Universität Rostock

Traditio et Innovatio

# QUANTIFIED VEHICLES:
# DATA, SERVICES, ECOSYSTEMS

DISSERTATION

to obtain the academic degree of

Doktor-Ingenieur (Dr.-Ing.)

of the Faculty of Computer Science and Electrical Engineering

at the University of Rostock

submitted by

**Christian Kaiser**

Dipl.-Ing., B.A.

Matriculation no. 217100002

born on 02.08.1987 in Vöcklabruck,

Austria

Rostock, 04.06.2021

https://doi.org/10.18453/rosdok_id00003515

**Reviewer:**

Prof. Dr. Michael Fellmann
*University of Rostock*
*Faculty of Computer Science and Electrical Engineering*

Prof. Dr. Simon Trang
*University of Göttingen*
*Faculty of Business and Economics*

Prof. Dr. Kurt Sandkuhl
*University of Rostock*
*Faculty of Computer Science and Electrical Engineering*

Date of submission: 04.06.2021
Date of defense: 21.01.2022

An activity as big as a dissertation is only possible if the environment allows it. This here is my opportunity to say thank you to all those who have contributed to this project.

First and foremost I would like to thank my wonderful family, my wife Marion, and my children Nina and Lydia. Thank you for respecting this project and always supporting me. I would also like to thank my parents Maria and Ernst as well as my brother René, without whom I would probably never have studied at all.

A successful dissertation, however, usually does not only include a family to watch your back, but also a mentor, a sparring partner, with whom you can exchange ideas in depth. I was lucky to have two of them, Alexander Stocker and Prof. Michael Fellmann: Working with you both was and is fantastic. Sincere thanks, Alex! Even though we argue in word fights like an old married couple (at least that's what the colleagues in the office say), we are making steady progress. For me, you are a very important piece of the puzzle of this dissertation, I can't thank you enough. And many thanks to Michael and your team at the University of Rostock, Chair of Business Information Systems. As a vehicle enthusiast, you were immediately on fire for new ideas, always motivating me from the topic identification phase to the end. You always supported me proactively and taught me the scientific background I needed. This great support is anything but self-evident, for which I would like to thank you very much.

Furthermore, I would like to thank all the co-authors I was allowed to support or who supported me in the writing of publications. In particular, I would like to mention Alex Stocker, Michael Fellmann, Andreas Festl, Gernot Lechner, Gianluigi Viscusi, and Alexander Richter.

I would also like to thank Tom Szilagyi, Friedrich Lindow, and Laura Höll, who worked on sub-topics of my dissertation in their seminar works / theses at the University of Rostock. Many thanks also to Benjamin Fischer for the hardware design and development of the Data Logger during his time at Virtual Vehicle Research GmbH.

Finally, I would also like to thank my company and all my colleagues at Virtual Vehicle Research GmbH who have supported me. This allowed me to work on the very same topic for several years in different projects (especially AEGIS, SCOTT, EVOLVE, InSecTT). Their official acknowledgement statements can be found in Section III Acknowledgement, p.275.

# I. Abstract

Advancing digitalization has highlighted the potential of so-called *Quantified Vehicles* for gathering valuable (vehicle) sensor data about the vehicle itself and its environment. Consequently, (vehicle) *Data* has become an important resource of the automotive industry, which can pave the way to (*Data-driven*) *Services*. There are multiple roles to occupy in service generation, from data provider via service developer and service provider to end user. The (*Data-driven Service) Ecosystem* of actors that collaborate to ultimately generate services, has thus only shaped up in recent years.

This dissertation was started in 2016 when literature and research in this field were still scarce. In retrospect, the rise of automotive *Data-driven Services* was accompanied. Vehicle *Data*, *Data-driven Services*, and the corresponding *Data-driven Service Ecosystems* have become an important market in recent years, as reports of analysts (e.g. Capgemini, Deloitte, KPMG, McKinsey, and PWC) show. The expectations even extend to statements about radical changes in automotive business models due to data-driven service possibilities. And such potentially radical changes of an economic driving force of Germany, where 830,000 employees generated around 435 billion Euros in 2019, are undoubtly a relevant research topic.

And so it is a variety of influences that characterize the service development or have so far prevented *Data-driven Services* from breaking through that could be investigated. However, the five objectives of this work are to (i) better understand how *vehicle data* becomes a relevant artifact for business and innovation, (ii) define and describe *Quantified Vehicles* as a form of digitalization in the automotive domain, (iii) develop concepts and *Data-driven Services* prototypically, that represent added value for consumers to enhance the understanding of challenges in service development, (iv) better understand the process and actors of value generation, and the interplay of the actors with each other in the *ecosystem*, by conducting empirical research involving automotive domain experts, and (v) conduct design activities backed by empirical research to conceptually model data-driven value generation and *Data-driven Service Ecosystem* building.

To address the objectives, three research questions were defined, which were worked through in the research process with the help of eight subtopics. Within the framework of the cumulative dissertation, the author of this thesis contributed to a total of 14 publications (ten

as corresponding / main author) that can be assigned to these eight subtopics and that contribute to answer the three research questions: Two publications for the subtopic *Definition and Introduction of "Quantified Vehicles"*, two publications for the subtopic *Analysis of the Market: Services, Start-ups, OEMs, Business Models, and Trends*, one publication for the subtopic *Definition of a Research Agenda for the Information Systems Community*, two publications for the subtopic *Analysis and Definition of the VDVC* (Vehicle Data Value Chain), three publications for subtopic *Concepts along the VDVC, two* publications for subtopic *Prototypical Implementations along the VDVC*, and one publication each for the subtopics *Analysis of Data-driven Service Ecosystems* and *Conceptual Model for Value Creation in Data-driven Services.*

The 14 publications consist of three journal publications (E&I[1], BISE, and IJIM), two book-series contributions (LNMOB, and LNBIP), eight conference contributions (NBM, i-Know, ECIS, ICVES, WEBIST, CAiSE, AMCIS, and VEHITS), and a contribution in a BITKOM position paper (BITKOM). Thereby, the journal publications IJIM (Impact Factor 8.21), BISE (Impact Factor 5.83), the ECIS conference publication (VHB-JQ3: B) are particularly outstanding, due to their high impact and prestige in the IS community.

1 Short name of the publication medium. See the List of Abbreviations (page 309 ff) for the long name.

# Zusammenfassung

Die fortschreitende Digitalisierung hat das Potenzial für sogenannte *Quantified Vehicles* aufgezeigt, welche wertvolle (Fahrzeug-)Sensordaten über das Fahrzeug selbst und seine Umgebung zu sammeln. Folglich sind (Fahrzeug-)*Daten* mittlerweile zu einer wichtigen Ressource der Automobilindustrie geworden, da sie auch (datengetriebenen) *Services* ermöglichen können. Bei der Service-Generierung sind mehrere Rollen zu besetzen, vom Datenlieferanten über den Service-Entwickler und den Service-Anbieter bis hin zum Endnutzer. Das (datengetriebenene Service) *Ökosystem* von Akteuren, die zusammenarbeiten, um letztlich Services zu generieren, hat sich erst in den letzten Jahren herausgebildet.

Diese Dissertation wurde im Jahr 2016 begonnen, als es noch wenig Literatur und Forschung in diesem Bereich gab. Rückblickend wurde der Aufstieg von datengetriebenen Services in der Automobilindustrie begleitet. Fahrzeugdaten, datengetriebene Services und die entsprechenden Ökosysteme sind in den letzten Jahren zu einem wichtigen Markt geworden, wie Berichte von Analysten (z.B. Capgemini, Deloitte, KPMG, McKinsey und PWC) zeigen. Die Erwartungen reichen bis hin zu Aussagen über radikale Veränderungen der automobilen Geschäftsmodelle durch datengetriebene Servicemöglichkeiten. Und solche potenziell radikalen Veränderungen eines Wirtschaftsstandortes Deutschland, in dem 830.000 Beschäftigte im Jahr 2019 rund 435 Milliarden Euro erwirtschafteten, sind zweifelsohne ein relevantes Forschungsthema.

Es gibt eine Vielzahl von Einflüssen, die die Serviceentwicklung kennzeichnen oder bisher den Durchbruch datengetriebener Services gehemmt haben, die untersucht werden könnten. Die fünf Ziele dieser Arbeit sind, (i) besser zu verstehen, wie Fahrzeugdaten zu einem relevanten Artefakt für Unternehmen und Innovationen werden, (ii) *Quantified Vehicles* als eine Form der Digitalisierung im Automobilbereich zu definieren und zu beschreiben, (iii) datengetriebene Dienste exemplarisch zu entwickeln, die einen Mehrwert für den Verbraucher darstellen, um das Verständnis für die Herausforderungen bei der Entwicklung von Diensten zu verbessern, (iv) den Prozess und die Akteure der Wertschöpfung sowie das Zusammenspiel der Akteure untereinander im Ökosystem besser zu verstehen, indem empirische Forschung unter Einbeziehung von Domänenexperten aus dem Automobilbereich durchgeführt wird, und (v) durch empirische Forschung gestützte Designaktivitäten durchzuführen, um die datengetriebene Wertschöpfung und den Aufbau eines datengetriebenen Service-Ökosystems konzeptionell zu modellieren.

Um die Ziele zu adressieren, wurden drei Forschungsfragen definiert, die im Forschungsprozess in Rahmen von sechs Unterthemen abgearbeitet wurden. Im Rahmen der kumulativen Dissertation hat der Autor dieser Arbeit an insgesamt 14 Publikationen (zehn als korrespondierender Autor / Hauptautor) mitgewirkt, die sich den sechs Unterthemen zuordnen lassen und respektive dazu beitragen die drei Forschungsfragen zu beantworten: Zwei Publikationen zum Unterthema *Definition und Einführung von "Quantified Vehicles"*, zwei Publikationen zum Unterthema *Analyse des Markts: Services, Start-ups, OEMs, Business Models, und Trends*, eine Publikation zum Unterthema *Definition einer Forschungs-Agenda für die Informationssystem-Community*, zwei Publikationen zum Unterthema *Analyse und Definition der VDVC* (Vehicle Data Value Chain), drei Publikationen zum Unterthema Konzepte *entlang der VDVC,* zwei Publikationen zum Unterthema *Prototypische Implementierungen entlang der VDVC* und je eine Publikation zu den Unterthemen *Analyse datengetriebener Service-Ökosysteme* und *Konzeptionelles Modell zur Wertschöpfung in datengetriebenen Services*.

Die 14 Publikationen setzen sich aus drei Journal-Publikationen (E&I[2], BISE und IJIM), zwei Buchreihenbeiträgen (LNMOB und LNBIP), acht Konferenzbeiträgen (NBM, i-Know, ECIS, ICVES, WEBIST, CAiSE, AMCIS und VEHITS) und einem Beitrag in einem BITKOM-Positionspapier zusammen. Dabei sind die beiden Journal-Publikationen IJIM (Einflussfaktor 8.21), BISE (Einflussfaktor 5.83), und die ECIS-Konferenzpublikation (B im VHB-JQ3 Ranking) aufgrund ihres hohen Einflusses und Prestiges in der IS-Community besonders hervorzuheben.

2 Kurzname des jeweiligen Publikationsmediums. Die Langbezeichnung findet sich im Abkürzungsverzeichnis (List of Abbreviations, S.309).

# II. Table of Contents

# 1. Synopsis

## 1.1 Introduction

The automotive industry is one of the economic driving forces in Germany and Austria. Around 830,000 employees in Germany generated around 435 billion Euros in 2019 (Austria: 35,000 employees, 15 billion Euros) (bmwi, 2020; Statista, 2020; WKO, 2020), while many more people are involved along the entire value chain (FV, 2020). Globally, vehicle sales reached 65.5 million units (fell by four percent) in 2019 (Bekker, 2020). However, the automotive industry is currently facing many challenges. To name just a few, automated driving, reduction of $CO_2$ emissions (e.g. through renewable energy sources for drive technologies), and digitalization are urged topics in automotive in the period from 2016 to 2020, which create a race among manufacturers for the fastest and most technologically advanced innovation development. PWC (Kuhnert et al., 2018) even expects “radical change” and states, that the “car of the future is electrified, autonomous, shared, connected and yearly updated“, while the focus will be “on the user” because these technological developments are accompanied by a changed understanding of mobility and the changed purchasing and mobility behavior of customers (Kessler and Buck, 2017), which additionally increases pressure. Especially among young adults in industrialized countries (e.g. Germany, Japan, and the USA), mobility patterns have been changing in recent years, and “the idea of freedom” has been decoupled “from the idea of car ownership” (Kessler and Buck, 2017, p.109). Consequently, vehicle ownership decreases (Kuhnimhof et al., 2012; Kessler and Buck, 2017; Fulthorpe, 2015), while young adults prefer the “ability to conveniently request, track, and pay for trips via mobile devices” (Murphy, 2016), thus e.g. “shared mobility [..] is growing as a mobility model” (Bertoncello et al., 2016, p. 10) among young adults, which are described as always being connected, ambivalent towards the car and adopting new transport technologies (Sakaria & Stehfest, 2013; Delbosc and Ralph, 2017). At the same time, analysts including Gartner (Davenport et al., 2020), McKinsey (Bertoncello et al., 2016), and Accenture (Seiberth and Gründinger, 2018) argue, that vehicle manufacturers will have to compensate for this loss of revenues from vehicle sales with additional revenues, e.g. stemming from vehicle data monetization. And one way to monetize vehicle data are *Data-driven Services* which are based on vehicle data (Viereckl et al., 2016).

As in many other industries, data has become an important resource, which can pave the way to novel services and optimization (Schüritz et al., 2017). Vehicle data represents this resource in the automotive industry (Kaiser et al., 2021), made available through the connected car, “a communications hub that transmits as well as receives data and information from its surroundings” (Fulthorpe, 2015). Connected cars can provide data for data-driven services, and new players like ICT companies are joining the value creation chain to develop and provide such services. For instance, HIS (Fulthorpe, 2015) not only put the *connected car* in their list of top five automotive challenges (in 2015 and beyond) but report on “raised concerns that carmakers face the risk of becoming subordinate to the business models of other industries as new types of firms enter”. In the meantime, there is already a competition between startups, large-scale ICT enterprises, and vehicle manufacturers to address this vehicle data (and data-driven service) market successfully (Probst et al., 2017; Viereckl et al., 2016), and to occupy the important roles present in value creation (Kaiser et al., 2017). “Where to play in the connected vehicle value chain” is an important decision, which should be “one of the first decisions for companies aiming to monetize vehicle data”, Deloitte (Hood et al., 2019) states. PWC (Kuhnert et al., 2018) expects that the “comprehensive and rapid reorganization of the automotive sector, as we predict, will have far-reaching consequences for the entire industry and its value chains. Elementary structures and attitudes will have to change fast in order to cope with the developments by 2030 and beyond.” While predicting increasing collaborations for vehicle manufacturers (i.e. with ICT companies), Kuhnert et al. (2018) state that it will be essential to link the “hardware” (i.e. the vehicle) with the “software” (i.e. the services)”.

To remain competitive, the automotive industry (in Germany as in all of Europe) thus is facing a period of digital transformation, restructuring and change, to no longer just primarily offer goods (e.g. selling manufactured vehicles as the main product) and product-related services (e.g. selling spare parts and conducting maintenance work) (Kaiser et al., 2021; Viereckl et al., 2016). Obviously, this restructuring is a particularly interesting research context as to its economic importance, and to its long tradition in catering to a basic human need – mobility (Piccinini et al., 2015).

To go into the topics of the challenge in more detail, the currently urged topics in the automotive industry of automated driving, $CO_2$ emission reduction, and digitalization are briefly described hereafter, as they already demonstrate (i) the need for and the increasingly important role of data, (ii) how *Data-driven Services* emerge and (iii) how collaboration between multiple players in the *Data-driven Service Ecosystem* is changing.

### 1.1.1 Data

Data is becoming increasingly important in all industries, which turn to “collect data on everything” (Mayer-Schönberger and Cukier, 2013). Recently, artificial intelligence, which requires data as input, has also become increasingly relevant. This is also the case in the automotive industry (Bertoncello et al., 2016), where much of the innovation comes from automation, based on data (Szalavetz, 2019). And this is how the challenges mentioned (automated driving, $CO_2$ emission reduction, and digitalization) fit in, as they also contribute to generating even more data, which can be used in data-driven services. Thereby, the automotive industry also contributes to the hypothesis of an ever-increasing amount of data generated. And more and more value is seen in the data, as the following statements indicate: Krzanich (2016) states, that “data is the new oil in the future of automated driving”. Soley et al. (2018) argues, that “connected vehicle data has higher value than people and companies are aware of.” and thus joins the opinion of researchers because the data enables data analysis, artificial intelligence, etc., which in turn can lead to e.g. revenues and sustainability (Lioutas and Charatsari, 2020). McAfee and Brynjolfsson (2012) explain, that data is an important source for decision making, as “data-driven decisions are better decisions”, while Li et al. (2019) show that “free” digital goods can create a lot of value out of data, “because consumers lack knowledge regarding the value of their own data.” And so, data also enables data security breaches (Hirsch, 2013). However, even the European Commission (2020) expects that “data-driven innovation will bring enormous benefits for citizens” and includes a “common European mobility data space” in its *European strategy for data* which should “facilitate access, pooling and sharing of data from existing and future transport and mobility databases”. Hence, vehicle data eventually paves the way for new types of data-driven services.

And so, it is also data that, through data analysis, should even enable fully automated driving. Fully automated driving is currently a much-discussed topic in the automotive industry, as it aims to turn drivers into passengers who no longer have to worry about driving. Not having to drive yourself would be a big change in the lives of many private individuals (especially commuters), but also in the economy, if for example in logistics drivers are no longer needed. There is also speculation that private individuals will use a means of transport rather as a service (quasi a robot taxi), and therefore do not necessarily have to own it anymore (vehicles are after all one of the most expensive purchases in peoples’ lives) (Fulthorpe, 2015). But this vision is still a long way off. Fully automated vehicles should be able to integrate themselves in mixed traffic with non-automated or partially automated vehicles without being an annoyance, must therefore exchange data with other vehicles and the infrastructure. Besides, users expect automated vehicles to be safer than the human driver in order to build the necessary trust to be chauffeured. To achieve this, automated vehicles are equipped with additional sensors, e.g. RADAR, LiDAR, and video cameras, to perform a comprehensive 360° analysis of the environment and understand it by means of algorithms, and consequently collect more and more data, which could also be of interest for third parties (e.g. find parking lots, detect road surface damage, evaluate regional driving styles, etc.).

In addition to automated driving, reducing $CO_2$ emissions is also a currently urged topic, which interestingly has a strong connection to vehicle data. The vehicle's on board diagnostic (OBD-2) interface originally was introduced, to capture vehicle data "which is relevant for testing whether the vehicle's emissions are still within tolerance" (Kaiser et al., 2020b). In general, in order to reduce $CO_2$ emissions into the environment, newly registered vehicles must increasingly comply with stricter emission classes (e.g. EURO 1 to EURO 6), and manufacturers must increasingly reduce the average fuel consumption of a vehicle fleet (briefly termed *fleet consumption*). Reducing fleet consumption is a difficult undertaking, as consumers currently prefer to buy larger and heavier vehicles, e.g. SUVs. To reduce fleet consumption, many vehicle manufacturers offer vehicles with different drive technologies including renewable energies, such as hybrid vehicles, electric vehicles, or hydrogen vehicles. Therefore, in some cases, a vehicle is developed for several different drive systems in parallel (e.g. petrol, diesel, electric, hydrogen, hybrid versions), which leads to an increase in the number of vehicle variants that are quickly brought to market in smaller lot sizes. And mistakes and immature technology are quickly punished nowadays. For example, the recent release of the electrified Volkswagen ID.3 was criticized because it still had "serious problems with the software development" (Mayr and Slavik, 2020; Kane, 2020), while news on an electric vehicle that started to burn (Sun et al., 2020) spread quickly in the past, resulting in image damage. Consequently, it is increasingly important, also in the case of electrified vehicles, to have mass vehicle data from the field available for analysis and artificial intelligence, e.g. to analyze and optimize battery performance (You et al., 2017).

Concerning vehicle data, that enables analysis and artificial intelligence to be used in data-driven services, brings us back to the last urgent topic in the automotive industry: digitalization. An industry that once provided physical transportation exclusively through mechanics and engineering by offering goods (e.g. selling manufactured vehicles as the main product) is now facing market pressure to offer digital features. Physical driving itself is taken for granted, and the quality differences between European manufacturers in terms of e.g. driving dynamics and crash behavior are shrinking, partly due to norms and standards tested in the *European New Car Assessment Programme* – Euro NCAP (Van Ratingen et al., 2016). Outstanding vehicles today are those, that are innovative in digitalization. From their working environment, consumers are increasingly used to the fact that the tools they use are digital and flexible. Thus, widgets often allow users to arrange user interfaces in the way they want. This was and is a challenge for traditional vehicle manufacturers, as suddenly functioning systems and systems of systems that were outsourced to suppliers were to change radically. For example, the vehicle cockpit, which in 2020 often appears as a digital touch screen with a personalized user profile, can also integrate smartphone applications (e.g. via Android Auto, Android Automotive or Apple Car Play (Apple, 2017)). Furthermore, digitization also brought new possibilities, since a sensor value can now be exchanged relatively easily with other software elements via software interfaces. Thus, and this is of special interest in this dissertation, a new field opened up since vehicle data (generated by vehicle sensors while driving) is potentially of interest to third parties. For instance, suppliers want to know how their components are used in the field (Childerhouse et al., 2003; Farahani et al., 2017). Insurance companies want to know how drivers drive (Tselentis et al., 2016). Traffic planners want to know how roads are used in reality (Kong et al., 2018). Road users want to be informed in case of safety-critical situations of a vehicle, and so on. Unfortunately, in 2016, when this dissertation started, not a single vehicle manufacturer offered the possibility to directly use vehicle data, e.g. in integrated smartphone applications. However, emerging start-ups were already collecting data and offering services, using either a gateway device (e.g. the start-up *Dash*) connected to the vehicles' OBD-II interface (that was never intended to be used like that), or smartphone sensors (e.g. the start-up *Zendrive*), e.g. by exploiting the smartphones' acceleration sensor and GPS position, which allow to calculate an approximation of vehicle speed and driving behavior (Stocker and Kaiser, 2016). The rich set of vehicle data, which would have been available on the internal bus system (e.g. CAN), was accessible with special measuring devices and interpretable with a decryption file (dbc file, "describes the communication of a single CAN network" (Vector, 2007)) only at that time, and making changes to the internal bus system is generally not allowed, except in test drives. But that has changed in recent years, so service providers (with the consent of the vehicle owner) can now obtain data from the vehicle (a filtered dataset of internal bus system messages) via an OEM backend.

### 1.1.2 Data-driven Services

In general, the utilization of data and data analysis is expected to “offer new ways for growth and competitive advantage” (Schüritz et al., 2017; Davenport and Harris 2007; Davenport and Harris 2017). While the increasing data offer even opens up opportunities for the creation of entirely new (data) services (Manyika et al., 2011), data-driven business models that use data as their key resource and exploit the possibilities of analytics emerge (Hartmann et al., 2016). Schüritz et al. (2017) refer to the “offerings of these business models as data-driven services” (but “digital services” is also common), which are seen as the main characteristic of modern, digitalized mobility (Stocker et al., 2021). Probst et al. (2017) even expect that “in the long run, 30% to 40% of the value in the automotive value chain will be captured by digital services”.

Data-driven services (in the automotive domain) may exploit vehicle sensor data. Modern passenger vehicles are equipped with sensors to (ideally) detect and react to every possible condition of the vehicle and its surrounding environment. The sensor data is generated by the sensor and is typically made available via a bus system to all electronic control units (ECUs) that need the information. Triggers use the input data for control loops or the call of actions, while persistent storing of the sensor data was not considered until a few years ago – the data volumes involved were simply considered too large. As a result, almost all the data disappeared again immediately after it was generated. This has changed in the last years, as some modern vehicles (in 2020) already provide the technical equipment to possibly transmit the data to a data-driven service. And the data already has its own market, e.g. data marketplaces (e.g. by the companies *Caruso*, or *Otonomo*) evolved in the last years, which have access to vehicle data and resell this data to service developers at a certain surcharge (Stocker et al., 2021; Kaiser et al., 2021; Spiekermann, 2019). Data marketplaces are even expected to “moving to the center of the data economy by providing an infrastructure for trading data and data-related services” (Spiekermann, 2019). Thereby, it is also attempted to pool data from several manufacturers to increase utility. In this respect, the EU project AutoMat (2018d) states in one of its deliverables, that “the higher the number of users on a given [data] package, the higher the return on it for the marketplace operator and OEMs.” And this can be seen in reality, for instance, the data marketplace by Caruso offers “harmonized multi-brand in-vehicle data”, this means vehicle data from several manufacturers (Audi, BMW, Ford, Mercedes-Benz, MINI, Volkswagen) can be integrated via a uniform interface (Caruso, 2020).

The probably much faster ongoing quantified-self trend of other domains (e.g. sports) has shown that sensor data, which can be used to draw conclusions about the use of the vehicle and the environment, has definite potential (Swan, 2013). A real hype arose in recent years, for example, even hobby runners know best about pace, the average time for one kilometer or calorie consumption and can compare themselves with each other (i.e. Runtastic with 80 million registered users (Runtastic, 2016)), while cyclists even duel each other virtually on parts of the route like climbing a hill to know their place in the ranking (Strava, 2017). And ever since the consumer electronics industry took up the quantified-self subject, it was clear that vehicles, which had been equipped with a lot of sensor technology for quite some time, would sooner or later become an area of interest – and this is what has happened in the last five years. Consequently, the behavioral patterns of self-tracking can be transferred to vehicles, which capture sensor data about themselves and their environment (Stocker et al., 2017a). Possible fields of application in automotive are manifold. For instance, driver and driving statistics present quantified, comparable behavior to the driver and try to increase safety, a topic which is always in the attention of the automotive industry, which means there is a lot of investment available for this. Furthermore, the driving style of the driver also interests the driver himself to be able to increase awareness and compare himself with other drivers (gamification approach). Several smartphone applications have emerged, which use vehicle data, collected from an OBD gateway device, and analyze it for safe driving relevant events like harsh braking, which are then presented to the user (Kaiser et al., 2020c).

However, software used to be not the core business of many vehicle manufacturers (AutoMat, 2016b; Volkswagen, 2019), and data-driven services are still not very widespread across them, but, perhaps since the sales figures for new cars have been stagnating or falling (Bekker, 2020), it is becoming increasingly interesting for vehicle manufacturers to make money via new possibilities, such as data-driven services (Probst et al., 2017). A *Head of Data Services* employed at a German Car Manufacturer once confirmed in an interview with the author of this thesis that they are also aware that they lack resources and skills in this area (software and data-driven service developers) and therefore deliberately contract with external software companies (Kaiser et al., 2019b), who have already been trying for some time to penetrate the market of data-driven services in the automotive domain.

Thus, the automotive industry ecosystem is slowly changing, in which software companies are increasingly becoming relevant players (Broy, 2006), as was also stated in a panel discussion at the conference *monetizing car data* in 2020, where the author of this thesis was a panelist together with representatives of BMW and Peregrine (service development company). Haghighatkhah et al. (2017) even state, that “the automotive industry is going through a fundamental change by moving from a mechanical to a software-intensive industry in which most innovation and competition rely on software engineering competence.”

### 1.1.3 Data-driven Service Ecosystems

In general, an *ecosystem* describes the relationships and interactions between living organisms and their environment (Schulze et al., 2005; Briscoe and De Wilde, 2006), and is recently increasingly used by researchers as “a new way to depict the competitive environment” (Jacobides et al., 2018). To distinguish an artificial ecosystem from a natural one, some authors add further attributes to the term to qualify it, e.g. software ecosystem, business ecosystem, or digital service ecosystem (Immonen et al., 2015). However, a commonly agreed definition does not yet exist. (Kaiser et al., 2021)

Considering the automotive industry, Teece (2007, p. 1325) defines an ecosystem as “the community of organizations, institutions, and individuals that impact the enterprise and the enterprise's customers and supplies” including “complementors, suppliers, regulatory authorities, standard-setting bodies, the judiciary, and educational and research institutions”. With a different point of view, Jacobides et al. (2018) define it as “a set of actors with varying degrees of multilateral, non-generic complementarities that are not fully hierarchically controlled” (p. 2264). Furthermore, Adner (2016, p. 40) defines an ecosystem as “the alignment structure of the multilateral set of partners that need to interact in order for a focal value proposition to materialize”. And already Stocker et al. (2017a) describe, that the traditional balance of power with a strong Original Equipment Manufacturer (OEM) is currently challenged by digitalization, which encourages OEMs to cooperate with new partners in the market (Subramaniam et al., 2019). (Kaiser et al., 2021)

Taking business into account, Nischak et al. (2017) mention three essential components of *digital business ecosystems*: value exchange (innovation, information, products/services), resources (digital and non-digital), and actors (organizations, individuals, societies). This definition can be adapted and specialized for digital automotive ecosystems (Kaiser et al., 2021), or more specifically, data-driven service ecosystems. A data-driven service ecosystem contains actors that in this case are original equipment manufacturers, data marketplaces, or data service providers. These actors have access to resources (e.g. infrastructure), for generating, transmitting, and storing data. Leveraging these resources, the actors participate in value exchanges by providing or consuming data. (Kaiser et al., 2021)

Nevertheless, research on *Data-driven Service Ecosystems* is still limited (plural, as each manufacturer can have its own ecosystem). Particularly in connection with vehicle data and the process of creating data-driven services, the literature repeatedly refers to data-based business ecosystems (Kitsios et al., 2017; Curry, 2016; Nachira et al., 2007). For instance, Immonen et al. (2014) outline the open data ecosystem from a business viewpoint and define ecosystem actors such as application users, data and service providers, application developers, and infrastructure providers along with their role in the data-based ecosystem. Also, in many cases, the authors refer more to technical ecosystems (e.g. Kolbe et al., 2017; Gerloff and Cleophas, 2017; Kuschel, 2008; Martínez de Aragón et al., 2018). In these technology-oriented perspectives, an analysis of the business relations enabled through the digitalization of the vehicle and the feasibility of new data-driven services is largely missing, while e.g. Athanasopoulou et al. (2019) report on digital technologies that “disrupt the existing business models within the automotive industry”. (Kaiser et al., 2021)

However, vehicle manufacturers employ many developers with vehicle development, e.g. Volkswagen mentions about 20,000 developers (Reuters, 2019). The company's own developers, according to Diess, Volkswagen CEO, “90 percent hardware-oriented” (Reuters, 2019), focus mainly on ensuring that the vehicles drive well (Athanasopoulou et al., 2019), with the development of software or even digital services being outsourced to supplier software companies. In the VW Group, the percentage of all software in vehicles is “less than 10 percent” (Volkswagen, 2019). Thus, innovations for data-driven services come from outside, from small, agile startups (Homfeldt et al., 2019).

In retrospect, starting with innovative start-ups that developed a gateway device for the standardized OBD interface to collect data, develop first data-driven services (pay as you drive insurance, car monitoring, fleet tracking, tutoring for smarter / safer / greener driving, etc.) and make business, vehicle manufacturers have also discovered the topic for themselves (Kaiser et al., 2017b). One background was certainly that they did not want to be deprived of this new business area. Through research projects funded by the European

Commission, which develop concepts and define data formats for the data exchange from vehicles to cloud servers (e.g. project *AutoMat* coordinated by Volkswagen and project Cross-CPP), as well as investments in start-ups, vehicle manufacturers have approached the topic, and in some cases developed and launched their own services in parallel under the connected car brand (e.g. BMW's brand *BMW ConnectedDrive* with BMW CarData to show key vehicle data to the user). An example of investments in start-ups is BMW's (German vehicle manufacturer) venture capital subsidiary BMW i Ventures, which has more than 45 companies in their portfolio, including the start-ups *Nauto* and *Zendrive*, which analyze the driver behavior based on sensor data. As an example for developing and launching its own services, for instance, Audi developed a service called *Green Light Optimal Speed Advisory* (GLOSA), which predicts the green phases of traffic lights and informs the driver about the optimal speed to pass the next traffic light within the green phase. Audi, therefore, teamed up with cities, to get access to the traffic light data. Another example, which shows that for applications cooperation between different actors of the data-driven service ecosystem is necessary.

Unfortunately, so far there seems to be no application, neither from manufacturers nor from third party service developers, which is outstanding, which really taps the full potential (Stocker et al., 2017a; Starepravo, 2019). In general, "complexity and dynamism [..] has made it difficult to make decisions regarding where to play and how to win" (Hood et al., 2019) for actors of the ecosystem, mentioning issues like data ownership, willingness to pay, and decisions like 'make or buy', cooperation options, and data protection (privacy and trust). E.g. several vehicle manufacturers argued recently, that service development is not profitable for them, (stated by representatives of Audi, BMW, and Daimler at the conference *monetizing car data* in 2020), thus, manufacturers currently emphasize the "added value" for users that is created by the offer and attempts are being made to build so-called ecosystems in which third party service developers and service providers can develop services. Thus, vehicle manufacturers could cover its costs in the role of a data and platform provider. Additionally, there are efforts to achieve cross-manufacturer cooperation, e.g. a uniform data transfer concept of the German automotive industry called NEVADA ("Neutral Extended Vehicle for Advanced Data Access") concept should enable data intermediaries and service providers to make better use of vehicle data (VDA, 2017b; VDA, 2017c). A core element of the NEVADA concept is a backend of the respective vehicle manufacturer, which forwards the data (depending on the use case class) to a neutral server (Reich et al., 2018). Furthermore, it is interesting, that i.e. BMW and Daimler, which officially cooperate long term to develop automated driving (BMW, 2019), also both contract with data marketplaces like Otonomo and Caruso, players of the data-driven service ecosystem.

The data-driven service ecosystem, in which several companies cooperate to generate the value that is being created, is particularly interesting for research because it is unclear whether some players will survive in the long run. A senior executive at a premium OEM even stated in a McKinsey report, that “no single player can succeed on a stand-alone basis in establishing the digital ecosystem around the car, and multiple stakeholders need to work together” (Bertoncello et al., 2018). Especially start-ups that have brought in the innovation have the risk to be bought up (e.g. by bigger companies, like OEMs) or to disappear again (e.g. Automatic). However, the developments around data-driven services will have an impact on the automotive market, and it is precisely the uncertainty of how the data-driven service market will develop that makes it so interesting and relevant. At the same time, there is no comparable work available in the automotive domain, so research in this field can be considered pioneering.

So, while the automotive market has recently developed strongly (w.r.t. vehicle data utilization and the entrance of new players (Cäsar et al., 2019)), it is part of this dissertation to analyze and understand exactly this market development within the *Data-driven Service Ecosystem* from a research perspective.

## 1.2 Motivation

The topic around *Quantified Vehicles* and *Data-driven Services* actually has quite a lot of aspects that one could study, and many different disciplines can provide valuable insights. Research on *Quantified Vehicles* can be partitioned in three areas: (i) *Data*, (ii) *Services*, and (iii) *Ecosystems*. These areas are highly interdependet and can all have a major impact on the future and can even support higher targets such as Vision Zero – the objective “to move close to zero deaths by 2050” (European Commission 2019) – since data provides the foundation for new services, and services are enabled and consumed by various stakeholders that are part of ecosytems. One topic where this is particularly evident is value generation. After an introduction to the relevance and novelty of the topics, the vitality of the topic will be briefly discussed, followed by the identification of automotive constraints and research gaps on the topics of data, services, and ecosystems, to motivate this dissertation.

### 1.2.1 Topic Relevance and Novelty

The dissertation was started in 2016 when literature and research in this field were still scarce. In retrospect, the rise of the subject was accompanied. Vehicle data, data-driven services, and the corresponding ecosystems have become an important market in recent years, as multiple exemplary statements of analysts in chronological order show in Table 1. The expectations even extend to statements about radical changes in automotive business models due to data-driven service possibilities, e.g. KPMG stated that the “automotive industry is shifting from asset-based to a service and software-driven business model. In 2025 we drive on data and data drives us” (KPMG N.V., 2017).

As a business information scientist, the author of this dissertation has observed that work on data-driven service solutions has been underway in the automotive domain in 2016 already, but even if its potential was (and still is) estimated high, market penetration is rather low in 2021. For example, Dash achieved more than 400,000 downloads and Automatic was acquired 2017 for about $100 million (TechCrunch, 2017), however, both popular services Automatic (c.f. Automatic Twitter, 2020) and Dash (homepage dash.by disappeared in summer 2020) vanished recently.

**Table 1 Exemplary statements of analysts on connected vehicles, data monetization, and trends.**

| Analyst | Exemplary statement(s) |
|---|---|
| KPMG (2016) | “In times of digitalization and connected vehicles, the customer, their data and revenues generated while driving a connected vehicle and using personal gadgets and apps are likely to be more significant than market share based on sold units. That means that in the future, 5,000 connected cars could be more valuable than 50,000 traditional, unconnected vehicles due to valuable revenue streams that can be generated in a connected car by customers providing information about their entire lifecycle.” |
| McKinsey (Bertoncello et al., 2016) | “Companies representing the high-tech, insurance, telecommunications, and other sectors that at once seemed, at most, “automotive adjacent” will play critical roles in enabling car data-related services that customers may be willing to pay for.”<br>“With increasing proliferation of new features and services, car data will become a key theme on the automotive industry agenda and – if its potential is fully realized – highly monetizable.”<br>“Together, these use cases have the potential to result in a total revenue pool of USD 450 - 750 billion by 2030” |

| | |
|---|---|
| PWC (Viereckl et al., 2016) | "In many industries, such as retail, banking, airline, and telecom, companies have long used the data they gather from customers and their connected devices to improve products and services, develop new offerings, and market more effectively. The auto industry has not had the frequent digital touch points to be able to do the same. The connected car changes all that."<br>"As revenues and profits shift from hardware to software, from products to services, and from the old economy to the new one, some players will succeed and others will falter." |
| KPMG (KPMG N.V., 2017) | "We started with the statement "Cars are data-generating engines" and we concluded after our investigations, research and interviews that the statement is definitely true and, although they are already generating a lot of data, cars will be even more data driven in the near future. Apart from generating data, more and more data will also be received and processed to support the driver and interact with his environment."<br>"OEMs do have access to more detailed data, e.g. for product improvement applications."<br>"New regulations in 2018 could possibly open the platform of the connected car to other non-OEM parties"<br>"The automotive industry is shifting from asset-based to a service and software-driven business model. In 2025 we drive on data and data drives us." |
| Ovum (Zoller, 2018) | "The connected car ecosystem is starting to generate data from a wide variety of complementary sources" |
| Deloitte (Hood et al., 2019) | "The amount of data a vehicle generates is set to explode"<br>"OEMs can find opportunities in data, but challenges exist" |
| Capgemini (Cäsar et al., 2019) | "In 2023, there will be 352.9 million connected cars on the road – around 24% of all cars worldwide" [2018: 8%]<br>"AI-powered use cases have high impact on driver experience and mobility services"<br>"Collaborating with the right partners – big players, startups, and other OEMs – will be critical for success with connected vehicles."<br>"An important element of future mobility is an open service platform" |
| Intellias (Haydin, 2020) | "Connected and autonomous cars create new business opportunities for personal data monetization. However, organizational, technological, social and regulatory issues make strategic partnerships critical for OEMs." |
| KPMG (2020) | 1154 automotive executives were surveyed for automotive key trends until 2030:<br>- "50% rated *connectivity & digitalization* as extremely important", which thus is ranked #2 in 2020 (2016 & '19 ranked #1, 2017 & '18 ranked #2).<br>- "45% rated *Understanding the mobility ecosystem* as extremely important", which thus is ranked #6 in 2020 (like 2019).<br>- "41% rated *(Big) data monetization (e.g. vehicle & user data)* as extremely important", thus is ranked #9 in 2020 (2016, '17 & '19 ranked #7, 2018 ranked #6).<br>"[Vehicle] Data is the raw material for ICT company business models" [..] "For years, automotive companies have been trying to find out in which areas data-driven business models have the highest probability of success." [..] "Will data one day become so valuable that mobility becomes free of charge?" |

As the analyst-statements above show, the potential is increasingly recognized, but there are still open points to be explored, which will be addressed in Subsection 1.2.3. First, however, the upcoming Subsection will demonstrate that data-driven services in the automotive sector have developed into a vital research topic during the dissertation project.

### 1.2.2 Vitality of Research

On the national level in Germany and Austria, as well as on the European level, several research projects and even a research program (BMWi program “Smart Service Welten” with the dedicated cluster “Mobilität”) are aiming to develop solutions to ease data-driven service development using vehicle data. The author of this dissertation had the chance to contribute to the projects AEGIS, EVOLVE (both funded by EU H2020), SCOTT, InSecTT (both EU ECSEL JU), and D-TRAS (bilateral Germany and Austria funded by BMWi and FFG). Further projects of other consortia worth noting include e.g. AutoMat, Cross-CPP (both EU H2020), CAR-BITS.de, and StreetProbe (both BMWi).

The increase in research activities is also reflected by the number of scientific publications in this area. To provide an ex-post overview about scientific research activity in the field of value creation in data-driven services in the automotive domain, Figure 1 is intended to provide an overview on how research on value creation through data-driven services in the automotive domain is steadily increasing. Gusenbauer and Haddaway (2020) concluded, that ScienceDirect, Scopus, and ACM DL are appropriate as principal search systems for systematic reviews or meta-analyses, while IEEE Xplore is listed as a supplementary search system, thus they were added. AISeL was added to this overview as well, as the information systems (IS) community is deemed to be relevant for the research of this dissertation, and they call themselves the “central repository for research papers and journal articles relevant to the information systems academic community” (AISeL, 2020). In total, 303 papers were published since 2011, visualized in Figure 1 by the number of papers per search system per year. Articles that used the following terms were included:

```
(value creation OR value)
AND (data-driven services OR data-based services)
AND (automotive OR vehicle OR car OR mobility)
```

More than 52% of the resulting papers have been published in the last two years (2019 and 2020), an indicator that research in this area is increasing and that the topic itself is relevant to research, even if the overall number of related publications declined slightly in

2020 (possibly due to the COVID pandemic situation). Figure 1 also emphasizes, that the dissertation was started when literature and research in this field were still scarce and that the rise of the subject was accompanied.

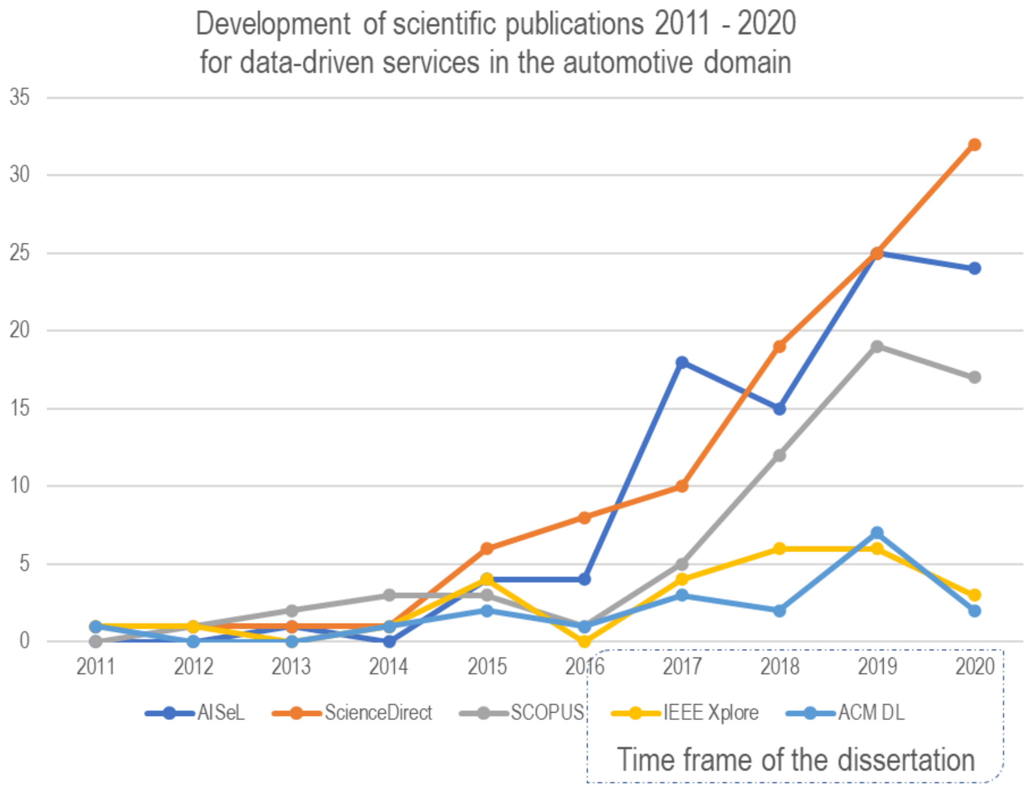


**Figure 1 Search results (all fields) in AISeL, ScienceDirect, Scopus, IEEE and ACM (2011-2019).**

### 1.2.3 Issues and Research Gaps

Although there have been efforts by vehicle manufacturers for several years now to develop data-driven services, as well as research projects and research, relatively few services are offered even in 2020. This is related to how the automotive industry is structured and where the industry comes from. Hence, in the following the issues of long automotive development cycles and the vehicle manufacturers' position on software development and data sharing are exemplified to illustrate the constraints for data-driven service development in the automotive industry. Furthermore, several research gaps are introduced and explained which in consequence serve as a basis for the research questions, which will be presented in Subsection 1.4.1. At the end of each sub-subsection a grey box summarizes the issues or the research gaps to provide an overview.

#### 1.2.3.1 Automotive Development Cycles

One reason why there is still relatively little service on offer is certainly related to the "*issue of long development cycles*" in the automotive industry (Kaiser et al., 2019b), meaning that vehicle development takes about three to five years (Berggren and Magnusson, 2012) from the idea to the start of production. The long automotive development cycles are a problem for software development, e.g. KPMG (2016) states: "As customers increasingly aim to be always connected, relationships are shifting to a much more service-oriented and new data-driven business model for which the traditional automotive industry is rather unprepared compared to other industries, as companies from the information and communication technology sector (ICT)" [..] "In our survey, few respondents show an understanding that some components require a shorter product development cycle. More than 40% of executives from OEMs still say that IT hardware and software components can be developed in the longer product development cycle, which at the end is not solving the clockspeed dilemma." However, long automotive development cycles result in the consequence that "*if it is not foreseen/enabled already today in the development of the vehicle to share specific data, then it will not be possible until about 2025 to have it present in a vehicle on the street*" (Kaiser et al., 2019b), as a representative of a German automotive manufacturing company stated. To summarize, it will thus take a few more years before the developments of data-driven services are available in series vehicles. Whereas the question remains whether expectable changes such as standardization of data provision (interfaces), available signals and data frequency to enable holistic traffic analyses, for example, will again take several years to be available in series vehicles. On the other hand, the major European vehicle manufacturers are not yet properly structured to manage the software development of such data-driven services on their own, while ICT companies still have to overcome a number of hurdles in order to be able to engage accordingly.

> Due to the long development cycles in the automotive industry,
> it takes years for innovations to find their way into series vehicles.

#### 1.2.3.2 Vehicle Manufacturers' Position on Software Development and Data Sharing

Vehicle manufacturers have traditionally tended to develop and sell vehicles that can drive well (Athanasopoulou et al., 2019), and there are still comparatively few departments at vehicle manufacturers that deal with software or even digital services, as the following statements indicate. "Today our 20,000 developers are 90 percent hardware-oriented" Herbert

Diess, CEO of German carmaker Volkswagen said in 2019 (Reuters, 2019). Christian Senger (Head of Software Volkswagen Group) even admits, that “today, our share [of self-developed software as a portion of all software in Volkswagen Group vehicles] is less than 10 percent” (Volkswagen, 2019), which in turn means that 90% were outsourced to supplier software companies in 2019. And so, it is no surprise that innovations for data-driven services came from outside here, from small, agile start-ups (Homfeldt et al., 2019).

Furthermore, vehicle manufacturers were and are unclear about enabling data-driven services. While the ACEA (European Automobile Manufacturers Association) promoted car data sharing (ACEA, 2017), the VDA (Verband der Automobilindustrie e.V.) opposed it. The VDA states, that “modern vehicles have up to one hundred on board control units that constantly communicate with each other to ensure the correct driving and customer functionality” (VDA, 2016), and sees any intervention in these systems (for example, write and read activities on the OBD-2 interface or devices that read the CAN-bus data traffic) as a safety risk, since one could influence the systems. In addition, there is the challenge that vehicles are often on the road for more than 15 years (or much longer as oldtimer), a long time in which the possibilities of the IT industry continue to develop, for example once secure encryption techniques may no longer be considered secure in the future. And to prevent possible hacking of vehicles, VDA suggest that car manufacturers “have to hold a stronger position in the future and may limit the capabilities of third parties to freely access car data.”

> The developers of the most important European vehicle manufacturers are traditionally hardware-oriented, while software development is largely outsourced to suppliers.

After the issues presented, which are intended to show that software innovations have traditionally had a difficult situation with prominent European vehicle manufacturers, research gaps in the development of services are highlighted in the following, which have not yet been adequately addressed by research.

#### 1.2.3.3 Vehicle Data Monetization and Vehicle Data Demand

Concerning *Data,* the first question is whether there is a need that can be met with vehicle data that makes it necessary to generate and collect the data at all, e.g. with *Quantified Vehicles*. In the case of vehicle data, sensor data has been available for a long time because it is needed for the “driving” function. Start-ups then used OBD dongles to collect and analyze data and showed that there is a need. For the vehicle manufacturers, however, it is an investment question. After all, vehicle manufacturers first have to modify the vehicles to enable data collection, e.g. an ECU (configurable from outside) connected to the vehicles' CAN bus which is capable of collecting and transmitting vehicle data to a cloud server. A *Head of Data Services* employed at a German vehicle manufacturer stated in an interview with the author of this dissertation, that "*vehicles are equipped with more expensive technology to enable data sharing*", and that these investments should later be covered by revenues (Kaiser et al., 2021).

Consequently, a highly relevant question is for what data there is a demand, because this influences requirements for vehicle data. Analysts tend to state which areas are monetizable and thus business relevant for companies, e.g. KPMG (2020) identified "four main areas in which companies are investing to make use of data..": "1) Optimization of internal processes, 2) Predictive maintenance, 3) Customer journey mapping, 4) Revenue streams from parallel industries (e.g., insurance, infrastructure, healthcare)".

Researchers like Bauer et al. (2019) investigate how consumers could benefit, e.g. if there is even a market for trusted car data, as vehicle data could increase trust for buyers when buying a used vehicle, and what effect this would have on sales prices. Since potentially even big (vehicle) data can be created (Kaiser et al., 2020b), how to process and persistently store the data is also a topic, e.g. Zhang et al. (2017) mention that "designing large-scale IoV [Internet-of-Vehicles] systems has become a critical task that aims to process big data uploaded by fleet vehicles and to provide data-driven services".

Furthermore, even the European Commission raises the topic of data generated by cars in an article and also highlights the topic of customer engagement: "While value creation will soon concentrate on data generated by cars, critical control points include the HMI, Digital Platforms, real-time geospatial information, and car sensor data. Three main areas of competition are thus emerging: data management, Human-Machine interface and customer engagement" (European Commission; 2017). Especially since the introduction of GDPR, there is also a need to clearly define what the data is collected for, and the questions arise how vehicle data can be accessed and collected, to whom the data belongs, and what quality the vehicle data has, to find out whether the data is reliable.

What are Quantified Vehicles?

What is the demand for vehicle data?

What are the requirements for vehicle data?

How can vehicle data be accessed and collected?

Who owns vehicle data?

What quality does vehicle data have?

#### 1.2.3.4 Data-driven Service Demand and Development

Vehicle data can enable data-driven services*. Data-driven Services* are defined as "services that support the decision-making process of customers through the provision of data and analytics" (Schüritz et al., 2019) which, "coupled with the generation and collection of big data [..] are becoming of great importance to business, economy and society" (Rizk et al., 2018).

However, as explained in the last sub-subsection, someone must also want to use the added value that a service provides to generate revenues and cover the costs (Piparsania, 2019), which raises the question, which user applications can be served with results from the data analysis, and which players in the market are interested in paying for them (KPMG, 2020; Hood et al., 2019). Piparsania (2019) mentions, that Counterpoint (analyst company) "estimates connectivity revenues will exceed 500 billion US dollars by 2030", and lists available services through connected cars including remote features (e.g. door unlocking), emergency assistance alerts, remote diagnostics (e.g. maintenance alerts), onboard features (e.g. vehicle tracker, geo-fencing, parking information) and infotainment. Seiberth and Gründinger (2018) (estimation based on six sources from analysts and researchers) or Koch et al. (2018) even state, that in 2050, 50% of the vehicle manufacturers' revenue will be based on data-driven services. Albertsson and Edström (2013) did a comprehensive study on how a company like Volvo could create revenues from connected car data, including pricing strategies and how they affect value capturing. Thereby, Albertsson and Edström mention direct revenues, i.e. subscription fees and indirect revenues, i.e. an improved customer relationship. Unfortunately it is not published how much the manufacturers currently earn with connected car services, not even in BMW's annual reports, for example. There are only estimates of globally connected car revenues, e.g. “revenues of USD 18.4 billion in 2018 will grow to USD 30.6 billion in 2023” (Statista, 2019).

Nevertheless, there are still relatively few service offerings, and there seemed to be certain hurdles and problems for software companies to develop services. For example, there was no well-documented specification of how a service could be developed, hence publications from the community as well as research carried out in the frame of this dissertation project have shown that several points around Data, Services and Ecosystems were unclear, as explained in detail in the following. For instance, KPMG (2016) concludes, that “around 70% of the executives state that across all corporate functions, data use is at an early stage, or even desired, but its realization and application is to be defined. And some go even further, stating that it is not in use at all”. This entails the question of how the data must be preprocessed so that it can be used properly (Andrienko et al., 2016). For instance, Kolarova et al. (2017) underpin that in stating: “Despite the advantages of vehicle data, there are still some challenges related to data collection and interpretation. The data are not faultless and requires elaborative pre-processing steps to cope with gaps or implausible records. Also, data sets gathered using different methods of data collection differ in their structure, content or variable units.” When it comes to data-driven service development, questions arise on which analyses on the data are valid at all, because poor data quality could lead to wrong assumptions. Zhang et al. (2017) questioned how usable vehicle data really are, by characterizing Internet-of-Vehicle data as “large volume with a low density of value and low data quality” which “pose challenges for developing real-time applications”.

For which data-driven services (based on vehicle data) is there a demand?

How can data-driven services (based on vehicle data) be developed?

How can collected vehicle data be processed and persistently stored?

Which analyses on vehicle data lead to meaningful results?

#### 1.2.3.5 Data-driven Service Ecosystem and Value Generation

With the general shift toward software development, an increase in cooperation and competition between automotive and ICT companies is expected, which makes it interesting to investigate because it is unclear how ICT companies will be integrated into the processes, what roles they will take on, how the balance of power will shift, and how the automotive ecosystem is changing. “In business the digital transformation brings greater efficiency and effectiveness of existing value chains, the realignment of value chains, and opportunities to create new value” (Reddy and Reinartz, 2017). As described above, the automotive industry is currently facing a paradigm shift towards new data-driven services based on vehicle usage data, as they are expected to provide added value to consumers. Thus, the automotive industry will sooner or later establish a process (or several ones), which will transform vehicle usage data into services. Such paradigm shifts are also investigated in other domains like the software industry, and examples from literature exist, which consistently focus on the value creation process to do so. For instance, Hilkert et al. (2010) state, that the “possible paradigm shift in the software industry [..] is to be considered on the level of the involved market players and hence aims at the analysis of value creation structures.” This also applies to data-driven services in the automotive domain, where *Data-driven Service Ecosystems* are the environments in which data-driven services are developed and delivered. While the sum of the actors involved in data-driven service delivery and their relationships form the foundation of data-driven service ecosystems, they are about to change when ICT companies takeover certain roles in the value creation process.

Cooperation and competition between vehicle manufacturers and ICT companies are experiencing ups and downs in recent years. According to KPMG (2020) “competition between automotive manufacturers and ICT companies has increased” in 2020, while the trend was rather towards cooperation in 2019. A representative from Flixbus explains, that “IT companies are closer to the customer than OEMs (e.g. Android on the smartphone). OEMs need to build an ecosystem to observe client behaviors, and IT companies have a large advantage here. As of today, they will always come faster into information”, while also a representative from Audi AG agrees: “When it comes to data services, collaboration is key. No car maker can tackle the challenges on its own” (Seiberth and Gründinger, 2018). However, from a research perspective and related to the ecosystem topic, it is yet unclear which actors exist on the market, and who takes over which role in the ecosystem to jointly develop data-driven services. Hence, the question is, which cooperations the existing and new actors have to form (Cäsar et al., 2019; Kaiser et al., 2019b; Haydin, 2020; Kaiser et al., 2021) to jointly develop data-driven services.

Furthermore, the question of cooperation also raises the question of where ICT companies should contribute their expertise, which in turn raises the underlying question of how value can actually be generated from vehicle data. If the data and services do not generate added value (for example, revenues or increased safety and satisfaction), they will not be considered in the long automotive development cycles and thus not be integrated into / enabled by future vehicles. But even if future vehicles are going to collect data and enable services, it is necessary to clarify how value is generated in data-driven service development in order to make the fundamental and tacit knowledge explicit and thus available to all interested parties. If it is, for instance for many ICT companies, unclear how to get the data, how to process it, or if the data quality is insufficient and does not allow for valuable analyses, it will be difficult to develop services that convince customers.

To answer these questions, a description of value creation would be helpful, but even more prominent representatives of data-driven value creation seem to have their problems with this. Comparing vehicle data with big data, the literature review of Furtado et al. (2017) shows that there is no clear view on value creation with big data yet, but many similar approaches. Nevertheless, two central application areas of big data value creation mentioned by Davenport (2014) should be mentioned: (i) data as input for explanatory and predictive models, to improve decision making, and (ii) product and service improvement through insights from data analysis (Davenport, 2014; Furtado et al., 2017). However, large quantities of (vehicle) data (big vehicle data) was used to promise great potential for value creation in the past. But this assumption, *big data => great potential*, is exactly what is not yet working for big data. “Yet this big data revolution has so far fallen short of its promise” Huberty (2015) states while adding: “Close examination shows that firms have largely used big data to improve on existing business models, rather than adopt new ones; and that those improvements have relied on data to describe and predict activity in worlds largely of their own making.” Huberty therefore criticizes the fact that big data is mainly used to optimize existing data businesses. But he also suggests what should change: “The big gains from big data will require a transformation of organizational, technological, and economic operations on par with that of the second industrial revolution. Then, as now, firms had to invest heavily in industrial research and development to build the foundations of entirely new forms of value creation. Those foundations permitted entirely new business models [..]” (Huberty, 2015). And also in the automotive sector, for example in 2020, presentations at the conference “monetizing car data” have been pushing back in order to lower expectations, as these new business models have yet to be explored. Therefore, this dissertation also focuses on value creation and investigates the process of how value is created, to illustrate it by means of a value chain. Hence, especially the questions related to Data and Services posed in the previous sub-subsections are not in random order but involve a certain process along the value chain, from data via analysis to service, along a *Vehicle Data Value Chain*. While such value chains already exists for big data (e.g. *Big Data Value Chain* by Curry, 2016), there is no such value chain for data-driven services in automotive yet, a clear research gap.

At the same time, it was noticed that it is perhaps not yet clear who plays what role in the ecosystems and the value creation process, for example, in accessing data created from vehicle sensors ("vehicle data"). Since in 2016, when this thesis was started, there was usually no data interface from the vehicle available, that was intended for data exchange. Thus, many software companies used the vehicle's on-board diagnostic interface (OBD-II), which is standardized but, as experienced, implemented differently for each vehicle and manufacturer. The possibility of accessing vehicle data via the OBD interface without even involving the vehicle manufacturer led vehicle manufacturers to argue that this should be prevented for safety reasons (VDA, 2016). A well-thought-out idea in times of successful platforms providers, as this would put vehicle manufacturers in a position where they are the only ones who can provide access to data, and thus determine the price. Some companies (e.g. Zendrive) use the smartphone and its sensors to be independent of the vehicle. In addition to software companies (vehicle data service providers) and vehicle manufacturers, there were and are other players who want to find their place in the ecosystem, such as external gateway providers, data intermediaries, data marketplaces, external data source providers, and many potential consumers, as we have learned. Consumers can also play a decisive role in this respect, as the services could highlight various safety-related problems (potholes, driver distraction, danger zones, etc.) that are of public interest (e.g. traffic planning, traffic radios, etc.), which could lead to regulation that these data must be available, probably free of charge. Such regulations could restart the process of ecosystem formation in upcoming years. One more reason why the relevant sub-themes should be examined accordingly.

To summarize, we have shown in the last three sub-subsections, divided into the topics of Data, Services and Ecosystems, that around the topic of *Quantified Vehicles* there are still many unanswered questions at the beginning of the dissertation (gray boxes). It is unclear what *Quantified Vehicles* are, what data they can provide, what restrictions exist, and what the quality of the data is. Furthermore, there are open points in the development and provision of data-driven services, which makes the development of services complex and challenging. This in turn is due to the lack of knowledge of the ecosystem, because neither the

Which cooperations do existing and new actors in the data-driven service ecosystem need to enter into in order to develop data-driven services (based on vehicle data)?

How can value be generated from vehicle data?

What is the underlying data value chain that enables data-driven services based on vehicle data?

Who plays what role in the data-driven service ecosystems and in the value creation process?

players and their relationships are clear, nor is the value chain known. This dissertation seeks to address this particular research gap in order to enable and support the design and exploration of *Quantified Vehicles*, the data they generate, the data-driven services they enable, and the ecosystem of stakeholders involved.

## 1.3 Focus and Objectives of this Dissertation

Digitalization turns vehicles into *'Quantified Vehicles'* when they collect data about themselves and their environment. *Quantified vehicles*, which are often described in the literature as connected vehicles (although this term is more associated with autonomous driving functions) thus produce data with their vehicle sensors.

Consequently, *Data* is the first main aspect of this dissertation, as data is essentially the enabler for any services. The set of available signals and their maximum sampling rate, in fact, determines possible analyses and, as a result, possible target markets and customers. And it is just these target markets and customers who influence, through their demand, which service offering will be successful. *Services* are the second main aspect of this dissertation. *Quantified Vehicles* enable such data-driven services (through the data generated). The hypothesis that the data of millions of vehicles on the road must not only serve driving, but as a by-product that can support higher targets such as Vision Zero, was the main reason to choose this topic as a thesis topic. At the beginning of the dissertation, there were initiatives and services on the market where it was unclear how they are structured and how they function and what purpose they serve. Consequently, there were open questions that are to be addressed here. And even the first look at services on the market showed that development by a single company is possible, but scales poorly. For larger solutions, it was foreseeable that cooperation between several players would be required, for example, vehicle manufacturers and service providers. And with these cooperations, it will of course be exciting to see who takes on which role and who has which business models. This brings us to the third aspect of this dissertation, *Ecosystems*. *Quantified Vehicles* lead to the emergence of ecosystems in service creation processes. Figure 2 summarizes in an overview, how the three key constituents of the topic *Quantified Vehicles – Data, Services and Ecosystems* – belong together. To repeat, *Quantified Vehicles* generate *Data*. The data of *Quantified Vehicles* enables *Data-driven Services*, while the sum of actors involved in service delivery and their relationships constitute the *Ecosystems*.

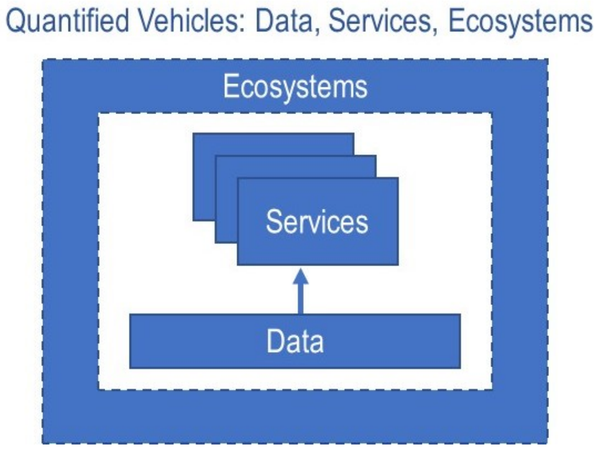


**Figure 2** ***Quantified Vehicles* provide *Data* for data-driven *Services*. The sum of actors involved in service delivery and their relationships constitute the *Ecosystem*.**

To summarize, the three key constituents *Data, Services, and Ecosystems* are relevant and fundamental topics related to the business informatics discipline and design- and knowledge-oriented contributions within this dissertation will ultimately support also other disciplines in their work. It is the goal of this dissertation to create knowledge and insights on the key constituents of *Quantified Vehicles*: *Data, Services and Ecosystems*, to ultimately support the answering of the related research questions (will be presented in the following Section in Subsection 1.4.1).

This analysis could be done involving several disciplines, which all provide added value for research: In addition to psychology, whether indications of wrong driving behavior actually help or additionally distract, it is on the one hand computer science with architectures and the software development to develop services, and on the other hand mechanical engineering in order to be able to correctly interpret sensor data, which is actually used for purposes other than those for which it was intended and was never meant to be passed on. Last but not least, business informatics provides the glue between all relevant disciplines. Business

Informatics connects IT and business, combining technical, business, and sociological knowledge for the recognition and optimization of economic processes (Gabler Wirtschaftslexikon, 2020; Fink et al., 2006). The business information scientist understands both economic relationships and software code. As an essential contribution, tools and methods are developed to understand the problems and to solve them most economically, e.g. to develop the services together with a software team and accordingly to work out concepts how such services can be scaled. As a business information scientist, the author of this dissertation focused on the latter, being aware that other disciplines are equally important. Thereby, the objectives are to:

(i) better understand how Vehicle Data becomes a relevant artifact for business and innovation
(ii) define and describe *Quantified Vehicles* as a form of digitalization in the automotive domain
(iii) develop concepts and data-driven services prototypically, that represent added value for consumers to enhance the understanding of challenges in service development
(iv) better understand the process and actors of value generation, and the interplay of the actors with each other in the ecosystem, by conducting empirical research involving automotive domain experts
(v) conduct design activities backed by empirical research to conceptually model data-driven value generation and data-driven service ecosystem building

This section is now followed by Section 1.4, which presents the research approach used. The focus on the three key points Data, Services and Ecosystems is also reflected in the three research questions, which are presented in detail in Subsection 1.4.1. In the process of answering the research questions, both design-oriented and knowledge-oriented contributions were developed, as is shown in subsections 1.4.2 'Research Process' and 1.4.3. 'Research Methods'. Section 1.5 then provides an overview and a detailed description of all contributions to answer the three research questions. The individual contributions themselves (all of which were also published in very similar form as a peer-reviewed publication) are then presented in detail in chapters 2-15. Finally, a conclusion, a summary, and an outlook of the dissertation are drawn in Chapter 16.

## 1.4 Research Approach

### 1.4.1 Research Questions

To investigate the undoubtedly very recent and exciting topic of *Quantified Vehicles*, the data generated by them, the services they enable, and the actors and relationships in the ecosystem of value creation, research questions for this thesis are set up. Since the topic is new and there is no relevant preliminary work that exactly matches the topic, the central question is, what are vehicle data-driven service ecosystems, and why are they relevant for business informatics. This superordinate topic was already addressed in previous parts of this Synopsis, which allows us to move on to more specific questions, each of which focuses on one of the three concepts within the dissertation title: *Data*, *Services* and *Ecosystems*.

A digital ecosystem, as defined by Gartner Research (2019) "is an interdependent group of enterprises, people and/or things that share standardized digital platforms for a mutually beneficial purpose (such as commercial gain, innovation or common interest)." We apply this to the *Data-driven Service Ecosystems*, which are thus formed by the organizations involved in service delivery and their relationships to each other, which in turn partially use vehicle data from *Quantified Vehicles*. In the context of this dissertation, the interest was particularly high to find out, how data-driven services can be technically created from vehicle data and how this is managed resulting in data-driven service ecosystems. In the *Vehicle Data Value Chain* steps, first of all, it requires vehicles, which, with the appropriate hardware and software, generate vehicle data and enable the acquisition of vehicle data. Such vehicles can be termed *Quantified Vehicles* (Stocker et al., 2017a). Thus, the first research question deals in detail with what *Quantified Vehicles* actually are and how important they are for ecosystem development. Thereby it covers the concept of *Data*.

***Research Question 1**: What are Quantified Vehicles and why are they important for Data-driven Service Ecosystems?*

As soon as it has been clarified what data *Quantified Vehicles* can make available for service developers, what standards and developments exist for accessing vehicle data, the question of the corresponding concrete technical implementation of services arises. A broad spectrum from i) the definition of value-creating steps (Vehicle Data Value Chain), ii) conceptual software architectures to shed light on aspects such as efficiency, effectiveness, privacy, etc. up to iii) the description of concrete prototypical implementations of data-driven services address the second research question, which thereby covers the concept of *Services*.

***Research Question 2**:* *How can vehicle data-driven services be engineered in an efficient and effective way?*

And since such services, which in the future are expected to be integrated into every new vehicle and to have enormous business potential, the hypothesis is made that several actors with interrelationships form the data-driven service ecosystem. To advance research in this respect, it is important to explore how such ecosystems are constituted. Hence, the third research question covers the concept of *Ecosystems*:

***Research Question 3**:* *What are important actors and relationships for service delivery in Vehicle Data Service Ecosystems?*

To answer the three research questions, an iterative research process using a mixture of different research approaches and methods was started in 2016 and is described in detail in the following.

### 1.4.2 Research Process

Congram and Epelman (1995) proposed the so-called Structured Analysis and Design Technique (SADT), originally introduced by Ross (1977; Dickover et al., 1977), to describe activities from a process perspective. Similar to Ahmed (2016), the author of this dissertation sees the methodological framework SADT as flexible in terms of notations and steps, thus suitable for describing research processes of scientists, such as those of a dissertation, for example. Figure 3 visualizes the generic structured activity box of SADT. An activity has three inputs and one output. The three inputs are 1) *Control*, which represents external influences, such as regulations, standards, etc., 2) *Input*, which represents what was added to the activity, and 3) *Mechanism*, which represents how the input was generated or used. Consequently, *Output* represents the output of the activity described.

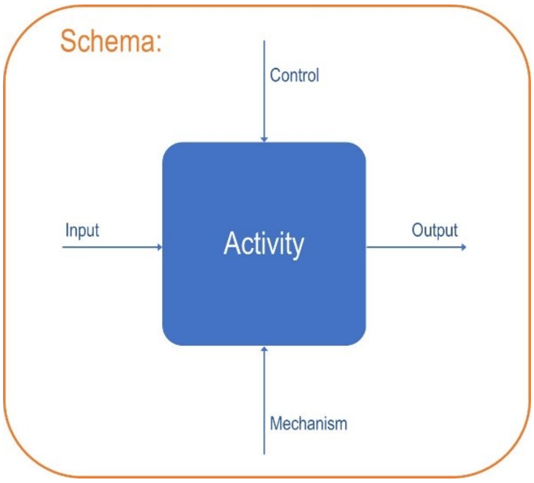


**Figure 3 The generic structured activity box of the Structured Analysis and Design Technique (SADT) (Own representation, originals to be found in Congram and Epelman (1995) or Dickover et al. (1977)).**

In Figure 4, the research process carried out in this dissertation is visualized according to the Structured Analysis and Design Technique (SADT). Starting activity "*Formulate Research Questions*" on the upper left corner acts as a control element for all other activities. In accordance with the research questions on the main topics *Data, Services and Ecosystems*, the three activities i) "*Define and introduce 'Quantified Vehicles' (QV)*", ii) "*Define the Vehicle Data Value Chain (VDVC)*" and iii) "*Analyze the Vehicle Data-driven Service Ecosystem*" emerged. With increasing knowledge about the value chain and the ecosystem, hardware and software artifacts were developed with the activity "*Design and Develop Artifacts along Value Chain Steps*" within research projects (AEGIS, SCOTT, EVOLVE), which the author worked on during his dissertation. The findings were also used to show the information system community with a research agenda where there is still potential for further research. This is included in the activity "*Define a Research Agenda for the IS community*". The individual activities were carried out using a wide variety of mechanisms (which have been colored in each case), in this case rather research methods, such as "*Literature Review*" (black), "*Conceptual / Reference Modeling*" (yellow), "*Inductive Research*" (orange) "*Case Study*" (purple), "*Prototyping*" (grey), "*Qualitative-empirical Cross-sectional Analysis*" (green), "*Empirical Intervies / Field Study*" (dark red). Activity outputs can be found on the right side, e.g. "*Data-driven Service Ecosystem Model*".

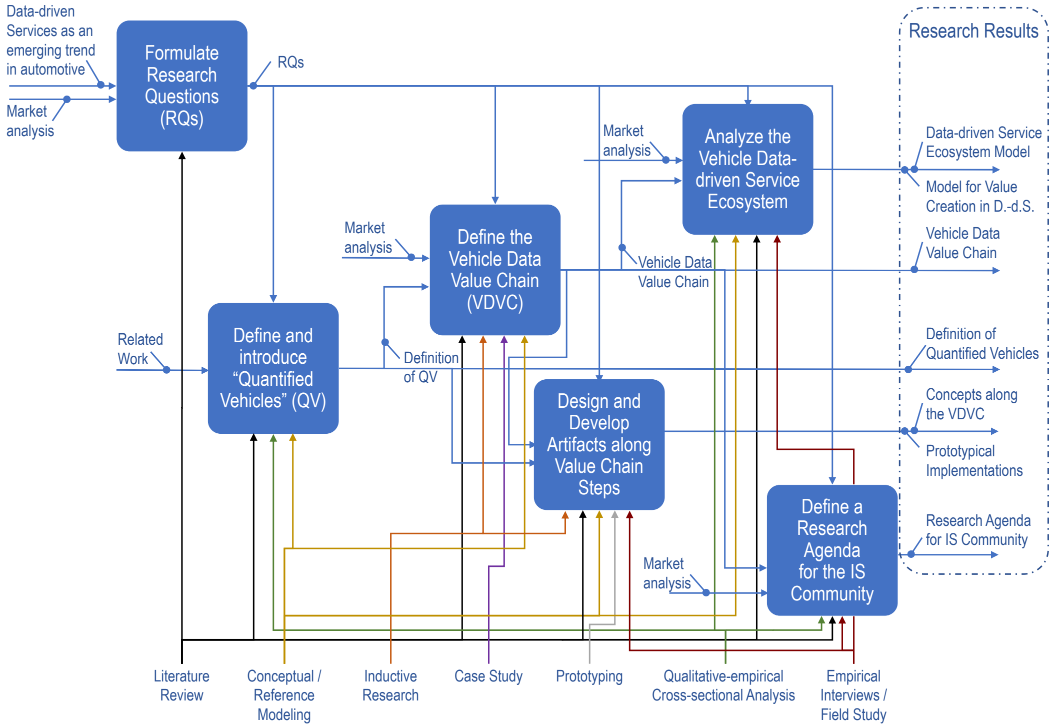


**Figure 4** **The research process carried out in this dissertation visualized with the Structured Analysis and Design Technique (SADT).**

In order to look at the process from a different perspective too, it will be presented in chronological and simplified form in the following, visualized in Figure 5. Again, to answer the research questions RQ1-RQ3, several process steps were run through in the course of a research process, and contributions were developed which answer the research questions. In 2016, when literature on the exploitation of vehicle data for data-driven services was scarce, the term *Quantified Vehicles* was derived from Quantified Self and defined accordingly. In a subsequent step, a market analysis provided an overview of data-driven services offered by vehicle manufacturers and third-party companies such as start-ups. Based on this market analysis and to advance research in the field of information systems, a research agenda for the IS community was developed. Since value creation followed a certain pattern, which was also evident in the activities conducted by the reviewed start-ups, the *Vehicle Data Value Chain* (VDVC) was first introduced as a conceptual model to sketch data-based value creation and further improved in several iterations using the knowledge of automotive domain experts. Furthermore, the development of concrete examples for data-driven services and concepts was started to better understand not only the business challenges but also the technical challenges. Finally, as of 2018, research revealed that several actors are involved in the creation of data-driven services, creating vehicle data-driven service ecosys-

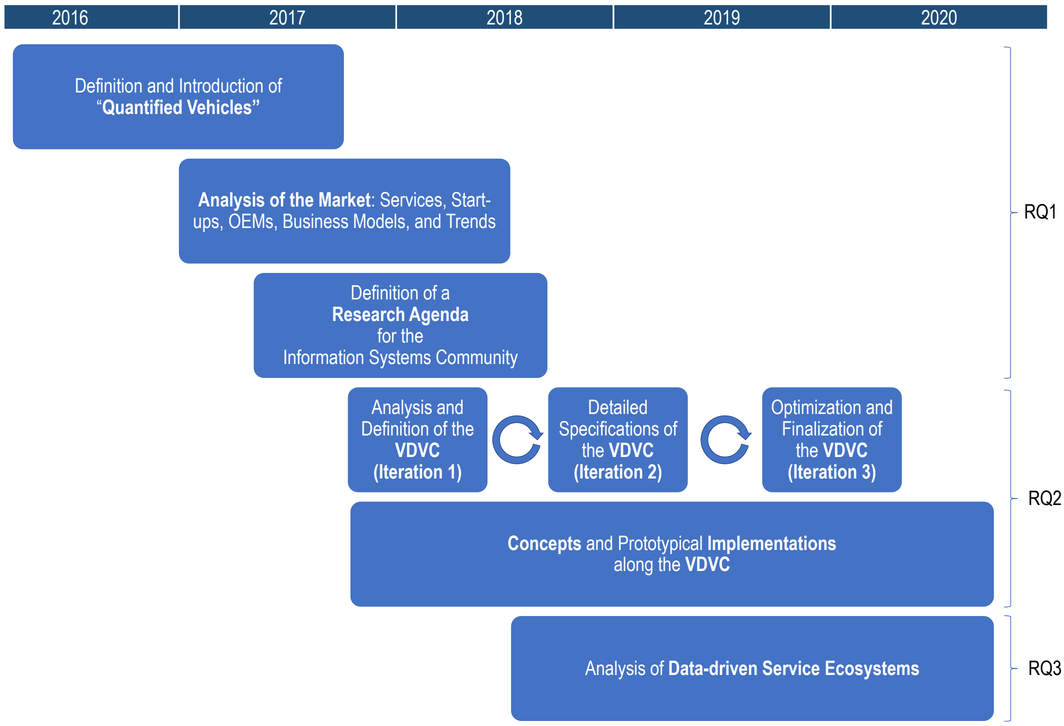


**Figure 5 The research process carried out in this dissertation (chronological and simplified).**

tems with dependencies and strong business relationships. Consequently, work was conducted until mid of 2020 to better understand such ecosystems and the role of data-driven value generation. However, it should be noted that these steps were interwoven and that it was moved back and forth between the activities, which is obvious due to the strong link between data, services and ecosystems.

In the course of the research process, the results already briefly mentioned were obtained, with the next section briefly showing an overview of the set of methods used for this purpose.

### 1.4.3 Research Methods

This dissertation was written as a cumulative dissertation. A cumulative dissertation is a dissertation consisting of several scientific articles, with each article measured against the quality standards of the corresponding international scientific community. Thereby, each article has a specific objective, uses a proper method (or even mixed-methods by combining qualitative and quantitative research, c.f. Jogulu and Pansiri, 2011), and presents findings, which contribute to one ore more research questions. Consequently, there is no methodology chapter in this dissertation that describes the one specific method that was used. However, a brief overview of which methods were used in the individual contributions / publications will be shown in the following, while they are then presented in detail in the chapters that introduce and present the publications.

In scientific articles it is common to show what the current state of knowledge is by means of a *Literature Review* (e.g. summarized in a related work section). Here, too, a literature review (sometimes more or less detailed, depending on the publication) was carried out in all fourteen publications to build the scientific basis, in which the author of the dissertation was involved in thirteen out of fourteen cases. For the further content in the publications, a comprehensive set of methods was used. For instance, *Inductive Research* was used to develop theories on which there has been little research, including e.g. the *Vehicle Data Value Chain (VDVC)*, and the *Open Vehicle Data Platform (OVDP).* Furthermore, design-oriented methods were used to create new artifacts, for instance the method of *Conceptual Modeling* was used to design five published iterations of the *Data-driven Service Ecosystem* (once based on *Design Science)*, while *Reference Modeling* was used to design the *VDVC* model, and *Prototyping* was used to develop concrete *Data-driven Services*.

In addition, behaviouristic methods were used to study behavior and trends in relation to IT usage. In this respect, the method *Qualitative-empirical Cross-sectional Analysis* was used to introduce the topic of *Quantified Vehicles*, analyse the positions and actions of popular start-ups and vehicle manufacturers, to provide a research agenda, and to analyse the *Data-driven Service Ecosystem.* The *Case Study* method was used to evaluate the VDVC, while *Empirical Field Studies* were used to demonstrate and evaluate a prototypical setup and to iterate a model on *privacy levels in sharing vehicle data*[3] based on user feedback. Finally, also empirical interviews with experts were the basis for in total three publications, including a research agenda and a *Data-driven Service Ecosystem* model.

There are also some methods that are relevant to the topics covered by the dissertation but have not been explicitly mentioned or applied in the publications. These will also be briefly listed here, along with an explanation of why they were not used or how they influenced the work. For instance, in the automotive industry, the focus has long been "on tangible output", a mindset described in *Product-Dominant Logic* (P-D Logic), "which arose from the success of the *Industrial Revolution*" (Spohrer et al., 2008a). In the meantime, a second worldview or mindset has been added, the *Service-Dominant Logic* (S-D logic) for marketing. Two coexisting mindsets in conflict, or a "product–service continuum" (Olivia and Kallenberg, 2003). Since the dissertation is also about data-driven services, the author has considered

[3] Planned and supervised by the author of this thesis, but carried out by student as part of his master's thesis.

the theoretical framework of *S-D logic*, for example the ten “service-dominant logic foundational premises” (Vargo and Lusch, 2008), to understand services. However, the most cited scientific *S-D logic* contributions are in the field of marketing and retail, and so it is understandable that their focus is on understanding the service world, not how to build services. This can already be recognized in the formulation of the *S-D logic* foundational premises (FP), e.g. FP4: “Knowledge is the fundamental source of competitive advantage” and FP7: “The enterprise can only make value propositions”. Unfortunately, *S-D logic* does not help with the concrete development of artifacts, because it is not a method and does for instance not provide a template on how to access or process (vehicle) data. Consequently, also *Service Science*, “the study of value co-creation interactions among entities” (Spohrer et al., 2008b) known as service systems, which is built on the foundation of the above mentioned *S-D logic*, is a theory to better understand service innovation (Maglio and Spohrer, 2008), which has its origins in marketing too, and is based on ten foundational concepts too, for instance resources, service system entities, and access rights (Spohrer et al., 2008a). In line, as a dissertation in the field of business informatics, this thesis aimed also about finding out how services are created, though less theoretically, rather what the technical hurdles are and what the technical development steps are. Thus, *S-D logic* as well as *Service Science* were dealt with at the beginning of this dissertation, but it was interpreted more as a basic literature. In the rather technical contributions of the “technology-driven subject” *Quantified Vehicles* (Stocker et al. 2017a), which also performs analysis of *Data-driven Service Ecosystems* from the perspective of the data and does not analyse the money (value) flows (Kaiser et al., 2021), *Service Science* is mentioned only once (in Kaiser et al., 2019b) in the fourteen publications included in this dissertation.

Similarly, *Product-Service Systems* (PSS) Engineering is more technical than *Service Science*, but still provides rather theoretical contributions. Müller (2013) defines *PSS* as “customer, lifecycle, and sustainability oriented socio-technical systems, solutions, or as offers, integrating products and services” and adds that “*PSS* is a system level approach although components are affected by requirements broken down from the system level”. Consequently, the most frequently used methods in association with *PSS*, according to Cavalieri and Pezzotta (2012), are precisely methods that support requirements engineering or concept development on a meta-level, and do not directly support the development. E.g. the method *TRIZ* is used “to optimize the idea generation process”, while the method *Quality*

*Function Deployment* (QFD) is used “to translate customer requirements into engineering characteristics” (Cavalieri and Pezzotta, 2012).

Moreover, also a differentiation to hybrid products should be given. Hybrid products, or "hybrid value creation strategy" (Velamuri et al., 2011), different terms exist, is a strategy that is also present in the automotive industry. It is “i.e. the process of generating additional value by innovatively combining products (tangible component) and services (intangible component)” (Velamuri et al., 2011). In case of the automotive domain, the vehicle is supplemented by services or offered as a service. The literature review by Velamuri et al. (2011, Table 3) shows that publications in the field of hybrid value creation strategy take a macro / strategic / marketing / business level / innovation / design / organization perspective, or examine sustainable aspects. However, it was much more the goal of this dissertation to find out how data-driven value creation is structured in the automotive industry (a rather technical / data perspective), to make it explicit / transparent, and to understand it, and not to develop for example a strategy how a more sustainable ecosystem would look like. That is why the publications never mention hybrid products or closely related concepts.

To summarize, each peer reviewed publication answers a specific sub-research question mentioned in the paper, and for the rather data-driven / technical or understanding-oriented contents the respectively most appropriate methods were chosen, which in retrospect also withstood the scientific discourse and were positively reviewed. In the following section, the research results are now presented in detail, structured according to research questions to which the publications have made the greatest contribution.

## 1.5 Research Results

As shown (c.f. Figure 2), *Data, Services and Ecosystems* are strongly interrelated. For instance, in data-driven service ecosystem studies, the services and their stakeholders are considered, which in turn are based on specific vehicle data. Therefore, the majority of the contributions which are presented in chapters 2 to 15 address not only one topic, but two or even all of the three topics (*Data, Services and Ecosystems*), and consequently contribute to answering several research questions.

To reduce complexity, each contribution and the related publication in which the contribution has been communicated to the scientific community is linked only to the research question to which it makes the main contribution. The following Figure 6 provides an overview of which publications address which research question. Each star figure in column *Scientific Publications* stands for a publication, for instance, *E&I* is the abbreviation for (German-speaking) Journal Elektrotechnik und Informationstechnik. Each contribution / scientific publication corresponds to one chapter in this dissertation.

| Research Question | Contributing Chapters | Main Findings | Scientific Publications |
|---|---|---|---|
| ***RQ1**: What are Quantified Vehicles and why are they important for Data-driven Service Ecosystems?* | *2-6* | *- Definition "Quantified Vehicles"*<br>*- Analysis of Data-driven Services*<br>*- Research Agenda for IS Community* | E&I, BISE, NBM, i-K now, ECIS |
| ***RQ2**: How can vehicle data-driven services be engineered in an efficient and effective way, and how can they be compared?* | *7-13* | *- Vehicle Data Value Chain (VDVC)*<br>*- Concepts along the VDVC*<br>*- Prototypical Implementations along the VDVC* | WEB IST, LN BIP, IC VES, LN MOB, CAi SE, BIT KOM, VEH ITS |
| ***RQ3**: What are important actors and relationships for service delivery in Vehicle Data Service Ecosystems?* | *14-15* | *- Data-Driven Service Ecosystem Model*<br>*- Conceptual Model for Value Creation in Data-driven Services* | AMC IS, IJIM |

**Figure 6 Research questions, chapters contributing to their answers, main findings, and scientific publications that address them.**

After this overview, the next sections provide summarized insights into the particular scientific publications and the contributions of the author of this thesis. In addition, for each research question, a table provides an overview of the chapters answering the question, including title of the publication (which is equal to the chapter title, or translated into English), results and research methods (from the author of this thesis), and the medium (a specific journal, conference, or series) of the scientific publication. For contributions in well-known journals and conferences, the current ranking in the following three ranking lists is provided, if available:

- *VHB Jourqual 3: business informatics* [4] (short "VHB-JQ3" in the tables)
  - applicable for both, Journals and Conferences
- *Computing Research & Education* [5] (short "CORE`18" in the tables)
  - applicable for Conference Proceedings only
- *ABDC Journal List* [6] (short "ABDC-JQ" in the tables)
  - applicable for Journals only

4 https://vhbonline.org/fileadmin/user_upload/JQ3_WI_01.xlsx

5 http://portal.core.edu.au/conf-ranks/?by=all&source=CORE2018

6 https://abdc.edu.au/wp-content/uploads/2020/04/abdc_jql_2019_0612-1.1.xlsx

### 1.5.1 Results for Research Question 1

First, targeting Research Question 1, w*hat are Quantified Vehicles and why are they important for Data-driven Service Ecosystems*, Chapters 2 and 3 contribute to answer it with a deduction and definition of the terms *Quantified Car* and *Quantified Vehicles* from Quantified Self and also introduce surrounding topics, including services, business models, and digital ecosystems. In addition, Chapters 4 and 5 provide an analysis of the market by examining the approaches of the two main groups of players – vehicle manufacturers and tech start-ups – including a business model analysis of their existing services. Furthermore, Chapter 6 demonstrates the relevance of *Vehicle Information Systems* (Vehicle IS) for the IS community. *Vehicle IS* are thereby defined as “a class of software applications processing vehicle data and/or other relevant data from different sources to finally provide valuable and action-relevant information to the vehicle driver and/or to other stakeholders.” The chapter closes with a set of example research questions for selected research directions, to show that there are still many unresolved research questions. Table 2 provides an overview of all contributions to answer Research Question 1, showing methods and results of the author of the dissertation, and the publication medium. After the table, the text goes into more detail about the contributions of each chapter, structured into the corresponding subtopics of the research process (c.f. Figure 5): (i) *Definition and Introduction of “Quantified Vehicles”*, (ii) *Analysis of the Market: Services, Start-ups, OEMs, Business Models, and Trends*, and (iii) *Definition of a Research Agenda for the Information Systems Community*.

#### 1.5.1.1 Definition and Introduction of “*Quantified Vehicles*”

As described in the introduction, publications and research were still scarce when this dissertation was started in 2016 and 2017 (c.f. Figure 1). Therefore, the dissertation project started with a definition and introduction to the topic as well as the establishment of the name *Quantified Vehicles*. This was done with two initial publications in scientific journals: *E&I* (Chapter 2) and *BISE* (Chapter 3). The idea for this dissertation project was still emerging in these two initial publications, therefore the author of this dissertation was at not the main author of these two publications. After that, however, the author of the dissertation took over the lead.

In Chapter 2, it is described that the continuous collection of vehicle data facilitates the generation of innovative products, services, and business models. This is underpinned with three case studies of start-ups – Automatic, Mojio and Dash – and investment figures.

**Table 2 Chapters contributing to RQ1, their title, author results, research methods, and medium it was published in.**

| Ch. | Title | Author Result(s) (AR) & Research Method(s) (RM) | Medium and Rank |
|---|---|---|---|
| 2 | Quantified Car: Potentials, business models and digital ecosystems (translated into English) | *AR*: Introduction to the Quantified-Car-phenomenon, Definition of Quantified-Cars, Analysis of business models of startups providing Data-driven Services, and a Data-driven Service Ecosystem for Quantified Cars.<br>*RM*: Literature Review, Qualitative-empirical Cross-sectional Analysis, and Conceptual Modeling. | (German-speaking) Journal Elektrotechnik und Informationstechnik (e&i) |
| 3 | Quantified Vehicles: Novel Services for Vehicle Lifecycle Data | *AR*: Introduction and Definition of Quantified Vehicle as a topic and its relevance to the Information Systems (IS) community. Market analysis (e.g. five start-ups) w.r.t. business models.<br>*RM*: Literature Review, Qualitative-empirical Cross-sectional Analysis, and Conceptual Modeling. | Journal Business & Information Systems Engineering (BISE)<br>VHB-JQ3: **B**<br>ABDC-JQ: **A** |
| 4 | Quantified Cars: An exploration of the position of ICT start-ups vs. car manufacturers towards digital car services and sustainable business models | *AR*: Investigation of the status of *Data-driven Services*, including stakeholders, start-up examples and their value propositions, two approaches to collect data (OBD-2 vs. smartphone), and the current position of vehicle manufacturers.<br>*RM*: Literature Review, and Qualitative-empirical Cross-sectional Analysis. | Intern. Conference on New Business Models (NBM) |
| 5 | Digital Vehicle Ecosystems and New Business Models: An Overview of Digitalization Perspectives | *AR*: Investigation of the status of *Data-driven Services* including a brief investigation of selected vehicle manufacturers and their digital service strategies.<br>*RM*: Literature Review, and Qualitative-empirical Cross-sectional Analysis. | Intern. Conference on Knowledge Technologies and Data-driven Business (i-KNOW) |
| 6 | A Research Agenda for Vehicle Information Systems | *AR*: Introduces Vehicle Information Systems (Vehicle IS) as a new class of information systems (IS). Relevant research directions and example research questions are provided for the IS community.<br>*RM*: Literature Review, Empirical Interviews, and Qualitative-empirical Cross-sectional Analysis. | European Conference on Information Systems (ECIS)<br>VHB-JQ3: **B**<br>CORE`18: **A** |

Thereby, the author of this dissertation contributed by examining the quantified car startups' (from the US) business models, which use vehicle data for data-driven services (see Table 3, an excerpt of the table which is published in Stocker and Kaiser, 2016). In addition the dissertation author was responsible for (i) the data-driven service ecosystem model presented, which shows how ICT players will contribute to creating a new data-driven service ecosystem within the automotive domain, and (ii) for the conclusion and discussion of the introduction of *Quantified Vehicles* (referred to in the article as *Quantified Car*). The ecosystem model includes important actors, such as cloud providers, service providers, primary and secondary end-users.

