IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 24 May 2024, accepted 30 July 2024, date of publication 16 August 2024, date of current version 2 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3445166

RESEARCH ARTICLE

A Conceptual Framework for Mobility Data Science

ALEXANDER STOCKER[1], CHRISTIAN KAISER[2], GERNOT LECHNER[3], AND MICHAEL FELLMANN[4]

[1]Virtual Vehicle Research GmbH, 8010 Graz, Austria
[2]KTM AG, 5230 Mattighofen, Austria
[3]Institute für Operations und Information Systems, University of Graz, 8010 Graz, Austria
[4]Institute of Informatics, Chair of Business Information Systems, University of Rostock, 18055 Rostock, Germany

Corresponding author: Alexander Stocker (alexander.stocker@v2c2.at)

This work was supported by European Commission under the Promoting and Incentivising Federated, Trusted, and Fair Sharing and Trading of Interoperable Data Assets (PISTIS) Project under Grant 101093016.

ABSTRACT The rapid digitalization of the mobility and transport ecosystem generates an escalating volume of data as a by-product, presenting an invaluable resource for various stakeholders. This mobility and transport data can fuel data-driven services, ushering in a new era of possibilities. To facilitate the development of these digitalized mobility services, we propose a novel conceptual framework for Mobility Data Science. Our approach seamlessly merges two distinct research domains: 1) mobility and transport science, and 2) data science. Mobility Data Science serves as a connective tissue, bridging the digital layers of physical mobility and transport artefacts such as people, goods, transport means, and infrastructure with the digital layer of data-driven services. In this paper, we introduce our conceptual framework, shaped by insights from domain experts deeply immersed in the mobility and transport ecosystem. We present a practical application of our framework in guiding the implementation of a driving style detection service, demonstrating its effectiveness in translating theoretical concepts into real-world solutions. Furthermore, we validate our framework's versatility by applying it to various real-world cases from the scientific literature. Our demonstration showcases the framework's adaptability and its potential to unlock value by harnessing mobility and transport data, enabling the creation of impactful data-driven services. We believe our framework offers valuable insights for researchers and practitioners: It provides a structured approach to comprehend and leverage the potential of mobility and transport data for developing impactful data-driven services, which we refer to as digitalized mobility services.

INDEX TERMS Mobility data science, mobility and transport, data science, digitalized mobility services, digitalization, digital innovation, conceptual framework.

I. INTRODUCTION

The evolution of the mobility and transport landscape is marked by transformative changes, giving rise to what we term as *Mobility Data Science*. This emerging approach is shaped by a confluence of factors, notably the swift progress in technology, accelerated innovation, the ubiquity of digital technologies, the surge in available mobility and transport data, and the establishment of interconnected mobility and transport data ecosystems (cf. e.g., [1], [2], [3], [4]). In the following subsections, we delve into these developments, before providing an in-depth exploration of Mobility Data Science in Section II.

The associate editor coordinating the review of this manuscript and approving it for publication was Yanli Xu.

A. DIGITALIZATION AND PHYSICAL MOBILITY

Rapid technological advancement and the pursuit of innovation are driving society more and more towards digitalization. *Digital technology*, as highlighted by various researchers [1], [4], [5], [6], holds immense potential for innovation across all fields. A key aspect of this potential lies in the creation of hybrid product architectures, combining the modular architecture of the physical product with layered digital components. These digital layers extend the capabilities of physical products through sensors and software-based functionalities, paving the way for new business models that enhance the interconnectedness of products, processes, and services [7]. Furthermore, technology-based innovation contributes to a shift in entrepreneurial culture [3]. While attaining the transformative impact of digital technologies requires time, effort, and determination [8], these technologies hold the potential to revolutionize mobility and transportation across all facets—actors, products, services, and business models [9], [10].

In the realm of mobility and transport, traditionally focused on *physical movement of people and goods*, a notable shift is occurring as physical assets such as passenger cars, trucks, buses, and trains become enriched with a digital layer. This digitization wave facilitates improved connectivity, user interaction, safety, and comfort. A fundamental change brought about by digital technologies is the evolving concept of ownership, with easy access to products increasingly perceived as the new form of ownership [11]. This shift is further emphasized by a decreasing willingness to pay solely for the potential use of a product and an increasing willingness to pay for the convenient fulfillment of a specific need.

The digital layer serves as a crucial instrument in enhancing access to physical mobility assets and improving their efficiency. This is exemplified by the extended utilization of passenger cars, which can now be in use for up to eight hours or even more per day, a significant contrast to the limited usage patterns of traditional car owners who, on average, drive 1-2 hours per day, if at all [12]. The conventional concept of physical ownership, often associated with substantial resource waste, such as parking spaces occupied by seldom-used vehicles, is being redefined.

The significance of digital technologies becomes even more evident when examining developments over the past decade (cf. [1], [13]). During this period, digital technology has been a catalyst for a diverse array of solutions, including digital extensions of physical products, exemplified by innovations like vehicle information systems [14] and automated driving systems [15]. These advancements have not only introduced new functionalities but have also played a pivotal role in shaping novel product architectures. Digital technologies have given rise to hybrid product architectures, combining the modular structure of physical products with the multi-layered architecture of digital technology. This synthesis allows the extension of the functionality of physical products through software-based functions [1], [11]. The transformative impact is far-reaching, ushering in a new era of product design and capabilities.

Furthermore, the influence of digital technologies extends to the development of various mobility services, benefiting a range of mobility actors. Examples include car sharing services like ShareNow from BMW Group and Daimler AG, ride-on-demand platforms like Uber, ride-hailing services such as RideConnect, and the advent of self-driving taxis exemplified by Waymo from Google. These examples underscore the undeniable impact of digital technologies on the mobility and transport landscape, emphasizing their indispensable role in shaping the present and future of these domains.

B. MOBILITY DATA AND DATA ECOSYSTEMS

The influence of digital technologies is intricately linked to the accessibility and utilization of data. Within the digital mobility ecosystem, data has emerged as a pivotal component [2], [16]. It stands as a fundamental resource, playing a vital role in powering and optimizing digitalized mobility services [17]. This underscores the transformative role data assumes in shaping the landscape of modern digitalized mobility, where its availability not only drives technological advancements but also serves as a foundational resource for innovative and efficient mobility solutions.

Several sources within the mobility and transport ecosystem generate data: First, data is generated throughout the whole lifecycle of physical and digital mobility assets, i.e., in the design, manufacturing, operation, and service phases. In the *design phase*, physical transport assets such as passenger cars or trucks are designed and therefore engineering data such as computer aided design (CAD), simulation results, or test data are generated by vehicle engineers in a collaborative development effort. In the *manufacturing phase*, data is generated by humans and machines that assemble the product from parts. In the *operation phase*, data is generated by sensors installed in vehicles to ensure the functionality and safety of vehicle operation [18]. Moreover, data processing systems integrated into users' mobile devices, such as smartphones or smartwatches, contribute to data generation by capturing user interactions and vehicle-related information [18], [19]. In the *maintenance phase*, additional data, such as diagnostic fault codes and performed repairs, is gathered, facilitating services like predictive maintenance [20].

Furthermore, the mobility and transport sector is experiencing a progressive shift towards increased interconnectedness, wherein various transport modes actively gather and exchange extensive datasets. Within emerging mobility and transport ecosystems [2], [16], an expanding network of mobility and transport artefacts, such as cars, trains, and the road infrastructure, engage in communication with data processing centers. These centers transmit a diverse array of information, including operational identifiers, location data, and various other mobility, traffic, and infrastructure-related datasets. This trend has given rise to a myriad of services that leverage such data, often in conjunction with additional contextual information.

A notable illustration of these services can be found in the insurance industry, where drivers now can pay premiums based on their individual driving behavior rather than a fixed rate [21]. Another instance is evident in in-vehicle navigation services, which not only provide users with route guidance but also furnish real-time traffic information. Additionally, the transportation infrastructure actively collects data on traffic movements, utilizing roadside units (RSUs) and static sensors on roadways. This data serves as a foundation for services supporting short-term traffic planning processes in newsrooms and long-term infrastructure planning by providers.

Digital technologies also play a pivotal role in delivering services that enhance mobility for individuals with disabilities [22], promote inclusive mobility [23], and address transportation poverty [24].

Although data capturing has become ubiquitous and is actively utilized to facilitate diverse services [11], [17], it is notable that substantial volumes of data are generated merely as by-products of transport artefacts or through the use of mobility services, frequently remaining underutilized and unprocessed [10]. As a result, there arises a pressing need for novel conceptual approaches to fully unlock the potential value embedded within the vast datasets generated within the mobility and transport ecosystem. Addressing this challenge requires innovative strategies that go beyond conventional data utilization, aiming to extract meaningful insights and capitalize on the untapped potential within this rich source of information.

C. A DISCOURSE ON SMART SERVICE SYSTEMS

In accordance with service science, a service is conceptualized as a system comprising interacting and interdependent elements, encompassing individuals, technology, and business activities, all oriented outward to attain and sustain a competitive advantage [25]. In this domain, the evolution of physical products that are digitally networked with other products and information systems, so-called smart products, is being explored as smart service systems [26]. In this context, the term 'digitalized' is being supplanted by 'smart,' emphasizing the 'embedded intelligence' within these new physical products. A smart service system encompasses service providers, consumers, and the smart product itself, serving as boundary objects that facilitate these services.

Over time, service systems have evolved towards greater intelligence, leveraging Big Data analytics to generate information and automate operations, thereby creating enhanced value for individuals [27]. Another prevalent concept in this domain is that of product-service systems [28], representing an integrated blend of products and services wherein customers pay for the utilization of assets rather than outright ownership. Notably, these product-service systems need not necessarily incorporate digital technologies, highlighting the diversity in their implementation and the flexibility in adapting to varying contexts.

While we lean towards the terminology 'digital' over 'smart' to resonate with the discourse within the digital transformation community, it is essential to acknowledge the shared principles that underlie both domains. Smart service systems emphasize intelligent and adaptive solutions, often harnessing technologies such as artificial intelligence and the Internet of Things. In the realm of Mobility Data Science, our focus underscores a commitment to employing data-driven methodologies for the development and enhancement of digitalized mobility services. This approach aligns with the broader goals of leveraging advanced technologies to optimize and transform the landscape of mobility services in a digitally connected world.

By synthesizing insights from the literature on smart service systems with our emphasis on Mobility Data Science, our aim is to enhance our understanding of how digital technologies can revolutionize the landscape of mobility and transportation. This holistic perspective enables us to leverage the principles inherent in smart service systems while tailoring our approach to address the distinct challenges and opportunities within the digitalized mobility domain. In doing so, we strive to develop a nuanced and effective framework that aligns seamlessly with the evolving dynamics of the mobility and transportation sector in the digital era.

In the realm of service science, a digitalized mobility service refers to an advanced and technologically integrated solution within the mobility sector. This service employs digital technologies, data analytics, and connectivity to transform traditional modes of transportation into dynamic and intelligent systems. Digitalized mobility services leverage information and communication technologies to enhance the efficiency, accessibility, and user experience of transportation offerings. This integration extends beyond the mere digitization of processes, involving the comprehensive application of digital technologies to create a seamless, interconnected, and data-driven mobility ecosystem. The aim is to optimize service delivery, improve decision-making processes, and cater to the evolving needs of users and stakeholders within the broader service framework.

In this context, we want to make a distinction between the terms "digitized" and "digitalized." While these terms are often used interchangeably in the literature [59], we specifically refer to "digitalized mobility services" (not "digitized mobility services"). This distinction emphasizes a type of service that not only involves the conversion of analog information into digital form but, more importantly, fulfills a clear need within the realm of mobility and transport, utilizing digital technologies to enhance and optimize services.

