DEPARTMENT OF INFORMATICS

TECHNICAL UNIVERSITY OF MUNICH

Master's Thesis Information Systems

Opportunities and Barriers for Advancing the API Economy within the Automotive Industry

Fridolin Jakob Koch

DEPARTMENT OF INFORMATICS

TECHNICAL UNIVERSITY OF MUNICH

Master's Thesis Information Systems

Opportunities and Barriers for Advancing the API Economy within the Automotive Industry

Chancen und Hinderungsgründe für die Weiterentwicklung der API Economy in der Automobilbranche

Author:	Fridolin Jakob Koch
Supervisor:	Prof. Dr. rer. nat. Florian Matthes
Advisor:	Gloria Bondel, M. Sc.
Submission Date:	15.02.2019

I confirm that this master's thesis information systems is my own work and I have documented all sources and material used.

Munich, 15.02.2019 Fridolin Jakob Koch

Acknowledgments

First and foremost, I would like to express my gratitude to my advisor Gloria Bondel for giving me the opportunity to work on this particular topic and providing me with guidance and support throughout the course of this thesis. Furthermore, I want to express my uttermost appreciation for the continuous constructive feedback and especially the assistance provided for the conduction of the interviews.

I also want to thank my supervisor Prof. Dr. Florian Matthes for providing me with the topic and the opportunity to write this thesis at his chair for Software Engineering for Business Information Systems (SEBIS). Furthermore, I want to express my thankfulness for the advisory when specifying the topic.

In addition, my gratitude is given to all interview participants who essentially made this work possible, by providing us with their time and expertise. This work would could not exist without their extensive input.

Finally, I want to thank my family, and colleagues who supported me continuously. Additionally, I want to express my deepest gratitude to Isabel, for her enduring and unwavering support during the last six months.

Abstract

Over the last decade, a vast number of public APIs have been created across many different industries. This development leads to powerful ecosystems and new business models enabling companies to develop new revenue streams. This trend is also referred to as API Economy. Within the Automotive Industry, little or no such trend can be observed. Especially modern vehicles with myriads of sensors could provide valuable data for new business models around the automotive industry. In this master thesis, current collection methods for vehicle-generated data are explored. Further, the value creation process for vehicle-generated data is analyzed in detail. The current state of the API Economy within the Automotive Industry is analyzed by gathering use-cases for vehicle-generated data from an extensive literature review and semi-structured expert interviews. Furthermore, it is analyzed which barriers are inhibiting the prevalent adoption of an API Economy within the Automotive Industry. The work eventually analyzes potential measures for advancing an API economy within the automotive industry which were gathered from the conducted interviews.

The target of this work is to provide a snapshot of current use cases for vehicle-generated data, the provisioning of vehicle-generated data through APIs provided by the automotive manufacturer and the state of a potential API economy.

Keywords: API, API Economy, Vehicle-Generated Data, Automotive Industry, Value Chain, Value Network, Platform Ecosystem

Contents

1. Introduction

1.1. Motivation

The term Application Programming Interface or API can be traced back to the year 1968 [1], until two decades ago the terms predominant usage was in the context of programming libraries. Today the colloquial meaning of API mostly refers to web APIs [2]. According to *ProgrammableWeb*, a web API directory, the number of APIs provided is still growing, especially APIs categorized as data, financial or analytics [3]. Many of the APIs available today provide access to large data sets. One example is the Google Maps Platform[1] with its Places API providing access to over 150 million places around the globe[2]. This kind of APIs enables other businesses to build business models around the provided data and add value to its customers. Many businesses use external APIs to enhance their product, enrich their data or to provide an API themselves. This ecosystem is often referred to as API Economy and has grown exponentially over the past decade [4], [5].

Starting in the late 1970s, automotive manufacturers started to equip their vehicles with microprocessor-based engine control modules [6]. Within the last two decades the number of sensors increased heavily. Fleming estimated in 2001 that the number of sensors for engine control applications would increase from approximately ten in 1995 to thirty in 2010. In a second review, conducted in 2008, the same author predicted an average of 70 sensors per vehicle in 2013 [7]. From this trend, it can be derived that over the last decade the amount of vehicle-produced data has increased. According to one of the interviewees, modern vehicles have over 140 built-in sensors. These sensors range from mandatory safety sensors to driving convenience sensors. All these sensors could deliver valuable data which could enable new business models. Albeit several standards and techniques are allowing the extraction of sensor data via physical interfaces from inside the vehicle (e.g., *OBD-II*, *CAN*), almost no Original Equipment Manufacturer (OEM) provides access to vehicle generated data through APIs or in

[1] `https://cloud.google.com/maps-platform`

[2] `https://cloud.google.com/maps-platform/places/`

any other way. A variety of different business models have been proposed based on *vehicle-generated data*, of which some have been implemented, but for the most part, they rely on third-party devices. Hence a larger API Economy trend within the automotive industry cannot be observed.

While other automotive market participants offer existing approaches for provisioning *vehicle-generated data* through APIs, the focus of this work are OEMs. This works baseline hypothesis is that if OEMs would provide *vehicle-generated data* through APIs, a powerful ecosystem would emerge and enable a similar API economy within the automotive industry. Many business models and use-cases utilizing *vehicle-generated data* have been proposed, but only a few have been implemented, predominately using third-party devices inside the vehicle. Among others, this is a strong indication that multiple factors inhibit the usage of *vehicle-generated data*. The goal of this work is to explore potential use-cases and business models that depend on *vehicle-generated data*. Further, in this work, it will be analyzed what barriers are inhibiting the widespread availability of *vehicle-generated data* through APIs provided by the OEMs and what measures could be applied to overcome these barriers.

1.2. Research Questions & Approach

The following section is outlining the research questions which where defined to examine and validate the hypothesis that if OEMs would provide programmatic access to vehicle-generated data an API economy would emerge within the automotive industry.

RQ1 *How is vehicle generated data currently collected and how is value created with it?*

This question is intended to be answered by providing an overview about state of the art methods and standards how *vehicle-generated data* is collected and processed. Furthermore it will elucidate how *vehicle-generated data* is monetized within the automotive industry. Hence the relevant stakeholders, roles and the processes are described.

RQ2 *What are state of the art use cases and business models that exist or are proposed for vehicle generated data?*

The next step is to provide a detailed insight about use cases and business models, either already existing or proposed. A extensive literature-review and expert interviews were conducted to answer this question.

RQ3 *What are barriers for use-cases not being implemented?*

As stated before, the widespread use of *vehicle-generated data* through OEM provided APIs is currently inhibited. Within the scope of the conducted interviews, barriers for the proposed use cases and business models were also explored.

RQ4 *What could be measures to overcome identified barriers and advance the API Economy within the automotive industry?*

Based on the interview results it was derived what measures would be feasible to enable an API Economy within the automotive industry, especially with the focus on *vehicle-generated data*.

1.3. Thesis Structure

Chapter 2 starts with the introduction of some definitions in section 2.1 which are relevant for the remaining work. Section 2.2 answers the first part of **RQ1** by examining how *vehicle-generated data* is currently collected. In section 2.3 the rest of **RQ1** will be answered by deriving the value creation process from similar domains. Chapter 2 concludes by elucidating on related work.

In the succeeding chapter 3 the research design of this thesis will be illustrated by providing a brief methodological overview of the techniques used to collect the relevant data and the subsequently conducted analysis. The chapter also provides details about the interview process and the participants.

Chapter 4 presents several use cases for *vehicle-generated data*. The chapter starts by presenting the results of the extensive literature review in section 4.1. Subsequently, use cases which have been discovered during the conducted interviews are presented in section 4.2. The section elaborates on both existing and proposed use cases which have been mentioned by interview participants. The chapter therefore answers **RQ2**.

RQ3 is discussed in chapter 5 in which the barriers encountered throughout the interviews are expounded. The chapter introduces a taxonomy to classify the discovered barriers into four distinct categories.

Chapter 6 answers **RQ4** by summarizing the measures which have been mentioned during the interviews. The chapter is subdivided in two subsections. In section 6.1 explicit potential measures, which have been explored in the interviews, are described.

The subsequent section 6.2 elucidates what other barriers the mentioned measures could tackle.

In chapter 7 the results of all four research questions are summarized and further implications based on the results are described.

The thesis eventually concludes in chapter 8 by describing the limitations of this work and potential topics for future research.

2. Foundations

This chapter outlines some common definitions. Albeit the defined terms are not new they are mostly ambiguously defined, therefore this chapter will provide definitions how the particular terms are to be interpreted in the context of this work.

2.1. Definitions

The following section will discuss and define several terms which are necessary to understand the remainder of this thesis.

2.1.1. API

API is the abbreviation for *Application Programming Interface*, the term was first used in an article authored by Cotton in 1968 [1]. While the term API also refers to the interoperability with frameworks, library or operating systems the colloquial meaning has changed today, so that the term API is most often a synonym for web API [2]. In the remaining work the term API will be used as a synonym for web API or at least in the meaning that the API is accessible remotely over a public network. Today the term web API is often equated to *REST API*. *REST* is an abbreviation for *Representational State Transfer*, the term was defined by Roy Fielding in his dissertation [8]. Today the majority of the web APIs are REST APIs. Besides REST there are several other web API protocols like *gRPC*[1], *GraphQL*[2] or *JSON-RPC*[3] but for this work the protocol or technology is not relevant. In this thesis API is furthermore defined as an interface that allows programmatic access over a state of the art network using a state of the art protocol. This means that there is no differentiation between a *gRPC-API* and a REST-API, if they provide access to the same kind of data in a similar manner.

[1] `https://grpc.io`
[2] `https://graphql.org`
[3] `https://www.jsonrpc.org/specification`

2.1.2. API Economy

After having defined the Term *API*, the term *API Economy* is now further examined, unfortunately this term is also ambiguously defined. In the following subsection an overview of different definitions will be provided and eventually conclude with the definition which has been used for this work. From a holistic view the term can be defined as *"[...] the commercial exchange of business functions, capabilities, or competencies as services using web APIs"* [9]. Historically the term API Economy has evolved from the concept of service-oriented architecture (SOA) which emerged in the early 2000s. Back then more and more companies started to build APIs for internal usage or as a data exchange method on a business-to-business level. Today companies are externalizing their business assets through APIs to monetize data and services. This leads to an ecosystem where interconnected APIs are more valuable as isolated APIs, this ecosystem is also referred to as API economy [10]. The providing of APIs to partners or to the public enables companies not only to become an active participant in the API economy, but to generate new revenue streams based on already existing digital assets. The monetization is supported through different business models like subscription-based, licence-based, freemium or pay-as-you-go [11].

Because there is no universal definition for API economy, the following definition was derived based on the explanation above:

> *API Economy describes the generation of value by providing access to digital services through APIs either publicly or only to partners. This process is independent of technology, protocol or functionality. The term also describes API lifecycle managements processes and tools.*

This definition was also used in the introduction of the guideline used for the interview which is summarized in subsection 3.2.1.

2.1.3. Vehicle Generated Data

As already mentioned modern vehicles have a large amount of sensors which are continuously producing data. In this work this kind of data is referred to as *vehicle-generated data*. No formal definition of the term was encountered, but based on the contextual usage of the term in [12]–[14] the following definition is specified.

This definition does not restrict how the data is acquired from the vehicle as there are different methods for accessing *vehicle-generated data*. Data generated by sensors of a third party device, which is installed inside the vehicle, is not considered as *vehicle-generated data*. If a third party device is used to acquire sensor data through an interface of the vehicle it is considered as *vehicle-generated data*.

> *Data which can be obtained from sensors inside or outside a vehicle e.g. odometer, tyre pressure or ambient temperature. Data obtained from a user's device while he is driving is not considered as vehicle generated data (e.g. cellphone sensor data).*

2.2. State of the art data collection

Starting in the early 1980s *OEMs* started to equip new vehicles with so called *engine control units* (ECU) and sensors which replaced pure mechanical or vacuum-based systems. The use of these electronically controlled components increased quickly which subsequently increased the complexity of identifying potential problems and errors with the vehicle. To resolve this issue *OEMs* started to integrate *on-board diagnostic* (OBD) software into the vehicles ECUs, allowing themselves and mechanics to pinpoint potential problems with the vehicle. *The California Air Resources Board* (CARB) became aware of such systems and eventually adopted *OBD* as a requirement in 1985. The objective of the regulation was to improve the emission compliance throughout the lifespan of the vehicle and to inform the driver about vehicle malfunctions. The first generation *OBD* (OBD-I) requirements were relatively simple and included mainly emission and performances related monitoring functions. The *OBD-I* regulation did not require standardized connectors, connector locations (i.e. under the hood v.s. passenger compartment) or even fault codes. This resulted in very heterogeneous implementations which were different from *OEM* to *OEM* or even within the *OEM* [15].

In 1988 the *CARB* started to revise the *OBD-I* standard with the goal to monitor all emission-related components, eventually a new *OBD-II* regulation became effective in 1994. The *CARB* revises the regulations approximately every two to six years. Over the years the requirements changed to include more diagnostic information. Unclear requirements regarding the communication protocol lead to new regulations, requiring the use of only one protocol, namely *Controller Area Network* (CAN) from 2008 onwards [15].

The *European On-Board Diagnostics* (EOBD) standard was approved by the European Parliament and the Council of the European Union with directive *70/220/EC* in 1998. The standard was intended to monitor anti-pollution vehicle equipment. The requirements became effective for gasoline-engine vehicles from January 1, 2000 onwards and for diesel-engine vehicles from January 1, 2003. Due to stricter regulation in the US manufactures mainly implemented the *United States Environmental Protection Agency's*

(EPA) *OBD* requirements [15].

With the *OBD-II* regulation multiple aspects have been standardized ensuring that it is possible to extract fault codes and parametric data using generic scan tools. Further it is required to provide access to an interface placed inside the passenger compartment of the vehicle. The connection is established through a standardized *Data Link Connector* (DLC) [15].

The *OBD-II* regulation does not enforce a minimum implementation standard, therefore there are implementation differences between different *OEMs*. Furthermore *OEMs* extend the standard with their own proprietary codes [16], [17]. For example data about the anti-lock breaking system are proprietary and different for each *OEM*, this kind of data is intentionally kept concealed from the general public for safety concerns [18]. It has been shown that through proprietary vehicle *CAN* protocols an order of magnitude more information can be accessed [19].

Originally the *OBD-II* standard was laid out with the intent to allow the manufacturer or repair shops to access diagnostic information, mainly if there are malfunctions with the vehicle. As the *OBD-II* interface is a mandatory standard, an abundance of different devices exists which are capable of accessing *vehicle-generated data*. One type are devices which are intended for mechanics to diagnose problems with the car, these devices are often intended to be used while the vehicle is stationary. The concept of the connected car also lead to the emergence of another device, often referred to as *OBD-II* dongle. These devices are mainly intended for the consumers who want to equip their vehicle with additional functionality. Most of these devices are intended to be used in conjunction with the user's smartphone via some kind of wireless connection (e.g., Bluetooth). But there are also standalone solutions which use the cellular network to collect extracted data [4] [20]. The concept of accessing the *OBD-II* interface remotely is not new, in 2004 Baltusis described several concepts to access diagnostic data via a cellular network [15], but the technological progress over the last decade made the concept more feasible and cheaper to implement. In the next years a strong market growth is expected for these so called retrofit solutions, indicating a high demand for connected vehicle based business models [20].

Some *OEMs* are already discussing to close the *OBD* interface while the vehicle is driving as the *OBD* interface was not designed for the continuous data extraction. Instead the data shall be accessible through a neutral server [21]. This concept will be discussed in detail in section 2.4.

[4] `https://vimcar.de/fahrtenbuch`

2.3. Vehicle-Generated Data Value Creation

As introduced in **RQ1** it is necessary to understand how value can be generated by using *vehicle-generated data*. The following section will provide an overview about how value is generated, who the involved stakeholders are and how they are involved in the value creation process. The section will introduce two concepts, the big data value chain and value networks which will be used to derive a *value chain* and a *value network* for *vehicle-generated data* presented in subsection 2.3.2.

2.3.1. Big Data Value Chain

While there are several fuzzy definitions for big data, a consensual definition is that there is a large volume of information available which is generated in a high velocity and variety. Furthermore, the subsequent creation of economic value and the transformation of insights from the data are attributed to the term big data [22]. Due to the vast amount of sensors in a modern vehicle and the general amount of vehicles on the road, *vehicle-generated data* can be considered as big data. This section will derive how value can be created by providing *vehicle-generated data* through APIs by comparing the automotive domain to similar domains.

In the field of business management a value chain is a decision support tool that is used to describe the series of activities a firm or organization has to perform to deliver a product or service to the market. This model is also applicable to information systems in the form of a virtual value chain. The data value chain is described as a series of events that generate value from data. The European Commission considers the data value chain as a key concept for digitally transforming more traditional sectors. Figure 2.1 illustrates The Big Data Value Chain which is a high level view on certain information system related activities in the context of big data [23].

Data Acquisition describes the collection and preprocessing (filtering, cleaning) of data before it is put into a storage solution (e.g. Data Warehouse). A major challenge of this process are the infrastructure requirements. The infrastructure must be capable of storing the expected amount of data while still providing a predictable latency [23].

Data Analysis is the activity of processing the raw data to be able to use it for decision-making or domain specific uses. The goal is to extract useful hidden information that might be useful from a business point of view [23].

Data Curation can be described as data life cycle management with the goal to ensure a defined level of data quality. Activities like content creation, selection, classification, transformation and validation are part of the process [23].

Data Storage describes persisting and managing the data in a scalable manner satisfying the needs of applications consuming the data [23].

Data Usage are business actives that require access to the data, its analysis and the relevant tools to integrate the data. Using the data for business decision-making can reduce costs, add value or influence other existing performance criteria and therefore enhance the competitiveness [23].

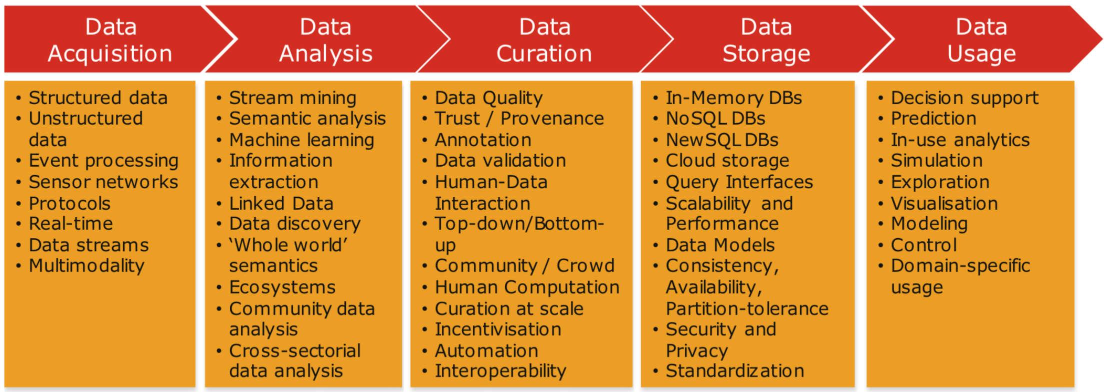

Figure 2.1.: The Big Data Value Chain as described within [23].

2.3.2. Vehicle-Generated Data Value Network

The core concept of a value network is the co-creation of value through a combination of multiple actors. The units of such a network are relatively autonomous and can be managed independently. Cooperation within the network is performed based on common principles. Actors participating in such a network are incrementally adding value to an overall offering, therefore the value is created at network level [24]. The value network "[...] visualizes business activities and sets of relationships from a dynamic whole systems perspective". Each node of the network is a different stakeholder, each edge is a relationship between two stakeholders [25].

