abida
ASSESSING BIG DATA

Digitale Transformation

BIG DATA IM ÖFFENTLICHEN PERSONENVERKEHR GUTACHTEN

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

01IS15016A-F

ConPolicy – Institut für Verbraucherpolitik

Dr. Julius Rauber
Prof. Dr. Christian Thorun

gemeinsam mit:

Innovationszentrum für Mobilität und gesellschaftlichem Wandel (InnoZ) GmbH

Dr. Robert Schönduwe
Benno Bock
Marc Schelewsky

ABIDA - ASSESSING BIG DATA

PROJEKTLAUFZEIT 01.03.2015-28.02.2019

Westfälische Wilhelms-Universität Münster,
Institut für Informations-, Telekommunikations- und
Medienrecht (ITM), Zivilrechtliche Abteilung

Karlsruher Institut für Technologie,
Institut für Technikfolgenabschätzung
und Systemanalyse (ITAS)

Leibniz Universität Hannover
Institut für Rechtsinformatik
(IRI)

Technische Universität Dortmund,
Wirtschafts- und Sozialwissenschaftliche
Fakultät (WiSo) Techniksoziologie

Ludwig-Maximilians-Universität München,
Forschungsstelle für Information, Organisation
und Management (IOM)

Wissenschaftszentrum
Berlin für Sozialforschung

ABIDA - Assessing Big Data
Über das Gutachten

Das Gutachten wurde im Rahmen des ABIDA-Projekts mit Mitteln des Bundesministeriums für Bildung und Forschung erstellt. Der Inhalt des Gutachtens gibt ausschließlich die Auffassungen der Autoren wieder. Diese decken sich nicht automatisch mit denen des Ministeriums und/oder der einzelnen Projektpartner.

ABIDA lotet gesellschaftliche Chancen und Risiken der Erzeugung, Verknüpfung und Auswertung großer Datenmengen aus und entwirft Handlungsoptionen für Politik, Forschung und Entwicklung.

www.abida.de
© 2018 – Alle Rechte vorbehalten

ZUSAMMENFASSUNG

Dieses Gutachten zum Thema *Digitale Transformation – Big Data im öffentlichen Personenverkehr* fasst bestehende Big Data-basierte Anwendungen im öffentlichen Personenverkehr (ÖPV) zusammen und analysiert die Einstellung der Nutzer des ÖPV gegenüber möglichen Innovationen in diesem Bereich. Basierend auf beiden Erkenntnissträngen werden Thesen zum Einsatz von Big Data im ÖPV formuliert und potentielle Handlungsempfehlungen für die Betreiber des ÖPV als auch für die Politik als regulierende Instanz abgeleitet.

Dazu werden zunächst im ersten Kapitel als Grundlage die Begriffe *Big Data* und *öffentlicher Personenverkehr* definiert und vorgestellt sowie ein Überblick über das methodische Vorgehen gegeben.

Im zweiten Kapitel werden potentielle Einsatzmöglichkeiten von Big Data im Bereich des ÖPV sowie anderen Verkehrsbereichen aufgezeigt. Dabei wird zwischen anbieterseitigen Anwendungen – z.B. bei der Planung und Wartung – und kundenorientierten Anwendungen – z.B. bei der Informationsvermittlung, Intermodalitätsförderung oder Marketing – unterschieden. Es wird deutlich, dass sich in der ÖPV-Branche im Vergleich zur Flugbranche oder Sharing-Economy noch wenige Beispiele für Big Data-basierte Anwendungen finden lassen.

Im dritten Kapitel wird eine repräsentative Befragung zum Thema des Gutachtens vorgestellt, die die Einstellung der Nutzer widerspiegelt. Dabei wird deutlich, dass die Nutzer bestehende digitale Angebote des ÖPV kennen und nutzen, aber die Aspekte Datenschutz, Datensicherheit, technische Sicherheit, Preis und soziale Inklusion im Zweifel über die Einführung von Big Data-basierten Innovationen stellen würden. Generell zeigen die Ergebnisse der Befragung aber auch, dass Big Data-basierte Anwendungen wirtschaftliche und ökologische Vorteile durch eine Erhöhung der Nutzerzahl für den ÖPV bieten können. Hierbei sind insbesondere Szenarien wie eine automatische Rückerstattung oder Entschädigungszahlungen bei Verspätungen sowie eine Anpassung des Angebots an das Verkehrsaufkommen in Echtzeit zu nennen.

Im vierten Kapitel werden dann die gewonnenen Erkenntnisse des Gutachtens reflektiert und Handlungsempfehlungen abgeleitet. Es wird deutlich, dass Investitionen in Big Data nachgeholt und – basierend auf durch Benchmarking zu identifizierende Best-Practice-Beispiele – entsprechende Entwicklungen gefördert werden sollten. Hierbei ist allerdings die Heterogenität des ÖPV-Sektors insbesondere bezüglich des Digitalisierungsgrads der verschiedenen Anbieter zu berücksichtigen. Auch müssen Wege für eine effiziente Datenintegration und -bereitstellung für Big Data-basierte Anwendungen im ÖPV – beispielsweise in Form von frei zugänglichen Plattformen mit einheitlichen Standards – gefunden werden. Von besonderer Bedeutung bei der Entwicklung solcher neuen Anwendungen sind aus Nutzersicht die Aspekte Datenschutz und Datensicherheit. Daher sollten die Prinzipien *Privacy and Security by Design* in der Entwicklung neuer Anwendungen von Anfang an berücksichtigt werden. Aus struktureller und regulatorischer Sicht sollten neue Möglichkeiten im Dreiklang von Besteller, Ersteller und privaten Unternehmen – etwa durch die Ausweitung von Ausschreibungen um digitale Komponenten oder den Ausbau der Förderung öffentlich-privaten Partnerschaften – entwickelt werden.

Inhaltsverzeichnis

1. Einleitung ... 6

1.1 Hintergrund ... 6

1.2 Zielsetzung und Fragestellung des Gutachtens ... 7

1.3 Methodisches Vorgehen ... 8

1.3.1 Angebotsseite ... 9

1.3.2 Nutzerseite ... 9

1.4 Definitionen von Big Data und Öffentlichem Personenverkehr ... 9

1.4.1 Big Data ... 10

1.4.2 Öffentlicher Personenverkehr ... 13

2. Betrachtung der Angebotsseite ... 16

2.1 Eigenschaften und Technologien ... 16

2.1.1 Eigenschaften ... 17

2.1.2 Big Data-Technologie Stack ... 17

2.1.3 Kategorisierung von Mobilitätsdaten ... 20

2.2 Big Data-Anwendungen im ÖPV ... 22

2.2.1 Anbieterseitige Anwendungen ... 24

2.2.1.1 Langfristige Angebotsplanung ... 24

2.2.1.2 Kurzfristige Nachfrageprognose ... 26

2.2.1.3 Fahrzeug- und Infrastrukturwartung ... 28

2.2.1.4 Ticketing und Fahrgeldmanagement ... 30

2.2.2 Kundenorientiere Anwendungen ... 32

2.2.2.1 Informationsvermittlung ... 32

2.2.2.2 Mobilitätsplattformen und MaaS ... 34

2.2.2.3 Kundenkommunikation und Marketing ... 35

2.2.3 Verwertung durch Dritte ... 36

3. Betrachtung der Nutzerseite ... 37

3.1 Bisherige Erhebungen zur Nutzermeinung über Big Data im ÖPV ... 37

3.2 Konzeption der Befragung ... 38

3.2.1 Interviews mit Vertretern von ... 38

Fahrgastverbänden und Anbietern als Grundlage für den Fragebogen ... 38

3.2.2 Aufbau des Fragebogens und Durchführung der Befragung ... 39

3.3 Ergebnisse der Befragung ... 39

3.3.1 Verteilung der Stichprobe 39
3.3.2 ÖPV wird genutzt, aber nicht geliebt – Junge, gebildete Großstädter nutzen ÖPNV am häufigsten 40
3.3.3 Customer Journey – Ticketpreise und Service bei Beschwerden im Fokus der Kritik 42
3.3.4 Bestehende Digitale Angebote – Routenplanung und Online-Tickets sind bekannt und werden genutzt 44
3.3.5 Innovationen im ÖPV – Befragte sind risikoavers bei Big Data-basierten Neuerungen 46
3.3.6 Zukunft von Big Data im ÖPV – Befragte wünschen flexibleres Angebot und einfachere Rückerstattungsmöglichkeiten 47
4. Reflexion der Ergebnisse und Handlungsempfehlungen 51
5. Glossar 57
6. Quellen 59
7. Anhang 62

ABKÜRZUNGEN

AWS: Amazon Web Services
BMVI: Bundesministerium für Verkehr und digitale Infrastruktur
BVG: Berliner Verkehrsbetriebe
DB AG: Deutsche Bahn AG
EC2: Elastic Compute Cloud
ECS: Elastic Container Service
FCD: Floating-Car-Data
FGSV: Forschungsgesellschaft für Straßen- und Verkehrswesen
FIN: Fakultät für Informatik
GPS: Global Positioning System
GTFS: General Transit Feed Specification
HAFAS: HaCon Fahrplan-Auskunfts-System
IAO: Institut für Arbeitswirtschaft und Organisation
IoT: Internet of Things
ITS: Intelligent Transport Systems
KNN: Künstliche neuronale Netzwerke
MaaS: Mobility as a Service
Marego: Magdeburger Regionalverkehrsverbund
MIV: Motorisierter Individualverkehr
NFC: Near-Field Communication
ÖPFV: Öffentlicher Personenfernverkehr
ÖPNV: Öffentlicher Personennahverkehr
ÖPV: Öffentlicher Personenverkehr
PBefG: Personenbeförderungsgesetz
RDS: Relational Database Services
RFID: Radio-Frequency Identification
RMV: Rhein-Main-Verkehrsbund
RNV: Rhein-Neckar-Verkehr
SNS: Simple Notification Service
SQS: Simple Queue Service
SSB: Stuttgarter Straßenbahnen
TfL: Transport for London
UDA: Teradata Unified Data Architecture
VCD: Verkehrsclub Deutschland e.V.
VDV: Verband Deutscher Verkehrsunternehmen
VRS: Verkehrsverbund Rhein-Sieg
ZOD: ZERO.ONE.DATA

ABBILDUNGEN

Abb. 1: Eigenschaften von Big Data 12
Abb. 2: Beispiele für Datenvolumen im Verkehrsbereich 17
Abb. 3: Anbieter und Technologien im Kontext von Big Data 19
Abb. 4: Akteurskonstellation am Beispiel von Telefonica NEXT und flinc 26
Abb. 5: Stimulation der Nachfrage beim Carsharing-Anbieter DriveNow 27
Abb. 6: Akteurskonstellation am Beispiel von MotionTag und DB Vertireb 31
Abb. 7: Akteurskonstellation am Beispiel von HaCon und ÖPNV-Anbieter 33
Abb. 8: Big Data-gestützte Kapazitätsabschätzung bei der SBB 33
Abb. 9: Akteurskonstellation am Beispiel von MOIA 35

TABELLEN

Tab. 1: Big Data-Charakteristika 11
Tab. 2: Abgrenzung öffentlicher und privater Verkehr 14
Tab. 3 Big Data-Technologie-Stack 18
Tab. 4: Verteilung der Stichprobe 40
Tab. 5: Nutzung von Verkehrsmitteln (geordnet nach häufigster Nutzung in Prozent) 40
Tab. 6: Zufriedenheit mit Verkehrsmitteln (geordnet nach Zufriedenheit in Prozent) 41
Tab. 7: Nutzung des ÖPNV (geordnet nach Bildungsabschluss) 41
Tab. 8: Nutzung des ÖPNV (geordnet nach Stadtgröße) 42
Tab. 9: Zufriedenheit mit Aspekten des ÖPV (geordnet nach Unzufriedenheit in Prozent) 43
Tab. 10: Zufriedenheit mit der Häufigkeit der Abfahrten/Taktung des ÖPV (geordnet nach Stadtgröße in Prozent) 44
Tab. 11: Bekanntheit und Nutzung bestimmter digitaler Angebote des ÖPV (geordnet nach Bekanntheit in Prozent) 44
Tab. 12: Unterschiede nach Geschlecht in der Nutzung bestimmter digitaler Angeboten des ÖPV (geordnet nach Nutzung in Prozent) 45
Tab. 13 Risiken und Vorteile von Big Data im ÖPV im Vergleich (geordnet nach „sehr große Zustimmung“ in Prozent) 46
Tab. 14: Geschätzte Bedeutung der Szenarien zu Big Data im ÖPV (geordnet nach Bedeutung „sehr wichtig“ in Prozent) 47
Tab. 15: Geschätzte Nutzungsänderung bei Realisierung der Szenarien zu Big Data im ÖPV (geordnet nach zunehmender Nutzung in Prozent) 49

Tab. A 1: Nutzung von ÖPNV (geordnet nach Altersklassen) 62
Tab. A 2: Nutzung von Bahnfernverkehr (geordnet nach Altersklassen) 62
Tab. A 3: Nutzung von Busfernverkehr (geordnet nach Altersklassen) 62
Tab. A 4: Nutzung von Bahnfernverkehr (getrennt nach Bildungsabschluss) 62
Tab. A 5: Nutzung von Busfernverkehr (getrennt nach Bildungsabschluss) 63
Tab. A 6: Nutzung von Bahnfernverkehr (geordnet nach Stadtgröße) 63
Tab. A 7: Nutzung von Busfernverkehr (geordnet nach Stadtgröße) 64
Tab. A 8: Zufriedenheit mit dem Preis des ÖPV (getrennt nach Geschlecht) 64
Tab. A 9: Zufriedenheit mit der Funktionstüchtigkeit des ÖPV (getrennt nach Geschlecht) 64
Tab. A 10: Zufriedenheit mit der Sauberkeit des ÖPV (getrennt nach Geschlecht) 64
Tab. A 11: Zufriedenheit mit den Umsteigezeiten des ÖPV (getrennt nach Geschlecht) 65
Tab. A 12: Zufriedenheit mit Informationen bei Verspätungen (getrennt nach Geschlecht) 65
Tab. A 13: Zufriedenheit mit dem Umgang mit Rückerstattungsforderungen im ÖPV (geordnet nach Altersklassen) 65
Tab. A 14: Zufriedenheit mit der Bearbeitungsdauer bei Rückerstattungsforderungen im ÖPV (geordnet nach Altersklassen) 65
Tab. A 15: Zufriedenheit mit den Umsteigezeiten im ÖPV (geordnet nach Stadtgröße) 66
Tab. A 16: Zufriedenheit mit der Anzahl der Anschlussmöglichkeiten (geordnet nach Stadtgröße) 66
Tab. A 17: Nutzung von digitalen Angeboten (geordnet nach Altersklassen) 67

Tab. A 18: Kenntnis von digitalen Angeboten (getrennt nach Bildungsabschluss) ... 68
Tab. A 19: Nutzung von digitalen Angeboten (getrennt nach Bildungsabschluss) ... 68
Tab. A 20: Einstellung zum Datenschutz (getrennt nach Geschlecht) ... 69
Tab. A 21: Einstellung zu unausgereiften Innovationen (getrennt nach Geschlecht) ... 69
Tab. A 22: Einstellung zum Datenschutz (geordnet nach Altersklassen) ... 69
Tab. A 23: Einstellung zu Datensicherheit (geordnet nach Altersklassen) ... 70
Tab. A 24: Einstellung zu Preiserhöhung (geordnet nach Altersklassen) ... 70
Tab. A 25: Einstellung zu Benachteiligung von Nutzergruppen (geordnet nach Altersklassen) ... 70
Tab. A 26: Einstellung zur Ausgereiftheit von Innovationen (geordnet nach Altersklassen) ... 71
Tab. A 27: Einstellung zum Datenschutz (getrennt nach Bildungsabschluss) ... 71
Tab. A 28: Einstellung zu Datensicherheit (getrennt nach Bildungsabschluss) ... 71
Tab. A 29: Einstellung zu Preiserhöhung (getrennt nach Bildungsabschluss) ... 71
Tab. A 30: Einstellung zur Benachteiligung von bestimmten Nutzergruppen (getrennt nach Bildungsabschluss) ... 72
Tab. A 31: Einstellung zur Ausgereiftheit von Innovationen (getrennt nach Bildungsabschluss) ... 72
Tab. A 32: Geschätzte Bedeutung des Szenarios zu personalisiertem Entertainment- und Informationsangebot (geordnet nach Altersgruppen) ... 72
Tab. A 33: Geschätzte Bedeutung des Szenarios zu personalisierter Sitzplatzreservierung (geordnet nach Altersgruppen) ... 72
Tab. A 34: Geschätzte Bedeutung des Szenarios zur Anpassung der Anzahl der eingesetzten Fahrzeuge an das Fahrgastaufkommen (getrennt nach Bildungsabschluss) ... 73
Tab. A 35: Geschätzte Bedeutung des Szenarios eines personalisierten Entertainment- und Informationsangebots im WLAN des ÖPV (geordnet nach Stadtgröße) ... 73
Tab. A 36: Nutzungsänderung bei Realisierung des Szenarios zu personalisiertem Entertainment- und Informationsangebot (geordnet nach Altersgruppen) ... 73
Tab. A 37: Nutzungsänderung bei Realisierung des Szenarios zu personalisiertem Entertainment- und Informationsangebot (getrennt nach Geschlecht) ... 74
Tab. A 38: Nutzungsänderung bei Realisierung des Szenarios zu flexiblen Busrouten (getrennt nach Geschlecht) ... 74
Tab. A 39: Nutzungsänderung bei Realisierung des Szenarios zu an die aktuelle Nachfrage angepassten Preisen (getrennt nach Geschlecht) ... 74
Tab. A 40: Nutzungsänderung bei Realisierung des Szenarios zu flexiblen Busrouten (geordnet nach Stadtgröße) ... 75

1. EINLEITUNG

1.1 HINTERGRUND

Mit rasanter Geschwindigkeit transformiert die Digitalisierung unterschiedlichste Lebens- und Wirtschaftsbereiche. Diese Transformation hat auch den Verkehrssektor erfasst und ermöglicht neuartige Mobilitätsangebote und -lösungen. Zentral sind hierbei die Verschmelzung von Information und Navigation, Buchung, Ticketing, Abrechnung und Clearing sowie Automatisierung. Über Passivdaten, Trackinginstrumente, Management-Dashboards sowie kundenbezogene Dispo- und Informationssysteme entsteht gleichzeitig eine neue Sicht auf den Puls der Mobilität, die in Echtzeit immer umfassender beobachtet und geplant werden kann.

Eine wesentliche technologische Grundlage für diese Veränderungen stellen Big Data-Methoden und Technologien dar. Es überrascht daher auch nicht, dass die Diskussion über Anwendungsmöglichkeiten von Big Data im öffentlichen Diskurs in den vergangenen Jahren stark zugenommen hat. Experten prognostizieren vor allem im Servicebereich ein starkes Umsatzwachstum von 1,3 Milliarden Euro in 2017 auf 2,2 Milliarden Euro in 2020.[1] Auch im Bereich Mobilität werden derzeit die Potentiale von verschiedenen Stakeholdern untersucht, da es hier interessante Möglichkeiten von Big Data-Anwendungen, wie z.B. für die Optimierung der Verkehrsplanung des öffentlichen Personenverkehrs (ÖPV), gibt. Gerade im Bereich des ÖPV sind große Potentiale zu vermuten, da über 10 Millionen deutsche Bürger[2] angeben, den öffentlichen Personennahverkehr (ÖPNV) täglich oder fast täglich zu nutzen.[3] Laut der Statistik des Verbandes Deutscher Verkehrsunternehmen (VDV) im Jahre 2016[4] sind im öffentlichen Personenverkehr deutliche Zuwächse sowohl im Hinblick auf das Verkehrsaufkommen als auch der Verkehrsleistung zwischen den Jahren 2003 und 2014 zu verzeichnen. Dies gilt ebenso für den Radverkehr. Im Unterschied zum monomodalen motorisierten Individualverkehr (MIV) besteht im ÖPV allerdings wenig Kenntnis über die spezifischen Muster und Verknüpfungen der Mobilitätsangebote. Die Potentiale von Big Data bestehen somit unter anderem in der Verbesserung der bestehenden Verkehrsangebote und deren Verknüpfung (z.B. zur Optimierung inter- oder multimodaler Angebote) sowie der Entwicklung korrespondierender Applikationen. Dies hätte nicht nur eine Auswirkung auf die Betreiber der öffentlichen Personenverkehrsangebote zur Folge, sondern auch auf einen großen Teil der Nachfrageseite öffentlicher Verkehrsangebote. Auch gehen einige von positiven Auswirkungen auf die Umwelt durch den Einsatz von Big Data aus, wenn beispielsweise der ÖPV

1 Experton. (2015). Prognose zum Umsatz mit Big Data-Lösungen in Deutschland von 2015 bis 2020 nach Segment. Abgerufen am 21. Juli 2017 von https://de.statista.com/statistik/daten/studie/603793/umfrage/umsatz-mit-Big Data-loesungen-in-deutschland-nach-segment/.

2 Auf eine geschlechtsindikative Doppelung oder Endung wird aus stilistischen Gründen verzichtet. Der Plural meint immer beide: Bürgerinnen und Bürger, Nutzerinnen und Nutzer etc.

3 IfD Allensbach. (2017). Bevölkerung in Deutschland nach Nutzungshäufigkeit öffentlicher Nahverkehrsmittel (Bus, Straßenbahn, U-Bahn, S-Bahn) von 2012 bis 2015. Abgerufen am 21. Juli 2017 von https://de.statista.com/statistik/daten/studie/170884/umfrage/nutzungshaeufigkeit-oeffentlicher-verkehrsmittel/.

4 Vgl. VDV-Statistik 2016 (2016). Danach stieg der Anteil der beförderten Personen im Segment "Eisenbahn" von 2003 bis 2014 um 33 %, gleichzeitig konnte ein Anstieg der Verkehrsleistung um fast 20 Mrd. Personenkilometer verzeichnet werden. Abgerufen am 19.02.2018 von https://www.vdv.de/jahresbericht---statistik.aspx.

dadurch weiter an Attraktivität gewinnen und in Folge dessen der Anteil an Autofahrten zurückgehen würde.

Den positiven Erwartungen an den Einsatz von Big Data im ÖPV stehen aber auch einige kritische Überlegungen gegenüber. So wirft beispielsweise die potentielle Verarbeitung persönlicher Mobilitätsdaten aus datenschutzrechtlicher Sicht einige Fragen auf. Auch ist noch nicht klar, welche Innovationen kommen werden und ob diese überhaupt von den entsprechenden Nutzergruppen akzeptiert, als sinnvoll betrachtet werden und sich in Routinen überführen lassen. Aus diesen Überlegungen ergeben sich aus wissenschaftlicher Sicht zwei Fragestellungen: Welche Einsatzmöglichkeiten bietet Big Data im Angebot des ÖPV und – soweit die Nutzer von Big Data betroffen sind – wie groß ist die Akzeptanz der Nutzer für die verschiedenen Einsatzmöglichkeiten von Big Data-Technologien.

1.2 ZIELSETZUNG UND FRAGESTELLUNG DES GUTACHTENS

Das Gutachten zielt darauf ab, gegenwärtige und zukünftige Entwicklungen des Einsatzes von Big Data im ÖPV-Bereich zu identifizieren und dabei insbesondere technologische, ökonomische und sozio-demographische Aspekte zu adressieren. Hierbei werden sowohl die Auswirkungen auf Verkehrsanbieter als auch auf die Nutzer beleuchtet.

Zum Zweck einer technologischen Verortung des Status quo wird die gängige Praxis zum Einsatz digitaler Anwendungen im ÖPV untersucht. Die Darstellung von innovativen Produktlösungen der wichtigsten IT-Dienstleister sowie neuartiger Dienste ausgewählter Start-Ups soll darüber hinaus den aktuellen Impuls in Richtung Big Data, Digitalisierung und Automatisierung umreißen und dabei beispielhaft die Rolle von Big Data-Technologien verdeutlichen. Darauf aufbauend werden Optimierungspotentiale aufgedeckt, indem z.B. der Einsatz von *Predictive Modelling* in der kurzfristigen Prognose von Fahrgastaufkommen, die Ausübung intelligenter Wartung als *condition-based-maintainance*, oder der Einsatz von Business Intelligence im Management untersucht werden. Mögliche Änderungen in der Marktstruktur (z.B. neue Intermediäre zur Datenbereitstellung oder neue Dienstleister mit Software-as-a-Service-Lösungen) sollen durch den Einsatz von Big Data für die Anbieterseite, also die Betreiber des öffentlichen Personennahverkehrs sowie des Fernverkehrs, aufgezeigt werden. Wichtig für die Umsetzung einer Big Data-Strategie für den ÖPV ist dabei auch die Bewertung der rechtlichen Rahmenbedingungen, die sowohl Innovationshürden darstellen oder aber Innovationen in eine gewünschte Richtung lenken können, die politisch gewollt und sozial gewünscht ist.

Komplementär zur Betrachtung der Anbieterseite werden Implikationen für die Nutzer, also die Nutzerseite der öffentlichen Verkehrsangebote, ins Auge gefasst. Deren Verhalten und Reaktion auf den Einsatz von Big Data im öffentlichen Personenverkehr wird von immenser Bedeutung für den Erfolg verschiedener anbieterseitiger Maßnahmen sein. Deshalb wurden im Rahmen einer Online-Befragung die Einschätzungen und die Wünsche von Nutzern an den Einsatz von Big Data im öffentlichen Nah- und Fernverkehr untersucht. Hierbei stehen Probleme in der Akzeptanz bezüglich der Nutzung der personenbezogenen Verkehrsdaten, Bedürfnisse von Nutzern an neuen Angeboten sowie mögliche Auswirkungen auf das Nutzungsverhalten (u.a. Rebound-Effekte und soziodemographische Unterschiede) im Mittelpunkt.

Die Ergebnisse der Anbieter- und Nutzerseite werden zusammengefasst betrachtet und darauf aufbauend erfolgt ein Ausblick auf mögliche Entwicklungen der Anbieter- und Nachfrageseite, die kontrastiert und gegenüberstellt werden. Diese Entwicklungen werden auf Wirtschaftlichkeit,

ökologische Nachhaltigkeit, politische Umsetzbarkeit sowie Akzeptanz durch die Nutzer diskutiert. Im Ergebnis werden Handlungsempfehlungen formuliert, um die Potentiale von Big Data nutzbar zu machen. Diese reflektieren dabei die Spezifika der öffentlichen Verkehrsbranche und beachten normative sowie regulatorische Grundsätze wie Nachhaltigkeit und Daseinsvorsorge.

Zur Erreichung dieser Zielsetzung werden die folgenden Themenkomplexe mit zugehörigen Leitfragen adressiert:

1. Konzeptioneller Rahmen und Abgrenzung des Themas: Was genau ist Big Data? Was genau versteht man unter öffentlichem Verkehr? Wie können die Begriffe abgegrenzt werden? Wie wird Big Data im Bereich des öffentlichen Personenverkehrs genutzt? Welche Unterschiede kann es im Bereich des urbanen Verkehrs gegenüber Fernreisen geben, die Einfluss auf die Nutzung von Big Data haben könnten?
2. Betrachtung der Angebotsseite: Welche Erkenntnisse liefert die bestehende Literatur zur Mobilitätsforschung hinsichtlich des Einsatzes von Big Data im öffentlichen Personenverkehr? Gibt es Unterschiede in den Einsatzmöglichkeiten zwischen urbanem Verkehr und Fernreisen? Welche Probleme bestehen derzeit beim Einsatz von Big Data insbesondere im Bereich öffentlicher Personenverkehr? Gibt es innovative Geschäftsmodelle, die sich auf Big Data im Bereich öffentliche Mobilität spezialisiert haben?
3. Betrachtung der Nutzerseite: Welche Erkenntnisse gibt es zur Akzeptanz von Big Data in der Verbraucherforschung? Welche spezifischen Probleme können im Bereich des öffentlichen Personenverkehrs für die Nutzer beim Einsatz von Big Data entstehen? Wie stehen Nutzer dem Einsatz von Big Data im Bereich öffentliche Mobilität gegenüber und welchen Einfluss auf ihr Nutzungsverhalten könnte der Einsatz von Big Data haben? Gibt es hierbei Unterschiede zwischen soziodemographischen Gruppen? Sind potentielle Rebound-Effekte zu erwarten? Gibt es aus Sicht der Nutzer Unterschiede zwischen der Nutzung des Nah- und Fernverkehrs?
4. Reflexion und Diskussion der Ergebnisse: Wie passen die Ansprüche von Anbietern öffentlicher Personenverkehrsangebote mit den Bedürfnissen der Nutzer an den Einsatz von Big Data in diesem Bereich zusammen? Welche möglichen Entwicklungen für den Einsatz von Big Data im Bereich öffentliche Mobilität können durch die Ergebnisse identifiziert werden? Welche Implikationen können diese Entwicklungen haben? Gibt es dabei Unterschiede zwischen Anbieter- und Nutzerseite? Sind die Entwicklungen wirtschaftlich, sozial und ökologisch sinnvoll und nachhaltig?
5. Ableitung von Handlungsempfehlungen: Welche Handlungsempfehlungen resultieren aus der Analyse des Einsatzes von Big Data durch Anbieter des öffentlichen Personenverkehrs unter Berücksichtigung der Interessen der Nutzer? Welche Entwicklungen sollte die Politik fördern und wie sollte sie das tun? Welche rechtlichen und regulatorischen Hemmschwellen gilt es hierbei zu überwinden?

1.3 METHODISCHES VORGEHEN

Um adäquate Erkenntnisse über den Einsatz von Big Data im öffentlichen Personenverkehr zu gewinnen, bedarf es sowohl einer Analyse des bestehenden Angebots an Big Data-basierten Technologien im ÖPV (Kapitel 2 des Gutachtens, geschrieben von InnoZ) als auch einer Untersuchung der Einstellung der Nutzer zu diesem Thema (Kapitel 3 des Gutachtens, geschrieben von ConPolicy). Von den

Projektpartnern wurden unterschiedliche methodische Vorgehensweisen gewählt, die im Folgenden erklärt werden.

1.3.1 ANGEBOTSSEITE

Für die Analyse der bestehenden Big Data-basierten Technologien auf Seiten von Mobilitätsanbietern wurde ein Überblick zu Technologien, Akteuren und Anwendungen erstellt. Die Zusammenstellung der Informationen basiert auf den Ergebnissen eines internen Workshops, einer umfangreichen Literaturrecherche, der Auswertung von Datenbanken und Expertengesprächen.

Um auf einem gemeinsamen Verständnis zentraler Begrifflichkeiten aufzubauen, wurden in einem initialen Workshop die wesentlichen Begriffe *Big Data* und *öffentlicher Personenverkehr* im Projektteam diskutiert. (Die Ergebnisse finden sich in den Kapiteln 1.4.1 und 1.4.2). Aufbauend auf diesem Workshop wurde der konzeptionelle Rahmen zur Analyse des Angebots an Big Data-basierten Technologien im ÖPV festgelegt. Die Bestandsaufnahme bestehender Anwendungen, Technologien und illustrierender Beispiele wurde mittels Literaturrecherche in einschlägigen Literaturdatenbanken und internen Wissensdatenbanken durchgeführt. Zudem wurden für die Recherche relevanter Akteure Produktdatenbanken und Konferenzprogramme ausgewertet. Im Mittelpunkt standen dabei Marktteilnehmer aus Deutschland, aber auch internationale Beispiele, die innovative Ansätze verdeutlichen, wurden berücksichtigt. Schließlich wurden stichprobenartig Stellenanzeigen in der Verkehrsbranche analysiert. Die Stellenanzeigen vermitteln dabei einen aktuellen Eindruck zur Relevanz von Big Data-Anwendungen in Unternehmen der Verkehrsbranche. Die Angaben zum Anforderungsprofil der Bewerber lassen Rückschlüsse darauf zu, welche Big Data-Technologien und IT-Lösungen derzeit im Fokus stehen und welche Anwendungsbereiche derzeit für Big Data vorgesehen sind.

1.3.2 NUTZERSEITE

Um die Kenntnisse, Wünsche und Befürchtungen der Nutzer beim Einsatz von Big Data im ÖPV zu ermitteln, wurde ein mehrstufiger Ansatz gewählt. Zunächst wurden bestehende Befragungen zum Nutzungsverhalten im ÖPV sowie zu ihrer Einstellung gegenüber dem Einsatz von Big Data im ÖPV recherchiert. Daraufhin wurde ein Fragebogen für eine repräsentative Befragung erstellt. Um alle potentiellen Wünsche und Befürchtungen der Nutzer abzufragen und um Versäumnissen vorzubeugen, wurden Fachgespräche mit drei Experten aus dem Bereich ÖPV geführt. Die Erkenntnisse dieser Gespräche flossen in den Fragebogen ein. Nachfolgend wurde der Fragebogen mit verschiedenen fachfremden Personen getestet und mit den Projektpartnern und Auftraggebern abgestimmt. Die Befragung wurde von einem professionellen Marktforschungsinstitut durchgeführt. Die erhobenen Daten wurden von ConPolicy ausgewertet und bilden die Grundlage für die Erkenntnisse in Kapitel 3.