D. MOBILITY DATA SCIENCE-A NOVEL APPROACH

The effective utilization of generated mobility and transport data remains a formidable challenge, necessitating a fresh approach for resolution, which we term *Mobility Data Science*. The domain of mobility and transportation presents distinctive features that warrant in-depth exploration and discussion. These unique characteristics contribute to the complexity of handling and deriving meaningful insights from the wealth of data generated within this sector. In the pursuit of overcoming these challenges, Mobility Data Science emerges as a specialized discipline that seeks to harness advanced analytical techniques, computational methods, and domain-specific expertise to unlock the full potential of mobility and transport data. In our exploration, we aim to delve into these distinctive features, shedding light on the nuances that make Mobility Data Science a crucial catalyst for transformative advancements within the realm of mobility and transportation.

Mobility and transport constitute a complex application domain characterized by a specialized language that experts employ, posing challenges for outsiders to comprehend and interpret. Gaining proficiency in this domain demands a considerable investment of time. Adding to the complexity is the diverse array of data formats, many of which are proprietary and pose challenges for processing and analysis, as highlighted by various experts [29].

Capturing mobility and transport data can be very challenging. For instance, access to crucial data sources like the vehicle CAN bus (Vehicle Controller Area Network) is often restricted to product and embedded system manufacturers due to security considerations. This limitation frequently necessitates additional information to decipher hexadecimal formatted data from the CAN bus. Vehicles typically incorporate multiple bus systems of varying types (CAN, LIN, FlexRay, etc.) serving different purposes such as diagnostics, multimedia, and engine control. These systems process data from an assortment of sensors, contributing to the already substantial and exponentially growing volume of data within the mobility and transport ecosystem [10]. This surge in data is particularly evident in high-tech applications integrating vehicle sensor technology, including driver assistance systems, automated driving, and in-cabin monitoring. The enormity and intricacy of data in this domain underscore the critical role of specialized expertise and advanced data processing techniques in extracting meaningful insights and navigating the challenges. Addressing environmental and social challenges encounters persistent issues related to data availability, as noted in various studies [30], [31]. A critical social concern is the concept of mobility data injustice, where individuals who do not actively generate data find themselves excluded from data collection efforts. This exclusion results in mobility and transportation services that may not cater to their specific needs, contributing to heightened inequalities [32].

Moreover, accessibility disparities between rural and (peri-)urban areas persist, impacting the availability and accessibility of existing mobility solutions, such as public transportation or sharing concepts [33]. These challenges are particularly pronounced in rural areas, despite the potential benefits they could bring to disadvantaged populations [34]. A commitment to inclusive mobility and transportation for all should also extend to individuals with impairments, a principle underscored in Article 20 of the United Nations' "Convention On The Rights Of Persons With Disabilities."

In the context of environmental challenges, there exists a considerable degree of uncertainty surrounding mobility and transport. While the undeniable impact of these sectors on the Earth's climate is acknowledged [35], the inherent complexity of this domain introduces uncertainty regarding the precise environmental implications. This uncertainty spans issues such as the accurate calculation of environmental impact, especially concerning the transition between different transport modes and diverse behavioral patterns [37], or the use of recycled materials like asphalt pavement [38]. These challenges emphasize the need for comprehensive, inclusive, and environmentally conscious strategies in the evolution of mobility and transportation systems.

Furthermore, the mobility and transport sectors are undergoing a transformation into intricate ecosystems, as observed in various studies [10], [39]. This evolution necessitates collaboration among numerous stakeholders to facilitate the emergence of novel products and services. Notably, the landscape is marked by the proliferation of proprietary information systems, and a standardized framework for data access and sharing is yet to be established. The existence of numerous influential players in the ecosystem, such as car manufacturers or Tier-1 suppliers, poses challenges in convincing them to provide access to critical vehicle operation data.

Finally, the mobility and transport domain has witnessed the emergence of numerous start-ups that offer compelling data-driven services [18], [40], particularly in the realm of connected vehicles. These start-ups often find themselves in competition with traditional players, contributing to a dynamic and competitive landscape within the sector. This complex interplay among established entities and emerging innovators underscores the need for effective collaboration frameworks, standardized data access protocols, and strategic approaches to foster innovation while maintaining a balance between traditional and novel players.

II. A BETTER UNDERSTANDING OF MOBILITY DATA SCIENCE

In an initial attempt to better understand *Mobility Data Science*, it can be characterized as the *convergence or synergistic fusion of two burgeoning research domains: Data Science and Mobility and Transport*. Both of these fields are currently experiencing significant growth and stand to benefit substantially from a collaborative and efficient integration.

Data Science, as a discipline, is primarily focused on the development of novel approaches and methodologies, validated through real-world data. Mobility and Transport contribute by providing data scientists with the essential volume, variety, and veracity of data originating from diverse mobility assets and services. Given the central role of mobility and transport in society, they play a pivotal role in achieving the United Nations' sustainable development goals.

The mobility and transport domain actively aspires to enhance sustainability across social, environmental, and economic dimensions, aligning with the triple bottom line (3BL) framework [41]. Data Science emerges as a valuable ally in this transformation process, offering crucial support to advance sustainability goals through the development and optimization of digitalized mobility services. The collaboration between these two domains holds the potential to drive innovations that not only improve the efficiency and effectiveness of mobility solutions but also contribute significantly to broader societal and environmental objectives.

In recent years, Data Science has emerged as a transformative discipline that translates raw data into valuable insights for individuals, organizations, and society at large [42]. The term Data Science encapsulates the systematic study of data, often emphasizing the analysis of extensive amounts of structured and unstructured data, including measurements, texts, images, and videos [43], [44]. According to Waller and Fawcett [45] Data Science is characterized by the application of both quantitative and qualitative methods to address pertinent problems and predict outcomes, acknowledging that expertise and analysis are inherently intertwined.

At a fundamental level, Data Science comprises a set of principles guiding the extraction of information and knowledge from data. However, a proficient data scientist must also possess the ability to approach significant business challenges from a data-centric perspective [46]. The typical workflow of a data scientist involves well-defined practices for extracting information and knowledge, facilitating data-driven decision-making. This process involves stages such as data acquisition, data preprocessing, data transformation, data exploration, data modeling, model evaluation, and finally, data visualization and the application of results. While these stages are not strictly linear, they undergo continuous evaluation, regression, and improvement.

A notable aspect of Data Science is the diverse types of data it deals with, presented in various formats and structures. Consequently, data scientists invest a significant portion of their problem-solving efforts in data preparation and processing, recognizing the critical role this phase plays in ensuring the accuracy and relevance of subsequent analyses and insights [46].

Mobility and Transport present a diverse array of challenges, and Data Science emerges as a crucial tool for addressing them. These challenges span a spectrum from understanding travel behavior and human mobility at a broader scale [47], analyzing bus transit and passenger travel behavior [48], detecting driver behavior [49], monitoring driver activities, and addressing distractions [50], [51], to optimizing resource-efficient mobility solutions.

Moreover, the mobility and transport sector is currently undergoing a significant transformation, reshaping traditional business models and products and services. The concept of "mobility-as-a-service" (MaaS) exemplifies this shift, emphasizing the purchase of mobility services tailored to consumers' needs rather than the acquisition of physical mobility assets [52]. These services are typically facilitated through digital platforms, websites or mobile applications, providing users with comprehensive access to trip planning, ticketing, payment, and real-time information [53].

Furthermore, research has given rise to data-driven applications that enhance mobility and transport infrastructure. Examples include the use of mobile sensors or smartphones for detecting potholes and road surface damage [54], [55], [56], as well as applications for identifying pavement patch defects [57] and determining travel conditions using data from drivers' devices [58]. These innovations highlight the transformative potential of Data Science in shaping the future of mobility and transport, fostering efficiency, sustainability, and a seamless user experience.

In the context of the various applications and services discussed, the application of Data Science methods plays a pivotal role in their development and operation. Considering the significant transformations in mobility and transport over the past decade, such as technological advancements, digitalization, and the formation of mobility and transport ecosystems, ***wedefine Mobility Data Science as the practical implementation of Data Science within the mobility and transport domain***.

Mobility Data Science represents Data Science in action within the mobility and transport domain. It involves the application of tailored methods to enable the development of novel, digitalized mobility services. These specialized methods, referred to as Mobility Data Science methods, form a systematic approach. This approach utilizes the data collected within the mobility and transport system by applying a data science process. The ultimate goal of Mobility Data Science is to facilitate the emergence of digitized mobility services characterized by their data-driven nature, providing added value for users and contributing to the ongoing evolution of the mobility and transport landscape.

These methods intricately process data derived from various *mobility and transportation artefacts*, including people and goods to be transported, *transportation assets* like cars or trains, and the overall *transportation infrastructure*. This data processing, in turn, serves as the catalyst for the development of innovative, *digitalized mobility services*.

To visually capture the essence of Mobility Data Science, Figure 1 outlines its scope. It represents the application of Data Science methods, establishing bidirectional links between the digital layers of mobility and transportation artefacts, such as vehicles or transport infrastructure, and the digital layer of digitalized mobility services.

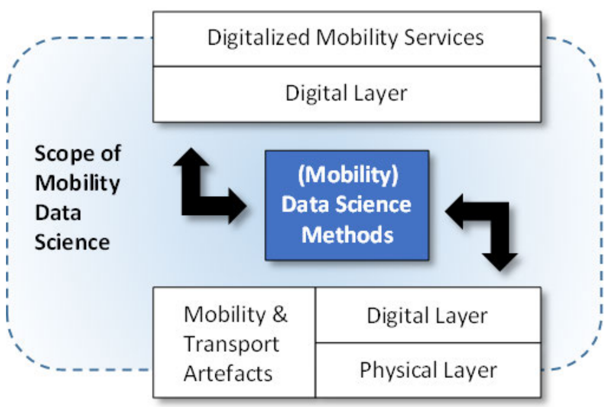

FIGURE 1. Scope of mobility data science: linking the digital layers.

III. A CONCEPTUAL FRAMEWORK FOR MOBILITY DATA SCIENCE

To enhance our comprehension of the scope and constituents of Mobility Data Science, we introduce a conceptual framework (Figure 2). Conceptual frameworks serve as abstract representations of a specific topic, facilitating communication and shared understanding among stakeholders to enhance the likelihood of successful development and utilization [60]. Typically graphical in nature, conceptual frameworks incorporate visual arrangements of modeling constructs in the form of graphical symbols and text [61]. Beyond aiding communication, these models contribute to a deeper understanding of a particular domain and provide insights into the system design process [62].

In accordance with our conceptual framework, we elucidate how Mobility Data Science contributes to both enabling digitalized mobility services and augmenting knowledge about mobility and transport artefacts and their utilization. Mobility Data Science centers on harnessing the value of data derived from mobility and transport assets and services, offering insights and facilitating better data-driven decision-making. Within this framework, the digital layer encompasses both mobility and transport artefacts and digitalized mobility services, generating a plethora of data.

For instance, physical entities to be moved, such as people connected via smartphones or wearables and goods connected through sensors or RFID tags, form an integral part of this digital layer. Physical transportation assets, including bicycles, cars, buses, trains, planes, ships, and potentially drones in the future, are also seamlessly connected to the digital layer through sensors or actuators. This interconnected digital layer forms the foundation for the wealth of data generated, laying the groundwork for effective Mobility Data Science applications.

