A Socio-Technical Analysis of the Importance of primary data for the Decarbonisation in road logistics

The impact of onboard sensors on road logistics to improve the estimation accuracy of emission factors

Master Thesis

Nicolas Will

Delft University of Technology

TUDelft

A Socio-Technical Analysis of the Importance of primary data for the Decarbonisation in road logistics

The impact of onboard sensors on road logistics to improve the estimation accuracy of emission factors

by

Nicolas Will

to obtain the degree of Master of Science
in Complex System Engineering & Management.
Faculty of Technology, Policy and Management
at the Delft University of Technology,
to be defended publicly on Monday, August 28, 2023, at 9:30 AM.

Chair:	Prof. Dr Ir. LA. Tavasszy (Section: TLO)
First Supervisor:	Dr JA. Annema (Section: TLO)
Second Supervisor:	Dr J. Ubacht (Section: ICT)
Company Supervisor:	T. Bohnhoff (shipzero)
Faculty:	Technology, Policy & Management, Delft

Cover:	shutterstock.com from metamoreworks, modified by Nicolas Will
Style:	TU Delft Report Style, with modifications by Nicolas Will

Acknowledgements

This thesis marks the end of my academic journey at the TU Delft, a time that I will hold very dear to my heart due to the personal friendships and academic connections I was able to make. Many people and friends influenced my time in Delft, making it so fun and challenging. First, I would like to thank my thesis committee, which constantly supported and challenged me throughout the research and topics, for the great feedback and guidance throughout the last semester.

My thanks go to my chair Lóri Tavasszy for his conceptual and constructive feedback throughout our meetings, his initial contact with the Smart Freight Centre that kick-started this research, and the industry contacts that proved vital for the design of this research. Next to that, I would like to thank my two supervisors Jan Anne Annema and Jolien Ubacht. Thank you both for your time, knowledge and commitment throughout the time and our biweekly meeting to support me with the needed feedback. Thank you, Jan Anne, for your deep understanding of the transport and logistical sector and your help while developing the methodological approach. Thank you, Jolien, for your insights into the design research approach and your technical know-how of digital system architectures. The combination of knowledge from both of my supervisors proved to be a great advantage for this multi-disciplinary study.

Next to that, I would also like to thank Violetta Matzoros from the Smart Freight Centre for connecting me with my collaborating company, for her feedback on my approach and research direction and for the invitation to the Smart Freight Week in Amsterdam. The conference was not just a fantastic experience to see all the efforts in the transport sector to reach carbon neutrality but also proved vital as a knowledge source throughout the research with the connections I made during the two-day event. Next, I would also like to thank all my interview partners who participated and helped me develop and evaluate the system and designs.

I would also like to express my gratitude to shipzero for collaborating with me on this interesting research topic. It was great to see how my home town Hamburg is a hub for state-of-the-art GHG emissions calculations in this new and emerging sector. Here I would especially like to thank Tobias Bohnhoff for his constant efforts and detailed insight into the topic that only a few know as detailed as he does. Next, I would like to thank my „company buddy " Tristan, who made my onboarding as smooth as possible and had excellent knowledge of data management structures and GHG emission calculations, and Olga, who gave me new perspectives on sustainability in the transport sector. However, I would also like to thank the entire team of shipzero, who all made my time there so fun, and I cannot wait for the next Darts match in the summer.

Lastly, I would also like to thank my friends and family, who supported me through this time and always had my back. Thank you for being so supportive and also telling me when to step back to see the bigger picture. A special thanks go to my parents, who supported me in any way imaginable, enabling me to devote my time to my passion and my academic journey. I know that this is by no means the norm, and I am truly grateful for that. Vielen Dank das ihr mir dies ermöglicht habt.