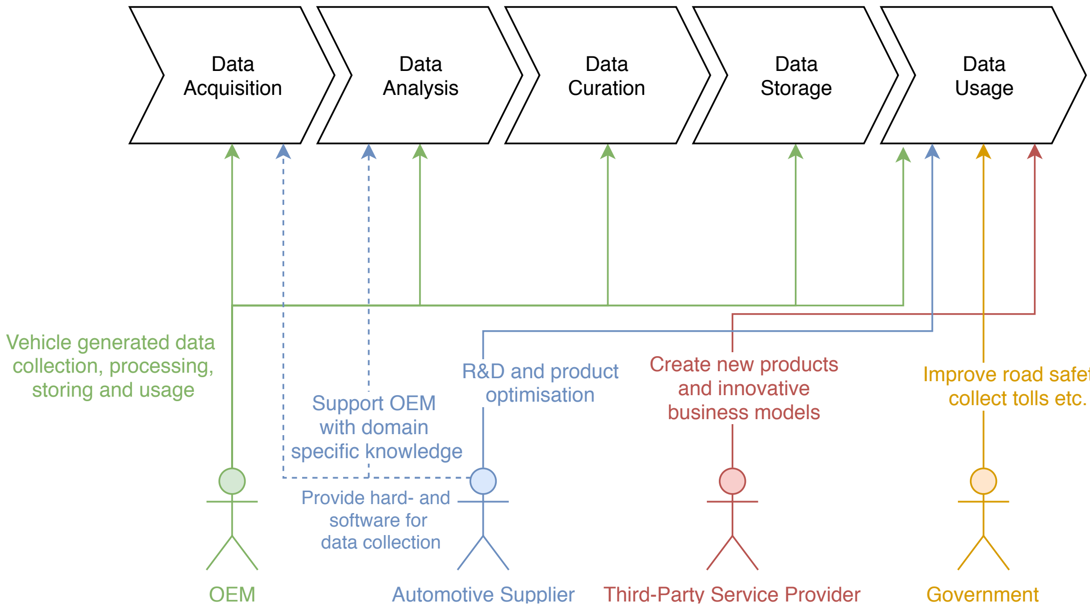

Figure 2.2.: The Vehicle Generated Data Value Chain with Stakeholders derived from [17], [23], [26].

To create a value network for *vehicle-generated data*, the involved stakeholders have to be identified. The manufacturer of the vehicle, in this work referred to as **OEM**, is already selling digital services along with the vehicle to his customers (e.g. infotainment, navigation). Furthermore the **OEM** is already analyzing the data of its customers to improve his product [26]. **Automotive Suppliers** are the developers of the hardware and software required to capture and/or analyze the data. They are also interested in the data to improve their own product and service portfolio [26]. **Third-Party Service Providers** are all stakeholders whose intend is to use *vehicle-generated data* within their business model and provide a service of some kind, either digital (e.g. driver's logbook) or physical (roadside assistance) [17]. This stakeholder category includes several company classifications. Insurers are interested in offering usage-based insurance contracts and to improve their understanding of the customer's behavior. Mobility providers (e.g. car sharing) already require *vehicle-generated data* for their business case, their interest is to improve their fleet operations and vehicle allocation. Start-ups are entering the economy from different angles like the provisioning of new applications, the development of new hardware or through innovative monetization schemes. Roadside assistance provider can use *vehicle-generated data* to collect and process distress calls in realtime to optimize the dispatching of rescue vehicles [26]. The **Government** has a binary function as legislator and data consumer. As a legislator the **Government** has to enforce competition and antitrust laws to ensure equal access to *vehicle-generated data* for all stakeholders participating in the ecosystem [17]. As a data consumer the **Government** uses the data to improve public infrastructure like roads, public transportation or to collect tolls [26]. Figure 2.2 illustrates the *vehicle-generated data* value chain which has been derived from [17], [23], [26]. The figure includes the involved stakeholders for each activity. The dashed line indicates that a stakeholder is supporting an activity.

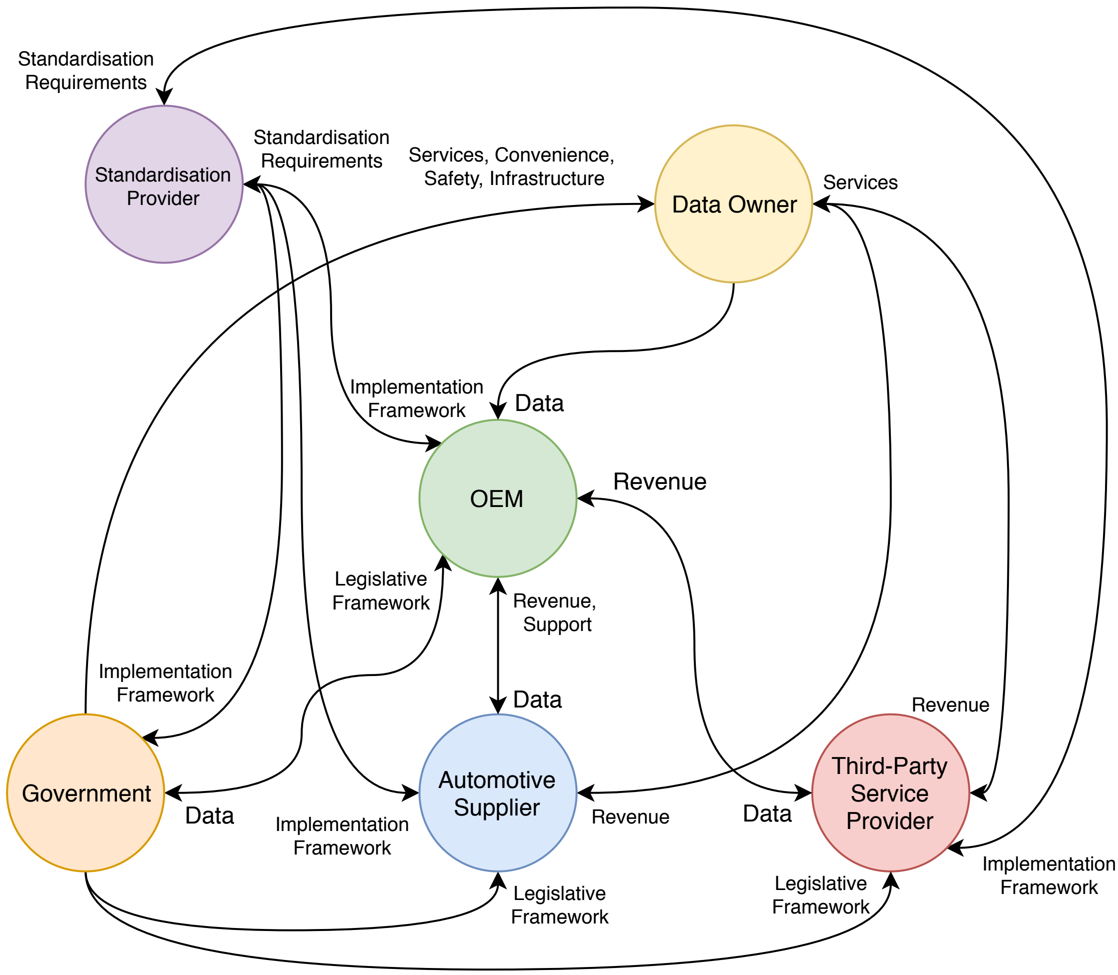

Figure 2.3.: Value Network Model for Vehicle Generated Data [17], [25], [26].

In Figure 2.3 a value network for *vehicle-generated data* is illustrated which was derived from [17], [25], [26]. When deriving the value network several assumptions have been made. First it is assumed that the **Data Owner** is the user of the vehicle and not the **OEM**, albeit chapter 5 will elucidate that the data ownership is contentious. Further it is assumed that some form of standardization by an independent organization is required to enable an API economy. The standardization organization is supposed to facilitate formal standards between the stakeholders and to constitute an implementation framework. In the following the relations between the stakeholders, illustrated as edges in the value network, are described in detail in their logical order according to the value chain illustrated in Figure 2.2. The data is generated from sensors inside a vehicle and has a **Data Owner** (i.e. the driver). The **OEM** is collecting the data from the **Data Owner**. From the **OEM** the Data is passed to either the **Government**, an **Automotive Supplier** or a **Third-Party Service Provider** in return the **OEM** receives pecuniary compensation from either the **Automotive Supplier** or **Third-Party Service Provider**. It is assumed that the **OEM** provides the data to the **Government** free of charge or due to some kind of legislative regulation. **Automotive Suppliers** also support the OEM with hardware, software and domain specific knowledge required for the data collection process. The **Government**, **Automotive Suppliers** and **Third-Party Service Providers** all make services available to the **Data Owner**. Additionally the **Government** uses the data to improve road safety, improve the public infrastructure or collect tolls and taxes. In return for providing services to the **Data Owner**, **Automotive Suppliers** and **Third-Party Service Providers** receive revenue, either in monetary form or other indirect compensation (e.g. user data). The **Government** defines the legislative framework (e.g. data privacy law, competition law, antitrust law) which applies to all entities involved in the data collection and usage.

2.4. Related Work

The following section will provide an overview over existing international and national standards, concepts, projects and APIs which are involved in the provisioning of *vehicle-generated data*.

2.4.1. ISO Standards

Since summer 2014 the automotive industry is collaborating with the *International Organization for Standardization* (ISO) as part of a large-scale standardization project named *Extended Vehicle* (ExVe). The project has the goal to create a web-based platform which shall provide secure access to *vehicle-generated data* for third party service providers. Automotive *OEMs* are currently implementing the technological prerequisites [27].

The relevant *ISO* standards can be grouped into three categories, Figure 2.4 provides an overview over the standards which are currently in development.

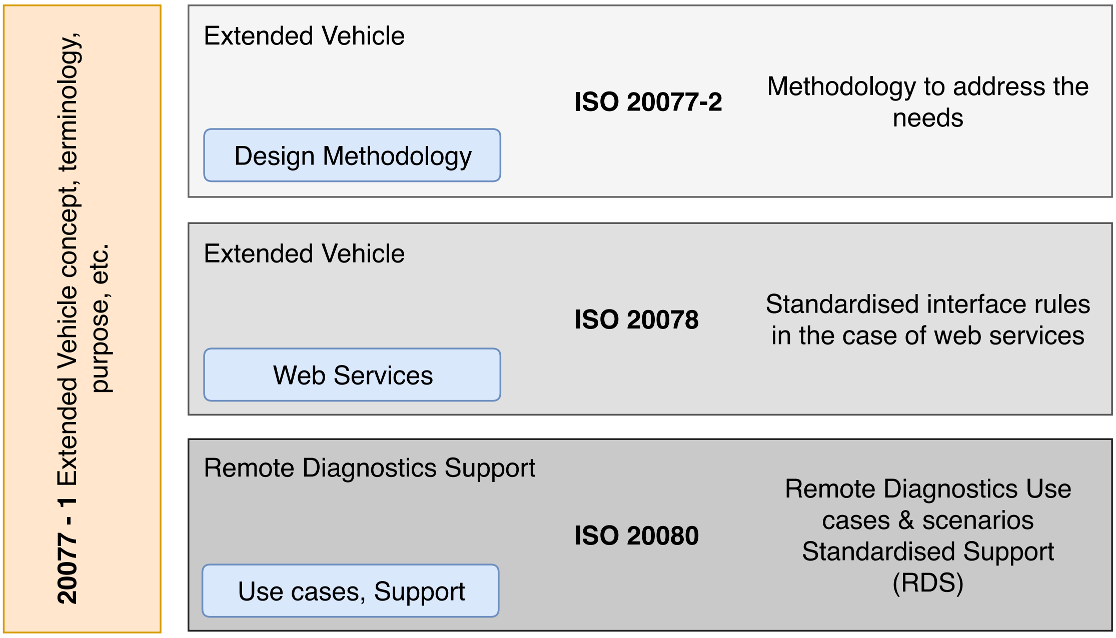

Figure 2.4.: ISO 20077, 20078, 20080 Standards and Projects adapted from [28].

The **ISO 20077** series of standards consists of two parts containing diverse generic specifications for the extended vehicle. The first part **ISO 20077-1** contains fundamental definitions, concepts and examples as well as related standards. The second part **ISO 20077-2** contains the methodology to design an extended vehicle. Neither of the two parts contains any technical specifications. **ISO 20077-2** specifies a set of general principles and rules which are intended for *OEMs* to derive their own methods and processes for designing an extended vehicle which supports a certain set of use cases and scenarios. The **ISO 20077-2** standard provides general guidance for the following points [28]:

- A template describing the usage the extended vehicle is intended to address, which shall be used by the requesting party (e.g. *Third-Party Service Provider*).
- Guidance for the vehicle *OEM* for not omitting any design steps in the design process.
- Guidance for the *OEM* for expressing the technical results of the design to a requesting party.

OEMs following the **ISO 20077-2** standard shall implement appropriate measures in their own design methods and procedures. The following list of principles are meant to be taken into account during or when changing the design of an *Extended Vehicle* [28].

1. The *OEM* is responsible for the design of an *Extended Vehicle*.
2. The *OEM* is responsible for the design of all interface that will allow communication with the *Extended Vehicle*.
3. The *OEM* decides which functionality is implemented into an *Extended Vehicle*.
4. The *OEM* is responsible for performing an impact assessment for new *Extended Vehicle* functionality over the life-cycle of an *Extended Vehicle*.
5. The *OEM* is responsible for the management of additional risks related to making existing *Extended Vehicle* functionality remotely available.
6. When adding additional remote functionality to the *Extended Vehicle*, the *OEM* is responsible for the management of ramifications by taking into account the existing design.
7. The *OEM* is responsible for the prioritization of all functionalities of the *Extended Vehicle*.
8. When implementing additional functionality the *OEM* is responsible for ensuring backward compatibility of the *Extended Vehicle*.
9. The *Extended Vehicle* design methodology is independently applicable from the physical transmission medium.
10. The *OEM* is responsible for defining and designing *Extended Vehicle* interfaces for a certain use case or use case scenario in a manner that it can be supported by any requester and without discriminating a certain party.
11. The *OEM* is responsible for the validation of the complete *Extended Vehicle* system design, which also includes additional or modified functionality.
12. The *OEM* is responsible for ensuring that the *Extended Vehicle* functionality is neither monitoring the vehicle owner for the means of competition purposes nor violating data protection.
13. The *OEM* is responsible for ensuring that the *Extended Vehicle* functionality is neither monitoring the *Third-Party Service Provider* for the means of competition purposes nor violating data protection.

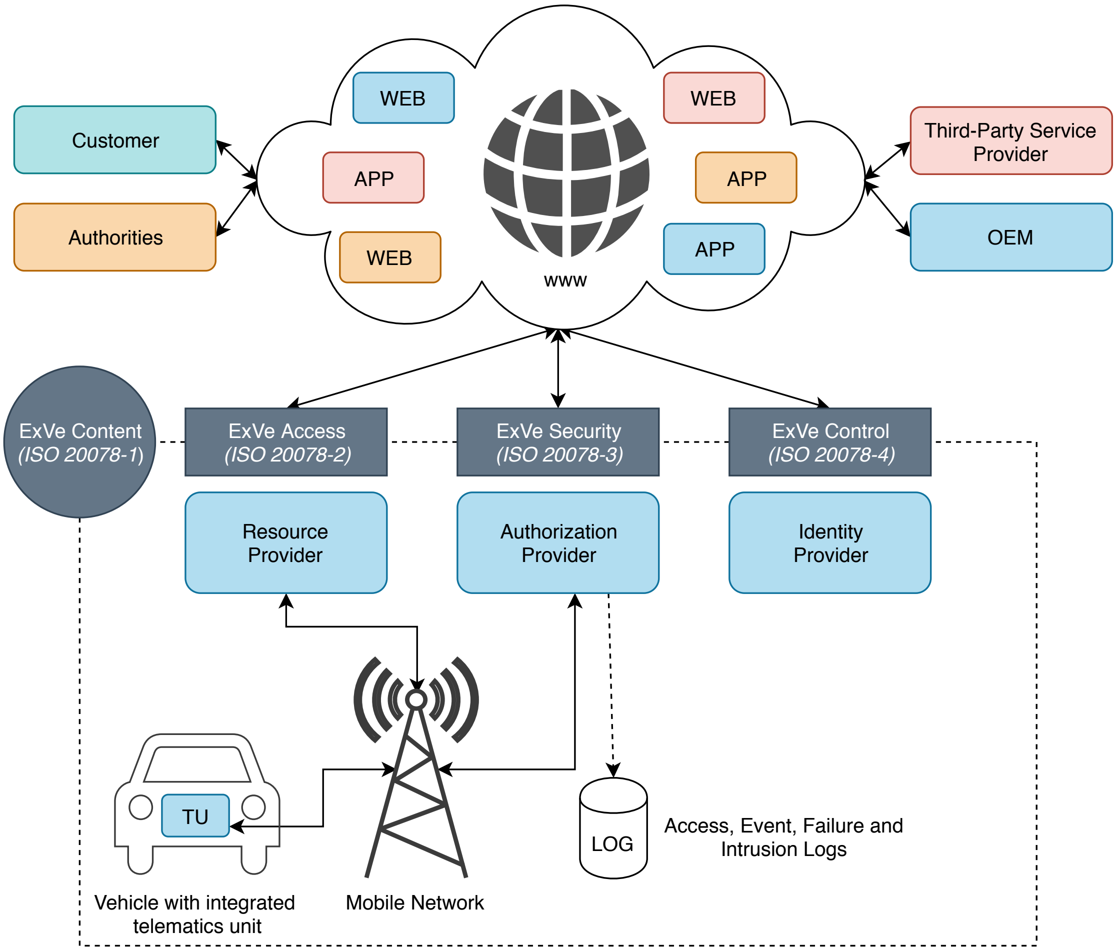

Figure 2.5.: Schematic presentation of the vision of ISO 20078 adapted from [29].

ISO 20078 is a series of standards which is currently under development. The purpose of the standard is to ensure interoperable interface for the access to *vehicle-generated data*. It defines standardized web services for accessing the *Extended Vehicle* defined in the **ISO 20077** standard series [30]. The standard defines requirements on the structure and format of resources (*vehicle-generated data*, aggregated information and functionality from connected vehicles), guideline how unique resources of an individual application shall be defined, entities and roles and how resources are accessed including how defined and referenced technologies shall be used. The standard does not include requirements for specific applications or resource definitions, which are defined in separate standards (e.g. **ISO 20080**). The **ISO 20078-1** standard states minimum requirements, recommendations, permissions and external constraints for interoperable web service based access to *vehicle-generated data*. The document defines all entities and roles that are used over the **ISO 20078** standard series. It defines how an offering party (referred to as *OEM* in this work) defines resources as well as different ways of representing resources in web services (e.g. *JSON* or *XML*). **ISO 20078-2** defines a communication protocol that enables accessing resources. It standardizes how an accessing party (e.g. a Third-Party Service Provider) is able to access resources through a web service. The standard uses the *Hypertext Transfer Protocol* (HTTP) over *Transport Layer Security* (TLS) as communication protocol. Further the *REST* paradigm is specified to be used for the implementation of APIs providing access to resources [29]. The standard also defines asynchronous resource request e.g. for forcing readouts from a connected vehicle [31]. The security model of the web service is standardized in **ISO 20078-3**, the standards includes the different entities and roles involved in the authorization process. The defined role model is a reference implementation of *OAuth 2.0* [5] and *OpenID Connect 1.0* [6]. In **ISO 20078-4** the preceding standards (**ISO 20078-1**, **ISO 20078-2**, **ISO 20078-3**) are summarized with logical processes describing the interaction between all defined roles and entities [29]. Figure 2.5 illustrates a schematic of the **ISO 20078** standard's vision.

Figure 2.5 outlines the intended data acquisition and usage for *vehicle-generated data*. The *Extended Vehicle* web service consists of the road vehicles combined with the telematics backend system provided by the *OEM*. For both Third-Party Service Providers and *OEMs*, acting as a service provider, access to *vehicle-generated data* is offered via the internet. Even if event logging (LOG in Figure 2.5) is an important part of any system, it is not within the scope of **ISO 20078** due to strong implications for certain backend infrastructures [29].

The **ISO 20080** standard defines a first application for the *Extended Vehicle* concept. The standard defines a remote access interface for diagnostic data intended for vehicle repair shops [32].