The knowledge derived from examining how transportation assets are operated plays a pivotal role in shaping product design, ultimately leading to the development of better-designed transportation systems tailored to specific use cases. In this context, the physical transportation infrastructure delineates the space within which entities are moved by physical transportation assets. Similar to physical entities and assets, the physical transport infrastructure is increasingly connected through sensors and actuators, further enhancing the interconnectedness of the entire system.

These three key entities—physical entities to be moved, physical transportation assets, and the physical transport infrastructure—collectively generate and utilize data. This data serves as invaluable input for mobility data scientists, empowering them to enable and implement novel digitalized mobility services. During the operational phase of digitalized mobility services, substantial amounts of data are generated, offering insights into various aspects of the transportation infrastructure, such as capacity utilization, traffic flows, and more. In essence, this reciprocal flow of data and knowledge between entities, assets, infrastructure, and digitalized mobility services is dynamic. It not only fuels the development and enhancement of digitalized mobility services but also contributes to an iterative process of refining transportation systems and infrastructure design, ultimately fostering a more efficient and informed mobility and transport landscape.

We present our conceptual framework for Mobility Data Science in Figure 2 and describe all model elements in detail in the following subsections.

A. MOBILITY & TRANSPORT ARTEFACTS

The shared essence among various definitions of the term mobility lies in the ability to move or be moved freely, signifying its significant impact on society. Litman [63] specifically characterizes mobility as the physical movement of people or goods, involving the act of taking or carrying them from one location to another, whether by means of a vehicle, aircraft, ship, or on foot. In the context of digitalized mobility, we conceptualize a group of physical components, potentially extended by a digital layer, as "Mobility & Transport Artefacts." In terms of digitalized mobility, we refer to group of physical components – whether or not extended by a digital layer – as *Mobility & Transport Artefacts*, consisting of *Physical Entities to be transported*, *Physical Means of Transport*, and *Physical Transport Infrastructure* [64].

Physical entities to be transported encompass people or goods intended for movement. In contemporary scenarios, these entities are increasingly extended and complemented by a digital layer. For individuals, this augmentation is often manifested through personal devices such as smartphones or smartwatches. In the case of goods, radio-frequency identification (RFID) tags and specialized sensors contribute to this digital layer. The ubiquity of personal devices among mobility users means that journeys from point A to point B are frequently recorded, whether intentionally (as seen in quantified self-applications) or unintentionally (e.g., for improving mapping services by providing the user's position in the global navigation satellite system - GNSS).

Beyond GNSS capabilities, smartphones, and wearable devices are equipped with a variety of sensors (acceleration, gyroscope, temperature), rendering them rich sources of data. Additionally, bonus cards with Near Field Communication (NFC) functionality or RFID tags serve as representatives of the digital layer for individuals, concurrently providing diverse data and potential information. Analogously, goods in transit or shipment also integrate a digital layer. In logistics, standards such as barcodes or RFID have been developed to track and manage the movement process. Other approaches focus on the real-time determination of the state of goods, employing tools like acceleration sensors [65]. This integration of digital layers enhances the visibility, traceability, and data richness associated with both individual and goods mobility.

Transport serves as the instrumental means for the tangible realization of mobility. The second artefact, *physical means of transport*, encompasses all entities utilized to transport people and goods. These modes can range from privately owned or shared passenger vehicles, taxis, buses, trucks, bicycles, trains, planes, cargo vessels, to pipelines—each playing a role in facilitating physical movement. Like entities to be transported, physical means of transport are frequently

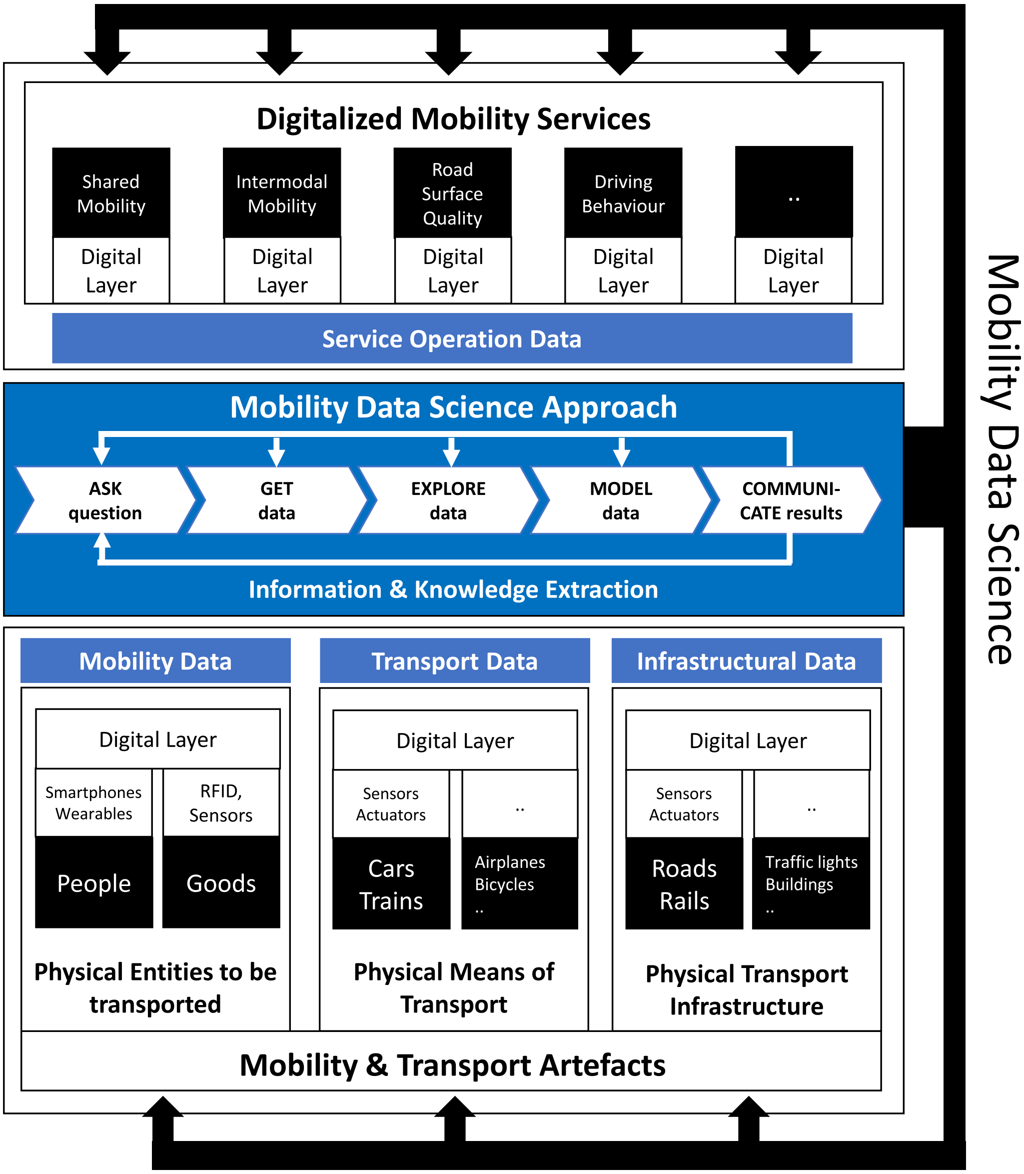

FIGURE 2. **A conceptual framework for mobility data science.**

equipped with or extended by sensors or actuators, enabling the translation of analog relationships into digital data.

For instance, in the case of passenger cars or trucks, generated data includes information such as vehicle speed, acceleration and braking patterns, position, catalytic converter status, and potentially videos from dash cams or even radar and/or lidar data [66]. Conversely, data about bicycles, such as speed, position, and unusual acceleration in the case of evasive actions, can be collected using the cyclists' smartphones, for example [67]. In the current era of technology-driven mobility and transport systems, characterized by a push for increased automation, further integration of sensors and actuators in transportation assets is anticipated. This integration aims to capture contextual information, draw conclusions, and make decisions based on the acquired data. Consequently, this ongoing development is expected to result in a growing number of potential data sources and an increase in data volume.

Applications based on mobility and transport data are diverse and include areas such as maintenance for transportation assets [20] or real-time traffic planning [68]. As technology continues to advance, the wealth of data generated by transportation assets presents numerous opportunities for improving efficiency, safety, and overall performance within the mobility and transport ecosystem.

A crucial enabling factor for mobility and transport is a well-suited *physical transport infrastructure*. The construction and operation of roads, (air)ports, metros, traffic lights, etc., play a pivotal role in facilitating and shaping mobility and transport. It's evident that the planning of mobility and transport infrastructure, whether in urban or rural areas, is intricately linked to the economy [69], the environment in terms of emissions resulting from mobility and transport [70], or urban sprawl in spatial development [71]. Additionally, it significantly impacts society, as access to transportation influences the health of individuals [72].

The infrastructure within the realm of mobility and transport contributes an increasing amount of data through its digital layer. Traffic lights and traffic counters can be connected in real-time to traffic planning systems, while video and audio data support decision-making processes, especially in terms of planning and safety. Local weather data further enables early information and intervention. In summary, the physical transport infrastructure serves as the third artefact in our conceptual framework for Mobility Data Science. This infrastructure not only forms the backbone of mobility systems but also acts as a dynamic source of data that, when integrated with the digital layer, enhances the efficiency, safety, and overall functioning of the entire mobility and transport ecosystem.

B. MOBILITY DATA SCIENCE METHODS: INGREDIENTS AND PROCESSES

The examples mentioned earlier showcase a vast array of applications and the underlying heterogeneous data in the realm of mobility and transport. Data Science methods play a crucial role in extracting knowledge from this diverse data, forming the foundation for an exchange between digitalized mobility services and mobility and transport artefacts. However, to effectively utilize data from these varied sources, data must undergo a progressive transformation into information, knowledge, and ultimately wisdom in terms of valued understanding [73].

To achieve this transformation, we adopt the Data Science process introduced by Pfister and Blitzstein from Harvard University. This approach, which is also applicable in the context of mobility and transport, typically involves the five steps (1) asking questions, (2) collecting data, (3) examining data, (4) modeling data, and (5) communicating results. These steps may not necessarily follow a linear process and can include single or multiple feedback loops. While other process models such as CRISP-DM [74] or ASUM-DM DM [75] can also be applied in the context of Mobility Data Science, Pfister and Blitzstein's concept is particularly accessible for interdisciplinary teams comprising both domain experts and non-experts. It encourages the clarification of the context before delving into technical and data challenges.

The initial step in this process is to formulate and define a question to be addressed with data. These defined questions heavily influence the acquisition of the required data, guiding which data may be relevant for analysis. When integrating multiple data sources, synchronization becomes essential to enable joint analysis [76]. Furthermore, the relevant data must be processed and stored, irrespective of its size, posing a challenge in the field of mobility and transport where the volume of data can be substantial. This initial phase sets the stage for the subsequent stages in the Data Science process: Several key steps follow the initial question formulation. These steps include data cleaning, normalization, transformation, exploration, modeling, validation, and communication of results:

Data Cleaning involves the removal of irrelevant data and processing of incomplete or noisy data to enhance the quality and reliability of the dataset. *Data Normalization* aims to bring numerical values to a common scale, facilitating consistent comparisons and analyses. *Data Transformation* encompasses the conversion of pre-processed data from one format or structure to another, optimizing it for subsequent analysis. *Data exploration* analyses the collected, preprocessed, and transformed data to identify specific characteristics using methods such as data visualization, tabular reports, or descriptive statistics. Patterns, relationships, and anomalies can be uncovered during this exploration. In the exploration phase, *machine learning* methods can be employed to identify relationships in data that may not be observable with more traditional methods, especially in the context of multidimensionality. The *modeling phase* involves determining whether the available data are sufficient to create a model that can answer the posed question. It includes selecting the types of models to be used, creating data-based models, and validating their effectiveness. In the modeling phase, decisions are made on how to *answer the question posed*, whether through explanatory, predictive, or descriptive analysis. The final step involves summarizing the findings and *communicating the results* obtained from the validated models. This can be achieved through storytelling, visualizations, and applications. Before dissemination, it is crucial to ensure that the results make sense in the specific, often domain-dependent context and effectively answer the desired question.