1.4 DEFINITIONEN VON BIG DATA UND ÖFFENTLICHEM PERSONENVERKEHR

Ziel dieses Berichts ist es, einen Überblick zum Status quo der Anwendung von Big Data-Technologien im Bereich des öffentlichen Personenverkehrs zu geben und die Einstellung der Nutzer hierzu zu ermitteln. Deshalb ist es zunächst notwendig, die Verwendung der zentralen Begriffe *Big Data* und *öffentlicher Personenverkehr* für diesen Bericht zu spezifizieren.

1.4.1 BIG DATA

Aktuelle Diskurse in der ÖPV-Branche richten sich sehr stark auf die disruptiven Effekte von Digitalisierung auf Unternehmensstrategien, Geschäftsmodelle, Service und veränderte Wettbewerbsbedingungen aus. Verkehrsunternehmen ohne Fahrzeuge und Infrastrukturen sind heute nicht nur denkbar, sondern bereits ein Stück Realität geworden, betrachtet man zum Beispiel FlixBus oder Uber. Diese Unternehmen zeigen, dass mit der Digitalisierung neue Geschäftsmodelle möglich werden, die nicht mehr auf „Hardware" wie Fuhrparks angewiesen sind, sondern diese von Subunternehmern anmieten oder portionsweise einkaufen. Diese Veränderungen der Wirtschaft und Gesellschaft werden mit den Begriffen *Digitalisierung* oder *digitale Transformation* beschrieben, wobei ein einheitliches Verständnis, was unter Digitalisierung zu verstehen ist, wie es sich zu anderen Phänomen abgrenzt und in welchem Verhältnis der Begriff Big Data dazu steht, bislang nicht herausgebildet hat. Allgemein kann man den Begriff Digitalisierung mit Bezug zum Binärsystem des Computerzeitalters fassen: Dinge werden abzählbar und dadurch maschinenlesbar gemacht. Bisher nicht quantifizierbare Sachverhalte werden so zu Daten und können gespeichert, neukombiniert und ohne Zeitverlust an jeden Ort der Welt übertragen werden. Daraus können Wettbewerbsvorteile entstehen, wenn auf Basis vielfältiger Informationen und Wissensbestände schneller fundierte Entscheidungen getroffen werden können. Hier setzt der Begriff Big Data an. Je mehr Daten vorhanden sind – so die einfache Prämisse – desto zielgerichteter und belastbarer können Entscheidungen getroffen und Dienstleistungen gestaltet werden.

Die Miniaturisierung von Rechnern und Sensoren und die Entwicklung neuer Algorithmen, Methoden und Softwareprodukte sind Treiber der Digitalisierung und somit Treiber des Datenwachstums. Durch ein kontinuierliches Wachstum der Datenmengen wird ein neuer Umgang mit Daten notwendig. Big Data ist vor diesem Hintergrund zunächst ein Begriff, der bereits bekannte Phänomene in einer handlichen Terminologie zusammenfasst. In der Verkehrsbranche und der Verkehrsforschung sind es vor allem die Bereiche Verkehrstelematik (Evers & Kasties 1999) und Intelligent Transport Systems (ITS) (Wootton, García-Ortiz, & Amin 1995), die Möglichkeiten zur intelligenten, datenbasierten Verkehrssteuerung bereits seit mehreren Jahrzehnten erarbeiten. In vielen Städten sind Verkehrsleit- und Verkehrsmanagementzentralen wichtige Einrichtungen zur Steuerung des Verkehrs und zur Sicherstellung des Verkehrsflusses auf Straßen und Schienen (Sandrock & Riegelhuth 2014). Durch die technologische Entwicklung und stetig steigende Rechenleistung bei sinkenden Kosten sollen die bestehenden Konzepte und Technologien der Verkehrstelematik und ITS weiterentwickelt werden und ganz neue Anwendungen erlauben.

Big Data als Beschreibung von Dateneigenschaften: Volume, Velocity, Variety und mehr

Wichtigste Neuerung und zentrale Eigenschaft von Big Data ist das zunehmend steigende Volumen (in Echtzeit verfügbarer) Daten. Big Data ist all das, was nicht von der Tabellenkalkulation Excel verarbeitet werden kann[5], meint Batty (2013, S.274) in einer sehr schlichten Begriffserklärung und greift damit diese zentrale Eigenschaft von Big Data auf.

Historisch lässt sich die Verwendung des Begriffs Big Data bis in die 1990er-Jahre zurückverfolgen (Diebold 2012). Bereits zu diesem Zeitpunkt stand der Bezug auf die Verarbeitung und Analyse großer

5 „(A)ny data that cannot fit into an Excel spreadsheet" (Batty 2013, S.274).

Datensätze im Mittelpunkt (Volume). Seit Anfang der 2000er-Jahre nahm die Anzahl der Studien und Projekte zum Themenfeld Big Data zu und die Definitionen wurden stetig erweitert (Kitchin & McArdle 2016). Im Vergleich von Definitionen aus unterschiedlichen Disziplinen kristallisieren sich zwei weitere zentrale Eigenschaften heraus: die Geschwindigkeit der Datenbereitstellung (Velocity) und die Vielfalt der Dateninhalte (Variety). Die zahlreichen weiteren Charakteristika, die in unterschiedlichen Publikationen genannt werden sind nur bedingt für eine eindeutige Beschreibung von Big Data geeignet. Tab. 1 zeigt eine Übersicht einiger weiterer häufig genannter Eigenschaften.[6] Diese deskriptiven Beschreibungen sind nur bedingt geeignet, um einen Datensatz eindeutig als Big Data bezeichnen zu können. Sie vermitteln aber vor allem einen Einblick darin, mit welchem Mindset Big Data verbunden ist und woran sich Kritik an Big Data orientiert.

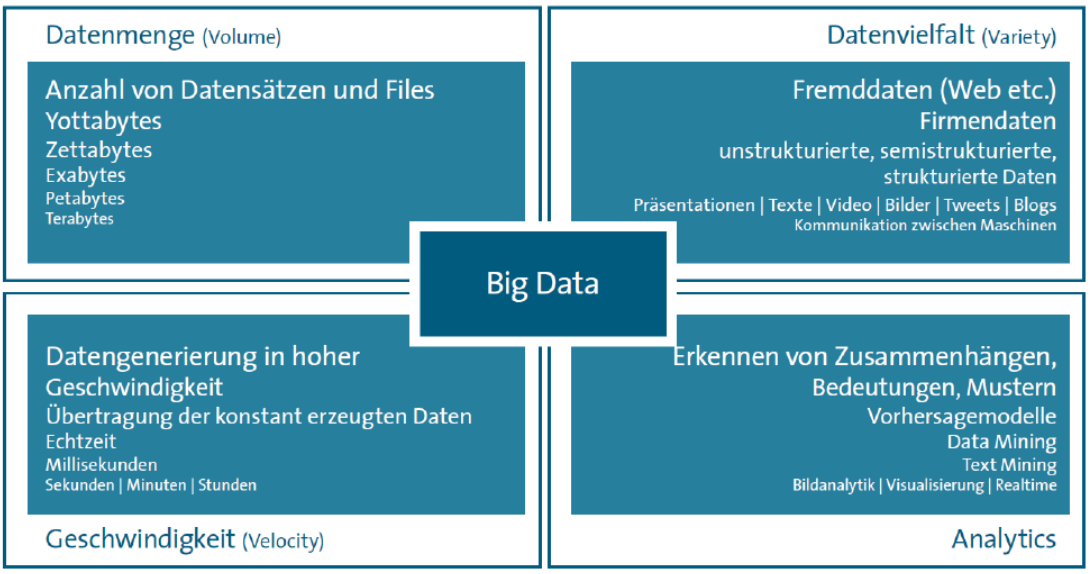
Tab. 1: Big Data-Charakteristika

Eigenschaft	Beschreibung	Quelle
Vollständigkeit (Exhaustivity)	Die Daten entsprechen (nahezu) einer Vollerhebung. Eine möglichst vollständige Erfassung aller Untersuchungsobjekte wird angestrebt, nicht das Sampling optimiert.	Mayr-Schonberger, Cukier (2013)
Hoher Detailgrad (fine-grained) und eindeutig indizierbar (uniquely indexical)	Die Daten sind feinkörnig aufgelöst und können durch eindeutige Indizierung identifiziert werden.	Dodge und Kitchin (2005)
Relationalität (Relationality)	Verschiedene Datensätze können über gemeinsame Variablen verknüpft werden.	Boyd und Crawford (2012)
Erweiterbarkeit (Extensionality) und Skalierbarkeit (Scaleability)	Es können Variablen auf einfache Weise hinzugefügt und die Datenbankgröße unkompliziert erweitert werden.	Marz and Warren (2012)
Sinnhaftigkeit (Veracity)	Die Sinnhaftigkeit und Vertrauenswürdigkeit der Daten sollte jeweils hinterfragt werden. Die Daten können ungeordnet und verrauscht sein sowie Unsicherheiten und Fehler enthalten.	Demchenko et al. (2013)
Werthaltigkeit (Value)	Aus den Daten können sehr viele Rückschlüsse gezogen und die Daten in anderen Kontexten angewandt werden.	Demchenko et al. (2013)
Variabilität (Variability)	Die Daten können ganz unterschiedliche Bedeutungen haben und Aussagen produzieren je nachdem in welchem Kontext sie genutzt werden.	Gandomi und Haider (2015)

6 Darüber hinaus nennt Uprichard (2013) zahlreiche weitere Beschreibungen der Eigenschaften von Big Data: Vielseitigkeit (versatility), Volatilität (volatility), Virtuosität (virtuosity), Vitalität (vitality), visionär (visionary), schlagkräftig (vigour), Brauchbarkeit (viability), Resonanz (vibrancy), Potenz (virility), Unergiebig (valueless), vampirartig (vampire-like), bösartig (venomous), vulgär (vulgar), verletztend (violating and very violent). Ähnlich gelagert ist der Vorschlag von Lupton, die V-Wörter durch P-Wörter zu ersetzen. Lupton nennt folgende Adjektive zur Beschreibung von Big Data: portentous, perverse, personal, productive, partial, practices, predictive, political, provocative, privacy, polyvalent, polymorphous and playful.

Die existierenden Definitionen für Big Data sind vielfältig, überwiegend beschreibend und relational sowie wenig konkret. Die im Rahmen dieser Studie verwendete Definition orientiert sich an den drei genannten zentralen Eigenschaften *Datenvolumen (Volume)*, *Geschwindigkeit der Datenbereitstellung und -verarbeitung (Velocity)* und *Vielfalt der Datenquellen (Variety)*. Zusätzlich sollten jeweils Aussagen zur Datenqualität möglich sein. Diese Definition entspricht auch der Verwendung des Begriffs durch den Branchenverband Bitkom (2015) (vgl. Abb. 1). Die Definition ist geeignet, das Thema einzugrenzen, ist aber gleichzeitig nicht zu eng gefasst, so dass eine gute Beschreibung des Status quo im Verkehrssektor möglich ist.

Abb. 1: Eigenschaften von Big Data



Quelle: BITKOM (2015, S.13)

Big Data als Bündel von Methoden und Dienstleistungen: Full stack data science solution

Big Data ist aber nicht allein durch die Beschreibung der Eigenschaften von Daten gekennzeichnet. Eine besondere Qualität der Ergebnisse wird dadurch erzielt, dass bisher nicht aufeinander bezogene Daten miteinander kombiniert werden können. Ermöglicht wird dies durch ein Bündel an Methoden und Technologien für die Erfassung, Aufbereitung, Speicherung, Analyse und Interpretation strukturierter und unstrukturierter Daten. Neben der beschreibenden Ebene ist es diese methodische Ebene, die Big Data-Anwendungen auszeichnet (vgl. Abschnitt 3).

Big Data als Produktionsfaktor: Strategisches und kompetitives Asset

Neben der beschreibenden und methodisch-technologischen Ebene wird Big Data schließlich als Produktionsfaktor angesehen. Damit wird die wirtschaftliche Bedeutung betont, die Big Data zugeschrieben wird. Diese wirtschaftliche Bedeutung wird als so groß angesehen, dass Big Data neben Arbeitskraft, Ressourcen und Kapital als vierter Produktionsfaktor beschrieben wird. Daten werden in dieser Sichtweise zum strategischen und kompetitiven Asset, das strategische Entscheidungen auf breiter Wissensgrundlage ermöglicht.

In Unternehmen wird allgemein davon ausgegangen, dass sich aus der Nutzung und Analyse von Big Data große wirtschaftliche Potentiale ergeben, wenn es gelingt, diese in Unternehmensprozesse zu integrieren (Hölzle, Schoder, Spiri, & Götz 2017). Aus diesem Grund wird der Begriff immer stärker auch vom Unternehmensmarketing aufgegriffen.

Big Data als Metapher in öffentlichen Diskursen: Chancen und Risiken der Digitalisierung

Schließlich dient der Begriff Big Data in öffentlichen Debatten als eine Metapher für allgemeine Chancen und Risiken der Digitalisierung. Big Data ist dabei ein Begriff, der polarisiert, da er auf der einen Seite Goldgräberstimmung und technokratische Allmachtphantasien hervorruft und auf der anderen Seite Proteststürme und Ängste schürt.

Die genannten Aspekte werden in der Betrachtung von Beispielen für Big Data-Anwendungen in Kapitel 2 aufgegriffen. Dabei wird jeweils darauf eingegangen, welchen Grad die Geschwindigkeit der Bereitstellung erreicht, welches Volumen die verwendeten Daten aufweisen und wie viele unterschiedliche Datenquellen genutzt werden. Zudem wird auf verwendete Methoden und Technologien eingegangen und die wirtschaftliche Bedeutung bewertet.

1.4.2 ÖFFENTLICHER PERSONENVERKEHR

Die digitale Durchdringung des Mobilitätsmarktes macht den Pulsschlag der mobilen Gesellschaft messbar. Dadurch können Kundenbeziehungen gestärkt, Mehrwertdienste entwickelt, die Auslastung von Fahrzeugen erhöht und der Transport von Gütern und Personen effizienter gestaltet werden. Es entsteht die Chance, Mobilitätsdienstleistungen einfacher, transparenter und besser zugänglich zu machen oder sogar völlig neuartige Angebote zu entwickeln. Dadurch wird sich aber auch der öffentliche Personenverkehr in Zukunft sehr stark verändern (Gerdes, Schelewsky, Knie, Schönduwe, & Jahn 2015). Aus diesem Grund ist es notwendig, die traditionelle Definition des Begriffs *Öffentlicher Personenverkehr (ÖPV)* zu erweitern.

Traditionelle Definition des Begriffs Öffentlicher Personenverkehr

Eine in der Verkehrsbranche und -forschung gebräuchliche Definition des Begriffs *Öffentlicher Personenverkehr* (manchmal auch nur *Öffentlicher Verkehr*) orientiert sich an der gesetzlichen Regelung im Personenbeförderungsgesetz (PBefG) (vgl. Heinze, Fehling, & Fiedler 2014). Dabei wird zunächst zwischen *öffentlichem Personenfernverkehr (ÖPFV)* und *öffentlichem Personennahverkehr (ÖPNV)* unterschieden. ÖPNV ist die allgemein zugängliche Personenbeförderung mit Bussen, Straßenbahnen, Eisenbahnen, Taxen, S- und U-Bahnen sowie Wasserfahrzeugen. Um als ÖPNV zu gelten, darf die Mehrzahl der Beförderungsfälle eine Reiseweite von 50 km bzw. eine Reisezeit von einer Stunde nicht übersteigen. Im PBefG sind fünf Pflichten für Verkehrsleistungen im ÖPNV definiert:

- Linienpflicht (es muss eine bestimmte, vorher festgelegte und kommunizierte Linie bedient werden)
- Betriebspflicht (es muss auch dann gefahren werden, wenn keine Nachfrage existiert)
- Fahrplanpflicht (in einem Fahrplan muss ersichtlich sein, zu welchem Zeitpunkt gefahren wird)
- Beförderungspflicht (alle Personen, die den Tarif entrichten, müssen befördert werden)
- Tarifpflicht (es muss ein Tarif für die Beförderungsdienstleistung festgelegt sein)

Ausnahmen von diesen Pflichten können über die Sonderformen des Gelegenheits-ÖPNV geregelt werden, im sog. Bedarfsverkehr. Damit können topographische und soziostrukturelle Ausprägungen des ländlichen Raums bei der Ausgestaltung des ÖPNV berücksichtigt werden, indem z.B. auf die Linienbindung verzichtet wird (Umsetzung eines Bedarfsverkehrs im Flächen- oder Richtungsbandbetrieb) oder Haltestellen nur bei Bedarf (telefonischer Voranmeldung) bedient werden.

Erweiterte Definition des Begriffs Öffentlicher Personenverkehr

Für die Analyse des bestehenden Angebots in Kapitel 2 wird eine erweiterte Definition des *öffentlichen Personenverkehrs* vorgeschlagen und genutzt. Die Digitalisierung hat eine Entwicklung in Gang gebracht, die zu einer Auflösung der traditionellen Trennung von öffentlichem und privatem Verkehr führen wird. Verantwortlich dafür ist die zunehmende Verfügbarkeit neuer Mobilitätsdienstleistungen. Die Entwicklung dieser Dienstleistungen wird primär durch die Verfügbarkeit von Daten und die Verbesserung der Zugänglichkeit über datenbasierte Anwendungen ermöglicht. Sharing-Dienstleistungen und Rideselling-Angebote zeigen bereits heute, wie durch die geteilte Nutzung von Verkehrsmitteln ein halb-öffentlicher Verkehr entstehen kann. In einem Definitionspapier der Forschungsgesellschaft für Straßen- und Verkehrswesen (FGSV) wird auf diese Entwicklung eingegangen. Es werden Verkehrsmittel und Mobilitätsdienstleistungen nach der Verkehrsart und dem Zugang unterschieden (vgl. Tab. 2).

Tab. 2: Abgrenzung öffentlicher und privater Verkehr

			Zugang			
			Privater Zugang	Öffentlicher Zugang		
Verkehrsart	*Modus*	*Erweiterter Modus*	individuelle Nutzung	individuelle Nutzung nacheinander	kollektive Nutzung während der Fahrt	
Nicht motorisierter Verkehr	Fußverkehr	Fußverkehr	Fußverkehr			
	Radverkehr	Radverkehr (nicht geteilt)	Fahrrad, Pedelec	Rikscha, Velotaxi		
		Geteilte Verkehrsmittel		Bikesharing (stationsunabhängig und stationsbasiert), Mietfahrrad		
Motorisierter Verkehr	MIV		peer-to-peer Autoteilen, private Fahrgemeinschaft	Carsharing (stationsunabhängig und stationsbasiert), ggf. auch Motorrad-, Moped-, Mofa-, eBike-, Mietwagen	Fahrgemeinschaft, Mitfahrgelegenheiten = Ridesharing (Fahrer bestimmt den Fahrtablauf; i.d.R. privat), Rideselling (Fahrgast bestimmt den Fahrtablauf; i.d.R. gewerblich)	
		MIV (nicht geteilt)	Pkw, Motorrad, Moped, Mofa, eBike (jeweils Fahrer oder Mitfahrer)			
	ÖV	ÖPNV		Taxi	Linienbus, Bürgerbus, Stadt- u. Straßenbahn inkl. Bahnen besonderer Bauart, U-Bahn, S-Bahn Nahverkehrszug, flexible Angebotsformen des ÖV, Fähre, Sammeltaxi, Reisebus (Stadtrundfahrt)	
		ÖPFV			Fernzug, Fernbus, Flugzeug, Schiff, Reisebus (Ausflugsfahrt, Urlaub)	

Quelle: FGSV (2017)

Ist der Zugang öffentlich, so werden diese Angebote und Mobilitätsdienstleistungen für die Analyse des bestehenden Angebots an Big Data-basierten Anwendungen in Kapitel 2 dieses Berichts berücksichtigt. *Klassischer* oder *traditioneller* ÖPV (bzw. ÖPNV oder ÖPFV) meint in diesem Kapitel den ÖPV im Sinne des PBefG in Abgrenzung zu den neuen, geteilten Angeboten.

Bei der Analyse der Nutzerseite in Kapitel 3 wird mit ÖPV (bzw. ÖPNV oder ÖPFV) hingegen ausschließlich der traditionelle öffentliche Personennah- und Fernverkehr gemeint, da Sharing-Angebote in der Befragung gesondert adressiert wurden.

2. BETRACHTUNG DER ANGEBOTSSEITE

Die Digitalisierung wird viele Bereiche des ÖPNV verändern, seien es innerbetriebliche Prozesse wie Wartung, Personaleinsatz oder auch der direkte Kundenkontakt über die Bereiche Angebotsgestaltung, Kommunikation, Marketing und Vertrieb. Die Nutzung von passiv oder aktiv generierten Daten kann dabei als Ausgangspunkt dieser Entwicklung betrachtet werden. Die zentrale Herausforderung besteht somit für die Unternehmen des ÖPV im Umgang mit anfallenden großen Datenmengen. Die Kombination von sinkenden Kosten in Erhebung, Verarbeitung und Speicherung sowie die allgegenwärtige Erfassung über Sensoren und Medien – häufig über private Endgeräte – ermöglicht die digitale Abbildung realer Prozesse, Kommunikationen, Betriebsabläufe und des Nutzerverhaltens. Die Auswirkungen auf die bestehenden Angebote im Verkehrssektor, die mit der Digitalisierung einhergehen (werden), werden dabei oftmals mit den Begriffen *radikal* oder *disruptiv* beschrieben. Dabei bleibt allerdings unklar, ob diese Beschreibungen eher Ausdruck großer Verunsicherung einer ungewissen Zukunft sind oder die tatsächlichen zu erwartenden Veränderungen beschreiben.

Darum ist ein Blick auf die sich andeutenden und prognostizierten Veränderungen angebracht, insbesondere in einem Bereich mit traditionell eher langsameren Innovationszyklen, wie dem ÖPV, um sich der Frage zu nähern, welche Auswirkungen der Einzug der Digitalisierung in die Welt der Mobilität tatsächlich mit sich bringen wird.

Die Ausgangssituation scheint klar: Für immer mehr Prozesse können kontinuierlich Daten bereitgestellt werden. Aufbauend auf diesen Daten können bestehende Angebote verbessert, neue Dienstleistungen entwickelt und unternehmensinterne Prozesse optimiert werden. In diesem Teil des Gutachtens wird anhand von Beispielen aufgezeigt, welche Big Data-Anwendungen bereits heute im Bereich des ÖPV existieren, Eingang in die Praxis gefunden haben und (End-)Nutzern zur Verfügung stehen. Dazu ist es zunächst notwendig, auf wesentliche Technologien und Methoden in der Verarbeitung von Big Data einzugehen und Eigenschaften von Mobilitätsdaten sowie Datenquellen zu nennen (Abschnitt 2.1). Welche datenbasierten Anwendungen bereits heute im ÖPV eingesetzt werden, wird in Abschnitt 2.2 dargestellt. Dabei wird nach anbieterseitigen Anwendungen (Abschnitt 2.2.1), kundenorientierten Anwendungen (Abschnitt 2.2.2) und Verwertungen der Daten durch Dritte (Abschnitt 2.2.3) unterschieden.

2.1 EIGENSCHAFTEN UND TECHNOLOGIEN

Fünf Entwicklungen aus den Bereichen Datenerhebung, -verarbeitung und -speicherung haben Big Data-Technologien erst ermöglicht:

- Allgegenwärtige Möglichkeiten der Datenerfassung
- Datenanalysen in Echtzeit
- Neuartige Analyse-Frameworks
- Verbesserung der Datenspeicherung sowie
- Cloud-Computing

Die zunehmende Leistungsfähigkeit und abnehmenden Kosten beschleunigten die Nutzung von Big Data-Technologien. Die mit den fünf genannten Entwicklungen korrespondierenden Technologien sind in aktuellen Big Data-Anwendungen bereits oft stark verzahnt. Aus Sicht des Mobilitätssektors gibt es hier Technologien, die sich eher als infrastruktureller Rahmen verstehen lassen und somit

stärker im Sinne eines Technologietransfers stehen, als auch Technologien die näher an den angestrebten Dienstleistungen angelehnt sind und somit auch stärker technologische Anstöße aus dem hier betrachteten Sektor erhalten. Diese betreffen vor allem Technologien zur Datenerfassung und zu Analysetechnologien wie Machine Learning oder Predictive Modelling. Zu den infrastrukturellen Rahmenbedingungen sind bspw. Datenbanktechnologien wie MapReduce, NoSQL und graphenorientierte Strukturen oder externe Dienstleistungen für Speicher- und Rechendienste wie bei Amazon Web Services (AWS), Google Cloud Plattform oder Microsoft Azure zu nennen.

2.1.1 EIGENSCHAFTEN

Die Verfügbarkeit von Daten nimmt auch im Verkehrsbereich exponentiell zu, zum einen durch den immer stärkeren Einsatz von Sensoren und zum anderen durch die zunehmende Prozessprotokollierung durch Anwendungs- und Prozesssoftware. Stark sinkende Sensorkosten bei immer kleineren Sensoren mit stetig verbesserter Sensorleistung erzeugen Datenmengen in bis vor wenigen Jahren noch unvorstellbaren Größenordnungen. In diesem Umfeld des Internet of Things (IoT) (Borgia 2014) oder gar Internet of Everything (Bradley, Barbier & Handler 2013) entsteht eine tägliche, weltweite Datenproduktion von mehreren Petabytes. Die weiterhin zunehmende Datenerfassung und -protokollierung wird in diesem Bericht für spezifische verkehrs- und mobilitätsbezogene Anwendungen skizziert. Die Verarbeitung und Speicherung kann beispielhaft in den Anwendungsbeispielen aus Kapitel 2.2 entnommen werden. Beispiele für Datenvolumen im Verkehrsbereich zeigt Abb. 2.

Abb. 2: Beispiele für Datenvolumen im Verkehrsbereich

	Beispiele
1 Byte (10^0)	
1 Kilobyte ~ 1.000 Byte (10^3)	
1 Megabyte ~ 1.000.000 Byte (10^6)	
1 Gigabyte ~ 1.000.000.000 Byte (10^9)	25 Gigabytes: Pro Stunde in einem Ford Fusion Energi Plug-in hybrid verarbeitete Daten
1 Terabyte ~ 1.000.000.000.000 Byte (10^{12})	60 Gigabytes: Pro Stunde in Googles Autonomen Shuttle verarbeitete Daten
1 Petabyte ~ 1.000.000.000.000.000 Byte (10^{15})	140 Gigabytes: Datenmenge produziert an einem Tag in einem Auto, das von Nokia HERE zum Kartieren eingesetzt wird
1 Exabyte ~ 1.000.000.000.000.000.000 Byte (10^{18})	30 Terabytes: Auf einem Transatlantikflug von einer Boeing 777 produzierte Daten
1 Zetabyte ~ 1.000.000.000.000.000.000.000 Byte (10^{21})	Einige Petabytes: Insgesamt von INRIX gespeicherte Datenmenge zum Verkehrsfluss (z.B. von Google Maps genutzt)
1 Yotabyte ~ 1.000.000.000.000.000.000.000.000 Byte (10^{24})	1 Zettabyte: Gesamtmenge der Daten aller visuellen Informationen aller Menschen weltweit an einem Tag im Jahr 2013
	4,4 Zettabytes: Geschätzte Datenmenge im „digitalen Universum“ im Jahr 2013
	44 Zettabytes: Geschätzte Datenmenge im „digitalen Universum“ im Jahr 2020

Quelle: (ITF/OECD 2015)

2.1.2 BIG DATA-TECHNOLOGIE STACK

Big Data-Methoden und -Technologien bilden ein Bündel (*stack*) an Tools und Infrastrukturen, die die gesamte Bandbreite der Datenverarbeitung abdecken. Ein Big Data-Technologie Stack unterstützt unterschiedliche Aspekte des Umgangs mit großen Datenmengen. In einem Satz zusammengefasst liegt die zentrale Aufgabe darin, Technologien und Methoden bereitzustellen, die es erlauben, Daten zu erheben, zu verteilen, zu speichern, zu transformieren, zu analysieren und zu visualisieren. Dieser

gesamte Prozess sollte möglichst in Echtzeit erfolgen können. Für alle Prozessschritte gilt, dass diese ausreichend dokumentiert sein sollten.

Die Auswertung großer Datenmengen erfordert spezielle Technologien und Algorithmen. Tab. 3 gibt einen Überblick zu den wichtigsten Elementen eines Big Data-Technologie Stacks und zeigt einige Beispiele für entsprechende Produkte und Anwendungen. Neuartige Frameworks, also Programmiergerüste, die eine kohärente Struktur für die Entwicklung von Anwendungen liefern, ermöglichen erst die effektive Bearbeitung großer Datenmengen. Die oftmals als Open-Source veröffentlichten Frameworks brechen dabei häufig aufwändige Rechenvorgänge für unstrukturierte, komplexe Datensätze in mehrere, besser koordinierbare Recheneinheiten auf, die in parallelen Prozessen auf dezentrale Serverstrukturen deutlich schneller bearbeitet werden können (MapReduce). Zudem sind diese Prozesse skalierbar, da sie ein Stück weit entkoppelt sind von der Bereitstellung von Speicherkapazität.

Tab. 3 Big Data-Technologie-Stack

Stack element	**Nutzung**	**Beispiele**	
		Open Source	**Kommerziell**
Datenerhebung und primäre Datenverarbeitung	Echtzeit-Datenverarbeitung, Datentransformation	Hadoop (und Erweiterungen), Anwendungen auf Basis von MapReduce,	Splunk, Oracle, kommerzielle Hadoop-Forks
Datenspeicherung und Datenmanagement	Speichern von Daten, Metadaten, Rohdaten, Indexbildung, Normalisierung	Neo4j, MongoDB, MySQL	Oracle DB
Analyse	Predictive tools, Descriptive tools, Machine learning, Statistische Methoden	R, Python, Pandas	SPSS, SAS, Mathlab
Visualisierung	Benutzerschnittstelle, Dashboard, Webbasierte Tools	D3.js, Gephi	Tableau, Mapbox
Infrastruktur	Server, Sensoren, Computer, Netzwerk, Scheduling, Monitoring, Betriebssystem	Linux, OpenStack, OpenShift, Puppet, Zookeeper, Oozie, HDFS, Kafka, JBoss, Xymon	AWS, Azur, Cloudera

Datenerhebung und primäre Datenverarbeitung

An die Datenerhebung und primäre Datenverarbeitung[7] werden mittlerweile sehr hohe Ansprüche gestellt. Bis vor wenigen Jahren wurden in der Verkehrsbranche noch hauptsächlich statische Datensätze verarbeitet, die in größeren zeitlichen Abständen aktualisiert wurden. Es etablierten sich immer kürzere Updatezyklen dieser statischen Datensätze. Später kamen sogenannte *Live-Feeds* hinzu, die im weiteren Verlauf auch allmählich Echtzeitdaten umfassten. Um diese Echtzeitdaten effizient erheben und verarbeiten zu können, sind neue Methoden und Technologien notwendig. Im Prozessschritt der primären Datenverarbeitung werden auch eine automatisierte Extraktion,

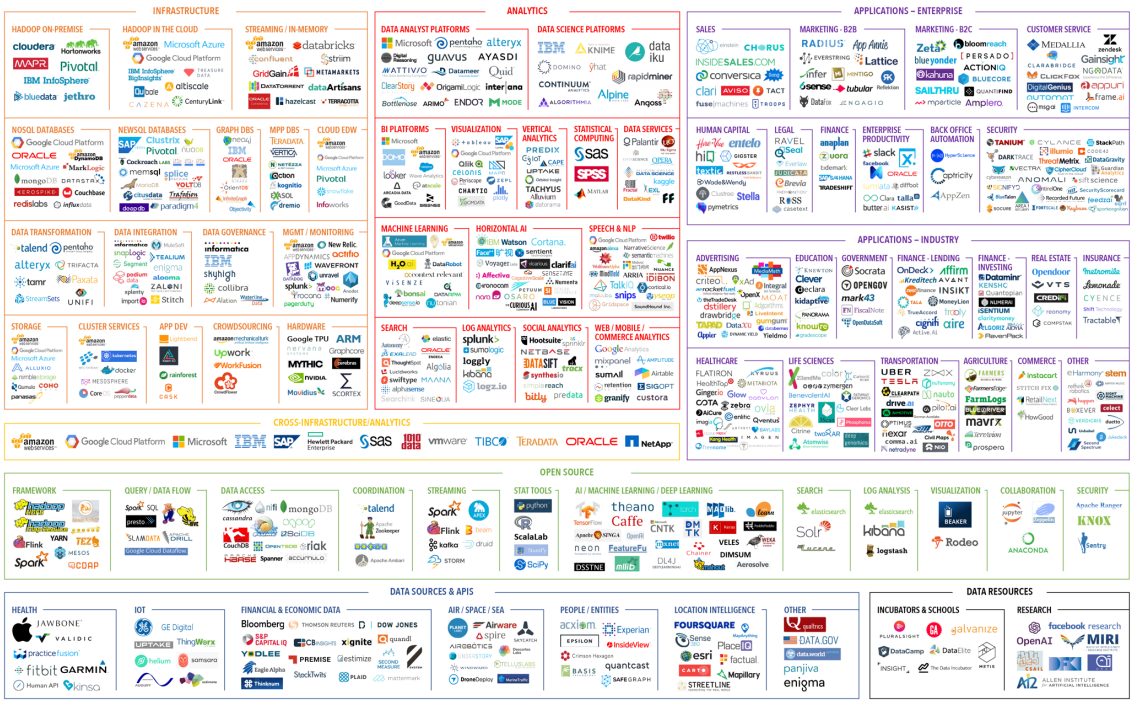
7 Die Arbeitsschritte der Datenerhebung und primären Datenverarbeitung werden auch als „data ingestion" bezeichnet.

