in: Transportation, Online First, DOI: 10.1007/s11116-024-10552-y.

Lawson, C. T., Krans, E., Rentz, E. (Green) & Lynch, J. (2023): Emerging trends in household travel survey programs, in: Social Sciences & Humanities Open, Bd. 7, Nr. 1, DOI: 10.1016/j.ssaho.2023.100466.

Loder, A., Dahmen, V., Waldorf, I., Álvarez-Ossorio Martinez, S. & Bogenberger, K. (2024): The Mobilität.Leben Study: A 20-Month Mobility-Tracking Panel, DOI: 10.13140/RG.2.2.10552.35843.

Masse, F., Gumy, A., Schultheiss, M.-E., Kaufmann, V. & Drevon, G. (2025): The Lemanic Panel, a five-year monitoring of populations' lifestyles—focus on the first wave: recruitment phase, mobility survey and GPS tracking, in: Transportation, Bd. 52, S. 2215–2248, DOI: 10.1007/s11116-025-10634-5.

Meister, A., Bashan, N. F., Basu, R., Shen, X., Wang, R. Q. & Sevtsuk, A. (2025): The BostonWalks study: a longitudinal travel survey using smartphone tracking, in: Transportation, Bd. 52, S. 2249–2279, DOI: 10.1007/s11116-025-10637-2.

Molloy, J., Castro, A., Götschi, T., Schoeman, B., Tchervenkov, C., Tomic, U., Hintermann, B. & Axhausen, K. W. (2023): The MOBIS dataset: a large GPS dataset of mobility behaviour in Switzerland, in: Transportation, Bd. 50, Nr. 5, S. 1983–2007, DOI: 10.1007/s11116-022-10299-4.

Piras, F., Sottile, E., Tuveri, G. & Meloni, I. (2025): Exploring factors influencing car commuters' willingness to participate to an app-based Voluntary Travel Behaviour Change programme, in: Transportation Planning and Technology, Online First, DOI: 10.1080/03081060.2025.2518118.

United Nations Statistics Division (2008): Designing Household Survey Samples: Practical Guidelines, in: Statistical Papers, Series F, Nr. 98, United Nations, Department of Economic and Social Affairs, Statistics Division.

Autor*innenangaben

Dr.-Ing. Mariana Batista
motiontag GmbH
Cuvrystr. 53 (Haus F, Scaling Spaces)
10997 Berlin

E-Mail: mariana.batista@motiontag.com

Journal für Mobilität und Verkehr

ISSN 2628-4154

www.dvwg.de

Generieren, verarbeiten, veröffentlichen: Einheitlicher Umgang mit Mobilitätsdaten in interdisziplinären Forschungsclustern

Svetlana Zubareva*, Louis Stille-Hönig*, Markus Lienkamp

Siehe Autor*innenangaben

Abstract

Mit der zunehmenden Verfügbarkeit digitaler Mobilitätsdaten eröffnen sich neue Möglichkeiten für die empirische Verkehrsforschung und die evidenzbasierte Verkehrsplanung. Gleichzeitig wächst der Bedarf an transparent dokumentierten, qualitativ hochwertigen und datenschutzkonformen Dateninfrastrukturen, die heterogene Datenströme langfristig nutzbar machen. Dieser Beitrag zeigt am Beispiel des Münchner Forschungsclusters für die Zukunft der Mobilität (MCube) die Vorteile einer gemeinsamen Datenstruktur zwischen den beteiligten Projekten. Anhand von digitalen Mobilitätstagebüchern, in denen die zurückgelegten Wege erfasst werden, sowie Carsharing-Flottendaten aus unterschiedlichen Forschungsprojekten wird die zentrale Verarbeitung von Mobilitätsdaten beschrieben. Dabei wird ein mehrstufiger Bereinigungsprozess zur Harmonisierung von Formaten, zum Umgang mit fehlenden Werten sowie zur Validierung der Datensätze anhand bekannter Mobilitätsmuster dargelegt. Auch der Workflow zur Anonymisierung, der den Schutz der Teilnehmenden gewährleistet und gleichzeitig den analytischen Nutzen für Querschnitts- und Längsschnittanalysen bewahrt, wird erläutert. Der Beitrag liefert ein praxisnahes Vorgehensmodell für den Aufbau und den Betrieb einer Mobilitätsdatenbank und regt dazu an, die Verarbeitung von Mobilitätsdaten innerhalb von Forschungsclustern zu zentralisieren. Die Erkenntnisse und Methoden des Beitrags richten sich an Verkehrsforschende, Dateninfrastrukturbetreibende und Verkehrsplanende. Die dargelegten Prozesse erleichtern die Veröffentlichung von Daten sowohl für wissenschaftliche Publikationen als auch für die Bürger*innenbeteiligung und die Öffentlichkeitsarbeit.

Schlagwörter:

Datenmanagement, Mobilität, Interdisziplinarität, Forschungscluster, Datenschutz

1 Einleitung

Durch die zunehmende Verfügbarkeit digitaler Mobilitätsdaten ergeben sich stetig neue Möglichkeiten der empirischen Verkehrsforschung. Die Daten werden dabei durch unterschiedlichste Geräte wie Smartphones, Sensoren und Fahrzeuge erfasst und lassen eine detaillierte Untersuchung des Mobilitätsverhaltens zu. Allerdings ist dies nur möglich, wenn die unterschiedlichen Daten harmonisiert und kombiniert bereitgestellt werden. Denn so können Verkehrsangebote und deren Nutzung räumlich sowie zeitlich differenziert analysiert werden. Die gewonnenen Erkenntnisse können zudem als Grundlage für eine evidenzbasierte Verkehrsplanung dienen (Waldorf et al., 2026).

Die Zusammenführung dieser Daten ist allerdings mit verschiedenen Herausforderungen verbunden, da die Quellen sich hinsichtlich ihrer Formate, zeitlichen Auflösung, räumlichen Genauigkeit und inhaltlichen Struktur unterscheiden. Weiterhin wird die Datenverarbeitung häufig für jedes Forschungsprojekt neu aufgesetzt und durchgeführt. So entstehen voneinander getrennte, heterogene Datenbestände und projektspezifische Verarbeitungsschritte, welche eine spätere Zusammenführung

*korrespondierende Autor*innen

und Wiederverwendung erschweren. Dies ist insbesondere dann problematisch, wenn verschiedene Forschungsprojekte einem Cluster angehören und somit auf ein übergeordnetes Ziel hinarbeiten.

Darüber hinaus enthalten Mobilitätsdaten häufig einen direkten oder indirekten Personenbezug und müssen dementsprechend nach hohen Datenschutz-Standards verarbeitet werden. Neben der Anonymisierung ist eine nachvollziehbare Dokumentation der Verarbeitung erforderlich. Durch eine zentrale Dateninfrastruktur lassen sich einheitliche Qualitätsanforderungen umsetzen und Daten projektübergreifend verfügbar machen.

Der vorliegende Beitrag untersucht daher, wie heterogene Mobilitätsdaten innerhalb eines interdisziplinären Forschungsclusters zentral bereinigt, anonymisiert und bereitgestellt werden können und welche Vorteile sich daraus in der Praxis ergeben. Hierzu wird ein praxisnahes Vorgehensmodell für den Aufbau und Betrieb einer clusterinternen Mobilitätsdatenbank vorgestellt. Die Umsetzung wird anhand digitaler Mobilitätstagebücher und Carsharing-Flottendaten erläutert. Abschließend werden die Vorteile und Grenzen der zentralen Datenverarbeitung diskutiert. Zudem wird untersucht, inwiefern sich das Vorgehensmodell auf andere Forschungscluster und Organisationen übertragen lässt.

2 Hintergrund und Stand der Forschung

Die systematische Erfassung, Aufbereitung und Bereitstellung von Mobilitätsdaten ist ein zentrales Thema der aktuellen Verkehrsforschung (Bellini et al., 2024).

Auf nationaler Ebene verfolgt die Nationale Forschungsdateninfrastruktur (NFDI) das Ziel, Forschungsdaten disziplinenübergreifend zu sichern und zugänglich zu machen (DFG, o. J.). Eine auf den Verkehrsbereich zugeschnittene Infrastruktur fehlt in diesem Rahmen bislang, wodurch Cluster bei der Datenverarbeitung kaum auf übergeordnete Standards zurückgreifen können. Ohne diesen Bezug entwickeln sich Formate, Metadaten und Aufbereitungsregeln in jedem Projekt eigenständig, was eine spätere Zusammenführung erschwert.

Der Mobility Data Space (MDS, ehemals mobiDS), entstanden im mFUND-Förderprogramm des Bundesverkehrsministeriums, verfolgt einen anderen Ansatz. Er baut auf der Referenzarchitektur der International Data Spaces auf und regelt den kontrollierten Austausch von Daten zwischen Kommunen, Unternehmen und Dienstleistenden unter Wahrung der Datensouveränität (Fraunhofer IVI, 2022). Im Zentrum steht damit das Teilen und Aufbereiten bereits vorhandener Daten zwischen verschiedenen Akteuren, nicht die interne Harmonisierung und mehrstufige Bereinigung heterogener Rohdaten innerhalb eines Forschungsclusters.

Ein Beispiel für eine langfristig etablierte, domänenspezifische Erhebung ist die German In-Depth Accident Study (GIDAS), die seit 1999 in Hannover und Dresden sowie seit 2023 auch in München detaillierte Unfalldaten erfasst und für die Wissenschaft aufbereitet (GIDAS, 2024). GIDAS zeigt, wie konsequente, über Jahre hinweg konsistente Datenerhebung in einer klar abgegrenzten Domäne methodisch gelingen kann. Der Ansatz ist jedoch auf Unfalldaten beschränkt und nicht darauf angelegt, unterschiedliche Mobilitätsdatenquellen zusammenzuführen. Andere Mobilitätsdaten, etwa Mobilitätstagebücher oder Flottendaten, folgen anderen Erhebungslogiken und lassen sich nicht ohne Weiteres an ein solches Schema anbinden.

Die genannten Ansätze verdeutlichen drei wiederkehrende Herausforderungen. Erstens sind die Quellen heterogen, was Formate, Auflösung und Semantik betrifft. Zweitens muss die Verarbeitung personenbezogener Mobilitätsdaten datenschutzkonform erfolgen. Drittens fehlt es an einheitlichen Standards für die Erhebung und Aufbereitung. Damit bleibt offen, wie eine Forschungsinfrastruktur konkret aussehen muss, die heterogene Mobilitätsdaten innerhalb eines interdisziplinären Forschungsclusters zentralisiert, nachnutzbar macht und dabei reproduzierbar dokumentiert. Gegenüber den skizzierten Initiativen verbindet das vorgestellte Modell übergeordnete Infrastrukturkonzepte mit der konkreten, reproduzierbaren Datenverarbeitung in der Verkehrs- und Mobilitätsforschung.

3 Daten im Kontext des Forschungsclusters MCube

Das Münchner Cluster für die Zukunft der Mobilität (MCube) hat das Ziel, Mobilitätsinnovationen zu entwickeln und in die Umsetzung zu bringen. Das Vorhaben basiert auf der Initiative „Clusters4Future", welche von dem Bundesministerium für Forschung, Technologie und Raumfahrt getragen wird.



Abbildung 1: Übersicht der MCube Projekte und Standorte in den Phasen 1 und 2 (MCube, o. J.)

In diesem Zusammenhang werden über neun Jahre bis zu 45 Millionen Euro Fördermittel in drei Umsetzungsphasen bereitgestellt (BMFTR, o. J.-a). Im Jahr 2021 ist MCube mit 14 Forschungsprojekten in die erste Umsetzungsphase gestartet, welche im November 2024 durch die zweite Umsetzungsphase mit neun Forschungsprojekten abgelöst wurde. Die Projekte haben das übergeordnete Ziel, die Qualität von Luft, Raum und Zeit zu verbessern, und bauen über die Projektphasen hinweg aufeinander auf (Clusterplattform Deutschland, o. J.; MCube, o. J.). Die Forschungsprojekte sowie deren Standorte werden in Abbildung 1 dargestellt.

In München werden in diesem Zusammenhang von Carsharing über Verkehrssimulationen bis hin zu Beteiligungs- und Partizipationsverfahren verschiedenste Facetten der Mobilität berücksichtigt. Das ist eine Chance für die interdisziplinäre Entwicklung von Innovationen, bringt aber auch Herausforderungen, wie das Management heterogener Daten, mit sich. MCube zeichnet sich dabei durch seinen Anwendungsbezug aus (MCube, o. J.). Forschungs- und Innovationsansätze wie die Incentivierung von klimafreundlicher Mobilität oder das Teilen privater Fahrzeuge werden in den Projektphasen umgesetzt und datenbasiert evaluiert (BMFTR, o. J.-b). Dieser Fokus auf lokale Umsetzungs- und Transformationsprozesse wird auch im wissenschaftlichen Diskurs als wichtiges Instrument für das Schaffen und Verstetigen von Innovationen beschrieben (Engels et al., 2019).

Die in der ersten Umsetzungsphase gewonnenen Daten wurden bereits im Rahmen von wissenschaftlichen Arbeiten untersucht und veröffentlicht (Loder et al., 2024; Nachtigall et al., 2025; Waldorf et al., 2026). Die Daten wurden in diesem Zusammenhang clusterübergreifend verarbeitet und für die Veröffentlichung aufbereitet. Dies wird auch in der zweiten Umsetzungsphase fortgeführt und soll in diesem Artikel am Beispiel der Projekte „Komfortables und effizientes Fahrzeug-Sharing auf Privatgrund (ComfficientShare 2.0)" und „Reallabor Future Mobility Campus Garching (ReFuMoLab_Garching)" näher dargestellt werden. In dem Forschungsprojekt ComfficientShare 2.0 werden Carsharing-Fahrten mithilfe von GPS-Loggern und Buchungsdaten aus einer Carsharing-App erfasst. Darüber hinaus werden mithilfe der TUM-Mobility-App digitale Mobilitätstagebücher der Nutzenden erstellt, in denen die zurückgelegten Wege fortlaufend mit Zeitpunkt, Länge und Verkehrsmittel erfasst werden. Die Kombination dieser Datensätze ermöglicht es, die Nutzung des Carsharings im Zusammenhang mit der Alltagsmobilität zu untersuchen. ReFuMoLab_Garching stellt den Projektteilnehmenden ebenfalls die TUM-Mobility-App zur Erstellung digitaler Mobilitätstagebücher zur Verfügung (MCube, o. J.).

4 Aufbau und Betrieb der Mobilitätsdatenbank

Bei der Mobilitätsdatenbank handelt es sich um einen selbst gehosteten PostgreSQL-Server mit der räumlichen Erweiterung PostGIS (PostgreSQL Global Development Group, 2024; PostGIS Project Steering Committee, 2024). Dieser ist modular über projektspezifische Schemata erweiterbar und wird durch einen zentralen Admin verwaltet.

4.1 Datenintegration und Harmonisierung

Die erhobenen Daten liegen in unterschiedlichen Formaten vor und werden über verschiedene Schnittstellen bereitgestellt (Tabelle 1). Ziel der Integration ist es, diese heterogenen Quellen in einer einheitlichen Datenstruktur abzubilden, in der jede Fahrt (Trip) als harmonisierter Datensatz mit Geometrie, Zeit und Länge vorliegt.

Tabelle 1: Daten, Schnittstellen und Formate

Daten	Schnitt-stelle	Inhalt
Carsharing Buchungen	Betreiber-plattform	Start / Ende / User-ID / Fahrzeug / Rolle
Fahrzeug-positionen	Betreiber-plattform	≈ 1 Position/Minute; Position / Uhrzeit / Fahrzeug
Fahrzeug-positionen	GPS-Logger	1 Hz; Position / Uhrzeit / Fahrzeug
Mobilitätstage-buch – Nutzende	TUM-Mobility-App-API	User-ID / E-Mail
Mobilitätstage-buch – Fahrten	TUM-Mobility-App-API	Startzeit / Endzeit / Länge / Verkehrsmittel
Online-Umfrage	Lime Survey	E-Mail / Soziodemographische Daten

Den Ausgangspunkt der Datenerhebung durch das Mobilitätstagebuch bildet eine Online-Umfrage über Lime Survey. Dort melden sich die Teilnehmenden an und erhalten anschließend die Zugangsdaten für die TUM-Mobility-App. Neben den Kontaktdaten werden dabei soziodemographische Merkmale erhoben. Über eine pseudonyme Kennung (UUID) können diese Angaben anschließend mit den erfassten Mobilitätsmustern verknüpft werden. Hierdurch lassen sich Unterschiede im Mobilitätsverhalten verschiedener Personengruppen untersuchen. Die Datenquellen sind direkt an die zentrale Datenbank angebunden. Die technische Umsetzung der Anonymisierung und Bereitstellung wird im Abschnitt „Datenbereitstellung und Nutzung“ erläutert.

Für ComfficientShare 2.0 umfasst das Schema drei Tabellen und eine Ansicht: eine Tabelle mit den Buchungsinformationen, eine Tabelle mit den Fahrzeugpositionen aus GPS-Loggern und Betreiberplattform sowie eine Tabelle, in der Buchungen und Positionen zu Fahrten verknüpft werden. Da die hochfrequenten Positionsdaten Rückschlüsse auf Personen sowie Wohn- und Arbeitsorte zulassen, werden für Analysen ausschließlich anonymisierte Ansichten bereitgestellt. Der Datenabruf und das Laden in die Datenbank erfolgen weitgehend automatisiert über einen weiteren Server im selben Netzwerk.

Analog dazu werden die Daten der digitalen Mobilitätstagebücher täglich von einem externen Dienstleister bereitgestellt und in einem eigenen Schema gespeichert. Dieses enthält eine Tabelle mit den registrierten Nutzenden und eine Tabelle mit den Fahrten. Auch hier werden die Fahrten nach denselben Maßstäben anonymisiert und über eine Ansicht bereitgestellt.

4.2 Mehrstufiger Bereinigungsprozess

Nach der erfolgreichen Bereitstellung der Daten muss deren Qualität geprüft und sichergestellt werden. Dies gilt insbesondere für die Carsharing-Daten, da die digitalen Mobilitätstagebücher bereits in aufbereiteter Form bereitgestellt werden. Die Methodik zur Qualitätssicherung der Carsharing-Daten erfolgt in drei Schritten:

Schritt 1:

Zunächst werden unvollständige Buchungen und fehlende Positionswerte behandelt. Als Fahrt gilt nur eine abgeschlossene Buchung mit vorhandener Start- und Endzeit, wobei die Endzeit nach der Startzeit liegen muss. Zusammenhängende Positionsfolgen werden anhand von Standzeitlücken von mehr als 15 Minuten in einzelne Fahrsegmente getrennt (Schuessler & Axhausen, 2009).

Schritt 2:

Fahrten mit leerer oder ungültiger Geometrie, einer Länge unter 100 m oder einer Dauer unter zwei Minuten werden als Fehler verworfen. Die Schwellenwerte wurden anhand des vorliegenden Datensatzes sowie der Daten aus ComfficientShare 1.0 entwickelt. Da nicht alle Fahrzeuge mit einem GPS-

Logger ausgestattet sind, liegen für einzelne Fahrzeuge nur die geringer aufgelösten Positionsdaten der Betreiberplattform vor. Die Positionsdaten werden zunächst technisch bereinigt. Hierzu werden Koordinaten entfernt, deren Breitengrad außerhalb des Intervalls von −90° bis 90° oder deren Längengrad außerhalb des Intervalls von −180° bis 180° liegt (Zheng, 2015).

Schritt 3:

Die Punkte beider Quellen werden mit demselben Map-Matching-Verfahren verarbeitet: Ein auf einem Server im lokalen Netzwerk betriebener GraphHopper-Dienst ordnet die GPS-Punkte dem dort vorgehaltenen OpenStreetMap-Straßengraphen zu und rekonstruiert den Fahrtverlauf (GraphHopper GmbH, o. J.; Newson & Krumm, 2009; OpenStreetMap contributors, 2024). Fehlen vollständig verwertbare Positionen, bleibt die Buchung ohne Geometrie erhalten.

Die Ergebnisse des Bereinigungsprozesses werden in Abbildung 2 dargestellt. In der oberen Karte werden die Rohdaten dargestellt, während die untere Karte die bereinigten Trajektorien zeigt. Die räumliche Verteilung lässt erkennen, dass sich die Fahrten entlang der überregionalen Hauptverkehrsachsen konzentrieren.