C. DIGITALIZED MOBILITY SERVICES

Digitalized Mobility Services (DMS), exemplified by platforms like Uber or Flixbus, already leverage information and knowledge derived from Data Science. At a more abstract level, digitalized mobility services [17], [77] add an additional digital layer to physical transport and mobility services, aiming to enhance efficiency, often referred to as smart mobility. In this context, smart mobility is a complex set of projects and actions, different in goals, contents, and technology intensity [78]. We want to highlight the pivotal role of digital technology in achieving the objectives of "smart mobility," including reducing pollution, traffic congestion, noise pollution, transfer costs, and enhancing passenger safety and transfer speed.

The European Commission's roadmap on "Smart Mobility Systems and Services" [79] underscores the importance of the proactive integration of smart mobility services with existing public transport and utility systems in future European innovation actions. Digitalized mobility services, as a critical component, play a crucial role in the shift toward smart mobility by enabling better coordination and decision-making based on processed information.

Researchers are increasingly exploring how digital technologies impact mobility and transport, resulting in various approaches and taxonomies, but without a final and universally agreed-upon conclusion (cf. e.g., [77] or [80]). Following Docherty et al. [81], examples such as shared mobility, on-demand mobility, intermodal mobility, or integrated mobility are introduced as digitalized mobility service categories, though a definitive and validated list is not provided.

Shared mobility involves the sharing of a transportation mode among two or more individuals, such as bicycles or vehicles. Shared mobility platforms typically operate without a predetermined form of ownership concerning shared transportation assets.

On-demand mobility is considered a subset of shared mobility that provides rides on demand for a fee. Subcategories include ride-hailing (including taxis as well as services like Uber or Lyft) and ride-splitting, defined as a ride-sourcing service that matches riders with similar origins or destinations to the same driver and vehicle in real time [82].

Intermodal mobility is a specific form of multimodality. In the context of freight, multimodality refers to the transport of goods using more than one mode (e.g., rail and sea), while intermodality involves transporting goods in the same transport unit, such as a container. For passenger transport, multimodality enables access to multiple modes during a trip, aiming to facilitate seamless travel in a combined travel chain using various transport modes like cars, trains, and bicycles [83].

Integrated mobility refers to the integration of different services, such as information, payment, and multimodality, through a single or common interface. These services aim to optimize personal travel across different transport modes, offering a combination of all available modes (bike, car, bus, train, etc.) based on cost and/or time considerations for specific routes and schedules [80].

These categories highlight the diverse ways in which digitalized mobility services are shaping and transforming the landscape of transportation, offering users more flexible, efficient, and interconnected options for their journeys. It's crucial to highlight that digitalized mobility services not only consume data but can also contribute by providing valuable data for further use. This data may encompass information on the mobility behavior of individual users or aggregated data regarding the type of transport demanded. This two-way data exchange plays a pivotal role in enhancing the overall understanding and optimization of mobility services.

IV. FRAMEWORK APPLICATION

In this section, we apply our conceptual framework for Mobility Data Science. Firstly, we utilize the framework to guide the design of a digitalized mobility service for driving style detection (subsection A). Then, we employ the framework to gain structured insights into existing digitalized mobility services from the literature by organizing them according to our framework (subsection B).

A. APPLICATION OF THE FRAMEWORK TO DESIGN DIGITALIZED MOBILITY SERVICES

The application of Mobility Data Science can result in the creation of innovative data-driven services. While familiar examples such as shared mobility and intermodal mobility were outlined in the previous section, in this section, we will elaborate on the development of a concrete digitalized mobility service designed to enhance driving and road safety. Aggressive driving stands out as a significant cause of accidents, where harsh acceleration and braking can immediately impact driving safety. Detecting such events in vehicle data forms the basis for a digitalized mobility service designed to identify and transparently communicate harsh driving behavior to relevant stakeholders, such as drivers. The ultimate goal is to bring about a positive change in driving behavior and contribute to overall road safety [10].

In the development of such a digitalized mobility service, service developers can tap into a wealth of valuable data from the mobility and transport ecosystem. Human drivers, for instance, bring their smartphones into their vehicles. Given that smartphones are equipped with a diverse array of sensors, they can offer insights into factors like time, location, acceleration, and rotation. Vehicles, serving as physical means of transport, come with a variety of sensors—steering angle sensors, radar, lidar, wheel speed sensors—utilized for vehicle functionality and safety. Furthermore, the physical transport infrastructure is increasingly fitted with technologies like cameras, parking sensors, or roadside units, capable of providing pertinent data such as the speed and distance of passing vehicles, or the road temperature.

To illustrate how a driving style detection service can be implemented (cf. Figure 3), we will conduct Mobility Data Science according to the proposed framework and detail the results in the following subsections.

1) ASK QUESTION

In our initial step, we formulate a pertinent research question aimed at enhancing driving safety: *How can data generated in the mobility and transport ecosystem facilitate the detection of harsh driving styles*? Driving styles refer to habitual driving behaviors characteristic of groups of drivers [85] and represent a significant area of research for advanced vehicle automation in the future. As per Sagberg [85], global driving styles encompass multiple driving indicators (such as aggressive, calm, or careful driving), while specific driving styles are measured by one or two indicators. Various methods, including surveys, driver interviews, online analysis of vehicle data, and offline analysis of vehicle data, can be employed to identify driving styles. In this context, we want to focus our attention on three maneuvers—harsh acceleration, braking, and turning—as indicators of a harsh driving style. These

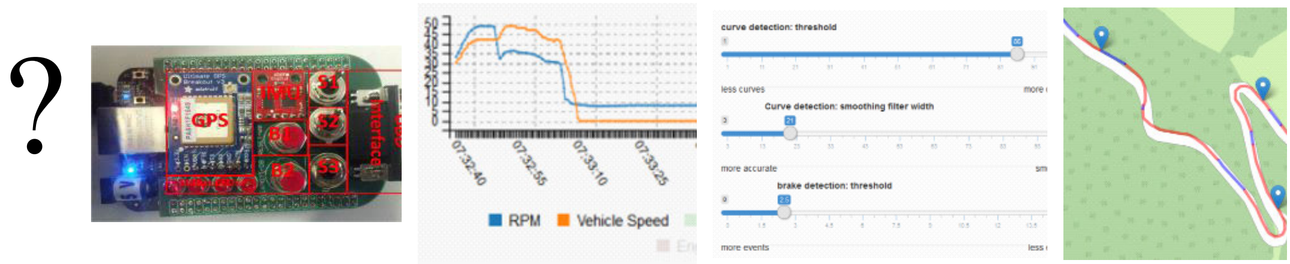

FIGURE 3. Applying the conceptual framework for mobility data science to guide the design of a digitalized mobility service for driving style detection.

specific indicators will be explored further in the subsequent analysis.

2) GET DATA

In the second step, we need to collect the necessary vehicle data. We have three possible approaches to vehicle data acquisition, utilizing the driver's smartphone mounted in the vehicle to record vehicle movements and other contextual information, connecting a data logger to the vehicle's on-board diagnostic (OBD) interface to collect vehicle data, such as speed, RPM, or location, or installing a professional logger connected to the vehicle's communication system (e.g., Controller Area Network or CAN) to obtain more extensive vehicle data, including the status of vehicle assistance systems, steering wheel angle, or wheel speed.

While the first option is straightforward, it can only capture contextual data and track the vehicle's movement without accessing vehicle sensors. The second option may provide access to some additional vehicle sensor data relevant to driving style detection. The third option theoretically allows access to all vehicle sensor signals, but decoding the raw data from the CAN bus requires information provided by the vehicle manufacturer or the relevant vehicle Electronic Control Unit (ECU) manufacturer (known as CAN DBC files). We have chosen the second option, using a car-mounted vehicle data logger to capture data such as speed, location, acceleration, and rotation for further exploration.

3) EXPLORE DATA

In the third step, we conducted an experiment involving ten drivers tasked with driving in city traffic for two hours to generate vehicle data. We noted instances of harsh acceleration, braking, or turning maneuvers, recording the maneuver type, location, and time for labeling purposes. Subsequently, we exported the logged time series data from all drivers to our computer for visual data exploration, including data plotting.

We focused on plotting speed, acceleration, and rotation at the times and locations where harsh driving maneuvers occurred. Changes in these signals were of particular interest. After storing the required vehicle data, several steps were performed to prepare the data for analysis. This involved addressing missing values, incorrect values, and outliers through methods like imputation or removal. Some signals with significant noise needed smoothing. Additionally, each signal was resampled to a common sampling rate for efficient data analysis, considering the specific analysis question.

Alignment of coordinate axes between the data logger and the vehicle was another critical step. Many signals are vector-valued, with acceleration being a notable example. Expressing these vectors in the reference frame of the car, such as aligning x-acceleration with the driving direction, simplifies analyses and interpretations. Multiple signals can be relevant depending on the type of event being analyzed.

4) MODEL DATA

The data prepared in this way can be used to intensify work on further data science questions to search for interesting events. In a fourth step we created a model using data from the first five drivers to detect each type of harsh driving behavior, harsh acceleration, braking and turning. The signals vehicle speed, acceleration in the driving direction, and rotation around the lateral axis (''pitching'') are particularly suitable for detecting strong acceleration and braking maneuvers. Pitching is caused by the change in weight distribution during speed changes: when accelerating, more weight moves to the rear axle - the rear drops and the front rises. When braking, it is the other way round. These movements can be detected. However, since detection using only one single signal can be prone to error, we use several signals in our model, which all must exceed certain thresholds simultaneously to trigger a detection of a harsh acceleration, driving or curving event. We created a model to detect the three types of driving styles from the collected data by using the following procedure, detect safe driving events when an 'event-signal' exceeds a certain threshold, and then store them together with associated information in a dataset for further processing and visualization. We finally used the data from the second group of drivers for testing and validating our model.