Integritätsprüfung, Säuberung und erste Aggregationen durchgeführt. Ein Anwendungsbeispiel in diesem Bereich ist Hadoop. Mittels Hadoop können große Datenmengen schnell, dezentral und parallel bearbeitet werden.

Datenspeicherung und Datenmanagement

Zur Datenspeicherung und dem Management der erhobenen Daten muss eine geeignete Datenbank ausgewählt werden. Dabei ist es nicht nur eine Herausforderung, die Datenvolumen zu verwalten, sondern auch die in hoher Geschwindigkeit anfallenden Daten verarbeiten zu können. Aufgrund sinkender Kosten zur Erfassung und Speicherung großer Datenmengen kann neben den kontinuierlich – quasi in Echtzeit – anfallenden Daten eine immer größere Menge historischer Daten bereitgehalten werden. Dabei werden allein aufgrund der Verfügbarkeit große Datenmengen gespeichert, die zum Zeitpunkt ihrer Erhebung noch keinen spezifischen Verwendungszweck besitzen, und die quasi auf Vorrat für potentielle zukünftige Anwendungen aufgehoben werden. Diese werden dabei dezentral und somit abseits vom eigentlichen Erzeugungsort gespeichert. Cloud-Computing ist hierbei die relevante Technologie von Netzwerkprotokollen, wie auch die Bezeichnung der dazugehörigen Dienstleistungen. Beispiele hierfür sind Amazon Web Services, Google Cloud Platform, Microsoft Azure und andere (vgl. Abb. 3). Wichtiger vorbereitender Schritt für die Datenanalyse ist auch die Fusion unterschiedlicher Datenquellen.

Abb. 3: Anbieter und Technologien im Kontext von Big Data



Quelle: Turck (2017)

Datenanalyse und Visualisierung

Es schließen sich die Arbeitsschritte der Analyse und Visualisierung an, für die wiederum sehr unterschiedliche Optionen zur Verfügung stehen. Die Visualisierung nimmt im Prozess der Analyse und Interpretation der Daten eine zunehmend wichtige Rolle ein. Insbesondere bei großen Datenmengen können geeignete Visualisierungen dazu beitragen, komplexe Sachverhalte so darzustellen, dass

Zusammenhänge schneller erkannt werden können. Zudem sind Visualisierungen sehr wichtig für die Kommunikation der Ergebnisse (Stabellini, Remondino, & Tamborrini 2017).

Die Entwicklung der Datenanalyse in der Verkehrsbranche zeigt Parallelen zur Datenverfügbarkeit auf. Mit dem Beginn der Big Data-Entwicklung Mitte der 2000er-Jahre finden sich die ersten Automatisierungsprozesse für die datengestützte Aufbereitung und Analyse von Unternehmenskennzahlen und Nutzerstatistiken (Antoniou 2016). Aktuell werden diese Tools zum Monitoring um weitere Module zur kurz- und mittelfristigen Prognose erweitert. Die Prognose wird durch die Datenhistorie ermöglicht.

Die Entwicklung geht dorthin, dass diese Prognosen zur automatisierten Prozessoptimierung eingesetzt werden. Die Interpretation der Ergebnisse sollte letztlich zur Entscheidungsfindung beitragen. Um die Ergebnisse interpretieren zu können, ist entsprechendes Fachwissen notwendig.

Big Data-Tools

Die Angebotslandschaft im Bereich Big Data ist sehr vielfältig. Eine aktuelle Übersicht zu Infrastrukturen, Analysetools, Datenbanken, Big Data-Anwendungen und Datenquellen zeigt Abb. 3. Je nach Anwendung, Datenquelle, Kompetenzen der Mitarbeiter und Fragestellung können ganz unterschiedliche Big Data-Technologie-Stacks zusammengestellt werden. Allgemeine Fragestellungen, an denen sich die Auswahl orientiert, lauten bspw.:

- Welcher Datentyp soll verarbeitet werden? Strukturierte, unstrukturierte, relationale oder graphen-basierte Daten?
- Welche Größe werden die Datensätze haben? Große Datensätze (Videos, hochauflösende Fotos) oder kleine Datensätze (Text)?
- Wie gelangen die Daten in die Datenbank? Streaming, Batch-File?
- Welche Kosten fallen für die Datenbank an? Lizenzen und Kosten für die Einrichtung der Datenbank?
- Welche Hardware wird benötigt?
- Wie skalierbar ist die Datenbank? Gigabytes, Terabytes, Petabytes, Exabytes, Yottabytes?
- Kann die Datenbank für Analysen genutzt werden?
- Welche Latenz wird für Abfragen und Analysen benötigt? Millisekunden? Tage?
- Ist das System für bestimmte Features optimiert? Schnelles Schreiben? Schnelles Lesen? Nutzerfreundlichkeit?
- Wie viele Nutzer sollen gleichzeitig auf das System zugreifen? Die Anzahl der Nutzer erhöht die Datenmenge enorm.
- Ist die Sicherheit der genutzten Datenbank gewährleistet? Auf welchem Level?
- Für welche Zeiträume sollen die Daten gespeichert werden?

2.1.3 KATEGORISIERUNG VON MOBILITÄTSDATEN

Im Bereich der Mobilität fallen Daten an verschiedenen Stellen und in unterschiedlichen Kontexten an, die für Big Data-Anwendungen im ÖPV genutzt werden können. Ein grundlegender Trend der Digitalisierung ist die Allgegenwärtigkeit von Elementen, die primär Daten produzieren oder bei sekundärer Datenverarbeitung weitere Prozessdaten generieren. Dieser Trend bewirkt auch, dass in diesem Abschnitt nur die Datenkategorien benannt werden können, die aktuell am relevantesten erscheinen und keine umfassende bzw. abschließende Auflistung möglich ist. Auch die Bedeutsamkeit

der verschiedenen Quellen unterliegt dem Wandel der Dienstleistungen und Technologien. Entsprechend ist in diesem Abschnitt nur eine Momentaufnahme für den hier betrachteten Sektor möglich.

Es existieren zahlreiche Versuche, Kategorisierungen von Mobilitätsdaten vorzunehmen. Häufig werden Mobilitätsdaten nach Anwendungskontext und Erhebungsmethodik kategorisiert (z.B. Calvert 2014). Hierbei wird vor allem zwischen konventionellen (z.B. Haushaltsbefragung) und neuartigen Erhebungsformen (z.B. automatisierter Sensor) unterschieden. Bei den Anwendungskontexten werden bspw. die betriebliche Nutzung, die Notfallreaktion, kundenorientierte Informationsdienste sowie langfristige Infrastrukturplanung unterschieden. Im Folgenden werden drei Dimensionen von Mobilitätsdaten betrachtet, die im Kontext von Big Data-Anwendungen besonders relevant sind.

Räumliche Dimension: ortsbasierte und ortsungebundene Daten

Daten können an einem festen Ort oder ortsungebunden erfasst werden. Zur ortsgebundenen (fixed) Erhebung zählen automatische Verkehrszählstellen und manuelle Verkehrszählungen (Fahrgastzählungen, Straßenprofilzählungen), automatische Fahrgastzählsysteme an Haltestellen oder ÖPV-Zugangsdaten über tragbare Zugangsmedien (Karten, Radio-Frequency Identification (RFID), Near-Field Communication (NFC)) und Parkdaten der digitalisierten Parkraumbewirtschaftung. Diese Datenquellen eignen sich, um z.B. aggregierte Informationen über Verkehrsflüsse zu erheben. Zu ortsungebundenen (Probe-) Bewegungsdaten zählen Floating-Car- und Floating-Phone-Daten sowie teilweise Daten aus den Betriebsabläufen des öffentlichen Verkehrs. Hinzu kommen passiv generierte Daten aus automatischen Fahrgastzählsystemen, Ticketverkäufen, Buchungsdaten und LogFiles der zugehörigen Applikationen. Die fortlaufende Aus- und Aufwertung von Passivdaten wird zukünftig an Bedeutung zunehmen (Schmidt & Mänel 2017). Hierfür sind Big Data-Technologien besonders relevant, da diese Daten häufig unstrukturiert und heterogen sind und sich hier ein hohes Potential für die Verschneidung unterschiedlicher Datenquellen ergeben.

Eine weitere ergiebige ortsungebundene Datenquelle sind inzwischen auch Sentimentdaten, die über soziale Medien erfasst werden oder individuelle Bewegungsdaten, die bspw. über Smartphone-Tracking aufgezeichnet werden. Daneben werden umfangreiche, aber klassisch erhobene Datensätze verwendet wie z.B. Haushalts- oder Fahrgastbefragungen sowie ÖPV-Fahrplandaten (z.B. im Format des HaCon Fahrplan-Auskunfts-Systems (HAFAS) oder der General Transit Feed Specification (GTFS), Tarifinformationen (z.B. Rhein-Main-Verkehrsbund (RMV)-Zonendatenbank) und Ticketverkaufszahlen. Der ÖPV bedient sich also auf der einen Seite durchaus ähnlicher Datensätze wie der motorisierte Individualverkehr, da betriebliche Fahrzeugdaten und Verkehrsaufkommensdaten eine vergleichbar bedeutende Rolle spielen. Darüber hinaus ergeben sich durch den stärker organisatorisch geregelten Betrieb eine Reihe von weitergehenden Prozess- und Kundendaten.

Eigentumsdimension: proprietäre und öffentliche Daten

Daten können auf einer Skala von proprietären bzw. privaten Daten zu öffentlich verfügbaren Daten eingeteilt werden. Diese Einteilung nach Eigentums- und Nutzungsrechten ist eine sinnvolle Möglichkeit, um zu spezifischeren Kategorien zu gelangen, da eine solche Einteilung viel über die mögliche Verwertung und Wertschöpfung verrät. Es lässt sich zunächst ableiten, dass bisher nur ein Teil der Daten außerhalb von Unternehmen für Big Data-Anwendungen zur Verfügung steht. Der Großteil relevanter Daten besteht aus wirtschaftlichen System- und Prozessdaten und befindet sich im proprietären Besitz von Unternehmen. Diese Untermenge kann einen qualitativen oder auch quantitativen Charakter haben.

Zweckdimension: Opportunistische und zweckgebundene Daten

Eine weitere Unterscheidung betrifft den Zweck der Datenerhebung. Die Skala reicht dabei von Daten, die für einen bestimmten Zweck erhoben wurden und solchen, die opportunistisch für originär nicht vorgesehene Zwecke verwendet werden. Ein Beispiel opportunistischer Daten im Mobilitätsbereich ist die Verwertung von Bewegungsdaten, die von Telekommunikationsanbietern auf Basis der Kommunikation von Mobiltelefon und Mobilfunkmast erhoben werden. Diese Daten werden erhoben, um die Kommunikation mit einem Mobiltelefon zu ermöglichen, können aber auch für die Bestimmung zellenbasierter Verkehrsflüsse genutzt werden. Ein weiteres Beispiel ist die Nutzung der zahlreichen Sensoren, die in Smartphones verbaut sind. Diese Daten können bspw. in Verbindung mit Crowdsourcing-Kampagnen kollektiv erhoben werden und z.B. für Lärmpegelmessungen genutzt werden. Interessant ist dabei die Differenzierung nach *Overt Crowdsourcing*, also eine offensichtliche Form der dezentralen Erhebung sowie das Gegenstück des versteckten Crowdsourcing: *Covert Crowdsourcing*. Die opportunistische Erhebung wird bestärkt durch die kostengünstige Speicherung und die zunehmende Wahrnehmung von Zweit- und Drittverwertungsmöglichkeiten. Diese Form der Verwertung ist bei personenbezogenen Daten mit hohen Ansprüchen an die Anonymisierung und Pseudonymisierung verbunden. Aufgrund sinkender Kosten können für viele Anwendungsfälle Sensoren für spezifische Erhebungszwecke eingesetzt werden. Wichtigste Eigenschaft hierbei ist, dass möglichst zielgerichtete Datensätze und wenig Rauschen produziert werden. Eingebettete Technologien in der Infrastruktur von Smart Cities bieten eine flächendenkende Datenbasis für weitere Datenverschneidungen.

2.2 BIG DATA-ANWENDUNGEN IM ÖPV

Big Data-Anwendungen im Verkehr können allgemein in zwei Bereiche unterteilt werden (vgl. IBM 2013). Anbieterseitige Anwendungen zur Optimierung interner Prozesse auf Anbieterseite und kundenorientierte Anwendungen, die oftmals darauf ausgerichtet sind, das Kundenerlebnis durch optimierte und schnellere Informationen, einfachere Zugänge oder Abrechnungen sowie Mehrwertdienste zu verbessern. Zu ersteren sind Anwendungen aus dem Bereich Business Intelligence zu zählen. Diese Anwendungen werden bspw. eingesetzt, um Preismodelle zu dynamisieren und die Auslastung von Fahrzeugen zu optimieren, indem Prognosen auf Basis historischer Daten erstellt werden. Ebenfalls zum Bereich anwenderseitiger Anwendungen zu zählen, sind Lösungen zur Optimierung der Wartung und Instandhaltung von Fahrzeugen (predictive maintenance). Insbesondere im Verkehrsbereich ist dies ein wichtiges Thema, da Fahrzeuge und Infrastrukturen teilweise sehr wartungsintensiv sind und die Lagerung und Bereitstellung von Ersatzteilen einen hohen Kostenfaktor darstellen. In einer branchenübergreifenden Befragung zum Einsatz von Big Data in Unternehmen wurde festgestellt, dass datenbasierte Anwendungen bisher meist im Vertrieb und Marketing eingesetzt wurden (Wrobel 2013). Im Mittelpunkt standen Marktmonitoring und die Identifikation von Möglichkeiten zur Einsparung von Kosten und zur Steigerung von Umsätzen. Als nachrangige Ziele wurden Qualitätssteigerung und Innovation genannt.

Beispiele für kundenorientierte Anwendungen sind Lösungen für eine stärkere Personalisierung von Dienstleistungen und Anwendungen zur Steigerung der Kundenbindung. Ohne datenbasierte Anwendungen wären Sharing-Dienstleistungen, intermodales Routing und Mobilitätsplattformen im Bereich der Mobility as a Service (MaaS) nicht umsetzbar. Insbesondere bei der Bereitstellung von intermodalen Verbindungsoptionen – also der Kombination verschiedener Verkehrsmittel auf einem Weg – liegt eine der Herausforderungen darin, aus der Vielzahl der möglichen Verbindungen eine relevante Auswahl anzuzeigen, die dem Nutzer einen wirklichen Mehrwert bietet. Neben der klassischen Routenwahl für Busse und Bahnen wird dabei immer stärker auch auf die Einbindung neuer

Mobilitätsdienstleister gesetzt. Solche Verfahren bieten auch die Grundlage für aktuelle Entwicklungen im Bereich vereinfachter Vertriebswege. Dazu gehören bspw. Check-In/Check-Out oder Be-In/Be-Out-Ticketingsysteme. Diese Angebote bilden auch die Basis für Daten und Technologien, die aus einer Verschneidung unterschiedlichster Datensätze eine Nachfrageprognose für die unmittelbare Zukunft für Kunden liefern. Bspw. nutzt der Carsharing-Anbieter DriveNow Nachfrageprognosen um über Preisreduktionen gezielt die Nachfrage zu stimulieren.

Die zur Anwendung kommenden Verfahren werden als selbstlernende oder intelligente Algorithmen bezeichnet und sind dem Teilgebiet "Künstliche Intelligenz" (AI – artificial intelligence) der Informatik zuzuordnen. Es gibt unterschiedliche Methoden und Verfahren, die als intelligente Algorithmen bezeichnet werden können. Drei wichtige Ansätze sind das maschinelle Lernen, Decision Tree-/Random Forrest-Ansätze und Neuronale Netze.

Das Prinzip des *Maschinellen Lernens* lässt sich vereinfacht mit dem Begriff *Mustererkennung* beschreiben. Es behandelt eine Klassifikationsproblematik, die mit mathematischen Verfahren (Algorithmen) gelöst werden sollen. Klassifizierungsalgorithmen sind mathematische Verfahren, die in der Lage sind aus einer Fülle an Daten Muster respektive Regeln automatisiert zu erkennen bzw. zu lernen. Dies geschieht im Allgemeinen dadurch, dass sie erst auf einem Trainingsdatensatz „lernen“ ein „typisches“ Muster zu erkennen.

Ein weiteres Verfahren befasst sich mit Entscheidungsbäumen und wird entsprechend als *Decision Tree-/Random Forrest*-Ansatz bezeichnet. Dabei wird der Raum der Datenpunkte durch binäre Ja/Nein-Fragen in Bereiche unterteilt. So wird z.B. gefragt, ob ein bestimmtes Feature oberhalb eines bestimmten Wertes liegt. Auf diese Art wird der Raum fragmentiert. Einer der simpelsten Algorithmen mit Baumstruktur ist der *Decision Tree*. Hier wird genau ein Baum angelegt (im Gegensatz zum Random Forest – siehe nächster Abschnitt), der die Struktur des Raumes analysiert und speichert. Ein wichtiger Parameter bei Bäumen ist die Tiefe. Je tiefer der Baum ist, desto genauer strukturiert er den Raum. Eine zu hohe Tiefe führt jedoch fast zwangsläufig zu Overfitting und damit zu einem zu starken Zuschneiden des Classifiers auf die Trainingsdaten.

Ein *Random Forest* funktioniert ähnlich wie ein Decision Tree, jedoch werden hier im Unterschied zu letzterem mehrere Bäume erzeugt (forest). Außerdem wachsen und lernen die Entscheidungsbäume in diesem Verfahren auf randomisierte Art und Weise. Zur Klassifizierung trifft jeder einzelne Baum am Ende eine Entscheidung und die Klasse mit den meisten Stimmen entscheidet über die Klassifikation. Ein Random Forest trainiert deutlich schneller. Die Trainingszeit des Klassifikators steigt dabei linear mit der Anzahl der Bäume. Random Forests sind robuster gegenüber redundanten Features, jedoch gibt es entscheidende Parameter wie die maximale Tiefe eines Baumes und die Anzahl der Features, die in jedem Baum berücksichtigt werden. Diese Parameter zu optimieren ist entscheidend, um eine hohe Genauigkeit zu erreichen.

Künstliche neuronale Netzwerke (KNN) sind ein Netzwerk aus Neuronen, die aus mehreren Ebenen aufgebaut sind. Die wichtigen Parameter die in der Lernphase herausgebildet werden, sind die Verbindungen (Gewichtung) zwischen den Neuronen und der Schwellenwerte der Neuronen. Für KNN gibt es diverse Anwendungsgebiete, so z.B. Schrifterkennung, Data-Mining, KI-Programmierung. Ein KNN kann schematisch über mehrere Ebenen von Neuronen dargestellt werden, auf denen jeweils spezifische Rechenoperationen durchgeführt werden. Ein entsprechender Input (Datenset) wird dann durch das Netzwerk transportiert, wobei auf jeder Ebene von jedem Neuron eine mathematische Operation auf den Daten ausgeführt wird. Der Output des letzten Neurons bestimmt dann die

ermittelte Klasse, der Datensatz wurde entsprechend klassifiziert. Es gibt unterschiedliche Arten der KNN zu lernen, am häufigsten wird der Backpropagation Algorithmus verwendet.

Im Folgenden wird anhand von Big Data-Beispielen aus der Praxis veranschaulicht, welche Anwendungen derzeit existieren, welche Optimierungspotentiale vorliegen und welche Mehrwerte generiert werden, aber auch wo Grenzen bestehen und wie diese Hürden überwunden werden sollen.

2.2.1 ANBIETERSEITIGE ANWENDUNGEN

In diesem Abschnitt wird anhand von Beispielen gezeigt, wie Big Data-Anwendungen unternehmensintern eingesetzt werden. Es werden vier Kategorien berücksichtigt: Anwendungen zur langfristigen und kurzfristigen Angebotsplanung, Anwendungen aus dem Bereich der Fahrzeug- und Infrastrukturwartung und -instandhaltung sowie Anwendungen aus dem Bereich Ticketing und Fahrgeldmanagement.

2.2.1.1 LANGFRISTIGE ANGEBOTSPLANUNG

Die Angebotsplanung im öffentlichen Verkehr der größeren Städte basiert traditionell meist auf daten- und rechenintensiven Verkehrsmodellen. Diese Verkehrsmodelle entstanden in den 1970er-Jahren (z.B. Kutter 1973) und wurden seitdem sukzessive zu marktfähigen Produkten weiterentwickelt. Die Nutzung unterschiedlicher Datenquellen ist dabei von je her zentraler Bestandteil. Zu den neueren Ansätzen zählen Multi-Agenten-Simulationen (Buliung & Kanaroglou 2007). Diese aktivitätenbasierten Modelle sind deutlich datenhungriger und können auch stark disaggregierte Datensätze verarbeiten wie zum Beispiel Gebäude- und Wohnungsdaten oder detaillierte Verkehrstagebücher. Multi-Agenten-Simulationen werden zunehmend mit klassischen Planungstools gekoppelt, können aber auch neuartige Datensätzen wie Smartphone Tracking Data oder Floating Phone Data verarbeiten. Validiert werden die Ergebnisse typischerweise mit Daten aus automatisierten Zählstellen und weiteren Verteilungswerten aus Verkehrserhebungen. Damit sind optimale Voraussetzungen gegeben, um Skalierungseffekte in der Erstellung von Verkehrsmodellen zu ermöglichen.

Langfristige Angebotsplanung mittels Multi-Agenten-Simulation

Dass Verkehrsmodelle zur langfristigen Angebotsplanung mittlerweile hinsichtlich des Datenvolumens, der Verarbeitungsgeschwindigkeit und der Datenvariabilität an die Big Data-Definition heranreichen, zeigt die Anwendung der Multi-Agenten-Simulation MATSim. Das Züricher Unternehmen Senozon nutzte MATSim, um ein dynamisches Nachfrage- und Verkehrsflussmodell für die Berliner Verkehrsbetriebe (BVG) zu erstellen und zu betreiben. Das Modell wird genutzt, um die Haltestellenplanung, Linienführung und Taktung von Buslinien mit starken Verspätungen zu verbessern (Neumann, Balmer, & Rieser 2012). Darüber hinaus dienen die Ergebnisse dazu, Effekte zwischen den verschiedenen Verkehrsmodi besser abzuschätzen. So kann bspw. berechnet werden, welche Haltestellen besonders stark in Konkurrenz zum Radverkehr stehen (Rieser 2016).

Verkehrsprognosen in Echtzeit

Aber auch die Anbieter eher traditioneller Planungstools setzen zunehmend auf Echtzeitverarbeitung von Mobilitätsdaten. So bietet der Weltmarktführer PTV mit dem Produkt *Optima* eine modellbasierte Lösung für Verkehrsprognosen in Echtzeit an. Dabei werden zahlreiche Datenquellen genutzt: Informationen aus automatisierten Zählstellen, Floating Car Data (FCD) und weitere Live-Daten der

Unternehmen INRIX, MotionLogic und TomTom. Die Daten werden in Echtzeit verarbeitet, um eine Nachfrageprognose für die nächste Stunde zu generieren. Aktuell werden diese Prognosen vor allem zum Verkehrsmanagement in Großstädten eingesetzt. Mit dem Zusatzmodul Public Transport ETA soll es aber auch ÖPV-Unternehmen ermöglicht werden, proaktiv auf die Kundennachfrage zu reagieren.[8] Im Rahmen eines Pilotprojekts wurde auf Basis von Floating Phone Data des Anbieters MotionLogic eine Quelle-Ziel-Matrix für die Stadt Karlsruhe generiert. Laut Nobert Weber von MotionLogic wurden dazu Verkehrsbewegungen in einem Zeitraum von sechs Wochen zwischen 136 Verkehrszellen ausgewertet. Aus Datenschutzgründen werden dabei keine Daten von einzelnen Personen verwendet. Die Daten werden aggregiert und Gruppen von mindestens 30 Personen analysiert.

Angebotsoptimierung und Standortplanung mittels predictive modelling

Neben den beschriebenen Ansätzen – die im Grunde klassische Planungsprozesse aufgreifen und inkrementell weiterentwickeln – werden Business Intelligence Lösungen eingesetzt, die auf andere Prämissen zurückgreifen. Das Unternehmen GeoSpin setzt nach eigener Aussage Methoden wie Deep Learning, neuronale Netze und maschinelles Lernen ein, um die Angebotsplanung zu optimieren. In Zusammenarbeit mit dem Carsharing-Anbieter Car2Go wurden diese Methoden genutzt, um das Bediengebiet des Free-Floating-Carsharing-Systems zu optimieren. Dazu wurden mehr als eine Million Buchungen analysiert und das Angebot auf Basis dieser Analyse angepasst.[9] Ein weiteres Anwendungsfeld der Methodik im Verkehrsbereich ist die Standortplanung für Ladeinfrastrukturen für Elektrofahrzeuge. Nach Aussage von GeoSpin werden die eingesetzten Big Data-Methoden neben der Angebotsplanung nun zunehmend für die Standortplanung von Filialen und ortsbezogenen Dienstleistungen genutzt. Im mFUND-Projekt ANODAS wird die Machbarkeit einer weiterführenden Analyse unter Verknüpfung von Unternehmensdaten mit Open-Data-Angebote in diesem Bereich betrachtet.

Mobile Network Data

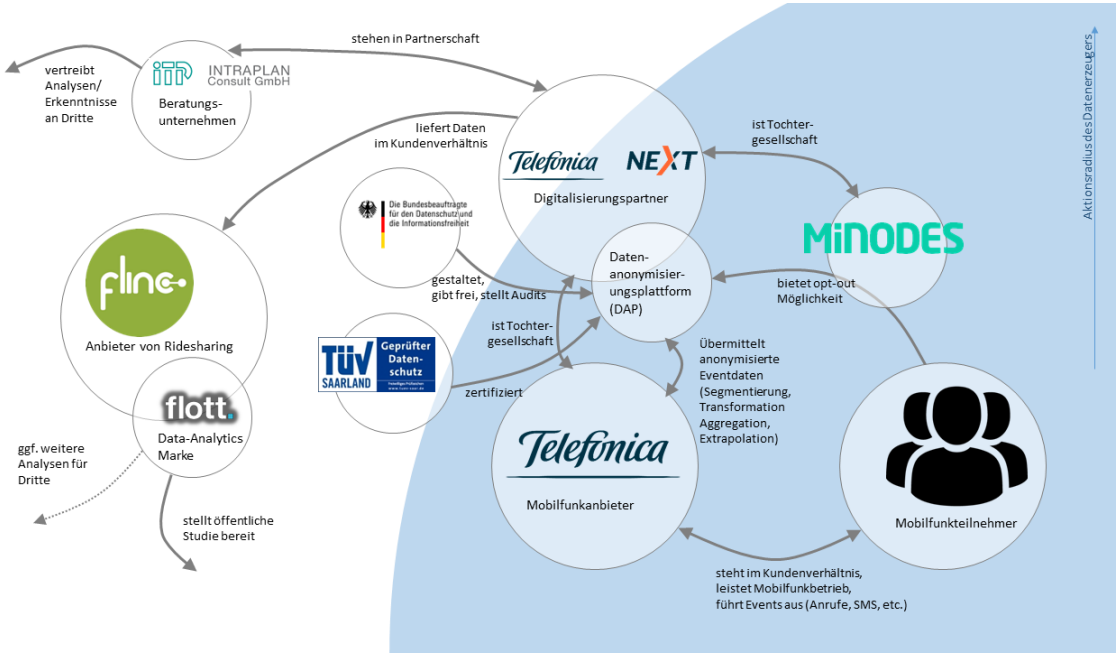
Einen wahren Hype erfährt die Nutzung von Floating-Phone-Daten für die Angebotsplanung im ÖPNV. So bietet die Telefonica Next GmbH Unternehmen des ÖPV die Bereitstellung umfassender Quell-Ziel-Analysen in Aussicht. Über anonymisierte Daten der Mobilfunkkunden können sehr umfassende regionale Bewegungsmuster mit hoher zeitlicher und räumlicher Auflösung bereitgestellt werden (vgl. Abschnitt 2.1.3). Dabei erzeugen laut Anbieter circa 43 Millionen Kunden ungefähr 170 Millionen ‚Events' in der Stunde – also Ereignisse, die Aussagen zum Aufenthaltsort ermöglichen. Die Verschneidung mit weiteren Datenquellen, z.B. sozialdemographischen Kennwerten sowie die schnelle Prozessierung aus Datenschutzgründen, machen das Verfahren zu einem Big Data-Prozess. Mit einzelnen ÖPV-Anbietern wird bereits zusammengearbeitet.[10]

Auch im Umfeld neuer Mobilität wird FloatingPhone Data für die Planung eingesetzt (vgl. Abb. 4). Für die Stadt Hamburg hat der Ridesharing-Anbieter flinc GmbH Verkehrsszenarien mit autonomen Shuttles im Simulationsmodell analysiert. Basis bildeten 35 Millionen Bewegungen für zwei Monate des Jahres 2015. Eine Validierungsstudie für den Stuttgarter Raum, die in Kooperation mit dem Fraunhofer Institut für Arbeitswirtschaft und Organisation (IAO) durchgeführt wurde, führte zu neuen

8 Vgl. http://vision-traffic.ptvgroup.com/en-uk/products/ptv-optima/use-cases/real-time-public-transport-eta/.
9 Vgl. https://www.dialog-energie-zukunft.de/datengetriebenane-mobilitaet/.
10 Vgl. https://www.vdv.de/digitales-forum---startseite.aspx.

Erkenntnissen bezüglich bestimmter Pendlerströmen, deren Verortung sowie deren Aufteilung nach Verkehrsmodi. Die Ergebnisse wurden mit weiteren, vergleichbaren Datensätzen wie Fahrgastbefragungen, automatischen Zählstellen oder GPS-Daten validiert. Die Verschneidung mit solchen Daten ist eine wichtige zukünftige Arbeit, um Verkehrsmittel mit einer höheren Qualität zu identifizieren.

Abb. 4: Akteurskonstellation am Beispiel von Telefonica NEXT und flinc



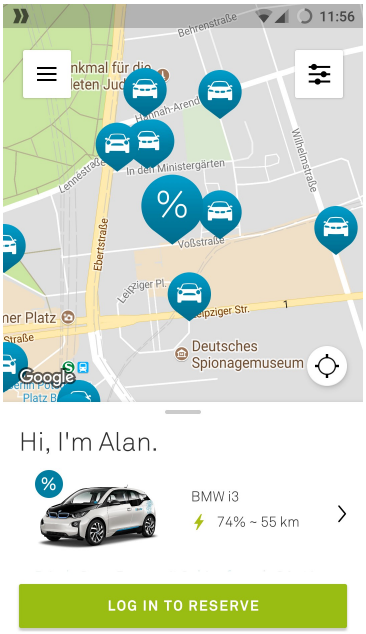
2.2.1.2 KURZFRISTIGE NACHFRAGEPROGNOSE

Die zuvor genannten Analysen zur Nachfrageabschätzung sind unabhängig vom Datenvolumen der Nutzer skalierbar, da sie nicht auf die Übertragung geräteseitig erfasster Sensordaten angewiesen sind, sondern passiv erzeugte Events (im Hintergrund erzeugter Datensatz des Mobiltelefons zur Identifikation der umliegenden Funkzellen) oder andere passiv erzeugte Datentypen nutzen. Dadurch sind sie für kurzfristige Prognosehorizonte einsetzbar und können als Grundlage für tagesaktuelle Angebotsplanungen dienen.