Die Plausibilität der Carsharing-Fahrten wird anhand der erfassten Daten von Nachtigall et al. (2025) hinsichtlich der charakteristischen Fahrtlängen und -dauern geprüft. Weichen die beobachteten Muster deutlich von diesen Referenzen ab, so deutet dies auf verbleibende Fehler im Bereinigungsprozess hin. Dies wird gleichermaßen für die Mobilitätstagebücher mithilfe der Daten aus dem Vorgängerprojekt Mobilität.Leben durchgeführt und in Abbildung 3 exemplarisch anhand der Wegehäufigkeit im Tagesverlauf dargestellt. Es zeigt sich, dass die beobachteten Fahrtlängen und -dauern den aus großräumigen Erhebungen wie Mobilität in Deutschland bekannten Mustern entsprechen (Nobis & Kuhnimhof, 2018). Bei der Wegehäufigkeit über den Tag werden zwar Unterschiede sichtbar, doch dies lässt sich durch den Fokus auf die Pendlermobilität in der aktuellen MCube-Studie erklären.

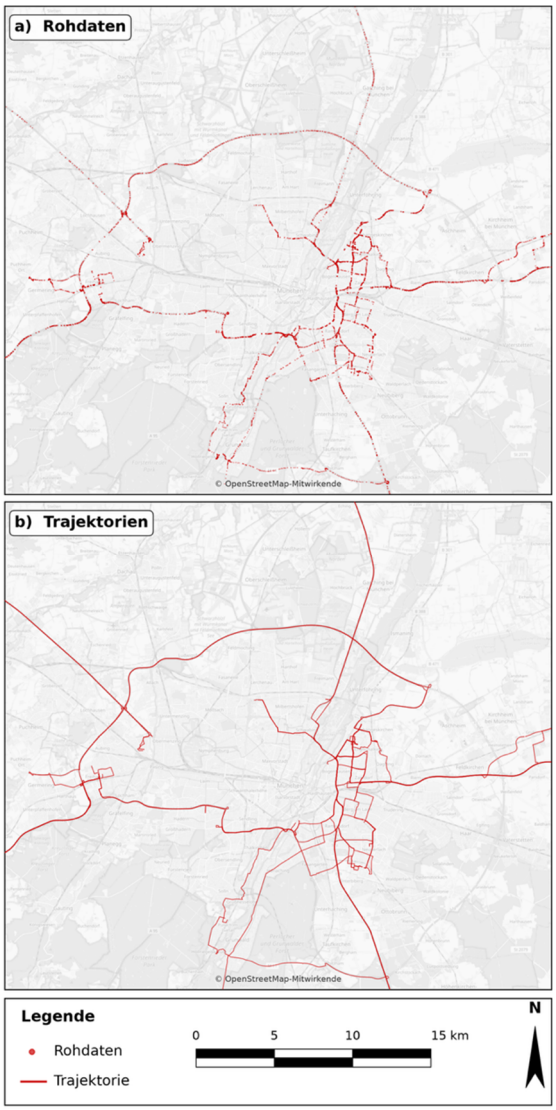

Abbildung 2: Aufbereitung der ComfficientShare 2.0 Fahrten in München

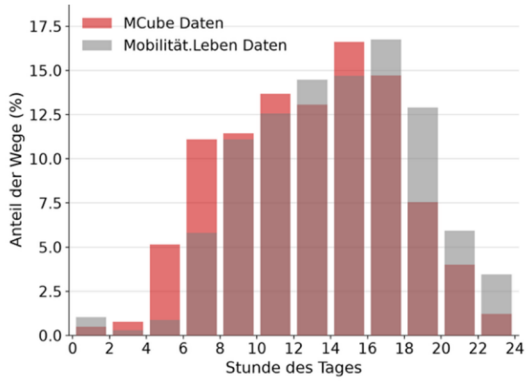

Abbildung 3: Verteilung der Wege im Tagesverlauf nach den zugrunde liegenden Datensätzen

4.3 Anonymisierung und Datenschutz

Im Rahmen des Projekts werden hochsensible Mobilitätsdaten erhoben. Daher ist eine Anonymisierung erforderlich, bevor diese für Forschungszwecke genutzt werden können. Die konkrete Ausgestaltung der Anonymisierung ist projektzielabhängig, da jedes Projekt unterschiedliche Forschungsfragen verfolgt. So erfordert beispielsweise ein Projekt zum Parkverhalten eine andere Datenaggregation als ein Projekt zur Mitarbeitendenmobilität. Um dennoch eine vergleichbare Datenbasis über alle Projekte hinweg zu gewährleisten, wurden die projektübergreifenden Anonymisierungsstufen noch vor der Datenerhebung definiert. Dadurch steht jedem Projekt dieselbe Datenmenge und -granularität zur Verfügung.

Die MCube-Datenbank setzt die Anonymisierung nicht als Einzelschritt um, sondern abgestuft, beginnend bei der Erhebung. Der gesamte Workflow der Datenaufbereitung ist in Abbildung 4 dargestellt. Grundlage ist die Pseudonymisierung. Alle Analysebestände sind über UUIDs referenziert, während direkte personenbezogene Merkmale wie E-Mail-Adressen ausschließlich in der separaten Nutzenden-Tabelle liegen. Sie ist von den analytischen Datensätzen getrennt, sodass Auswertungen ohne direkten Personenbezug erfolgen.

Das übergeordnete Ziel der Anonymisierung besteht darin, die Nachvollziehbarkeit einzelner Teilnehmender auszuschließen. Es soll verhindert werden, dass sich die Bewegungen einzelner Personen rekonstruieren lassen. Daher wurde jedem Teilnehmenden vor der Datenerhebung eine eindeutige UUID zugeordnet. Sämtliche weiteren Datenbestände lassen sich ausschließlich über diese UUID miteinander verknüpfen, ohne dass ein Rückschluss auf die jeweilige Person möglich ist.

Auf Basis dieser Pseudonymisierung werden drei Datenformate bereitgestellt. Zum einen die Umfragedaten, bei denen personenbezogene Angaben (E-Mail-Adresse sowie Vor- und Nachname) entfernt und durch die UUID ersetzt wurden. Zum anderen die GPS-Daten, die in Trajektorien und das Mikromobilitätsverhalten unterteilt werden. Trajektorien sind Datensätze mit der Geometrie einzelner Fahrten. Alle Trajektorien werden anonymisiert, indem ihre Start- und Endpunkte als Geohash der Genauigkeitsstufe 7 codiert werden (Balkić et al., 2012). Zusätzlich werden Dauer und Länge der Fahrt sowie Start- und Endzeitpunkt, jeweils auf 15 Minuten gerundet, bereitgestellt. Durch die räumliche und zeitliche Aggregation werden Rückschlüsse auf einzelne Personen verhindert.

Das Mikromobilitätsverhalten umfasst im Gegensatz zu den Trajektorien einzelne GPS-Punkte innerhalb eines definierten Radius. Diese liegen somit als Punktdaten und nicht als trajektorienbasierte Verläufe vor. Diese Daten sind ausschließlich für Teilnehmende innerhalb des jeweils betrachteten Projekts relevant. Die Nutzung ist somit projektspezifisch und auf den entsprechenden Kreis der Teilnehmenden beschränkt.

4.4 Datenbereitstellung und Nutzung

Abschließend werden die aufbereiteten Daten zentral vorgehalten und projektbezogen für die Nut-

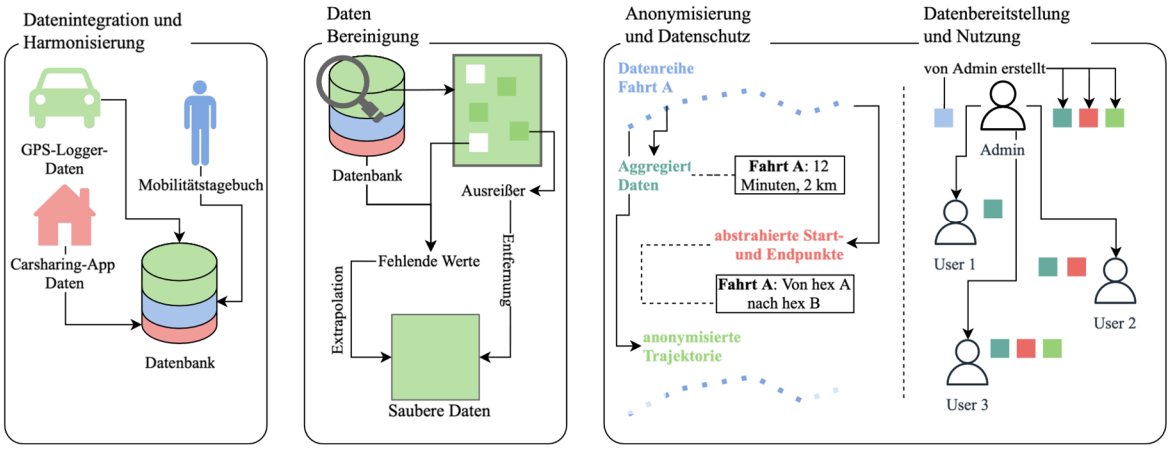

Abbildung 4: Workflow zur Datenaufbereitung

zung bereitgestellt. Die für die Weitergabe vorgesehenen, anonymisierten Datenbestände (vgl. Abschnitt Anonymisierung) werden dabei als Datenbankansichten (Views) anstatt physischen Tabellen angelegt. Dieser Ansatz hat den Vorteil, dass die Anonymisierungs- und Aggregationslogik zentral in den Views gepflegt wird. Sobald neue Rohdaten in die Datenbank importiert werden, stehen sie automatisch in der korrekt aggregierten, anonymisierten Form bereit, wodurch pro neuem Datensatz kein manueller Aufbereitungsschritt erforderlich ist. Die zuvor eingeführten Datenaggregationsstufen werden projektbezogen als entsprechende Ansichten abgebildet, sodass jedes Projekt auf die für seine Forschungsfrage vorgesehenen Datenformate zugreift. Die Zugriffskontrolle erfolgt als rollenbasierte Zugriffskontrolle (Role-Based Access Control, RBAC) (Sandhu et al., 1996): Ein Administrator weist den Datenbanknutzenden Rollen zu und regelt auf dieser Grundlage, welche Ansichten von wem eingesehen werden dürfen, wobei sich die Berechtigung an der jeweiligen Projektzugehörigkeit ausrichtet.

Die über die Views bereitgestellten Daten dienen nicht nur der projektbezogenen Forschungsauswertung, sondern finden darüber hinaus in unterschiedlichen Nutzungsszenarien Verwendung. So fließen sie in Publikationen ein, in denen Ergebnisse aus der Mobilitätsforschung der wissenschaftlichen Gemeinschaft sowie der Fachöffentlichkeit zugänglich gemacht werden. Zudem bilden sie eine Grundlage für die Bürger*innenbeteiligung. Auf Basis der anonymisierten und aggregierten Bestände lassen sich verkehrliche Sachverhalte evidenzbasiert und verständlich aufbereiten, um Bürgerinnen und Bürger in Planungs- und Entscheidungsprozesse einzubinden. Schließlich kommen die Daten auch in der Öffentlichkeitsarbeit zum Einsatz, etwa um Forschungsergebnisse und daraus abgeleitete Erkenntnisse einer breiteren Öffentlichkeit zu vermitteln. In allen Fällen greifen die Nutzenden ausschließlich auf die anonymisierten Views zu, sodass die Trennung zwischen Rohdatenhaltung und Weiterverwendung stets gewahrt bleibt.

5 Diskussion

Das vorgestellte Vorgehensmodell zeigt, dass sich heterogene Mobilitätsdaten innerhalb eines interdisziplinären Forschungsclusters zentral integrieren, bereinigen und nachnutzbar aufbereiten lassen. Am Beispiel der Projekte ComfficientShare 2.0 und ReFuMoLab_Garching wird deutlich, wie Carsharing-Flottendaten und appgestützte Mobilitätstagebücher in einem gemeinsamen PostgreSQL/PostGIS-basierten Datenmodell zusammengeführt werden. Der Mehrwert einer solchen Zentralisierung liegt weniger in der Datenhaltung als in der einheitlichen Bereinigungs- und Anonymisierungskette: Sie reduziert Doppelarbeiten zwischen Cluster-Projekten, vereinfacht die Einhaltung der clusterweiten Datenschutzvereinbarung und macht die Prozessschritte für Publikationen und Nachnutzende transparent nachvollziehbar. Es zeigt sich, dass die Datenverarbeitung über die Projekte hinweg einheitlich gelingt, wodurch die Funktionsfähigkeit des Prozesses bereits im laufenden Betrieb deutlich wird.

Dem stehen Grenzen gegenüber. Der Integrationsaufwand wächst mit jeder neuen Datenquelle. E-Scooter-Telematik, Maut- oder mobilfunkbasierte Verkehrsdaten erfordern jeweils eine eigene Integration, Schema-Mappings und Bereinigungsregeln. Die Validierung an Referenzdaten ist nur dort möglich, wo geeignete Erhebungen wie MiD, SrV oder Mobilität.Leben vorliegen. Darüber hinaus können rechtliche Fortschreibungen etablierte Anonymisierungs-Workflows nachträglich infrage stellen. Zudem setzt das Modell eine funktionierende Cluster-Governance voraus: Nur wenn alle Beteiligten ihre Daten tatsächlich zentral einspeisen, entfaltet sich der skizzierte Mehrwert.

Die Übertragbarkeit auf vergleichbare Strukturen, etwa andere BMFTR-Zukunftscluster oder kommunale Datenräume, ist grundsätzlich gegeben, da die Prozesse generisch gehalten sind. Bei stark kommerziell getriebenen Datenräumen mit Datensouveränität zwischen eigenständigen Akteuren stößt das Modell jedoch an paradigmatische Grenzen (vgl. Mobility Data Space). Als zentrale Erkenntnis lässt sich festhalten: Die frühzeitige Festlegung auf ein gemeinsames Datenmodell, Datenschutzvereinbarungen und Anonymisierungs-Workflow ist entscheidend für einen effizienten Forschungsbetrieb.

6 Fazit und Ausblick

Der Beitrag hat ein Vorgehensmodell für den Aufbau und Betrieb einer clusterinternen Mobilitätsdatenbank am Beispiel von MCube vorgestellt. Es verbindet die Integration heterogener Quellen in ei-

nem einheitlichen Datenmodell mit einer mehrstufigen Bereinigung und einem Anonymisierungs-Workflow, der den Schutz der Teilnehmenden mit dem analytischen Nutzen für Querschnitts- und Längsschnittanalysen in Einklang bringt. Die Validierung anhand bekannter Mobilitätsmuster bestätigt, dass die zentral aufbereiteten Daten den erwarteten Mustern entsprechen. Zugleich werden Grenzen sichtbar, die aus der automatisierten Erkennung und der Heterogenität der Quellen erwachsen.

Daraus ergibt sich die Empfehlung, Mobilitätsdaten innerhalb von Organisationen und Forschungsclustern zentral zu verarbeiten und Datenbereinigung sowie Anonymisierung als gleichberechtigte Bestandteile des Forschungsdatenzyklus zu verankern, nicht als nachgelagerte Schritte.

Weitere Studien sollten die Erweiterung um neue Datenquellen (ÖPNV-Soll/Ist-Daten, weitere Sharing-Angebote, mobilfunkbasierte Daten), die Automatisierung von Schema-Mapping sowie die Ausreißererkennung etwa durch ML-gestützte Verfahren adressieren. Auch die Verankerung der FAIR-Prinzipien über Metadatenstandards und die Anbindung an übergeordnete Initiativen wie NFDI und Datenräume sind relevante Forschungsfelder.

Danksagung

Die vorliegende Arbeit wurde durch die Förderung des Bundesministeriums für Forschung, Technologie und Raumfahrt (BMFTR) ermöglicht. Die Förderung ist Teil der Initiative Clusters4Future und ihres Mobilitätsclusters MCube (Förderkennzeichen: 03ZU2105CA).

Literaturverzeichnis

Balkić, Z., Šoštarić, D., & Horvat, G. (2012). GeoHash and UUID identifier for multi-agent systems. In G. Jezic, M. Kusek, N.-T. Nguyen, R. J. Howlett & L. C. Jain (Hrsg.), Agent and Multi-Agent Systems: Technologies and Applications (LNCS 7327, S. 290–298). Springer. https://doi.org/10.1007/978-3-642-30947-2_33

Bellini, P., Bilotta, S., Collini, E., Fanfani, M., & Nesi, P. (2024). Data sources and models for integrated mobility and transport solutions. Sensors, 24(2), 441. https://doi.org/10.3390/s24020441

Bundesministerium für Forschung, Technologie und Raumfahrt [BMFTR]. (o. J.-a). Eckpunkte der Förderung. Clusters4Future – Zukunftscluster-Initiative. https://www.clusters4future.de/foerderinitiative/eckpunkte-der-foerderung

Bundesministerium für Forschung, Technologie und Raumfahrt [BMFTR]. (o. J.-b). Zukunftscluster MCube. Clusters4Future – Zukunftscluster-Initiative. https://www.clusters4future.de/die-zukunftscluster/die-zukunftscluster-der-ersten-wettbewerbsrunde/mcube

Clusterplattform Deutschland. (o. J.). Münchner Cluster für die Zukunft der Mobilität in Metropolregionen (MCube). https://www.clusterplattform.de/CLUSTER/Redaktion/DE/Cluster/MCube_Zukunftscluster.html

Deutsche Forschungsgemeinschaft [DFG]. (o. J.). Nationale Forschungsdateninfrastruktur (NFDI). https://www.dfg.de/de/foerderung/foerderinitiativen/nfdi

Engels, F., Wentland, A., & Pfotenhauer, S. M. (2019). Testing future societies? Developing a framework for test beds and living labs as instruments of innovation governance. Research Policy, 48(9), 103826. https://doi.org/10.1016/j.respol.2019.103826

Fraunhofer-Institut für Verkehrs- und Infrastruktursysteme IVI. (2022). Mobility Data Space: Ein sicherer Datenraum für die souveräne und plattformübergreifende Bewirtschaftung von Mobilitätsdaten (Whitepaper, 2. Aufl.). https://www.mobility-dataspace.de/content/dam/ivi/mobility-data-space/documents/Mobility_Data_Space_DE_20220603_web.pdf

GIDAS – German In-Depth Accident Study. (2024). About GIDAS. https://www.gidas.org/about-en.html

GraphHopper GmbH. (o. J.). GraphHopper Routing Engine [Software]. https://github.com/graphhopper

Loder, A., Dahmen, V., Waldorf, I., Álvarez-Ossorio Martinez, S., Cantner, F., & Bogenberger, K. (2024). The Mobilität.Leben study: A 20-month mobility-tracking panel (Arbeitsberichte Mobility Policy, Bd. 2). Technische Universität München. https://media-tum.ub.tum.de/1751562

MCube. (o. J.). MCube – Münchner Cluster für die Zukunft der Mobilität in Metropolregionen. Technische Universität München. https://mcube-cluster.de/

Nachtigall, N., Dietermann, B., & Lienkamp, M. (2025). Design and evaluation of a closed-group, electric residential car-sharing system – Findings from a case study in Munich, Germany. Cleaner and Responsible Consumption, 19, 100340. https://doi.org/10.1016/j.clrc.2025.100340

Newson, P., & Krumm, J. (2009). Hidden Markov map matching through noise and sparseness. In Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (S. 336–343). ACM. https://doi.org/10.1145/1653771.1653818

Nobis, C., & Kuhnimhof, T. (2018). Mobilität in Deutschland – MiD: Ergebnisbericht. Studie von infas, DLR, IVT und infas 360 im Auftrag des Bundesministeriums für Verkehr und digitale Infrastruktur. https://www.mobilitaet-in-deutschland.de/archive/pdf/MiD2017_Ergebnisbericht.pdf

OpenStreetMap contributors. (2024). Planet dump [Datensatz]. https://planet.openstreetmap.org

PostGIS Project Steering Committee. (2024). PostGIS documentation. https://postgis.net/documentation/

PostgreSQL Global Development Group. (2024). PostgreSQL 16 documentation. https://www.postgresql.org/docs/16/

Sandhu, R. S., Coyne, E. J., Feinstein, H. L., & Youman, C. E. (1996). Role-based access control models. Computer, 29(2), 38–47. https://doi.org/10.1109/2.485845

Schuessler, N., & Axhausen, K. W. (2009). Processing raw data from Global Positioning Systems without additional information. Transportation Research Record, 2105(1), 28–36. https://doi.org/10.3141/2105-04

Waldorf, I., Loder, A., Müller, J., Dahmen, V., & Bogenberger, K. (2026). Potential for commuting demand management: Analysis of day-to-day modal variability based on a 1-year GPS-tracking dataset. Transportation Research Record. https://doi.org/10.1177/03611981251356498

Zheng, Y. (2015). Trajectory data mining: An overview. ACM Transactions on Intelligent Systems and Technology, 6(3), Artikel 29. https://doi.org/10.1145/2743025

Autor*innenangaben

Svetlana Zubareva
Technische Universität München
TUM School of Engineering and Design
Lehrstuhl für Fahrzeugtechnik
Boltzmannstraße 15
85748 Garching b. München

E-Mail: s.zubareva@tum.de

Louis Stille-Hönig
Technische Universität München
TUM School of Engineering and Design
Lehrstuhl für Fahrzeugtechnik
Boltzmannstraße 15
85748 Garching b. München

E-Mail: louis.stille-hoenig@tum.de

Markus Lienkamp
Technische Universität München
TUM School of Engineering and Design
Lehrstuhl für Fahrzeugtechnik
Boltzmannstraße 15
85748 Garching b. München

E-Mail: lienkamp@tum.de

Autor*innenbeitrag

Svetlana Zubareva und Louis Stille-Hönig haben zu gleichen Teilen zu dieser Arbeit beigetragen und teilen sich die Erstautorenschaft. Sie waren verantwortlich für Konzeptualisierung, Methodik, Software, formale Analyse, Datenkuratierung und Visualisierung, verfassten den Originalentwurf und wirkten an der Überprüfung und Bearbeitung des Manuskripts mit. Markus Lienkamp hat zur Konzeption und wissenschaftlichen Einordnung der Arbeit beigetragen, die Untersuchung im Rahmen des Forschungsprojekts wissenschaftlich betreut und das Manuskript kritisch überarbeitet. Alle Autor*innen haben die eingereichte Fassung gelesen und ihr zugestimmt.