5) COMMUNICATE RESULTS

In a fifth and final step, to communicate the results effectively to the drivers, we developed a dashboard that offers statistical insights into their trips, including information such as trip duration and the number of safety-related events

DigMob Service	Authors	Ask question	Get data	Explore data	Model data	Communicate results	Data characteristics
Driving Style Detection	this paper	How can a driver's driving style be quantified?	In-vehicle data logger that records vehicle speed, acceleration, rotation, and position	Comparison of time series data with observed events based on manual event logging	Signal processing and definition of thresholds, statistical analyses	Geo-spatial visualization in a web dashboard. Relative score.	High-frequency and high-volume vehicle sensor data
Driving Distraction Detection	Kashevnik et al. 2020	How can a driver's distraction be detected?	Camera recordings of the driver's head and upper body, recording of the driver's voice	Determination of parameters such as eye opening, yaw angle of the head, head pitch angle and mouth opening.	Extraction of visual features of the driver and classification of dangerous states	Smartphone app that detects driver distraction and warns them	High-frequency and high volume camera and audio data of drivers
Road surface quality analysis	Chatterjee et al. 2018	How can defects in the road surface be identified?	HD camera that captures images from the road and GPS sensor that captures the position of the vehicle	Computer vision; labeling images, creating training and validation datasets.	Model training (classifiers such as gradient boosting and random forest showed the best performance)	Demonstration application for displaying defects on street images	High frequency and high volume camera data of road surfaces
Public transport performance monitoring	Lock et al. 2020	How can transport performance management be supported?	Real-time Big Data from large fleets of public transport vehicles (time, position and delays).	Design of approaches to transform data into a suitable format for analysis and visualization	Complex data pipelines (filtering, data aggregation, clustering, and machine learning)	Web-dashboard	High frequency and high volume sensor data
Indoor wayfinding service	Fusco and Coughlan 2020	How can visually impaired travelers navigate indoors?	Smartphone (held by user), camera in walking direction; 2D map of room	Cascade-based adaboost approach to detect signs; integration indoor map with walls, barriers and signs	Computer-vision based localization algorithm (adaboost , ray-tracing , particle filter, visual inertial odometry)	Mobile application for smartphones for localization/direction finding indoors	High-frequency and high volume camera data, static map data
REDTag Service – integrated framework to track and monitor parcels	Proto et al. (2020)	How can shipped packages comprehensively be tracked and monitored, i.e. in terms of the status (broken / safe) of parcels?	Parcel-related data obtained through an embedded RFID Tag, e.g., sender/receiver, package, shipping order, location	3D-visualization after Singular Value Decomposition for dimensionality reduction, visualization of pairwise feature correlation	Feature extraction, feature engineering and selection, training / application / tuning of classifiers; various classification algorithms	Front-end service (including multi-dimensional view of historical data and real-time updates) uses back-end services which compute predictions	(Historical) Low-frequency and low volume, well-structured parcel-related data
Reliable route planning for real-world intermodal planning	Ruß and Gust (2023)	How can a planner be designed that identifies routes in urban intermodal transportation networks? What are travel time savings of route planning?	Real-time departure data of public transport, GPS-based floating car data / taxi trip data / etc., route planning data for bikes / e-scooters / pedestrians, transfer time for switching mode of transport	Preprocessing with iterative data cleaning and enhancement steps: changes of schedules, approximation for missing data, etc.	Representation of a real-world intermodal network in a graph model which is solved through an route planning algorithm	Visualizations of travel time savings obtained through intermodal transport, and impact of parameters like trip complexity / traffic level / trip distance / etc.	(Historical) Low-frequency and high volume, structured data
Quantification of impact of mutualized mobility in urban environments	Sun and Ertz (2021)	What is the environmental impact of mutualized mobility solutions in terms of actual mobility behavior?	Share of different modes of mutualized mobility, environmental / lifecycle-related data for lifecycle assessment	Descriptive statistics of share of mutualized mobility solutions	Several LCA-scenarios based on plethora of different data	Visualization of environmental development over years related to usage of mutualized mobility offers	(Historical) Low-frequency and low volume, well-structured data on mutualized mobility and environmental impact, residual uncertainty due to measurement method

FIGURE 4. Applying the conceptual framework for mobility data science on existing digitalized mobility services from the literature to better structure and understand them.

detected during their journeys. Moreover, we implemented a marker-based visualization to display the three different events using markers overlaid on a geographical map, leveraging the location information. This visualization aims to assist drivers in comprehending their driving style and encourage them to adopt safer driving practices. Additionally, drivers can compare their driving style with that of others, fostering motivation to become the safest driver within their peer group.

B. APPLICATION OF THE FRAMEWORK TO BETTER UNDERSTANDING DIGITALIZED MOBILITY SERVICES

The application of our conceptual framework for Mobility Data Science to the initial case, which led to the development of a driving style detection solution, has already demonstrated its usefulness. In a subsequent step, we further apply our framework model to additional cases of digitalized mobility solutions presented in the scientific literature, showcasing its utility in classifying cases and establishing its validity as a general procedural model for Mobility Data Science.

In this process, we extract relevant information, including the mobility topic addressed, the research question, details on data collection, exploration, modeling, communication of results, and the digitized mobility service developed, from various scientific papers ([54], [86], [87], [88], [89], [90], [91]) (cf. Figure 4).

In the first work examined [86], researchers developed a *digitalized mobility service* called "Drive Safely," which is a smartphone application designed to monitor driving behavior. The application focuses on *detecting road-safety relevant inattentive and distracted driving* by utilizing data collected from the driver's personal smartphone. The smartphone's sensors and cameras record the driver's head and upper body movements, as well as the driver's voice, thus capturing mobility data from individuals connected via their personal smartphones. The service processes high-frequency and high-volume time series data, specifically facial images and upper body movements of forward-facing drivers. It extracts parameters such as eye opening, head yaw angle, head tilt angle, and mouth opening from the recorded data. By analyzing visual features of the driver, including head posture, eye movements, and dangerous gazes, the service classifies dangerous situations. The core functionality of the application lies in its ability to detect driver distraction and subsequently alert drivers, prompting them to regain focus on the road.

In the second work examined [54], researchers developed a *digitalized mobility service* focused on *road damage detection*. This service utilizes machine learning techniques and is designed to analyze transport data collected from bicycle trips. The data is recorded using an HD camera mounted on a bicycle, capturing images of the road surface, and a GPS sensor that provides the vehicle's position with high frequency and volume. The data pre-processing phase involved image labeling, employing computer vision approaches to extract relevant features, and creating training and validation datasets. The study explored the performance of various classifiers, with gradient boosting and random forest demonstrating the best overall effectiveness in the road damage detection application.

In the third work examined [87], researchers developed a *digitalized mobility service* focused on *public transport performance monitoring*. This service takes advantage of transport data sourced from public transport fleets operating in a metropolitan area. Specifically, the system utilizes data obtained from the high-volume, real-time general transit feed specification (GTFS) feed provided by a public transit provider. The GTFS feed includes information about vehicle locations and vehicle delays. The authors implemented complex data pipelines to process substantial amounts of geo-spatial data, transforming it into a suitable format for analysis and visualization. The processing involved diverse techniques such as filtering, data aggregation, clustering, and machine learning approaches. For communicating the results, the system employs various visualization elements like heat maps, point clouds, and network flow animations, displayed on geographic maps using the KeplerGL library.

In the fourth work examined [90] the researchers introduced the REDTag Service, a *digitalized mobility service* designed for *tracking and monitoring the status of parcels*. The service leverages embedded Radio-Frequency Identification (RFID) tags in parcels, equipped with sensors, batteries, memory, processor, and a network module. These tags collect various data, including information about sender and receiver details, package specifications, hardware tag ID, assigned worker, events, shipping order, and the current segment indicating progress towards the destination. The authors employed 3D visualization techniques after Singular Value Decomposition and visualized pairwise feature correlation to explore potential dimensionality reduction. The back-end service using tuned classification algorithms computes predictions. The front-end service offers real-time information on the status of parcels and visualizes multidimensional data. The REDTag Service contributes to improving parcel delivery services by enhancing knowledge on product faults and service disruptions.

In fifth examined work, the design of an *intermodal route planning model* for urban intermodal transportation networks [88], the researchers utilized low-frequency, high-volume field data collected over 5 weeks. The data sources included GPS-based floating car data, information on 412 central public transit stops, records of available shared vehicles at 30-minute intervals, optimal routes for 120 different origin-destination pairs collected every 60 minutes, and optimal routes for bicycles, pedestrians, and e-scooters. The research involved iterative preprocessing of data, such as accounting for changes in schedules during the data recording period and approximating missing data. A graph model was then established and solved through a route planning algorithm. The resulting model, a proof-of-concept, serves as a digitalized mobility service for reliable and time-saving real-world intermodal planning. In their *digitalized mobility service*, the authors also visualize travel time savings achieved through intermodal transportation and discuss the impact of parameters such as trip complexity, traffic level, and trip distance.

In the sixth work examined, a study focusing on the *environmental impact of an increased share of ride-hailing in cities*, with a particular emphasis on mobility behavior was conducted a [89]. The research involved gathering data on mobility behavior and the environmental impact of various mobility options, enabling a Life Cycle Assessment (LCA) of mutualized mobility behavior. The authors explored the data using descriptive statistics and subsequently created various LCA scenarios based on different datasets with low-frequency, low-volume, and well-structured information. These scenarios were used to visualize the environmental development over the years concerning the usage of mutualized mobility offers. The resulting (proof-of-concept) model serves as a *digitalized mobility service* providing insights into the ecological and mobility aspects related to the environmental impact of mutualized mobility in urban contexts.

In the seventh work examined [91] the focus is on creating a *digitalized mobility service* for *improving wayfinding for visually impaired travelers* to enhance accessibility. The research addresses the question of how visually impaired individuals can better navigate indoors. The proposed approach utilizes video data captured by a smartphone and a 2D map of a room. To recognize standard "EXIT" signs, the authors employ an Adaboost cascade-based approach. Simultaneously, the orientation in the room is determined through a localization algorithm that uses the indoor map, considering the locations of walls and barriers, as well as the position and orientation of signs. The outcome is a mobile application for smartphones that enables indoor localization and wayfinding, serving as a *digitalized mobility service* for enhanced accessibility.

In conclusion, the utilization of our conceptual framework as a characterization and categorization scheme offers a systematic approach for providing a comprehensive description of how Mobility Data Science contributes to the development of digitalized mobility services. Applying our conceptual framework to existing digitalized mobility services from the literature enables us to better structure and understand these services. By systematically analyzing these services according to our framework, we can identify common patterns, key components, and areas for improvement. This structured approach enhances our understanding of how digitalized mobility services are conceptualized, designed, and implemented, leading to insights that can inform the

Contribution	Description
Conceptual Framework	Introduces a conceptual framework for Mobility Data Science, integrating theoretical insights from mobility and transport science with data science principles.
Theoretical Foundations	Establishes theoretical underpinnings for understanding the interplay between mobility, data science, and digitalized services.
Methodological Guidance	Provides methodological guidance for researchers to apply the framework in analyzing and designing digitalized mobility services.
Epistemological Insights	Offers epistemological insights into the nature of data-driven mobility research and its implications for theory development.
Paradigmatic Shift	Represents a paradigmatic shift in the conceptualization of mobility services, emphasizing data-centric approaches for innovation.

FIGURE 5. Summary of our contributions.

development of future services and drive innovation in the mobility domain.

V. DISCUSSION, CONTRIBUTION, AND LIMITATIONS

A. NOVELTY AND CONTRIBUTIONS TO THEORY

Data lies at the core of data-driven services within the mobility and transport ecosystem. Presently, a wealth of valuable data is generated as a by-product. However, there exists untapped potential in harnessing this data to usher in new data-driven services [10]. In response to this, we introduce the concept of Mobility Data Science, a novel intersection of two distinct research domains: Data Science and Mobility and Transport. This innovative approach combines the realms of mobility and transport with data science, addressing a theoretical gap by providing conceptual models that illustrate how data from the mobility and transport ecosystem can give rise to innovative data-driven services. Mobility Data Science presents a pioneering perspective on the challenges and possibilities associated with utilizing emerging data to create digitalized mobility services, employing methodologies such as big data analytics, time series data analytics, machine learning, and more.

Our conceptual framework for Mobility Data Science acts as a cohesive framework that bridges the digital layer of physical transport artefacts—such as people or goods, means of transport, and transport infrastructure—with the digital layer of data-driven services. This framework serves as a unifying element for researchers and practitioners keen on harnessing the potential of mobility and transport data.