Gezielte Stimulation der Nachfrage

Kurzfristige Nachfrageprognosen können genutzt werden, um über Preisreduktionen gezielt die Nachfrage zu stimulieren. So bietet der Carsharing-Anbieter DriveNow Fahrzeuge beobachteter schwacher Auslastung (in Nachfragesenken) zu verringerten Fahrpreisen an, um diese schneller in Gebiete mit erhöhter Nachfrage zu überführen. Diese Preisnachlässe werden dem Kunden über Prozentsymbole anstelle der Fahrzeugsymbole angezeigt (vgl. Abb. 5). Eine entsprechende Technologie wurde von der Bundeswehruniversität München mit Methoden der automatischen Optimierung auf Basis von Shared-Mobility-Buchungsdaten entwickelt (Weikl & Bogenberger 2015). Werden historische Datensätze der Anbieter eingesetzt, so werden überschaubare Mengen an Daten verwendet. Das Volumen entsteht durch die Einbindung von externen Faktoren wie zum Beispiel Wetterdaten oder Eventdaten sowie der Verknüpfung mit örtlichen Geodaten.

Abb. 5: Stimulation der Nachfrage beim Carsharing-Anbieter DriveNow



Diese Anreizsysteme können gedanklich weitergeführt werden, insbesondere im Bereich des Geomarketings. Das StartUp-Unternehmen Car2Ad nutzt Informationen über Aufkommen und Mobilitätsnachfrage und platziert gezielt Marketinganreize über Displays, die auf Autodächern angebracht sind. Im mFUND-Projekt Smarttop wird ein Konzept entwickelt, mit dem Dispositionsfahrten von elektrischen Carsharing-Fahrzeugen zu Ladesäulen und zu Orten mit hoher Nachfrage als Freifahrten angeboten werden können. Die Informationen werden über eine dezentrale Monitor-Infrastruktur vermittelt, die gleichzeitig enorme Datenmengen zum location-based targeting über ein live WiFi-Profiling am Carsharing-Standort ermöglicht.

Die kundenseitige Fahrzeugdisposition bei neuartigen Mobilitätsangeboten kann also durch Big Data-Technologien als eine Form der betrieblichen Optimierung erfolgen. Das klassische Beispiel hierfür ist UBER, das seine Taxiflotte so positioniert, dass ein bestimmter Richtwert für die Kundenwartezeit erreicht wird. Abweichungen werden über ein datengestütztes Yieldmanagement abgefedert.

Derzeit versuchen Anbieter, u.a. CleverShuttle, die DB AG oder auch die BVG, in Deutschland derartige nachfrageorientierte Angebote zu platzieren. Dies geschieht im Unterschied zu UBER in Kooperation mit Kommunen und örtlichen Verkehrsanbietern. Das Start-Up-Unternehmen CleverShuttle bietet ein System, das Fahrgast und Beförderer zusammenbringt. Dabei setzt der Anbieter auf Elektromobilität und strebt die datengestützte Optimierung der Kosten und Einnahmen mit einem nachfragegesteuerten Angebot (On-Demand Ridesharing), einem intelligenten Pooling, einer optimierten Routenführung sowie der automatisierten Disposition an. Die Disposition der Fahrzeuge berücksichtigt dabei ebenso die vorgeschriebenen Wartezeiten der Fahrer sowie die betriebsbedingten Ladezeiten der Fahrzeuge.

2.2.1.3 FAHRZEUG- UND INFRASTRUKTURWARTUNG

Neben der Disposition ist die Wartung von Fahrzeugen und Infrastruktur ein Einsatzfeld mit zum Teil bestehender Praxis der Big Data-Technologie. Die Herausforderung liegt dabei in der Optimierung der vorrausschauenden Wartung von Fahrzeugen und Infrastrukturelementen.

Predictive Maintenance im Luftverkehr

Vorreiter im Bereich datenbasierter Fahrzeug- und Infrastrukturwartung ist die Luftverkehrsbranche. Unternehmen wie die Lufthansa nutzen Methoden des predictive maintanence, um kostenintensive Standzeiten ihrer Flotte zu optimieren und gleichzeitig die Flugsicherheit zu maximieren. Bei einer Flotte von über 1.000 Flugzeugen entstehen sehr große Mengen an Sensordaten, Log-Daten, Prozesseinträgen und Fehlermeldungen, die für eine vorhersehende Wartung eingesetzt werden können. Eine der Lösungen, die dafür eingesetzt werden, ist die RapidMiner Data Science Plattform.[11] Diese Software ermöglicht Datenanalysten die Datenaufbereitung und Analyse mittels Machine Learning. Damit werden Echtzeitanalysen von Daten ermöglicht. Mittels Mustererkennung und Klassifizierung werden Wartungsaufträge zu optimalen Zeitpunkten vergeben. Auch Plattformen auf Basis von SAP BW-on-HANA und SAP BO werden für eine datengestützte Wartung bei Lufthansa eingesetzt.[12]

Predictive Maintenance im Schienenverkehr

Im Schienenverkehr bietet die Siemens AG als einer der größten Schienenfahrzeughersteller weltweit ihren Kunden eine Predictive-Maintenance-Lösung an. Auch hier sollen Wartungsvorgänge datengestützt optimiert werden. Das klassische Vorgehen des *Reactive Maintenance* – also dem Reagieren auf technische Ausfälle – sowie der Regelwartung soll dabei durch den Einsatz von Big Data abgelöst werden. In Anlehnung an den Begriff *Internet of Things* spricht der Hersteller dabei vom *Internet of Trains*. Neben Pilotprojekten in Spanien und Großbritannien setzt sich die Technologie allmählich auch im deutschen Markt durch. Bei der Bestellung von ICX der Deutschen Bahn AG (DB AG) wie auch den ca. 100 Regionalzügen der Rhein Ruhr Express spielten die Predictive-Maintenance-Lösungen laut Herstellerangaben als Zusatzdienste eine wichtige Rolle. Wichtigste Datenquelle bei diesem Dienst ist der Zufluss aus der Fahrzeugsensorik, der nahezu in Echtzeit realisiert wird. Zur vorbereitenden Analyse wird ein vergleichsweise übersichtlicher Datensatz von ca. einer Million Sensorlogdaten verwendet. Diese wurden im Rahmen eines Pilotversuchs im Laufe eines Jahres im Fünfminutentakt von 300 Sensoren erhoben. Verglichen mit weiteren betrieblichen Prozessdaten können mittels datengestützter Mustererkennung die relevantesten Ereignisse identifiziert werden. Die entsprechenden Grenzwerte werden nun im Realbetrieb fortlaufend und automatisiert überprüft und führen nach einer Validierungszeit entsprechend zu Wartungsaufträgen. Siemens nutzt als Technologie ein angepasstes Format der Unified Data Architektur (UDA) des Systemlieferanten Teradata. Diese Entwicklung bedient sich u.a. der Hadoop-Technologie. Für Siemens wird dabei die Tür für eine Reihe neuer Produkte für den Verkehrsmarkt aufgestoßen. So seien nun laut Gerhard Kress (Direktor Mobility

11 https://rapidminer.com/resource/lufthansa-predictive-maintenance/.

12 https://www.lufthansa-industry-solutions.com/de-de/loesungen-produkte/sap/neue-reporting-plattform-erleichtert-flugzeugueberholung/.

Data Services) neue Produkte und Geschäftsmodelle denkbar wie Betriebszeitgarantien, Risikoteilungsmodelle oder Vergabeaufträge, die auf der Performanz der Züge basieren.[13]

Predictive Maintenance im Infrastrukturbereich

Im Schieneninfrastrukturbereich stehen insbesondere Weichen, Signale und Gleise im Fokus. Der größte Netzbetreiber in Deutschland, die DB Netz AG, setzt dabei u.a. auf Big Data-Lösungen aus dem Start-Up Bereich. Hier wurden in den vergangenen Jahren einige Initiativen über ein bahninternes Accelerator-Programm (DB Mindbox) gefördert. Nennenswert ist das Angebot von Konux, das eine umfassende Lösung vom Sensor über die Datenanalyse bis zur interaktiven Überwachung für die Weicheninstandhaltung anbietet. Wartungen werden mit der Lösung vorrausschauend planbar und Störungen reduziert. Gleichzeitig können die Sensoren auch einen Beitrag zur Bewertung der Schienenfahrzeuge leisten. Auf künstlicher Intelligenz basierende prädiktive Algorithmen werden genutzt, um wiederkehrende Muster zu erkennen und relevante Abweichungen zu identifizieren. Das Ergebnis ist die automatische Erzeugung einer Wartungsempfehlung. Eine solche Datenanalyse wurde zusammen mit der DB AG pilotiert, um in Echtzeit den Zustand der Weichen zu ermitteln.[14] Weitere Unternehmen, die im Bereich Predictive Modelling mittels Sensordaten aktiv sind und im Rahmen des Accelerator-Programms gefördert werden, sind Senvisy, Neuron Soundware oder iNDTact. Das erste Beispiel verarbeitet mit dem Produkt SensorLink Schallwellenaufnahmen, die durch die Integration im Netzwerk für die Auswertung im Bereich Predictive Maintenance genutzt werden. Auch Neuron Soundware nutzt Schallquellen, um auf einer Analytics-Plattform die Wartung zu verbessern. *iNDTact* liefert hochsensible, aber robuste Sensoren, die zur problemorientierten Analyse bei der Prozess- und Zustandsüberwachung genutzt werden können.

Predictive Maintenance im Nahverkehr

Der Stuttgarter Nahverkehrsbetrieb *Stuttgarter Straßenbahnen (SSB) AG* setzt laut Hersteller Microsoft die Azure Plattform als Cloud-Anwendung für die Instandhaltung der Oberleitungen ein. Hier geht es weniger um eine Form des Predictive Maintenance, sondern eher um eine Form der Crowd-Sourced-Datenerhebung über die Mitarbeiter. Die entsprechende Inspektion wird mit einer mobilen Dateneingabe über eine WebApp abgeschlossen. Die Cloud-Anwendung dient dabei zur Datensynchronisation. Erwähnt wurde, dass im Piloten die Inspektion von ca. 5.000 Masten mit Stammdaten zu Bauart, Durchmesser und Position sowie variablen Daten zum aktuellen Zustand durchgeführt werden musste. Bei diesem Beispiel erfolgt keine direkte Big Data-Anwendung, aber der Einsatz einer zentralen Datenplattform ebnet den Weg für einen entsprechenden Einsatz, da eine entsprechende skalierbare Dateninfrastruktur zur späteren Einbindung z.B. von Sensordaten vorhanden ist.[15]

Betriebliche Optimierung

Bei der Optimierung von sonstigen betrieblichen Prozessen finden sich wiederum vermehrt Beispiele aus der Luftfahrt – insbesondere international aber auch auf nationaler Ebene. Am Flughafen Frankfurt werden Business Intelligence Plattformen verwendet, um relevante Managementaussagen aus den komplexen Datenquellen einer der weltgrößten Flughäfen zu gewinnen. Das SAS-basierte System

13 http://www.teradata.at/Resources/Case-Studies/Siemens-The-Internet-of-Trains.

14 https://www.konux.com/de/fallstudien/.

15 https://customers.microsoft.com/de-DE/story/mobile-masten-inspektion-mit-schlanker-lsung-aus-der-azure-cloud.

dient dem Betreiber Fraport als Messinstrument für das Gesamtbild und bindet Flugbewegungen, Passagier- und Personaldaten, Verspätungen und Wetterdaten in nahezu Echtzeit ein, denn die Datengrundlage wird alle fünf Minuten aktualisiert. Die Betrachtung von extrem komplexen Wechselwirkungen wird dabei gekoppelt mit dem Anspruch quantifizierbare Planungsgrundlagen heranzuziehen, so dass nach Aussage des Systemlieferanten eine hohe Qualität beim Datenverarbeitungsprozess vorliegen muss. Eine historisch gewachsene IT-Infrastruktur mit extrem heterogenen Prozessen muss verlässlich zu einem Big Picture zusammengeführt werden. Dieses wird aus relevanten Metadaten gebildet und – in Form von Key Performance Indikatoren aufbereitet – den entsprechenden fachlich kompetenten Mitarbeitern kommuniziert. Die Anwendung soll von klassischen Nutzungskontexten, wie Controlling und Reporting, auch auf weitere Felder wie z.B. die Parkraumbewirtschaftung ausgeweitet werden.[16]

Weniger umfassende Beispiele finden sich im Bereich des schienengebundenen Verkehrs. Die Schweizer Bundesbahn nutzt eine SAP-HANA-Lösung, um betriebliche Kosten zu reduzieren. Ganz konkret wird bei diesem Anwendungsfall der Frage nachgegangen, inwiefern teure Energiespitzen abgefedert werden können, ohne dass der Endkunde davon etwas mitbekommt. Die Big Data-Lösung ermöglicht es dem Betreiber, in Echtzeit auf Lastspitzen zu reagieren und unmerklich Heizkörper für Züge und Weichen zu drosseln, um den Energieverbrauch im notwendigen Maß zu senken. Gleichzeitig können dadurch Infrastrukturinvestitionen vermieden und die Sicherheit der Energieversorgung erhöht werden.[17]

Die Big Data-Anwendungen im Bereich predicive maintenance erfahren derzeit besonders große Aufmerksamkeit, da mit dem Einsatz dieser Technologien große Kosteneinsparungen erwartet werden. Allerdings tritt nach den ersten Erfahrungen in der Praxis etwas Ernüchterung ein, da nicht alle Erwartungen vollumfänglich erfüllt werden. Nur 6 % der deutschen Industrieunternehmen sehen derzeit in der vorausschauenden Instandhaltung einen nachhaltigen Nutzen, 75 % sind von dem technologischen Konzept derzeit noch nicht überzeugt.[18]

Für deutsche ÖPNV-Anbieter finden sich in der Literatur aktuell nur wenige Beispiele für Big Data-Analysen. Eine wichtige Voraussetzung für eine standardisierte Durchführung wird aktuell durch die VDV-Vernetzungsinitiative gelegt. Das Beispiel wird im Kapitel 2.2.2.2 näher erläutert.

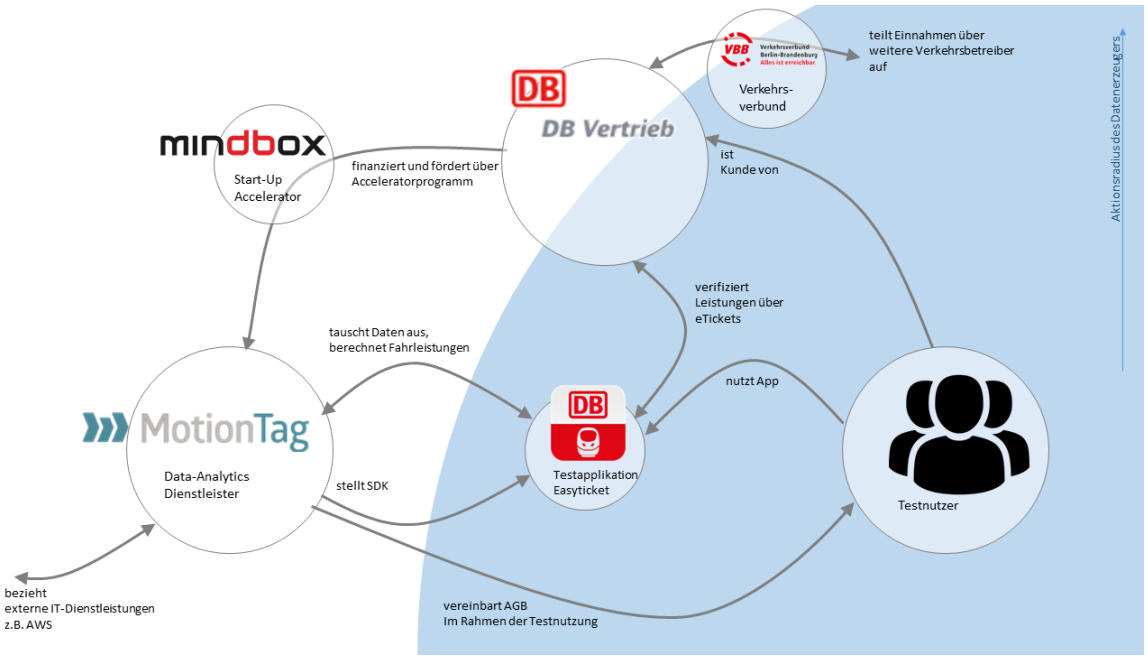
2.2.1.4 TICKETING UND FAHRGELDMANAGEMENT

Die Analyse der Daten des öffentlichen Verkehrs kann dazu genutzt werden, die Fahrgewohnheiten der Verkehrsnutzer im gesamten Verkehrsnetz zu verstehen und darauf aufbauend den Betrieb zu optimieren. Mittels unterschiedlicher Methoden wird erfasst, wohin die Nutzer reisen, welchen Verkehrsträger sie wählen, wie häufig sie reisen und wie zuverlässig die Fahrten sind. Datenquelle sind bspw. Smartcards. Diese werden seit Anfang der 2000er-Jahre in einigen Ländern eingesetzt. Diese Zugangsmedien sind bspw. mit RFID-Sensoren ausgestattet. Mit Hilfe der Sensoren wird der Zeitpunkt und Ort des Ein- und Ausstiegs erfasst. In Deutschland haben sich diese Systeme nicht etabliert. Beispiele aus anderen Ländern zeigen aber, dass Verkehrsunternehmen, die Smartcards einsetzen,

16 https://www.sas.com/de_de/customers/fraport.html.
17 https://www.sap.com/documents/2017/11/368f3d5a-df7c-0010-82c7-eda71af511fa.html.
18 http://www.automotiveit.eu/predictive-maintenance-enttaeuscht-erwartungen/news/id-0060652.

einen Erfahrungsvorsprung in der Arbeit mit Daten vorweisen können, der auch für die Arbeit mit Big Data vorteilhaft ist. Zu diesen Unternehmen zählt bspw. Transport for London (TfL). TfL organisiert als eine Art Dachorganisation den Verkehr in London. Neue Wege geht auch das Start-Up MotionTag, das über Machine Learning und Smartphone Tracking automatisch Wege und Verkehrsmittel erkennt und damit eine Verifizierung für die jeweiligen Vertriebsstellen des ÖPV-Partners auslöst (Vgl. Abb. 6).

Abb. 6: Akteurskonstellation am Beispiel von MotionTag und DB Vertireb



Smartcard

TfL verwendet Big Data Tools zur Analyse der Smartcard-Daten und Daten zu Fahrzeugstandorten. Die Ergebnisse werden für die Netzwerkplanung, für die betriebliche Organisation und zur Kundenbindung genutzt. Eine wichtige Datenquelle ist die lokale Smartcard-Lösung mit dem Namen Oystercard. Mit der Oystercard werden Informationen über ca. 8 Millionen Wege pro Tag erfasst. Für einen Zeitraum von acht Wochen sind die Daten einzelner Fahrten individuellen Oystercards zuzuordnen. Nach Ablauf dieser Zeit werden die individuellen Informationen gelöscht, die Angaben zu den genutzten Routen bleiben jedoch für Forschungs- und Entwicklungszwecke vorhanden. Zusätzlich sind Busse mit GPS-Sensoren ausgestattet, so dass Informationen zum Standort der Busse ebenfalls vorliegen. Durch die Verknüpfung von Nutzer- und Bewegungsdaten können detaillierte Informationen zum Mobilitätsverhalten unterschiedlicher Zielgruppen erhoben werden (Parliamentary Office of Science & Technology 2014).

Fahrgeldmanagement

Der Magdeburger Regionalverkehrsverbund (Marego) und die Otto-von-Guericke-Universität in Magdeburg teilten im August 2017 mit, eine Kooperation zur Verbesserung des Fahrgeldmanagements einzugehen. Die Studenten der Fakultät für Informatik (FIN) arbeiten hierzu in einem projektbezogenen Kurs an einer datenbankbasierten Software. Dieser Kurs wurde im Sommersemester 2017 bereits zum zweiten Mal als *Digital Engineering Project - Transportation Data Management for Marego* angeboten und umfasste neben der agilen Programmierung auch Entwicklungssprints.

2.2.2 KUNDENORIENTIERE ANWENDUNGEN

Die Auswertung von Kundendaten soll Unternehmenstätigkeiten revolutionieren und traditionelle Methoden der Marktforschung und Kundenbindung ablösen (Bloching, Luck & Ramge 2012). Gerade im öffentlichen Verkehr ist der Kunde nach wie vor das unbekannte Wesen. Mittels Daten kann kundenzentrierter agiert werden. Dazu müssen Kundendaten jedoch intelligent genutzt und Datenschutzbedenken sachgemäß berücksichtigt werden.

2.2.2.1 INFORMATIONSVERMITTLUNG

Der Taxifahrtenvermittler mytaxi nutzt Big Data-Anwendungen, um extreme Nachfrageschwankungen zu bewältigen und setzt dafür auf Services und Technologien von AWS. Dabei geht es darum, die Stabilität der Dienstleistung in Zeiten hoher Nachfrage, zum Beispiel zum Jahreswechsel, zu gewährleisten. Ziel ist ein zuverlässiger Service für den Endkunden, wenn sich bspw. zu Spitzenlastzeiten die Nachfrage an einem Abend nahezu vervierfacht. mytaxi betreibt eine der weltweit am meisten genutzten Taxi-Applikation mit ca. 10 Millionen Nutzern sowie 45.000 Taxis in 40 Städten aus sechs Ländern. Die Servicekontinuität wird mit einer elastischen Infrastruktur basierend auf einer Microservices-Architektur auf Docker-Containern gewährleistet. Damit kann auch das sekundäre Ziel einer wachstumsbasierten Skalierung adressiert werden. Inwiefern mytaxi die AWS nutzt, um datengestützte Prozesse zu ermöglichen, ist unklar. Für die App wird ein Big Data-Einsatz auf Basis der AWS-Technologie angedeutet, um dem Kunden eine genauere Fahrzeuganzeige anzubieten. Jedoch können die genutzten Technologien auch für weitergehende Big Data-Analysen genutzt werden. Die Architektur mit sogenannten Docker-Containern wird über die AWS-Dienste Amazon Elastic Compute Cloud (EC2) und Amazon Elastic Container Service (ECS) ermöglicht. Dabei dienen die EC2-Anwendungen aufgrund der Cluster-Unterstützung auch für Big Data-Anwendungen. Für die Datenspeicherung und -verwaltung werden die AWS-Dienste Amazon S3 und Amazon Relational Database Services (RDS) als relationale Datenbankstruktur eingesetzt. Für die in S3 verwalteten Daten können mit Zusatzmodulen mehrere Petabytes für Datenbankabfragen herangezogen werden. Die Mensch-Maschine-Schnittstelle im Beispiel mytaxi erfolgt jedoch über Amazon Simple Queue Service (SQS) sowie Amazon Simple Notification Service (SNS), welche nicht spezifisch für Big Data-Anwendungen sind.[19]

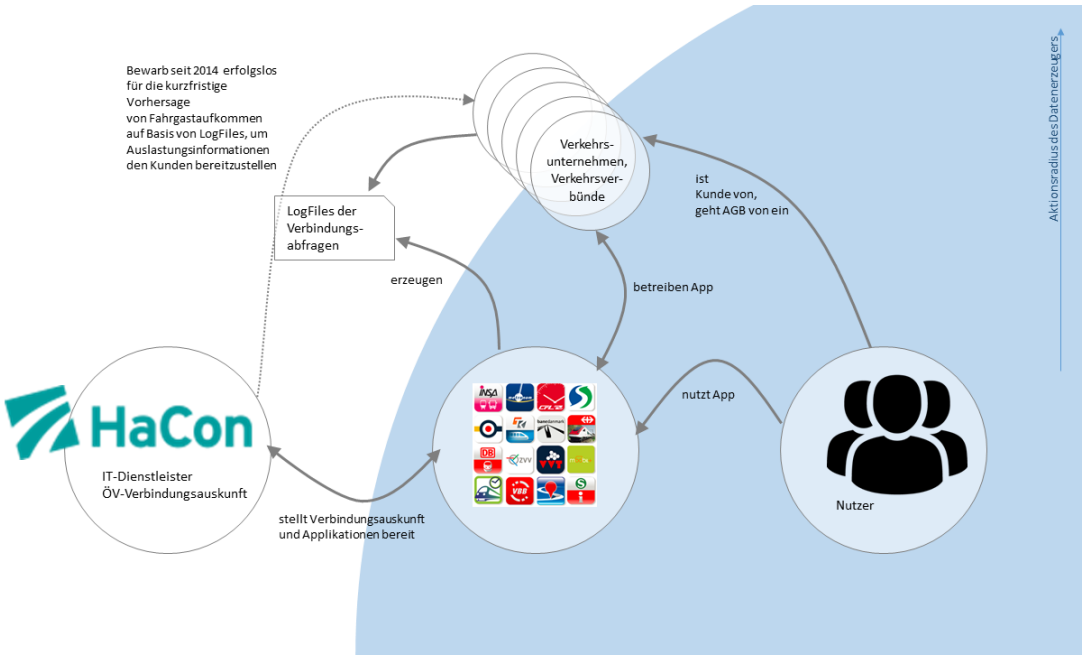
Informationen zur Kapazitätsauslastung

Um den Fahrgästen Informationen zur Auslastung zu vermitteln, erprobte die HaCon Ingenieurgesellschaft mbH die Berechnung der voraussichtlichen Fahrgastaufkommen mittels einer Verschneidung von verschiedensten Datenquellen. Ein wesentlicher Teil der Berechnung bildeten die LogFiles der Verbindungsauskünfte der kooperierenden Unternehmen, die das HaCon-Produkt HAFAS nutzen (vgl. Abb. 7). Die LogFiles geben neben Hinweisen zu den gewünschten Verbindungen und Uhrzeiten, die die Nutzer für die Verbindungssuche aufrufen, auch Hinweise auf die spezifischen Verbindungen, für die weitere Details recherchiert wurden. Zusammen mit weiteren Datensätzen – etwa den Ticketverkäufen und spezifischen Verbindungsabonnements, wie sie typischerweise von Pendlern mit ganz

19 https://aws.amazon.com/de/solutions/case-studies/mytaxi/.

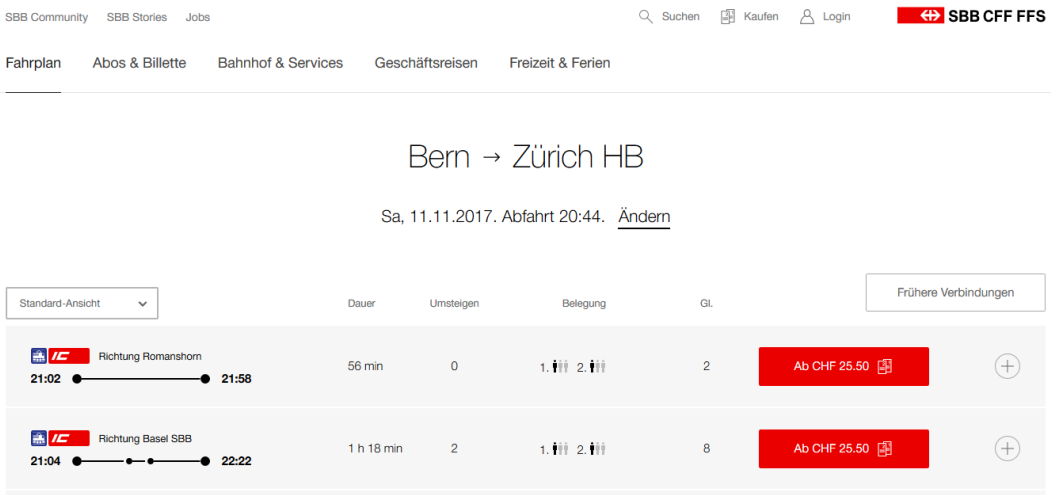
spezifischen Streckenrelationen genutzt werden – kann die HaCon für die nahe Zukunft Aufkommensabschätzungen erzeugen.

Abb. 7: Akteurskonstellation am Beispiel von HaCon und ÖPNV-Anbieter



Diese Technologie wurde bereits 2013 auf der HAFAS-Anwenderkonferenz vorgestellt. Eine Einbindung einer entsprechenden Prognose erfolgte bei der Schweizer Bundesbahn auf Basis einer alternativen Lösung (vgl. Abb. 8). Allerdings ist dabei zu beachten, dass die rein datenbasierten Prognosen deutliche Genauigkeitsschwankungen haben, da die Entsprechung *Informationsabfrage* und *Fahrgastaufkommen* durch Pendlerströme verzerrt werden, da diese aufgrund des ausgeprägten Routineverhaltens weniger Informationen anfragen. So werden die Auslastungsprognosen im Nachhinein mit Erfahrungswerten gewichtet und mit Testfahrten überprüft, um in der Praxis die erforderlichen Genauigkeiten zu erreichen.

Abb. 8: Big Data-gestützte Kapazitätsabschätzung bei der SBB



Quelle: SBB.ch

2.2.2.2 MOBILITÄTSPLATTFORMEN UND MAAS

Mobilitätsplattformen mit Metasuchfunktion über verschiedene Anbieter hinweg sind im Bereich des Luftverkehrs sehr weit verbreitet. Auf Flugvergleichsportalen werden Flugangebote von zum Teil hunderten Informationsquellen bezogen und für den Kunden in verständlicher und nutzbarer Form aufbereitet. Beispiele dafür sind die Portale Fluege.de, wegoreise.de oder Amadeus.net, die als weltweit führende Metasuchdienste im Flugverkehr gelten. Als Kern solcher Dienste muss eine leistungsfähige Datenplattform vorhanden sein. Im Falle von wegoreise.de wird dabei das Software-Tool Apache Storm verwendet, um Daten in Echtzeit zu prozessieren. Dabei wird extrem flexibel mit den Datensätzen hantiert, so dass große Datenmengen über verschiedene Kanäle und Analysen weiterverwertet werden.[20]

VDV Vernetzungsinitiative

Auch im klassischen ÖPV werden umfassende (intermodale, multimodale) Suchen inkl. Preisinformationen und Buchungs- sowie Abrechnungsfunktionalität nach dem Vorbild des Flugverkehrs angestrebt. Beispiele sind dafür moovel, Qixxit oder free2move, die jeweils unterschiedliche Schwerpunkte und Integrationstiefen aufweisen. Bemerkenswert ist auch die Brancheninitiative des VDV, die derzeit unter dem Label *Vernetzungsinitiative* kommuniziert wird.[21] Für die beteiligten Verkehrsunternehmen und -verbünde bedeutet dieser Schritt eine Ausweitung digitaler Prozesse von einer lokalen Ebene innerhalb eines Verbundes auf eine besser skalierbare Ebene bundesweit. Hierdurch sollen Türen zu Maßnahmen der Digitalisierung geöffnet werden, die für die einzelnen Akteure bisher nicht denkbar waren. Das Vorgehen ist dabei nicht zuletzt dadurch begründet, dass sich bislang branchenfremde Unternehmen im Verkehrsbereich als neue Akteure etablieren und den Innovationsdruck dadurch erhöhen. Angestrebt wird eine Standardisierung im Umgang mit Fahrplan-, Haltestellen-, Tarif- und sogar Kundendaten. Daneben steht die Errichtung einer übergreifenden Plattformarchitektur im Fokus. Unmittelbares Ziel ist die Bereitstellung eines anbieterübergreifenden Ticketing. In weiteren Schritten sollen auch die Angebote neuer Mobilitätsdienstleistungen sowie mobilitätsferne Elemente eingebunden und eine Skalierung nach unten, also eine Systemübertragung auf kleinere Anbieter, umgesetzt werden. Diese Schritte dienen dazu, nicht nur Skaleneffekte auszuschöpfen, sondern auch neue Einnahmemöglichkeiten durch digitale Business-Cases zu finden. Zudem ist der VDV darauf angewiesen, die Interessen aller im Verband organisierten Unternehmen zu vertreten.