Journal für Mobilität und Verkehr

ISSN 2628-4154

www.dvwg.de

Mobilitätsdaten als Grundlage für eine ganzheitliche ÖPNV-Planung: Die PIMO-Plattform

Marlene Grebe, Michael Ortgiese*, Pradeep Burla, Kerstin Stark

Siehe Autor*innenangaben

Abstract

Mobilitätsdaten sind ein zentraler Hebel für die Weiterentwicklung der Verkehrsplanung und bilden die Grundlage für effiziente, integrierte und evidenzbasierte Planungsprozesse. Technologische Entwicklungen in den Bereichen Sensorik, Vernetzung von Mobilitätssystemen und IoT sowie neue gesetzliche Anforderungen, beispielsweise im Rahmen der IVS-Richtlinie, führen zu einer wachsenden Verfügbarkeit offener Mobilitätsdaten. Vor diesem Hintergrund braucht es Tools und Plattformen, die Daten integrieren, verarbeiten und für die Planung niedrigschwellig nutzbar machen. Die im Forschungsprojekt entwickelte ÖPNV-Planungsplattform PIMO zielt darauf ab, die strategische mit der operativen Planung zu verknüpfen und die Bereitstellung eines attraktiven, nachhaltigen und resilienten ÖPNV-Angebots zu unterstützen. Hierfür stellt sie ein technisches System zur Verfügung, das zentrale Komponenten integriert: (1) ein Datenmanagement zur Integration und Verarbeitung von Daten, zur Sicherung ihrer Qualität sowie zur Nutzung von Datenräumen, um die Souveränität der Akteure zu wahren; (2) die Integration von Verkehrsmodellen verschiedener Planungsebenen (makroskopisch, mesoskopisch, mikroskopisch); (3) die Bereitstellung von Daten für verschiedene Akteure über angepasste Nutzer-oberflächen einschließlich der Anbindung bestehender Fahr- und Dienstplansysteme. Die PIMO-Plattform wird auf der Grundlage von Anwendungsfällen in der Region Oberspreewald-Lausitz entwickelt. Erste Erkenntnisse aus der Anwendung zeigen, dass die Plattform in Kombination mit einer konsistenten Datenbasis und insbesondere der Integration von Auslastungsdaten dazu beitragen kann, die Lücke zwischen der strategischen Netzplanung und der operativen Angebotsplanung zu schließen und eine ganzheitliche, zukunftsfähige ÖPNV-Planung zu unterstützen.

Schlagwörter:

ÖPNV-Planung, digitale Planungsplattform, Mobilitätsdaten

1 Einleitung: Treiber für eine neue Planungsinfrastruktur

Die Verkehrs- und Mobilitätsplanung steht zunehmend vielfältigen und komplexen Anforderungen gegenüber. Leitbilder und Rechtsvorschriften auf EU- und Bundes- und Landesebene fordern die Umsetzung von Klimaschutzzielen, die Förderung nachhaltiger, intelligenter Mobilität sowie die Etablierung integrierter Planungsansätze. Besonders hervorzuheben ist die EU-Verordnung zum transeuropäischen Verkehrsnetz (TEN-V), die städtische Knoten zur Aufstellung von nachhaltigen urbanen Mobilitätsplänen (SUMPs) und zur Übermittlung von Mobilitätsdaten verpflichtet und gleichzeitig einen ganzheitlichen Planungsansatz für das gesamte „funktionale Stadtgebiet“ und über Sektoren hinweg fordert (Verordnung (EU) 2024/1679). Insbesondere eine vertikal und horizontal integrierte Planung, die über Stadt- und Kreisgrenzen hinausgeht und Wirtschaftsverkehr, Energieinfrastruktur und Raumplanung mit einbezieht, erfordert ein digital vernetztes Planungssystem. Die Komplexität und wachsende Anzahl der zu beteiligenden Akteure in der Verkehrsplanung werden zudem durch die Integration neuer

*korrespondierender Autor

Mobilitätslösungen wie automatisierter Fahrzeuge, Sharing-Angebote und Bedarfsverkehre verstärkt. Eine weitere Anforderung mit zunehmender Relevanz besteht in der Resilienz und Sicherheit von Verkehrssystemen. Zunehmende Extremwetterereignisse und eine hybride geopolitische Bedrohungslage erfordern die Integration der zivilen mit der militärischen Planung, damit im Krisenfall schnell reagiert werden kann.

Mit diesen zunehmenden und neuen Anforderungen muss sich auch das Planungssystem weiterentwickeln. Erforderlich ist ein agiler, integrierter und digitaler Planungsprozess, der Akteure aus unterschiedlichen Bereichen einschließt und flexibel auf aktuelle Entwicklungen reagieren kann. Mobilitätsdaten bilden hier eine zentrale Grundlage. Sie ermöglichen es, das Verkehrssystem zunehmend digital abzubilden und in Kombination mit Modellen und Plattformen Szenarien zu simulieren, Wirkungen zu antizipieren und so evidenzbasierte Entscheidungen zu treffen. Die Verfügbarkeit von Mobilitätsdaten steigt zunehmend. In den vergangenen Jahren wurden insbesondere auf EU-Ebene mit dem Europäischen Datengesetz (Verordnung (EU) 2023/2854), der IVS-Richtlinie (Richtlinie (EU) 2023/2661), der TEN-V-Verordnung und der Daten-Governance-Verordnung (Verordnung (EU) 2022/868) Vorschriften zum Umgang mit Mobilitätsdaten erlassen, die Standards für die Erhebung, die Übermittlung und den Austausch von Daten festlegen. Die IVS-Richtlinie wird in Deutschland durch das Intelligente-Verkehrssysteme-Gesetz (IVSG) umgesetzt. Es sieht die Einführung definierter intelligenter Verkehrssysteme (IVS) vor, verpflichtet die zuständigen Stellen zur Bereitstellung definierter Mobilitätsdaten nach festgelegten Standards und regelt den Betrieb des nationalen Zugangspunktes, über den Daten bereitgestellt werden müssen (IVSG §§ 3-9, 2026). Die dynamische Entwicklung im regulatorischen Bereich trägt dazu bei, dass eine immer breitere Basis offener Mobilitätsdaten zur Verfügung steht, wodurch sich neue Potenziale für die digitale Verkehrsplanung eröffnen.

Um die vorhandenen Daten für die Planung nutzbar zu machen, sind entsprechende Infrastrukturen und Tools notwendig. Aktuelle Herausforderungen bestehen darin, heterogene Datenbestände zu harmonisieren, die Interoperabilität zwischen Systemen herzustellen und die Arbeitsprozesse von Akteuren aus unterschiedlichen Bereichen der strategischen und operativen Planung zu verzahnen. Einen Lösungsansatz bietet PIMO, eine Software-Plattform zur kollaborativen Planung von ÖPNV-Angeboten, die im Rahmen eines Forschungsprojekts vom Deutschen Zentrum für Luft- und Raumfahrt (DLR) entwickelt wird.

2 Herausforderungen in der ÖPNV-Planung

Der ÖPNV leistet einen entscheidenden Beitrag zur Daseinsvorsorge (BMV, 2025). Neben der Sicherstellung der alltäglichen Erreichbarkeit und dem Beitrag zur Mobilitätswende, spielt er auch in Krisenzeiten eine zentrale Rolle, indem er als potenziell kritische Infrastruktur die grundlegende Mobilität aufrecht erhält (BSI-KritisV § 9, 2016). Der ÖPNV steht vor der Herausforderung, trotz begrenzter finanzieller Mittel und neuer gesetzlicher Anforderungen ein attraktives, nachhaltiges und resilientes Angebot zur Verfügung zu stellen. ÖPNV-Planungen erfordern die Zusammenarbeit verschiedener Organisationen, Verwaltungsebenen und technischer Systeme. Demnach ist eine Vielzahl an Akteuren wie Aufgabenträger, Verkehrsunternehmen, Straßenbaulastträger sowie Stadt- und Regionalplanende beteiligt, die jeweils in ihrem eigenen Zuständigkeitsbereich arbeiten aber zur Erfüllung gemeinsamer Ziele aufeinander angewiesen sind. Die Zahl der relevanten Akteure nimmt durch das Aufkommen neuer Sharing- und On-Demand-Anbieter, die Elektrifizierung und Automatisierung von Angeboten sowie die Digitalisierung von Mobilitätsdiensten (beispielsweise in Form von Mobility as a Service (MaaS)-Plattformen) weiter zu. Insbesondere bei der Planung über kommunale und Landkreisgrenzen hinweg verstärkt die Akteursvielfalt die Herausforderung, trotz unterschiedlicher Datengrundlagen, Verkehrsmodelle und organisatorischer Strukturen zusammenzuarbeiten.

Hinzu kommt, dass die strategische und die operative Ebene der ÖPNV-Planung oft nicht ausreichend verknüpft sind. So werden die Verkehrsnachfrage bzw. Nachfragepotenziale und die Wirtschaftlichkeit des Angebots bei der übergeordneten strategischen Planung oft nur unzureichend berücksichtigt. Gleichzeitig liegt der Fokus bei der operativen Fahr- und Dienstplanung in erster Linie auf der Optimierung des vertraglich festgelegten Angebots des

Verkehrsunternehmens. Zwischen den beiden Ebenen fehlt eine systematische Rückkopplung, die strategische Zielsetzungen stärker mit operativen Umsetzungserfahrungen verbindet.

Neben der Verteilung der Zuständigkeiten über mehrere Organisationen und Verwaltungsebenen hinweg und der Notwendigkeit der Kopplung von Planung und Betrieb kommen die Heterogenität der Planungsfälle und die nur teilweise mögliche Formalisierbarkeit planerischer Abwägungen als Herausforderungen hinzu, die Komplexität und Unsicherheit in der Planung erhöhen können.

Diese Herausforderungen erfordern ein geteiltes Verständnis darüber, wie Rollen und Zuständigkeiten verteilt sind und welche Analysen zur Beantwortung einer konkreten Fragestellung sinnvoll sind. Um diesen Bedarf an Koordination und Transparenz zu decken, braucht es integrierte Planungsumgebungen, die kollaborative Prozesse und eine flexible, organisationsübergreifende Planung ermöglichen.

3 Die PIMO-Plattform: Konzept und Mehrwerte

Die Planungsplattform Intermodale Mobilität (PIMO) wird mit dem Ziel entwickelt, Akteure und Anwendungen aus der strategischen und operativen ÖPNV-Planung über eine flexible Plattformarchitektur mit offenen Schnittstellen zu verknüpfen, sodass eine integrierte Planung ermöglicht wird. Die Zielgruppe umfasst ÖPNV-Aufgabenträger, Verkehrsunternehmen und private Mobilitätsdienstleister. Auf der Ebene von Regionen bzw. funktionalen Stadtgebieten ist die Bereitstellung des ÖPNV-Angebots oft heterogen organisiert und es wird mit unterschiedlichen Systemen, Modellen und Datengrundlagen gearbeitet. Mit PIMO wird eine zentrale Softwareumgebung bereitgestellt, an die bestehende Fahrplan-, Dienstplan- und Leitsysteme sowie Datenbestände über offene Schnittstellen angeschlossen werden können. Zudem stellt die Plattform Verkehrsmodelle auf verschiedenen Planungsebenen bereit. Dies ermöglicht eine ganzheitliche, integrierte Planung bei der, je nach konkretem Planungsfall, Akteure und technische Infrastrukturen flexibel kombiniert werden können. Die gemeinsame Entwicklung und Bewertung von Szenarien, Ableitung von Maßnahmen und Visualisierung von Vorhaben stärkt die Transparenz, reduziert den Abstimmungsaufwand und beschleunigt Entscheidungen, beispielsweise in der Abstimmung von Stadt und Landkreis. Außerdem wird die Überführung strategischer Maßnahmen (beispielsweise Taktverdichtungen oder neue Bedarfsverkehre im Umland) in die Praxis erleichtert. Durch diese stärkere Verknüpfung der strategischen und operativen Ebene werden mögliche Konflikte zwischen Strategie- und Umsetzungsplanung frühzeitig erkannt und neue Angebote können schneller umgesetzt werden.

Die Vorteile einer datengestützten, kollaborativen Planung liegen vor allem darin, dass Entscheidungen auf einer breiten Grundlage an Daten, Modellen und interdisziplinären Abwägungen basieren und dadurch evidenzbasiert und nachvollziehbar sind. Zudem werden während des Planungsprozesses Organisationsaufwand und Ressourcen gespart, was insbesondere vor dem Hintergrund begrenzter finanzieller Mittel, gerade in ländlichen Regionen relevant ist. Durch die Simulation von Szenarien kann außerdem die Wirkung von Maßnahmen und Maßnahmenbündeln (z.B. Wirtschaftlichkeit, Nachfrageentwicklung oder Verlagerungseffekte) vorab ermittelt werden, sodass kostenintensive Fehlplanungen vermieden werden. Auch das indikatorbasierte Monitoring und damit die Rückführung operativer Daten in die strategische Planung kann über die digitale Planungsplattform stattfinden.

Neben der Optimierung des Regelbetriebs ist die PIMO-Plattform auch für zukunftsrelevante Planungsfälle geeignet, die mit einer höheren Unsicherheit verbunden sind. Dies betrifft beispielsweise die Einführung neuer Angebotsformen, wie bedarfsgesteuerter Verkehrsangebote (Demand Responsive Transport, DRT), deren Wirtschaftlichkeit, potenzielle Nachfrage und Netzintegration sich mit herkömmlichen Methoden deutlich schwerer abbilden lassen. Außerdem ermöglicht die Systemarchitektur die Kopplung und integrierte Planung mit weiteren Sektoren wie Wirtschaftsverkehr und Energie. Letzteres ist besonders für die Operation elektrifizierter Flotten relevant, deren Umfang, Routen und Ladezeiten eine direkte Abstimmung mit Netzkapazitäten erfordern. Über die in die Plattform integrierten Simulationsmodelle können zum Beispiel Vehicle2Grid-Konzepte geplant werden, bei denen Fahrzeugbatterien als Speicher zur Netzstabilisierung dienen. Besonders relevant ist die flexible Architektur der Plattform auch in Krisensituationen, in

denen eine schnelle, auf aktuellen Daten basierende Anpassung des ÖPNV-Systems notwendig wird, um eine grundlegende Mobilität aufrechtzuerhalten. Im Fall von Naturkatastrophen, Störungen der Infrastruktur oder militärischer Bedrohung muss innerhalb kürzester Zeit unter Einbezug verschiedener Akteure ein alternativer Betriebsplan erstellt werden, der ausgefallene Netzabschnitte umgeht und Ressourcen und Aufgaben verteilt. PIMO unterstützt dies durch die Möglichkeit, ad-hoc Szenarien zu simulieren, die Auswirkungen der eingeschränkten Infrastruktur unmittelbar zu bewerten und unter Einbezug aller relevanten Entscheidungsträger Handlungsmöglichkeiten abzuwägen.

4 Anforderungen an die Plattformarchitektur

Um die beschriebenen Mehrwerte in vollem Umfang realisieren zu können, muss eine Planungsplattform unterschiedliche Anforderungen erfüllen. Die Plattformarchitektur von PIMO kombiniert die folgenden Funktionen und Komponenten um einen kollaborativen und effizienten ÖPNV-Planungsprozess zu unterstützen.

Integration bestehender Anwendungen

Aufgabenträger und Verkehrsunternehmen arbeiten bereits mit etablierten Softwaresystemen, beispielsweise zur Fahr- und Dienstplanung und zur Verkehrsmodellierung. Eine zentrale Anforderung an die Plattformarchitektur besteht deshalb darin, bestehende Anwendungen zu integrieren statt sie zu ersetzen. PIMO ermöglicht dies durch eine modulare Architektur mit offenen Schnittstellen (Konnektoren), die die Interoperabilität und Anbieterunabhängigkeit sicherstellt. Über diese Konnektoren können Softwaresysteme, regionale Verkehrsmodelle und Datenbestände angebunden und Medienbrüche vermieden werden. Gleichzeitig bleibt jeder Akteur für seinen Zuständigkeitsbereich verantwortlich und behält die Hoheit über die eigenen Daten. Auch operative Leitsysteme (z.B. ITCS, RTL) können über die Plattform integriert werden, sodass Betriebsdaten kontinuierlich in die strategische Planung zurückfließen. Insgesamt ermöglichen der modulare Aufbau und die offenen Schnittstellen eine dynamische Anpassung und Erweiterung der Plattform in Abstimmung mit aktuellen und zukünftigen Planungsfragen.

Verkehrsmodellierung, Simulation und Analyse

Die Verkehrsmodellierung über mehrere Planungsebenen bildet den methodischen Kern der Plattform. Durch die Kopplung makroskopischer, mesoskopischer und mikroskopischer Modelle (z.B. TAPAS, MATSim, SUMO) können Angebote systematisch geplant und gesteuert und strategische und operative Planungsschritte verknüpft werden. Die Modelle ermöglichen es, Maßnahmen wie Taktverdichtungen oder Bedarfsverkehre vorab auf ihre betriebliche Machbarkeit und Wirkung im Gesamtnetz zu prüfen. Der direkte Vergleich und die Visualisierung von Planungsszenarien durch Simulationen ermöglichen eine schnelle, kollaborative und evidenzbasierte Entscheidungsfindung. Zudem wird die Beteiligung externer Akteure wie Entscheidungsträger und Bürger*innen erleichtert und die Akzeptanz von Maßnahmen gesteigert. Basierend auf der breiten Datenbasis und den Ergebnissen von Modellierungen und Simulationen können detaillierte Analysen durchgeführt werden, die Indikatoren zu Kosten, Qualität und Nachfrage von Angeboten liefern. Neben Verkehrsmodellen ermöglicht die Architektur der Plattform außerdem die Integration von Modellen aus weiteren Sektoren wie Wirtschaft und Energie, beispielsweise für die Planung elektrisierter Flotten und Vehicle2Grid-Konzepten.