We define Mobility Data Science as the practical application of data science principles within the realm of mobility and transport. The development of our conceptual framework stands as a key theoretical contribution, derived from insights provided by domain experts. These insights were gathered through interviews with stakeholders in the mobility and transport ecosystem and collaborative projects funded by the European Commission under the Horizon 2020 framework.

The novelty of our Mobility Data Science (cf. Figure 5) approach lies in its integration of two distinct research domains: mobility and transport science, and data science. This unique combination creates a conceptual framework that serves as a bridge between the physical mobility and transport artefacts and the digital layer of data-driven services. The approach provides a systematic and comprehensive model for understanding, processing, and leveraging the wealth of data generated within the mobility and transport ecosystem to develop impactful digitalized mobility services. It not only addresses technical aspects but also emphasizes a conceptual, non-technical perspective to guide the development of innovative solutions in the realm of mobility and transport. This integrative approach is novel and contributes to the advancement of research and practice in the field.

Our framework offers unique insights and methodologies that haven't been explored in the existing literature. By combining general data science principles with domain-specific considerations, such as those found in the mobility sector, our framework provides a specialized approach that can enhance the analysis and application of data in this context. Additionally, our work contributes to the advancement of scientific knowledge by providing a structured framework that can be applied and validated in real-world scenarios, ultimately leading to further innovation and progress in the field of mobility data science.

Most existing data science methodologies are predominantly technical, emphasizing the development and application of algorithms for data-driven problem-solving. In the context of mobility and transportation, these methodologies often center on algorithmic approaches such as object recognition, vehicle trajectory planning, traffic flow analysis, or driver classification. However, data science encompasses not only these technical dimensions but also crucial conceptual aspects that are frequently overlooked. Our work addresses this conceptual perspective within data science, aiming to transition from the physical realm of mobility and transport artefacts to the generated data, relevant questions, and actionable data-driven services that deliver societal value. By emphasizing the conceptual dimension, we contribute to a more comprehensive understanding of data science in the context of mobility and transportation.

The second key contribution of our paper lies in its non-technical, conceptual approach to Mobility Data Science. Through extensive evaluation and application of our conceptual framework in various cases (cf. figure 4 and figure 6), we assert that our model can offer significant value to both researchers and practitioners, providing insights into how mobility and transport data can be utilized to develop diverse data-driven services. Initially, we applied our conceptual framework to demonstrate its effectiveness in guiding the development of a specific digitized mobility service, focusing on driving style recognition. This practical application showcases how our framework facilitates the translation from data to technical implementation through specific approaches. In a subsequent step, we further validate the utility of our framework by applying it to several published cases from the literature, highlighting its broad applicability and effectiveness. By doing so, we aim to demonstrate the versatility and robustness of our framework in organizing and making sense of various digitalized mobility services. This approach not only validates the framework's applicability across different scenarios but also highlights its potential to provide clearer insights and more cohesive understanding of existing research. By structuring published cases of digitalized mobility services according to our framework, we can identify common patterns, gaps, and opportunities in the field, ultimately contributing to more informed and effective development of future services.

We intentionally crafted Mobility Data Science as a high-level framework, recognizing that it may not encompass all intricacies involved in designing digitalized mobility services. Despite its limitations, we foresee significant implications for research. Our framework can serve as a valuable structuring tool for designing, comparing, and analyzing cases of mobility and transportation service development. Additionally, it can aid academics in gaining a clearer understanding of the opportunities and challenges inherent in the development of digitalized mobility services. While we believe the framework's transferability to other domains, it is important to note that demonstrating this transferability extends beyond the scope of our current paper.

Mobility and transport play pivotal roles in advancing technologies related to sensors, actuators, mechatronics, and information and communication technologies. The integration of sensors in vehicles, urban infrastructure, and roads produces an abundance of data, forming the foundation for innovative data-driven services. Our conceptual framework aims to assist individuals keen on harnessing this data wealth. Mobility and transport, being a multifaceted application domain, involve diverse areas of expertise and distinct languages that can pose challenges for those outside these communities to interpret. The complexity extends to the data itself. In the realm of physical transport artefacts like vehicles, data can be categorized across various stages of the product life cycle, spanning from inception to utilization and eventual disposal. In the context of infrastructure, data is typically associated with the usage phase, offering insights into the condition and operation of infrastructure through sensor-based mechanisms, such as traffic monitoring or air quality assessments. This complexity forms intricate data ecosystems that have the potential to fuel additional applications, contingent upon the availability of data for such endeavors. To facilitate this, Mobility Data Spaces have been established in the European Union, serving as collaborative platforms where stakeholders can share mobility-generated data to foster the development of innovative services. These data spaces act as interconnected hubs for entities interested in collaborative data sharing to drive advancements in mobility applications.

B. CONTRIBUTIONS TO PRACTICE

Our research holds notable implications for the realm of data governance, particularly in the intricate landscape of conceptualizing value creation in data-driven services [10]. In this context, our work stands to empower the industry in harnessing the potential of mobility and transport data by offering a comprehensive guide for the development of data-driven services. Numerous initiatives have been initiated to provide access to specific data sources, ranging from data marketplaces to data brokers.

The insights from our paper aim to aid the industry in crafting more innovative use cases grounded in shared data. To exemplify the applicability of our work, we present a straightforward case example. The utilization of our framework facilitated the development of a digitalized mobility service for driving style detection. This application case not only served to assess the proposed conceptual framework but also underscored its practical value and utility. Furthermore, it can function as a step-by-step reference example, offering practical guidance for service development. Finally, applying our framework to existing digitalized mobility services enhances the understanding of services also for practitioners, aiding in their analysis and improvement.

VI. CONCLUSION AND FUTURE WORK

In conclusion, we introduce a pioneering approach, Mobility Data Science, designed to propel the development of data-driven services within the mobility and transport ecosystem by leveraging the data it generates. We present a conceptual framework for Mobility Data Science, illustrating how it can establish connections between the digital layers of mobility and transport artefacts, including people, goods, means of transport, and transport infrastructure, with the digital layer of digitalized mobility services.

Moreover, we showcase the practical application of our framework by expediting the development of a specific digitalized mobility service tailored to driving style detection. Additionally, we demonstrate its value in structuring and understanding existing digitalized mobility services. Our research makes a significant contribution to both transport research and practice by offering a framework grounded in the insights of transport domain experts, revealing how data from the mobility and transport ecosystem can drive the evolution of digitalized mobility services.

Certainly, our work provides a foundation for future research in several ways. Firstly, researchers can utilize our conceptual framework when embarking on the development of digitalized mobility services, enabling them to navigate

Topic	DigMob Service	Authors	Ask question	Get data	Explore data	Model data	Communicate results	Data characteristics
Shared mobility	Real-time repositioning of bicycles in a bike-sharing system.	Zhang et al. 2021	How should bikes be positioned to minimize costs when demand is uncertain?	Historical data on observed bicycle use	Preparation of the data (e.g. adding virtual nodes representing sharing stations)	Neural network to predict bike-sharing demand, hybrid meta-heuristic approach to optimize repositioning.	Simulation model that delivers optimized results	(Historical) Low-frequency and low volume, well-structured demand data for bicycles
Transport optimization	Service for real-time dispatching, routing and regrouping of ambulances	Aringhieri et al. 2018	How can the number of emergency calls handled be maximized while minimizing wait time?	Data of an emergency department, instance generator based on distributions	Planar graph with n nodes and m arcs, visualization of graphs	Real-time policies implemented as online algorithms in a discrete-event simulation model	Simulation model that delivers optimized results	(Artificial) low frequency and low volume, well-structured emergency requests.
Accessibility	Prediction of response of neighborhood to natural disasters based on mobility behavior	Hong et al. (2021)	How can a neighborhood's response to and recovery from natural disasters be measured and evaluated?	Geolocation data of about 1m unique users in Houston during a 2-month period (hurricane Harvey) through mobile phones	Preprocessing - Analysis of spatial distribution of mobility / travel patterns with subsequent reduction of dataset to 829,350 unique users	Agglomerative hierarchical clustering to identify groups with similar characteristics, between-group comparison using ANOVA / Tukey's multi-comparison method	Visualization of different groups of neighborhoods, their resilience capacities, and mobility / travel patterns of residents	(Historical) High-frequency and high volume, well-structured geolocation data obtained from mobile phones
Social Distancing	Capture social distancing beliefs	Procher and Renault (2021)	How does the content of tweets relate to mobility behavior during COVID-19 pandemic?	Covid-19-related tweets from Twitter including geo-tags, epidemiological data	Descriptive statistics, derived indicators (share of social distancing-tweets, impact of tweets)	Ordinary least squares (OLS) model	OLS model and interpreted results	(Historical) High-frequency and medium volume message data, well-structured tweets

FIGURE 6. Applying the conceptual framework for mobility data science on existing digitalized mobility services from the literature to better structure and understand them (additional examples).

the intricate landscape of data-driven innovations. Secondly, our framework can serve as a valuable tool for comparative analyses, allowing researchers to systematically assess and compare various cases of digitalized mobility services. Finally, we anticipate that our framework's applicability extends beyond the mobility and transport domain, offering insights for other domains where non-digital artefacts generate data during their utilization. This opens up further avenues for cross-disciplinary applications and further exploration of data-driven service development across diverse domains.

APPENDIX A: SELECTED STUDIES

Further case studies of digitalized mobility services are detailed in Figure 6.

REFERENCES

[1] Y. Yoo, O. Henfridsson, and K. Lyytinen, "Research commentary—The new organizing logic of digital innovation: An agenda for information systems research," *Inf. Syst. Res.*, vol. 21, no. 4, pp. 724–735, Dec. 2010.

[2] F. Lindow, C. Kaiser, M. Fellmann, and A. Stocker, "A system dynamics model-based simulation of the data-driven automotive service ecosystem," in *Proc. Amer. Conf. Inf. Syst.*, 2021. [Online]. Available: https://aisel.aisnet.org/amcis2021/adopt_diffusion/adopt_diffusion/15/

[3] R. F. Ciriello, A. Richter, and G. Schwabe, "Digital innovation," *Bus. Inf. Syst. Eng.*, vol. 60, no. 6, pp. 563–569, Dec. 2018, doi: 10.1007/s12599-018-0559-8.

[4] J. Lee and N. Berente, "Digital innovation and the division of innovative labor: Digital controls in the automotive industry," *Org. Sci.*, vol. 23, no. 5, pp. 1428–1447, Oct. 2012, doi: 10.1287/orsc.1110.0707.

[5] S. Nambisan, K. Lyytinen, and Y. Yoo, *Handbook of Digital Innovation*. Cheltenham, U.K.: Edward Elgar Publishing, 2020.

[6] D. Nylén and J. Holmström, "Digital innovation strategy: A framework for diagnosing and improving digital product and service innovation," *Bus. Horizons*, vol. 58, no. 1, pp. 57–67, Jan. 2015.

[7] E. University, A. Bharadwaj, O. A. El Sawy, P. A. Pavlou, and N. Venkatraman, "Digital business strategy: Toward a next generation of insights," *MIS Quart.*, vol. 37, no. 2, pp. 471–482, Feb. 2013.

[8] M. Fitzgerald, N. Kruschwitz, D. Bonnet, and M. Welch, "Embracing digital technology: A new strategic imperative," *MIT Sloan Manag. Rev.*, vol. 55, no. 2, p. 1, 2014.

[9] P. Leviäkangas, "Digitalisation of Finland's transport sector," *Technol. Soc.*, vol. 47, pp. 1–15, Nov. 2016.