2.4.2. Neutral Extended Vehicle for Advanced Data Access

The *German Association of the Automotive Industry* (VDA) has developed a concept named *Neutral Extended Vehicle for Advanced Data Access* (NEVADA) which allows the secure exchange of *vehicle-generated data* with third parties or the government. This enables new digital business models as well as an improved road safety. The concept intends to provide a sustainable protection of the vehicles security and the business interests of

[5] `https://tools.ietf.org/html/rfc6749`
[6] `https://openid.net/connect/`

all economic actors who want to access *vehicle-generated data*. The data can be accessed either through *OEM* provided server or through neutral servers. The concept provides a framework for the collection and immediate transmission of *vehicle-generated data* [33]. Essentially, the *NEVADA* concept consists of two building blocks which are data categories and a concept for the provisioning of *vehicle-generated data* towards service providers. To ensure the concept is viable, key elements and other relevant aspects where assessed regarding their effectiveness and feasibility by implementing proof of concepts. The *NEVADA* concept defines five distinct categories for data classification [17]. Table 2.1 provides an overview over the proposed data categories.

Category	Description
1	**Data for road safety improvements** are provided anonymized to public services. Road traffic control centers, police and fire departments can use the provided data to improve road safety [34].
2	**Data for brand independent services** are provided as foundation for the development of new services and innovative new business models as well as for road traffic optimization [34].
3a	**Data for brand specific services** are only available for the *OEM* or selected partners of the *OEM*. The data in this category is generated through competitive innovations in the vehicle, to protect the innovations and investments of the *OEM* the data is only transferred to third parties based on B-to-B contracts [17], [34].
3b	**Data for component analysis and product optimization** is used for component quality assurance and is protected in the same way as category **3a**. The data will be provided to suppliers and development partners of the *OEM* to allow them to meet the requested quality constraints as well as to fulfill potential monitoring obligations [17], [34].
4	**Personal data** is used to provide personalized services to the user of the vehicle. These personally identifiable information are only made available after the explicit consent of the user [34].

Table 2.1.: *NEVADA* Data Categories as described in [17], [34].

Independent of the data category the exchange is based on contracts to ensure solid, controlled, and fair framework conditions. When using *vehicle-generated data* the underlying use case is relevant. Data records can be assigned to multiple categories. The ambient temperature could either be used for detecting black ice (Category 1) or to provide weather information on a map (Category 2). The determining factor is the defined use case class (i.e. data category) in the contract between the data provider (*OEM*) and the data consumer (e.g. Third-Party Service Provider). No specific use case description is required for the defined use case class and is therefore not required to be communicated to the *OEM*. When using data from category 1, 2 or 4 the data format has to be specified alongside with data quality constrains. This ensures that services using the provided data have a solid data pool as foundation. The association of data to the mentioned categories is no onetime process rather the association is adjusted periodically. This prevents a time consuming process where unnecessary work is conducted as it is currently unclear what kind of data records and use cases are requested in a few years. The process of initially defining the data records to be exchanged has been started. For example in cooperation with the *Gesamtverband der Deutschen Versicherungswirtschaft* (GDV) which is the umbrella organization of the german insurance industry [17].

The second building block is the provisioning of *vehicle-generated data*. The data provisioning is exclusively performed through the IT backend of the *OEM*. The main reason is to ensure strong end-to-end security for accessing *vehicle-generated data* and vehicle operation. To ensure a secure and resilient overall system and to mitigate potential threats or attacks, the *OEM* has full control over the vehicle board network as well as over the IT backend infrastructure. Anonymous write access to vehicle function is generally precluded due to product liability regulations and safety risks. Limited write access to specific functions is still possible based on bilateral contracts. The *OEM* data access interface is compliant with the two *ISO* standards **ISO 20077** and **ISO 20078** which are described in subsection 2.4.1. The *OEM* is not allowed to deduce the functionality of a service provided by a third party by analyzing the forwarded data. Another element of the *NEVADA* concept is the possibility to deploy a so called neutral server. Within the scope of bilateral contracts the neutral server collects data from multiple *OEMs* to broker the data to Third-Party Service Providers. This enables comprehensive access to the data of multiple *OEMs*. This reduces the coordination effort Third-Party Service Providers have to conduct, furthermore it allows accessing *vehicle-generated data* without directly contacting the *OEM*. This is possible for anonymized data as well as personally identifiable data. The required permission and data flow have been jointly defined and audited in regards to data privacy laws. Multiple proof of concepts have been successfully implemented to validate the neutral server concept. Due to product liability write access to the APIs is not possible via a neutral server. It is assumed that multiple competing neutral server will be deployed, ensuring an efficient market and fair competition [17]. Figure 2.6 illustrates the neutral server concept.

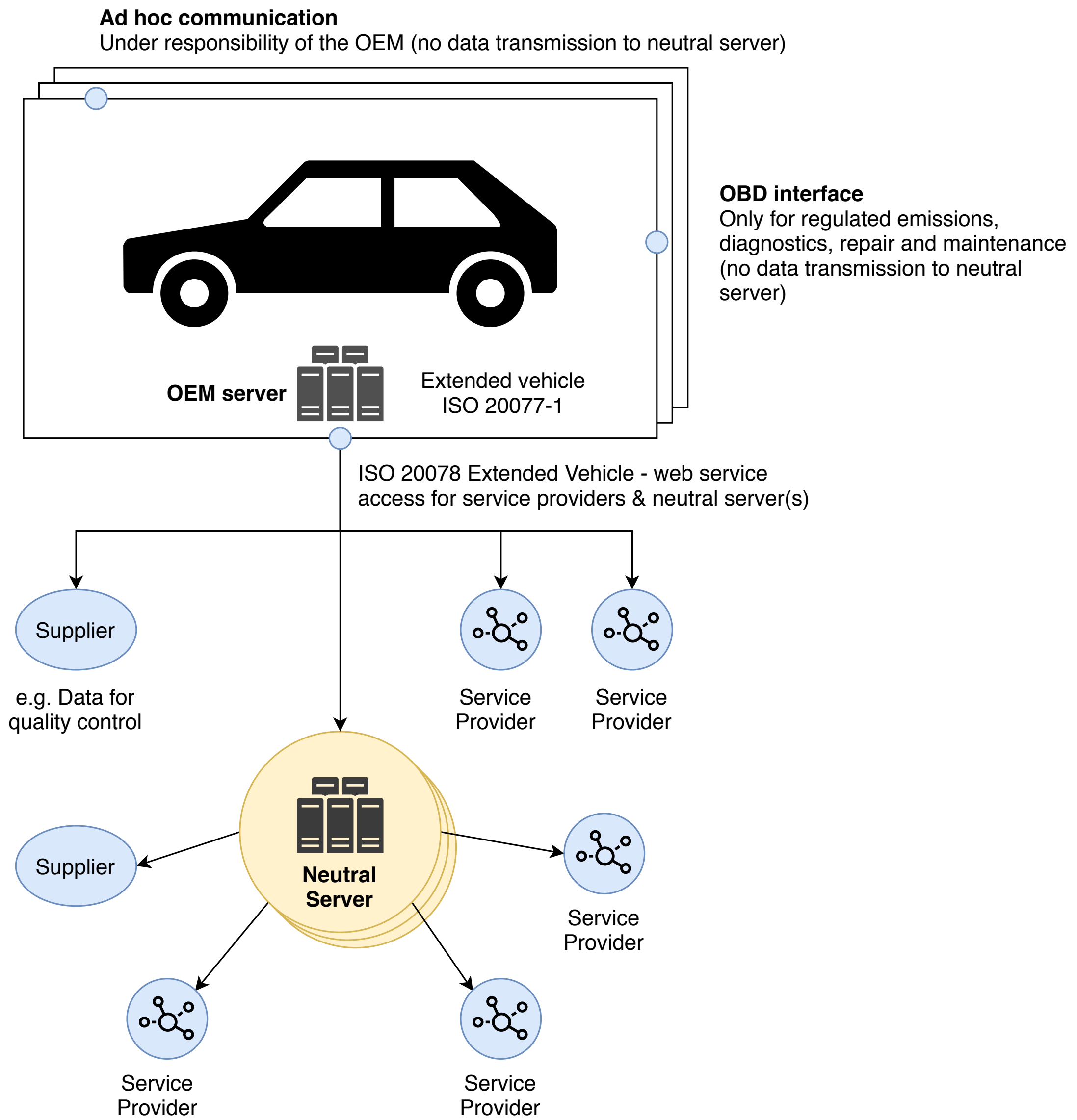

Figure 2.6.: *NEVADA* Neutral Server Concept adapted from [17].

2.4.3. Automat Project

The *Automat Project* is a project funded through the *Horizon 2020 research and innovation programme* [7] of the European Union. The goal of the project is *"[...] to establish a novel and open ecosystem in the form of a cross-border Vehicle Big Data Marketplace that leverages currently unused information gathered from a large amount of vehicles from various brands"*. The bedrock concept of the project is an *OEM* independent data model named *Common Vehicle Information Model* (CVIM) which makes aggregated *vehicle-generated data* available to Third-Party Service Providers. The project has several goals [35]:

- Create an open ecosystem for *vehicle-generated data*
- Single point of access for Third-Party Service Providers
- Definition of standardized and open interfaces
- Definition of the *CVIM* data format which enables a standardized access to *vehicle-generated data*
- Provisioning of a broad spectrum of *vehicle-generated data*
- Definition and prototypical implementation of a *vehicle-generated data* big data marketplace
- Proof of concept based evaluation of the prototype

The project consists of a consortium with different stakeholders from the automotive industry and research community. The consortium includes two *OEMs*, namely *Volkswagen AG* and *Renault SAS* [36].

The project does not rely on the *ISO* standards in subsection 2.4.1 and defines its own data format (i.e. *CVIM*). The project defines no specific data collection process and also relies on the *OBD-II* interface in conjunction with dongles. In contrast to the *NEVADA* concept the project defines more technical details like the *CVIM* Data Package structure and puts its focus on the a central marketplace for the extracted data [37]–[39].

2.4.4. Existing APIs

Today selective *APIs* provided by *OEMs* are available for use. *BMW Car Data* [8] is the first implementation of the *NEVADA* concept and therefore also complies with

[7] `https://ec.europa.eu/programmes/horizon2020/en`
[8] `https://aos.bmwgroup.com/de/web/oss/apps/otp-public`

ISO 20077 and **ISO 20078** [17], [40]. *BMW Car Data* currently implement two use cases. The first use case is for the customer to request all *vehicle-generated data* that has been transmitted to *BMW*. The second use case is, given the customers consent, the transmission of *vehicle-generated data* to a third party [40].

Mercedes Benz provides a developer portal [9] exposing multiple *APIs* which are compliant with the *ISO* standards described in subsection 2.4.1. The *Vehicle Status API* provides state information about the windows, sun root and lights. The *API* is **ISO 20078** compliant. The *Vehicle Lock Status API* is a **ISO 20078** compliant *API* which allows to determine if a connected vehicle is connected or not [41]. Another available **ISO 20078** compliant *API* is the *Pay As You Drive Insurance* which provides access to the odometer. The goal of the *API* is to enable insurances to provide contracts where the customer is charged based on the milage he is driving [42]. The *Fuel Status API* and the *Electric Vehicle Status* are both **ISO 20078** compliant and provide access to the remaining range of the vehicle and the fuel level or the state of charge [43], [44]. All *APIs* require the consent of the vehicle owner as well as their consent to share the data with a third party. The *Remote Diagnostic Support* is a **ISO 20080** compliant *API* allowing the retrieval of which diagnostic capabilities are available for the selected vehicle, which hardware and software features the selected vehicle has, *diagnostic trouble codes* (DTCs) for once specific or all *ECU* and snapshot data associated with a specific *DTC* (e.g. the milage when the error occurred). The service is only available to *"[...] independent operators in the context of the European type approval framework [...]"* [45].

Porsche is hosting the *Porsche NEXT OI Competition* which *"[...] is an open innovation competition to find groundbreaking ideas and bring them to Porsche cars"*. In the scope of the competition, developers are invited to develop applications that interact with simulated *Porsche* sports-car *APIs*. Several SDKs for different platforms are provided by *Porsche* [46]. It remains unclear if the provided data is actually gathered from real vehicles.

[9] `https://developer.mercedes-benz.com/`

3. Research Design

This chapter outlines the research design of this work. As mentioned beforehand two distinct approaches where chosen to answer **RQ2**. Proposed and existing use cases where explored by conducting an extensive literature review, which is described in section 3.1. Alongside the extensive literature review, semi structured expert interviews have been conducted to further explore existing and proposed use cases. Furthermore the results of the interview were used to answer **RQ3** and derive **RQ4**.

3.1. Literature Review

To gain an overview of existing uses cases for *vehicle-generated data* an extensive literature review was conducted. The primary goal of the review was to gather a rough overview about existing and proposed use cases. The secondary objective has been to discover proposed taxonomies used for classifying use cases which are based on *vehicle-generated data*. An initial set of keywords for the search was defined including the terms *vehicle-generated data*, *car data*, *vehicle data*, *connected vehicle*, *floating car data*, *usage-based insurance*, *OBD*, *OBD-II*, *in-vehicle sensors* and *car sensor data*. Scientific search engines, like Google Scholar [1] and Scopus [2], have been used to find an initial set of literature. Subsequently, the results have been scanned for relevant content. If a result was relevant it either was used directly, or a forward and backward search has been conducted as proposed by Webster & Watson (2002) [47]. Furthermore relevant results have been triaged for other relevant keywords which in turn have been used to conduct a further web search. The search was halted after no more significant different use cases have been discovered. The results of the extensive review are presented in section 4.1.

[1] `https://scholar.google.de`
[2] `https://www.scopus.com/`

3.2. Interview Methodology

The qualitative research approach of this work is to some extent based on the *grounded theory methodology* (GTM) [48]. The target of *GTM* is aimed at theory construction by providing a set of systematic, but flexible, guidelines [49]. Different schools of thought influenced the *GTM* over the years leading to a certain divergence in methodology [49]–[51]. This work follows the relatively newer constructivist approach by Charmaz (2000) [52]. In contrast to the assumption that data or theories are discovered, the approach assumes that the researcher constructs both. The research process is accompanied by a researcher who plays an active role developing a dialog with the data (i.e., interviewees) [49]. In the field of information systems, *GTM* is frequently used for qualitative research [51], [53]. *GTM* aims to contribute to research in a different form, of which one is the *rich description* of empirical observations without providing an abstract explanation [51], [53]. A *rich description* is based on a systematic exploration of different depictions of the phenomena (e.g. interviews, archival materials, etc.) [53]. Hence the *rich description* is a good methodology fit for this work as it describes use cases, barriers, and solutions based in the context of the topic by the conduction of a literature review (use cases) and performing expert interviews (use cases, barriers, potential measures).

For the qualitative data collection, semi-structured interviews with 19 participants were conducted. When performing semi-structured interviews, the researcher asks the interviewees a set of predefined but open-ended questions. This method gives the researcher more control over the interview than unstructured interviews, while still allowing answers which are not in a predefined set of responses which is the case for interviews that use closed questions. Before conducting the interviews a written *"interview guide"*, is developed. This guide is used during the interview by the interviewer to direct the conversation and move back and forth based on the interviewee's responses. The topics of the guide are based on the research questions of the work the interviews are contributing to. The questions should not imply the response or lead to a simple binary response (i.e., yes or no) [49], [54]. The interview guideline which was used for this work is elucidated in subsection 3.2.1. The interview guideline and format was evaluated by conducting test interviews with three members of the chair. The feedback was subsequently implemented into the guideline to improve it.

Theoretical sampling refers to the continuous data collection process in *GTM* based research. The data collection is stopped when the theoretical saturation is reached. The goal of the process is to understand the nature and dimensions of the concept rather than identifying a representative population. The concept of theoretical saturation supports the researcher to focus on the field and to proceed the data collection until new construct cease to occur and, the next analysis steps are feasible [53]. Due to the strict time constraints of this work this strategy was not viable to its full extent, still, the initial set of interview participants was selected by utilizing the expertise and past experiences of members of the chair. The focus of the sampling was distinctly targeted on persons associated with the automotive industry. After conducting the first interviews, the results were discussed, and the format of the interview was further improved. New concepts, which were discovered within an interview, were further examined in succeeding interviews to discover how prevalent certain concepts and theories are within the automotive industry and its stakeholders (e.g., *NEVADA*). Besides, the sample was extended with contacts referred by the interviewees.

Except for one interview, an audio recording has been created for every interview. The recording has subsequently been transcribed and anonymized regarding personal information or any information that allows inferring the interviewee or the organization the interviewee is working for. The analysis of the data follows the *iterative coding* approach used in *GTM* research. In this context, coding refers to the systematic process of building theory from data. The level of detail for the coding process is defined by the research interest and the nature of the data [53]. The coding process helps to structure and analyze the transcribed interviews. The software *MAXQDA2018* [3] was used to conduct the coding. The data was coded by conducting multiple iterations with the goal of discovering all use cases, barriers and potential measures mentioned by the interviewees and eventually a taxonomy was derived from the results.

Constant comparison is used in *GTM* to analyze the data from different viewpoints, this is done by writing memos that act as a pivotal point for the comparison, emergence, sampling, and theoretical densification. Memos help to understand the data and its relations. The form of the memos is based on the preference of the researcher (e.g., diagrams, mind maps, text narrative, etc.) [53]. This thesis uses memos to summarize or paraphrase the statements of the interviewees, derive taxonomies, compare different viewpoints of the participants. The results of the analysis are summarized in chapter 4, chapter 5 and chapter 6.

3.2.1. Interview Guideline

The interview guideline used by the interviewers consists of three coarse parts. The first part contains an introduction and formal information about the interview. In the second part the interviewee is inquired to provide basic information about him/her

[3] `https://www.maxqda.de/`

and the organization he/she is working for (see Table 3.1). The third part contains the topic related question guideline. The guideline was not provided to the interviewee before or during the interview.

The first part starts with a short explanation of the problem statement. Afterwards two working definitions are introduced to ensure the interviewers and the interviewee are on the same level of knowledge. The two defined terms are *API economy* and *vehicle-generated data*, the detailed definition can be found in subsection 2.1.2 resp. subsection 2.1.3. This is followed by formal information for the interviewee like the terms of confidentiality, consent for tape recording and contact information of the interviewers.

In the second part the interviewee is inquired to provide non-personal information about him/her and his/hers organization. The initial question was to choose a classification for the organization he/she is working for. Afterwards the size of the interviewee's organization was prompted and classified. Eventually the interviewee was asked to state his role in the previously determined organization, the interviewers tried to match the role as closely as possible to a set of common roles.

The actual content related questions are in the last part of the guideline, which itself can be separated into three different sections. The first section is concerned with past experiences of the interviewee in the context of *vehicle-generated data* and specific use cases. Despite binary questions should be avoided [54] the first question is used to check if the interviewee has experience with *vehicle-generated data* at all. The answer is used to select a different branch of questions in the first part. The second part of the guideline is about future expectations to *vehicle-generated data*, potential use cases and barriers. The interviewee is asked for barriers which are specific for use cases he/she mentioned and for general barriers. Depending on the extend of the interviewees response, the interviewers provided four different fields of barriers **a)** Legal **b)** Social Acceptance **c)** Technology **d)** Economic to help the interview participant to view the problem in a more holistic manner. The question was succeeded by questions related to rating and elucidating the severity of the barriers mentioned before. The final section of the guideline contains questions about potential measures and the corresponding timeframe for an implementation of the measure. The complete interview guideline can be found in the appendix of this work.

3.2.2. Interview Participants

A total of 18 interviews with 19 distinct interviewees has been conducted. With the exception of one interview all interviews have been recorded and transcribed. The total duration of all recorded interviews was 10 Hours 15 Minutes 56 Seconds, interview 1 was the shortest interview conducted and lasted 23 Minutes 10 Seconds. The longest interview took 57 Minutes and 49 Seconds. The average interview duration was 36 Minutes and 14 Seconds with a standard deviation of approximately 11 Minutes and 18 Seconds. Table 3.1 contains an overview over all interview which have been conducted. The interviews where undertaken by the author of this work (abbreviated as *FK* in Table 3.1) and the advisor of this work (abbreviated as *GB* in Table 3.1).