**Table 3 An excerpt overview of *Quantified Vehicle* start-ups. (Source: Stocker and Kaiser, 2016)**

| Company | Value proposition | Possible use | Business applications | Costs | Inves-tors |
|---|---|---|---|---|---|
| **auto-matic.com** San Francisco (USA), founded 2011 | Connects car with the digital life of the driver. Enables drivers with knowledge about themselves and the vehicle to drive safer and smarter. | Connects the vehicle to many apps, for example for problem diagnosis, consumption optimization, location and emergency services. Includes a web dashboard with comprehensive statistics. | Cloud-based services for insurance companies, fleet operators and vehicle manufacturers, provision of the data analysis infrastructure. | $99.95 for the OBD2 adapter to enter the digital ecosystem, no information on business service pricing | $24 million in 3 rounds from 13 investors |
| .. | .. | .. | .. | .. | .. |
| **zubie.com** Sullivans Island (USA), founded 05/2012 | We make driving safer, easier & cheaper. Connects your vehicle to the Internet for real-time information on your smartphone. | Find driving skills, behavioral warnings, rankings, maintenance tips, engine diagnostics, battery warnings, emergency services, tracking, motion recording, perks and gas stations. | Solution for insurance companies and vehicle providers, fleet tracking, GPS tracking, vehicle condition monitoring, driver behavior. | Different price models, of once $99 for OBD2 adapters and $10 per month. Business use from $17.95 per month. | $25.87 million in 5 rounds from 8 investors |

Chapter 3 is based on the publication in the *BISE* (Business & Information Systems Engineering) journal, in which the term *Quantified Vehicles* is introduced and defined for the first time as "the behavioral patterns of self-tracking [that] can be transferred to vehicles, which capture sensory data about themselves and their environment, thus becoming *'Quantified Vehicles'*" (Stocker et al., 2017a). Besides, it shows which trends exist internationally and in Europe, their relevance for research, and the existing research gap and the relevance of *Quantified Vehicle* research for the *BISE* journal community. The *BISE* journal publishes scientific research on the effective and efficient design and utilization of information systems and is, therefore, a relevant medium to publish scientific work in the field of business informatics. The relevance of *Quantified Vehicle* research for BISE is expressed by proposing a research framework and a discussion of its elements.

The author of this dissertation mainly contributed to Chapter 3 (which is based on this *BISE* publication) with an investigation of the business models of the quantified vehicle startups and the development of the data-driven service ecosystem model (a refinement of the ecosystem model from Chapter 2, see Figure 7). Moreover, the dissertation author was also involved in the creation of the other contents, for example, through brainstorming, discussions and revisions, thus contributed to the introduction and definition of *Quantified Vehicles*.

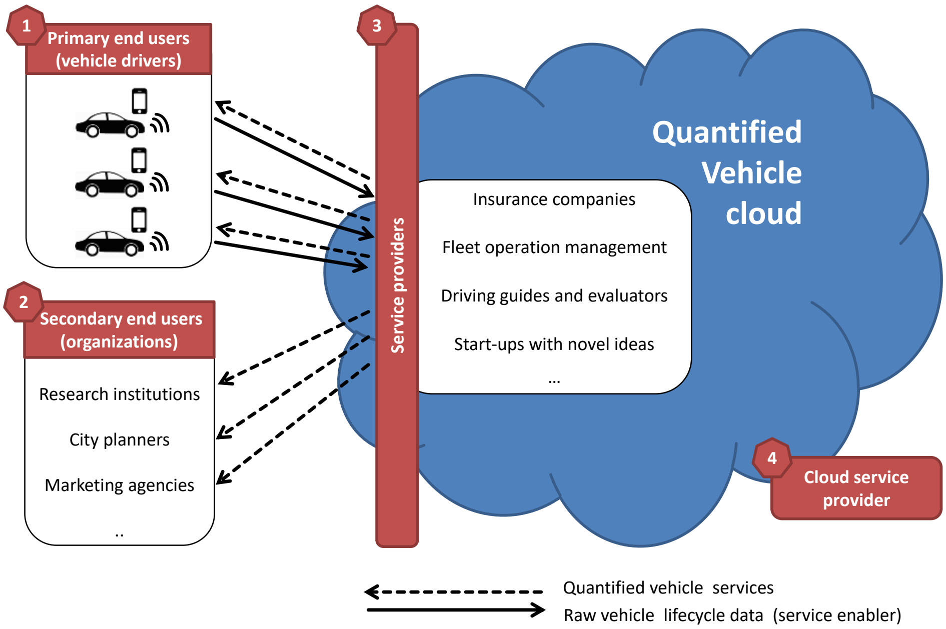


**Figure 7** **A preliminary model of the *Data-driven Service Ecosystem.* (Source: Stocker et al., 2017a)**

#### 1.5.1.2 Analysis of the Market: Services, Start-ups, OEMs, Business Models, and Trends

After the introduction to *Quantified Vehicles*, in which a number of start-ups were already named and analyzed in terms of business models in two publications, the next step was to conduct an even more intensive market analysis, in which the positions of vehicle manufacturers were also examined and compared. This was published in two publications in the year 2017, which are related to each other. Chapters 4 and 5 are based on these publications.

In detail, Chapter 4 provides insights into the *Quantified Vehicle* phenomenon and explores the approaches of the two major stakeholder groups, car manufacturers and tech start-ups, on their journey to develop novel digital services and sustainable business models. The chapter examines stakeholders who have an interest in the data, lists trending start-ups and investigates their visions, goals, and business models, shows developments in the behavior of German vehicle manufacturers, and attempts to derive future trends from this.

Chapter 5 provides a short overview of the digitalization phenomenon in general, the impact of digitalization in the automotive domain through quantified vehicle start-ups and new business models, as well as a brief investigation of the position of vehicle manufacturers and their digital service strategies – all of them concluded in a comparison of value creation for business model elements.

As the two chapters / publications are related to each other, also the contribution of the dissertation author is related to both. Thereby, the author of this dissertation conducted the literature review (with second author of the two publications the chapters are based on) and contributed as the main author within discussions and revisions to the overview on stakeholders with interest in the services, start-up examples and their value proposition, the two approaches to collect data (OBD-2 vs. smartphone), and the current position of vehicle manufacturers.

#### 1.5.1.3 Definition of a Research Agenda for the Information Systems Community

Once an initial overview of the market was available, the aim was to create a scientifically sound overview of how relevant the topic is for a target community (in this case the IS community) and which research questions still need to be answered by it. This was successfully accomplished with a publication at the prestigious *ECIS* conference presented in Chapter 6.

In detail, the chapter introduces *Vehicle Information Systems* (Vehicle IS) as a new class of *Information Systems* (IS). Furthermore, this chapter investigates existing literature on *Vehicle IS* published by the academic IS community, provides a definition of the term '*Vehicle Information System*' and gives an overview of relevant research directions with a set of example research questions, to assist the academic IS community to advance the state-of-the-art in designing *Vehicle IS*. The article finally proposes a research agenda for vehicle information systems with a set of research questions targeted as IS researchers.

Thereby, the author of this dissertation conducted the literature review, co-authored the introduction and motivation, was responsible for Subsection "Scope and Examples of Vehicle IS", was also involved in defining the concept of *Vehicle IS*, managed the development of the research agenda and sample research questions which is based on six conducted in-depth interviews, and co-authored the conclusions and future work section. Figure 8 shows the research directions which were used to frame the research questions in this publication.

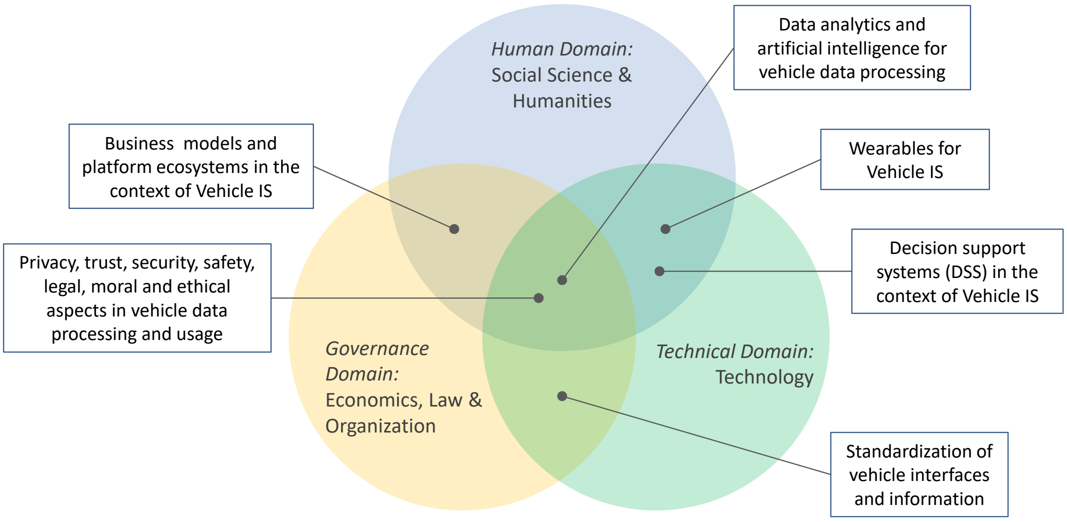


**Figure 8** **Research directions for Vehicle IS. (Source: Kaiser et al., 2018b)**

### 1.5.2 Results for Research Question 2

Second, targeting Research Question 2, *how can vehicle data-driven services be engineered in an efficient and effective way*, structured into the corresponding subtopics of the research process (c.f. Figure 5), chapters 7-8 answer it with an "*Analysis and Definition of the VDVC*", while chapters 9-11 answer it by presenting *"Concepts along the VDVC",* and while chapters 12-13 answer it by presenting *"Prototypical Implementations along the VDVC"*. The vehicle data value chain on the one hand shows the structure of service development, and on the other hand makes services comparable. Examples are a concept for data sharing using an *Open Vehicle Data Platform*, and the description a *Vehicle Data Logger*, an IoT platform, a data-driven vehicle telematics service, a service to detect individual driving behavior and potholes on the street, and a smartphone application. Table 4 provides an overview of all contributions to answer Research Question 2, showing the methods used and the results of the author of the dissertation, and in which medium the paper was published. Hereafter follows a description of the contributions of each chapter in more detail.

#### 1.5.2.1 Analysis and Definition of the VDVC

"One of the greatest challenges in operation management is creating a process that will transform input into services, and product value to internal, as well as external, customer added value" (Gertner, 2013). This process is called value creation and can be described with a descriptive model (Carlucci et al., 2004) of the so-called "value chain". According to Kaplinsky and Morris (2000), the "value chain describes the full range of activities which are required to bring a product or service from conception, through the different phases of production (involving a combination of physical transformation and the input of various producer

**Table 4 Chapters contributing to RQ2, their title, author results, research methods, and medium it was published in.**

| Ch. | Title | Author Result(s) (AR) & Research Method(s) (RM) | Medium and Rank |
|---|---|---|---|
| 7 | The Vehicle Data Value Chain as a Lightweight Model to Describe Digital Vehicle Services. | *AR*: Definition of the Vehicle Data Value Chain (VDVC), Application of the VDVC to describe and compare existing digital vehicle services.<br>*RM*: Literature Review, Inductive Research and Reference Modeling. | Intern. Conference on Web Information Systems and Technologies (WEBIST)<br>WEB IST CORE`18: **C** |
| 8 | Digital Services Based on Vehicle Usage Data: The Underlying Vehicle Data Value Chain | *AR*: Refinement and definition of the Vehicle Data Value Chain (VDVC) using eight characteristics per step. In addition, the VDVC is used to classify data-driven services.<br>*RM*: Literature Review, Inductive Research, Reference Modeling, and Case Study. | Series on Lecture Notes in Business Information Processing (LNBIP)<br>LN BIP VHB-JQ3: **C** |
| 9 | Towards a Generic IoT platform for Data-driven Vehicle Services | *AR*: Development of the concept to collect, process, and use vehicle data for data-driven services with the focus on the IoT platform for data-driven vehicle services under my leadership in project SCOTT. Development (telemetry data visualization) and co-development (Vehicle Data Logger) of certain artifacts.<br>*RM*: Literature Review, and Prototyping. | IEEE Intern. Conference on Vehicular Electronics and Safety (ICVES)<br>IC VES |
| 10 | Towards a Privacy-preserving Way of Vehicle Data Sharing – a Case for Blockchain Technology? | *AR*: Co-development of the concept for a blockchain-based privacy-preserving way of vehicle data sharing, and the concept of an Open Vehicle Data Platform (OVDP). Development of an e3value model about actors and value flows of a data-driven service ecosystem.<br>*RM*: Literature Review, Inductive Research, and Conceptual Modeling. | Part of book series Lecture Notes in Mobility (LNMOB)<br>LN MOB |
| 11 | A Lightweight Framework for Multi-device Integration and Multi-sensor Fusion to Explore Driver Distraction. | *AR*: Analysis of related work, and planning and implementation of the field study.<br>*RM*: Literature Review, and Empirical Field Study. | Intern. Conference on Advanced Information Systems Engineering (CAiSE)<br>CAi SE VHB-JQ3: **C**<br>CORE`18: **A** |
| 12 | Use of Automotive Big Data for the Development of Two New Applications | *AR*: Administration of the description of developed prototypical applications. Thereby, the individual steps of the Vehicle Data Value Chain are run through to develop two applications for users.<br>*RM*: Prototyping. | BITKOM position paper “Practical use cases of artificial intelligence & big data in industry”<br>BIT KOM |
| 13 | A Vehicle Telematics Service for Driving Style Detection: Implementation and Privacy Challenges | *AR*: Preparatory work (definition of privacy levels (c.f. 10.3.2), and deduction of the “willingness to share data” model) and supervision of empirical work on privacy related content. In addition, co-development of the proof-of-concept implementation.<br>*RM*: Literature Review, Inductive Research, Reference Modeling, Empirical Field Study, and Prototyping. | Intern. Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS)<br>VEH ITS |

services), delivery to final consumers, and final disposal after use”. However, those traditional value chains of incumbent companies are challenged by digital transformation.

To better understand and capture the ongoing digital transformation in the automotive domain, the *Vehicle Data Value Chain* (VDVC) is introduced as a lightweight model to describe and examine *Data-driven Services.* The first version of the *Vehicle Data Value Chain* was developed back in 2017, and was once submitted to the TRA (Transport Research Arena) 2018 conference. An abstract was then accepted by the ETC (European Transport Conference) 2018, where it was also presented. The abstract (considered as VDVC - Iteration 1, c.f. Figure 5, page 31) was published publicly (Kaiser et al., 2018a), but as an abstract it was not specifically included in the list of publications of this dissertation. In the course of the ETC submission, a full paper could also be submitted for review for the Special Issue of Transport Research Procedia, but unfortunately this Journal ultimately selected a different topic for the Special Issue. After a revision, the full paper was accepted for *WEBIST 2019* (considered as VDVC - Iteration 2), and also got on the short list of candidates to win the *WEBIST 2019* best student paper award. In detail, the *WEBIST 2019* paper presented in Chapter 7 introduces the VDVC as a lightweight model, grounded on big data, to describe and examine data-driven services, to better understand and capture the ongoing digital transformation in the automotive domain. Additionally, the VDVC is applied to data-driven services from the market to identify commonalities and differences.

Thereby, the author of this dissertation – as the corresponding author – contributed to writing and incorporated the experience from the development and analysis of *Data-driven Services* (in research projects) that the value creation of services is subject to a certain pattern, comparable to the *Big Data Value Chain*. Thus, Subsections 7.2.2 and 7.2.3, where the VDVC is proposed and described in detail (c.f. Figure 9) as a lightweight model and applied to existing *Data-driven Services*, originated under the responsibility of the author of this dissertation, based on discussions with the second author and the last author.

Since the author of this dissertation was invited to expand on the *WEBIST* paper (Kaiser et al., 2019a) for submission to the *LNBIP* series, the VDVC definition was extended accordingly (considered as VDVC - Iteration 3) and now describes the value creation process in more detail using eight characteristics. In particular, each value-creating step (from Generation to Usage) is described with a set of characteristics, e.g. the scope for the step, the input data, the output data generated, typical actors involved, typical architectures, relevant trends and tools and, finally, the contribution of a particular step to value creation, as shown in Figure 10. Hence, Chapter 8 presents this iterated and extended version of VDVC published in Kaiser et al. (2019a) and uses it as a model for better structuring, describing, and

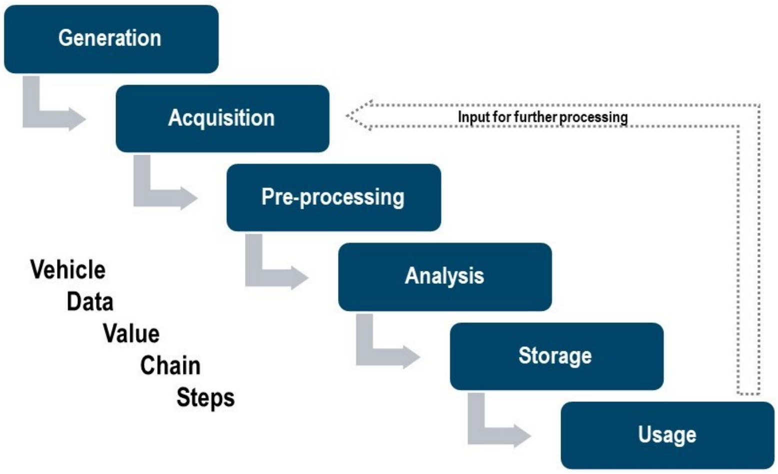


**Figure 9** **The Vehicle Data Value Chain (VDVC) based on Kaiser et al. (2018a; 2019a). (Source: Kaiser et al., 2020b)**

testing digital services based on vehicle usage data. To show the general applicability and usefulness of the VDVC in a practical context and to evaluate it, two cases are shown, where the VDVC is used to classify two implemented data-driven services, (i) an intermodal mobility service, and (ii) a pothole and driving style detection service.

As described above, theoretical insights such as the *Vehicle Data Value Chain* were also derived from own experience in the conception, development and application of *Data-driven Services* in which the author of this dissertation was involved. During the years of the dissertation project, the author had the chance to work on several research projects, to gain experience, for example, in the international research projects *AEGIS*, *SCOTT*, *InSecTT*, *EVOLVE, D-TRAS* and in internal strategic projects (at Virtual Vehicle Research GmbH) *CLOud conNEcted CAR (CloneCAR), Wearables4Drivers,* and *Lightweight Digital Mobility Assistance*. Research projects, which are funded with EU funding and/or national funding, are expected to disseminate their findings, thus there are also several (rather) technical papers that follow now, starting with contributions that present a concept, and concluding with contributions that focus on prototypical implementations, where the transition between the categories is fluid and some publications show both a concept and prototypical implementations.

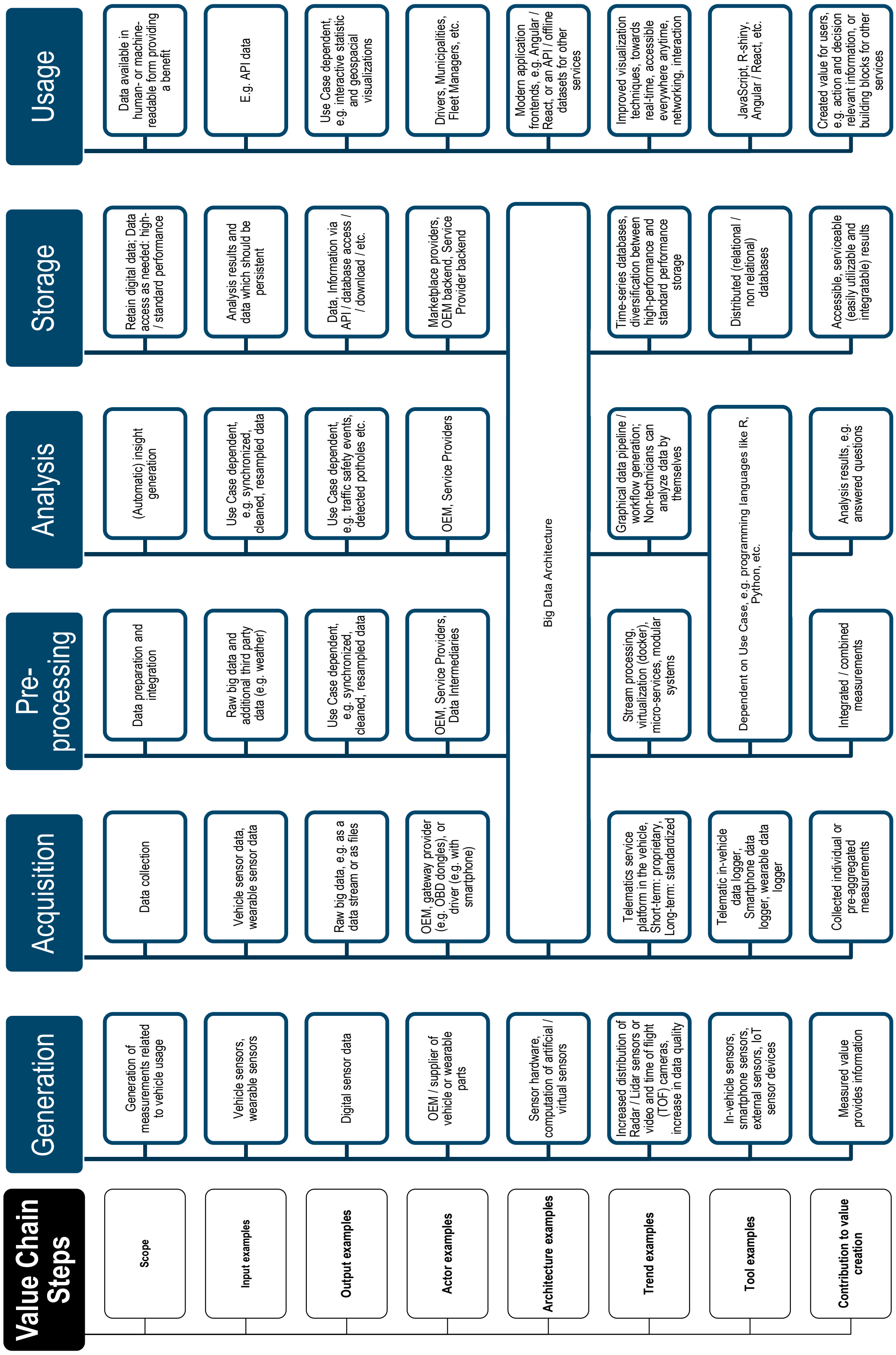


**Figure 10** **The Vehicle Data Value Chain (VDVC) – Iteration 4, derived from Curry (2016) extended with characteristics. (Source: Kaiser et al., 2020b)**

#### 1.5.2.2 Concepts along the VDVC

In the case of project SCOTT, the author of this dissertation led the task (WP11 Task 3) in which a data-driven service was developed, together with colleagues and partners from the company RISE from Sweden (and JIG from Spain, but they did not join publication writing). As part of the dissemination, the concept was published in a peer-reviewed paper. Chapter 9 presents this concept to collect, process, and use vehicle data to enable data-driven services. Challenges like independence, scalability, and flexibility while ensuring e.g. privacy, and accountability must be considered for an IoT platform supporting data provision for services. Thus, this chapter presents a conceptual architecture of a generic IoT platform for enabling such data-driven services and describes how such a platform can be implemented, ranging from the gateway device capturing and transmitting the vehicle data, to a vehicle data-driven service application with added value for the driver.

Thereby, the author of this dissertation was responsible for Subsections 9.2.1 and 9.4.1 on related work for *Quantified Vehicles* and the description of the *Vehicle Data Logger*. Furthermore, as task leader in the SCOTT research project (WP11, Task 3) in which this approach was developed, the concept was also developed under the leadership of the dissertation author based on joint discussions and brainstorming sessions. Furthermore, the list of “potential applications for the end users” in Section 9.3 was added, and the telemetry data visualization on which Figure 11 (Figure 45 in the corresponding chapter) is based, was developd by the author of this dissertation in the internal project *CloneCar* and made available for this publication.

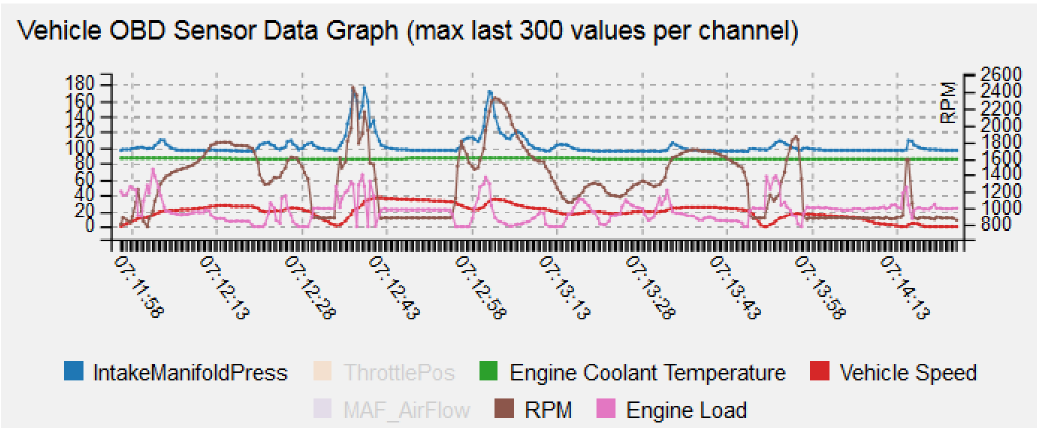


**Figure 11 Telemetry data visualization utilizing OBD-II data. (Source: Papatheocharous et al., 2018)**

In the next publication, based on an internal company meeting, a concept idea on the topic of vehicle data sharing was created together with the second and fifth authors of the publication, which Chapter 10 is based on. Vehicle data, a valuable source for digital services, allows analysis that touches personal privacy, thus, this chapter deals with the question of what would be a privacy-preserving way of vehicle data exploitation. To answer this question the blockchain-based *Open Vehicle Data Platform* concept is presented, followed by a discussion on unsolved technical and non-technical issues.

Thereby, the author of this dissertation stepped into the role as the first and corresponding author when M. Steger left the company and wrote the original draft in the concept section (10.3) based on the joint considerations. Figure 12 illustrates, how a broker entity using Blockchain technology would manage the data exchange with *Data-driven Services*, if the driver/vehicle owner agrees.

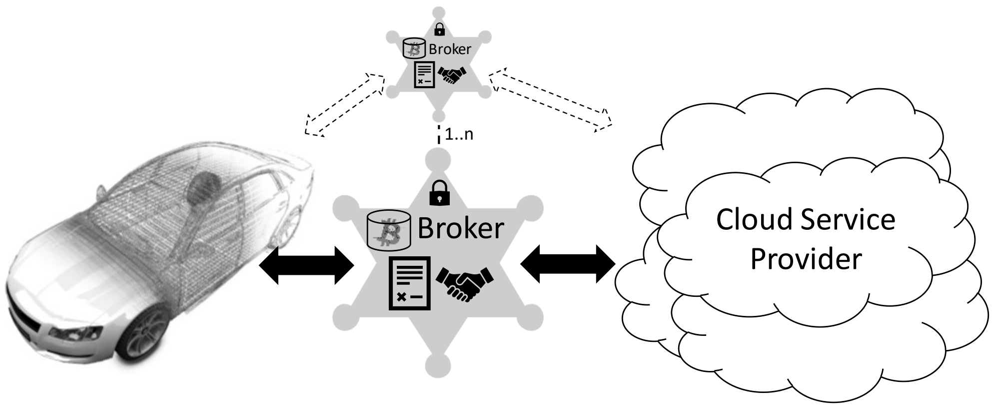


**Figure 12** **Data exchange between origin (vehicle) and target (service providers) is managed by a broker using Blockchain technology for smart contracts. (Source: Kaiser et al., 2019b)**

In addition, the author of this dissertation designed the e3value model that shows the actors and value flows of a (vehicle data sharing) data-driven service ecosystem, illustrated in Figure 13 (Figure 48 in the corresponding chapter), and co-authered all other sections of the publication.

Furthermore, in the internal project *Wearables4Drivers*, the aim was to use sensor fusion of four consumer devices, to analyze driver behavior. A *Vehicle Data Logger* provided vehicle data (e.g. vehicle speed), a smartwatch provided GPS position, heart rate data, and acceleration data (to analyze hand movements), a *MicroSoft HoloLens* provided head movement data, and a smartphone was used as central hub for data exchange. As a result, the heart rate from the smartwatch and the current vehicle position were shown in the *HoloLens* while driving. Based on this setup, the author of this dissertation conducted an empirical field study with ten volunteers, which were asked to perform sequences of specific motions which simulate a variety of distraction tasks. The concept, the empirical field study and the evaluation have been published in a generic way as *a lightweight technical framework for the real-time*

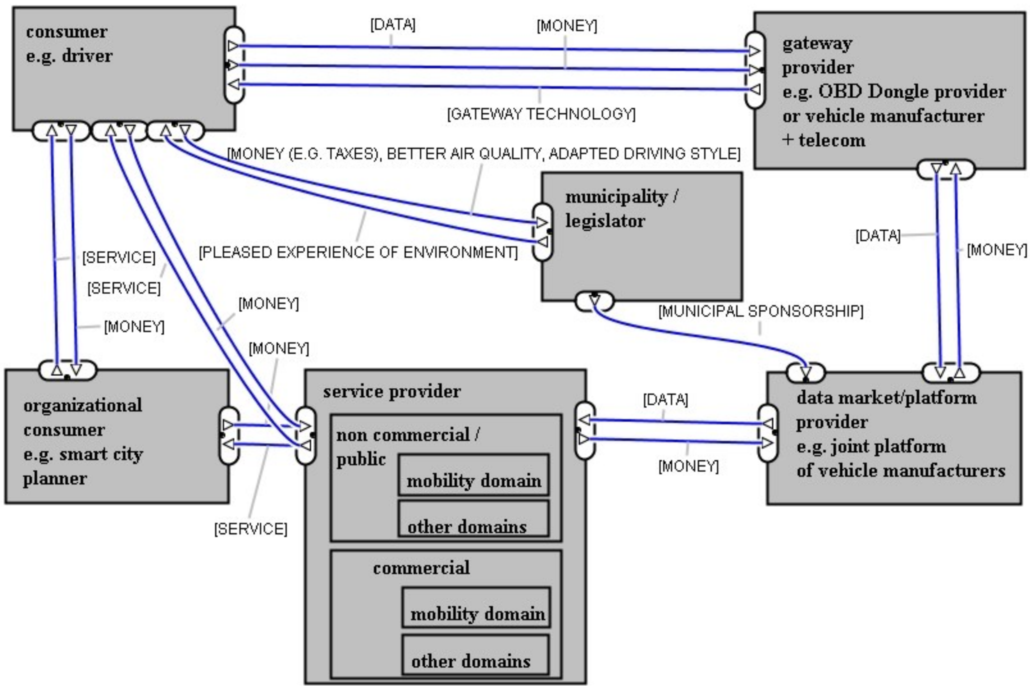


**Figure 13** **Actors and value flows (e3value model) of a data-driven service ecosystem. (Source: Kaiser et al., 2019b)**

*fusion of vehicle data and other contextually relevant data*, as will be shown in Chapter 11. In the publication, the proposed framework is used to assess the driver's status with appropriate measurement equipment, to detect driver distraction and driver inattention, which are both major challenges in road traffic and major causes of accidents.

#### 1.5.2.3 Prototypical Implementations along the VDVC

The last two publications answer Research Question 2 by means of prototypical implementations. The the first of the two publications is also based on project work, this time from the AEGIS and EVOLVE projects. Here, for instance, vehicle data (collected with the *Vehicle Data Logger*) was pre-processed and analysed accordingly to detect potholes on the road from the data, and to calculate a risk/safety-driving-score that indicates how safely the driver is driving. With this content, there was an application for a presentation at the annual *BITKOM Big-Data.AI Summit (#BAS19)*, which was accepted, and conducted by the author of the thesis. As a result of the presentation at the *Big-Data.AI Summit 2019*, the author of the thesis was invited to submit a contribution to a BITKOM position paper. In this paper, together with the co-authors, developments from several research projects are described, while partly content from project deliverables was reused, in which the author of the thesis has collaborated. In detail, Chapter 12 shows, on a level targeted at practitioners, how the individual steps of the VDVC are run through to finally provide two applications for end-users,

(i) the detection of individual driving behavior (c.f. Figure 14, for example the trip with ID "Trip_069" is less risky than 56.72% of all trips) and (ii) the detection of potholes on streets. In particular, details such as the correct alignment of the coordinate system, that this is essential for further data analysis and the data pipeline, which shows the sequence and branching of the implementation, are shown.


Driver1
Total riskscore (%): 50.37

| | trip ID | risk score (%) | duration (h) | accevents (n) |
|---|---|---|---|---|
| 0 | 20181119, Trip_069 | 56.72 | 0.44 | 12.0 |
| 1 | 20181119, Trip_066 | 60.40 | 0.34 | 7.0 |
| 2 | 20181119, Trip_065 | 48.55 | 0.40 | 16.0 |
| 3 | 20181119, Trip_064 | 93.57 | 0.45 | 5.0 |
| 4 | 20181119, Trip_063 | 37.06 | 0.41 | 21.0 |
| 5 | 20181119, Trip_062 | 49.54 | 0.13 | 4.0 |
| 6 | 20181119, Trip_061 | 43.40 | 0.16 | 7.0 |
| 7 | 20181119, Trip_059 | 23.81 | 0.34 | 26.0 |
| 8 | 20181119, Trip_058 | 15.32 | 0.11 | 9.0 |
| 9 | 20181119, Trip_056 | 45.94 | 0.14 | 6.0 |
| 10 | 20181119, Trip_055 | 89.25 | 0.66 | 3.0 |
| 11 | 20181119, Trip_054 | 46.04 | 0.19 | 6.0 |
| 12 | 20181119, Trip_053 | 31.62 | 0.26 | 19.0 |
| 13 | 20181119, Trip_052 | 89.48 | 0.48 | 6.0 |
| 14 | 20181119, Trip_051 | 34.35 | 0.28 | 16.0 |

**Figure 14 Map visualization of a trip (left) and driver trip details (right). (Source: Kaiser et al., 2020a)**

The (chronologically) last publication of prototypical implementations again deals with the results of the SCOTT project. In contrast to chapter 9, in which the concept was presented, chapter 13 rather describes results of the project work under the leadership of the author of the dissertation. In order to bring the topic of privacy into focus once again (because here too there were repeated considerations and ideas in the course of the dissertation project), the results of the master's thesis by Tom Szilágyi, whose content was supervised by the author of the dissertation, were also included.

In detail, Chapter 13 shows how individual components of the VDVC have been implemented for a proof-of-concept resulting in a data-driven service for driving style detection and what has to be considered in terms of privacy. This chapter also presents the results of an empirical study on privacy, e.g. which data, for which services, and under which circumstances the survey participants would log and provide the data. Thereby, the contents of the section on data privacy (13.1) are based on joint considerations by the author of the dissertation and the last author of the publication. These considerations, such as the "preliminary model of the willingness to share data" and a preliminary version of the "privacy levels for vehicle data sharing" (c.f. Figure 15 or Figure 60 in the corresponding chapter), served as input for Tom Szilágyi's master's thesis (Szilágyi, 2019), in which he investigated the level system for data sharing in an empirical study under the guidance of the author of the dissertation. Furthermore, as leader of the corresponding task in the SCOTT research project (WP11, Task 3), through regular meetings with the partners, the author of the dissertation was involved in the conception of all steps described in the service implementation section

(13.2). Thereby, he co-developed / supported the iteration of the *Vehicle Data Logger* (13.2.1) used in project SCOTT together with the second and third authors, based on the initial design and development of the former colleague Benjamin Fischer.

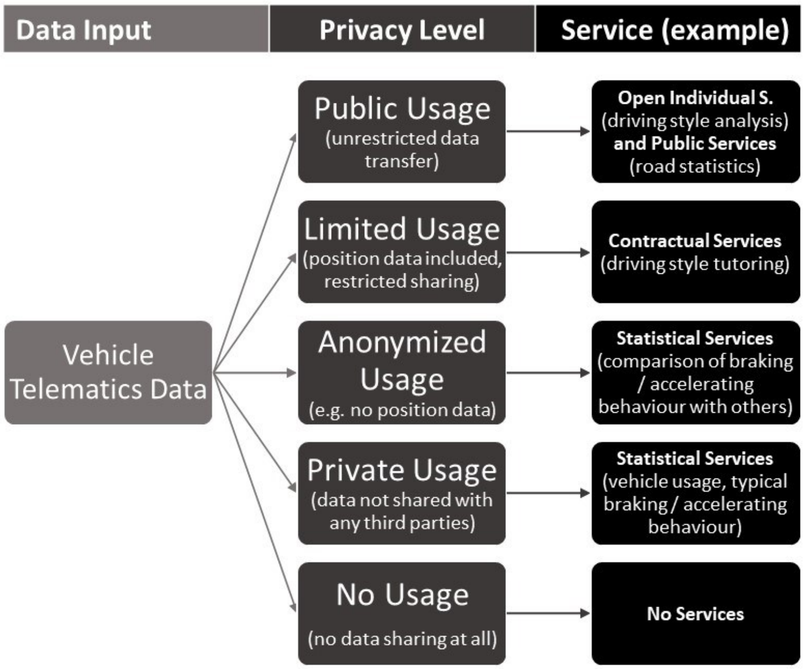


**Figure 15 Privacy levels for vehicle data sharing. (Source: Szilágyi, 2019; Kaiser et al., 2020c)**

### 1.5.3 Results for Research Question 3

Third, targeting Research Question 3, *what are important actors and relationships for service delivery in Vehicle Data Service Ecosystems*, Chapters 14-15 contribute to answering it. Thereby, conceptual models of *Data-driven Service Ecosystems* in the automotive domain are presented, which were developed based on expert interviews. Table 5 provides an overview of all contributions to answer Research Question 3, showing the methods used and the results of the author of the dissertation, and in which medium the two papers were published. After the table, the text goes into more detail about the contributions of both chapters.

#### 1.5.3.1 Analysis of Data-driven Service Ecosystems

In the last two publications that are part of this cumulative dissertation, the data-driven service ecosystem is analyzed to better understand the value creation process. Both publications are based on interviews with experts from the automotive domain.

Chapter 14 (based on the publication in *25th Americas Conference on Information Systems – AMCIS 2019*) sheds light on the actors involved in the creation of data-driven services in the automotive domain. Thus, it examines the ecosystem transformed by the emergence of data-driven services. To improve the understanding, it analyses both the actors and the

**Table 5 Chapters contributing to RQ3, their title, author results, research methods, and medium it was published in.**

| Ch. | Title | Author Result(s) (AR) & Research Method(s) (RM) | Medium and Rank |
|---|---|---|---|
| 14 | Understanding Data-driven Service Ecosystems in the Automotive Domain. | *AR*: Development of a conceptual, high-level *Data-driven Service Ecosystem* model and the related data ecosystem in the Automotive Domain, based on expert interviews and a literature review conducted.<br>*RM*: Literature Review, Empirical Expert Interview Study, Qualitative-empirical Cross-sectional Analysis, and Conceptual Modeling. | Americas Conference on Information Systems (AMCIS)<br>VHB-JQ3: **D**<br>CORE`18: **A** |
| 15 | Conceptualizing Value Creation in Data-driven Services: The Case of Vehicle Data | *AR*: Development of a conceptual model for value creation in vehicle *Data-driven Services* – a model that connects key actors with value-adding data sharing processes – based on expert interviews and a literature review conducted.<br>*RM*: Literature Review, Empirical Expert Interview Study, Qualitative-empirical Cross-sectional Analysis, and Conceptual Modeling. | International Journal of Information Management (IJIM)<br>VHB-JQ3: **C**<br>ABDC-JQ:**A*** |

role of vehicle sensor data from an ecosystem perspective. Based on literature analysis, results from eleven expert interviews are used for the examination to gain relevant actors in the ecosystem as well as their relationships, data flows, and services. Results thus provide a fundamental understanding of data-driven service ecosystems in the automotive domain and form the basis for future IS research on (big) data flows and analytics within such ecosystems.

Thereby, as corresponding author, the author of this thesis conducted the literature review to derive a preliminary model of the *Data-driven Service Ecosystem*, developed the interview guide, contacted interview participants, conducted four interviews in person, and seven interviews online. Furthermore, the interviews and the related work were analyzed and the original draft of (i) the high-level model of the automotive *Data-driven Service Ecosystems*, (ii) the detailed model of the data ecosystem as shown in Figure 16 below (c.f. Figure 66 in the corresponding chapter), and (iii) the overview of the providers and consumers of selected data-driven services were developed by the author of the thesis.

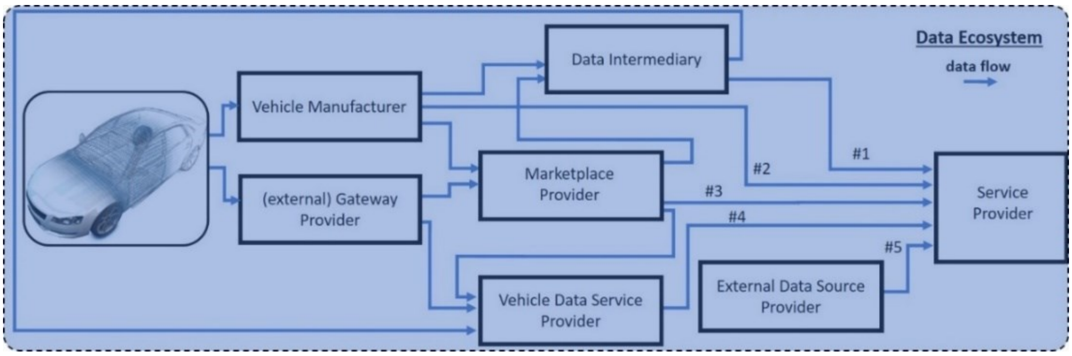


**Figure 16 Detailed model of the data ecosystem. (Source: Kaiser et al., 2019b)**

#### 1.5.3.2 Conceptual Model for Value Creation in Data-driven Services

Furthermore, based on the AMCIS publication, an article for *International Journal of Information Management* (*IJIM*) was formed in seven iterations. It states, that the digitalization of the automotive industry brings fundamental changes to how value is created and by whom. As part of this transformation, the creation of data-driven services generates new value streams, thus leading to the emergence of new actors and ultimately, new market configurations. Eventually, vehicle data paves the way for new types of data-driven services. Chapter 15, based on the *IJIM* article, provides a model suitable to support academics and practitioners in the identification of the actors that will play a key role in data-driven service generation and resources involved in value creation processes. Based on interviews with eleven prominent experts of the central European automotive industry, a conceptual model that connects these key actors with value-adding data sharing processes is developed. To validate the model, it is applied to a real-life case: the design of a data-driven service for road surface quality detection. Furthermore, the model's implications to both theory and practice are discussed.

Thereby, based on eleven domain-expert interviews, the author of this thesis has continued to work with the co-authors on this journal publication, which presents a conceptual model for value creation in *Vehicle Data-driven Services*. As the corresponding author, he wrote, reviewed and edited the original draft with the help of the co-authors. Specifically, the author of this thesis conducted the literature search in established scientific electronic databases with the second author, developed the interview guide, and conducted four interviews in person and seven interviews online. He critically compared and analysed the interviews and related work, and developed the original draft of the conceptual model in discussions with the second author. Based on discussions with the co-authors, the author of this thesis iterated the results and improved the model and the article accordingly based on several reviewer feedbacks (reviews from four top-ranked conferences and journals). In Figure 17 (Figure 72 in the corresponding chapter) below the conceptual model is illustrated. In addition, the author of this thesis participated in the development and writing of the model validation and in the writing of the discussion and conclusion sections.

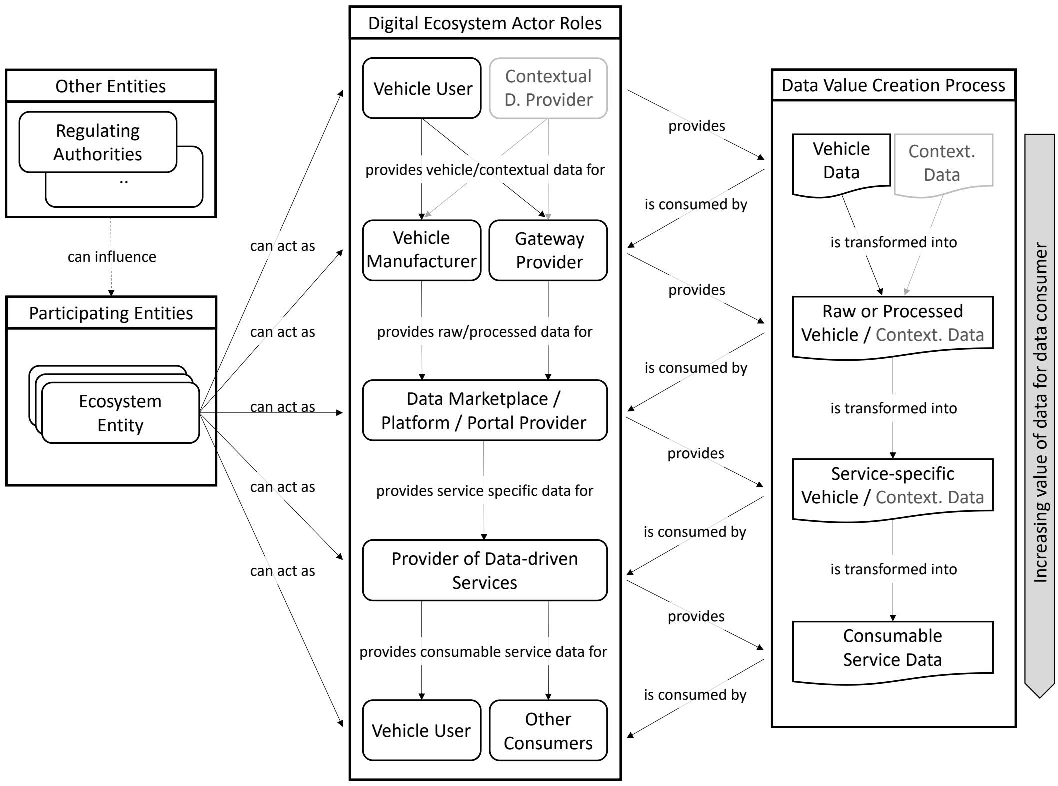


**Figure 17 A conceptual model for value creation in vehicle data-driven services. (Source: Kaiser et al., 2021)**

## 1.6 Summary

To summarize, a synthesized overview of the contents of the Synopsis is given in the following.

The ongoing digitalization of the incumbent industries opens up new opportunities for value creation. Thereby, data gets increasingly important as a source, to develop and provide data-driven services. In the introduction, it has been shown that this can be applied to the automotive industry (e.g. exploiting vehicle usage data) and actually is a currently urged topic in the industry and is slowly reaching the scientific community. The introduction and analysis of the term "*Quantified Vehicles*" in 2017 showed, that it is quite new and that there is almost no research on it, as it was in the emerging phase when the dissertation was started, and even in 2020 the harmonization processes are not yet complete. This allows the conclusion, that findings within this dissertation form a new and relevant contribution to the ongoing research, as not everything is already set and defined yet.

In the automotive industry, data-driven services based on vehicle usage data offer the opportunity for (i) value creation for the automotive industry (e.g. predictive maintenance, product analysis, and improvement), and (ii) other consumers (e.g. pothole analysis for road maintenance, traffic analysis for traffic planners), and (iii) added value for the driver (e.g. tutoring services, information on driving style, comparison possibilities). Thereby the value creation opportunity represents an important factor in maintaining innovation in the automotive industry, to find new ways of generating value for other consumers, and to increase / maintain customer loyalty of vehicle drivers.

With data analysis, new players from the ICT sector are also pushing into the value chains, as the automotive industry traditionally tends to outsource software development. The cooperation of the different companies from the automotive and ICT sector for the development of data-driven services based on vehicle data is presented as an ecosystem, which is just establishing at this time of formation.

Thus, the aim of this dissertation, entitled '*Quantified Vehicles: Data, Services, Ecosystems'*, is to increase understanding on, define and explore *Quantified Vehicles* that generate data about themselves and their environment, and its three key constituents: i) (vehicle) *Data*, ii) the *Data-driven Services* that can be generated from them, and iii) the *Data-driven Service Ecosystem* that develops data-driven services. The lack of research on this topic and the importance of the automotive industry for the European and global economy underlines the relevance of this dissertation.

To investigate the recent and exciting topic of *Quantified Vehicles*, the data generated by them, the services they enable, and the actors and relationships in the ecosystem of value creation, three research questions were set up. The research questions were then worked through in the presented research process by fourteen individual contributions, all of which were also published to and accepted by the scientific community. The results, analyses, and prototypes provide answers and solutions to the three very relevant and larger research questions that were motivated at the beginning. The individual contributions were summarized above, and are presented in full detail in the upcoming chapters 2-15. However, to provide not only a view on scientific publications from a research question viewpoint (c.f. Figure 6), a topic-related overview including all 14 publications is provided in Figure 18 below. In retrospect, the idea for this dissertation project developed in the two initial publications *E&I* (Chapter 2) and *BISE* (Chapter 3), therefore the author of the dissertation was not the main author (lighter blue color) of these two publications. After that, however, the author of the dissertation took over the lead and is the main author (darker blue color) of 10 of the following 12 publications.

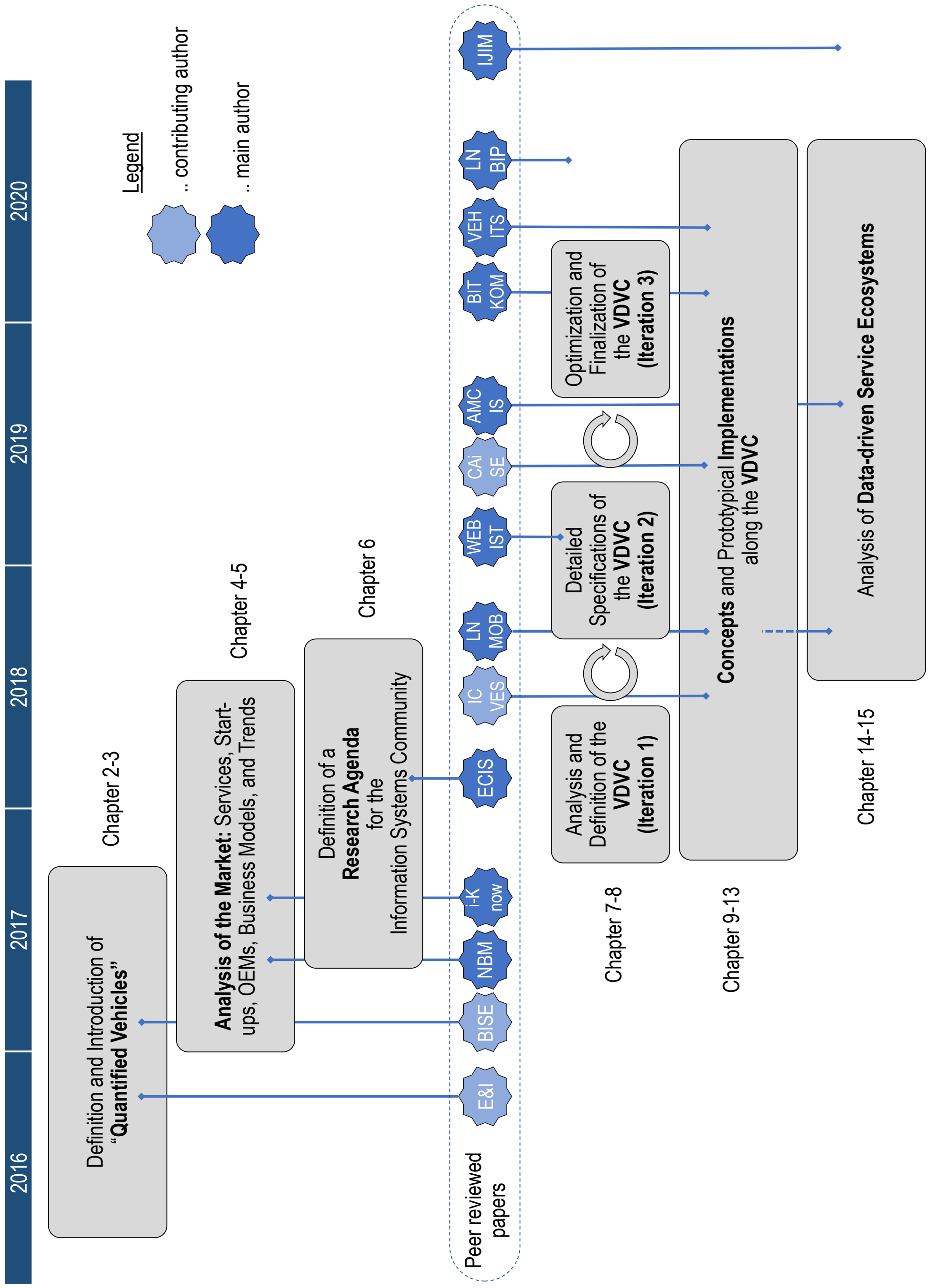


**Figure 18** **A topic-related overview of the results – fourteen peer reviewed papers.**

In brief, first, the topic of *Quantified Vehicles* was defined and introduced, and then the state-of-the-art of data-driven services based on vehicle usage data was investigated in the market with several market analyses. From this, a research agenda was successively derived, which confirmed, among other things, that value creation and value generation are insufficiently investigated. Subsequently, the VDVC was defined and iteratively improved, for example with findings from proof-of-concept implementations, literature, and expert interviews. In the final part, primarily the cooperation of actors in the ecosystem was examined. Hence, two tangible key results of this dissertation are i) the design of a value chain for data-driven services in the automotive domain, and ii) the design of an data-driven service ecosystem model.

The 14 publications consist of three journal publications (*E&I*, *BISE*, and *IJIM*), two book-series contributions (*LNMOB*, and *LNBIP*), eight conference contributions (*NBM*, *i-Know*, *ECIS*, *ICVES*, *WEBIST*, *CAiSE*, *AMCIS*, and *VEHITS*), and a contribution in a *BITKOM* position paper (*BITKOM*). Thereby, the journal publications *IJIM* (Impact Factor 8.21), *BISE* (Impact Factor 5.83), the *ECIS* conference publication (VHB-JQ3: B), and the publication in the book series *LNBIP* (VHB-JQ3: C) are particularly outstanding, due to their high impact and prestige in the IS community, as shown in Table 6.

**Table 6 Outstanding publications where the author of this dissertation contributed.**

| Medium | Statistics |
|---|---|
| IJIM | Impact Factor: **8.21** (Feb. 2021)<br>**2.88** in SJR 2019, **Q1** (best quartile for MIS)[7]<br>ranked **A*** in the ABDC journal list ranking<br>**C** in VHB -JOURQUAL3 (ranking from 2015) |
| BISE | Impact Factor: **5.83** (Feb. 2021)<br>**1.31** in SJR 2019, **Q1** (best quartile for IS)<br>**A** in the ABDC journal list ranking<br>**B** in VHB -JOURQUAL3 |
| ECIS | **A** in CORE`18<br>**B** in VHB -JOURQUAL3 |
| LNBIP | **C** in VHB -JOURQUAL3<br>**0.26** in SJR 2019, **Q3** (third quartile for IS/MIS) |

The next few years will show how much data-driven services will become an important factor in the automotive industry. In any case, the work of this dissertation on "*Quantified Vehicles: Data, Services, Ecosystems*" has contributed to a better understanding of the value creation steps of VDVC and has shown, for example, how the ecosystem currently looks and thus also has a share in future changes. The author hopes that the results have created a fruitful basis on which other researchers or even industry players can build further. Finally, Chapter 16 provides a summary from the author's personal perspective.

7 https://www.scimagojr.com/journalsearch.php?q=15631&tip=sid

# 2. Quantified Car: Potentials, Business Models and Digital Ecosystems [8]


**Summary and Author Contribution**

**Definition and Introduction of "*Quantified Vehicles*"**
**(Paper 1/2)**

In this journal article, the term Quantified Car is introduced and derived from the Quantified Self trend. Modern vehicles enable such Quantified Cars – a continuous collection of vehicle lifecycle data which facilitates the generation of innovative products, services, and business models. The article describes how start-ups from the USA use vehicle data for data-driven services, and that this will also become relevant for the European region. The ICT players will create a new digital ecosystem within the automotive domain. In this digital ecosystem there are several important actors: cloud providers, service providers, primary and secondary end users. With three case studies – Automatic, Mojio and Dash – use cases and the potential of start-ups are shown, and many similarities are identified. The potential is discussed based on investment figures, which reach $20 million in some cases. To sum up, this article provides an introduction into the Quantified-Car-phenomenon and analyses the business models of three different Quantified-Car-startups.

Here, I focussed on the introduction to and the definition of *Quantified Vehicles* (termed *Quantified Cars* in the paper) as an emerging trend for Europe. In doing so, I tried to shed light on business models of start-ups in this field, to increase understanding of how startups utilize vehicle data for *Data-driven Services*, and how they commercialize the services. In addition, the discussion on the constitution of the *Data-driven Service Ecosystem* is started (termed *Quantified Car Ecosystem* in the paper).


## 2.1 From Quantified Self to Quantified Car

In a networked world, physical objects of daily life collect more and more data about themselves and their environment and slowly transform into "Smart, Connected Products" (Porter and Heppelmann, 2014; Porter and Heppelmann, 2015). Products become a source of data and data scientists evaluate the ever-increasing amount of data collected over the entire life cycle of these intelligent products in order to gain interesting insights into user or product behavior. In the private sector, for example, the Quantified Self Movement uses the tools of modern data-based analysis to gain multi-layered insights into the own organism.

---

8 The content of this chapter was translated into English for integration into this English-language dissertation and is based on

Stocker, A., and **Kaiser, C.** (2016). Quantified car: potentials, business models and digital ecosystems (original german title: Quantified Car: Potenziale, Geschäftsmodelle und Digitale Ökosysteme). *e&i Elektrotechnik und Informationstechnik, 133(7),* 334-340. DOI 10.1007/s00502-016-0429-3.

As a term, Quantified Self refers to the growing willingness of many people to collect data about themselves, their behavior and their environment, whether biological, behavioral or physical (Swan, 2009; Swan, 2013; Swan, 2015). Today, millions of Quantified-Selfers want to gain multilayered insights into personal, health or sports issues by systematically collecting, analyzing and evaluating data. They often use their personal smartphone, which is equipped with a variety of sensors and tools and accompanies the Quantified-Selfers on a daily basis. The Quantified-Self-Trend, which was already awarded its own term in 2007, shows its economic relevance primarily through recent acquisitions of mobile smartphone apps in the sports sector. In 2015, for example, Adidas acquired a majority stake in the Austrian company Runtastic for around 220 million EUR (Runtastic, 2015). The ecosystem built by Runtastic now offers consumers a wide range of health and fitness products, services and content. As of April 2016, Runtastic's popular app already has over 80 million registered users (Runtastic, 2016). This example shows very clearly that established industry leaders are now investing large sums of money to participate in innovative technology start-ups in the quantified self environment. Analyst Gartner even expects a market volume of 5 billion USD for quantified self devices in 2016 (McIntyre, 2013).

In the lives of many people, vehicles are a big purchase and are associated with high value and interest for their owners. The interest in insights gained by quantifying vehicle lifecycle data, such as diagnosing the condition of objects, analyzing and optimizing driving style, or improving one's own driving safety by integrating environmental data, can be seen as particularly high, especially among vehicle enthusiasts. This is because vehicles also have a high potential as a data source for enabling new services. If the term Quantified Self refers to the recording, analysis and evaluation of data generated in the course of one's own activities, the term Quantified Car, by analogy, covers the systematic collection of vehicle lifecycle data in the usage phase and subsequent intelligent analysis in order to provide benefits for various stakeholders.

## 2.2 Quantified Car as Enabler of Digital Ecosystems

Due to constant changes in the competitive environment caused by the digital transformation, there is a high pressure on companies to adapt. The current debate on digitization focuses primarily on the networking of technical devices used in everyday life (Yoo et al., 2010). Digital ecosystems represent the habitat of digital content by forming a technically separate system that networks hardware, software, content and services with each other (Ammon and Brem, 2013). According to this classification, the much-cited Apple ecosystem, for example, consists of iPod, iPhone, iPad, Mac Desktop, MacBooks and peripherals (hardware), iOS/Mac OS, Office Suite & Core Apps, iTunes & iBooks, tools for media developers (software), music podcasts, audio books, music video, TV series, movies, eBooks, iBooks, textbooks (content), iCloud and iTunes (services). The opportunities created by such digital ecosystems are minimized by risks, not least because providers are constantly confronted with the need to rethink their services as well as the underlying business models in order to take strategic measures, if necessary, that will result in a classic, innovative or disruptive reorientation (Matt et al., 2015).

Digital ecosystems have also emerged around the quantified self-trend: In addition to classic smartphones, a range of smart products such as smart glasses, smart watches, or fitness wristbands are available to enable further facets of continuous data collection and data analysis. For this reason, the marketplaces of Google and Apple offer an unmanageable number of quantified self applications. Some manufacturers of the smart, connected products mentioned above, already offer their own marketplaces where third parties can also provide applications, such as Recon Instruments (Recon Instruments, 2016), a company acquired by Intel.

Compared to Quantified Self, the Quantified Car movement is still in its infancy, and user numbers are correspondingly lower. There is a lot of discussion today, particularly with regard to the development of driver assistance systems, about networking the vehicle with other vehicles (“Car2Car”) as well as with the infrastructure ("Car2X"). However, an innovative use of the life cycle data collected from vehicles with a focus on generating added value for the driver – in analogy to the quantified self-movement – has not yet reached a high level of importance among European vehicle manufacturers. This leaves the question unanswered as to how the Quantified Self phenomenon can be successfully transferred to the vehicle.

Modern vehicles are high-performance computers on four wheels. Equipped with extensive and complex sensor technology, they already collect a vast amount of data about themselves and, with the widespread availability of driver assistance systems, increasingly about their environment. According to the EU project AutoMat (AutoMat, 2016a), which is coordinated by Volkswagen and deals with technical issues relating to the establishment of a marketplace for vehicle life cycle data, a modern vehicle processes up to 4,000 signals per second in the Controller Area Network (CAN) bus system (ISO, 2015) established in vehicles, making it a much more comprehensive and interesting sensor measuring node and data generator than the smartphones and wearables used by quantified self-sufficient data collection. Today, the continuous collection and ongoing analysis of vehicle data serves a central purpose, the guarantee or monitoring of vehicle functions. However, the possibilities for further, intelligent use of this data go far beyond this original purpose, as completely new products, services and even digital ecosystems could be developed.

## 2.3 Quantified Car Ecosystems

### 2.3.1 Actors of a Digital Ecosystem for Quantified Cars

The following actors can be defined in a quantified car ecosystem:

- **Primary end users**, as individual service consumers, are drivers/owners of vehicles, who benefit directly and immediately from innovative products, visualizations, statistics, gamification elements and driving style optimization recommendations, which they have created by providing their data.
- **Secondary end users** are organizations or organizational units such as city planners, insurance companies or fleet operators who indirectly benefit from collected and evaluated vehicle life cycle data by consuming the services provided by service providers.
- **Service providers**, in turn, are organizations that offer products/services to primary and/or secondary end users and thereby generate revenue. These include, for example, fleet management service providers, service providers for driving style-dependent insurance, or service providers for preventive vehicle maintenance services. All services are based on data provided by primary end users.
- **Cloud providers** (platform operators) are responsible for operating the entire infrastructure of a digital ecosystem and making it available to service providers. Primary and secondary end users, as individual or organizational service consumers, are users of the services provided by service providers in this cloud infrastructure.

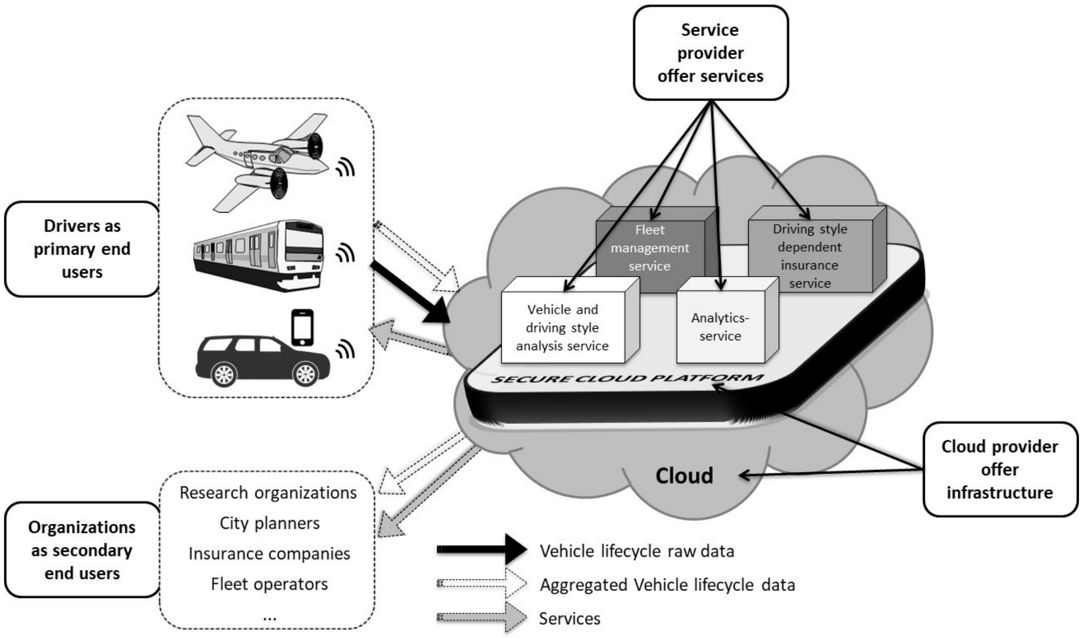


**Figure 19 Actors of a Digital Ecosystem for Quantified Cars. (Stocker and Kaiser, 2016)**

### 2.3.2 Innovation driver USA - quo vadis Europe?

While enormous efforts have been made in the USA for some years now to develop digital ecosystems for quantified car, this development almost completely bypasses Europe. Apart from the already mentioned EU project AutoMat, only isolated and comparatively small activities without significant impact are taking place in Europe. In contrast, a very lively start-up scene has established itself in the USA, financed with enormous venture capital investments of sometimes more than USD 20 million, as Table 7 shows. According to CrunchBase, a portal with information on innovative technology companies and related investor information, companies such as Magna International, Continental ITS, and BMW i Ventures also invest alongside the IT scene.

**Table 7 Quantified Car Start-ups in the USA (Venture capital and investment information from Crunchbase.com, last accessed 25.07.2016). (Stocker and Kaiser, 2016)**

| Company | Value proposition | Possible use | Business applications | Costs | Investors |
|---|---|---|---|---|---|
| **automatic.com** San Francisco (USA), founded 2011 | Connects car with the digital life of the driver. Enables drivers with knowledge about themselves and the vehicle to drive safer and smarter. | Connects the vehicle to many apps, for example for problem diagnosis, consumption optimization, location and emergency services. Includes a web dashboard with comprehensive statistics. | Cloud-based services for insurance companies, fleet operators and vehicle manufacturers, provision of the data analysis infrastructure. | $99.95 for the OBD2 adapter to enter the digital ecosystem, no information on business service pricing | $24 million in 3 rounds from 13 investors |
| **moj.io** Vancouver (CAN), founded 10/2012 | Empowerment of the driver. Connect your devices to your vehicle at any time. | Access to a marketplace for vehicle networking apps (location tracking, vehicle diagnostics, driving analyses, ...), open platform for accessing the vehicle, developer center, API. | Not explicitly mentioned | $149 for the OBD2 adapter to enter the digital ecosystem. Includes 3G/4G connectivity. | $10.3 million in 3 rounds from 6 investors |
| **dash.by** New York (USA), founded 06/2012 | Enabling smart, safe, green and affordable driving. | Connected Car platform, services for vehicle diagnostics, fuel efficiency, emergency services, driver comparisons, gamification and community services | Not explicitly mentioned | Also works with OBD2 adapters from other manufacturers, which can be purchased from the dash webshop | $1.9 million in 3 rounds from 8 investors |
| **vin.li** Dallas (USA), founded 2014 | Your vehicle - your way. Brings countless apps to your vehicle, from security to entertainment to WIFI. | High-speed WIFI, app store for the vehicle, all kinds of smart vehicle apps and services for download in a dedicated app store, developer portal. | Not explicitly mentioned | $199.99 for the OBD2 adapter to get started in the digital ecosystem. Includes 3G/4G connectivity. Additional cost for Internet/Wifi depending on data volume required. | $6.5 million in 2 rounds from 6 investors |
| **zendrive.com** San Francisco (USA), founded 2013 | Safe drivers - safe roads, smartphone based safety on roads in cities, for fleets and individuals | Analysis of smartphone sensor data through machine learning. Driver and fleet analyses. Collision detection, insurance support, etc. Developer kit and API. | Fleet services, insurance, car sharing. | Different pricing models, from free to $4 per driver/month for fleets. Business rates are not listed. | $15 million in 3 rounds from 15 investors |
| **zubie.com** Sullivans Island (USA), founded 05/2012 | We make driving safer, easier & cheaper. Connects your vehicle to the Internet for real-time information on your smartphone. | Find driving skills, behavioral warnings, rankings, maintenance tips, engine diagnostics, battery warnings, emergency services, tracking, motion recording, perks and gas stations. | Solution for insurance companies and vehicle providers, fleet tracking, GPS tracking, vehicle condition monitoring, driver behavior. | Different price models, of once $99 for OBD2 adapters and $10 per month. Business use from $17.95 per month. | $25.87 million in 5 rounds from 8 investors |

These developments show very clearly that the automotive industry also estimates the market value of a digital ecosystem for Quantified Car as enormous, although from today's perspective it is not the innovation driver there. The EU project Automat coordinated by Volkswagen cites three main reasons why the automotive industry in particular, with its Connected efforts, is not yet in a position to establish an open and comprehensive digital ecosystem (AutoMat, 2016b):

- At present, offerings relating to the Connected Vehicle are characterized by brand-specific business approaches that have resulted in proprietary and closed individual solutions. Original Equipment Manufacturers (OEMs) are entering completely new markets as vehicle manufacturers, which do not necessarily correlate with their core business.
- Current Connected Services focus on the individual vehicle buyer, which inevitably leads to data protection concerns. At present, no consideration is being given to how anonymous vehicle lifecycle data could be used in other contexts that do not concern the individual driver.
- The associated risk of cooperation between competing OEMs with regard to a common, standardized provision of vehicle lifecycle data in a digital ecosystem is a major hurdle why such a system has not yet been established.

## 2.4 Case Studies for Quantified Car Ecosystems

In the following, the business models of the three quantified car case studies Automatic, Mojio and Dash are analyzed according to the methodology of Stähler (2002), where business models are roughly differentiated according to the three aspects value proposition, architecture of value creation and revenue model.

### 2.4.1 Quantified Car Case Study 1: Automatic

Value proposition: According to the slogan "Connect your car to the rest of your digital life", the San Francisco-based company Automatic Labs (Automatic, 2017) offers applications for end customers and business users. In order to be able to use them, a special adapter must first be operated on the standard diagnostic interface (OBD) of a vehicle, which provides vehicle data via a smartphone paired via Bluetooth to various apps as a database, which can be obtained via a separate marketplace, the Automatic Gallery.

Architecture of value creation: Automatic has been offering a range of services for private users since 2011. These include, for example, extensive statistics on trips on the smartphone and in a browser dashboard, functions for diagnosing engine and ECU problems, feedback on the respective driving style, functions for finding a parked vehicle, collision detection with emergency services, and the possibility of linking vehicle functions with other digital services from the web via Automatic using IFTTT (If This, Then That) (IFTTT, 2016). Automatic Labs also offers services for business customers. These include the operation of an automotive cloud, cloud-based vehicle insurance, services for the intelligent maintenance of vehicles and for increased customer loyalty in after-sales, fleet management and data analysis. Specific solutions for OEMs are in planning.

Revenue model: The Automatic OBD-II Adapter is available in the USA for USD 99.95 and is required for entry into the Automatic quantified car ecosystem (lock-in effect). All services are limited to the USA. Prices for business solutions are not actively communicated on the Automatic Labs website.

### 2.4.2 Quantified Car Case Study 2: Mojio

Value proposition: Similar to Automatic, Mojio (Mojio, 2016) also wants to "empower" the driver by allowing him to connect to his vehicle at any time using a single device, the smartphone. With Mojio, too, the interface between smartphone and vehicle consists of an adapter operated on the vehicle's OBD-II port.

Architecture of added value: Mojio also offers a wide range of apps and services for drivers. These range from location tracking, vehicle diagnostics, driving analytics and driving style analysis to the use of mobile apps provided by third parties in the Mojio digital ecosystem. With the Developer Center, Mojio also offers an open Connected Car platform with Application Programming Interfaces (APIs) and Software Development Kits (SDKs) to enable third parties to easily develop apps.

Revenue model: Mojio operates an online shop where the OBD-II adapter including a built-in sim card for the AT&T mobile network can be purchased for USD 149 in the USA. This means that Mojio has direct connectivity of the vehicle to the Internet – and not, as with Automatic Labs, only via a paired smartphone.

### 2.4.3 Quantified Car Case Study 3: Dash

Value proposition: In line with the slogan "Smarter.Driving.Everyday.", the app from the New York company dash (Dash, 2017), founded in 2012, uses an OBD-II adapter to once again collect vehicle lifecycle data to inform drivers in real time about interesting events. In the business version, vehicle lifecycle data is also presented in aggregated form for fleets.

Architecture of added value: The Dash app is freely available for Android and IOS and can communicate with several OBD-II adapters available on the market, which can be purchased for example via Dash's online shop. Highlights of Dash are analyses of driving behaviour and vehicle condition, ratings and rankings (community functions), personal trend analyses and a map function to find the parked vehicle again. Dash also offers a platform for developers with the Dash Chassis API under the keyword "Internet of Cars".

Earnings model: From TechCrunch, a popular online news portal for technology and Internet companies, referred to as "FitBit for Cars", Dash intends to generate interesting insights from the collected vehicle life cycle data via its own analytics platform ("Dash IQ"), which will also be offered to other organizations. Dash is also a project partner in the DriveSmart project of the New York City Department of Transportation (New York City DOT, 2016), which aims to help drivers save time and money while driving more safely through feedback from the app. For example, drivers receive rewards when they drive in New York outside rush hour or use less congested routes. However, the website does not provide any information on specific sources of revenue.

## 2.5 Conclusion and Discussion

After an introduction to the quantified car phenomenon and to the efforts of US start-ups to establish digital ecosystems in this field, the business models of the quantified car case studies, Automatic, Mojio and Dash were described according to the methodology of Stähler (2002). It is striking that all three players pursue similar application scenarios. They each aim to generate relevant information from the vehicle life cycle data collected during the usage phase and to visualize this information for drivers accordingly. According to an analysis of the number of downloads from Android installations on Google Play, Dash (100,000-500,000 downloads) seems to be the most widespread, ahead of Automatic (10,000-50,000 downloads) and Mojio (1,000-5,000 downloads). The following illustration shows exemplary screenshots of the respective basic apps, which allows the similarities of the application scenarios to be deduced from the design of the user interface.


**Figure 20 Exemplary screenshots of the GUIs of Quantified Car Apps. (Stocker and Kaiser, 2016)**

In summary, the fuel for successful quantified car ecosystems is the vehicle life cycle data provided by drivers. Only when a critical mass of drivers voluntarily provide a critical mass of data can these Digital Ecosystems come into being. But this requires a variety of incentives, which can probably only be generated by interesting and free services with added value for drivers.

The chicken and egg problem is that drivers provide a large amount of data for third parties to develop applications based on, but many drivers will probably only provide data when interesting and beneficial services already exist. Parallels can be drawn between the chicken-and-egg problem of the "Web of Cars" and the chicken-and-egg problem of the Web of Data (Latif et al., 2009; Stocker et al., 2010). In order to overcome such a chicken-and-egg problem at all, the start-ups researched in this article were partly started with very simple applications that should allow drivers to quickly perceive a benefit.

Compared to the already established Quantified-Self-Movement, the Quantified-Car-Movement is still in its infancy, especially in Europe. The issue of data protection is very strongly anchored in politics, society and industry, particularly in the German-speaking countries. Accordingly, projects relating to data protection aspects must be carried out with a very sensitive approach. Not only for this reason, it can be assumed that a radical data-driven innovation in the Quantified Car environment will not necessarily be carried out by European vehicle manufacturers from the perspective of technology development.

Nevertheless, initial activities for data collection and transmission in connection with new vehicles have already taken place in Europe, as described in an article by Heise Online as reporting on an ADAC experiment (ADAC, 2016). However, it is not clear from these activities how collected vehicle lifecycle data is intended to generate benefits for the driver, nor how drivers can configure or prevent partial or total data transfer through selective data protection settings.

It is precisely in such a sensitive environment that startups from the USA, which do not suffer from a "privacy burden", are coming up with fresh ideas. Comparable to the activities of IT giants such as Google, Apple and Facebook, the topic of data protection is being pushed "a little" into the background for the time being. The competition between IT companies and established industry giants from the automotive industry for dominance in the establishment of digital ecosystems around vehicle life cycle data will certainly be exciting, as a current discussion paper from the BVDW describes (BVDW, 2016). The long-term exploitation strategy of some quantified car start-ups is certainly to sell technology and its user base to major players in the automotive and IT industries. Such a strategy can be seen in the very high risk capital investments on the one hand and the intransparent revenue models of the start-ups on the other.

Finally, it should be explicitly pointed out at this point that a number of start-ups focusing on the development of services related to mobility and smart cities have also emerged in Europe in recent years. However, these have far less risk capital than their "competitors" from the USA and are therefore probably not competitive in the long term. In Austria, for example, Parkbob (Parkbob, 2016) is an innovative manufacturer of a smart parking system, which has attracted an investment of EUR 250,000 in 2016. German vehicle manufacturers have already recognized the relevance of the Smart Parking topic. The German premium manufacturer BMW, for example, is working together with the US company and traffic data analyst INRIX (USD 143 million in 7 rounds of 6 investors according to Crunchbase.com) on a corresponding solution for intelligent parking as part of the ConnectedDrive initiative (INRIX, 2016).

# 3. Quantified Vehicles: Novel Services for Vehicle Lifecycle Data [9]

| Summary and Author Contribution |
| --- |
| **Definition and Introduction of "*Quantified Vehicles*"<br>(Paper 2/2)**<br>This journal article builds on the journal article from Chapter 2 (Stocker and Kaiser, 2016) and extends it with a research framework for the BISE community. The term *Quantified Vehicles* is derived and defined. In addition, it shows which trends exist internationally and in Europe, and their relevance for research. The article shows the existing research gap and the relevance of Quantified Vehicle research for BISE (Business & Information Systems Engineering) journal community by means of a research framework and discusses its contents. The BISE journal is a journal that publishes scientific research on the effective and efficient design and utilization of information systems and is therefore a relevant medium to publish scientific work in the field of business informatics.<br><br>Here, I focussed on the introduction to and the definition of *Quantified Vehicles as an* emerging topic to the BISE/IS community, the community I primarily want to address with my research. In doing so, I tried to shed light on business models of start-ups in this field, to increase understanding of how startups collect vehicle data for the services, and how they commercialize the services. I added the European start-up ulu.io to the overview of start-ups and their business models, to show that the trend is slowly reaching Europe, but is still being driven by enormous investments in the US. Furthermore, I wanted to include the high investment numbers of those start-ups to demonstrate how high investors perceive the market value. In addition, the discussion on the constitution of the *Data-driven Service Ecosystems* is started (termed *Digital Ecosystems* in the paper) and the reader learns why *Quantified Vehicles* are relevant for the BISE/IS community. |

## 3.1 Quantified Vehicles

Three trends have shown significant impact in recent years: (1) The Internet of Things (Wortmann and Flüchter, 2015) has become an enabler for a connected world full of smart objects equipped with sensors and supplies enormous and still rising amounts of (2) Big Data (Mayer-Schönberger and Cukier, 2013), which can be analyzed and then turned into business value in various areas, including (3) the Quantified Self movement as a popular example for everyday life big data analytics (Swan, 2009). On a more abstract level, capturing real

[9] The content of this chapter is based on

Stocker, A., **Kaiser, C.**, and Fellmann, M. (2017). Quantified vehicles. *Business & Information Systems Engineering (BISE), 59(2),* 125-130. DOI 10.1007/s12599-017-0465-5 world events and digitizing them into machine-readable data to satisfy needs or assist humans and machines in decision making, evaluation and comparison of physical world events has become increasingly important. Even a new branch of business has emerged through such big data analytics for data-driven innovations, while information overload has wiped off its negative image and has become the beautiful bride everybody wants to dance with. Mayer-Schönberger and Cukier's (2013) pragmatic book on the capacity of big data to change the world has become an international bestseller and was referenced by researchers more than 1000 times according to Google Scholar.

In line with these developments, consumer products are increasingly connected to the Internet and have become a major source of data, too. So-called smart, connected products (Porter and Heppelmann, 2014 & 2015) are capable of capturing an increasing amount of data about their product life through all kinds of embedded sensors. The archetype of a smart, connected product is the well-known, widely used, and constantly switched on smartphone. The smartphone has become an interesting hub for sensors of all kinds, and has therefore kicked off the development of new services encapsulated in mobile applications. Some of those applications promise an additional value for the smartphone user by applying algorithms for sensor data analysis if the user is willing to share the required sensor data.

In the age of computing humans have become data-generating subjects, because they consciously or unconsciously leave behind 'electronic traces' when using their computer (Wolf, 2013). 'Quantified self-tracking' (later shortened to 'quantified self') is a more current term, referring to an intended collection of any data about the self that can be measured, including biological, physical, behavioral, or environmental information (Swan, 2009). Quantified-selfers are a diverse group of early adopters including life hackers, data analysts, computer scientists, health enthusiasts, gamers, productivity gurus, and patients, who track many kinds of data about themselves (Choe et al., 2014). Making use of this data collected through smartphones or wearables in the private domain to learn more about one's body and leisure behavior is an emerging topic and has become a major creator of value (PWC, 2016).

This quantified self-movement, or in more general terms the pattern of collecting personal data via consumer devices and making sense of this data thereafter, has increasingly gained attention throughout many areas – not only in domain of quantified self-tracking services (Swan, 2009). Given the fact that vehicles have turned into 'computers on wheels' (Haberle et al., 2015) and the intimate relation between a driver and a car, it is quite obvious and straightforward to also interpret cars as customer devices. In this way, the pattern of self-tracking via customer devices adopted by the quantified-self community can be transferred to vehicles which in this sense become Quantified Vehicles. From this perspective, the data modern vehicles collect holds a huge potential for further exploitation and value creation. The automotive industry is still in the early beginning of leveraging this potential. Quantifying vehicles in terms of analyzing gathered data and developing innovative applications and third party services for the consumer context is currently not the state of the art in the vehicle domain, especially from the perspective of Original Equipment Manufacturers (OEMs). Although a vast amount of data for the purposes of steering, regulating and informing has already been collected, up to now this data is mainly processed to ensure vehicle functionality. Exploring novel application scenarios complementing their core mode of use is hence among the most important topics in the automotive industry. Lifecycle data generated and transmitted by vehicles may further be a part of a future connected world (Swan, 2015). They will enable a wide range of application scenarios and business models and therefore have to be considered a valuable area for BISE research.

## 3.2 Current Developments: Towards Digital Ecosystems for Quantified Vehicles

The EU funded project termed AutoMat and coordinated by Volkswagen kicked off in April 2015 and is one of the first approaches to establish an automotive big data marketplace for innovative cross-sectorial vehicle big data services (AutoMat, 2016a). According to the project outline, more than 4000 signals are processed per second inside modern vehicles, and the amount of data transferred by the Controller Area Network (CAN)-Bus (ISO, 2015) inside a single vehicle accumulates to about 500 MB per hour. This continuously generated vehicle lifecycle data embodies a significant business potential not only for vehicle OEMs, but for all kinds of cross-sectoral industries.

However, this potential to use vehicle lifecycle data for purposes other than driving currently remains almost untapped by automotive OEMs. According to AutoMat, the automotive industry has not yet been able to successfully establish an ecosystem for Quantified Vehicle apps equivalent to that of smartphone manufacturers. In its problem statement, the AutoMat project mentions three reasons why OEMs are currently struggling:

- brand-specific business approaches dominate, and as a consequence there is a lack of brand-independent vehicle lifecycle data,

- current proprietary vehicle services focus on the individual customer, which results in privacy concerns, and few ideas exist how anonymized vehicle data can be used to establish other services, and
- the implied or required collaboration between OEMs on vehicle data and services is considered risky in terms of competition.

Apart from this individual major research project, digital innovation in the Quantified Vehicle domain seems rather to be pushed forward by a steadily growing number of innovation-friendly start-up companies, the majority of them located in the USA. They aim to create market demand by first providing novel platforms, APIs, apps and services. Their market approaches share a few commonalities: All of them provide a basic branded hardware required for capturing vehicle data from the CAN bus and transferring it directly (via embedded 3G/4G modem) or indirectly (via the vehicle driver's smartphone) to a cloud platform. This task is usually conducted by providing a proprietary adapter connected to the standardized vehicle On Board Diagnostic interface (OBDII standard), which is available in any modern vehicle. Some of these startups even allow third party apps and services to be built on top of vehicle data, which is not gathered by just one single vehicle but by a plethora of vehicles. In the best case, all vehicles would supply a cloud platform vendor with their lifecycle data to enable a digital ecosystem with interesting applications and services for drivers and other stakeholders similar to those ecosystems which Apple and Google have created for smartphone apps. To increase the customer value of such services, vehicle lifecycle data could eventually be enriched by data from other sources including weather data or map data. This creates synergies with the open data movement.

Applications and services provided via an established Quantified Vehicle cloud can generate value for the individual vehicle driver (e.g., assessing personal driving style and offering suggestions how to improve it), for an organization (e.g., easing insurance contracts or supporting fleet management), or even for both, to ensure a sustainable business model. Table 8 below presents a preliminary overview of Quantified Vehicle start-ups and services obtained through desktop research (using a combination of the terms “quantified”, “connected”, “vehicle”, “car”, and “startup”). Information regarding funding was collected via crunchbase.com.

**Table 8 Overview of Quantified Vehicle start-ups. (Stocker et al., 2017a)**

| Company | Value Proposition | Personal Use | Business Use | Revenue Model | Funding |
|---|---|---|---|---|---|
| **auto-matic.com** San Francisco (USA), founded 2011 | Connect car to the rest of the digital life. Empower drivers with knowledge about themselves and their cars so they can drive safer and smarter | Connect the car to a world of apps, car problem diagnosis, Fuel efficiency, location and emergency service, Web dashboard, keep track of the car, third party apps | not explicitly mentioned | Light version: $49.55 for car adapter and free apps for iPhone and Android Pro version: $129.95 for car adapter, unlimited 3G syncing for 5 years, crash alert service | $24M in 3 round from 12 investors |
| **moj.io** Vancouver (CAN) & Palo Alto (USA), founded 10/2012 | Be a smarter Driver. Open platform for connected cars, enables a suite of apps to empower and inform car drivers. | Access to marketplace for connected car apps (location tracking, vehicle diagnostics, recalls and maintenance, driving analytics, accerlerometer), API | Aiming to launch carrier-braned car services | Not explicitly mentioned | $10.3M in 2 rounds from 6 investors |
| **ulu.io** Eindhoven (NED), founded 12/2014 | Create a more sage and better driving experience by opening up the car ecosystem and connect all cars to the internet. | Car insigts, car monitoring, assistants, learning to improve driving behavior, saving money while driving | Solutions for insurance, leasing, car dealership and repair shops | Personal use: €129 excl. VAT per year; business use: €177 excl. VAT per year | $558.75k in 2 rounds from 2 investors |
| **vin.li** Dallas (USA), founded 2014 | Your car, your way: Brings an endless range of apps to your car, from safety to entertainment to onboard wifi. | Highspeed wifi, all kinds of smart car apps and services downloaded via an own appstore. More than 40 apps and integrations. App ecosystem for the car. | not explicitly mentioned | Not explicitly mentioned | $7M in 2 rounds from 6 investors |
| **zendrive.com** San Francisco (USA), founded 2013 | Improving driving for everyone through better data and analytics: Use smartphone sensors to measure a driver's behavior | Powerful analytics using machine learning algorithms, driver and fleet analytics. Collision detection, Insurance support, etc. Provider of SDK and API | Insurers, cashare, valet, platform partner | Free for 1-4 drivers, $4 per driver/month for 5-249 drivers, no prices communicated for 250+ drivers (fleets, insurance partners, platform partners) | $20M in 3 rounds from 14 investors |
| **zubie.com** Sullivans Island (USA), founded 05/2012 | We make driving safer and worry free: Connect car to the Internet to deliver real-time information to the smartphone. | Driving insigts, behavior alerts, leaderboard, maintenance alerts, engine diagnostics, low battery alert, roadside assistance, live map, trip tracking, motion monitor, perks, fuel finder | Solutions for insurance and car dealers, low cost fleet tracking, GPS tracking, vehicle health, driver performance | Personal use for $99.95 per year Business use starts at $17.95 per month (including $49.95 for device) | $25.87M in 5 rounds from 8 investors |

In the following, the authors abstract from concrete product and service offerings available so far (cf. Table 8) and sketch the concept of a Quantified Vehicle cloud, which could serve as a basis for novel applications and services. In order to make the concept of "Quantified Vehicles" achieve its full potential, at least four different types of stakeholders have to be considered:

1. Primary end users (individual service consumers) are vehicle drivers who directly benefit from innovative products, visualizations, statistics, gamification aspects, and recommendations based on assessing their shared vehicle lifecycle data. Value generated on the individual level is an incentive to share personal driving data.
2. Secondary end users (organizational service consumers) are organizations which indirectly benefit through collected and assessed vehicle lifecycle data from multiple vehicles by consuming special services (e.g. engineering, city planning, advertising).
3. Service providers are organizations which provide Quantified Vehicle services for primary and secondary end users, thereby generating revenues (e.g., providing fleet management services, traffic-style dependent insurance services, vehicle maintenance prediction services).
4. The cloud service provider (platform provider) operates the required infrastructure for the Quantified Vehicle ecosystem and allows service providers to establish their services based on vehicle lifecycle data as well as primary and secondary end users to consume these services and share their vehicle lifecycle data in return.

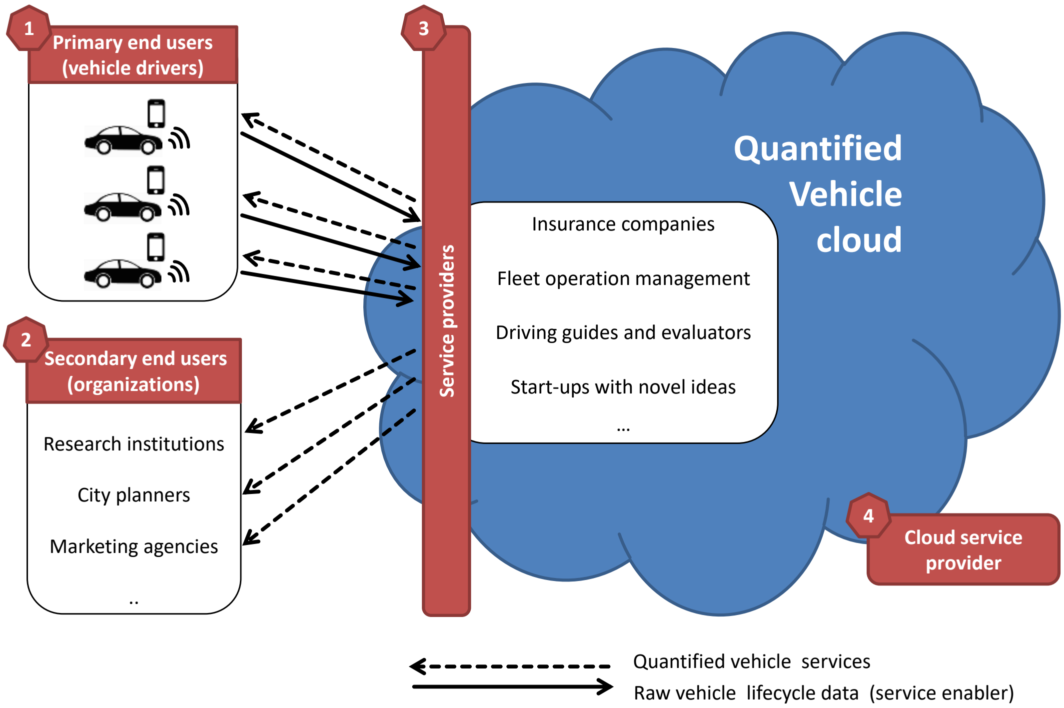


**Figure 21 A digital ecosystem for quantified cars. (Stocker et al., 2017a)**

## 3.3 Relevance of Quantified Vehicles for BISE Research

As an interdisciplinary subject, conducting research in the Quantified Vehicle domain requires the integration of different disciplines, ranging from vehicle enthusiasts, mechanical engineers, electrical engineers, information designers, and computer scientists. As recent advances in information technology are a core enabler for Quantified Vehicles, BISE faces a wide range of research challenges through the emergence of Quantified Vehicles, from business model design to the ideation, conceptualization and development of Quantified Vehicle apps and services.