[10] C. Kaiser, A. Stocker, G. Viscusi, M. Fellmann, and A. Richter, "Conceptualising value creation in data-driven services: The case of vehicle data," *Int. J. Inf. Manage.*, vol. 59, Aug. 2021, Art. no. 102335, doi: 10.1016/j.ijinfomgt.2021.102335.

[11] A. Stocker, G. Lechner, C. Kaiser, and M. Fellmann, "Digitalized mobility," in *Proc. Amer. Conf. Inf. Syst.*, 2021. [Online]. Available: https://aisel.aisnet.org/amcis2021/adv_info_systems_general_track/adv_info_systems_general_track/4/

[12] P. Barter, "Cars are parked 95% of the time," Parking Reform Netw., 2013. [Online]. Available: https://www.reinventingparking.org/2013/02/cars-are-parked-95-of-time-lets-check.html

[13] S. Hendler and H. Boer, "Digital-physical product development: A review and research agenda," *Int. J. Technol. Manage.*, vol. 80, nos. 1–2, p. 12, 2019.

[14] C. Kaiser, A. Stocker, A. Festl, G. Lechner, and M. Fellmann, "A research agenda for vehicle information systems," in *Proc. 26th Eur. Conf. Inf. Syst., Beyond Digitization, Facets Socio-Tech. Change*, 2018. [Online]. Available: https://aisel.aisnet.org/ecis2018_rip/33/

[15] T. E. Kalayci, E. G. Kalayci, G. Lechner, N. Neuhuber, M. Spitzer, E. Westermeier, and A. Stocker, "Triangulated investigation of trust in automated driving: Challenges and solution approaches for data integration," *J. Ind. Inf. Integr.*, vol. 21, Mar. 2021, Art. no. 100186, doi: 10.1016/j.jii.2020.100186.

[16] C. Kaiser, A. Stocker, and M. Fellmann, "Understanding data-driven service ecosystems in the automotive domain," in *Proc. Amer. Conf. Inf. Syst.*, 2019. [Online]. Available: https://aisel.aisnet.org/amcis2019/org_transformation_is/org_transformation_is/14/

[17] M. Schreieck, C. Pflügler, D. S. Setzke, M. Wiesche, and H. Krcmar, "Improving urban transportation: An open platform for digital mobility services," in *Digital Marketplaces Unleashed*. Berlin, Germany: Springer, 2018, pp. 479–489.

[18] A. Stocker, C. Kaiser, and M. Fellmann, "Quantified vehicles: Novel services for vehicle lifecycle data," *Bus. Inf. Syst. Eng.*, vol. 59, no. 2, pp. 125–130, Apr. 2017, doi: 10.1007/s12599-017-0465-5.

[19] C. Kaiser, A. Stocker, A. Festl, M. Petrovic, E. Papatheocharous, A. Wallberg, G. Ezquerro, J. Orbe, T. Szilagyi, and M. Fellmann, "A vehicle telematics service for driving style detection: Implementation and privacy challenges," in *Proc. 6th Int. Conf. Vehicle Technol. Intell. Transp. Syst.*, 2020, pp. 29–36, doi: 10.5220/0009329400290036.

[20] R. Prytz, S. Nowaczyk, T. Rögnvaldsson, and S. Byttner, "Predicting the need for vehicle compressor repairs using maintenance records and logged vehicle data," *Eng. Appl. Artif. Intell.*, vol. 41, pp. 139–150, May 2015.
[21] M. F. Carfora, F. Martinelli, F. Mercaldo, V. Nardone, A. Orlando, A. Santone, and G. Vaglini, "A 'pay-how-you-drive' car insurance approach through cluster analysis," *Soft Comput.*, vol. 23, no. 9, pp. 2863–2875, May 2019.
[22] R. M. Aileni, G. Suciu, V. Suciu, S. Pasca, and J. Ciurea, "Smart systems to improve the mobility of people with visual impairment through IoM and IoMT," in *Technological Trends in Improved Mobility of the Visually Impaired*. Cham, Switzerland: Springer, 2020, pp. 65–84.
[23] T. Fian and G. Hauger, "Composing a conceptual framework for an inclusive mobility system," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 960, no. 3, Dec. 2020, Art. no. 032089.
[24] S. Ranchordás, "Smart mobility, transport poverty and the legal framework of inclusive mobility," in *Smart Urban Mobility*. Berlin, Germany: Springer, 2020, pp. 61–80.
[25] S. Barile and F. Polese, "Smart service systems and viable service systems: Applying systems theory to service science," *Service Sci.*, vol. 2, nos. 1–2, pp. 21–40, Jun. 2010.
[26] D. Beverungen, O. Müller, M. Matzner, J. Mendling, and J. V. Brocke, "Conceptualizing smart service systems," *Electron. Markets*, vol. 29, no. 1, pp. 7–18, Mar. 2019.
[27] P. P. Maglio and C.-H. Lim, "Innovation and big data in smart service systems," *J. Innov. Manage.*, vol. 4, no. 1, pp. 11–21, May 2016.
[28] T. S. Baines et al., "State-of-the-art in product-service systems," *Proc. Inst. Mech. Eng. B, J. Eng. Manuf.*, vol. 221, no. 10, pp. 1543–1552, 2007.
[29] C. Kaiser, A. Festl, G. Pucher, M. Fellmann, and A. Stocker, "Digital services based on vehicle usage data: The underlying vehicle data value chain," in *Proc. Int. Conf. Web Inf. Syst. Technol.*, 2019, pp. 22–43.
[30] J. Preston and F. Rajé, "Accessibility, mobility and transport-related social exclusion," *J. Transp. Geogr.*, vol. 15, no. 3, pp. 151–160, 2007, doi: 10.1016/j.jtrangeo.2006.05.002.
[31] V. D. Pyrialakou, K. Gkritza, and J. D. Fricker, "Accessibility, mobility, and realized travel behavior: Assessing transport disadvantage from a policy perspective," *J. Transp. Geogr.*, vol. 51, pp. 252–269, Feb. 2016, doi: 10.1016/j.jtrangeo.2016.02.001.
[32] F. Behrendt and M. Sheller, "Mobility data justice," *Mobilities*, vol. 19, no. 1, pp. 151–169, Jan. 2024, doi: 10.1080/17450101.2023.2200148.
[33] L. Frank, N. Dirks, and G. Walther, "Improving rural accessibility by locating multimodal mobility hubs," *J. Transp. Geogr.*, vol. 94, Jun. 2021, Art. no. 103111, doi: 10.1016/j.jtrangeo.2021.103111.
[34] S. Wappelhorst, M. Sauer, D. Hinkeldein, A. Bocherding, and T. Glaß, "Potential of electric carsharing in urban and rural areas," *Transp. Res. Proc.*, vol. 4, pp. 374–386, Jan. 2014, doi: 10.1016/j.trpro.2014.11.028.
[35] K. Calvin et al., *Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III To the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, H. Lee and J. Romero, Eds., Geneva, Switzerland: Intergovernmental Panel on Climate Change, Jul. 2023, doi: 10.59327/ipcc/ar6-9789291691647. [Online]. Available: https://www.ipcc.ch/report/ar6/syr/downloads/report/IPCC_AR6_SYR_FullVolume.pdf
[36] A. La Notte, S. Tonin, and G. Lucaroni, "Assessing direct and indirect emissions of greenhouse gases in road transportation, taking into account the role of uncertainty in the emissions inventory," *Environ. Impact Assessment Rev.*, vol. 69, pp. 82–93, Mar. 2018, doi: 10.1016/j.eiar.2017.11.008.
[37] D. J. Reck, H. Martin, and K. W. Axhausen, "Mode choice, substitution patterns and environmental impacts of shared and personal micro-mobility," *Transp. Res. D, Transp. Environ.*, vol. 102, Jan. 2022, Art. no. 103134, doi: 10.1016/j.trd.2021.103134.
[38] S. Bressi, M. Primavera, and J. Santos, "A comparative life cycle assessment study with uncertainty analysis of cement treated base (CTB) pavement layers containing recycled asphalt pavement (RAP) materials," *Resour., Conservation Recycling*, vol. 180, May 2022, Art. no. 106160, doi: 10.1016/j.resconrec.2022.106160.
[39] T. Schulz, M. Böhm, H. Gewald, Z. Celik, and H. Krcmar, "The negative effects of institutional logic multiplicity on service platforms in intermodal mobility ecosystems," *Bus. Inf. Syst. Eng.*, vol. 62, no. 5, pp. 417–433, Oct. 2020.
[40] A. Rizk, B. Bergvall-Kåreborn, and A. Elragal, "Towards a taxonomy for data-driven digital services," in *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, 2018. [Online]. Available: https://aisel.aisnet.org/hicss-51/da/ict_enabled_services/7/
[41] T. F. Slaper and T. J. Hall, "The triple bottom line: What is it and how does it work," *Indiana Bus. Rev.*, vol. 86, no. 1, pp. 4–8, 2011.
[42] W. Van Der Aalst, "Data science in action," in *Process Mining*. Berlin, Germany: Springer, 2016, pp. 3–23.
[43] R. Agarwal and V. Dhar, "Editorial—Big data, data science, and analytics: The opportunity and challenge for IS research," *Inf. Syst. Res.*, vol. 25, no. 3, pp. 443–448, 2014.
[44] V. Dhar, "Data science and prediction," *Commun ACM*, vol. 56, no. 12, pp. 64–73, 2013.
[45] M. A. Waller and S. E. Fawcett, "Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management," *J. Bus. Logistics*, vol. 34, no. 2, pp. 77–84, 2013.
[46] F. Provost and T. Fawcett, "Data science and its relationship to big data and data-driven decision making," *Big Data*, vol. 1, no. 1, pp. 51–59, Mar. 2013.
[47] C. Chen, J. Ma, Y. Susilo, Y. Liu, and M. Wang, "The promises of big data and small data for travel behavior (aka human mobility) analysis," *Transp. Res. C, Emerg. Technol.*, vol. 68, pp. 285–299, Jul. 2016.
[48] S. Tao, J. Corcoran, I. Mateo-Babiano, and D. Rohde, "Exploring bus rapid transit passenger travel behaviour using big data," *Appl. Geogr.*, vol. 53, pp. 90–104, Sep. 2014.
[49] F. Lindow, C. Kaiser, A. Kashevnik, and A. Stocker, "AI-based driving data analysis for behavior recognition in vehicle cabin," in *Proc. 27th Conf. Open Innov. Assoc. (FRUCT)*, Sep. 2020, pp. 116–125.
[50] A. Kashevnik, I. Lashkov, A. Ponomarev, N. Teslya, and A. Gurtov, "Cloud-based driver monitoring system using a smartphone," *IEEE Sensors J.*, vol. 20, no. 12, pp. 6701–6715, Jun. 2020, doi: 10.1109/JSEN.2020.2975382.
[51] A. Kashevnik, R. Shchedrin, C. Kaiser, and A. Stocker, "Driver distraction detection methods: A literature review and framework," *IEEE Access*, vol. 9, pp. 60063–60076, 2021.
[52] M. Kamargianni, W. Li, M. Matyas, and A. Schäfer, "A critical review of new mobility services for urban transport," *Transp. Res. Proc.*, vol. 14, pp. 3294–3303, Jan. 2016, doi: 10.1016/j.trpro.2016.05.277.
[53] P. Jittrapirom, V. Caiati, A.-M. Feneri, S. Ebrahimigharehbaghi, M. J. A. González, and J. Narayan, "Mobility as a service: A critical review of definitions, assessments of schemes, and key challenges," *Urban Planning*, vol. 2, no. 2, pp. 13–25, Jun. 2017.
[54] S. Chatterjee, P. Saeedfar, S. Tofangchi, and L. M. Kolbe, "Intelligent road maintenance: A machine learning approach for surface defect detection," in *Proc. ECIS*, 2018, p. 194.
[55] J. Eriksson, L. Girod, B. Hull, R. Newton, S. Madden, and H. Balakrishnan, "The pothole patrol: Using a mobile sensor network for road surface monitoring," in *Proc. 6th Int. Conf. Mobile Syst., Appl., Services*, Jun. 2008, pp. 29–39.
[56] A. Mednis, G. Strazdins, R. Zviedris, G. Kanonirs, and L. Selavo, "Real time pothole detection using Android smartphones with accelerometers," in *Proc. Int. Conf. Distrib. Comput. Sensor Syst. Workshops (DCOSS)*, Jun. 2011, pp. 1–6.
[57] S. E. Christodoulou, C. Kyriakou, and G. Hadjidemetriou, "Pavement patch defects detection and classification using smartphones, vibration signals and video images," in *Mobility Patterns, Big Data and Transport Analytics*. Amsterdam, The Netherlands: Elsevier, 2019, pp. 365–380.
[58] C. Dobre and F. Xhafa, "Intelligent services for big data science," *Future Gener. Comput. Syst.*, vol. 37, pp. 267–281, Jul. 2014.
[59] A. Frenzel, J. C. Muench, M. T. Bruckner, and D. Veit, "Digitization or digitalization?—Toward an understanding of definitions, use and application in IS research," in *Proc. 27th Amer. Conf. Inf. Syst.*, Montreal, QC, Canada, 2021. [Online]. Available: https://aisel.aisnet.org/amcis2021/adv_info_systems_general_track/adv_info_systems_general_track/18/
[60] Y. Wand and R. Weber, "On the ontological expressiveness of information systems analysis and design grammars," *Inf. Syst. J.*, vol. 3, no. 4, pp. 217–237, Oct. 1993.
[61] P. Bera, P. Soffer, and J. Parsons, "Using eye tracking to expose cognitive processes in understanding conceptual models," *MIS Quart.*, vol. 43, no. 4, pp. 1105–1126, 2019.
[62] Y. Wand and R. Weber, "Research commentary: Information systems and conceptual modeling—A research agenda," *Inf. Syst. Res.*, vol. 13, no. 4, pp. 363–376, Dec. 2002.
[63] T. Litman, "Measuring transportation: Traffic, mobility and accessibility," *ITE J.*, vol. 29, p. 27, Oct. 2003.
[64] J.-P. Rodrigue, C. Comtois, and B. Slack, *The Geography of Transport Systems*. Evanston, IL, USA: Routledge, 2016.
[65] S. Janssen, I. Pankoke, K. Klus, K. Schmitt, U. Stephan, and J. Wöllenstein, "Two underestimated threats in food transportation: Mould and acceleration," *Phil. Trans. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 372, no. 2017, Jun. 2014, Art. no. 20130312.