#	**Organization Classification**	**Role**	**#Employees**	**Duration**	**Participants**	**Interviewers**
1	Consulting	Technical Architect	251 - 500	0:23:10	IV1	FK, GB
2	Mobility Start-up	CTO	1-10	0:56:05	IV2	FK, GB
3	Consulting	Head of Department Automotive	501 - 1000	0:57:49	IV3	FK, GB
4	Consulting	Digital Innovation Officer	251 - 500	0:32:10	IV4	FK
5	Mobility Start-up	Senior Partner Manger	51 - 250	0:27:10	IV5	FK
6	Insurance	Technical Architect	>100.000	0:28:54	IV6	FK, GB
7	Automotive supplier	Both Business Development	>100.000	0:34:58	IV7.1, IV7.2	FK
8	Automotive Association	Head of Department IT	51 - 250	0:42:53	IV8	FK
9	Automotive supplier	Project Manager	50.001 - 100.000	-	IV9	FK, GB
10	Consulting	Head of Department IT	251 - 500	0:44:17	IV10	FK
11	Automotive supplier	Head of Department IT	>100.000	0:48:32	IV11	FK, GB
12	OEM	Technical Architect	50.001 - 100.000	0:44:16	IV12	FK, GB
13	Mobility Start-up	Head of Department IT	11 - 50	0:30:16	IV13	FK, GB
14	OEM	Business Analyst	>100.000	0:20:14	IV14	FK, GB
15	Automotive Association	Product Manager	5001 - 10.000	0:37:50	IV15	FK, GB
16	Mobility Start-up	Data Analyst	11 - 50	0:36:25	IV16	FK, GB
17	Finance	Managing Director	251 - 500	0:27:25	IV17	FK
18	Insurance	Technical Architect	>100.000	0:23:32	IV18	FK, GB

Table 3.1.: Interview Participants

The distribution of the interviewees into the organization classifications introduced in subsection 3.2.1 is illustrated in Figure 3.1. Two interviewees work for an *Automotive Association*, three for an *Automotive Supplier*, four for a *Consulting* Firm, one for a *Finance* Venture, two for an *Insurance*, four for a *Mobility Start-up* and two for an *OEM*.

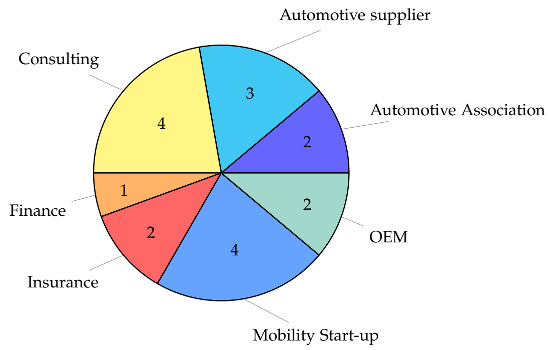

Figure 3.1.: Organization Classification distribution of the interview participants

The distribution of the number of employees of the organization for which the interviewees work is illustrated in Figure 3.2. Five interviewees work for organizations with more than *100.000* employees, two work for organizations with more than *50.001* and less than *100.000* employees, two works for an organization with more than *5001* and less than *10.000* employees, two work for an organization with more than *500* and less than *1000* employees, four work for organizations with more than *251* and less than *500* employees, two work for organizations with more than *51* and less than *250* employees, two work for organizations with more than *11* and less than *50* employees and one works for an organization with less than *10* employees.

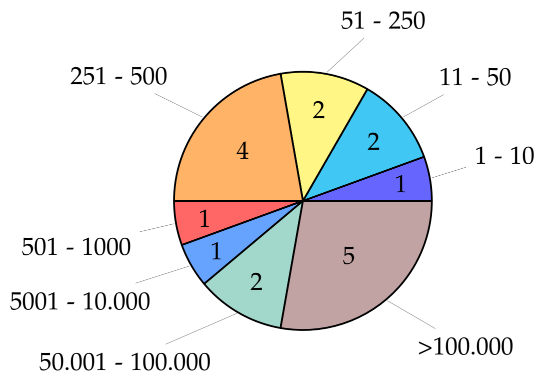

Figure 3.2.: Distribution of the number of Employees of the interviewees organization

Figure 3.3 illustrates the interviewees role distribution. If the interviewee stated a role which was in the list described in subsection 3.2.1 the corresponding role was used, otherwise the stated role was used. Four interviewees each stated that their role was *Head of Department IT* and *Technical Architect*. Two interviewees had a role in the *Business Development* of the organization. All other roles were only represented once.

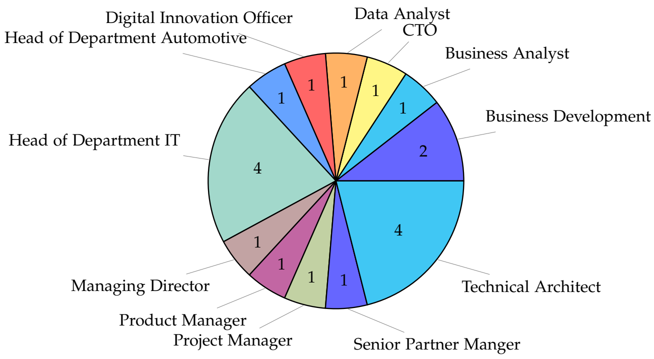

Figure 3.3.: Role distribution of the interview participants

4. Use Cases

This chapter discusses several use cases for *vehicle-generated data*,independently of the data-collection method (e.g., *OBD-II* dongle). In section 4.1 the results of the extensive literature review conducted, are presented. The succeeding section 4.2 discusses the results of the semi-structured expert interviews.

4.1. Literature Review Results

A very concise overview about use cases for vehicle generated data can be found in a report of the consulting firm *McKinsey & Company* from 2016. The report distinguishes between three macro categories of value creation models for *vehicle-generated data*. The **Generating revenues** category includes use cases that can be directly monetized and are related to tailored advertising or to selling data. Use cases that lead to R&D and material cost reduction, reduction of the customer's costs or the improvement of the customer satisfaction are categorized as **Reducing costs**. The category **Increasing safety and security** includes use cases that reduce the time for intervention by sharing critical information in advance [26]. Table 4.1 provides an overview over the proposed use cases.

De describes in his paper from 2018 three different supercategories based on the intended benefit for the users of the vehicle. **Passenger safety and vehicle security** includes basic use cases like stolen vehicle recovery, emergency call, and car tracking as well as location-based services like vehicle navigation, traffic alerts, weather forecast based traffic advisory or generic geo-fencing applications. The second category **service and vehicle quality** use cases and services are related to after-sales services and vehicle quality. Typical use cases in this category are remote diagnostics, service reminders & updates, eco-driving feedback, battery charge monitoring or the provisioning of location-based air pollution data. The third category is **comfort and convenience services for users**. This category includes, amongst others, services like automatic tolling, usage-based insurances, in-vehicle purchase and payment and real-time parking availability info [55].

Generating revenues	**Reducing costs**	**Increasing safety and security**
• Over-the-air software add-ons • Usage-based tolling and taxation • Networked parking service • Gamified/social-like driving experience • Tracking/theft protection service • Fleet management solutions • Vehicle usage monitoring and scoring • Remote car performance configuration • Connected navigation service • In-car hot spot • Onboard delivery of mobility-related contents/services • Onboard platform to purchase non-driving-related goods • Predictive maintenance • Targeted advertisements and promotions	• Warranty costs reduction • Data-/feedback-based R&D optimization • Traffic-data-based retail footprint and stock level optimization • Usage-based insurance: Pay-As-You-Drive/Pay-How-You-Drive • Driving style suggestions • E-hailing • Early recall detection and software updates • Car pooling • P2P car sharing • Trucks platooning	• Driver's condition monitoring service • Improved road/infrastructure maintenance and design • Breakdown call service • Emergency call service • Aggregated car data-based CCTV service • Road laws monitoring and enforcement

Table 4.1.: Vehicle-generated data use cases described in [26]

Willke et al. (2009) propose a tree bases taxonomy for inter-vehicle communication applications. The branches of the tree describe the common reason for communication with the assumption that applications in this category have common communication requirements. The leaves of the tree map directly to use cases. Figure 4.1 contains the proposed taxonomy [56].

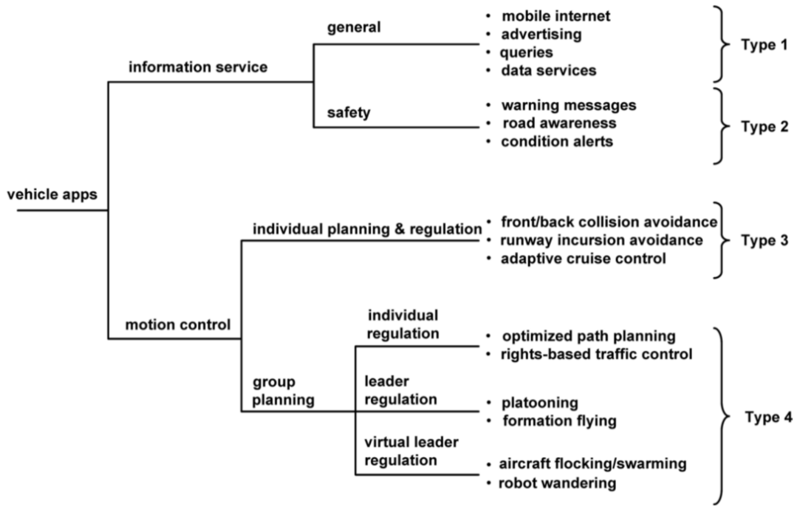

Figure 4.1.: Taxonomy for Inter-Vehicle Communication Applications as described within [56]

Siegel et al. (2018) adopted Willke's taxonomy to use it in the context of connected vehicles. The paper differentiates between **Telematics Applications**, which collect vehicle data locally and transmit it to remote services, and **Vehicle-to-Vehicle and Vehicle-to-Infrastructure Applications** which directly communicate with other vehicles or the infrastructure (e.g., Roads) itself. **Telematics Applications** include consumer related safety and security services like roadside assistance, collision notification, or remote door unlocking. The data can also be used to collect information about the traffic flow. Furthermore, the paper mentions Pay-As-You-Drive as a potential insurance use case and vehicle miles traveled tracking, which is a model where the driver is charged based on the mileage. For use cases related to the **Vehicle-to-Vehicle and Vehicle-to-Infrastructure Applications** category, the paper uses a modified taxonomy of Willke's, that distinguishes between four categories, namely **Information Services**, **Safety Services**, **Individual Motion Control** and **Group Motion Control**. Another approach is to distinguish between **Non-Critical** and **Operation Critical**. Applications within the **Information Services** category include, among others, fault prediction and response (i.e., predictive maintenance), autonomous driving, internet sharing between vehicles, traffic flow analysis, automated traffic routing & tolling. **Safety Services** are intended to mitigate the risk of hazards. This category includes applications for collision avoidance, hazard reporting, and driver monitoring (e.g., driver drowsiness detection). Applications associated with the **Individual Motion Control** describes applications that issue warnings to the vehicle operator or directly control a single vehicle (e.g., collision avoidance, assisted lane switching). The last category **Group Motion Control** includes application like vehicle platooning and intersection control to maximize vehicle throughput and reduce air pollution [57].

In a literature review from 2007, Cassias and Kun described several use cases for *vehicle-generated data*. The use cases are subdivided into seven different categories. Services in the **Navigation** category include *Global Positioning System* (GPS) related applications like turn-by-turn navigation or the identification of new roads by continuously analyzing collected *GPS* information and incorporating it into street maps. **Remote Diagnostic** services include maintenance-related applications like monitoring the tire pressure of vehicle fleets or remotely accessing *DTCs*. The **Fleet Management** category comprises use cases like e-hailing, utilization optimization (e.g., for logistics) and location tracking of individual vehicles. **Safety** related use cases include hazard avoidance, crash detection, cooperative collision avoidance (in the context of platooning). Services which provide **Information Access** to the vehicle occupants include entertainment, safety, and travel-related applications. The category of **Context Awareness** includes use cases like environmental monitoring (e.g., pollution maps or weather information), traffic information services. The last category **Mobile Commerce** describes use cases allowing the facilitation of business transactions, e.g. usage-based insurance, tolling, parking fees [58].

Different authors have proposed safety-related use cases over time. In 2009 Hauschild described a system for advanced hazard warnings. The system uses *vehicle-generated data* to detect weather-related road hazards (e.g., black ice), road properties (winding, inclining or declining), road properties (e.g., potholes), infrastructure (e.g., traffic light, intersections or one-way roads) or points of interest (e.g., gas station, tollbooth). The results of various, so-called, primary detection modules is then used with secondary detection modules to detect the traffic state and send the information to nearby vehicles or a central system [59]. Abdelhamidad et al. (2014) also suggest using the environmental sensors of a vehicle to detect the weather status or road and traffic conditions. The processing could either be performed distributed by each vehicle or centrally at the collecting system [13]. Using a combination of third-party devices (i.e., *OBD* dongle and smartphone) to perform road condition monitoring is proposed by AbuAli (2015). The proposed system can monitor three distinct types of events.

Driver behavior events like speeding, hard braking, drifting or weaving through traffic can be detected using *vehicle-generated data*. The system also detects road artifacts like slippery road conditions, sand drifts, speed bumps, and road deterioration. Further, the system can perform accident detection by monitoring sudden stops or vehicle rollovers [60]. Pillmann et al. (2017) suggest that the large scale collection and analysis of *vehicle-generated data* increases the quality of road quality monitoring. In contrast to other solutions that only use the smartphone's accelerometer, the usage of multiple vehicle sensors significantly increases the data quality [37]. Correlating sensor data with accident databases could also further improve road hazard warning systems [61]. Ni et al. (2017) propose that road surface monitoring should be used to improve the infrastructure by providing transportation agencies with the acquired data. The data could be stored at a central system which also prioritizes road repairs according to their severity. Another proposed use case, by the same authors, is traffic collision reconstructions. Post-Accident investigations of law enforcement agencies could be supported not only with data from the involved vehicles but also with data from nearby vehicles (e.g., videos) [62]. Another law enforcement related use case is described by Lee et al. (2006). The MobEyes system uses vehicular resources (i.e., cameras) to scan license plates of nearby vehicles. The gathered information is distributed within the sensor network if certain predefined conditions match (e.g., if the license plate matches a wanted vehicle) [63].

Guerrero-Ibáñez et al. (2018) describe several *Intelligent Transportation System* (ITS) related use cases. A taxonomy with six distinct categories for ITS related applications is defined in the article. The **Safety** category enfolds use cases which increase the safety of the vehicle occupants. Applications include lane keeping assistance, adaptive cruise control, road hazard warning, intersection collision warning. Use cases classified as **Traffic Management** are intended to improve the traffic flow. Applications include automatic tolling, lane management, surveillance, parking management, and intersection management. The **Diagnostic** category includes use cases related to the detection of component failures of the vehicle. The authors suggest using a *"personalized vehicle registry"* keeping track of the status of each vehicle part. This approach would enable predictive maintenance related use cases. Use cases which are related to the monitoring of road conditions are categorized as **Environment**. It is suggested to use various sensors of the vehicle to create a map with road surface anomalies like potholes or speed bumps. The obtained information can also be used to schedule infrastructure maintenance. The **User** category enfolds use cases related to the monitoring of the driver's performance and behavior. Applications include drowsy warning systems, health monitoring, and emotion recognition. Use cases which increase the convenience for the driver are associated with the **Assistance** category. Example applications are tourist guide applications, parking spot locator applications or dynamic navigation routing based on traffic [64].

Using the environmental sensors of the vehicle for creating local microclimate urban emission models is proposed by Massaro et al. (2017). Using this data could help to analyze the climate change at a larger scale. Using *vehicle-generated data* for weather measurements could significantly improve the accuracy of weather and environmental sensing [62]. Pillmann et al. (2017) also suggest using *vehicle-generated data* for providing hyper-local weather predictions. Local weather phenomena are often unnoticed by observations sites as they are usually multiple kilometers apart. Therefore, the crowd-sourcing of such data could enhance certain weather models [37].

Using low-cost third-party hardware Türk and Challenger (2018) describe as fleet management system allowing to monitor the speed and position of connected vehicles [65]. Yun et al. (2011) propose an *XML* based data exchange protocol for remote inspection and maintenance. The system uses the standardized *OBD* interface for data extraction. The authors state that using remote *OBD* helps to identify problems quickly and therefore leads to faster repairs [63]. Khorsravinia et al. (2017) present an electric vehicle monitoring system which provides the user of an electric vehicle with more detailed feedback about the vehicle's status. The system also uses the *OBD-II* interface the acquire the relevant data [66]. A related use case is presented by Tseng et al. (2016), the authors use the *OBD-II* interface of electric vehicles to evaluate the carbon footprint of the vehicle [67].

Ni et al. (2017) propose to use *vehicle-generated data* to implement the use case of parking navigation. Vehicles searching for a place to park in congested areas, especially in peak hours, increase traffic congestions, air pollution, vehicle accidents, and waste fuel. The authors propose to use video cameras and in-vehicle fog nodes to detect available parking spaces. The data is stored in a central system, allowing other vehicles to access the data when searching for a vacant parking spot [62].

Many authors have proposed usage-based insurances, furthermore, there are already offers for consumers on the market, albeit they all rely on third-party devices (i.e., smartphone or *OBD* dongle). The current models in Germany offer discounts for safe driving rather than charging more for reckless driving [68]. The prevalent differentiation for usage-based insurances is made between Pay-As-You-Drive and Pay-How-You-Drive. The Pay-As-You-Drive model charges the user based on the mileage driven, while Pay-How-You-Drive based insurances charge or offer discounts based on the driving style [26], [69]. Husnjak et al. (2015) emphasize the importance of rich telematics data to enable the precise and resilient creation of custom insurance policy models [69]. Streich (2018) analyzed the different usage-based insurance models by conducting a survey among 215 people, with the result that there are still privacy concerns inhibiting the user acceptance of such models [70].

4.2. Interview Summary

The following section elucidates the different use cases that were mentioned in the interviews. This section is subdivided into existing use cases and use cases that have been proposed by the interview participants for the future.

4.2.1. Existing Use Cases

Throughout conducting the interviews several use cases which are already implemented have been identified. This subsection provides a detailed overview of the use cases.

Five interviewees mentioned insurance related use cases which are already implemented. Two different usage models where mentioned. The Pay-How-You-Drive model charges the driver based on his driving style [IV03], [IV12], [IV13], [IV16]. The Pay-As-You-Drive charges the user based on the driven milage [IV03], [IV06], [IV13], [IV16]. For the mentioned use cases the data retrieval is conducted through third party devices [IV03], [IV06], [IV12], [IV13], [IV16].

Use cases in the area of optimizing the research and development process where mentioned by three distinct interview participants [IV09], [IV13], [IV14]. The obtained data is either used by automotive suppliers to optimize their components [IV09], [IV13] or directly by the OEM itself [IV14].

In the context of carsharing *vehicle-generated data*, is used for a multitude of use cases. The data is used to detect when for example the windshield wiper water or the engine coolant has to be refiled. This use case is commonly referred to as predictive maintenance. Furthermore, the data is used for billing related use cases [IV01], [IV04]. Very similar to carsharing, but more generic, are fleet management use cases. Internally an OEM uses *vehicle-generated data* to monitor and track test vehicles [IV01]. Another interview participant is using data extracted via the *OBD-II* interface to offer fleet management and driver's logbook applications. This is done using an *OBD* dongle with an integrated subscriber identification module (SIM) to transmit the data via the cellular network [IV05]. Data obtained from a vehicle fleet also allows inferring if a vehicle could be replaced with an electric vehicle [IV16].

One *OEM* is using the data of the ultrasonic sensors to detect free parking spaces while the vehicle is driving along the street. At least one other *OEM* is using *vehicle-generated data* to infer the traffic situation around the vehicle. This information is then used to provide traffic information in the *OEMs* navigation system [IV11].

Of course, *vehicle-generated data* is also used in its original purpose to provide diagnostic information through the *OBD-II* interface. The primary use case described in the interview is to provide roadside assistance [IV15].

In the context of electromobility *vehicle-generated data*, is used to determine where the electricity supplier should place electric vehicle charging stations. This is done by analyzing the position and the time when a vehicle is parked [IV16].

Another use case is milage driving leasing which is a different leasing model where the monthly leasing installment is not fixed but depends on either the driving style or the mileage driven [IV17].

