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.

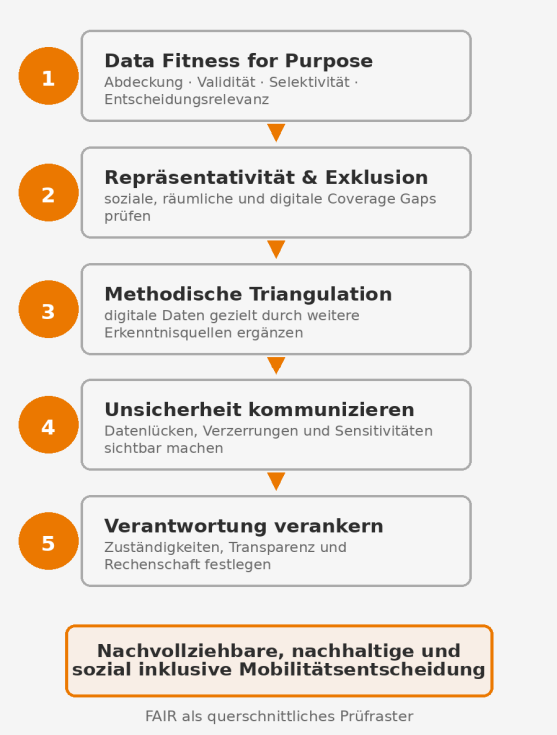


Abbildung 2: Governance-Prozess für eine inklusive datenbasierte Mobilitätsplanung. (Eigene Darstellung)

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