Spezialisierte Nutzeroberflächen

Da Planungsprozesse heterogene Rollen umfassen, braucht es Nutzeroberflächen (User Interfaces, UI), die auf die individuellen Anforderungen und Aufgaben zugeschnitten und niedrigschwellig nutzbar sind. Die Plattform stellt nutzerfreundliche Dashboards bereit, über die Parameter für Szenarien eingestellt und Ergebnisse visualisiert werden können, ohne dass ein tiefgreifendes Wissen über die zugrundeliegenden Modelle notwendig ist. Kooperatives Arbeiten wird durch eine Client-Server-Architektur unterstützt, die einen gemeinsamen Zugriff auf Planungsstände ermöglicht. Durch die übersichtliche Darstellung von KPIs (Key Performance Indicators) und Lagebildern wird die Interaktion zwischen Fachplaner*innen, Entscheidungsträger*innen und Bürger*innen erleichtert. Das System kann außerdem zwischen Regelbetrieb und Notstandsfall unterscheiden und in Krisensituationen entsprechende Ansichten bereitstellen und andere Prioritäten, Akteure und Zuständigkeiten adressieren.

Datenmanagement

Sämtliche Modelle und Analysen sind nur so belastbar wie die Daten, auf denen sie beruhen. Das Datenmanagement stellt daher das Fundament der Plattform dar und umfasst die Integration, Qualitätssicherung und Annotation heterogener Datenbestände. Relevant sind dabei unterschiedliche Datentypen: Betriebsdaten (z.B. Fahrplanleistung, Ist-Abfahrten), Auslastungsdaten (z.B. Fahrgastzahlen, Kapazitätsnutzung), Netzdaten (z.B. Haltestellen), Fahrplandaten, Verkehrsverhaltensdaten sowie soziodemografische Daten (z.B. Siedlungsstruktur, Einwohnerverteilung). Ein Datenmanagement mit standardisierten Schnittstellen und einem einheitlichen Datenmodell sichert einen durchgängigen Datenfluss ohne Medienbrüche. Da Zuständigkeiten und Datenbestände auf mehrere Organisationen verteilt sind, setzt PIMO nicht auf eine zentrale Datenhaltung. Stattdessen erfolgt der Datenaustausch über Datenräume, die es ermöglichen, dass jeder Akteur die Datenhoheit über seine Bestände behält und Daten unter definierten Bedingungen bereitstellt, statt sie abzugeben. So wird die Datensouveränität der Akteure gewahrt, was eine Voraussetzung für die Akzeptanz der Plattform bei Aufgabenträgern und Verkehrsunternehmen ist.

Semantische Schicht (Ontologie)

Die Ontologie bildet den semantischen Kern der Plattform. Sie verknüpft Planungsprozesswissen (z.B. Akteurskonstellationen, rechtliche Rahmenbedingungen, Planungsabläufe und Zuständigkeiten) mit Methoden- und Technologiewissen (z.B. über Verkehrsmodelle, Simulationsdienste und erforderliche Datengrundlagen). So werden fachliche Fragestellungen in ausführbare Prozessketten übersetzt und es wird sichergestellt, dass technische Komponenten kontextgerecht und nachvollziehbar eingesetzt werden. Dazu werden die Beziehungen zwischen organisatorischen Rollen und technischen Diensten maschinenlesbar abgebildet, sodass das System automatisch zuordnen kann, welche Werkzeuge für eine gegebene Planungsaufgabe geeignet und welche Akteure einzubinden sind. Ohne diese vermittelnde Schicht müsste jede Kombination aus Akteur, Planungsfall und Fachmodell einzeln und statisch verdrahtet werden – ein in einem verteilten, heterogenen und potenziell mit Krisenfällen konfrontierten System nicht tragfähiger Ansatz.

5 Funktionsweise der PIMO-Plattform

Die beschriebenen Komponenten und Funktionen der Plattform werden durch die technische Architektur umgesetzt und verbunden (Abb. 1). Die technische Funktionsweise von PIMO wird im Folgenden anhand eines konkreten Anwendungsfalls veranschaulicht: der Einführung eines bedarfsgesteuerten Verkehrs (DRT) in einer ländlichen Region. Diese Fragestellung ist repräsentativ, da sie sowohl strategische Überlegungen (Nachfragepotenziale, Wirtschaftlichkeit) als auch operative Anforderungen (Planbarkeit, Fahrzeugeinsatz) umfasst und die Zusammenarbeit mehrerer Akteure erfordert.

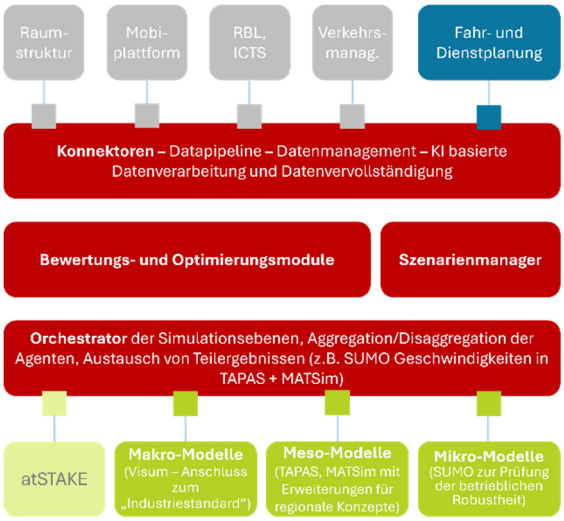

Abbildung 1: technische Architektur der PIMO-Plattform (Eigene Darstellung, 2026)

Ausgangspunkt ist die **Definition des Planungsszenarios**. Über ein nutzerfreundliches Dashboard geben die Planenden die Parameter für die Angebotsszenarien ein. Ausgewählt werden können beispielsweise das Untersuchungsgebiet (z.B. Landkreis oder Teilregion), ein Nachfrage-Schwellenwert (unter dem der Linienverkehr unwirtschaftlich wird), die DRT-Flottengröße, das gewünschte Service-Level (z.B. maximale Wartezeit, Fußweg zur Haltestelle) sowie die Strategie (Ersatz einzelner oder aller Linienabfahrten durch DRT). Diese Eingaben werden im **Szenarienmanager** verwaltet, der verschiedene Szenarien konsolidiert und für die weitere Verarbeitung vorbereitet. Die Ontologie interpretiert die Eingaben semantisch: sie erkennt, dass es sich um eine Angebotsplanung im ländlichen Raum handelt, identifiziert die relevanten Akteure und bestimmt, welche technischen Komponenten für die Fragestellung relevant sind.

Auf Basis dieser semantischen Zuordnung greift die Plattform über **Konnektoren** auf verteilte Datenbestände zu. So können über den Datenraum zum Beispiel Betriebsdaten, Auslastungsdaten, Netzdaten oder soziodemografische Daten abgerufen werden. Die Datenverarbeitung wird durch KI-basierte Verfahren unterstützt, die beispielsweise fehlende Werte ergänzen oder Muster in Nachfragedaten erkennen.

Anschließend erfolgt die Modellierung, die von einem **Orchestrator** koordiniert wird. Dieser koppelt die verschiedenen **Verkehrssimulationsmodelle** und stellt sicher, dass sie konsistent zusammenarbeiten. Im Beispiel der DRT-Planung können makroskopische Modelle wie Visum (PTV GmbH, 2026) beispielsweise die Nachfragepotenziale im Gesamtnetz berechnen und Abschnitte mit geringer Auslastung identifizieren. Mesoskopische Modelle wie MATSim (Horni et al., 2016) oder TAPAS (Heinrichs et al., 2016) können das DRT-Angebot auf individueller Ebene simulieren und abbilden, wie sich Fahrgastströme verlagern, wenn Linienverkehre durch DRT ersetzt werden. Mikroskopische Modelle wie SUMO (Alvarez Lopez et al., 2018) können die betriebliche Robustheit prüfen, etwa über den Verkehrsfluss an Haltestellen oder Fahrgastwechselzeiten.

Die Simulationsergebnisse werden von **Bewertungs- und Optimierungsmodulen** in entscheidungsrelevante Indikatoren überführt. Diese Module berechnen KPIs zu Nachfrageffekten (z.B. Verlagerung, Neugewinnung), Kosten (z.B. Fahrzeugeinsatz, Personalaufwand), Qualität (z.B. Reisezeit, Umsteigebeziehungen) und Angebotseffekten (z.B. Abdeckung, Taktung). Spezifische Optimierungsmodule für DRT können die Ausgestaltung des bedarfsgesteuerten Angebots unterstützen. Die Ergebnisse werden visuell vergleichbar aufbereitet, sodass Planende die Wirkungen von Szenarien im gesamten Netz betrachten und vergleichen können.

Im nächsten Schritt werden die Ergebnisse über **Konnektoren zur Fahr- und Dienstplanung** an angeschlossene Systeme übergeben. So kann geprüft werden, welche Abfahrten sich betrieblich sinnvoll ergänzen lassen. Berücksichtigt werden können je nach System zum Beispiel Fahrzeug- und Personaleinsatz, Umlaufplanung, Dienstplanoptimierung und mögliche Leerlaufzeiten. Diese Rückmeldung fließt über die Konnektoren in die Plattform zurück, sodass der Szenarienmanager die Szenarien um ihre operative Machbarkeit erweitern kann.

Schließlich werden die Szenarien in der Nutzeroberfläche zur **kollaborativen Entscheidungsfindung** bereitgestellt (Abb. 2). Die Ontologie stellt sicher, dass alle Akteure Zugang zu den für sie relevanten Informationen haben. Durch die übersichtliche Visualisierung von Indikatoren und Szenarien können auch externe Akteure wie Entscheidungsträger*innen oder Bürger*innen ohne Fachkenntnisse in die Entscheidungsfindung einbezogen werden. Durch dieses Zusammenspiel technischer Komponenten ermöglicht PIMO eine evidenzbasierte Entscheidung darüber, ob im Untersuchungsgebiet DRT-Angebote umgesetzt werden sollen, welche klassischen Linien diese ersetzen sollen und wie die konkrete Ausgestaltung des Angebots aussehen kann.

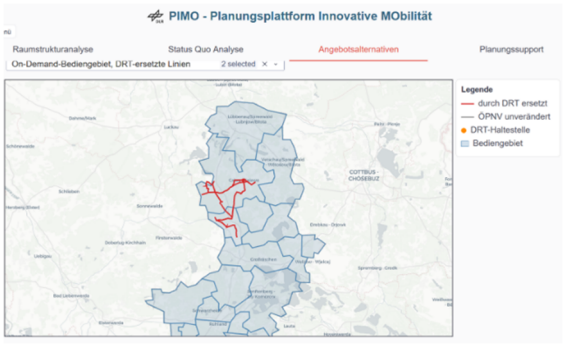

Abbildung 2: Visualisierung eines beispielhaften DRT-Angebotsszenarios in der PIMO-Nutzeroberfläche (Eigene Darstellung, 2026)

6 Anwendung und Pilotimplementierung

Die Architektur und Funktionsweise der PIMO-Plattform ermöglichen eine Vielzahl an Anwendungsfällen in unterschiedlichsten Bereichen der ÖPNV-Planung. Die durch die TEN-V-Verordnung vorgeschriebene Erstellung von SUMPs fordert explizit eine kooperative, zuständigkeitsübergreifende und ganzheitliche Planung für das „funktionale Stadtgebiet“ sowie die indikatorenbasierte Maßnahmenbewertung und datengestützte Entscheidungsfindung (Rupprecht Consult, 2026). PIMO unterstützt durch den modularen Aufbau, die offenen Schnittstellen und die nutzerspezifische Oberfläche die notwendige Integration von Planungsebenen, Planungssektoren, Zuständigkeitsbereichen, technischen Systemen und Datenbeständen. Die Plattform kann verwendet werden, um das vorhandene ÖPNV-

Angebot zu bewerten und zu optimieren und neue Angebote zu konzipieren, indem die Wirkungen von Maßnahmen (z.B. auf Nachfrage, Auslastung, Kosten) auf verschiedenen Ebenen simuliert werden. Dabei umfasst der Anwendungsbereich sowohl die strategische Planung, beispielsweise die Definition von Angebotsformen und -zielen oder die Erschließung neuer Bediengebiete, als auch die operative Planung, beispielsweise die Anpassung von Fahrplänen oder den Einsatz von Fahrzeugen und Personal.

Der Einsatzbereich von PIMO geht zudem über den Regelbetrieb hinaus und umfasst auch Krisen- und Notfallszenarien. Im Fall von Krisensituationen wie Naturkatastrophen, Infrastrukturausfällen oder Sicherheitslagen verändert sich das Verkehrsverhalten sowie die Verfügbarkeit von Infrastrukturen grundlegend und Datengrundlagen aus dem Regelbetrieb verlieren schnell ihre Aussagekraft. Die Plattform ermöglicht hier das Hinterlegen von Notfallplänen, die schnelle Integration neuer Datenquellen und die Generierung von Szenarien, um auch bei unvorhersehbaren Entwicklungen handlungsfähig zu bleiben. Zudem können weitere relevante Akteure und Systeme (beispielsweise aus dem Bereich der Bundeswehr) leicht eingebunden werden.

Die Pilotimplementierung in der Region Oberspreewald-Lausitz konkretisiert die beschriebenen Anwendungsbereiche unter realen Bedingungen. Der Landkreis Oberspreewald-Lausitz in Brandenburg ist ländlich geprägt, weist eine hohe Auspendlerstruktur auf und steht, wie viele weitere Landkreise, vor der Herausforderung, mit begrenzten finanziellen Mitteln ein attraktives ÖPNV-Angebot zu gewährleisten. Die PIMO-Plattform liegt aktuell als webbasierter Demonstrator vor und wird in Zusammenarbeit mit dem Transferpartner highQ Computerlösungen GmbH und dem Praxispartner Verkehrsgesellschaft Oberspreewald-Lausitz (VGOSL) umgesetzt und weiterentwickelt. Ziel der Zusammenarbeit mit der VGOSL ist es, die Praxistauglichkeit der Plattform unter realen Bedingungen zu testen und die Anwendung genau auf die Anforderungen der Planenden abzustimmen, damit ein möglichst hoher Mehrwert erzielt werden kann. Als kommunaler Verkehrsbetrieb stellt die VGOSL Regional- und Stadtbuslinien auch über die Landkreisgrenzen hinaus bereit und ist repräsentativ für zahlreiche Aufgabenträger in ländlich geprägten Regionen. Als Praxispartner gibt die VGOSL Anforderungen und Planungsfälle vor, die bei der Entwicklung berücksichtigt werden, stellt ausgewählte Planungsdaten zur Verfügung und wird den Prototyp im Realbetrieb testen sowie Feedback zu Bedienbarkeit, Praxistauglichkeit und Ergebnisqualität geben.

Im Folgenden werden exemplarische Anwendungsfälle beschrieben, die in der Region Oberspreewald-Lausitz mithilfe von PIMO umgesetzt werden sollen.

Ein zentraler Anwendungsfall ist die Einführung von DRT-Verkehren. Im Gebiet Calau/Großräschen soll ab 2027 testweise On-Demand-Verkehr statt liniengebundenen Angeboten erprobt werden. PIMO soll die VGOSL dabei unterstützen, weitere geeignete Gebiete für die Umstellung auf bedarfsgesteuerte Angebote zu identifizieren, die Auswirkungen auf den verbleibenden Linienverkehr, Umläufe, Dienste und Fahrzeugbedarf zu analysieren und nach der Implementierung zu evaluieren, ob das Angebot die gewünschte Wirkung erzielt hat. Dabei fließen unter anderem Betriebsdaten (Fahrplanleistung, Ist-Abfahrten), Auslastungsdaten (Fahrgastzahlen) und soziodemografische Daten (Einwohnerverteilung, Siedlungsstruktur) in die Modellierung ein.

Ein weiterer Use Case ist die Unterstützung des jährlichen Fahrplanwechsels. PIMO ermöglicht hier den Variantenvergleich verschiedener Fahrplanszenarien, die Versionierung von Arbeitsständen und die kartenbasierte Darstellung von Netzeffekten. Dies ist insbesondere für die Kommunikation mit dem Landkreis als Aufgabenträger und dem Kreistag als politisches Gremium relevant, da Karten und die Visualisierung von Linienstärken greifbarer sind als reine Zahlen.

Im Bereich der Notfallplanung besteht insofern Handlungsbedarf, als zwar Notfallpläne existieren, die jedoch manuell per Anruf oder E-Mail aktiviert werden müssen. PIMO soll vorbereitete Notfallszenarien bereitstellen, die bei Unwettern, Nebel oder anderen Störungen schnell aktiviert werden können. Priorisierte Points of Interest (POI) wie Schulen, Ärztehäuser oder Krankenhäuser sowie Mindestbedienanforderungen werden datenbasiert hinterlegt, sodass Ausdünnungsentscheidungen objektiv getroffen werden können.

Die Datenintegration im Landkreis Oberspreewald-Lausitz umfasst unter anderem Fahrplandaten,

Linien- und Haltestelleninformationen, Abo- und Ticketzahlen sowie perspektivisch auch Auslastungsdaten aus Zählsystemen, die ab 2027 in Bussen der VGOSL verfügbar sein werden. Insbesondere diese Auslastungsdaten sind von hoher Relevanz, da sie eine belastbare Grundlage für die Bewertung von Angeboten liefern. Auch Bahnfahrpläne und Fremdlinien sollen für die Anschlussplanung eingebunden werden. Über ein differenziertes Rollen- und Rechtekonzept wird sichergestellt, dass Akteure nur Zugriff auf für sie relevante Daten haben. So benötigt die VGOSL Bearbeitungsrechte, der Landkreis Einsichts- und Entscheidungsrechte, während externe Akteure wie privatwirtschaftliche oder wissenschaftliche Einrichtungen als Inputgeber mit begrenzter Sicht auftreten.

Technisch ist in der Pilotregion zudem die Verzahnung von PIMO mit der Software PlanB relevant, die zur Fahr- und Dienstplanung verwendet wird. Ein bidirektionaler Datenfluss ermöglicht es, dass PIMO Analyse- und Szenarienergebnisse an PlanB übergibt, und PlanB wiederum betriebliche Rückmeldungen (z.B. Fahrzeug- und Personaleinsatz, Leerlaufzeiten) an PIMO zurückspielt. So wird die Lücke zwischen strategischer Planung und operativer Machbarkeit geschlossen.

Die Pilotimplementierung im Landkreis Oberspreewald-Lausitz zeigt beispielhafte Anwendungsfälle der PIMO-Plattform und dient ihrer Validierung und Überführung in ein marktfähiges Produkt. Die Erkenntnisse aus der Praxis fließen kontinuierlich in den Entwicklungsprozess zurück, wodurch die Software genau mit den Anforderungen der Planungspraxis abgestimmt wird.

7 Fazit und Perspektive

Die PIMO-Plattform demonstriert, wie eine datengetriebene, kollaborative Planungsinfrastruktur dabei unterstützen kann, die ÖPNV-Planung an aktuelle und zukünftige Anforderungen anzupassen – von der Erfüllung gesetzlicher Vorgaben wie der TEN-V-Verordnung über die effiziente Nutzung finanzieller Mittel bis hin zur regionen- und zuständigkeitsübergreifenden Zusammenarbeit. Die Architektur der Plattform ermöglicht es, heterogene Datenbestände zu integrieren, Wirkungen auf verschiedenen Ebenen zu modellieren und Szenarien zu simulieren und vergleichend darzustellen, sodass evidenzbasierte Entscheidungen unter Einbezug aller relevanten Akteure getroffen werden können. Ein Kernmerkmal von PIMO ist dabei die niedrigschwellige Nutzbarkeit und die flexible Anpassbarkeit der Oberfläche und Funktionen, je nach Nutzergruppe, Planungsfall oder lokalen Gegebenheiten.

Die modulare Architektur ermöglicht es, PIMO über die klassische ÖPNV-Planung hinaus zu erweitern. Anwendungsbereiche wie Logistik und Güterverkehr oder Zivilschutz und Verteidigung können zukünftig an die Plattform angeschlossen werden, um eine ganzheitliche und resiliente Planung zu ermöglichen. Zudem ermöglicht der offene Aufbau der Plattform die schrittweise Integration sich weiterentwickelnder Technologien wie Künstlicher Intelligenz, die an verschiedenen Stellen der Plattform eingesetzt werden kann.

Die nächsten Schritte bestehen in der fortlaufenden Pilotimplementierung der Plattform und darauf basierenden Weiterentwicklung. Im Verlauf des Forschungsprojekts wird die Plattform von TRL 4 (Versuchsaufbau im Labormaßstab) auf TRL 7 (System im operationellen Umfeld) weiterentwickelt. Am Ende steht die Überführung in ein marktfähiges Produkt, das in der breiten Praxis angewendet werden soll.