Security-related use cases include using *vehicle-generated* data to implement theft protections applications [IV03]. In the domain of personal security one interviewee described a use case where *vehicle-generated* is used to monitor the location of company executives [IV01].

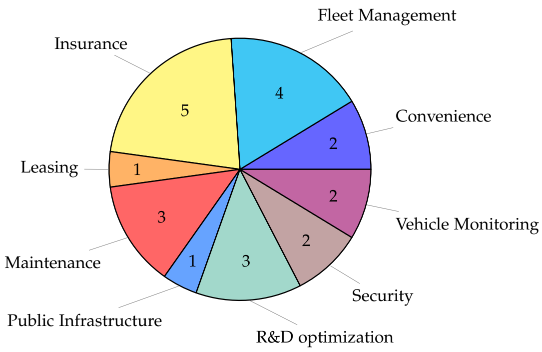

Figure 4.2.: Macro Categories of existing use cases

Nine distinct macro categories have been identified based on the interview responses. Figure 4.2 illustrates the severity of each category. In this context, the severity is defined as a unique use case mentioned at least once by a distinct interview participant. The **Insurance** category clusters all insurance related use cases (e.g., Pay-As-You-Drive, Pay-How-You-Drive). **Fleet management** includes use cases like carsharing, test fleet management, and driver's logbook. The **Maintenance** category covers predictive maintenance and other diagnostic applications. **R&D optimization** includes all use cases related to product optimization as mentioned above. The **Vehicle Monitoring** category includes use cases which enable the user to gain detailed insights into the vehicle's sensor readings. Use cases like parking place detection, in-vehicle payment services or traffic flow maps are classified as **Convenience**.

4.2.2. Proposed Use Cases

Along the existing use cases the interviewees were asked to elaborate on potential future use cases that would be possible if there would be widespread access to *vehicle-generated data*.

One prevalent area of use cases is research and development related. For the *OEM* as well as for the automotive suppliers it is helpful for future product iterations to know which flaws the current generation has or how the attrition of certain components is progressing [IV09], [IV10]. It is challenging to diagnose the failure reason due to the attrition of certain parts of the vehicle, because the driving behavior of the vehicle user is mostly unknown. Using *vehicle-generated data* could improve the diagnosis quality and root cause analysis. Furthermore, it makes it possible to notify the driver if some part is about to fail [IV11]. *OEMs* intend to use *vehicle-generated data* to track how the driver is using the vehicle to improve the customer experience and ergonomics of the vehicle interior [IV12], [IV14]. From the viewpoint of one automotive association, the data could also be used to analyze typical patterns of damage and provide certain recommendations to the consumer [IV15]. The customer side of this area is the use case of predictive maintenance. Using *vehicle-generated data* to predict failures and notify the driver or the fleet manager could significantly reduce repair costs (i.e., by preventing cascading failures) and decrease the mean time between failures [IV01], [IV07.1], [IV09], [IV10], [IV11], [IV15], [IV16].

Several safety-related use cases have been proposed. Aggregating *vehicle-generated data* of multiple vehicles in the same area could be utilized to create a warning about potential road hazards like black ice or oil on the road. The data could also be used to inform road maintenance operators which then can deploy appropriate measures [IV08]. In the context of fleet management, accident detection is another proposed use case [IV01]. An automotive association's representative suggests that the data could also be used for roadside assistance-related use cases, this could be in the form of remote diagnostic support, enabling the roadside assistance provider to dispatch an appropriate support vehicle (e.g., a tow truck) [IV15].

Using *vehicle-generated data* to offer precisely fitting insurance products to the consumer is another major use case. The Pay-As-You-Drive and Pay-How-You-Drive models have also been mentioned as potential use cases [IV01], [IV15], [IV18]. It is also speculated that by analyzing the driving style of a particular person its risk attitude could be derived, this information can, in turn, be used to determine the monthly fees of certain property insurances or even life insurance fees. Therefore the data is desirable for insurance companies [IV06].

For the development of autonomously driving vehicles, a huge amount of data is necessary to safeguard its operation. Therefore the collection of *vehicle-generated data* is especially important for *OEMs* developing autonomous vehicles [IV12], [IV14]. Further, it is speculated that it is necessary for *OEMs* to share the data among themselves to advance autonomous vehicles [IV01]. In the context of fleet management *vehicle-generated data* could be used to predict failures (i.e., predictive maintenance) and in turn direct the autonomous vehicles to drive to the next car repair shop. This could for example help to increase the utilization of carsharing fleets and therefore increase the revenue [IV12].

Multiple use cases directed at the consumer have also been proposed. *Vehicle-generated data* could be used to analyze the driving style of a consumer and generate suggestions to drive more economically or more ecologically [IV15], [IV16]. This could also be used to suggest alternate routes for commuters based on the current air pollution [IV16]. Another use case are reward-programs based on how employees use company cars if an employee is using a vehicle more carefully than other employees he/she could be rewarded by the employer for saving costs. The driving style could also be used by ride-sharing applications providing a ranking of the driver for the fellow passengers [IV01]. Another use case is location-based advertising which could be used for a variety of different scenarios. Using the *GPS* location of a vehicle could be used to advertise for nearby offers [IV03], [IV08], [IV07.1] or in combination with predictive maintenance if some component of the vehicle is about to fail nearby car shops with matching offers could be advertised. Furthermore, the data could be used to track the conversion rate of the advertisement (i.e., if the consumer is driving to the advertised location) [IV16]. Another use case is intended for the second-hand car market, *vehicle-generated data* could be used to create a journal capturing all events occurring over the lifecycle of a vehicle (e.g., battery charging cycles, accidents or repairs) [IV13]. The journal could either be kept at some neutral central entity [IV04], or it could be stored distributed inside a blockchain [IV16]. The sensors of a vehicle (e.g., ultrasonic) could also be used to detect parking spaces. This data could be aggregated and integrated into navigation systems to provide the driver with information about places to park [IV01], [IV04], [IV07.1].

Vehicle-generated data related use cases have also been proposed to use for the optimization of public infrastructure. The data could be integrated into intelligent transportation systems to make the road network more resilient. For example the data could be used to improve the traffic light circuits [IV01], [IV12], [IV17]. This data could also be used by the cities and municipalities to optimize the placement of road signs, roads and prevent accidents by mitigating for example dangerous intersections [IV04]. Carsharing providers could use the data to infer the demand for certain areas and adopt their areas of operation and fleet size accordingly [IV14], [IV17].

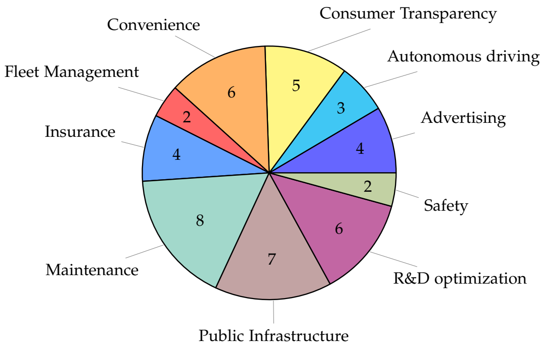

Figure 4.3.: Macro Categories of the proposed use cases

Based on the use cases mentioned in the interviews ten distinct macro categories have been identified. Figure 4.3 provides an overview over uniquely named use cases which have been associated with the corresponding macro category. **Maintenance** is the prevalent use case category and also includes use cases in the area of predictive maintenance. Use cases like infrastructure monitoring, carsharing and traffic flow optimization are consolidated in the **Public Infrastructure** category. **R&D optimization** includes all use cases related to product optimization as mentioned above. Use cases like driving style suggestions, parking place detection or traffic jam avoidance are consolidated into the **Convenience** category. Insurance-related usage of the data is grouped in the **Insurance** category. The **Consumer Transparency** category summarizes all use cases that provide information about the vehicle or its data and the corresponding history (i.e., car usage log or driving style feedback). Roadside assistance or accident detection are classified as **Safety** related. The numbers in Figure 4.3 indicates proposed use cases mentioned by unique interview participants.

5. Barriers

The following chapter discusses the barriers which inhibit the implementation of the use cases described in chapter 4 as well as general barriers which obstruct the advancement of an API economy within in the automotive industry. Accordingly this chapter answers research question three. The insights presented in this chapter have been obtained from the results of the expert interviews described in section 3.2.

5.1. Taxonomy

In subsection 3.2.1 four distinct fields of barriers have been described, which where used as guidance mechanism for the interview participants. The four categories, **a)** Legal **b)** Social Acceptance **c)** Technology **d)** Economic , also served as a pivot when coding the interviews. Hence a taxonomy has been developed based on that four categories, which was continuously extended while coding the interviews. Figure 5.1 illustrates the final taxonomy. The original four categories are marked in green, nodes with a dashed border serve as an intermediary and have not been used to directly associate codes. The blue circles in the top right corner indicate how many distinct interviewees mentioned the particular barrier. Circles with a dashed border are the sum of mentioned barriers under the specific subtree.

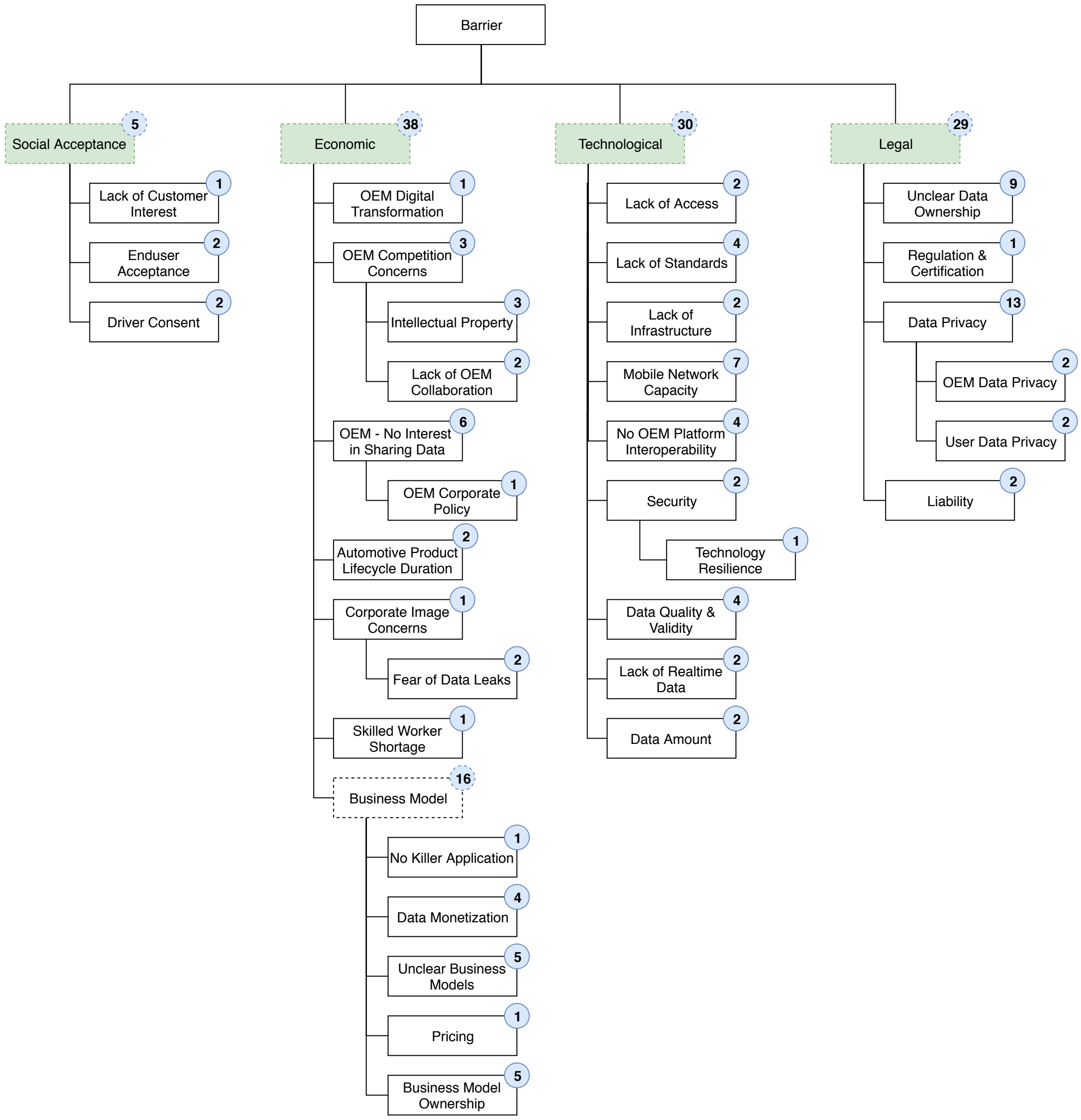

Figure 5.1.: Taxonomy of the Barriers discovered during the Expert Interviews

5.2. Discovered Barriers

The following section provides a detailed insight about the barriers inhibiting the advancement of an API economy within the automotive industry. The section is subdivided into the four categories described in section 5.1. During the interviews, 34 distinct barriers have been mentioned across all four categories. Figure 5.1 illustrated the top then barriers mentioned by distinct interviewees.

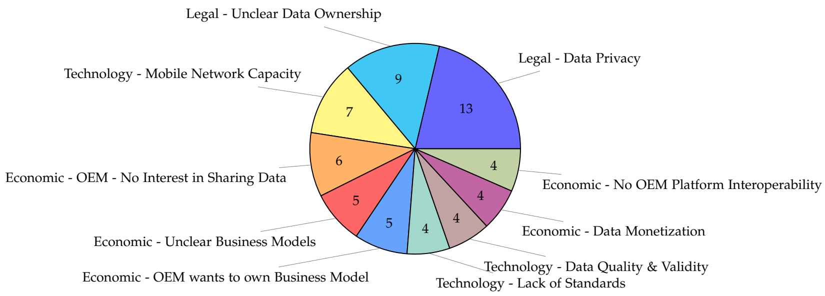

Figure 5.2.: Top ten barriers discovered during the interviews

5.2.1. Legal

The following subsection elucidates on various barriers which have been associated with the legal domain. The discovered barriers are predominantly concerned with data privacy and data ownership.

Within the legal category and also across all categories **data privacy** was the most significant barrier which was mentioned by 18 distinct interviewees. *Vehicle-generated data* is considered at least partially user related, therefore complying with data privacy laws while providing the data is considered a general barrier [IV01], [IV03], [IV05], [IV07.1], [IV07.2], [IV17], [IV18]. The implementation of data collection software which is compliant with current data protection regulation is also considered a barrier by one *OEM*. A potential solution is currently in a test phase. Data privacy has a high significance, especially in Germany and for German *OEMs*, in conjunction with the high penalties of the *General Data Protection Regulation* (GDPR), this is considered a barrier for the faster adoption [IV12]. The European *GDPR*, which became enforceable from 25th of May 2018 onwards, has also been described as a major barrier [IV09], [IV10]. In general, the collection of data has to be conducted for a specific purpose, which makes it harder for *OEMs* to collect data generically. The German legislators overshoot the mark while failing to define the three roles, data owner, data producer, and data user, which are considered essential for data-driven business models. Moreover, the interviewee considers current data collection terms of certain *OEMs* as not *GDPR* compliant as they are not bound to a specific purpose [IV03]. Due to the regulation the user has to provide his consent more specifically than before, convincing the user is hence also considered a barrier related to the **data privacy** [IV11]. Due to the broad media coverage of the *GDPR*, the aspect of **data privacy** has also gained greater awareness in public, which leads to a higher suspicion towards companies collecting data, especially if the data is shared with for the user unknown third parties [IV13]. From the perspective of another *OEM* the **data privacy** is important, but not considered a major barrier as current applications collecting data have high consent rates of between 60 and 70 percent. Albeit the topic is considered complex as misconduct could have severe consequences towards user acceptance [IV14]. One interviewee considers **data privacy** a general barrier, but also considers the *GDPR* a good framework for implementing data privacy law compliant data collection systems [IV16]. It is not clear what kind of *vehicle-generated data* is considered *personally identifiable information* this is therefore considered an open issue for further discussion [IV08]. From the perspective of the user (i.e., data producers) the aspect that the collection of large amounts of *vehicle-generated data* could be used to create extensive profiles could lead to an inhibited adoption and is therefore considered another barrier [IV02], [IV04].

The **unclear data ownership** was described as another barrier by several interview participants. Multiple interviewees described an ongoing discussion about the data ownership, the core question is if the data is owned by the *OEM* who produced the vehicle, the owner of the vehicle or the driver of the vehicle [IV02], [IV03], [IV04], [IV06], [IV09], [IV11], [IV13]. It is even considered if the data has high relevancy for society and should consequently be owned by it [IV06]. Assuming the owner of the data is not the *OEM*, the *GDPR* requires each processor of the data to allow the removal of the data if requested by the data producer (e.g., the driver). Furthermore, the *OEM* would be required to inform the data producer to whom the data was transmitted [IV02]. One interviewed *OEM* has the position that the customer owns the data, albeit the data ownership is still considered a barrier inhibiting the large scale availability of *vehicle-generated data* [IV12].

The collection of *vehicle-generated data* also poses **liability** concerns. By an interviewed *OEM* it is speculated that it is required by law to inform the user about potential threats to the vehicle safety. An example would be sensor data about the brake pad. If it is indicated that the brake pad thickness leaves a certain threshold, the driver could be endangered. If collecting data, the *OEM* would be required to monitor this issue and inform the user, otherwise the *OEM* would be potentially liable in the case of an accident. This discussion is very diverse, some viewpoints are that if it is possible to gather insight about the vehicle safety from *vehicle-generated data*, the *OEM* would be obliged to collect the data. This is considered a major barrier by the *OEM* [IV12]. The liability issue is also mentioned by a third-party service provider, who indicates that also a third-party service provider might be liable for potential damages to the vehicle [IV13].

Strict **regulation and certification** requirements within the automotive industry are also considered a barrier especially for start-ups entering the market. This barrier is particularly relevant if the company is offering third-party hardware for the vehicle [IV13].

5.2.2. Economic

In this subsection economic barriers are described. This subsection covers all barriers related to business models and other economic factors that inhibit the advancement of an API economy within the automotive industry.

The economic barrier that was most frequently mentioned by the interview participants was that *OEMs* have **no interest in sharing** *vehicle-generated data*. The reasoning behind that is very diverse. It is pointed out that the *OEMs* want to keep the data to be the sole **owner of the business model**. Furthermore, especially German **OEM** have serve **competition concerns** when it comes to data. Data-Driven companies like Google are considered a serious threat when it comes to data-driven business models. General data privacy concerns of the *OEMs* are also mentioned as a reason for the lack of sharing interest by the interviewee [IV01]. Another viewpoint is that the *OEMs* have no interest in sharing *vehicle-generated data* as the data is considered very valuable by the *OEMs* [IV06].

A very similar stance is shared by [IV10], who speculates that there is almost a *"war for data"* because every stakeholder of the automotive industry realizes the high strategic value of the data. Furthermore, it was mentioned that the lack of sharing-interest is also caused by a customer lock-in strategy implemented by the *OEMs*. The issue is compared to the *OBD-II* interface of which only a small subset is standardized by law, but the majority of the data is only accessible through proprietary protocols [IV13]. This lock-in strategy is also mentioned by another interview participant and compared to the patented *Nespresso System*[1], where the car is the coffee machine and digital services are the capsules [IV09]. Another interviewee states that *OEMs* are not interested in an open service ecosystem, as they all try to implement their own

[1] https://en.wikipedia.org/wiki/Nespresso

proprietary ecosystems and build business models on top of the ecosystem. This is even the case for tier-one automotive suppliers, which do not have access to data generated by components the have built [IV16]. The lack of data sharing is also justified by the cooperate policy of the *OEMs* [IV18].