Relevant aspects and their relations are illustrated in the next figure that may serve as an initial research framework for Quantified Vehicles. Quantified Vehicles will require new infrastructures such as interfaces for sensor data and data storage capabilities. Based on this foundation, vehicle lifecycle data can be collected. With this data, new processes and new insights will emerge, which use this data in conjunction with appropriate data analytics techniques. On top of this, new business models and strategies are possible. Though Quantified Vehicles as such is a technology-driven subject, cross-cutting concerns such as human aspects (e.g., needs, intentions, expected benefits) as well as security and privacy issues cannot be neglected. Finally, the overall design space for Quantified Vehicles will be constrained by different legal aspects and by standards representing commonly agreed technologies and practices.

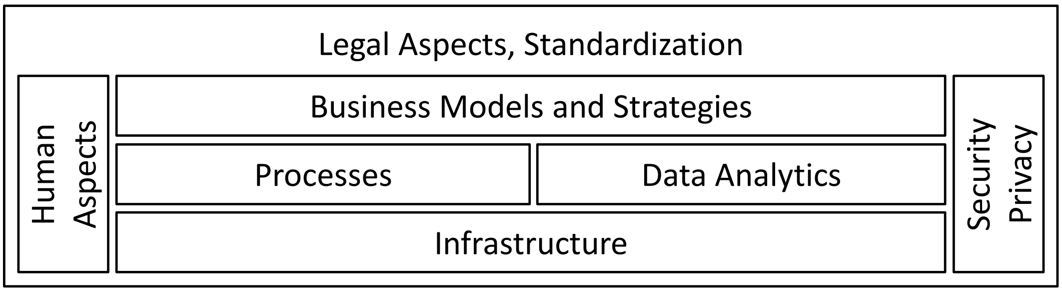


**Figure 22** A research framework for Quantified Vehicles. (Stocker et al., 2017a)

### 3.3.1 Infrastructure, Processes, and Data Analytics

Modern vehicles gather an enormous amount of data and information, raising manifold challenges concerning storing, securely transferring and analyzing this enormous amount of lifecycle data, which have to be solved. Moreover, problems of information integration will arise. In this regard, the integration of Quantified Vehicle data and services into existing enterprise information systems (EIS) such as Enterprise Resource Planning (ERP) systems will require novel approaches, e.g., for business analytics and visualization. To leverage the potential of vehicle lifecycle data, e.g., to improve decisions of the stakeholders, the way in which they are presented is of utmost importance. Intuitive visualizations, different point of views on the data, comparison possibilities and statistics amongst others, should be integrated into EISs alongside the administration of the vehicles itself.

### 3.3.2 Business Models and Strategies

Quantified vehicles will enable novel apps and services and cause vast opportunities for business model innovation in various fields (Cichy et al., 2014). However, the ideation, design, and evaluation process of useful and valuable services and apps as well as standardization issues of data and interfaces are a major challenge to be tackled in an interdisciplinary way, which can be supported by BISE researchers through providing appropriate methodology. Novel innovation approaches including, e.g., hackathons (Briscoe and Mulligan, 2014) can provide means for quickly transforming ideas into experienceable demonstrators for Quantified Vehicle apps and for evaluating their business potential together with end users.

Vehicle lifecycle data can be used to enable a broad portfolio of value-added consumer services including vehicle diagnostics, driving dashboards, or concierge services. Novel applications and services can generate value on the level of the individual driver as the primary beneficiary (e.g., benchmark driving style and offer suggestions on how to improve it), on the level of an organization as the secondary beneficiary (e.g., traffic prediction for smart cities), and on that of the society (e.g., reduced fuel consumption or sharing approaches).

Future research is required to determine the willingness of drivers to pay for Quantified Vehicle services and apps and to identify incentives for all stakeholders to share their vehicle lifecycle data. For example, it may be the case that the benefit of other stakeholders such as vendors or insurance companies dominates the value proposition – i.e., that individual users are not willing to pay for new data collection features of vehicles. Hence it has to be investigated in how far incentives to buy a “Quantified-enabled Vehicle” via transfer payments are possible.

### 3.3.3 Human and Societal Aspects

On the level of the individual driver, quantified vehicles offer a series of new possibilities to investigate driving behavior and drivability with respect to enhancing driving style, driving safety, and security. Quantified Vehicles enable new ways to investigate driver reaction and driver emotions by using computational approaches. In this way, also pleasure of use or additional stress and distraction may be measured. Obtaining easier access to vehicle lifecycle data would encourage a larger group of researchers to use this data for better understanding driver experience and driver behavior (Wilfinger et al., 2013).

On the level of the society, it has to be investigated what the merit of Quantified Vehicles could be in terms of safety, environmental impact as well as possible dangers such as increased surveillance possibilities that create new potentials for misuse and unethical behavior.

### 3.3.4 Security, Privacy, Legal Aspects, and Standardization

Quantified Vehicle ecosystems can only be successful if a critical mass of drivers shares their driving data. Hence, privacy concerns have to be mastered to support the emergence of third-party services with sufficient data to create representative statistics. If no data is shared, no value is generated. Raising awareness in the society on what kind of data vehicles generate, process, store, and potentially transmit to a vehicle manufacturer is an important task which can be supported by researchers.

The ‘My Car My Data’ campaign (MyCar MyData, 2016) started by Fédération Internationale de l’Automobile (FIA) educates car drivers about potentials and pitfalls of connectivity. One strategy is to let drivers decide about if and what data should be shared with whom to be used in what kind of third party services. This raises the question if new data protection laws are needed or if existing regulations in the IT-domain are still sufficient. The legal status of Quantified Vehicle data (e.g., in legal procedures) has to be determined, which creates cross-disciplinary research opportunities for IT-oriented researchers cooperating with legal scholars. New legal questions will emerge, e.g., whether vehicle data is trustworthy in a legal sense (as replacement for the "driver’s logbook") and can serve as basis for vehicle tax calculations.

Finally, in terms of technology, standardization may be the key to progress and may prevent competing and incompatible solutions. In this regard, it has to be investigated if current standards such as CAN and OBD that were developed in the late 1990s of the previous century are still sufficient to establish digital ecosystems for Quantified Vehicles. Researchers should look into the degree to which information services should be standardized to foster the development of Quantified Vehicle ecosystems. The W3C for instance has established an own automotive working group to create Web standards for the automotive industry (W3C, 2016).

# 4. Quantified Cars: An Exploration of the Position of ICT Start-ups vs. Car Manufacturers Towards Digital Car Services and Sustainable Business Models [10]


Summary and Author Contribution

**Analysis of the Market: Services, Start-ups, OEMs, Business Models, and Trends (Paper 1/2)**

This chapter provides insights into the quantified car phenomenon and explores the approaches of the two major stakeholder groups, car manufacturers and tech start-ups, on their journey to develop novel digital services and sustainable business models. While providing a playground for innovative US tech start-ups backed with risk capital, in contrast to that, especially German-speaking car manufacturers have been rather reluctant to reap the value of 'their' car operation data in delivering successful digital services to stakeholders. The chapter examines stakeholders who have an interest in the data, lists trending start-ups and investigates their visions, goals, and business models, shows developments in the behavior of German vehicle manufacturers, and attempts to derive future trends from this.

Here, I focused to draw the attention of the business model community to the *Data-driven Service* trend in the U.S., while there are still no well-known services in Europe and no business model studies on this. I show that there are several stakeholder groups that should be interested in it. I also discovered open key questions concerning vehicle data (What data do I have, How is their quality? Etc.), which is a big risk for start-ups with the OBD-2 dongle approach, and that there is also an approach using smartphone data. Furthermore, the reader learns that the German vehicle manufacturers tend to react negatively to the trend and want to restrict the activities of start-ups.


Capturing real world events and digitizing them into machine-readable information has become increasingly important. The digital age has transformed humans to data generators, while they consciously or unconsciously have left behind their 'electronic traces' (Wolf 2013). 'Quantified self' is a term coined to describe the intended collection of any measurably characteristics about a person, including biological, physical, behavioral, or environmental aspects (Swan 2009). Usually data is collected through the consumer devices of 'Quantified Selfers', most notably through smartphones.

---

[10] The content of this chapter is based on

**Kaiser, C.**; Stocker, A.; Viscusi, G.; Festl, A.; Moertl, P.; and Glitzner M. (2017). Quantified Cars: An Exploration of the Position of ICT Start-ups vs. Car Manufacturers Towards Digital Car Services and Sustainable Business Models. In *Proceedings of 2nd international conference on new business models (NBM)*, pp. 336-350.

Quantified self has become a major creator of value through Android/iOS mobile applications. One example is the Austrian start-up Runtastic which provide a smartphone application (105 million registered users) to analyze how users perform when they run, bike, etc. Runtastic was acquired by Adidas in 2015 for about 220 million EUR. Adidas now holds the power to the knowledge of designing sportswear combined with the digital knowledge on usage gained through the user base of Runtastic, which can offer new insights for individual product development. This acquisition demonstrates that big industrial players invest into innovative quantified self start-ups with an exploitable mass of collected data from a broad user base. So what can this acquisition of a quantified-self start-up contribute to the car domain? During the last two decades, passenger cars have slowly turned into computers on wheels (Haeberle 2015) equipped with many sensors used for functionality, safety and joy. Taken into account that cars capture sensory data about themselves and about their environment, the behavioral patterns of self-tracking can be transferred to cars (and vehicles in general), which in this sense become 'Quantified Vehicles' (Stocker et al., 2017a).

Obviously, quantifying cars in terms of analyzing driving data and developing innovative applications is a comparably new phenomenon. The continuous collection of car operation data can enable the analysis of both car- and driver behavior and thereby facilitate the generation of innovative digital products, services as well as sustainable business models for many beneficiaries, including e.g. drivers and organizational customers. There are many opportunities to reduce emissions by stimulating safer driving and improving road safety while caring more about the natural environment by using novel digital services as the following Table 9 suggests.

**Table 9 Stakeholders and their interest towards digital services. (Kaiser et al., 2017b)**

| STAKEHOLDERS | INTEREST FOR DIGITAL SERVICES |
|---|---|
| Individual drivers | Individual drivers may be empowered to assess their personal driving style and get improvement suggestions to drive more safely or economically friendly. |
| Various organizational customers | Insurance companies, to name a typical beneficiary heavily investing into quantified car start-ups, can provide new kinds of insurance contracts for safer drivers and will be provided with new means to infer driving risks.<br>Driving schools can be supported in supervising students based on digitally monitored driving styles, teaching them to drive safer and economically friendly. |

| | |
|---|---|
| Governmental authorities | Road traffic departments of cities can be empowered to make informed decisions based on their gained knowledge about traffic patterns, thereby increasing road safety and reducing driving emissions in urban environments. |
| Automotive industry | Car manufacturers may use digital services to optimize powertrain calibration for special usage behavior (e.g. in postal delivery transport).<br>Automotive engineers may improve the accuracy of driver models and testing for advanced driver assistance systems. |

There are many stakeholders who have an interest in exploiting the data generated by cars either supporting their current business processes and models or adapting them towards establishing digital ecosystems. However, in the scope of this chapter, two stakeholder groups are of particular interest: Car manufacturers as the owners of the underlying technology and ICT start-up companies especially from the tech savvy USA, who are keen to develop new digital ecosystems in the automotive domain. Against this background, the chapter outlines the following research question:

> *What are the roles of start-ups vs. car manufacturers in delivering novel digital services and sustainable business models built on car operation data analysis?*

After this introduction, the chapter will introduce background, motivation and the research approach in Section 4.1. Based on that, the chapter will present approaches towards novel digital services and elaborate on the positions of ICT start-ups versus car manufacturers, exploring suggestions of the VDA in Section 4.2. The chapter will conclude with a discussion on future opportunities of these two stakeholders in Section 4.3.

## 4.1 Background and Research Approach

### 4.1.1 Background and Motivation

Digitalization is a sociotechnical process on encoding analog information into a digital format (*digitizing*) applied to social and institutional contexts, transforming their sociotechnical structures, thus rendering digital technologies infrastructural (Tilson et al., 2010). Digitalization results in digital artifacts characterized by *editability*, *interactivity*, *reprogrammability/openness*, *distributedness* (Kallinikos et al., 2013), also implying a shift in product design moving from modularity to generativity (Yoo et al., 2012). Quantified cars are one of the key results of digitalization in the automotive industry, where incumbents have to face competing concerns systematically interrelated, as shown by Svahn et al. (2017) through the case of Volvo Cars: *innovation capability*, *innovation focus*, *innovation collaboration*, and *innovation governance*. Thus, from a strategy perspective, digitalization enforces *internetworking* considered as “those business processes/activities conducted or mediated online between employees, customers, suppliers and partners of firms, using internet-based technologies accessed through internet-based infrastructures.” (Brews & Tucci, 2007, p.224). Besides incumbents digitalization enables digital entrepreneurship providing less bounded entrepreneurial processes and outcomes and less predefinition in entrepreneurial agency (Nambisan, 2016), as shown also, e.g., by the rising number of start-ups focusing on quantified cars. Accordingly, the role of information value is a central challenge in the competitive scenarios emerging from digitalization as well as information capacity of companies (Viscusi & Batini, 2014). Key questions here are: What data do I have? How is their quality? Can I exploit them in their full potential? What can I infer from them? While ICT start-ups have already started to apply the quantified self phenomenon to cars, launching apps and services to generate a new market, car manufacturers are currently in the transition process from vehicle manufacturers towards integrated mobility and data service providers. According to Dedrick (2010) researchers have framed the impacts of IT on environment as first-order (impacts from IT hardware during the product lifecycle), second-order (impacts of ICTs on other processes such as transportation or industrial production), and third-order effects (changes in lifestyles and economic structures). The latter are relevant when considering the increased use of social media transformative potential of ‘green’ information systems on the demand side, encouraging practices such as, e.g., carpooling and ridesharing applications coupled with the Internet of things (Malhotra et al., 2013). According to Malhotra et al. (2013) this two-way, sensor-driven communication is blurring the boundaries between the production side and the demand side. Furthermore, information systems and interdisciplinary strategies for quantified cars may provide models to assess the value of information, in particular the social value of related open data (Viscusi et al., 2014).

Taking the above issues into account, business models are a key element for competing in markets characterized by extensive use of ICTs and currently transformed by digitalization. In general terms, a business model describes the rationale of how an organization creates, delivers, and captures value (Osterwalder & Pigneur, 2010). Massa et al. (2017) provides a systematic view on the different perspectives on business model research pointing out that business models can be considered as i) attributes of real firms, ii) cognitive/linguistic schema, and iii) formal conceptual representations/descriptions of the former two issues. As for conceptual representations, Al-Debei & Avison (2010) identified four key dimensions of business model: *value proposition*, *value architecture*, *value network*, *value finance*. As for business models as attributes of real firms it is worth mentioning the definition by Zott & Amit (2010), who conceptualize a business model as "a system of interdependent activities that transcends the focal firm and spans its boundaries. The activity system enables the firm, in concert with its partners, to create value and also to appropriate a share of that value" (2010, p. 216). This definition is useful for understanding business models of companies interested in quantified cars when linked to the above concept of internetworking and current pervasiveness and strategic relevance of digital platforms (Parker et al., 2016). Considering quantified cars, despite the "analyzing" stance of the main market players, a set of traditional and new business models can be applied (Shipilov, 2016), in particular the *infomediary* one (Afuah & Tucci, 2000) can be adopted under a utility perspective and extended from data collection for, e.g., marketing purpose to data useful for social value, as capability and functioning they enable (Viscusi & Batini, 2016), and finally for sustainability issues.

Besides environmental and societal issues, business sustainability refers to "business models and managerial decisions that creates value over the short, medium, and long terms, based on mutually beneficial interactions between the company's value chain and the social and environmental systems on which it depends" (Lüdeke-Freund et al., 2016, p. 18). Furthermore, Schaltegger et al. (2016, p. 6) points out that a business model for sustainability "helps describing, analyzing, managing, and communicating (i) a company's sustainable value proposition to its customers and all other stakeholders, (ii) how it creates and delivers this value, (iii) and how it captures economic value while maintaining or regenerating natural, social, and economic capital beyond its organizational boundaries." Still, business model innovation in automotive industry asks for understanding the different ways the various actors can follow to innovate their business models; in particular, as pointed out by Massa & Tucci (2014, p. 424), business model design in newly formed organizations, which refers to their "entrepreneurial activity of creating, implementing and validating a business model", and business model reconfiguration in incumbent firms, encompassing the reconfiguration and eventual acquisition of organizational resources to change an existing business model.

### 4.1.2 Research Approach

This chapter is aimed to provide a first exploration of the position of innovative ICT start-ups vs. car manufactures towards establishing new services and sustainable business models. Although the topic quantified car per se is in the domain of car manufacturers, an increasing number of ICT start-ups have used their innovation capabilities to develop own means for capturing this valuable data source.

Against this background the authors of this chapter conducted a desk research approach analyzing information available on the Web to further explore the activities of quantified car start-ups and car manufacturers. They have used a combination of the terms quantified, connected, vehicle, car, and start-up in search engines to capture the current developments. Furthermore they have used crunchbase.com to capture additional meta-information on company location, business and funding.

After having identified the major quantified car start-ups, which are listed in Table 10, two authors have studied start-up websites in detail to find out more about their visions and goals as well as about their business models, products and services. Both authors have reviewed the websites of all start-ups and discussed their knowledge with the other person afterwards to come to a common understanding. The information was then validated by two additional persons, which are co-authors of this chapter as well.

## 4.2 Results: An Exploration of Novel Services and Business Models

### 4.2.1 The Position of ICT Start-ups in the USA towards Exploiting Car Data

In analogy to the quantified self movement, the dominating IT/Web industry of the USA has already lined up a series of quantified car start-up companies backed by risk capital, reaching more than 20 million USD in some cases, demonstrating how high investors perceive the market value of a car data ecosystem for quantified cars (Stocker et al., 2017a). The start-ups exploit data generated by cars while driving. The crucial source for any data-driven start-up is data and this statement also holds for quantified car start-ups. The following table lists start-ups which have been identified in desk research. It provides an overview and includes the company names, their URL as well as their value propositions provided on the website.

**Table 10 Quantified Car Start-ups. (Kaiser et al., 2017b)**

| COMPANY | URL | VALUE PROPOSITION |
|---|---|---|
| Automatic | automatic.com | Unlimited car monitoring, zero fees. The only connected car adapter with unlimited 3G included. |
| Automile | automile.com | #1 Fleet & Asset Tracking |
| Dash | dash.by | Connect your car to Dash, to make driving smarter, safer, greener and more affordable. |
| Metromile | metromile.com | Metromile's pay-per-mile insurance offering saves low-mileage drivers a ton of money. |
| Mojio | moj.io | The Leading Open Platform for Connected Cars |
| Vinli | vin.li | …the leading car platform for bringing smart car functionality to any car on any lot, in any fleet, or in any shop. |
| Zendrive | zendrive.com | Smartphone-powered road safety for cities, fleets, and individuals. |
| Zubie | zubie.com | …connects your car to the internet to deliver real time location, trip history, maintenance alerts, engine diagnostics and driving insights. |

For data acquisition purposes, start-ups must involve car drivers/owners somehow. The car driver creates data while driving his car and therefore has to be understood as the owner of the driving data. While car manufacturers have a comparatively easy technical access to the data a car generates during its operation, tech start-ups have to identify novel ways on how to capture this data before they can use it in applications. The conducted research has identified two major data acquisition approaches pursued by the start-ups:

- The first approach, pursued by the majority of quantified car start-ups including e.g. Automatic or Mojio, is to utilize a branded hardware plug connected to the OBD-II interface, a standardized interface for all modern cars. This allows them accessing certain car sensor data, e.g. speed or rpm. They may equip their devices with additional sensors including GPS or accelerometer to collect additional relevant data describing the movement of the car. Both plug and internet connectivity are usually not free of charge. They produce a lock-in effect to the particular business model of a quantified car start-up and are one way to safeguard revenues.
- The second approach, e.g. pursued by Zendrive, is to use the sensors built into modern smartphones, e.g. GPS, accelerometer or luminance, to capture data while

driving. This makes smartphones suitable devices to track car trips, too. However, smartphones lack information provided by car sensors including emission data or rpm. Nevertheless, smartphones have the advantage of an 'always on' connectivity for event extraction and service provision.

The reviewed start-ups have specialized in capturing, storing, and analyzing large quantities of car data and offering services in smartphone applications to motivate for sharing valuable driving data. The majority of start-ups are capable of extracting interesting driving events including e.g. hard brake, hard acceleration or speeding to name a few. These events are hidden in the field data and have to be revealed through applying data analytics. Mobile applications running on the user's smartphone then pull the results and visualize them on the driver's smartphone. Figure 23 provides snapshots of such mobile app user interfaces.

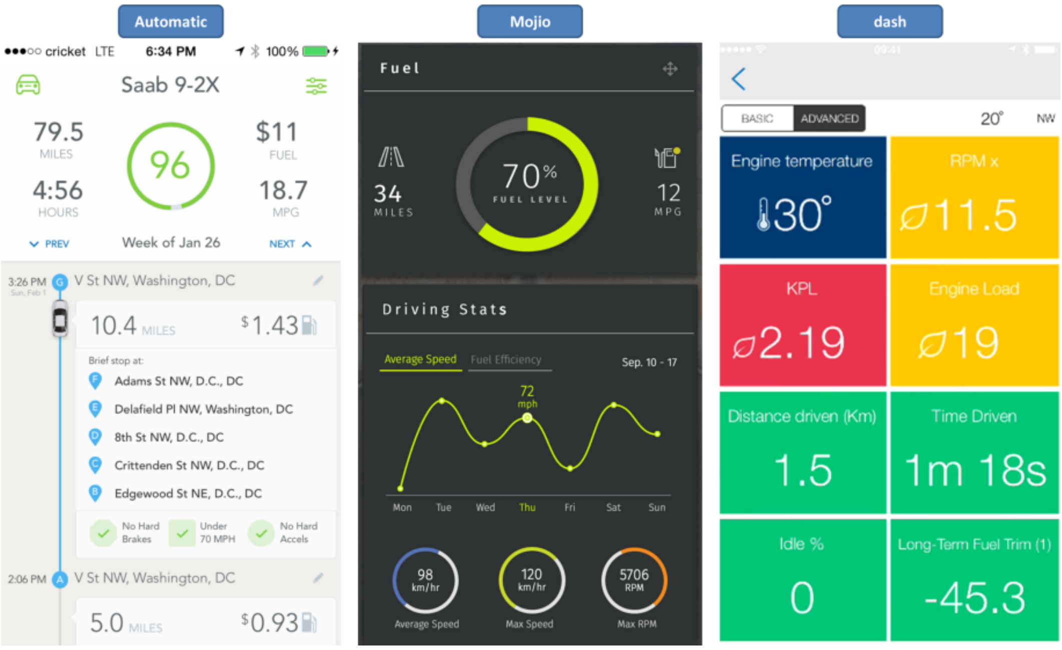


**Figure 23** **Start-up Apps: Automatic, Mojio and dash. (Kaiser et al., 2017b)**

Some start-ups even provide APIs and software development kits to software developers in order to increase their market reach through third party apps or even to become the most important car data service platform, through a business model comparable to Apple iTunes. Some may even pursue the strategy of being bought by a big player in the future. These start-ups are definitely eager to increase market penetration. Their main competencies are applying novel data analytics on large quantities of trip data, storing large quantities in their datacenters, providing innovative mobile applications to the user including gamification apps as well as dashboards and interfaces for other parties (e.g. fleet managers) to allow analytics on fleet data.

### 4.2.2 The Position of Car Manufacturers towards the Quantified Car Phenomenon

Regarding access and ownership of car generated data, car manufacturers are in a comparably lucky position. However, they were not very successful in exploiting this market yet. The potential to exploit car lifecycle data for purposes other than driving currently remains almost untapped by automotive OEMs (Stocker et al., 2017a). According to the EU research project AutoMat (2016c), the automotive industry has not yet been able to successfully establish an ecosystem for apps and services equivalent to that of smartphone manufacturers. The project mentions three reasons why OEMs are currently struggling: Brand-specific business approaches dominate, and as a consequence there is a lack of brand-independent car lifecycle data. Current proprietary car services focus on the individual customer, what leads to privacy concerns, and few ideas exist how anonymized car data can be used to establish other services. The implied or required collaboration between OEMs on car data and services is considered risky in terms of competition.

However, success and interest in car data start-ups seem to have made an impact on OEMs business strategies. As the AutoMat system structure (Figure 24) illustrates, OEMs are interested in taking over the data provider role and to establish a car data ecosystem: Data acquisition systems will be integrated into cars. Cat data is transmitted to an OEM backend, where it can be cleaned and enriched with further relevant information before publishing it to service providers, e.g. tech-start-ups, which then can provide third party applications on a marketplace. Though there is an OEM backend, the Automat System terms its approach an 'open ecosystem' in its deliverables, where the willingness of other stakeholders to pay for digital business models is an important topic, too.

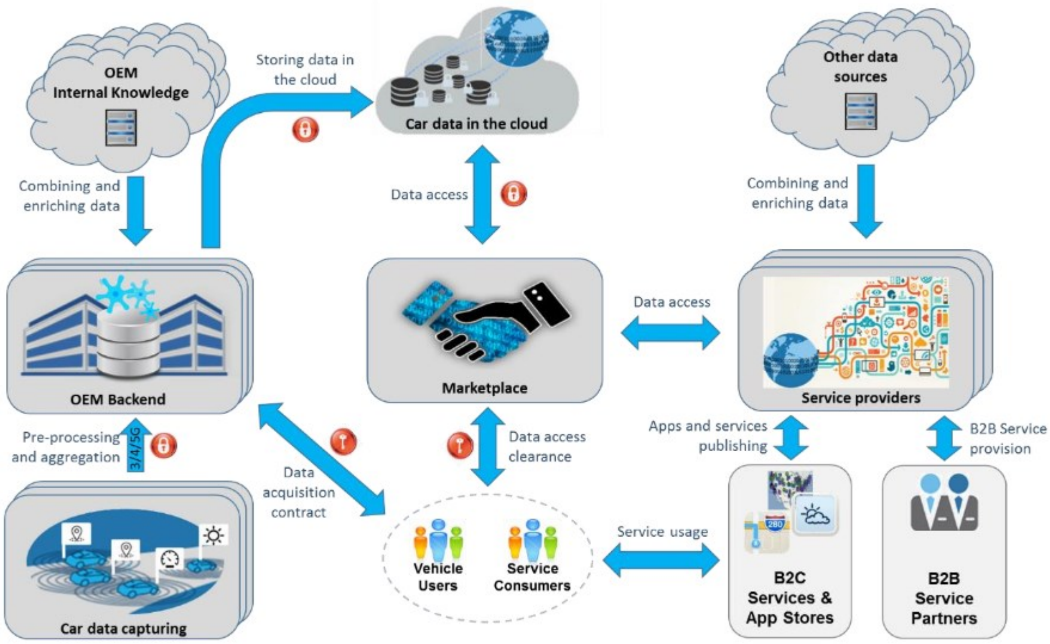


**Figure 24 The AutoMat System. (Source: AutoMat, 2016c)**

Tech start-ups heavily depend on the OBD-II interface yet. If access to this interface would be limited or denied, their business models would be endangered. A recent eeNews Automotive (2017) article titled ‘German car industry plans to close OBD interface’ emphasizes the hypothesis that car manufacturers want to take over the data provider role, it states:

> *“Instead, the data will be made accessible to interested third parties through a neutral server, and basically under control of the automotive industry.”*

There are two relevant recent position papers from VDA concerning the role of the German-speaking car manufacturers towards digital car data ecosystems. The position paper ‘data protection principles for connected vehicles’ (VDA, 2014) refers to the continuous transformation of vehicles towards ‘connected vehicles’ with a permanent uplink to the internet and the feasibility to connect various data sources for establishing new services. The position paper suggests three principles for VDA members to handle the advancements in connectivity and the new services associated with respect to responsible data handling as well as with data protection:

- Transparency: VDA members strive for adequate information about the data in connected vehicles and the use of these data.
- Self-determination: VDA members are striving to enable customers to determine themselves the processing and use of personal data through various options.

- Data-security: VDA members strive to implement the strong safety culture in connected vehicles.

The short paper closes with a chart of data categories in connected vehicles and their relevance for protection:

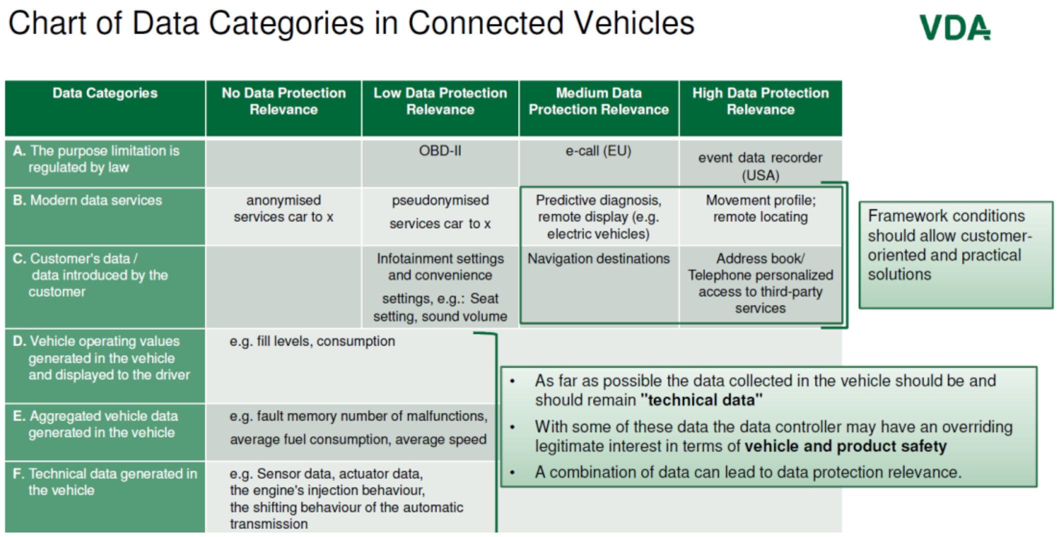
Chart of Data Categories in Connected Vehicles

VDA

| Data Categories | No Data Protection Relevance | Low Data Protection Relevance | Medium Data Protection Relevance | High Data Protection Relevance |
|---|---|---|---|---|
| A. The purpose limitation is regulated by law | | OBD-II | e-call (EU) | event data recorder (USA) |
| B. Modern data services | anonymised services car to x | pseudonymised services car to x | Predictive diagnosis, remote display (e.g. electric vehicles) | Movement profile; remote locating |
| C. Customer's data / data introduced by the customer | | Infotainment settings and convenience settings, e.g.: Seat setting, sound volume | Navigation destinations | Address book/ Telephone personalized access to third-party services |
| D. Vehicle operating values generated in the vehicle and displayed to the driver | e.g. fill levels, consumption | | | |
| E. Aggregated vehicle data generated in the vehicle | e.g. fault memory number of malfunctions, average fuel consumption, average speed | | | |
| F. Technical data generated in the vehicle | e.g. Sensor data, actuator data, the engine's injection behaviour, the shifting behaviour of the automatic transmission | | | |

Framework conditions should allow customer-oriented and practical solutions

- As far as possible the data collected in the vehicle should be and should remain **"technical data"**
- With some of these data the data controller may have an overriding legitimate interest in terms of **vehicle and product safety**
- A combination of data can lead to data protection relevance.

**Figure 25 Data categories in connected vehicles. (Source: VDA, 2014)**

The second position paper titled 'access to the vehicle and vehicle generated data' (VDA, 2016) discusses data-centric requirements for security, privacy, and discrimination free innovation. According to this report, each OEM has the role of a system administrator and is hence responsible for the safe and secure transfer of car data to a business to business (B2B) OEM interface. Third parties can access this car data directly over the OEM B2B interface or via neutral servers, which gather the data from the cars.

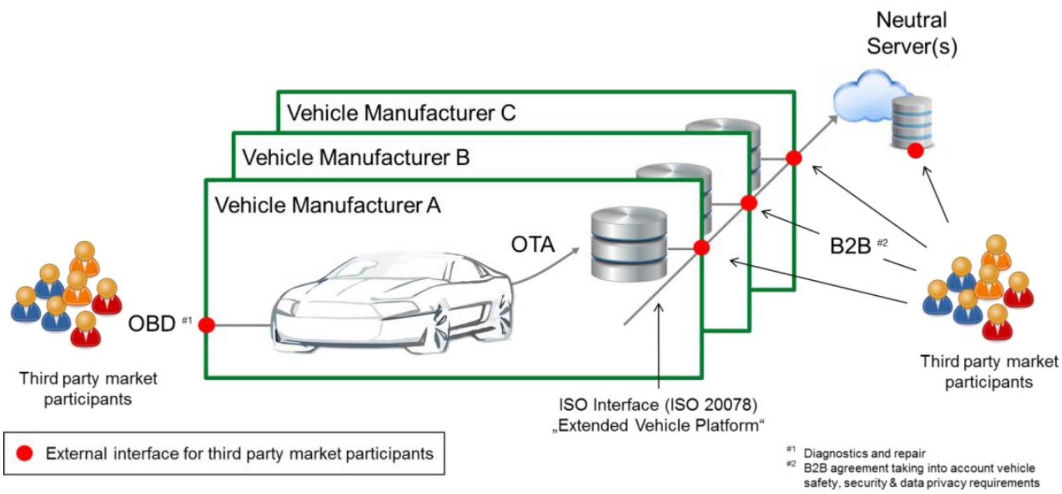


**Figure 26 Access to vehicle generated data. (Source: VDA, 2016)**

According the information contained in the position paper, direct access to this data for third parties will be disabled. As a result, access to car data may be very limited in the future for start-ups, because OEMs want to increase their influence on what stakeholders can do what with the data cars generate.

## 4.3 Conclusion and Outlook

Modern cars have become data generators. Hence, their data can be collected, stored in databases, analyzed, and finally aggregated to generate new products, digital services, and business models. In analogy to the quantified self phenomenon which is about capturing the data about oneself to provide new insights to people's behavior, the authors have coined the trend described above with the term quantified vehicles (Stocker et al., 2017a).

The authors expect that many stakeholders have an interest in exploiting car data to provide digital services. There are certainly a lot of benefits to achieve if this is done accordingly, which will increase individual driving safety as well as road safety. Furthermore many of these activities will have a direct influence on the environment as safer driving through less speeding, fewer braking, and smoother accelerating will positively correlate with reducing emissions while driving.

Two concrete stakeholder groups were focused, US ICT start-ups and German car manufacturers. While ICT start-ups adopt either an OBD-II interface plug or smartphone sensory to capture data, car manufacturers would have a direct access to car data – at least from a technical perspective. US ICT start-ups have occupied the market in a new freshness by creating new services and business models based upon analyzing large quantities of car data. They have built up an enormous expertise in gathering and exploring field data, detecting patterns and events in the data or providing analyses which are of interest to drivers. However, recent articles and reports from VDA suggest that German car manufacturers have become aware of the huge market which is at loose to the ICT industry. Hence, OEMs start to advance own projects and discuss restricting the OBD-II interface. Car manufacturers are seeking new opportunities and may establish a data market for third party services. If German car manufacturers will pursue the approach described in both VDA reports, then the battle on setting up a successful car data-service-ecosystem will progress to a very exciting next round.

# 5. Digital Vehicle Ecosystems and New Business Models: An Overview of Digitalization Perspectives [11]

| Summary and Author Contribution |
| --- |
| **Analysis of the Market: Services, Start-ups, OEMs, Business Models, and Trends (Paper 2/2)**<br>This chapter investigates new business models created within the automotive industry based on vehicle data-driven services, causing an interesting power struggle between the various stakeholders. A bunch of innovation-friendly IT start-ups – the majority of them from outside Europe – has already put energy into the development of novel quantified vehicle services for various beneficiaries, including drivers as well as third parties, challenging the traditional role of vehicle manufacturers. They gather data on how vehicles are used and offer digital products and services exploiting this data. This chapter provides a short overview on the role of the digitalization phenomenon in general, the impact of digitalization in the automotive domain through quantified vehicle start-ups and new business models, as well as a brief investigation of the position of vehicle manufacturers and their digital service strategies – all of them concluded in a comparison of value creation for business model elements.<br><br>This paper was written at the same time as the paper for the NBM conference (Kaiser et al., 2017b), which makes them partly related in content. In contrast, with this contribution, a conference in the field of knowledge / information was addressed, and the focus was adopted accordingly. I contributed to this article as the main author by incorporating my knowledge gained from the lightweight online market research approach, the discussions with the co-authors and members of research project AEGIS. We aimed to illustrate various developments (digitalization in the automotive domain, the emerging topic of *Data-driven Services*, and new business models) that all indicate that *Data-driven Services* will be a very relevant topic in the future. Finally, we also compared the business models of the *Data-driven Services* of vehicle manufacturers with those of start-ups. |

Digitalization is an important driver of service and business model innovation in the vehicle domain. Digitalization challenges are currently often subsumed under the more popular term ‘connected vehicles’. Taking a look on digital services and mobile applications, it emerges that connecting vehicles closer to their human drivers is an increasingly requested topic of research. Such include e.g. the detection of behavioral patterns from data collected during

[11] The content of this chapter is based on

**Kaiser, C.**; Stocker, A.; Viscusi, G.; (2017). Digital Vehicle Ecosystems and New Business Models: An Overview of Digitalization Perspectives. Platform Economy & Business Models workshop at i-KNOW ’17. Link: http://ceur-ws.org/Vol-2025/paper_ddds_1.pdf self-tracking activities (e.g. conclude from heart bpm value and GPS positions if a sport activity is performed), which can be transferred to vehicles, too. As vehicles capture sensory data about themselves and their environment during operation and thereby reveal how they are used by a driver, they can become ‘Quantified Vehicles’ (Shipilov, 2016). Quantified vehicles represent one key result of digitalization in the automotive industry, where incumbents have to face a key set of competing systemic challenges spanning from innovation capability, focus, and governance. In particular, digitalization has enabled digital entrepreneurship (Nambisan, 2016) providing less bounded entrepreneurial processes and outcomes and less predefinition in entrepreneurial agency, as shown also by the rising number of start-ups focusing on quantified vehicles (Kaiser et al., 2017b; Stocker et al., 2017a).

Considering the development process of successful business models based on exploiting vehicle operation data, there is currently a friendly competition between the major players from Information and Communication Technology (ICT) industry against the vehicle manufacturers on the supremacy of digital ecosystems. ICT start-ups have already successfully transferred the *quantified-self* (Neff and Nafus, 2016; Swan, 2013) phenomenon to vehicles and have launched apps and services to generate a whole new market, while vehicle manufacturers are currently in the transition process from vehicle manufacturers towards integrated mobility and data service providers (Kaiser et al., 2017b). Thus, the digital ecosystem for the incumbents and start-ups, which are competing in the quantified vehicles market, have to consider industry transformation in general, impacting on the overall digital information assets available as well as on the sustainability of business models.

This chapter aims to explore different approaches on developing platform oriented and sustainable business models for the occupation of digital services, probably the most lucrative components in the future automotive industry, between the incumbent vehicle manufacturers and the emerging tech start-ups. In particular, this chapter aims to identify the main actors and roles in emerging digital ecosystems of vehicle usage data platforms and their contribution to eventually define an information infrastructure for an automotive industry centered around ‘quantified vehicles’.

After this introduction, the chapter will provide a summary on digitalization as well as its implications for the automotive domain ranging in Section 5.1, whereas Section 5.2 outlines the applied research approach. Section 5.3 discusses relevant digital challenges of vehicle manufacturers and the increased competition on the way to digital services including recent activities within the startup ecosystem. Section 5.4 concludes with an outlook and a classification of value drivers and business model elements for each stakeholder investigated in this chapter.

## 5.1 Background and Motivations on Digitalization

Digitalization is a sociotechnical process that leverages the technical process of the encoding of analog information in a digital format (digitizing) applied to broader social and institutional contexts, transforming their sociotechnical structures, thus rendering digital technologies infrastructural (Tilson et al., 2010; Yoo et al., 2010). Furthermore, digitalization depends and results in digital artefacts characterized by attributes such as *editability*, *interactivity*, *reprogrammability / openness*, *distributedness* (Kallinikos et al., 2013), also implying a shift in product design moving from modularity to generativity (Lyytinen et al., 2016; Yoo et al., 2012; Yoo, 2013).

Quantified vehicles represent one of the key results of digitalization in the automotive industry, where incumbents have to face a key set of competing concerns systematically interrelated as shown by Svahn et al. (2017) through the case of Volvo: *innovation capability* (existing versus requisite), *innovation focus* (product versus process), *innovation collaboration* (internal versus external), and *innovation governance* (control versus flexibility). In particular, it is worth noting that from a strategy perspective, the digitalization enforces *internetworking* to be considered as "those business processes/activities conducted or mediated online by and between employees, customers, suppliers and partners of firms, using internet-based technologies accessed through internet-based infrastructures." (Brews and Tucci, 2007, p.224). Besides incumbent's digitalization enabled by digital entrepreneurship providing less bounded entrepreneurial processes and outcomes and less predefinition in entrepreneurial agency (Nambisan, 2016), as shown also in the case of the specific industry, the authors consider a rising number of start-ups focusing on quantified vehicles.

Taking these issues into account, the role of information and its value are a central challenge in the competitive scenarios emerging from digitalization as well. Consequently, a key issue is related to evaluation of the digital information asset of a company as well as its information capacity defined as the current stock of understandings informed by a given installed base (Viscusi and Batini, 2014). The key questions are:

*What data do I have?*

*How is the quality of data?*

*Can I exploit it in their full potential?*

*What can I infer from it given my current systems?*

According to Viscusi & Batini (2014) information capacity represents the potential of a digital information asset defined and evaluated independently from its usage, determining not only its economic value, but also the enabling capabilities.

Considering the development process of successful business models based on exploiting vehicle data, there is a competition between the major players from ICT industry against the vehicle manufacturers on the supremacy of digital ecosystems. ICT start-ups have already started to apply the quantified-self phenomenon to vehicles and have and launched apps and services to generate a new market, while vehicle manufacturers are currently in the transition process from vehicle manufacturers towards integrated mobility and data service providers (Kaiser et al., 2017b; Stocker et al., 2017a). Yet, the digital ecosystem for the incumbents and start-ups competing in the quantified vehicles market have to consider another industry transformation, impacting on the overall digital information asset available as well as the sustainability of business models.

At the state of the art, the need to build on an appropriate ICT infrastructure, the open system integration of the energy landscape led to its definition as “Internet of Energy” (Appelrath et al., 2012; Karnouskos and Terzidis, 2007). As a consequence of the “Internet” metaphor, the energy related challenges and concerns have been alternatively interpreted as a consequence of lack of information “to enable and motivate economic and behaviorally driven solutions” (Watson et al., 2010). Consequently, “energy informatics” (EI) has emerged as a new field within information systems (IS) research to analyze, design, and implement systems increasing the efficiency of energy demand and supply infrastructure (Watson et al., 2010). According to Dedrick (2010) researchers have framed the impacts of IT on the environment as first, second, and third-order effects:

- *First-order effects*: direct impacts from IT hardware during the product lifecycle, including production, use and disposal of computer equipment.
- *Second-order effects*: impacts of ICTs on other processes such as transportation or industrial production, influencing their environmental impact.
- *Third-order effects*: are longer term and more dynamic impacts, occurring when widespread use of ICTs leads to changes in lifestyles and economic structures.

Third order effects are relevant when considering the increased use of social media transformative potential for green IS on the demand side, encouraging better practices reducing the burden on the environment such as, e.g., the emerging carpooling and ridesharing applications impact on transportation coupled with the Internet of things (Malhotra et al., 2013). According to Malhotra et al. (2013) this two- way, sensor-driven communication is blurring the boundaries between the production side and the demand side. Furthermore, EI and IS and interdisciplinary strategies for quantified vehicles may provide models to assess the value of information, in particular the social value of related open data, adopting classification frameworks such as the one proposed by Viscusi et al. (2014).

Over the last twenty years, actually after the massive access to the Internet and the World Wide Web, the interest in the strategy concept of business model has grown, thus becoming a key element for competing in markets characterized by extensive use of ICTs and currently transformed by digitalization. In general terms a business model describes the rationale of how an organization creates, delivers, and captures value (Osterwalder and Pigneur, 2010; Boons and Lüdeke-Freund, 2013).

However business model is a multifaceted concept, still raising debate in academia as to its definition, Massa et al. (Massa et al., 2017) provided a systematic view on the different perspectives pointing out that business models can be considered as i) attributes of real firms (*how firms do business*), ii) cognitive/linguistic schema (*how the way firms do business is interpreted by organizational members*) and iii) formal conceptual representations/descriptions of the former two issues. As for formal/ conceptual representations/descriptions, Al-Debei & Avison (2010) identified four key dimensions of a business model: *value proposition, value architecture, value network, value finance*. Whereas as for business models as attributes of real firms it is worth mentioning, especially for the case the authors consider of automotive industry and quantified vehicles, the definition by Zott and Amit (2010), who conceptualize a business model as "*a system of interdependent activities that transcends the focal firm and spans its boundaries. The activity system enables the firm, in concert with its partners, to create value and also to appropriate a share of that value*" (Zott and Amit, 2010, p.216). This definition is particularly useful for understanding business models of companies interested in quantified vehicles when linked to the above concept of internetworking and current pervasiveness and strategic relevance of digital platforms (Eisenmann et al., 2006; Eisenmann et al., 2011; Parker et al., 2016; Yoo et al., 2012).

As argued by Nambisan (Nambisan, 2013, p.217), IT can act as either an *operand resource* "(often tangible and static) that an actor acts on to obtain support for executing a task", or as an *operant resource*" (often intangible and dynamic) that act on other resources to produce effects". Accordingly, in a case a digital platform can be considered an enabler of innovation processes and outcomes; whereas, in the other case, it acts as a trigger, informing rather than being informed by the users. Considering the automotive sectors and especially quantified vehicles, despite the "analysing" stance of the main market players, a set of traditional and new business models can be applied (Shipilov, 2016), in particular the infomediary one (Afuah and Tucci, 2000; Rappa, 2001) can be adopted under a utility perspective (Rappa, 2004) and extended from data collection for, e.g., marketing purpose to data useful for social value, as capability and functionings they enable (Viscusi and Batini, 2016; Viscusi et al., 2014), and finally for sustainability issues. Besides environmental and societal issues, business sustainability refers to "business models and managerial decisions that create value over the short, medium, and long terms, based on mutually beneficial interactions between the company's value chain and the social and environmental systems on which it depends" (Lüdeke-Freund et al., 2016, p.18). Furthermore, according to Boons & Lüdeke-Freund (2013, p.14) a business model perspective may contribute to a sustainable innovation agenda to overcome internal and external barriers. Also, Schaltegger et al. (2016, p.6) points out that a business model for sustainability "helps describing, analyzing, managing, and communicating (i) a company's sustainable value proposition to its customers, and all other stakeholders, (ii) how it creates and delivers this value, (iii) and how it captures economic value while maintaining or regenerating natural, social, and economic capital beyond its organizational boundaries." Thus, considering the automotive industry and the potential transformation enforced by digital innovation and quantified vehicles, business model innovation can open different horizons and path impacting the sustainability of companies, although how they can innovate their business models toward greater sustainability still need to be significantly addressed in research and practice (Foss and Saebi, 2016, p.221). Still, business model innovation in automotive industry asks for understanding the different ways the various actors can follow to innovate their business models; in particular, as pointed out by Massa & Tucci (2014, p.424), *business model design* in newly formed organizations, which refers to their "entrepreneurial activity of creating, implementing and validating a business model", and *business model reconfiguration* in incumbent firms, encompassing the reconfiguration and eventual acquisition of organizational resources to change an existing business model.

## 5.2 Research Approach

This chapter is aimed to provide an overview of innovative ICT start-ups towards establishing new services and sustainable business models. Besides that, selected digitalization initiatives of German and Italian vehicle manufactures are analyzed. The authors conducted a lightweight online market research approach analyzing information available on the Web. They used a combination of the terms car, vehicle, connected, quantified, start-up and vehicle in search engines to capture the current developments in the start-up domain.

After having identified major quantified vehicle start-ups, two authors studied the start-up websites in detail to find out more about their visions and goals as well as about their business models, products and services. Two out of the three authors reviewed the websites of all start-ups and discussed their knowledge with the third author afterwards to come to a common understanding. The information was then validated in discussions within the AEGIS project consortium as well (AEGIS, 2017). Furthermore the authors used crunchbase.com to capture additional meta-information on business and funding, where the authors revealed interesting facts on investments into tech start-ups as well as of strategic partnerships.

As for the manufacturers for the countries considered, the authors carried out a market research on the websites of industry organizations, press news, and magazines, apart from policy documents from European Union organizations. They especially paid attention to review the current digital services of vehicle manufacturers by taking into account the content on the various product websites, which is expected to be up to date.

## 5.3 Results: Digitalization in the Automotive Industry

### 5.3.1 Overview

Digitally-enhanced driving is an emerging topic, which can be counted to the heavily used umbrella terms ‘connected cars’/’connected vehicles’ or ‘connected driving’. According to a definition from PWC (Viereckl et al., 2016) in their 2016 connected car report, connected vehicles are defined as vehicles that have access to the Internet and a variety of sensors, and that are thus able to send and receive signals, sense the physical environment around them, and interact with other vehicles or entities. According to this report, revenues in the automotive industry (will) shift from hardware to software, from products to services, and from old to new economy. The report highlights four trends changing the automotive competition:

- *radically new technology at low price* (increase in vehicle to infrastructure connectivity (e.g. through 5G), increase in computing speed to operate artificial intelligence for self-driving, evolution of low cost sensors to make a vehicle aware of its surroundings),
- *shorter innovation cycles by new high-tech entrants* (non-traditional tech companies to offer new services as add-ons to automobiles and disrupt traditional vehicle value chain, Apple to invest $10 billion into an iCar, Google's self-driving vehicles to drive more than 1,5 million miles, data-centric business models depend on revenues from ongoing services and the sale of information),
- *new mobility concepts and increasingly urban customers* (urban residents to lose interest in owning vehicles, millennials in cities to face affordability issues, movements towards vehicle sharing and ride sharing, expectations of highly sophisticated levels of connectivity and services), and
- *evolving regulatory and policy constraints* (policy and regulations to lack behind the technological process, expect regulators to respond with laws ensuring the safety of driverless vehicles, cities to discourage the use of non-electric private vehicles).

The PWC report further differentiates three main categories for technologies and services for connected vehicles:

- Consumer services such as internet and cloud based digital services that add to driving experiences,
- Connected vehicle packages with feature to improve or help managing the vehicle's operation, and
- Supply-side technologies as underlying systems connecting the vehicle to the wider world.

Another report from McKinsey (Habeck et al., 2014) on the connected vehicle trend estimates a global market size of 170 billion EUR to 180 billion EUR for vehicle connectivity in 2020. This report foresees connectivity to trigger a redistribution of automotive revenue pools except vehicle operations – based on a survey of 2000 vehicle buyers. McKinsey perceives human-machine-interface (larger screens, multiple screens, innovative UIs, Augmented Reality), vehicle condition data (for offering maintenance and insurance services), and dynamic real-time geo-information (oligopoly of TomTom, Here, and Google) to become key control points in this redistribution. This report has identified five different groups of vehicle buyers:

maxed-out vehicle enthusiasts, integration and entertainment lovers, safe and secure navigators, purists/minimalists and price conscious traditionalists, each group with own preferences and attitudes towards connectivity-related features. Along with the preliminary market analysis a 1st competition analysis is provided. According to the McKinsey report, the connected vehicle ecosystem of the future will be highly influenced by additional players including digital players, telecom players and insurers.

- Vehicle manufacturers explore ways to exploit the connected vehicle into the provisioning of a software operating system to serve as a platform for a potential app store, as well as into the development of specific apps and services.
- Automotive suppliers want to establish direct relationships with the end customers of vehicles produced by vehicle manufacturers who they supply.
- Digital players adapt their smartphone platforms to vehicle-specific customer needs and to integrate their infotainment into vehicle systems.
- Telecom players see new opportunities in terms of infrastructures, while SIM cards are installed in vehicles.
- Insurers expect new opportunities while e.g. offering telematics-based coverage options.

The authors can expect a huge power struggle between all players – including vehicle manufacturers – on who will reap most value from the connected vehicle market. According to the problem statement of the AutoMat project (Automat, 2017) coordinated by Volkswagen Research, the automotive industry has not yet been able to successfully establish an ecosystem for smart driving applications equivalent to that of smartphone manufacturers (Kaiser et al., 2017b). In its prior problem statement, the AutoMat project mentions three reasons why vehicle manufacturers are currently struggling: (i) Brand-specific business approaches dominate, and as a consequence there is a lack of brand-independent vehicle lifecycle data, (ii) current proprietary vehicle services focus on the individual customer, which results in privacy concerns, and few ideas exist how anonymized vehicle data can be used to establish other services, and (iii) the implied or required collaboration between vehicle manufacturers on vehicle data and services is considered risky in terms of competition.

### 5.3.2 The Business Models of US Start-ups

The IT industry of the USA has already lined up a series of tech start-ups backed by risk capital, reaching more than 20 million USD in some cases, demonstrating how high investors perceive the market value of an ecosystem built on exploiting vehicle data (Kaiser et al., 2017b). The majority of start-ups including automatic.com, automile.com, dash.by, moj.io, vin.li, zubie.com to name a few capture vehicle operation data through the On-board diagnostics (OBD) interface of the vehicle which is originally intended to provide a repair technician access to the status of the various vehicle subsystems and diagnostic information.

These connected vehicle start-ups have specialized in capturing, storing, and analyzing large quantities of vehicle operation data and offering digital services in smartphone applications to motivate the driver for sharing valuable driving data. The majority of start-ups are currently capable of automatically extracting interesting driving events using computational intelligence including e.g. hard brakes, hard accelerations or speeding to name a few. These events are hidden in the gathered vehicle operation data (time series data) and have to be revealed through applying big data analytics. Mobile applications running on the driver's smartphone then pull the results from the start-up's datacenter and then visualize them on the driver's smartphone or tablet. The majority of these start-ups provide mobile apps connected to the OBD interface of the vehicle via Bluetooth with very similar functionality to the driver: These include for instance means to drive smarter by unlocking diagnostics and real-time data and greener by using the app for gaining an overview on how driving habits influence fuel consumptions (dash.by), make the vehicle smarter by revealing insights of vehicle data and providing driving stats (automatic.com), or delivering real time location, trip history, maintenance alerts, engine diagnostics and driving insights (zubie.com).

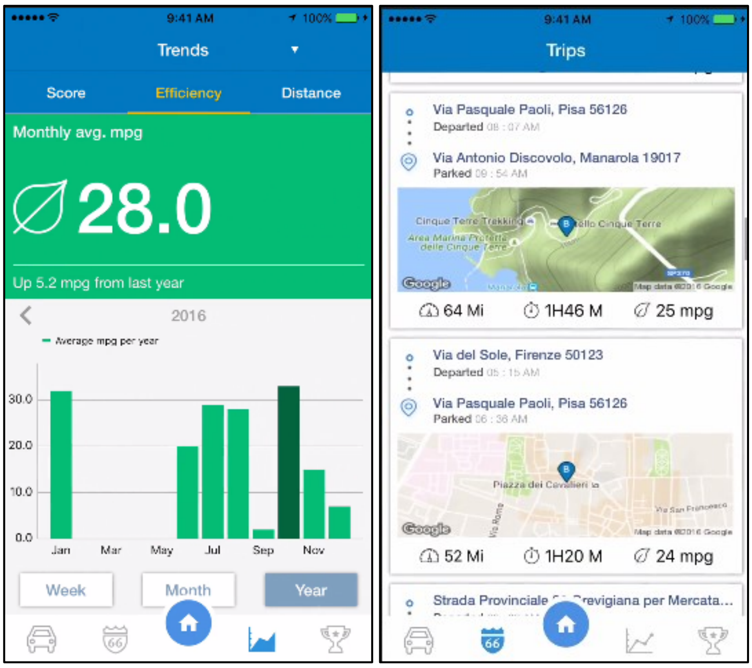


**Figure 27** **Exemplary snapshots of dash.by's mobile app. (Kaiser et al., 2017a)**

Another interesting tech-startup recently receiving amongst others a huge investment from BMW i Ventures is Zendrive.com (BMW Group, 2014). Zendrive is taking advantage of the sensors built into modern smartphones to capture smartphone sensor data while driving, and provide cloud-enabled driving analytics aiming at safer drivers as well as on safer roads by using gamification features. According to the business information platform crunch-base.com a bulk of investors of these start-ups stems from the insurance industry, too.

### 5.3.3 Digital Services of Vehicle Manufacturers

German vehicle manufacturers including AUDI, BMW, DAIMLER, and VOLKSWAGEN currently offer some digital/connected services. These services allow for instance accessing some vehicle functions through the drivers' smartphone via mobile apps (e.g. lock/unlock the vehicle), vehicles conducting (semi-) automatic calls for emergencies in case of a detected accident, roadside assistance, or allowing (stolen) vehicle location via mobile apps to name a few.

Their offer surely is expected to increase, taking also into account new strategic partnerships as well as investments into connected vehicle start-ups. For instance BMW i Ventures (BMW, 2017a) heavily invests into tech start-ups aiming to facilitate safer driving including e.g. Zendrive.com (providing smartphone-powered driving analytics including statistics and gamification), or Nauto.com (multi-sensor device to monitor the driving behavior including statistics especially for safety relevant events. Moreover BMW (BMW, 2017b) has recently teamed up with IBM in its activity to deploy the vehicle data platform and to enhance it with analytics features (IBM, 2017).

The following subsubsections summarize information on USP and the different services captured from the various product websites of German vehicle manufacturers and published studies like Karlsson et al. (2016):

**AUDI**

"*Innovative services and functions that connect drivers with their Audi and the world: that is Audi connect. myAudi and the Audi connect services make driving even more relaxing and safer*." Source: (Audi, 2017a)

"The term 'Audi connect' covers all applications and developments that connect Audi vehicles to their drivers, the internet, transportation infrastructure and other vehicles. Audi is continually building up its lineup of products and services in this technical area – with new solutions such as the Audi connect SIM and the traffic light information service for the US market." (Audi, 2017b)

**BMW**

"*BMW Connected is a personal mobility assistant which facilitates everyday mobility and assists drivers in reaching their destinations relaxed and on time. Mobility-relevant information such as recommendations for optimal departure times are available remotely via smartphone or smartwatch and can be seamlessly transferred into the vehicle.*"

BMW ConnectedDrive (BMW, 2017c):

- Remote Services: Locking and unlocking the vehicle, indicate the vehicle's location by honking the horn or flashing the lights, or on a map in the app. Activate the vehicle's climate control immediately or on schedule.
- Concierge Services: Select travel destinations and get information, connect with call-center agents to look for nearest services or to book services, addresses sent directly to navigation system.
- Real Time Traffic Information: Information about the current traffic situation, calculate expected delays and recommend detours, on street parking information.
- Intelligent Emergency Call: If an airbag is deployed, the BMW Call Centre is contacted via an accident-proof telephone unit permanently installed in the vehicle, precise position of vehicle is communicated including relevant accidental data.
- Digital Services: With BMW CarData a vehicle owner can view the key vehicle data and share them with third parties if required.

**DAIMLER**

"Mercedes me is your package of innovative services, products and lifestyle offers from Mercedes-Benz, Daimler and our cooperation partners – including access to your vehicle via smartphone, of course."

Mercedes me' connect services (Mercedes, 2017a; Mercedes, 2017b):

- Vehicle Setup: Remote Retrieval of Vehicle Status, Remote Door Locking and Unlocking, Programming of Charge Settings and Pre-Entry Climate Control, Personalization.
- Vehicle Monitoring: Geofencing, Vehicle Tracker, Parked Vehicle Locator, Route Planning for plug-in hybrid vehicles.

- Parking using a smartphone app: Geofencing, Vehicle Tracker, Parked Vehicle Locator, Route Planning for plug-in hybrid vehicles.

**VOLKSWAGEN**

"VW Car-Net® makes your Volkswagen more like a friend. It gives advice, helps you along the way, entertains you, and watches out for you. It connects you to the world outside all from the comfort of your driver's seat. VW Car-Net is your partner in drive."

VW Car-Net (Volkswagen, 2017):

- Via app-connected drivers can access select smartphone apps right from their dash.
- Guide & Inform Services via SiriusXM® Traffic subscription and SiriusXM® Travel Link

**Fiat Chrysler Automobiles (FCA)**

In order to not exclusively look at German vehicle manufacturers, at the glance some of the main initiatives by the Italian-controlled multinational corporation Fiat Chrysler Automobiles (FCA), oriented towards consumers and dealers are discussed in the following. As for consumers, it is worth mentioning the Uconnect® navigation, entertainment (with CarPlay (Apple, 2017) to use iPhone while driving by putting his applications and functions on the vehicle's built-in display) and communication system that allows drivers to being connected while driving and paying attention to the road and related events (FCA, 2017a). Moreover, apart from CarPlay and iPhone, FCA is collaborating with Google to integrate the Android open-source platform with the Uconnect 8.4-inch connected system (Audi, 2017a).

Considering now dealers, The FCA Dealer Digital Programme is a collection of tools, process, and support channels aiming at coordinating the action by the different partners and acting in the digital space in order to enable the engagement of local in-market shoppers to increase selling (FCA, 2017b). Furthermore, the FCA Dealer Digital Websites are the only sites having traffic directly from the brand website's dealer locator, thus increasing through this connection their visibility in search results on the main search engines such as e.g. Google also through the support of FCA Dealer Digital Advertising (DDA) Programme (FCA, 2017b), a one-stop shop for digital advertising campaigns connecting dealers to Certified Providers with national coverage and extensive digital Automotive industry expertise helping dealers in digital marketing and sales. This connection to the brand websites is aimed not only to increase marketing and selling activities, but also to improve integration with FCA in terms of timely updates of assets, campaigns, pricing and inventory. However, this integration, each Dealer Digital Website site, can be customized according to the value proposition of the dealer brand (FCA, 2017b).

The FCA effort in digitization of customers' experience and dealers services represents the basis for moving from connected to self-driving vehicles as shown by the partnership with Google and the announcement in the spring of 2016, that they would build 100 self-driving Chrysler Pacifica hybrids minivans, formerly tested in Arizona, California and Michigan. Furthermore, Google's self-driving vehicle project, Waymo, has announced in April 2017 a program on larger scale in Phoenix program using the FCA 500 Pacifica minivans, allowing hundreds of people in Phoenix applying on Waymo's website to ride in the vehicles in order to get feedback on the experience. (Associated Press Fiat Chrysler, 2017)

## 5.4 Conclusion

Digitalization is an unstoppable trend in the automotive industry in general and increasingly observable by the driver. In modern passenger vehicles, drivers can connect to the cloud, where services to drivers and other stakeholders are provided. Thereby three approaches have been discovered:

- Brand dependent assistance services, which provide access to vehicle functions and services via smartphone. Users thereby get access to vehicle functions via apps.
- Brand-independent apps and services, often as components of data ecosystems with several stakeholders, which provide transparency on driving data to be used e.g. in driving behavior analytics.
- Strategic alliances of vehicle manufacturers with ICT firms (e.g. BMW teams up with IBM) to establish services & business models on how to make value out of vehicle data.

As pointed out by Zott & Amit (2017) digitalization is strictly connected to product innovation and its acting both at business and customers/users side asks for new ideas and business models design in the case of start-ups and/or *reconfiguration* for vehicle manufacturer (Massa and Tucci, 2014). Following the perspective by Amit and Zott (2001; 2010; 2012) of business model as an activity system that defines the way a company does business, whose elements are *content* (the '*What*?'), *structure* (the '*How*?'), and *governance* (the '*Who*?') and its value drivers are *novelty*, *lock-in*, *complementarities*, and *efficiency*,

Table 11 identifies business model's elements and value drivers for start-ups and vehicle manufacturers (termed OEM for Original Equipment Manufacturers, in gray) discussed in previous sections. Vehicle manufacturers seem more oriented towards governance and the exploitation of complementarities as value drivers (thus confirming a platform orientation as business model (Parker et al., 2016)), with Volkswagen targeting novelty for customers through business model innovation at content level and FCA using it for efficiency and lock-in at dealers level. As to this issue, the resulting ecosystems show a relevance of digital (entrants) players as partners and a focus of FCA on the inclusion of dealers as a key part of its value constellation (Normann and Ramìrez, 1994) rather than value chain, through internetworking. As for the start-ups, the ones considered appears to focus on 'structure' as business model element by mostly targeting (quite surprisingly) efficiency as value driver instead of novelty, thus having an execution orientation rather than a differentiation one to digital business (Viscusi, 2015).

**Table 11 Business model elements and source of value creation for vehicle manufacturers and start-ups in automotive. (Kaiser et al., 2017a)**

| Company / Initiative | Type | Value drivers | Business model element |
|---|---|---|---|
| Audi / Audi connect | OEM | Complementarities | Governance |
| Automatic | Start-up | Efficiency | Structure |
| Automile | Start-up | Efficiency | Governance |
| BMW / BMW Connected | OEM | Complementarities | Governance |
| Dash | Start-up | Efficiency | Structure |
| FCA / Uconnect | OEM | Complementarities | Governance |
| FCA / Dealer Digital Programme/ Digital Websites/Dealer Digital Advertising | OEM | Efficiency / Lock-in | Content |
| Mercedes Benz-Daimler / Mercedes Me | OEM | Complementarities / Novelty | Governance |
| Metromile | Start-up | Efficiency | Structure |
| Mojio | Start-up | Complementarities | Structure |
| Vinli | Start-up | Complementarities | Structure |
| Volkswagen / VW Car-Net | OEM | Novelty | Content |
| Zendrive | Start-up | Efficiency | Content |
| Zubie | Start-up | Efficiency / Complementarities | Structure |

# 6. A Research Agenda for Vehicle Information Systems [12]

**Summary and Author Contribution**

**Definition of a Research Agenda for the Information Systems Community (Paper 1/1)**

This chapter introduces vehicle information systems (Vehicle IS) as a new class of information systems (IS). Vehicle IS are enabled through the data generated by a plethora of different sensors within modern vehicles, meshed up with data from a variety of different other sources. Expecting the awareness on and the needs for Vehicle IS to steadily increase in the future, this chapter investigates existing literature on Vehicle IS published by the academic IS community. A definition of the term 'vehicle information system' and an overview of relevant research directions with a set of example research questions is provided, to assist the academic IS community to advance the state-of-the-art in designing Vehicle IS.

The aim of this paper was to present my impressions, that *Data-driven Services* are becoming increasingly important and that there is a need for research, in one of the top conferences of the IS community in order to increase the international awareness and reach of our work. For this purpose, I administered the developments (e.g. definition of Vehicle IS, finding research directions and example questions) and developed content myself (e.g. literature review, in-depth interviews, certain research questions, summary). With the example research questions we also wanted to show other researchers from the IS community that application-oriented *Data-driven Services* still contain many aspects on which there is no IS research yet, to advance the state-of-the-art in designing Vehicle IS in the long run.

According to numerous studies and reports published by business analysts from Gartner (Ramsey, 2017), IBM (IBM CAI, 2015), and McKinsey (Gao et al., 2016) digitalization is an important driver of service and business model innovation in the automotive domain. Modern passenger vehicles have slowly become 'computers on four wheels' equipped with a plethora of different types of sensors, generating and utilizing enormous amounts of data (Haeberle et al., 2015). While currently exclusively utilized for vehicle functionality and safety, the continuous collection of such vehicle data can facilitate the generation of novel IS for vehicle drivers and other stakeholders, even from beyond the automotive domain including e.g. insurers, meteorologists, or city planners. Due to another practitioners' report, future connected vehicles will heavily interact with ecosystems of automotive data (that may

[12] The content of this chapter is based on

**Kaiser, C.**; Stocker, A.; Festl, A.; Lechner, G.; Fellmann, M.; (2018). A research agenda for vehicle information systems. In *Proceedings of European Conference on Information Systems (ECIS 2018).*

be generated by other road users' digital devices), with the digital infrastructure, and also with the data provided by other services (Strategy Engineers and fka, 2017).

Obviously, current up-to-date vehicles already capture manifold types and amounts of data about themselves and their environment and, and in this way, can even be coined *Quantified Vehicles*. This term is borrowed from the quantified-self movement sharing similar thoughts and then transferred to the vehicle domain (Stocker et al., 2017a). Quantifying a vehicle represents one interesting field of digitalization in the vehicle domain, to which many tech-startups have already dedicated huge investments (Kaiser et al., 2017b). Meshing-up the data a modern vehicle generates during its operation phase with environmental data, weather data, governmental data, or data from complementary businesses is a hot topic in the vehicle industry. It may pave the way for new IS and open the door to currently untapped potentials and opportunities, as a recent report from VDA (2016) outlines.

So, how can the IS community finally contribute to the digitalization of the automotive industry? Marabelli et al. (2017) state that the IS community has so far devoted relatively little attention to what they call *sensor-based technologies in the vehicle domain*. The preliminary investigation of related work within the Association for Information Systems Electronic Library (AISeL) has confirmed this statement and the comparably limited interest of the academic IS community in vehicle-related IS. Applying the search string *"vehicle information systems"* on abstracts from the AISeL retrieved just 3 scientific papers. Widening-up the scope of search by using the search string

*(vehicle OR automotive) AND information systems*

retrieves at least 67 publications, which still shows the limited awareness of the IS community in vehicle-related IS research directions. While the authors think the central topic *Vehicle Information System* has not yet fully arrived in the academic IS community, the authors nevertheless assume the IS community to contribute greatly to the digitalization of the automotive industry by providing appropriate models, methods and guidelines, and explaining or even predicting the behavior of the various road users involved.

Though this research-in-progress is not intended to be a systematic literature review, it nevertheless investigates at least existing research from and explores which topics have been addressed so far by the academic IS community on the state of the art of Vehicle IS. By using AISeL the authors ensure the integration of quality assured, peer-reviewed scientific knowledge but avoid sources tending to be biased, as for example industry-driven technology roadmaps. In a further step, the authors aim for a definition of the term *vehicle information system*. Then the authors propose a research agenda for directions, which the authors deem of importance to the IS community in an emerging field of Vehicle IS. These research directions have been identified as a result of in-depth interviews with six experts from different engineering domains (business informatics, business studies, data science, informatics, mechanical- and electrical engineering) employed at an Austrian research center, hosting several vehicle engineering disciplines, and three experts in management positions for product development from associated industrial partners, then synthesized and summarized by four authors and finally quality assured by the fifth author, a professor for IS.

This chapter is structured as follows: After this introduction and motivation, the authors continue in Section 6.1 with a definition of the term *vehicle information systems* and further report on the results of this preliminary literature investigation on Vehicle IS in the AIS Electronic Library. In Section 6.2, the authors provide a research agenda for Vehicle IS as an emerging class of information systems. This research agenda includes relevant research directions and example research questions to explore without putting them in a temporal order yet as it would be the case for a scientifically sound technology roadmap (Garcia and Bray 1997, Phaal et al. 2003). The chapter closes with a summary, a limitation of research, and an outlook of future work in Section 6.3.

## 6.1 Vehicle Information Systems (Vehicle IS)

### 6.1.1 Towards a Definition of Vehicle IS

It is the mission of the IS academic community to *"advance the knowledge and excellence in the study and profession of information systems"* (Association of Information Systems, 2017). The IS discipline has a more than 40-year history evolving through four eras with considerable diversity amongst its members in terms of research interests and believes what belongs and what does not belong into the field, spanning a wide variety of themes like decision support systems, organizational impact of IS, IS adoption, IS evaluation, or knowledge management to name a few (Hirschheim and Klein, 2012). According to Nunamaker and Briggs (2011), one major purpose of the IS discipline is to *"to understand and improve the ways people create value with information"*, while studying the *"understandings people require so they can create new value, and of the analysis, design, development, deployment, operation, and management of systems to inform these understandings"*.

While there has been a lot of attention within the IS community to investigate how IS – as an academic discipline – has evolved over time, there still seems to be no ultimate definition of what the tangible part of an information system (i.e. the system – not the discipline) factually is. Surprisingly, even many renowned scientific papers on IS dealing with the tangible part – the information system –neither review prior definitions of this term nor provide an own definition. For instance, both highly cited papers from Delone and McLean (1992 and 2003) on measuring the success of an information system within an organization introduce an IS success model without providing a sound definition on what the information system actually is. What seems to be a common practice in IS makes it more than challenging for us to provide a sound scientific definition of a *vehicle information system*.

However, some rather practical definitions in the literature have proven to be helpful: For example, from a Management Information Systems (MIS) perspective, Laudon and Laudon (2013) define an information system *"technically as a set of interrelated components that collect (or retrieve), process, store, and distribute information to support decision making and control in an organization"*. Another example is from Neumann et al. (2014) who define a business information system as a *"socio-technical system containing human beings and machines, which use and produce information to support and enable the processes and operations of an enterprise"*. Taking into account the cited work, the authors understand information systems as *socio-technical systems, supporting users to execute tasks by providing task-relevant information*. Information provision typically is supported by hardware and software capable to process digitized input efficiently (which increasingly is available due to digitalization) that in turn creates opportunities for increased automation or new business models.

Before going towards a definition of a "vehicle information system", the authors deem it important to distinguish Vehicle IS from a series of *vehicle automation systems* (Stanton and Young, 2010), including in particular automotive safety systems (e.g. anti-lock braking system – ABS, or electronic stability control – ESP) and advanced driver assistance systems – ADAS (e.g. adaptive cruise control – ACC, or lane assist), which directly influence the driving process increasing safety and/or comfort. Vehicle automation systems even influence vehicle dynamics while keeping the driver fully out of the loop. In contrast the factual interaction of the driver with the information processed by an information system is one fundamental property of Vehicle IS. Hence, the authors understand Vehicle IS as a class of *software* *applications processing vehicle data and/or other relevant data from different sources to finally provide valuable and action-relevant information to the vehicle driver and/or to other stakeholders*.

### 6.1.2 Scope and Examples of Vehicle IS

In order to describe the transformation and enrichment of (vehicle) data to enable Vehicle IS a data value chain can be applied. The Vehicle Data Value Chain adapted from Curry et al. (2016) shown in the figure below outlines the data process from vehicle data generation to usage in Vehicle IS.

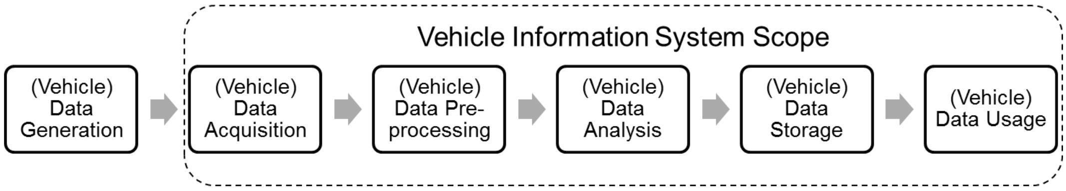


**Figure 28 Stages of a Vehicle Data Value Chain. (Kaiser et al., 2018b)**

A Vehicle IS is an IS providing information to users (e.g. the vehicle owner, the vehicle driver, or co-drivers who are granted the rights to access this information) during different phases of vehicle operation, most notable before a trip, during a trip, and/or after a trip has been completed. Depending on the particular vehicle operation phase, Vehicle IS can be used from inside a car (often referred to as *in-vehicle information systems*, e.g. by Ryder et al. (2016), Peng et al. (2014) and in a patent of Banski and Faenger (2017)) as well as from outside the car. Vehicle IS may directly use the dashboard of the vehicle, but may also extend the system border of the vehicle and establish a second dashboard (Stevens et al., 2017). The following visualization provides two examples of Vehicle IS for each operation phase.

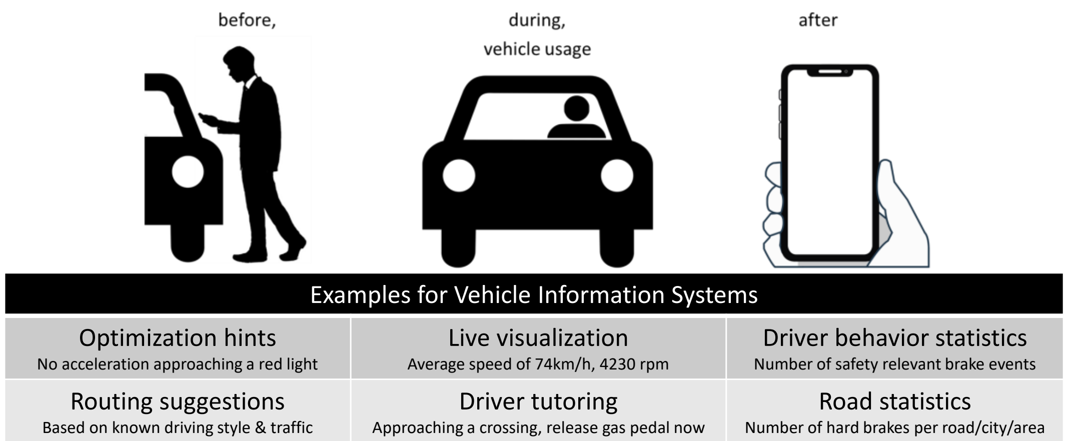


**Figure 29 Examples for Vehicle IS to be used before, during, and after a trip. (Kaiser et al., 2018b)**

## 6.2 A Research Agenda for Vehicle IS

Our preliminary literature investigation of the AIS Electronic Library has retrieved few relevant publications from the academic IS community related to Vehicle IS: Ryder et al. (2016) present an in-vehicle information system prototype for drivers providing warnings of upcoming accident hotspots based on data collected from service users. Nastjuk et al. (2016) investigate the impact of in-vehicle IS on perceived range stress (fear of a discharged battery). Brandt (2013) reviews the past, present and future of IS in automobiles, especially paying attention to IS linked to electric vehicles. Taking a wider viewpoint, Brandt provides a categorization for Vehicle IS in convenience, communication, and entertainment (CCE), vehicle monitoring, geo IS and navigation, and finally safety and collision avoidance. Kolbe et al. (2015) investigated the influence of technological and sociodemographic factors on perceived stress, resulting from human interaction with Vehicle IS. Wacker et al. (2014) investigated what information green IS should provide to the individual users of electric vehicles.

Our preliminary analysis has shown that research is scattered and diverse. In accordance to Rehm et al. (2017), the authors therefore argue that structuring research directions in three different domains, the technical domain (e.g. the technology enabling the Vehicle IS), the governance domain (e.g. a Vehicle IS has to be designed in accordance to legal and ethical standards), and the human domain (e.g. a Vehicle IS has to provide value to the human driver in order to be used) is a feasible approach when aiming towards a research agenda.

So, in which areas within these three domains can the academic IS community finally contribute to the design of Vehicle IS? To answer this question, the authors apply an approach similar to Yoo et al. (2010) and comparable to Abbasi et al. (2016) briefly describing relevant research directions and providing a set of example research questions after an introduction of the concept in scope. Those are presented in Table 12 at the end of this section. The identified research directions for Vehicle IS are classified in the proposed domain scheme from Rehm et al. (2017) as it can be seen in Figure 30. From this diagram it is apparent, that the authors are dealing with highly interdisciplinary research directions as all of them are part of at least two domains.

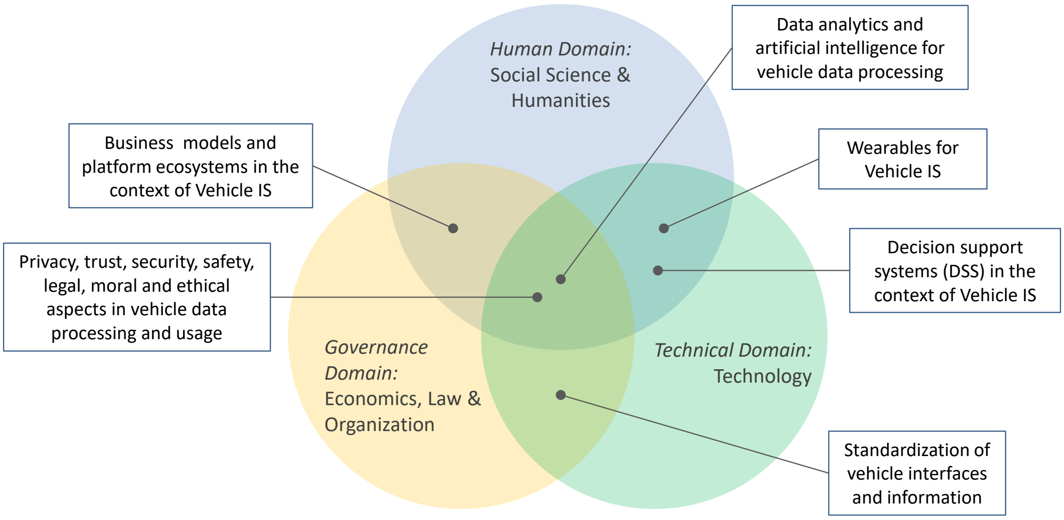


**Figure 30** **Characterizing research directions for Vehicle IS. (Kaiser et al., 2018b)**

### 6.2.1 Data Analytics and Artificial Intelligence for Vehicle Data Processing

Data is one key source for Vehicle IS – and data analytics is the key to leverage its value. The data generated by modern vehicles is of enormous size and often describes very volatile processes (AutoMat, 2017). It can thus not easily be interpreted by humans in raw form. Instead, it is necessary to either transform the data, i.e. to compute meaningful and interpretable properties of the data (e.g. fuel consumption) or to assess the driving metric of interest by statistically describing it with a (machine learning) model. The metric "aggressiveness of driving" may serve as an example for the latter case: It cannot easily be computed directly, but only by complex interactions of many other, simpler parameters (Toledo et al., 2008). In contrast, many of the parameters in-vehicle IS present to the driver are exactly computable by some formula derived from physical or chemical system properties and in most cases easily interpretable. Machine learning models, on the other hand, seem to be rarely used.

Algorithms to build ("train") such machine learning models are typically not directly leveraging the raw data, but use so called *features*, i.e. calculated properties of data which are related to the parameter of interest. For many applications, especially in the context of vehicles, it is not straightforward to choose the right set of features: If you want to model a property of the driver, you need to ensure that the chosen features do indeed capture those properties and non-properties of the vehicle or of environmental conditions as road type or traffic volume. While other disciplines, like image recognition have developed de-facto standard sets of robust features for different tasks (Lowe, 1999; Rosten and Drummond, 2006; Alcantarilla et al., 2012), comparative feature sets for Vehicle IS are still missing.

When machine learning models for Vehicle IS become more common, additional challenges will arise, especially ones related to safety and reliability. To ensure proper, correct and save behavior of the models, the development of suitable testing procedures will become inevitable. It is however uncertain if such procedures can be based on classical machine learning quality metrics which give information on one model at a time only. Suitable testing will need to take the possible interactions between multiple models into account, and, to make things even harder, a ground truth to compare the results against is often not available or hard to obtain.

### 6.2.2 Wearables for Vehicle IS

Besides data directly linked to vehicles, also appliances usually rather related to "quantified self" (QS) can act as data source and operating infrastructure for a Vehicle IS. Wearables, including smartwatches, fitness trackers, head-mounted displays, smart clothing or jewelry, or even implantables, deliver a huge variety of data with different quality levels. Although De Moya and Pallud (2017) denote QS as an immature domain of research, their literature review revealed that profound research was conducted in technological, health-related and, social domain. Applications based on wearables range from finger/hand gesture recognition using smartwatches (Xu et al., 2015) to emotion recognition systems via electrocardiography (Zhao et al., 2016) through to drowsiness detection of drivers (Warwick et al., 2015). Sun et al. (2017) explore challenges and future directions in the view of combined smart wearables and intelligent vehicles. The authors classify further research potential into *Communications and Services* like interference mitigation, *Security and Privacy*, *System-Level Considerations* in terms of, e.g. communication protocols or networks, and *Other Issues* covering for example dynamic channel modelling or power supply issues.

### 6.2.3 Privacy, Trust, Security, Safety, Legal, Moral and Ethical Aspects in Vehicle Data Processing and Usage

Whenever data from human behavior is captured and leveraged in IS, ethical aspects have to be discussed, as the data might be exploited for other purposes, too. For example, thousands of smartphone applications are available for “free”, if users agree terms and conditions, which include access to personal information like search terms, or even the ability to record sound from the microphone. As a result, users of free applications often pay indirectly through data they provide to the application, which then can be sold on the digital data market. At the moment, a majority of users does not care about privacy. However, this situation might change if their initial trust is destroyed.

Especially with behavior-revealing information including speeding, accelerations in kick-down-mode, or hard breakings, it is important that users can trust the IS to only use the data to improve their end user experience, e.g. to defuse dangerous crossroads instead of exploiting it for other purposes like automating the detection of speeders for the police.

Consequently, regarding vehicle data and Vehicle IS, privacy and trust are related to each other, while security has to ensure that data and information is kept within defined boundaries, e.g. that no intruder can get access to a vehicle. This is safety relevant as well, as there will be services which also write data to the vehicle interface, and thereby possibly – depending on the setup – can have a negative impact on vehicle behavior.

### 6.2.4 Standardization of Vehicle Interfaces and Information

Vehicle usage data is produced if a vehicle is operated, therefore it could be concluded that this data belongs to the person operating (be it manually or autonomously) and/or owning it. Currently, most of the produced raw data is not accessible to the driver at all. In turn, the authors believe that vehicle usage data, collected from a mass of vehicles, can lead to the development of novel services for various stakeholders, if it were publicly accessible. Consequently, it has to be decided which stakeholders (e.g. vehicle manufacturers or public transport departments) are in charge of playing the governmental role in order to push standardization and execution.

According to Pillmann et al. (2017a), standardization is required since in the current state parameters vary from engine type to engine type and from manufacturer to manufacturer. The amount of signals which are currently accessible and in fact available across all passenger vehicle types and manufacturers is quite small (e.g. the signals defined in OBD-II standard to be found in ISO 15031-5) and thus not much greater than what one can find out using sensory of a smartphone mounted in the vehicle.

Another critical point is the anticipated amount of data to be sent and the querying frequency used by the data logging device, as stressing the vehicles' bus system for information exchange with this low-priority information retrieval might hinder more important actions.

Standardization on higher levels of aggregated information is another topic: Different manufactures are typically no longer an issue as the collected data is more or less source-agnostic. Nevertheless, it is still challenging to create a suitable data model, as the complexity and variety of the computed information is huge. There are already ongoing research projects, which aim to create such models, specialized in representing transportation and traffic related data and its exchange. The DATEX II multi-part standard (http://www.datex2.eu) and the EU project AutoMat proposing a Common Vehicle Information Model (CVIM) (Pillmann et al., 2017a) may serve as examples of such.

### 6.2.5 Business Models and Platform Ecosystems in the Context of Vehicle IS

According to Rehm et al. (2017), platform ecosystems conceptualize a platform as a "set of shared core technologies and technology standards underlying an organizational field that support value co-creation through specialization and complementary offerings" (Thomas et al. 2015). Consequently, platform owners provide the platform, manage, and control the ecosystem (e.g. Google with the Android platform), and according to Svahn et al. (2017) platform ecosystems based on vehicle data recently attracted vehicle manufacturers that seek to improve "end user experience and open up new revenue streams" with digital technologies.

Kuschel and Dahlbom (2007) stated that leveraging vehicle sensor data for services will not be profitable unless manufacturers make the sensor data open. Since then, vehicle sensor data still is not publicly available or accessible. However, some ICT start-ups from the US exploit the OBD interface or the smartphone sensory for their Vehicle IS, which in turn now forces manufacturers to react and develop ideas how they can enter this promising market themselves, as the EU project AutoMat (AutoMat, 2017) shows. In case of the ICT start-ups, many stakeholders from different domains, especially insurance, made investments in this topic to develop and explore new business models (Kaiser et al. 2017b). Hence, Mikusz and Herter (2016) mention, there is a "research gap on value propositions in business models for the Connected Car".

### 6.2.6 Decision Support Systems (DSS) in the Context of Vehicle IS

Over the last decades, a shift from pure human-made decisions to more and more computerized decision support systems could be observed. This is particularly true in the context of vehicles. Decision support or even automated decisions are provided for drivers with respect to (re)routing (Santos et al., 2011), braking in dangerous situations (Broggi et al., 2009), automated overtaking (Richter et al., 2016) or refueling (Suzuki et al., 2014). A recent article (Ryder et al., 2017) studies the impact of accident hotspot warnings on driver behavior. In this regard, increased driving data availability can even provide new options for stakeholders or infrastructure, like in the case of insurance risk selection processes (Baecke and Bocca, 2017) or charging infrastructure planning (Dong et al., 2014).

In the more distant future, fully automated vehicles will automatically obtain and interpret data from sensors correctly (positioning system, acceleration sensor, cameras, radar, etc.), aggregate and process this data, and decide on context-related information. However, until a sufficient level of automation is achieved, human interaction is required. Thus, humans will probably still play a major role in decision-making in foreseeable future.

In the view of a transition phase towards automated vehicles, major questions regarding vehicles' DSS appear, concerning drivers as well as stakeholders.

### 6.2.7 A Summary of Research Directions and Example Questions

In the context of Vehicle IS, the academic IS community may pay attention to the outlined research directions. To support this, the authors provide a list of example research questions per research direction listed in the following table.

**Table 12 Example research directions and questions for Vehicle IS. (Kaiser et al., 2018b)**

| Research direction | Example research questions per direction |
|---|---|
| Data analytics and artificial intelligence for vehicle data processing | - How can data analytics and artificial intelligence be leveraged for the design of Vehicle IS?<br>- What added value can be generated for Vehicle IS users through applying machine learning and artificial intelligence on data their vehicle generates?<br>- Which requirements will Vehicle IS pose on the quality of analysis algorithms?<br>- What are the most relevant features in vehicle data that are useful for generating data-driven services? Which features describe the driver, which the environment?<br>- What are relevant and available data sources for Vehicle IS?<br>- How intrusive should information about possible dangers be presented to the driver in order to achieve optimal response?<br>- How can the quality and reliability of complex models and information based on their interaction be assessed? |
| Wearables for Vehicle IS | - How can wearables contribute to optimized information provision? What type of information should be extracted from provided data?<br>- Under which conditions can wearables provide additional insights in the context of a Vehicle IS? Which data and functionalities provide added value to an IS considering current status and future development of wearables?<br>- Which requirements demand data of persons instead of data of vehicles? How can the supply of person-related data (e.g. level of stress) be organized while driving / co-driving?<br>- How can wearables contribute to more safety, both for occupants and non-drivers, in the context of a Vehicle IS?<br>- How can wearables be integrated into the quantified vehicle? How can the integration be ensured, in particular between the priorities of rapid software development / maximizing safety?<br>- To which extent are security-related as well as data privacy issues concerned, if wearables are integrated in the quantified car? How can unintended transfer of personalized data be prevented? What is needed in order to guarantee information security? |
| Privacy, trust, security, safety, legal, moral and ethical aspects in vehicle data processing and usage | - What is the impact of privacy and trust to the design and appropriation of Vehicle IS?<br>- How can Vehicle IS be designed in order to respect privacy and trust?<br>- What kind of privacy and trust labelling is required to better judge the risk of using Vehicle IS?<br>- How can IS design assure that the security of vehicles and as a consequence the safety of the vehicle driver is protected?<br>- What are ethical issues linked to Vehicle IS and how can ethical concerns be considered in a best possible way?<br>- How can the data owner of data generated by vehicle usage be defined?<br>- How can Vehicle IS provide information on how and by whom data is used? |
| Standardization of vehicle interfaces and information | - What is the influence on standardization on the design of Vehicle IS?<br>- How can the IS community support standardisation in the domain of Vehicle IS?<br>- How can standardisation accelerate the emergence of Vehicle IS and new business models based on Vehicle IS?<br>- What factors influence standardisation processes in the digital ecosystem of automotive industry?<br>- What are roles of actors within the vehicle ecosystem with respect to standardisation processes?<br>- How and by whom could a governmental authority be installed to orchestrate standardization respecting all stakeholders equally? |
| Business models and plat- | - How and by whom should platform ecosystems and business models be managed and controlled in order to enable innovativeness and fairness?<br>- What are different roles of stakeholders within a digital ecosystem for the Vehicle industry? |

| | |
|---|---|
| form ecosystems in the context of Vehicle IS | - What are individual value propositions in business models of Vehicle IS?<br>- How can IS contribute to a better understanding of digital ecosystems in the vehicle domain?<br>- How can ecosystems based on vehicle data be sustained?<br>- What is the role of ICT start-ups in the design and sustainability of platform ecosystems in the context of Vehicle IS? |
| Decision support systems (DSS) in the context of Vehicle IS | - What are driving factors of public acceptance for semi or fully automated driving without the need for human supervision?<br>- Under which conditions and to which extent is the transformation from human decision-making to automated decision-making accepted?<br>- How can interaction between manually-controlled and (semi-)automated vehicles be organized?<br>- What kind of information needs to be provided by a Vehicle IS for decision support regarding vehicles? What are sufficient data to provide information? How can the processing of information into decisions be designed?<br>- How to cope with situations not represented in underlying decision support system's models? How can human decision-making be ensured if required?<br>- In which way can DSS support non-drivers/stakeholders' decision-making? |

## 6.3 Conclusion and Future Work

To conclude, this chapter presents a research agenda for Vehicle IS with relevant research directions, including a set of example research questions per direction which the authors deem important for the academic IS community to advance the state-of-the-art in designing Vehicle IS.