Nicolas Will
August 2023, Delft

Executive Summary

The transport sector is one of the most significant contributors to European GHG emissions and is expected to grow even further in the coming years. The target of the European Union to reduce transport-related GHG emissions by 55% by 2030 seems to be a challenging task, keeping in mind the projected growth rates of the sector. In order to achieve the set-out targets, transparent and standardised GHG emissions quantification methodologies and reporting schemes must exist to monitor the GHG emissions of vehicles and reduce them where possible. Vehicle emissions used to be approximated with standard emission factors per vehicle class. However, the accuracy of these traditional emission factors is limited, especially on a larger scale, such as the European Union, leading to inaccurate vehicle emissions. Thus, primary fuel consumption data could be used to calculate the emissions of vehicles accurately by collecting the primary data via onboard sensors. The combination of data management structures for primary data in road logistics with the possibility of improving the accuracy of emissions calculations and reporting from a socio-technical perspective is currently not sufficiently addressed in scientific literature. The study is also highly relevant from a private interest, as new legislative decisions force private companies to declare their emissions on only validated and correct emissions factors.

The research aimed to investigate the current practices of using vehicles' primary data to improve GHG emissions' accuracy by creating a system architecture for the data flow from the vehicle to the final visualisation of the GHG emissions. For that, the design science research methodology (DSRM) approach was chosen to answer the main research question of:

"What onboard sensor systems architecture can enable road logistics operators to gather primary data from their fleet to accurately determine their vehicle emissions?"

Four sub-questions were formulated in line with the used DSRM design cycle to answer the main research question and to contribute to the existing body of knowledge via a socio-technical analysis, system requirements, system architecture, and an evaluation.

The research utilised interviews, scientific literature, and informal conversations with industry players as the main knowledge source. The expert interviews were first used during the design process to understand the environment, derive system requirements for the later design, and to create a stakeholder overview. In the second phase of the research, the experts were utilised to evaluate the created design. Interview partners were selected based on their role in the system and potential expertise to help steer the design and evaluation. The feedback received was directly implemented in the designs.

The first step of the analysis was the socio-technical analysis. It revealed the first requirements and design principles for the later design phase, based on the institutional setting and stakeholder demands derived from interviews and the available literature. It also showcased the active and influencing role of the EU in the system, which underlined the need for a socio-technical analysis. Lastly, the stakeholder overview visualised the transport sector's highly fragmented and multi-stakeholder domain, which industry experts evaluated and approved.

The complete system requirements were established and finalised in the second phase of the research. They were separated into three clusters: institutional, stakeholder, and technical-related requirements and were further categorised into functional and non-functional system requirements. Design principles were also created to steer the design process. Six functional and ten non-function system requirements were derived and four design principles. The requirements were evaluated and approved by the expert interviews and were used as the main input for the design phase. The main conclusion was the stakeholder-specific characteristic of some of the requirements due to the different needs of logistical actors and related IT companies that calculated GHG emissions.

The third phase was the designing of the system architecture. It was separated into two parts. First, the creation of a list of possible design options to address the derived system requirements with another evaluation round with the expert. Second, the creation of the system architecture by creating system architecture components, which incorporate the most fundamental design options from the design phase. The experts again evaluated these system architecture components before they were incorporated into stakeholder-specific system architecture, which captured the overarching processes and data flows of an IT company that specialised in the quantification of GHG emissions of vehicles in road logistics. The main conclusion was that, due to the diversity of system stakeholders, a general system architecture that adresses all stakeholder needs is less feasible than the creation of stakeholder-specific system architectures.

The fourth phase was the evaluation, which happened throughout all stages of this research, and concluded the general correctness of the derived stakeholder-specific system architecture by the experts. It also pointed out potential limitations of the designs. The specific knowledge needed to validate such designs of the technical domain (data management structures), the policy and institutional knowledge, and the specific details of state-of-the-art GHG quantification methodologies makes evaluating the entire system more challenging. Thus, a broad sample of interview partners was needed. Moreover, selecting a design option, especially in the perception and physical layer of the system architecture, can create path dependencies and narrow down the design space. The evaluation phase was concluded by addressing the general success factors of the proposed design. Here the willingness of the logistical operators to adopt the GHG emissions reporting, the importance of methodology alignments and the need for truly value-adding services were especially highlighted as success factors of the system architecture.