[66] D. Göhring, M. Wang, M. Schnürmacher, and T. Ganjineh, ''Radar/LiDAR sensor fusion for car-following on highways,'' in *Proc. 5th Int. Conf. Autom., Robot. Appl.*, Dec. 2011, pp. 407–412.
[67] J. Strauss, L. F. Miranda-Moreno, and P. Morency, ''Mapping cyclist activity and injury risk in a network combining smartphone GPS data and bicycle counts,'' *Accident Anal. Prevention*, vol. 83, pp. 132–142, Oct. 2015.
[68] T. Liebig, N. Piatkowski, C. Bockermann, and K. Morik, ''Dynamic route planning with real-time traffic predictions,'' *Inf. Syst.*, vol. 64, pp. 258–265, Mar. 2017.
[69] H. Meersman and M. Nazemzadeh, ''The contribution of transport infrastructure to economic activity: The case of Belgium,'' *Case Stud. Transp. Policy*, vol. 5, no. 2, pp. 316–324, Jun. 2017.
[70] R. Xie, J. Fang, and C. Liu, ''The effects of transportation infrastructure on urban carbon emissions,'' *Appl. Energy*, vol. 196, pp. 199–207, Jun. 2017.
[71] L. Tian, Y. Li, Y. Yan, and B. Wang, ''Measuring urban sprawl and exploring the role planning plays: A Shanghai case study,'' *Land Use Policy*, vol. 67, pp. 426–435, Sep. 2017.
[72] R. L. Mackett and R. Thoreau, ''Transport, social exclusion and health,'' *J. Transp. Health*, vol. 2, no. 4, pp. 610–617, Dec. 2015.
[73] R. L. Ackoff, ''From data to wisdom,'' *J. Appl. Syst. Anal.*, vol. 16, no. 1, pp. 3–9, 1989.
[74] P. Chapman et al., ''CRISP-DM 1.0: Step-by-step data mining guide,'' SPSS, Chicago, IL, USA, Tech. Rep., 2000, vol. 9, p. 13.
[75] S. Angée, S. I. Lozano-Argel, E. N. Montoya-Munera, J.-D. Ospina-Arango, and M. S. Tabares-Betancur, ''Towards an improved ASUM-DM process methodology for cross-disciplinary multi-organization big data & analytics projects,'' in *Proc. Int. Conf. Knowl. Manage. Org.*, 2018, pp. 613–624.
[76] G. Lechner, M. Fellmann, A. Festl, C. Kaiser, T. E. Kalayci, M. Spitzer, and A. Stocker, ''A lightweight framework for multi-device integration and multi-sensor fusion to explore driver distraction,'' in *Proc. Int. Conf. Adv. Inf. Syst. Eng.*, in Lecture Notes in Computer Science: Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics. Cham, Switzerland: Springer, Jun. 2019, pp. 80–95, doi: 10.1007/978-3-030-21290-2_6.
[77] G. Cledou, E. Estevez, and L. S. Barbosa, ''A taxonomy for planning and designing smart mobility services,'' *Government Inf. Quart.*, vol. 35, no. 1, pp. 61–76, Jan. 2018, doi: 10.1016/j.giq.2017.11.008.
[78] C. Benevolo, R. P. Dameri, and B. D'auria, ''Smart mobility in smart city,'' in *Empowering Organizations*. Cham, Switzerland: Springer, 2016, pp. 13–28.
[79] F. Dotter, F. Lennert, and E. Patatouka, ''Smart mobility systems and services: Roadmap 2019,'' Eur. Commission, 2019. [Online]. Available: https://trimis.ec.europa.eu/system/files/2021-04/stria_roadmap_2019_smart-mobility-systems-and-services.pdf
[80] J. Sochor, H. Arby, I. C. M. Karlsson, and S. Sarasini, ''A topological approach to mobility as a service: A proposed tool for understanding requirements and effects, and for aiding the integration of societal goals,'' *Res. Transp. Bus. Manage.*, vol. 27, pp. 3–14, Jun. 2018.
[81] I. Docherty, G. Marsden, and J. Anable, ''The governance of smart mobility,'' *Transp. Res. A, Policy Pract.*, vol. 115, pp. 114–125, Sep. 2018.
[82] W. Li, Z. Pu, Y. Li, and X. J. Ban, ''Characterization of ridesplitting based on observed data: A case study of chengdu, China,'' *Transp. Res. C, Emerg. Technol.*, vol. 100, pp. 330–353, Mar. 2019.
[83] S. Shaheen and M. Christensen, ''Is the future of urban mobility multi-modal & digitized transportation access,'' New Cities Found., Paris, France, Tech. Rep., 2014. [Online]. Available: https://newcities.org/wp-content/uploads/PDF/Research/Cities-on-the-move/Is-the-future-of-urban-mobility-multi-modal-and-digitized-transportation-access-Susan-Shaheen-Matthew-Christensen-New-Cities-Foundation-Cities-On-The-Move.pdf
[84] C. Willing, T. Brandt, and D. Neumann, ''Intermodal mobility,'' *Bus. Inf. Syst. Eng.*, vol. 59, no. 3, pp. 173–179, Jun. 2017.
[85] F. Sagberg, S. Selpi, G. F. B. Piccinini, and J. Engström, ''A review of research on driving styles and road safety,'' *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 57, no. 7, pp. 1248–1275, Nov. 2015, doi: 10.1177/0018720815591313.
[86] A. Kashevnik, I. Lashkov, and A. Gurtov, ''Methodology and mobile application for driver behavior analysis and accident prevention,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 6, pp. 2427–2436, Jun. 2020.
[87] O. Lock, T. Bednarz, and C. Pettit, ''The visual analytics of big, open public transport data—A framework and pipeline for monitoring system performance in greater Sydney,'' *Big Earth Data*, vol. 5, no. 1, pp. 134–159, Jan. 2021.
[88] M. Ruß and G. Gust, ''Reliable route planning and time savings in real-world urban intermodal transportation networks: Evidence from Hamburg, Germany,'' *Expert Syst. Appl.*, vol. 227, Oct. 2023, Art. no. 120196, doi: 10.1016/j.eswa.2023.120196.
[89] S. Sun and M. Ertz, ''Environmental impact of mutualized mobility: Evidence from a life cycle perspective,'' *Sci. Total Environ.*, vol. 772, Jun. 2021, Art. no. 145014, doi: 10.1016/j.scitotenv.2021.145014.
[90] S. Proto, E. Di Corso, D. Apiletti, L. Cagliero, T. Cerquitelli, G. Malnati, and D. Mazzucchi, ''REDTag: A predictive maintenance framework for parcel delivery services,'' *IEEE Access*, vol. 8, pp. 14953–14964, 2020, doi: 10.1109/ACCESS.2020.2966568.
[91] G. Fusco and J. M. Coughlan, ''Indoor localization for visually impaired travelers using computer vision on a smartphone,'' in *Proc. 17th Int. Web All Conf.* New York, NY, USA: Association for Computing Machinery, Apr. 2020, pp. 1–11, doi: 10.1145/3371300.3383345.

ALEXANDER STOCKER received the Ph.D. degree in business administration from the University of Graz, Austria. He has been a Key Researcher with Joanneum Research, a Senior Researcher with the Know-Center, and a Consultant for information management with Datev. Since 2013, he has been a Key Researcher for contextual information systems with Virtual Vehicle Research GmbH, Europe's largest research and development center for virtual vehicle technology. His research interests include data-driven information systems and services in the mobility and transport domain.

CHRISTIAN KAISER received the Ph.D. degree in business informatics from the University of Rostock, Germany. He was a Senior Researcher with Virtual Vehicle Research GmbH. Since 2021, he has been an IT Business Consultant for digital IT solutions with KTM AG, Europe's largest motorcycle manufacturer. His research interests include vehicle data, services that can be created with it, and their associated automotive ecosystems.

GERNOT LECHNER received the Ph.D. degree in business administration from the University of Graz. He is currently an Assistant Professor with the Department of Information and Operations Systems, University of Graz. He has worked in the industry for more than eight years in various positions in operations and supply chain management (OSCM) and has about ten years of academic experience. His research interests include data-driven research on OSCM challenges and sustainable operations and circular supply chains.

MICHAEL FELLMANN received the Ph.D. degree from the University of Osnabrück, in 2012. He is currently a Professor of business informatics, in particular business information systems with the Faculty of Computer Science and Electrical Engineering, University of Rostock. He previously worked with German Research Centre for Artificial Intelligence, Institute for Business Informatics. His current research interests include assistance systems for knowledge-intensive work and innovation management.