The investigation of literature indicates increased research activities in the field of Vehicle IS, but at the same time emphasizes the need for definition and standardization of concepts and terms. To the best of the authors’ knowledge, the approach is the first one aiming at bridging the gap by defining a vehicle-centred IS. Furthermore, Vehicle IS is indicated as a new class of IS, which are currently probably considered too little by current research published in the AIS Electronic Library. Though the AISeL database has a broad coverage of IS literature relevant for the IS discipline, including no additional academic libraries in the investigated related work clearly represents a limitation of the chapter. Evidently, the limitation to exemplary research questions is inherent to the current approach.

Hence, for future work it is considered to conduct a systematic literature review in additional major scientific databases including e.g. ACM, ScienceDirect, Scopus, and Springer Link. Furthermore, to extend the research agenda (including the vehicle data generation stage) and develop a scientifically sound technology roadmap to even better support the IS community in finding appropriate and relevant research topics. In addition, it is planned to conduct a qualitative, system-oriented study to add perspectives from science, business, government and society in order to provide a holistic view of Vehicle IS. This study ensures overcoming the limitation of exemplary research questions by offering a complete research agenda including a multi-dimensional set of relevant and precise research questions.

# 7. The Vehicle Data Value Chain as a Lightweight Model to Describe Digital Vehicle Services[13]

**Summary and Author Contribution**

**Analysis and Definition of the VDVC**

**(Paper 1/2)**

In this chapter, to better understand and capture the ongoing digital transformation in the automotive domain, the Vehicle Data Value Chain (VDVC), grounded on big data, is introduced as a lightweight model to describe and examine *Data-driven Services* (termed digital vehicle services in the article). Current *Data-driven Services* are applied to the VDVC to identify commonalities and differences within three crucial steps: data generation, acquisition, and usage to evaluate the VDVC and show its general applicability in a practical context.

I contributed to this article by incorporating my experience gained during the development and analysis of *Data-driven Services.* While researching data-driven services, it became increasingly clear that the development of services follows a certain pattern that is comparable to the *Big Data Value Chain*. Therefore, I wanted to define and describe the so called *Vehicle Data Value Chain* (VDVC) in this paper. The paper is based on the abstract I presented at the European Transport Conference (ETC) 2018 (Kaiser et al., 2018a). In this paper, I derive the VDVC from the Big Data Value Chain and specify the individual steps. Furthermore, I show that the VDVC is also suitable to analyze and compare existing services. The paper was on the short list of candidates to win the WEBIST 2019 best student paper award, but unfortunately did not win. However, it did result in an invitation to extend the existing paper for a submission to the *Lecture Notes in Business Information Processing (LNBIP) series*, to be found in the upcoming Chapter 8.

Digitalization is an important driver of innovation within all industries, including the automotive industry (Accenture, 2016). While many digitalization challenges in the automotive industry are currently focused on bringing highly automated driving into practice (McKinsey and Company, 2016), it is also a crucial topic of research to explore how and which digital vehicle services can improve the current practice of manual driving or even enable novel applications for other stakeholders and other markets outside the automotive domain.

---

[13] The content of this chapter is based on

**Kaiser, C.**, Festl, A., Pucher, G., Fellmann, M., and Stocker, A. (2019). The Vehicle Data Value Chain as a Lightweight Model to Describe Digital Vehicle Services. In *Proceedings of the 15th International Conference on Web Information Systems and Technologies* (WEBIST), ISBN 978-989-758-386-5, pages 68-79. DOI: 10.5220/0008113200680079

The ongoing digitalization of passenger cars could even rearrange stakeholder power relations in the automotive industry. In the last decade, numerous IT start-ups from outside Europe have created several interesting digital services, exploiting data gained from the vehicle on-board diagnostic (OBD) interface, from additional sensors built into a connected OBD plug-in device and/or from the driver's smartphone. This could lead to new business models emerging in the automotive domain, some of which have already attracted the attention of car manufacturers. One prominent example is BMW i Ventures and its recent investments in start-ups such as Zendrive (2017) and Nauto (2017).

Two current key drivers of digitalization in the automotive domain are the ever-increasing amount of vehicle data generated and the capability of modern information and communication technologies (ICT) to transform these data into business value for various stakeholder groups. These may include individual stakeholders (e.g. vehicle drivers) as well as organizational stakeholders (e.g. insurance companies, infrastructure operators, or traffic operators). "Modern vehicles have up to one hundred on board control units that constantly communicate with each other to ensure the correct driving and customer functionality" (VDA, 2016). Hence, vehicles are already generating vast amounts of data using in-vehicle sensors. Certain parts of these data are safety-critical and will therefore not be allowed to leave the passenger car, while the remainder can and will be utilized to establish novel digital vehicle services (as indicated by the European Parliament in Directive 2010/40/EU (EU 2013)), which can go far beyond merely assuring driving functionality and safety.

Digital Vehicle Services are data processing services operating inter alia on vehicle data, which can provide added value to those consuming them. In this context, the term 'service' can be considered from two different points of view: On the one hand, a 'service' is understood as a piece of software applying approaches from computer science to transform and merge different sources of data (be it raw or pre-processed) into new, enriched forms of aggregated data. If done correctly, the value of these enriched data is inherently higher than the sum of values of the single datasets which were combined in the process. On the other hand, a 'service' is understood as something of economic relevance, providing an added value to one or more stakeholder groups as a service offering.

While the enormous amount of data available today enables the creation of valuable digital services in the first place, it also poses a great challenge with regard to data processing. To create value, data must be acquired, transformed, anonymized, annotated, cleaned, normalized, aggregated, analyzed, appropriately stored and finally presented to the end user in a meaningful way. This implies that an entire data value chain needs to be created, implemented and monitored. With this in mind, the authors analyze, summarize and provide insights into how existing initiatives on the market tackle this challenge. Hence, the authors aim to answer the following research question from the emerging field of digital vehicle services research: *What is the underlying data value chain enabling digital vehicle services and how can it be applied to describe existing services?*

To answer this research question, the authors first review the literature on relevant concepts for digital vehicle services, including Quantified Self, Big Data, and the Internet of Things. Based on the Big Data Value Chain as described by Curry (2016), the authors derive a *Vehicle Data Value Chain (VDVC)* that is intended to provide a structure and a frame of reference allowing to systematically describe the transformation of data into valuable services, to compare digital vehicle services and to understand and explain the data-related challenges associated with them in a second step. In a third step, the authors apply the developed VDVC and use it to finally classify current digital vehicle services offered by four selected start-ups and five car manufacturers.

After this introduction and a description of motivation, the authors continue with the relevance of (big) data from both a general point of view and a vehicle-specific perspective in Section 7.1. In Section 7.2, the authors describe how vehicle data are turned into digital vehicle services, introducing the vehicle data value chain as the underlying process of value generation. The authors then apply this vehicle data value chain to visualize and compare the digital product innovations by selected start-ups (Automatic, Dash, Vinli, Zendrive) and car manufacturers (BMW, Honda, Mercedes, Porsche, Opel) in Subsection 7.2.3, before the authors use the VDVC to analyze the digital mobility service MoveBW in detail in Subsection 7.2.4. Finally, in Section 7.3, the chapter is concluded and provides an outlook on various current activities including standardization and other activities of the European Commission and car manufacturers and the ongoing research project AEGIS, which aims to ease data fusion and the linking of data artifacts from multiple data sources.

## 7.1 From Data to Big Vehicle Data

### 7.1.1 Data: One Aspect of Digitalization

More than a decade ago, Tim O'Reilly formulated his extensively cited principles of the Web 2.0 (O'Reilly, 2005) including one principle about the emerging value of data according to which "data is the next intel inside". Since then, the hype on how to generate added value from all kinds of available data has been building. Data is the new buzzword. A book by Mayer-Schönberger and Cukier (2013) on how Big Data is changing our world has become an international bestseller and been cited by researchers more than 2964 times according to Google Scholar. Big Data has received considerable attention from multiple disciplines, including information systems research (Abbasi et al., 2016) and database management (Batini et al., 2015), to mention two of them.

The volume of data is growing exponentially. It is expected that there will be more than 16 zettabytes (16 Trillion GB) of useful data by 2020 (Turner et al., 2014). It is just a logical consequence that data generation, data analysis, data usage –

and related new business models – have found their way into all areas of life. Homes are increasingly equipped with smart meters, a replacement for mechanic measurement of electricity usage, enabling the emergence of digital services to assist home monitoring and to optimize electricity usage. Smartwatches can track the wearer's behavioral data and calculate periodic statistics such as daily, weekly, or monthly walking distances including burned calories per day, week, or month. Many people use their smartphones when exercising to gather information on their workout.

Smartphone apps such as Runtastic (2017a) and Strava (2017) help to monitor how and where people run or cycle, automatically calculating route, pace and periodic statistics including mean speed, time per kilometer, and calories burned. These apps even allow sharing the aggregated data via social networks, thus enabling benchmarking with peers and increasing the joy of exercise. The pattern of collecting, analyzing, and sharing data constitutes the baseline for individual improvements. Instantly calculated and visualized behavioral statistics are easy to compare or share with peers on social media. The collected information per se is not new to these communities. For instance, experienced runners started comparing their real and average time per kilometer using stopwatches a long time ago. However, the simplicity of digital services and the fact that many friends on social media regularly post about their exercising routine has motivated a whole digital generation to track themselves, as 210 million Runtastic app downloads demonstrate (Runtastic, 2017b). 30 million app sessions per month in Europe produce a reasonable amount of big movement data, which is sufficient for performing representative data analyses and attracts various stakeholders including Adidas.

To summarize, digitalization has greatly simplified data collection and analysis methods which used to be too complex and/or only available to experts. Hence, more and more people are joining the self-tracking movement and, in turn, produce more and more data which can be exploited using novel digital services.

### 7.1.2 The Big Data Value Chain

The internet age has spawned far more data on anything than any other technical or organizational innovation. Big Data refers to the current conglomerate of newly developed methods and information technologies to capture, store and analyze large and expandable volumes of differently structured data. In a definition by Demchenko et al. (2013), the defining properties of Big Data are Volume, Velocity, Variety, Value and Veracity, as shown in Figure 31. Exploiting the new flows of data can even improve the performance of companies, if the decision-making culture is appropriate (McAfee and Brynolfsson, 2012).

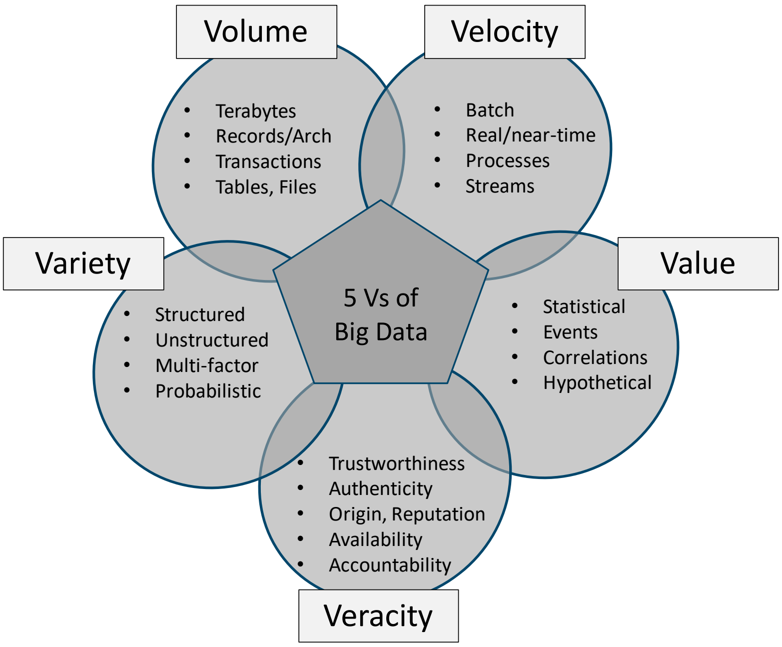


**Figure 31 The 5 Vs of Big Data (derived from Demchenko et al., 2013). (Kaiser et al., 2019a)**

Big Data and intelligent things seem to have an intimate relationship. While in the Web 2.0 era data was mainly generated by humans sharing user-generated content on portals including YouTube, Wikipedia, or Facebook, the Internet of Things has led to new patterns of data generation driven by machines. Smart, connected objects equipped with all kinds of sensors have now taken over this task (Porter and Heppelmann, 2014 and 2015). The Quantified Self phenomenon is making use of these data generated by things (Swan, 2009 and 2015). Quantified Self refers to the intention to collect any data about the self that can be tracked, including biological, physical, behavioral, and environmental information. Making use of these data to establish applications and services has become a major creator of value. This value is created through multiple activities which are chained together, while the value of the output is steadily increasing.

The concept of a value chain was originally introduced by Porter to describe a series of activities of a company to create and build value (Porter and Millar, 1985). This value chain concept can also be applied to the data domain to describe activities ranging from data generation to the usage of data in data-driven services for the customer. Data value chains are a model to describe data flows as a series of steps, each of them transforming the value of data. The concept of data value chains has already been used to describe the value of Linked Data (Latif et al., 2009) as well as of Big Data by Curry et al. (2014) as illustrated in Figure 32. The Big Data Value Chain mentions several steps of Big Data transformation in the process of generating the data-driven result with the maximum business value.

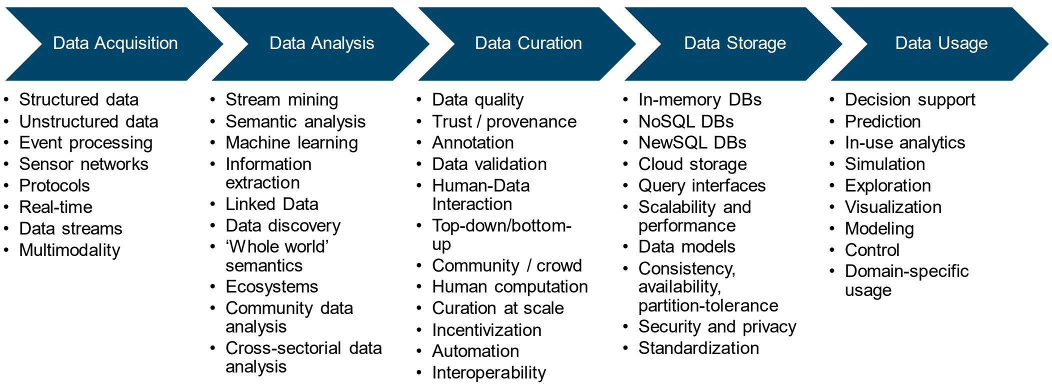


**Figure 32 The Big Data Value Chain of Curry et al. (2014) / Curry (2016). (Kaiser et al., 2019a)**

### 7.1.3 Big Data in the Context of Vehicles

Decades ago, vehicles were merely equipped with mechanical components such as mechanic handbrakes. However, electrification and comfort requirements continuously led to an electrically operated handbrake. The handbrake status (applied or released) and its process status (handbrake is applying/releasing) can be captured and used as input for vehicle safety checks and other features. An applied handbrake will automatically be released if the driver starts driving to prevent damage. The data generated through all these vehicle functions can be captured and used within other scenarios, e.g. to create statistics on how often a window is opened/closed or how often somebody is wedged in.

Many vehicle sensors are currently only used to offer functionality and/or to increase comfort and safety. As sensors and car features may widely differ from manufacturer to manufacturer and even per car variant, there is not only one single truth about how much data is effectively generated by a modern vehicle today. For instance, the participants from the EU project Automat (2017 and 2018c) state that about 4000 CAN bus signals (one signal could be one measurement value) per second create up to 1 GB of data per CAN bus (without mentioning a sample rate). According to Pillmann et al. (2017b), there are “usually 4-12 CAN busses in one car” (with varying amounts of input signals).

Considering the current hype around bringing highly automated driving into practice, several camera, radar and LiDAR (light detection and ranging) systems are additionally implemented within vehicles to capture each angle of the vehicle’s environment. Autonomous vehicles are forced to exchange information with other vehicles (V2V) and with the infrastructure (V2I), which will boost the amount of available vehicle data enormously in the future. Considering different countries and different patterns of individual driving and mobility behavior, bringing highly automated driving into practice can be seen as a grand digitalization challenge.

However, while only some of these data can be exploited for digital vehicle services (e.g. because the sampling rate is too high or because some values are simply not relevant) and while only a portion of these data will be made accessible due to safety reasons (EU, 2013), the remainder of accessible sensor data from modern vehicles will most likely be sufficient to design and develop a reasonable number of novel digital vehicle services for various stakeholder groups, including individual drivers, various organizational customers, government authorities, and the automotive industry (Kaiser et al., 2017b). To sum up, modern vehicles already constitute big vehicle data generators.

## 7.2 Generating Business Value: From Vehicle Data to Digital Vehicle Services

### 7.2.1 Generating Business Value by Leveraging the Self-tracking Trend

Many digital natives enjoy generating data anytime and anywhere using mobile devices including smartphones and smart watches. Increasing the knowledge about oneself and eventually enabling new discoveries while performing physical activities including running or cycling has turned into a business-relevant phenomenon. The behavior of turning collected data about oneself into actionable knowledge and insight which is valuable for other stakeholders, too, has been termed Quantified Self. Interestingly, the quantified self phenomenon has recently been successfully transferred to the automotive industry by US-based start-ups. In this sense and quite analogously, Quantified Vehicles (Stocker et al., 2017a) imply a successful transformation of data from different kinds of sensors related to the vehicle (in-vehicle sensors, smartphone and wearable sensors used by the driver) into actionable knowledge, e.g. on the behavior of the vehicle. This way, they generate value for different kinds of stakeholders that are part of digital vehicle data service ecosystems such as insurance or fleet management providers, finally resulting in novel digital vehicle services in various domains (Kaiser et al., 2018b; Kaiser et al., 2019b).

The pattern of self-tracking using consumer devices, as portrayed by the Runtastic example, can be easily transferred to vehicles: By default, vehicles gather a plethora of vehicle operation data through sensors and control units safeguarding a vehicle's functionality. However, these vehicle Big Data could be used to enable a series of apps and services. In the case of Runtastic, the combination of the company and the high volume of generated data, i.e. knowledge on where, how and how often users engage in physical activity such as running, was considered worth €220 million by the Adidas Group, which acquired Runtastic in 2015 (Runtastic, 2015).

The market value for vehicle data is considered to be even higher due to the importance of vehicles in first world countries. A number of US-based ICT start-ups seized this opportunity, now offering smartphone and web applications providing insights into vehicle-generated data, after they received up to €25 million of funding from investors (Stocker et al., 2017a). Interestingly, while some car manufacturers and suppliers (e.g. Magna International, Continental ITS, and BMW i Ventures) are among the investors, forming strategic partnerships with start-ups, others participate in research projects and try to keep data-related business in their own area of influence. This holds for Volkswagen, for example, which coordinates the EU project Automat to develop a marketplace for vehicle lifecycle data (Stocker and Kaiser, 2016). Furthermore, recent reports from the German automotive industry association (VDA) suggest that car manufacturers "have to hold a stronger position in the future and may limit the capabilities of third parties to freely access car data." To summarize, the potential of vehicle data seems to be such that it has become a battle worth fighting (Kaiser et al., 2017b). But how can vehicle data actually generate value?

### 7.2.2 The Vehicle Data Value Chain

In order to provide a structure and a frame of reference allowing to systematically describe the transformation of data into valuable services, to compare digital vehicle services and to understand and explain the data-related challenges associated with them, a value chain for vehicle data can be used. In this regard, the authors propose the Vehicle Data Value Chain (VDVC) as a lightweight model. The authors derived the VDVC from the Big Data Value Chain (Curry et al., 2016, illustrated in Figure 32). The authors adapted Curry's value chain regarding the name, number and order of stages to reflect the authors' experiences from research projects in the automotive domain. The stage of (Vehicle) Data Generation was added as a separate stage to explicitly reflect the origin of the data (e.g. in-vehicle or related sensors). The stage (Vehicle) Data Acquisition corresponds to Curry's Data Acquisition. Moreover, the authors have changed the order of Curry's stages of analysis and curation since the authors interpret the terminology differently. For example, Curry seems to include normalization procedures implicitly within machine learning in the stage of Data Analysis, whereas the authors consider this as an important separate pre-processing step which correlates with Curry's stage of Data Curation. Hence, the authors have re-named Curry's stage of Data Curation (Vehicle) Data Pre-processing, which is followed by the stages (Vehicle) Data Analysis, (Vehicle) Data Storage, and (Vehicle) Data Usage (see Figure 33).

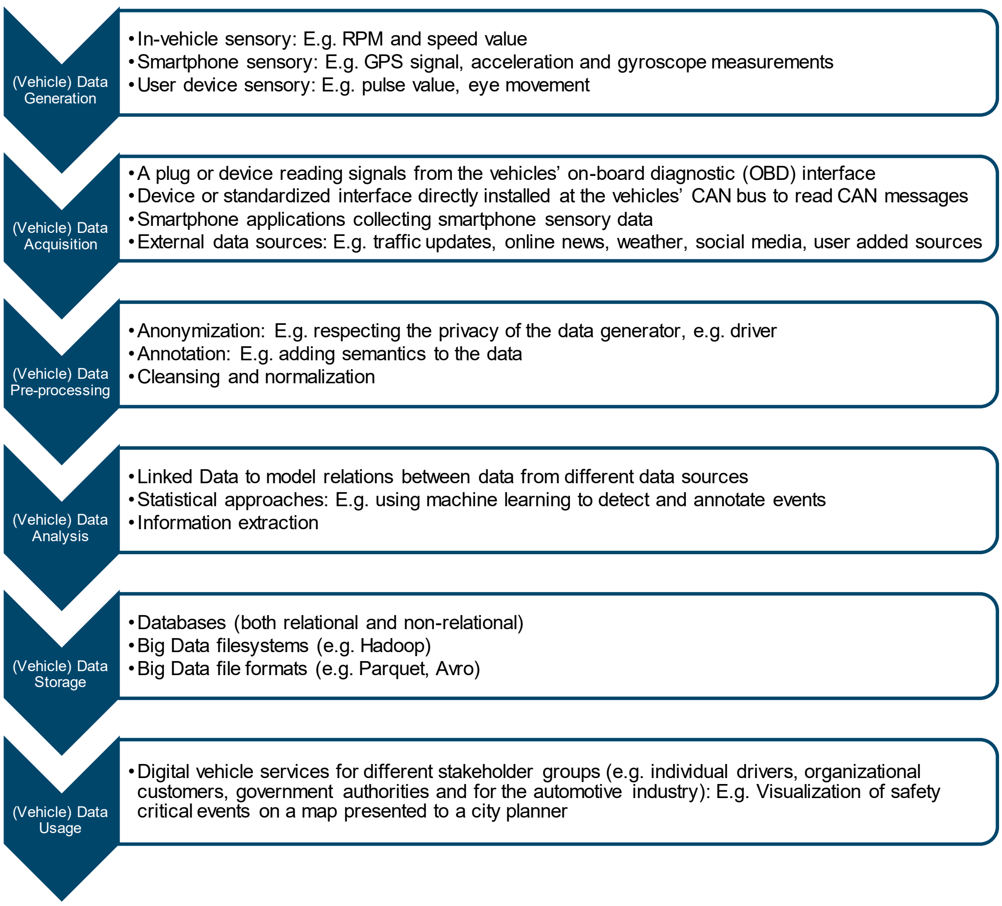


**Figure 33 The Vehicle Data Value Chain derived from Curry (2016) and based on Kaiser et al. (2018a). (Kaiser et al., 2019a)**

**(Vehicle) Data Generation** summarizes any sensors which can capture data directly (throttle pedal position) or indirectly (road surface condition). In the case of direct influence, the authors mainly see three data sources: In-vehicle sensors, smartphone sensors and individual user device sensors (e.g. a pulse transmitter belt). An indirectly influencing data source can be literally any relevant data source, for instance a road operator camera to indicate traffic flow. This process step is not included in the Big Data process described in Subsection 7.1.3, however, it is essential for the vehicle data value chain, as the data origin indicates the reliability and the influence type (direct, indirect).

**(Vehicle) Data Acquisition** is the process of gathering the generated data. In-vehicle sensor data per se is not directly accessible, as it is captured with the purpose of safeguarding a vehicle's functionality and therefore only shared between the various electronic control units via one of the vehicle's CAN buses. However, a filtered amount of these sensor data is already accessible via the On-board diagnostic (OBD) interface (Turker and Kutlu, 2015), which is intended to be used by service staff to read generated error messages. It is however possible to develop plug-in devices with internet connection, to effectively use the OBD-port as a sensor data source. There are already some professional solutions with data acquisition devices installed in the vehicle, which directly read signals from the CAN bus in an unfiltered way. To meet the requirements of the EU Directive 2010/40/EU inter alia on the costless provision of universal, road safety-related minimum traffic information (EU, 2013), a standardized interface would be feasible sooner or later. Data from smartphones is acquired by using specific applications, which are capable of gathering and transmitting data. In the case of external data sources restricted to sources accessible via the Internet, the main issue are the different availability and quality levels of the data. For example, APIs commonly limit the number of requests allowed per time interval, meaning that the acquisition process must be adapted to respect these thresholds. Gathered data is stored for further processing; the chosen storage and format heavily depend on the following processing steps.

**(Vehicle) Data Pre-processing** describes any anonymization, annotation, cleansing and normalization activities before any data analysis is conducted. Sensor values may include private user information or may be erroneous, different sensors may have their own sampling frequency and so on. Data quality highly influences service quality. For instance, if the GNSS signal accuracy is low, a trip may not be linked to the correct road and may lead to false conclusions.

**(Vehicle) Data Analysis** with the purpose of extracting useful hidden information involves linking data from different data sources, exploring data, performing statistical analyses, using machine learning algorithms, and, if needed, detecting events, etc. For instance, weather data can be linked with the vehicle speed on a certain road to detect if the driver drives differently when the road is wet or icy.

**(Vehicle) Data Storage** “is the persistence and management of data in a scalable way that satisfies the needs of applications that require fast access to the data” (Curry, 2016). In the case of vehicle sensor data, persistent storage is usually achieved by using a combination of classic relational databases (for meta-data), Big Data file systems (for raw input data) and so called “time series databases”, which allow fast analyses on the stored contents.

**(Vehicle) Data Usage** covers all ways of user or software interaction with the collected data and any conclusions derived from it in the above-mentioned process. The accessed data could either be regarded as the end result of the process, in which case it will be presented more or less directly to end users, or it could serve as input for further processing steps, forming a circular path in the processing chain.

### 7.2.3 Applying the VDVC to Describe Digital Vehicle Services Offered by US Start-ups and Prominent Car Manufacturers

The Vehicle Data Value Chain (VDVC) introduced in the previous section describes a set of activities to create value out of vehicle data. Consequently, a “vehicle data to service”-process can be derived from the above mentioned VDVC. In this section, the authors aim to apply the VDVC as a lightweight model to characterize selected public digital vehicle services offered by four start-ups and five car manufacturers. The stages of (Vehicle) Data Curation to (Vehicle) Data Storage of the value chain are part of the respective digital vehicle service providers’ business asset and are therefore not publicly available. In addition, not all digital vehicle service providers can be expected to publish a full list of third-party stakeholders which have access to the vehicle data acquired. However, in a second step the authors add a detailed description of a single service called MoveBW, which was co-developed by one of the authors, so that the authors can give insights into the value chain of this service.

Digital vehicle service providers the authors chose are presented in Table 13. This table focuses on services for individual drivers and explicitly observes the following three process steps: (i) (vehicle) data generation, (ii) (vehicle) data acquisition and (iii) (vehicle) data usage.

(Vehicle) data usage is structured using four categories: (a) Recommendation specifies all digital vehicle services that give recommendations to the user, e.g. how to improve fuel efficiency; (b) Vehicle status & trip statistics lists services which represent the status of the vehicle (e.g. remaining fuel) and statistics from recent trips (e.g. a score representing the driver's cautiousness); (c) Access to vehicle features gives a list of services which enable vehicle features to be accessed using a smartphone application (e.g. controlling the air conditioning); (d) Other contains all services which go beyond the aforementioned categories.

The resulting table shows that the various digital vehicle services provided by start-ups and car manufacturers (termed OEM for Original Equipment Manufacturer) vary in terms of data generation, data acquisition and data usage. For instance, start-ups access in-vehicle data mainly by exploiting the OBD interface, except for Zendrive, which relies on smartphone sensors only. The OBD plug-in devices used by the start-ups differ, as they have additional sensors built in to capture additional data and hardware to establish UMTS/WIFI connections for transmitting data to the storage. The only exception is Honda, which also uses the OBD plug solution. Car manufacturers use the advantage they have as the vehicle developer and rely on an integrated CAN bus device that can capture vehicle data from far more sensors than OBD-based devices. It is surprising that the offered digital vehicle services somehow resemble one another.

Due to limited information access, the applicability of the VDVC for US tech start-ups and prominent car manufacturers has been shown using the steps Data Generation, Acquisition, and Usage only. However, in the following section, the authors analyze one mobility service where the authors have insights into the full process using each step of the VDVC.

**Table 13 A digital vehicle service overview focusing on Data Generation, Acquisition and Usage. (Kaiser et al., 2019a)**

| Service Provider | Service Purpose | Data Generation | Data Acquisition | Data Usage for drivers including business customers |
|---|---|---|---|---|
| Automatic (Start-up) | Driving statistics to infer behavior | In-vehicle sensors & device sensors | OBD device | *Vehicle status & trip statistics* (driving behavior insights using a score; event detection e.g. hard brakes or speeding; location & trip tracking; business tagging; fill up logging; vehicle error messages; crash alert)<br>*Other*: IFTTT (IF This Then That) |
| dash (Start-up) | Driving statistics to infer behavior; Gamification | In-vehicle sensors & device sensors | OBD device | *Recommendation* (improve fuel efficiency; ranked refueling places)<br>*Vehicle status & trip statistics* (driving behavior insights using a score; location & trip tracking; fuel quantity; live speed, rpm, engine load, etc.; vehicle error messages)<br>*Other*: gamification (rewards, leaderboard) |
| Vinli (Start-up) | Ecosystem with 40 Apps: individual purposes | In-vehicle sensors & device sensors | OBD device | Some known services: Amazon 'Alexa': trip statistic integration when asked for ("Where is my car?"); IFTTT (IF This Then That) programmable functionality; 'Home Connect': e.g. close garage door when leaving with passenger car; etc. |
| Zendrive (Start-up) | Gamification, fleet mgmt. | Smart-phone sensors | Smart-phone app | *Recommendation* (coach safe driving)<br>*Vehicle status & trip statistics* (driving behavior insights using scores, e.g. Caution Score; location & trip tracking) |
| BMW (i) Connected Drive (OEM) | Personal mobility assistant | In-vehicle sensors & external sources | CAN bus device | *Recommendation* (Navigation based on traffic status; parking place recommend.)<br>*Vehicle status & trip statistics* (location & trip tracking; emergency call; fuel quantity)<br>*Remote access to vehicle features* (unlocking; honking; air conditioning) |
| Honda (OEM) | Driving statistics to infer behavior | In-vehicle sensors & device sensors | OBD device | *Vehicle status & trip statistics* (location & trip tracking even in car parks; emergency call; fuel quantity; vehicle error messages)<br>*Other*: geo-fence alarms (leaves area) |
| Mercedes me (OEM) | Personal mobility assistant | In-vehicle sensors & external sources | CAN bus device | *Recommendation* (Navigation based on traffic status; fleet mgmt.)<br>*Vehicle status & trip statistics* (location & trip tracking; emergency call; trip logging; fleet mgmt.)<br>*Remote access to vehicle features* (remote parking; unlocking; air conditioning) |
| Opel OnStar (OEM) | Personal mobility assistant | In-vehicle sensors & external sources | CAN bus device | *Vehicle status & trip statistics* (location & trip tracking; emergency call; fuel quantity, tire pressure, etc.; vehicle error messages)<br>*Remote access to vehicle features* (unlocking; honking)<br>*Other*: Privacy options |
| Porsche Connect (OEM) | Personal mobility assistant | In-vehicle sensors & external sources | CAN bus device | *Recommendation* (Navigation based on traffic status, including hints for frequently used routes if navigation is turned off; parking place recommend.)<br>*Vehicle status & trip statistics* (location & trip tracking; emergency call; fuel quantity)<br>*Remote access to vehicle features* (Smartphone and vehicle navigation are connected; unlocking; honking)<br>*Other*: smart home features (close garage door when leaving with passenger car) |

### 7.2.4 Applying the VDVC to Describe MoveBW, a Digital Vehicle Service

MoveBW is a regional, intermodal mobility service offered by a European industry consortium and which is currently being developed to increase the compliance rate of transport users (e.g. the percentage of people using a park and ride option) with regard to the current transport strategy of the region. The strategy mainly aims at meeting air quality targets and reducing traffic jams all over the federal province of Baden-Württemberg (Germany), including its provincial capital Stuttgart.

Stuttgart is geographically located in a valley basin, which has a negative effect on air pollution with particulate matter. Thus, the city of Stuttgart continuously develops transport strategies to better comply with air quality regulations. In the past, these strategies were communicated to the public using radio traffic messages or electric traffic signs only. However, the compliance rate and thus success were comparably low. The new MoveBW mobility service smartphone application aims to increase the compliance rate, especially that of visitors new to the region. It does so by including easy-to-use routing functionalities which are connected to rewards: Bonus points are granted if a user follows the recommended route. Collected bonus points can later be exchanged for immaterial or monetary values.

Users of the MoveBW smartphone application can plan their trips in advance using the intermodal journey planner. They can pick their preferred combination of transport modes from different options suggested to them. Additional information is displayed, not only showing travel time, but also eco-friendliness, travel costs and incentives gained (e.g. public transport vouchers and $CO_2$ savings). Moreover, it is possible to directly book tickets for the different modes of transport included in their preferred journey and yet to receive only one bill. In this way, transport services such as public transportation, car sharing, bike sharing, and parking space management are integrated conveniently, encouraging users to make efficient use of all modes of transport. The application also provides on-trip navigation and information on traffic obstructions such as construction works or accidents.

The MoveBW services are currently monitored and evaluated in an extensive trial phase. Based on the findings, both the digital service and traffic control strategies will be revised, aiming to maximize favored effects on the individual mobility behaviors of traffic participants, for example by applying different strategies for daily commuters and visitors. The smartphone application is planned to be released in the first quarter of 2019. Mock-ups of the current design are shown in Figure 34.

**Figure 34 The MoveBW smartphone application. (Source: Strukhoff et al., 2017)**

Taking a wide range of data sources into account for the intermodal routing algorithms in the MoveBW App, data management becomes a challenge. The Vehicle Data Value Chain introduced in this section helps to provide a clearer view. Its application to the underlying data transformation process, from Data Generation to Data Usage, is shown in Table 14.

**Table 14 An overview of digital vehicle service MoveBW including all VDVC steps (Kaiser et al., 2019a)**

| VDVC step | Description of MoveBW-Service |
|---|---|
| Data Generation | Various sensor data and basic reference data is considered, e.g.<br>- floating car data: average mean travel time per road segment based on anonymized GNSS data of vehicles,<br>- stationary traffic measurement: rate of flow for single measurement locations,<br>- public transport: schedule and sometimes occupancy rate,<br>- car park interfaces: occupancy rate,<br>- park & ride interfaces: occupancy rate,<br>- air quality measurement units: air quality measurements and forecast (includes weather forecast); |
| Data Acquisition | Querying the web APIs from the various data sources. Additionally, the smartphone App which is used in Data Usage provides GNSS information, as this is used for on-trip routing and to detect which means of transport the user actually uses to be able to reward if the recommended option is used. |
| Data Pre-processing | Annotation, normalization and semantic extraction of data. Transformation of data to meet a common reference basis (in this case a public transport grid, no typical geo-coordinates). Furthermore, GNSS data from the smartphone App is anonymized (start- and end-trajectories are truncated). In this step the data is hosted in a distributed database system (e.g. PostgreSQL cluster) |
| Data Analysis | A dynamic routing algorithm which also takes the provided intermodal transport strategy, CO2 savings, and personal preferences into account. A self-developed algorithm which utilizes pgRouting (an open source project to extend PostGIS/PostgreSQL to provide geospatial routing functionality) and the popular Dijkstra algorithm (to find the shortest path between nodes)<br>Thus, the algorithm provides routing recommendations (weightings for routes) |
| Data Storage | A distributed database system, e.g. a PostgreSQL cluster |
| Data Usage | The MoveBW App currently being developed should help the commuter to choose a mode of transport and guides the commuter to the selected destination in compliance with environmentally-oriented traffic management strategies. |

## 7.3 Conclusion and Outlook

Digitalization has become an important driver of innovation in the automotive industry, enabling a plethora of digital vehicle services. The authors have presented a review of available digital vehicle services offered by startups and car manufacturers and described them applying the Vehicle Data Value Chain (VDVC). Many of them were originally motivated by the self-tracking phenomenon, which has been transferred to the vehicle domain, constituting quantified vehicles.

As an outlook, it should be mentioned that digital vehicle services and the required technological infrastructure to facilitate data acquisition, pre-processing, analysis and storage, are currently urged topics in the automotive domain. There are already initial ideas using blockchain technology and brokers to make data sharing transparent and secure, as described in Kaiser et. al (2019c). Yet, while some car manufacturers invest in start-ups, others limit access to data via the OBD interface, arguing that they are not suitable for digital vehicle services (VDA, 2017; ACEA, 2016). In contrast, the European Automobile Manufacturers Association ACEA promotes car data sharing (ACEA, 2017).

One reason for activities in this area is the Commission Delegated Regulation (EU) No 886/2013 (regarding Directive 2010/40/EU on Intelligent Transport Systems – ITS) published by the European Commission. It regulates the costless provision of universal, road safety-related minimum traffic information to users and requests car manufacturers to provide safety-relevant data to the public by making it accessible through national contact points (EU 2013).

Furthermore, the International Organisation for Standardisation (ISO 2017) has set up a standardization work group titled ISO/TC 22/SC 31/WG 6 Extended Vehicle/Remote diagnostics (ISO 2018) to inter alia define access, content, control and security mechanisms for the provision of vehicle data for web services (VDA, 2017).

In parallel, a joint initiative of 17 EU Member States and road operators is launching a solution for C-ITS services in order to transmit information from infrastructure (e.g. road side units) to the vehicle cockpit, e.g. to inform about slow or stationary vehicle(s), traffic jams ahead, weather conditions, speed limits, etc. (C-ROADS, 2017).

Additionally, current EU-funded projects such as the AEGIS Big Data project or EVOLVE are developing solutions to ease the integration and fusion of multiple data sources for the purpose of service and business development using Linked Data (AEGIS, 2017; EVOLVE, 2019; Latif et al., 2009). “Linked data is a lightweight practice for exposing and connecting pieces of data, information, or knowledge using basic web standards. It promises to open up siloed data ownership and is already an enabler of open data and data sharing” (Rusitschka and Curry, 2016).

To conclude, the authors expect the market of digital vehicle services to grow tremendously in the future, as the combination of vehicle data with data from external sources (e.g. weather data, traffic data, open data) will enable new scenarios for digital vehicle services.

# 8. Digital Services Based on Vehicle Usage Data: The Underlying Vehicle Data Value Chain [14]


**Summary and Author Contribution**

**Analysis and Definition of the VDVC (Paper 2/2)**

The quantify-everything trend has reached the automotive sector while digitalization is a still the major driver of innovation. New digital services based on vehicle usage data are being created for different actors and purposes. As a side effect, a growing number of ICT start-ups from outside Europe have entered the automotive market to work on innovative use cases. Their digital services are based on the availability of vehicle data on a large scale. To better understand and capture this ongoing digital change in the automotive sector, this chapter presents an extended version of the Vehicle Data Value Chain (VDVC) originally published in Kaiser et al. (2019a) and uses it as a model for better structuring, describing and testing digital services based on vehicle usage data. Furthermore, the VDVC is used to classify digital services of two projects, an intermodal mobility service and a pothole and driving style detection service, to evaluate the VDVC and show its general applicability and usefulness in a practical context.

In this expansion based on Kaiser et al. (2018b) and Kaiser et al. (2019a), I wanted to describe the individual steps of the VDVC more precisely to give the reader an even better understanding of what each step entails and constitutes. Thereby, I describe the Scope, give examples for Inputs, Outputs, Actors, Architecture, Trends, and Tools and descibe the contribution to value creation. In addition, we expanded the paper with an evaluation, which includes two cases. The newly added second evaluation case describes a *Pothole and Driving Style Detection Service*, which we (first, second and last author) have developed together with other colleagues as part of an internal project under my leadership and in the EU projects SCOTT and EVOLVE.


Modern mobility is an important driver of the increasingly global economy: raw materials are transported around the globe and processed into products in value-added processes until they finally find their way to the customer via many intermediate stations. Passenger cars and trucks are assembled in a complex supply chain consisting of many small parts and

[14] The content of this chapter is based on

**Kaiser, C.**, Festl, A., Pucher, G., Fellmann, M. and A. Stocker (2020). Digital Services Based on Vehicle Usage Data: The Underlying Vehicle Data Value Chain. In: Bozzon A., Domínguez Mayo J.F., Filipe J. (eds) *Web Information Systems and Technologies*. WEBIST 2019. *Lecture Notes in Business Information Processing (LNBIP), vol 399*. Springer, Cham. https://doi.org/10.1007/978-3-030-61750-9_2 components and finally manufactured in several value-added steps on a production line before they are delivered to customers. This basic business principle was very successful in many domains for a long time, until digitalization added another business aspect, which is becoming an important driver and has even become the decisive criterion in many sectors, including the automotive industry (Accenture, 2016). Similar to smartphones, where the focus is no longer on the original innovation, i.e. telephoning, but on digital apps, it is becoming increasingly important for vehicles, too, which digital functionalities they offer – from the Bluetooth connectivity with smartphones to Vehicle-to-Infrastructure (V2I) and Vehicle-to-Vehicle (V2V) services or to third-party services that someone can install. And in the context of vehicle use, services based on vehicle usage data have the potential to go beyond the usual application focus of quantified self-applications, namely self-optimization, learning about oneself, social comparison and interaction or gaming, as they can even be extended in a life-saving manner. Driving style analysis, for example, is able to detect driver fatigue and distraction (Lechner et al., 2019), two of the most common causes of accidents. Thus, it is crucial for research to explore how digital services based on vehicle data can improve the practice of driving or enable novel applications for other stakeholders and other markets outside the automotive domain (Stocker et al., 2017a).

The basis for digitalization in the automotive domain are the ever-increasing amount of vehicle usage data generated (e.g. modern vehicles are increasingly equipped with radar, lidar and video to support ADAS functionalities) and the ever-increasing capacity of information and communication technologies (ICT) to convert this data into business value for different stakeholder groups. These may include individual stakeholders (e.g. vehicle drivers) as well as organizational stakeholders (e.g. car manufacturers, fleet managers, infrastructure maintainers, or traffic planners). Utilizing "up to one hundred on board control units that constantly communicate with each other" (VDA, 2016), modern vehicles are already generating big data using in-vehicle sensors. Certain parts of this data are safety-critical and must there-fore not leave the car, while the rest can and will be used for the establishment of novel digital services based on vehicle usage data, which can go far beyond ensuring driving functionality and safety and opens up a multitude of possibilities.

As IT companies enter the automotive market with their services, the balance of power between the players in the automotive industry may also change. IT start-ups have already created several interesting digital services based on data from the vehicle's on-board diagnostic (OBD) interface or from the driver's smartphone (Kaiser et al., 2017b). This has led to the emergence of new business models in the automotive sector and has even attracted the attention of car manufacturers already. A prominent example is BMW i Ventures and its recent investments in start-ups such as Nau-to (improving the safety of commercial fleets, investment made in 2017) and Zūm (providing technologies for reliable child transportation, investment made in 2019). Now the point is reached where it has to be decided how to go on: Either the large vehicle manufacturers will buy in / redevelop the most promising digital services of the start-ups, or, to the vehicle could merely become an exchangeable device/ plat-form on which digital services run, similar to the smartphone.

Digital services based on vehicle usage data are data processing services which, among other things, work with data related to vehicle driving and can offer added value to users. In this context, the term ‘service’ can be viewed from two different angles: On the one hand, a ‘service’ is understood as a piece of software applying approaches from computer science to transform and merge different sources of data (be it raw data or pre-processed data) into new, enriched forms of aggregated data (Lechner et al, 2019). When performed correctly, the value of these enriched data is inherently higher than the sum of values of the single datasets which were combined in the process. On the other hand, a ‘service’ is understood as something of economic relevance, providing an added value as a service offering to one or more stakeholder groups.

However, the market entry of start-ups has already created a new *data-driven service ecosystem* in the automotive sector, leading to new data flows and collaborations in service development, as Kaiser et al. (2019b) describes. In the high-level view of this empirically obtained ecosystem with experts from the field, there is a data flow from data providers to service providers, who offer services on the market that are consumed by service consumers at the end of the value chain. On closer inspection, for example, there are five ways in which a service provider can already obtain relevant data on a car trip, i) from a market place (e.g. otonomo.io), ii) directly from the car manufacturer (e.g. BMW), iii) from data intermediaries (e.g. HERE Technologies, which has a close relationship to BMW and Daimler), iv) from the results of other service providers and v) from external data sources (e.g. weather services, con-gestion warnings).

The enormous amount of data available today makes the creation of valid digital services possible in the first place, but also poses a major challenge with regard to data processing (Xu et al., 2017). To create value, data must be acquired, trans-formed, anonymized, anno-tated, cleaned, normalized, aggregated, analyzed, appropriately stored and finally presented to the end user in a meaningful way. This implies that an entire data value chain must be created, implemented and monitored. With this in mind, Kaiser et al. (2019a) derived the *Vehicle Data Value Chain (VDVC)* from the Big Data Value Chain as described by Curry (2016) and a literature review on relevant concepts for digital services based on vehicle usage data, including Quantified Self, Big Data, and the Internet of Things. This VDVC is intended to provide a structure and a framework allowing to systematically describe the transformation of data into valuable services, to compare existing digital vehicle services with each other and to understand and explain the data-related challenges associated with them. Hence, the VDVC was used to analyze, summarize, and provide insights into existing start-up and vehicle manufacturer initiatives on the market. As a result, the authors decided to apply the VDVC in the development of services in two case studies, the intermodal mobility service MoveBW (case A) and a pothole and driving style detection service (case B). Finally, this chapter is an extended version of Kaiser et al. (2019a), elaborating the VDVC and using another case study of a digital service based on vehicle data for evaluating the improved VDVC.

During the development of digital services based on vehicle data it will always be necessary to obtain an overview of certain characteristics of the individual data value chain steps, e.g. the scope of each step, the input data received in a particular step, the output data generated in a step, typical actors involved, typical architectures that are relevant, relevant trends and tools and, finally, the contribution of a particular step to value creation. For this reason, the authors subsequently extend the VDVC presented in Kaiser et al. (2019a) by adding relevant characteristics to each data value chain step and thus aim to answer the following research question:
*What are the relevant steps in developing digital vehicle services that should be part of a data value chain and how can the contribution to value creation be described with characteristics?*

After this introduction and motivation, the authors continue the chapter with a review of background information in Section 8.1. In Section 8.3, the extension of the Vehicle Data Value Chain is presented and described. The authors then apply this value chain to analyze the intermodal mobility service MoveBW (case A) as well as a pothole and driving style detection service (case B) in Section 8.2. Finally, a conclusion and an outlook of the chapter is drawn in Section 8.4.

## 8.1 Background

### 8.1.1 Data as Business Enabler

Tim O'Reilly formulated his extensively quoted principles of Web 2.0 (O'Reilly, 2005) including one about the emerging value of data more than a decade ago. Since then, the hype on how to generate added value from all kinds of available data has continued to grow. Data has become the new buzzword. A book by Mayer-Schoenberger and Cukier (2013) on *how Big Data is changing our world* has become an international bestseller and been cited by researchers more than 5360 times according to Google Scholar. Big Data has received considerable attention from multiple disciplines, including information systems research (Abbasi et al., 2016) and database management (Batini et al., 2015), to name but two.

Due to the exponential growth in the amount of data, for example, an amount of 16 zettabytes (16 trillion GB) of useful data is expected in 2020 (Turner et al., 2014). It is just a logical consequence that data generation, data analysis, data usage – and the new business models associated with it – have found their way into all areas of life. Homes are increasingly equipped with smart meters, a replacement for mechanic measurement of electricity usage, enabling the emergence of digital services to assist home monitoring and to optimize electricity usage. Smartwatches can track the wearer's movements and, create behavioral data and calculate periodic statistics such as daily, weekly, or monthly walking distances including burned calories per day, week, or month. Many people use their smartphones when exercising to gather extra information on their workout effectivity.

Smartphone apps such as Runtastic (2017a) and Strava (2017) help to monitor how and where people run or cycle, automatically calculating route, pace and period-ic statistics including mean speed, time per kilometer, and calories burned. These apps even allow sharing the aggregated data via social networks, thus enabling benchmarking with peers and increasing the joy of exercise. The pattern of collecting, analyzing, and sharing data constitutes the baseline for individual improvements. Instantly calculated and visualized behavioral statistics are easy to compare or share with peers on social media. The collected information per se is not new to these communities. For instance, experienced runners started comparing their real and average time per kilometer using stopwatches a long time ago. However, the simplicity of digital services and the fact that many friends on social media regularly post about their exercising routine has motivated a whole digital generation to track themselves, as 300 million downloads of the Runtastic app (recently renamed to Adidas running) demonstrate (Runtastic, 2020). 30 million app sessions per month in Europe produce a reasonable amount of big movement data, which is sufficient for performing representative data analyses and have led to an acquisition by the sports clothing company Adidas. To summarize, digitalization has greatly simplified data collection and analysis methods which used to be too complex and/or only available to experts. Hence, more and more people are joining the self-tracking movement and, in turn, produce more and more data which can be exploited using novel digital services.

### 8.1.2 A Value Chain for Big Data

In contrast to all previous technical or organizational innovations, the Internet age has made it possible for data volumes to reach undreamt-of dimensions. Big Data refers to the current conglomerate of newly developed methods and information technologies to capture, store and analyze large and expandable volumes of differently structured data. In a definition by Demchenko et al. (2013), the defining properties of Big Data are Value, Variety, Velocity, Veracity and Volume as shown in Figure 35. Exploiting the new flows of data can even improve the performance of companies, if the decision-making culture is appropriate (McAfee and Brynolfsson, 2012).

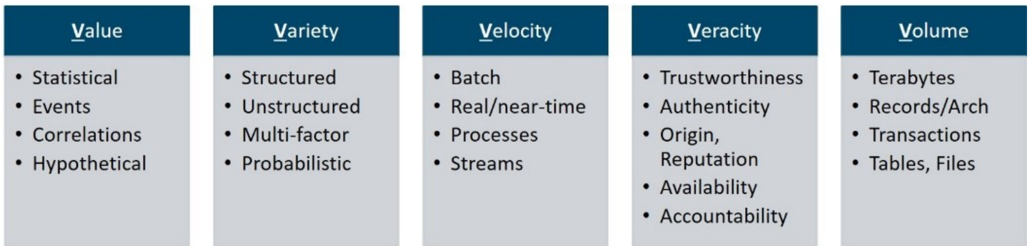


**Figure 35 The 5 Vs of Big Data (Demchenko et al., 2013). (Kaiser et al., 2020b)**

It seems that smart things are increasingly based on big data analysis, which makes it possible to speak of an intimate relationship between those two. While in the Web 2.0 era data was mainly generated by humans sharing user-generated content on portals including YouTube, Wikipedia, or Facebook, the Internet of Things has led to new patterns of data generation driven by machines. Smart, connected objects equipped with all kinds of sensors have now taken over this task (Porter and Heppelmann, 2014 and 2015). The Quantified Self phenomenon is making use of these data generated by things (Swan, 2009 and 2015). Quantified Self refers to the intention to collect any data about the self that can be tracked, including biological, physical, behavioral, and environmental information. Making use of these data to establish applications and services has become a major creator of value. This value is created through multiple activities which are chained together, while the value of the output is steadily increasing.

A company's activities to create and build value were once described by Porter and Millar (1985) with the so-called concept of the value chain. However, this value chain concept can be applied to the data domain to describe activities ranging from data generation to the usage of data in data-driven services for the customer. Data value chains are a model to describe data flows as a series of steps, each of them transforming the value of data. Recently, Åkerman et al. (2018) described a data value chain in the context of production, where data analytics leads to regulations of a production system like in a closed loop control system. Furthermore, the concept of data value chains has been used to describe the value of Linked Data (Latif et al., 2009) and Big Data (Curry et al., 2014) as illustrated in Figure 36. As modern vehicles are likely to produce big data (e.g. from and for (semi-)automated vehicles), the Big Data Value Chain including several steps of Big Data transformation in the process of generating the data-driven result with the maximum business value is of high relevance to the automotive sector (Xu et al., 2017).

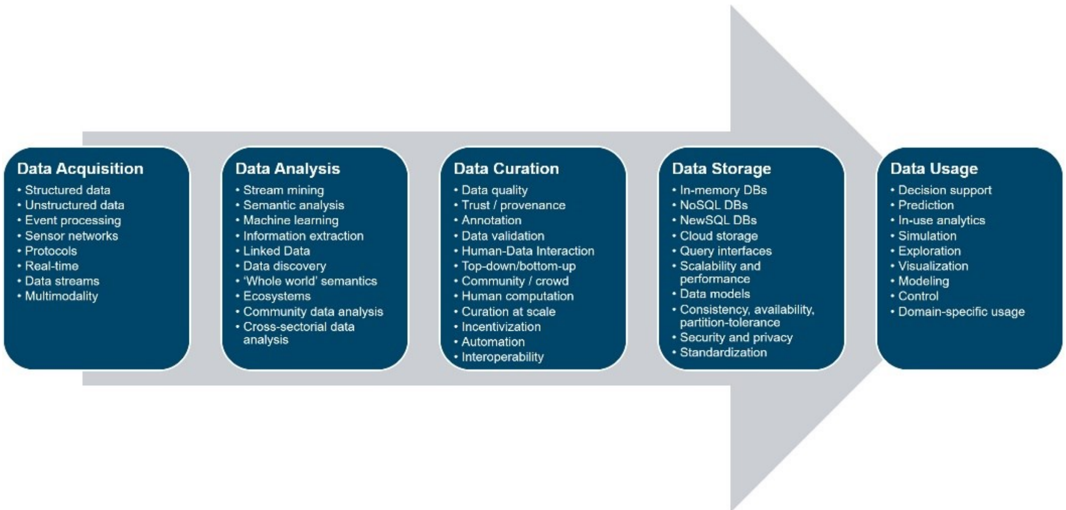


**Figure 36 The Big Data Value Chain of Curry et al. (2014) / Curry (2016). (Kaiser et al., 2020b)**

## 8.2 Evaluation of the VDVC

### 8.2.1 Case A: Description of the intermodal mobility service MoveBW

A regional, intermodal mobility service called MoveBW helps to increase the compliance rate of transport users (e.g. the percentage of people using a park and ride option) in relation to the current transport strategy of the region. The strategy offered by an European industry consortium mainly aims at meeting air quality targets and reducing traffic jams all over the federal province of Baden-Württemberg (Germany), including its provincial capital Stuttgart.

Geographically situated in a valley basin, Stuttgart, like all cities in valley basins (e.g. Graz), struggles with air pollution through fine dust. Thus, the city of Stuttgart continuously develops transport strategies to better comply with air quality regulations. In the past, these strategies were communicated to the public using radio traffic messages or electric traffic signs only. However, the compliance rate and thus success were comparably low. The MoveBW mobility service smartphone application aims to increase said compliance rate, especially that of visitors new to the region. It does so by including easy-to-use routing functionalities which are connected to rewards: Bonus points are granted if a user follows the recommended route. Collected bonus points can later be exchanged for immaterial or monetary values.

The intermodal journey planner allows users of the MoveBW smartphone application to plan their trips in advance. They can pick their preferred combination of transport modes from different options suggested to them. Additional information is displayed, not only showing travel time, but also eco-friendliness, travel costs and incentives gained (e.g. public transport vouchers and $CO_2$ savings). Moreover, it is possible to directly book tickets for the different modes of transport included in their preferred journey and yet to receive only one bill. In this way, transport services such as public transportation, car sharing, bike sharing, and parking space management are integrated conveniently, encouraging users to make efficient use of all modes of transport. The application also provides on-trip navigation and information on traffic obstructions such as construction works or accidents.

The MoveBW services are currently being monitored and evaluated in an extensive test phase. Based on the findings, both the digital service and traffic control strategies will be revised, aiming to maximize favored effects on the individual mobility behaviors of traffic participants, for example by applying different strategies for daily commuters and visitors. The smartphone application is planned to be released in the first quarter of 2019. Mock-ups of the current design are shown in Figure 37.

**Figure 37 The MoveBW smartphone application provides functions for intermodal journey planning, traffic information, ticketing and on-trip navigation. (Source: Strukhoff et al., 2017)**

A special challenge regarding data management is the multitude of data sources for the intermodal routing algorithms in the MoveBW App. The Vehicle Data Value Chain introduced in Section 8.3 helps to provide a clearer view. Its application to the underlying data transformation process, from Data Generation to Data Usage, is shown in Table 15.

**Table 15 An overview of the MoveBW-Service. (Source: Kaiser et al., 2019a) (Kaiser et al., 2020b)**

| VDVC step | Description of MoveBW-Service |
|---|---|
| Data Generation | Various sensor data and basic reference data is considered, e.g.<br>- floating car data: average mean travel time per road segment based on anonymized GNSS data of vehicles,<br>- stationary traffic measurement: rate of flow for single measurement locations,<br>- public transport: schedule and sometimes occupancy rate,<br>- car park interfaces: occupancy rate,<br>- park & ride interfaces: occupancy rate,<br>- air quality measurement units: measurements and forecast (includes weather forecast); |
| Data Acquisition | Querying web APIs from the various data sources. Additionally, the smartphone App which is described in Data Usage provides GNSS information, which is used for on-trip routing and to detect which means of transport the user actually uses to be able to reward them if the recommended option is used. |
| Data Pre-processing | Annotation, normalization and semantic extraction of data. Transformation of data to meet a common reference basis (in this case a public transport grid, no typical geo-coordinates). Furthermore, GNSS data from the smartphone App is anonymized (start- and end-trajectories are truncated). In this step the data is hosted in a distributed database system (e.g. PostgreSQL cluster) |
| Data Analysis | A dynamic routing algorithm which also takes the provided intermodal transport strategy, $CO_2$ savings, and personal preferences into account. A self-developed algorithm which utilizes pgRouting (an open source project to extend PostGIS/PostgreSQL to provide geospatial routing functionality) and the popular Dijkstra algorithm (to find the shortest path between nodes). Provision of routing recommendations (weightings for routes) through this algorithm. |
| Data Storage | A distributed database system, e.g. a PostgreSQL cluster |
| Data Usage | The MoveBW App currently being developed should help the commuter to choose a mode of transport and guides the commuter to the selected destination in compliance with environmentally-oriented traffic management strategies. |

In case of MoveBW, where all steps of the MoveBW service are known to the authors, the VDVC provides a framework to describe the service layer by layer and thus also helps others to understand the service and its underlying value chain.

In the next subsection, the development of a pothole and driving style detection service is described using the VDVC.

### 8.2.2 Case B: Description of a Pothole and Driving Style Detection Service

Generating value out of vehicle data is a challenging task: For this purpose, vehicle data analytics has become an important technique in identifying the value of generated vehicle data. However, to exploit this value in products and services, several steps must be performed, and several (not only technical) challenges have to be solved. In the beginning, an appropriate analytics question must be identified such as e.g. identify the driving style of the driver from vehicle data, detect the road surface quality, identify potholes on roads, or predict the engine's wear.

Then, vehicle data must be captured: Three different approaches for data capturing are possible: the installation/use of own sensors within the vehicle to record vehicle movements and other contextual information, the connection of a vehicle data logger to the vehicle's on board diagnostic (OBD) interface to capture vehicle data such as vehicle speed or RPM, or the installation of a professional Controller Area Network (CAN) logger to obtain even more vehicle data from the vehicle's CPUs such as for example the state of vehicle assistance systems or the steering wheel angle. While the first option is probably the simplest one, it can only record contextual data and track the movement of the vehicle, but it does not allow access to vehicle sensors. The second option can provide already access to some vehicle sensor data such as vehicle speed or engine temperature, which is relevant for testing whether the vehicle's emissions are still within tolerance. The third option in theory provides access to all vehicle sensor signals, but only if the device listening to the CAN bus can decode the streamed raw CAN bus data to readable data, requiring either the vehicle manufacturer or the respective vehicle CPU manufacturer to provide the necessary decoding information (usually referred to as CAN-DBC files).

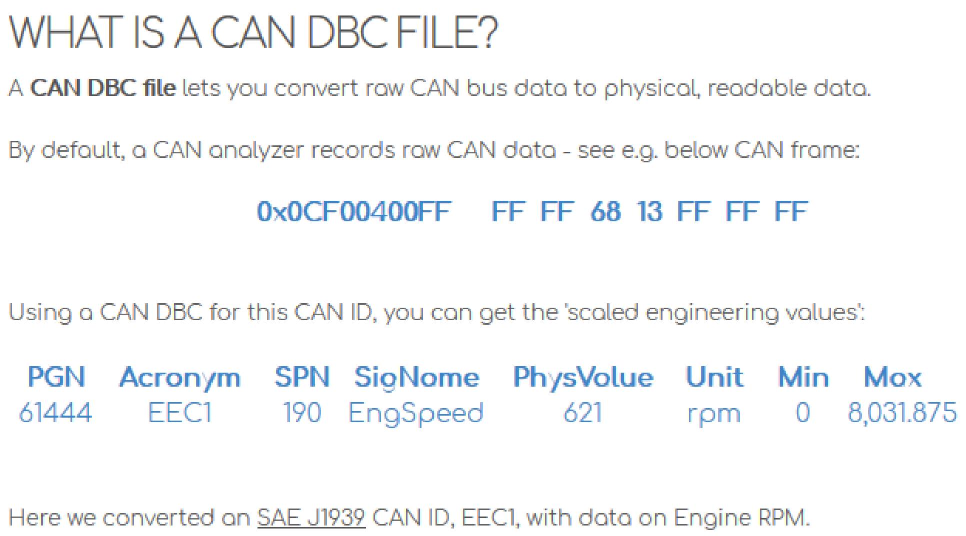


**Figure 38 CAN DBC files. Source: CSS electronics (2020). (Kaiser et al., 2020b)**

Different data loggers may store the data in different formats. Typically, they can collect multiple signals at once, which are all transmitted on the same wire. Thus, the logger needs to know and save at least three different properties of the data: What was measured, what was its value and when was it measured. This naturally leads to a tabular format very similar to the example depicted in Table 16.

While this format is convenient for the logger to store data, it is much less suited for a statistical analyses or automated processing of the data. There are three main difficulties: First, several signals are mixed together in one column, creating the need for grouping and filtering even before very simple operations. Second, there can be multiple signals that were measured at the same time, requiring the analyst to investigate multiple rows at once to check a single instance in time. The third difficulty lies in the varying sampling rates of the signals. Each signal may have been captured with a different rate and even within a single signal, smaller deviations of the sampling rate are possible and common. Clearly pre-processing of the captured vehicle data is needed to make it better explorable for data analysts.

**Table 16 Vehicle raw data structure (example). (Kaiser et al., 2020b)**

| Timestamp | Signal Name | Signal Value |
|---|---|---|
| 2019-9-13 5:28:36.206089 | RPM | 1500 |
| 2019-9-13 5:28:36.226331 | Acceleration-X | 0.476 |
| 2019-9-13 5:28:36.245312 | Vehicle Speed | 39 |
| 2019-9-13 5:28:36.268915 | Engine oil temperature | 90 |
| .. | .. | .. |

After the required vehicle data is stored, a series of further steps must be performed to prepare the data for analysis. This data (pre-)processing process can be quite comprehensive and depends very much on the respective analysis question to be solved, e.g. the detection of potholes from vehicle data. A crucial step in this process is the alignment of the coordinate system of data logger and vehicle. Many signals are vector valued, with acceleration as the maybe most prominent example. To simplify analyses and interpretations, it is highly desirable to express these vectors in the reference frame of the car, i.e. x-Acceleration should be the component in the x-direction of the car / the driving direction. In general, one cannot assume that the logger was mounted such that its internal coordinate system corresponds to the one of the vehicle. This is especially true when cheap devices that are mounted by end-users are used. Any misalignment of the reference frames needs to be detected and corrected prior to analysis.

As with most other data types, vehicle data signals should be searched for missing values, wrong values, and outliers and these should be removed. Some signals may contain a lot of noise and must be smoothed. To separate the signals into different columns the data should be transformed using the ‘signal name’ as pivot. Simultaneously, it makes sense to resample each signal to a common sampling rate from the analysis’ viewpoint. The “right” sampling rate again depends on the question the be answered. The result is than in a similar form as depicted in Table 3. Now each row corresponds exactly to a point in time and the time interval between the rows is constant, in this example 0.1s / 10Hz.

**Table 17 Structure of pre-processed vehicle usage data (example). (Kaiser et al., 2020b)**

| Timestamp | Engine Speed | Acceleration-X | Vehicle Speed | .. |
|---|---|---|---|---|
| 2019-9-13 5:28:36.20000 | 1500 | 0.477 | 39 | |
| 2019-9-13 5:28:36.30000 | 1501 | 0.479 | 40 | |
| 2019-9-13 5:28:36.40000 | .. | .. | .. | |

The data prepared in this way can now be used to work on the vehicle data analysis question and/or to search for interesting events (such as potholes for example). Depending on the type of event, multiple signals can be relevant. Events should usually be post-processed to combine separate events, which are only divided by a short-time interruption, into a single event. The recorded events may be linked with weather and position data, so that for each event the time and place of occurrence as well as the prevailing weather is known.

For different types of events, different detection methods need to be employed. One can detect a pothole event (driving over a pothole) by investigating acceleration values and rotation rates as follows: Consider the acceleration normal to the road, as well as the vehicle's rotation around its lateral axis ('pitch') The acceleration readings will exhibit a distinct spike, while a certain pattern is simultaneously visible in the rotation rate: When the front tires are in the pothole, the front of the vehicle is lower than the rear, if the rear tires are in the pothole, it is vice versa, causing a rotation around the lateral axis. This results in a typical "pitch" movement that can be detected. In a last step, the results of the analysis – in this case the detected potholes – can be visualized on a map. In this case it supports drivers in not choosing bad roads, or support road operators in better maintaining roads.

To detect strong acceleration and braking events, the signals vehicle speed, acceleration in the direction of travel and rotation around the lateral axis ("pitching") are particularly suitable. The "pitching" is caused by the change in weight distribution when the speed changes: when a vehicle is accelerating, more weight moves to the rear axle – the rear drops and the front rises. When a vehicle is braking, the opposite is true. These movements can be detected. However, since detection using only a single signal can be prone to error, the developers always use several signals in the algorithm, which must all deflect simultaneously to trigger detection.

The driving styles of drivers can differ in many facets (e.g.: comfort level, gear choice, aggressiveness). Depending on the type of vehicle the driving style may have a large influence on fuel/power consumption, component-wear and road safety. In an attempt to quantify this influence, the developers use all calculated events to calculate a 'risk score' that indicates how unsafe a single trip was. The more safety-related events per time unit occurred in a trip, the higher the value. Furthermore, the developers consider the influence of environmental conditions in their calculations. For example, heavy braking in rain will result in a higher risk than the same braking on a dry road. To make the risk score interpretable, the developers normalize it using the scores from all available trips as a basis. The developers then present the value as statistical rank, for example a value of 56.72% means that this trip is safer than 56.72% of all trips in the database. In a map visualization, the driver is presented the trip with markers indicating start and stop positions, as well as locations for safety-relevant events.

Based on this methodology, a smartphone application, shown in Figure 39, has been developed for drivers interested in monitoring their driving style.

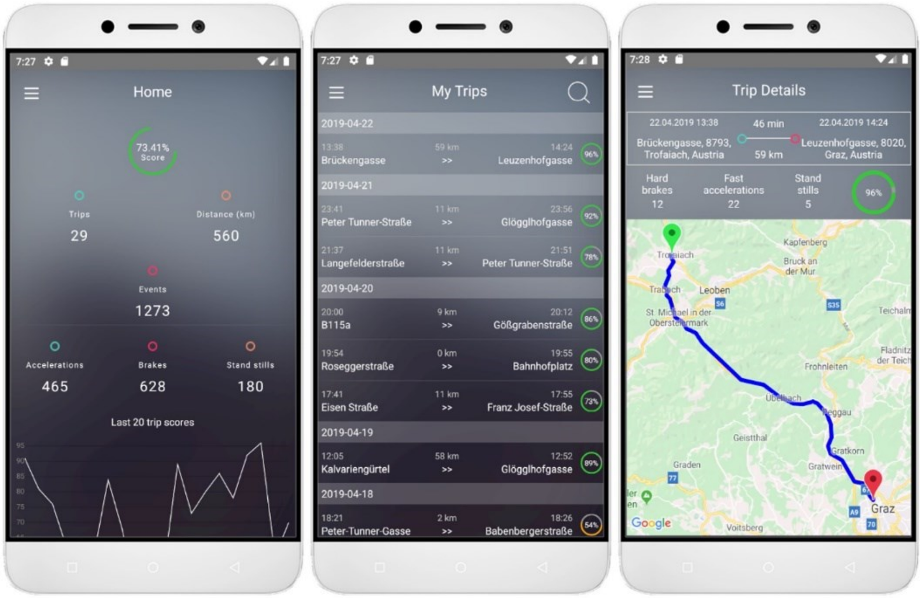


**Figure 39** **A smartphone application for driving style detection. (Kaiser et al., 2020b)**

On the left screen named *Home*, the driver has an overview of his trips. In the presented figure, his overall score is 73.41%; he has 29 trips with a total distance of 560 kilometers. In these trips 1273 events have been detected, which are composed of 465 acceleration events, 628 brake events and 180 stand-still events. On a second screen named *My Trips,* which is displayed in the center, a list of the most recent trips, grouped by date, is shown. For each trip, the information on which location and at which time the trip started and ended is displayed together with the trip score and the trip distance. Selecting one of the trips opens a third screen named *Trip Details*, where additionally the events are decomposed into categories and the trip is visualized on a map.

Now that the idea of this service has been described, the authors want to show in the following table how clear and comparable the service becomes by using the VDVC.

**Table 18 An overview of the pothole and driving style detection service. (Kaiser et al., 2020b)**

| VDVC step | Description of pothole and driving style detection service |
|---|---|
| Data Generation | Vehicles are equipped with data loggers that record the signals required for pothole and driving style detection (e.g. speed, acceleration, rotation, position, etc.) These data loggers are connected to the on-board diagnostic interface of the vehicle and additionally generate acceleration, rotation and GPS data. |
| Data Acquisition | Vehicle movement data including OBD measurements as well as acceleration, rotation and position measurements is periodically recorded and imported as raw vehicle data into a local PostgreSQL database on the data logger. The collected data is made available as a data stream or as manually exported files in a PostgreSQL database running in the cloud. |
| Data Pre-processing | The pre-processing of the vehicle data includes the alignment of the datalogger's coordinate axis with the trajectories of the vehicle, the search for missing and incorrect values and outliers and their elimination, the smoothing of the signals to reduce noise and the interpolation of all signals to a useful sampling rate. Additionally, contextual weather data is integrated. |
| Data Analysis | For pothole detection, the acceleration perpendicular to the road and the "pitching" of the vehicle (i.e. the rotation around the transverse axis) are used. If these exceed certain threshold values, a pothole event is generated. In comparison, vehicle speed, acceleration in the direction of travel and rotation around the transverse axis ("nodding") are used to detect events relevant to driving safety, such as strong acceleration, braking and cornering maneuvers. If these exceed certain threshold values, a harsh acceleration, braking and cornering event occurs. |
| Data Storage | The events calculated in the analysis phase (harsh acceleration, braking, cornering as well as potholes) are stored in the PostgreSQL database together with their GPS locations and the corresponding weather information to visualize them on maps and perform additional statistical analyses, such as calculating a risk score for a single trip, taking into account the amount and severity of detected events per trip length as well as the respective weather conditions and a cumulative risk score for a driver. |
| Data Usage | Drivers should be provided with information to improve their driving. The application shown in the figure above should help the driver to monitor his own driving and compare it with the driving of other drivers in order to improve driving safety. Finally, the application can visualize detected potholes so that the driver can avoid driving into these potholes. |

### 8.2.3 Big Data Based on Vehicle Usage Data

The automotive industry is also constantly finding innovations for its vehicles as a result of electrification and comfort requirements. For example, mechanical components such as hand brakes or window lifters are increasingly being changed to electronic versions, such as the electric hand brake and electric window lifters. The status (handbrake is applied or released) and its process status (handbrake is applying/releasing) can be captured and used as input for vehicle safety checks and other features. An applied handbrake will automatically be released if the driver starts driving to prevent damage. The data generated through all these vehicle functions can be captured and used within other scenarios, e.g. to create statistics on how often a window is opened/closed or how often somebody is wedged in.

Also due to the common practice of vehicle development to purchase many components from suppliers, many vehicle sensors have so far only been used to provide and support a specific functionality and to increase comfort and safety, although these vehicle sensor data may also be interesting for third parties. As sensors and car features may widely differ from manufacturer to manufacturer and even per car variant, there is not only one single truth about how much data is effectively generated by a modern vehicle today. For instance, the participants from the European research project AutoMat (Automat, 2017) state in a deliverable (Automat, 2018c) that about 4000 CAN bus signals (one signal could be one measurement value) per second create up to 1 GB of data per CAN bus (without mentioning a sample rate). According to Pillmann et al. (2017b), there are "usually 4-12 CAN busses in one car" (with varying amounts of input signals). This clearly shows the high amount of data generated as a by-product during vehicle use.

For highly automated driving, several camera, radar and LiDAR (Light Detection and Ranging) systems are currently being implemented in the vehicles to cover every corner of the vehicle environment. Autonomous vehicles may be forced to exchange information with other vehicles (V2V) and with the infrastructure (V2I), which will boost the amount of available vehicle data enormously in the future. Considering different countries and different patterns of individual driving and mobility behavior, bringing highly automated driving into practice will be a grand digitalization challenge.

Although only part of this data is available for digital vehicle services (e.g. the high sampling rates generate such large amounts of data that the limits of data transmission are exceeded, which would require re-sampling at a lower rate or some signals are simply not relevant) and while only a portion of these data will be made accessible due to safety reasons (EU, 2013), the remainder of accessible sensor data from modern vehicles will most likely be sufficient to design and develop a reasonable number of novel digital vehicle services for various stakeholder groups, including individual drivers, various organizational customers, government authorities, and the automotive industry (Kaiser et al., 2017b). To sum up, modern vehicles already constitute impressive generators of big vehicle usage data.

## 8.3 A Value Chain for Vehicle Usage Data

### 8.3.1 Quantified-Self

Digital natives like to have access to services anytime and anywhere and are therefore willing to let their mobile devices such as smartphones and smart watches generate data around the clock. Increasing the knowledge about oneself and eventually enabling new discoveries while performing physical activities including running or cycling has turned into a business-relevant phenomenon. The behavior of turning collected data about oneself into actionable knowledge and insight which is valuable for other stakeholders, too, has been termed Quantified Self. Interestingly, the quantified self-phenomenon has recently been successfully transferred to the automotive industry by US-based start-ups. In this sense and quite analogously, Quantified Vehicles (Stocker et al., 2017a) imply a successful transformation of data from different kinds of sensors related to the vehicle (in-vehicle sensors, smartphone and wearable sensors used by the driver) into actionable knowledge, e.g. on the behavior of the vehicle. This way, they generate value for different kinds of stakeholders that are part of digital vehicle data service ecosystems such as insurance or fleet management providers, finally resulting in novel digital services based on vehicle data in various domains (Kaiser et al., 2018b; Kaiser et al., 2019b).

Self-tracking with consumer devices, as shown in the example of Runtastic (Adidas running), can also be transferred to vehicles: Vehicles already collect a large amount of operating data via sensors and control units that ensure the functionality of the vehicle. However, these big vehicle data could be used to enable a series of apps and services for different customer groups. The market value for vehicle usage data is considered to be even higher than for other markets due to the importance of vehicles in first world countries. A number of US-based ICT start-ups seized this opportunity, now offering smartphone and web applications providing insights into vehicle-generated data, after they received up to €25 million of funding from investors (Stocker et al., 2017a). Interestingly, while some car manufacturers and suppliers (e.g. Magna International, Continental ITS, and BMW i Ventures) are among the investors, forming strategic partnerships with start-ups, others participate in research projects and try to keep data-related business in their own area of influence. This holds for Volkswagen, for example, which coordinates the EU project AutoMat to develop a marketplace for vehicle lifecycle data (Stocker and Kaiser, 2016). Furthermore, recent reports from the German automotive industry association (VDA) suggest that car manufacturers “have to hold a stronger position in the future and may limit the capabilities of third parties to freely access car data." To summarize, the potential of vehicle usage data seems to be such that it has become a battle worth fighting (Kaiser et al., 2017b). How vehicle usage data generates value leads us to the next subsection in which the Vehicle Data Value Chain is described.

### 8.3.2 The Vehicle Data Value Chain (VDVC)

To systematically describe the transformation of data into valuable services, the concept of value chain can create a suitable structure and framework. In this regard, the Vehicle Data Value Chain (VDVC) is proposed as a lightweight model. The VDVC was derived from the Big Data Value Chain (Curry et al., 2016, illustrated in Figure 40). The authors adapted Curry's value chain regarding the name, number and order of stages to reflect the authors' experiences from research projects in the automotive domain. The stage of *Generation* (of vehicle usage data) was added as a separate stage to explicitly reflect the origin of the data (e.g. in-vehicle or related sensors). The stage *Acquisition* (of vehicle usage data) corresponds to Curry's Data Acquisition. Moreover, the authors have changed the order of Curry's stages of analysis and curation since the authors interpret the terminology differently. For example, Curry seems to include normalization procedures implicitly within machine learning in the stage of Data Analysis, whereas the authors consider this as an important separate pre-processing step which correlates with Curry's stage of Data Curation. Hence, the authors have re-named Curry's stage of Data Curation, *Pre-processing*, which is followed by the stages *Analysis*, *Storage*, and *Usage* (in each case: of vehicle usage data), as visualized in Figure 40. As the result of the processing could be the input for further analysis, an arrow back to Acquisition indicates the possible a circular path.

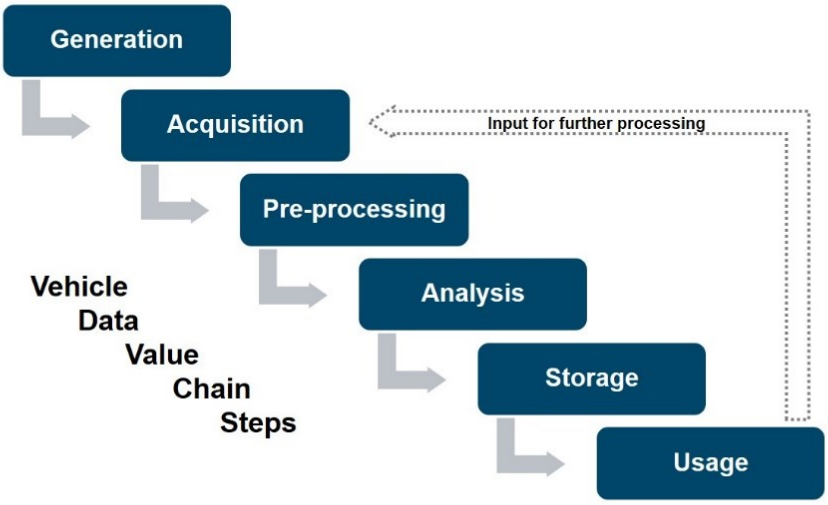


**Figure 40 The Vehicle Data Value Chain derived from Curry (2016) and based on Kaiser et al. (2018b) and Kaiser et al. (2019a). (Kaiser et al., 2020b)**

Furthermore, to compare digital services based on vehicle data and to understand and explain the data-related challenges associated with them, the authors added eight characteristics to each value chain step: *i) Description / Scope* to describe the scope of the step, *ii) Input examples* and *iii) Output examples* to name possible inputs and outputs per step, *iv) Actor examples* to name relevant actors in this step, *v) Architecture examples* to describe which architecture usually is used in a specific step, *vi) Trend examples* to name current trends in the specific value chain step, *vii) Tool examples* to name possible tools and *viii) Contribution to value creation* to summarize the contribution of this step to value creation. The single value chain steps are shown in Figure 41 and are described in the following sub-subsections.

#### 8.3.2.1 Generation (of vehicle usage data)

This step has the scope of generating measurements through any sensors which can capture condition data directly (engine RPM or vehicle speed) or indirectly (road surface). In the case of direct influence, the authors see three main data sources: In-vehicle sensors, smartphone sensors and sensors in individual user devices (e.g. a pulse watch). An indirect data sources can be literally any data source that provides information on the state of a vehicle, its driver or surroundings; an example could be a road operator camera to display traffic flow. This process step is essential for the vehicle data value chain, since the data origin determines the reliability and the type of influence (direct, indirect). The current trend to equip modern vehicles with ADAS functionalities (e.g. through the use of radar and lidar sensors for better detection of the driving environment) increases the amount of data generated and the possibilities for use cases once more.

#### 8.3.2.2 Acquisition (of vehicle usage data)

This step describes the process of collecting the generated data. In-vehicle sensor data is not directly accessible as it is secured in order to safeguard vehicle functionality and is therefore only exchanged between the various electronic control units via one of the vehicle's internal bus systems, e.g. CAN bus. However, a filtered quantity of this sensor data is accessible via the On-board diagnostic (OBD) interface (Turker and Kutlu, 2015), which is intended to be used by service staff to read out the generated error messages. It is therefore possible to develop plug-in devices with an internet connection, thereby effectively using the OBD-port as a source of sensor data. There are already some professional solutions with data acquisition devices built into the vehicle, which read signals directly from the CAN bus in an unfiltered way. To meet the requirements of the EU Directive 2010/40/EU – establishing inter alia the costless provision of universal, road safety-related minimum traffic information (EU, 2013) – a standardized interface would be feasible sooner or later. Data from smartphone sensors is acquired using specific applications, capable of gathering and transmitting data. In the case of external data sources, the main issues are the varying availability and quality levels of the data. For example, APIs usually limit the number of requests allowed per time interval, so the acquisition process must be adapted to meet these thresholds. Gathered data is stored for further processing; the chosen storage and format heavily depend on the subsequent processing steps.

#### 8.3.2.3 Pre-processing (of vehicle usage data)

This step consists of the process of data preparation and integration. It is the sum of any anonymization, annotation, cleansing and normalization activities before any data analysis is conducted. Sensor values including private user information, erroneous sensor readings, different sensor sampling frequencies or unsynchronized data are examples of issues addressed in this stage. Data quality has a high impact on service quality. For instance, if the accuracy of the GNSS signal is low, a trip may not be linked to the correct road and may lead to wrong conclusions.

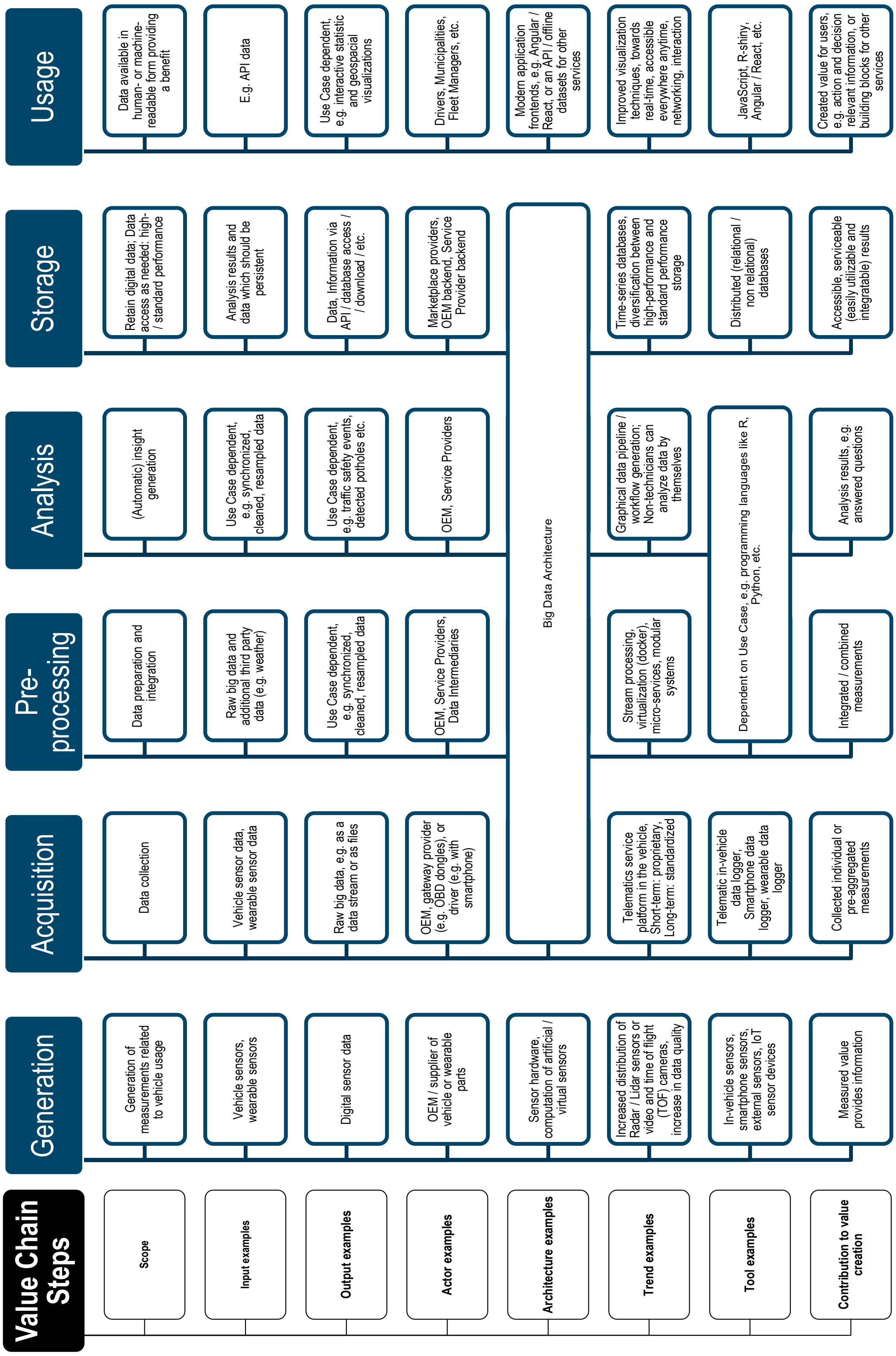


**Figure 41** **The Vehicle Data Value Chain derived from Curry (2016) and based on Kaiser et al. (2018b) and Kaiser et al. (2019a) extended with characteristics. (Kaiser et al., 2020b)**

#### 8.3.2.4 Analysis (of vehicle usage data)

This step is the process of automatic insight generation, with the purpose of extracting useful hidden information. This involves linking data from different data sources, exploring the data, performing statistical analyses and using machine learning algorithms to detect latent information hidden in the data. For instance, weather data can be linked to vehicle speed on a particular road to determine whether the driver is driving differently in wet or icy conditions. Weather data can be linked to acceleration data to determine whether a driver is driving aggressively in bad weather conditions.

#### 8.3.2.5 Storage (of vehicle usage data)

In this step of the value chain, proper data access is established. It is already defined in the Big Data Value Chain as “the persistence and management of data in a scalable way that satisfies the needs of applications that require fast access to the data” (Curry, 2016). In the case of vehicle sensor data, persistent storage is usually achieved by using a combination of classical relational databases (for metadata), Big Data file systems (for raw input data) and so called “time series databases” to store data that changes with time, which allow fast analyses on the stored contents.

#### 8.3.2.6 Usage (of vehicle usage data)

The final step deals with making the data available in human- or machine-readable form (or both, as required). It includes all kinds of user or software interaction with the collected data and any conclusions derived from it in the above-mentioned process. The retrieved data could either be regarded as the end result of the process, being presented more or less directly to end users, or it could serve as input for further processing steps, thus forming a circular path in the processing chain.

## 8.4 Conclusion and Outlook

An increasing number of digital services based on vehicle usage data are offered on the market and are increasingly used and demanded by users. Digitalization has not only become an important driver of innovation in the automotive sector, but may also change the balance of power in the automotive sector in the long term. With the background that society is strongly driven by mobility, it is almost the authors duty to examine the emergence of digital services based on vehicle usage data more closely. Consequently, in this chapter the authors have looked at a way of better describing and structuring digital services based on vehicle usage data. After a comprehensive analysis of related work, the authors have reviewed two different digital services by using the VDVC for a better structured description of how value is created. Using the VDVC model, the authors explicitly describe which activities must be carried out in the individual steps of the value chain in order to finally enable these two services.

As an outlook, it should be mentioned that digital vehicle services and the required technological infrastructure to facilitate data acquisition, pre-processing, analysis and storage, are currently a hot topic in the automotive domain. There are already ideas for using blockchain technology and brokers to make data sharing more transparent and secure, as described in Kaiser et. al (2019c). Yet, while some car manufacturers invest in start-ups, others limit access to data via the OBD interface, arguing that they are not suitable for digital vehicle services (VDA, 2016; ACEA, 2016). In contrast, the European Automobile Manufacturers Association ACEA promotes car data sharing (ACEA, 2017).

Regulation (EU) No. 886/2013 (concerning the Directive 2010/40/EU on Intelligent Transport Systems ITS), published by the European Commission, has actually been regulating the provision of universal, road-safety relevant minimum traffic information to users free of charge for years and calls on car manufacturers to make safety-relevant data available to the public via national contact points (EU 2013). While the vehicle manufacturers have long referred to the no longer up-to-date transmission standard based on WLAN technology (e.g. G5), several EU-wide initiatives (such as the C-ROADS initiative) have not given up, extending the development to telecommunications technologies (e.g. 4G, 5G) and presenting a concrete implementation plan for C-ITS services with Day 1 Applications. Since the end of 2019, the latest Volkswagen Golf is the first series-production vehicle on the market to use this data exchange standard. The C-ROADS initiative of several EU member states and road operators aims to use C-ITS services to enable the transmission of infrastructure information (e.g. roadside units) to the vehicle cockpit, e.g. to inform about dangerous situations, e.g. a vehicle backing out or pedestrians in the crosswalk behind the next bend. (C-ROADS, 2017).

At the same time the International Organisation for Standardisation (ISO 2017) has set up a standardization work group titled *ISO/TC 22/SC 31/WG 6 Extended Vehicle/Remote diagnostics* (ISO 2018) to inter alia define access, content, control and security mechanisms for the provision of vehicle data for web services (VDA, 2016).

Additionally, current EU-funded projects such as EVOLVE are developing solutions to ease the integration and fusion of multiple data sources for the purpose of service and business development using Linked Data (EVOLVE, 2019; Latif et al., 2009). "Linked data is a lightweight practice for exposing and connecting pieces of data, information, or knowledge using basic web standards. It promises to open up siloed data ownership and is already an enabler of open data and data sharing" (Rusitschka and Curry, 2016).

To conclude, the authors expect the market of digital services based on vehicle usage data to grow tremendously in the future, as the combination of vehicle data with data from external sources (e.g. weather data, traffic data, open data) will enable new scenarios for digital vehicle services.