Literaturverzeichnis

Alvarez Lopez, Pablo; Behrisch, Michael; Bieker-Walz, Laura; Erdmann, Jakob; Flötteröd, Yun-Pang; Hilbrich, Robert; Lücken, Leonhard; Rummel, Johannes; Wagner, Peter; Wießner, Evamarie (2018): Microscopic Traffic Simulation using SUMO, in: 2019 IEEE Intelligent Transportation Systems Conference (ITSC), Seiten 2575-2582. IEEE. The 21st IEEE International Conference on Intelligent Transportation Systems, 2018-11-04 - 2018-11-07, Maui, USA. doi: 10.1109/ITSC.2018.8569938

Bundesministerium für Verkehr (BMV) (2025): ÖPNV - Öffentlicher Personennahverkehr. https://www.bmv.de/SharedDocs/DE/Artikel/G/oeffentlicher-personennahverkehr.html

Gesetz über intelligente Verkehrssysteme im Straßenverkehr und deren Schnittstellen zu anderen Verkehrsträgern und über die Datenbereitstellung über den Nationalen Zugangspunkt (Intelligente-Verkehrssysteme-Gesetz - IVSG) vom 15. Mai 2026. (2026). Bundesgesetzblatt Teil I, Nr. 148. https://www.gesetze-im-internet.de/ivsg_2026/

Heinrichs, Matthias; Krajzewicz, Daniel; Cyganski, Rita; von Schmidt, Antje (2018): Disaggregated car fleets in microscopic travel demand modelling, in: The 7th International Conference on Ambient Systems, Networks and Technologies (ANT 2016). The 7th International Conference on Ambient Systems, Networks and Technologies, 2016-05-23 - 2016-05-26, Madrid, Spanien. doi: 10.1016/j.procs.2016.04.111.

Horni, Andreas; Nagel, Kai.; Axhausen, Kay. W. (Hrsg., 2016): The Multi-Agent Transport Simulation MATSim. London: Ubiquity Press. doi: http://dx.doi.org/10.5334/baw. License: CC-BY 4.0

PTV Planung Transport Verkehr GmbH (2026): PTV Visum. https://www.ptvgroup.com/de/produkte/ptv-visum

Richtlinie (EU) 2023/2661 des Europäischen Parlaments und des Rates vom 22. November 2023 zur Änderung der Richtlinie 2010/40/EU zum Rahmen für die Einführung intelligenter Verkehrssysteme im Straßenverkehr und für deren Schnittstellen zu anderen Verkehrsträgern. (2023). Amtsblatt der Europäischen Union, L 2023/2661. https://eur-lex.europa.eu/eli/dir/2023/2661/oj

Rupprecht Consult - Forschung & Beratung GmbH (2026): Guidelines for Developing and Implementing a Sustainable Urban Mobility Plan, 3. Auflage. https://urban-mobility-observatory.transport.ec.

europa.eu/document/download/3fb33992-f63e-4c9e-b5cf-d18a576489f9_en?filename=SUMP%20

Guidelines_third%20edition_2026_final%20cover.pdf

Verordnung zur Bestimmung kritischer Anlagen nach dem BSI-Gesetz (BSI-Kritisverordnung - BSI-KritisV) vom 22. April 2016. (2016). Bundesgesetzblatt Teil I, 958. https://www.gesetze-im-internet.de/bsi-kritisv/

Verordnung (EU) 2022/868 des Europäischen Parlaments und des Rates vom 30. Mai 2022 über europäische Daten-Governance und zur Änderung der Verordnung (EU) 2018/1724 (Daten-Governance-Rechtsakt). (2022). Amtsblatt der Europäischen Union, L 152, 1-44. https://eur-lex.europa.eu/eli/reg/2022/868/oj

Verordnung (EU) 2023/2854 des Europäischen Parlaments und des Rates vom 13. Dezember 2023 über harmonisierte Vorschriften für einen fairen Datenzugang und eine faire Datennutzung und zur Änderung der Verordnung (EU) 2017/2394 und der Richtlinie (EU) 2020/1828 (Datenverordnung). (2023). Amtsblatt der Europäischen Union, L 2023/2854. https://eur-lex.europa.eu/eli/reg/2023/2854/oj

Verordnung (EU) 2024/1679 des Europäischen Parlaments und des Rates vom 13. Juni 2024 über Leitlinien der Union für den Aufbau eines transeuropäischen Verkehrsnetzes, zur Änderung der Verordnungen (EU) 2021/1153 und (EU) Nr. 913/2010 und zur Aufhebung der Verordnung (EU) Nr. 1315/2013. (2024). Amtsblatt der Europäischen Union, L 2024/1679. https://eur-lex.europa.eu/eli/reg/2024/1679/oj

Autor*innenangaben

Marlene Grebe
Wissenschaftliche Mitarbeiterin
DLR Institut für Verkehrsforschung
Rudower Chaussee 7, 12489 Berlin

E-Mail: marlene.grebe@dlr.de

Prof. Dr. Michael Ortgiese
Professor für Verkehrs- und Mobilitätsmanagement
DLR Institut für Verkehrsforschung
Rudower Chaussee 7, 12489 Berlin

E-Mail: michael.ortgiese.@dlr.de

Pradeep Burla
Wissenschaftlicher Mitarbeiter
DLR Institut für Verkehrsforschung
Rudower Chaussee 7, 12489 Berlin

E-Mail: pradeep.burla@dlr.de

Dr. Kerstin Stark
Wissenschaftliche Mitarbeiterin
DLR Institut für Verkehrsforschung
Rudower Chaussee 7, 12489 Berlin

E-Mail: kerstin.stark@dlr.de

Journal für Mobilität und Verkehr

ISSN 2628-4154

www.dvwg.de

Entwicklung eines konzeptionellen Frameworks zur Erfassung, Qualitätssicherung und Bereitstellung urbaner Umweltdaten als Grundlage für ein umweltsensitives Mobilitätsmanagement im Projekt AIAMO

Greta Schmalhorst, Timo Houben, Ema Vosgerau, Florian Berger, Michael Weger, Joost Hemmen, Alex Romanjuta, Paul Remmler, David Schäfer, Jan Bumberger, Thomas Trabert*

Siehe Autor*innenangaben

Abstract

Die ab 2030 geltenden, verschärften Grenzwerte der EU-Luftqualitätsrichtlinie stellen Kommunen vor die Herausforderung, verkehrsbedingte Immissionsspitzen gezielt und vorausschauend zu steuern. Herkömmliche Modellierungen und vereinzelte amtliche Messstationen bieten jedoch weder die räumlich-zeitliche Auflösung noch die operationelle Datenanbindung, die für ein dynamisches, umweltsensitives Verkehrsmanagement erforderlich sind. Dieser Beitrag stellt ein ganzheitliches methodisches Framework vor, das ein Netzwerk aus ca. 30 Sensoren im Stadtgebiet Leipzig mit Verkehrsfluss- und Wetterdaten verbindet. Die Systemarchitektur basiert auf dem Digital Ecosystem for FAIR Time Series Data Management und integriert ein Sensor Management System (SMS), die Plattform time.io sowie automatisierte Qualitätskontrollen mittels SaQC. Fehlende Daten in den Zeitreihen werden über ein vortrainiertes Foundation-Modell imputiert, wodurch ein konsistenter, FAIR-konformer und AI-ready Datenbestand entsteht. Auf dieser Grundlage ermöglicht ein KI-Vorhersagemodell eine 48-Stunden-Prognose der lokalen Stickstoffdioxid-Belastung (NO2).

Die verkehrswissenschaftliche Auswertung im Stadtgebiet Leipzig belegt eine ausgeprägte Koppelung zwischen Verkehrsfluss und Immissionen mit deutlichen Peaks zu Rush-Hour-Zeiten. Die Analyse differenziert zwischen strukturellen Dauerbelastungen im Leipziger Norden und temporären Spitzen durch Veranstaltungsverkehre. Im Rahmen des Projekts AIAMO (Artificial Intelligence and Mobility) fließen die Daten über eine interoperable OGC SensorThings API in automatisierte Wochenberichte für die Kommune ein und bilden das Fundament für umweltsensitive Routingkonzepte. Der Beitrag zeigt damit den konzeptionellen Weg von der reinen Datenerhebung bis zur Bereitstellung verlässlicher Entscheidungsgrundlagen für die Verkehrsplanung und umweltsensitive Verkehrssteuerung auf.

Schlagwörter:

Umweltsensitives Verkehrsmanagement, Sensorik, FAIR-Datenökosystem, Urbaner Raum, Monitoring

1 Einleitung / Problemstellung

Das anhaltend hohe Verkehrsaufkommen in verdichteten urbanen Räumen stellt Kommunen vor die komplexe Herausforderung, Mobilitätsbedürfnisse und Umwelt- sowie Gesundheitsschutz miteinander in Einklang zu bringen (Deng et al., 2021). Insbesondere im innerstädtischen Hauptverkehrsstraßennetz führt Verkehr zu Rush-Hour-Zeiten zu ausgeprägten Immissionsspitzen. Mit den ab 2030 geltenden Grenzwerten der neuen EU-Luftqualitätsrichtlinie werden die zulässigen Werte für verkehrsrelevante Schadstoffe wie Stickstoffdioxid (NO2) und Feinstaub weiter abgesenkt, sodass in zahlreichen

* korrespondierender Autor

europäischen Ballungsräumen ohne gezielte Interventionen mit signifikanten Grenzwertüberschreitungen zu rechnen ist (EU-Kommission, 2024). Um Kommunen zukünftig in die Lage zu versetzen, wirksame und dynamische Maßnahmen zu ergreifen, bedarf es einer hochaufgelösten und belastbaren Datenbasis sowie standardisierter Strukturen für deren Bereitstellung (Bibri et al., 2024).

Im Zuge der fortschreitenden Digitalisierung des Verkehrssektors eröffnen sich durch automatisierte Sensornetzwerke, vernetzte Fahrzeuge und Methoden der Künstlichen Intelligenz neue Möglichkeiten zur Erfassung und Analyse urbaner Mobilitätsdaten (Ismagilova et al., 2019; Minerva et al., 2020). Während datengetriebene Ansätze und digitale Zwillinge zunehmend als Schlüsseltechnologien für eine nachhaltige Mobilitätswende diskutiert werden (Weil et al., 2023; Supianto et al., 2024), stützt sich die Praxis häufig auf stark vereinfachte Modelle, ohne verkehrliche Simulations- und Messdaten mit realistischen meteorologischen Bedingungen zusammenzuführen (Trabert et al., 2026 und Weger et al., 2026). Zudem finden durchgängig validierte Systemarchitekturen, die Umweltsensornetze unter realen Betriebsbedingungen automatisiert an operative Verkehrssteuerungen koppeln, bisher kaum Anwendung (Weil et al., 2023). Ungeprüfte Rohdaten oder messnetzbedingte Datenlücken sind mit dem Risiko schwerwiegender Fehlsteuerungen im nachgelagerten Verkehrssystem verbunden.

Der vorliegende Beitrag begegnet diesen Herausforderungen, indem er den Aufbau und die praktische Implementierung eines durchgängigen methodischen Frameworks zur Erfassung, Qualitätssicherung, Imputation und verkehrlichen Integration von Umwelt- und Mobilitätsdaten vorstellt. Das Framework basiert auf den praktischen Erfahrungen des zweijährigen Aufbaus und Betriebs eines urbanen Sensornetzwerks im Rahmen des Forschungsprojekts AIAMO – Artificial Intelligence and MObility, in dem das Helmholtz-Zentrum für Umweltforschung (UFZ) das urbane Sensornetzwerk in der Pilotregion Leipzig aufbaut und betreibt. Die Sicherung einer belastbaren Datenqualität hinsichtlich Vollständigkeit, Konsistenz und Vergleichbarkeit sowie die Kalibrierung der Hardware bilden dabei die zentrale Voraussetzung, um Umweltdaten verlässlich in operative Verkehrssysteme einzuspeisen.

Am konkreten Anwendungsfall der AIAMO Pilotregion Leipzig wird demonstriert, wie die heterogenen Zeitreihen (Luftqualitätsdaten) von knapp 30 stationären Sensoren über offene Schnittstellen kontinuierlich aggregiert, qualitätsgesichert und für zielgerichtete Maßnahmen des umweltsensitiven Verkehrsmanagements operationalisiert werden können.

Durch die Bereitstellung über standardisierte Schnittstellen (OGC, STA-SensorThings API) und die Anbindung an städtische Verkehrsdaten zeigt der Beitrag exemplarisch auf, wie datenbasierte Ansätze von der reinen Messwerterhebung bis hin zur praktischen Umsetzung umweltsensitiver Steuerungskonzepte im urbanen Raum beitragen können. Abbildung 1 demonstriert an dieser Stelle das aufgebaute Umweltmessnetz.

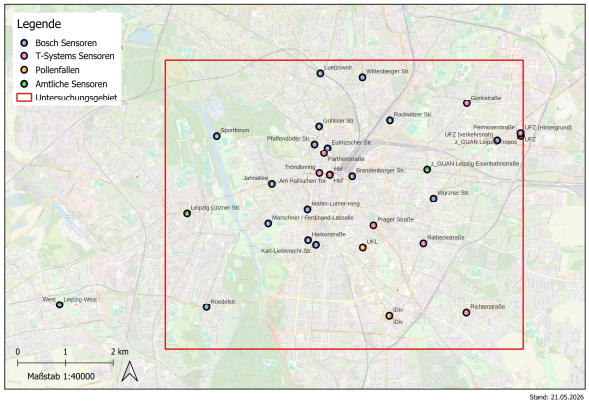

Abbildung 1: Sensornetzwerk in Leipzig (Eigene Darstellung UFZ, 2026)

2 Methodisches Framework

Zur Beschreibung des entwickelten Frameworks (Abbildung 2) gliedert sich der methodische Aufbau in drei Schritte: (1) Identifikation und Schaffung der Datenbasis, (2) Datenmanagement, (3) Datenbereitstellung und -integration in KI-Lösungen.

2.1 Identifikation und Schaffung der Datenbasis

Um eine valide Grundlage für das umweltsensitive Verkehrsmanagement zu schaffen, erfordert das Framework eine synchrone Erfassung von Luftqualitäts- und Verkehrsdaten sowie meteorologischen Parametern. Die Datenbasis setzt sich aus drei komplementären Teilbereichen zusammen.

Den Kern bildet ein installiertes Messnetzwerk aus ca. 30 stationären Luftqualitätssensoren, welche an straßenseitiger Infrastruktur wie Laternenmasten und Trägern des Parkleitsystems montiert wurden.

Die Festlegung der Sensorstandorte erfolgte in enger Abstimmung mit der Kommune anhand verkehrsplanerischer Kriterien. Dabei wurden gezielt Hauptverkehrsstraßen im Stadtzentrum und Verkehrsknotenpunkte, an denen häufige Halte- und Anfahrvorgänge zu ausgeprägten Emissionsspitzen führen, gewählt. Ergänzend wurden Hintergrundmessungen in verkehrsberuhigten Bereichen eingerichtet. Das Sensornetz erfasst primär die Schadstoffe Stickstoffdioxid (NO2) und Feinstaub (PM10 und PM2.5) sowie Ozon, begleitet von Temperatur und relativer Luftfeuchte in minütlicher zeitlicher Auflösung.

Um den bekannten Messunsicherheiten der Sensoren, insbesondere Sensordrifts sowie Querempfindlichkeiten gegenüber Temperatur- und Feuchtigkeitsschwankungen, wirksam zu begegnen, durchlaufen alle Geräte ein Kalibrierungsverfahren. In einem ersten Schritt werden die Sensoren unter kontrollierten Bedingungen an der Referenzmessstation des Helmholtz-Zentrums für Umweltforschung vorab kalibriert, um sensorindividuelle Parameter für den Nullpunkt-Offset und die Sensitivität zu bestimmen. Daran schließt sich eine feldbezogene Co-Location-Kalibrierung im realen Straßenraum an. Feste Re-Kalibrierungszyklen im März 2025 sowie im Sommer 2026 stellen zusammen mit einer bedarfsgesteuerten Wartung bei auffälligen Messabweichungen die langfristige Datenqualität sicher. Dass eine fundierte Kalibrierung für Low-Cost-Sensorik zwingend erforderlich ist, belegen Zhivkov et al. (2025), die zeigen, dass sich das Bestimmtheitsmaß R2 durch Kalibrierungsverfahren von 0,29 auf bis zu 0,95 steigern lässt (Zhivkov et al., 2025).

Ihre verkehrswissenschaftliche Relevanz erhalten diese Luftqualitätsdaten durch die synchrone Verknüpfung mit detaillierten Verkehrsflussinformationen. Hierfür werden Detektordaten aus dem städtischen Straßennetz herangezogen, die über Induktionsschleifen oder Radardetektoren fahrstreifen- und richtungsgetrennte Zähldaten liefern.

Da die Schadstoffausbreitung im Straßenraum maßgeblich durch Witterungsbedingungen bestimmt wird, werden parallel amtliche Wetterdaten eingebunden. Parameter wie Windgeschwindigkeit, Windrichtung, Luftdruck und Niederschlag ermöglichen es, meteorologisch bedingte Konzentrationsschwankungen und Inversionslagen rechnerisch von den rein verkehrsbedingten Effekten zu isolieren.

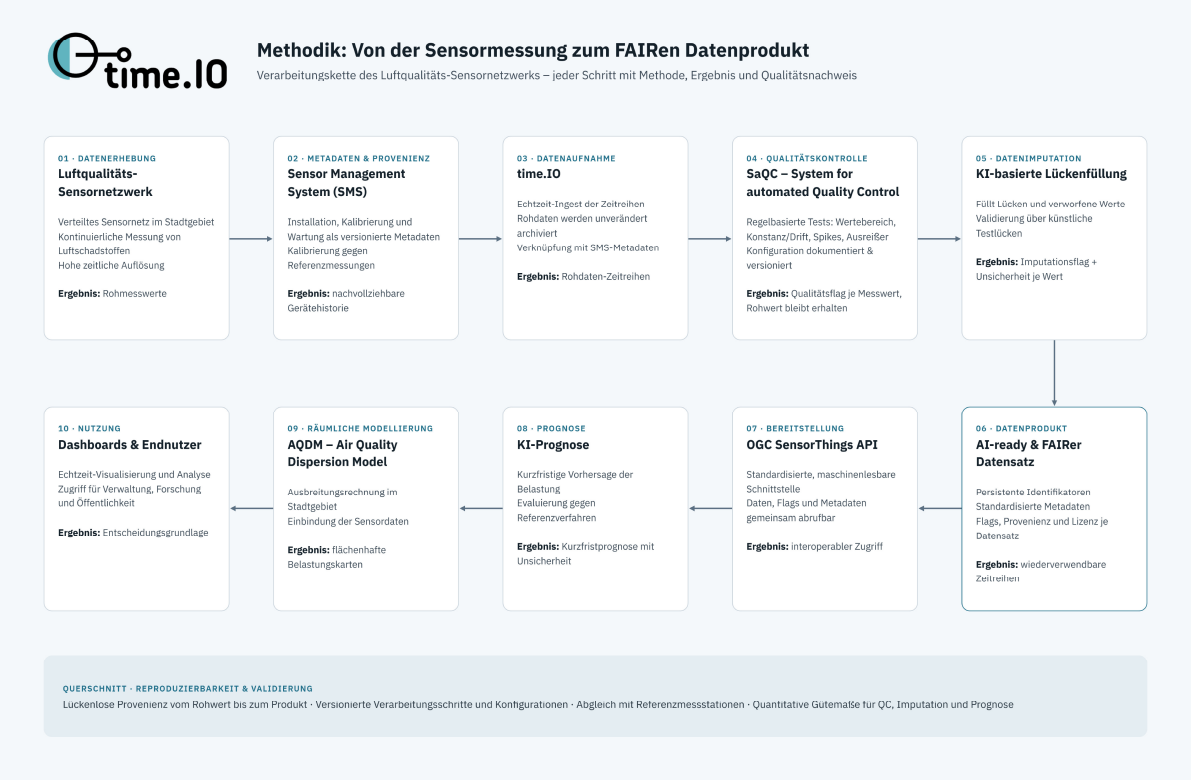

Abbildung 2: Konzeptionelles Framework (eigene Darstellung UFZ, 2025)

2.2 Datenmanagement

Die kontinuierliche Verarbeitung und Bereitstellung der erhobenen Messwerte stützt sich auf eine modular aufgebaute Datenpipeline, die eine automatisierte Übernahme, Qualitätsprüfung und Veredelung der heterogenen Datenströme nach den FAIR-Prinzipien (Findable, Accessible, Interoperable, Reusable) gewährleistet. Die grundlegende Dateninfrastruktur basiert dabei auf dem Framework eines „Digital Ecosystem for FAIR Time Series Data Management in Environmental System Science“ (Bumberger et al., 2024), welches aus drei eng miteinander gekoppelten Kernkomponenten besteht.