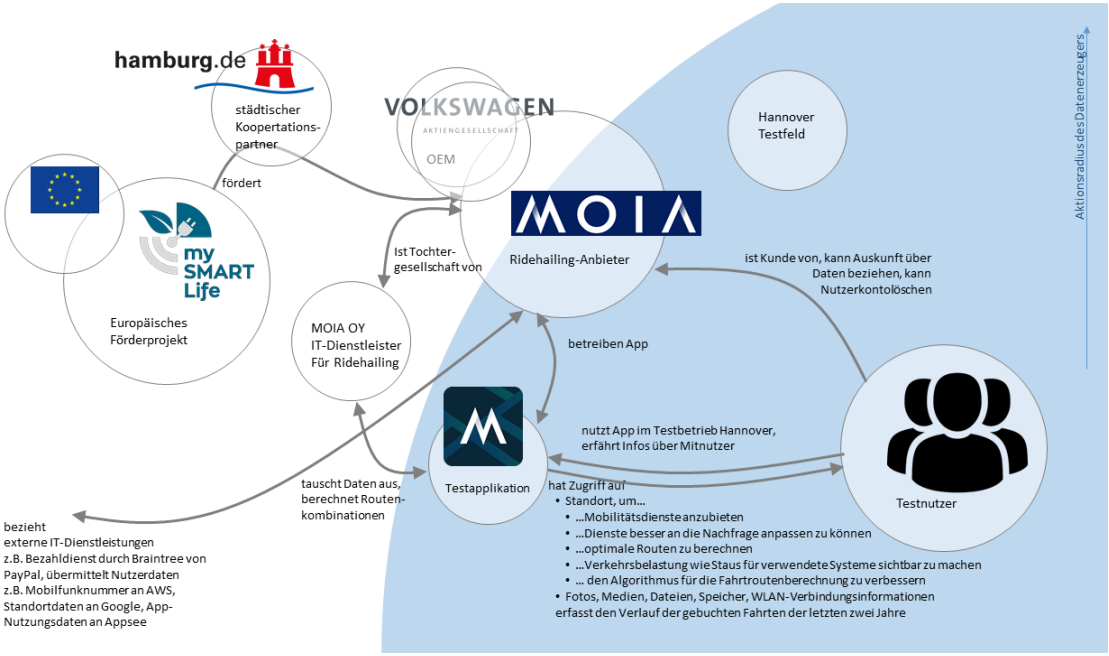
Integrierte Mobilitätsdienstleistungen

Daneben finden sich auch IT-Dienstleister, die den vielfältigen Anbietermarkt mit kundenspezifischen MaaS-Diensten bedienen. Die österreichische Fluidtime Data Services GmbH richtete zum Beispiel dem Aachener Verkehrsverbund eine Applikation für eine integrierte Mobilitätsdienstleistung ein. Die Verarbeitung von großen Datenmengen erfolgt über die Echtzeiteinbindung von Systeminformationen über eine cloudbasierte Mobilitätsplattform. Die App avv connect ermöglicht Fahrplanauskünfte und Routenplanung für Bus- und Bahnverbindungen in Nordrhein-Westfalen und Teilen von Belgien

20 https://storm.apache.org/Powered-By.html.
21 http://www.mobilityinside.de.

und den Niederlanden.[22] Ein weiterer MaaS-Anbieter, der aktuell im Feldversuch datengestützte Fahrtenbündelung erprobt, ist die neue Marke *MOIA* im Volkswagenkonzern (vgl. Abb. 9).

Abb. 9: Akteurskonstellation am Beispiel von MOIA



2.2.2.3 KUNDENKOMMUNIKATION UND MARKETING

Chatbots sind einer der am weitesten verbreiteten Kundendienste auf Basis von Big Data Technologien. Auch im ÖPV finden sich hierfür Beispiele wie die bahninterne Entwicklung assistify. Das Konzept sieht vor, Mitarbeiter bei der Kundenberatung mit künstlicher Intelligenz zu unterstützten. Dabei werden in Echtzeit die Chatkonversationen ausgelesen und ausgewertet. Hervorgegangen ist die Entwicklung aus einem Assistenten für die Reise, der Natural Language Processing bei der Kundenkonversation einsetzten sollte.[23]

Stomt ist ein weiterer Anbieter, der für Verkehrsunternehmen Feedback-Widgets erstellt. Die Zuordnung des Kundenfeedbacks erfolgt dabei auch über Algorithmen. Die Kundenangaben werden in Echtzeit mit Metadaten aus dem System ergänzt.[24]

Die Einbindung sozialer Netzwerke und der damit verbundenen Schwarmintelligenz für Innovationsprozesse ist Geschäftsgrundlage von PhantoMinds, das gleich mehrere Referenzen aus dem Mobilitätsbereich vorzuweisen hat. Im Auftrag von ÖPV-Anbietern wie der Hamburger Hochbahn, den BVG oder der DB AG nutzte Phantominds soziale Netzwerke als Ressource um Lösungsansätze zu spezifischen Problemstellungen zu erarbeiten.[25] Die datenbasierte Besonderheit

22 https://www.fluidtime.com/de/referenzen/avv-connect.
23 https://skydeck.deutschebahn.com/index.php/demoday/.
24 https://www.stomt.com/features.
25 http://PhantoMinds.com.

liegt darin, dass spezifische Experten für die Problematik automatisiert identifiziert und eingebunden werden können. Im Falle der Hochbahn konnten über 70 Lösungsansätze für passende Services für die Zielgruppe der 15- bis 30-Jährigen konzipiert werden.

2.2.3 VERWERTUNG DURCH DRITTE

Eine Kommerzialisierung von unternehmensintern erzeugten Daten, wie sie zum Beispiel durch Telekommunikationsanbieter vorgenommen wird, zeichnet sich im Bereich des ÖPV bisher nur rudimentär ab. Der Umfang an potentiell vermarktbaren Daten ist auf regionaler Ebene durchaus beachtlich. Hier stellt sich bei öffentlichen Betreibern und Verbünden jedoch die Frage nach dem Unternehmensauftrag und ob eine Querfinanzierung über Datenverkäufe dazu gehört. Trotz dieses Rahmens, finden sich deutliche Zeichen, dass eine Öffnung bestimmter Daten über Open Data Portale zur gängigen Strategie von ÖPNV-Betreibern werden könnte. Die Bereitstellung von Daten ist dabei nicht unbedingt eine Big Data-Praxis, bietet aber die Möglichkeit, entsprechende Anwendungsbereiche überhaupt erst zu identifizieren und damit eine Grundlage für Praxisanwendungen zu schaffen.

Ein Beispiel für die kommerzielle Verwertung eigener Daten ist das Corporate Start-up ZERO.ONE.DATA (ZOD) der DB AG. Mit der Gründung von ZOD wurde eine zentrale Einheit zur Förderung datengetriebener Arbeitsweisen geschaffen. ZOD hat u.a. den Auftrag, neue Umsatzpotentiale im Kontext von Big Data zu erschließen. Zunächst geht es darum, Know-how zu neuen Methoden der Datenanalyse im Hinblick auf bestehende Daten aufzubauen, neue Datenquellen zu erschließen und eine Digitalisierungsplattform mit Analysewerkzeugen bereitzustellen.

3. BETRACHTUNG DER NUTZERSEITE

Wie die Betrachtung der Anbieterseite zeigt, gibt es einige Big Data-Anwendungen, die auch im ÖPV eingesetzt werden bzw. eingesetzt werden könnten. Insbesondere bei kundenorientierten Angeboten ist es von zentraler Bedeutung, dass sich die Anbieter an den Wünschen und Bedürfnissen der Nutzer orientieren. Deshalb stellt Kapitel 3 die Ergebnisse einer repräsentativen Online-Befragung zur Einstellung von Nutzern gegenüber dem Einsatz von Big Data-basierten Technologien im ÖPV vor.

Nach einem Überblick zu relevanten Befragungen zu den Themenfeldern Mobilität und Big Data (Kapitel 3.1.), wird zunächst kurz das methodische Vorgehen vorgestellt (Kapitel 3.2). Hier werden die Konzeption des Fragebogens und die Durchführung der Befragung erläutert.

Der empirische Ergebnisteil (Kapitel 3.3) beginnt dann mit der Darstellung der Stichprobe (3.3.1) und den Daten zur Nutzung und Zufriedenheit mit dem ÖPV (3.3.2) – auch im Verhältnis zu anderen Verkehrsmitteln – bevor die größten Mängel des ÖPV aus Sicht der Befragten vorgestellt werden (3.3.3). Nach dieser grundsätzlichen Einordnung des ÖPVs wird auf die Bekanntheit und die Nutzung bestehender digitaler Angebote im ÖPV eingegangen (3.3.4). Dies vermittelt einen Eindruck, ob digitale Angebote bekannt sind und von welchen Personengruppen sie genutzt werden. Die Einstellung der Befragten zu bestimmten Spannungsfeldern wie z.B. zwischen Big Data-basierten Innovationen und Datenschutz wird anhand von Fragen zu Innovationen und Problemen untersucht (3.3.5). Darauffolgend wird untersucht, welche möglichen Big Data-Technologien im ÖPV in Zukunft erwartet werden und welche Ansätze einen Einfluss auf das Nutzungsverhalten haben könnten. Dazu werden die Ergebnisse zur Abfrage von zehn potentiellen Anwendungsszenarien von Big Data im ÖPV vorgestellt (3.3.6).

3.1 BISHERIGE ERHEBUNGEN ZUR NUTZERMEINUNG ÜBER BIG DATA IM ÖPV

Die Recherche ergab, dass es bisher noch keine öffentlich zugänglichen Erhebungen oder Befragungen über die Einstellungen der Nutzer zum Einsatz von spezifischen Big Data-basierten Anwendungen im ÖPV gibt. Insofern stellt die im Rahmen dieses Gutachtens durchgeführte Befragung ein Novum dar. Im Folgenden wird daher ein kurzer Überblick zu bestehenden Untersuchungen zur Nutzung und zu Einstellungen der Deutschen gegenüber dem ÖPV sowie ihrer Kenntnis und ihren Einstellungen bezüglich der Digitalisierung im Allgemeinen gegeben.

Das Nutzungsverhalten und die Einstellung der Nutzer gegenüber dem ÖPV in Deutschland ist ein gut erforschtes Feld. Wie viele andere Erhebungen zeigen auch die Ergebnisse des Deutschen Mobilitätspanels aus dem Jahr 2016, dass mit öffentlichen Verkehrsmitteln nach dem Auto die meisten Wege in Deutschland zurückgelegt werden. Eine Befragung von Research Now (2017) zeigt, dass vor allem junge Leute den ÖPV überproportional häufig nutzen. Trotz der häufigen Nutzung des ÖPV gibt es Aspekte, mit denen die Nutzer unzufrieden sind: Laut einer Studie des Verkehrsclubs Deutschland e.V. (VCD) aus dem Jahr 2015 sind die Nutzer des ÖPV zwar mit 10 von 13 abgefragten Kriterien des ÖPV eher zufrieden als unzufrieden, bei den Kriterien *Preis-Leistungsverhältnis*, *Verknüpfung mit dem Fahrrad* und *Verknüpfung mit dem Auto* überwiegt aber die Unzufriedenheit der Befragten. Es besteht also noch Handlungsbedarf bezüglich der Intermodalität und der Preisstruktur. Die bestehenden digitalen Angebote des ÖPV – Internetauftritte und Smartphone Apps – werden laut der Studie

von der Mehrheit der Befragten als positiv empfunden, es besteht also in der Breite keine Ablehnung gegen die Digitalisierung des ÖPV.

Die allgemeine Einstellung der Deutschen gegenüber der Digitalisierung analysiert die Studie *D21 – Digital Index 2017/2018*, die von Kantar TNS 2017 durchgeführt und von der Bundesregierung gefördert wird. Insgesamt steigt laut der Studie zwar der Digitalisierungsgrad, aber die Kenntnis von Fachbegriffen der Digitalisierung hängt stark von Alter, Bildungsniveau und auch Geschlecht ab. Den Begriff *Big Data* kennt nur jeder fünfte Befragte bzw. kann ihn einigermaßen schlüssig erklären. In Bezug auf die Mobilität wird deutlich, dass digitale Dienstleistungen hier signifikant häufiger von jüngeren und höher gebildeten Menschen genutzt werden. Zudem zeigt die Studie auch eine Skepsis der Befragten gegenüber der Nutzung und Weitergabe von Daten – einer Voraussetzung für viele Anwendungsfälle von Big Data – auf. Nur 47% der Befragten stimmen der Aussage „Es ist für mich in Ordnung, wenn meine persönlichen Daten anonymisiert ausgewertet werden, damit die Gesellschaft davon profitiert" zu. Die Zustimmung sinkt noch weiter, wenn auf Basis von Daten zum Nutzungsverhalten individuelle Tarife angeboten werden (26%) oder Anbieter von Dienstleistungen als Gegenleistung für den Service Zugriff auf persönliche Daten verlangen (12%).

3.2 KONZEPTION DER BEFRAGUNG

Diese Analyse bestehender Erhebungen macht verschiedene relevante Punkte für die Befragung im Kontext dieser Studie deutlich: Erstens gibt es noch keine dezidierten Befragungen zum Einsatz von Big Data-Anwendungen im ÖPV, weshalb die Umfrage im Rahmen dieses Gutachtens einen ersten Beitrag darstellt, diese Blackbox zu öffnen. Zweitens gibt es im ÖPV ein großes Potential für den Einsatz von Big Data-Technologien, da schon viele Menschen den ÖPV nutzen und digitalen Innovationen prinzipiell positiv gegenüberstehen, aber vor allem mit dem Preis und der Intermodalität unzufrieden sind. Vor allem bei letzterem Punkt könnte Big Data den ÖPV voranbringen, wobei der Datenschutz immer berücksichtigt werden sollte. Eine weitere wichtige Erkenntnis für die Konzeption der Befragung war zudem, dass die fehlende Kenntnis der Nutzer über technische Begriffe wie *Big Data* berücksichtigt und der Fragebogen verständlich formuliert werden muss.

Um die Bedürfnisse und Wünsche der Nutzer an den zukünftigen Einsatz von Big Data im ÖPV zielgenau untersuchen zu können, wurde ein mehrstufiger Ansatz für die Befragung gewählt: Zunächst wurden grundlegende Themenblöcke und Schwerpunkte für einen Fragebogen entworfen. Dann wurden drei semi-strukturierte Fachgespräche mit Mobilitätsexperten geführt, um relevante Fragestellungen für den Fragebogen zu ermitteln. Der Fragebogen wurde dann innerhalb des Projektteams in mehreren Schleifen weiterentwickelt, im Rahmen von Pre-Tests auf Verständlichkeit geprüft und mit den Auftraggebern abgestimmt, bevor mit der Befragung begonnen wurde. Durch diesen mehrstufigen Validierungsansatz sollte ein umfassender, strukturierter und verständlicher Fragebogen ermöglicht werden.

3.2.1 INTERVIEWS MIT VERTRETERN VON FAHRGASTVERBÄNDEN UND ANBIETERN ALS GRUNDLAGE FÜR DEN FRAGEBOGEN

Um sicherzustellen, dass der Fragebogen, die aus Nutzersicht relevanten Themenbereiche abdeckt, wurden drei Fachgespräche mit zwei Vertretern der ÖPV-Nutzer (dem Fahrgastverband PROBAHN

sowie dem Verbraucherzentrale Bundesverband) und einem Vertreter eines ÖPV-Anbieters (dem Verkehrsverbund Berlin-Brandenburg) geführt. Diese Gespräche haben gezeigt, dass ein nutzerorientierte Fragebogen zumindest die folgenden drei Themenbereiche adressieren sollte:

- Kenntnis und Nutzung bestehender digitaler Anwendungen im Bereich Mobilität
- Spannungsverhältnis zwischen Innovation und (Daten-) Sicherheitsbedürfnis
- Szenarien zur potentiellen Anwendung von Big Data-basierten Technologien im ÖPV

Der Leitfaden zu den Interviews sowie eine kurze Zusammenfassung der Erkenntnisse aus den Interviews sind in Anhang 2 und 3 zu diesem Bericht zu finden.

3.2.2 AUFBAU DES FRAGEBOGENS UND DURCHFÜHRUNG DER BEFRAGUNG

Der finale Fragebogen bestand aus den folgenden Teilen:

- Öffentlichen Personenverkehr – Nutzung und Bewertung
- Öffentlichen Personenverkehr – Eigenschaften (Customer Journey)
- Bestehende digitale Angebote – Wissen und Nutzung
- Einstellungen zu möglichen Problemen durch Big Data
- Szenarien zur Anwendung von Big Data – Bedeutung im ÖPV und Einfluss auf die Nutzung

Der Fragebogen befindet sich in voller Länge in Anhang 4 zu diesem Bericht. Die Befragung wurde vom 23.11.2017 bis 30.11.2017 durch das Marktforschungsinstitut Omniquest durchgeführt. Die erhobenen Daten wurden darauffolgend von ConPolicy analysiert.

3.3 ERGEBNISSE DER BEFRAGUNG

Im Folgenden werden die zentralen Ergebnisse der Befragung vorgestellt. Dabei wird nach einer kurzen Vorstellung der Stichprobe zunächst auf das Nutzungsverhalten und die Zufriedenheit der Befragten mit dem ÖPV als Mobilitätsalternative eingegangen. Danach werden Bekanntheit und Nutzung bestehender digitaler Angebote des ÖPV sowie die Einstellung zu Bedenken gegenüber Big Data-basierten Innovationen analysiert. Abschließend werden die Einschätzung der Bedeutung von zehn Szenarien zu Big Data-basierten Innovationen sowie die zu erwartende Auswirkung auf das Nutzungsverhalten erörtert.[26]

3.3.1 VERTEILUNG DER STICHPROBE

Die Stichprobe ist für die deutsche Bevölkerung online-repräsentativ bezüglich Alter und Geschlecht. Zudem wurde der Bildungsstand der Befragten erhoben und ob sie in ländlich oder städtisch geprägten Gebieten wohnen. Tabelle 4 stellt die Verteilung der Stichprobe dar.

26 Es werden die Ergebnisse für die gesamte Stichprobe jeweils in einer Tabelle präsentiert. Zudem werden ausgewählte Tabellen (z.B. signifikante Unterschiede bzgl. Geschlecht, Alter, Bildungsniveau oder Wohnort) im Text dargestellt. Weitere Tabellen sind im Anhang zu finden.

Tab. 4: Verteilung der Stichprobe

Geschlecht	Anzahl befragte Personen	Anteil in %
weiblich	513	51,30 %
männlich	487	48,70 %
Alter		
0-29 Jahre	163	16,30 %
30-59 Jahre	515	51,50 %
60+ Jahre	322	32,20 %
Bildung		
Abitur	398	39,80 %
kein Abitur	602	60,20 %
Städtegröße		
unter 20.000 Einw.	121	12,10 %
20.000 bis unter 50.000 Einw.	95	9,50 %
50.000 bis unter 100.000 Einw.	84	8,40 %
100.000 bis unter 500.000 Einw. (Verdichtungs - & Übergangsbereich)	140	14,00 %
100.000 bis unter 500.000 Einw. (Kernbereich)	167	16,70 %
500.000+ Einw. (Verdichtungs- & Übergangsbereich)	82	8,20 %
500.000+ Einw. (Kernbereich)	311	31,10 %
Total	1000	100,00 %

3.3.2 ÖPV WIRD GENUTZT, ABER NICHT GELIEBT – JUNGE, GEBILDETE GROßSTÄDTER NUTZEN ÖPNV AM HÄUFIGSTEN

Nach dem privaten PKW stellt der traditionelle ÖPNV das am meisten genutzte Verkehrsmittel in Deutschland dar (Siehe Tabelle 5). So nutzt die Hälfte der Befragten den ÖPNV mindestens monatlich und fast 30 % der Befragten sogar mindestens einmal in der Woche. Der öffentliche Fernverkehr in Form von Bahn und Bus wird nicht ganz so häufig genutzt, jedoch nutzen immerhin 65 % der Befragten zumindest gelegentlich die Bahn und fast die Hälfte Fernbusse. Sharing-Angebote werden weiterhin nur von einer geringen Anzahl an Early-Adoptern genutzt. Diese Ergebnisse stehen in Einklang mit den bestehenden Studien.

Tab. 5: Nutzung von Verkehrsmitteln (geordnet nach häufigster Nutzung in Prozent)

Verkehrsmittel (N=1000)	(fast) täglich	1-3 Tage pro Woche	1-3 Tage pro Monat	seltener als monatlich	nie
Privater PKW	56,2	24	5,2	3	11,6
ÖPNV	14,4	14,5	20,1	31,2	19,8
Eigenes Fahrrad	13,7	20,1	19,4	17,1	29,7
Fernverkehr Bus	3,1	3,4	7,6	32,9	53
Fernverkehr Bahn	2,5	4,4	11,8	46,1	35,2
Motorrad etc.	1,5	4,4	6	8,2	79,9
Bike-Sharing	1,1	1,5	3,3	6,1	88

Fortsetzung Tabelle 5: Verkehrsmittel (N=1000)	(fast) täglich	1-3 Tage pro Woche	1-3 Tage pro Monat	seltener als monatlich	nie
Roller-Sharing	0,9	1,3	2	4,9	90,9
Car-Sharing	0,8	1,9	4,9	9,6	82,8
Taxi	0,4	1,8	10,1	54	33,7

Der häufigen Nutzung des ÖPNV steht jedoch eine vergleichsweise geringere Zufriedenheit der Befragten mit den Angeboten des öffentlichen Personenverkehrs im Vergleich mit anderen Verkehrsmitteln gegenüber. Während 66 % der PKW-Nutzer sehr zufrieden mit dem Auto sind, geben dies im traditionellen ÖPV nur 10-12 % der Nutzer an. Insgesamt sind jedoch auch hier die Zufriedenheitswerte gut und weniger als 20 % sind unzufrieden mit den Angeboten des traditionellen ÖPV, wobei kein Unterschied zwischen Nah- und Fernverkehr besteht.

Tab. 6: Zufriedenheit mit Verkehrsmitteln (geordnet nach Zufriedenheit in Prozent)

Verkehrsmittel	Sehr unzufrieden	2	3	4	Sehr zufrieden	N=
Privater PKW	2,6	2,6	6	23,2	66,1	884
Eigenes Fahrrad	3,3	5,7	15,5	28,2	47,4	703
Motorrad etc.	3	7	30,3	28,4	31,3	201
Bike-Sharing	4,2	7,5	36,7	30,8	20,8	120
Car-Sharing	4,1	7	40,7	29,1	19,2	172
Taxi	2,9	12,2	32,9	34,2	17,8	663
Roller-Sharing	7,7	8,8	37,4	29,7	16,5	91
ÖPNV	5,9	12,6	36	33,2	12,3	802
Fernverkehr Bus	5,5	11,7	39,1	33	10,6	470
Fernverkehr Bahn	5,7	11	37,3	35,3	10,6	648

Soziodemographische Unterschiede

In Einklang mit bestehenden Erkenntnissen nutzen Personen unter 30 Jahren den traditionellen öffentlichen Nah- und Fernverkehr deutlich häufiger als ältere Personen (siehe Tabelle A1-A3 im Anhang). Zudem gibt es auch einen Unterschied im Nutzungsverhalten zwischen Personen mit und ohne Abitur, da insbesondere der ÖPNV deutlich häufiger von Personen mit Abitur genutzt wird (siehe Tabelle 7). Erwartungsgemäß nutzen Bewohner von Großstädten insbesondere in deren Kernbereich, also den Innenstädten, den ÖPNV deutlich häufiger als Bewohner aus Städten mit weniger als 100.000 Einwohnern wie Tabelle 8 zeigt. Bei der Nutzung von Fernbussen oder Fernzügen besteht hingegen kein signifikanter Unterschied zwischen Großstädten und anderen Stadtgrößen (Siehe Tabellen A6 und A7 im Anhang).

Tab. 7: Nutzung des ÖPNV (geordnet nach Bildungsabschluss)

Bildung (N=1000)	(fast) täglich	1-3 Tage /Woche	1-3 Tage /Monat	seltener als monatlich	nie
kein Abitur	11,1	12,3	17,6	33,6	25,4
Abitur	19,3	17,8	23,9	27,6	11,3

Tab. 8: Nutzung des ÖPNV (geordnet nach Stadtgröße)

Städtegröße (N=1000)	(fast) täglich	1-3 Tage /Woche	1-3 Tage /Monat	seltener als monatlich	nie
bis unter 20.000 Einwohner	8,3	4,1	11,6	38,8	37,2
20.000 50.000 Einwohner	5,3	5,3	15,8	31,6	42,1
50.000 100.000 Einwohner (Kern, Verdichtungs- und Übergangsbereich)	7,1	10,7	17,9	38,1	26,2
100.000 - 500.000 Einwohner (Verdichtungs- und Übergangsbereich)	5	9,3	19,3	38,6	27,9
100.000 - 500.000 Einwohner (Kernbereich)	18	12,6	27,5	31,1	10,8
Mehr als 500.000+ Einwohner (Verdichtungs- und Übergangsbereich)	15,9	9,8	18,3	39	17,1
Mehr als 500.000+ Einwohner (Kernbereich)	23,5	27	22,2	20,9	6,4

3.3.3 CUSTOMER JOURNEY – TICKETPREISE UND SERVICE BEI BESCHWERDEN IM FOKUS DER KRITIK

Wie gezeigt wird der ÖPV häufig genutzt (Tabelle 5), aber die Zufriedenheit ist – sowohl bei den Angeboten im Nah- als auch im Fernverkehr – noch verbesserungsfähig (Tabelle 6). Diese Diskrepanz deutet darauf hin, dass viele Befragte auf den ÖPV angewiesen sind, auch wenn sie nicht sehr zufrieden mit bestimmten Aspekten des ÖPV zu sein scheinen. Dieses Unterkapitel beleuchtet deshalb die Zufriedenheit mit verschiedenen Bereichen einer Customer Journey im ÖPV.

Die Ergebnisse in Tabelle 9 zeigen, dass vor allem der Preis für Tickets – nur knapp 25 % der Befragten sind damit zufrieden – sowie die Bearbeitungsdauer und der Umgang mit Rückerstattungen für Unzufriedenheit sorgen. Beim Kauf von Tickets, der Erreichbarkeit von Haltestellen, dem Sicherheitsgefühl, den Fahrplaninformationen, der Anzahl der Sitzplätze und der Funktionstüchtigkeit der Einrichtung überwiegt die Anzahl an zufriedenen Kunden die Anzahl der Unzufriedenen hingegen deutlich.

Tab. 9: Zufriedenheit mit Aspekten des ÖPV (geordnet nach Unzufriedenheit in Prozent)

Aspekt	Sehr unzufrieden	2	3	4	Sehr zufrieden
Preis für Tickets	24,6	23	27,6	17,1	7,6
Bearbeitungsdauer von Rückerstattungen	18,4	19,6	31,3	18,8	11,8
Umgang mit Rückerstattungsforderungen	17,5	19,7	31,1	20,1	11,6
Information bei Verspätungen	17,4	21,8	31	20,8	8,9
Beschwerdemöglichkeiten	14,7	19,2	30,1	24,3	11,6
Häufigkeit der Abfahrten/Taktung	10,8	15,4	31,7	29,2	12,8
Zustand der Haltestellen (Lifte, Bänke etc.)	9,4	16,4	33,8	28,2	12,1
Anzahl der Anschlussmöglichkeiten	8,8	16,4	33,7	29,3	11,7
Einfachheit des Ticketkaufs	8,7	14,4	27	29,2	20,6
Umsteigezeiten bei den Anschlussmöglichkeiten	8,5	16,6	34	29	12
Sauberkeit	7,8	14,8	37,9	27,8	11,8
Erreichbarkeit von Haltestellen	7,2	11,3	27,8	33,7	20
Sicherheitsgefühl	7,1	11,8	25,9	35,2	20
Fahrplaninformationen	6,9	10,9	28,2	35,2	18,7
Anzahl der Sitzplätze im Verkehrsmittel	4,7	11,3	28,7	38	17,4
Funktionstüchtigkeit der Einrichtung in den Fahrzeugen	4,6	8,9	30,2	39,1	17

Soziodemographische Unterschiede

Während mehr als die Hälfte der Frauen mit dem Preis des ÖPNV unzufrieden sind, sind es bei den Männern nur knapp über 40 % (Tabelle A8). Auch mit den Aspekten *Funktionstüchtigkeit*, *Sauberkeit*, *Umsteigezeiten* und *Information bei Verspätungen* sind Frauen signifikant weniger zufrieden als Männer (Tabellen A9-A12). Personen über 60 Jahren sind zudem signifikant unzufriedener mit dem Umgang und der Bearbeitungsdauer von Rückerstattungen (Tabellen A13 und A14). Zwischen städtischen und ländlichen Räumen gibt es signifikante Unterschiede in der Zufriedenheit bezüglich der Taktung wie Tabelle 10 zeigt. In Städten unter 50.000 Einwohnern sind nur ca. 30 % der Einwohner zufrieden mit der Häufigkeit der Abfahrten während es in größeren Städten in jeder Kategorie 40 % und mehr sind. Der gleiche Effekt ist bei den Aspekten *Umsteigezeit* und *Anschlussmöglichkeiten* zu beobachten (Tabellen A15 und A16).

Tab. 10: Zufriedenheit mit der Häufigkeit der Abfahrten/Taktung des ÖPV (geordnet nach Stadtgröße in Prozent)

Stadtgröße (N=914)	sehr unzufrieden	2	3	4	sehr zufrieden
bis unter 20.000 Einwohner	15,7	26,9	26,9	20,4	10,2
20.000 bis unter 50.000 Einwohner	9	20,5	41	21,8	7,7
50.000 bis unter 100.000 Einwohner (Kern, Verdichtungs- und Übergangsbereich)	11	12,3	24,7	37	15,1
100.000 - unter 500.000 Einwohner (Verdichtungs- und Übergangsbereich)	16	10,9	31,9	31,1	10,1
100.000 - unter 500.000 Einwohner (Kernbereich)	7,7	14,1	31,4	35,3	11,5
500.000+ Einwohner (Verdichtungs- und Übergangsbereich)	13,3	22,7	21,3	32	10,7
500.000+ Einwohner (Kernbereich)	8,5	11,5	35,4	27,9	16,7

3.3.4 BESTEHENDE DIGITALE ANGEBOTE – ROUTENPLANUNG UND ONLINE-TICKETS SIND BEKANNT UND WERDEN GENUTZT

Es gibt schon einige digitale Angebote im traditionellen ÖPV, die auf dem Einsatz von großen Datenquellen basieren, jedoch noch nicht allen Kriterien von Big Data entsprechen. Hier sind insbesondere Routenplaner-Apps, Online-Buchungen und WLAN-Zugang durch die Anbieter unter den Befragten bekannt, wie aus Tabelle 11 deutlich wird. Dabei werden die Angebote von einer Mehrheit auch genutzt, wenn sie denn bekannt sind. Schlusslicht bildet hierbei das kontaktlose Zahlen, das in Deutschland noch nicht sehr verbreitet ist: Nur circa jeder Dritte, der es kennt, nutzt es auch.

Tab. 11: Bekanntheit und Nutzung bestimmter digitaler Angebote des ÖPV (geordnet nach Bekanntheit in Prozent)

Digitales Angebot	Bekannt (in %) (N=1000)	Nutzung (in %, wenn bekannt)
Routenplanung über Websites oder Apps	73,3	79,7 (von 733)
Buchung von Online-Tickets über Websites oder Apps der Anbieter des öffentlichen Personenverkehrs	63,9	56 (von 639)
WLAN in Verkehrsmitteln oder Haltestellen von Anbietern des öffentlichen Personenverkehrs	53,2	59,4 (von 532)

(Fortsetzung Tab.11) Digitales Angebot	Bekannt (in %) (N=1000)	Nutzung (in %, wenn bekannt)
Suche nach Sparangeboten für Bahn- oder Busreisen auf digitalen Plattformen von Drittanbietern (my-train.de, Sparpreisfinder.de, Itur.com etc.)	46,6	68,9 (von 466)
Online-Auftritt, Twitter-Auftritt, Facebook-Auftritt etc. von Anbietern des öffentlichen Personenverkehrs	39,5	60,5 (von 395)
Kontaktloses Bezahlen (NFC) in Fahrzeugen z.B. mit der Kreditkarte oder dem Handy	38,1	30,7 (von 381)

Soziodemographische Unterschiede

Routenplaner-Apps und digitale Sparangebotsplattformen werden sowohl von Männern als auch Frauen häufig genutzt, während der Online-Auftritt und das WLAN-Angebot der Anbieter sowie Onlinebuchung von Tickets signifikant häufiger von Männern genutzt werden wie Tabelle 12 zeigt.

Tab. 12: Unterschiede nach Geschlecht in der Nutzung bestimmter digitaler Angeboten des ÖPV (geordnet nach Nutzung in Prozent)

Digitales Angebot	Frauen	Männer
Routenplanung über Websites oder Apps	79,7% (von 369)	79,7% (von 364)
Suche nach Sparangeboten für Bahn- oder Busreisen auf digitalen Plattformen von Drittanbietern (mytrain.de, Sparpreisfinder.de, Itur.com etc.)	68,8% (von 237)	69% (von 229)
Online-Auftritt, Twitter-Auftritt, Facebook-Auftritt etc. von Anbietern des öffentlichen Personenverkehrs	54% (von 187)	66,3% (von 208)
WLAN in Verkehrsmitteln oder Haltestellen von Anbietern des öffentlichen Personenverkehrs	53,8% (von 260)	64,7% (von 272)
Buchung von Online-Tickets über Websites oder Apps der Anbieter des öffentlichen Personenverkehrs	50% (von 322)	62,1% (von 317)
Kontaktloses Bezahlen (NFC) in Fahrzeugen z.B. mit der Kreditkarte oder dem Handy	28% (von 182)	33,2% (von 199)

Wie zu erwarten nutzen jüngere Personen die meisten Angebote häufiger (Tabelle A17). Außerdem geht ein höheres Bildungsniveau mit einer höheren Kenntnis und Nutzung digitaler Angebote einher (Tabellen A18 und A19). Die Stadtgröße hat keinen eindeutigen Einfluss auf die Kenntnisse und Nutzung der digitalen Angebote. Dies ist angesichts der deutlich höheren Nutzung des ÖPNV in Städten erstaunlich.