# 9. Towards a Generic IoT Platform for Data-driven Vehicle Services [15]

| Summary and Author Contribution |
|---|
| **Concepts and Prototypical Implementations along the VDVC (Paper 1/5)** |
| This chapter presents a concept to collect, process and use vehicle data for data-driven services. Challenges like independency, scalability, and flexibility while ensuring e.g. privacy, and accountability must be considered for such an IoT platform. Thus, this chapter presents a conceptual architecture of a generic IoT platform for enabling such *Data-driven Services* and describes how this platform can be implemented, ranging from the gateway device (Vehicle Data Logger) capturing the vehicle data, to a vehicle data-driven service application for the driver. |
| As part of the dissemination of our work in research project SCOTT (all authors of the paper were involved), we published the concept of our own *Data-driven Service* based on vehicle data developed under my administration in this paper in a generalized form. Here, besides the contribution to related work for *Quantified Vehicles*, I introduced the hardware to collect vehicle data from the OBD-2 interface, the *Vehicle Data Logger,* as it was its data that made analyses possible for us in this and other research projects in the first place. The Vehicle Data Logger hardware was initially developed by the former student employee Benjamin Fischer, supervised by the last author and me. When Benjamin left, I took over the assembling and the maintenance of the hardware and repeatedly persuaded colleagues to install the hardware in their vehicles to collect data. While author four developed the software of the data logger, he and I both individually used the collected data to build visualizations for end-users. For example, Figure 46 shows my overview of telemetry data, and Table 19 show the bill of material for the main components of the data logger. Moreover, the procedure presented in "Evaluation through a demonstrator" (Section 9.4) describes steps of the VDVC, because the experiences we made in SCOTT, which are presented here, have contributed to the development of the VDVC. |

## 9.1 Motivation and Goals

Automotive electronic systems nowadays expand rapidly and massively in becoming fully connected and capable of extended functionalities. Modern vehicles typically consist of several dozens Electronic Control Units (ECUs) linked to communication networks processing data to ensure a vehicle's functionality. This rapid development is even more visible in the advances in vehicle software which are increasing both in complexity and functionality, relying on platforms developed entirely for the automotive domain (e.g., the standard Automotive

[15] The content of this chapter is based on

Papatheocharous, E., Frécon, E., **Kaiser, C.**, Festl, A., and Stocker, A. (2018). Towards a generic IoT platform for data-driven vehicle services. In *2018 IEEE International Conference on Vehicular Electronics and Safety (ICVES)* (pp. 1-6). IEEE. DOI: 10.1109/ICVES.2018.8519505

Open System Architecture – AUTOSAR). These denote extended software programming development environments with underlying subsystems, including language, runtime, components and other associated libraries.

These recent developments offer a range of possible service types classified in Axelsson et al. (2014) as:

i. product services, to extend a product with new or improved functionality (usually targeting an end-user group),
ii. process services, to improve the operation of a product as part of a larger process (often oriented to the product owners and having limited control over the actual product development),
iii. lifecycle services, to use data in ways that improve the associated lifecycle processes, including predictive and preventive maintenance (usually targeting the manufacturer of the product), and,
iv. extended services, use data from products to improve the operation of decoupled to the product products or services (targeting administrators and governments).

Taking into account the new advances and trends in the research field of vehicle engineering, this work lists a number of challenges starting from a trend, coined as the "Quantified-self" (QS) or "Quantified Vehicles" (QV), which occurs as the world is becoming more and more equipped of smart connected things with sensors and actuators. Nevertheless, the business and technical benefits are not entirely obvious, i.e., how can they be reaped in a successful way, and several contenders (from start-ups to automotive suppliers, manufacturers up to large software companies) are in the search of the optimum platform and position to host and release it. Thus, a popular trend is the search of a universal Internet of Things (IoT) platform from a variety of emerging commercial and open source solutions. There are specific requirements that need to be satisfied, including offering functionality that opens up to new opportunities and more rigid challenges, such as privacy, safety and complying with regulations, e.g., the EU regulation of data privacy enforced recently in all businesses (General Data Protection Regulation – GDPR) and the General Safety Regulation (EC) No 661/2009 in the automotive sector.

This chapter describes the conceptual architecture of a generic IoT platform to support security, safety, automation, operations and management, based on the concepts mentioned above and the new advances and trends outlined, enabling data-driven services in the vehicle domain. As a preliminary proof-of-concept, the authors have developed a demonstrator which supports two-way connectivity needs of third-party cloud services for quantified connected vehicles. The proposed solution accepts, stores and processes vehicle data obtained from a Vehicle Data Logger connected to the OBD-II interface of the vehicle. Raw vehicle data is then analyzed on the proposed cloud-based platform with a focus on extracting driving styles and behavior. The chapter shows how the solution enables the operation and management of third-party access to collect, store, transform, process and visualize data, and then feeding back the result to human drivers or other stakeholders.

The main contribution of this chapter is the presentation of a generic IoT platform, building on open source technologies and containers, for seamlessly enabling novel data-driven connected vehicle services. The authors then validate the platform through a demonstrator, that is feeding interesting information, related to trips and driving behavior, gained from processing real data from vehicles to drivers and other stakeholders through a mobile and a web application.

The chapter is organized as follows: Section 9.2 covers the background work and challenges in the area, Section 9.3 analyses the proposed conceptual architecture and solution walk-through, Section 9.4 is about the proof-of-concept implementation through a demonstrator, Section 9.5 provides a discussion of the main business and technical benefits of the solution, and, finally, Section 9.6 describes future research steps and conclusions.

## 9.2 Background and Challenges

In this section, the authors introduce the notion of Quantified Vehicles, as an extension of the Quantified-self concept, and the authors also describe the authors' journey towards realizing a generic IoT platform that proposes a technical solution to support security, safety, automation, flexibility and scalability.

### 9.2.1 From Quantified-self to Quantified Vehicles

The trends of digitizing real-world events and self-digitization have become increasingly important and have transformed humans to data generators (McFedries, 2013). The "Quantified-self" (QS) is a term coined to describe the quantitative assessment of measurable characteristics about a person, including biological, physical, behavioral, and/or environmental aspects (Swan, 2009).

QS has become a major creator of value with the increase number of applications on smartphones and other consumer devices. E.g., Strava and Runtastic are popular applications to analyze and compare sports behavior (running, biking, etc.) and therefore possess detailed knowledge on what, how, where and when their users perform sports, which is highly privacy relevant. Runtastic was acquired by Adidas in 2015 for about 220 million EUR, which underpins the high potential of exploitable volumes of QS data (Kaiser et al., 2017).

However, carrying out complex computing tasks with QS data on consumer devices, as well as sending it to a cloud platform can be challenging in terms of bandwidth and latency due to the amount of data produced (autonomous vehicles are expected to produce up to 4TB per day (Krzanich, 2016). Modern vehicles are equipped with many sensors already and even more will be added for enabling assistive driving functionalities (i.e., LIDAR/RADAR – LIght/RAdio Detection And Ranging, video). Several start-ups (including otonomo.io and moj.io), and research initiatives (including automat-project.eu and aegis-bigdata.eu) have identified and targeted this promising field, aiming to create value out of QS data of drivers and vehicles, termed as “Quantified Vehicles” (QV) (Stocker et al., 2017a; Swan, 2015; Kaiser et al., 2018). However, vehicle manufacturers ensure that data, generated for and used for vehicle functionality, will not reveal any internal knowledge or access or intellectual property for safety reasons, e.g. to not provide to outsiders with the information on how data is actually processed and used within a vehicle (AutoMat, 2018b).

### 9.2.2 The Outlook for a Universal IoT Platform

The development of cloud platforms to enable QV applications is a major trend in current technology start-ups and large companies, looking to reap the business opportunities created by data.

The increasing prevalence of several software platforms today calls for the ability to augment solutions and support an emerging portfolio of leading technology solutions and trends. It is unthinkable to design or use any software technology without standing on the shoulders of a multitude of layers of existing platforms. Popular contenders include Google Cloud, Microsoft Azure, Amazon Web Services, IBM Cloud and many more. Industrial vendors ride the Everything-as-a-Service (XaaS) (Duan et al., 2015) wave, through the provision of computing resources and platform services for the realization of “Internet-first” applications, offering different possibilities for management, operation and orchestration of the infrastructure.

Especially relevant to this chapter is a wide range of cloud platforms tuned to the specific needs of IoT applications. A documentation analysis of the commonalities between the architectures of three IoT platforms from Google, Amazon and Microsoft and, to a lower extent, from Ericsson AppIoT was carried out by a subset of the authors of this chapter, with the overall aim to realize a generic platform with as few restrictions as possible. The motivation is wrapped around the emerging opportunities of:

i. lower entry barriers (i.e., allow for new entrants to easily use and contribute the platform),
ii. increased shared responsibility (i.e., provide joint responsibility and to not just one party),
iii. increased framework support, and,
iv. lower technological lock-in.

Our analysis resulted in the generic design of an IoT platform encompassing the common features supported by the platforms analyzed, shown in Figure 42. The analysis has revealed a number of common functionalities and usage of same high-level concepts, even though different APIs and implementations occur. One of the contributions of this chapter is the design of this generic IoT platform and its instantiation in a demonstrator consisting of data-driven vehicle services for drivers and other stakeholders.

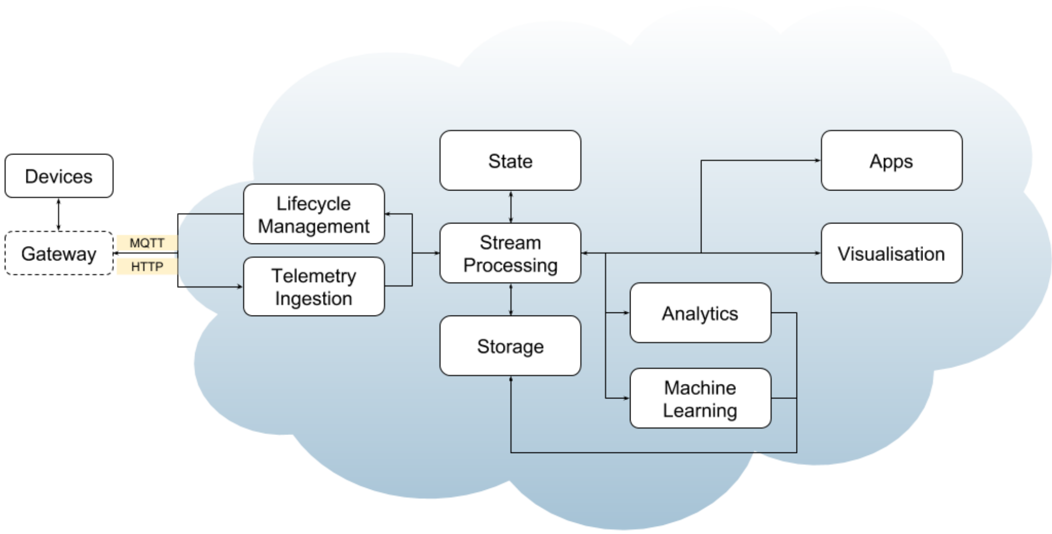


**Figure 42 Generic IoT platform design. MQTT is preferred when communicating back to devices. (Papatheocharous et al., 2018)**

In the generic IoT platform design Figure 42, devices (any smart device with connectivity) go through the gateway (a device or software) to accomplish exchange of data with the platform. Some devices can also bypass the gateway. The exchange is carried out through MQTT

(Banks and Gupta, 2015) or HTTP connections, even though there is a plethora of other similarly purposed standards. The platform offers:

i. telemetry ingestion (accepts data),
ii. stream processing (data flows are processed and converted to unified formats),
iii. storage (data is stored in one or several databases), analytics (data is statistically and semantically analyzed to extract information),
iv. machine learning (data is processed with machine learning algorithms to extract knowledge and intelligence),
v. visualization (data is depicted in meaningful charts and graphs to extract summarized information, generalizations, locate anomalies, etc.),
vi. lifecycle management (consists of supporting functions for the management of devices, such as software updates or (re)configuration),
vii. state (consists of storing the state of devices at all given times), and,
viii. apps (consist of extended applications and services that can extend the platform and some offer additional functionality or end-user value).

### 9.2.3 Provisions for Privacy and Regulations

Another challenge encountered is performing secure data operations in cloud platforms provisioning for security, privacy, safety, trust, etc. which is challenging for the following main reasons: (a) the various system types composing platforms and their purposes are becoming more and more variable in time, (b) infrastructures span over diverse geographical locations, (c) rapid distribution in type as well as in growing size and complexity of components and technologies, and, (d) there is a growing need in elasticity (resilience) and dynamic in the system's structure, behavior as well as interactions. Thus, developing and maintaining cloud platforms with end-to-end security and privacy is challenging but at the same time mandatory and a requirement from legislations and governments (i.e., GDPR, eCall – the European initiative for emergency rapid assistance to motorists involved in a collision anywhere in the EU). To cover this gap, cloud platforms need to provision for security, privacy, safety and control over data in an efficient way. The solution the authors propose in this work mitigates some of these challenges, as explained next.

## 9.3 Conceptual Architecture and Solution Walk-through

In this section the conceptual architecture of the solution followed by a solution walk-through is described.

The architecture of the proposed solution is composed of a generic IoT platform to support data-driven services for the Quantified Vehicles scenario and is illustrated in Figure 43. The architecture is designed in three layers providing: 1) a cloud platform to support novel QV services, 2) operations, and, 3) management provisions.

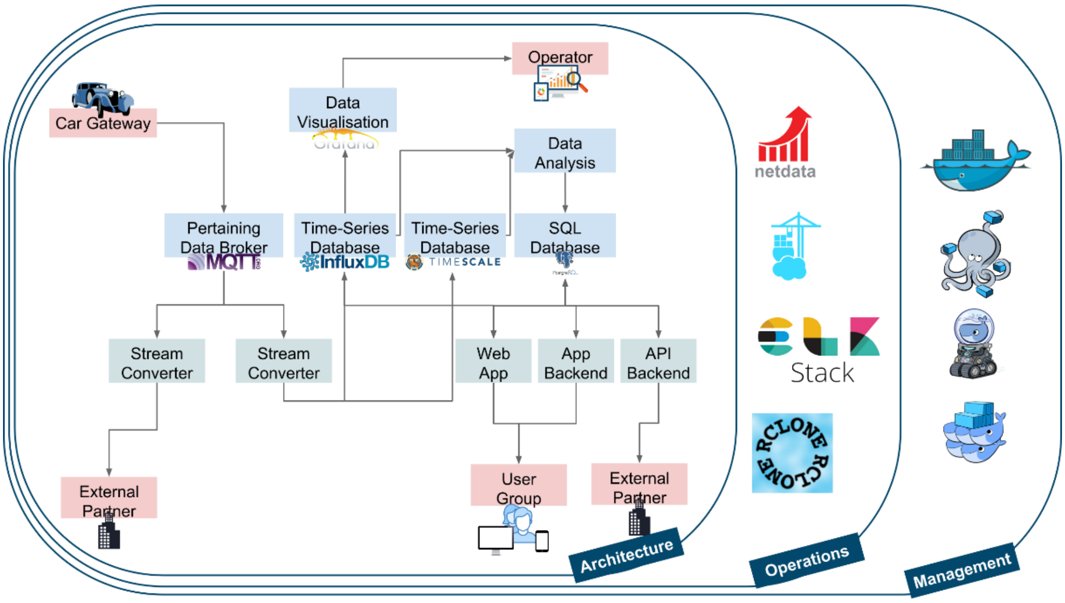


**Figure 43** **Conceptual architecture: Docker (rightmost) manages infrastructure, operated in the cloud. Architecture maps onto generic IoT platform, using e.g. MQTT (telemetry ingestion), databases (storage) and Grafana (visualization). (Papatheocharous et al., 2018)**

A vehicle logger acts as the vehicle gateway device, and is designed as a generic data logger to capture OBD-II (ISO, 2016) (or CAN-bus (ISO, 2015)) data from vehicles as well as GPS, rotation, and acceleration data. The data is then stored directly on the logger waiting for favorable mobile connection conditions. When connection is established via the mobile network, the logger sends the data to the cloud platform.

The platform supports exchange of data with devices (either directly or via a gateway) and can accommodate cloud services deployed to ingest data, store, process and manage data in ways to create end-user value.

The end-users are broadly illustrated in Figure 43 as connected vehicles, user-groups, operators (e.g., data operators, fleet managers), and, external partners (organizations or governments).

Potential applications for the end-users' support, among other:

i. driver scores (calculations of risk, financial/cost, efficiency, or safety based on amplitudes of braking, accelerating, cornering, distraction, etc.),

ii. driver tutoring (suggestions, recommendations, warnings to improve e.g. safety),

iii. driving route recommendations (alternative and recommended driving paths based on other drivers' behaviors, e.g. usual "stand-still" times, diversions),

iv. entertainment (games: earn points for driving safe, or become the major of a street),

v. benchmarking (compare with other drivers and award "safest drivers"), and,

vi. city planning (suggest alternative routes based on heat map overlays and identify potholes or obstacles).

## 9.4 Evaluation through a Demonstrator

As a proof-of-concept the authors have developed a demonstrator comprising of an own-designed Vehicle Data Logger mounted on real vehicles, a cloud platform, collecting data and which in turn enables two data-driven services: one for individual drivers and one for external stakeholders (e.g., companies and governments). These are released through a mobile application and a web application.

### 9.4.1 Vehicle Data Logger

Our own-designed and implemented Vehicle Data Logger solution (shown in Figure 44), is used in multiple vehicles and by different drivers.

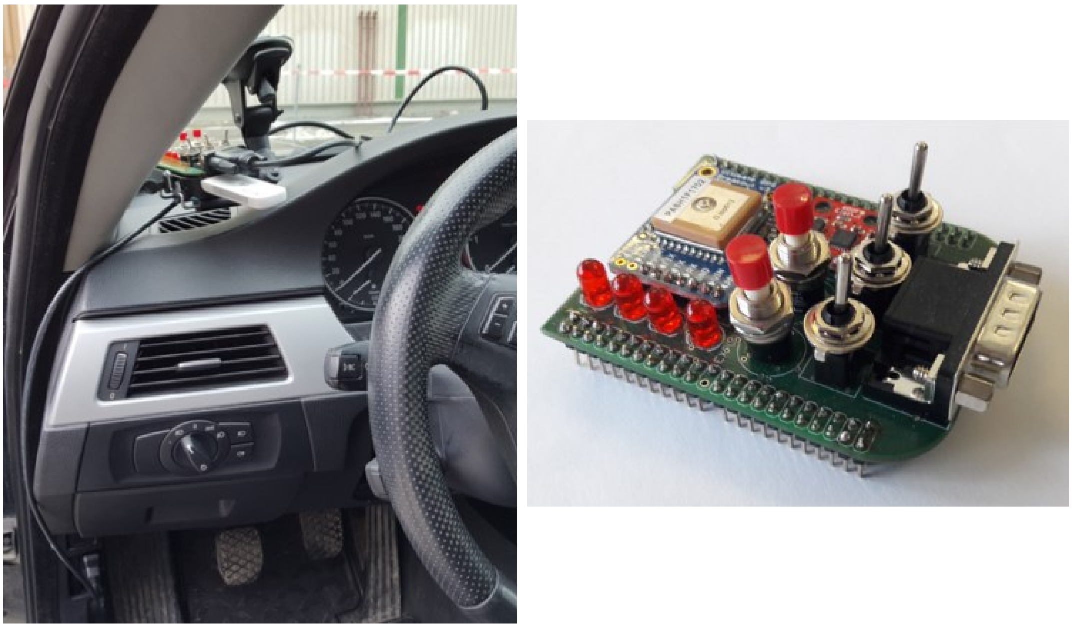

**Figure 44** **Vehicle Data Logger mounted in a vehicle (connected to OBD-II, left) and the cape with GPS / IMU sensory (right). (Papatheocharous et al., 2018)**

The Vehicle Data Logger implementation is based on a BeagleBone Black single-board computer, a low-cost, community-supported development platform. The hardware runs Debian Linux and is extended by a cape, an own-developed Printed Circuit Board (PCB) equipped with a CAN chip, and with GPS and IMU sensory, switches and buttons. The cape is mounted on the BeagleBone's plug connector. It provides a software capable of reading telemetry data from the vehicles' OBD-II interface (or directly from the CAN bus system of the vehicle) when connected to it. Table 19 provides a list of the main components used in this low-cost device.

**Table 19 Bill of material for the main components of the data logger. (Papatheocharous et al., 2018)**

| Type | Actual Component Used |
|---|---|
| Base System Hardware | BeagleBone Black |
| Acceleration Sensor | ADXL345 (via SparkFun 6 DoF IMU) |
| Rotation rate Sensor | ITG-3200 (via SparkFun 6 DoF IMU) |
| GPS positions | Adafruit Ultimate GPS Breakout |
| CAN/OBD-II data Chip | MCP2551 CAN chip |

The sensor data collected includes about five measurements per second for:

i. vehicle data, e.g. available OBD-II data like vehicle speed,

ii. acceleration and gyroscope measurements (in x, y, and z directions) from the IMU sensor, and

iii. the GPS position including the current time-stamp. The data is stored on a local buffer and sent to a MQTT broker via a 4G cellular network modem, thus enabling applications for various stakeholders. An example of visualization of telemetry data for one driver is shown in Figure 45.

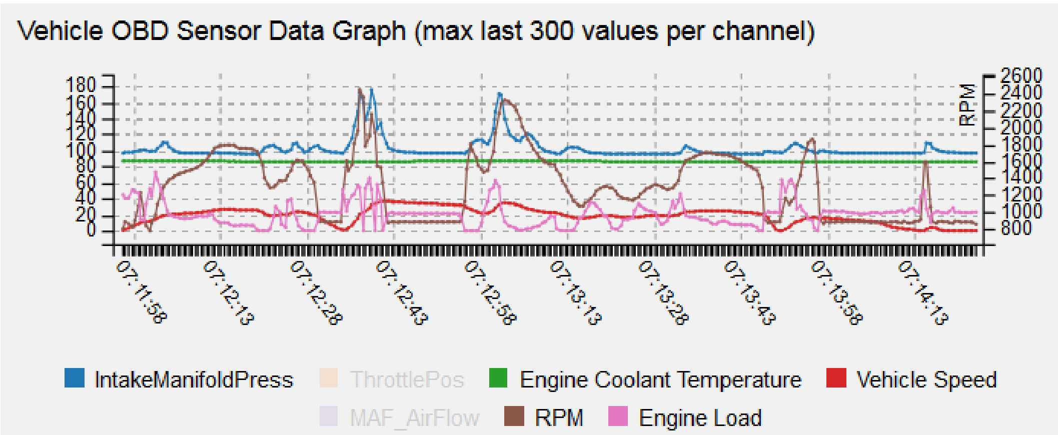


**Figure 45 Telemetry data visualization from OBD-II interface, i.e., intake manifold pressure, engine RPM, vehicle speed, engine load, etc. (Papatheocharous et al., 2018)**

Using a standard single-board computer running Linux is a clear benefit, while software and building instructions for the own-developed cape are planned to be published in the future as well to enable scientists globally to participate in this research. A public dataset produced by the Vehicle Data Logger, demonstrating data quality, is already published at Zenodo (Stocker et al., 2017b).

### 9.4.2 Cloud Services

The cloud services architecture deployed for the demonstrator is already described in Figure 43. Data ingestion from remote vehicles is made primarily through a MQTT broker, and formatted as JSON (Jennings et al., 2018). Use of the broker and the publish-subscribe pattern (Birman and Joseph, 1987) makes it possible for remote and external trusted partners to receive raw data if necessary. Data is transformed and pushed into two separated time-series databases: InfluxDB and Timescale, a module of PostgreSQL, two competing solutions that the authors are still evaluating. Upon trip detection, raw data is analyzed, and features are extracted in order to be inserted into the PostgreSQL database so they can serve as a base for the extended services described in the next subsection. Raw data can be visualized through Grafana, even though this is best achieved through application-specific user interfaces. The cluster provides web APIs for accessing analyzed data, for example from remote smartphone applications. All access to the services provided by the cluster, including the MQTT broker, is forcefully encrypted with Transport Layer Security (TLS) and certificates from Let's Encrypt, a certificate authority that provides free certificates for TLS encryption via an automated process (https://letsencrypt.org/).

Through a substantial use of the various Docker (Merkel, 2014) tools (Engine, Compose, Swarm, Machine, etc. – see the Management layer in Figure 43) and of Machinery, the architecture can easily be deployed and redeployed at any of the available cloud providers. Machinery is an open source software architecture management tool (available at: https://github.com/efrecon/machinery). It also encompasses a number of containers and solutions for daily operations of the applications, including data backups or application supervisions.

### 9.4.3 End-User Applications

For a proof-of-concept, the authors have used the data provided by the Vehicle Data Logger and the Cloud Services to enable two end-user applications, one for the driver and another one for a city planner. While the first application provides interesting driving statistics including safety-relevant events as an overlay on a geographic map in a mobile application, the latter application produces a heat map of safety-relevant hotspots in a city to help decision-makers improving the transport flows in a web dashboard. Figure 46 shows a snapshot of the driver end-user application highlighting the events occurring during the trip and providing summary statistics for the trip (duration, speed, fuel consumption, etc.).

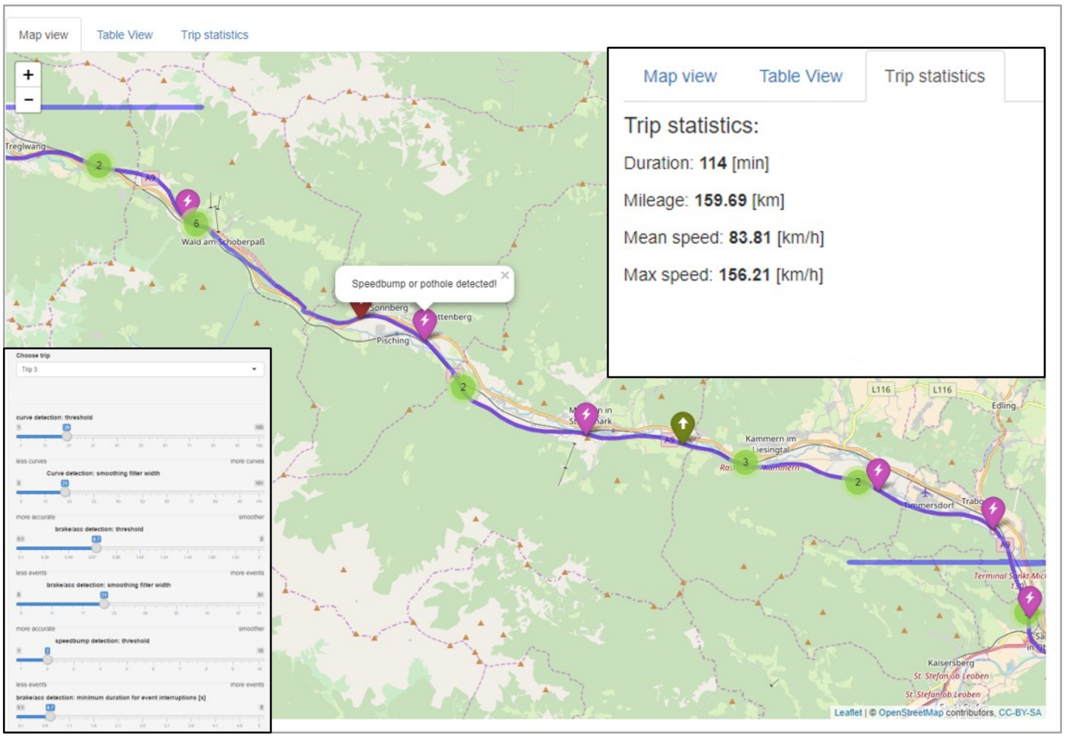


**Figure 46** **End-user application: Analysis of trip data and visualization. (Papatheocharous et al., 2018)**

## 9.5 Benefits and Discussion

The following benefits, divided to business and technical, resulting from the taken approach are worth to discuss.

### 9.5.1 Business Benefits

Quantified Vehicles excavate a hidden business treasure by simply harvesting sensor data already produced for other purposes. The dreamlike scenario having live data streaming from every single vehicle in operation, attracts many data scientists and creative minds, not only from the automotive industry. Hence, several stakeholders (e.g. start-ups, vehicle manufacturers, automotive suppliers, IT companies, service company providers) battle for obtaining the supremacy of this topic, whereas the lucrative role of the platform provider is already identified as a key element in the landscape. In times of increasing awareness and governance on privacy (GDPR), the drivers producing data may benefit from the innovation perspective, e.g. by innovative driver assistance systems, but they will require to have a strong incentive to share their data.

The proposed architecture based on open source technologies empowers IT companies and freelancers to set up individual platforms and to develop their own services in order to create an even more competitive business landscape. The need for innovation, productivity and reduced time-to-market in products may dramatically increase and create a difficult collaborative environment for key players of the automotive industry striving for competitive advantage. Many risks still need to be mitigated, e.g., related to licensing, rights, data ownership, accountability, negotiations, agreements and business models.

## 9.5.2 Technical Benefits

### 9.5.2.1 Security

The proposed architecture takes a security-by-design approach through a widespread use of network encryption techniques and careful management of sensitive data within the platform. For example, data is ingested directly, through secure APIs and using renewable TLS certificates. Within the cloud cluster, network communication occurs within encrypted overlay networks and keys are regularly and automatically rotated. Secrets that need to be shared between architectural components are stored in an encrypted and replicated key-value store, and made available as memory mapped files to relevant processes only. Finally, containerization techniques provide for a high level of encapsulation between the components, making connections and dependencies explicit and traceable.

### 9.5.2.2 Automation

The cloud infrastructure and application are deployed and managed by Machinery, a high-level infrastructure management tool at the top of the Docker tooling pyramid. Machinery takes a declarative approach to cluster management, resulting in describing the entire architecture (machines, services, networks, etc.) as configuration and code that can be managed as a set of project files placed under a version control system – whereas keeps secrets apart.

### 9.5.2.3 Open APIs

For the realization of smartphone applications, or other external services, the platform provides a REST interface to the trip data. PostgREST (https://postgrest.com/) turns any PostgreSQL database directly into a RESTful API. This interface is application-independent and generic, and it facilitates access to the underlying database from modern back-end services or devices. All other components of the architecture have well-documented APIs to support communication between the internal services.

#### 9.5.2.4 Standardized Data Formats

While MQTT provides a standard for loosely decoupled communication between interested parties, it leaves several open design decisions, such as topics organization and data formats. The proposed architecture uses SenML (Jennings et al., 2018) serialized to JSON as a common data streaming format. JSON was selected because of its ubiquitous implementations across platforms, toolkits and languages, but SenML can also be serialized to, e.g. CBOR for improved compactness. Integrating further formats, such as the CoRE link format (Shelby, 2012), could provide an embryo to a standardized management of devices and their capabilities.

#### 9.5.2.5 Scalability

Scalability of the solution is principally achieved through the careful selection of architectural components that have a known migration path to horizontal scalability and high availability. In most cases, this means choosing and configuring components and technologies so that they will be able to scale out when necessary. Containers are orchestrated through Docker Swarm, thus providing a route to increasing replicas under load.

#### 9.5.2.6 Flexibility

The major benefit of the declarative approach is to facilitate migration between cloud vendors through redeployments in other premises, provided that the appropriate credentials are available. By being built on top of Docker Machine, Machinery is not only able to interface with all major cloud vendors for the creation of virtual machines, but also to facilitate onboarding of existing bare-metal servers, if necessary. This increases the robustness of the solution by making it independent of any cloud provider. Continuous integration techniques permit to constantly improve and update application-specific containers, and redeploy as necessary.

## 9.6 Conclusion and Future Work

This chapter has presented a generic IoT architecture for enabling vehicle services and a demonstrator with end-user applications as a proof-of-concept. The results demonstrated, enable useful and usable set of data-driven services for vehicle drivers and other stakeholders. The business and technical benefits are discussed.

Next, as future work, the authors plan to investigate improvements and extensions to the proposed conceptual architecture presented. An example would be the offer to host private containerized Docker repositories to facilitate security and privacy configurations for the realization and integration of end-user applications. Moreover, the authors will investigate how to elastically raise up or destroy machines and containers to adapt to the demand of vehicle fleets. Similarly, the authors plan to extend the implementations to facilitate continuous deployments of applications or other back-end services. The authors would also like to integrate deeper privacy concerns and regulations, so as to enable the automatic erase of user-related historical content from all media and/or increase anonymization and minimize data inter-relations. Also, the authors' vision is to further develop cooperative vehicle-infrastructure applications, eco-driving and assisted driving coaches and propose novel human-centric interfaces and displays for these applications. Such implementations will be further investigated in future work and compared to existing implementations which enable data-driven vehicle services, in order to identify actors and their application domains which participate in such platforms, and their value models.

# 10. Towards a Privacy-preserving Way of Vehicle Data Sharing – a Case for Blockchain Technology? [16]

## Summary and Author Contribution

**Concepts and Prototypical Implementations along the VDVC (Paper 2/5)**

This chapter is about vehicle data sharing. Vehicle data is a valuable source for digital services, and the exchange of vehicle and driving data will massively increase. Since vehicle data from the field allows inferences and analysis which are in personal privacy, this chapter deals with the question on what would be a privacy-preserving way of vehicle data exploitation? To answer this question the blockchain based *Open Vehicle Data Platform* concept is presented, as well as a discussion on unsolved technical and non-technical issues.

The origin of this paper was the discussion at that time about data ownership of vehicle data, as well as an internal presentation of the second author on the topic of blockchain. While the Automat project (Automat, 2017), in which vehicle manufacturers are also involved, published a video[17] explaining that manufacturers should also have rights on vehicle data, in contrast, in this paper we present a blockchain-based concept that allows vehicle users to decide with whom they share their data. As corresponding author, I contributed to this article by developing the concept idea jointly together with the second and the fifth author, by writing parts of the related work, the concept, and the conclusion sections, and by developing the e3value model showing actors and value flows of a vehicle data sharing ecosystem. Overall, with this publication we would like to continue and renew the discussion about data ownership, and show a way to give vehicle users the rights they deserve in our view.

---

16 The content of this chapter is based on

**Kaiser, C.**, Steger, M., Dorri, A., Festl, A., Stocker, A., Fellmann, M., and Kanhere, S. (2019). Towards a Privacy-Preserving Way of Vehicle Data Sharing – A Case for Blockchain Technology? In *Proceedings of Advanced Microsystems for Automotive Applications (AMAA) 2018. Part of the Lecture Notes in Mobility book series (LNMOB)*. Springer, Cham. pp 111.122. DOI: 10.1007/978-3-319-99762-9_10



17 https://www.youtube.com/watch?v=W3kxHd3CdL0&feature=emb_logo especially in the range of 4:21 - 5:03 [last accessed February 2021]

## 10.1 Scope

### 10.1.1 Motivation

Future smart vehicles will provide advanced autonomous driving functions and will be highly connected to other vehicles, roadside infrastructure and to various cloud services. The information gained through these wireless interconnections will be used by any smart vehicle to enrich its own information gathered by built-in sensors such as cameras and radar sensors to further increase the reliability of its autonomous driving functions. However, it will also assist to solve automotive research topics like detection of driver fatigue or driver distraction. These research topics will receive additional focus at the time autonomously driven vehicles will face real world problems on the street and will have to force the driver to takeover. However, the data collected within current vehicles of limited smartness can be used beyond assisting their drivers in driving. Moreover, vehicle data is valuable for third parties (Stocker and Kaiser, 2016; Stocker et al., 2017a; Kaiser et al., 2018b) including e.g. vehicle manufacturers (i.e., OEMs), suppliers, and traffic managers to name three stakeholders, although, there are still many open issues connected to the exchange of vehicle usage data. One dominant challenge for vehicle and driving data exploitation is how to safeguard the privacy of the driver. Despite the privacy regulation has gotten stricter in Europe with the General Data Protection Regulation (GDPR) (European Commission, 2018b), the authors argue that the exchange of vehicle usage data will increase a lot in the future due to two recent developments, tech startups pushing artificial intelligence technologies and the rising interest of the automotive industry to foster the automated driving paradigm.

Shortcomings of current vehicle data provisioning approaches are: Data, information, and services are mostly exchanged within proprietary closed environments, as collected vehicle usage data is usually directly sent from the smart vehicle to a single service provider (e.g., by a device connected to the OBD-II interface of the vehicle or via the drivers' smartphone). As a result, a vehicle owner willing to share data with multiple service providers will have to provide the data multiple times while collecting the data with different devices in parallel. This can be critical due to the large amount of data collected by smart vehicles (up to 4TB of data per day are expected (Krzanich, 2016)), and because a significant portion of current service providers (e.g., Automile and Zubie) is using dedicated OBD-II dongles to gather data from smart vehicles. Thus, it is currently not feasible or at least not practical to use several services at the same time. Finally, these closed systems certainly disrespect the vehicle owner's privacy, as they do not make it transparent how they further monetize the gathered data nor with whom they share it. They typically do not allow the end user to control what data is transferred and shared. And most of them have a lock-in effect, i.e. they use the vehicle data for their own purposes. Finally, their business models do not scale yet as their user community is still composed mostly of early adopters (Stocker et al, 2017).

### 10.1.2 Contributions and Structure

Sharing data always holds the risk of violating one's privacy. So, what is a privacy-preserving way of vehicle data exploitation? Can the Blockchain technology act as an enabler?

Blockchain technology is currently revolutionizing the way smart contracts between parties will be managed due to its outstanding advantages namely *decentralization* and *transparency* per design. The application of Blockchains as a solid basis for a secure data exchange platform seems to be promising to solve the challenge of monetizing vehicle usage data while protecting the data owner's privacy. In contrast to closed systems, a so-enabled *Open Vehicle Data Platform* for vehicle usage data based on smart contracts maintained within Blockchains would allow the user to choose which service providers can access certain vehicle data for which exploitation purpose. Thus, end users can make use of services from various service providers at the same time, while being in full control over the collected data, which will also be crucial for autonomous driving. Full control can be achieved by employing privacy settings for each authorized service provider. The user can decide whether to share only anonymized data (e.g., as required by traffic management systems), vehicle-specific data (e.g., for OEMs for continuous improvement), or even user-specific data (e.g., as required by insurance companies to provide flexible insurance rates in Pay-As-You-Drive (PAYD) models (Husnjak et al., 2014)). Such a platform will be able to support a wide range of service providers and allow different benefit/business models advantageous for both the users and the service providers.

Towards proposing a concept for an *Open Vehicle Data Platform*, in Section 10.1, the authors reviewed existing solutions for vehicle data sharing, highlight strengths and weaknesses, and particularly focused on potential privacy issues. Thereafter, in Section 10.2, the authors provide related work and background for Blockchain technology in the automotive domain and for connected vehicles. Consequently, the authors discuss the actors and roles of a vehicle data sharing ecosystem, the underlying privacy challenge and propose possible privacy setting schemes protecting the privacy of the involved users, followed by a concept for a Blockchain-based *Open Vehicle Data Platform* in Section 10.3. In the latter, Blockchain technology ensures a trustworthy data exchange between all involved entities and users.

After providing a description of a conceptual workflow, the authors discuss open issues and related aspects required to realize the proposed data sharing platform and thereby conclude the chapter with a discussion and outlook in Section 10.4.

## 10.2 Related Work and Background

### 10.2.1 Blockchain Technology (in automotive)

Blockchains were first introduced as underlying technology of Bitcoin in 2008 (Nakamoto, 2008). In this initial form, single transactions are used to describe a cash flow from one entity to another. Every new transaction is distributed to the entire Block-chain system and in a subsequent step a predefined amount of these transactions is compiled into a block, and finally this block is then stored in the Blockchain. The latter can be seen as a distributed database, where blocks are immutably chained to each other. The immutable property on block and on transaction level is ensured by using cryptographic hash functions and digital signatures. Every entity within the Blockchain system can easily verify a transaction as well as a block without requiring any trusted party within the system.

Newer versions of Blockchain allow, besides the exchange of simple transactions, also the creation of smart contracts. The latter can be seen as executable “if-then” condition which is stored on the Blockchain and can e.g. be used to trigger a cash flow by an event (e.g., transfer the flat rent to the landlord on the 1st day of each new month). Besides simple examples, smart contracts also allow describing more complex relations between companies, governmental bodies, etc. and thus is a promising technology to realize a wide range of distributed ser-vices and applications in various industrial domains and especially w.r.t. IoT solutions.

Thus, Blockchains and especially smart contracts can potentially be used to solve certain open issues in the automotive industry due to its capability to preserve privacy; in particular w.r.t. long-term research topics like detection of driver attention/fatigue and current topics like utilizing vehicles as distributed comprehensive environmental sensors, thereby connecting vehicles to each other (V2V) as well as to surrounding infrastructure (V2I).

As a result of this, Blockchain technology raised enormous attention in research, academia and industry. Various projects and initiatives covering different industrial domains were started in the last months with the goal of identifying real business opportunities for the use of Blockchain in future products, or even to develop concrete (distributed) applications where the use of Blockchain technology can be beneficial, including the automotive industry which has identified potential areas for the use for Blockchains. Recently, automotive car manufacturers BMW, GM, Ford and Renault started the Mobility Open Blockchain Initiative (MOBI) together with other industrial and academic partners such as Bosch, Blockchain at Berkeley, Hyperledger, Fetch.ai, IBM and IOTA (Russel, 2018). Also, other vehicle manufacturers are evaluating Blockchains or are already working on concrete projects: In 2017, Daimler started a project where Blockchain technology is used to manage financial transactions (Dotson, 2017). Furthermore, the automotive supplier ZF teamed up with IBM and UBS to work on a Blockchain-based automotive Platform called Car eWallet with the goal of paving the way for autonomous vehicles by allowing automatic payments and by providing other convenience features (Kilbride, 2017).

Hence, Blockchain definitely gained attention in the automotive industry. How-ever, concrete ideas, products and services are needed to show that Blockchain is more than a hyped technology but rather allows the development of new business cases.

### 10.2.2 Connected Vehicles and Data Exploitation

Future vehicles will communicate with each other as well as with surrounding road infrastructure to collect valuable information about road conditions and to sense the current traffic situations (e.g., very relevant in traffic intersection scenarios). Furthermore, vehicles will increasingly be connected to the Internet to provide a wide range of convenience services to the users, to gather latest traffic and map information, the current city traffic strategy or even to report an accident (i.e., eCall).

This Internet connection could of course also be used to transfer environmental data collected by the vehicle itself (e.g., camera, Radar, or Lidar data) to the cloud. Intel recently released a statement saying that future (self-driving) vehicles will collect up to 4 TB of data each day (Krzanich, 2016). A wide range of different ser-vice providers (not restricted to automotive) would be interested in using the collected data in various ways. Sharing the collected data could/should also be beneficial for the owner/driver of the vehicle (see Section 10.3.1) and, on the down side, will raise serious privacy issues, as the exchanged information could be used to e.g. track down the user's location or analyze the user's behavior (see also Section 10.3.2).

Several tech startups such as Automile, Dash, and Zendrive, as well as large initiatives driven by vehicle manufacturers such as AutoMat (coordinated by Volkswagen), started initiatives with the goal to collect and utilize data from single vehicles up to entire fleets following different purposes (Stocker et al, 2017):

i. Provide specific services in order to generate a benefit for the driver or the vehicle/fleet owner in return for sharing data.
ii. Create value by monetizing the collected data coming from a mass of vehicles to third parties, which in turn use it as input for algorithms.
iii. Further improve the business offerings of service providers and develop new services.

Furthermore, in times of a shift of the automotive industry towards digitalization, in times to manage different SAE levels of autonomous driving on the road simultaneously, and in times of the Internet of Things where sensors are increasingly connected to the Internet, the automotive industry still tries to solve many long-known phenomena. These phenomena include for example the detection of the driver's distraction, fatigue and trust or the vehicles security and safety, which will increasingly be done in the cloud, by feeding the algorithms with sensitive and privacy relevant data from vehicle usage.

Data ownership of vehicle sensor data seems to be yet unclear from a legal perspective. Driver, vehicle owner, passengers, and the vehicle manufacturer may claim their right on certain data. In the AutoMat project, coordinated by Volkswagen, it is argued that as usual in other domains, e.g. in the music show business, "the copyright is distributed proportionally among the members of the value chain" (AutoMat, 2018a). This copyright distribution would give vehicle manufacturers the right to use the data a driver produces without charge, and thus would bring vehicle manufacturers into the profitable data platform provider role (as they can integrate a data interface in their cars easily). However, from a driver's/vehicle owner's/passenger's perspective, copyright should not be distributed as there would not be any data without them driving the vehicle. This is usual in many domains e.g. digital camera manufacturers do not have a copyright on produced photos, and a competitive market with open data platforms will force innovative solutions and offer more benefits to the data owner to attract data provision.

## 10.3 Towards Privacy-Preserving Vehicle Data Sharing

### 10.3.1 A Vehicle Data Sharing Ecosystem

A series of stakeholders including vehicle developers, vehicle manufacturers, insurers, and even smart cities could benefit a lot from an open privacy-preserving vehicle data sharing platform, and thus participate in a vehicle data sharing eco-system. The following Figure 47, sketches such a vehicle data sharing ecosystem and highlights the connections between the different stakeholders. The figure illustrates stakeholders and advantages for their businesses (based on shared vehicle data), as well as advantages for vehicle owners (using the service the stakeholder provides based on their shared data). Thereby different connection types and privacy levels are envisaged, as different stakeholders are interested in different aspects of the data collected by connected vehicles.

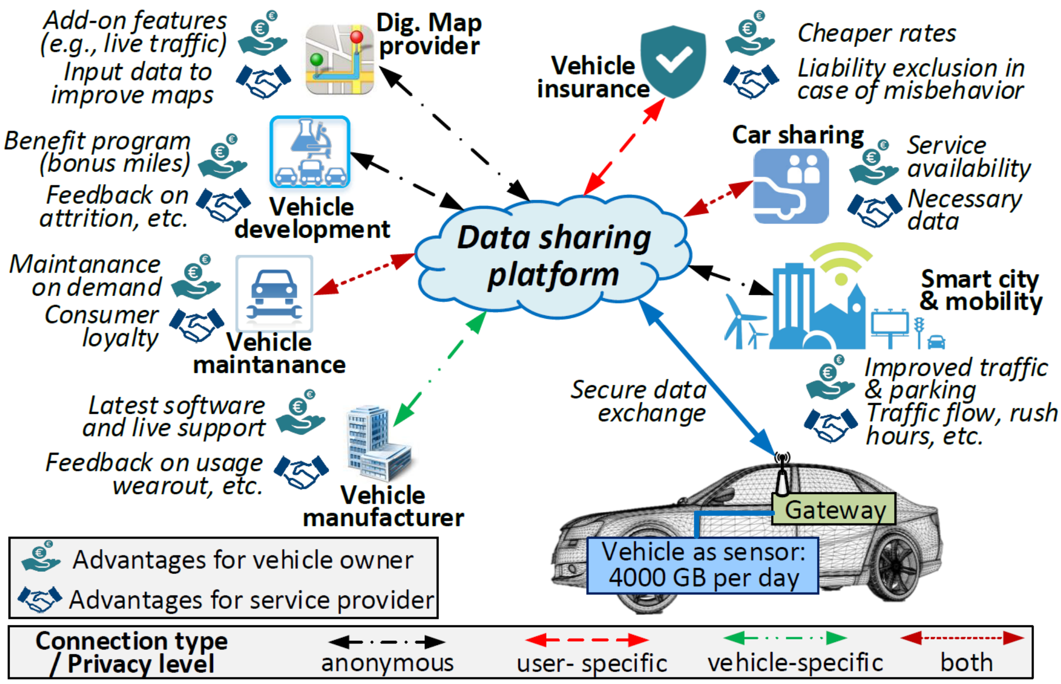


**Figure 47** Vehicle usage data can be used for various services and by different entities and bring ad-vantages to vehicle owner/user and service provider/consumer. (Kaiser et al., 2019b)

As indicated in Figure 47, certain service providers such as city planners or map providers are not interested in who is driving (i.e., do not need driver-specific in-formation) or what specific type of vehicle (i.e., do not need vehicle-specific in-formation such as brand, color, or model). Thus, these services can be satisfied by providing anonymized vehicle usage data. Other (automotive) services targeting on the vehicle development lifecycle (e.g., predictive maintenance or wearout of vehicle components), will only require vehicle-specific data, whereas other services will be mainly interested in user-specific information (i.e., who is/was driving).

The proposed Open Vehicle Data Platform will address the fact that different services require a different kind of data and allow specifying which components of the collected data is shared to enable services. Thereby, privacy is especially addressed as a connected vehicle will not necessarily have to share an entire dataset with a service provider but rather only the data which is really needed by the service provider to provide a specific service. In the simplified model of a vehicle data ecosystem four types of data sharing might be distinguished: sharing anonymous data, driver-specific data, vehicle specific data, or a combination of them.

From a more abstract point of view, a vehicle data sharing ecosystem can have several types of actors linked by value flows, as indicated in the e3value model in Figure 48. For instance, a driver can share driving and vehicle data with a gateway provider who then forwards this data to a data platform provider. In return the driver may receive money but will probably have to mount a vehicle data gate-way device in his vehicle. A service provider may use driving data from the data market/platform to establish a preventive maintenance service for drivers. While drivers may pay service providers a fee for consuming this service, the data market receives another fee from the service provider for providing the technical da-ta, which is the baseline for this service.

Consequently, the ecosystem has mutual dependencies and thus allows scenarios where e.g. a driver uses an attractive service which is offered for free, because an organizational consumer (in current scenarios from the market usually without the knowledge of the driver) pays the service provider for the development and service provision in the background, in order to get the data or access to a valuable service based on this data.

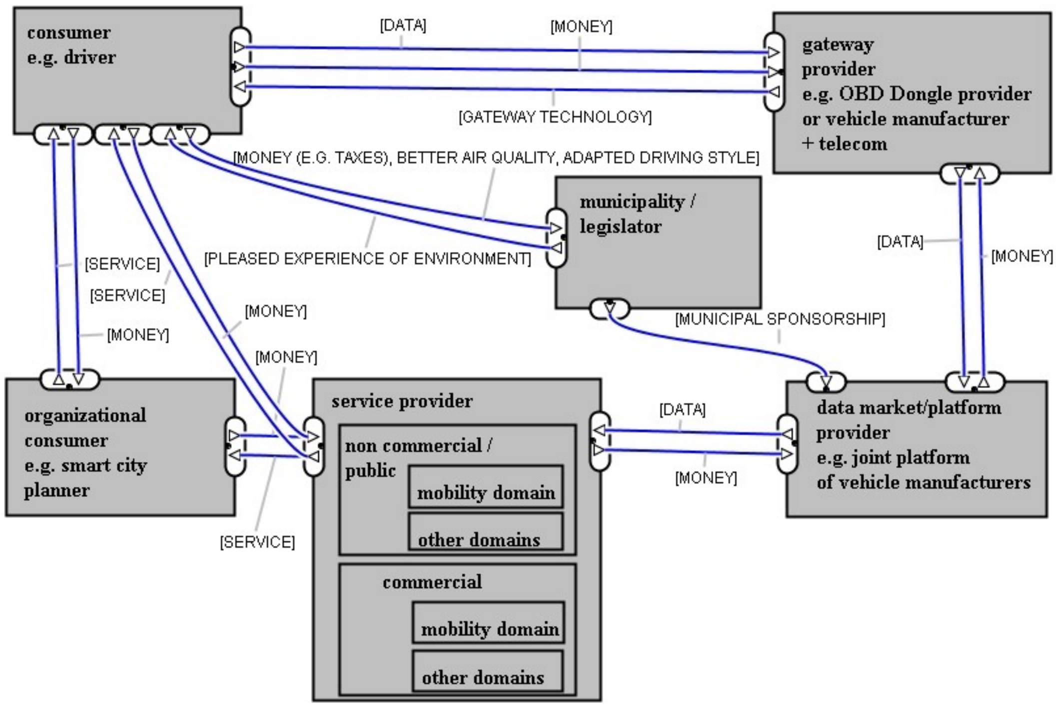


**Figure 48 Actors and value flows (e3value model) of a vehicle data sharing ecosystem. (Kaiser et al., 2019b)**

## 10.3.2 The Privacy Challenge for Data Sharing

As discussed before, service providers will monetize data collected by connected vehicles and thus should reward drivers providing the data with certain benefits. In case that the exchange of data between the connected vehicle and the service provider is insecure (Valasek and Miller, 2015) (or the service provider itself is compromised / acts malicious), privacy issues ranging from tracking down the user to stealing sensitive information can arise. Hence, security and privacy must be addressed when de-signing a vehicle usage data platform, and, as general rule, a service provider should only be allowed to access relevant (i.e., for providing a specific service) data collected by a connected vehicle.

The driver may conduct a driving behavior which could be interpreted in a negative way and might not be willing to share the so generated driving data with others as this would either imply legal, social or ethical consequences. For instance, an aggressive driving behavior might cause social (if shared with friends while benchmarking) or even legal consequences (if captured by the police). Drivers becoming aware of this fact may not want to contribute to any data sharing platform at all if their shared vehicle data could allow to cause negative consequences for them. This fact is also reflected in current studies and surveys, where users are asked about trust and privacy w.r.t. connected vehicles. In one of these studies, Walter et al. (2018) details the user concerns regarding connected vehicles and highlights the needs for a privacy-aware data sharing mechanism.

Defining a privacy configuration mechanism w.r.t usability and transparency brings up different opportunities:

One approach is a distinction between vehicle specific and driver specific data, where one can opt to share both of them either anonymized or not, just one or none.

Another approach would be to have four easy understandable levels with decreasing privacy:

i. don't share, where simply no data is shared at all,
ii. private, where data is provided e.g. to calculate some basic individual statistics, but can-not be used for anything else,
iii. anonymized for public usage, where data can be used like in private level and additionally is provided to public in an anonymized way, and
iv. public, where all data is provided to public.

However, this approach would raise awareness of drivers and service providers would have to adopt the concept, hence it limits possibilities and perhaps opens legal loopholes and at the end of the day it lacks transparency which specific data a service has access to.

Therefore, the authors argue that it is feasible to adopt the approach of Android smartphone applications, which clusters the access to certain data into topics (i.e. An app needs access to one's contacts and images). The level of detail is a decisive factor for such clusters: emission values can be clustered under a huge topic named vehicle sensor data or be seen as an individual emission values category, while using quite granular categories would require basic technical understanding of every user. The authors still see improvement potential as this solution has somehow a touch of too much information, comparable to terms and conditions no one really reads carefully.

### 10.3.3 A Concept for a Blockchain-based Open Vehicle Data Platform

The concept provided in this section sketches a privacy preserving *Open Vehicle Data Platform*. Instead of going into detail and arguing for certain tools and architectures, we'd rather spread the idea by describing the workflow.

A vehicle is capable of acquiring a lot of valuable data and the driver of the connected vehicle shall be able to decide if and how this data is shared with service providers, as discussed earlier. In the proposed concept and as indicated in Figure 49, smart contracts based on Blockchain technology are used to specify whether a service is allowed to access data from a certain vehicle and also which kind of data will be shared.

Once an agreement between the connected vehicle and the service provider (i.e., smart contract) is signed, Blockchain technology is exploited to i) make sure that the smart contract cannot be tampered with, as well as ii) to make the smart con-tract available to so called Brokers. The latter provides an online storage, where data collected by connected vehicles is stored securely, and it is also responsible to handle the access of a specific service on data stored on its online storage according to existing smart contracts. Furthermore, the Broker will maintain secure data connections between its online storage and connected vehicles as well as service providers by using suitable protection mechanisms (e.g., TLS).

In the proposed concept, several Brokers will take over the aforementioned tasks, and thereby also allow connected vehicles to switch between different Brokers or even to store data on different locations. The Blockchain will thereby fulfill two essential tasks. Firstly, the Blockchain provides tamperproof storage for smart contracts as well as other transactions, and secondly also provides a way to ensure the authenticity of data collected by a connected vehicle and stored on an online storage, as the hash of a collected dataset is integrated in a transaction and then stored on the Blockchain. Such a transaction can also be seen as a trigger for service providers informing them about the latest available dataset.

Please note that storing data directly on the Blockchain is not advisable from technological point of view. Also note that existing contracts on the Blockchain can simply be revoked or changed by filing a new contract between the connect-ed vehicle and the concerned service provider.

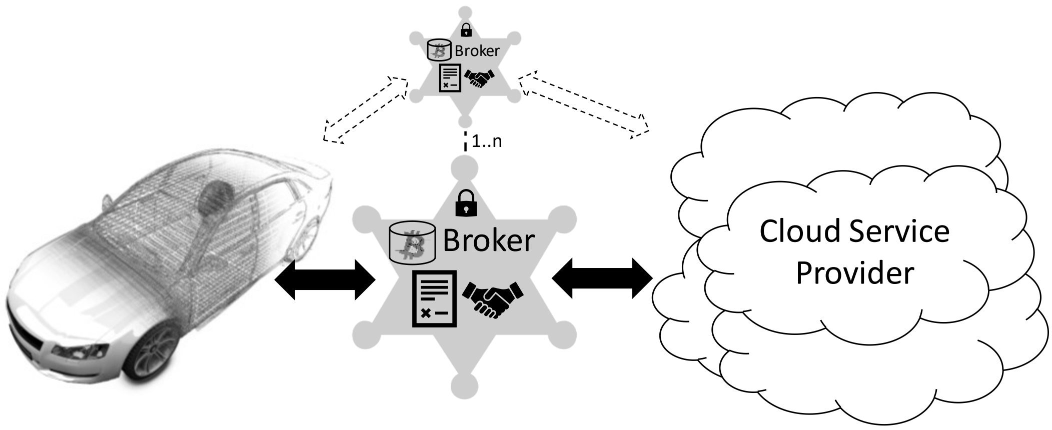


**Figure 49** **Data exchange between origin (vehicle) and target (service providers) is managed by a broker using Blockchain technology for smart contracts. (Kaiser et al., 2019b)**

The proposed concept will rely on two different entities which are stored on the Blockchain, namely

i. Smart contracts, describing which data is shared with a certain ser-vice provider and also specifies the corresponding reward. It will contain information about the Broker that is used to store the collected data, and the timespan in which a certain service is allowed to access the collected data. Each smart contract will be signed by the connected vehicle (is owner) and the service provider before it is stored on the Blockchain;
ii. Dataset transactions, containing the hash of a dataset stored on the online storage of a Broker. Every transaction is signed by the connected vehicle (or its owner), and also by the Broker once the dataset was successfully transferred (and verified) to its online storage.

The proposed concept is able to securely interconnect connected vehicles and services providers in a privacy-preserving way, by utilizing Blockchain as tamperproof, decentralized database, as well as by using dedicated Brokers providing a secure online storage and handling access control w.r.t. the stored data. In the following, the authors summarize seven steps required to share data between a connected vehicle and a service provider and use this example to highlight the benefits of the propose vehicle data sharing platform:

1. Initially, the owner of a connected vehicle wants to use a certain service and, in further consequence, will get into contact with the responsible service provider. In this initial step, the user will be informed about the type of the data the service provider requires to provide a specific service.

2. If the user agrees to this terms, a smart contract specifying the relation between the connected vehicle, its owner, and the service provider is created and signed by the vehicle owner (representing the connected vehicle) and the service provider.
3. Once the smart contract is finalized, it will be stored on the Blockchain.
4. While being used, the connected vehicle will continuously collect valuable data, which is divided into datasets (e.g., after a predefined time or once a certain amount of data is collected) and sent encrypted to the online storage of the Broker. Each transferred dataset is accompanied by a dataset transaction containing the hash of the dataset as well as the digital signature of the connected vehicle (its owner).
5. Hence, the Broker on the one hand can verify that the dataset was not altered while being transferred, and on the other is held from changing the dataset itself as this would invalidate the digital signature already included in the dataset transaction. Once the currently received dataset is verified, the Broker will add its signature (thus completes the transaction) to the transaction and broadcast it on the Blockchain network.
6. Service providers can monitor the Blockchain and will be directly notified about the latest available dataset by looking for relevant dataset transaction. In case such a transaction was found, the service provider requests the dataset by establishing a connection with the Broker.
7. Next, the latter looks for a suitable smart contract on the Blockchain and provides access to data as specified in the smart contract or declines the request in case no smart contract was found or it was revoked.

## 10.4 Conclusion, Discussion and Outlook

This chapter was aimed to launch the discussion on how the Blockchain technology may help to establish an open vehicle data sharing platform, respecting the privacy of both the vehicle owner and the vehicle driver. Thereby smart contracts are introduced as a mode to fully digitize the data sharing relationship between a consumer (e.g. a driver, who provides his data with the purpose to use services) and a service provider (e.g. a provider of a preventive maintenance service). They describe what kind of data will be provided by whom and for what data exploitation purpose. While these smart contracts are stored on the Blockchain to increase the trust between the vehicle data sharing ecosystem stakeholders, the shared data itself will not be stored on the Blockchain, but for instance on a separate data platform and a data market.

However, a series of issues and research topics remain open and will be targeted in future work:

There are certain pre-requisites vehicles would need for the provided concept. For example, a standardized vehicle data interface across manufacturers, where in general all vehicle data can be provided to extern (to be stored on SD card or on a hard drive if used for private purposes, or to be sent to online destinations), would ease data acquisition. Only data which is marked to be stored/sent to somewhere should be captured, all other data should be deleted or continuously overwritten.

In order to participate, users need to be able to authorize themselves (e.g. to use their privacy settings in every vehicle they use) to the vehicle and the Broker, so they need to register and have an identity.

Using Blockchain technology ensures a privacy preserving way to securely share the data from the vehicle to the service provider. If a service provider gets access to one's data, then this indicates that he is not allowed to resell it unless this is explicitly mentioned in the contract. However, in praxis this can not be prevented with the presented concept, thus privacy can not fully be ensured.

As mentioned in Section 10.3.2, how to cluster data in useful groups and in which granularity is a topic for future research. An initial version could be as follows:

- Emission data
- Vehicle data (e.g. base weight, number of passengers, year of manufacture, type, brand)
- Environment data (e.g. road topography, temperature outside, rain)
- Traffic data (e.g. detected entities around the vehicle including humans and vehicles, information about the streets throughput rate)
- Driver data (e.g. Driver ID, music channel, mood, fatigue level, driving score, heart rate)
- Ride data (e.g. GPS position, temperature inside, start datetime, target)
- Other data

# 11. A Lightweight Framework for Multi-device Integration and Multi-sensor Fusion to Explore Driver Distraction [18]


**Summary and Author Contribution**

**Concepts and Prototypical Implementations along the VDVC (Paper 3/5)**

This chapter examines how a lightweight technical framework for the real-time fusion of vehicle data and other contextually relevant data could look like. Such a framework could be used to assess driver's status with appropriate measurement equipment, e.g. to detect driver distraction and driver inattention. This is particularly of interest, as driver distraction and driver inattention are major challenges in road traffic and major causes of accidents. Especially novel quantification approaches combining data from different sensors and devices are necessary for comprehensively determining causes of driver distraction. Thus, a generic architecture is presented and its application in a proof-of-concept implementation is shown. This application is also demonstrated and examined by means of an empirical study, where drivers (study participants) perform distracting tasks, like using multimedia interfaces or glancing at co-drivers. Preliminary results of our analysis have indicated a high accuracy of distraction detection for individual distraction tasks and thus the framework's usefulness.

With this paper, we wanted to disseminate our findings from an internal project where we recorded fused sensor data from a smartphone, a smartwatch, and a smart glass in addition to vehicle data to detect distraction. In this work, I was responsible for setting up and conducting the empirical field study, where we tested the solution and collected data for the evaluation.


Driver distraction is a major challenge in road traffic, resulting in an enormous number of accidents and fatalities every year (NHTSA, 2018a; EU, 2018; NHTSA, 2018b). In a study based on the analysis of 997 crashes, Thomas et al. state that 11% of crashed drivers were distracted and 8% were inattentive (Thomas et al., 2013). However, distraction does not only concern manual passenger vehicle driving but is also a key issue towards increasingly automated driving functions, Advanced Driver Assistance Systems (ADAS) and autonomous driving. When driving scenarios are not covered by an automation, the advanced driving

---

18 The content of this chapter is based on

Lechner, G., Fellmann, M., Festl, A., **Kaiser, C.**, Kalayci, T. E., Spitzer, M., Stocker, A., (2019). A Lightweight Framework for Multi-device Integration and Multi-sensor Fusion to Explore Driver Distraction. In *International Conference on Advanced Information Systems Engineering (CAiSE)* (pp. 80-95). Springer, Cham.

functions fail, and a manual take-over is required (Payre et al., 2017; Zeeb et al., 2015). For this purpose, termed “to keep the *Human-in-the-loop*”, drivers must be responsive and prepared all the time: drivers must end activities unrelated to driving and give enough attention to take manual control of the vehicle in an adequate manner.

Apart from the maturity level of the vehicle with respect to autonomous driving, one key point is correct measurement of driver distraction to assess driver’s status. Further development of (partly) automated driving is highly dependent on qualitative and quantitative measurement of driver distraction to trigger appropriate measures to get drivers back into the loop. In addition, safety of drivers at the present time will be increased significantly through these measures by improving mechanisms to avoid distraction.

A major issue of measuring driver distraction is that in most existing approaches the sensors and devices are obtrusive, and already their application would distract the driver. For example, the methods of neuroimaging or eye tracking depend on bulky, distracting, and cost-intensive technical equipment and are therefore not applicable for daily use in real traffic but suitable only for experimental setups (ABC Science, 2018; American Psychological Association, 2018). Prototypical audiovisual systems for distraction measurement are in use but rarely in combination with other sensors (Eskenazi et al., 2018). However, available consumer-grade hardware such as smartwatches, wristbands, smartphones, or other types of wearables are already used by many drivers. Data captured by their sensors bears a huge potential to be used in driver distraction detection – a potential currently not fully exploited in research. A recent practice-oriented driver distraction study published by Zendrive.com – a startup dedicated to better inferring distraction by using smartphone sensors investigating 100 billion miles – found out that on an average day, over 60 percent of people use their phones at least once while being behind the wheel (Zendrive, 2018). This determines the potential of using popular consumer electronics owned by drivers for distraction research. But how can data from multiple sensors in different devices be captured and made available for analysis in a convenient way? A literature review has indicated a lack of lightweight frameworks that integrate multiple sensors from multiple consumer devices in order to fuse data and enable a comprehensive analysis and detection of multiple distraction tasks.

Hence, the objective of the research presented in this chapter is to answer the following research question: *What is a lightweight software framework for multi-device integration and multi-sensor fusion for comprehensive driver distraction detection?* Furthermore, this framework should serve as a baseline for potential interventions in the case of driver distraction and allow the measurement of interventions' success. Apart from being lightweight and cost-efficient, an important aspect of the proposed framework is easy extendibility, i.e. adding newly developed or improved sensors must be practical. This ensures the framework's adaption to various (distraction-related) data sources.

Concerning driver distraction and driver inattention, the authors follow the definitions provided by Regan et al. (2011) who aim for better distinguishing driver distraction from other forms of inattention. They conclude that *Driver Inattention* means *insufficient or no attention to activities critical for safe driving and driver distraction (they refer to Driver Diverted Attention as a synonym)* as just one form of driver inattention.

## 11.1 Background

In the following sections, an overview of topics related to the presented research is provided. The authors discuss *driver inattention/distraction* followed by ways how to *quantify drivers and driving*, and finally the authors conclude with an *overview of frameworks*.

Driving a vehicle is a complex task, and *driver inattention and distraction* increase the risk of having a crash. Transport researchers have paid a lot of attention to explore both topics in review articles, simulation studies and field trials. Most of the highly cited articles are review papers: Sussman et al. (1985) review research into driver attentional processes about safety implications of inattention, related psychological and physiological indices, and in-vehicle instrumentation for detection. Young and Regan (2007) provide a literature review for driver distraction, which is a priority issue in road safety in many countries worldwide. They explicitly highlight the effect of in-vehicle devices (in particular mobile phones) on driving performance. Dong et al. (2011) review technologies for driver inattention monitoring (distraction and fatigue) and discuss the application of hybrid measures to give more reliable solutions compared to single driver physical measures or single driving performance measures. Simulator studies and field trials are important for validating developed technical approaches. Horberry et al. (2006) present the findings of a simulator study which examined the effects of driver distracted while having to perform in-vehicle distraction tasks. Klauer et al. (2006) conducted a widely cited comprehensive study on the impact of driver inattention on near crash and crash risk, using data from the 100 Car naturalistic driving study. Their findings indicate higher risks when driving drowsy, or when engaging in complex tasks while driving. D'Orazio et al. (2007) propose a visual approach based on image recognition for monitoring driver vigilance, detecting if a driver's eyes are open or closed while evaluating the temporal occurrence of eyes open to estimate the driver's attention level validated in experiments. Fletcher and Zelinsky (2009) present a prototype system which can estimate a driver's observations and detect driver inattentiveness based on eye gaze – road event correlation validated in laboratory experiments and road trials.

Driving data resulting from *quantification of drivers and driving* is a valuable source for knowledge generation and can boost research in driver inattention and distraction. The quantification of human behavior termed *quantified self* (Swan, 2009) has sped-up the generation of data and is a very popular example for everyday life data analytics. Individuals engaged in self-tracking any kind of biological, physical, behavioral, or environmental information (Swan, 2013) have led to a multitude of data that can be used for different purposes including detecting health issues or developing personalized training plans. This pattern of self-tracking by using consumer devices can be easily transferred to vehicles, and in this sense, may be termed as *quantified vehicles* (Stocker et al., 2017a). Sensors and electronic control units (ECUs) within modern vehicles create a plethora of vehicle operation data, but this data is hardly accessible for the reasons of safety and espionage. Although examples for real-time streaming of Controller Area Network (CAN) bus data are available (e.g. CSS electronics, 2018), conversion of CAN data is restricted due to proprietary/manufacturer-specific standards (like SAE 2018). Yet, drivers and driving can be quantified either by using data accessed through the vehicles' on-board diagnostics (OBD) interface, or by utilizing the driver's wearables. While much of the data gained from the OBD interface is not suitable to detect driver distraction (e.g. ambient temperature), there are a few valuable measurements including vehicle speed or vehicle RPM (revolutions per minute) that can be explored. Additionally, smartphones provide a set of useful data for quantifying drivers and driving, for example through GPS, acceleration and gyroscope sensors. Analogously, many wearables like smartwatches have similar sensors and generated data may be used to quantify certain tasks of drivers.

Data collected by a driver's wearables using *Lightweight Frameworks for Multi-Device Integration and Multi-Sensor Fusion* can be used to enable driver distraction and inattention detection mechanisms which are independent from a car's built-in functionality. For instance, in the zendrive.com case which was introduced previously, two sensors from the driver's smartphone have been used to quantify safety-critical smartphone usage during driving. But, for implementing such use cases, in-depth-knowledge about technical infrastructures for (multi-)sensor data acquisition, pre-processing, synchronization, and fusion is crucial. Much past research in multi-sensor integration has focused on hardware-based approaches for data acquisition (e.g., for physiological monitoring (Pandian et al., 2008) or context awareness (Gellersen et al., 2002), to name two examples), which seems logical due to the former lack of devices with considerable computation power for dealing with software-based solutions. In search of such software solutions, a review of the scientific literature has shown that approaches for multi-device integration and multi-sensor fusion with a focus on wearables are not sufficiently reported. While some researchers (for example Ramos et al., 2016; Shoaib et al., 2015; Vilarinho et al., 2015; Casilari et al., 2016) have already combined smartphone with smartwatch sensor data for recognizing human activities outside the driving domain, their papers lack the detailed description of applied technical frameworks for capturing and synchronizing measurement data. It is obvious that these papers are focused on the analytics part, which is obviously more interesting for data scientists, but lack on the technical part which should also be considered relevant. With respect to detecting unsafe driving, researchers (Eskenazi et al., 2018; Giang et al., 2015; Liu et al., 2015) report on studies using sensor data from wearables. However, a description of their technical framework enabling the reported analysis is missing, too. Finally, de Arriba-Pérez et al. (2016) present a research article elaborating on a technical framework, though embedded in an educational context.

## 11.2 A Technical Framework for Real-time Fusion and Logging of Sensor Data

### 11.2.1 Requirements

In order to provide the basis for driver distraction detection, the framework must be able to record and integrate various data in a timely manner. Hence, requirements exist regarding which data should be recorded and how it should be stored and integrated. In the following, the authors describe fundamental requirements regarding these aspects.

- **Requirement 1**: The framework should be able to record data about the driver, vehicle state and the current GPS-position.

Regarding the driver, physiological data is important such as movements of the driver's head (e.g. to detect if the driver is looking on the street or not), movements of arms/wrist (e.g. to detect if hands are placed on the steering-wheel) as well as the driver's heartrate (e.g. to detect situations of excessive stress). Regarding the vehicle state, data such as RPM or usage of the entertainment system is relevant (e.g. high RPM variation in combination with frequent gear reductions may indicate a more focused driving style, while excessive entertainment system usage may indicate the opposite). In order to detect the movement of the vehicle in the real world to obtain contextual information, GPS positioning data should also be recorded.

- **Requirement 2**: The framework should provide access to integrated data via a common query interface as well as the capability to store real-time time series data.

Access to integrated data via a common interface is an important requirement to facilitate the analysis of diverse sensor data. The database should moreover be capable of processing large volumes of insert operations in almost real time (e.g. to store acceleration, gyroscope and GPS data). Regarding analysis, the database should provide support for time series data, i.e. querying large amounts of sensor values with a timestamp.

- **Requirement 3**: The framework should provide a mechanism for synchronizing data originating from various sensors from multiple devices.

In order to synchronize sensor data such as GPS position, heartrate, head movement and RPM at a certain point in time, the timestamp of each sensor value could be used. As internal clocks of different sensors are not synchronized all sensors must agree on a common time. For this purpose, a GPS time stamp can be used since GPS sensors are available on many hardware devices such as smart glasses, smartwatches and inside vehicle data loggers used for experiments.

### 11.2.2 Generic Architecture

The generic architecture consists of an arbitrary number of *generic components (GC)* with various functions, thereby considering the requirements stated in the previous subsection. The GCs can acquire data from one or more sensors, but also – uni- or even bi-directionally – from other generic components through a receiver (GC reader). Such *data collection* is the primary and simplest function of a generic component. The *data transfer*-functionality makes the interchange of data between generic components possible and is built upon data collection, as collection of data is required before its transfer. For this purpose, a transmitter acting as a GC writer is integrated. *Data integration and processing* is an optional function separated from transfer and performed in a *data forwarder*. The data forwarder of the GC may thus act as an integrator, thereby processing data from one or more different data sources (individual sensors or generic components). In summary, as shown in Figure 50, the GC is capable to perform following functions: (1) data collection, (2) data transfer, and (3) data integration and processing. In addition, the GC has a configuration layer providing basic configuration options which is accessible by users for specific configuration requirements.

These characteristics of the generic architecture allow to consider the framework as a general one which can be extended in various ways: as sensors quantify information from environment and bring them into a structured, machine-readable format, they can easily be connected to GCs. In particular, the structure of data in terms of measurement frequency and data types provided by different sensors is similar for many different sensors (as, e.g., accelerometers or gyroscopes), independent from specific brands or devices. However, the concrete implementation and respective performance of hardware determines both the number of performed tasks and sensors that can be added.

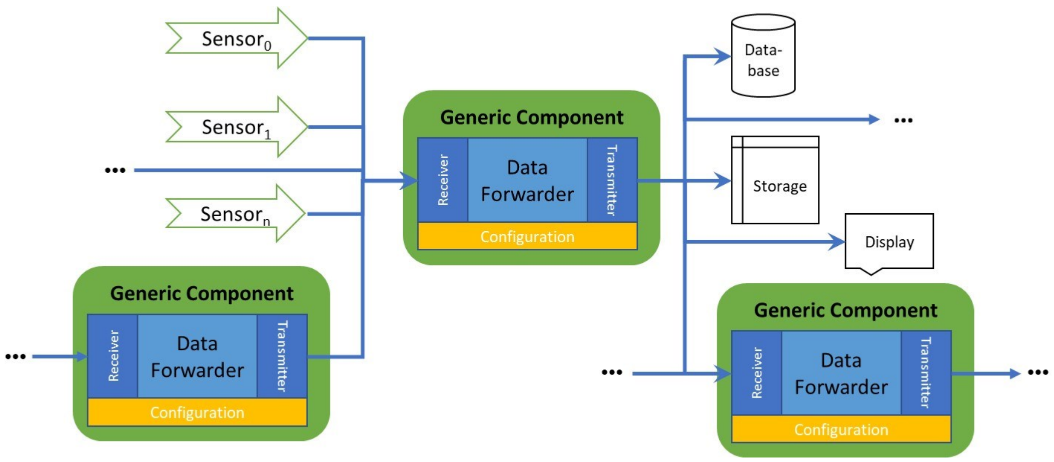


**Figure 50** **Generic Architecture for real-time fusion and logging of sensor data. (Lechner et al., 2019)**

### 11.2.3 Implementation

In order to explore driver distraction, an instance based on the generic architecture is implemented, satisfying the requirements stated previously (cf. Figure 51). A smartwatch (top left) that delivers physiological data such as heartrate, but also movement-related data such as gyroscope- and accelerometer data, is connected via Bluetooth to a smartphone (top right). Smart glasses (bottom right) are connected to a smartphone via WiFi in order to record head movements. The smartphone forwards data both from the smartwatch and the smart glasses to the central database (center) using a 4G LTE connection. Data about the vehicle state is collected by a vehicle data logger (bottom left) that also transfers data via a 4G LTE connection to the database. Results can be investigated in a user frontend on a web page. Regarding the concrete implementation of the system, four main components have been created. (1) An Android Wear App was developed that consists of one component running on the smartphone (Samsung Galaxy A5) interlinked with another component running on the smartwatch (Huawei Watch 2). This design is also referred to as "companion app". It provides the advantage that energy-intense tasks such as GPS data logging and writing data to the database over a network connection are delegated to the more powerful (in terms of battery and computing power) smartphone. One or more smartwatch-smartphone pairings can be used for data acquisition at the same time. (2) A smart glasses App for Microsoft HoloLens was developed for logging head movements. (3) A Vehicle Data Logger that accesses the ODB-II interface of the car provides real-time car data. (4) The time-series database InfluxDB was installed on a server to store all collected data and provide data for further analysis. A high-precision GPS-timestamp was added to all these data sources for synchronization. For more information on used hardware and sensors acting as data sources, the authors refer to Table 20.

The software architecture of the whole system follows the generic architecture. Figure 51 shows the system architecture. This architecture was implemented on several devices in several programming languages. The smart watch/smart phone implementation is in Java, the HoloLens implementation in C# and the Vehicle Data Logger implementation on the BeagleBone in C++.

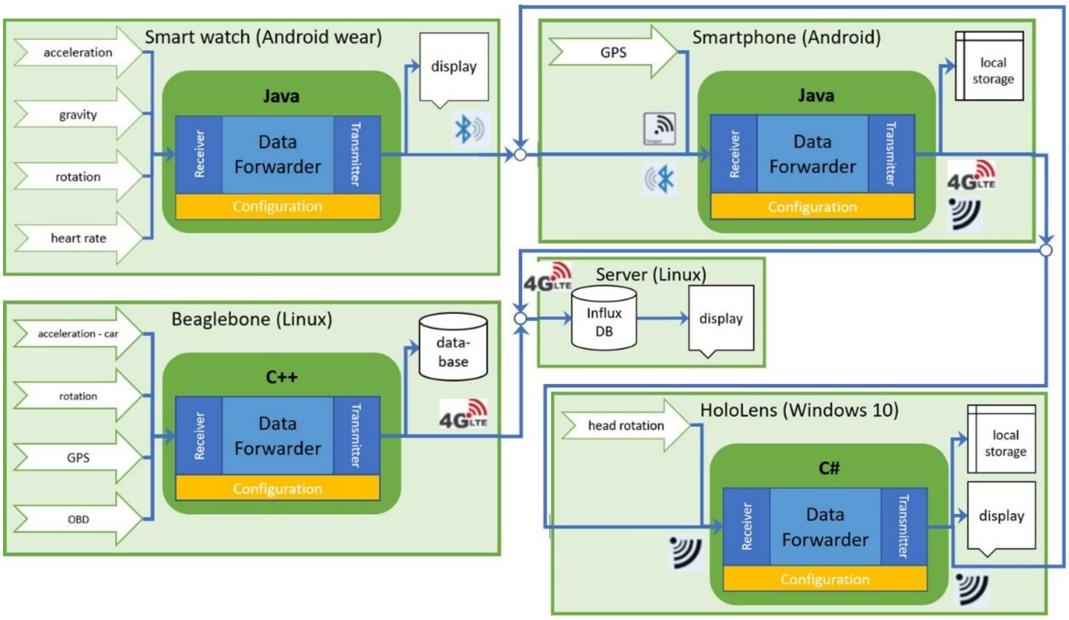


**Figure 51** **System architecture. (Lechner et al., 2019)**

Figure 52 shows the interface specification of the C# implementation for the HoloLens. This principle was also applied to the other involved components written in C++ and Java. The *ISensorDataReceiver* is responsible for acquiring the current log point from the attached sensor. This could be done with a predefined sampling rate, but also asynchronous by triggering the *LogPointAvailable* event when a new sensor value is present. The *ISensorDataReceiver* reflects the receiver of the generic architecture. Every measurement regardless of the sensor type is stored in a *LogPoint*. A log point consists of a timestamp, an array of captured values and names for the captured values. The *IDataForwarder* interface reflects the data forwarder defined in the general architecture. It is responsible for connecting the *ISensorDataReceiver* (receiver) with any appropriate *IDataTransmitter* (transmitter). Additionally, start/stop of the data acquisition is controlled by the *IDataForwarder*. Various configuration settings such as URL endpoints, sampling rates, Sensor IDs, are also handled by the *IDataForwarder* interface. The *IDataTransmitter* interface reflects the transmitter of the general architecture. It is responsible for processing the acquired sensor value, e.g. store the value in a database, a local CSV file or to send the value over a WebSocket connection. All other devices and implementations follow the same generic architecture.

From a practical viewpoint, stability of connections between devices has to be tested intensively in order to avoid unwanted disconnections. In addition, battery runtime of devices must be considered when planning test setups, as continuous receiving and transmitting of data consumes much more power than standard operation. The easy-to-realize regular measurement of timestamps from GNSS-devices and related synchronization of multiple devices facilitates the joint usage of acquired data, as otherwise the devices in instances are prone to clock drift, what renders measurement data unusable.

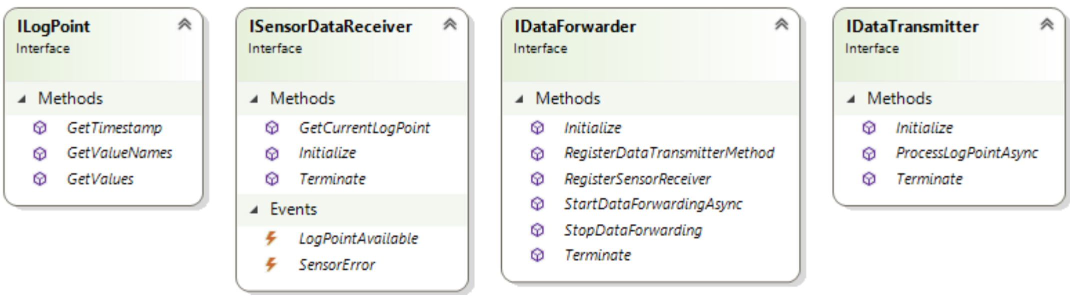


**Figure 52 Interface definition in C# for HoloLens. (Lechner et al., 2019)**

### 11.2.4 Hardware and Equipment

The applied types of hardware and equipment were determined according to the requirements of the experimental setup. A thorough market and requirements analysis led to the application of the hardware and equipment listed in Table 20. In addition, details concerning the sensors of each device from which data was acquired are described. Please note that for all measurements along x-/y-/z-axis the reference system is provided by the sensor itself.

**Table 20 Overview of applied hardware/equipment and acquired sensors. (Lechner et al., 2019)**

| Hardware & Equipment | Sensor description & acquired data from sensors |
|---|---|
| Huawei Watch 2 (2 units) | Off-the-shelf smartwatch based on Android Wear 2.0, offering a variety of sensors for data acquisition; one watch per wrist.<br>- Along x-/y-/z-axis: acceleration including/excluding gravity, measured acceleration (with/without bias compensation), force of gravity<br>- Around x-/y-/z-axis: Euler angles, rate of rotation (with/without drift compensation), estimated drift<br>- Scalar component of the rotation vector<br>- Heart rate in beats per minute |
| Samsung Galaxy A5 (2 units) | Off-the-shelf smartphone serving running app for data acquisition (operating system: Android 7.0), one per smartwatch.<br>- Latitude/Longitude of smartphone's position & GPS timestamp |
| Microsoft HoloLens (1 unit) | Off-the-shelf mixed reality glasses (Operating system: Windows Mixed Reality).<br>- Current absolute position of HoloLens-camera on x-/y-/z-axis<br>- Component of quaternion (x/y/z/w): current rotation of camera<br>- Angle between two sequenced quaternions<br>- Component of Euler angle (x/y/z) representation of rotation |
| BeagleBone Vehicle Data Logger (1 unit) | Non-commercial OBD-II Vehicle Data Logger based on BeagleBoard-platform, including a sensor measurement cape, designed and manufactured at Virtual Vehicle Research Center. Although being a standalone data logger, data is synchronized with other acquired data by GPS timestamp.<br>- Acceleration x-/y-/z-axis including gravity [unit g]<br>- Measured altitude above sea level<br>- Rate of rotation around the x-/y-/z-axis<br>- Latitude/Longitude of measured position & GPS timestamp<br>- Various data according to OBD-II specification |
| GoPro cameras (2 units) | Mounted in the in-cabin room for videos about drivers. As only required for reasons of traceability during analysis phase, video data was not synchronized with remaining data in real-time.<br>- In-cabin video recording |

## 11.3 Demonstration

The implemented lightweight framework was applied in a real-world study. The field study allows to develop improved systems for detecting driver distraction, as the results allow to better infer concrete distraction tasks. In addition, it makes the preliminary examination of the level of obtrusion possible.

### 11.3.1 Study Setup

The implemented framework was used for logging and fusing data from drivers and vehicles to gain insights into detection of distraction during driving. In an experimental series, ten drivers (5 female, 5 male) were asked to perform sequences of specific motions which simulate a variety of distraction tasks. The used vehicle was a Ford Mondeo with automatic transmission with a customized front video recorder capturing (non-synchronized) context of driving for reasons of traceability. Age of drivers ranged from 26 to 43 years, with a mean of 33.6 years (median: 34 years; standard deviation: 5.71). Depending on age, subjects were categorized as *digital natives* (age 18-30, 2:2), *middle age* (born in transition time to digital natives, age 31-40, 2:2), and *digital immigrants* (age 41+, 1:1). Six of the participants wore glasses, and 3 had prior experience with Smart Glasses. Test participants had to carefully read and sign an informed consent prior to the performance, where test purpose and procedure have been described without biasing the test participants for the tests. They could quit the study at any time, nevertheless, all participants finished all their tasks.

A comprehensive distraction detection requires the recording of a variety of different tasks. Thus, two different scenarios were tested: performing tasks in a stationary vehicle (static scenario) and while driving (dynamic scenario), with 10 repetitions for each. The sequence of motions in the static scenario consisted of (S1 – task not related to driving) turn a knob at the center console; (S2 – task related to driving) release and put on hand brake; (S3 – task not related to driving) look for at least 2 seconds into the eyes of the co-driver. In the dynamic case the tasks were (D1 – task related to driving) switch to Drive Mode, release the brake and accelerate to 5 km/h; (D2 – task not related to driving) turn a rotary knob at the center console; (D3 – task not related to driving) look for at least 2 seconds into the eyes of the co-driver; (D4 – task related to driving) stop the vehicle and switch to parking mode. The motions of the drivers were recorded and transmitted over the air to the database, applying the implemented framework. Drivers used smartwatches on both hands for acceleration and rotation of hand motions, and – despite the potential to be obtrusive – drivers wore smart glasses (HoloLens-device) in order to capture head motion data. A smartphone recorded the

GPS position of the vehicle, and a data logger built on the BeagleBone Black platform (beagleboard.org/black) recorded acceleration, rotation as well as data from the OBD interface of the vehicle (e.g. RPM and speed) to collect contextual information.

### 11.3.2 Evaluation of the Framework: Detection of Driver Distraction Tasks

In a preliminary, post-experimental step, data was analyzed and assessed with the objective to detect motion patterns related to driver distraction by applying advanced methods of data science. For this purpose, measurement data obtained from smartwatches, smartphones and smart glasses was used, while contextual information provided by the vehicle data logger was not required. Processing and manipulation of time series data was performed in *Jupyter Notebook* (Kluyver et al., 2016) with *Python* libraries *pandas* (McKinney, 2010) and *numpy* (Oliphant, 2006), while *tsfresh* (Christ et al., 2018) for feature extraction and *scikit-learn* (Pedregosa et al., 2011) for classification were used. Data quality (e.g. plausibility, completeness, …) was checked at first. Subsequently, the data was cleaned and manually split into labelled sequences by the help of the videos recorded by the GoPro cameras mounted in the car. Each sequence represented time series data of a specific distraction task. Sequences of the same type were bundled in sets of identical tasks. In order to have positive and negative examples in each bundle, the authors added time series sequences of comparable length that did not represent the distraction task to the bundle. These were chosen randomly from the available non-distraction data. Feature extraction from these extended bundles of time series data provided input values for machine learning algorithms. In the first step, importance of features was estimated using *ExtraTreesClassifiers*. Next, features were sorted by their estimated importance, and a limited number (2, 3, …, 33, 34) were provided to the machine learning algorithms. Algorithms applied on data covered *Support Vector Classification (SVC)* with linear/RBF kernels, *ExtraTreesClassifier*, *RandomForestClassifier*, *KNeighborsClassifier*, and *AdaBoostClassifier*, all implemented in *scikit-learn* and selected based on experience from similar approaches. Due to the limited amount of data the authors focused on a *k*-fold cross validation approach for each of the scenarios. Therefore, data was randomly partitioned into 10 subsamples of equal size. In each of the 10 training runs, 9 (different) subsamples acted as training data, while the model was tested and validated in terms of prediction accuracy using the remaining subsample.

Table 21 provides an overview of the results of these tests. In addition, the number of (positive and negative) input sequences, used sensor data, best performing algorithm and number of features used for obtaining the best prediction are given. The results show a satisfying accuracy in terms of distraction task classification. For all scenarios except (D4), the best average classification accuracy was greater or equal 90%, and for all except two, it was even greater than or equal to 95%. In general, classification was slightly worse in dynamic scenarios than in static ones, a fact that can be explained by less distinct data induced by vehicle movements. Apart from the finding that individual devices deliver sufficient information to detect simple movements, considering joint sensors like accelerometer and gyroscope does not necessarily lead to better results in such experiments with limited amount of data (see (S2), (D4)): although additional dimensions can increase the amount of available information, the available data becomes sparse due to the increased dimension of the feature space. In order to overcome this potential issue, more data must be provided. Please note that the implemented framework can detect single movements using sensor data from a single device but is also capable to detect different types of distraction by a combination of devices. In order to provide a comprehensive distraction detection, such a combination is required.

**Table 21 Performance on classification of distraction tasks (A/G1: Accelerometer/Gyroscope smartwatch, G2: Gyroscope smart glasses). (Lechner et al., 2019)**

| Scenario | No. of Sequences (pos. + neg.) | Sensor: source of data | Mean accuracy [%] (Std. Dev.) | Best Algorithm | No. of Features (best algorithm) |
|---|---|---|---|---|---|
| (S1) | 29 + 29 | A + G1 | 100% (0.000) | ExtraTreesClassifier | 9 |
| | 29 + 29 | A | 98.0% (0.060) | SVC (linear) | 21 |
| | 29 + 29 | G1 | 100% (0.000) | ExtraTreesClassifier | 3 |
| (S2) | 29 + 29 | A + G1 | 97.5% (0.075) | SVC (linear) | 4 |
| | 29 + 29 | A | 98.0% (0.060) | SVC (linear) | 2 |
| | 29 + 29 | G1 | 97.5% (0.075) | ExtraTreesClassifier | 4 |
| (S3) | 62 + 62 | G2 | 97.5% (0.038) | SVC (linear) | 19 |
| (D1) | 60 + 60 | A + G1 | 94.7% (0.043) | AdaBoostClassifier | 33 |
| | 60 + 60 | A | 93.0% (0.077) | SVC (linear) | 33 |
| | 60 + 60 | G1 | 94.7% (0.043) | AdaBoostClassifier | 33 |
| (D2) | 55 + 55 | A + G1 | 95.4% (0.084) | RandomForestClassifier | 2 |
| | 55 + 55 | A | 95.4% (0.084) | RandomForestClassifier | 2 |
| | 55 + 55 | G1 | 94.2% (0.077) | ExtraTreesClassifier | 21 |
| (D3) | 62 + 62 | G2 | 95.8% (0.042) | SVC (rbf) | 28 |
| (D4) | 52 + 52 | A + G1 | 88.8% (0.058) | ExtraTreesClassifier | 7 |
| | 52 + 52 | A | 90.0% (0.045) | SVC (linear) | 9 |
| | 52 + 52 | G1 | 82.4% (0.074) | ExtraTreesClassifier | 31 |

### 11.3.3 Participant Survey and Results

To evaluate comfort, usability and user experience of applied devices, all participants were asked to fill a questionnaire after the experiment to generate data for a qualitative evaluation of the experiment. In addition, drivers were also requested to provide answers on questions in a break between static and dynamic scenarios.

All participants had to answer some qualitative questions after both, the static and the dynamic scenarios. For example, regarding the wearing comfort of the smart glasses, on a scale from 1 (very uncomfortable) to 5 (very comfortable) all participants answered between 2 and 4 (standard deviation 0.6), on average with 3.2 after the static scenarios. Ratings had a greater variance and were more negative on overall in the dynamic scenario (2.8 on average), e.g. participant 7 rated the wearing comfort very uncomfortable, while participant 3 rated it as very comfortable.