The research concluded by recognising the great potential of primary data to improve the accuracy of emission factors in road logistics. Seven main conclusions and contributions to the field of logistics were made:

1. The inclusion of a multi-domain designer perspective when designing an abstract system architecture in the logistical sector - should be mandatory in the scoping of any project.
2. The identification of relevant stakeholder clusters and an abstract stakeholder analysis to be considered when designing in the socio-technical environment.
3. The categorization of systems requirements into institutional, stakeholder and technical requirements to represent the multi-domain character of the system.
4. The need for financial quantification tools of the CO_2 reduction for logistical operators to validate their investment decisions.
5. Compliance with leading European GHG emissions quantification methodologies and data regulations - should also be implemented, e.g. in the EU taxonomy.
6. A stakeholder cluster-specific system architecture, which incorporated the derived requirements and outlined business relationships and data flows in the system, evaluated and approved by industry experts.
7. The issue of the stakeholder-specific requirements for the system leading to multiple co-existing system architecture specifications.

Finally, the research concluded with a short and long-term outlook of how the sector might develop and presented potential future research topics, such as the possibility of using other forms of data sharing, such as data spaces or blockchain applications, for secure and trusted data sharing.

Contents

List of Figures

List of Tables

Nomenclature

Abbreviations

Abbreviation	Definition
3 & 4PL	Third & Fourth Party Logistical Operators
ACEA	European Automobile Manufactures Association
CO	Carbon monoxide
CO2	Carbon dioxide
COPERT model	Calculation of air pollutant emissions from road transport
CORSIA	Carbon Offsetting and Reduction Scheme for International Aviation
DACH-region	German-speaking countries within Europe (Germany (D), Austria (A), and Switzerland (CH)).
EEA	European Environment Agency
EF	Emission factors
ERMES	European research on mobile emissions sources
FEDeRATED	European project for interoperability of data platform communication
FMS	Fleet Management System
GHG	Greenhouse gases
GLEC framework	Global Logistics Emissions Council
HC	Hydrocarbons
HDV	Heavy-Duty Vehicles
ISO	International Organisation for Standardisation
JRC	European Research Centre
LAT	Aristotle University - Laboratory of Applied Thermodynamics
LSP	Logistical Service Providers
NOx	Nitrogen Oxides
OBD systems	Onboard diagnostic systems
OBFCM	On-board fuel consumption meter
OEM	Original Equipment Manufacturer
PEMS	Portable emissions measurement system
PM	Particulate matter
RDE	Real-driving emissions
rFMS	Remote Fleet Management System
S1, S2, S3	Scope 1, 2 and 3 emissions
SaaS	Software as a Service
TOC	Transport Operation Characteristic
VVA	Verification, Validation and Assurance companies
WLTP	Worldwide Harmonised Light Vehicle Test Procedure

1
Introduction

This chapter will introduce the research by providing a first overview of emission factors in road logistics by establishing the relevance and the background to the topic (see section 1.1). It is followed by an introduction to the determined knowledge gap and the importance of accurate emission quantification to meet future legislative and commercial demands. The first drawbacks of current emission estimation methodologies are considered. Section 1.2 will give a brief overview of the coming research, and section 1.3 will mention the involved parties in the research. Finally, section 1.4 will outline the structure of the thesis.

1.1. Introduction to Emission Factors & Context

Reducing emissions to tackle and mitigate the effects of global warming is one of the major challenges for the next decades. Multiple national and international agreements and policy documents, such as the Paris Agreement or the European Green Deal, are aimed at reducing global and European CO_2 emissions to limit global warming to below 2°C [37, 35]. The European Green Deal goes even further by aiming to reduce the carbon emissions of the European transport industry by 55% by 2030 [35]. Global trade and logistics are one of the major contributors to carbon emissions, with 27% of all European emissions being caused by them in 2018, while the need for trade is continuously growing [39, 74]. Within the transportation sector (definitions and scope will be defined in section 2.2), road logistics contributes the most to the total emissions output with more than 71% in 2018 (see Fig. 1.1), which even increased to 77% in 2020 [39, 40]. Emissions caused by logistical road operations via light and heavy-duty trucks and buses are the second largest part of those emissions, with 9% and 19% respectively [39, 40]. Fig. 1.1 summarises the emissions contribution per logistical sector in the EU. The significance of road logistics vehicles for the European ambition to reduce emissions is further supported by the legislative decisions to make Europe a centre for innovative and techno-economically driven logistical operations [38].