Another barrier that has been mentioned multiple times is that potential **business models are unclear** for the *OEMs*. Within the automotive industry, the business is considered very *"classically"* [IV09]. It is speculated that the *OEMs* are still in an identification stage regarding potential business models. This is also related to the **long automotive product cycle**. The development phase of a new car takes about seven years, which is considered a reason for the delayed adoption of new business models within the automotive industry [IV13]. The vast amount of *vehicle-generated data* available to the *OEMs* is mentioned as the underlying cause for the barrier of **unclear business models**. *OEMs* are currently in an exploration phase regarding the data and potential business models [IV12]. According to another interviewee, the issue of **unclear business models** is also caused by other market actors not thinking far enough how they could participate in such an ecosystem. Concerning diagnostic use cases the same interviewee speculates that there is currently no viable business model for such use cases [IV16]. In the context of usage-based insurances, it is mentioned that this kind of business cases would violate the insurance principle. Therefore the business models are not yet clear and viable [IV18].

Moreover, the *OEMs* try to implement many business models and use cases themselves, even if small start-ups have more expertise on the subject as they are highly specialized. This kind of **OEMs policy** is also described as a significant barrier by [IV05]. Along with the **OEM policy** and the desire to own the business models, strong competition concerns are described. This restrictive policy is considered as a severe threat to the competitive advantage of the German and European automotive industry as *OEMs* from China, or the U.S. are more open to data-driven business models [IV03].

No clear strategy regarding the **monetization of the data** has also been mentioned as another factor inhibiting the widespread availability of *vehicle-generated data*. The lack of data-monetization strategies is explained with the lack of empirical studies and case studies showing that the revenue from selling data could significantly improve the overall revenue of *OEMs* [IV03]. While it is evident that *OEMs* want to generate revenue from selling the data, viable pricing models that enable an API economy are still to explore. Moreover, the interviewee assumes that a *vehicle-generated data* economy could only work if it generates a reliable revenue stream for all stakeholders [IV10]. The topic of **data monetization** is also considered a barrier by an *OEM* himself. The value of the unstructured and unprocessed data is unclear, without knowing if the data is of any value for the *OEM*, it is just generating costs for storing it [IV12].

Furthermore, the data collection would also generate additional costs for the *OEM*. The costs are composed of general infrastructure costs for collecting and storing the data, as well as for per-vehicle costs for extra hardware and data transfer through the cellular network. When developing vehicles the *OEM* expects the vehicle to have a particular contribution margin, but the added value of collected data is hardly quantifiable. When developing new vehicles, this calculation method results in the withdraw of components that would be responsible for the data collection. Moreover, the long **long automotive product lifecycles** require the *OEMs* to plan three to five years in advance when it comes to collection *vehicle-generated data* [IV14].

Competition concerns either between the *OEMs* or in general, have also been identified as an economic barrier. While there are selective initiatives where *OEMs* cooperate, each *OEM* is still staking out his territory, which makes it harder to establish a collaborative API economy [IV04]. This **lack of OEM collaboration** is also mentioned by another interviewee, who is suspecting that the allocation of market segments to each *OEM* is very static and therefore no larger collaboration interest exists between the *OEMS* [IV10]. Competitive concerns regarding **intellectual property** have also been mentioned as a barrier. The large scale provisioning of *vehicle-generated data* by the *OEM* could provide too much insight into proprietary technologies and is therefore inhibited [IV02], [IV16].

Corporate Image concerns of the *OEM* have also been stated as barriers. It is hypothesized that *OEMs* are afraid of a negative corporate image due to selling the data of their customers, which could lead to a decrease in vehicle sales outweighing the additional revenue stream from selling data. Particular use cases also do not fit to corporate image the *OEM* is cultivating for himself. The interviewee referred to a milage-driven leasing project with an *OEM*, which was canceled because the concept did not fit to the corporate image the *OEM* is facilitating towards its customers [IV17]. As *vehicle-generated data* also includes data which related to the driver, another barrier mentioned by two interviewees is the **fear of data leaks**. It is suspected that the *OEMs* are concerned that providing *vehicle-generated data* through APIs could lead to data leaks and subsequently damage their corporate image and reduce the trust of their customers [IV03], [IV10].

Most German *OEMs* have a long company history. Hence their thinking and processes are more mechanical than digital. In this context, one interviewee describes the process of **digital transformation** as another severe barrier for *OEMs*. This again inhibits the emergence of new data-driven business models within the automotive industry and therefore also a potential API economy. The same interviewee also states that the automotive industry has a **skilled worker shortage** which inhibits new business models [IV09].

Regarding one of the few existing *vehicle-generated data* APIs, one interviewee mentioned the pricing model as a barrier. According to the interviewee, the **pricing** model is between 5 Cent and 10 Cent per data point. Depending on the sample rate of the sensor this pricing model could cause substantial costs for any data user or service provider [IV16].

5.2.3. Social Acceptance

The following subsection describes social acceptance related barriers, which all refer to enduser related issues.

In the context of usage-based insurances, it is speculated that there is a **lack of customer interest** for this kind of insurance products. After conducting field trials, the potential customers did not accept the product as desired as a flat-rate approach seems more attractive [IV12].

For B2B fleet management vendors, the customer and the driver are often different persons. This leads to potential barriers when it comes to the acquisition of *vehicle-generated data* from a vehicle of the fleet. The interviewee assumes that the driver is the owner of the data and is therefore required to consent to the data usage [IV05]. Another interview participant also mentions the barrier of **driver consent** in a broader context. While people are generally willing to share the data, they want to explicitly know what happens to their data and for what the data is used. Furthermore, the user wants the possibility to revoke his data sharing consent. For service providers, wanting to use *vehicle-generated data*, it poses, therefore, a severe barrier to acquire the consent of the user and convince her/him that she/he benefits from sharing the data with the service provider [IV15].

This barrier is also reflected in the mentioned barrier of **enduser acceptance**. Especially in Germany people are more conservative regarding data sharing and new data-driven business models [IV06]. The barrier of **enduser acceptance** is described as very serve barrier as many different factors, like media and politics, influence the decision making of people [IV11].

5.2.4. Technological

Technological barriers are elucidated in the following subsection. The barriers described in the following all relate to either direct technological issues (e.g., lack of infrastructure) or barriers which are implicitly related (e.g., data quality).

In the technology category the available **mobile network capacity** (i.e., the cellular network) was the most frequently mentioned barrier. Multiple different nuances of this barrier were described. One interviewee considers the cellular network only as a minor problem, which is only a problem for areas with many vehicles at the same time [IV03]. Two other interview participants consider the current cellular network capacity a general problem, both state that the vast amount of data, assuming *vehicle-generated data* would be collected extensively, would just be too much, even for next-generation cellular networks like *5G*[2] [IV04], [IV07]. Another participant elucidates that the network capacity barrier is depending on the use case. A simple parking space detection system that uses an ultrasonic sensor would also be possible using a *2G* network. Transmitting the video stream for thousands of cars in real-time (e.g., when they are leaving a parking garage after an event) would overload not only the network, but also the backhaul of the base station [IV11]. The mobile broadband prices in Germany are also considered a barrier by one interviewee. Transmitting data in realtime would currently go entirely beyond the scope of any Machine-To-Machine LTE contract [IV13]. From the perspective of one *OEM*, the mobile network capacity is currently no issue but will be an issue for autonomous driving use cases, as the amount of data required for such applications is an order of magnitude higher [IV14]. Another interview participant also thinks that this barrier is currently no issue, but will be in the future when more and more services are provisioned to connected vehicles [IV08].

Two interviewees consider that the potential **amount of data** is a barrier. Currently, insurances do not have the required infrastructure to process the data. One side of the problem is the transmission of large amounts of data between different data centers. The other side is a lack of processing infrastructure as insurance companies currently do not have use cases that require processing of such large amounts of data [IV06]. From an *OEM* perspective, the data amount is problematic because it generates costs for transmitting and storing it. Furthermore, it would require backend infrastructure for the collection and analysis of the data [IV12].

The two interviewees, who are employees for an *OEM*, both consider the *OEMs* IT infrastructure as a barrier. Currently, the *OEMs* have no infrastructure capable of

[2] `https://ec.europa.eu/research/press/2013/pdf/ppp/5g_factsheet.pdf`

handling the blind collection of all vehicles with support for over-the-air data collection [IV12]. The heterogeneity of the internal applications which are currently used for storing such data is also considered a barrier [IV14].

The **data quality and validity** is also an open issue mentioned by multiple interview participants. One interviewee considers the data quality a general issue [IV07.1]. The quality of the data which is available today through the *OBD-II* interface is very heterogeneous and therefore requires extensive testing and data correction [IV13]. From an *OEMs* perspective, the data quality is also described as a barrier which correlates with the heterogeneity of the system landscape and the lack of data governance. The issue goes that far that different departments calculate different values for key performance indicators even though they are based on the same kind of data [IV14]. From an insurance perspective, the validity and non-repudiation of the data is an essential aspect. Therefore, it is considered a barrier when the data is stored at the *OEM* [IV06].

Another frequently mentioned barrier is the **lack of standards**, which define the data exchange formats and are used by multiple *OEMs* [IV07.2]. The lack of extensive standards is also a barrier for current use cases using the *OBD-II* interface for data acquisition, as the data is even different for different models of the same *OEM* [IV04]. This lack of standards creates uncertainty for potential users and therefore inhibits the adoption [IV13]. Another interviewee points out that the term *"web service"* is defined to arbitrary as there are too many different protocols [IV10].

The **absence of platform interoperability** between each *OEM* is also considered a serve barrier, even for current applications. Even though the *OBD-II* interface is standardized, it fails to deliver consistent data points across multiple *OEMs*, the implementation of the standard is described as *"higgledy-piggledy"* [IV16]. Another interviewee relates to the same issue and elaborates further that operations for the *OBD-II* interface are not supported the way they are described in the *OEM's* documentation [IV13]. Interoperability is considered a general problem within the automotive industry, due to historical reasons each *OEM* uses its own protocols for the communication between the components of a vehicle. The situation is described as *"nothin fits together"* by an interviewee [IV11]. Furthermore, the approach of each *OEM* to establish his platform as a standard is considered a barrier related to the interoperability [IV03]. The **general lack of access** is also described as a barrier inhibiting the adoption of *vehicle-generated data* [IV06], [IV07.1], [IV07.2].

Some use cases, especially if they are related to maintenance or diagnostics, require **realtime data** from the vehicle. *Vehicle-generated data* available today is predominantly historic data [IV07.1]. To the knowledge of one interviewee, current concepts where *vehicle-generated data* is provisioned through an API, are not capable of delivering real-time data, which are often required to diagnose problems with vehicle components [IV15].

Another barrier is the higher **security** requirements for road vehicles. The vehicle has to provide access to *vehicle-generated data* while preventing unauthorized access [IV07.1]. Still maintaining the same security standards as before is described as a difficult challenge for *OEMs* [IV08]. Furthermore, the software systems of a vehicle have to be more **resilient** than the average smartphone or desktop computer as a failure might have catastrophic effects. This is also an inhibiting factor for the slow advancement of innovative platforms providing *vehicle-generated data* [IV11].

5.2.5. Summary & Assessment

The results of the interviews indicate that most barriers are in the fields **legal** and **economic**. The prevailing opinion though all interviews was that the technical barriers would be solvable rather quickly. If not already stated, the interview participants were asked what, in their opinion, is the most challenging barrier. Six interviewees mentioned that the lack of willingness to cooperate with third-party service providers by the *OEM* is the principal barrier [IV02], [IV05], [IV06], [IV13], [IV16], [IV18]. The interviewee of an *OEM* also indicated that the current mindset of the *OEMs* has to change. The understanding that data is valuable is currently not present at all hierarchy levels of the *OEMs* [IV14]. A similar perspective was that currently, every market participant is waiting for the other to make a first move, particularly the *OEMs* [IV09]. Five interviewees identified **Data Privacy** as a major inhibitor for *vehicle-generated data* not being available at a larger scale [IV03], [IV04], [IV07.1], [IV12], [IV17]. The lack of standardization was also indicated as a very challenging barrier [IV07.1]. Without being more specific, the fields of **legal** and **economics** were described as most challenging by [IV10]. Furthermore, the topic of social acceptance was also described as the most challenging barrier as the end-user (i.e., the driver) has to understand the benefit he/she gains by sharing the data [IV02], [IV11].

6. Potential Measures

The following chapter describes potential measures to overcome the barriers described in chapter 5. As explained earlier this chapter is answering research question four. The results presented in the following have also been gathered from the conducted interviews. This chapter is subdivided into two parts. The first part describes potential measures which have explicitly been associated with a certain barrier by an interviewee. The second part describes implicit barrier-measure associations that have been derived based on the previous findings of this work. The eleven distinct potential measures which have been suggested during the interviews are illustrated in Table 6.1.

Barrier	Potential Measures										
	Critical Data Mass	Targeted Data Collection	Demonstrate Sustainability	Neutral Data Provider	Government Regulation	Incentives for Endusers	OEM Organization Policy	OEM Collaboration	Enduser Transparency	Edge Computing	Market Pressure
Legal											
Unclear Data Ownership	-	-	-	-	●	-	○	-	-	-	○
Regulation & Certification	-	-	-	-	-	-	-	-	-	-	-
Data Privacy	-	●	-	-	○	-	-	-	●	-	-
OEM Data Privacy	-	-	-	-	-	-	-	-	-	-	-
User Data Privacy	-	-	-	-	-	-	-	-	-	-	-
Liability	-	-	-	-	○	-	-	-	-	-	-

Table 6.1.: Barriers and Potential Measures.
● = Measure has been explicitly mentioned by an interviewee.
○ = Measure has been derived by the author of this work.

Barrier	Critical Data Mass	Targeted Data Collection	Demonstrate Sustainability	Neutral Data Provider	Government Regulation	Incentives for Endusers	OEM Organization Policy	OEM Collaboration	Enduser Transparency	Edge Computing	Market Pressure
Economic											
OEM Digital Transformation	-	-	-	-	-	-	○	-	-	-	-
OEM Competition Concerns	-	-	-	-	-	-	○	-	-	-	-
Intellectual Property	-	-	-	-	-	-	-	-	-	-	-
Lack of OEM Collaboration	-	-	-	-	-	-	-	●	-	-	●
OEM - No Interest in Sharing Data	-	-	-	-	●	-	●	-	-	-	●
OEM Corporate Policy	-	-	-	-	-	-	-	●	-	-	-
Automotive Product Lifecycle Duration	-	-	-	-	-	-	-	-	-	-	-
OEM image	-	-	-	-	-	-	-	-	-	-	-
fear of leaks	-	-	-	-	-	-	-	-	-	-	-
skilled worker shortage	-	-	-	-	-	-	-	-	-	-	-
fair use	-	-	-	-	●	-	○	○	-	-	-
Business Model											
No Killer Application	-	-	-	-	-	-	-	-	-	-	-
Data Monetization	-	○	●	-	-	-	-	○	-	-	-
Unclear Business Models	●	-	●	-	-	-	-	○	-	-	-
Pricing	-	-	-	○	-	-	○	-	-	-	-
Business Model Ownership	-	-	-	-	-	-	○	-	-	-	-

Table 6.1.: Barriers and Potential Measures.
● = Measure has been explicitly mentioned by an interviewee.
○ = Measure has been derived by the author of this work.

Barrier	Critical Data Mass	Targeted Data Collection	Demonstrate Sustainability	Neutral Data Provider	Government Regulation	Incentives for Endusers	OEM Organization Policy	OEM Collaboration	Enduser Transparency	Edge Computing	Market Pressure
Social Acceptance											
Lack of Customer Interest	-	-	-	-	-	●	-	-	●	-	-
Enduser Acceptance	-	-	-	-	-	●	-	-	●	-	-
Driver Consent	-	-	-	-	-	●	-	-	●	-	-
Technology											
Lack of Access	-	-	-	-	●	-	○	○	-	-	●
Lack of Standards	-	-	-	-	-	-	-	●	-	-	-
Lack of Infrastructure	-	-	-	○	-	-	-	-	-	-	-
Mobile Network Capacity	-	○	-	-	-	-	-	-	-	●	-
No OEM Platform Interoperability	-	-	-	-	-	-	-	○	-	-	-
Security	-	-	-	-	●	-	-	-	-	-	-
Technology Resilience	-	-	-	-	○	-	-	-	-	-	-
Data Quality & Validity	-	-	-	●	-	-	-	-	-	-	-
Lack of Realtime Data	-	-	-	-	-	-	-	-	-	-	-
Data Amount	-	●	-	●	-	-	-	-	-	●	-

Table 6.1.: Barriers and Potential Measures.
● = Measure has been explicitly mentioned by an interviewee.
○ = Measure has been derived by the author of this work.

6.1. Explicit Potential Measures

The following subsection describes potential measures to overcome specific barriers mentioned in chapter 5. The interviewees have explicitly mentioned the measures in this subsection.

Government Regulation has been mentioned by seven different interviewees as a potential measure. It is proposed that the access to *vehicle-generated data* could be enforced by the government as the data has particular importance to society as a whole [IV06]. Another interviewee considers that **Government Regulation** is the only potential measure that would ensure fair access to *vehicle-generated data* for all economic actors. Furthermore, the interviewee does not believe that the market is self-regulating regarding *vehicle-generated data*. Therefore a statutory regulation is necessary to ensure fair access, especially for small companies and start-ups. The fair access to the user interface of the vehicle (i.e., head unit) must also be warranted under the same conditions, as the *OEM*, for all actors. This means that third-party service providers must be allowed to also run their applications within the vehicle (i.e., terminal mode) [IV15]. The barrier of unclear data ownership is also suggested to be approached through **Government Regulation**. There is a regulatory need for the question who owns the data of a vehicle after it is purchased, either implicit through judicial decisions or explicit through legislation [IV10]. For data ownership, an isolated national legislative regulation is considered not sufficient. Such regulation should at least be implemented homogeneously throughout the European Union, which is considered a complex and tedious process [IV17]. This kind of regulatory approach is also suggested by another interviewee, who compares the situation of *vehicle-generated data* in the automotive industry to the finance sector. The *Payments Service Directive 2 (Directive (EU) 2015/2366)* promotes innovative services within the finance sector by enforcing rules for open banking [71]. This is considered a possible example of a similar legislative framework which could be a blueprint for a similar regulation in the domain of *vehicle-generated data* and APIs [IV18]. A stricter legislative regulation could also be a measure for data security-related concerns [IV07.1].

Another potential measure which was also mentioned by seven distinct interview participants is summarized under the term **Enduser Transparency**. The measure is suggested for multiple barriers with a particular prevalence towards data privacy. The Enduser of the vehicle (i.e., the driver) must be able to understand what happens to the data of his/her vehicle. Furthermore, he/she must have the possibility to influence what kind of data is shared with whom [IV04], [IV05]. Moreover, the security of the systems must be shown to them , and it must be shown what happens to the data and how it could influence the private life of the end user [IV07.1]. Another interviewee points out the *GDPR* is a good initial framework for establishing trust towards the end user, as the regulation allows the user to request specific details about data he/she provided to a third party [IV13]. Besides, it is suggested that all end users should be able to decide if the vehicle is transmitting data at all, but excluding safety-related data required for providing for example automated emergency services. It is suspected that this kind of mechanism would increase the trust of the end users towards a *vehicle-generated data* economy and therefore enable it [IV15]. The general end-user acceptance of data-driven services around the vehicle could also be improved by demystifying the vehicle. The interviewee suspects that the general public perceives the vehicle as a mystery. Showing that the vehicle is a logically structured machine might help the end user accepting new functionality within a vehicle like data sharing [IV11]. A similar stance is taken by another interview participant, who suggests showing the added value for the end user by third-party service providers could improve the general acceptance [IV13].

The willingness to share data with the *OEM* and third-parties by the user could also be improved by providing **Incentives for Endusers**. For usage-based insurances contracts, this could be a very simple pecuniary incentive. For example, an end user who provides *vehicle-generated data* to his insurer pays less than compared to a user with a classic flat-rate-based insurance model [IV06]. Acquiring the end user consent for sharing data is described as a barrier which applies to all kind of data-sharing models. Therefore, the proposed measure is that service-providers have to engage the customer through different sales channels and create innovative products that ease the life of the potential customer [IV15]. An interviewee who considered the user acceptance only a minor barrier, suggested that services improving the convenience for the end user are a crucial element for the broader adoption of *vehicle-generated data* based services [IV09].