All answers to the question if any technical or external issues occurred during the study were related to the smart glasses, but no other device was mentioned. Despite that, regarding the question if participants can imagine using the smart glasses regularly while driving using a 4-level Likert item (No=1 to Yes=4), results were rather towards Yes after the static scenario (mean: 3, std. dev.: 0.77), and surprisingly even more positive after the dynamic scenario (mean: 3.2, std. dev.: 0.75). Furthermore, participants were asked to answer the question "Describe the smart glasses experience. What do you think?" after static and dynamic scenarios. The answers showed a high variability in terms of weight, value of additional information, and interaction with the smart glasses. Selected positive statements after the static scenario include for example that "additional information using AR is good, as they are always in the visual field" (Participant 1), that the "weight is not annoying" (P2), that the "Experience is better than expected" (P3), or that it is "rather futuristic, but one could get used to it" (P5). In contrast, selected negative statements after the static scenario include for example an experienced "heavy weight" (P7), that the "Experience is worse than expected" (P8) and that it is "heavy, uncomfortable, difficult to interact" (P10). Participants 1, 2, 7 and 10 reported "heavy weight" after the dynamic scenario, too. However, positive statements after the static scenario include that "one gets used to it easily" (P2), that the "experience is better than expected" (P3), that the experience is "good in general. Visualization flickered often but did not disturb (in contrast to the expectation)" (P4), and that the map visualization is "nice, as you always know where you are" (P6).

## 11.4 Conclusion

In this work, the authors focused on exploring comprehensive detection of driver distraction based on a lightweight framework for multi-device integration and multi-sensor fusion. The research on similar frameworks in scientific literature revealed a lack of comparable approaches, what encouraged us to develop a generic architecture for connecting devices and respective sensors in order to achieve the research objective. The proposed generic architecture allows flexible configuration and linkage of generic components, thus being easily extendable by adding/removing generic components. In view of the very heterogeneous landscape, the generic architecture provided a standard frame of reference for all stakeholders involved in the development process. Thus, instantiating the generic architecture was straightforward. Beyond that, the architecture supported the communication with a remote software developer: a significant improvement in terms of software quality could be observed when comparing the first version without using the architecture with the second version based on the generic architecture. Although not being validated, the instance of the framework is expected to work also as baseline for potential interventions in the case of driver distraction: acquired data can directly serve as input for an analysis system with subsequent information of drivers. Of course, this gives rise to research on required real-time capabilities of such a system.

The evaluation of the study in terms of a proof-of-concept has indicated that the implemented framework is capable to acquire sufficient data with respect to quality and quantity from different sensors built into off-the-shelf consumer hardware to detect driver motions as a signal for driver distraction. Driver distraction mechanisms implemented in modern vehicles have a limited feasibility to fully capture all events to finally warn the driver and keep the driver in the loop, as they rely on data generated by the vehicles' own sensory and do not integrate data from external devices like wearables (Kaiser et al., 2018b). With a view to improvement of such systems, data from multiple wearables of a driver may enable a better detection of distraction-related events and thereby contribute to increasing driving safety. Obviously, combining both approaches can turn out to be an effective way. In view of this, fusing data of wearables with other (additional) sources' data would facilitate a more comprehensive distraction analysis. Although the study size and subsequently the size of the collected data are a limitation of the presented research, the small amount of data generated and analyzed has delivered surprisingly accurate results concerning distraction task classification. Thus, a complete analysis, recognition and prediction of individual patterns is conceivable.

The developed framework has a wider application potential than being limited to the presented type of research. Besides the recognition of driver distraction through the classification of critical tasks, further applications based on data acquired by such a lightweight framework could cover, e.g. driving style detection and classification, storage of (driver-related) driving events in the case of accidents like black boxes in aviation, or the personalized adaption of features to habits.

# 12. Use of Automotive Big Data for the Development of Two Applications [19]


**Summary and Author Contribution**

**Concepts and Prototypical Implementations along the VDVC (Paper 4/5)**

This chapter shows concretely how the individual steps of the Vehicle Data Value Chain are run through to finally provide two applications for users. The two applications are the detection of individual driving behaviour and the detection of potholes on the street. In particular, details such as the correct alignment of the coordinate system, as this is essential for further data analysis, and the data pipeline, which shows the sequence and branching of the implementation, are shown.

We applied for a presentation at the BITKOM event "Big-Data.AI Summit" (BAS) in 2018 and 2019 with an abstract, which was accepted in each case. In 2019, I held the presentation, through which we were subsequently invited to elaborate on our experience in a BITKOM position paper on "*Practical Use Cases of Artificial Intelligence & Big Data in Industry*", which is presented here. In doing so, we wanted to show how the individual steps of the vehicle data value chain are concretely run through in order to finally provide two applications for users. The descriptions show developments from several projects (including AEGIS, EVOLVE) in which I have collaborated. I administered the project of writing the paper, reviewed and edited the manuscript.


A vehicle is a computer on four wheels. In a modern vehicle, numerous control units an enormous amount of data, which is collected by the installed sensor technology during vehicle usage is generated to determine the function of the vehicle and the systems. Even the data generated by a single vehicle is the development of data-driven applications and services is highly exciting. The true However, the potential of vehicle data can only be tapped when many vehicles from fleets or even all vehicles on the road at all provide their Big Data treasure. Combining this vehicle data with data from other domains is another desirable goal.

This chapter shows how the data generated by sensors can be used in two concrete applications: the detection of the driving behavior of individual drivers to show the driving risk and the detection of potholes in an urban road network across several drivers and vehicles. The chapter describes the necessary steps from the definition of the applications, the

---

19 The content of this chapter is based on

**Kaiser, C.**; Festl, A.; and Stocker, A. (2020). Nutzung von Automotive Big Data zur Entwicklung zweier neuer Anwendungen. In *BITKOM position paper on Practical Use Cases of Artificial Intelligence & Big Data in Industry (original german title: Konkrete Anwendungsfällevon KI & Big-Data in der Industrie),* pages 10-19. Link: https://bit.ly/2VsqTkv acquisition of vehicle data, data pre-processing, the actual data analysis and the generation of results to support the decision making of drivers and infrastructure managers.

For this purpose, a Big Data Analytics platform was used, which was developed in the project funded by the European Commission and combines numerous open source tools relevant for data scientists such as Spark, Kafka, Flink, Elastic, or Jupyter in one system. The two applications described in the article were realized as demonstrators on this big data analytics platform.

## 12.1 Data Acquisition and Data Structure

For data collection, an in-house developed data logger is used, which was implemented on the basis of a low-cost single-board computer (BeagleBone Black, comparable to a Raspberry Pi). This data logger is connected to the on-board diagnostic interface (OBD) of a vehicle and automatically starts data storage when the vehicle is started and operated. It automatically ends data storage and shuts down when the engine is switched off. In addition to tapping classic vehicle data such as speed or rpm, the logger also contains additional sensors for determining position and for recording rotation and acceleration.

The raw data produced by the data logger is in tabular form and always contains exactly the measured value of a signal at any given time. In particular, several lines can describe different signals at the same time (Figure 53). In addition, the rate at which the values are acquired varies between the signals. Within a signal, this sampling rate is approximately constant, but smaller deviations are possible and common. Since the data logger can be mounted by the drivers at different positions in the vehicle, the coordinate system of the acceleration and rotation sensors of the data logger is basically not automatically aligned with the vehicle – the x-axis of the acceleration sensor, for example, does not have to be parallel to the x-axis of the vehicle.

| Timestamp | Signal Name | Signal Value |
|---|---|---|
| 2020-8-13 8:28:36.206089 | Vehicle Speed | 39 |
| 2020-8-13 8:28:36.226331 | Acceleration-X | 0.476 |
| 2020-8-13 8:28:36.245312 | RPM | 2470 |
| 2020-8-13 8:28:36.268915 | Engine oil temperature | 89 |
| .. | .. | .. |

**Figure 53 Structure of the raw data. (c.f. Kaiser et al., 2020a)**

In order to be able to use the data generated by the vehicle at all for the development of data-driven applications, an extensive data processing process is necessary, which is described in the following sections.

## 12.2 Data Handling Process

In order to extract the information required for the services from the data, three major steps must be taken: In data pre-processing, the data is prepared in such a way that meaningful further processing becomes possible. In the next step, relevant events such as strong acceleration, sharp braking or driving over a pothole are detected in the prepared data. In the last step, the actual analysis, the already calculated components are combined and displayed in interpretable visualizations.

### 12.2.1 Data Pre-processing

In this step, all signals are first scanned for outliers and these are removed. The signals of some sensors such as the acceleration sensor and the gyroscope contain a lot of noise and have to be smoothed. Additionally, the structure of the data is changed by giving each signal “its own” column. To make this possible, all signals have to be interpolated and sampled with a regular frequency. The result is again in tabular form, but now each row corresponds to exactly one point in time and the time interval between the rows is constant, for example 0.1s / 10Hz (Figure 54).

| Timestamp | Engine Speed | Acceleration-X | Vehicle Speed | Oil Temperature | .. |
|---|---|---|---|---|---|
| 2020-8-13 8:28:36.20000 | 1500 | 0.477 | 39 | 89 | |
| 2020-8-13 8:28:36.30000 | 1501 | 0.479 | 40 | 89 | |
| 2020-8-13 8:28:36.40000 | .. | .. | .. | .. | |

**Figure 54 Structure of the pre-processed data. (c.f. Kaiser et al., 2020a)**

A particular challenge is the unknown position of the data logger in the vehicle. For a successful data analysis it is essential to know the directions of the measured accelerations and rotations – but these directions depend on the unknown data logger position. The measurements must therefore be “rotated” parallel to the coordinate axes of the vehicle (Figure 55). To do this, the authors determine the direction of travel, the road normal and the lateral direction from the measurements and calculate a rotation matrix from this.

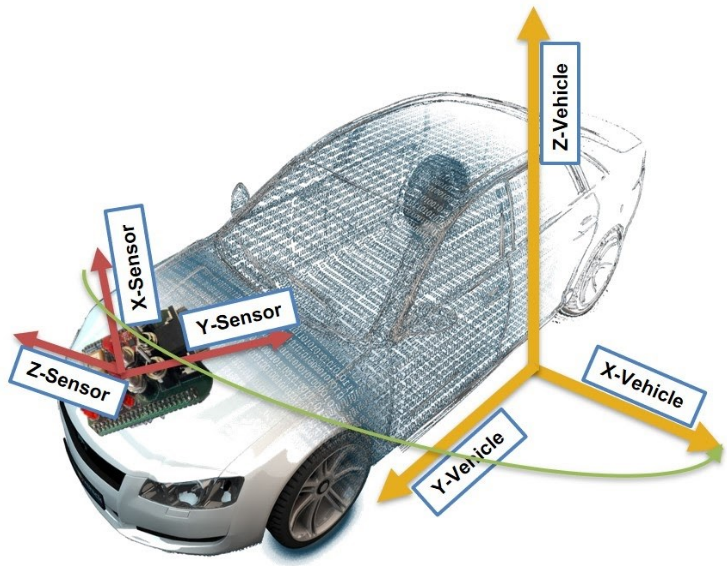


**Figure 55** **Measurements shall be aligned with the coordinate axis of the vehicle. (Kaiser et al., 2020a)**

### 12.2.2 Event Detection

The data prepared in this way can now be searched for events relevant to the services. Different signals are relevant depending on the event type. All events are still post-processed, for example, in order to combine events that are close together in time and are only separated by a short interruption into a single event. The detected events are linked with weather and position data, so that for each event the time and place of occurrence as well as the weather prevailing at that time is known.

#### 12.2.2.1 Potholes Events

This type of event indicates that a pothole (or similar road damage) has been driven over. For detection, the acceleration normal to the road and the “nodding” of the vehicle (i.e. the rotation around its lateral axis) are considered. If the front tyres are in the pothole, the front of the vehicle is lower than the rear, if the rear tyres are in the pothole, it is the other way round. This results in a typical “pitching” movement which can be detected.

#### 12.2.2.2 Acceleration and braking events

The signals of vehicle speed, acceleration in the direction of travel and rotation around the lateral axis (“nodding”) are particularly suitable for detecting strong acceleration and braking processes. The “pitching” is caused by the change in weight distribution when the speed changes: when accelerating, more weight moves to the rear axle – the rear drops and the front rises. When braking, the opposite is true. These movements can be detected. However, since detection using only a single signal can be susceptible to errors, the authors always use several signals in the algorithm, which all have to deflect simultaneously to trigger detection.

#### 12.2.2.3 Fast cornering events

Fast cornering is characterized by high lateral acceleration. However, in order to use this information, one must first detect the curves themselves. For this purpose, an approximate curve radius at each point is calculated from the vehicle speed and lateral acceleration. With the help of this information, the curves themselves can be found and the fast cornering can be detected in further subsequent steps.

### 12.2.3 Implementation

The entire data processing process was implemented on a big-data platform that provides Apache Hadoop, Apache Spark and appropriate user interfaces, among others. All steps there can be executed fully automatically when new data arrives. Intermediate results are stored in separate datasets to enable use in other services or sharing with other users. Figure 56 shows an overview of the data processing chain.

## 12.3 Applications

By collecting, pre-processing and analyzing the data appropriately, applications can be created that can support users from different domains and areas. The authors would like to pick out two examples. One is the detection of the driving behaviour of individual drivers, a use case that is intended to support individual drivers personally in the analysis of their own driving behaviour, and the other is the detection of potholes, which is of particular interest to infrastructure operators.

### 12.3.1 Detection of Individual Driving Behaviour

The driving style of motorists can be very different in many facets (e.g.: comfortable/foresight/aggressive, gear choice, average distance and speed, attention, fatigue, etc.), and depending on the type of vehicle (petrol, diesel, electric with and without recuperation, unladen weight, power) has a small or large influence on consumption, wear and tear and road safety. To illustrate this, the authors use all the events calculated in Section 12.2 and the information to which trip and to which driver they belong, to calculate a 'risk score' that indicates how safe a single trip was. The more safety-related events per trip (in relation to the track length), the lower the value. However, not only the number of safety-relevant events has an influence, but also the circumstances, especially the weather. For example, heavy braking

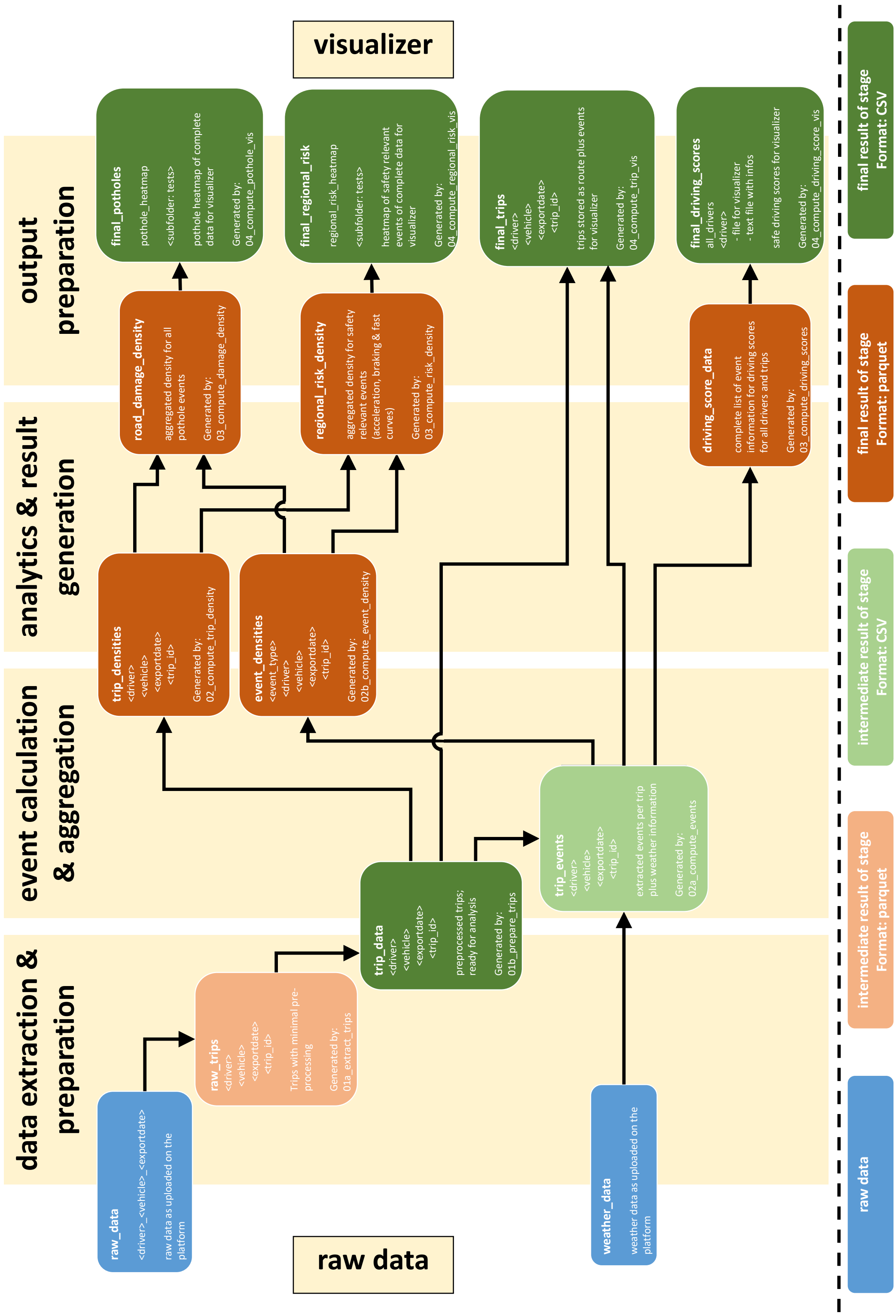


**Figure 56** **The data processing chain to identify, for example, heavy braking processes from raw vehicle data and display them in a visualization. (Kaiser et al., 2020a)**

in rain will result in a greater reduction in the value than the same braking on a dry road. The value itself is a statistical rank, for example a value of 56.72% means that this trip is better than 56.72% of all trips in the database. In a map visualization (Figure 57), for example, the driver is then shown all events of the trip with a marker. Blue markings correspond to strong accelerations, while red markings correspond to strong braking.


Driver1
Total riskscore (%): 50.37

| | trip ID | risk score (%) | duration (h) | accevents (n) |
|---|---|---|---|---|
| 0 | 20181119, Trip_069 | 56.72 | 0.44 | 12.0 |
| 1 | 20181119, Trip_066 | 60.40 | 0.34 | 7.0 |
| 2 | 20181119, Trip_065 | 48.55 | 0.40 | 16.0 |
| 3 | 20181119, Trip_064 | 93.57 | 0.45 | 5.0 |
| 4 | 20181119, Trip_063 | 37.06 | 0.41 | 21.0 |
| 5 | 20181119, Trip_062 | 49.54 | 0.13 | 4.0 |
| 6 | 20181119, Trip_061 | 43.40 | 0.16 | 7.0 |
| 7 | 20181119, Trip_059 | 23.81 | 0.34 | 26.0 |
| 8 | 20181119, Trip_058 | 15.32 | 0.11 | 9.0 |
| 9 | 20181119, Trip_056 | 45.94 | 0.14 | 6.0 |
| 10 | 20181119, Trip_055 | 89.25 | 0.66 | 3.0 |
| 11 | 20181119, Trip_054 | 46.04 | 0.19 | 6.0 |
| 12 | 20181119, Trip_053 | 31.62 | 0.26 | 19.0 |
| 13 | 20181119, Trip_052 | 89.48 | 0.48 | 6.0 |
| 14 | 20181119, Trip_051 | 34.35 | 0.28 | 16.0 |

**Figure 57** **Map visualization of a trip (left) and details of the trips of driver 1, for example the trip with trip_ID "Trip_069" is less risky than 56.72% of all trips (risk score). (Kaiser et al., 2020a)**

### 12.3.2 Potholes Detection

Roads also age, get asphalt cracks, holes, deform. In addition to the legal obligation of road maintenance by municipalities and cities (liability in the event of damage in the event of intent and gross negligence), one naturally also wants to offer the inhabitants and visitors of the municipality/city an adequate infrastructure. The city of Graz consumes approx. 25,000 tons of asphalt and 40,000 tons of frost-proof material per year and thus renovates 100,000 m² of roads annually (Gestrata, 2009). But how do you find out which roads are damaged to what degree and are in most urgent need of repair or renewal? A business enterprise of the city of Graz writes that in 2008 "90% of the road network was visually recorded with three own recording teams" and evaluated "according to a systemized damage catalogue" using the school grading system. This is where the problem already becomes apparent: such an analysis is costly, time-consuming, not completely objective and does not record all roads. This is where the second application case comes in, the automatic detection and display of potholes/road damage on a map. Specifically, the detected damages are displayed in a heat map – the color corresponds to the safety with which a pothole was detected (Figure 58).

The data originate from several vehicles that are used privately in the Graz area. These were equipped with sensors by us: Thus, the authors have already achieved a high coverage of the inner city of Graz. Potholes are detected as an event, as described in subsection 12.2.2, and stored together with GPS position, date and time. The greater the number of trips over the same section of road where a pothole is detected, the more serious the problem is. The map shows this as follows: Smaller problems are violet, medium problems are green to yellow, and big problems are red. This can also be displayed as a table to compare the road condition on the basis of objective figures, even if the damping of the passenger compartment where the sensor is placed distorts the result a little. The individual problem areas can be communicated and repaired with pinpoint accuracy (as far as GPS accuracy allows).


**Figure 58** **Heatmap of detected potholes/road damage in a part of Graz, Austria. (Kaiser et al., 2020a)**

## 12.4 Conclusion

The present chapter has shown how data generated by a vehicle via a data processing pipeline enables two interesting applications, the detection of the individual driving behavior of individual drivers and the aggregation of these data into a risk heatmap, as well as the detection of potholes in an urban road network and the visualization of the result in a heatmap. The article refers to the numerous challenges in data processing and shows possible solutions.

# 13. A Vehicle Telematics Service for Driving Style Detection: Implementation and Privacy Challenges [20]

| Summary and Author Contribution |
|---|
| **Concepts and Prototypical Implementations along the VDVC (Paper 5/5)**<br>In this chapter it is shown how the authors have implemented individual components of the Vehicle Data Value Chain for a proof-of-concept. In line with the previous chapter, the concept of the data logger includes the possibility to set a privacy level. Therefore, this chapter also presents results of an empirical study on privacy, e.g. which data, for which services and under which circumstances the survey participants would log and provide the data.<br><br>Here, I focused on two things: (i) to mention the privacy topic again in connection with vehicle data, as I consider this to be very important for acceptance among users, as well as (ii) the concrete implementation example to help interested software developers understand how they themselves can develop a service step by step (along the VDVC). Thereby, I was responsible for Section 13.1, I administered the implementation illustrated in the paper as task leader (project SCOTT), and I homogenized and revised the inputs of the co-authors. |

Increasing road safety is a major worldwide challenge. Though road safety in the EU has improved in the last decades, still more than 25.000 people have lost their lives on EU roads in 2017 (European Commission, 2018a). Harsh driving remains one of the major causes of accidents. A report from the NSTSCE (Camden et al., 2015) lists violating speed limits, excessive speed and lateral acceleration on curves, unplanned lane departures, frequent hard braking, close following distances, lateral encroachment, failure to yield at intersections, and general disobedience of the road rules as risky driving behavior. The NSTSCE report continues that a reduction in such risky driving should lead to a reduction in accidents and related deaths and injuries. Hence, making harsh and risky driving better visible to drivers and other stakeholders such as traffic planners or public authorities is a useful tool to develop better strategies for road safety improvement. In order to make it visible, vehicle telematics

[20] The content of this chapter is based on

**Kaiser, C.**; Stocker, A.; Festl, A.; Djokic Petrovic, M.; Papatheocharous, E.; Wallberg, A.; Ezquerro, G.; Ortigosa Orbe, J.; Szilagyi, T.; and Fellmann, M. (2020).
A Vehicle Telematics Service for Driving Style Detection: Implementation and Privacy Challenges. In *Proceedings of the 6th International Conference on Vehicle Technology and Intelligent Transport Systems - Volume 1 (VEHITS),* ISBN 978-989-758-419-0, p.29-36. DOI: 10.5220/0009329400290036

data of so-called Quantified Vehicles (Stocker et al, 2017) provides the baseline of data needed for the analysis. However, in the current age of glass people, the road to total monitoring, such as automated penalties, is not far away. Hence, privacy and trust are among research relevant topics in that field (Kaiser et al., 2018b) and must be achieved to get drivers to join in.

In the following sections, an empirical study on vehicle telematics data sharing is presented, which results into a preliminary model of the willingness to share data and five privacy levels that users would like to have to choose from. Although there is empirical evidence in the literature on actors of a service ecosystem (e.g. Kaiser at al., 2019b) and the value-adding steps, descriptions of concrete implementations are still missing. Hence, an actual implementation of a vehicle telematics service is described afterwards, by outlining the required data acquisition, the data analytics process from data collection, the data computing in the cloud, and data use within an information system running on a smartphone developed along the steps of the so called Vehicle Data Value Chain (Kaiser et al., 2019a). The chapter concludes with a discussion of the results and their benefits to drivers and other stakeholders and a brief outlook.

## 13.1 Empirical Study on Privacy in Vehicle Data Sharing

For a long time, the industry was told that one would have large data treasures lying around if one only had to lift them. That this is not the case is shown by many practical examples where it is found that large amounts of data are available but not the right data to derive profitable findings. The situation is similar with vehicle telematics data: Exciting applications require big amounts of detailed data from a range of vehicles and drivers. Unfortunately, after several scandals in recent years where data was stolen or misused, many users lost their basic trust and are now more sensitive about who they give the data to.

To investigate background in this field, the authors conducted a literature review and came up with the search string

```
Data Sharing OR Data Sharing Theories AND
(Automotive OR Automobile OR Vehicle OR
Car OR Vehicle Data)
```

which the authors applied to popular scientific search engines (SCOPUS, Google scholar, AISel) to identify 16 relevant results with data sharing theories. As a summary, the majority of the 16 papers focus on technologies and application possibilities and give just little insights why someone would or would not share his driving data.

In a next step, models and theories widely used for technology acceptance were investigated by the authors. However, neither the *Technology Acceptance Model* (e.g. TAM3) (Venkatesh and Bala, 2008), nor the three theories *Unified Theory of Acceptance and Use of Technology* (e.g. UTAUT2) (Venkatesh et al., 2012), the *Theory of Reasoned Action* (TRA) (Fishbein and Ajzen, 1975), or the *Theory of Planned Behavior* (TPB) (Ajzen, 1991) seem to fit ideally. In contrast, Ju and Mou (2018) show in their research model hypotheses that the willingness to disclose personal information depends not only on *Controls*, e.g. age or gender, which influence willingness, but also on the *Costs* and *Rewards* for disclosing privacy, an interesting approach.

Based on the literature analysis, two of the authors compared their practical knowledge with the above-mentioned models and theories, and finally derived their own model, which is described in the following.

### 13.1.1 A Data Sharing Willingness Model

The authors found out, that the *willingness to share vehicle data* depends on the *intended usage*, which in turn depends on a mix of *Benefits* and *Efforts*, as visualized in Figure 59. Per intended usage, different benefits have a positive effect and can range from self-awareness, optimization, rewards, image, comfort to predictive maintenance and thus tempt a potential user to consider sharing vehicle telematics data for the intended usage. In contrast, per intended usage, different efforts have a negative effect, e.g. costs (acquisition), the technical effort for installation, ongoing expenses (operation, mobile phone costs), irritation through advertising/spam and lower privacy speak against a use.

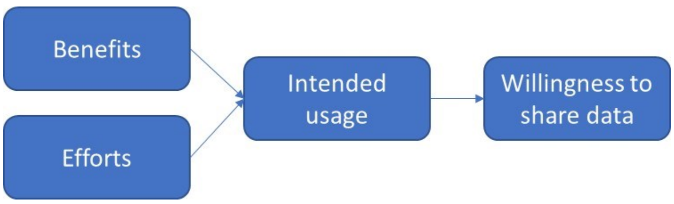


**Figure 59 A preliminary model of the willingness to share data, e.g. vehicle telematics data. (Kaiser et al., 2020c)**

On this basis, the authors conducted an empirical online survey, which was distributed to members of the Faculty of Computer Science at the University of Rostock and to researchers at Virtual Vehicle Research GmbH. With the 42 survey participants, the authors tried to find out whether someone would pass on their vehicle telematics data, for which application cases they would do so, and whether they would change this situationally, for example to block data transfer in certain periods of time. For the situational adaptation of the data transfer, it was particularly interesting for us how many levels there should be here. Levels can range from, e.g. a binary level system that is either on or off, up to a fine-granular system with several levels which offer anonymization options and forwarding for selected service providers/services only.

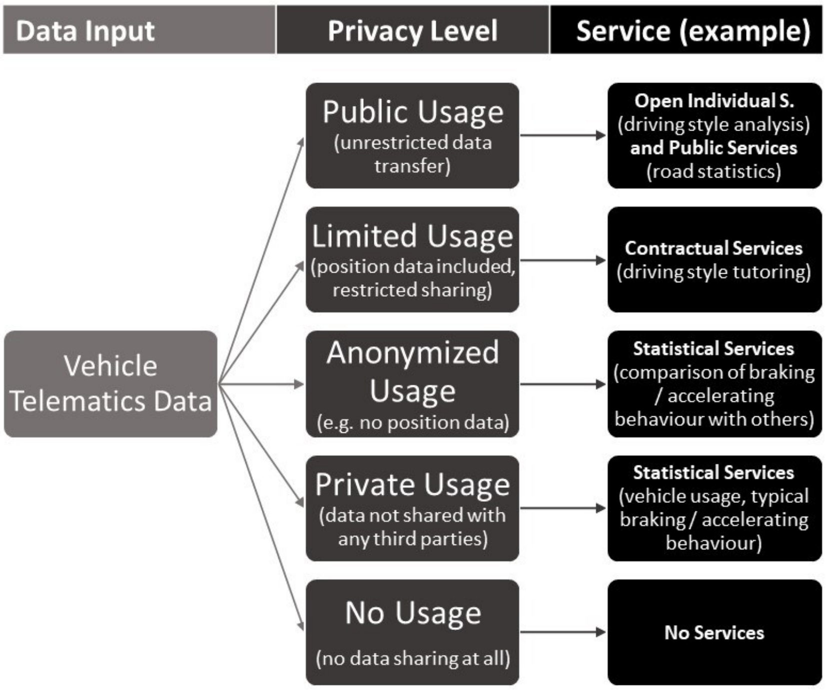


**Figure 60 Empirical result: privacy levels for vehicle data sharing. (Szilágyi, 2019; Kaiser et al., 2020c)**

### 13.1.2 Empirical Results

Although the average of the 42 study participants proposed to provide four privacy levels (average 3.97, standard deviation 1.15) and described them, the two researchers who analyzed the results and synthesized the individual statements into the model shown in Figure 60 detected five privacy levels, as only five levels include all viewpoints mentioned, namely contractual services vs. open services, anonymized data vs. not-anonymized data, private usage vs. public usage. However, if it has to be four levels, then the privacy levels *Private Usage* and *Anonymized Usage* can be merged, as this aspect has the lowest priority. The individual levels are described in the following.

**Level *No Usage*** does not allow any collection or sharing of vehicle telematics data, and thus prevents any services.

**Level *Private Usage*** uses collected vehicle telematics data locally in the vehicle to create e.g. statistics on driving behavior, which only the driver/owner can see in order to interpret and optimize oneself. However, no data is shared with any third parties, thus no services, other than installed services in the vehicle, can be used.

**Level *Anonymized Usage*** includes the services installed in the vehicle, and additionally sends small amounts, e.g. statistics or histograms, of anonymized data to chosen third-party services. The driver can not be identified, due to anonymization, e.g. location data is not shared.

**Level *Limited Usage*** is intended to optimize traffic for everybody, thus road specific data like traffic jams, potholes, accidents, slipping wheels, etc., is shared with other drivers on this road through a service. Hence, also a bigger amount of vehicle telematics data is shared, but still not all of them. and again, anonymized for third-party services.

**Level *Public Usage*** does not restrict data transfer – all data will be shared using a proper sampling rate per signal (perhaps on demand). Third parties will be able to use this data without anonymization, e.g. to enable the comparison between friends or services which analyze regional differences in driving behavior.

The survey participants also were asked to state, how interested they are in sharing their data for a particular domain, ranging from 1 (not likely) to 5 (very likely). In general, the survey participants' willingness to share their vehicle telematics data for each domain (c.f. Table 22) were lower than in their interest. To summarize, the majority would provide data for traffic improvement and emergency services, while all the other mentioned domains would have to offer an individual added value (benefit) so that users give their data for it.

**Table 22 How willing are survey participants to share their data for a given set of domains. (Kaiser et al., 2020c)**

more likely ⇔ more unlikely

| Domain | Average (1 to 5) | Standard deviation |
|---|---|---|
| Community games | 1.64 | 1.06 |
| Automobile club | 1.86 | 1.18 |
| Pay as you drive insurance | 2.02 | 1.33 |
| Services for drivers | 2.74 | 1.43 |
| Vehicle improvement | 2.86 | 1.41 |
| Public governance | 2.86 | 1.39 |
| Weather detection | 2.98 | 1.56 |
| Research (novel services) | 3.29 | 1.49 |
| Traffic improvement | 3.67 | 1.44 |
| Emergency services | 4.00 | 1.40 |

Since different privacy levels lead to an increased effort for the service development if one offers a reduced solution for the privacy levels Private Usage, Anonymized Usage and Limited Usage, functionalities for setting privacy levels are difficult to find or not implemented at all in reality, although the customers would approve of this.

Hence, in the following subsections of Section 13.2, the authors show how selected steps of an actual implementation approach of a vehicle telematics service for the driver can be done, and thereby reflect where and how privacy levels have to be taken into account.

## 13.2 Vehicle Telematics Service Implementation

In order to develop a smartphone application prototype which informs the driver about his recent driving style, several steps along a vehicle data value chain are involved and thus explained in the following subsections, to provide an overview of complexity and dependencies. According to (Kaiser et al., 2019a), the value chain consists of the steps *Data Generation*, *Data Acquisition, Data Pre-Processing*, *Data Analysis*, *Data Storage* and *Data Usage*.

In the following implementation example, the Vehicle Data Logger (*Data Acquisition*) collects data generated by vehicle sensors from the vehicle's bus system via the OBD interface and additional data generated from sensors at the logging device (*Data Generation*). A Cloud Platform receives the data and acts as temporary raw data storage and platform for data pre-processing and analysis (*Data Pre-Processing and Data Analysis*), e.g. use of an algorithm to detect harsh brake events. The processing results are then stored permanently (*Data Storage*) and provided to end users in a proper form (*Data Usage*), e.g. using a smartphone application.

Privacy should play a role in data acquisition, so that only authorized data is collected. Per privacy level, different services are made possible with the data, meaning that individual data pre-processing and data analysis processes are needed per privacy level.

In the service, the driver wants to learn about his driving style, e.g. get a score per trip which indicates if it was good (100), bad (0) or somewhere in-between, and wants to be able to check where events like harsh braking or harsh accelerating have been detected. While event detection and route recording can be done locally in the vehicle with a low privacy level, at least privacy level Anonymized Usage is needed to calculate the driving score, as in this case the amount of events are compared with the data from other drivers.

### 13.2.1 Vehicle Data Logger

The first building block of this service is a data acquisition system, called vehicle data logger, which acts as gateway device to collect vehicle telematics data. The vehicle data logger is based on a BeagleBoard single platine computer featuring an additional “sensor cape” stacked onto it with GPS, rotation and acceleration sensors. The time-series data captured by the logger is stored on a MariaDB database on the logger. As soon as a connection is established via the mobile network, the logger can send captured data to the cloud platform. A rotary switch on the hardware device can be used to set the privacy level. To reduce the workload of mobile network connections and to increase the throughput, SenML data format is used for transmitting the data. SenML is a compromised data format especially developed for IoT device data. A TPM module is added via another stackable “cape” to provide encryption possibilities. A configuration file on the SD card can be used to configure database name, username, password, which sensors are recorded and the online API the data is sent to. A more detailed specification of the logger is provided in Papatheocharous et al. (2018) or Lechner et al. (2019).

### 13.2.2 Cloud Platform

The data logger described in the previous subsection sends data to a defined channel of a message broker, in this case a MQTT (Message Queuing Telemetry Transport) Broker. One of the MQTT listeners is triggered, parses and formats the data if needed and forwards it to a cloud platform hosted by the company RISE. The cloud platform aims to support the exchange of data between devices and accommodate the deployment of cloud computing services. Connection between the cloud platform and devices occurs either directly or through a gateway. Any authorized smart device with connectivity can go through a gateway (a device or software designed for the purpose) to exchange data with the cloud platform. Devices may also choose to bypass the gateway and exchange data with the cloud platform directly. The data exchange can be carried out through MQTT or HTTP connections.

The cloud platform offers telemetry ingestion (accepts data), stream processing (data flows are processed and converted to unified formats), storage (data is stored in one or several databases), analytics (data is statistically and semantically analyzed to extract information), machine learning (data is processed with machine learning algorithms to extract knowledge and intelligence), visualization (data is depicted in meaningful charts and graphs to extract summarized information, generalizations, locate anomalies, etc.), lifecycle management (consists of supporting functions for the management of devices, such as software updates or (re)configuration), state (consists of storing the state of devices at all given times), and, finally, apps (consist of extended applications and services that can extend the platform, and offer some additional functionality or end-user value).

### 13.2.3 Cloud Computing Services

Different types of cloud computing services can be deployed on the cloud platform. Foremost the solution provisions for edge and cloud computing services for safe and secure connected mobility applications. The services accommodate data ingestion, storage, processing and management.

Data ingestion is made primarily through an MQTT broker, formatted as SenML JSON (Jennings et al., 2018). Use of the broker and the publish-subscribe pattern (Birman and Joseph, 1987) makes it possible for remote and external trusted partners to receive raw data, if necessary. Additionally, to increase trust in privacy, users should be able to listen to the defined channel (decrypted for them) to be able to check which data is sent.

Data is stored through deployed databases, after any required preprocessing is carried out. Timescale (a module of PostgreSQL) for time-series data is used. Access to the databases is encrypted with Transport Layer Security (TLS) and certificates from Let's Encrypt. Let's Encrypt (Internet Security Research Group (ISRG), 2019) is a certificate authority that provides free certificates for TLS encryption via an automated process.

Management is accomplished through the use of several Docker (Merkel, 2014) tools, i.e., Engine, Compose, Swarm, Machine, and Machinery (Frécon, 2018). They offer efficient system architecture deployments for any type of cloud provider and provision for the daily operations of a number of containers and solutions necessary for the applications, such as data backup, restore and application supervision.

### 13.2.4 Processing of Data

Docker containers were set up in this prototype to process the data. Pre-processing and data analysis are dependent on the privacy level chosen, as each service has specific requirements for sampling rate or the need of position data. However, in this case, to inform the driver about his recent driving style, the two pre-processing steps (i) resampling and (ii) coordination system alignment of vehicle and logging device start the processing, before algorithms detect four event types (harsh brakes, harsh accelerations, standstills and potholes) in the data. Later, they are used to calculate an indicator how safe a driver’s trip was, compared to other trips in the database.

Hence, the initial phase in the pre-processing of data is the resampling of the raw data, namely the measurement signals (e.g. acceleration, speed, GPS, etc.) which were recorded with individual sample rates on the data logging device. In data analytics this step is a challenge, as some measurement signals are recorded at irregular time intervals. For example, to receive data collected from the vehicles OBD interface, the Vehicle Data Logger is posting a request to the OBD interface. As the OBD device has low priority, while all other ECUs in the vehicle have a higher priority, it might happen that time intervals between two values for one signal type increase up to seconds. For each signal the recorded values must be interpolated/extrapolated using polynomial functions (e.g. natural splines), so there are no discontinuities in curves, and they are smooth. Hence, in this case a resampling of the signal values at the regular time interval of 10 Hz (1/10 sec) provides the data for the further analysis.

The next pre-processing step is to align the coordinate system of sensor with coordinate system of the vehicle. It is usually unknown, how the Vehicle Data Logger was exactly mounted in the vehicle. Hence this is an important step to e.g. detect forward driving as forward driving if the logging device was mounted in the wrong direction, but also a few degrees shift would already make an impact in detecting i.e. hard accelerations and hard brakes. For solving this data analytics task, the following assumptions are adopted: the position of sensor is fixed during the trip and on average the vehicle Z-direction coincides with gravity vector, due to the fact that the vehicle drives horizontally. Then the following steps can be taken: identify Z-direction of the vehicle as direction of gravity, identify periods of deceleration and acceleration in the measurement using OBD data, identify driving direction as vector between the mean values for acceleration and deceleration, orthogonalize the driving direction and gravity vectors, compute vector in lateral direction as cross product of driving direction and gravity, compute rotation matrix from the driving direction, gravity and lateral direction vectors and finally rotate accelerometer and gyroscope measurements.

From the pre-processed measurement data, four different event types are extracted: brake, acceleration, standstill and pothole. Categorizing brake and accelerate events is based on the vehicle speed in combination with acceleration and deceleration values. Figure 61 (right) shows a detected harsh acceleration event, where the driver accelerated from 22.28 km/h to 37.28 km/h within five seconds. Identifying a pothole event is based on detecting acceleration in Z direction and rotation around Y-axis (pitch). For example, both signals indicate short peaks at the beginning and the end of a pothole.

The safe driving score is based on statistical ranks. For each trip and each event type, event-rate per time unit is calculated (e.g. a trip has 0.1 hard brakes per hour). The trip-event-score is also calculated as the percentage of trips with the lower event-rate, for the current event-type. The score for one trip, trip-score, is calculated as the mean of all trip-event-scores for that trip. Finally, the driver-score is the latest value of the exponentially smoothed time series of trip-scores for that driver. The values for driver-score and trip-score are scaled from 0 to 100. Hence, a safe driving driver-score of 97 would mean, that this driver is currently better than 97% of all drivers in the database. A low safe driving score indicates a risky driver.

The results of data processing can be obtained on trip level (trip meta-data like start time and end time, trip specific events with GPS location and meta-data, and a safe driving trip-score), or on driver level (overall safe driving driver-score, summed up statistics like kilometers driven or events for a requested time-period like last month). A PostgREST API takes data requests of authenticated users, and provides the data, e.g. for the smartphone application described in the following subsection.

### 13.2.5 Smartphone Application

The Android *Offline Trip Analyser* (OTA) mobile application, will present to the users the information produced in the trips they conducted. The application collects the trip and event information from the PostgREST API. The purpose of the application is to present the user detailed information per trip with a focus on safe driving relevant events.

Once a user is logged into the application, the user can switch between four menu items *Home*, *History*, *Cars* and *Profile* (c.f.Figure 61, left, on the bottom).

The *Home* page, visualized in Figure 61 (left), visualizes a general summary and a summary of the events that have occurred during a selected time period, configurable with the filter on the right top, e.g. last day, last week, a specific selected timeframe or always.

On the *History* page, users will find the history of their trips along with brief details, e.g. starting position, ending position, trip-score and privacy level per trip, sorted from the most recent to the oldest. Clicking a trip, if applicable, a sub-page on details of the individual trip is shown, including graph visualizations of the course of vehicle speed, RPM, etc., and an event overview of the trip per event type. The application user can also switch to a sub-page visualizing a map of the individual trip (c.f. Figure 61, right), to see the trip route on a map. Markers represent the detected events at the event location and allow interactive analysis of the events, as a tooltip pops up on click providing detailed information, e.g. duration, start- and end-speed of the acceleration event in Figure 61. Hence, the user can zoom and navigate through the map and click markers. Furthermore, below the map, four tables (one per event type: brake, acceleration, standstill, pothole) list all event occurrences of the specific event type in this trip, to provide another viewpoint on the data.

The SCOTT OTA aims to make it easier for the drivers to keep detailed control of trips, learn from it in order to improve their driving behavior. The safe driving score per trip gives a quick indicator and an objective evaluation of the driving style, while it is possible to analyze every event in detail as well if needed.

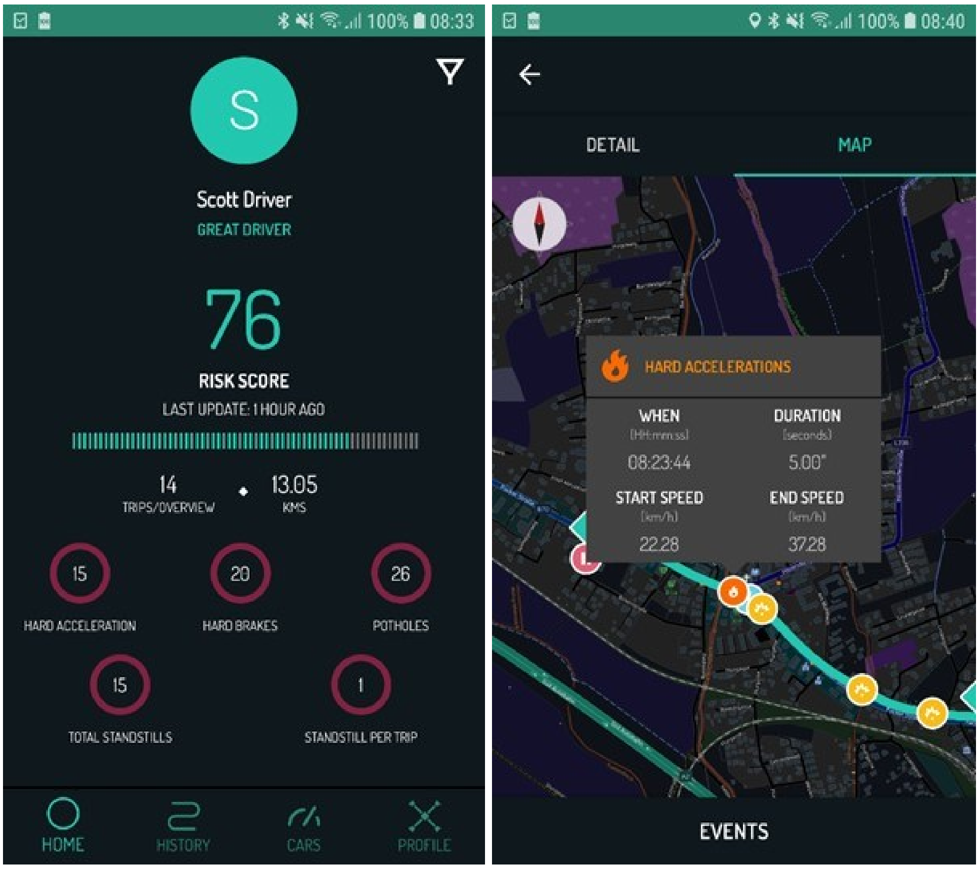


**Figure 61 Smartphone App for drivers: Home (left) and Trip Map (right). (Kaiser et al., 2020c)**

## 13.3 Conclusion and Outlook

In this chapter, potentials and issues of vehicle telematics data sharing is investigated. Hence, the authors show a preliminary model of the willingness to share vehicle data, before the authors conduct an empirical study on the topic of privacy. Furthermore, the authors show how an actual implementation of a vehicle telematics service can look like, and where privacy has to be taken into account.

The results clearly show the single development steps along the vehicle data value chain, namely data collection, data computing in the cloud, and data use within an information system running on a smartphone, to provide a safe and secure connected mobility smartphone application for drivers based on vehicle data. Furthermore, for every step a privacy-preserving way of a vehicle telematics service is discussed.

While the potential of data-driven connected mobility services as well as the potential of driver statistic services is already proven by literature (Kaiser et al., 2018b) and a bunch of start-ups operating in this field (Kaiser et al., 2017b), this chapter misses a structured literature analysis for that topic, which is a clear limitation. Furthermore, the presented results, the data collection, the computing in the cloud and the secure connected mobility smartphone application need to be evaluated for scalability, to prove if hundreds of users can use it simultaneously.

As an outlook, the mentioned privacy issues to be tackled, which are now discussed in each implementation step, will be implemented to evaluate this as well.

# 14. Understanding Data-driven Service Ecosystems in the Automotive Domain [21]


Summary and Author Contribution

**Analysis of Data-driven Service Ecosystems (Paper 1/2)**

Vehicle data-driven services are becoming increasingly relevant, but still little is known about the actors involved. This chapter examines the ecosystem transformed by the emergence of vehicle data-driven services. To improve the understanding, it analyses both the actors and the role of vehicle sensor data from an ecosystem perspective. Based on a literature analysis, results from expert interviews are used for the examination. Eleven experts are involved in the latter, including representatives of service providers, authorities, data market providers, research institutions and vehicle manufacturers. By combining both results relevant actors in the ecosystem as well as their relationships, data flows and services are gained. Results thus provide a fundamental understanding of data-driven service ecosystems in the automotive domain and form the basis for future IS research on (big) data flows and analytics within such ecosystems.

In this paper, I wanted to demonstrate in one of the most important IS conferences that *Data-driven Service Ecosystems* are not sufficiently researche, and that even domain experts have different views on what the ecosystem currently looks like, what makes it difficult for new actors (e.g. ICT companies) to participate. To do this, I analyzed related work, conducted eleven expert interviews, and synthesized a high-level model for data-driven service ecosystems and a detailed data ecosystem model from the results to increase understanding about the emergence of *Data-driven Services*. The interviews also revealed the dilemma that vehicle manufacturers do not want to/are not able to develop the services themselves, but also find it difficult to give external service providers access to the data.


Digital innovation, digitization and digital transformation are on everyone's lips today. Both leading business analysts and strategy consultants such as Accenture (2016), Bain (2017), Deloitte (2017), KPMG (2017; 2018), and McKinsey (2016; 2017) and researchers (Kessler and Buck, 2016; Mikusz and Herter, 2016; Kohl et al., 2017; Mocker and Fonstad, 2017b) report on the increasing digitization of the automotive industry. In addition to the importance of current technological trends in shaping the automotive future, such as autonomous driving (Kung and Lin 2018), connected vehicles (Gerloff and Cleophas, 2017), intelligent manufacturing and maintenance (Laubis et al., 2016; Gerloff and Cleophas, 2017), blockchain (Kaiser

[21] The content of this chapter is based on

**Kaiser, C.**, Stocker, A., Fellmann, M., (2019). Understanding Data-driven Service Ecosystems in the Automotive Domain. In *Proceedings of 25th Americas Conference on Information Systems (AMCIS 2019)*.

et al., 2019), and big data analytics (Dremel et al., 2017b), analysts are increasingly focusing on new relationships within digital ecosystems.

Coined by ecologists, the 'ecosystem' term has been taken up in literature on economics (Pilinkienė and Mačiulis, 2014) and seems to have gained some attention in the IS community over the last few years (Nischak et al., 2017; Ceccagnoli et al., 2012), too. In recent decades, both hardware-based platforms, such as PCs, mobile computing systems, and video game consoles, and software-based platforms, such as smartphone app stores and online marketplaces, have become increasingly important with regard to the creation of ecosystems (Evans et al., 2006). However, many traditional organizations do still not think in terms of ecosystems. Instead they are thinking of participating in or even controlling a linear value chain (Weill and Woerner, 2015). In contrast, there are ecosystem drivers such as Amazon, Apple or Google, which provide platform technology to other participants to enable new business opportunities. In such platform ecosystems, technology owners co-create business value with other organizations that use their platform (Ceggagnoli et al., 2012). In the automotive domain, vehicles, too, can be understood as a platform technology. This way, vehicle manufacturers can build ecosystems together with other stakeholders, e.g. actors from information technology (Riasanow, 2017). Following this analogy further, vehicle manufacturers could even promote the creation of platform ecosystems. As an example of an ecosystem in the automotive domain, Vaia et al. (2012) present a case of one of the largest Italian motor insurers together with a system integrator for telematics-based insurance.

However, vehicles do not only establish platforms, but, as computers on wheels, they also generate valuable data (Stocker et al., 2017a). Vehicle data are the main source of data-driven service ecosystems as they can be linked to other data sources to enable innovative uses. As an example, the technology startup Automatic (Automatic, 2018) has begun to establish a data-driven service ecosystem based on its core technologies. Automatic acquires a limited set of vehicle sensor data using a branded (external) gateway device (in this case, an OBD dongle) connected to a vehicle's on-board diagnostics (OBD) interface. This enables the synchronization of driving data with a connected smartphone, allowing the driver to keep track of driving behavior via the Automatic smartphone app. On the way to an Automatic data-driven service ecosystem, third-party apps that also use the vehicle data captured by OBD dongles from Automatic have been developed for the Automatic platform (Automatic, 2018). Third-party applications in this ecosystem include, for example, the creation of invoices based on mileage (FreshBooks, 2018), vehicle cost management (Xero, 2018), and management of track earnings and warnings during ride sharing (SherpaShare, 2018). The Automatic example shows the potential of data-driven service ecosystems in the automotive sector.

The ecosystem concept seems to be useful to better understand current and future relationships between actors and services. However, it has rarely been applied in the automotive domain as a tool for understanding data and value flows and cooperation between different actors. Furthermore, the concept of data-driven service ecosystems in the context of automotive innovation has not been considered very relevant in IS-related literature. In line with the research agenda for Vehicle Information Systems (Kaiser et al., 2018b), the authors have formulated the following research question: *What are data-driven service ecosystems in the automotive domain and what are their actors and value flows?* To answer this research question, the authors have conducted an analysis of IS literature to better understand the concept of 'ecosystems' in general. In a second step, the authors derived a preliminary model of an ecosystem with actors, value flows and services from the literature to understand the genesis of data-driven service ecosystems in the automotive domain. The authors finally discussed this model and elaborated it using knowledge of eleven experts in an expert interview study.

In the remainder of this chapter, the authors begin by explaining the ecosystem concept and how it can be applied to the automotive sector. This is followed by findings, including a list of actors, services, a high-level model of the data-driven service ecosystem and detailed views on the underlying data and service ecosystems as described in detail in Section 14.2. The results show that the ecosystem concept is suitable to understand the challenges of data-driven service ecosystems in the automotive domain.

## 14.1 The Ecosystem Concept and Related Work

Since the goal of the chapter is to shed light on the components and relations that form data-driven service ecosystems, the authors first present related work on ecosystems that aims to further explain and clarify the ecosystem concept in the automotive domain. The authors then refer to related streams of research that address important aspects of such ecosystems. In general, the concept of an ecosystem is inspired by natural ecosystems, where it describes the relationships and interactions between living organisms and their environment (Schulze et al., 2005). The concept of ecosystems has been transferred as an analogy to various scientific disciplines, including Social Sciences, Computer Science and Natural Science (Briscoe and De Wilde 2006). In order to delineate artificial ecosystems from natural ecosystems, some authors add further attributes to the term to qualify it, e.g. *software* ecosystems, *business* ecosystems or *digital service* ecosystems (Immonen et al., 2015). However, a commonly agreed definition of the term does not yet exist. With this in mind, Nischak et al. (2017) conducted an analysis of peer-reviewed articles on ecosystems and information systems that introduced the ecosystem concept based on how it was originally used by biologists. The authors emphasize that three components are essential elements of digital business ecosystems, namely *Value Exchange* (innovation, information, products/services), *Resources* (digital and non-digital) and *Actors* (organizations, individuals, societies). This definition can easily be adapted and specialized to data-driven service ecosystems in the automotive domain (cf. Figure 62, the two leftmost ecosystems are depicted according to Nischak et al. (2017)).

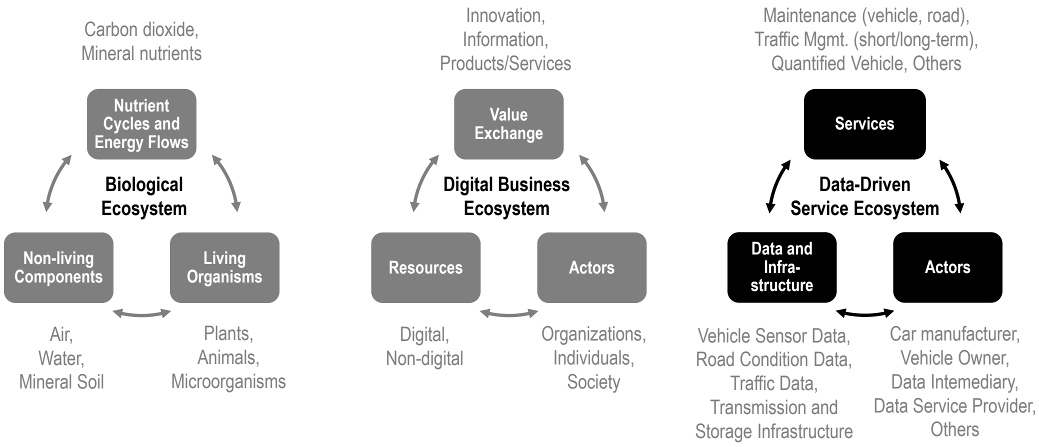


**Figure 62 Biological, Digital Business and Data-Driven Service Ecosystem. (Kaiser et al., 2019b)**

Similar to a Digital Business Ecosystem, a *Data-driven Service Ecosystem* in the automotive domain contains *Actors* (e.g. car manufacturers, vehicle data service providers, etc.). These actors have access to resources which, in the case of a data-driven ecosystem in the automotive domain, are *Data and Infrastructure* for generating, transmitting and storing data (e.g. vehicle sensor data, road condition data, etc.). With these resources, the actors participate in value exchange by providing or consuming *Services* (e.g. data-driven services for vehicle maintenance, short-term traffic management, etc.).

In general, the concept of service ecosystems is often defined in literature as “relatively self-contained, self-adjusting systems of resource-integrating actors connected by shared institutional logics and mutual value creation through service exchange” (Akaka et al., 2012; Lusch and Nambisan, 2015). Despite these works on service ecosystems that also involve data-driven services, research concerning data-driven service ecosystems in the automotive domain is still scarce: The query

```
(automotive OR vehicle OR car) AND (digital service
   ecosystem OR data-driven service ecosystem)
```

executed on Google Scholar produced 22 results (Nov. 2018), but after looking at the title and abstract, the authors concluded that no research work sufficiently addressed service ecosystems in the automotive domain. Consequently, and to the best of the authors' knowledge, no research work explicitly addresses data-driven service ecosystems in the automotive domain. However, further related research can be found for smaller sub-systems of such an ecosystem. First, the combination of *Data and Infrastructure* with *Services* is subject of many research works that analyze how vehicle sensor data can be aggregated and combined with other data to create meaningful information services along a big data value chain for vehicle data (Kaiser et al., 2018a; Kaiser et al., 2018b). Second, the interplay of *Services* and *Actors* has been extensively investigated in the fields of Value Networks (Pagani 2013) and, more recently, Service Science, where the theory of the Service-Dominant Logic (S-D Logic) is of central importance (Vargo and Lusch, 2004). The theory implies that almost any product can be interpreted as a service, meaning that the physical component is no longer the most important aspect.

To sum up, the identified research streams emphasize either technical or business aspects but do not investigate the interrelationships and interdependencies between data, services and actors in an integrated way. Furthermore, a systematic account of which actors are involved in which service and which data they can access is still missing. The authors close this gap by identifying relevant actors and their relationships, data flows and services. In doing so, the authors apply the perspective of automotive data-driven service ecosystems.

## 14.2 Data-driven Service Ecosystems in the Automotive Domain

### 14.2.1 Data Collection and Analysis

After analyzing the relevant literature on ecosystems, the authors designed a model for data-driven service ecosystems in the automotive domain, which the authors used as the main input for an expert interview study involving eleven experts in data-driven services from large industries (e.g. automotive manufacturers), small and medium enterprises (e.g. data-driven startups) and research entities. Figure 63 lists background information on experts, e.g. expert 1 is involved in the standardization of Vehicle-to-Infrastructure (V2I) communication and acts as a representative in international committees in the automotive domain. Expert 2 leads a computer science degree program focusing on the automotive sector and deals with data-driven services. Expert 7 leads a project to develop and launch a data marketplace, where the provision of vehicle data for data-driven services is a major issue, while expert 9 leads an automotive engineering company that develops hardware and software components for vehicle manufacturers dealing with vehicle sensor data. In summary, each individual expert has a strong background in data-driven services in the automotive domain.

The authors had several meetings with these experts, either in person or virtually. The authors asked each of them for qualitative input on relevant actors and their influences, and on value flows between actors, and for feedback on the originally developed graphical ecosystem model. Semi-structured interviews, each lasting about 90 minutes, started with a set of questions about their background, i.e. their general professional experience and expertise area. The experts were asked to provide a comprehensive description of one data-driven service in the automotive domain they know of. Finally, they were requested to think five years into the future and briefly describe the changes they see. All interviews were transcribed, apart from three cases where audio recording was not allowed, which is common in the automotive sector for confidentiality reasons.

| Expert No. | Type of Actor | Personal Expertise w.r.t. data-driven services in the automotive domain | Work Exp. [years] | Expertise Level [1-5] |
|---|---|---|---|---|
| 1 | Public Authority | representative in international committees | 22 | NC |
| 2 | Research Organization | leader of computer science degree program | 24 | 2-3 |
| 3 | Research Organization | researcher involved in vehicle data analytics projects | 8 | 2 |
| 4 | Research Organization | researcher involved in projects with vehicle manufact. | 4 | 3 |
| 5 | Vehicle Data Service Provider | senior manager of data-driven service provision | 13 | 2 |
| 6 | Public Authority | representative in international committees | 25 | NC |
| 7 | Data Marketplace Provider | leader of a data marketplace development | 20 | 2 |
| 8 | Vehicle Data Service Developer | consultant involved in projects with public authorities | 6 | 3 |
| 9 | Engineering Service Provider | owner of engineering service company | 23 | 1-2 |
| 10 | Vehicle Data Service Provider | developer involved in data-driven service provision | 4 | 2 |
| 11 | Vehicle Manufacturer | head of a data service department | 18 | 1 |

**Figure 63 Background information on the expert interview study participants. (Kaiser et al., 2019b)**

The results of the expert interview study were graphical sketches of eleven models of actors and value flows in a data-driven service ecosystem and qualitative statements on cooperation within the ecosystem. While some experts even outlined their own views on paper, others provided stories about such an ecosystem. Some experts mentioned very concrete, data-driven services and how they create value, while others remained on a higher level. Further results were related to a graphical model of the Data-driven Service Ecosystem designed by us and shown to the experts, which gave feedback.

The data generated in the expert interview study were analyzed by two different researchers in a one-day workshop. The input for this workshop were the individual experts' views on the ecosystem. In order to bring together eleven individual views of experts into a unified model, the researchers had to put themselves in the shoes of the experts and try to understand their specific points of view. This also included the coordination of actors (if possible) and the search for neutralized names for actors. At first, both researchers individually attempted to generalize eleven models into one model for data-driven service ecosystems. They then discussed their results with each other before defining the generalized model shown in Figure 65. They assessed whether all eleven expert opinions were adequately reflected in their generalized model, especially by assigning actors and services to the model referred to by the experts. A suitable graphical ecosystem model would make it possible to map common data-driven services, identify typical relationships and data flows between stakeholders or even identify services that have too many service providers or are monopolies. The main result of this design process is a graphical ecosystem model consisting of two interconnected parts to increase readability and usefulness: The first part is a graphical model that relates to the perspective of data acquisition, sharing, and provision, which is similar for all services. The second part focuses on service development, provision, and consumption: The service (something of value) offered by a provider can vary per consumer, making graphical visualization difficult. Each actor's relevance with regard to offering or consuming a data-driven service, as illustrated in the second part of the model, was derived from study statements that were aligned with the literature.

### 14.2.2 A Model for Data-driven Service Ecosystems

The first result of this chapter, as shown in Figure 64, is a summary of the relevant actors in the ecosystem, derived from the individual views of experts on ecosystems. In order to obtain a holistic ecosystem model with actors from different categories, the authors have not limited the type of actor to a single category (e.g. service providers). Instead, the authors included all the actors mentioned by the experts, while in some cases, two or more actors were aligned to one actor, which was then represented by a more general term. For example, cloud platform provider, database hosting provider, and web space provider were grouped under the term *backend service provider*; data marketplace, private data platform, public data platform, automotive data platform were summarized under the term *data marketplace* *provider*; decision taker, EU/EC, and national regulation were summarized under the term *public authority*. In total, the 17 resulting and neutralized actors named by the experts are shown in Figure 64. For example, expert #1 mentioned six actors including vehicle data service provider, vehicle manufacturer, etc., as shown by an 'x' in the matrix. The two right-most columns indicate whether the actor is included in the sub-parts of the high-level model as introduced later in Figure 65-Figure 67.

A definition was created for all actors to provide a solid basis for discussion. As an example, the *backend service provider*, e.g. aws.amazon.com, is defined as *a company offering IT infrastructure services (e.g. servers, a cloud platform, databases management and hosting, load balancing, etc.)*, while the *data intermediary*, e.g. the company HERE owned by the vehicle manufacturers Audi, BMW and Daimler, is defined as *a data aggregator with special relationships to vehicle manufacturers, uses data from various sources, aggregates it and provides it to contractual partners.*

| Expert Number / Actor Name | #1 | #2 | #3 | #4 | #5 | #6 | #7 | #8 | #9 | #10 | #11 | Mentions | Participant in Data Ecosystem | Participant in Service Ecosystem |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Vehicle Data Service Provider | x | x | | x | x | x | | x | x | x | x | 9 | yes | yes |
| Vehicle Manufacturer | x | x | x | x | x | x | x | | x | | x | 9 | yes | yes |
| Marketplace Provider | x | x | | x | | x | x | x | x | x | | 8 | yes | |
| Road User (vehicle driver, biker, public transport user) | x | x | x | x | x | x | | x | | | x | 8 | yes | yes |
| Public Authority | | | x | | x | x | | x | x | x | | 6 | | yes |
| Technical Infrastructure Provider (e.g. telecommunication provider) | x | | x | | x | | | x | x | | | 5 | | |
| Backend Service Provider | | | x | x | x | | | | x | x | | 5 | | |
| Vehicle Hardware/Software Supplier, (external) Gateway Provider | | | x | x | x | | | | x | x | | 5 | yes | yes |
| Data Intermediary | | | x | | | x | | | x | | | 3 | yes | yes |
| Fleet Operator | | | x | | x | | | | | x | | 3 | | yes |
| Road Operator | x | x | x | | | | | | | | | 3 | | yes |
| Vehicle Owner | | x | | | x | x | | | | | | 3 | | yes |
| Workshop Operator | | x | | | x | x | | | | | | 3 | | yes |
| Traffic Manager / Operator | | x | x | | | | | | | | | 2 | | yes |
| Automobile Club | | x | | | | | | | | | | 1 | | yes |
| Research Organization | | | | | | | | x | | | | 1 | | yes |
| Insurance Company | | | | | | x | | | | | | 1 | | yes |

**Figure 64 Ranking of actors in data-driven service ecosystem based on expert interviews. (Kaiser et al., 2019b)**

In a next step, a high-level, generalized model of actors in a data-driven service ecosystem in the automotive domain was designed based on the contributions of the eleven experts, shown in Figure 65. This model consists of three main elements, the *Data Ecosystem,* which encapsulates actors and data flows (focusing on sensor data from the vehicle only, therefore no data input from drivers such as age or sex is taken into account), the *Service Ecosystem*, which encapsulates actors, service provision, and consumption activities, and finally the two major external influencing factors as indicated by the experts, the *public authority* with all regulations and the *Technical Infrastructure Providers,* which form the baseline for data-driven services.

To shed more light on the perspective of the *Data Ecosystem* (the light blue area in Figure 65), the more detailed model in Figure 66 outlines actors and data flows between actors. It is assumed that every data flow in one direction generates a backflow in the form of money or another tangible or intangible asset to ultimately maintain a business relationship. For the sake of simplicity, only data flows are visualized in the model. The basic idea is that a set of data transformations on different sources and a provision mechanism across different actors are required to finally enable a service development cycles process.

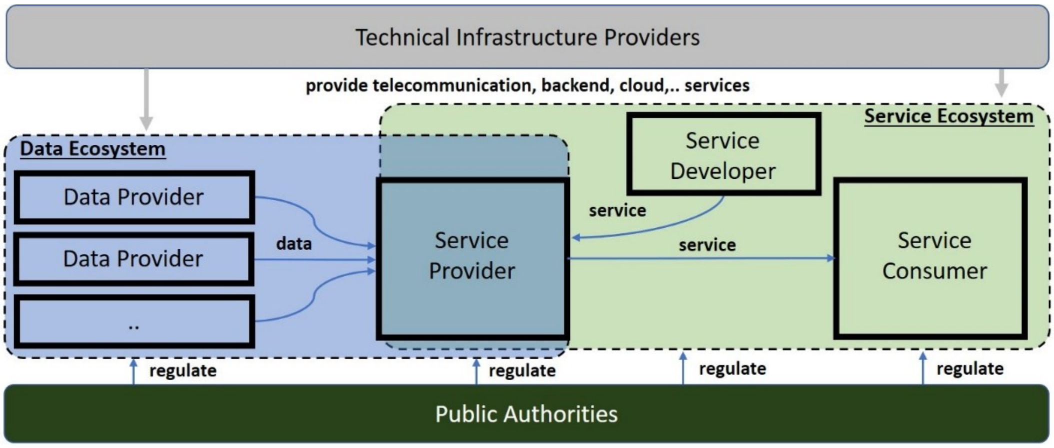


**Figure 65** **High-level model of data-driven service ecosystems in the automotive domain. (Kaiser et al., 2019b)**

First, the vehicle data are collected from a vehicle controlled by a driver (actor: road user) before they are transferred, either directly to a vehicle manufacturer, or via an external gateway provider to a marketplace provider or to a vehicle data service provider. The vehicle manufacturer can transfer all data (or selected data points) to a data intermediary, a marketplace provider or directly to a provider of a data-driven service. The marketplace provider can transfer data to a data intermediary, a service developer, or even a vehicle data service provider. A service provider can use five different data sources to establish a data-driven service (#1 to #5 in Figure 66) from a data intermediary, a vehicle manufacturer, a marketplace provider, a vehicle data service provider (e.g. Automatic), or an external data source provider (e.g. a weather data provider). The graph shows that if there is no data interface in the vehicle for external gateway providers to use (e.g. no OBD interface), the vehicle manufacturer will have a dominant position in the ecosystem and thus data acquisition is contested. This is supported by the expert statements. Expert 6, for example, says that "*OBD dongles are for vehicle manufacturers like a red rag for a bull*". Expert 9 states that *"startups are more like groundbreaking shooting stars that will disappear when big players come onto the market"*, while expert 3 argues that *"all international public authorities should unite at* *least on European level to create a reasonable counterweight to vehicle manufacturers and data-intermediaries”*.

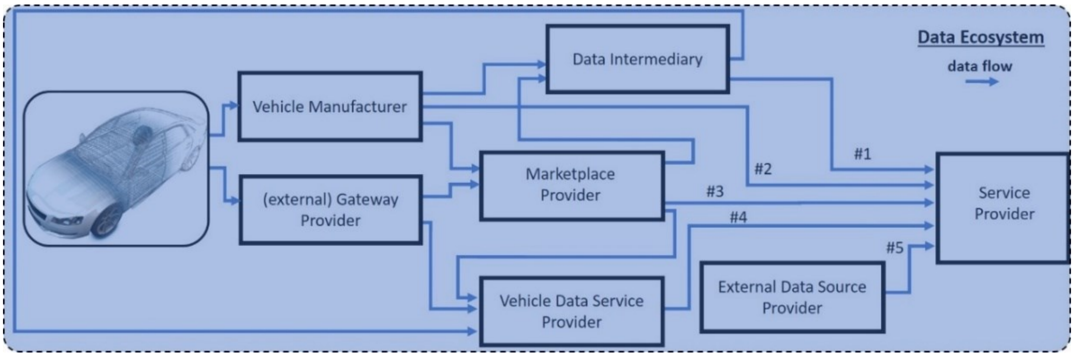


**Figure 66 Detailed model of the underlying data ecosystem with actors and data flows. (Kaiser et al., 2019b)**

However, vehicle manufacturers have the feeling that they are not in the best position, as illustrated by expert 11, who complains about the “*issue of long development cycles*”, meaning that “*if it is not foreseen/enabled already today in the development of the vehicle to share specific data, then it will not be possible until about 2025 to have it present in a vehicle on the street*”,. Expert 1 adds that there is “*sensor raw data in the vehicle, which will never leave the vehicle, not even to be transferred to the vehicle manufacturer*”.

The *Service Ecosystem* part (the light green area in Figure 65) highlights five services that were mentioned by more than one expert during the expert interview study and were also included in the related literature that was analyzed (e.g. Dremel et al. 2017b; Brandt 2013), namely maintenance for vehicles and roads, short term traffic management, quantified vehicle information service, and automotive product improvement. Expert 3, for instance, states that *“traffic planning and management, road operation companies and infrastructure providers will realize the value of vehicle data sooner or later”*. Other services mentioned only once are excluded from the next figure. These additional services include long-term traffic management, autonomous driving functionalities, tutoring, safety and collision avoidance, vehicle monitoring (e.g. in critical areas), and navigation. It should be noted that interview participants were asked to name and describe a prominent service they knew, rather than listing all the service ideas that came to mind. The authors did this because capturing the full range of services was less important since the main objective was to identify the actors and connections between actors in data-driven services, which is a prerequisite for understanding the underlying ecosystem in the automotive domain.

The first column of Figure 67 shows possible relationships between actors who provide a service (left part) or consume a service (right part) and can be read as follows: Data-driven services that are offered, such as vehicle maintenance, are displayed in the table header. If an actor decides to offer a service, the IDs of possible service consumers are listed in the table cell: For example, actor *ID 1 automobile club* can *offer a data-driven service for vehicle maintenance* (e.g. statistics on known problems of a certain vehicle type) *to actor IDs 4, 5, 6, (7), and 9*. If there is no actor ID listed in the table cell, the respective actor may not be able to offer this service. Actors, e.g. automobile clubs and vehicle manufacturers may address different customers, which is why the cells of their table rows contain different actor IDs. If it is not clear whether an actor is a targeted consumer, the actor ID is placed in brackets, e.g. “(7)”. The shading of the color of the table cell indicates the relevance of an actor for providing a service (each ID is counted as a full point, each ID in brackets as 0.2 points for calculating the color tone) and shows actors who play a major or minor role per service in the ecosystem. Based on results on the left side, the information about which actor can consume a service is presented in the right part of Figure 67 (three columns are cut off due to space constraints), revealing the targeted consumers of a specific service. For example, a fleet operator (ID 5) can consume a data-driven vehicle maintenance service from actors with the ID 1, 2, 3, 7, (10), and (11). The darker the background color of the table cells, the more relevant an actor is as a consumer of services offered by other actors. All actors in the data ecosystem part can also be actors in the service ecosystem part as service consumers, e.g. a vehicle manufacturer, a data intermediary, an (external) gateway provider, etc. can use services provided by external service providers as input for their own purposes.

| ID | Name (Actor / data-driven Service) | offers data-driven service for vehicle maintenance | offers data-driven service for road maintenance | offers data-driven service for (short term) traffic management | offers data-driven quantified vehicle information service | offers data-driven service for automotive product improvement |
|---|---|---|---|---|---|---|
| | | to ID | to ID | to ID | to ID | to ID |
| 1 | Automobile Club | 4,5,6,(7),9, | | | 4,5,6,(7), 11,12,13 | |
| 2 | Data Intermediary | (1),3,5,(6), (7),(9),10,11 | (1),3,(5),(6), (7),8,9,11 | 3,(5),(7), (8),9,11 | 6,8,9,11, (13),(14) | |
| 3 | Vehicle Data Service Provider | 1,4,5,(6),7,(9), (10),(11),12,14 | 1,4,5,6,7,8, 9,(10),11,12 | 1,4,5,(6),(7), 8,9,11,12 | (1),4,5,(6),7,(8-1 0),11,12,(13-14) | 4,5,10,11, 12,(13) |
| 4 | Road User (vehicle driver, biker, public transport user) | | | | | |
| 5 | Fleet Operator | | | (4),(12) | 3,4,7,12,(13) | |
| 6 | Public Authority | | | | (1),(4),(7), 8,9,(12) | |
| 7 | Research Organization | (1),(3-4),5,6, (9),10,11,(13) | (1),(3-4),8,(9), (10),11,(12) | 1,(3),(4),5,6, 8,9,(11),(12) | (1),(3-6), (8-13) | 3,4,5,8,9, 10,11,12 |
| 8 | Road Operator | | 1,3,4,5,6,7, 9,(10),11,12 | (1),(3),4,5,(6), (7),9,11,12 | (1),3,4,(5), 6,(7),9,12 | (4),9,10,11, (12),(13) |
| 9 | Traffic Manager / Operator | | | (1),3,4,5,6, (7),8,(11),12 | (1),3,4,5,(6), (7),8,11,12 | |
| 10 | Vehicle Hardware/Software/ external Gateway Supplier | (1),(5),(7),11, (12),13,(14) | | | | 11 |
| 11 | Vehicle Manufacturer | (1),(4),(5),(7),(9), 10,(12),13,(14) | 4,5,6,(8), (9),(10),12 | 2,(3),4,(5),12 | 4,5,(7),12, (13),(14) | (1),(3),4,5,(6-9),10,(12),(13) |
| 12 | Vehicle Owner | | | | | |
| 13 | Workshop Operator | | | | | |
| 14 | Insurance Company | | | | (1),3,5,(6-7), (9),12 | |

Legend: low — actor relevance in offering the service — high

| ID | Name (Actor / data-driven Service) | consumes data-driven service for vehicle maintenance | consumes data-driven service for road maintenance | consumes d driven service f (short ter traffic managem |
|---|---|---|---|---|
| | | from ID | from ID | from ID |
| 1 | Automobile Club | (2),3,(7), (10),(11) | (2),3,(7),8 | 3,7,(8),( |
| 2 | Data Intermediary | | | 11 |
| 3 | Vehicle Data Service Provider | 2,(7) | 2,(7),8 | 2,(7),(8 9,(11) |
| 4 | Road User (vehicle driver, biker, public transport user) | 1,3,(7),(11) | 3,(7),8,11 | 3,(5),(7 8,9,11 |
| 5 | Fleet Operator | 1,2,3,7, (10),(11) | (2),3,8,11 | (2),3,7, 8,9,(11 |
| 6 | Public Authority | 1,(2),(3),7 | (2),3,8,11 | (3),7,(8), |
| 7 | Research Organization | (1),(2),3, (10),(11) | (2),3,8 | (2),(3),(8), |
| 8 | Road Operator | | 2,3,7,(11) | (2),3,7,9 |
| 9 | Traffic Manager / Operator | 1,(2),(3), (7),(11) | (2),3,(7),8,(11) | 2,3,7,8 |
| 10 | Vehicle Hardware/Software/ external Gateway Supplier | 2,(3),7,11 | (3),(7),(8),(11) | |
| 11 | Vehicle Manufacturer | 1,2,(3),7,10 | 2,3,7,8 | 2,3,(7),8,( |
| 12 | Vehicle Owner | 1,3,(10),(11) | 3,(7),8,11 | 3,(5),(7), 8,9,11 |
| 13 | Workshop Operator | 1,(7),10,11 | | |
| 14 | Insurance Company | 3,(10-11) | | |

Legend: low — actor relevance in consuming the service — high

**Figure 67 Selected services and actors offering/consuming them (figure truncated). (Kaiser et al., 2019b)**

## 14.3 Discussion

The automotive industry is undergoing a digital transformation, from selling vehicles and the provision of related services to the provision of new data-driven services. However, new actors entering the automotive sector are colliding with established players, creating new ecosystems. Expert 8 argues that *"there will be a shift towards a sharing economy"* with a broad range of new players and mobility services challenging the traditional business models of car manufacturers. However, the question raised by expert 10 about *"how many service providers will be able to find services monetizable to all"*, will remain unanswered.

Our results show that the provision of services (left part of Figure 67) is very competitive in most cases, e.g. a data-driven service for vehicle maintenance is provided by four different actors (automobile clubs, data intermediaries, vehicle data service providers, and research organizations) which in some cases indicates an overlapping consumer audience (e.g. actor 5 – fleet operators). Figure 67 allows interested parties to participate in the ecosystem to explore service issues and make better-informed decisions about service development and investments. An improved version of this figure, including additional (and, ideally, all possible) services, should be regularly updated by public authorities to detect market anomalies such as monopolies. Surprisingly, vehicle manufacturers and data intermediaries are key actors with regard to providing vehicle data, but not the main actors with regard to data-driven service provision. According to the expert's inputs, this role is covered best by vehicle data service providers. In a more detailed view, actors such as road operators, traffic managers/operators, or automobile clubs can also play an important role in offering data-driven services. With regard to consuming data-driven services (see right part of Figure 67) road users, fleet operators, public authorities, road operators, traffic managers/operators, vehicle manufacturers and vehicle owners are the target group and thus relevant actors. Workshops only consume certain services, while insurance companies and engineering service providers are not regarded as consumers of any of the five examined data-driven services.

By combining the results of a literature review with data from eleven expert interviews, the authors have gained a basic understanding of data-driven service ecosystems. The main result of this chapter is a model of such ecosystems including stakeholders (actors), data and infrastructure (resources) and data-driven services (value exchange) that helps to understand emerging business relationships. The model consists of two interconnected parts – the data ecosystem and the service ecosystem. The data ecosystem shows that the vehicle manufacturer is a critical actor with regard to data collection as vehicle measurements can only be retrieved from an external gateway provider or a vehicle manufacturer. If a vehicle manufacturer does not grant direct access to the vehicle data via an interface, data-driven service developers must negotiate directly with the vehicle manufacturer to gain access. This allows manufacturers to control data-driven services without regulation. *"Granting others access to vehicle data via the OBD interface is a safety risk",* as expert 11 indicates.

Controversial statements about the cooperation dilemma between vehicle manufacturers and vehicle data service providers were made. Expert 5, CEO of a vehicle data service provider, states that "*vehicle manufacturers are currently not doing the services on their own, they look for service providers, as they don't have the hardware and the resources*", while expert 4, a researcher cooperating with vehicle manufacturers, argues that "*vehicle manufacturers find it difficult to give third parties access to vehicle sensors on deeper levels*", meaning that "*vehicle manufacturers will highly regulate the ecosystem. [The ecosystem] will not be as open as other ecosystems we know from the B2C area. Partners will be chosen carefully preferring trusted long-term partners".* According to expert 9, it is challenging for third parties to obtain data from vehicle manufacturers, as *"we have been discussing with OEMs [vehicle manufacturers] to get access to data for years, not yet a success, although they seem interested"*. In the service ecosystem part of the model, the authors indicate which services can be provided or consumed by which actors. The authors focus on five main services only, which is another limitation of this research.

## 14.4 Conclusion and Outlook

In this chapter, relevant actors, relationships, and data flows in a data-driven service ecosystem in the automotive domain are presented. A relevant aspect of this research is how such ecosystems emerge and how vehicle data enable new services. To improve understanding, the authors reviewed the literature on the ecosystem concept and enriched the results with a subsequent expert interview study. The fact that the authors only interviewed eleven experts is a limitation of this research. In order to investigate ecosystem mechanisms in the automotive domain, further studies are needed to determine the ability of two or more actors to cooperate when providing a service, e.g. how likely is it that a vehicle manufacturer (in collaboration with other stakeholders) will provide a data-driven vehicle maintenance service to a fleet operator? In line with further literature analysis, the authors want to show how external factors such as standardization, data protection, autonomous driving, or taxation can influence the ecosystem and the willingness of an actor to participate.

# 15. Conceptualizing Value Creation in Data-driven Services: The Case of Vehicle Data [22]

Summary and Author Contribution

**Analysis of Data-driven Service Ecosystems**
**(Paper 2/2)**

The digitalization of the automotive industry brings fundamental changes to how value is created and by whom. As part of this transformation, the creation of data-driven services generates new value streams, thus leading to the emergence of new actors and ultimately, new market configurations. Eventually, vehicle data paves the way for new types of data-driven services. This chapter aims to provide a (multi-actor) model suitable to support academics and practitioners in the identification of the actors that will play a key role in data-driven service generation and resources involved in value creation processes. Based on interviews with eleven prominent experts of the central European automotive industry, a conceptual model that connects these key actors with value-adding data sharing processes is developed. To validate the model, it is applied to a real-life case: the design of a data-driven service for road surface quality detection. Furthermore, the model's implications to both theory and practice are discussed.

With this journal paper in the highly-ranked *Journal of Information Management* (Impact factor 8.21 in February 2021) I wanted to address the research gap that key concepts and relations regarding data-driven value creation are insufficiently explored today, as the authors observe it in the research on vehicle data-driven services. Guided by the design-science paradigm, together with the co-authors I conducted a literature review to set the context in the relevance cycle; a participatory approach to build a practice-based, conceptual model, capturing the individual views of eleven experts from the automotive domain on the ecosystem in the design cycle; and a conceptual validation by applying our model to an application case in the rigour cycle. In particular, I conducted all interviews of the eleven experts, and developed the conceptual model and validated the model together with the second author. As a result, our identification of the most relevant ecosystem actors provides the foundation for better understanding their data sharing relationships and interdependencies. The model connects actors to particular steps within the data value creation process and describes a data value chain to generate useful insights from vehicle data. Such a model might help organizations that choose to become part of the automotive ecosystem to better understand their role, relationships, and opportunities for *Data-driven Service* provision.

[22] The content of this chapter is based on (minor changes)

**Kaiser, C.**, Stocker, A., Viscusi, G., Fellmann, M. and Richter, A. (2021). Conceptualizing Value Creation in Data-driven Services: The Case of Vehicle Data. *International Journal of Information Management (IJIM), Volume 59 (2021). https://doi.org/10.1016/j.ijinfomgt.2021.102335*

The ongoing transition towards a digitalised world also affects primarily physical industries (Hanelt et al., 2015). Due to its long tradition in catering to a basic human need – mobility, the automotive domain stands out in particular (Piccinini et al., 2015). Traditionally, businesses within the automotive domain were geared towards offering goods (e.g. selling manufactured vehicles as the main product) and product-related services (e.g. selling spare parts and conducting maintenance work). However, digitalisation has led the automotive industry to think differently, as vehicles become increasingly connected and capture a lot of data about themselves and their environment (Swan, 2013; Swan, 2015). This captured vehicle data eventually paves the way for new types of data-driven services (Pillmann et al., 2017; De Winter et al., 2019; Bridgelall et al., 2018; Pütz et al., 2019).

Consequently, vehicles are increasingly becoming part of an automotive ecosystem that includes not only drivers and passengers but also other road users, vehicle manufacturers or service developers. Connected vehicles enable the possibility to develop data-driven services such as remote vehicle diagnostics or interactive trip analytics (Papatheocharous et al., 2018; Kuschel, 2008). Thus, the digital transformation offers new players outside the automotive sector the opportunity to enter this traditionally closed ecosystem (Athanasopoulou et al., 2016). Among those, we find major companies like Tesla, Google, or Apple (Wittmann, 2017) and start-ups like vin.li and Zendrive.com who create data-driven services related to digital asset tracking, vehicle health or driving safety (Stocker et al., 2017a). Yet, it remains a challenge for those start-ups to translate their technical innovations into commercially successful product offerings.

Despite these disruptive changes caused by digitalisation, the core industrial product of the automotive industry, the vehicle, cannot be digitised entirely (Piccinini et al., 2015). Instead, it will be complemented by both traditional and data-driven services (Kaiser et al., 2019a). Declining revenues from vehicle sales can be compensated by additional income from the monetarization of vehicle data (Bertoncello et al., 2016; Seiberth and Gründinger, 2018; Davenport et al., 2020). However, it remains no less of a challenge for incumbent companies in the automotive industry to fully embrace such digital innovation (Svahn et al., 2017).

*Data-driven services* are services that support customers' decision-making processes by providing data and analytics to create value for the customer (Schüritz et al., 2019). The provision of data-driven services is often, but not necessarily, accompanied by a physical product equipped with sensors for digital connection to other products and information systems-IS (Beverungen et al., 2018; Tomiyama et al., 2019). Although this digital transformation in the automotive domain is underway (Kuhnert et al., 2018; McKinsey, 2016), little is known about the most relevant actors and their data sharing relationships to deliver value-added services based on exploiting vehicle data. Especially in the advent of big data, it is even more important than ever to understand the characteristics of data-based or data-driven value creation (Lim et al., 2018; Schüritz et al., 2019).

We put our focus on the automotive domain as their industrial-age core product cannot be digitized completely (Piccini et al., 2015). Furthermore, automotive is one of the most important industries related to non-digital artefacts (i.e. vehicles) (Henfridsson et al., 2009). It is, however, worth noting that the automotive sector has begun to experiment with vehicle telematics solutions and connected car initiatives since a few years (Svahn et al., 2017).

The primary goal of our research is to investigate ways through which (small and big[23]) vehicle data can spawn new data-driven services and to provide a framework to structure and evaluate vehicle data-driven value creation. We argue that improved knowledge about key actors and their data sharing relationships will contribute to a better understanding of vehicle data-driven value creation. Accordingly, a fundamental starting point for our research is to map those actors that will have a crucial role in data sharing and then design how data sharing relationships can connect them. Thus, this chapter addresses the following three research questions:

- *Which actors play a key role in vehicle data-driven service generation?*
- *How do data sharing relationships connect those actors to enable value creation?*
- *How can a conceptual model illustrate both actors and their data sharing relationships?*

The identification of the most relevant ecosystem actors provides the foundation for better understanding their data sharing relationships and interdependencies, allowing us to design a conceptual model of data-driven value creation. Conceptual models are abstract represen-

23 Although big data and big data analytics are definitely important, we would like to emphasize that we do not focus on research on the adoption on big data analytics. Many vehicle data-driven services are based on "small data" (using only a few data points of a single signal): For instance, services that can detect safety critical situations inside a vehicle and forward this information to operational organizations like emergency services do not rely on big data analytics. In many cases, a few data points from a few signals are sufficient to generate a data-driven service with high added value as for instance mentioned by the interviewed data marketplace provider.

tations of some subject matter, which serve to promote communication and common understanding between stakeholders, thereby improving the prospects for successful information system development and use (Wand and Weber, 1993). Conceptual models are mostly of a graphic nature and usually contain a visual arrangement of modelling constructs in the form of graphical symbols and text (Bera et al., 2019). Besides supporting communication, they contribute to a better understanding of a particular domain and provide input for the information systems design process (Wand and Weber, 2002).

The theoretical gap addressed here is the lack of conceptual models that can unravel the underlying value chain (actors and data sharing relationships) when establishing vehicle data-driven services. We thereby address the calls of researchers (Parvinen et al., 2020) to closely examine data-driven value creation and ecosystems (Wamba et al., 2015). Our designed conceptual model aims to link actors with specific steps within the data value chain. Thereby, it seeks to help organisations that choose to become part of the automotive ecosystem to better understand their role, relationships, and opportunities for data-driven service provision better. Thus, our model supports the design of vehicle data-driven services by introducing the most relevant actors, and data flows, ultimately leading to data-driven value creation. As of now, the model preserves the different perspectives of key actors while addressing the research gap that key concepts and relations regarding data-driven value creation are nowadays insufficiently explored, as we observe it in the research on vehicle data-driven services. As to these arguments, they are grounded in our research within a large-scale research project [Omitted for blind review] funded by the European Commission in the Horizon 2020 framework programme involving 19 key partners from 11 European countries from the automotive, big data, cloud and high-performance computing worlds aiming at better exploitation of big data.

The remainder of this chapter is structured as follows: Section 15.1 presents the framework that guided our research approach and chapter structure. Section 15.2 embraces the theoretical foundations of this chapter. After this, in Section 15.3, we elaborate on the data collection including results from eleven expert interviews and their sketching activities. These views are unified and serve as the basis for our conceptual model presented in Section 15.4. We provide the results of the final evaluation of our model in Section 15.5, and discuss our findings in Section 15.6 before we summarise and conclude the chapter in Section 15.7.

## 15.1 Research Framework and Chapter Structure

We address the lack of conceptual models that can unravel the underlying vehicle data value chain (actors and data sharing relationships) in establishing data-driven services. Our research framework is guided by the design-science paradigm (Hevner et al., 2004), with its three research cycles (Hevner, 2007): *relevance cycle*, *design cycle*, and *rigor cycle* (Figure 68). Design-science extends "the boundaries of human and organizational capabilities by creating new and innovative artifacts" (Hevner et al., 2004). In our case, the innovative artifact is the conceptual model for *value creation in vehicle data-driven services.* This research framework allowed us to obtain the different perspectives of key stakeholders (researchers, users, clients, sponsors, and practitioners) while studying complex problems.

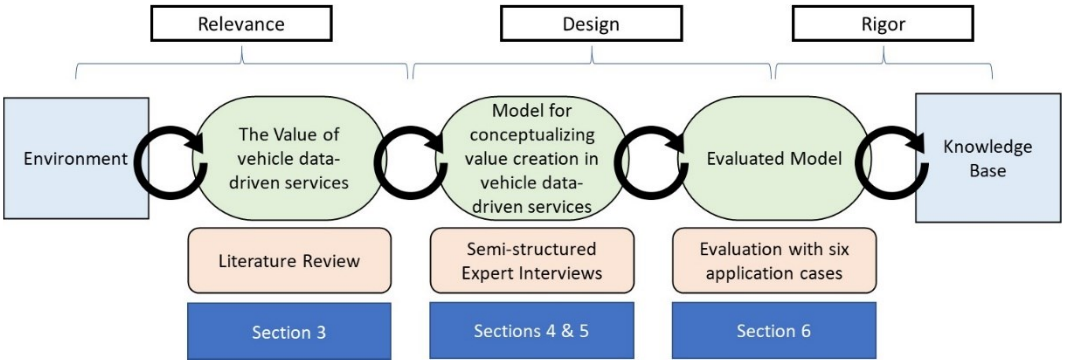


**Figure 68 Research Approach and Chapter Structure. (Kaiser et al., 2021)**

As part of the relevance cycle *(Section 15.2),* we conducted a literature review of well-regarded scientific electronic databases extended through backward and forward search regarding our application context (value creation in vehicle data-driven services) and theoretical lens (ecosystems). Existing theory on the value of data-driven services and data value chains was used as theoretical input within the design phase.