A first step towards decarbonising the road transport sector is to accurately quantify Greenhouse gas emissions (GHG) to identify emissions hotspots. Detailed insights into the sources of emissions in road logistics via primary data have three advantages. From a policy perspective, they enable the creation of meaningful policies to facilitate emission reduction by directing the policies towards areas with the greatest potential for emission reductions. Additionally, coming policies such as the mandatory reporting of impacts of corporate activities on the environment for private companies from 2025 via the Corporate Sustainability Reporting Directive (CSRD) will require the EU to have a better understanding of the emitted emissions to control the reporting [51]. From a commercial perspective, greater availability of emissions data can improve logistical operations by optimising fuel consumption or driving behaviour. Moreover, the mandatory reporting directives require companies to have detailed access to their logistical data and accreditation schemes can lead to an improved brand image. Additionally, they also foster innovative practices from companies, which is also a defined goal of the EU, and thirdly improve the possibilities for end consumers to make sustainable decisions in their daily life. These three aspects of accurate emission factors via primary data are summarised in the following thought experiment:

Since we are all consumers, let's take "goods we buy" as an example. Besides price, quality, and overall offer, we would want to know which product is less harmful to our environment to make a conscious choice. But, if we have only generalised/average industry or sector data, all products would have the same emission scores. This leads to 2 dilemmas: First, we as consumers cannot make a conscious choice. The principle of the survival of the better is levered out. Second, innovation is not attractive for businesses since it does not bring a competitive advantage if their effort is not reflected in the sustainability score of the product or in the worst case, not even visible to the end consumers. To overcome this challenge, two things are needed: Accurate primary emission data for the specific goods purchased. In the context of my research, an accurate way to measure the specific emission of the transport of the purchased good combined with a widely accepted way to split the emissions for all the goods that were transported simultaneously.

The thought experiment highlights how primary data can contribute to the decarbonisation of road logistics while facilitating innovative practices and giving consumers more decision power. However, the status quo of using primary data to determine emission looks different.

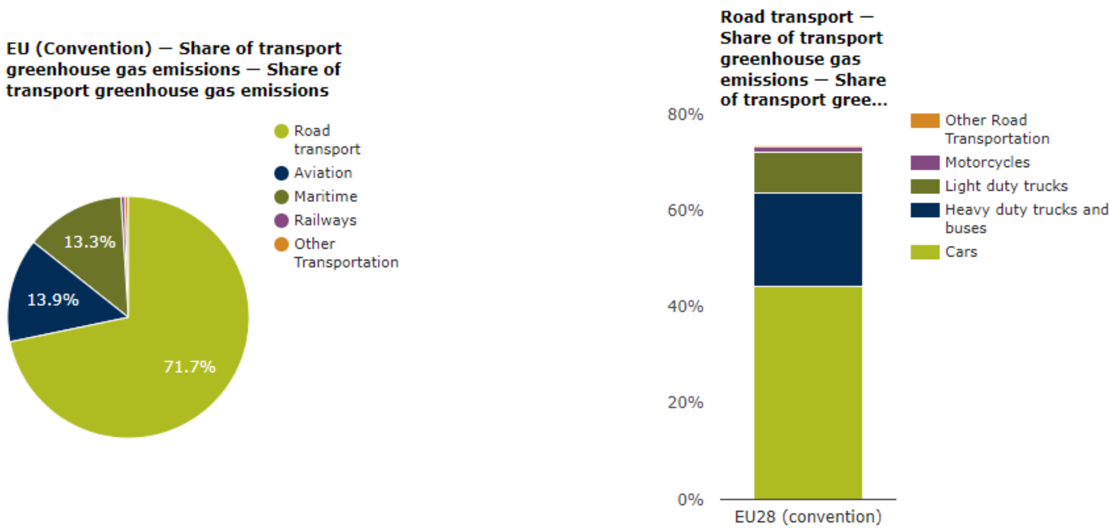

Figure 1.1: Emissions caused by the transport sector in the European Union in 2018, Share of caused emissions by transport mode (left), Share of road transport modes (right). Retrieved from [39]