3.3.5 INNOVATIONEN IM ÖPV – BEFRAGTE SIND RISIKOAVERS BEI BIG DATA-BASIERTEN NEUERUNGEN

Häufig gibt es ein Spannungsfeld zwischen dem Wunsch nach Innovationen im ÖPV und dem Sicherheitsbedürfnis bei der Nutzung. Bei Big Data-basierten Innovationen im Bereich des ÖPV scheint die Mehrheit der Befragten Risiken durch Innovationen nicht eingehen zu wollen.[27] Wie Tabelle 13 zeigt, sollten Big Data-basierte Innovationen nach Ansicht der Befragten nicht eingesetzt werden, wenn sie mit einer Ausgrenzung bestimmter Nutzergruppen, Risiken für die Datensicherheit, mit einer Weitergabe von Daten sowie höheren Preise einhergehen. Auch beim Einsatz unausgereifter Innovationen ist die Zustimmung geringer als die Ablehnung.

Tab. 13 Risiken und Vorteile von Big Data im ÖPV im Vergleich (geordnet nach „sehr große Zustimmung“ in Prozent)

Aussage (N=1000)	Sehr große Ablehnung	2	3	4	sehr große Zustimmung
Eine Benachteiligung bestimmter Nutzergruppen (z.B. ältere Personen oder Personen ohne Internetzugang) muss ausgeschlossen werden, wenn Innovationen durch Big Data eingeführt werden.	10,3	7,1	20,3	23,2	39,1
Mir ist es wichtig, dass Innovationen durch Big Data möglichst bald im öffentlichen Personenverkehr eingesetzt werden. Dafür würde ich notfalls auch in Kauf nehmen, dass diese Innovationen noch nicht komplett ausgereift und zuverlässig sind.	24,9	17,2	31,4	18,2	8,3
Beim Einsatz von Big Data für Innovationen im öffentlichen Personenverkehr müssen Nutzer bestimmte Risiken bei der Datenübertragung und Datensicherheit in Kauf nehmen.	32,6	18,3	26,3	15,2	7,6
Es ist in Ordnung, dass personenbezogene Daten von den Betreibern des öffentlichen Personenverkehrs verarbeitet und ggf. an Dritte weitergegeben werden, wenn dafür Innovationen durch Big Data im öffentlichen Personenverkehr ermöglicht werden können.	31,7	18,3	28,6	15,5	5,9
Ich bin bereit, mehr Geld für den öffentlichen Personenverkehr zu zahlen, wenn dadurch Innovationen durch Big Data ermöglicht werden können.	40,7	18,5	23,8	12,4	4,6

Soziodemographische Unterschiede

Der Einsatz unausgereifter Innovationen und die Verarbeitung bzw. Weitergabe personenbezogener Daten wird von Frauen deutlich häufiger abgelehnt als von Männern (Tabellen A20 und A21). Alle

[27] Auch wenn man die Fragestellung umgekehrt formuliert (siehe Item zu Benachteiligung sozialer Gruppen) – also das Framing ändert – ist eine starke Präferenz für die Sicherheitsaspekte zu beobachten.

Altersgruppen stellen das Sicherheitsbedürfnis über mögliche Vorteile von Big Data-basierten Innovationen, wobei diese Einstellung mit steigendem Alter zunimmt (Tabellen A22-A26). Personen mit Abitur stimmen den Aussagen zwar signifikant häufiger zu, jedoch ist auch unter ihnen die Skepsis sehr ausgeprägt (Tabellen A27-A31). Bezüglich der Stadtgröße konnten keine strukturellen Unterschiede ermittelt werden.

3.3.6 ZUKUNFT VON BIG DATA IM ÖPV – BEFRAGTE WÜNSCHEN FLEXIBLERES ANGEBOT UND EINFACHERE RÜCKERSTATTUNGSMÖGLICHKEITEN

Nach der Analyse des derzeitigen Nutzungsverhaltens und der allgemeinen Einstellung zu Big Data-basierten Innovationen, stellt sich fast zwangsläufig die Frage, welche zukünftigen Innovationen die Befragten für wichtig erachten und ob diese einen Einfluss auf ihre Nutzungshäufigkeit haben könnten.

Geschätzte Bedeutung der Szenarien für den ÖPV

Die Bedeutung von zehn Zukunftsszenarien für den Einsatz von Big Data aus Sicht der Befragten wird in Tabelle 14 dargestellt. Die geringe Zufriedenheit mit den Beschwerdemechanismen und Rückerstattungsmöglichkeiten bei traditionellen ÖPV-Anbietern spiegelt sich hier wider: Knapp 69% der Befragten schätzen eine automatische Rückerstattung ohne Antragsstellung als überdurchschnittlich wichtig ein. Ebenfalls sehr wichtig finden die Befragten eine Anpassung der eingesetzten Fahrzeuge des ÖPV an die Nachfrage, während zusätzlicher Annehmlichkeiten wie spezielle Sitzplatzreservierungen oder ein personalisiertes Entertainment nur von einer Minderheit als zukünftig wichtig erachtet werden.

Tab. 14: Geschätzte Bedeutung der Szenarien zu Big Data im ÖPV (geordnet nach Bedeutung „sehr wichtig“ in Prozent)

Szenario: Big Data wird dafür genutzt, … (N=1000)	Sehr unwichtig	2	3	4	Sehr wichtig
dass den Fahrgästen im öffentlichen Personenverkehr bei entsprechend großen Verspätungen automatisch Rückerstattungen oder Entschädigungen auf ihr Konto überwiesen werden, ohne dass es dafür einer Antragsstellung bedarf.	5,9	4,9	20,6	31,7	36,9
die Anzahl an eingesetzten Fahrzeugen im öffentlichen Personenverkehr in Echtzeit flexibel an das Fahrgastaufkommen anpassen.	6,4	6,3	21,7	30,8	34,8
die Funktionsfähigkeit der Ausstattung von Bussen und Bahnen durch eine datenbasierte Optimierung der Wartung zu verbessern.	7,9	7,5	27,7	31	25,9
Routen von Buslinien flexibel an die Nachfrage an Haltestellen anzupassen.	9,8	9,1	26,3	31,2	23,6

(Fortsetzung Tab.14) **Szenario: Big Data wird dafür genutzt, …** (N=1000)	**Sehr unwichtig**	**2**	**3**	**4**	**Sehr wichtig**
Fahrgästen eine individuelle Reiseplanung ermöglichen.	9,1	7,7	32,1	31	20,1
mit einer einzigen App alle für die Reiseroute erforderlichen Verkehrsmittel (z.B. Bus, Bahn, Carsharing, Bikesharing) buchen und bezahlen zu können.	12,4	8,8	30,4	28,9	19,5
je nach Fahrgast und seiner individuellen Zahlungsbereitschaft unterschiedliche Preise für eine Fahrt anzubieten.	24,6	12,2	25,2	22	16
an das aktuelle Fahrgastaufkommen angepasste Preise im öffentlichen Personenverkehr anzubieten.	27,2	15,3	30	16,3	11,2
Fahrgast-Gruppen mit unterschiedlichen Bedürfnissen bei der Reservierung von Sitzplätzen zu trennen.	22,3	13,9	32,4	21,2	10,2
den Fahrgästen ein personalisiertes Entertainment- und Informationsangebot im WLAN des öffentlichen Personenverkehrs anzubieten.	27,3	16,1	29,2	17,6	9,8

Soziodemographische Unterschiede bezüglich der Bedeutung der Szenarien

Insgesamt wird die Bedeutung der Szenarien von älteren Personen als geringer eingeschätzt. Insbesondere das Entertainment-Angebot finden über 40 % der unter 30-Jährigen wichtig, aber nur knapp 16 % der über 60-Jährigen (Tabelle A32). Ein ähnlich stark ausgeprägter Unterschied zwischen Alt und Jung ist auch beim Szenario zur zielgenauen Sitzplatzreservierung zu finden (Tabelle A33). Zudem schätzen Personen mit Abitur einige Szenarien wichtiger ein als Personen mit niedrigeren Bildungsabschlüssen. Der größte Unterschied besteht hier beim Szenario zur Anpassung der Anzahl an Fahrzeugen an die Nachfrage in Echtzeit (Tabelle A34). Interessanterweise sind bezüglich der Stadtgröße zwar einzelne Auffälligkeiten, aber keine strukturellen Unterschiede zu beobachten. Personen aus Städten mit 20.000-50.000 Einwohnern (38 %) finden die Einführung eines personalisierten Entertainment- und Informationsangebots im WLAN des ÖPV deutlich wichtiger als Personen aus dem Verdichtungs-und Übergangsbereich von Städten mit mehr als 500.000 Einwohner (20 %) (Tabelle A35). Zwischen den Geschlechtern gibt es hier keine signifikanten Unterschiede.

Potentielle Nutzungsänderungen des ÖPV bei Realisierung der Szenarien

Eine Realisierung der als am wichtigsten erachteten Szenarien hätte auch den größten positiven Einfluss auf das Nutzungsverhalten der Befragten. So geben 47 % der Befragten an, den ÖPV bei einer automatischen Rückerstattung des Fahrpreises bei entsprechend großen Verspätungen häufiger nutzen zu wollen (Tabelle 15). Eine Anpassung des Angebots in Echtzeit an das Fahrgastaufkommen wird ebenfalls sehr positiv gesehen und könnte die Anzahl an potentiellen Fahrgästen erhöhen. Eine zeitlich bedingte Anpassung des Fahrpreises an das Fahrgastaufkommen wird von den Befragten hingegen als sehr kritisch angesehen und könnte einen Rückgang in den Fahrgastzahlen bewirken.

Auffällig ist, dass Personen ohne beabsichtigte Nutzungsänderung – Ausprägung „3“ – die größte Gruppe bei allen Szenarien darstellen. Das weist darauf hin, dass bestimmte zukünftige Szenarien von

den Befragten zwar als wichtig für die Gemeinschaft erachtet werden, jedoch ein großer Teil der Befragten ihr eigenes Nutzungsverhalten auch bei einer erfolgreichen Einführung von bestimmter Big Data-basierter Innovationen nicht ändern würde.

Tab. 15: Geschätzte Nutzungsänderung bei Realisierung der Szenarien zu Big Data im ÖPV (geordnet nach zunehmender Nutzung in Prozent)

Szenario: Big Data wird dafür genutzt, … (N=1000)	Geringere Nutzung	2	3	4	Häufigere Nutzung
dass den Fahrgästen im öffentlichen Personenverkehr bei entsprechend großen Verspätungen automatisch Rückerstattungen oder Entschädigungen auf ihr Konto überwiesen werden, ohne dass es dafür einer Antragsstellung bedarf.	7,3	5,9	40,1	27,1	19,6
die Anzahl an eingesetzten Fahrzeugen im öffentlichen Personenverkehr in Echtzeit flexibel an das Fahrgastaufkommen anpassen.	8,6	6,6	43,3	24,9	16,6
Routen von Buslinien flexibel an die Nachfrage an Haltestellen anzupassen.	10,1	6,1	46,3	22,6	14,9
je nach Fahrgast und seiner individuellen Zahlungsbereitschaft unterschiedliche Preise für eine Fahrt anzubieten.	22,2	9,2	38,8	17,2	12,6
Fahrgästen eine individuelle Reiseplanung ermöglichen.	10,8	8,7	45,6	23	11,9
mit einer einzigen App alle für die Reiseroute erforderlichen Verkehrsmittel (z.B. Bus, Bahn, Carsharing, Bikesharing) buchen und bezahlen zu können.	11,4	10,1	45,5	21,4	11,6
die Funktionsfähigkeit der Ausstattung von Bussen und Bahnen durch eine datenbasierte Optimierung der Wartung zu verbessern.	9,8	7,8	51,4	19,9	11,1
den Fahrgästen ein personalisiertes Entertainment- und Informationsangebot im WLAN des öffentlichen Personenverkehrs anzubieten.	16,9	8,7	51,8	14,1	8,5
Fahrgast-Gruppen mit unterschiedlichen Bedürfnissen bei der Reservierung von Sitzplätzen zu trennen.	13,1	11,4	51,7	15,8	8
an das aktuelle Fahrgastaufkommen angepasste Preise im öffentlichen Personenverkehr anzubieten.	25,3	14,6	38,4	14,1	7,6

Soziodemographische Unterschiede bezüglich der Nutzungsänderungen durch die Szenarien

Alle Szenarien würden eher jüngere als ältere Befragte zu einer häufigeren Nutzung des ÖPV bewegen, wobei die Unterschiede bei personalisiertem Entertainment- und Informationsangebot und zielgenauer Sitzplatzreservierung besonders groß sind (Tabelle A36). Ein personalisiertes Entertainment- und Informationsangebot, flexiblere Buslinien und eine Anpassung der Preise an die Nachfrage in

Echtzeit würde zudem eher Männer als Frauen zu einer häufigeren Nutzung des ÖPV veranlassen (Tabelle A37-A39). Bei neun von zehn Szenarien würden Personen mit Abitur im Vergleich zu Personen ohne Abitur den ÖPNV bei der Umsetzung der Szenarien signifikant häufiger nutzen, lediglich beim Szenario *Angebot unterschiedliche Preise je nach individueller Zahlungsbereitschaft* gibt es keinen signifikanten Unterschied zwischen den beiden Gruppen. Wie schon bei der Einschätzung der Bedeutung der Szenarien ist auch bei der potentiellen Nutzungsänderung kein struktureller Unterschied zwischen den Städtegrößen zu erkennen. Es gibt bei 5 der 10 Szenarien zwar signifikante Unterschiede, jedoch hängt die Wirkungsrichtung stark von den spezifischen Szenarien ab. Das auffälligste Ergebnis besteht beim Szenario zu flexiblen Busrouten. 64 % der Befragten aus Städten mit weniger als 20.000 Einwohnern geben an bei einer Realisierung des Szenarios den ÖPV häufiger nutzen zu wollen während es in den anderen Größenklassen nur zwischen 33 % und 41 % sind (Tabelle A40).

4. REFLEXION DER ERGEBNISSE UND HANDLUNGSEMPFEHLUNGEN

In den ersten drei Kapiteln dieses Gutachtens wurden Big Data und öffentlicher Personenverkehr definiert und vorgestellt (Kapitel 1), potentielle Einsatzmöglichkeiten von Big Data im Bereich des ÖPV und anderen Verkehrsbereichen aufgezeigt (Kapitel 2), sowie die Einstellung und Akzeptanz der Bevölkerung gegenüber möglichen Big Data-basierten Innovationen im ÖPV untersucht (Kapitel 3). Im Folgenden werden die gewonnen Erkenntnisse in Thesen zusammengefasst und gesamtheitlich reflektiert. Überdies werden Handlungsempfehlungen abgeleitet.

These 1: Trotz zunehmender Digitalisierung existieren noch vergleichsweise wenige Big Data-Anwendungen im ÖPV – Nutzer des ÖPV sind mit bestehenden digitalen Angeboten vertraut

Während Fluggesellschaften, Carsharing-Anbieter, Taxi-Dienste oder App-Anbieter bereits auf Big Data-Verfahren setzen und im Realbetrieb verwenden, ergab unsere Analyse nur wenige hervorstechende Beispiele im klassischen ÖPV. Zwar konnte im Rahmen dieser Studie lediglich eine Desk-Recherche durchgeführt werden, dennoch deuten die fehlenden Literatur- und Internetquellen darauf hin, dass der öffentliche Personenverkehr eine Branche ist, die tendenziell eher als Nachzügler im Bereich der Big Data-Anwendungen eingestuft werden kann. Widersprüchlich ist dabei, dass die Akteure schon seit längerem in ihren komplexen technischen Systemen viele digitale Abläufe organisieren und auch die langfristige Planung mit IT-basierten Datenmodellen erfolgt, jedoch keine frühe Adaption der Entwicklungen im Bereich Big Data erfolgte. Vielmehr scheinen die Akteure des klassischen ÖPV in der Findungsphase bzgl. des Umgangs und der Anwendung der Technologie zu sein.

Für die weitere Entwicklung könnte sich der ÖPV an der Luftverkehrsbranche orientieren, die sich in einer deutlich weiter fortgeschrittenen Phase der Technologiediffusion befindet. Beispiele aus den Anwendungsfeldern Predictive Maintenance und Mobilitätsplattformen verdeutlichen dies. Auch im Betrieb von Großflughäfen werden umfassende Big Data-Lösungen eingesetzt. Die Übertragbarkeiten der im Flugverkehr erprobten Ansätze in andere Kontexte erfordert jedoch eine genaue Prüfung, welche Elemente übernommen und welche angepasst werden müssten. Dies umso mehr als deutliche Unterschiede zwischen der Luftfahrt und der ÖPV-Branche bestehen. Die ÖPV-Branche ist gekennzeichnet durch eine ausgeprägte Heterogenität der Unternehmen, die sich auch im Grad der Digitalisierung ausdrückt. Diese Unterschiede machen eine einfache Übertragbarkeit der Konzepte schwierig, da der Ansatzpunkt zunächst auf einen nachholenden Aufbau der IT-Infrastrukturen zielen sollte. Dabei ist jedoch zu empfehlen, bei den erforderlichen Investitionen von vornherein auf Mehrwertanwendungen zu achten, die Big Data-Ansätze mit sich bringen. Das bedeutet, dass auf Integrierbarkeit in und Anschlussfähigkeit an übergeordnete MaaS-Systeme sowie auf die Einhaltung zugehöriger Standards geachtet werden muss. Dabei sollte gleichermaßen an den Aufbau von entsprechenden Datenpunkten und Sensoren gedacht werden, die auch eine Predictive Maintenance ermöglichen. Ebenso ist in diesem Bereich zu empfehlen, dass über brancheninterne Kommunikationen Nachzügler von den Erfahrungen der Vorreiter profitieren und über ein branchenumfassendes Benchmarking Best-Practice-Beispiele identifiziert werden, die als Maßstab für wirtschaftliche Effizienzgewinne oder neue Geschäftsmodelle bzw. Vertriebsstrukturen dienen können.

Auf Seiten der Kunden scheint keine prinzipielle Ablehnung gegenüber neuen digitalen Technologien zu bestehen, wie die Ergebnisse der Befragung in Kapitel 3 zeigen. Innovative digitale Angebote für Kunden des ÖPV – z.B. Routenplaner-Apps, Online-Tickets oder WLAN in Verkehrsmitteln – sind

mehrheitlich bekannt und werden auch von vielen genutzt. Auch ältere Menschen kennen und nutzen diese Angebote.

Handlungsempfehlung 1: Die Digitalisierung des ÖPV ist sehr unterschiedlich fortgeschritten. Um die Möglichkeiten von Big Data-Anwendungen für die gesamte Branche nutzbar zu machen, sind entsprechende nachholende Investitionen und Entwicklung anzustoßen. Dabei sollten verfügbare Standards eingehalten werden. Ebenso sollten innerhalb der Branche Best-Practice-Beispiele über ein Benchmarking identifiziert und als Grundlage für weitergehende Investitionen genutzt werden.

These 2: Datenintegration und -bereitstellung sind die Grundlage für die Anwendung von Big Data im ÖPV

Eine der größten Herausforderungen beim Einsatz von Big Data-Methoden im ÖPV liegt naturgemäß in der Integration und Konsolidierung von Daten aus unterschiedlichen Quellen, von verschiedenen Providern und Sensoren. Über welche Daten wird verfügt? Welchen Wert haben diese Daten? Welche Innovationen lassen sich daraus ableiten? Gibt es neue Geschäftsmodelle oder Einnahmequellen und wenn ja, dürfen diese als öffentliches oder kommunales Unternehmen überhaupt exploriert werden? Diese Fragen sind weitgehend ungeklärt und ein gemeinsames Verständnis zum Umgang mit den in der Branche existierenden Datenbeständen ist erst im Entstehen. Wie mit geteilten (und damit zumindest teilweise frei verfügbaren) und heterogenen Daten umgegangen werden könnte, lässt sich im Bereich von Open-Data-Portalen ablesen. So veröffentlicht der Verkehrsverbund Rhein-Sieg (VRS) das ÖPNV-Liniennetz und die Sollfahrplandaten von Bus und Bahn über die Open-Data-Plattform der Stadt Bonn. Begleitet wird diese Entwicklung von wiederkehrenden Workshops oder Hackathons. Die Plattform basiert auf dem freien Open Data Framework dkan und ist somit prinzipiell auf Anwendungen in anderen Regionen übertragbar. Ausgehend von solchen Erfahrungen befürworten viele größere Verkehrsverbünde die Umsetzung einer branchenweiten Open-Data-Strategie. Geäußerte Vorbehalte richten sich nicht grundsätzlich gegen den Ansatz, sondern betonen im Einzelnen die erforderlichen Qualitätsstandards bzw. die Kennzeichnung der Datenqualität in Metadatenbeschreibungen – so ein Diskussionsstrang des Expertennetzwerks "Digitale Netze und Mobilität" des Bundesministeriums für Verkehr und digitale Infrastruktur (BMVI).

Einen Lösungsansatz bietet hier die Verwendung der Daten durch eine neutrale Stelle (Clearingstelle), die deren Qualität und Vollständigkeit überprüft und Nutzungsbedingungen nach definierten Kriterien festlegt. Diese Kriterien können bspw. auf Indikatoren aufgebaut sein, die sich am Begriff der Nachhaltigkeit orientieren und ökonomische, ökologische und soziale Effekte erfassen.

Handlungsempfehlung 2: Es müssen Plattformen für einen schnellen und effizienten Austausch von Daten geschaffen werden. Dafür bedarf es eines klaren regulatorischen Rahmens. Daten sollten möglichst kostenlos öffentlich zugänglich gemacht werden. Die Bereitstellung von Daten und die Festlegung der Nutzungsbedingungen könnten dabei über eine unabhängige Clearingstelle durchgeführt werden, wie sie die DLR[28] bereits betreibt.

28 http://www.dlr.de/cs/DesktopDefault.aspx/1177_read-2160/.

These 3: Bedeutung von Datenschutz und Datensicherheit im ÖPV wird durch den Einsatz von Big Data wachsen – Nutzer stellen Sicherheitsbedürfnis über den Wunsch nach Innovationen

Die Bereitstellung und Nutzung von Daten für Verbesserungen im Verkehrsbereich bringt auch Herausforderungen mit sich. Dies umso mehr, falls die Daten personenbezogen oder personenbeziehbar sind und Aufschluss über das individuelle Verkehrsverhalten bieten sowie möglicherweise auch über Wohnort, Arbeitsplatz und Aufenthalte. Schnell können über diese Daten individuelle Profile angelegt werden, die sich um weitere Informationen aus anderen Datenquellen anreichern lassen (Datenfusion) und für eine Vielzahl – auch unerwünschter – Zwecke genutzt werden können. Für Unternehmen kann dies durchaus mit negativer Publicity verbunden sein, wenn solche Fälle öffentlich werden. Einzelne Beispiele zeigen bereits die Entfaltung dieses EmpörungsPotentials im Hinblick auf die nicht ausreichend kommunizierte Verwendung von Daten.[29]

Die Herausforderung liegt darin, geeignete Methoden zu entwickeln, um diese Daten hinreichend zu anonymisieren bzw. zu pseudonymisieren, gleichzeitig aber für verkehrsbezogene Fragestellungen verwenden zu können. Hierfür angemessene Lösungen zu finden, ist aus Sicht der Nutzer von großer Bedeutung: Wie in Kapitel 3 gezeigt, lehnen es 50 % der Befragten ab, beim Einsatz von Big Data für Innovationen im ÖPV Risiken bezüglich Datenschutz und Datensicherheit in Kauf zu nehmen, während nur 22 % bereit sind, hierfür Risiken einzugehen. Diese Erkenntnis gilt – bei leichten Unterschieden – für alle soziodemographischen Gruppen. Nutzer des ÖPV könnten also im Zweifelsfall Sicherheitsbedenken über mögliche Vorteile von Big Data-Anwendungen stellen.

Ein effektiverer Schutz der personenbezogenen Standortdaten muss deshalb im Vorfeld von Technologien, Algorithmen und Prozessen gestaltet werden. Die Anpassung des Datenschutzrahmens an immer mehr durchdringende und präzise Standortdaten ist schwierig, zumal der Datenschutz bisher von vornherein oftmals nicht als Gestaltungselement berücksichtigt wurde. Sowohl freiwillige als auch regulatorische Initiativen sollten einen "Privacy and Security by Design"-Ansatz verfolgen, der sicherstellt, dass ein starker Datenschutz und strenge Kontrollen im Vorfeld in die Datenerfassungs- und Datenverarbeitungsprozesse integriert werden. Technologische Fortschritte, einschließlich der Einführung von System-on-a-Chip-Sensoren, können dies durch die Möglichkeit der On-the-fly-Datenverschlüsselung unterstützen. Zu den weiteren Fortschritten könnten auch Protokolle gehören, die den Nutzern die Kontrolle und Zuweisung von Rechten an ihren Daten ermöglichen. Wenn ein starker Schutz der Privatsphäre nicht gewährleistet ist, kann dies zu einem regulatorischen Gegenspiel gegen die Erfassung und Verarbeitung von Standortdaten führen. Dies könnte Innovationen behindern, das Wohlbefinden der Nutzer beeinträchtigen und den sozialen und wirtschaftlichen Nutzen, den die Verwendung solcher Daten bietet, bremsen.

Handlungsempfehlung 3: Bei der Gestaltung des regulatorischen Rahmens müssen der Datenschutz und die Datensicherheit in den Vordergrund gestellt werden. Optimalerweise sollte die Industrie durch passende Ko- und Selbstregulierungen hierfür gute Wege finden, um möglichen Vertrauensverlusten zuvorzukommen. Wichtige Prinzipien in diesem Kontext sind die des Privacy and Security by Design.

---

[29] Ein Beispiel ist die Verwendung von Mobilfunkdaten durch die VAG Nürnberg: https://netzpolitik.org/2015/data-analytics-nuernberger-nahverkehr-beendet-pilotprojekt-zur-rasterung-von-mobilfunk-vorratsdaten/.

These 4: Big Data-basierte Innovationen bieten Chancen aus wirtschaftlicher und ökologischer Sicht – Soziale Aspekte müssen berücksichtigt werden

Die gemeinsame Betrachtung der Erkenntnisse der Analysen von Angebot- und Nachfragseite macht deutlich, dass es beim Einsatz von Big Data viel Potential gibt, den ÖPV bezüglich Performance, Nutzerinformationen, Kundenfreundlichkeit, Maintenance und internen Betriebsabläufe zu verbessern.

Aus wirtschaftlicher Sicht könnten die Anbieter des ÖPV einerseits durch bestimmte Big Data-basierte Innovationen Kunden gewinnen und damit ihre Auslastung erhöhen. So zeigt die Analyse der Nutzerseite, dass der Einsatz von Big Data-basierten Innovationen im ÖPV große Auswirkungen haben könnte. 40 % der Befragten würden den ÖPV öfter nutzen, wenn eine automatische Rückerstattung bzw. Entschädigung bei Verspätungen umgesetzt würde – ein Vorhaben, dass sich auch im Entwurf des Koalitionsvertrags zwischen CDU/CSU und SPD wiederfindet (CDU, CSU, SPD 2018). Auch eine Anpassung des Angebots an das Verkehrsaufkommen in Echtzeit würde für viele der Befragten den ÖPV attraktiver machen. Der Einsatz von Big Data-Methoden für eine flexiblere Anpassung der Routen von Buslinien an die aktuelle Nachfrage, eine individuellere Reiseplanung oder eine verkehrsmittelübergreifende Zahlungsmöglichkeit durch eine einzige App könnte ebenfalls zu einer höheren Nutzung des ÖPV durch die Bevölkerung beitragen. Anbieterseitig könnte – wie in Kapitel 2 gezeigt – durch den Einsatz von Big Data noch große Effizienzgewinne in der Angebotsplanung oder der Wartung erzielt werden. Von großer Bedeutung für den Erfolg von Innovationen auf der Grundlage von Big Data-Technologien wird auch eine preisneutrale Umsetzung sein, da die Nutzer des ÖPV den Preis als größtes Problem ansehen (vgl. Tabelle 9) und preiselastisch reagieren könnten. Die Einführung von Big Data-basierten Innovationen sollte also nicht über eine Erhöhung der Fahrpreise finanziert werden.

An dieser Stelle ist zu betonen, dass die Ideen und Vorhaben nicht nur im Hinblick auf das technisch Machbare bewertet werden sollten, sondern auch, inwieweit die Konzeption des öffentlichen Verkehrs als Element der Daseinsvorsorge weiterhin diesen Charakter behält. Nachhaltigkeit besitzt neben der ökonomischen und ökologischen auch eine soziale Komponente; gleichermaßen nimmt die Daseinsvorsorge diejenigen Menschen in den Blick, die über keine Alternativen verfügen. Ein ÖPV, der vollständig flexibilisiert und nur noch über Smartphones nutzbar ist, steht vielen Bevölkerungsgruppen nicht mehr zur Nutzung zur Verfügung und exkludiert bestimmte soziale Schichten. Zu denken ist dabei an Kinder, ältere Menschen, kognitiv beeinträchtigte Personen oder Menschen, die sich aus ökonomischen Gründen kein Smartphone leisten können. Diese Meinung teilt auch eine Mehrheit der befragten Nutzer: 62 % lehnen Big Data-basierte Innovationen ab, falls diese zur Benachteiligung bestimmter Nutzergruppen führen (vgl. Tabelle 13).

Nimmt man die ökologischen Aspekte in den Bick, dann bietet der Einsatz von Big Data große Chancen, da viele Innovationen zu einer erhöhten Nutzung des ÖPV und somit zu einer Abnahme bei anderen weniger umweltfreundlichen Mobilitätsangeboten führen könnten. Allerdings sollten hier verschiedene Rebound-Effekte im Voraus berücksichtigt werden. Im Rahmen dieses Gutachtens wurde deutlich, dass z.B. die Anpassung des Fahrpreises in Echtzeit an das aktuelle Verkehrsaufkommen zu einer geringeren Nutzung des ÖPV führen könnte – 40 % der Befragten in Kapitel 3 geben dies an. Andere Rebound-Effekte ergeben sich aus einer Vergrößerung der Nachfrage zu bestimmten Bedienzeiten. Hier sind die bestehenden Beförderungskapazitäten bereits vollständig ausgeschöpft und die bestehenden Infrastrukturen lassen weder eine Verlängerung der eingesetzten Fahrzeuge noch eine Erhöhung der Taktfrequenz zu. Eine weitere Steigerung des Fahrgastaufkommens ließe sich nur durch Investitionen in die Infrastruktur erzielen, wodurch hohe sprungfixe Kosten entstehen,

die entweder durch öffentliche Mittel oder eine Erhöhung der Beförderungsentgelte gegenfinanziert werden müssten. Beides stößt auf Hindernisse, wie oben bereits beschrieben.

Handlungsempfehlung 4: Es muss genau geprüft werden, welche Big Data-basierten Innovationen im ÖPV eingeführt werden, um positive wirtschaftliche sowie ökologische Effekte zu erzielen und sozialpolitisch unerwünschte Folgen auszuschließen. Innovationen zulasten eines höheren Fahrpreises sollten vermieden werden. Ebenso sollte der Charakter der Daseinsvorsorge im ÖPV beibehalten werden und gleichermaßen alle Säulen des Nachhaltigkeitsdiskurses bei der Implementierung von Big Data-Anwendungen oder -Lösungen berücksichtigt werden.