In the design cycle, we first took a participatory approach to build a practice-based, conceptual model, capturing the individual views of eleven experts from the automotive domain on value creation in vehicle data driven services. We complemented our interviewing approach with simple graphical design activities, letting experts draw sketches on key actors and their data sharing relationships *(Section 15.3)*. We consolidated the individual expert views in a conceptual model of value creation in vehicle data-driven services, applied conceptual modelling (Wand and Weber, 2002) inspired by the concept of data value chains (e.g. Latif et al., 2009; Curry, 2016; Faroukhi et al., 2020) and presented our artifact: a unified conceptual model for the data value creation process consisting of three parts, (i) actors involved, (ii) key ecosystem actors and (iii) data sharing relationships between the actors.

We established proof-of concept before we proceeded to evaluating our model with real-life cases (*Sections 15.3 and 15.4)*.

In the rigour cycle (*Sections 15.4 and 15.5*), we performed conceptual model evaluations by applying our model to in total six real-life application cases enabled by vehicle data and established its proof-of-value. After each design, we conducted an evaluation of the model that resulted in model revision. This resulted either in a change in model actors, a change in data-sharing relationships, or a change in both, while the general structural design of the model remained unchanged. Our chapter only includes the sixth evaluation of our model within a real-world case, the development of a data-driven service for road surface quality detection to underpin its practical applicability.

## 15.2 Theoretical Foundations

This This section places our research in the context of the relevant existing literature. First, we illustrate the concept of data-driven services in the automotive domain and the rationale behind it. Second, we take a look at the literature on ecosystems which we use as a theoretical lens for our study.

Table 23 shows how research on data-driven services is steadily increasing. In total, we identified 222 papers published since 2011 in established scientific electronic databases as AISeL, ScienceDirect, Scopus, IEEE Xplore and ACM DL (Falagas et al., 2008; Gusenbauer and Haddaway, 2020). More than 36% of these papers were published in 2019.

**Table 23 Search results (all fields) in AISeL, ScienceDirect, Scopus, IEEE and ACM (2011-2019). (Kaiser et al., 2021)**

| Year \ Database | AISeL | ScienceDirect | SCOPUS | IEEE Xplore | ACM DL |
|---|---|---|---|---|---|
| 2011-2013 | 1 | 3 | 3 | 2 | 1 |
| 2014-2016 | 8 | 15 | 7 | 5 | 4 |
| 2017-2019 | 54 | 56 | 35 | 16 | 12 |

We included articles that used the following terms: “value creation” or “value”, and “data-driven services”, or “data-based services”, and “automotive”, or “vehicle”, or “car”, or “mobility”. We added further papers on the value of data-driven services in general and on vehicle data-driven services in particular by applying backwards and forward search. In what follows, we thematically discuss the main concepts from a representative sample of 48 of the retrieved papers on the value of (vehicle) data-driven services.

### 15.2.1 Value of Data-driven Services

In the last two decades, the service sector has seen an unprecedented development, also due to the expansion of the application of Information and Communication Technologies – ICTs (Berkley and Gupta, 1994; Rai and Sambamurthy, 2006) and the subsequent digital transformation of businesses and society (Lim et al., 2018; Lusch and Nambisan, 2015; Bharadwaj et al., 2013; Spohrer and Maglio, 2008; Chesbrough and Spohrer, 2006). Among the different available definitions, we adopt here the concise summary provided by Spohrer and Maglio (2008, p. 241) defining *service* as “pay for performance in which value is coproduced by client and provider”. This is true, for example, when considering information intensive services (IIS) where the “value is created primarily via information interactions, rather than physical and interpersonal interactions, between the customer and the provider” (Lim et al., 2018, p.121). Moreover, these services rely on the data that generate the information driving the activities, making them valuable for the final customer (Kumar et al., 2013; Azkan et al., 2020; Maass et al., 2018). Consequently, value creation based on data should take into account the data value chain as well as key factors, such as for example the data, the data source, data collection, data analysis, information delivery, information on the user, the value in information use, and the provider network (Lim et al., 2018, p.122).

Taking these issues into account, the role of data and information value (Brennan et al., 2019; Attard and Brennan, 2018; Batini et al., 2018) is a central challenge in the competitive scenarios emerging from digitalisation, in particular for understanding what concerns the evaluation of the information capacity suitable to allow companies to the create and capture value by digital assets and data-driven services (Batini et al., 2018). Furthermore, according to Dedrick (2010) researchers have framed the impacts of the IT on environment as first-order (impacts of ICT hardware during the product lifecycle), second-order (impacts of ICT on other processes such as transportation or industrial production), and third-order effects (changes in lifestyles and economic structures). The latter are relevant when considering the increased use of the media's transformative potential of 'green' IS on the demand side, encouraging practices such as, e.g., carpooling and ridesharing applications coupled with the Internet of Things (Malhotra et al., 2013).

Moreover, scholars from computer science and IS have also questioned, which business models could be suitable to capture the value of data-driven services (Lim et al., 2018; Schüritz et al., 2017; Zolnowski et al., 2017; Zolnowski et al., 2016; Schüritz and Satzger, 2016).

These contributions complement the questions advanced in the field of technology management (Sorescu, 2017; Hartmann et al., 2016) about the role of data-driven services in business model innovation. Additionally, IS scholars have investigated the antecedent factors of value creation in connection with the big data analytics phenomenon (Günther et al., 2017; Mikalef et al., 2020; Surbakti et al., 2020; Wiener at al., 2020). Central questions concern the big data analytics capabilities that companies require to a) enhance organisational performance (Wamba et al., 2017; Akter et al., 2016), b) create business value (Wamba et al., 2015; Grover et al., 2018; Conboy et al., 2020), and c) enable service innovation (Lehrer et al., 2018), as well as d) which barriers may prevent their adoption.

For example, Dremel et al. (2017) discuss in a case study of AUDI how traditional manufacturing organizations can introduce big data analytics and master related organizational transformations. Dremel et al. (2020) identify big data analytics (BDA) actualization mechanisms from a revelatory case of a vehicle manufacturer. Akter et al. (2016) aim at improving the organizational performance of a company through big data analytics and proposes a hierarchical model. Grover et al. (2018) explore the success of big data analytics projects with respect to creating strategic business value, i.e., by addressing intra-organizational aspects. Lehrer et al. (2018) propose a theoretical model of big data analytics service innovation developed from multiple cases from insurance, banking, telecommunications, and e-commerce that have all implemented big data analytics. Mikalef et al. (2017) recommend that more attention should be paid to the organizational changes that big data analytics brings and how big data analytics should be adopted strategically. Mocker and Fonstad (2017) discuss AUDI's challenges towards the sharing economy and how AUDI has transformed its organizational structure, processes, and architecture. Svahn et al. (2017) address, how incumbent firms embrace digital innovation proposing the Volvo case study and identifying four concerns, but focusing on the perspective of the vehicle manufacturer, only. Wamba et al. (2015) emphasize a lack of empirical research to assess the potential of big data and provides both a literature review and case studies to present an interpretive framework to analyze the different perspectives of big data as well as a taxonomy to better understand the role of big data in value creation. Wamba et al. (2017) propose a big data analytics capability model, extend previous research by examining the direct effects of big data analytics on firm performance. Woźniak et al. (2015) introduce a practical example for big data value creation from Volvo and share the story of building a big data service for the automotive industry in a case study. These papers focus heavily on the big data analytics phenomenon and the strategic and organizational capabilities required to create value from big data analytics. They all take a single-actor (i.e., micro) perspective.

### 15.2.2 Data-driven Services based on Vehicle Data

The recent advances in computing infrastructure, including the Internet of Things (IoTs) as well as the data generation/processing capabilities in products, have boosted the development of data-driven services. However, those who generate and collect the data are not necessarily those who develop and provide data-driven services. The systematic use of the data generated in connection with vehicle use happens in practice within complex actor-networks and ecosystems. Vehicle data paves the way to enabling novel data-driven services (Stocker et al., 2017a) and with the current increase in connected vehicles this data can finally be exploited. Connected vehicles are equipped with hardware and software to connect them to the cloud, collect data from sensors (e.g., the vehicles speed, acceleration, and steering wheel angle at a certain time), and send these data to the vehicle manufacturers' servers; this allows obtaining insights on, e.g., driving pattern analysis or estimated time of arrival in the case of fleets. Thus, an ecosystem for such services emerges (Venkataram, 2019; Dhungana et al., 2016).

In general, vehicle manufacturers seek to leverage the value of the data collected through their vehicles to better meet customer needs (Stocker et al., 2017a; Kaiser et al., 2018b). According to Gissler (2015), all new passenger vehicles sold in 2025 will be connected, forcing vehicle manufacturers to define their role and determine where they can best benefit from connectivity. Volkswagen, Daimler and BMW all recently announced major investments in data-driven services like “Volkswagen We”, “Mercedes me” or “BMW CarData” (Volkswagen, 2018; Daimler, 2020; BMW, 2020). However, there are also approaches for vehicle data collection and use in data-driven services that bypass vehicle manufacturers. These are the ones pursued by tech start-ups such as dash.by, vin.li, or pace.car who bring their own solutions into vehicles (to create a gateway to sensor data) and thereby compete with the activities of vehicle manufacturers in vehicle data collection (Stocker et al., 2017a).

Also, emerging data marketplaces, such as caruso-dataplace.com, high-mobility.com or otonomo.io, provide another approach to leverage vehicle data (Pillmann et al., 2017). Data marketplaces are digital platforms on which data products are traded, acting as neutral intermediaries, and allowing others to sell their data products (Spiekermann, 2019). The aim of vehicle data marketplaces is to make available vehicle data collected by different brands of connected vehicles, vehicle manufacturers, fleet operators and other data providers to interested data-driven service developers directly or indirectly through a single point of access.

### 15.2.3 Theoretical Lens: the Ecosystem Concept

In general, an ecosystem describes the relationships and interactions between living organisms and their environment (Schulze et al., 2005; Briscoe and De Wilde, 2006). To differentiate an artificial ecosystem from a natural one, some authors add further attributes to the term to qualify it, e.g. software ecosystem, business ecosystem or digital service ecosystem (Immonen et al., 2015). However, a commonly agreed definition does not yet exist.

Considering the field of strategy as relevant for the focus of this research on the automotive industry, an early definition has been provided by Teece (2007, p. 1325), who considers an ecosystem as "*the community of organisations, institutions, and individuals that impact the enter-prise and the enterprise's customers and supplies*" including "*complementors, suppliers, regulato-ry authorities, standard-setting bodies, the judiciary, and educational and research institutions*". Focusing on modularity and coordination for different types of complementarities (in production vs. in consumption), Jacobides et al. (2018) have proposed a consolidated perspective on the ecosystem concept. They define it as "*a set of actors with varying degrees of multilateral, non-generic complementarities that are not fully hierarchically controlled*" (p. 2264). Furthermore, Adner (2016, p. 40) define an ecosystem as "*the alignment structure of the multilateral set of partners that need to interact in order for a focal value proposition to materialize*". Considering 'partners' in the automotive industry, while Original Equipment Manufacturer (OEM) traditionally exerted a strong influence on ecosystems, this configuration is currently challenged by the digitalisation characterising new breeds of quantified vehicles and new actors on the market (Stocker et al., 2017a).

Nischak et al. (2017) emphasise that three components are essential elements of digital business ecosystems: value exchange (innovation, information, products/services), resources (digital and non-digital) and actors (organisations, individuals, societies). This definition can be adapted and specialised for digital automotive ecosystems. Similar to a digital business ecosystem, a digital automotive ecosystem contains actors that in this case are original equipment manufacturers (OEM), data intermediaries or data service providers, for example. These actors have access to resources, such as data and infrastructure, for generating, transmitting and storing data. Leverag-ing these resources, the actors participate in value exchanges by providing or consuming data.

Nevertheless, research on digital automotive ecosystems is still limited. Particularly in connec-tion with vehicle data and the process of creating data-driven services, the literature repeatedly refers to data-based business ecosystems (Kitsios et al., 2017; Curry, 2016; Nachira et al., 2007). For instance, Immonen et al. (2014) outline the open data ecosystem from a business viewpoint and define ecosystem actors such as application users, data and service providers, application developers and infrastructure providers along with their role in the data-based ecosystem. Also, in many cases the authors refer more to technical ecosystems (e.g. Kolbe et al., 2017; Gerloff and Cleophas, 2017; Kuschel, 2008; Martínez de Aragón et al., 2018). In these technology-oriented perspectives, an analysis of the business relations enabled through the digitalisation of the vehicle and the feasibility of new data-driven services is largely missing.

Researchers focusing on the exploration of actors and relationships between actors have often taken a different perspective, e.g. describing automotive engineering as an automotive ecosys-tem of interacting organisations (Knauss and Damian, 2014), or presenting a strategically moti-vated approach to discover business models in traditional industries and apply them to the mo-bility sector without empirically substantiating their findings (Remane et al., 2017). Researchers have also used data from automotive investment and partnering activities to better understand the ecosystem: Riasanow et al. (2017) have used data from crunchbase.com to derive roles, de-sign the automotive value network, and discuss the model with five experts. Nischak and Hanelt (2019) have used data about alliances, joint ventures, mergers and acquisitions along with net-work visualisation techniques for a longitudinal analysis of the automotive ecosystem. Although vehicle data paves the way to ecosystem-building activities, none of the reviewed articles con-tains a focus on vehicle data-based ecosystems.

## 15.3 An Expert Perspective on the Value of Vehicle Data-driven Services

The literature review showed that actors and their data sharing relationships were only marginally considered with regard to the development of services based on the data generated by connected vehicles. Also, the majority of the reviewed contributions do not address the specifics of vehicle data-driven ecosystems, which we aim to elicit by conducting interviews with eleven automotive domain experts with an average professional experience of more than 16 years, all of them be-ing opinion leaders for the Central European market.

### 15.3.1 Data Collection

From From May 2018 to December 2018, we have conducted in-depth interviews with experts from the Central European automotive industry. More specifically, we combine two instruments, cap-turing automotive experts' general views on value creation in vehicle data-driven services by conducting semi-structured interviews, and then aiming towards gaining a deeper understanding of the data-driven value creation process through experts' graphical models of actors and data sharing relationships. Two of the authors conducted the interviews and the fieldwork, while the other three authors acted as critical and reflexive actors (Gioia et al., 2013) during the monthly online meetings for discussing the material added to the emerging corpus of interviews, memos, and archival documents.

According to findings of Scholte et al. (2015) from ecosystem research, expert-based approach-es hold the potential that experts can be asked to express their own opinions and values starting with what they find important, while in-depth (unstructured or semi-structured) interviews can be used to gain a deeper understanding on ecosystems. Interviews have been used in the past by previous related research (e.g. Riasanow et al., 2017; Beverungen et al., 2019) to conceptualise service (eco)systems. However, we argue that conducting interviews alone may not be sufficient to gain a deep understanding of the complex data-sharing relationships of identified actors. Therefore, we complemented our interviewing approach with simple graphical design activities to let experts visualise the value creation process from their perspective.

Involved experts had on average more than 16 years professional work experience (cf. Table 24) and included large industries (e.g. automotive manufacturers), small and medium enterprises (e.g. data marketplaces, suppliers, and data-driven start-ups), public authorities and automotive research organisations. Due to the reputation of the experts, it sometimes took months before an appointment was possible. Interviews lasted between 60 to 90 minutes and were divided into several parts:

- *Part 1*: We covered the experts' background, professional experience, and attitude towards using data-driven applications.
- *Part 2*: We asked them to describe vehicle data-driven services they knew and have already used to judge their experience better.
- *Part 3*: We showed experts an existing ecosystem model from the media domain built by Gordijn et al. (2006) and asked them to attempt to sketch their view on vehicle-date driven value creation, which we assumed to be a cognitively challenging task. To guide

experts, we asked them to start their personal design process by first naming relevant actors before de-signing data-sharing relationships. Finally, we asked them to describe the changes they ex-pect in the digital automotive service ecosystem over the next 5 years.

We have conducted in total eleven expert meetings, four of them face to face with experts who were using pen and paper to sketch their views (experts 2, 3, 8, and 9). The seven remaining meetings were conducted online, using a video conferencing service with screen sharing enabled. For the virtual meetings, we prepared a special online document for eco-system design in which the experts had to list the relevant actors before linking them with data sharing relationships. In total, eight experts gave their consent to have their voices recorded during the meeting, while the remaining three experts refused recording, due to strict automotive confidentiality policies.

**Table 24 Information on the background of the experts involved in the design process. (Kaiser et al., 2021)**

| Expert No. | Organisation | Expert profile | Work Exp. |
|---|---|---|---|
| 1 | Public authority | Responsible for a metadata service for accessing vehicle data | 23 yrs. |
| 2 | Automotive research | Research manager dealing with vehicle data and data-driven services | 25 yrs. |
| 3 | Automotive research | Senior data scientist involved in vehicle data analytics projects | 9 yrs. |
| 4 | Automotive research | Senior researcher involved in projects with vehicle manufactur-ers that deal with data-driven services | 5 yrs. |
| 5 | Provider of data-driven service | Senior manager of a provider of vehicle data services | 14 yrs. |
| 6 | Public authority | Representative in international committees in charge of a vehicle data provision service | 26 yrs. |
| 7 | Data marketplace provider | Leader of a national research project on data marketplaces | 21 yrs. |
| 8 | Provider of data-driven services | Senior consultant involved in development of vehicle data-driven services | 7 yrs. |
| 9 | Automotive and soft-ware engineering | Owner and managing director of an automotive engineering ser-vice company | 24 yrs. |
| 10 | Provider of data-driven services | Senior developer of vehicle data-driven services | 5 yrs. |
| 11 | Vehicle manufacturer | Head of a data-driven service department | 19 yrs. |

In the following section, we present and discuss the various ecosystem models designed by the experts, with a detailed example from expert No. 3 and a summary of all other experts.

### 15.3.2 Case Vignette

We present the output of one expert interview as a case vignette. This is a representative example to illustrate that all experts have a particular context from their field of expertise, but an excellent overview of the automotive and mobility sector in general.

Expert 3, Frank, doctor of technical mathematics, has more than 9 years of experience as a senior data scientist for an industrial research company. He was jointly responsible for the development and operation of a data-driven service based on Floating Car Data (FCD), which is used by a traffic control center to provide information for road users, traffic planners, and state governments. Therefore, Frank has a particular view on the sharing of vehicle data, which is characterized by his own working context.

Stakeholders relevant to Frank are decision-makers, infrastructure providers, vehicle manufacturers, suppliers of vehicle manufacturers, road users, data intermediaries, traffic news offices and traffic management. Frank identified several data-sharing relationships between these actors and presented them as connecting lines on the drawing board (cf. Figure 69). As a data scientist, Frank began designing the ecosystem around *'data'* that is the basic ingredient for data-driven services: *"The problem that I have is that this data is separated. Data packets go from the data intermediary to the road user, and that does not necessarily have to be the same data that the road user sends somewhere else."* During this design process Frank also starts to explain and interpret what he has achieved so far in the ecosystem model, related to different *actors*, their needs and relationships in the ecosystem: "*Infrastructure providers would like to [get data from vehicle manufacturers], but they don't get it [the data]."* Therefore, infrastructure providers, as service users, appear to be actors that would benefit greatly from data on vehicle movements and would even start collecting such data by using stational roadside units to detect passing vehicles, e.g. to measure and predict traffic flows. However, their willingness to pay other actors for vehicle data is still questionable: "*This is still in the making, that infrastructure providers really pay data intermediaries for data",* and adds, *"INRIX, TomTom, or HERE – these are the classic [data intermediary] players."*

Data intermediaries emerge as *new players* who are beginning to establish a powerful position within the ecosystem. *"These are institutions that penetrate the market from outside and deal a lot with data. They are rather atypical. What is now very immanent in this system is that someone enters the traffic data market that actually has nothing to do with it originally."* Other new players are about to enter the market for data-driven service generation and are seeking cooperation with existing players. Some actors seem to have developed their own *practices to gather data* for decision making, e.g. traffic planners are used to collect their own data manually, instead of cooperating with other players: *"...traffic planning is still a point, but now they are still outside. These would already be relevant, but they now usually* *collect the data using standard methods."* Infrastructure providers started to make their data available to traffic planning and management: *"Vehicle measuring stations on the motorways belong to the infrastructure provider who makes the data available to traffic management. [...] They would also make this data available to the decision-makers, which would be classic loop-data."* The cooperation of actors outside the closed automotive ecosystems will only slowly take shape. There is still a lack of cooperation at national and European level, which would benefit both policy makers and traffic managers. "*Decision-makers, infrastructure providers and traffic management - which is often the same institution - have to join forces and network at least at European level in order to achieve a critical mass in order to represent the interests of data intermediaries, who currently have a very high power."*

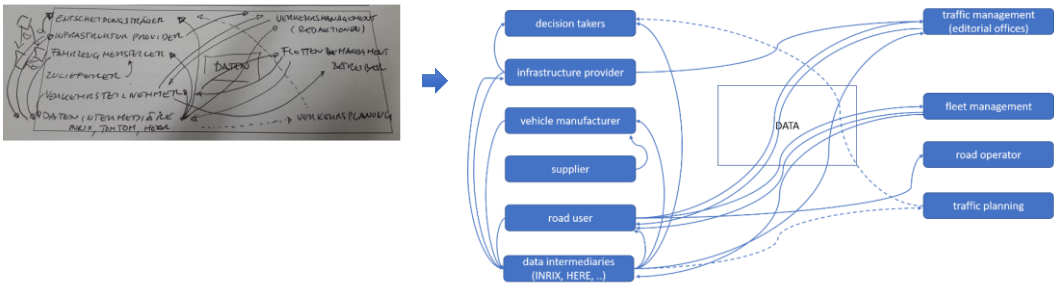


**Figure 69** **Expert 3 hand drawing and digitized vehicle data-driven service sketch. (Kaiser et al., 2021)**

In summary, from expert 3's view point, the main actors for the design of data-driven services are data intermediaries (10 relationships), road users (7 relationships), traffic management (4 relationships) and infrastructure providers (4 relationships). Among decision makers, traffic management, road operator and traffic planning, four actors have been identified who are relat-ed to or usually funded by national authorities reflecting the research background of the expert.

### 15.3.3 Summary of the Individual Design Processes of the Remaining Experts

This sub-section summarises the results of the individual design activities of the remaining experts. Since their sketches are space-consuming, we present only the sketches of the experts No. 5 and 6. We are aware that each expert argues from his or her own perspective, also depending on the organization in which the expert is employed, so there are discrepancies in the interview statements. The aim of the empirical data collection, however, was to gain an overview as complete as possible of the actors and data exchange relationships.

*Expert 1* was responsible for a metadata service to access vehicle sensor data and identified the vehicle (driver) as the main actor transmitting generated vehicle data via a telecommunication provider to either the OEM, a private- or a public data platform provider. A metadata provider, an actor in which he is personally involved, could provide the interface to a service provider or road operator to search and automate access to vehicle data.

*Expert 2* understands the ecosystem as a network of relationships between actors around a data marketplace. Data collection is mentioned several times and seems to be an unresolved problem, as expert 2 is uncertain who is currently deciding on data sharing: The expert assumes that data could be shared with a service provider without the knowledge of vehicle users. A total of 11 actors were included in the ecosystem sketch, 10 of which were connected with data sharing relationships. The eleventh actor, the vehicle user, "*does not receive any data, but actually only the services*". Main actors are OEM (7 links), vehicle owner (5 links), service provider (3 links), vehicle (3 links) and data marketplace (3 links).

*Expert 4* was involved in several large-scale projects with vehicle manufacturers, mentions six actors and adds that there are literally data sharing relationships between them all. The expert adds that a user can generally pay for services either "*with data or with money*". Furthermore, the expert argues that the data marketplace will be a "*closed platform [of OEMs]*", as the "*access to useful vehicle data is too critical to be open*", meaning that the information could be exploited to launch a cyber-attack on vehicles.

*Expert 5* (cf. Figure 3) is a senior manager of a vehicle service provider, designs an in-depth model and presents the vehicle as a central actor that passes on vehicle data to seven actors. The expert mentions an external influence through national and European regulations to positively influence OEMs to provide access to vehicle data for innovative service developments of other actors. The expert concluded by saying that "*trust is the key to the whole ecosystem*".

*Expert 6* (cf. Figure 3) sketches a data value chain from the vehicle driver via a data enricher, to a service provider who provides a service to the OEM, who in turn provides a service to any service user such as workshops, statistic services, enablers such as data marketplaces, insurance companies, or public authorities.

*Expert 7* is involved in the development of a data marketplace and has sketched a data flow from the OEMs to a data market provider that makes the data available to potential buyers and service developers. The expert describes "*while* the *OEM servers will host the* *data, the data markets will only do the contracting and data access and will be able to mesh data [from different data providers]*” and predicts that “*data markets will succeed and take hold [in the ecosystem]. Many of them are just beginning, and some successful ones will survive*”.

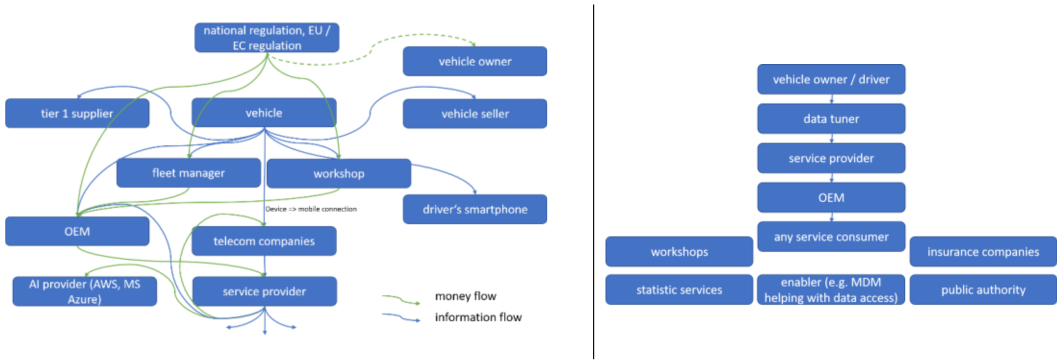


**Figure 70 Digitized vehicle data-driven service sketches of experts 5 (left) and 6 (right). (Kaiser et al., 2021)**

*Expert 8* is a consultant who sketches the model based on his own experiences in developing data-driven services with SMEs and public authorities. The main actor is “data”, which can be interpreted as a data platform or portal, but automotive (as data supplier) and infrastructure providers (who receive data from three other actors) play an important role, too. The actor ‘automotive’ (a synonym for car/vehicle manufacturers) “*also retrieves the data for own services, which is probably the main application for car manufacturers*”. Service providers, IT infrastructure providers and academic research are all relevant players in service provision, with access to the data remaining the key element.

*Expert 9* is the managing director of an automotive company and did not sketch direct data-sharing relationships but mentioned eleven actors. He sees OEMs in a stronger position, which is suggested by the statement that “*start-ups will disappear when larger players [such as OEMs] enter the [service] market”.* He doubts that external players will enter the value chain between the data source and the data enricher, because “*data should not simply be passed on to external parties, [..], CAN data must be interpreted correctly*”. He argues that “*the balance of power between technology companies vs. OEMs vs. public authorities will be crucial [for the future of the data-driven service ecosystem], and a balanced situation would be best*” for all stakeholders.

*Expert 10*, who is employed by a service provider, mentioned eight ecosystem actors. Vehicle data flows logically through the gateway provided by a gateway provider to a data platform provided by a hosting provider, to a service provider, and then to customers and fleet operators. The expert mentioned several data-sharing relationships during the interview but did not sketch them explicitly. The expert also mentioned the EU as an external influencing factor.

*Expert 11* is head of service development at a vehicle manufacturer and outlined three different actors in two different scenarios depicting the dominant role of vehicle manufacturers in the ecosystem. In the first scenario, where a customer uses a service from the OEM, the vehicle user allows data access, "*the consumer has the right to say no*", and pays the OEM for the service, which provides the technical infrastructure such as the mobile connection installed in the vehicle. The OEM in turn provides vehicle data to a contractually bound service provider and provides the vehicle user with the developed service. In the second scenario, the vehicle user buys a service from a third-party service provider and thus grants the service provider access to vehicle data. Due to strict European data protection legislation the vehicle user can "*already decide, which parties can be granted access to the data*". The service provider, in turn, uses the OEM's technical infrastructure, such as the mobile connection installed in the vehicle, and pays the OEM for its use, which the expert underlines by the statement that "*vehicles are equipped with more expensive technology to enable data sharing*".

The statements made by the eleven experts clearly show the influence of their own work on the designation of key actors and data sharing relationships. Experts working in the classic automotive industry (e.g. experts 9, 10 and 11) see the vehicle manufacturer in a dominant role in the data-driven service ecosystem, while scientific actors and those working in service development take a more differentiated view on the ecosystem..

## 15.4 A Conceptual Model for Value Creation in Vehicle Data-driven Services

### 15.4.1 Design Process

We have used two data sources, interview statements and expert sketches, to derive key actors and their data sharing relationships. We carefully examined the transcribed interviews and the individual conceptual models sketched by experts and extracted terms that had been used to describe the different actors. We ended up with a list of 90 terms, some of which were men-tioned more than once, and were finally able to identify 64 different actors. As experts tend to use different terms, levels and descriptions for the same type of actors (e.g.

'OEM', 'automotive manufacturer', and 'vehicle manufacturer') we have renamed some actors in order to create a consistent terminology for our conceptual model. Another challenge was the distinction between specific actors, such as cloud service providers, and providers of data-driven services to end-users. As a result, one group of actors was referred to as 'provider of cloud computing services', while another one was referred to as 'provider of data-driven services. We then categorised indi-vidual actors into groups and placed terms such as 'AI provider', 'cloud provider', and 'database provider' into the actor group 'platform provider'. After the ninth expert interview, it became apparent that no previously unknown actors were named by the experts who could not be classi-fied into the actor groups described below. Following the principle of theoretical saturation (Saunders et al. 2018), we therefore judged our sample to be complete. As a result, the cleaned set included 25 actor groups. In order to make the actor groups more tangible for our model de-sign, we identified the six highest ranked actors named by the experts and classified them into groups (Table 25).

**Table 25 Top ranked actors (left) and actor groups (right). (Kaiser et al., 2021)**

| Top ranked actors from expert interviews (N=64) | Quantity |
|---|---|
| OEM | 7 |
| Service provider | 6 |
| Infrastructure provider | 3 |
| Public authority | 3 |
| Road operator | 3 |
| Vehicle | 3 |
| [58 further actors] | 1-2 |

| Top ranked actor groups (N=25) | Quantity |
|---|---|
| Vehicle manufacturer | 10 |
| Data marketplace | 9 |
| Vehicle data service provider | 9 |
| Vehicle user | 7 |
| Consumer | 6 |
| Platform provider | 5 |
| [19 further actor groups] | 1-4 |

We built our first conceptual model on the expert interview statements and their model sketches. We thereby carefully examined transcribed interviews and their sketched models and extracted terms (actors, actor roles, types of data sharing relationships) to create a consistent terminology for the conceptual model (cf. Table 25). We then designed the first conceptual model of a unified model using only the main actor groups from the consolidated actor group list and upon the re-viewed literature, thus establishing proof-of-concept (Nunamaker et al. 2015). In a further de-sign step, we linked actors with data supply and data consumption activities to outline the data transformation process. The process from data supply to data use is often referred to in the scientific literature as the data value chain (Curry, 2016, Kaiser et al., 2019a, Latif et al., 2009, Miller and Mork, 2013). Our first design of a conceptual model was inspired by structuring ap-proaches that linked actors with data transformation steps (Latif et al., 2009).

After each design, we conducted an evaluation of the model that resulted in a model revision, either in a change in model actors, a change in data-sharing relationships, or a change in both, while the general structural design of the model remained unchanged. We carried out a total of six such iterations of the model in order to evaluate it against the individual views of the en-gaged automotive experts and against several real-life use cases of value creation in data-driven services. In doing so, we follow the suggestions of design researchers such as Gregor and Hevner (2013) to use case studies as a technique for conceptual model evaluation. With regard to Sonnenberg and vom Brocke (2011) and Venable et al. (2016), our evaluation can be seen as an ex-post evaluation, while we referred to the evaluation criteria model completeness, fidelity with real-world phenomena, internal consistency, level of detail and robustness as published by March and Smith (1995). In our ex-post evaluations, we also demonstrated the usefulness of the model to describe value creation in data-driven services, establishing proof of value (Nunamaker et al. 2015).

Using our designed conceptual model to describe value creation referring to concrete vehicle-data driven services led to several improvements of the model. We have already included the vehicle user as an essential element in the first conceptual model. However, as we found that many vehicle data-driven services are not enabled by vehicle data alone, but also use contextual data such as weather data or aggregated traffic data that obviously cannot be provided by the vehicle user alone, we added the actor role 'contextual data provider' in a second design itera-tion to the model to refer to actors providing other data as part of the value creation process. Furthermore, we have learned from several cases that the main beneficiary of vehicle-data driv-en services can be the vehicle user, e.g. by offering services such as intelligent parking while driving. However, vehicle data can also lead to services whose beneficiaries go beyond the vehicle user, e.g. by providing a dynamic map of traffic density to urban traffic managers. Hence, we have added the actor role 'other consumers' in a third iteration of our model. For space reasons, we will only show the final evaluation of our model with an exemplary real-life application case in Section 15.5.

### 15.4.2 Conceptual Model Description

The presented conceptual model is a result of iteratively designing a conceptual model. In our design process, we performed six iterations of the model, to evaluate it against the views of the interviewed automotive experts and against several real-life application cases of value creation in data-driven services. Existing theory on value of data-driven services (cf. Subsection 15.2.1) and data-driven value chains (e.g. Latif et al., 2009; Curry, 2016; Faroukhi et al., 2020) was used as theoretical input within this design phase. Our structural design of the model was informed by Latif et al. (2009), referring to entities that can act as ecosystem roles connected by data sharing relationships, i.e. consuming or providing data. Figure 71 shows the metamodel of our conceptual model. It outlines that each participating entity (i.e. organisations or persons) can act in one or more actor roles, thereby either providing data to the data-driven value creation process, consuming value-added data, or doing both (if more than one role is taken by the same ecosystem entity).

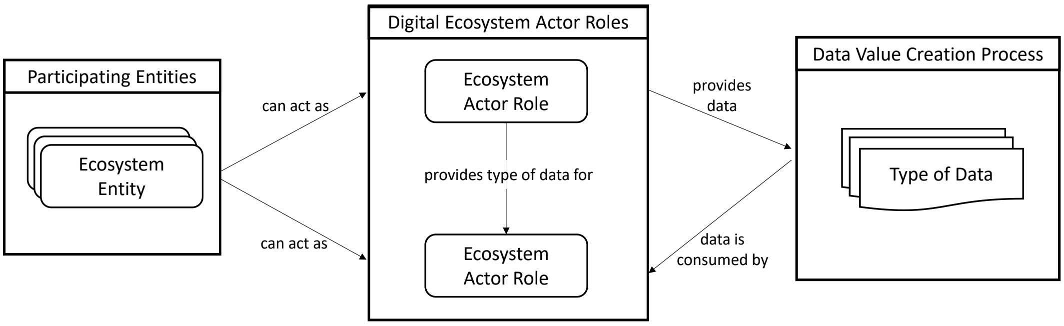


**Figure 71** **Metamodel of our conceptual model: Participating entities, digital ecosystem actor roles and types of data in the data-driven value creation process. (Kaiser et al., 2021)**

Participating entities can be individuals, organisations or organisational units that can take on one or more of the following actor roles: vehicle users as primary data providers, contextual data providers offering additional data for service design, vehicle manufacturers that can exploit access to vehicle sensors, gateway providers collecting vehicle data with their own equipment, data marketplace or portal providers allowing access to data via their application programming interfaces (APIs), data-driven service providers, and finally vehicle users as well as other consumers. We will now take a closer look into these ecosystem actor roles and illustrate their data sharing relationships.

A *vehicle user* is a professional or private actor that decides to provide *vehicle data* (i.e. data generated while vehicle operation by sensors and electronic control units) to be used in data-driven services in any format and in any level of aggregation to the related *vehicle manufacturer* directly, or to other actors via a *gateway provider* indirectly. Vehicle users must give their consent to the sharing of vehicle data to other ecosystem actors.

A *contextual data provider* is any organisation that has additional contextual data that is relevant to the provision of data-driven services and is willing to share this data for service development. Examples of contextual data providers are companies that can provide geo-data, weather data, traffic data or map data, but also governmental actors that publish open data.

A *vehicle manufacturer* is an actor that develops, manufactures, and maintains vehicles as its main industrial product. Vehicle manufacturers have equipped vehicles with advanced sensors that collect and process a wealth of data to ensure the driving function, optimize the vehicle's internal functions and facilitate safety. Most vehicle manufacturers have equipped their latest vehicles already with telematics software and connectivity to allow use of the data generated in data-driven services. Various types of vehicle dynamics data such as vehicle speed, acceleration, rotation, position as well as other data such as information on fuel, battery, service, and window status, wheel rotation, or steering wheel angle can be provided at different sampling rates.

A *gateway provider* is an actor that collects either raw or processed vehicle data or other contextual data such as weather data or map data for the development of data-driven services. Gateway providers may collect vehicle data through deploying a data capturing device connected to the vehicle's on-board diagnostics interface (OBD), to the controller area network bus (CAN), or through the use of a dedicated independent sensor and connectivity device. Besides, vehicle dynamics data can also be collected by the gateway provider using a special sensor kit that is not connected to the vehicle's bus systems or a mobile application installed on a smartphone which captures data from smartphone sensors while the vehicle is moving.

A *data marketplace, platform or portal provider* is an actor that receives data from various vehicle manufacturers, contextual data providers, and/or gateway providers and performs data harmonisation, transformation, and storage activities, either with the distinct purpose of selling service-specific vehicle data and/or service-relevant data from third parties (*marketplace, platform*) to enable the development of data-driven services, or to provide such data for free (*portal*). Data market place providers may provide data to the developers of data-driven services who only need to integrate once via their APIs, instead of having to enter into many different relationships with OEMs and other data suppliers, while at the same time having to deal with diverse (and changing) data formats.

A *provider of data-driven services* is an actor that consumes service-specific (vehicle) data from a data marketplace or data portal provider and provides consumable service data to a service user, which in turn may be either a vehicle user or another consumer, i.e. any other type of end-user or organisation wishing to consume a data-driven service enabled by vehicle data and probably enriched with other contextual relevant data. Vehicle data service providers ultimately offer data-driven services, such as road surface quality detection, harsh driving detection, or predictive maintenance.

Finally, both *vehicle users* and *other consumers* may be consumers of data-driven services offered by data-driven service providers. As the final actor in the data-driven value chain, these professional or private actors are end users and main beneficiaries of the entire data transformation process. Examples of data-driven services are road surface quality detection (consumed by municipalities or a road infrastructure managers) or harsh driving detection (consumed drivers to improve their driving style or by insurance companies to provide a 'pay as you drive' insurance that calculates the insurance premium based on the driving style).

Actors provide and consume different types of data within the data value creation process. First, in order to comply with data protection regulations such as the General Data Protection Regulation (GDPR) in Europe, vehicle users should grant access to the data their vehicle generates before *vehicle data* may be used in services. *Contextual data* relevant for the development of a particular data-driven service, such as weather data, traffic data or data on accident hotspots, are provided by providers of contextual data for the data value creation process. This data can be provided as *raw vehicle/contextual data* (e.g. as data that is measured and collected directly from vehicle sensors without any kind of pre-processing) or as *processed vehicle/contextual data* (i.e. including some kind of data cleaning, transformation, resampling and conversion into a data format that is better suitable for service development). Service-specific *vehicle/third party data* is provided by a data marketplace, platform or portal provider that has been transformed from raw or processed vehicle data into a form that can be used by data-driven service developers within data-driven services. Finally, *consumable service data* is provided by data-driven service developers to vehicle users and third parties within provided applications (services), creating value for the end-users.

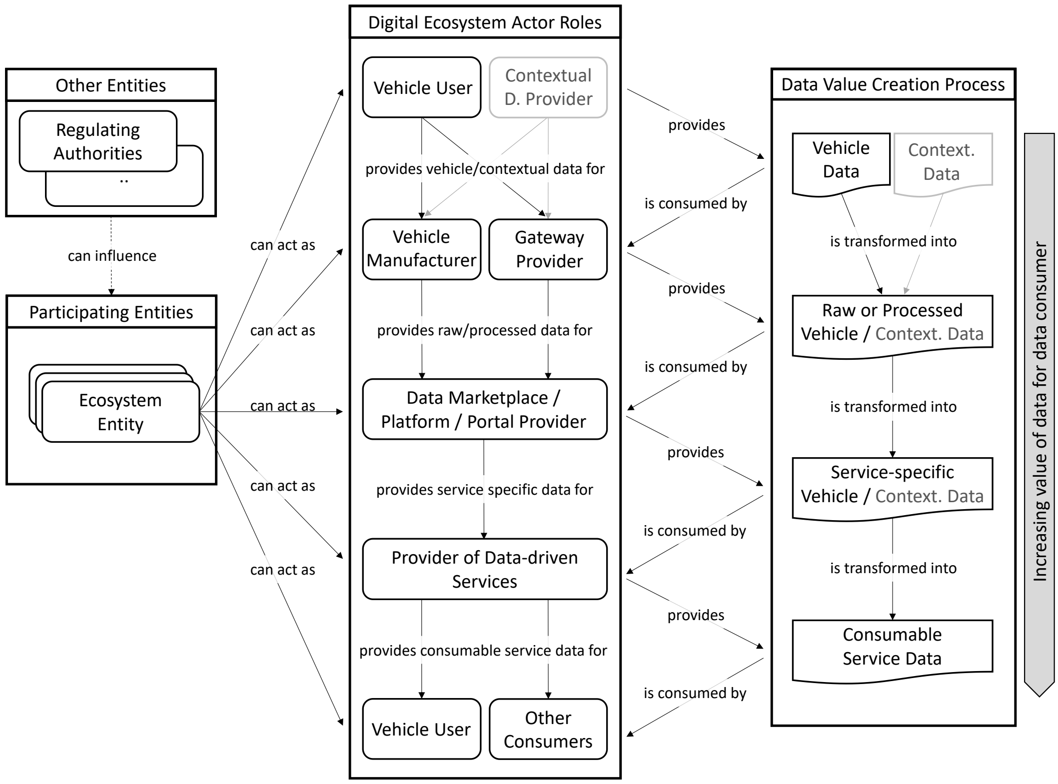


**Figure 72** **A conceptual model for value creation in vehicle data-driven services. (Kaiser et al., 2021)**

The conceptual model, as shown in Figure 72, outlines individual actors and their steps in vehicle data-driven value creation. The value concept we used in the model is added value for the data consumer. From the perspective of end-users, consumable service data is the most valuable data. Therefore, end users may be willing to provide monetary or non-monetary consideration for this type of data.

## 15.5 Evaluation

We evaluated our conceptual model (the artifact) ex-post by applying it to six *real-life cases* such as designing a *data-driven service for road surface quality detection*, to identify actors and data sharing relationships as shown in Figure 73. After each design, we conducted an ex-post evaluation (Sonnenberg and vom Brocke 2011; Venable et al. 2016) of the model that resulted in a model revision, either in a change of model actors, a change of data-sharing relationships, or a change of both, while the general structural design of the model remained unchanged. The presented case is the sixth and final evaluation of the model and based on real experiences of two authors working on the project mentioned in the introduction. After this last evaluation, the model remained stable.

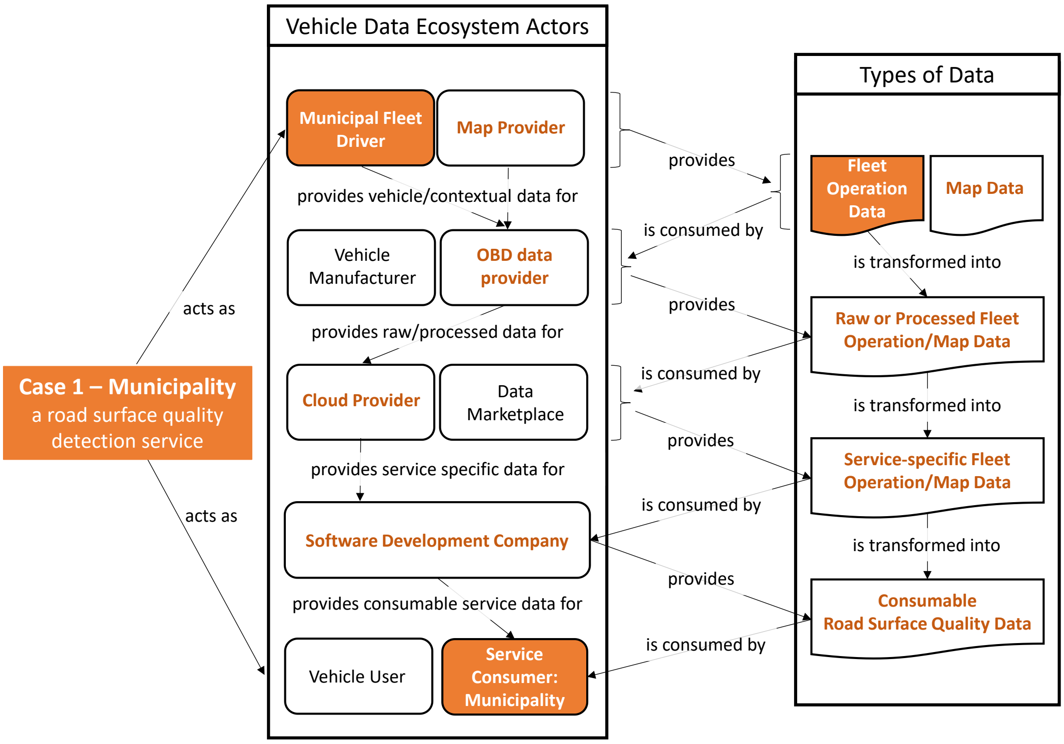


**Figure 73 Actors and data sharing relationships in the design of a data-driven service for road surface quality detection. (Kaiser et al., 2021)**

A *data-driven service for road surface quality detection* can be envisaged by the municipality of a city, responsible for a road network (e.g. the City of Vienna with a road network of 3.000 kilometres). The municipality operates a vehicle fleet and has an infrastructure management department which orchestrates road maintenance work. Thus the municipality acts as *vehicle user (collecting fleet operation data)* and *other consumer (consuming road surface quality data)* in this case (orange background colour in Figure 73).

Our model indicates drivers from the municipality fleet as data creators who can opt in to deliver service-relevant data such as vehicle speed, acceleration, rotation and position. Furthermore, map data from a contextual data provider (Map provider, orange font colour) must be used for georeferencing detected road surface quality from recorded vehicle data. Vehicle data can be captured directly by vehicle manufacturers in case they already operate vehicles in this city that can capture and transmit those data, which requires special contracts with selected manufacturers. Vehicle data can also be collected via gateway providers that provide devices for installation in vehicles equipped with the necessary sensors, which presupposes that a sufficient number of gateways are installed in vehicles moving in the area where road surface quality is to be measured. A municipality may already operate own fleets with different vehicle brands, consisting of all employees' business vehicles. In this case, the municipality may form a business relationship with a gateway provider (OBD data provider, orange font colour) to support the data collection by equipping all vehicles in the fleet independent of their brand with gateways. Vehicle operation data being collected could be prepared for further processing and then either be made available to the provider of the data-driven service (a company in charge of developing the road surface quality detection service) by a platform / portal provider or a data marketplace ideally in combination with corresponding map data from a map provider. As the provider of the data-driven service already cooperates with the cloud provider AWS (orange font colour), AWS is also chosen as platform provider. Finally, a contracted software development company (orange font colour) is responsible for developing the road surface quality detection service and takes vehicle data and map data required for service provision, applies a data processing approach, extracts events indicating a particular road surface quality such as potholes from identified deviations of the processed vertical acceleration and pitch together with their positions (vehicle data), and visualises the geographic position of identified potholes in a web dashboard (using map data) to create value for the service consumer, the infrastructure management department of the municipality (orange background colour). In addition, the software development company prepares a table with prioritised repair lists and interaction possibilities to investigate the worsening or improvement of road surface quality for the infrastructure management department.

This real-life application case illustrates the complexity of developing a road surface quality detection service and shows the usefulness of our model for better understanding the roles of concrete ecosystem actors and their data sharing relationships in the development of a vehicle data-driven service.

## 15.6 Discussion

Vehicles are increasingly equipped with advanced sensors to ensure driving functionality, optimise the vehicle's functions, and facilitate safety and comfort through increased automation such as providing adaptive driving assistance systems (Stocker et al., 2017a). Moreover, most vehicle manufacturers have additionally equipped their latest vehicles with advanced software and connectivity to make use of the data generated and to provide additional services to drivers. The data generated during vehicle use can enable new types of data-driven services addressing many interesting use cases (if drivers opt in to vehicle data sharing) that go far beyond supporting the operation of vehicles, especially through intelligent linking of vehicle sensor data with other contextual data such as weather data or data on the traffic situation. This raises the important question of which ecosystem actors can and will contribute to use cases that can only be implemented if data is shared between multiple actors.

### 15.6.1 Implications for theory

The theoretical gap addressed in this chapter is the lack of conceptual models that can unravel the underlying value chain (actors and data sharing relationships) when establishing vehicle data-driven services. In this chapter, we have therefore presented a novel conceptual model that includes multiple actors and their data sharing relationships (i.e. in terms of a data value chain) that are relevant for vehicle data-driven value creation. As such, our multi-actor model shows data and information flows as a series of data sharing and data transformation steps that are needed to finally generate value and useful insights to service consumers, establishing proof-of-concept (Nunamaker et al., 2015). Following Baskerville et al. (2018) we present a novel and useful conceptual model and thus generate a significant contribution. While previous research in (big) data has shown a clear focus on data users (Wiener at al., 2020), we also emphasize the importance of data providers and intermediaries and their interactions in a multi-actor model, thus extending the perspective to the ecosystem where the value creation is enacted. Consequently, we emphasize that data-driven value creation in the automotive ecosystem must be achieved through collaboration among various stakeholders, thus contributing to the debate on realizing value from (big) data (cf. Günther et al., 2017) by stressing a multi-actor perspective.

Several researchers in the field of information systems have also been engaged in the study of value creation from big data because big data is a comparatively new phenomenon and the organizational implications of big data are of great interest to them (e.g., Akter et al., 2016; Dremel et al., 2017; Dremel et al., 2020; Grover et al., 2018; Lehrer et al., 2018; Mikalef et al., 2017; Svahn et al., 2017; Wamba et al., 2015; Wamba et al., 2017). However, they focus on the impact of big data analytics on the level of an individual organization (e.g., on organizational performance, strategic business value, strategic use, organizational change, or required organizational capabilities) and exclude the network and ecosystem perspectives for creating data-driven services. They focus on an intra-organizational (i.e., micro) perspective, whereas we want to look at value creation in a multi-actor ecosystem (i.e., macro) perspective. While their research specifically targets the big data phenomenon, we want to emphasize that value can also be created from services enabled by the exchange of small data between actors.

Several studies investigate how vehicle usage data can lead to novel services, such as location-based services for carsharing vehicles (Wagner et al., 2015), predictive maintenance of connected vehicles (Gerloff and Cleophas, 2017), or eco-feedback on driving behaviour (Bätz et al., 2020). Yet, these studies focus rather on data analytics approaches to exploit vehicle data than on the data ecosystem perspective. Also considering the state of the art gaps discussed in previous sections, we argue that our proposed conceptual model would allow relevant actors to be identified and mapped in order to eventually achieve periods of stability and change (Nischak et al., 2017, p. 17) and the interactions that ultimately lead to the envelopment (Eisenmann et al., 2011) of other emerging digital business ecosystems. Furthermore, our model indicates choices for how the value chain can evolve and, above all, which other actors are needed, because the development of a data-driven service and the selection of suitable actors is a decision-making task.

Our model shows that actors are involved in a multi-party data value creation process to ultimately provide sustainable data-driven services to service customers such as vehicle drivers and therefore contributes to a better understanding of vehicle data-driven value creation in general. Based on our interviews with experts, all of whom have a connection to vehicle data-driven value creation and some of whom are developing these vehicle data-driven services themselves, we have learned that the successful development and provision of data-driven services in the automotive domain and thus the successful monetisation of vehicle operation data will require new partnerships between individual ecosystem actors, as no actor will bear the service development risk alone. We argue that our conceptual model provides a solid understanding of the ecosystem actors and their role in data sharing and in the creation of data-driven services, thus supporting strategic decisions, e.g., in terms of partnerships and sourcing. In doing so, we are contributing to research on data monetization, responding to the call by Parvinen et al. (2020) for a better understanding of the role of data aggregators and refiners in data monetization, how they create value and how different parties can capture it.

We have developed our model empirically, drawing on the knowledge of automotive domain experts who have an average of more than 16 years of professional experience in the mobility industry. Laying emphasis on actors that have a stake in data generation and sharing, we differ methodically from the approaches of other researchers who study ecosystems in the mobility domain, including Riasanow et al. (2017) using crunchbase.com data to visualise the current automotive ecosystem in a generic value network, Remane et al. (2017)

focusing on the identification of business model types of start-ups, or Kolbe et al. (2017) creating an IoT framework and focusing on semantic interoperability.

Our background is in the field of data-driven service development in the automotive domain, and we stress that our conceptual model is inspired by research on data-driven value creation published by Curry et al. (2016), Miller and Mork (2013), or Latif et al. (2009). Our concept of connecting automotive ecosystem actors with data sharing and enrichment processes is new. We understand our model as a descriptive tool that shows the process towards providing a data-driven service from both an actor and a data perspective. Furthermore, we believe that our presented research is also helpful in better describing and classifying existing data-driven services. Our model can support ecosystem actors to better recognise and understand their interdependencies with other actors or even to understand what interdependencies exist at all.

It is worth mentioning that actors within the ecosystem for vehicle data-driven value creation are different from the classical actors within the vehicle supply chain. For instance, although vehicle manufacturers (OEMs) are heavily dependent on original equipment suppliers in the supply chain, these Tier-1 (module or system suppliers) and Tier-2 (component suppliers) are not specifically addressed in our model. However, they have an indirect relevance within the creation of vehicle-data driven services: First, they can supply the vehicle telematics device to the vehicle manufacturer, which enables data acquisition and data transfer to the manufacturer's backend servers. However, suppliers do not have a direct role within the process "from data to service", as they do not have direct access to the vehicle data transmitted by their supplied telematics units to the vehicle manufacturer. Second, suppliers may act as service developers providing not only hardware but also data-driven services to vehicle manufacturers. If suppliers choose to do so, they are included in the model in the actor role "Provider of data-driven services". We have deliberately avoided an actor role "supplier" in the model. For example, Tier-1 Robert Bosch GmbH not only designs vehicle telematics devices but also offers data-driven services, such as road condition-based services (Bosch, 2020a) or Connected Horizon (Bosch, 2020b). The development and provision of both services can be well described by the use of our model, and both cases served within the conceptual model evaluation process. Third, suppliers can act as users of a data-driven service, and in this case, are included in the model as "other consumers". A prominent example case is the provision of a data-driven service for ECU health, that is made available to suppliers. This service can help suppliers to monitor the functionality of ECUs they have designed and delivered to vehicle manufacturers and that are installed in the vehicle by the OEM. Suppliers can also take advantage of driving style recognition or environmental condition monitoring services that will both help them to improve their ECU designs as well.

### 15.6.2 Implications for practice

In addition, we see several implications of our work for business practice. Based on a specific role of an ecosystem actor, we have shown in the evaluation that our conceptual model is useful to practitioners to better understand their own position in the ecosystem.

For example, a manager responsible for digitalisation can identify which actors are relevant to provide data-driven services. In addition, service developers may recognise the special role of a vehicle user, without whose consent to the provision of collected data the development of a data-driven service will not be possible. Vehicle manufacturers may be able to better communicate their own position in the value chain as the one who can technically store, interpret, and forward generated vehicle data. The manufacturer may recognise that a scaling provision of certain data-based services will only be possible if other actors are granted access to the vehicle's bus information systems or if the manufacturer stores, transmits and makes vehicle data available to others via its own datacentre.

Start-ups interested in producing data-driven services may realise that they can also turn to data marketplaces that have already signed contracts with vehicle manufacturers and do not need to negotiate individually with each manufacturer to access the necessary data. The provision of vehicle data to data marketplaces can also lead to new ways for vehicle manufacturers to monetise vehicle data, namely when others use it to develop services that generate value independent of their core product, the vehicle. Those who wish to design data-driven services can better identify the key players in the ecosystem they need to deal with, and those who want to be part of the service delivery process can better understand who they need to work with. Since one of the first decisions for organisations seeking to monetize vehicle data is to figure out, where to play in the value chain (Hood et al., 2019), we consider the knowledge contained in our model to be a significant contribution.

### 15.6.3 Limitations

In our concept phase, we tried to generalise the expert's individual mental models on data-driven value creation in order to eliminate individual perspectives as much as possible. Furthermore, we have involved eleven experts from Central Europe in the data collection, who also work together with specific players in the automotive ecosystem and thus contribute their own views. All interviewed experts are opinion leaders for the Central European market (the location of some of the largest vehicle manufacturers in the world), and therefore we believe that the interviewed experts represent an impressive amount of knowledge. The interviews and individual sketching activities of the experts showed that there was a consensus on many important patterns (i.e., on the roles of the actors and their data sharing relationships). This seems to show that our sample is appropriate for our research purpose. It is also worth noting that two of the authors have been working in the automotive sector for eight years each. Their contextual bias is mitigated by closely involving the other three authors in the research process in order to adopt an external and critical perspective, and by reflecting the results of the design process with them, so that "the higher-level perspective necessary for informed theorizing" is maintained (Gioia et al., 2013, p. 5). Finally, we have evaluated the model in total six times ex-post by applying it to real-life cases, establishing proof-of-value (Nunamaker et al., 2015). Furthermore, we established proof-of-use by successfully applying the model in a research proposal that was granted with funding.

## 15.7 Conclusion

In this chapter, the authors adopt an ecosystem (i.e., macro) perspective and propose a novel conceptual, multi-actor model for value creation in vehicle data-driven services consisting of ecosystem actors and their data sharing relationships, establishing proof-of-concept. They thereby illustrate how different key actors such as vehicle users, manufacturers, data marketplaces, and service providers have to engage in data sharing relationships to create value from vehicle data (i.e., data that is collected by the vehicle's sensors) and other relevant contextual data. They evaluated their model ex-post by applying it to six real-life application cases, such as the development of a vehicle data-driven service for road surface quality detection, which they also present in this chapter, establishing proof-of-value.

The theoretical gap addressed in this chapter is the lack of conceptual multi-actor models that can unravel the underlying value chain (actors and data sharing relationships) when establishing (vehicle) data-driven services and consider an ecosystem perspective. Many of the researchers cited have focused on the perspective of a single organization, with an emphasis on deciphering the phenomenon of big data analytics and its implications at the intra-organizational (i.e., micro) level. As our evaluation has shown the conceptual model contributes to a better understanding of the (data-driven) value creation logic and reveals critical actors and their data sharing activities that ultimately lead to created value.

While the authors of this chapter designed their conceptual model as a high-level model to reduce the complexity of the whole automotive ecosystem and focus on vehicle data provision and use, they are aware that their model cannot represent and explain all relevant aspects, value flows and power relations. We have focused on the data value chain and have therefore only included the most important actors in terms of data sharing. Nevertheless, the authors see numerous practical implications as our model could be used as a governance and/or creativity tool to influence data sharing regulation (e.g., to better understand the dominant role of the OEM in enabling vehicle data-driven services) or even for the design of data-driven services outside the automotive domain. In addition, at an academic level, the authors see their research as a first contribution to the systematic design of a multi-actor model for vehicle data-driven value creation in the automotive sector that can help to guide next research endeavours in data-driven service development.

Finally, the authors expect their article to have further implications on research such as becoming a structuring tool to design, compare and / or analyse cases of data-driven service development, or simply help future researchers to better understand potentials and pitfalls in the development of data-driven services. The authors even believe that the presented model is transferable to other domains where non-digital artefacts are the core product that generate data during use (such as the aircraft industry), although proving this claim would go beyond the scope of this chapter. Going after this claim, however, may spur future research endeavours.

# 16. Summary, Conclusion, and Outlook

First of all, from an economic point of view, the automotive industry is one of the most important industries not only in Germany, Austria, and Europe but worldwide. The digital transformation opens up new opportunities for value creation in the automotive industry, but in some cases forces radical changes (e.g. from vehicle manufacturers) to remain competitive on the market. In this context, data is becoming an increasingly important asset, enabling data-driven services that can be used to open up new customer segments. For the development of these, new players are also pushing into the ecosystems in which such data-driven services are developed in the cooperation of several roles (e.g. data provider, data marketplace, data-driven service developer, etc.).

Since these data-driven services were only increasingly noticed by vehicle manufacturers and the traditional automotive industry with the start of this dissertation and since there was hardly any related work, the investigations of this dissertation clearly extend the body of academic knowledge, potentially have a high impact to practice and serve to inform and shape future automotive research and practice. Thereby, contributions were made in many individual areas, which were presented in chapters 2-15. In a nutshell and to close the summary, the following core contributions of this dissertation can be mentioned:

(i) Definition and description of *Quantified Vehicles* as a form of digitalization in the automotive domain
(ii) Understanding how *Vehicle Data* becomes a relevant artifact for business and innovation
(iii) Understanding and prototypical development of concepts and *Data-driven Services* along the *Vehicle Data Value Chain* that represent added value for consumers
(iv) Understanding of the process and actors of value generation, and the interplay of the actors with each other in the *Data-driven Service Ecosystem*
(v) Development of data-driven value generation and *Data-driven Service Ecosystem* models

The intended recipients of the conducted applied research are equally the actors of the industry, especially the actors of the *Data-driven Service Ecosystem* (e.g. OEMs, ICT companies), as well as the scientific community, e.g. the IS community or automotive research in general. I hope that the results have created a fruitful basis on which other researchers or even industry players can build further.

The fact that the implementation concepts and implementations were only ever prototyped and never matured into real products is a clear limitation, even if they were developed on the basis of existing literature or empirically collected expert know-how. The same applies to the *Data-driven Service Ecosystem* models, that lack continuous evaluation. It is worth mentioning that the ecosystems of vehicle manufacturers (who play a dominant role as data providers) are still in the formation phase and therefore in a state of transition. Further limitations are the study sizes, because field studies, surveys, case studies, etc. were mostly conducted in small to medium scales. Furthermore, analogies from other industries (e.g. platform techniques of the software industry) could have been investigated more intensively, as well as to check whether the findings might be valid for domains other than the vehicle domain (e.g. for the agricultural sector).

In the end, however, I was able to improve the understanding of what *Quantified Vehicles* actually are, how (technically) vehicle data can be turned into valuable *Data-driven Services* for end customers, which questions are included in a research agenda the *Information Systems* (IS) community could answer, how the underlying *Vehicle Data Value Chain* is structured, which could be a *frame of reference* regarding the processes, how services can be developed concretely, which privacy challenges there are and how they could be solved (e.g. with the *Open Vehicle Data Platform*), and how data about vehicle use can be collected without vehicle manufacturers with a *Vehicle Data Logger* and how commodity hardware can be integrated.

Second, I would like to draw a conclusion about what happened in the years 2016 to 2020 / 2021. Working in automotive research has already given me an insight into a relatively conservative and mechanical engineering-driven automotive industry before, which is suddenly facing major challenges and a shift towards an even more software-intensive industry as a result of digital transformation. Data-driven services were praised by analysts and others as a great opportunity to generate value and revenue in the future in the automotive domain. So this topic was very appropriate for a business information scientist, and supporting this first phase turned out to be very exciting. The different contributions made in many individual areas are expected to have an impact on both industry and science.

However, some of the start-ups (Automatic, dash), which were pioneers in the field with their data-driven services (mostly based on OBD data and based in the USA) and introduced this innovation, have already disappeared from the market again. To go into more detail, Automatic for example, founded in 2011, was once financed with enormous venture capital investments of more than USD 24 million (Stocker and Kaiser, 2016), was bought up by SiriusXM for over USD 115 million in April 2017 (TechCrunch, 2017), and decided to discontinue its services in May 2020 (SiriusXM, 2020). In parallel, through standardization processes (ISO, C-ITS, NEVADA) as well as many (research and development) projects (e.g. AutoMat, Cross-CPP), most vehicle manufacturers in Europe have now taken up the topic of *Data-driven Services*. New vehicles (e.g. from BMW, Daimler) often already offer now the technical infrastructure to theoretically collect all sensor data (e.g. wheel speeds per wheel, speed, tyre pressure, engine load, steering angle, ABS / ESP on or off, ACC in use, etc.) of the vehicle and make them available for algorithms of data-driven services, if the driver or vehicle owner wants to do so. This was not the case when the dissertation was started. Manufacturers in 2020 even are calling data-driven service developers to become part of their own ecosystem to develop manufacturer-specific data-driven services that customers (e.g. BMW or Daimler) can then install and use in their vehicles, similar to the app stores on smartphones. At the same time, newly developed data marketplaces (such as Caruso and Otonomo) offer the possibility of accessing data from several manufacturers via the platform. So it can be seen how the industry is developing and it was great to be able to accompany this development and shed light on some chosen topics like the value creation within the *Vehicle Data Value Chain*.

So, what will happen next with this research and the topics involved in *Quantified Vehicles*: *Data, Services, Ecosystems*? As an outlook, with (higher levels of, e.g. SAE levels 2-5) *automated driving*, there is a further driver to make data available for Vehicle-2-Vehicle (V2V) and Vehicle-2-Infrastructure (V2I) communication. For a smooth traffic flow between a mixture of automated, semi-automated, and non-automated vehicles, one cannot rely solely on sensor technology within vehicles to detect the environment correctly, and so the exchange of vehicle and driving data is an important component here too, from which many other areas can benefit. Hence, the possibilities for data-driven services based on vehicle usage data are still increasing, be it traffic planning, traffic safety, driver tutoring, gamification, product optimization, or new products (e.g. navigation to the next available parking space), etc, as soon as the vehicle usage data is available in an adequate form. Whereas the question remains whether the proprietary solutions of individual manufacturers (e.g., BMW and Daimler) with their own ecosystems in which service developers can develop and offer data-driven services will be the final state, or whether the provision of data (interfaces), selection of available signals, and data frequency will be standardized in the future to enable holistic traffic analyses or automated driving, for example, and whether several years of the typically very long automotive development cycles will then again pass until a standardized solution is available in series vehicles.

However, e.g. at the conference *monetizing car data* in 2020, vehicle manufacturers like Audi, BMW, and Daimler stated that there is no longer any talk of profit based on vehicle data from vehicle manufacturers. The pools of ideas are full at all manufacturers (they say), but they can hardly be monetized, was the common opinion. It is expected that there will be services that stand out and bring profit, but the outstanding idea has not yet been implemented. However, since it is possible, that such a service will still be found, manufacturers emphasized the "added value" that is created by the offer to justify additional hardware and software costs. However, OEMs have learned that it may take many attempts to find a profitable application, which is why the analysts, among others, are taking up the issue of monetizability and offering to advise the OEMs, e.g. Capgemini Invest (Winkler et al., 2020) stated that "early achievements lag behind expectations". B-2-C services, for example, seem to have little potential at present, as this customer group (e.g. private drivers) is only willing to pay for applications to a limited extent (Cäsar et al., 2020) or to even share data (Winkler et al., 2020). A Capgemini Invent study showed, that "out of 23 use-case categories investigated, safety- and security-related services are valued most [e.g. Collision warning, hazard warning, theft detection, automatic pedestrian detection, and eCall] while in-car delivery and commerce are valued least" (Cäsar et al., 2020), this is why safety and security services are still the most likely to be paid for, followed by practical supports such as intelligent route planning, real-time parting information, automatic distance control. The B-2-B sector, on the other hand, is different. Startups such as *Carployee.com*, for example, sell their commuting solution to the company, and all employees can then use the app for ridesharing. In the meanwhile, OEMs attempt to build so-called ecosystems in which third party service developers and service providers can develop services. In this situation, vehicle manufacturers can cover their costs in the role of a data and platform provider. In parallel, there are efforts to achieve cross-manufacturer cooperation in Germany.

The next few years will show which services will establish themselves on the market. In any case, the work of this dissertation on "*Quantified Vehicles: Data, Services, Ecosystems*" has made a contribution to better understanding the value-creating steps of the VDVC and, for example, has shown what the ecosystem currently looks like and thus also has a share in future changes.

My work will continue seamlessly, including newly funded research projects, albeit with a slightly different focus. In the beginning of 2021 the bilateral (with consortium partners from Germany and Austria) research project named 'D-TRAS' (Digital Platform for Traffic Safety-Risk Prediction in Rural Areas) was launched, which uses vehicle data to detect and warn motorcyclists of hazards like a slippery turn. In addition, there are plans to leverage the emerging vehicle data marketplaces (e.g., Caruso Dataplace), and add driver smartphone data to them to expand the capabilities. Smartphones are well suited due to their widespread use and the elimination of hardware costs. There is already existing research and literature to use smartphone data (e.g. GPS, camera, IMU sensor, microphone, radio signals), for example to detect driver distraction, one of the major causes of road accidents, or to automatically calculate commuting suggestions to ultimately reduce congestion at peak times, and thus also unnecessary $CO_2$ emissions. So even if vehicle manufacturers here do not yet reap big profits from vehicle data in the short term, vehicle data can pave the way to reduce congestion, pollution, stress and, most importantly, fatalities, and help road users in areas such as parking, safety, etc., so that we can get from A to B more happily in the future. And so I am sure vehicle data will continue to play a big role in the future.

# III. Acknowledgement

This dissertation work was partially conducted within and supported by EU H2020 research projects (especially AEGIS, SCOTT, EVOLVE, and InSecTT), a national research project (D-TRAS) and the employer, Virtual Vehicle Research GmbH with its COMET K2 funding. Enclosed are the official declarations of acknowledgement:

**AEGIS**: The AEGIS project (https://www.aegis-bigdata.eu/) has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 732189. The document reflects only the author's views and the Commission is not responsible for any use that may be made of information contained therein.

**SCOTT**: This work is partially funded by the SCOTT project (http://www.scott-project.eu). SCOTT has received funding from the Electronic Component Systems for European Leadership (ECSEL) Joint Undertaking under grant agreement No 737422. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and Germany, Austria, France, Czech Republic, Netherlands, Belgium, Spain, Greece, Sweden, Italy, Ireland, Poland, Hungary, Portugal, Denmark, Finland, Luxembourg, Norway, Turkey. In Austria, the project was also funded by the program "ICT of the Future" and the Austrian Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK).

**EVOLVE**: The EVOLVE project (www.evolve-h2020.eu) has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 825061. The document reflects only the author's views and the Commission is not responsible for any use that may be made of information contained therein.

**InSecTT**: This work is partially funded by the InSecTT project (https://www.insectt.eu/). InSecTT has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 876038. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Austria, Sweden, Spain, Italy, France, Portugal, Ireland, Finland, Slovenia, Poland, Netherlands, Turkey. In Austria, the project was also funded by the program “ICT of the Future” and the Austrian Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK).

**D-TRAS:** The D-TRAS project has received funding by the Austrian Research Promotion Agency (FFG) in the ICT of the Future programme under project number 880046. ICT of the Future is a research, technology and innovation funding programme of the Republic of Austria, Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK). The Austrian Research Promotion Agency (FFG) has been authorised for the programme management.

**COMET K2:** As an employee of VIRTUAL VEHICLE Research GmbH in Graz, I also acknowledge the financial support of the COMET K2 – Competence Centers for Excellent Technologies Programme of the Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK), the Federal Ministry for Digital and Economic Affairs (bmdw), the Province of Styria (Dept. 12) and the Styrian Business Promotion Agency (SFG). The Austrian Research Promotion Agency (FFG) has been authorised for the programme management.

# IV. References


Abbasi, A., Sarker, S. and R.H. Chiang (2016). “Big data research in information systems: Toward an inclusive research agenda.” *Journal of the Association for Information Systems*, *17*(2).

ABC Science (2018). Driver Distraction, https://www.youtube.com/watch?v=yVbPGmsG5sI [last accessed April 2020]

Accenture (2016). Digital Transformation of Industries: Automotive Industry. https://www.accenture.com/t20170116T084448__w__/us-en/_acnmedia/Accenture/Conversion-Assets/WEF/PDF/Accenture-Automotive-Industry.pdf [last accessed April 2020]

ACEA (2016). ACEA (European Automobile Manufacturers Association) Position Paper: Access to vehicle data for third-party services https://www.acea.be/publications/article/position-paper-access-to-vehicle-data-for-third-party-services [last accessed April 2020]

ACEA (2017). ACEA (European Automobile Manufacturers Association) CAR DATA FACTS http://cardatafacts.eu/ [last accessed April 2020]

ADAC (2016). Autohersteller sammeln Daten in großem Stil: www.heise.de/newsticker/meldung/ADAC-Untersuchung-Autohersteller-sammeln-Daten-in-grossem-Stil-3227102.html [last accessed May 2020]

Adner, R. (2016). "Ecosystem as Structure: An Actionable Construct for Strategy." *Journal of Management*, *43*(1), 39–58.

AEGIS (2017). https://www.aegis-bigdata.eu/ [last accessed May 2020]

Afuah, A. and C. Tucci. (2000). Internet Business Models and Strategies. McGraw-Hill College, Boston.

Ahmed, F. (2016). *Using structured analysis and design technique (SADT) for simulation conceptual modelling* (Doctoral dissertation, Loughborough University).

AISeL (2020). Association for Information Systems (AIS) eLibrary, https://aisel.aisnet.org/ [last accessed October 2020]

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes, 50(2)*, 179-211.

Akaka, M.A., Vargo, S.L. and R.F. Lusch (2012). “An exploration of networks in value co-creation: a service ecosystems view.” Review of Marketing Research 9:13-50.

Akter, S., Wamba, S.F., Gunasekaran, A., Dubey, R. and S.J. Childe (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*, 182, 113–131.

Albertsson, M., & Edström, J. (2013). Generating Revenue-Streams from Vehicle Connectivity Solutions-A case study of Volvo Car Corporation. https://gupea.ub.gu.se/bitstream/2077/33374/1/gupea_2077_33374_1.pdf [last accessed October 2020]

Alcantarilla P.F., Bartoli, A. and A.J. Davison (2012). “KAZE Features.” In: Fitzgibbon A., Lazebnik S., Perona P., Sato Y., Schmid C. (eds) *Computer Vision – ECCV 2012. Lecture Notes in Computer Science*, vol 7577. Springer, Berlin, Heidelberg

Al-Debei, M.M. and D. Avison (2010). Developing a unified framework of the business model concept. European Journal of Information Systems 19: 359–376. https://doi.org/10.1057/ejis.2010.21

American Psychological Association (2018). Understanding Driver Distraction, https://www.youtube.com/watch?v=XToWVxS_9lA [last accessed May 2020]

Amit, R. and C. Zott (2001). Value creation in E-business. Strategic Management Journal 22, 6–7: 493–520.

Amit, R. and C. Zott (2012). Creating value through business model innovation. MIT Sloan Management Review 53, 3.

Ammon, T. and A. Brem (2013): Digitale Ökosysteme und deren Geschäftsmodelle: Analyse und Implikationen für klassische Buchverlage. In Keuper et al. (Hrsg.): Digitalisierung und Innovation. Planung, Entstehung, Entwicklungsperspektiven. Wiesbaden: Springer Gabler, Seite. 93 – 121.

Andrienko, G., Andrienko, N., & Fuchs, G. (2016). Understanding movement data quality. *Journal of location Based services, 10(*1), 31-46.

Appelrath, H.-J., Terzidis, O. and C. Weinhardt (2012). Internet of Energy - ICT as a Key Technology for the Energy System of the Future. Business & Information Systems Engineering 4, 1: 1–2.

Apple (2017). CarPlay. https://www.apple.com/ios/carplay/available-models/ [last accessed May 2020]

Associated Press (2017). Fiat Chrysler, Google Begin Offering Rides in Self-driving Cars. VOA. https://www.voanews.com/a/fiat-chrysler-self-driving-cars/3824581.html [last accessed May 2020]

Association of Information Systems (2017). *About AIS*. http://aisnet.org/ [last accessed May 2020]

Athanasopoulou, A., Bouwman, H., Nikayin, F. and M. de Reuver (2016). "The disruptive impact of digitalisation on the automotive ecosystem: a research agenda on business models, platforms and consumer issues." In *Proceedings of 29th Bled eConference*, pp.597-604.

Athanasopoulou, A., de Reuver, M., Nikou, S. and H. Bouwman (2019). What technology enabled services impact business models in the automotive industry? An exploratory study. *Futures*, 109, 73-83.

Attard, J. and R. Brennan (2018). Challenges in value-driven data governance. In OTM Confederated International Conferences "On the Move to Meaningful Internet Systems" (pp. 546-554). Springer, Cham.

Audi (2017a). Audi Connect. http://www.audi.com/en/innovation/connect.html [last accessed May 2020]

Audi (2017b). Audi connect and infotainment. https://audi-illustrated.com/en/CES-2016/Audi-connect-und-Infotainment [last accessed 2017, not available any more, information to be found at https://www.audi.com/en/experience-audi/models-and-technology/digital-services.html]

AutoMat (2016a). About AutoMat. https://www.automat-project.eu/content/about-automat [last accessed May 2020]

AutoMat (2016b). Objectives. www.automat-project.eu/content/objectives [last accessed May 2020]

AutoMat (2016c). http://automat.atosresearch.eu/sites/default/files/automat/public/content-files/articles/Automat-D2%203_Public%20System%20Concept_0.pdf [last accessed May 2020]

Automat (2017). http://automat-project.eu/ [last accessed May 2020]

Automat (2018a). Automat: Connected car data - the unexcavated treasure. https://www.youtube.com/watch?v=uRjvnahJ-9o [last accessed May 2020]

AutoMat (2018b). Deliverable D2.5 Cyber Security Framework: 4.1 Security and Privacy Engineering Objectives, RFC6690, https://goo.gl/a51DAs [last accessed May 2020]

Automat (2018c). Deliverable D5.3 Full Prototype of Cross-Sectorial Vehicle Data Services http://automat-project.eu/sites/default/files/automat/public/content-files/articles/AutoMat%20D5%203_Full%20Prototype%20of%20Cross-Sectorial%20Vehicle%20Data%20Services_final.pdf [last accessed May 2020]

Automat (2018d). Deliverable D6.7 Value Chain, Value Network, Economic model and Marketplace P&L simulation. https://www.automat-project.eu/sites/default/files/automat/public/content-files/articles/AutoMat-D6.7%20Push%20and%20effectual%20business%20model_final.pdf [last accessed September 2020]

Automatic (2017). https://www.automatic.com/ [last accessed May 2020. Service termination with May 2020, probably no longer accessible afterwards]

Automatic (2018). “Automatic: Connect Your Car to Your Digital Life.” URL 1: https://automatic.com/ and URL 2: https://developer.automatic.com/ [last accessed May 2020. Service termination with May 2020, probably no longer accessible afterwards]

Automatic Twitter (2020). https://twitter.com/automatic/status/1257341497550266368 [last accessed October 2020]

Axelsson, J., Papatheocharous, E. and J. Andersson (2014) “Characteristics of software ecosystems for federated embedded systems: A case study,” *Information and Software Technology*, vol. 56, no. 11, pp. 1457–1475.

Azkan, C., Iggena, L., Gür, I., Möller, F. and B. Otto (2020). A Taxonomy for Data-Driven Ser-vices in Manufacturing Industries. In *PACIS* (p. 184).

Åkerman, M., Lundgren, C., Bärring, M., Folkesson, M., Berggren, V., Stahre, J., Engström, U. and M. Friis (2018) “Challenges Building a Data Value Chain to Enable Data-Driven Decisions: A Predictive Maintenance Case in 5G-Enabled Manufacturing.” *Procedia Manufacturing*, 17, 411-418. (2018).

Baecke P. and L. Bocca (2017). The value of vehicle telematics data in insurance risk selection processes. *Decision Support Systems*, 98, 69-79.

Bain & Company (2017). “Changing Gears 2020 – How digital is transforming the face of the automotive industry.” http://www2.bain.com/bainweb/PDFs/BAIN_REPORT_Changing_Gears_2020.pdf [last accessed May 2020]

Banks, A. and R. Gupta (2015). “MQTT Version 3.1.1 Plus Errata 01,” OASIS standard, Dec. 2015. http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/mqtt-v3.1.1.html [last accessed May 2020]

Banski H. and J. Faenger (2017). “Mobile electronic device integration with in-vehicle information systems”. Google Patents. https://www.google.com/patents/US9589533 [last accessed May 2020]

Baskerville, R., Baiyere, A., Gregor, S., Hevner, A., and M. Rossi (2018). Design Science Re-search Contributions: Finding a Balance between Artifact and Theory. *Journal of the Associa-tion for Information Systems (19*:5), pp. 358–376.

Batini, C., Rula, A., Scannapieco, M. and G. Viscusi (2015). From data quality to big data quality. *Journal of Database Management*, *26*(1), 60-82.

Batini, C., Castelli, M., Viscusi, G., Cappiello, C. and C. Francalanci (2018). Digital Information Asset Evaluation: A Case Study in Manufacturing. *SIGMIS Database*, *49*(3), 19–33. https://doi.org/10.1145/3242734.3242737

Bauer, I., Zavolokina, L., & Schwabe, G. (2019). Is there a market for trusted car data?. *Electronic Markets*, 1-15.

Bätz, A., Gimpel, H., Heger, S., and M. Wöhl (2020). Driving Sustainably–The Influence of IoT-based Eco-Feedback on Driving Behavior. In *Proceedings of the 53rd Hawaii International Conference on System Sciences*.

Bekker, H. (2020). 2019 (Full Year) International: Worldwide Car Sales. https://www.best-selling-cars.com/international/2019-full-year-international-worldwide-car-sales/ [last accessed August 2020]

Bera, P., Soffer, P., & Parsons, J. (2019). Using Eye Tracking to Expose Cognitive Processes in Understanding Conceptual Models. *MIS Quarterly, 43*(4), 1105-1126.

Berggren, C., and T. Magnusson (2012). Reducing automotive emissions—The potentials of combustion engine technologies and the power of policy. *Energy Policy*, *41*, 636-643.

Berkley, B. J. and A. Gupta (1994). Improving service quality with information technology. *International Journal of Information Management*, *14*(2), 109–121. https://doi.org/https://doi.org/10.1016/0268-4012(94)90030-2

Bernstein, D., Vidovic, N. and S. Modi (2010) A cloud PAAS for high scale, function and velocity mobile applications. Proceedings of 2010 Fifth International Conference on Systems and Networks Communications

Bertoncello, M., Camplone, G., Gao, P., Kaas, H. W., Mohr, D., Möller, T. and D. Wee (2016). Monetizing car data—new service business opportunities to create new customer benefits. McKinsey & Company.

Bertoncello, M., Camplone, G., Husain, A. and D. Möller (2018). FROM BUZZ TO BUCKS – AUTOMOTIVE PLAYERS ON THE HIGHWAY TO CAR DATA MONETIZATION. McKinsey & Company.

Beverungen, D., Lüttenberg, H., and V. Wolf (2018). Recombinant service systems engineering. *Business & Information Systems Engineering*, *60*(5), 377-391.

Beverungen, D., Müller, O., Matzner, M., Mendling, J. and J. Vom Brocke (2019). Conceptualizing smart service systems. *Electronic Markets*, 29(1), 7-18.

Bharadwaj, A., El Sawy, O., Pavlou, P. and N. Venkatraman (2013). Digital Business Strategy: Toward A Next Generation Of Insights. *MIS Quarterly*, *37*(2), 471–482.

Birman, K. and T. Joseph (1987). “Exploiting virtual synchrony in distributed systems,” SIGOPS Oper. Syst. Rev., vol. 21, no. 5, pp. 123–138. http://doi.acm.org/10.1145/37499.37515

BMW (2017a). BMW i Ventures. https://www.bmwiventures.com/portfolio-3 [last accessed May 2020]

BMW (2017b). BMW Cardata.

BMW (2017c). BMW Connected Drive. https://www.bmw.com/en/topics/fascination-bmw/connected-drive/overview.html [last accessed May 2020]

BMW (2019). Contract signed: BMW Group and Daimler AG launch long-term development cooperation for automated driving. https://www.press.bmwgroup.com/global/article/detail/T0298266EN/contract-signed:-bmw-group-and-daimler-ag-launch-long-term-development-cooperation-for-automated-driving?language=en [last accessed Sept. 2020]

BMW (2020). BMW and MINI CarData. https://www.bmwgroup.com/en/innovation/technologies-and-mobility/cardata.html [last accessed May 2020]

BMW (i) ConnectedDrive (2017). https://www.bmw-connecteddrive.at/app/at/index.html#/portal/store [last accessed May 2020]

BMW Group (2014). BMW i Ventures announces strategic investment in Zendrive. Retrieved August 14, 2017 from https://www.press.bmwgroup.com/usa/article/detail/T0196902EN_US/bmw-i-ventures-announces-strategic-investment-in-zendrive?language=en_US [last accessed May 2020]

bmwi (2020). https://www.bmwi.de/Redaktion/DE/Textsammlungen/Branchenfokus/Industrie/branchenfokus-automobilindustrie.html [last accessed July 2020]

Boons, F. and F. Lüdeke-Freund (2013). Business models for sustainable innovation: state-of-the-art and steps towards a research agenda. Journal of Cleaner Production 45: 9–19. https://doi.org/https://doi.org/10.1016/j.jclepro.2012.07.007

Bosch (2020a). Predictive road condition services. https://www.bosch-mobility-solutions.com/en/products-and-services/passenger-cars-and-light-commercial-vehicles/automated-driving/predictive-road-condition-services/ [last accessed September 2020]

Bosch (2020b). Connected Horizon. https://www.bosch-mobility-solutions.com/en/products-and-services/passenger-cars-and-light-commercial-vehicles/connectivity-solutions/connected-horizon/ [last accessed September 2020]

Brandt T. (2013). “Information Systems in Automobiles - Past, Present, and Future Uses.” *In Proceedings of AMCIS 2013*: Association for Information Systems. doi: 10.13140/RG.2.1.3337.9688

Brennan, R., Attard, J., Petkov, P., Nagle, T. and M. Helfert (2019). Exploring Data Value Assessment: A Survey Method and Investigation of the Perceived Relative Importance of Data Value Dimensions. In *Proceedings of the 21st International Conference on Enterprise Information Systems - Volume 1: ICEIS* (pp. 200–207). https://doi.org/10.5220/0007723402000207

Brews, P.J. and C.L. Tucci (2007). The structural and performance effects of internetworking. *Long Range Planning 40*, 2: 223–243.

Bridgelall, R., Lu, P., Tolliver, D. D. and T. Xu (2018). Mining Connected Vehicle Data for Beneficial Patterns in Dubai Taxi Operations. *Journal of Advanced Transportation*, *2018*.

Briscoe, G. and C. Mulligan (2014). Digital innovation: the hackathon phenomenon. Creativeworks, London, working paper No. 6

Briscoe, G. and P. De Wilde (2006). “Digital Ecosystems: Evolving service-oriented architectures.” In: Conf. on Bio Inspired Models of Network, Information and Computing Systems. IEEE Press.

Broggi, A., Cerri, P., Ghidoni, S., Grisleri, P. and H.G. Jung (2009). A New Approach to Urban Pedestrian Detection for Automatic Braking. *IEEE Trans. Intelligent Transportation Systems*, 10, 594-605.

Broy, M. (2006). Challenges in automotive software engineering. In *Proceedings of the 28th international conference on Software engineering* (pp. 33-42).

BVDW (2016). Connected Cars – ein Diskussionspapier zum Thema Services des BVDW–Bundesverband Digitale Wirtschaft: www.bvdw.org/presseserver/ConnectedCars/Final-version_Diskussionspapier_Services_15.06.pdf [last accessed July 2016, not available any more]

Cäsar, M., Mehl, R., Tschödrich, S., et al. (2019). Connected Vehicle Trend Radar: The drive to add value: lessons from China. https://www.capgemini.com/wp-content/uploads/2019/08/Connected-Vehicle-Trend-Radar.pdf [last accessed October 2020]

Cäsar, M., Mehl, R., Tschödrich, S., et al. (2020). Connected Vehicle Trend Radar 2: The road towards profitability for automotive connected services. Capgemini Invent. https://www.capgemini.com/wp-content/uploads/2020/09/CapgeminiInvent_VehicleDataMonetization_POV_Sep2020.pdf [last accessed May 2021]

Camden, M.C., Hickman, J.S. and R.J. Ha (2015). *United States Department of Transportation: Federal Motor Carrier Safety Administration*. https://bit.ly/2YpbD9H [last accessed May 2020]

Carlucci, D., Marr, B. and G. Schiuma (2004). The knowledge value chain: how intellectual capital impacts on business performance. *International Journal of Technology Management, 27*(6-7), 575-590.

Caruso (2020). "What we Offer" and "Our Data Suppliers". https://www.caruso-dataplace.com/ [last accessed July 2020]

Casilari, E., Santoyo-Ramón, J.A. and J.M. Cano-García (2016). Analysis of a Smartphone-Based Architecture with Multiple Mobility Sensors for Fall Detection. PLOS ONE 11(12)

Cavalieri, S., & Pezzotta, G. (2012). Product–Service Systems Engineering: State of the art and research challenges. *Computers in industry*, *63*(4), 278-288.

Ceccagnoli, M., Forman, C., Huang, P., and D.J. Wu (2012). "Cocreation of value in a platform ecosystem: the case of enterprise software." MIS Q. 36, 1 (March 2012), 263-290.

Chesbrough, H. and J. Spohrer (2006). A research manifesto for services science. *Communications of the ACM*, *49*(7), 35–40.

Childerhouse, P., Hermiz, R., Mason-Jones, R., Popp, A., & Towill, D. R. (2003). Information flow in automotive supply chains–identifying and learning to overcome barriers to change. *Industrial Management & Data Systems*.

Choe, E.K., Lee, N.B., Lee, B., Pratt, W. and J.A. Kientz (2014). Understanding quantified-selfers' practices in collecting and exploring personal data. CHI 2014, One of a CHInd. Toronto, Canada

Christ, M., Braun, N., Neuffer, J. and A.W. Kempa-Liehr (2018). Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests. Neurocomputing 307, 72–77

Cichy, P., Salge, T.O. and R. Kohli (2014). Extending the privacy calculus: the role of psychological ownership. ICIS 2014, AIS, Auckland

Conboy, K., Dennehy, D. and M. O'Connor (2020). 'Big time': An examination of temporal complexity and business value in analytics. *Information & Management, 57*(1), 103077. https://doi.org/https://doi.org/10.1016/j.im.2018.05.010.

Congram, C., and M. Epelman (1995). How to describe your service. *International Journal of Service Industry Management*. 6 (2), pp. 6 – 23.

C-ROADS (2017). Detailed pilot overview report, https://www.c-roads.eu/fileadmin/user_upload/media/Dokumente/Detailed_pilot_overview_report_v1.0.pdf [last accessed May 2020]

CSS electronics (2018). CAN DBC FILE – CONVERT DATA IN REAL TIME, https://www.csselectronics.com/screen/page/dbc-database-can-bus-conversion-wireshark-j1939-example/language/en [last accessed May 2020]

Curry, E. (2016). The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches. In New Horizons for a Data-Driven Economy (pp. 29-37). Springer International Publishing.

Curry, E., Ngonga, A., Domingue, J., Freitas, A., Strohbach, M. and T. Becker (2014). D2.2.2. Final version of the technical white paper. Public deliverable of the EU-Project BIG (318062; ICT-2011.4.4).

D'Orazio, T., Leo, M., Guaragnella, C. and A. Distante (2007). A visual approach for driver inattention detection. Pattern Recognition 40(8), 2341–2355

Daimler (2020a). Digitalisation. https://www.daimler.com/innovation/digitalisation/ [last accessed May 2020]

Daimler (2020b). So far – and yet so near. https://www.daimler.com/innovation/case/connectivity/connected-with-the-vehicle.html [last accessed May 2020]

Dash (2017). https://dash.by/ [last accessed May 2020]

Davenport, T. H. and J. G. Harris (2007). Competing on Analytics: The new Science of Winning, *Harv. Bus. Rev*., vol. 84, no. 1, pp. 98–107, 2007.

Davenport, T. (2014). *Big data at work: dispelling the myths, uncovering the opportunities*. Harvard Business Review Press.

Davenport, T. and J. Harris (2017). *Competing on analytics: Updated, with a new introduction: The new science of winning*. Harvard Business Press.