Today's carbon emissions of road logistics are mainly estimated and determined by two methods. Either through real-life experiments in laboratories, which are proven to underestimate the actual CO_2 emissions [83] or via models and estimations [85]. Emission factors are constants used to determine CO_2 pollution in vehicles. They are dependent on multiple categories such as vehicle categories, fuel types, age of the vehicles, driving characteristics, speed of the vehicles, and many more (see Fig. 1.2 for a complete list of factors that affect the emitted emissions of vehicles) [85]. However, even the fuel consumption data used to calculate the emission factors (fuel consumption of vehicles) are based on experimental data. Thus, the emission factors are based on past measurements. However, the accuracy of those measurements is often limited, as the experiments are mainly done in a controlled laboratory environment, which leads to an inaccurate representation of driving styles and caused emissions. Other aspects are continuing improvements in fuels, aerodynamics or road surfaces, which are difficult to capture consistently in standard factors. The potential gap between real-life and estimated emissions can be insignificant on a vehicle level. However, it can be impactful if these results are extrapolated nationally or on a European level, especially if they are being used for policy decisions.

Another complexity of emission factors is the sheer volume and categories of both vehicle constants and possible models to calculate emissions, which increases the effort of accurately calculating emissions of vehicles. As an example, Ligterink, van Zyl & Heijne established in their report from 2016 for the Dutch CO_2 emissions for road vehicles a total of 325 different emission inventory categories for three different driving environments (urban, rural, and motorway), thus resulting in a total of 975 different emission factors (even excluding other parameters such as street slope, acceleration, or weight) [85]. Modern calculations of emissions factors are more and more data-management tasks, with carriers[1] needing detailed information on the load factors of their vehicles. However, uncertainties in the provided emissions data per vehicle category make the estimates even more complicated even if all other required information are present [26, 60]. Next to the problem of data uncertainty is the issue of constantly updating these various factors. The mentioned improvements must be accounted for in the constants and the resulting emissions calculations. Thus, emissions based on accurate and fuel-consumption-based primary data can reduce data uncertainties and enable detailed tracking and reporting of emissions, which is currently not adequately addressed next to the aforementioned legislative issues [67]. Fig. 1.2 visualises the external factors that impact the emitted emissions of a vehicle [2].

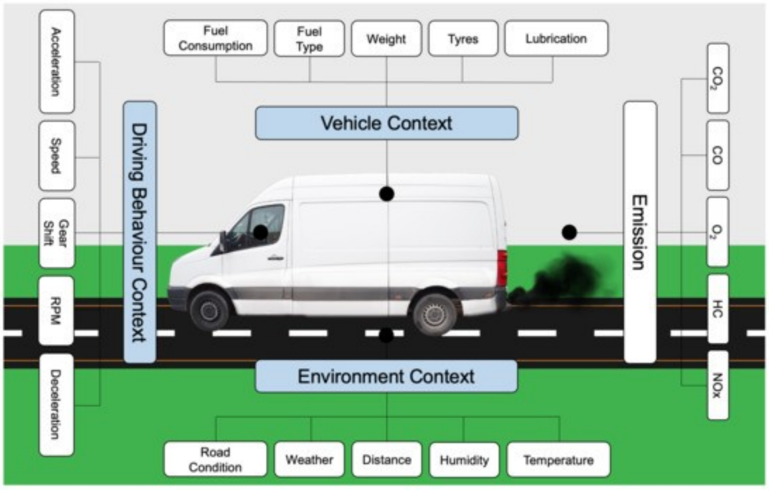

Figure 1.2: External factors impacting the emission output of road vehicles. Retrieved from Yavari et al. (2022) [125, p. 5].