Improving the **OEM Collaboration** is suggested as another measure to overcome multiple barriers. One interviewee considers that the attitude of German *OEM* is very competitive, instead of enlarging the market, the *OEMs* are only trying to increase their market share. The interviewees suspect that on strong collaboration between the *OEMs* could enable an API economy providing *vehicle-generated data*, but for that, the mindset of the *OEMs* has to change [IV02]. Another interviewee confirms this hypothesis and states that progress can only be achieved through collaboration. The reciprocal exchange of data will sooner or later enable an API economy within the automotive industry [IV04]. Another interview participant suggests that a change of **OEM Organization Policy** is required to enable an API economy. This change should shift the policy in the direction of open innovation and the willingness to cooperate with start-ups and other young companies. Furthermore, it is suggested that *OEMs* have to slim down their hierarchies and structures [IV05].

A measure which is not directly applicable is **Market Pressure**, which was mentioned by four interview participants. The current economic situation if the *OEMs* is described as very healthy, whereas the selling of data is described as a business model with many risks. Therefore, large enterprises will only change their objective through sufficient market pressure. This pressure could come from foreign competitors trying to disrupt the market or local competitors making a first move [IV03]. This is also described by another interviewee, suspecting that the realization that the exchange of data between different stakeholders could lead to reciprocal benefits, will be recognized by the *OEMs* as time passes. Further, it is suspected that a first-mover is required to accelerate the emergence of an API economy within the automotive industry [IV16]. The market pressure could also come from the consumer who wants to understand the vehicle he/she owns and want's to share the data with third-parties [IV10], [IV13].

An interview participant of an *OEM* suggests that there has to be a **Critical Data Mass** available to be able to derive business models. Currently, there is no killer application which could disrupt the market and lead to the emergence of a *vehicle-generated data* based API economy. The interviewee speculates that smart cities could be such an application [IV12]. Before an API economy could emerge the viewpoint of the *OEMs* regarding data has to change. Hence, another interviewee of an *OEM* proposes that it is necessary to **Demonstrate the Sustainability** of the data collection by showing how the data could be monetized and what kind of business models could be created based on *vehicle-generated data* [IV14].

Regarding technological barriers, the prevalent potential measure was **Edge Computing**. The concept was mentioned as a potential measure for the barrier of the data amount and the mobile network capacity [IV03]. The rising amount of sensors and the subsequently more substantial data amount will require a preprocessing in the vehicle. On the one hand to reduce the data amount and perform an initial data analysis directly in the vehicle (e.g., Black ice detection) on the other hand to improve the data quality by directly filtering invalid values and therefore reducing the load on the mobile network [IV04], [IV10]. Another approach, which is already being used by one *OEM*, is the **Targeted Data Collection**. By defining a subset of vehicles and sensors of which data shall be collected from, it is possible to reduce the data amount and only collect the data which is necessary for a predefined use case. The downside of this approach is that it must be established what kind of data the use case or business model requires [IV12].

Another potential measure is the establishment of a **Neutral Data Provider**. One interviewee from the insurance industry suggests that a neutral data provider could help increase the validity and trust of *vehicle-generated data*. This could also reduce the amount of data moved between different data centers as insurances do not have to acquire the data from multiple *OEMs* [IV06]. The concept of a **Neutral Data Provider** is also proposed to ensure fair market access for all stakeholders [IV15]. Another interviewee suspects that a central platform used by all *OEMs* is required for an API economy to emerge [IV18].

6.2. Derived Potential Measures

The following section describes which other barriers could be overcome by the measures mentioned before. The associations have been derived from the context of the interviews.

A **Targeted Data Collection** approach could also help *OEMs* finding a data monetization approach while keeping the infrastructure costs low. The targeted collection allows to test business models on a small scale and finding an approach for a sustainable pricing model. As for technological barriers, the targeted data collection could also be a measure for the mobile network capacity. Retrieving only data from specific sensors or from cars in defined geographical areas etc. This could dramatically reduce the required bandwidth.

Establishing a **Neutral Data Provider** could also have a positive influence on the pricing of the data. As the infrastructure for storing the data would be shared among all *OEMs*, the costs should automatically be lower than having multiple platforms. This would also help to overcome the barrier for the lack of infrastructure. The fact that data of multiple *OEMs* would be available through one platform could also influence the pricing as one *OEM* might lower the prices to attract new customers. A similar effect was observed for cloud storages prices between *Amazon* and *Google* [72].

In addition to the already mentioned barriers that **Government Regulation** would potentially solve, there are also a few more, that can be derived from the context of the interviews. Through the *GDPR* there is already a legislative framework available that could support the involved stakeholders. However, it is considered a moot point if the *GDPR* is helpful in terms of supporting the emergence of an API economy within the automotive industry or not. Liability concerns of the *OEMs* could easily be approached by defining a legal framework which moves the liability away from the *OEMs*. Technology resilience could also be enforced through **Government Regulation**. Defining certain technological thresholds that *OEMs* and third-party service providers must fulfill could prevent the compromising of the vehicle security or resilience. This could also be connected to mandatory external audits, conducted by an independent entity.

The majority of the responses of the interviewees implicate that changing the **OEMs Organization Policy** could drastically accelerate the emergence of an API economy. The conducted interviews showed that the question of data ownership is a very controversial subject. The *OEMs* could establish a principle that defines data ownership in compliance with the current legislation. Furthermore, the lack of standards could also be overcome by either implementing and documenting existing standards or by the collaborative development of new standards. It seems that this issue is already being approached by the *ISO* standards described in subsection 2.4.1. Changing the organizational focus towards a digital enterprise might help to tackle issues arising from the inhibited digital transformation of the *OEMs*. Besides, changing the mindset to think more openly when it comes to sharing data could also support the elimination of barriers related to competition concerns as well as concerns regarding the fair use and access for other economic actors. It is suspected that acting more openly towards other organizations could also resolve the barrier of the business model ownership.

The lack of collaboration between the *OEMs* has been mentioned multiple times during the conducted interviews. A fortified **OEM Collaboration** could help to resolve several barriers like the fair access to *vehicle-generated data*. As for economic barriers, a closer collaboration might enable the exchange of experiences between the *OEMs* and lead therefore to more viable pricing strategies and business models. In the technology area, a closer collaboration might also improve the interoperability between the different platforms of the *OEMs*.

A more vigorous **Market Pressure** towards the *OEMs* might also help to overcome the barrier of the unclear data ownership. Either other economic actor or the consumer could pursue the goal of laying out a definition of the data ownership. Moreover, the market demand for new digital services around the vehicle might lead to an increasing willingness to rapidly clarify the ownership of the data.

The previous two sections show that the prevalent two measures are **Government Regulation** and **Enduser Transparency**. This again indicates that the prevalent barriers are not related to the technology field. Furthermore, the majority of the barriers and corresponding measures are either located in the economic or legal domain.

7. Discussion

This chapter will discuss the results of this thesis. For that the results for each research question will be summarized in the following.

RQ1: How is vehicle generated data currently collected and how is value created with it?

The elucidation of current data collection methods has been presented in section 2.2. Currently the prevalent methods to collect *vehicle-generated data* are based on the standardized *OBD-II* interface. A large market for diagnosis hardware exists which is either targeted at professional or weekend mechanics. With the ubiquitous presence of smartphones a new category of devices has emerged. The so-called *OBD* dongles are intended to connect the vehicle with the smartphone or directly connect the vehicle to the internet through an integrated *SIM* card. This kind of devices are targeted at consumers wanting to gain more insights about their vehicle, but also at business customers for use cases like fleet management.

The value creation process for *vehicle-generated data*, described in section 2.3, has been derived by looking at similar domains, primarily the big data domain. Based on the finding a value chain with the involved stakeholders has been created. Furthermore, existing literature finding and the bedrock hypothesis of this work has been incorporated into a value network which illustrates how value is created from *vehicle-generated data*. The results of the interviews described in the subsequent section confirm the assumptions undertaken to derive the value chain.

The subsequent section 2.4 elaborates on related work by providing an overview of similar research projects and standardization initiatives. The findings of the sections indicate that there is currently high demand and stir regarding *vehicle-generated data*. The *ISO* standards are intended to be published throughout the current year. Furthermore, the first *NEVADA* based *APIs* are publicly available. It is expected that more APIs will be available by the end of the second quarter of 2019 [IV08].

RQ2: What are state of the art use cases and business models that exist or are proposed for vehicle generated data?

The second research question was answered in chapter 4. The first part of the chapter illustrates the results of the extensive literature review, which showed that an abundance of potential use cases had been proposed. Albeit, only a small subset has been implemented, most of the implementations are prototypical to demonstrate a particular use case or algorithm.

Afterwards, in section 4.2 use cases which have been mentioned by the interviewees are discussed. The section provides insight into existing use cases and proposed use cases. While several new use cases have been proposed, the implementation status is mostly congruent with the use cases described in the literature. The majority of the existing use cases relies on data extracted from the *OBD-II* interface. No third-party use cases with API based access to *vehicle-generated data* were described.

RQ3: What are barriers for use-cases not being implemented?

In chapter 5 the barriers gathered from the conducted interviews were described. A taxonomy has been developed to classify the barriers mentioned by the interviewees. A total of 34 distinct barriers has been discovered throughout the interviews. The majority of the barriers were either classified as related to legal topics or related to economic topics like potential business models. An interesting finding was that many barriers were described as very hard to solve as they are based on fundamental issues within the automotive industry.

RQ4: What could be measures to overcome identified barriers and advance the API Economy within the automotive industry?

The last research question was answered in chapter 6. The participants of the interviews mentioned the measures described in the chapter. The interviewees either connected the barriers explicitly to the proposed measures or the connection has been derived based on the overall context of the interviews. The chapter concludes that the prevalently mentioned measures are also directed at legal or economic matters. The two most frequently mentioned measures where the need for Government Regulation and more transparency towards the end user.

The findings of this work indicate that there is currently much stir around *vehicle-generated data* and potential ecosystems. Many use cases have been proposed, and only a few have been implemented, several potential barriers for the inhibited adoption have been identified. Based on the results of the interview it is hardly possible to provide a timeframe for the widespread availability of *vehicle-generated data*. The findings of this work show that first APIs are already available and that more APIs will be available within the next six month. Currently, only little information about existing and upcoming APIs is available. Also, the pricing is not yet very transparent as the conducted interviews have shown. The implications for potential users of the APIs are that there might be still much change in the current ecosystem. Early adopters are currently moving forward to provide programmatic access to *vehicle-generated data*, but the maturity level of most available APIs is described as experimental.

8. Conclusion

The bedrock hypothesis of this work is that if OEMs would provide programmatic access to *vehicle-generated data* an API economy would emerge within the automotive industry. Based on the results of this work, this hypothesis can be accepted. The prevalent state of the art collection method still is to use the *OBD-II* interface inside the vehicle. This approach solely relies on third-party hardware and requires its installation inside the vehicle by the owner or a mechanic acting on behalf of the vehicle owner. Further, a reasonable value chain and value network for *vehicle-generated data* was presented. The presented artifacts illustrate potential monetization models for *vehicle-generated data*. The results of the interviews have confirmed the assumptions that were made. The analysis of related work indicated that there is being worked on several solution approaches. The two most promising developments are the *ISO* standards 20077, 20078 and 20080 as well as the *NEVADA* concept of the *VDA*. The *ISO* standards are either already published or currently under publication. First *APIs* which implement the ISO standards are currently emerging. The same applies to the *NEVADA* concept which is built upon the *ISO* standards. A first *NEVADA* compliant implementation is publicly available and the emergence of more *NEVADA APIs* is expected in the next six months.

In this work, multiple use cases gathered from different sources have been described. The results of the extensive literature review indicate that most of the proposed use cases are not implemented or if they are implemented then mostly in a prototypical manner. A very similar observation can be made for existing use cases described during the conducted interviews. Almost all use cases rely on third-party devices to acquire the relevant *vehicle-generated data*. The overwhelming part of the use cases proposed by the interviewees require a large amount of data that has been gathered over a longer time frame. For use cases like predictive maintenance or product optimization, a more substantial amount of data is required to produce significant results. For this kind of use cases, the collection of data through optional third-party devices is not feasible. This once again confirms the base hypothesis of this work.

The results of the conducted interviews revealed 34 distinct barriers from four different categories (economic, legal, social acceptance, technological) inhibiting the advancement of an API economy within the automotive industry. The analysis of the data shows that the significant barriers are located in the legal and economic domain. The general topic of data privacy or the moot point of the data ownership where among the top ten barriers mentioned during the interviews. Another result of the conducted interviews were eleven potential measures to overcome the barriers elucidated in this work. The two most frequently mentioned barriers where the need for government regulation and to establish more transparency towards the end user.

In summary, it can be stated that the advancement of the API economy is inhibited by fundamental structural barriers which exist in the automotive industry. The proposed measures do not indicate that a near-term change in the current situation is feasible.

The remainder of this chapter discusses the limitations of this work as well as potential starting points for future work.

8.1. Limitations

Within the narrow time frame of this thesis it was not possible to conduct the *grounded theory methodology* to its full extent. In particular, the theoretical sampling process could have been extended in a larger time frame. While the theoretical saturation reached is well enough to validate the hypothesis of this thesis, more interview participants from *OEMs* would have been beneficiary. As only two participants were employed at *OEMs*, the results could be biased towards non-*OEM* viewpoints. Even though the distribution of the interviewees to the organization categories provides a very diverse snapshot of the current situation.

As one researcher only conducted the coding process of the interviews, it might be subject to cognitive biases of the researcher. Furthermore, one interview was not recorded and was therefore not transcribed. The coding of that interview only relies on the written notes of two researchers.

This work was focused on the German automotive industry. The majority of the interview participants were located in southern Germany and therefore, the results might be biased towards *OEMs* from this particular geographic area.

8.2. Future Work

The findings of this work can be used to conduct further research regarding *vehicle-generated data* and a potential API economy. The following section describes several potential research fields.

As mentioned previously the narrow time frame limited the extent of the *GTM* applied and therefore also the theoretical saturation reached. Future research could, therefore, continue by further conducting interviews within this work underrepresented groups like *OEMs*.

Proceeding on the assumption that multiple *APIs* providing *vehicle-generated data* will be available in the future, another interesting research topic would be to examine and evaluate those APIs. The APIs could be evaluated regarding common best practices, their compliance with the proposed *ISO* standards and regarding the available data.

With the prerequisite that multiple *APIs* providing *vehicle-generated data* exist, future research could also be conducted regarding the economic effects for the automotive industry, especially for *OEMs*. Different business models and pricing strategies could be evaluated for both the OEM and *vehicle-generated data* consumers. Another interesting topic could be to validate if the, in this thesis, hypothesized reciprocal benefits affect all economic actors.

Appendix

A. Interview Guideline (German)

Interview

Intro

In den letzten zehn Jahren wurde, über viele verschiedenen Branchen hinweg, eine große Anzahl von APIs öffentlich bereitgestellt. Diese Entwicklung hat zur Entstehung von leistungsfähigen Ökosystemen und neuen Geschäftsmodellen geführt, die es Unternehmen ermöglichen neue Einnahmequellen zu erschließen. Dieser Trend wird auch als API Economy bezeichnet. Innerhalb der Automobilindustrie können wir diesen Trend kaum oder gar nicht. Gerade moderne Fahrzeuge mit unzähligen Sensoren könnten wertvolle Daten für neue Geschäftsmodelle rund um die Automobilindustrie liefern.

Unser Ziel ist es, zu analysieren, warum es keinen solchen Trend gibt und was getan werden könnte, um eine API Economy in der Automobilindustrie zu ermöglichen.

Working Definitions

Vehicle generated data

Daten die von Sensoren innerhalb oder außerhalb eines Fahrzeuges ausgelesen werden können z.B. Temperatur oder Kilometerstand. (Nicht Daten, die vom Gerät eines Benutzers während der Fahrt gelesen werden, z.B. Google Maps Traffic Data).

API Economy

API Economy beschreibt die Wertschöpfung durch die Bereitstellung von digitalen Diensten über APIs entweder öffentlich oder nur an Partner. Dieser Vorgang ist unabhängig von Technologie, Protokoll oder Funktionalität. Der Begriff beschreibt auch API Lifecycle Management Prozesse und Tools.

Interview Purpose

In diesem Interview möchten wir mit Ihnen über den aktuellen Stand von fahrzeuggenerierten Daten und deren Einsatzmöglichkeiten sprechen.

Terms of confidentiality & Tape Recording

Dieses Interview wird anonym durchgeführt, die einzige Information, die wir mit Ihren Antworten in Beziehung setzen möchten, ist eine Klassifizierung Ihres Unternehmens.

a) OEM
b) Automotive supplier
c) Consulting
d) Industrial Association
e) Insurance
f) Research
g) Other :………………….

Um den größten Wert aus Ihrer Expertise zu schöpfen und eine korrekte Forschungsarbeit durchzuführen, bitten wir Sie um Ihre Erlaubnis, dieses Interview aufzunehmen.

Würden Sie uns erlauben, dieses Interview aufzunehmen? → YES / NO

If YES: Würden Sie uns auch gestatten, den Inhalt dieses Interviews direkt zu zitieren, möglicherweise in Verbindung mit der zuvor genannten Unternehmensklassifizierung?

Format of the Interview

Das folgende Interview besteht aus 6 offenen Fragen. Das bedeutet, dass Sie uns Ihre Sichtweise zum dem Thema mitteilen können. Das Interview dauert in der Regel 30 bis 60 Minuten.

Contact & Last Remarks

Wenn Sie nach dem Interview weitere Fragen oder Anmerkungen haben, können Sie sich jederzeit an Gloria (gloria.bondel@tum.de) oder mich (frido.koch@tum.de) wenden.

Haben Sie noch weitere Fragen, bevor wir mit dem Interview beginnen?

Information About Interviewee

1) Wie viele Mitarbeiter sind in Ihrem Unternehmen beschäftigt?
    a. 1 - 10
    b. 11 - 50
    c. 51 - 250
    d. 251 - 500
    e. 500 – 1.000
    f. 1.001 - 2.000
    g. 2.001 – 5.000
    h. 5.001 – 10.000
    i. 10.001 – 50.000
    j. 50.001 – 100.000
    k. Mehr als 100.000
2) Welche Rolle haben Sie in der aktuellen Firma?
    a. Technical Architect (Security Architect, Software Architect, Technical Architect)
    b. Enterprise Architect
    c. Domain Architect / Solution Architect
    d. Business Architect / Process Architect
    e. Internal Audit / External Audit / Revision
    f. External Partner / Consultant
    g. Controller (Business or IT)
    h. Head of Department IT
    i. Head of Department Business
    j. Project Manager IT
    k. Project Manager Business
    l. Portfolio Manager (Business of IT)
    m. Corporate Development / Corporate Governance
    n. Software Developer
    o. Requirements Engineer
    p. Business Analyst
    q. Process Owner
    r. Data Owner
    s. Business Owner
    t. Application Owner
    u. Others: ………………….

Questions

Past experiences

1. Haben Sie Erfahrung mit der Verwendung von fahrzeuggenerierten Daten aus aktuellen oder vergangen Projekten?

   1.1. If YES:

   1.2. Welche Art von Daten verwenden Sie und für welche spezifischen Anwendungsfälle verwenden sie diese Daten?

   1.3. Wie einfach oder schwer war es für Sie die, für den Anwendungsfall relevanten Daten, zu beschaffen?

   1.4. If NO:

   1.5. Warum haben Sie bisher keine fahrzeuggenerierten Daten verwendet?

Future expectation & outlook

1. Welche Art von fahrzeuggenerieten Daten wären zukünftig am relevantesten für Ihre Organisation und was wären potentielle Anwendungsfälle, für diese Art von Daten?
   1.1. Erwarten Sie Barrieren für die genannten Anwendungsfälle, falls Ja welche Barrieren erwarten Sie konkret?
2. Welche anderen Barrieren existieren wenn das Thema fahrzeuggenerierte Daten und potentielle Anwendungsfälle allgemein betrachtet wird?
   2.1. **CHECK Fields:** Legal, Social Acceptance, Technology, Economic
3. Welche, der genannten spezifischen und allgemeinen Barrieren sind besonders herausfordernd?
   3.1. Was macht diese Barrieren so herausfordernd?
4. **If no Barriers:** Warum kann kein weitläufigerer API-Economy-Trend in der Automobilindustrie beobachtet werden?