These 5: Strukturelle Anpassungen im Bereich der ÖPV-Regulierung sind notwendig

Die von den Aufgabenträgern in den Ausschreibungen formulierten Leistungsanforderungen an die ausführenden Unternehmen im ÖPV beinhalten oftmals nur die Umsetzung von Betriebsleistungen (Kapazitätsbereitstellung bzw. Beförderungsleistung), die im Sinne der Daseinsvorsorge zu erbringen sind. Weitergehende Entwicklungen und Investitionen werden im Rahmen von bestellten Verkehren nicht gefordert und von den leistungserbringenden Unternehmen nicht erbracht, da solche Investitionen weder vergütet noch gegenüber Wettbewerbern bei Ausschreibungen einen Vorteil bringen. Innovatives Handeln oder proaktives Investment bleibt dadurch wirtschaftlich unrentabel. Dieser Umstand stellt eine Finanzierungshürde dar, die zum Beispiel in der Vergangenheit dem Ausbau von WLAN im Regionalverkehr entgegenstand. In die Zukunft gedacht ist das sogenannte Besteller/Ersteller-Prinzip auch als Innovationshindernis für den Einsatz innovativer Technologien und Big Data-Anwendungen zu sehen. Hier sind unterschiedliche Lösungsansätze möglich, die den ÖPV aus technologischer Perspektive besser für die Zukunft vorbereiten.

Eine naheliegende Lösung ist es, entsprechende Anforderungen an die Unternehmen in den Ausschreibungen für die Leistungserbringung mit zu formulieren und entsprechend mit Mitteln zu hinterlegen. Allerdings bleibt hier der Anreiz selbst innovativ tätig zu werden auf der Strecke. Besonders die Unternehmen des ÖPNV verbleiben im sogenannten "goldenen Käfig", in dem sie zwar auskömmlich finanziert, aber wenig auf zukünftige Herausforderungen vorbereitet werden.

Andere Ansätze liegen in öffentlich-privaten Partnerschaften, die privaten Unternehmen Zugang zu den öffentlich finanzierten Verkehren ermöglicht, diese aber mit eigenen Investitionen modernisiert. Im Gegenzug könnten die privaten Unternehmen die Möglichkeit erhalten, die erhobenen Daten vertrieblich zu nutzen und zu verwerten. Entsprechende Ansätze verfolgt zum Beispiel das IT-Unternehmen Cisco, die in mehreren Kooperationen in Deutschland und Frankreich Verkehrsunternehmen bei der Digitalisierung finanziell unterstützen wollen, indem sie die digitale Infrastruktur bereitstellen. Im Gegenzug erhält das Unternehmen einen exklusiven Zugang zu den Nutzerdaten, die zum Beispiel über die Verbindung der Smartphones von Reisenden mit dem im Zug befindlichen WLAN anfallen. Partnerschaften zwischen dem öffentlichen und dem privaten Sektor müssen möglicherweise über die reine Lieferanten-Kunden-Beziehung hinausgehen und bedürfen neuer Regularien und vertraglicher Beziehungen, um auch hier öffentliche Interessen sowie die Interessen der Nutzer hinsichtlich ihrer personenbezogenen Daten zu wahren. Solche Kooperationen und damit einhergehenden exklusiven Zugänge und Verwertungsrechte erfordern – neben einer genauen datenschutzrechtlichen Kontrolle – eine Überprüfung der Marktstruktur, um ausgeprägte Marktkonzentrationen und Monopole zu verhindern. Auch Kosten-Nutzen-Bewertungen bei Investitionen und Zielformulierungen

öffentlicher Versorgungseinrichtungen im Sinne der Nachhaltigkeitsdiskurse sind erforderlich und sollten weiterhin handlungsleitend sein.

Handlungsempfehlung 5: Um Investitionen in die digitale Infrastruktur zu fördern, um damit die Grundlagen für Big Data-Anwendungen zu schaffen, sollten neue Möglichkeiten im Dreiklang von Besteller, Ersteller und private Unternehmen entwickelt werden. Ansätze bestehen darin, Ausschreibungen um digitale Komponenten zu erweitern oder aber in größerem Maße öffentlich-privaten Partnerschaften zuzulassen. Dabei sind regulatorische Rahmenbedingungen zu schaffen, die Monopolstellungen verhindern und Zielstellungen der Daseinsvorsorge, des Datenschutzes und der Nachhaltigkeit nicht obsolet werden lassen.

5.GLOSSAR

Beförderungspflicht: alle Personen, die den Tarif entrichten, müssen befördert werden.

Betriebspflicht: Ein ÖPV-Anbieter muss auch dann fahren, wenn keine Nachfrage existiert.

Crowd-Sourcing: Eine Technik zur kollektiven, digitalen Erhebung von Daten durch eine große Zahl an Beitragenden als Community (sprich „Crowd"). In „Covert" (versteckt) oder „Overt" (offengelegte) Erhebungen eingeteilt.

Data-Mining: Techniken, die eine Mustererkennung für große Datensätze ermöglichen und Methoden der Statistik (Assoziationsanalyse, Clusteranalyse, mathematische Klassifikation, Regressionsanalyse) mit automatisiertem Lernen kombinieren.

Echtzeit: Der Betrieb einer fortlaufenden Datenverarbeitung, so dass Verarbeitungsergebnisse innerhalb einer kurzen, fest definierten Zeitspanne verfügbar sind.

Exhaustivity (Vollständigkeit): Die Daten entsprechen (nahezu) einer Vollerhebung. Eine möglichst vollständige Erfassung aller Untersuchungsobjekte wird angestrebt, nicht das Sampling optimiert.

Fahrplanpflicht: Ein ÖPV-Anbieter muss in einem Fahrplan ersichtlich machen, zu welchem Zeitpunkt gefahren wird.

Floating Car Data (FCD): Ein Verfahren, um über Fahrzeugflotten Daten zum Verkehrsfluss und der Verkehrsinfrastruktur zu sammeln.

Floating Phone Data: Aggregierte Verbindungsprotokolle aus der Mobilfunkkommunikation, die Rückschlüsse über die Route des Smartphones geben.

Intelligent Transportation Systems (ITS): Ein ITS umfasst eine Systemintegration, die den intelligenten Betrieb und eine weitreichende, übergreifende Nutzung von Einzelsystemen im ÖPV ermöglicht. Das Ergebnis ist ein anbieter- und verkehrsmittelübergreifendes System mit höherer Planbarkeit und Zuverlässigkeit für den Nutzer.

Internet of Things (IoT): Digitales Netzwerk von Geräten des alltäglichen Lebens.

Linienpflicht: Es muss eine bestimmte, vorher festgelegte und kommunizierte ÖPV-Route bedient werden.

LogFiles: Digitale Prozesse werden in LogFiles dokumentiert. So können vergangene Vorgänge nachvollzogen werden und als Datenquelle z.B. von Nutzungshäufigkeiten bestimmter Funktionen oder Besucherzahlen bestimmter Inhalte genutzt werden.

MaaS: Kurz für *Mobility as a Service* welches zum Sammelbegriff für anbieterübergreifende Mobilitätsdienste geworden ist. Werden mehrere Angebote verschiedener Verkehrsbetreiber oder Dienstleister mit zum Teil unterschiedlichen Verkehrsmitteln kombiniert, um Informationen, Zugang oder Abrechnung zu ermöglichen, spricht man von MaaS.

Machine Learning: Unter diesen Begriff fällt die Entwicklung von Algorithmen, die automatisch komplexe Muster erkennen und erlernen. Ein Beispiel ist das Natural Language Processing.

MapReduce: Bei dem für Big Data-Anwendungen wichtigen Verfahren werden Rechenschritte parallelisiert und auf diverse Rechencluster verteilt (Map) und im Anschluss die Ergebnisse zusammengeführt (Reduce).

Öffentlicher Personenfernverkehr (ÖPFV): Siehe öffentlicher Personennahverkehr, Fahrweiten über 50 km oder Fahrdauern über 60 Minuten. Bis 100 km Entfernungen spricht man auch häufig vom Regionalverkehr.

Öffentlicher Personennahverkehr (ÖPNV): ÖPNV ist die allgemein zugängliche Personenbeförderung mit Bussen, Straßenbahnen, Eisenbahnen, Taxen, S- und U-Bahnen sowie Wasserfahrzeugen bei geregelter Fahrttaktung bis 50km Länge bzw. 60 Minuten Dauer.

Öffentlicher Personenverkehr (ÖPV): Klassischerweise werden unter diesem Begriff alle Angebote gefasst, die sich den Pflichten des PBefGs widmen, also über einen Fahrplan und festen Linien Fahrdienste diskriminierungsfrei für alle Personen anbieten. Durch das Aufkommen neuer Angebotslösungen wird auch zunehmend ein öffentlicher Zugang als entscheidendes Kriterium für die Definition von ÖPV genutzt.

Open Data: Daten werden als offen bezeichnet, wenn sie ohne Einschränkung der Nutzungsrechte öffentlich zur Verfügung gestellt werden und in der Nutzung nicht finanziell oder systemtechnisch Kontrolliert werden.

Open Source: Ein Programm, welches ohne Beschränkungen für Zugang einsehbar und ggf. abwandelbar und zur weiteren Nutzung zur Verfügung steht.

Personenbeförderungsgesetz (PBefG): Ist Teil des deutschen Verkehrsrechts und regelt die Beförderung von Personen. Es wird auf Bundesebene erlassen.

Predictive Modelling: Die daten- und rechnergestützte Bestimmung von Ereignissen, meist als Vorhersage für zukünftige Ereignisse. Hierbei können diverse statistische Verfahren oder Modelle eingesetzt werden.

Relationality (Relationalität): Verschiedene Datensätze können über gemeinsame Variablen verknüpft werden.

Semantic Web: Die automatische Auswertung von Semantiken digitaler Inhalte im Internet.

Sensoren: Ein Gerät, welches Eigenschaften seiner Umgebung erfassen kann und darüber hinaus eine Zustandskategorie oder einen konkreten Messwert digital kommunizieren kann.

Tarifpflicht: es muss ein Tarif für die Beförderungsdienstleistung festgelegt sein.

Value (Werthaltigkeit): Aus den Daten können sehr viele Rückschlüsse gezogen und die Daten in anderen Kontexten angewandt werden.

Variability (Variabilität): Die Daten können ganz unterschiedliche Bedeutungen haben und Aussagen produzieren, je nachdem in welchem Kontext sie genutzt werden.

Veracity (Sinnhaftigkeit): Die Sinnhaftigkeit und Vertrauenswürdigkeit der Daten sollte jeweils hinterfragt werden. Die Daten können ungeordnet und verrauscht sein sowie Unsicherheiten und Fehler enthalten.

Volume: Im Zusammenhang mit Big Data für die Datenmenge verwendet, die für die Übertragung oder Speicherung von Informationen verwendet wird.

6. QUELLEN

Antoniou, C. (2016). *Big Data for transportation systems analysis. Challenges and opportunities.* Paper presented at the Verkehr aktuell. München.

Batty, M. (2013). Big data, smart cities and city planning. *Dialogues in Human Geography*, *3*(3), 274-279. doi:10.1177/2043820613513390

BITKOM Bundesverband Informationswirtschaft Telekommunikation und neue Medien e.V. (Ed.) (2015). *Big Data und Geschäftsmodell-Innovationen in der Praxis: 40+ Beispiele. Leitfaden.* Berlin.

Bloching, B., Luck, L., & Ramge, T. (2012). *Data unser. Wie Kundendaten die Wirtschaft revolutionieren*. München.

Borgia, E. (2014). The Internet of Things vision: Key features, applications and open issues. *Computer Communications*, *54*(0), 1-31. doi: http://dx.doi.org/10.1016/j.comcom.2014.09.008.

Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, *15*(5), 662-679. doi:10.1080/1369118X.2012.678878

Bradley, J., Barbier, J., & Handler, D. (2013). *Embracing the Internet of Everything. White Paper*. San Jose.

Buliung, R. N., & Kanaroglou, P. S. (2007). Activity–Travel Behaviour Research: Conceptual Issues, State of the Art, and Emerging Perspectives on Behavioural Analysis and Simulation Modelling. *Transport Reviews*, *27*(2), 151-187. doi:10.1080/01441640600858649

Calvert, F. (2014). *Probes and Big Data: Opportunities and challenges.* Paper presented at the New ways to collect transport and traffic data and their potential applications. Melbourne.

CDU, CSU & SPD (2018). Ein neuer Aufbruch für Europa. Eine neue Dynamik für Deutschland. Ein neuer Zusammenhalt für unser Land. *Koalitionsvertrag zwischen CDU, CSU und SPD*. http://www.handelsblatt.com/downloads/20936422/4/koalitionsvertrag_final.pdf. Zugegriffen: 26. Februar 2018.

Demchenko, Y., Grosso, P., de Laat, C., & Membrey, P. (2013, 20-24 May 2013). *Addressing big data issues in Scientific Data Infrastructure.* Paper presented at the 2013 International Conference on Collaboration Technologies and Systems (CTS). San Diego.

Diebold, F. X. (2012). A Personal Perspective on the Origin(s) and Development of 'Big Data': The Phenomenon, the Term, and the Discipline. *PIER Working Paper, No. 13-003*.

Dodge, M., & Kitchin, R. (2005). Codes of Life: Identification Codes and the Machine-Readable World. *Environment and Planning D: Society and Space*, *23*(6), 851-881. doi:10.1068/d378t

Evers, H., & Kasties, G. (Eds.). (1999). *Kompendium der Verkehrstelematik. Technologien, Applikationen, Perspektiven* (Vol. 4.). Köln: TÜV-Verlag.

Fachverband Außenwerbung e.V. (2017). Trendanalyse Mobilität und Mediennutzung 2017. Studie durchgeführt von Research Now. https://www.faw-ev.de/wp-content/uploads/2017/11/Charts_Trendanalyse-Mobilität-und-Mediennutzung-2017.pdf Zugegriffen: 12. Dezember 2017.

Forschungsgesellschaft für Straßen- und Verkehrswesen Arbeitskreis Multi- und Intermodalität (FGSV). (2017). *Multi- und Intermodalität: Hinweise zur Umsetzung und Wirkung von Maßnahmen im Personenverkehr. Teilpapier 1: Definitionen*. Köln.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137-144. doi: https://doi.org/10.1016/j.ijinfomgt.2014.10.007.

Gerdes, F., Schelewsky, M., Knie, A., Schönduwe, R., & Jahn, V. (2015). *Die Digitalisierung des Schienenverkehrs. White Paper*. Berlin.

Heinze, C., Fehling, M., & Fiedler, L. H. (2014). *Personenbeförderungsgesetz: PBefG* (Vol. 2). München: C.H. Beck.

Hölzle, K., Schoder, T., Spiri, N., & Götz, V. (2017). Big Data und technologiegetriebene Geschäftsmodellinnovation. In D. Schallmo, A. Rusnjak, J. Anzengruber, T. Werani, & M. Jünger (Eds.), *Digitale Transformation von Geschäftsmodellen: Grundlagen, Instrumente und Best Practices* (pp. 355-374). Wiesbaden: Springer Fachmedien Wiesbaden.

IBM (2013). *Big data and analytics in travel and transportation. Beyond the hype: Solutions that deliver big value*. *White Paper*. Somers.

Initiative D21 (2018). *D21-Digital-Index 2017 / 2018 – Jährliches Lagebild zur Digitalen Gesellschaft*. ISBN: 978-3-9818331-4-0.

International Transport Forum (ITF), & Organisation for economic co-operation and development (OECD). (2015). *Big Data and Transport. Understanding and assessing options*. Paris.

Kitchin, R., & McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society*, *3*(1), 205395171663113. doi:10.1177/2053951716631130

Kutter, E. (1973). A model for individual travel behaviour. *Urban Studies*, *10*, 235-258.

Marz, N., & Warren, J. (2012). *Big Data: Principles and Best Practices of Scalable Realtime Data Systems*. Westhampton, NJ: Manning.

Mayer-Schonberger, V., & Cukier, K. (2013). *Big Data: A Revolution that will Change How We Live, Work and Think*. London: John Murray.

Neumann, A., Balmer, M., & Rieser, M. (2012). *Converting a static macroscopic model into a dynamic activity-based model to analyze public transport demand in Berlin*. Paper presented at the 13th International Conference on Travel Behaviour Research "Travel behaviour research: current foundations, future prospects". Toronto.

Parliamentary Office of Science & Technology. (2014). *Big and Open Data in Transport* (Vol. 472). London.

Ren, Y., Chen, G., Han, Y., & Zheng, H. (2017). *Extracting potential bus lines of customized city bus services based on public transport big data*. Paper presented at the 6th Digital Earth Summit IOP Conference Series: Earth and Environmental Science. Peking.

Rieser, M. (2016). *Anwendungen agentenbasierter Verkehrsmodelle*. Paper presented at the Verkehrsingenieurtag. Zürich.

Sandrock, M., & Riegelhuth, G. (Eds.). (2014). *Verkehrsmanagementzentralen in Kommunen. Eine vergleichende Darstellung*. Wiesbaden.

Schmidt, A., & Mänel, T. (2017). *Potentialanalyse zur Mobilfunkdatennutzung in der Verkehrsplanung*. Stuttgart: Fraunhofer IAO.

Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, *70*, 263-286. doi:10.1016/j.jbusres.2016.08.001

Stabellini, B., Remondino, C. L., & Tamborrini, P. (2017). Data Visualization Collection. How graphical representation can inspect and communicate sustainability through Systemic Design. *The Design Journal, 20*(sup1), S1673-S1681. doi:10.1080/14606925.2017.1352690

Turck, M. (2017). *Big Data Landscape 2017*. http://mattturck.com/bigdata2017. Zugegriffen: 11. Oktober 2017.

UN Global Pulse. (2017). *Using big data for improved public transport*. New York.

Uprichard, E. (2013). Focus : big data, little questions? *Discover Society*, *1*(1), 1-6.

VCD – Verkehrsclub Deutschland (2015). *Wie gut sind Bus und Bahn bei Ihnen vor Ort? – Umfrage ÖPNV-Zugänglichkeit*. https://www.vcd.org/themen/oeffentlicher-

personennahverkehr/oepnv-zugaenglichkeitscheck/umfrage-oepnv-zugaenglichkeit/. Zugegriffen: 12. Oktober 2017.

VDV - Verband Deutscher Verkehrsunternehmen (2016). *VDV-Statistik 2016*. https://www.vdv.de/jahresbericht---statistik.aspx. Zugegriffen: 12. Oktober 2017.

Weikl, S., & Bogenberger, K. (2015). Integrated Relocation Model for Free-Floating Carsharing Systems. *Transportation Research Record: Journal of the Transportation Research Board*, *2536*, 19-27. doi:10.3141/2536-03

Weiß, C., Chlond, B., Von Behren, S., Hilgert, T. & Vortisch, P. (2016). *Deutsches Mobilitätspanel (MOP) – Wissenschaftliche Begleitung und Auswertungen Bericht 2015/2016: Alltagsmobilität und Fahrleistung*. Durchgeführt im Auftrag des Bundesministeriums für Verkehr und digitale Infrastruktur. Forschungsprojekt FE-Nr. 70.923/2015.

Wootton, J. R., García-Ortiz, A., & Amin, S. M. (1995). Intelligent transportation systems: A global perspective. *Mathematical and Computer Modelling*, *22*(4–7), 259-268. doi: http://dx.doi.org/10.1016/0895-7177(95)00137-Q.

Wrobel, S. (2013). *Big Data - Perspectives for Germany. Seize the opportunity*. Paper presented at the 1st German-Indian Workshop "Big Data in Engineering and Science". Berlin.

7. ANHANG

Anhang 1: Tabellenanhang

Tab. A 1: Nutzung von ÖPNV (geordnet nach Altersklassen)

Altersklasse (N=1000)	(fast) täglich	1-3 Tage /Woche	1-3 Tage /Monat	seltener als monatlich	nie
0-29 Jahre	30,1	19,6	16	23,3	11
30-59 Jahre	11,8	14,6	22,3	28,9	22,3
60+ Jahre	10,6	11,8	18,6	38,8	20,2

Tab. A 2: Nutzung von Bahnfernverkehr (geordnet nach Altersklassen)

Altersklasse (N=1000)	(fast) täglich	1-3 Tage /Woche	1-3 Tage /Monat	seltener als monatlich	nie
0-29 Jahre	6,1	14,7	15,3	35	28,8
30-59 Jahre	2,5	3,7	13,2	44,7	35,9
60+ Jahre	0,6	0,3	7,8	54	37,3

Tab. A 3: Nutzung von Busfernverkehr (geordnet nach Altersklassen)

Altersklasse (N=1000)	(fast) täglich	1-3 Tage /Woche	1-3 Tage /Monat	seltener als monatlich	nie
0-29 Jahre	8,6	8	11	38,7	33,7
30-59 Jahre	2,3	3,1	10,1	31,8	52,6
60+ Jahre	1,6	1,6	1,9	31,7	63,4

Tab. A 4: Nutzung von Bahnfernverkehr (getrennt nach Bildungsabschluss)

Bildung (N=1000)	(fast) täglich	1-3 Tage /Woche	1-3 Tage /Monat	seltener als monatlich	nie
kein Abitur	2,2	1,5	8,5	43,4	44,5
Abitur	3	8,8	16,8	50,3	21,1

Tab. A 5: Nutzung von Busfernverkehr (getrennt nach Bildungsabschluss)

Bildung (N=1000)	(fast) täglich	1-3 Tage /Woche	1-3 Tage /Monat	seltener als monatlich	nie
kein Abitur	2,3	2,7	5,3	30,4	59,3
Abitur	4,3	4,5	11,1	36,7	43,5

Tab. A 6: Nutzung von Bahnfernverkehr (geordnet nach Stadtgröße)

Städtegröße (N=1000)	(fast) täglich	1-3 Tage /Woche	1-3 Tage /Monat	seltener als monatlich	nie
bis unter 20.000 Einwohner	2,5	2,5	11,6	38,8	44,6
20.000 50.000 Einwohner	3,2	0	6,3	53,7	36,8
50.000 100.000 Einwohner (Kern, Verdichtungs- und Übergangsbereich)	0	3,6	11,9	52,4	32,1
100.000 - 500.000 Einwohner (Verdichtungs- und Übergangsbereich)	0,7	5	10	45	39,3
100.000 - 500.000 Einwohner (Kernbereich)	3	6,6	12,6	44,9	32,9
Mehr als 500.000+ Einwohner (Verdichtungs- und Übergangsbereich)	1,2	2,4	13,4	40,2	42,7
Mehr als 500.000+ Einwohner (Kernbereich)	3,9	5,8	13,5	47,6	29,3

Tab. A 7: Nutzung von Busfernverkehr (geordnet nach Stadtgröße)

Städtegröße (N=1000)	(fast) täglich	1-3 Tage /Woche	1-3 Tage /Monat	seltener als monatlich	nie
bis unter 20.000 Einwohner	1,7	3,3	5	30,6	59,5
20.000 50.000 Einwohner	1,1	1,1	4,2	34,7	58,9
50.000 100.000 Einwohner (Kern, Verdichtungs- und Übergangsbereich)	0	3,6	7,1	34,5	54,8
100.000 - 500.000 Einwohner (Verdichtungs- und Übergangsbereich)	2,1	4,3	10,7	27,1	55,7
100.000 - 500.000 Einwohner (Kernbereich)	3,6	3,6	9	39,5	44,3
Mehr als 500.000+ Einwohner (Verdichtungs- und Übergangsbereich)	2,4	0	3,7	35,4	58,5
Mehr als 500.000+ Einwohner (Kernbereich)	5,5	4,5	8,7	31,2	50,2

Tab. A 8: Zufriedenheit mit dem Preis des ÖPV (getrennt nach Geschlecht)

Geschlecht (N=916)	1 - sehr unzufrieden	2	3	4	5 - sehr zufrieden
weiblich	25,5	27,3	25,1	15,2	6,9
männlich	23,6	18,7	30,2	19,1	8,4

Tab. A 9: Zufriedenheit mit der Funktionstüchtigkeit des ÖPV (getrennt nach Geschlecht)

Geschlecht (N=911)	1 - sehr unzufrieden	2	3	4	5 - sehr zufrieden
weiblich	5,2	11,4	29,7	39,2	14,4
männlich	4	6,3	30,9	39,1	19,7

Tab. A 10: Zufriedenheit mit der Sauberkeit des ÖPV (getrennt nach Geschlecht)

Geschlecht (N=915)	1 - sehr unzufrieden	2	3	4	5 - sehr zufrieden
weiblich	9,6	16,9	39	24,8	9,6
männlich	5,8	12,5	36,8	30,8	14,1

Tab. A 11: Zufriedenheit mit den Umsteigezeiten des ÖPV (getrennt nach Geschlecht)

Geschlecht (N=886)	1 - sehr unzufrieden	2	3	4	5 - sehr zufrieden
weiblich	9,7	19,3	34,6	27,2	9,2
männlich	7,3	13,8	33,3	30,8	14,7

Tab. A 12: Zufriedenheit mit Informationen bei Verspätungen (getrennt nach Geschlecht)

Geschlecht (N=894)	1 - sehr unzufrieden	2	3	4	5 - sehr zufrieden
weiblich	21,2	22,7	32,5	17,8	5,8
männlich	13,7	20,9	29,4	23,8	12,1

Tab. A 13: Zufriedenheit mit dem Umgang mit Rückerstattungsforderungen im ÖPV (geordnet nach Altersklassen)

Altersklasse (N=656)	1 - sehr unzufrieden	2	3	4	5 - sehr zufrieden
0-29 Jahre	14,1	15,6	32,8	22,7	14,8
30-59 Jahre	17,8	17,5	31	23,3	10,3
60+ Jahre	19,4	26,7	30	12,2	11,7

Tab. A 14: Zufriedenheit mit der Bearbeitungsdauer bei Rückerstattungsforderungen im ÖPV (geordnet nach Altersklassen)

Altersklasse (N=642)	1 - sehr unzufrieden	2	3	4	5 - sehr zufrieden
0-29 Jahre	13,6	13,6	33,6	20,8	18,4
30-59 Jahre	19,2	16,8	32,2	21,5	10,3
60+ Jahre	20,2	29,2	28,1	12,4	10,1

Tab. A 15: Zufriedenheit mit den Umsteigezeiten im ÖPV (geordnet nach Stadtgröße)

Städtegröße (N=886)	1 - sehr unzufrieden	2	3	4	5 - sehr zufrieden
bis unter 20.000 Einwohner	13,7	27,5	25,5	24,5	8,8
20.000 50.000 Einwohner	5,3	26,3	35,5	18,4	14,5
50.000 100.000 Einwohner (Kern, Verdichtungs- und Übergangsbereich)	5,6	12,5	38,9	30,6	12,5
100.000 - 500.000 Einwohner (Verdichtungs- und Übergangsbereich)	11,2	17,2	31	28,4	12,1
100.000 - 500.000 Einwohner (Kernbereich)	8,4	10,4	39	28,6	13,6
Mehr als 500.000+ Einwohner (Verdichtungs- und Übergangsbereich)	4,3	11,4	44,3	30	10
Mehr als 500.000+ Einwohner (Kernbereich)	8,1	15,5	31,4	33,1	11,8

Tab. A 16: Zufriedenheit mit der Anzahl der Anschlussmöglichkeiten (geordnet nach Stadtgröße)

Städtegröße (N=894)	1 - sehr unzufrieden	2	3	4	5 - sehr zufrieden
bis unter 20.000 Einwohner	15,4	23,1	28,8	24	8,7
20.000 50.000 Einwohner	7,8	28,6	35,1	20,8	7,8
50.000 100.000 Einwohner (Kern, Verdichtungs- und Übergangsbereich)	6,8	16,4	34,2	28,8	13,7
100.000 - 500.000 Einwohner (Verdichtungs- und Übergangsbereich)	10,5	15,8	35,1	29,8	8,8
100.000 - 500.000 Einwohner (Kernbereich)	8,3	10,9	35,9	30,8	14,1
Mehr als 500.000+ Einwohner (Verdichtungs- und Übergangsbereich)	6,8	13,7	35,6	32,9	11
Mehr als 500.000+ Einwohner (Kernbereich)	7,4	14,8	32,7	31,6	13,5

Tab. A 17: Nutzung von digitalen Angeboten (geordnet nach Altersklassen)

Digitale Aspekte	0-29 Jahre	30-59 Jahre	60+ Jahre
Routenplanung über Websites oder Apps	86,9% (von 122)	75,7% (von 379)	82,3% (von 232)
Buchung von Online-Tickets über Websites oder Apps der Anbieter des öffentlichen Personenverkehrs	61,6% (von 112)	55,6% (von 331)	53,6% (von 196)
Suche nach Sparangeboten für Bahn- oder Busreisen auf digitalen Plattformen von Drittanbietern (mytrain.de, Sparpreisfinder.de, ltur.com etc.)	66,7% (von 78)	68,9% (von 228)	70% (von 160)
Kontaktloses Bezahlen (NFC) in Fahrzeugen z.B. mit der Kreditkarte oder dem Handy	43,9% (von 66)	34,2% (von 190)	18,4% (von 125)
WLAN in Verkehrsmitteln oder Haltestellen von Anbietern des öffentlichen Personenverkehrs*	66,7% (von 96)	62,4% (von 282)	49,4% (von 154)
Online-Auftritt, Twitter-Auftritt, Facebook-Auftritt, etc. von Anbietern des öffentlichen Personenverkehrs	63,3% (von 79)	61,7% (von 206)	56,4% (von 110)

Tab. A 18: Kenntnis von digitalen Angeboten (getrennt nach Bildungsabschluss)

Digitale Aspekte (N=1000)	Kein Abitur	Abitur
Routenplanung über Websites oder Apps	68,9	79,9
Buchung von Online-Tickets über Websites oder Apps der Anbieter des öffentlichen Personenverkehrs	57,1	74,1
Suche nach Sparangeboten für Bahn- oder Busreisen auf digitalen Plattformen von Drittanbietern (mytrain.de, Sparpreisfinder.de, ltur.com etc.)	41,7	54
Kontaktloses Bezahlen (NFC) in Fahrzeugen z.B. mit der Kreditkarte oder dem Handy	33,1	45,7
WLAN in Verkehrsmitteln oder Haltestellen von Anbietern des öffentlichen Personenverkehrs*	49,5	58,8
Online-Auftritt, Twitter-Auftritt, Facebook-Auftritt, etc. von Anbietern des öffentlichen Personenverkehrs	31,4	51,8

Tab. A 19: Nutzung von digitalen Angeboten (getrennt nach Bildungsabschluss)

Digitale Aspekte	Kein Abitur	Abitur
Routenplanung über Websites oder Apps	74,9% (von 415)	85,8% (von 318)
Buchung von Online-Tickets über Websites oder Apps der Anbieter des öffentlichen Personenverkehrs	48,8% (von 344)	64,4% (von 295)
Suche nach Sparangeboten für Bahn- oder Busreisen auf digitalen Plattformen von Drittanbietern (mytrain.de, Sparpreisfinder.de, ltur.com etc.)	68,5% (von 251)	69,3% (von 215)
Kontaktloses Bezahlen (NFC) in Fahrzeugen z.B. mit Kreditkarte oder Handy	25,1% (von 199)	36,8% (von 184)
WLAN in Verkehrsmitteln oder Haltestellen von Anbietern des öffentlichen Personenverkehrs*	53,4% (von 298)	67,1% (von 234)
Online-Auftritt, Twitter-Auftritt, Facebook-Auftritt, etc. von Anbietern des öffentlichen Personenverkehrs	58,2% (von 189)	62,6% (von 206)

Tab. A 20: Einstellung zum Datenschutz (getrennt nach Geschlecht)

	Aussage: Es ist in Ordnung, dass personenbezogene Daten von den Betreibern des öffentlichen Personenverkehrs verarbeitet und ggf. an Dritte weitergegeben werden, wenn dafür Innovationen durch Big Data im öffentlichen Personenverkehr ermöglicht werden können.				
Geschlecht (N=1000)	**1 - sehr große Ablehnung**	**2**	**3**	**4**	**5 - sehr große Zustimmung**
weiblich	35,1	18,1	29,4	12,9	4,5
männlich	28,1	18,5	27,7	18,3	7,4

Tab. A 21: Einstellung zu unausgereiften Innovationen (getrennt nach Geschlecht)

	Aussage: Mir ist es wichtig, dass Innovationen durch Big Data möglichst bald im öffentlichen Personenverkehr eingesetzt werden. Dafür würde ich notfalls auch in Kauf nehmen, dass diese Innovationen noch nicht komplett ausgereift und zuverlässig sind.				
Geschlecht (N=1000)	**1 - sehr große Ablehnung**	**2**	**3**	**4**	**5 - sehr große Zustimmung**
weiblich	27,9	18,9	30,6	15,2	7,4
männlich	21,8	15,4	32,2	21,4	9,2

Tab. A 22: Einstellung zum Datenschutz (geordnet nach Altersklassen)

	Aussage: Es ist in Ordnung, dass personenbezogene Daten von den Betreibern des öffentlichen Personenverkehrs verarbeitet und ggf. an Dritte weitergegeben werden, wenn dafür Innovationen durch Big Data im öffentlichen Personenverkehr ermöglicht werden können.				
Alter (N=1000)	**1 - sehr große Ablehnung**	**2**	**3**	**4**	**5 - sehr große Zustimmung**
0-29 Jahre	19	17,2	38	19	6,7
30-59 Jahre	30,1	16,9	30,5	15,9	6,6
60+ Jahre	40,7	21,1	20,8	13	4,3