Davenport, J., Pacheco, P. and A. Priestley (2020). Market Trends: Monetizing Connected and Autonomous Vehicle Data. Gartner Information Technology Research

De Arriba-Pérez, F., Caeiro-Rodríguez, M., and J.M. Santos-Gago (2016). Collection and Processing of Data from Wrist Wearable Devices in Heterogeneous and Multiple-User Scenarios. Sensors 16(9)

De Moya, J.-F. and J. Pallud (2017). "QUANTIFIED SELF: A LITERATURE REVIEW BASED ON THE FUNNEL PARADIGM", *In Proceedings of the 25th European Conference on Information Systems (ECIS)*, Guimarães, Portugal, pp. 1678-1694.

De Winter, J.C.F., Dodou, D., Happee, R. and Y.B. Eisma (2019). Will vehicle data be shared to address the how, where, and who of traffic accidents?. *European Journal of Futures Research*, *7*(1), 1-9.

Dedrick, J. (2010). Green IS: Concepts and issues for information systems research. Communications of the Association for Information Systems, 27, 173–184.

Delbosc, A., and K. Ralph (2017). A tale of two millennials. *Journal of Transport and Land Use*, 10(1), 903-910.

Deloitte (2017). "Automotive Data Treasure – Vehicle digitalization and the question of data treasures." https://www2.deloitte.com/content/dam/Deloitte/de/Documents/risk/EN-Risk-Advisory-Risk-Automotive-Data-Treasure-Frage-nach-dem-Datenschatz-2017_s.pdf [last accessed May 2020]

DeLone, W.H. and E.R. McLean (1992). "Information systems success: The quest for the dependent variable." *Journal of Information Systems Research*, 3(1), 60–95.

DeLone, W.H. and E.R. McLean (2003). "The DeLone and McLean Model of Information Systems Success: A Ten-Year Update", *Journal of Management Information Systems* 19 (4), 9-30.

Demchenko, Y., Grosso, P., de Laat, C. and P. Membrey (2013). Addressing big data issues in Scientific Data Infrastructure. 2013 International Conference on Collaboration Technologies and Systems (CTS), San Diego, CA, pp. 48-55. doi: 10.1109/CTS.2013.6567203

Dhungana, D., Engelbrecht, G., Parreira, J.X., Schuster, A., Tobler, R. and D. Valerio (2016). Data-driven ecosystems in smart cities: A living example from Seestadt Aspern. In *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)* (pp. 82-87). IEEE.

Dickover, M. E., McGowan, C. L., & Ross, D. T. (1977, January). Software design using: SADT. In *Proceedings of the 1977 annual conference* (pp. 125-133).

Dong, J., Liu, C. and Z. Lin (2014). Charging infrastructure planning for promoting battery electric vehicles: An activity-based approach using multiday travel data, *In Transportation Research Part C: Emerging Technologies* 38, 44-55

Dong, Y., Hu, Z., Uchimura, K. and N. Murayama (2011). Driver Inattention Monitoring System for Intelligent Vehicles: A Review, IEEE Transactions on Intelligent Transportation Systems 12(2), 596–614 (2011)

Dotson, K. (2017). Daimler and LBBW issue $114M corporate bond using blockchain, SiliconAngle

Dremel, C., Herterich, M., Wulf, J., Waizmann, J.-C., and W. Brenner (2017). How AUDI AG Established Big Data Analytics in Its Digital Transformation, *MIS Quarterly Executive (16*:2), pp. 81-100.

Dremel, C., Herterich, M.M., Wulf, J. and J. Vom Brocke (2020). Actualizing big data analytics affordances: A revelatory case study. *Information & Management 57*(1):103121.

Dremel, C., Wulf, J., Maier, A. and W. Brenner (2017b). "Understanding the Value and Organizational Implications of Big Data Analytics – The Case of AUDI AG" In: Proceedings of the Thirty eight International Conference on Information Systems (ICIS 2017), as well as Journal of Infor-ation Technology Teaching Cases 2018, 8(2), 126-138.

eeNews Automotive (2017). http://www.eenewsautomotive.com/news/german-car-industry-plans-close-obd-interface?news_id=93237 [last accessed May 2020]

Eisenmann, T.R., Parker, G. and M.W. Van Alstyne (2006). Strategies for Two-Sided Markets. Harvard Business Review October: 1–12. https://doi.org/10.1007/s00199-006-0114-6

Eisenmann, T., Parker, G. and M. Van Alstyne (2011). Platform envelopment. Strategic Management Journal 32, 12: 1270–1285. https://doi.org/10.1002/smj.935

Eskenazi, M., Morency, L.-P., Black, A.W. and D. Yulun (2018). Detection of Driver Distraction, http://ppms.cit.cmu.edu/media/project_files/UTC_project_13_Multimodal_Detection_of_Driver_Distraction_-_final_report.pdf [last accessed May 2020]

EU (2013). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32013R0886 [last accessed May 2019]

EU (2018). Study on good practices for reducing road safety risks caused by road user distractions, https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/pdf/behavior/distraction_study.pdf [last accessed May 2020]

European Commission (2017). Digital Transformation Monitor: The race for automotive data. https://ec.europa.eu/growth/tools-databases/dem/monitor/sites/default/files/DTM_The%20race%20for%20automotive%20data%20v1.pdf [last accessed October 2020]

European Commission (2018a). *Annual Accident Report 2018.* https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/pdf/statistics/dacota/asr2018.pdf [last accessed May 2020]

European Commission (2018b). Data protection in the EU. 2018, https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en [last accessed May 2020]

European Commission (2019). EU Road Safety Policy Framework 2021–2030 – Next steps towards 'Vision Zero', https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/move-2019-01178-01-00-en-tra-00_0.pdf [last accessed July 2020]

European Commission (2020). A European strategy for data, https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf [last accessed August 2020]

Evans, D.S., Hagiu, A. and R. Schmalensee (2006). "Software platforms." Industrial Organization and the Digital Economy, 31.

EVOLVE (2019). https://www.evolve-h2020.eu/ [last accessed May 2020]

Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, web of science, and Google scholar: strengths and weaknesses. *The FASEB journal, 22*(2), 338-342

Farahani, P., Meier, C., & Wilke, J. (2017). Digital supply chain management agenda for the automotive supplier industry. In *Shaping the digital enterprise* (pp. 157-172). Springer, Cham.

Faroukhi, A. Z., El Alaoui, I., Gahi, Y., & Amine, A. (2020). Big data monetization throughout Big Data Value Chain: a comprehensive review. *Journal of Big Data, 7*(1), 1-22.

FV (Fachverband der Fahrzeugindustrie Österreichs) (2020). "Beschäftigte rund um das Automobil." https://www.fahrzeugindustrie.at/fileadmin/content/Zahlen___Fakten/Statistikjahrbuch/Seite5.7.pdf [last accessed July 2020]

FCA (2017a). u-connect. https://www.driveuconnect.eu/en/features/controls [last accessed May 2020]

FCA (2017b). FCA Dealer Digital Programme. http://www.fcadealerdigital.com/ [last accessed August 2017, not available any more]

Fink, A., Schneidereit, G., & Voß, S. (2006). *Grundlagen der Wirtschaftsinformatik*. Springer-Verlag.

Fishbein, M. and I. Ajzen (1975). Belief, attitude, intention, and behavior: an introduction to theory and research, *Addison-Wesley series in social psychology*. Addison-Wesley Pub. Co, Reading, Mass.

Fletcher, L. and A. Zelinsky (2009). Driver Inattention Detection based on Eye Gaze—Road Event Correlation. The International Journal of Robotics Research 28(6), 774–801

Foss, N.J. and T. Saebi (2016). Fifteen Years of Research on Business Model Innovation. Journal of Management 43, 1: 200–227. https://doi.org/10.1177/0149206316675927

Frécon, E. (2018). *Machinery*, http://www.docker-machinery.com/ [last accessed May 2020]

FreshBooks (2018). “Connect FreshBooks & Your Car | Automatic Works With FreshBooks.” https://automatic.com/apps/freshbooks/ [last accessed May 2020]

Fulthorpe, M. (2015). Five critical challenges facing the automotive industry. http://cdn.ihs.com/www/pdf/AUT-TL-WhitePaper-5.pdf [last accessed August 2020]

Furtado, L., Dutra, M., & Macedo, D. (2017). Value creation in big data scenarios: A literature survey. *Journal of Industrial Integration and Management, 2*(01), 1750002.

Gabler Wirtschaftslexikon (2020). https://wirtschaftslexikon.gabler.de/definition/wirtschaftsinformatik-51980 [last accessed October 2020]

Gao, P., Kaas, H.-W., Mohr, D. and D. Wee (2016). *Disruptive trends that will transform the auto industry*. McKinsey Report. https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/disruptive-trends-that-will-transform-the-auto-industry [last accessed May 2020]

Garcia, M.L. and O.H. Bray (1997). *Fundamentals of technology roadmapping* (No. SAND--97-0665). Sandia National Labs., Albuquerque, NM (United States).

Gartner Research (2019). www.gartner.com/technology/topics/digital-ecosystems.jsp [last accessed May 2020, not available any more]

Gellersen, H.W., Schmidt, A. and M. Beigl (2002). Multi-sensor context-awareness in mobile devices and smart artifacts. Mobile Networks and Applications 7(5), 341–351

Gerloff, C. and C. Cleophas (2017). “Excavating the Treasure of IoT Data: An Architecture to Empower Rapid Data Analytics for Predictive Maintenance of Connected Vehicles.” In: Proceedings of the Thirty eight International Conference on Information Systems (ICIS 2017).

Gertner, M. I. (2013). The value chain and value creation. *Advances in management, 6*(10).

Gestrata (2009). http://www.gestrata.at/publikationen/archiv-journal-beitrage/gestrata-journal-124/gehweg-radweg-und-strasenerhaltung-in-graz [last accessed May 2020]

Giang, W.C., Shanti, I., Chen, H.Y.W., Zhou, A. and B. Donmez (2015). Smartwatches vs. smartphones: A preliminary report of driver behavior and perceived risk while responding to notifications. In: Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 154–161. Nottingham

Gissler, A. (2015). Connected vehicle: Succeeding with a disruptive technology. *Accenture. http://www.automotivebusiness.com.br/abinteligencia/pdf/Accenture-digital-Connected-Vehicle.pdf* [last accessed May 2020]

Gioia, D.A., Corley, K.G. and A.L. Hamilton (2013). "Seeking qualitative rigor in inductive research: Notes on the Gioia methodology." *Organisational research methods* (16:1), pp.15-31.

Gordijn, J., Petit, M. and R. Wieringa (2006). "Understanding business strategies of networked value constellations using goal- and value modeling." In *Proceedings of 14th IEEE International Requirements Engineering Conference*, pp. 129-138.

Gregor S. and A.R. Hevner (2013). Positioning and presenting design research for maximum im-pact, *MIS Quarterly Vol. 37* No. 2, pp. 337-355, June 2013.

Grover, V., Chiang, R.H.L., Liang, T.-P. and D. Zhang (2018). Creating Strategic Business Val-ue from Big Data Analytics: A Research Framework. Journal of Management Information Systems, 35(2), 388–423. https://doi.org/10.1080/07421222.2018.1451951.

Gusenbauer, M. and N.R. Haddaway (2020). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research synthesis methods*, *11*(2), 181-217.b

Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M. and F. Feldberg (2017). Debating Big Data: A Literature Review on Realizing Value from Big Data, *The Journal of Strategic Information Systems (26*:3), pp. 191–209.

Habeck, A., Newman, J., Bertoncello, M., Kässer, M., Weig, F., Hehensteiger, M., Hölz, J., Plattfaut, R., Wegner, C., Guminski, M. and Z. Yan (2014). Connected Car, automotive value chain unbound.

Haeberle, T., Charissis, L., Fehling, C., Nahm, J. and F. Leymann (2015). The connected car in the cloud: a platform for prototyping telematics services. IEEE Softw 32(6):11-17

Haghighatkhah, A., Banijamali, A., Pakanen, O. P., Oivo, M., and P. Kuvaja (2017). Automotive software engineering: A systematic mapping study. *Journal of Systems and Software, 128*, 25-55.

Hanelt, A., Piccinini, E., Gregory, R.W., Hildebrandt, B. and L.M. Kolbe (2015). "Digital Transformation of Primarily Physical Industries - Exploring the Impact of Digital Trends on Business Models of Automobile Manufacturers." In *Proceedings of Wirtschaftsinformatik 2015*, pp.1313-1327.

Hartmann, P. M., Zaki, M., Feldmann, N. and A. Neely (2016). Capturing value from big data – a taxonomy of data-driven business models used by start-up firms. *International Journal of Operations & Production Management*, *36*(10), 1382–1406. https://doi.org/10.1108/IJOPM-02-2014-0098

Haydin, V. (2020). Vehicle Data Monetization: Challenging Yet Promising. https://www.intellias.com/vehicle-data-monetization-challenging-yet-promising/ [last accessed October 2020]

Henfridsson, O., Mathiassen, L. and F. Svahn (2009). Reconfiguring Modularity: Closing Capability Gaps in Digital Innovation. All Sprouts Content. 277.

Hevner, A.R., March, S.T., Park, J. and S. Ram (2004). Design science in information systems research. *MIS quarterly* (2004): 75-105.

Hevner AR. (2007). A three cycle view of design science research. *Scandinavian journal of information systems*. 2007 Jan 1;19(2):4.

Hilkert, D., Wolf, C. M., Benlian, A. and T. Hess (2010). The “as-a-service”-paradigm and its implications for the software industry–insights from a comparative case study in CRM software ecosystems. *In International Conference of Software Business* (pp. 125-137). Springer, Berlin, Heidelberg.

Hirsch, D.D. (2013). The glass house effect: Big Data, the new oil, and the power of analogy. *Maine Law Review*, 66, 373.

Hirschheim R., and H.K. Klein (2012). „A Glorious and Not-So-Short History of the Information Systems Field.", *Journal of the Association for Information Systems*, Volume 13, Issue 4, pp. 188-235.

Homfeldt, F., Rese, A., & Simon, F. (2019). Suppliers versus start-ups: Where do better innovation ideas come from?. *Research policy, 48*(7), 1738-1757.

Honda myHonda (2017). http://www.honda.de/cars/services/my-honda/my-honda.html [last accessed May 2020]

Hood, J., Hoda, O. and R. Robinson (2019). Monetizing data in the age of connected vehicles. Deloitte Insights. https://www2.deloitte.com/us/en/insights/industry/automotive/monetizing-data-connected-vehicles.html [last accessed September 2020]

Horberry, T., Anderson, J., Regan, M. A., Triggs, T. J. and J. Brown (2006). Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. Accident Analysis & Prevention 38(1), 185–191

Huberty, M. (2015). Awaiting the second big data revolution: from digital noise to value creation. *Journal of Industry, Competition and Trade, 15*(1), 35-47.

Husnjak, S., Perakovi, D., Forenbacher, I. and M. Mumdziev (2015). Telematics system in usage based motor insurance, vol. 100, no. January. Elsevier Ltd, 2015, pp. 816–825, conference of 25th DAAAM International Symposium on Intelligent Manufacturing and Automation, DAAAM 2014.

IBM (2017). IBM Integrates with BMW CarData to Enable New and Innovative Services for Drivers. News Releases. www-03.ibm.com/press/us/en/pressrelease/52595.wss [last accessed May 2020]

IBM CAI (Center for Applied Insights) (2015). *Digital disruption and the future of the automotive industry: Mapping new routes for customer-centric connected journeys.* https://www-935.ibm.com/services/multimedia/IBMCAI-Digital-disruption-in-automotive.pdf [last accessed November 2017, not available any more]

IFTTT (2016). https://ifttt.com/ [last accessed May 2020]

Immonen, A., Ovaska, E. and J. Kalaoja (2015). "A service requirements engineering method for a digital service ecosystem." Service Oriented Computing and Applications 10(2), pp. 151–172

Immonen, A., Ovaska, E. and T. Paaso (2017). Towards certified open data in digital service ecosystems. Software Quality Journal (26), 1257–1297 (2018) https://doi.org/10.1007/s11219-017-9378-2

Immonen, A., Palviainen, M. and E. Ovaska (2014). Requirements of an open data based business ecosystem. *IEEE access*, *2*, 88-103.

INRIX (2016). www.inrix.com [last accessed May 2020]

Internet Security Research Group (ISRG) (2019). *Let's Encrypt*. https://letsencrypt.org/ [last accessed May 2020]

ISO (2015). "Road vehicles – controller area network (CAN) – Part 1: Data link layer and physical signalling," ISO Standard 11898-1:2015, Dec. 2015. https://www.iso.org/standard/63648.html [last accessed May 2020]

ISO (2016). "Road vehicles diagnostic communication over controller area network (DoCAN) Part 4: Requirements for emissions-related systems," ISO Standard 15765-4:2016, Apr. 2016. https://www.iso.org/standard/67245.html [last accessed May 2020]

ISO (2017). https://www.iso.org/committee/5383568.html [last accessed May 2020]

Jacobides, M.G., Cennamo, C. and A. Gawer (2018). Towards a theory of ecosystems. *Strategic Management Journal*, 39(8), 2255–2276.

Jennings, C., Shelby, Z., Arkko, J., Keranen, A. and C. Bormann (2018). *Sensor Measurement Lists (SenML) draft-ietf-core-senml-16.* https://tools.ietf.org/html/draft-ietf-core-senml-16 [last accessed May 2020]

Jogulu, U. D., & Pansiri, J. (2011). Mixed methods: A research design for management doctoral dissertations. *Management research review*.

Ju, M. and J. Mou (2018). Privacy as a Commodity Is Not the Case: Privacy Calculus Model for Connected Cars. In *WHICEB 2018 Proceedings* (p. 44).

Kaiser, C., Festl, A., Stocker, A., Pucher, G. and M. Fellmann (2018a). "The Vehicle Data Value Chain as a Lightweight Model to Describe Digital Vehicle Services." In: Proceedings of European Transport Conference 2018 (ETC 2018). https://aetransport.org/past-etc-papers/conference-papers-2018?abstractId=5800&state=b [last accessed February 2021]

Kaiser, C., Festl, A., Stocker, A., Pucher, G. and M. Fellmann (2019a). The Vehicle Data Value Chain as a Lightweight Model to Describe Digital Vehicle Services. In *Proceedings of the 15th International Conference on Web Information Systems and Technologies* (WEBIST 2019), p. 68-79.

Kaiser, C., Steger, M., Dorri, A., Festl, A., Stocker, A., Fellmann, M. and S. Kanhere (2019c). "Towards a Privacy-Preserving Way of Vehicle Data Sharing – A Case for Blockchain Technology?" In: Dubbert J., Müller B., Meyer G. (eds) Advanced Microsystems for Automotive Applications 2018. AMAA 2018. *Lecture Notes in Mobility*. Springer, Cham, https://doi.org/10.1007/978-3-319-99762-9_10

Kaiser, C., Stocker, A. and G. Viscusi (2017a). Digital vehicle ecosystems and new business models: An overview of digitalization perspectives. i-KNOW '17 October, 11-12.

Kaiser, C., Stocker, A. and M. Fellmann (2019b). Understanding Data-driven Service Ecosystems in the Automotive Domain. In *Proceedings of 25th Americas Conference on Information Systems (AMCIS 2019)*.

Kaiser, C., Stocker, A., Festl, A., Lechner, G. and M. Fellmann (2018b). A research agenda for vehicle information systems. In *Proceedings of European Conference on Information Systems* (ECIS 2018).

Kaiser, C., Stocker, A., Viscusi, G., Festl, A., Moertl, P. and M. Glitzner (2017b). Quantified cars: an exploration of the position of ICT start-ups vs. car manufacturers towards digital car services and sustainable business models. In *Proceedings of 2nd international conference on new business models (NBM)* (pp. 336-350).

Kaiser, C. Festl, A. and A. Stocker (2020a). Nutzung von Automotive Big Data zur Entwicklung zweier neuer Anwendungen. In *BITKOM position paper on Practical Use Cases of Artificial Intelligence & Big Data in Industry (original german title: Konkrete Anwendungsfällevon KI & Big-Data in der Industrie),* pages 10-19. Link: https://bit.ly/2VsqTkv

Kaiser, C., Festl, A., Pucher, G., Fellmann, M. and A. Stocker (2020b). Digital Services Based on Vehicle Usage Data: The Underlying Vehicle Data Value Chain. In: Bozzon A., Domínguez Mayo J.F., Filipe J. (eds) Web Information Systems and Technologies. WEBIST 2019. *Lecture Notes in Business Information Processing (LNBIP), vol 399.* Springer, Cham. https://doi.org/10.1007/978-3-030-61750-9_2

Kaiser, C.; Stocker, A.; Festl, A.; Djokic Petrovic, M.; Papatheocharous, E.; Wallberg, A.; Ezquerro, G.; Ortigosa Orbe, J.; Szilagyi, T.; and Fellmann, M. (2020c). A Vehicle Telematics Service for Driving Style Detection: Implementation and Privacy Challenges. In *Proceedings of the 6th International Conference on Vehicle Technology and Intelligent Transport Systems - Volume 1 (VEHITS)*, ISBN 978-989-758-419-0, p.29-36. DOI: 10.5220/0009329400290036

Kaiser, C., Stocker, A., Viscusi, G., Fellmann, M. and Richter, A. (2021). Conceptualizing Value Creation in Data-driven Services: The Case of Vehicle Data. *International Journal of Information Management (IJIM), Volume 59 (2021). https://doi.org/10.1016/j.ijinfomgt.2021.102335*

Kallinikos, J., Aaltonen, A. and A. Marton (2013). The ambivalent ontology of digital artifacts. MIS Quarterly 37, 2: 357–370.

Kane (2020). Volkswagen ID.3 Software Problem: 'It's No Longer A Laughing Matter' https://insideevs.com/news/406375/vw-id3-software-problem-update/ [last accessed September 2020]

Kaplinsky, R. and M. Morris (2000). *A handbook for value chain research (Vol. 113)*. Brighton: University of Sussex, Institute of Development Studies.

Karlsson, P.-H., Ahn, H.K. and B. Choi (2016). Connected CAR. Retrieved from https://www.ipsos.com/sites/default/files/2016-06/022.1-connected-car-a-new-ecosystem.pdf [last accessed May 2020]

Karnouskos, S. and O. Terzidis (2007). Towards an information infrastructure for the future Internet of energy. In Communication in Distributed Systems (KiVS) 2007 ITG-GI Conference, 1–6.

Kemler, R. (2017). "Die automotive Zukunft Österreichs". https://www.automotive.at/kfz-wirtschaft/die-automotive-zukunft-oesterreichs-151915 [last accessed July 2020]

Kessler, T. and C. Buck (2017). "How digitization affects mobility and the business models of automotive OEMs." In: Phantom Ex Machina (pp. 107-118). Springer, Cham.

Kilbride, J. (2017). Secure Payments "On The Go" With Blockchain Technology From ZF, UBS and IBM

Kitsios, F., Papachristos, N. and M. Kamariotou (2017). Business models for open data ecosystem: challenges and motivations for entrepreneurship and innovation. In *2017 IEEE 19th Conference on Business Informatics (CBI)* (Vol. 1, pp. 398-407). IEEE.

Klauer, S., Dingus, T., Neale, V., Sudweeks, J. and D. Ramsey (2006). The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data. Virginia Tech Transportation Institute, Blacksburg

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B. et al. (2016). Jupyter Notebooks – a publishing format for reproducible computational workflows. In: Proceedings of the 20th International Conference on Electronic Publishing. IOS Press, pp. 87–90. http://dx.doi.org/10.3233/978-1-61499-649-1-87

Knauss, E. and D.E. Damian (2014). "Towards Enabling Cross-Organisational Modeling in Automotive Ecosystems." In *Proceedings of MD2P2@ MoDELS*, pp.38-47.

Koch, M., Wang, H., Bürgel, R and T. Bäck (2020). Towards Data-driven Services in Vehicles. In VEHITS (pp. 45-52).

Kohl, C., Mostafa, D., Böhm, M. and H. Krcmar (2017). "Disruption of Individual Mobility Ahead? A Longitudinal Study of Risk and Benefit Perceptions of Self-Driving Cars on Twitter." In Proceedings of 13. Internationale Tagung Wirtschaftsinformatik (WI 2017), p. 1220-1234.

Kolarova, V., Kuhnimhof, T., & Trommer, S. (2017). Assessment of real-world vehicle data from electric vehicles–potentials and challenges. In Proceedings of 11th International Conference on Transport Survey Methods ISCTSC.

Kolbe, L., Nastjuk, I. and S.T.-N. Trang (2015). „The Influence of Sociodemographic and Technology-Associated Factors on Stress from Human Interaction with Electric Vehicle Information Systems." *In Proceedings of AMCIS* 2015. Association for Information Systems.

Kolbe, N., Kubler, S., Robert, J., Le Traon, Y. and A. Zaslavsky (2017). "Towards semantic interoperability in an open IoT ecosystem for connected vehicle services." In *Proceedings of 2017 Global Internet of Things Summit (GIoTS)*, pp. 1-5.

Kong, X., Song, X., Xia, F., Guo, H., Wang, J., & Tolba, A. (2018). LoTAD: Long-term traffic anomaly detection based on crowdsourced bus trajectory data. *World Wide Web*, 21(3), 825-847.

KPMG (2016). "Global Automotive Executive Survey 2016". https://assets.kpmg/content/dam/kpmg/dk/pdf/DK-2018/02/kpmg-global-automotive-executive-survey-2016.pdf [last accessed October 2020]

KPMG (2017). "Global Automotive Executive Survey 2017." http://kpmg.com/gaes2017 [last accessed May 2020]

KPMG (2018). "Global Automotive Executive Survey 2018." http://kpmg.com/gaes2018 [last accessed May 2020]

KPMG (2020). "Global Automotive Executive Survey 2020." https://automotive-institute.kpmg.de/GAES2020/ [last accessed October 2020]

KPMG N.V. (2017). The Connected Car is here to Stay. KPMG NV. https://assets.kpmg.com/content/dam/kpmg/nl/pdf/2017/sector/automotive/the-connected-car-is-here-to-stay.pdf [last accessed October 2020]

Krzanich. B. (2016). Data is the new oil in the future of automated driving. https://newsroom.intel.com/editorials/krzanich-the-future-of-automated-driving/ [last accessed May 2020]

Kuhnert, F., Stürmer, C. and A. Koster (2018). Five trends transforming the Automotive Industry. PricewaterhouseCoopers GmbH Wirtschaftsprüfungsgesellschaft: Berlin, Germany, 1(1), 1-48.

Kuhnimhof, T., Armoogum, J., Buehler, R., Dargay, J., Denstadli, J. M., and T. Yamamoto (2012). Men shape a downward trend in car use among young adults—evidence from six industrialized countries. *Transport Reviews*, 32(6), 761-779.

Kumar, V., Chattaraman, V., Neghina, C., Skiera, B., Aksoy, L., Buoye, A. and J. Henseler (2013). Data-driven services marketing in a connected world. *Journal of Service Management*, *24*(3), 330–352. https://doi.org/10.1108/09564231311327021

Kung, L.-C. and C.-C. Lin (2018). “Autonomous vehicle services or ride-sharing services? A game theoretic investigation” In: Proceedings of Twenty-second Pacific Asia Conference on Information Systems (PACIS 2018).

Kuschel, J. and B. Dahlbom (2007). Mobile Services for Vehicles. *In Proceedings of ECIS 2007,* pp. 1863-1874.

Kuschel, J. (2008). "The vehicle ecosystem." In *IFIP Working Conference on Open IT-Based Innovation: Moving Towards Cooperative IT Transfer and Knowledge Diffusion*, pp. 309-322. Springer, Boston.

Latif, A., Saeed, A.U., Hoefler, P., Stocker, A. and C. Wagner (2009). The Linked Data Value Chain: A Lightweight Model for Business Engineers. In A. Paschke, H. Weigand, W. Behrendt, K. Tochtermann, T. Pellegrini (eds.), Proceedings of I-Semantics 2009. 5th International Conference on Semantic Systems (p./pp. 568--577), Journal of Universal Computer Science.

Laubis, K., Viliam S. and A. Schuller (2016). “Road condition measurement and assessment: A crowd based sensing approach.” In: Proceedings of the Thirty seventh International Conference on Information Systems (ICIS 2016).

Laudon K.C. and J.P. Laudon (2013). *Management Information Systems: Managing the Digital Firm.* 13th Edition. Boston: Pearson.

Lechner, G., Fellmann, M., Festl, A., Kaiser, C., Kalayci, T.E., Spitzer, M., and A. Stocker (2019). A Lightweight Framework for Multi-device Integration and Multi-sensor Fusion to Explore Driver Distraction. In *International Conference on Advanced Information Systems Engineering* (pp. 80-95). Springer, Cham.

Lehrer, C., Wieneke, A., vom Brocke, J., Jung, R. and S. Seidel (2018). How Big Data Analytics Enables Service Innovation: Materiality, Affordance, and the Individualization of Service. *Journal of Management Information Systems, 35*(2), 424–460.

Li, W.C., Nirei, M. and K. Yamana (2019). Value of data: there’s no such thing as a free lunch in the digital economy. *US Bureau of Economic Analysis Working Paper*, Washington, DC.

Lim, C., Kim, K. H., Kim, M. J., Heo, J. Y., Kim, K. J., and P.P. Maglio (2018). From data to value: A nine-factor framework for data-based value creation in information-intensive services. *International Journal of Information Management*, *39*, 121-135.

Lioutas, E. D. and C. Charatsari (2020). Big data in agriculture: Does the new oil lead to sustainability?. *Geoforum*, 109, 1-3.

Liu, L., Karatas, C., Li, H. and S. Tan et al. (2015). Toward Detection of Unsafe Driving with Wearables. In: Proceedings of the 2015 workshop on Wearable Systems and Applications, pp. 27–32. Florence

Lowe D.G. (1999). "Object recognition from local scale-invariant features," *In Proceedings of the Seventh IEEE International Conference on Computer Vision*. pp. 1150-1157.

Lüdeke-Freund, F., Massa, L., Bocken, N., Brent, A. and J. Musango (2016). Business models for shared value: Main Report. Network for Business Sustainability South Africa.

Lusch, R.F. and S. Nambisan (2015). “Service innovation: A service-dominant logic perspective.” MIS Quarterly 39(1), 155-175

Lyytinen, K., Yoo, Y. and R.J. Boland Jr (2016). Digital product innovation within four classes of innovation networks. Information Systems Journal 26, 1: 47–75. https://doi.org/10.1111/isj.12093

Maass, W., Parsons, J., Purao, S., Storey, V. C. and C. Woo (2018). Data-driven meets theory-driven research in the era of big data: opportunities and challenges for information systems research. *Journal of the Association for Information Systems, 19*(12), 1.

Maglio, P. P., and J. Spohrer (2008). Fundamentals of service science. *Journal of the academy of marketing science*, *36*(1), 18-20.

Malhotra, A., Melville, N.P. and R.T. Watson (2013). Spurring Impactful Research on Information Systems for Environmental Sustainability. MIS Quarterly 37, 4: 1265–1274.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., and A. H. Byers (2011). *Big data: The next frontier for innovation, competition* (Vol. 5, No. 6). and productivity. Technical report, McKinsey Global Institute.

Marabelli M., Hansen, S., Newell, S. and C. Frigerio (2017). "The Light and Dark Side of the Black Box: Sensor-based Technology in the Automotive Industry." *Communications of the Association for Information Systems* 40, Article 16.

March S.T. and G.F. Smith (1995). Design and natural science research on information technology. *Decision support systems*. 1995 Dec 1;15(4):251-66.

Martínez de Aragón, B., Alonso-Zarate, J. and A. Laya (2018). “How connectivity is transforming the automotive ecosystem.” *Internet Technology Letters* (1:e14).

Massa, L. and C.L. Tucci (2014). Business Model Innovation. In The Oxford Handbook of Innovation Management. https://doi.org/0.1093/oxfordhb/9780199694945.013.002

Massa, L., Tucci, C.L. and A. Afuah (2017). A Critical Assessment of Business Model Research. Academy of Management Annals, 11(1), 73–104. https://doi.org/https://doi.org/10.5465/annals.2014.0072

Matt, C., Hess, T. and A. Benlian (2015): Digital Transformation Strategies, Business Information Systems Engineering 57(5):339–343 (2015).

Mayer-Schönberger, V. and K. Cukier (2013). Big Data: A Revolution That Will Transform How We Live, Work, and Think. Boston: Houghton Mifflin Harcourt. ISBN: 0544002695 9780544002692

Mayr, S. and A. Slavik (2020). "Das ist nicht mehr zum Lachen" https://www.sueddeutsche.de/auto/bmw-daimler-vw-software-id3-1.4856930 [last accessed September 2020]

McAfee, A. and E. Brynjolfsson (2012). Big Data: The Management Revolution. *Harvard Business Review*, 90, 60--68.

McIntyre, A. (2013). Market Trends: Enter the Wearable Electronics Market With Products for the Quantified Self. Gartner Report 2013. https://www.gartner.com/doc/2537715/market-trends-enter-wearable-electronics [last accessed May 2020]

McKinney, W. (2010). Data Structures for Statistical Computing in Python. In: Proceedings of the 9th Python in Science Conference, pp. 51–56. Austin

McKinsey & Company (2016). “Automotive revolution - perspective towards 2030: How the convergence of disruptive technology-driven trends could transform the auto industry” https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/disruptive-trends-that-will-transform-the-auto-industry/de-de [last accessed May 2020]

McKinsey & Company (2017). "The automotive revolution is speeding up – perspectives on the emerging personal mobility landscape." https://mck.co/2VVQuSP [last accessed May 2020]

medianet (2020). https://medianet.at/news/automotive-business/oesterreichs-autoindustrie-setzt-43-mrd-euro-um-19756.html# [last accessed July 2020]

Mercedes (2017a). Mercedes Me. https://www.mercedes-benz.com/en/mercedes-me/ [last accessed May 2020]

Mercedes (2017b). Mercedes Me Connect. https://www.mercedes-benz.com/en/mercedes-me/connectivity/ [last accessed May 2020]

Merkel, D. (2014). "Docker: Lightweight linux containers for consistent development and deployment," Linux J., vol. 2014, no. 239, Mar. 2014. http://dl.acm.org/citation.cfm?id=2600239.2600241

Merkel, D. (2014). Docker: Lightweight Linux Containers for Consistent Development and Deployment. *Linux Journal*, Bd. 239, pp. 1075-3583, 2014.

Mikalef, P., Pappas, I. O., Krogstie, J. and P.A. Pavlou (2020). Big data and business analytics: A research agenda for realizing business value. *Information & Management, 57*(1), 103237. https://doi.org/https://doi.org/10.1016/j.im.2019.103237.

Mikalef, P., Pappas, I. O., Krogstie, J. and M. Giannakos (2017). Big Data Analytics Capabilities: A Systematic Literature Review and Research Agenda, *Information Systems and E-Business Management*, pp. 1-32. (https://doi.org/10.1007/s10257-017-0362-y).

Mikusz, M. and T. Herter (2016). "How Do Consumers Evaluate Value Propositions of Connected Car Services?" In: Proceedings of Americas Conference on Information Systems (AMCIS 2016).

Miller, H.G. and P. Mork (2013). "From data to decisions: a value chain for big data." *It Professional*, (1), 57-59.

Mocker, M. and N. Fonstad (2017b). "Driving Digitization at Audi." In: Proceedings of the Thirty eight International Conference on Information Systems (ICIS 2017).

Mocker, M. and N.O. Fonstad (2017.) How AUDI AG Is Driving Toward the Sharing Economy, *MIS Quarterly Executive (16*:4), pp. 279-293.

Mojio (2016). www.moj.io [last accessed May 2020]

Müller, P. (2013). *Integrated engineering of products and services: Layer-based development methodology for product-service systems* (Doctoral dissertation). Fraunhofer Verlag, 2014, ISBN: 3-8396-0549-0

Murphy, C. (2016). *Shared mobility and the transformation of public transit* (No. TCRP J-11/TASK 21).

MyCar MyData (2016). http://mycarmydata.eu [last accessed May 2020]

Nachira, F., Dini, P. and A. Nicolai (2007). A network of digital business ecosystems for Europe: roots, processes and perspectives. *European Commission, Bruxelles, Introductory Paper*, *106*.

Nakamoto, S. (2008) Bitcoin: a peer-to-peer electronic cash system, Whitepaper http://www.bitcoin.org/bitcoin.pdf [last accessed May 2020]

Nambisan, S. (2013). Information Technology and Product / Service Innovation: A Brief Assessment and Some Suggestions for Future Research. Journal of the Association for Information Systems 14: 215–226.

Nambisan, S. (2016). Digital Entrepreneurship: Toward a Digital Technology Perspective of Entrepreneurship. Entrepreneurship Theory and Practice

Nambisan, S. (2016). Digital Entrepreneurship: Toward a Digital Technology Perspective of Entrepreneurship. Entrepreneurship Theory and Practice: n/a-n/a. https://doi.org/10.1111/etap.12254

Nastjuk I., M. Marrone and L.M. Kolbe (2016). "Less is Sometimes More - The Impact of In-Vehicle Information Systems on Perceived Range Stress." *In Proceedings of ICIS 2016. Association for Information Systems*.

Nauto (2017). Nauto | Safe driving made simple. https://www.nauto.com/ [last accessed May 2020]

Neff, G. and D. Nafus (2016). Self-Tracking.

Neumann G., Sobernig, S. and M. Aram (2014). "Evolutionary Business Information Systems - Perspectives and Challenges of an Emerging Class of Information Systems". *Business & Information Systems Engineering*, 6, 33-38.

New York City DOT (2016). www.nyc.gov/html/dot/html/home/home.shtml [last accessed May 2020]

NHTSA (2018a). Risky Driving, https://www.nhtsa.gov/risky-driving/distracted-driving. [last accessed May 2020]

NHTSA (2018b). Research Note on Distracted Driving, https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812517 [last accessed May 2020]

Niederman, F. and S. March (2019). The "Theoretical lens" concept: We all know what it means, but do we all know the same thing?. *Communications of the Association for Information Systems, 44*(1), 1.

Nischak, F., Hanelt, A. and L.M. Kolbe (2017). "Unraveling the Interaction of Information Systems and Ecosystems - A Comprehensive Classification of Literature" In: Proceedings of the Thirty eight International Conference on Information Systems (ICIS 2017).

Nischak, F. amd A. Hanelt (2019). Ecosystem Change in the Era of Digital Innovation–A Longitudinal Analysis and Visualization of the Automotive Ecosystem. In *Proceedings of International Conference on Information Systems (ICIS 2019)*.

Normann, R. and R. Ramìrez (1994). Designing interactive strategy: from value chain to value constellation. Wiley, Chichester.

Nunamaker J.F., and R.O. Briggs (2012). "Towards a Broader Vision for Information Systems." *ACM Transactions on Management Information Systems*, Vol. 2, No. 4, Article 20.

Nunamaker J. F., Briggs, R. O., Derrick, D. C. and G. Schwabe (2015). The last research mile: Achieving both rigor and relevance in information systems research. *Journal of management information systems, 32*(3), 10-47.

O'Reilly, T. (2005). What is web 2.0. O'Reilly Media

Oliphant, T.E. (2006). Guide to NumPy. 2nd edn. Trelgol Publishing

Oliva, R. and Kallenberg, R. (2003), "Managing the transition from products to services", *International Journal of Service Industry Management, Vol. 14*(2), pp. 160-172. https://doi.org/10.1108/09564230310474138

Opel OnStar (2017). http://www.opel.at/onstar/onstar.html [last accessed May 2020]

Osterwalder, A. and Y. Pigneur (2010). Business Model Generation. John Wiley & Sons.

P. McFedries (2013). “Tracking the quantified self,” IEEE Spectrum, vol. 8, no. 50.

Pagani, M. (2013). “Digital Business Strategy and Value Creation: Framing the Dynamic Cycle of Control Points.” MIS Quarterly 37(2), pp. 617–632.

Pandian, P.S., Mohanavelu, K., Safeer, K.P., Kotresh, T.M., Shakunthala, D.T., Gopal, P. and V.C. Padaki (2008). Smart Vest: Wearable multi-parameter remote physiological monitoring system. Medical Engineering & Physics 30(4), 466–477 (2008)

Papatheocharous, E., Frécon, E., Kaiser, C., Festl, A. and A. Stocker (2018). Towards a Generic IoT Platform for Data-driven Vehicle Services. In *IEEE International Conference on Vehicular Electronics and Safety*, Bd. doi: 10.1109/ICVES.2018.8519505.

Parkbob (2016). www.parkbob.com [last accessed May 2020]

Parker, G.G., Van Alstyne, M.W. and S.P. Choudary (2016). Platform revolution: How networked markets are transforming the economy-and how to make them work for you. WW Norton & Company.

Parvinen, P., Pöyry, E., Gustafsson, R., Laitila, M. and M. Rossi (2020). Advancing Data Monetization and the Creation of Data-based Business Models. *Communications of the Association for Information Systems*, 47, https://doi.org/10.17705/1CAIS.04702

Payre, W., Cestac, J., Dang, N.-T., Vienne, F. and P. Delhomme (2017). Impact of training and in-vehicle task performance on manual control recovery in an automated car. Transportation Research Part F: Traffic Psychology and Behaviour 46(A), 216–227

Pedregosa, F., Varoquaux, G., Gramfort, A. and V. Michel et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12, 2825–2830

Peng Y., Boyle, L.N. and J.D. Lee (2014). “Reading, typing, and driving: How interactions with in-vehicle systems degrade driving performance.” *Transportation Research Part F: Traffic Psychology and Behaviour* 27(Part A), 182-191.

Phaal, R., Farrukh, C.J. and D.R. Probert (2004). “Technology roadmapping—a planning framework for evolution and revolution”. *Technological forecasting and social change*, *71*(1-2), 5-26.

Piccinini, E., Hanelt, A., Gregory, R. and L.M. Kolbe (2015). "Transforming industrial business: the impact of digital transformation on automotive organisations." In *Proceedings of Thirty Sixth International Conference on Information Systems (ICIS 2015)*.

Pilinkienė, V. and P. Mačiulis (2014). “Comparison of different ecosystem analogies: The main eco-nomic determinants and levels of impact.” Procedia - Social and Behav. Sciences 156, 365-370.

Pillmann, J., Wietfeld, C., Zarcula, A., Raugust, T. and D. C. Alonso (2017a). “Novel common vehicle information model (CVIM) for future automotive vehicle big data marketplaces.” In *Intelligent Vehicles Symposium (IV), 2017 IEEE* (pp. 1910-1915). IEEE.

Pillmann, J., Sliwa, B., Schmutzler, J., Ide, C. and C. Wietfeld (2017b). Car-To-Cloud Communication Traffic Analysis Based on the Common Vehicle Information Model. In IEEE

Vehicular Technology Conference (VTC-Spring) Workshop on Wireless Access Technologies and Architectures for Internet of Things (IoT) Applications. Presentation slides: http://automat-project.eu/sites/default/files/automat/public/content-files/articles/VTC_Presentation.pdf [last accessed May 2020]

Piparsania V. (2019). How Do Automakers Generate Additional Revenues from Connected Vehicles? https://www.counterpointresearch.com/automakers-generate-additional-revenues-connected-vehicles/ [last accessed October 2020]

Porsche Connect (2017). https://www.porsche.com/germany/connect/ [last accessed May 2020]

Porter, M.E. and J.E. Heppelmann (2014). How smart, connected products are transforming competition, Harvard Business Manager, November 2014.

Porter, M.E. and J.E. Heppelmann (2015). How smart, connected products are transforming companies, Harvard Business Manager, December 2015.

Porter, M.E. and V.E. Millar (1985). How information gives you competitive advantage.

Probst, L., Pedersen, B. and O.-K. Lonkeu (for European Commission): Digital Transformation Monitor (2017). *Digital Transformation Monitor: The race for automotive data.* https://ec.europa.eu/growth/tools-databases/dem/monitor/sites/default/files/DTM_The%20race%20for%20automotive%20data%20v1.pdf [last accessed September 2020]

Pütz, F., Murphy, F., Mullins, M. and L. O'Malley (2019). Connected automated vehicles and insurance: Analysing future market-structure from a business ecosystem perspective. *Technology in Society*, *59*, 101182.

PWC (2016). The wearable life 2.0. Connected living in a wearable world. PWC Report

Rai, A., and V. Sambamurthy (2006). Editorial Notes—The Growth of Interest in Services Management: Opportunities for Information Systems Scholars. *Information Systems Research*, *17*(4), 327–331.

Ramos, F.B.A., Lorayne, A., Costa, A.A.M. and R.R. de Sousa et al. (2016). Combining Smartphone and Smartwatch Sensor Data in Activity Recognition Approaches: An Experimental Evaluation. In: Proceedings of the 28th International Conference on Software Engineering and Knowledge Engineering SEKE 2016, pp. 267–272. Redwood City

Ramsey M. (2017). *Creating Innovations in the Automotive and Smart Mobility Industry Primer for 2017*. Gartner Report by Gartner Inc.

Rappa, M.A. (2001). Business models on the web: Managing the digital enterprise. Retrieved from digitalenterprise.org/models/models.html

Rappa, M.A. (2004). The utility business model and the future of computing services. IBM Systems Journal 43, 1: 32–42. https://doi.org/10.1147/sj.431.0032

Recon Instruments (2016). http://reconinstruments.com [last accessed May 2020]

Reddy, S. K. and W. Reinartz (2017). Digital transformation and value creation: Sea change ahead. *Marketing Intelligence Review, 9*(1), 10-17.

Regan, M.A., Hallett, C. and C.P. Gordon (2011). Driver distraction and driver inattention: Definition, relationship and taxonomy. Accident Analysis & Prevention 43(5), 1771–1781

Rehm S.-V., Faber, A. and L. Goel (2017). "Visualizing Platform Hubs of Smart City Mobility Business Ecosystems." *In Proceedings of ICIS 2017. Association for Information Systems.*

Reich, A., Krämer, N. A., & Lenninger, R. (2018). Fahrzeugdaten-Management Standardisierter Zugang als Basis für neue Geschäftsmodelle. *ATZelektronik, 13*(2), 42-47.

Remane, G., Hildebrandt, B., Hanelt, A. and L.M. Kolbe (2016). "Discovering New Digital Business Model Types-a Study of Technology Startups from the Mobility Sector." In *Proceedings of Pacific Asia Conference on Information Systems (PACIS 2016)*, p. 289

Reuters (2019). "Volkswagen CEO expects software to make up 90 percent of auto industry innovation." https://www.reuters.com/article/us-volkswagen-results-software/volkswagen-ceo-expects-software-to-make-up-90-percent-of-auto-industry-innovation-idUSKBN1QT1ZN [last accessed July 2020]

Riasanow, T., Galic, G. and M. Böhm (2017). "Digital Transformation in the Automotive Industry: Towards a Generic Value Network." In: Proceedings of European Conference on Information Systems (ECIS 2017).

Richter, T., Ruhl, S., Ortlepp, J. and E. Bakaba (2016). Prevention of Overtaking Accidents on Two-lane Rural Roads. *In Transportation Research Procedia*, Volume 14, 2016, 4140-4149

Rizk, A., Bergvall-Kåreborn, B., and A. Elragal (2018). Towards a Taxonomy for Data-Driven Digital Services. In Proceedings of the 51st Hawaii International Conference on System Sciences.

Ross, D.T. (1977). Structured Analysis (SA): A Language for Communicating Ideas. *IEEE Transactions on Software Engineering*, SE-3,16-34.

Rossi, L., Walker, J. and M. Musolesi (2015). Spatio-temporal techniques for user identification by means of GPS mobility data. *EPJ Data Science* 4.

Rosten, E. and T. Drummond (2006). "Machine Learning for High-Speed Corner Detection." In: Leonardis A., Bischof H., Pinz A. (eds) *Computer Vision – ECCV 2006. Lecture Notes in Computer Science*, vol 3951. Springer, Berlin, Heidelberg

Runtastic (2015). Adidas Group acquires Runtastic: www.adidas-group.com/en/media/news-archive/press-releases/2015/adidas-group-acquires-runtastic [last accessed May 2020]

Runtastic (2016). https://www.runtastic.com/mediacenter/corporate-assets/german/companyoverview/20160405_corporate_overview_de.pdf [last accessed July 2016, not available any more]

Runtastic (2017a). https://www.runtastic.com/en [last accessed May 2019]

Runtastic (2017b). https://www.runtastic.com/en/mediacenter/press-releases/20170503_de_office-opening-runtastic [last accessed May 2019, not available any more]

Runtastic (2020). https://www.runtastic.com/career/facts-about-runtastic/ [last accessed 2020/01/19]

Rusitschka, S. and E. Curry (2016). Big Data in the Energy and Transport Sectors. In New Horizons for a Data-Driven Economy (pp. 225-244). Springer International Publishing.

Russel, J. (2018). BMW, GM, Ford and Renault launch blockchain research group for automotive industry, Techcrunch, May, 2018.

Ryder B., Gahr, B. and A. Dahlinger (2016). "An in-Vehicle Information System Providing Accident hotspot Warnings." *In Proceedings of ECIS 2016*. Prototypes. 3.

Ryder, B., Gahr, B., Egolf, P., Dahlinger, A. and F. Wortmann (2017). "Preventing traffic accidents with in-vehicle decision support systems-The impact of accident hotspot warnings on driver behaviour." *Decision Support Systems*, 99, 64-74.

SAE (2018). SAE J1939 Standards Collection on the Web: Content, https://www.sae.org/standardsdev/groundvehicle/j1939a.htm [last accessed May 2020]

Sakaria, N., and N. Stehfest (2013). *Millennials and mobility: Understanding the millennial mindset and new opportunities for transit providers* (No. Task 17, TCRP Project J-11).

Santos L., Coutinho-Rodrigues, J.M. and C.H. Antunes (2011). A web spatial decision support system for vehicle routing using Google Maps. *Decision Support Systems*, 51(1), 1-9.

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H. and C. Jinks (2018). Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality & quantity*. 2018 Jul 1;52(4):1893-907.

Schaltegger, S., Lüdeke-Freund, F. and E.G. Hansen (2016). Business Models for Sustainability. Organization & Environment 29, 3: 264–289. https://doi.org/10.1177/1086026616633272

Scholte, S.S., van Teeffelen, A.J. and P.H. Verburg (2015). "Integrating socio-cultural perspectives into ecosystem service valuation: a review of concepts and methods." *Ecological economics*, 114, 67-78.

Schulze, E.-D., Beck, E., and K. Müller-Hohenstein (2005). "Plant Ecology." Berlin: Springer. ISBN 3-540-20833-X.

Schüritz, R., Farrell, K., Wixom, B. W. and G. Satzger (2019). Value Co-Creation in Data-Driven Services: Towards a Deeper Understanding of the Joint Sphere. In *Proceedings of International Conference on Information Systems (ICIS 2019)*.

Schüritz, R. and G. Satzger (2016). Patterns of data-infused business model innovation. In Business Informatics (CBI), 2016 IEEE 18th Conference on (Vol. 1, pp. 133–142). IEEE.

Schüritz, R., Seebacher, S. and R. Dorner (2017). Capturing Value from Data: Revenue Models for Data-Driven Services. In *Proceedings of the 50th Hawaii International Conference on System Sciences*.

Seiberth, G. and W. Gründinger (2018). Data-driven business models in connected cars, mobility services & beyond. *BVDW Research, 1*, 18.

Shelby, Z. (2012), "Constrained RESTful Environments (CoRE) Link Format," RFC6690, https://tools.ietf.org/html/rfc6690 [last accessed May 2020]

SherpaShare (2018). "Connect SherpaShare & Your Car | Automatic Works With SherpaShare." https://automatic.com/apps/sherpashare/ [last accessed May 2020]

Shipilov, A. (2016). Why Aren't Automakers Embracing Digital Business Models? Harvard Business Review.

Shoaib, M., Bosch, S., Scholten, H., Havinga, P.J. and O.D. Incel (2015). Towards detection of bad habits by fusing smartphone and smartwatch sensors. In: Proceedings of the 2015

IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops), pp. 591–596. St. Louis

SiriusXM (2020). Automatic Labs Update. https://listenercare.siriusxm.com/prweb/PRRestService/seo/km/help/SupportCenter/KC-2203 [last accessed November 2020]

Soley, A.M., Siegel, J.E., Suo, D. and S.E. Sarma (2018). Value in vehicles: economic assessment of automotive data. *Digital Policy, Regulation and Governance* 20(6), https://doi.org/10.1108/DPRG-05-2018-0025

Sonnenberg, C. and J. vom Brocke (2011). Evaluation Patterns for Design Science Research Artefacts, in *Practical Aspects of Design Science*, Springer Berlin Heidelberg, pp. 71-83. (https://doi.org/10.1007/978-3-642-33681-2_7).

Sorescu, A. (2017). Data-Driven Business Model Innovation. *Journal of Product Innovation Management*, *34*(5), 691–696. https://doi.org/10.1111/jpim.12398

Spiekermann, M. (2019). Data Marketplaces: Trends and Monetisation of Data Goods. *Intereconomics*, *54*(4), 208-216.

Spohrer, J., Anderson, L., Pass, N., Ager, T. and D. Gruhl (2008b) Service Science. *Journal of Grid Computing, Special Issue on Grid Economics and Business Models 6,* 313–324 (2008). https://doi.org/10.1007/s10723-007-9096-2

Spohrer, J., Anderson, L., Pass, N., and T. Ager (2008a). Service science and SD logic. In *Proceedings of the Otago Forum* (pp. 214-237).

Spohrer, J. and P.P. Maglio (2008). The Emergence of Service Science: Toward Systematic Service Innovations to Accelerate Co-Creation of Value. *Production and Operations Management*, *17*(3), 238–246. https://doi.org/10.3401/poms.1080.0027

Stähler, P. (2002). Geschäftsmodelle in der digitalen Ökonomie, EUL Verlag, 2. Auflage.

Stanton N.A., and M.S. Young (2010). Vehicle automation and driving performance, Ergonomics, 41:7, 1014-1028.

Starepravo I. (2019). From Device-Connected to Car-Connected: Unlock the Full Power of Connected Car Commerce. https://medium.com/datadriveninvestor/from-device-connected-to-car-connected-unlock-the-full-power-of-connected-car-commerce-deb0fcdfbef0 [last accessed September 2020]

Statista (2019). Connected Car Report 2019. https://de.statista.com/statistik/studie/id/43051/dokument/connected-car-2019/ [last accessed October 2020]

Statista (2020). https://de.statista.com/statistik/daten/studie/30703/umfrage/beschaeftigtenzahl-in-der-automobilindustrie/ [last accessed July 2020]

Stevens G., Bossauer, P., Jakobi, T. and C. Pakusch (2017). „Second Dashboard: Information Demands in a Connected Car." *In Proceedings of Mensch und Computer 2017 – Tagungsband, 215-226.*

Stocker, A., Kaiser, C. and M. Fellmann (2017a). "Quantified Vehicles: Novel Services for Vehicle Lifecycle Data." *Business & Information Systems Engineering* 59(2), 125-130. https://doi.org/10.1007/s12599-017-0465-5

Stocker, A. and C. Kaiser (2016). Quantified car: potentials, business models and digital ecosystems, *e & i Elektrotechnik und Informationstechnik*, vol. 133(7), 334-340.

Stocker, A., Kaiser, C. and A. Festl (2017b). “Automotive Sensor Data. An Example Dataset from the AEGIS Big Data Project,” Jun. 2017, research dataset of the AEGIS project. https://doi.org/10.5281/zenodo.820576

Stocker, A., Tochtermann, K. and P. Scheir (2010). Die Wertschöpfungskette der Daten. Eine Basis für zukünftige wirtschaftswissenschaftliche Betrachtungen des Web of Data, HMD Praxis der Wirtschaftsinformatik, October 2010, Volume 47, Issue 5, pp 94–104.

Stocker, A., Lechner, G., Kaiser, C. and M. Fellmann (2021). Digitalized Mobility. Currently under review (March 2021) by 27th Americas Conference on Information Systems (AMCIS).

Strategy Engineers and fka (Forschungsgesellschaft Kraftfahrwesen mbH, Aachen) (2017). *Monetizing automotive data: A perspective on how data drives revenues and growth in future automotive business models*. White Paper. https://www.fka.de/images/publikationen/2017/data-as-product-SE17.pdf [last accessed May 2020]

Strava (2017). https://www.strava.com/?hl=en [last accessed May 2020]

Strukhoff, R., Khizhniak, A. and C. Gutierrez (2017). “Mobile Devices Are Propelling Industrial IoT Scenarios”. https://www.altoros.com/blog/mobile-devices-are-propelling-industrial-iot-scenarios/ [last accessed July 2020]

Subramaniam, M., Iyer, B. and V. Venkatraman (2019). Competing in digital ecosystems. *Business Horizons, 62*(1), 83-94.

Sun, W., Liu, J. and H. Zhang (2017). When smart wearables meet intelligent vehicles: challenges and future directions. *IEEE wireless communications*, 24(3), 58-65.

Sun, P., Bisschop, R., Niu, H. and X. Huang (2020). A review of battery fires in electric vehicles. *Fire technology*, 1-50.

Surbakti, F. P. S., Wang, W., Indulska, M. and S. Sadiq (2020). Factors influencing effective use of big data: A research framework. *Information & Management, 57*(1), 103146. https://doi.org/https://doi.org/10.1016/j.im.2019.02.001.

Sussman, E.D., Bishop, H., Madnick, B. and R. Walters (1985). Driver Inattention and Highway Safety. Transportation Research Record 1047, 40–48

Suzuki Y., Montabon, F. and S.-H. Lu (2014). DSS of vehicle refueling: A new enhanced approach with fuel weight considerations. *Decision Support Systems*, 68, 15-25.

Svahn, F., Mathiassen, L. and R. Lindgren (2017). Embracing digital innovation in incumbent firms: How Volvo cars managed competing concerns. *MIS Quarterly 41*, 1: 239–253.

Swan, M. (2009). Emerging Patient-Driven Health Care Models: An Examination of Health Social Networks, Consumer Personalized Medicine and Quantified Self-Tracking, Int. J. Environ. Res. Public Health 2009, 6(2), pp 492-525; doi:10.3390/ijerph6020492

Swan, M. (2013). The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery, Big Data. June 2013, 1(2): 85-99.

Swan, M. (2015). “Connected car: Quantified self becomes quantified car,” Journal of Sensor and Actuator Networks, vol. 4, no. 1, pp. 2–29, 2015. http://www.mdpi.com/2224-2708/4/1/2

Szalavetz, A. (2019). Digitalisation, automation and upgrading in global value chains–factory economy actors versus lead companies. *Post-Communist Economies, 31*(5), 646-670.

Szilágyi, T. R. (2019). „Data Sharing von Fahrzeugnutzungsdaten – Akzeptanzfaktoren und Möglichkeiten der Anonymisierung“. Masterarbeit. Universität Rostock, Lehrstuhl für Wirtschaftsinformatik

TechCrunch (2017). SiriusXM acquires connected car company Automatic for over $100M. https://techcrunch.com/2017/04/27/siriusxm-acquires-connected-car-company-automatic-for-over-100m/?guccounter=1 [last accessed October 2020]

Teece, D. J. (2007). "Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance." *Strategic management journal* (28:13), pp.1319-1350.

Thomas, P., Morris, A., Talbot, R. and H. Fagerlind (2013). Identifying the causes of road crashes in Europe. Annals of advances in automotive medicine 57(13), 13–22

Tilson, D., Lyytinen, K. and C. Sørensen (2010). Digital Infrastructures: The Missing IS Research Agenda. Information Systems Research 21, 4: 748–759.

Toledo T., Musicant, O. and T. Lotan (2008). "In-vehicle data recorders for monitoring and feedback on drivers' behavior." *In Transportation Research Part C: Emerging Technologies* 16(3) 320-331.

Tomiyama, T., Lutters, E., Stark, R. and M. Abramovici (2019). Development capabilities for smart products. *CIRP Annals*, *68*(2), 727-750.

Tselentis, D. I., Yannis, G., & Vlahogianni, E. I. (2016). Innovative insurance schemes: pay as/how you drive. *Transportation Research Procedia*, 14, 362-371.

Turker, G.F. and A. Kutlu (2015). Methods of monitoring Vehicle's CAN data with mobile devices. Global Journal of Computer Sciences. 5(1), 36-42. doi: http://dx.doi.org/10.18844/gjcs.v5i1.31

Turner, V., Gantz, J.F., Reinsel, D. and S. Minton (2014). The digital universe of opportunities: rich data and the increasing value of the internet of things. Rep. from IDC EMC.

Vaia, G., Carmel, E., DeLone, W., Trautsch, H. and F. Menichetti (2012). "Vehicle telematics at an Italian insurer: new auto insurance products and a new industry ecosystem." MIS Quarterly Executive 11(3), 113-125.

Valasek, C. and C. Miller (2015). Remote Exploitation of an Unaltered Passenger Vehicle, White Paper, p. 93, 2015.

Van Ratingen, M., Williams, A., Lie, A., Seeck, A., Castaing, P., Kolke, R., Adriaenssens, G. and A. Miller (2016). The European new car assessment programme: a historical review. *Chinese journal of traumatology*, 19(2), 63-69.

Vargo, S.L. and R.F. Lusch (2004). "Evolving to a New Dominant Logic for Marketing." Journal of Marketing 68(1): 1-17.

Vargo, S. L. and R.F. Lusch (2008). Service-dominant logic: continuing the evolution. *Journal of the Academy of marketing Science*, *36*(1), 1-10.

VDA (2014). https://www.vda.de/en/topics/innovation-and-technology/network/data-protection-principles-for-connected-vehicles.html [last accessed May 2020]

VDA (2016). *Access to the vehicle (and vehicle generated data).* https://www.vda.de/en/topics/innovation-and-technology/network/access-to-the-vehicle.html [last accessed May 2020]

VDA (2017). Normung. https://www.vda.de/de/themen/sicherheit-und-standards/normung/normung.html [last accessed May 2020]

VDA (2017b). Protecting the vehicle - sharing data - creating added value – VDA. https://www.vda.de/en/topics/innovation-and-technology/data-security/protecting-the-vehicle [last accessed October 2020]

VDA (2017c). Access to the vehicle and vehicle generated data-"NEVADA Share and SecureConcept". https://www.vda.de/dam/vda/publications/2017/2017_10_24-Zusammenfassung-NEVADA-Share--Secure-English-translation_GS-clean/2017_10_24%20Zusammenfassung%20NEVADA-Share%20%20Secure%20English%20translation_GS%20clean.pdf [last accessed October 2020]

Vector (2007). DBC File Format Documentation. http://read.pudn.com/downloads766/ebook/3041455/DBC_File_Format_Documentation.pdf [last accessed January 2021]

Velamuri, V. K., Neyer, A. K., and K.M. Möslein (2011). Hybrid value creation: a systematic review of an evolving research area. *Journal für Betriebswirtschaft, 61*(1), 3-35.

Venable, J., Pries-Heje, J. and R. Baskerville (2016). FEDS: A Framework for Evaluation in Design Science Research, *European Journal of Information Systems (25*:1), pp. 77-89. (https://doi.org/10.1057/ejis.2014.36).

Venkataram, P.B. (2019). *Connected Vehicles-An Ecosystem for Services* (No. 2019-28-2447). SAE Technical Paper.

Venkatesh, V. and H. Bala (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision sciences, 39(2),* 273-315.

Venkatesh, V., Thong, J.Y. and X. Xu (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS quarterly*, 36(1), 157-178.

Viereckl, R., Ahlemann, D., Koster, A., Hirsh, E., Kuhnert, F., Mohs, J., Fischer, M., Gerling, W., Gnanasekaran, K., Kusber, J., Stephan, J., Crusius, D., Kerstan, H., Warnke, T., Schulte, M., Seyfferth, J. and E. H. Baker (2016). Connected car report 2016: Opportunities, risk, and turmoil on the road to autonomous vehicles. https://www.strategyand.pwc.com/gx/en/insights/2016/connected-car-2016/connected-car-report-2016.pdf [last accessed September 2020]

Vilarinho, T., Farshchian, B., Bajer, D.G. and O.H. Dahl et al. (2015). A Combined Smartphone and Smartwatch Fall Detection System. In: Proceedings of the 2015 IEEE International Conference on Computer and Information Technology, pp. 1443–1448. Liverpool

Vinli (2017). https://www.vin.li/ [last accessed May 2020]

Viscusi, G. (2015). Digital business innovation: roadmaps and attitudes from a Future Enterprise perspective. Génie Logiciel 115: 36–41.

Viscusi, G. and C. Batini (2014). Digital Information Asset Evaluation: Characteristics and Dimensions. In Smart Organizations and Smart Artifacts SE - 9 (Lecture Notes in Information Systems and Organisation), 77–86. https://doi.org/10.1007/978-3-319-07040-7_9

Viscusi, G. and C. Batini (2016). Information Production and Social Value for Public Policy: A Conceptual Modeling Perspective. Policy & Internet 8, 3: 334–353. https://doi.org/10.1002/poi3.121

Viscusi, G., Castelli, M. and C. Batini (2014). Assessing social value in open data initiatives: a framework. Future Internet 6, 3: 498–517. https://doi.org/10.3390/fi6030498

Volkswagen (2017). Volkswagen car-net. Retrieved August 14, 2017 from http://www.vwcar-netconnect.com/ [last accessed May 2020]

Volkswagen (2018). Volkswagen develops the largest digital ecosystem in the automotive industry https://www.volkswagenag.com/en/news/stories/2018/08/volkswagen-develops-the-largest-digital-ecosystem-in-the-automot.html [last accessed May 2020]

Volkswagen (2019). "Member of the Brand Board of Management Christian Senger explains the new Volkswagen software unit." https://www.volkswagenag.com/en/news/stories/2019/06/volkswagen-is-developing-more-of-its-own-software.html [last accessed July 2020]

W3C (2016). Automotive and Web at W3C: http://www.w3.org/auto/ [last accessed May 2020]

Wacker A., Jurisch, M., Jin, J., Wolf, P. and H. Krcmar (2014). „Identifying Main User Groups for Green IS - An Empirical Study of Electric Vehicles in China." *In Proceedings of ICIS 2014.* Association for Information Systems.

Wagner, S., Willing, C., Brandt, T. and D. Neumann (2015). Data analytics for location-based services: enabling user-based relocation of Carsharing vehicles. In *Proceedings of the Thirty Sixth International Conference on Information System (ICIS 2015)*.

Walter, J. and B. Abendroth (2018). Losing a Private Sphere? A Glance on the User Perspective on Privacy in Connected Cars.

Wamba, S.F., Akter, S., Edwards, A., Chopin, G. and D. Gnanzou (2015). How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, 165, 234–246. https://doi.org/https://doi.org/10.1016/j.ijpe.2014.12.031

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J., Dubey, R. and S.J. Childe (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, 70, 356–365. https://doi.org/https://doi.org/10.1016/j.jbusres.2016.08.009

Wand, Y. and R. Weber (1993). On the Ontological Expressiveness of Information Systems Analysis and Design Grammars. *Journal of Information Systems (3)*, pp. 217-237.

Wand, Y. and R. Weber (2002). Information Systems and Conceptual Modeling: A Research Agenda, *Information Systems Research (13*:4), pp. 363-376.

Wang, P. (2018). Taking the "Eco" Seriously: A Multilevel Model of Digital Innovation Ecosystems. *In Proceedings of Twenty-second Pacific Asia Conference on Information Systems* (PACIS 2018).

Warwick, B., Symons, N., Chen, X. and K. Xiong (2015). Detecting driver drowsiness using wireless wearables. *In Proceedings of the 12th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, 585-588.

Watson, R. T., Boudreau, M.-C. and A. J. Chen (2010). Information systems and environmentally sustainable development: energy informatics and new directions for the is community. MIS Quarterly 34, 1: 23–38. https://doi.org/Article

Weill, P. and S.L. Woerner (2015). "Thriving in an increasingly digital ecosystem." MIT Sloan Management Review 56(4), 27.

Wiener, M., Saunders, C. and M. Marabelli (2020). Big-Data Business Models: A Critical Litera-ture Review and Multiperspective Research Framework. *Journal of Information Technology (35*:1), SAGE Publications Ltd, pp. 66–91.

Winkler, M., Mehl, R., Matthies, M. and S. Monske (2020). Monetizing Vehicle Data: How to fulfill the promise. Capgemini Invent. https://www.capgemini.com/wp-content/uploads/2020/09/CapgeminiInvent_VehicleDataMonetization_POV_Sep2020.pdf [last accessed May 2021]

Wittmann J. (2017). *Electrification and Digitalization as Disruptive Trends: New Perspectives for the Automotive Industry?* In: Khare A., Stewart B., Schatz R. (eds) Phantom Ex Machina. Springer, Cham, pp.137-159.

WKO (2020). “Fahrzeugindustrie: Branchendaten.” http://wko.at/statistik/BranchenFV/B_217.pdf [last accessed July 2020]

Wolf, G. (2013). Tracking the quantified self. IEEE Spectr 50(8):24

Woźniak, P., Valton, R. and M. Fjeld (2015). Volvo Single View of Vehicle: Building a Big Data Service from Scratch in the Automotive Industry. *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, New York, USA, pp. 671-678. (https://doi.org/10.1145/2702613.2702972).

Xero (2018). “Connect Xero Mileage Tracker & Your Car | Automatic Works With Xero.” https://automatic.com/apps/xero/ [last accessed May 2020]

Xu, C., Pathak, P. H. and P. Mohapatra (2015). “Finger-writing with smartwatch: A case for finger and hand gesture recognition using smartwatch.” *In Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications*, 9-14.

Xu, W., Zhou, H., Cheng, N., Lyu, F., Shi, W., Chen, J. and X. Shen (2017). Internet of vehicles in big data era. *IEEE/CAA Journal of Automatica Sinica* 5(1), 19-35.

Duan, Y., Fu, G., Zhou, N., Sun, X., Narendra, N.C. and B. Hu (2015). “Everything as a Service (XaaS) on the Cloud: Origins, current and future trends,” in Cloud Computing (CLOUD), 2015 IEEE 8th International Conference on. IEEE, 2015, pp. 621–628.

Yoo, Y. (2013). The Tables Have Turned: How Can the Information Systems Field Contribute to Technology and Innovation Management Research? Journal of the Association for Information Systems 14: 227–236.

Yoo, Y., Boland, R. J., Lyytinen, K. and A. Majchrzak (2012). Organizing for Innovation in the Digitized World. Organization Science 23, 5: 1398–1408. https://doi.org/10.1287/orsc.1120.0771

Yoo, Y., Henfridson, O. and K. Lyytinen (2010). The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research, Information Systems Research, Volume 21, Issue 4, pp 724-735.

Yoo, Y., Lyytinen, K. J., Boland, R. J. and N. Berente (2010). The Next Wave of Digital Innovation: Opportunities and Challenges: A Report on the Research Workshop “Digital Challenges in Innovation Research” (June 8, 2010). SSRN. Retrieved May 12, 2017 from https://ssrn.com/abstract=1622170 or http://dx.doi.org/10.2139/ssrn.1622170

You, G. W., Park, S. and D. Oh (2017). Diagnosis of electric vehicle batteries using recurrent neural networks. *IEEE Transactions on Industrial Electronics*, 64(6), 4885-4893.

Young, K. and M. Regan (2007). Driver distraction: A review of the literature. In: Faulks, I. J., Regan, M., Stevenson, M., Brown, J., Porter, A., Irwin, J. D. (eds.) Distracted driving, pp. 379–405, Sydney, NSW: Australasian College of Road Safety

Zeeb, K., Buchner, A. and M. Schrauf (2015) What determines the take-over time? An integrated model approach of driver take-over after automated driving. Accident Analysis & Prevention 78, 212–221 (2015)

Zendrive (2017). Zendrive | Making Roads Safe Using Data and Analytics. https://www.zendrive.com/ [last accessed May 2020]

Zendrive (2018). Distracted Driving, https://d1x6dm64pjo2h2.cloudfront.net/casestudies/Zendrive_Distracted_Driving_2018.pdf [last accessed May 2020]

Zhang, M., Wo, T., Xie, T., Lin, X., and Y. Liu (2017). Carstream: an industrial system of big data processing for internet-of-vehicles. *Proceedings of the VLDB Endowment, 10*(12), 1766-1777.

Zhao M., F. Adib and D. Katabi (2016). "Emotion recognition using wireless signals." *In Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, 95-108.

Zoller, E. (2018). Connected Car Consumer Services: Winning at Data Monetization. http://images.intelligence.informa.com/Web/InformaUKLimited/%7B590f4685-85e1-4083-8b00-5fe732bed22f%7D_OvumIoT2018Ebook-ConnectedCar.pdf [last accessed October 2020]

Zolnowski, A., Anke, J., & Gudat, J. (2017). Towards a cost-benefit-analysis of data-driven business models. In *Proceedings of 13th International Conference on Wirtschaftsinformatik* (WI 2017).

Zolnowski, A., Christiansen, T., & Gudat, J. (2016). Business model transformation patterns of data-driven innovations. In *Proceedings of European Conference on Information Systems (ECIS 2016)*.

Zott, C. and R. Amit (2010). Business model design: an activity system perspective. Long range planning 43, 2–3: 216–226. https://doi.org/10.1016/j.lrp.2009.07.004

Zott, C. and R. Amit (2017). Business Model Innovation: How to Create Value in a Digital World. GfK Marketing Intelligence Review 9, 1: 18–23. https://doi.org/10.1515/gfkmir-2017-0003

# V. List of Abbreviations

ABDC........*Australian Business Deans Council*

ACM........*Association for Computing Machinery*

ADAC........*Allgemeiner Deutscher Automobil-Club e. V.*

ADAS........*Advanced Driver Assistance Systems*

AISeL........*Association for Information Systems Electronic Library*

AMCIS........*Americas Conference on Information Systems*

API........*Application Programming Interface*

BISE........*Journal Business & Information Systems Engineering*

BITKOM........*Bundesverband Informationswirtschaft, Telekommunikation und neue Medien*

CAiSE........*International Conference on Advanced Information Systems Engineering*

CAN........*Controller Area Network*

CEO........*Chief Executive Officer*

E&I........*Journal Elektrotechnik und Informationstechnik*

ECIS........*European Conference on Information Systems*

ECU........*Electronic control units*

EIS........*Enterprise Information Systems*

ERP........*Enterprise Resource Planning*

EU........*European Union*

FIA........*Fédération Internationale de l'Automobile*

GDPR........*General Data Protection Regulation*

GNSS........*Global Navigation Satellite System*

GPS........*Global Positioning System*

HTTP........*Hypertext Transfer Protocol*

ICT........*Information and Communication Technology*

ICVES........*IEEE International Conference on Vehicular Electronics and Safety*

IJIM........*International Journal of Information Management*

i-Know........*International Conference on Knowledge Technologies and Data-driven Business*

IMU........*Inertial Measurement Unit*

ISO........*International Organization for Standardization*

JSON........*JavaScript Object Notation*

LiDAR........*Light Detection And Ranging*

LNBIP........*Series on Lecture Notes in Business*

LNMOB........*Book Series Lecture Notes in Mobility*

MIS........*Management Information Systems*

MQTT ........ *Message Queuing Telemetry Transport*
NBM ........ *International Conference on New Business Models*
NCAP ........ *New Car Assessment Programme*
OBD ........ *On Board Diagnostic*
OEM ........ *Original Equipment Manufacturer*
OVDP ........ *Open Vehicle Data Platform*
PCB ........ *Printed Circuit Board*
RADAR ........ *RAdio Detection And Ranging*
RPM ........ *Revolutions Per Minute*
RQ ........ *Research Question*
SADT ........ *Structured Analysis and Design Technique*
SAE ........ *Society of Automotive Engineers, now SAE International*
SDK ........ *Software Development Kit*
SUV ........ *Sport Utility Vehicles*
TAM ........ *Technology Acceptance Model*
TLS ........ *Transport Layer Security*
TPB ........ *Theory of Planned Behavior*
TPM ........ *Trusted Platform Module*
TRA ........ *Theory of Reasoned Action*
US ........ *United States*
USA ........ *The United States of America*
USD ........ *US-Dollar*
UTAUT ........ *Unified Theory of Acceptance and Use of Technology*
V2I ........ *Vehicle-to-Infrastructure*
V2V ........ *Vehicle-to-Vehicle*
VDA ........ *Verband der Automobilindustrie e.V.*
VEHITS *International Conference on Vehicle Technology and Intelligent Transport Systems*
VHB ........ *Verband der Hochschullehrer für Betriebswirtschaft e.V.*
W3C ........ *World Wide Web Consortium*
WEBIST ........ *International Conference on Web Information Systems and Technologies*

# VI. List of Figures

# VII. List of Tables

# VIII. Contributions in Peer-reviewed Publications

In the following, for each of the fourteen peer-reviewed publications, a paragraph will discuss what contributions the author of this dissertation made to the respective publication.

1. Stocker, A., and **Kaiser, C.** (2016). Quantified car: potentials, business models and digital ecosystems (original german title: Quantified Car: Potenziale, Geschäftsmodelle und Digitale Ökosysteme). *e&i Elektrotechnik und Informationstechnik, 133(7),* 334-340. DOI 10.1007/s00502-016-0429-3.[24]

   I contributed to this article by examining the quantified car startups' business models to reveal that all startups are pursuing similar use cases by generating services from vehicle data. I was responsible for Section 2.3 “Quantified Car Ecosystems” and developed the ecosystem model together with the main author. Furthermore, I was responsible for Section 2.5 “Conclusion and Discussion”, where I have concluded and discussed the introduction of *Quantified Vehicles* (referred to in the article as *Quantified Car*) and our findings. I also supported the writing in all sections and all decisions were discussed and made together within the two authors. I also critically revised the manuscript and thus contributed to the introduction and definition of the Quantified Vehicles concept.

[24] Section numbers from the doctoral thesis. Corresponding paper chapters: Section 2.3 ≙ Chapter 3, Section 2.5 ≙ Chapter 5.

2. Stocker, A., **Kaiser, C.**, and Fellmann, M. (2017). Quantified vehicles. *Business & Information Systems Engineering (BISE), 59(2),* 125-130.DOI:10.1007/s12599-017-0465-5[25]

   First, my contributions are related to the work presented in Stocker and Kaiser (2016), as this journal article is based on it. Hence, the investigation of the quantified vehicle startups' business models and the development of the ecosystem model are also my contribution in this article, as I continued the evolution with support of the co-authors. Thereby I was responsible for Section 3.2 "Current Developments: Towards Digital Ecosystems for Quantified Vehicles". Furthermore, I also contributed as a discussion partner to the introduction and definition of *Quantified Vehicles*. Second, the journal article presents a research framework for *Quantified Vehicles* and discusses the relevancy of selected research directions of this research framework for the BISE/IS (Business & Information Systems Engineering / Information Systems) community. I also contributed as a discussion partner to the creation of the research framework for Quantified Vehicles and critically revised the manuscript.

3. **Kaiser, C.**; Stocker, A.; Viscusi, G.; Festl, A.; Moertl, P.; and Glitzner M. (2017). Quantified Cars: An Exploration of the Position of ICT Start-ups vs. Car Manufacturers Towards Digital Car Services and Sustainable Business Models. In *Proceedings of 2nd international conference on new business models (NBM)*, pp. 336-350.

   As the corresponding author, I wrote, reviewed and edited the original draft with the help of the co-authors. Specifically, the second author and I each conducted a desk research approach in which we analyzed information available on the Web. So we both studied in depth the websites of all the start-ups (included in the paper) to learn more about their visions and goals, as well as their business models, products and services. We then discussed our knowledge with the other authors to come to a common understanding. Consequently, the following content is also all based on these considerations and includes my contributions (I discussed the implications with the co-authors before writing the original draft, reviewed and edited the original draft): Stakeholders with interest in the services, start-up examples and their value proposition, the two approaches to collect data (OBD-2 vs. smartphone), and the current position of vehicle manufacturers. Finally, I critically revised the final manuscript and integrated improvements according to the reviewer comments.

---

25 Section number from the doctoral thesis. Corresponding paper chapter: Section 3.2 ≙ Chapter 2.

4. **Kaiser, C.**; Stocker, A.; Viscusi, G.; (2017). Digital Vehicle Ecosystems and New Business Models: An Overview of Digitalization Perspectives. Platform Economy & Business Models workshop at i-KNOW ’17. Link: http://ceur-ws.org/Vol-2025/paper_ddds_1.pdf

The article was written during the same period as the paper for the NBM conference (Kaiser et al., 2017b), so they are partially related in content. However, this paper addressed a conference in the knowledge/information domain, which changed the focus in my project administration and description of related work to this domain accordingly. As the corresponding author, I wrote, reviewed and edited the original draft with the help of the co-authors. As a research method, the second author and I each conducted a lightweight online market research approach in which we analyzed information available on the Web. I discussed the results with the co-authors and within the members of the AEGIS project consortium (EU project in which all three authors collaborated as well). Furthermore, I was responsible for analyzing Audi's and Volkswagen's service offerings for data-driven services in order to develop an overview of the vehicle manufacturers' service offerings together with the co-authors. Finally, I critically revised the manuscript.

5. **Kaiser, C.**; Stocker, A.; Festl, A.; Lechner, G.; Fellmann, M.; (2018). A research agenda for vehicle information systems. In *Proceedings of European Conference on Information Systems (ECIS 2018).*[26]

As the corresponding author, I wrote, reviewed and edited the original draft with the help of the co-authors. In particular, I conducted the literature review on AISeL, co-authored the introduction and motivation with the second author, and was responsible for Subsection 6.1.2 “Scope and Examples of Vehicle IS”. I was also involved in defining the concept of Vehicle IS through discussions and by editing the original draft. In addition, I managed the development of the research agenda and sample research questions (Section 6.2), synthesizing and summarizing the results of six in-depth interviews and discussing all content with the co-authors. In the process, I wrote the original draft of Subsection 6.2.5 “Business Models and Platform Ecosystems in the Context of Vehicle IS” and provided example research questions for all research directions. I also co-authored the conclusions and future work section with the second author. Finally, I critically revised the final manuscript and integrated improvements according to the reviewer comments.

6. **Kaiser, C.**, Festl, A., Pucher, G., Fellmann, M., and Stocker, A. (2019). The Vehicle Data Value Chain as a Lightweight Model to Describe Digital Vehicle Services. In *Proceedings of the 15th International Conference on Web Information Systems and Technologies* (WEBIST), ISBN 978-989-758-386-5, pages 68-79. DOI: 10.5220/0008113200680079[27]

In this article, I have incorporated my experience from the development and analysis of *Data-driven Services* that the value creation of services is subject to a certain pattern, comparable to the Big Data Value Chain. As the corresponding author, I wrote, reviewed and edited the original draft with the help of the co-authors. In particular, I contributed to writing part of the introduction, related work, and conclusions. I was also responsible for Subsections 7.2.2 and 7.2.3, where we propose and describe in detail the *Vehicle Data Value Chain* (VDVC) as a lightweight model and apply it to describe and compare existing *Data-driven Services.* I have designed the VDVC based on discussions with the second author and the last author, and contributions from the literature review. Finally, I critically revised the final manuscript.

---

26 Section numbers from the doctoral thesis. Corresponding paper sections: Subsection 6.1.2 ≙ Section 2.2, Section 6.2 ≙ Chapter 3, Section 6.2.5 ≙ Section 3.5.

27 Subsection numbers from the doctoral thesis. Corresponding paper sections: Subsection 7.2.2-7.2.3 ≙ Section 3.2-3.3.

7. **Kaiser, C.**, Festl, A., Pucher, G., Fellmann, M. and A. Stocker (2020). Digital Services Based on Vehicle Usage Data: The Underlying Vehicle Data Value Chain. In: Bozzon A., Domínguez Mayo J.F., Filipe J. (eds) Web Information Systems and Technologies. WEBIST 2019. *Lecture Notes in Business Information Processing (LNBIP), vol 399*. Springer, Cham. https://doi.org/10.1007/978-3-030-61750-9_2

Since I was invited to expand on the WEBIST paper (Kaiser et al., 2019a) for submission to the LNBIP series, with the help of my co-authors, I expanded on the VDVC definition accordingly and have now described it in more detail in this article using eight characteristics. The paper now also includes an evaluation of VDVC with two case studies. As the corresponding author, I wrote, reviewed and edited the original draft with the help of the co-authors. In particular, I was responsible for the development of the extended version of the VDVC (eight characteristics per step) and all related changes in the article. In addition, I wrote the text of Case A based on discussions with the third author, and I supervised and co-authored the writing of Case B, a real life case where I was involved in the development, and also led the developments within an internal project (project leader of the project *Lightweight Digital Mobility Assistance*) and an EU ECSEL JU project (responsible task leader of task 3 in work-package 11 of project SCOTT). Finally, I critically revised the final manuscript.

8. Papatheocharous, E., Frécon, E., **Kaiser, C.**, Festl, A., and Stocker, A. (2018). Towards a generic IoT platform for data-driven vehicle services. In *2018 IEEE International Conference on Vehicular Electronics and Safety (ICVES)* (pp. 1-6). IEEE. DOI: 10.1109/ICVES.2018.8519505[28]

Here, I was responsible for Subsections 9.2.1 and 9.4.1 on related work for *Quantified Vehicles* and the description of the *Vehicle Data Logger*. Furthermore, as task leader in the SCOTT research project (WP11, Task 3) in which this approach was developed, the concept was also developed under my leadership based on joint discussions and brainstorming sessions. Furthermore, I contributed related work, added the list of “potential applications for the end users” in Section 9.3, developed the telemetry data visualization on which Figure 45 is based, and contributed to the work presented in Subsection 9.4.3 (“End-User Applications”) through joint discussions with authors number four and five. Finally, I critically edited and revised the final manuscript.

---

28 Section/Figure numbers from the doctoral thesis. Corresponding paper chapters: Subsection 9.2.1 ≙ Section II.A, Subsection 9.4.1 ≙ Section IV.1, Subsection 9.4.3 ≙ Section IV.3, Section 9.3 ≙ Chapter III, Figure 45 ≙ Figure 4.

9. **Kaiser, C.**, Steger, M., Dorri, A., Festl, A., Stocker, A., Fellmann, M., and Kanhere, S. (2019). Towards a Privacy-Preserving Way of Vehicle Data Sharing – A Case for Blockchain Technology? In *Proceedings of Advanced Microsystems for Automotive Applications (AMAA) 2018*. *Part of the Lecture Notes in Mobility book series (LNMOB)*. Springer, Cham. pp 111.122. DOI: 10.1007/978-3-319-99762-9_10[29]

   This published concept idea was created together with the second and fifth authors. As the corresponding author, I wrote, reviewed and edited the original draft with the help of the co-authors. In particular, I contributed to Section 10.1 “Scope” with discussions, and in revising the text. In the related work section, I was responsible for 10.2.2 “Connected Vehicles and Data Exploitation” together with the fifth author. After M. Steger left the company, I stepped into the role as the first and corresponding author and wrote the original draft in the concept section (10.3), which was based on our joint considerations. And also in the conclusion section (10.4) some parts are written by me, such as the description of issues and research topics that are still open. In addition, I designed the e3value model that shows the actors and value flows of a vehicle data sharing ecosystem (Figure 48). Finally, I critically edited and revised the final manuscript.

10. Lechner, G., Fellmann, M., Festl, A., **Kaiser, C.**, Kalayci, T. E., Spitzer, M., Stocker, A., (2019). A Lightweight Framework for Multi-device Integration and Multi-sensor Fusion to Explore Driver Distraction. In *International Conference on Advanced Information Systems Engineering (CAiSE)* (pp. 80-95). Springer, Cham.

    I contributed to the paper in co-writing the parts in the background section about “ways how to *quantify drivers and driving”*, and by assembling and maintaining the Beaglebone hardware (gateway device for collecting vehicle data). In addition, I also planned and conducted the empirical driver study in which we tested the solution and collected data for evaluation. I also critically revised the manuscript.

[29] Section/Figure numbers from the doctoral thesis. Corresponding paper chapters: Section 10.1 ≙ Chapter 1, Sections 10.3-10.4 ≙ Chapters 3-4, Figure 48 ≙ Figure 2.

11. **Kaiser, C.**; Festl, A.; and Stocker, A. (2020). Nutzung von Automotive Big Data zur Entwicklung zweier neuer Anwendungen. In *BITKOM position paper on Practical Use Cases of Artificial Intelligence & Big Data in Industry (original german title: Konkrete Anwendungsfällevon KI & Big-Data in der Industrie),* pages 10-19. Link: https://bit.ly/2VsqTkv

As a result of my presentation at the Big-Data.AI Summit 2019, I was invited to submit a contribution to a BITKOM position paper. In this paper, together with my co-authors, we describe developments from several research projects (including AEGIS, EVOLVE) in which I have collaborated. We reused content from the project deliverables, which I also co-authored. As the corresponding author, I administered, partially wrote, reviewed and edited the original draft and critically revised the final manuscript.

12. **Kaiser, C.**; Stocker, A.; Festl, A.; Djokic Petrovic, M.; Papatheocharous, E.; Wallberg, A.; Ezquerro, G.; Ortigosa Orbe, J.; Szilagyi, T.; and Fellmann, M. (2020). A Vehicle Telematics Service for Driving Style Detection: Implementation and Privacy Challenges. In *Proceedings of the 6th International Conference on Vehicle Technology and Intelligent Transport Systems - Volume 1 (VEHITS),* ISBN 978-989-758-419-0, p.29-36. DOI: 10.5220/0009329400290036[30]

Both individual contents of this paper were developed under my leadership and with my contribution. First, the contents of the section on data privacy (13.1) are based on joint considerations by myself and the last author. Our considerations, such as the "preliminary model of the willingness to share data" (Figure 59) and a preliminary version of the "privacy levels for vehicle data sharing" (Figure 60), served as input for Tom Szilágyi's master's thesis (Szilágyi, 2019), in which he investigated the level system for data sharing in an empirical study under my guidance. I included the results in the paper to also expose it to scientific discourse and further disseminate the findings. Second, as leader of the corresponding task in the SCOTT research project (WP11, Task 3), through regular meetings with the partners, I was involved in the conception of all steps described in the service implementation section (13.2). Thereby, I co-developed the Vehicle Data Logger (13.2.1) together with the second and third authors. In addition, as the corresponding author, I homogenized and revised the partners' paper-contributions and critically revised the final manuscript.

[30] Section numbers from the doctoral thesis. Corresponding paper chapters: Sections 13.1-13.2 ≙ Chapters 2-3, Subsection 13.2.1 ≙ Section 3.1, Figure 59-Figure 60 ≙ Figure 1-2.

13. **Kaiser, C.**, Stocker, A., Fellmann, M., (2019). Understanding Data-driven Service Ecosystems in the Automotive Domain. In *Proceedings of 25th Americas Conference on Information Systems (AMCIS 2019)*.[31]

    As the corresponding author, I wrote, reviewed and edited the original draft with the help of the co-authors. In particular, I conducted the literature review to derive a preliminary model of the *Data-driven Service Ecosystem*, developed the interview guide, contacted interview participants, conducted four interviews in person, and seven interviews online. I analyzed the interviews and the related work and developed the original draft of (i) the high-level model of the automotive *Data-driven Service Ecosystems*, (ii) the detailed model of the data ecosystem, and (iii) the overview of the providers and consumers of selected *Data-driven Services*. In addition, based on the comments of the co-authors, I iterated the findings and improved the models and the paper accordingly. Together with the second author, I wrote the introduction and the sections 14.2 - 14.4 (*Data-driven Service Ecosystem*, Discussion, Conclusion and Outlook).

[31] Section numbers from doctoral thesis. Corresponding paper ch.: Sections 14.2 - 14.4 ≙ Chapters 3-5.

14. **Kaiser, C.**, Stocker, A., Viscusi, G., Fellmann, M. and Richter, A. (2021). Conceptualizing Value Creation in Data-driven Services: The Case of Vehicle Data. *International Journal of Information Management (IJIM)*[32]

Based on my expert interviews (c.f. AMCIS conference paper: Kaiser et al., 2019b), I have continued to work with the co-authors on this journal publication, which presents a conceptual model for value creation in *Vehicle Data-driven Services*. As the corresponding author, I wrote, reviewed and edited the original draft with the help of the co-authors. Specifically, I conducted the literature search in established scientific electronic databases outlined in 15.2 with the second author, developed the interview guide, and conducted four interviews in person and seven interviews online. I critically compared and analysed the interviews and related work (Section 15.3), and developed the original draft of the conceptual model in discussions with the second author (Section 15.4). Based on discussions with the co-authors, I iterated the results and improved the model and the article seven times accordingly based on several reviewer feedbacks (reviews from four top-ranked conferences and journals). In addition, I participated in the development and writing of the model validation (Section 15.5) and in the writing of the discussion and conclusion sections (15.6 and 15.7). Finally, I critically revised the manuscripts.

32 Section numbers from doctoral thesis. Corresponding paper chapters: Sections 15.2-15.7 ≙ Chapter 3-8.

# IX. Statutory Declaration (Eidesstattliche Erklärung)

**Eidesstattliche Erklärung**

Hiermit bestätige ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen sind, wurden unter Angabe der Quelle kenntlich gemacht.

Aurach am Hongar, 04.06.2021 (Christian Kaiser)

**Statutory Declaration**

I hereby certify that I have written this dissertation independently and have not used any aids other than those specified. The places of the work, which are taken from other works in the wording or sense of the text, have been marked with indication of the source.

Aurach am Hongar, 04.06.2021 (Christian Kaiser)