Possible measures

2. Was sind, Ihrer Meinung nach, Lösungsansätze für die genannten Barrieren?
   2.1. (Welchen Zeithorizont halten Sie realistisch für eine mögliche Umsetzung?)
3. (Können Sie sich eine “killer application” (Disruption) vorstellen die eine weitläufige Verfügbarkeit von fahrzeuggenerierten Daten beschleunigen würde?

Closer

Nachdem wir alle Interviews durchgeführt haben, werden wir die Ergebnisse zusammenfassen. Als nächsten Schritt möchten wir eine kurze Online-Umfrage durchführen, die nicht mehr als 10 Minuten dauern wird. Wir würden uns sehr freuen, wenn Sie an der Online-Umfrage teilnehmen würden.

List of Abbreviations

API .... Application Programming Interface
CAN .... Controller Area Network
CARB .... California Air Resources Board
CVIM .... Common Vehicle Information Model
DLC .... Data Link Connector
DTC .... Diagnostic Trouble Code
ECU .... Engine Control Unit
EOBD .... European On-Board Diagnostics
EPA .... United States Environmental Protection Agency
ExVe .... Extended Vehicle
GDPR .... General Data Protection Regulation
GDV .... Gesamtverband der Deutschen Versicherungswirtschaft
GPS .... Global Positioning System
GTM .... Grounded theory method
HTTP .... Hypertext Transfer Protocol
ISO .... International Organization for Standardization
ITS .... Intelligent Transportation System
JSON .... JavaScript Object Notation
NEVADA .... Neutral Extended Vehicle for Advanced Data Access
OEM .... Original Equipment Manufacturer
REST .... Representational State Transfer
SIM .... Subscriber Identification Module
TLS .... Transport Layer Security
XML .... Extensible Markup Language

List of Figures

List of Tables

Bibliography

[1] I. W. Cotton and F. S. Greatorex Jr, *"Data structures and techniques for remote computer graphics,"* in *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I*, ser. AFIPS '68 (Fall, part I), San Francisco, California: ACM, 1968, pp. 533–544. DOI: 10.1145/1476589.1476661.

[2] B. De, *API Management*. Apress, 2017. DOI: 10.1007/978-1-4842-1305-6.

[3] W. Santos. (2018). Research Shows Interest in Providing APIs Still High, [Online]. Available: https://www.programmableweb.com/news/research-shows-interest-providing-apis-still-high/research/2018/02/23 (visited on 01/13/2019).

[4] W. Tan, Y. Fan, A. Ghoneim, M. A. Hossain, and S. Dustdar, *"From the service-oriented architecture to the web api economy,"* *IEEE Internet Computing*, vol. 20, no. 4, pp. 64–68, Jul. 2016, ISSN: 1089-7801. DOI: 10.1109/MIC.2016.74.

[5] R. C. Basole, *"On the evolution of service ecosystems: A study of the emerging api economy,"* in *Handbook of Service Science, Volume II*, P. P. Maglio, C. A. Kieliszewski, J. C. Spohrer, K. Lyons, L. Patrício, and Y. Sawatani, Eds. Cham: Springer International Publishing, 2019, pp. 479–495, ISBN: 978-3-319-98512-1. DOI: 10.1007/978-3-319-98512-1_21.

[6] W. J. Fleming, *"Overview of automotive sensors,"* *IEEE Sensors Journal*, vol. 1, no. 4, pp. 296–308, Dec. 2001, ISSN: 1530-437X. DOI: 10.1109/7361.983469.

[7] ——, *"New automotive sensors—a review,"* *IEEE Sensors Journal*, vol. 8, no. 11, pp. 1900–1921, Nov. 2008, ISSN: 1530-437X. DOI: 10.1109/JSEN.2008.2006452.

[8] R. T. Fielding, *"Architectural Styles and the Design of Network-based Software Architectures,"* Dissertation, University of California, Irvine, 2000.

[9] K. Holley, S. Antoun, D. A. Arsanjani, W. A. " Brown, J. F. Costas, C. Cozzi, D. P. Goyal, S. Iyengar, H. Jamjoom, C. Jensen, J. Laredo, J. Maddison, R. Narain, A. Natarajan, J. Petriuc, K. Ramachandran, R. Ravishankar, R. Reinitz, S. Vaidya, and M. Vukovic, *"The power of the api economy - stimulate innovation, increase productivity, develop new channels, and reach new markets,"* IBM Corp, Tech. Rep., 2014.

[10] M. Vukovic, J. Laredo, V. Muthusamy, A. Slominski, R. Vaculin, W. Tan, V. Naik, I. Silva-Lepe, A. Kumar, B. Srivastava, and J. W. Branch, *"Riding and thriving on the api hype cycle,"* *Commun. ACM*, vol. 59, no. 3, pp. 35–37, Feb. 2016, ISSN: 0001-0782. DOI: 10.1145/2816812.

[11] P. C. Evans and R. C. Basole, *"Revealing the api ecosystem and enterprise strategy via visual analytics,"* *Commun. ACM*, vol. 59, no. 2, pp. 26–28, Jan. 2016, ISSN: 0001-0782. DOI: 10.1145/2856447.

[12] Verband der Automobilindustrie e. V., *"Access to the vehicle and vehicle generated data,"* Verband der Automobilindustrie e. V, Tech. Rep., 2016.

[13] S. Abdelhamid, H. Hassanein, and G. Takahara, *"Vehicle as a mobile sensor,"* vol. 34, 2014, pp. 286–295. DOI: 10.1016/j.procs.2014.07.025.

[14] S. Abdelhamid, H. S. Hassanein, and G. Takahara, *"Vehicle as a resource (vaar),"* *IEEE Network*, vol. 29, no. 1, pp. 12–17, Jan. 2015, ISSN: 0890-8044. DOI: 10.1109/MNET.2015.7018198.

[15] P. Baltusis, *"On board vehicle diagnostics,"* SAE International, Tech. Rep., 2004.

[16] A. X. A. Sim and B. Sitohang, *"Obd-ii standard car engine diagnostic software development,"* in *2014 International Conference on Data and Software Engineering (ICODSE)*, Nov. 2014, pp. 1–5. DOI: 10.1109/ICODSE.2014.7062704.

[17] A. Reich, N. A. Krämer, and R. Lenninger, *"Fahrzeugdaten-management standardisierter zugang als basis für neue geschäftsmodelle,"* *ATZelektronik*, vol. 13, no. 2, pp. 42–47, Apr. 2018, ISSN: 2192-8878. DOI: 10.1007/s35658-018-0006-0.

[18] D. J. Enriquez, A. Bautista, P. Field, S. Kim, S. Jensen, M. Ali, and J. Miller, *"Canopnr: Can-obd programmable-expandable network-enabled reader for real-time tracking of slippery road conditions using vehicular parameters,"* in *2012 15th International IEEE Conference on Intelligent Transportation Systems*, Sep. 2012, pp. 260–264. DOI: 10.1109/ITSC.2012.6338905.

[19] Z. Szalay, Z. Kánya, L. Lengyel, P. Ekler, T. Ujj, T. Balogh, and H. Charaf, *"Ict in road vehicles — reliable vehicle sensor information from obd versus can,"* in *2015 International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, Jun. 2015, pp. 469–476. DOI: 10.1109/MTITS.2015.7223296.

[20] P. G. Kleimann, D. W. Bernhart, J. P. Hasenberg, A. Brenner, T. Hunstock, F. Kaiser, and R. Eirich, *"Connected car - app based dongle solution as shortcut to connectivity,"* Roland Berger, Tech. Rep., 2016.

[21] C. Hammerschmidt. (2017). German car industry plans to close OBD interface, [Online]. Available: http://www.eenewsautomotive.com/news/german-car-industry-plans-close-obd-interface?news_id=93237 (visited on 01/20/2019).

[22] A. De Mauro, M. Greco, and M. Grimaldi, *"What is big data? a consensual definition and a review of key research topics,"* *AIP Conference Proceedings*, vol. 1644, no. 1, pp. 97–104, 2015. DOI: 10.1063/1.4907823. eprint: https://aip.scitation.org/doi/pdf/10.1063/1.4907823.

[23] E. Curry, *"The big data value chain: Definitions, concepts, and theoretical approaches,"* in *New Horizons for a Data-Driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe*, J. M. Cavanillas, E. Curry, and W. Wahlster, Eds. Cham: Springer International Publishing, 2016, pp. 29–37, ISBN: 978-3-319-21569-3. DOI: 10.1007/978-3-319-21569-3_3.

[24] M. Böhm, G. Koleva, S. Leimeister, C. Riedl, and H. Krcmar, *"Towards a generic value network for cloud computing,"* in *Economics of Grids, Clouds, Systems, and Services*, J. Altmann and O. F. Rana, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 129–140, ISBN: 978-3-642-15681-6.

[25] W. Schladofsky, J. Mitic, A. P. Megner, C. Simonato, L. Gioppo, D. Leonardos, and A. Bröring, *"Business models for interoperable iot ecosystems,"* in *Interoperability and Open-Source Solutions for the Internet of Things*, I. Podnar Žarko, A. Broering, S. Soursos, and M. Serrano, Eds., Cham: Springer International Publishing, 2017, pp. 91–106, ISBN: 978-3-319-56877-5.

[26] McKinsey & Company, *"Monetizing car data – new service business opportunities to create new customer benefits,"* McKinsey & Company, Tech. Rep., 2016.

[27] Verband der Automobilindustrie e. V. (). Normung, [Online]. Available: https://www.vda.de/de/themen/sicherheit-und-standards/normung/normung.html (visited on 01/23/2019).

[28] M. McCarthy, M. Seidl, S. Mohan, J. Hopkin, A. Stevens, and F. Ognissanto, *"Access to in-vehicle data and resources,"* European Commission, Tech. Rep., 2017.

[29] International Organization for Standardization, *"Road vehicles – Extended vehicle (ExVe) 'web services' – Part 1: ExVe content,"* International Organization for Standardization, Geneva, CH, Standard, 2018.

[30] European Automobile Manufacturers Association, *"Access to vehicle data for third-party services,"* European Automobile Manufacturers Association, Position Paper, 2016.

[31] International Organization for Standardization, *"Road vehicles – Extended vehicle (ExVe) 'web services' – Part 2: ExVe access,"* International Organization for Standardization, Geneva, CH, Standard, 2018.

[32] Verband der Automobilindustrie e. V., *"Jahresbericht 2016,"* Verband der Automobilindustrie e. V., Berlin, DE, Rnnual Report.

[33] ——, (). Was ist "NEVADA-Share & Secure"? [Online]. Available: `https://www.vda.de/de/themen/innovation-und-technik/datensicherheit/was-ist.html` (visited on 01/23/2019).

[34] ——, (). Datennutzungskategorien, [Online]. Available: `https://www.vda.de/de/themen/innovation-und-technik/datensicherheit/datennutzungskategorien.html` (visited on 01/24/2019).

[35] AutomatProject. (). Automotive Big Data Marketplace for Innovative Cross-sectorial Vehicle Data Services, [Online]. Available: `http://www.automat-project.eu/` (visited on 01/26/2019).

[36] ——, (). AutoMat Objectives, [Online]. Available: `http://www.automat-project.eu/consortium_list` (visited on 01/26/2019).

[37] J. Pillmann, C. Wietfeld, A. Zarcula, T. Raugust, and D. C. Alonso, *"Novel common vehicle information model (cvim) for future automotive vehicle big data marketplaces,"* in *2017 IEEE Intelligent Vehicles Symposium (IV)*, Jun. 2017, pp. 1910–1915. DOI: `10.1109/IVS.2017.7995984`.

[38] J. Pillmann, B. Sliwa, J. Schmutzler, C. Ide, and C. Wietfeld, *"Car-to-cloud communication traffic analysis based on the common vehicle information model,"* in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, Jun. 2017, pp. 1–5. DOI: `10.1109/VTCSpring.2017.8108664`.

[39] J. Pillmann, B. Sliwa, and C. Wietfeld, *"The automat cvim - a scalable data model for automotive big data marketplaces,"* in *2018 19th IEEE International Conference on Mobile Data Management (MDM)*, Jun. 2018, pp. 284–285. DOI: `10.1109/MDM.2018.00052`.

[40] S. Zimmermann and P. Zoller, *"A Neutral and Secure Approach for Accessing In-vehicle Data for 3rd Party Services,"* BMW Group, Presentation, 2018.

[41] Mercedes Benz. (). Vehicle Lock Status, [Online]. Available: `https://developer.mercedes-benz.com/apis/vehicle_lock_status/overview` (visited on 01/27/2019).

[42] ——, (). Pay As You Drive Insurance, [Online]. Available: `https://developer.mercedes-benz.com/apis/pay_as_you_drive_insurance/overview` (visited on 01/27/2019).

[43] ——, (). Fuel Status, [Online]. Available: `https://developer.mercedes-benz.com/apis/fuel_status` (visited on 01/27/2019).

[44] ——, (). Electric Vehicle Status, [Online]. Available: `https://developer.mercedes-benz.com/apis/electric_vehicle_status` (visited on 01/27/2019).

[45] ——, (). Remote Diagnostic Support, [Online]. Available: `https://developer.mercedes-benz.com/apis/remote_diagnostics_api` (visited on 01/27/2019).

[46] Porsche. (). Porsche NEXT OI Competition, [Online]. Available: `http://www.porsche-next-oi-competition.com/` (visited on 02/03/2019).

[47] J. Webster and R. T. Watson, *"Analyzing the past to prepare for the future: Writing a literature review," MIS Quarterly*, vol. 26, no. 2, pp. xiii–xxiii, 2002, ISSN: 02767783.

[48] B. G. Glaser and A. L. Strauss, *The Discovery of Grounded Theory - Strategies for Qualitative Research*. Aldine Publishing, 1967.

[49] L. M. Given, *The SAGE Encyclopedia of Qualitative Research Methods*. Thousand Oaks, CA: SAGE Publications, Inc., 2008, vol. 1-2, ISBN: 978-1-4129-4163-1. DOI: `10.4135/9781412963909`.

[50] N. Ralph, M. Birks, and Y. Chapman, *"The methodological dynamism of grounded theory," International Journal of Qualitative Methods*, vol. 14, no. 4, p. 1 609 406 915 611 576, 2015. DOI: `10.1177/1609406915611576`.

[51] M. Wiesche, M. C. Jurisch, P. W. Yetton, and H. Krcmar, *"Grounded theory methodology in information systems research," MIS Q.*, vol. 41, no. 3, pp. 685–701, Sep. 2017, ISSN: 0276-7783. DOI: `10.25300/MISQ/2017/41.3.02`.

[52] K. Charmaz, *"Grounded theory: Objectivist and constructivist methods,"* in. Jan. 2000.

[53] D. F. Birks, W. Fernandez, N. Levina, and S. Nasirin, *"Grounded theory method in information systems research: Its nature, diversity and opportunities," European Journal of Information Systems*, vol. 22, Nov. 2013. DOI: `10.1057/ejis.2012.48`.

[54] D. W Turner, *"Qualitative interview design: A practical guide for novice investigators," Qualitative Report*, vol. 15, May 2010.

[55] S. De, *"Next-gen business models for the automotive industry for connected cars and services," SAE Technical Papers*, vol. 2018-April, 2018. DOI: `10.4271/2018-01-0045`.

[56] T. L. Willke, P. Tientrakool, and N. F. Maxemchuk, *"A survey of inter-vehicle communication protocols and their applications," IEEE Communications Surveys Tutorials*, vol. 11, no. 2, pp. 3–20, Second 2009, ISSN: 1553-877X. DOI: `10.1109/SURV.2009.090202`.

[57] J. E. Siegel, D. C. Erb, and S. E. Sarma, *"A survey of the connected vehicle landscape—architectures, enabling technologies, applications, and development areas," IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 8, pp. 2391–2406, Aug. 2018, ISSN: 1524-9050. DOI: `10.1109/TITS.2017.2749459`.

[58] I. Cassias and A. L. Kun, *"Vehicle telematics : A literature review,"* 2007.

[59] M. Hauschild, *"Weather and road surface information and hazard warnings data content acquisition through advanced probe vehicle systems,"* vol. 3, 2009, pp. 1609–1619.

[60] N. AbuAli, *"Advanced vehicular sensing of road artifacts and driver behavior,"* in *2015 IEEE Symposium on Computers and Communication (ISCC)*, Jul. 2015, pp. 45–49. DOI: `10.1109/ISCC.2015.7405452`.

[61] E. Massaro, C. Ahn, C. Ratti, P. Santi, R. Stahlmann, A. Lamprecht, M. Roehder, and M. Huber, *"The car as an ambient sensing platform," Proceedings of the IEEE*, vol. 105, no. 1, pp. 3–7, 2017. DOI: `10.1109/JPROC.2016.2634938`.

[62] J. Ni, A. Zhang, X. Lin, and X. Shen, *"Security, privacy, and fairness in fog-based vehicular crowdsensing," IEEE Communications Magazine*, vol. 55, no. 6, pp. 146–152, 2017. DOI: `10.1109/MCOM.2017.1600679`.

[63] H.-J. Yun, S.-K. Lee, and O.-C. Kwon, *"Vehicle-generated data exchange protocol for remote obd inspection and maintenance,"* 2011, pp. 81–84.

[64] J. Guerrero-Ibáñez, S. Zeadally, and J. Contreras-Castillo, *"Sensor technologies for intelligent transportation systems," Sensors (Switzerland)*, vol. 18, no. 4, 2018. DOI: `10.3390/s18041212`.

[65] E. Turk and M. Challenger, *"An android-based iot system for vehicle monitoring and diagnostic,"* 2018, pp. 1–4. DOI: `10.1109/SIU.2018.8404378`.

[66] K. Khorsravinia, M. K. Hassan, R. Z. A. Rahman, and S. A. R. Al-Haddad, *"Integrated obd-ii and mobile application for electric vehicle (ev) monitoring system,"* in *2017 IEEE 2nd International Conference on Automatic Control and Intelligent Systems (I2CACIS)*, Oct. 2017, pp. 202–206. DOI: `10.1109/I2CACIS.2017.8239058`.

[67] C.-M. Tseng, W. Zhou, M. Al Hashmi, C.-K. Chau, S. G. Song, and E. Wilhelm, *"Data extraction from electric vehicles through obd and application of carbon footprint evaluation,"* in *Proceedings of the Workshop on Electric Vehicle Systems, Data, and Applications*, ser. EV-SYS '16, Waterloo, Ontario, Canada: ACM, 2016, 1:1–1:6, ISBN: 978-1-4503-4420-3. DOI: `10.1145/2939953.2939954`.

[68] A. Düsterhöft and E. Brandmayer. (2018). Telematik-Tarif bei der Kfz-Versicherung, [Online]. Available: `https://www.finanztip.de/kfz-versicherung/telematik-tarif/` (visited on 02/04/2019).

[69] S. Husnjak, D. Peraković, I. Forenbacher, and M. Mumdziev, *"Telematics system in usage based motor insurance,"* January, vol. 100, 2015, pp. 816–825. DOI: `10.1016/j.proeng.2015.01.436`.

[70] M. Streich, A. D'Imperio, and J. Anke, *"Bewertung von anreizen zum teilen von daten für digitale geschäftsmodelle am beispiel von usage-based insurance," HMD Praxis der Wirtschaftsinformatik*, vol. 55, no. 5, pp. 1086–1109, Oct. 2018, ISSN: 2198-2775. DOI: 10.1365/s40702-018-0432-8.

[71] *"Directive (EU) 2015/2366 of the European Parliament and of the Council of 25 November 2015 on payment services in the internal market, amending Directives 2002/65/EC, 2009/110/EC and 2013/36/EU and Regulation (EU) No 1093/2010, and repealing Directive 2007/64/EC," Official Journal of the European Union*, vol. L 337, pp. 35–127, Dec. 2015.

[72] C. Kirsch. (2014). Preiskampf: Amazon verbilligt Cloud-Dienste drastisch, [Online]. Available: https://www.heise.de/ix/meldung/Preiskampf-Amazon-verbilligt-Cloud-Dienste-drastisch-2155847.html (visited on 02/11/2019).