Die grundlegende Administration der Messinfrastruktur erfolgt über das Sensor Management System (SMS), welches die Metadaten der eingesetzten Hardware (Sensoren), deren physische Standorte im städtischen Straßennetz sowie historische Änderungen an Konfigurationen und Kalibrierungen im Zeitverlauf dokumentiert (Lorenz et al.2026). Außerdem wird hier die Zuordnung der individuellen Datenströme verwaltet. Durch die Umsetzung internationaler Standards wie OGC, SensorML und der JSON:API-Spezifikation sowie eindeutiger Persistent Identifiers (EUDAT B2INST) gewährleistet das SMS ein einheitliches Metadatenmanagement sowie die nahtlose, interoperable Auffindbarkeit und Integration aller Messstellen in übergeordnete Dateninfrastrukturen (Brinkmann et al. 2024, Bumberger et al. 2024).

Als zentrale Datenintegrationsplattform fungiert das vom Helmholtz-Zentrum für Umweltforschung entwickelte System time.io. Die Plattform übernimmt die Steuerung und Konsolidierung aller Daten, indem sie heterogene Datenströme, darunter die über MQTT und (S-)FTP eingehenden Rohmesswerte sowie vorverarbeitete Referenzdaten des Umweltbundesamtes und Wetterzeitreihen des Deutschen Wetterdienstes, automatisiert aggregiert und persistent speichert. Als technisches Bindeglied synchronisiert time.io die Metadaten aus dem SMS und leitet die Zeitreihen direkt an die nachgelagerte Qualitätskontrolle (Bumberger et al. 2024, Schäfer et al. 2025).

Auf den minütlich eingehenden Zeitreihendaten setzt unmittelbar eine automatisierte Qualitätskontrolle auf, die auf dem Framework SaQC (System for automated Quality Control, Schmidt et al., 2023) basiert. Das System greift auf statistische Verfahren zurück, um Ausreißer und unplausible Signalsprünge in Echtzeit zu detektieren. Anstatt die Rohdatenbestände irreversibel zu überschreiben, markiert das Framework auffällige Messwerte über ein standardisiertes Flagging-Schema (Data Annotation). Auf diese Weise wird eine lückenlose Nachvollziehbarkeit und Reproduzierbarkeit gemäß den FAIR-Prinzipien für alle nachgelagerten Analyse- und Steuerungsprozesse garantiert (Schäfer et al. 2025, Schmidt et al. 2023).

Auf dieser Datenbasis setzt die weiterführende Veredelung auf. Um sensorspezifische Datenlücken zu schließen, durchlaufen die bereinigten Zeitreihen eine nachgelagerte KI-basierte Imputation auf Basis des Foundation Modells TSPulse (Ekambaram et al., 2025). Foundation Modelle für Zeitreihen werden auf großen Datenmengen vortrainiert und lassen sich effizient auf Downstream-Aufgaben adaptieren (Wang et al., 2025). Sie erlernen allgemeine temporale Strukturen wie Saisonalität, Periodizität und Trends (Kottapalli et al., 2025). TSPulse ist auf Aufgaben wie Imputation, Anomalieerkennung und Klassifikation spezialisiert und kombiniert zeitliche mit spektraler Rekonstruktion (Ekambaram et al., 2025). Dies ist beispielsweise für Luftqualitätsdaten relevant, die tägliche, wöchentliche und saisonale Muster aufweisen. Das feinabgestimmte Modell rekonstruiert fehlende Messwerte unter Einbezug zyklischer Zeitvariablen, räumlicher Abhängigkeiten und meteorologischer Parameter. Auf diese Weise entsteht ein lückenloser, verlässlicher und konsistenter AI-ready Datensatz, der als fundierte Datenbasis für weiterführende Verkehrsmodelle fungiert.

Die operative Überwachung des Messnetzes und die Visualisierung der aufbereiteten (qualitätsgesicherten) Daten erfolgen in Echtzeit über webbasierte Dashboards. Diese ermöglichen die Darstellung der Immissionslagen im Straßennetz mittels Karten, die vergleichende Analyse von Zeitreihen einzelner Sensorstandorte sowie die direkte Sichtbarkeit der vergebenen Qualitätsflags. Dass die zeitreihenbasierte Datenhaltung in spezialisierten Datenbanken in Kombination mit Dashboards und KI-Modellen eine hochpräzise Echtzeit-Überwachung sowie die Vorhersage von Grenzwertüberschreitungen im urbanen Raum ermöglicht, demonstrieren Naveed et al. (2025).

Die konsolidierte Datenbasis bildet zudem das Fundament für ein nachgelagertes KI-Modell zur Vorhersage stationsspezifischer Schadstoffimmissionen mit einem Prognosehorizont von 48 Stunden. In der gegenwärtigen Ausbaustufe konzentriert sich das Modell auf den verkehrsrelevanten Leitschadstoff Stickstoffdioxid NO2. Als Trainingsgrundlage dient ein statischer Datensatz (ein Jahr) aus historischen Vor-Ort-Messreihen und meteorologischen Parametern.

2.3 Datenbereitstellung und Anbindung von Umweltdaten

Den abschließenden Schritt des Frameworks bildet die Etablierung eines Datenmanagementsystems, welches die aufbereiteten Umwelt- und Verkehrsdaten bündelt und über die zentrale AIAMO Integrationszone bereitstellt. Über die Integrationszone kann eine Anbindung an das Ausbreitungsmodell Air Quality Dispersion Model (AQDM) erfolgen. Durch die synchrone Bereitstellung von Echtzeit-Messwerten, Wetter- und Verkehrsdaten erlaubt das Modell eine räumlich hochaufgelöste Simulation der Schadstoffverteilung im gesamten Stadtgebiet und schließt damit die Lücke zwischen den verteilten Sensorstandorten und einer flächendeckenden Luftqualitätsbewertung.

3 Das Framework im verkehrlichen Kontext

Die Anwendbarkeit und der verkehrssystemische Mehrwert des vorgestellten konzeptionellen Frameworks werden nachfolgend am Beispiel des Stadtgebiets Leipzig demonstriert. Ziel dieses Kapitels ist es zu zeigen, wie aus der Verknüpfung von Umweltsensorik, Verkehrsdaten und Wetter fundierte Erkenntnisse über verkehrsbedingte Belastungsdynamiken gewonnen und in operative sowie planerische Anwendungsfälle des umweltsensitiven Verkehrsmanagements überführt werden können.

Die Bereitstellung über die AIAMO Integrationszone ermöglicht die Einbindung der Umweltmonitoring-Pipeline in die umweltsensitive Verkehrssteuerung. Während das vorgestellte methodische Framework die messtechnischen und datenseitigen Voraussetzungen schafft, entfaltet das Gesamtsystem seinen praktischen Nutzen mit der konkreten Einbindung in den Stadtverkehr.

3.1 Korrelation von Verkehrsfluss und Schadstoffdynamik

Die kontinuierliche Verknüpfung von Verkehrs- und Luftqualitätsdaten ermöglicht eine differenzierte räumliche Bewertung des Straßennetzes. Insbesondere an stark belasteten Verkehrspunkten treten während der morgendlichen und nachmittäglichen Verkehrsspitzen charakteristische Schadstoffkonzentrationen auf. Die Auswertung der hochaufgelösten Messdaten zeigt deutlich, wie eng Schadstoffbelastung und Verkehrsaufkommen im Tagesverlauf zusammenhängen.

Durch den Abgleich der Immissionsverläufe mit den Geschwindigkeitsdaten und Zählschleifenwerten lässt sich der direkte Einfluss von Stauerscheinungen und Stop-and-Go-Mustern auf die lokale Stickstoffdioxid- und Feinstaubexposition empirisch nachweisen und von Hintergrundeffekten isolieren.

Im Vergleich zu den städtischen Hintergrundmesspunkten, die ein weitgehend homogenes Tagesprofil ohne verkehrstypische Spitzen aufweisen, belegt dieser Kontrast die direkte Kausalität zwischen den gemessenen Immissionsspitzen und dem Verkehrsgeschehen im Straßenkorridor.

3.2 Identifikation von Hotspots

Die Identifikation räumlicher Belastungsschwerpunkte erfolgte in enger Abstimmung mit den Fachämtern der Kommune. Die Auswertung der kombinierten Messreihen offenbart dabei grundlegend verschiedene Typen verkehrsbedingter Hotspots im Stadtgebiet.

Einerseits lassen sich strukturell bedingte Dauerbelastungszonen lokalisieren, wie sie insbesondere im Leipziger Norden auftreten. Dort führen ein kontinuierlich hohes Verkehrsaufkommen und zentrale Straßenknotenpunkte sowie die Autobahnanbindungen zu einer dauerhaft erhöhten Schadstoffgrundlast.

Andererseits können temporäre, ereignisbezogene Belastungsspitzen identifiziert werden (z. B. Großveranstaltungen). An Veranstaltungstagen überlagert der An- und Abreiseverkehr des motorisierten Individualverkehrs das reguläre Verkehrsgeschehen, was kurzzeitig zu lokalen Überlastungen des Straßennetzes führen kann.

Durch den synchronen Abgleich von Geschwindigkeitsabfällen mit den gemessenen Immissionsanstiegen lassen sich gezielt Knotenpunkte im Stadtgebiet lokalisieren, an denen Stauerscheinungen und Stop-and-Go-Zustände die Grenzwerte der EU-Luftqualitätsrichtlinie akut gefährden. Diese differenzierte und evidenzbasierte Erfassung bildet die Voraussetzung, um Steuerungsmaßnahmen räumlich zu adressieren.

3.3 AI-ready Daten für das umweltsensitive Verkehrsmanagement

Der direkte Mehrwert für die Verkehrsplanung und die städtischen Akteure zeigt sich in zielgruppenspezifischen Anwendungen für die kommunale Praxis. Im Rahmen von AIAMO werden die Umweltdaten der Stadt Leipzig für das strategische und operative Mobilitätsmanagement kontinuierlich mit den Verkehrsfluss- und Detektordaten des Deutschen Zentrums für Luft- und Raumfahrt fusioniert und in Form automatisierter Wochenberichte zur Verfügung gestellt. Diese Berichte dienen als evidenzbasierte Entscheidungsgrundlage, um kommunale, verkehrssteuernde Maßnahmen einzuleiten oder verkehrsplanerische Eingriffe an hochbelasteten Achsen zu evaluieren.

Auf operativer Ebene ermöglichen die aktuellen Messwerte und 48-Stunden-NO2-Prognosen die Konzeption umweltsensitiver Navigationsstrategien. Das Ziel besteht darin, den motorisierten Individualverkehr frühzeitig um prognostizierte Hotspots herumzuleiten und den Fuß- sowie Radverkehr über emissionsarme bzw. gemeinwohlkonforme Routen zu führen, bei denen anstelle einer reinen individuellen Fahrzeitminimierung primär die Reduktion von Schadstoffexposition und lokaler Umweltbelastung im Vordergrund steht.

4 Diskussion

Die Ergebnisse aus dem zweijährigen Betrieb des Messnetzes in Leipzig belegen, dass die automatisierte Zusammenführung von Sensorik, amtlichen Wetterdaten und Verkehrsflussparametern eine praxistaugliche Grundlage für ein umweltsensitives Verkehrsmanagement schafft. Insbesondere vor dem Hintergrund der ab 2030 verschärften Grenzwerte aus der EU-Luftqualitätsrichtlinie bietet das konzeptionelle Framework Kommunen die notwendige Datendichte, um drohende Grenzwertüberschreitungen nicht mehr nur reaktiv festzustellen, sondern ihnen durch proaktive Netzbeeinflussung und umweltsensitive Verkehrssteuerung gezielt entgegenzuwirken. Da sich dieses Framework derzeit noch in der Erprobung befindet, werden weitere notwendige Schritte getestet. Das konzeptionelle Framework kann hierfür AI-ready Daten bereitstellen, um ein umweltsensitives Mobilitätsmanagement zu ermöglichen. Dahingehend werden mögliche Empfehlungen für Verkehrsmanagementmaßnahmen bereitgestellt, die sowohl situativ als auch perspektivisch Einfluss auf das Verkehrsgeschehen nehmen können.

5 Fazit und Ausblick

Zusammenfassend zeigt das hier vorgestellte Framework eine mögliche Vorgehensweise zur Bereitstellung qualitätsgeprüfter Daten auf – von der Erfassung und Verknüpfung von Umwelt-, Verkehrs- und Wetterdaten über Qualitätssicherung und KI-basierte Imputation bis zur standardisierten Bereitstellung und Integration in KI-Lösungen für das umweltsensitive Mobilitätsmanagement.

Im Gesamtkontext von AIAMO schafft diese Datenbasis die Voraussetzung, hochaufgelöste Umweltdaten gemeinsam mit Verkehrs- und Wetterdaten in ein umfassendes KI-gestütztes Mobilitätsmanagement zu integrieren und für die kommunale Praxis nutzbar zu machen. Damit erhalten Kommunen eine gemeinsame Datengrundlage für fachbereichsübergreifende Entscheidungen: Verkehrliche Maßnahmen können bereits vor ihrer Umsetzung hinsichtlich ihrer zu erwartenden verkehrlichen und ökologischen Wirkungen bewertet werden. Nach der Umsetzung lassen sich die tatsächlichen Wirkungen auf Basis der integrierten Daten umfassend analysieren und die Maßnahmen darauf aufbauend weiter optimieren.

Zugleich ist zu berücksichtigen, dass Messunsicherheiten und Sensordrifts auch das Training der KI-Anwendungen beeinflussen können. Eine vollumfängliche Validierung dieses Ansatzes ist dahingehend noch ausstehend.

Der strukturierte Aufbau dieses Frameworks zeigt einen Leitfaden zur Bereitstellung und Nutzung dieser Datensätze in verschiedenen Anwendungen. Besonderer Dank gilt der Pilotregion Leipzig, in der das Framework im Rahmen des vom Bundesministerium für Digitales und Staatsmodernisierung (BMDS)

geförderten Projekts AIAMO (Artificial Intelligence and Mobility) erprobt wird.

Literaturverzeichnis

Bumberger, J., Abbrent, M., Brinckmann, N., Hemmen, J., Kunkel, R., Lorenz, C., Lünenschloss, P., Palm, B., Schnicke, T., Schulz, C., van der Schaaf, H., Schäfer, D. (2025). Digital ecosystem for FAIR time series data management in environmental system science. SoftwareX, 29, 102038. https://doi.org/10.1016/j.softx.2025.102038.

Deng, Tianhu, Keren Zhang und Zuo-Jun (Max) Shen (2021). A Systematic Review of a Digital Twin City: A New Pattern of Urban Governance toward Smart Cities. Journal of Management Science and Engineering, 6(2), 125–134. https://doi.org/10.1016/j.jmse.2021.03.003.

Ekambaram, Vijay, Subodh Kumar, Arindam Jati, u. a. (2026). TSPulse: Tiny Pre-Trained Models with Disentangled Representations for Rapid Time-Series Analysis. arXiv:2505.13033. Preprint, arXiv, 4. März 2026. https://doi.org/10.48550/arXiv.2505.13033.

EU-Kommission (2024). EU air quality standards. https://environment.ec.europa.eu/topics/air/air-quality_en?prefLang=de. Abgerufen im Dezember 2024.

Ismagilova, Elvira, Laurie Hughes, Yogesh K. Dwivedi und K. Ravi Raman (2019). Smart Cities: Advances in Research—An Information Systems Perspective. International Journal of Information Management, 47, 88–100. https://doi.org/10.1016/j.ijinfomgt.2019.01.004.

Kaffashzadeh, Najmeh (2023). A Data-Driven Persistence Test for Robust (Probabilistic) Quality Control of Measured Environmental Time Series: Constant Value Episodes. Atmospheric Measurement Techniques, 16(12), 3085–3100. https://doi.org/10.5194/amt-16-3085-2023.

Kottapalli, Siva Rama Krishna, Karthik Hubli, Sandeep Chandrashekhara, u. a. (2025). Foundation Models for Time Series: A Survey. arXiv:2504.04011. Preprint, arXiv, 5. April 2025. https://doi.org/10.48550/arXiv.2504.04011.

Lorenz, C., Brinckmann, N., Bumberger, J., Hanisch, M., Kuhnert, T., Loup, U., Moorthy, R., Obersteiner, F., Schäfer, D. und Schnicke, T. (2026). Sensor management system (SMS): Open-source software for FAIR sensor metadata management in earth system

sciences. SoftwareX, 35, 102858. https://doi.org/10.1016/j.softx.2026.102858.

Minerva, Roberto, Gyu Myoung Lee und Noel Crespi (2020). Digital Twin in the IoT Context: A Survey on Technical Features, Scenarios, and Architectural Models. Proceedings of the IEEE, 108(10), 1785–1824. https://doi.org/10.1109/JPROC.2020.2998530.

Naveed, Khazina, Tariq Umer, Aamer Bilal Asghar, u. a. (2025). Machine Learning Assisted Predictive Urban Digital Twin for Intelligent Monitoring of Air Quality Index for Smart City Environment. Environmental Modelling & Software, 192, 106559. https://doi.org/10.1016/j.envsoft.2025.106559.

Schäfer, David, Martin Abbrent, Florian Gransee, u. a. (2025). time.IO – A Digital ecosystem for FAIR time series data management. Version 2025-09-12. Zenodo, veröffentlicht am 12. September 2025. https://doi.org/10.5281/ZENODO.8354839.

Schmidt, L., Schäfer, D., Geller, J., Lünenschloss, P., Palm, B., Rinke, K., Rebmann, C., Rode, M. und Bumberger, J. (2023). System for automated Quality Control (SaQC) to enable traceable and reproducible data streams in environmental science. Environmental Modelling & Software, 169, 105809. https://doi.org/10.1016/j.envsoft.2023.105809.

Trabert, Thomas, Elmar Brockfeld, Alexander Sohr, u. a. (2026). Urban Digital Twin for Environmentally Sensitive Mobility Planning: Conceptual Framework and Application in Pilot Region Leipzig. International Journal of Applied Earth Observation and Geoinformation, 152, 105477. https://doi.org/10.1016/j.jag.2026.105477.

Wang, Jun, Wenjie Du, Yiyuan Yang, u. a. (2025). Deep Learning for Multivariate Time Series Imputation: A Survey. arXiv:2402.04059. Preprint, arXiv, 20. Mai 2025. https://doi.org/10.48550/arXiv.2402.04059.

Weger, Michael, Thomas Trabert, Timo Houben, u. a. (2026). Hyperlocal Urban NO_2 Hotspot Modeling Driven by Microscopic Traffic Data. npj Climate and Atmospheric Science, 9(1), 150. https://doi.org/10.1038/s41612-026-01471-z.

Weil, Charlotte, Simon Elias Bibri, Régis Longchamp, François Golay und Alexandre Alahi (2023). Urban Digital Twin Challenges: A Systematic Review and Perspectives for Sustainable Smart Cities. Sustainable Cities and Society, 99, 104862. https://doi.org/10.1016/j.scs.2023.104862.

Zhivkov, Peter, Stefka Fidanova und Ivan Dimov (2025). Digital Twin-Based Framework for Real-Time Monitoring and Analysis of Urban Mobile-Source Emissions. Atmosphere, 16(6), 731. https://doi.org/10.3390/atmos16060731.