A different approach to tracking the emission of vehicles is based on collecting and monitoring real-life emission data via onboard sensors (Telematics, IoT, or other data-gathering devices). Thus, this method uses the actual emission data from the vehicles and does not rely on emission factors and vehicle constants to estimate the emissions. This can be done in different ways. One approach is the addition of onboard sensors close to the vehicle's exhaust to measure the CO_2 being emitted [100]. The exhaust data is then collected by the sensor and the sensor hardware to be stored on the cloud [100]. Another measurement option is to collect the fuel consumption data and combine it with the fuel type of the vehicle to calculate the exact emissions. The advantages of these accurate emission tracking options via the primary data of the vehicle come with a downside to data management strategies. Depending on the sensor system, large volumes of data can be collected and needs to be managed and stored. Thus companies must create plans for handling the new flood of data from their vehicle fleets [115, 96].

Bojan et al. (2014) presented one possible systems architecture for an IoT-based emission tracking sensor system, which creates a constant flow of raw data from the vehicle to the web servers and databases (see Fig. 1.3) [14]. The proposed setup had three different sensors (location, computer, and ambience)[3]. The location system was able to track the vehicle's location via a GPS receiver, a Global Systems for communication modem (GSM), and a micro-controller to track the vehicle's location [14]. The GPS also tracked the speed, time steps and longitudinal and latitudinal data [14]. Lastly, the ambient system tracked the vehicle's internal and external temperature and humidity [14]. The raw data was sent every 5 minutes, and the research concluded a positive result for the accuracy of the location, temperature, and humidity tracking [14]. Thus, applying the same strategy to a CO_2 or fuel

[1] Logistical operator that is responsible for the transportation of goods
[2] Article uses the total weight of the vehicle (including the parcel weights) as an input, thus not directly the load factors for the emissions calculations
[3] This researcher did not cover any edge computing application on the vehicle itself

consumption sensor could be possible. This raw data needs to be cleaned and analysed to be used and displayed. This systems architecture can result in large data volumes depending on the fleet size, IT capabilities or tracking frequency. This method is a typical example of tackling the challenge from a purely technical standpoint while excluding important institutional and socio-technical considerations of the different stakeholders.

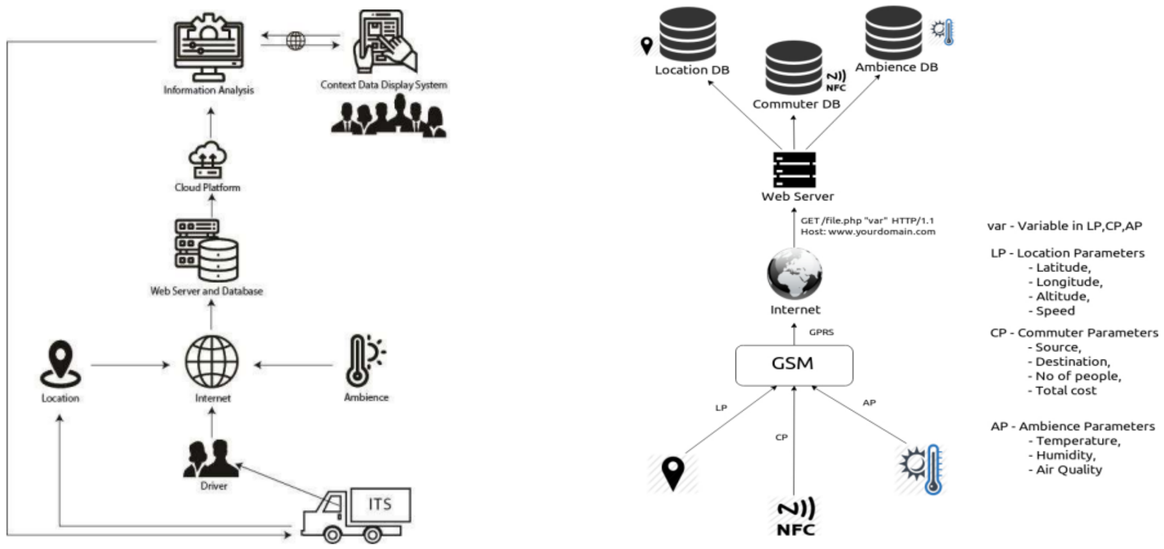

Figure 1.3: Proposed systems architecture by Bojan et al. (2014) [14, pp. 2-4].