Tab. A 23: Einstellung zu Datensicherheit (geordnet nach Altersklassen)

	Aussage: Beim Einsatz von Big Data für Innovationen im öffentlichen Personenverkehr müssen Nutzer bestimmte Risiken bei der Datenübertragung und Datensicherheit in Kauf nehmen.				
Alter (N=1000)	**1 - sehr große Ablehnung**	**2**	**3**	**4**	**5 - sehr große Zustimmung**
0-29 Jahre	22,1	13,5	35	20,2	9,2
30-59 Jahre	29,9	15,3	28,3	18,3	8,2
60+ Jahre	42,2	25,5	18,6	7,8	5,9

Tab. A 24: Einstellung zu Preiserhöhung (geordnet nach Altersklassen)

	Aussage: Ich bin bereit, mehr Geld für den öffentlichen Personenverkehr zu zahlen, wenn dadurch Innovationen durch Big Data ermöglicht werden können.				
Alter (N=1000)	**1 - sehr große Ablehnung**	**2**	**3**	**4**	**5 - sehr große Zustimmung**
0-29 Jahre	27	18,4	30,1	17,8	6,7
30-59 Jahre	39,8	16,7	24,7	14	4,9
60+ Jahre	49,1	21,4	19,3	7,1	3,1

Tab. A 25: Einstellung zu Benachteiligung von Nutzergruppen (geordnet nach Altersklassen)

	Aussage: Eine Benachteiligung von bestimmten Nutzergruppen (z.B. ältere Personen oder Personen ohne Internetzugang) muss ausgeschlossen werden, wenn Innovationen durch Big Data eingeführt werden.				
Alter (N=1000)	**1 - sehr große Ablehnung**	**2**	**3**	**4**	**5 - sehr große Zustimmung**
0-29 Jahre	7,4	4,9	32,5	26,4	28,8
30-59 Jahre	9,9	8	21,4	25,4	35,3
60+ Jahre	12,4	6,8	12,4	18	50,3

Tab. A 26: Einstellung zur Ausgereiftheit von Innovationen (geordnet nach Altersklassen)

	Aussage: Mir ist es wichtig, dass Innovationen durch Big Data möglichst bald im öffentlichen Personenverkehr eingesetzt werden. Dafür würde ich notfalls auch in Kauf nehmen, dass diese Innovationen noch nicht komplett ausgereift und zuverlässig sind.				
Alter (N=1000)	1 - sehr große Ablehnung	2	3	4	5 - sehr große Zustimmung
0-29 Jahre	15,3	16,6	36,8	23,3	8
30-59 Jahre	24,1	17,1	29,1	20,8	8,9
60+ Jahre	31,1	17,7	32,3	11,5	7,5

Tab. A 27: Einstellung zum Datenschutz (getrennt nach Bildungsabschluss)

	Aussage: Es ist in Ordnung, dass personenbezogene Daten von den Betreibern des öffentlichen Personenverkehrs verarbeitet und ggf. an Dritte weitergegeben werden, wenn dafür Innovationen durch Big Data im öffentlichen Personenverkehr ermöglicht werden können.				
Bildung (N=1000)	1 - sehr große Ablehnung	2	3	4	5 - sehr große Zustimmung
kein Abitur	35,9	18,1	27,1	13,3	5,6
Abitur	25,4	18,6	30,9	18,8	6,3

Tab. A 28: Einstellung zu Datensicherheit (getrennt nach Bildungsabschluss)

	Aussage: Beim Einsatz von Big Data für Innovationen im öffentlichen Personenverkehr müssen Nutzer bestimmte Risiken bei der Datenübertragung und Datensicherheit in Kauf nehmen.				
Bildung (N=1000)	1 - sehr große Ablehnung	2	3	4	5 - sehr große Zustimmung
kein Abitur	35,2	16,6	27,4	12,8	8
Abitur	28,6	20,9	24,6	18,8	7

Tab. A 29: Einstellung zu Preiserhöhung (getrennt nach Bildungsabschluss)

	Aussage: Ich bin bereit, mehr Geld für den öffentlichen Personenverkehr zu zahlen, wenn dadurch Innovationen durch Big Data ermöglicht werden können.				
Bildung (N=1000)	1 - sehr große Ablehnung	2	3	4	5 - sehr große Zustimmung
kein Abitur	44,4	18,6	23,3	9,5	4,3
Abitur	35,2	18,3	24,6	16,8	5

Tab. A 30: Einstellung zur Benachteiligung von bestimmten Nutzergruppen (getrennt nach Bildungsabschluss)

	Aussage: Eine Benachteiligung von bestimmten Nutzergruppen (z.B. ältere Personen oder Personen ohne Internetzugang) muss ausgeschlossen werden, wenn Innovationen durch Big Data eingeführt werden.				
Bildung (N=1000)	**1 - sehr große Ablehnung**	**2**	**3**	**4**	**5 - sehr große Zustimmung**
kein Abitur	12,8	7,8	19,9	22,1	37,4
Abitur	6,5	6	20,9	24,9	41,7

Tab. A 31: Einstellung zur Ausgereiftheit von Innovationen (getrennt nach Bildungsabschluss)

	Aussage: Mir ist es wichtig, dass Innovationen durch Big Data möglichst bald im öffentlichen Personenverkehr eingesetzt werden. Dafür würde ich notfalls auch in Kauf nehmen, dass diese Innovationen noch nicht komplett ausgereift und zuverlässig sind.				
Bildung (N=1000)	**1 - sehr große Ablehnung**	**2**	**3**	**4**	**5 - sehr große Zustimmung**
kein Abitur	27,9	16,8	31,9	15,9	7,5
Abitur	20,4	17,8	30,7	21,6	9,5

Tab. A 32: Geschätzte Bedeutung des Szenarios zu personalisiertem Entertainment- und Informationsangebot (geordnet nach Altersgruppen)

Altersgruppe (N=1000)	**1 - sehr unwichtig**	**2**	**3**	**4**	**5 - sehr wichtig**
0-29 Jahre	12,9	11	35,6	24,5	16
30-59 Jahre	21,4	16,1	31,8	20,8	9,9
60+ Jahre	44,1	18,6	21,7	9	6,5

Tab. A 33: Geschätzte Bedeutung des Szenarios zu personalisierter Sitzplatzreservierung (geordnet nach Altersgruppen)

Altersgruppe (N=1000)	**1 - sehr unwichtig**	**2**	**3**	**4**	**5 - sehr wichtig**
0-29 Jahre	12,9	13,5	34,4	27	12,3
30-59 Jahre	20,6	12,8	32	23,3	11,3
60+ Jahre	29,8	15,8	32	14,9	7,5

Tab. A 34: Geschätzte Bedeutung des Szenarios zur Anpassung der Anzahl der eingesetzten Fahrzeuge an das Fahrgastaufkommen (getrennt nach Bildungsabschluss)

Bildung (N=1000)	1 - sehr unwichtig	2	3	4	5 - sehr wichtig
kein Abitur	8	7,1	23,1	27,9	33,9
Abitur	4	5	19,6	35,2	36,2

Tab. A 35: Geschätzte Bedeutung des Szenarios eines personalisierten Entertainment- und Informationsangebots im WLAN des ÖPV (geordnet nach Stadtgröße)

Stadtgröße (N=1000)	1 - sehr unwichtig	2	3	4	5 - sehr wichtig
bis unter 20.000 Einwohner	21,5	16,5	29,8	20,7	11,6
20.000 bis unter 50.000 Einwohner	21,1	13,7	27,4	27,4	10,5
50.000 bis unter 100.000 Einwohner (Kern, Verdichtungs- und Übergangsbereich)	32,1	14,3	23,8	19	10,7
100.000 - unter 500.000 Einwohner (Verdichtungs- und Übergangsbereich)	30	16,4	26,4	15	12,1
100.000 - unter 500.000 Einwohner (Kernbereich)	24,6	18	35,3	12,6	9,6
500.000+ Einwohner (Verdichtungs- und Übergangsbereich)	30,5	24,4	25,6	14,6	4,9
500.000+ Einwohner (Kernbereich)	29,6	13,8	29,9	17,7	9

Tab. A 36: Nutzungsänderung bei Realisierung des Szenarios zu personalisiertem Entertainment- und Informationsangebot (geordnet nach Altersgruppen)

Altersgruppe (N=1000)	1 - geringere Nutzung	2	3	4	5 - häufigere Nutzung
0-29 Jahre	6,1	5,5	50,9	23,3	14,1
30-59 Jahre	15,1	8,2	50,9	16,3	9,5
60+ Jahre	25,2	11,2	53,7	5,9	4

Tab. A 37: Nutzungsänderung bei Realisierung des Szenarios zu personalisiertem Entertainment- und Informationsangebot (getrennt nach Geschlecht)

Geschlecht (N=1000)	1 - geringere Nutzung	2	3	4	5 - häufigere Nutzung
Weiblich	18,5	9,7	52,8	10,9	8
Männlich	15,2	7,6	50,7	17,5	9

Tab. A 38: Nutzungsänderung bei Realisierung des Szenarios zu flexiblen Busrouten (getrennt nach Geschlecht)

Geschlecht (N=1000)	1 - geringere Nutzung	2	3	4	5 - häufigere Nutzung
Weiblich	11,3	7	47	18,9	15,8
Männlich	8,8	5,1	45,6	26,5	14

Tab. A 39: Nutzungsänderung bei Realisierung des Szenarios zu an die aktuelle Nachfrage angepassten Preisen (getrennt nach Geschlecht)

Geschlecht (N=1000)	1 - geringere Nutzung	2	3	4	5 - häufigere Nutzung
Weiblich	26,3	17,3	37,8	11,7	6,8
Männlich	24,2	11,7	39	16,6	8,4

Tab. A 40: Nutzungsänderung bei Realisierung des Szenarios zu flexiblen Busrouten (geordnet nach Stadtgröße)

Stadtgröße (N=1000)	1 - geringere Nutzung	2	3	4	5 - häufigere Nutzung
bis unter 20.000 Einwohner	9,1	7,4	43,8	36,4	27,3
20.000 bis unter 50.000 Einwohner	16,8	5,3	45,3	20	12,6
50.000 bis unter 100.000 Einwohner (Kern, Verdichtungs- und Übergangsbereich)	19	4,8	39,3	21,4	15,5
100.000 - unter 500.000 Einwohner (Verdichtungs- und Übergangsbereich)	14,3	7,1	46,4	16,4	15,7
100.000 - unter 500.000 Einwohner (Kernbereich)	10,2	6,6	45,5	24,6	13,2
500.000+ Einwohner (Verdichtungs- und Übergangsbereich)	9,8	6,1	47,6	28	8,5
500.000+ Einwohner (Kernbereich)	4,2	5,5	49,5	24,1	16,7

Anhang 2: Leitlinien für das Interview zum Thema „Big Data und öffentlicher Personenverkehr“

- 1) Aktueller Stand des Einsatzes von BD im öffentlichen Personenverkehr
    - In welchen Bereichen wird BD bisher genutzt?
    - In welchem Ausmaß wird BD genutzt?
    - Welche Player sind dabei beteiligt (z.B. Datenerhebung durch Drittanbieter)?
- 2) Wie sehen Sie die Perspektive im ÖPV bezüglich BD in 2-3 Jahren?
- 3) Spezifische/wichtigste Probleme aus Sicht des Verbands beim Einsatz von BD im öffent lichen Personenverkehr (Jetzt und in Zukunft):
    - Datenschutz?
    - Datensicherheit?
    - Probleme bei der Analyse und Anwendung (VBB)
- 4) Chancen aus Sicht des Verbands bzw. des VBB bzgl. des Einsatzes von Big Data im öffentlichen Personenverkehr
    - Wie könnten sich die Anbieter verbessern? (VBB)
    - Wie könnten Nutzer/Verbraucher profitieren?
- 5) Forderungen des Verbands bzw. des VBB bzgl. des Einsatzes von Big Data im öffentlichen Personenverkehr an die Politik bzw. an die Anbieter:
    - Wie ist die Position des Verbandes bzw. des VBB zu BD?
    - Was wünschen sich die Verbände von den Anbietern und der Politik?
    - Was wünschen sich die Anbieter von der Politik?

Anhang 3: Zusammenfassung der Ergebnisse der Fachgespräche zur Gestaltung des Fragebogens

Die drei Fachgespräche mit zwei Vertretern der ÖPV-Nutzer – dem Fahrgastverband PROBAHN sowie dem Verbraucherzentrale Bundesverband - und einem ÖPV-Anbieter – dem Verkehrsverbund Berlin-Brandenburg – lieferten erste Erkenntnisse über die Einschätzung von Big Data im Bereich öffentlicher Personenverkehr. Diese konnten als Grundlage für die Gestaltung des Fragebogens genutzt werden.

Es wurde bei den drei Fachgesprächen klar, dass die Experten unter Big Data die Sammlung großer Datenmengen verstehen, wobei noch kein konkreter Nutzungszweck im Zuge der Sammlung verfolgt wird. In allen Gesprächen wurde darauf hingewiesen, dass die Relevanz von Big Data-Methoden in Zukunft als sehr groß eingeschätzt wird. Es wurde von mehreren Seiten darauf hingewiesen, dass die ÖPV-Anbieter inzwischen schon sehr viele Pilotprojekte zu Big Data fördern. Ein mehrfach genanntes Beispiel war hierbei die Nutzung der gesammelten anonymisierten Daten eines Telefonanbieters zur Bestimmung des Fahrverhaltens von Kunden des ÖPV durch die Ortung der Mobiltelefone. Jedoch waren die Experten der Meinung, dass die Umsetzung und Anwendung von Big Data-Methoden in der Breite im ÖPV noch 5-10 Jahre dauern wird.

Bezüglich der Probleme beim Einsatz von Big Data im ÖPV wurden vor allem das Spannungsverhältnis zwischen Datenschutz und Nutzung von Big Data genannt. Einerseits wird die Individualisierung der Angebote für die Nutzer angestrebt, andererseits ist das nur mit der Preisgabe einer Vielzahl von Daten möglich. Hier können unzureichende oder unklare Datenschutzpraktiken aus Sicht der Verbrauchervertreter als Problem empfunden werden. Eine vollständige Anonymisierung der Daten wird spätestens beim Verschneiden vieler Datensätze aus Sicht eines Experten nicht mehr möglich sein. Für den effizienten Einsatz von Big Data wird zudem die zerstückelte Landschaft der ÖPV-Anbieter in Deutschland als Problem genannt.

Als Chancen versprechen sich die Experten eine bessere Abstimmung von Angebot und Nachfrage. Eine getrennte Meinung herrscht hier aber bei der Frage, ob sich die Nachfrage an das Angebot angleichen wird oder umgekehrt. Vor allem eine bessere und genauere Information der Nutzer wird kurz- und mittelfristig als große Chance für Big Data im ÖPV gesehen. Als Wunsch an die Politik für eine angemessene zukünftige Entwicklung von Big Data-Technologien werden die Flexibilisierung der Kapazitätsplanung bei Ausschreibungen für Bahnstrecken („goldene Handschellen“), die Wahrung des Datenschutzes und der Datensicherheit sowie die Nutzungsmöglichkeit von Big Data zur Kontrolle der Anbieter (z.B. Pünktlichkeit) genannt.

Anhang 4: Fragebogen

Anmerkung: Der Fragebogen wurde online vom Marktforschungsunternehmen Omniquest im Rahmen der Online-Befragung umgesetzt. Der hier angehängte Fragebogen im Word-Format entspricht inhaltlich dem Online-Fragebogen.

Einführungstext:

Sehr geehrte Teilnehmerin, sehr geehrter Teilnehmer,

vielen Dank für Ihre Teilnahme an unserer Umfrage!

Diese Umfrage wird im Auftrag von ConPolicy – Institut für Verbraucherforschung – umgesetzt.

Bei der Befragung geht es um den Einsatz von Big Data (also der Zusammenführung und Nutzung verschiedenartiger, großer Datenmengen) im öffentlichen Personenverkehr. Hiermit meinen wir Fernzüge, Fernbusse, U-Bahn, S-Bahn, Straßenbahn, Stadtbusse etc.

Dazu werden Ihnen Fragen zu Ihrem Mobilitätsverhalten und Ihrer Einstellung zum öffentlichen Personenverkehr gestellt. Außerdem werden Sie gebeten, verschiedene Szenarien einzuschätzen sowie zu bewerten und Ihre Einstellung zur Anwendung von Big Data im ÖPV anzugeben.

Für die Umfrage sollten Sie sich ca. 20 Minuten Zeit nehmen.

Die Umfrage wird ausschließlich von den Mitarbeiterinnen und Mitarbeitern des Forschungsprojekts bei ConPolicy ausgewertet. Ihre Antworten sind hierbei vollkommen anonym und es sind keine Rückschlüsse auf Ihre Person möglich.

Falls Sie Fragen oder Anmerkungen zu dieser Studie haben, senden Sie uns eine E-Mail.

Öffentlicher Personenverkehr - Nutzung und Bewertung

Frage 1: Wie häufig nutzen Sie die folgenden Verkehrsmittel?

Verkehrsmittel	Häufigkeit
• Fahrrad (ohne Sharing)	Bitte anklicken
• Bikesharing	Bitte anklicken
• Privater PKW	Bitte anklicken
• Taxi	Bitte anklicken
• Car-Sharing	Bitte anklicken
• Motorrad, Roller, Mofa (ohne Sharing)	Bitte anklicken
• Roller-Sharing	Bitte anklicken
• Öffentlicher Nahverkehr (U-Bahn, Bus, Straßenbahn, S-Bahn)	Bitte anklicken
• Öffentlicher Fernverkehr mit der Bahn (ab 100km einfache Strecke)	Bitte anklicken
• Öffentlicher Fernverkehr mit dem Bus (ab 100km einfache Strecke)	Bitte anklicken

(Skala (von oben nach unten: „(fast) täglich"; „1-3 Tage pro Woche"; „1-3 Tage pro Monat"; „seltener als monatlich"; „nie")

Frage 2: Sie haben angegeben, die folgenden Verkehrsmittel zu nutzen. Wie zufrieden sind Sie mit diesen Verkehrsmitteln?

Verkehrsmittel	Zufriedenheit
• Fahrrad (ohne Sharing)	Bitte anklicken
• Bikesharing	Bitte anklicken

• Privater PKW	Bitte anklicken
• Taxi	Bitte anklicken
• Car-Sharing	Bitte anklicken
• Motorrad, Roller, Mofa (ohne Sharing)	Bitte anklicken
• Roller-Sharing	Bitte anklicken
• Öffentlicher Nahverkehr (U-Bahn, Bus, Straßenbahn, S-Bahn)	Bitte anklicken
• Öffentlicher Fernverkehr mit der Bahn (ab 100km einfache Strecke)	Bitte anklicken
• Öffentlicher Fernverkehr mit dem Bus (ab 100km einfache Strecke)	Bitte anklicken

(Skala von „sehr unzufrieden“ bis „sehr zufrieden“ auf einer Geraden mit 5 Ausprägungen unter (oder neben) dem Item. Nur die Endausprägungen werden angezeigt.)

Öffentlicher Personenverkehr – Eigenschaften

Frage 3: Wie zufrieden sind Sie mit folgenden Aspekten im öffentlichen Personenverkehr (ÖPV) im Allgemeinen? (Hinweis für die Befragten: Zum öffentlichen Personenverkehr gehören der öffentliche Personennahverkehr (U-Bahn, Bus, Straßenbahn, S-Bahn) in Ihrer Region, der öffentliche Personenfernzugverkehr und der öffentliche Personenfernbusverkehr.)

Aspekte	Zufriedenheit
• Fahrplaninformationen	Bitte anklicken
• Einfachheit des Ticketkaufs	Bitte anklicken
• Preis für Tickets	Bitte anklicken
• Häufigkeit der Abfahrten/Taktung	Bitte anklicken
• Erreichbarkeit von Haltestellen	Bitte anklicken
• Zustand der Haltestellen (z.B. von Liften, Bänken, etc.)	Bitte anklicken
• Funktionstüchtigkeit der Einrichtung in den Fahrzeugen (z.B. von Sitzen, Türen, Licht, etc.)	Bitte anklicken
• Anzahl der Sitzplätze im Verkehrsmittel	Bitte anklicken
• Sicherheitsgefühl	Bitte anklicken
• Sauberkeit	Bitte anklicken
• Information bei Verspätungen	Bitte anklicken
• Anzahl der Anschlussmöglichkeiten	Bitte anklicken
• Umsteigezeiten bei den Anschlussmöglichkeiten	Bitte anklicken
• Beschwerdemöglichkeiten	Bitte anklicken
• Umgang mit Rückerstattungsforderungen	Bitte anklicken
• Bearbeitungsdauer von Rückerstattungen	Bitte anklicken

(Skala von „sehr unzufrieden“ bis „sehr zufrieden“ auf einer Geraden mit 5 Ausprägungen unter (oder neben) dem Item sowie „keine Angabe möglich“. Nur die Endausprägungen werden angezeigt. „Keine Angabe möglich“.)

Digitale Angebote – Wissen und Nutzung

Frage 4: Welche der folgenden digitalen Angebote kennen und nutzen Sie?

(Anmerkung: Nur wenn „Kenntnis“ mit „Ja“ beantwortet wird, erscheint „Nutzung“ daneben.)

Digitale Angebote	Kenntnis?	Nutzung?
• Routenplanung über Websites oder Apps	Bitte anklicken	Bitte anklicken
• Buchung von Online-Tickets über Websites oder Apps der Anbieter des öffentlichen Personenverkehrs	Bitte anklicken	Bitte anklicken
• Suche nach Sparangeboten für Bahn- oder Busreisen auf digitalen Plattformen von Drittanbietern (mytrain.de, Sparpreisfinder.de, Itur.com etc.)	Bitte anklicken	Bitte anklicken
• Kontaktloses Bezahlen (NFC) in Fahrzeugen z.B. mit der Kreditkarte oder dem Handy	Bitte anklicken	Bitte anklicken
• WLAN in Verkehrsmitteln oder Haltestellen von Anbietern des öffentlichen Personenverkehrs	Bitte anklicken	Bitte anklicken
• Online-Auftritt, Twitter-Auftritt, Facebook-Auftritt, etc. von Anbietern des öffentlichen Personenverkehrs	Bitte anklicken	Bitte anklicken

(Antwortmöglichkeiten jeweils „Ja“ und „Nein“.)

Big Data – Bedeutung im öffentlichen Personenverkehr

Einführender Text zum Dateneinsatz, der zunächst einzeln auf einer Seite gezeigt wird:

Unter „Big Data“ versteht man die Erhebung, Auswertung und Zusammenführung großer, verschiedenartiger Datenquellen. Dadurch können neue Angebote im öffentlichen Personenverkehr geschaffen und dessen Attraktivität gesteigert werden.

Bei den hierfür verwendeten Daten kann es sich um Fahrzeugdaten (z.B. Fahrtdaten oder Wartungsdaten), personenbezogene Nutzerdaten (z.B. Bewegungsdaten oder Merkmale der Person wie Alter, etc.) oder Daten von Dritten (z.B. Wetterdaten oder Daten von Großereignissen wie Fußballspielen, Konzerten, etc.) handeln. Gerade die Verwendung von personenbezogenen Daten ist hierbei umstritten. Es gibt Vor- und Nachteile.

Im Folgenden werden Ihnen nacheinander einzelne Szenarien vorgestellt, die durch Big Data in Zukunft verwirklicht werden könnten.

Frage 5:

Bitte schätzen Sie die Wichtigkeit des folgenden Szenarios für den öffentlichen Personenverkehr allgemein ein. Geben Sie bitte auch an, ob Sie den öffentlichen Personenverkehr dann häufiger oder seltener nutzen würden.

(Anmerkungen: Alle Szenarien werden einzeln und randomisiert dargestellt. Über dem Szenario-Text befindet sich die Fragestellung von Frage 5, unter dem Szenario-Text befinden sich jeweils die beiden Fragen mit Schiebereglern.)

Szenario 1:
Big Data wird dafür genutzt, Fahrgästen eine individuelle Reiseplanung zu ermöglichen. Die vorgeschlagene Verbindung passt sich in Echtzeit an das Verkehrsaufkommen an und berücksichtigt dabei unterschiedliche Verkehrsmittel (z.B. Bus, Bahn, Carsharing, Bikesharing). Dadurch wird Ihnen also während Ihrer Reise immer die aktuell beste Verbindung auf Ihrem Weg von A nach B angezeigt.

Wie wichtig wäre aus Ihrer Sicht eine solche Neuerung für den öffentlichen Personenverkehr?

(Skala von „sehr unwichtig" bis „sehr wichtig" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Wie würde sich Ihre persönliche Nutzung des öffentlichen Personenverkehrs dann ändern?

(Skala von „geringere Nutzung" bis „häufigere Nutzung" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Szenario 2:
Big Data wird dafür genutzt, mit einer einzigen App alle für die Reiseroute erforderlichen Verkehrsmittel (z.B. Bus, Bahn, Carsharing, Bikesharing) buchen und bezahlen zu können.

Wie wichtig wäre aus Ihrer Sicht eine solche Neuerung für den öffentlichen Personenverkehr?

(Skala von „sehr unwichtig" bis „sehr wichtig" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Wie würde sich Ihre persönliche Nutzung des öffentlichen Personenverkehrs dann ändern?

(Skala von „geringere Nutzung" bis „häufigere Nutzung" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Szenario 3:
Big Data wird dafür genutzt, die Anzahl an eingesetzten Fahrzeugen im öffentlichen Personenverkehr in Echtzeit flexibel an das Fahrgastaufkommen anzupassen. Es fahren also mehr Busse und Bahnen, wenn die Nachfrage groß ist und weniger, wenn die Nachfrage gering ist.

Wie wichtig wäre aus Ihrer Sicht eine solche Neuerung für den öffentlichen Personenverkehr?

(Skala von „sehr unwichtig" bis „sehr wichtig" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Wie würde sich Ihre persönliche Nutzung des öffentlichen Personenverkehrs dann ändern?

(Skala von „geringere Nutzung" bis „häufigere Nutzung" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Szenario 4:
Big Data wird dafür genutzt, Routen von Buslinien flexibel an die Nachfrage an Haltestellen anzupassen. Busse fahren Haltestellen an, wenn per Datenübertragung übermittelt wird, dass an der Haltestelle Fahrgäste warten.

Wie wichtig wäre aus Ihrer Sicht eine solche Neuerung für den öffentlichen Personenverkehr?

(Skala von „sehr unwichtig" bis „sehr wichtig" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Wie würde sich Ihre persönliche Nutzung des öffentlichen Personenverkehrs dann ändern?

(Skala von „geringere Nutzung" bis „häufigere Nutzung" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Szenario 5:
Big Data wird dafür genutzt, Fahrgast-Gruppen mit unterschiedlichen Bedürfnissen bei der Reservierung von Sitzplätzen zu trennen (z.B. Geschäftsreisende gegenüber Familien). Dies geschieht durch die Berücksichtigung von Fahrgastprofilen.

Wie wichtig wäre aus Ihrer Sicht eine solche Neuerung für den öffentlichen Personenverkehr?

(Skala von „sehr unwichtig" bis „sehr wichtig" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Wie würde sich Ihre persönliche Nutzung des öffentlichen Personenverkehrs dann ändern?

(Skala von „geringere Nutzung" bis „häufigere Nutzung" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Szenario 6:
Big Data wird dafür genutzt, die Funktionsfähigkeit der Ausstattung von Bussen und Bahnen durch eine datenbasierte Optimierung der Wartung (z.B. Echtzeit-Information über defekte Türe, Toilette etc. durch einen Sensor) zu verbessern.

Wie wichtig wäre aus Ihrer Sicht eine solche Neuerung für den öffentlichen Personenverkehr?

(Skala von „sehr unwichtig" bis „sehr wichtig" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Wie würde sich Ihre persönliche Nutzung des öffentlichen Personenverkehrs dann ändern?

(Skala von „geringere Nutzung" bis „häufigere Nutzung" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Szenario 7:
Big Data wird dafür genutzt, den Fahrgästen ein personalisiertes Entertainment- und Informationsangebot (z.B. Anzeigen von Fußball-News für Fußball-Fans, Musikvideos der jeweiligen Lieblingsband) im WLAN des öffentlichen Personenverkehrs anzubieten.

Wie wichtig wäre aus Ihrer Sicht eine solche Neuerung für den öffentlichen Personenverkehr?

(Skala von „sehr unwichtig" bis „sehr wichtig" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Wie würde sich Ihre persönliche Nutzung des öffentlichen Personenverkehrs dann ändern?

(Skala von „geringere Nutzung" bis „häufigere Nutzung" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Szenario 8:
Big Data wird dafür genutzt, an das aktuelle Fahrgastaufkommen angepasste Preise im öffentlichen Personenverkehr anzubieten. Das bedeutet, dass die Preise steigen, wenn mehr Leute fahren möchten. Wenn weniger Leute fahren möchten, dann sinken die Preise hingegen.

Wie wichtig wäre aus Ihrer Sicht eine solche Neuerung für den öffentlichen Personenverkehr?

(Skala von „sehr unwichtig" bis „sehr wichtig" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Wie würde sich Ihre persönliche Nutzung des öffentlichen Personenverkehrs dann ändern?

(Skala von „geringere Nutzung" bis „häufigere Nutzung" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Szenario 9:
Big Data wird dafür genutzt, je nach Fahrgast und seiner individuellen Zahlungsbereitschaft unterschiedliche Preise für eine Fahrt anzubieten. Das bedeutet, dass der Fahrpreis beispielsweise aufgrund von Merkmalen wie Einkommen, verfügbarer Zeit oder Nutzungsverhalten individuell berechnet werden könnte.

Wie wichtig wäre aus Ihrer Sicht eine solche Neuerung für den öffentlichen Personenverkehr?

(Skala von „sehr unwichtig" bis „sehr wichtig" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Wie würde sich Ihre persönliche Nutzung des öffentlichen Personenverkehrs dann ändern?

(Skala von „geringere Nutzung" bis „häufigere Nutzung" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Szenario 10:
Big Data wird dafür genutzt, dass den Fahrgästen im öffentlichen Personenverkehr bei entsprechend großen Verspätungen automatisch Rückerstattungen oder Entschädigungen auf ihr Konto überwiesen werden, ohne dass es dafür einer Antragsstellung bedarf.

Wie wichtig wäre aus Ihrer Sicht eine solche Neuerung für den öffentlichen Personenverkehr?

(Skala von „sehr unwichtig" bis „sehr wichtig" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Wie würde sich Ihre persönliche Nutzung des öffentlichen Personenverkehrs dann ändern?

(Skala von „geringere Nutzung" bis „häufigere Nutzung" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Big Data – Einstellungen

Frage 6: Bitte geben Sie Ihre Zustimmung bzw. Ablehnung zu den folgenden Aussagen zum Einsatz von Big Data zur Ermöglichung von Innovationen (wie z.B. die zuvor gezeigten Szenarien) im öffentlichen Personenverkehr an.

Aussage	Zustimmung
• Es ist in Ordnung, dass personenbezogene Daten von den Betreibern des öffentlichen Personenverkehrs verarbeitet und ggf. an Dritte weitergegeben werden, wenn dafür Innovationen durch Big Data im öffentlichen Personenverkehr ermöglicht werden können.	Bitte anklicken
• Beim Einsatz von Big Data für Innovationen im öffentlichen Personenverkehr müssen Nutzer bestimmte Risiken bei der Datenübertragung und Datensicherheit in Kauf nehmen.	Bitte anklicken
• Ich bin bereit, mehr Geld für den öffentlichen Personenverkehr zu zahlen, wenn dadurch Innovationen durch Big Data ermöglicht werden können.	Bitte anklicken
• Eine Benachteiligung von bestimmten Nutzergruppen (z.B. ältere Personen oder Personen ohne Internetzugang) muss ausgeschlossen werden, wenn Innovationen durch Big Data eingeführt werden.	Bitte anklicken
• Mir ist es wichtig, dass Innovationen durch Big Data möglichst bald im öffentlichen Personenverkehr eingesetzt werden. Dafür würde ich notfalls auch in Kauf nehmen, dass diese Innovationen noch nicht komplett ausgereift und zuverlässig sind.	Bitte anklicken

(Skala von „sehr geringe Zustimmung" bis „sehr große Zustimmung" auf Schieberegler mit 5 Stufen. Nur die Endausprägungen werden angezeigt.)

Soziodemographischen Variablen

Frage 7: Bitte geben Sie ihr Geburtsjahr an.

Frage 8: Welches Geschlecht haben Sie?

Frage 9: Wie ist die Postleitzahl Ihres derzeitigen Wohnorts?

Frage 10: Was ist ihr höchster Bildungsabschluss?

Frage 11: Was ist ihr Beruf?