Autor*innenangaben

Greta Schmalhorst
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: greta.schmalhorst@ufz.de

Timo Houben
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: timo.houben@ufz.de

Ema Vosgerau
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: ema.vosgerau@ufz.de

Florian Berger
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: florian.berger@ufz.de

Dr. Michael Weger
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: michael.weger@ufz.de

Joost Hemmen
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: joost.hemmen@ufz.de

Alex Romanjuta
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: alex.romanjuta@ufz.de

Paul Remmler
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: paul.remmler@ufz.de

David Schäfer
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: david.schaefer@ufz.de

Dr. Jan Bumberger
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig
German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig
Puschstraße 4
04103 Leipzig

E-Mail: jan.bumberger@ufz.de

Dr. Thomas Trabert
Helmholtz Zentrum für Umweltforschung GmbH
Research Data Management – RDM
Department Monitoring and Exploration Technologies
Permoser Straße 15
04318 Leipzig

E-Mail: thomas.trabert@ufz.de

Journal für Mobilität und Verkehr

ISSN 2628-4154

www.dvwg.de

Mobilitätsdaten zwischen Effizienz und Gerechtigkeit: Ein Governance-Framework für eine nachhaltige und inklusive Mobilitätsplanung

Ann-Katrin Voit*

Siehe Autor*innenangaben

Abstract

Die zunehmende Nutzung von Mobilitätsapps, Plattformdaten, GPS-Tracking, Smartcards, Sensoren und automatisierten Auswertungsverfahren erweitert die empirische Grundlage der Verkehrsplanung erheblich. Digitale Mobilitätsdaten ermöglichen hochaufgelöste Analysen von Bewegungsmustern und können dazu beitragen, Verkehrsangebote effizienter, bedarfsgerechter und nachhaltiger zu gestalten. Große Datenmengen gewährleisten jedoch weder Repräsentativität noch gesellschaftliche Inklusivität. Insbesondere ältere Personen, einkommensschwache Haushalte, Menschen mit geringer digitaler Teilhabe oder mit Behinderungen sowie bestimmte nicht motorisierte Mobilitätsformen können in digital erzeugten Datensätzen systematisch unterrepräsentiert sein.

Der Beitrag untersucht daher, unter welchen methodischen und institutionellen Bedingungen Mobilitätsdaten für eine effiziente und zugleich sozial inklusive Verkehrsplanung eingesetzt werden können. Methodisch folgt der Beitrag einer integrativen interdisziplinären Literatursynthese zu Mobilitätsdatenerhebung, Datenqualität und Repräsentativität, digitaler Exklusion, Mobility Justice und Data Justice. Die identifizierten Problemfelder werden analytisch verdichtet und in ein FAIR-Mobility-Data-Framework mit den Dimensionen Fairness, Accuracy, Inclusion und Responsibility überführt. Das Framework dient als Prüfraster für Datenerhebung, Analyse und datenbasierte Entscheidungsprozesse.

Als zentrale Governance-Instrumente werden die Triangulation digitaler Daten mit Befragungs- und Verwaltungsdaten, systematische Repräsentativitätsprüfungen, die Dokumentation von Datenlücken und Unsicherheiten, partizipative Ergänzungsverfahren sowie klare Verantwortungs- und Transparenzstrukturen vorgeschlagen. Der Beitrag zeigt damit, wie Effizienzpotenziale datenbasierter Mobilitätsplanung genutzt werden können, ohne digitale Sichtbarkeit mit gesellschaftlicher Repräsentativität gleichzusetzen.

Schlagwörter:

Mobilitätsdaten, Data Governance, Repräsentativität, Mobilitätsgerechtigkeit, digitale Exklusion

1 Einleitung: Mobilitätsdaten zwischen Effizienz und Gerechtigkeit

Die Digitalisierung verändert nicht nur Mobilitätsangebote, sondern auch die Wissensgrundlage, auf der Mobilitätspolitik und Verkehrsplanung beruhen. Während klassische Verkehrserhebungen überwiegend auf Haushaltsbefragungen, Wegetagebüchern, manuellen Zählungen oder punktuellen Beobachtungen basieren, entstehen Mobilitätsdaten heute kontinuierlich aus einer Vielzahl digitaler Quellen wie Smartphones, GPS-Anwendungen, Navigationsdienste, Mobilitätsplattformen, Smartcards, Sharing-Systeme, Mobilfunknetze und stationäre Sensoren. Diese erzeugen hochaufgelöste Informationen über räumliche und zeitliche Bewegungsmuster. Damit können Mobilitätsprozesse in einer Detailtiefe analysiert werden, die traditionelle Erhebungsverfahren allein kaum ermöglichen

*korrespondierende Autorin

(González et al., 2008; Nitsche et al., 2014; Shen & Stopher, 2014; Zannat & Choudhury, 2019).

Die Attraktivität solcher Daten liegt insbesondere in ihrer räumlichen und zeitlichen Granularität sowie in der Möglichkeit kontinuierlicher oder zumindest wiederholter Beobachtungen. Smartphone- und GPS-basierte Verfahren können beispielsweise Wege erfassen, die in klassischen Wegetagebüchern vergessen oder ungenau angegeben werden, und ermöglichen automatisierte Rückschlüsse auf Verkehrsmittelwahl, Routen und teilweise auch Wegezwecke (Nitsche et al., 2014; Shen & Stopher, 2014; Xiao et al., 2016). In der Verkehrsplanung eröffnen diese Daten neue Möglichkeiten zur Nachfrageanalyse, Netzplanung, Angebotssteuerung und Evaluation. Hagen und Reinfeld (2021) weisen zugleich darauf hin, dass gerade die Kombination von Befragungs- und Trackingdaten für die Analyse neuer Mobilitätsangebote Vorteile gegenüber einer ausschließlichen Nutzung einzelner Datentypen bietet.

Die zunehmende Datenverfügbarkeit führt jedoch zu einer grundlegenden methodischen Herausforderung. Datenmenge ist nicht mit Repräsentativität gleichzusetzen. Digitale Mobilitätsdaten entstehen typischerweise nicht im Rahmen probabilistischer Stichproben, sondern als Nebenprodukt der Nutzung bestimmter Geräte, Dienste oder Plattformen. Ob eine Person in einem Datensatz sichtbar wird, hängt damit unter anderem von Gerätebesitz, digitaler Kompetenz, App-Nutzung, Einwilligungsentscheidungen, Mobilitätsverhalten und den Geschäftsmodellen privater Datenanbieter ab. Bereits bei smartphonebasierten Mobilitätserhebungen treten Unterschiede hinsichtlich technischer Ausstattung, Teilnahmebereitschaft und Datenqualität auf (Nitsche et al., 2014; Shen & Stopher, 2014). Untersuchungen kommerzieller Standortdaten zeigen darüber hinaus räumliche und soziodemografische Verzerrungen, die bei einer unmittelbaren Übertragung der Daten auf Gesamtpopulationen problematisch werden können (Li et al., 2024; Jardel & Delamater, 2024).

Die Problematik reicht über klassische Stichprobenfehler hinaus. Die Digitalisierung von Mobilität kann bestehende Formen von Transportbenachteiligung mit Formen digitaler Ungleichheit verbinden. Groth (2019) beschreibt dies als „multimodal divide“: Personen, die bereits über geringere Mobilitätsoptionen verfügen, können gleichzeitig schlechteren Zugang zu jenen digitalen Technologien haben, die zunehmend Voraussetzung für die Nutzung neuer Mobilitätsangebote sind. Neuere Forschung bestätigt, dass digitale Technologien zwar Mobilitätsmöglichkeiten erweitern können, zugleich aber insbesondere bei älteren Menschen, einkommensschwächeren Bevölkerungsgruppen, Menschen mit geringeren digitalen Kompetenzen oder Beeinträchtigungen zusätzliche Zugangshürden erzeugen können (Durand et al., 2022; Durand et al., 2023). Die technische Digitalisierung von Mobilität ist damit nicht automatisch sozial inklusiv.

Diese Problematik wird besonders relevant, wenn digital erzeugte Mobilitätsdaten nicht lediglich der Beschreibung von Mobilität dienen, sondern zur Grundlage von Planungs-, Investitions- und Angebotsentscheidungen werden. Was in den Daten häufig vorkommt, wird analytisch sichtbar; was nicht oder nur unzureichend erfasst wird, kann dagegen auch im Planungsprozess unsichtbar bleiben. Die Forschung zu Data Justice weist seit längerem darauf hin, dass Daten nicht nur gesellschaftliche Realität abbilden, sondern über Kategorisierungs-, Sichtbarkeits- und Entscheidungsprozesse selbst Verteilungswirkungen entfalten können (Taylor, 2017). Für den Mobilitätsbereich verbindet die jüngere Forschung deshalb Ansätze der Mobility Justice mit Data Justice und fragt danach, wessen Mobilität erfasst, interpretiert und in datenbasierten Entscheidungen berücksichtigt wird (Lucas, 2012; Martens, 2017; Sheller, 2018; Behrendt & Sheller, 2024). Der Begriff der Mobility Data Justice bringt diese beiden Perspektiven ausdrücklich zusammen und verweist darauf, dass die Datafizierung von Mobilität bestehende Ungleichheiten sowohl sichtbar machen als auch reproduzieren oder verstärken kann (Behrendt & Sheller, 2024).

Vor diesem Hintergrund untersucht der Beitrag zwei miteinander verbundene Forschungsfragen: Unter welchen methodischen und institutionellen Bedingungen können digital erzeugte Mobilitätsdaten so genutzt werden, dass ihre Effizienz- und Erkenntnispotenziale ausgeschöpft werden, ohne soziale und räumliche Ungleichheiten systematisch zu reproduzieren? Und wie kann ein Governance-Framework ausgestaltet werden, das Datenqualität, Repräsentativität, Inklusion und Verantwortlichkeit

systematisch in datenbasierte Mobilitätsentscheidungen integriert?

Zur Beantwortung dieser Fragen wird im Folgenden zunächst das methodische Vorgehen erläutert. Anschließend werden die Potenziale und Selektivitäten digitaler Mobilitätsdaten herausgearbeitet. Darauf aufbauend wird das FAIR-Mobility-Data-Framework als analytisches und anwendungsorientiertes Prüfraster entwickelt. Abschließend werden aus dem Framework konkrete Governance-Prinzipien für Kommunen, Verkehrsunternehmen und Mobilitätsplattformen abgeleitet.

2 Methodisches Vorgehen: integrative Literatursynthese und Framework-Entwicklung

Der Beitrag verfolgt ein konzeptionelles, literaturbasiertes Forschungsdesign. Ziel ist nicht die empirische Quantifizierung einzelner Verzerrungen, sondern die systematische Zusammenführung bislang teilweise getrennt geführter Forschungsstränge zu Mobilitätsdatenerhebung, Datenqualität, sozialer Inklusion und Data Governance. Anders als bei einer systematischen Literaturübersicht mit Anspruch auf vollständige Erfassung eines klar abgegrenzten Forschungsfeldes dient die Literaturauswertung damit der theoriegeleiteten Identifikation wiederkehrender Problem- und Governance-Dimensionen.

Methodisch lässt sich das Vorgehen als integrative Literatursynthese mit anschließender konzeptioneller Framework-Entwicklung charakterisieren. Integrative Reviews eignen sich insbesondere für Forschungsfragen, bei denen heterogene theoretische und empirische Literaturbestände zusammengeführt werden sollen, um neue konzeptionelle Perspektiven zu entwickeln (Snyder, 2019). Anders als bei einem systematischen Review steht damit nicht die erschöpfende Identifikation sämtlicher Publikationen zu einer eng abgegrenzten Forschungsfrage im Vordergrund, sondern die theoriegeleitete Synthese unterschiedlicher Forschungsstränge. Die Entwicklung des Frameworks orientiert sich ergänzend an der Conceptual Framework Analysis nach Jabareen (2009). Konzeptionelle Frameworks werden dabei durch die Identifikation, Kategorisierung und Integration wiederkehrender Konzepte aus multidisziplinären Wissensbeständen entwickelt.

Berücksichtigt werden vier Literaturkomplexe. Erstens wird Forschung zur digitalen Erhebung von Mobilität, insbesondere zu GPS-, Smartphone-, Mobilfunk-, Smartcard- und plattformbasierten Daten, herangezogen (González et al., 2008; Nitsche et al., 2014; Shen & Stopher, 2014; Zannat & Choudhury, 2019). Zweitens wird Literatur zu Datenqualität, Selektionsproblemen und Repräsentativität digitaler Standort- und Mobilitätsdaten ausgewertet (Li et al., 2024; Jardel & Delamater, 2024). Drittens werden Arbeiten zu Transportbenachteiligung, digitaler Exklusion, Transport Justice und Mobility Justice berücksichtigt (Lucas, 2012; Martens, 2012, 2017; Pereira et al., 2017; Sheller, 2018; Groth, 2019; Durand et al., 2022, 2023). Viertens wird die Literatur zu Data Justice und verantwortungsvoller Data Governance einbezogen (Floridi & Taddeo, 2016; Taylor, 2017; Heeks & Renken, 2018; Dencik et al., 2019; Behrendt & Sheller, 2024).

Die Synthese erfolgt entlang wiederkehrender Fragestellungen: Wer wird durch eine Datenquelle erfasst? Mit welcher Genauigkeit wird Mobilität abgebildet? Welche Gruppen oder Mobilitätspraktiken können fehlen? Und welche institutionellen Akteur*innen treffen auf Grundlage dieser Daten Entscheidungen und tragen dafür Verantwortung? Diese vier Problemkomplexe werden anschließend konzeptionell zu den Dimensionen Fairness, Accuracy, Inclusion und Responsibility verdichtet.

Die Verdichtung erfolgt in zwei Schritten. Zunächst werden aus den berücksichtigten Forschungsfeldern wiederkehrende Anforderungen an die Nutzung von Mobilitätsdaten identifiziert. Anschließend werden inhaltlich verwandte Anforderungen zu übergeordneten Kategorien gebündelt. Fragen distributiver Wirkungen werden der Dimension Fairness, Anforderungen an Validität und Datenqualität der Dimension Accuracy, Fragen gesellschaftlicher und räumlicher Abdeckung der Dimension Inclusion und Aspekte institutioneller Zuständigkeit, Transparenz und Rechenschaft der Dimension Responsibility zugeordnet. Das Framework ist damit nicht deduktiv vorgegeben, sondern stellt eine konzeptionelle Synthese der in den einbezogenen Forschungsfeldern wiederkehrenden Problemdimensionen dar. Abbildung 1 fasst den methodischen Synthese- und Verdichtungsprozess zusammen.



Abbildung 1: Methodisches Vorgehen der integrativen Literatursynthese und Framework-Entwicklung. (Eigene Darstellung)

Das daraus entwickelte FAIR-Mobility-Data-Framework ist als heuristisches Governance-Modell zu verstehen. Es soll keine statistische Qualitätsprüfung ersetzen, sondern unterschiedliche Qualitäts- und Gerechtigkeitsanforderungen in einem gemeinsamen Entscheidungsrahmen verbinden. Die Bezeichnung ist dabei von den etablierten FAIR Data Principles zu unterscheiden. Diese definieren Forschungsdaten anhand von Findability, Accessibility, Interoperability und Reusability und zielen primär auf gutes Datenmanagement und Wiederverwendbarkeit (Wilkinson et al., 2016). Das hier entwickelte FAIR-Mobility-Data-Framework ergänzt diese technisch-organisatorische Perspektive um die normative Frage, ob Mobilitätsdaten hinreichend fair, valide, inklusiv und verantwortungsvoll genutzt werden.

3 Mobilitätsdaten zwischen Effizienz und gesellschaftlicher Selektivität

Digitale Mobilitätsdaten besitzen gegenüber klassischen Erhebungsverfahren erhebliche Vorteile. Traditionelle Haushaltsbefragungen und Wegetagebücher sind kosten- und zeitintensiv und beruhen auf Erinnerungs- und Auskunftsleistungen der Befragten. GPS- und Smartphone-Verfahren können dagegen Wege kontinuierlich erfassen und reduzieren bestimmte Formen des Recall Bias. Zudem lassen sich räumliche Bewegungsmuster mit hoher zeitlicher Auflösung beobachten (Nitsche et al., 2014; Shen & Stopher, 2014). Bereits González et al. (2008) konnten anhand anonymisierter Mobilfunkdaten zeigen, dass großskalige digitale Datensätze neue Erkenntnisse über Regelmäßigkeiten individueller Mobilität ermöglichen. Auch für den öffentlichen Verkehr haben Smartcard-, Automatic-Vehicle-Location- und weitere operative Datenquellen das Spektrum möglicher Analysen deutlich erweitert (Zannat & Choudhury, 2019).

Die Stärke digitaler Daten liegt deshalb insbesondere darin, traditionelle Datengrundlagen zu ergänzen. Problematisch wird ihre Verwendung dagegen, wenn hohe Fallzahlen als Ersatz für Stichprobenqualität interpretiert werden. Bei kommerziellen Smartphone- und Mobilfunkdaten ist die Grundgesamtheit häufig nur eingeschränkt bekannt, weil die Datenerzeugung von Gerätebesitz, Anbieterstruktur, App-Nutzung, Standortfreigabe und weiteren Selektionsprozessen abhängt. Li et al. (2024) zeigen anhand kommerzieller Standortdaten, dass deren Repräsentativität sowohl räumlich und zeitlich als auch hinsichtlich demografischer und sozioökonomischer Merkmale variiert. Auch Jardel und Delamater (2024) identifizieren bei großskaligen Mobilfunkdaten systematische Unterschiede zwischen der erfassten Population und der zugrunde liegenden Bevölkerung. Gerade bei kleinräumigen Analysen können solche Abweichungen für planerische Entscheidungen relevant werden. Große Beobachtungszahlen lösen das Problem damit nicht automatisch. Dies entspricht einem grundlegenden Befund der Big-Data-Forschung, wonach Umfang und technische Granularität eines Datensatzes systematische Selektionsmechanismen nicht kompensieren können (Boyd & Crawford, 2012; Lazer et al., 2014).

Besonders deutlich wird das Problem bei Bevölkerungsgruppen, deren digitale Teilhabe geringer ist. Digitalisierte Mobilitätsangebote setzen häufig Smartphonebesitz, mobile Internetnutzung, Zahlungsoptionen, digitale Kompetenzen oder die Bereitschaft zur Datenfreigabe voraus. Groth (2019) zeigt für Deutschland, dass sich klassische Mobili-

tätsbenachteiligung und digitale Ungleichheit gegenseitig verstärken können. Digitale Exklusion im öffentlichen Verkehr betrifft zudem nicht nur den Besitz eines Geräts. Qualitative Untersuchungen zeigen Probleme durch fehlende digitale Kompetenzen, ungeeignete Geräte, komplexe Benutzeroberflächen oder die Notwendigkeit, unterwegs digitale Technologien nutzen zu müssen. Besonders betroffen sein können ältere Menschen, Personen mit geringerer Bildung, Menschen mit Beeinträchtigungen oder Migrationserfahrung (Durand et al., 2023).

Damit entsteht ein doppeltes Selektionsproblem. Personen können erstens von digitalen Mobilitätsangeboten ausgeschlossen oder in deren Nutzung eingeschränkt sein. Zweitens erzeugen gerade diese Angebote Daten, die wiederum zur Planung zukünftiger Angebote verwendet werden. Wird beispielsweise die Nachfrage nach Mobilitätsangeboten ausschließlich anhand von App-Nutzungsdaten bewertet, bildet der Datensatz definitionsgemäß vor allem diejenigen ab, die bereits Zugang zur App haben und diese aktiv nutzen. Nichtnutzung kann dann fälschlicherweise als Nichtbedarf interpretiert werden.

Auch konventionelle Mobilitätserhebungen sind keineswegs frei von Repräsentativitätsproblemen. Neuere Analysen von Haushaltsverkehrsbefragungen zeigen ebenfalls Unterrepräsentationen vulnerabler Gruppen und verdeutlichen damit, dass nicht zwischen „verzerrten digitalen“ und „unverzerrten klassischen“ Daten unterschieden werden kann. Vielmehr unterscheiden sich Datenquellen in Art und Richtung ihrer jeweiligen Fehler. Gerade deshalb sind multimethodische Ansätze relevant. Vergleichsstudien digital unterstützter Erhebungsverfahren betonen, dass unterschiedliche Methoden kombiniert werden sollten, um Erfassungsdefizite einzelner Verfahren auszugleichen (Safi et al., 2017; Stopher & Greaves, 2007).

Die zentrale Governance-Frage lautet daher nicht, welche einzelne Datenquelle die „richtige“ ist. Entscheidend ist vielmehr, welche Mobilitätsrealität eine Datenquelle sichtbar macht, welche sie systematisch ausblendet und wie diese Grenzen in Entscheidungen berücksichtigt werden.

4 Das FAIR-Mobility-Data-Framework

Aus der Literatursynthese lassen sich vier Anforderungen ableiten, die bei datenbasierten Mobilitätsentscheidungen gleichzeitig berücksichtigt werden sollten.

4.1 Fairness

Fairness betrifft die Verteilungswirkungen datenbasierter Entscheidungen. Aus Sicht der Transport- und Mobility-Justice-Forschung reicht es nicht aus, die durchschnittliche Leistungsfähigkeit eines Verkehrssystems zu verbessern, wenn Nutzen und Belastungen systematisch ungleich verteilt sind (Martens, 2012, 2017; Pereira et al., 2017; Sheller, 2018). Datenbasierte Planungen sollten daher nicht nur fragen, welche Maßnahmen aufgrund beobachteter Nachfrage effizient erscheinen, sondern auch, welche Bevölkerungsgruppen, Räume und Mobilitätsbedürfnisse von der Entscheidung profitieren oder belastet werden.

Operationalisiert werden kann diese Dimension beispielsweise durch räumlich und sozial differenzierte Wirkungsanalysen. Investitionen, Linienanpassungen oder Sharing-Standorte sollten nicht ausschließlich nach aggregierter Nachfrage beurteilt werden. Zusätzlich sollte geprüft werden, ob unterversorgte Räume oder Bevölkerungsgruppen systematisch schlechter abschneiden.

4.2 Accuracy

Accuracy bezeichnet die methodische Qualität und Validität der verwendeten Daten. Digitale Daten können sehr präzise sein, ohne deshalb notwendigerweise valide für die jeweilige Forschungsfrage zu sein. GPS-Daten können beispielsweise Position und Bewegung hochgenau erfassen, während Wegezwecke oder soziodemografische Eigenschaften häufig nur indirekt erschlossen werden müssen (Shen & Stopher, 2014; Xiao et al., 2016).

Vor datenbasierten Entscheidungen sollten deshalb Herkunft, Erfassungsverfahren, Abdeckung, zeitliche Stabilität, Fehlerraten und bekannte Selektionsmechanismen dokumentiert werden. Von besonderer Bedeutung ist die Triangulation mehrerer Datenquellen. Digitale Bewegungsdaten können beispielsweise mit Befragungsdaten, Zensus- beziehungsweise Verwaltungsinformationen, Verkehrszählungen oder qualitativen Erhebungen abgegli-

chen werden. Auch Hagen und Reinfeld (2021) empfehlen für die Mobilitätsforschung ausdrücklich Multimethodenansätze, die Befragungs- und Trackingdaten kombinieren.

4.3 Inclusion

Inclusion fragt danach, wer in einem Datensatz fehlt. Damit verschiebt sich die Aufmerksamkeit von der Größe eines Datensatzes auf dessen soziale und räumliche Abdeckung. Gerade bei digitalen Mobilitätsdaten sollten mögliche Coverage Gaps hinsichtlich Alter, Einkommen, Behinderung, digitaler Kompetenz, Raumtyp und Mobilitätsform geprüft werden (Groth, 2019; Durand et al., 2022, 2023).

Inklusion erfordert daher gegebenenfalls gezielte ergänzende Erhebungen. Werden beispielsweise bestimmte Bevölkerungsgruppen durch App-Daten unzureichend erfasst, können klassische Befragungen, telefonische oder persönliche Erhebungen sowie partizipative Formate eingesetzt werden. Mixed-Mode-Ansätze können Repräsentativitätsdefizite reduzieren, wenn sie gezielt auf schwer erreichbare Gruppen ausgerichtet werden (Safi et al., 2017; Stopher & Greaves, 2007).

4.4 Responsibility

Responsibility betrifft schließlich die institutionelle Verantwortung für den Übergang vom Datensatz zur Entscheidung. Datenbasierte Entscheidungen sind nicht allein technische Ergebnisse eines Modells. Auswahl der Datenquelle, Bereinigung, Gewichtung, Modellierung und Interpretation beruhen auf Entscheidungen von Organisationen und Personen.

Eine verantwortungsvolle Governance verlangt deshalb nachvollziehbare Zuständigkeiten und Transparenz über Annahmen, Datenlücken und Unsicherheiten. Dies gilt insbesondere bei Daten privater Plattformen, deren Erhebungs- und Auswahlmechanismen für öffentliche Planungsstellen möglicherweise nur teilweise sichtbar sind. Der Data-Justice-Ansatz macht deutlich, dass Fragen nach Kontrolle, Sichtbarkeit und Entscheidungsmacht integraler Bestandteil datenbasierter Governance sind (Taylor, 2017; Heeks & Renken, 2018; Dencik et al., 2019; Behrendt & Sheller, 2024).

Das FAIR-Mobility-Data-Framework verbindet damit technische Datenqualität und gesellschaftliche Legitimität. Eine Datenbasis ist für Planungsentscheidungen nicht allein deshalb geeignet, weil sie groß, aktuell oder hochaufgelöst ist. Sie muss zugleich hinreichend valide, inklusiv und hinsichtlich ihrer Verteilungswirkungen sowie Verantwortungsstrukturen überprüfbar sein.

5 Vom Datensatz zur Entscheidung: Governance für eine inklusive Mobilitätsplanung

Aus dem Framework lässt sich ein mehrstufiger Governance-Prozess ableiten. Ausgangspunkt jeder datenbasierten Entscheidung sollte eine Data-Fitness-for-Purpose-Prüfung sein. Dabei wird nicht abstrakt nach der „Qualität" eines Datensatzes gefragt, sondern danach, ob die Daten für die konkrete Entscheidungsfrage geeignet sind. Ein Datensatz, der Verkehrsströme auf Hauptverkehrsachsen sehr präzise abbildet, kann für die Analyse kleinräumiger Mobilitätsbedürfnisse älterer Menschen dennoch ungeeignet sein.

Erstens sollte jeder datenbasierten Entscheidung eine Data-Fitness-for-Purpose-Prüfung vorausgehen. Für die praktische Anwendung lässt sich die Data-Fitness-for-Purpose-Prüfung in vier aufeinander bezogene Prüfschritte übersetzen. Erstens ist die Abdeckung zu prüfen: Welche Population, welche räumlichen Bereiche und welche Mobilitätsformen können durch die Datenquelle überhaupt beobachtet werden? Zweitens ist die Validität zu bewerten: Welche Merkmale werden unmittelbar gemessen und welche lediglich modelliert oder aus anderen Merkmalen abgeleitet? Drittens ist eine Selektivitätsprüfung erforderlich: Welche technischen, sozialen oder ökonomischen Voraussetzungen müssen erfüllt sein, damit Personen im Datensatz sichtbar werden? Viertens ist zu bestimmen, welche Entscheidungsrelevanz verbleibende Datenlücken besitzen. Nicht jede Unterrepräsentation ist für jede Planungsfrage gleichermaßen problematisch. Kritisch wird sie insbesondere dann, wenn gerade diejenigen Gruppen oder Räume unzureichend erfasst werden, die von einer Entscheidung besonders betroffen sind (Shen & Stopher, 2014; Li et al., 2024; Jardel & Delamater, 2024).

Das FAIR-Mobility-Data-Framework ist deshalb nicht als binäre Zertifizierung in „geeignete" und „ungeeignete" Datensätze zu verstehen. Seine Funktion besteht vielmehr darin, die Voraussetzungen

und Grenzen der verwendeten Daten explizit zu machen. Ein Datensatz kann für eine Fragestellung hinreichend akkurat, für eine andere jedoch aufgrund fehlender sozialer oder räumlicher Abdeckung ungeeignet sein. Datenqualität wird damit als kontextabhängige Eigenschaft verstanden, die unmittelbar mit dem jeweiligen Planungszweck verbunden ist.

Zweitens sollte eine Repräsentativitäts- und Exklusionsprüfung folgen. Neben klassischen Stichprobenmerkmalen sollte untersucht werden, welche Voraussetzungen für die Entstehung der Daten erfüllt sein müssen. Dazu gehören beispielsweise Smartphonebesitz, Installation einer bestimmten App, aktive Standortfreigabe, digitale Zahlungsfähigkeit oder die Nutzung eines spezifischen Verkehrsmittels. Diese Voraussetzungen bilden eine potenzielle Selektionskette, die transparent dokumentiert werden sollte.

Drittens sollte bei entscheidungsrelevanten Datenlücken eine methodische Triangulation erfolgen. Dabei geht es nicht darum, digitale Daten durch klassische Befragungen zu ersetzen. Vielmehr können die jeweiligen Stärken kombiniert werden: Digitale Daten liefern hohe räumlich-zeitliche Auflösung; Befragungen erschließen Motive, Bedürfnisse und sozioodemografische Hintergründe; administrative Daten bieten Informationen über Bevölkerungsstrukturen; partizipative Verfahren können Bedürfnisse sichtbar machen, die in keiner bestehenden Datenquelle ausreichend erfasst werden.

Viertens sollte die Unsicherheit der Daten auch in der Entscheidungskommunikation sichtbar bleiben. Eine Scheingenauigkeit digitaler Kennzahlen ist insbesondere dann problematisch, wenn unbekannte Selektionsmechanismen oder algorithmische Klassifikationen hinter präzisen Prozentwerten verschwinden. Angaben zur Datenabdeckung, zu bekannten Verzerrungen und zur Sensitivität von Ergebnissen gegenüber alternativen Datenquellen sollten deshalb Bestandteil von Planungs- und Entscheidungsunterlagen sein.

Schließlich sollte Verantwortung institutionell verankert werden. Kommunen und Verkehrsunternehmen benötigen klare Zuständigkeiten dafür, wer die Eignung externer Datensätze prüft, Repräsentativitätsrisiken bewertet und gegebenenfalls ergänzende Erhebungen veranlasst. Für private Mobilitätsplattformen ergibt sich umgekehrt die Anforderung, Metadaten über Erhebung, Abdeckung und wesentliche Selektionsmechanismen so weit offenzulegen, dass öffentliche Stellen die Aussagekraft der Daten beurteilen können. Abbildung 2 visualisiert den daraus abgeleiteten Governance-Prozess vom Datensatz zur Planungsentscheidung.

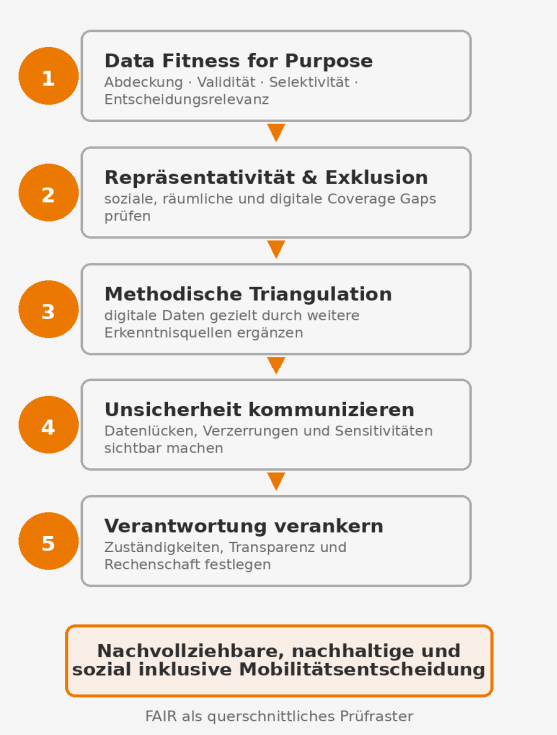

Abbildung 2: Governance-Prozess für eine inklusive datenbasierte Mobilitätsplanung. (Eigene Darstellung)

Ein solcher Governance-Ansatz verändert die Logik datenbasierter Mobilitätsplanung. Die zentrale Frage lautet dann nicht mehr ausschließlich: „Was zeigen die Daten?“, sondern zusätzlich: „Wer ist in diesen Daten sichtbar, wer fehlt und welche Konsequenzen entstehen daraus für die Entscheidung?“

6 Fazit

Digitale Mobilitätsdaten erweitern die Möglichkeiten der Verkehrsplanung erheblich. Sie können Bewegungen hochaufgelöst abbilden, Nachfrageveränderungen schneller sichtbar machen und klassische Erhebungsverfahren sinnvoll ergänzen. Ihre große Fallzahl und technische Präzision dürfen jedoch nicht mit gesellschaftlicher Repräsentativität gleichgesetzt werden.

Die Literatur zeigt, dass digitale Mobilitätsdaten durch technische, soziale und institutionelle Selektionsmechanismen geprägt sind. Insbesondere dort, wo Smartphone-, App- oder Plattformnutzung Voraussetzung der Datenerzeugung ist, können bestehende digitale und mobilitätsbezogene Ungleichheiten auch in die Datenbasis zukünftiger Entscheidungen eingehen.

Das entwickelte FAIR-Mobility-Data-Framework adressiert dieses Problem über die vier miteinander verbundenen Dimensionen Fairness, Accuracy, Inclusion und Responsibility. Sein zentraler Mehrwert liegt darin, Datenqualität nicht ausschließlich technisch zu verstehen. Gute Mobilitätsdaten müssen nicht nur korrekt und umfangreich sein, sondern auch hinsichtlich ihrer sozialen Selektivität, ihrer Verteilungswirkungen und der Verantwortung für darauf basierende Entscheidungen geprüft werden.

Für die Praxis folgt daraus kein Verzicht auf digitale Daten, sondern deren reflektierte Kombination mit anderen Erkenntnisquellen. Triangulation, Repräsentativitätsprüfungen, transparente Dokumentation von Unsicherheiten, gezielte Ergänzung unterrepräsentierter Gruppen und institutionelle Verantwortlichkeit sind deshalb zentrale Bestandteile einer nachhaltigen datenbasierten Mobilitätsplanung. Mobilitätsdaten werden damit nicht als vermeintlich neutrales Abbild von Mobilität verstanden, sondern als eine von mehreren Grundlagen für nachvollziehbare und gesellschaftlich inklusive Entscheidungen.

Literaturverzeichnis

Behrendt, F. & Sheller, M. (2024): Mobility data justice. Mobilities, 19(1), S. 151–169. https://doi.org/10.1080/17450101.2023.2200148

Boyd, d. & Crawford, K. (2012): Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. Information, Communication & Society, 15(5), S. 662–679. https://doi.org/10.1080/1369118X.2012.678878

Dencik, L., Hintz, A., Redden, J. & Treré, E. (2019): Exploring data justice: Conceptions, applications and directions. Information, Communication & Society, 22(7), S. 873–881. https://doi.org/10.1080/1369118X.2019.1606268

Durand, A., Zijlstra, T., van Oort, N., Hoogendoorn-Lanser, S. & Hoogendoorn, S. (2022): Access denied? Digital inequality in transport services. Transport Reviews, 42(1), S. 32–57. https://doi.org/10.1080/01441647.2021.1923584

Durand, A., Zijlstra, T., Hamersma, M., 't Hoen, A., van Oort, N., Hoogendoorn, S. & Hoogendoorn-Lanser, S. (2023): Fostering an inclusive public transport system in the digital era: An interdisciplinary approach. Transportation Research Interdisciplinary Perspectives, 22, 100968. https://doi.org/10.1016/j.trip.2023.100968

Floridi, L. & Taddeo, M. (2016): What is data ethics? Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2083), 20160360. https://doi.org/10.1098/rsta.2016.0360

González, M. C., Hidalgo, C. A. & Barabási, A.-L. (2008): Understanding individual human mobility patterns. Nature, 453, S. 779–782. https://doi.org/10.1038/nature06958

Groth, S. (2019): Multimodal divide: Reproduction of transport poverty in smart mobility trends. Transportation Research Part A: Policy and Practice, 125, S. 56–71. https://doi.org/10.1016/j.tra.2019.04.018

Hagen, T. & Reinfeld, N. (2021): Möglichkeiten und Grenzen der Analyse neuer Mobilitätsangebote mittels existierender Befragungen und neuartiger Erhebungsmethoden. Journal für Mobilität und Verkehr, 10, S. 15–25. https://doi.org/10.34647/jmv.nr10.id68

Heeks, R. & Renken, J. (2018): Data justice for development: What would it mean? Information Development, 34(1), S. 90–102. https://doi.org/10.1177/0266666916678282

Jabareen, Y. (2009): Building a conceptual framework: Philosophy, definitions, and procedure. International Journal of Qualitative Methods, 8(4), S. 49–62. https://doi.org/10.1177/160940690900800406

Jardel, H. V. & Delamater, P. L. (2024): Uncovering representation bias in large-scale cellular phone-based data: A case study in North Carolina. Geographical Analysis, 56(4), S. 723–745. https://doi.org/10.1111/gean.12399

Lazer, D., Kennedy, R., King, G. & Vespignani, A. (2014): The parable of Google Flu: Traps in big data analysis. Science, 343(6176), S. 1203–1205. https://doi.org/10.1126/science.1248506

Li, Z., Ning, H., Jing, F. & Lessani, M. N. (2024): Understanding the bias of mobile location data across spatial scales and over time: A comprehensive analysis of SafeGraph data in the United States. PLOS ONE, 19(1), e0294430. https://doi.org/10.1371/journal.pone.0294430

Lucas, K. (2012): Transport and social exclusion: Where are we now? Transport Policy, 20, S. 105–113. https://doi.org/10.1016/j.tranpol.2012.01.013

Martens, K. (2012): Justice in transport as justice in accessibility: Applying Walzer's 'Spheres of Justice' to the transport sector. Transportation, 39(6), S. 1035–1053. https://doi.org/10.1007/s11116-012-9388-7

Martens, K. (2017): Transport Justice: Designing Fair Transportation Systems. New York/London: Routledge. https://doi.org/10.4324/9781315746852

Nitsche, P., Widhalm, P., Breuss, S., Brändle, N. & Maurer, P. (2014): Supporting large-scale travel surveys with smartphones – A practical approach. Transportation Research Part C: Emerging Technologies, 43, S. 212–221. https://doi.org/10.1016/j.trc.2013.11.005

Pereira, R. H. M., Schwanen, T. & Banister, D. (2017): Distributive justice and equity in transportation. Transport Reviews, 37(2), S. 170–191. https://doi.org/10.1080/01441647.2016.1257660

Safi, H., Assemi, B., Mesbah, M. & Ferreira, L. (2017): An empirical comparison of four technology-mediated travel survey methods. Journal of Traffic and Transportation Engineering (English Edition), 4(1), S. 80–87. https://doi.org/10.1016/j.jtte.2015.12.003

Shen, L. & Stopher, P. R. (2014): Review of GPS travel survey and GPS data-processing methods. Transport Reviews, 34(3), S. 316–334. https://doi.org/10.1080/01441647.2014.903530

Sheller, M. (2018): Mobility Justice: The Politics of Movement in an Age of Extremes. London: Verso.

Snyder, H. (2019): Literature review as a research methodology: An overview and guidelines. Journal of Business Research, 104, S. 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Stopher, P. R. & Greaves, S. P. (2007): Household travel surveys: Where are we going? Transportation Research Part A: Policy and Practice, 41(5), S. 367–381. https://doi.org/10.1016/j.tra.2006.09.005

Taylor, L. (2017): What is data justice? The case for connecting digital rights and freedoms globally. Big Data & Society, 4(2), S. 1–14. https://doi.org/10.1177/2053951717736335

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., 't Hoen, P. A. C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. & Mons, B. (2016): The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018. https://doi.org/10.1038/sdata.2016.18

Xiao, G., Juan, Z. & Zhang, C. (2016): Detecting trip purposes from smartphone-based travel surveys with artificial neural networks and particle swarm optimization. Transportation Research Part C: Emerging Technologies, 71, S. 447–463. https://doi.org/10.1016/j.trc.2016.08.008

Zannat, K. E. & Choudhury, C. F. (2019): Emerging big data sources for public transport planning: A systematic review on current state of art and future research directions. Journal of the Indian Institute of Science, 99(4), S. 601–619. https://doi.org/10.1007/s41745-019-00125-9

Autor*innenangaben

Ann-Katrin Voit
Professur für Volkswirtschaftslehre, insb. Wirtschaftspolitik
Herkulesstr. 32
45127 Essen

E-Mail: ann-katrin.voit@rub.de