Institutional and interoperability requirements lay the foundation for the estimation/calculations of emissions factors, regardless of which approach is used. These can be due to data sharing and privacy implications, such as the GDPR or new industry standards for reporting and calculating emissions in the logistical sector, such as the new ISO 14083:2023 standard [75, 53]. Recent legislative developments demand companies to report their emission in greater detail, and current less detailed practices are insufficient. Thus, there is a need to shift towards methods that can capture these emissions in more detail, which is being facilitated from a policy and commercial perspective. Moreover, current discussions are mainly focused on technical considerations while leaving other logistical stakeholders' needs and the socio-technical environment untouched, even though these two domains play a key role in the legislative framework and functionality of the system.

1.2. Introduction to the research

The research is situated at the border of the socio-technical & institutional environment of GHG emissions calculations of logistical road vehicles and the technical domain of data management strategies to potentially manage and maintain an extensive system that can collect, store, share and process the primary vehicle data for the GHG emissions calculations. The impact of purely primary-based GHG emissions could reduce the gap between actual emissions and reported emissions of vehicles. This gap must be closed to address, on the one hand, the European ambition of decarbonising the (road) logistical sector with the help of impactful policies, which can only be established if emissions hotspots are identified. On the other hand, to optimise logistical operations with the available emissions data on fuel consumption. However, the scope of such a system with its various stakeholder, requirements and potential challenges is not yet fully understood, as key multi-domain perspectives are not incorporated yet (see Fig. 1.4).

The necessary knowledge base and technical insights will be obtained through a literature review in chapter 2. Twenty-one[4] articles were selected for this, focusing on both emission and emission factors

[4] Eighth additional external documents were also selected to also look at the policy and commercial perspective, explained in section 2.1

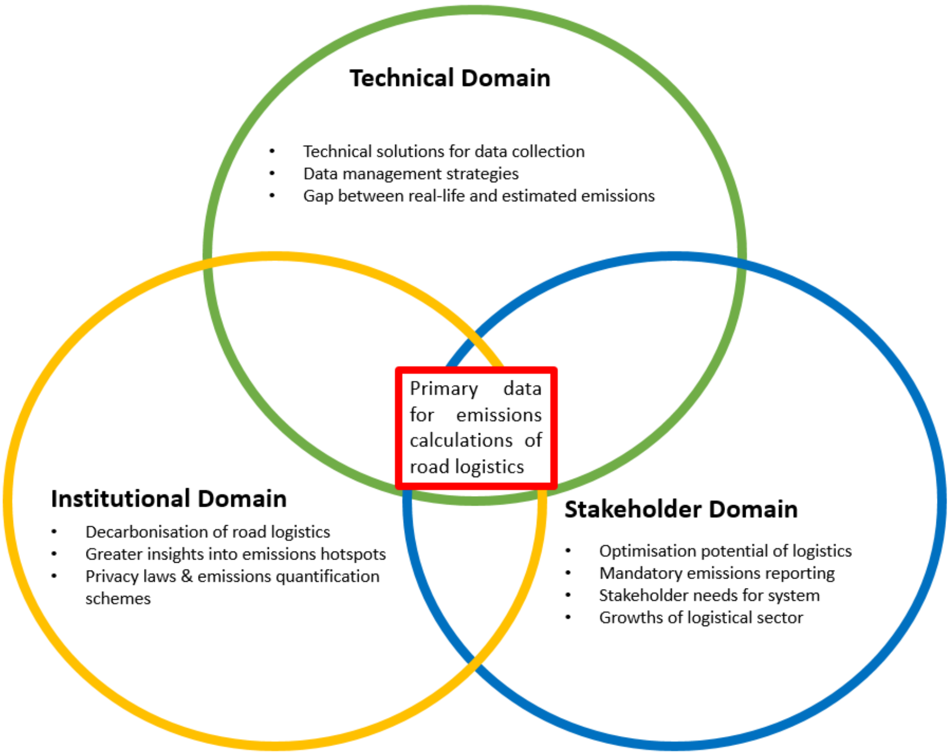

Figure 1.4: First understanding of the multi-domain influences on using primary data to calculate GHG emissions in road logistics.

in combination with onboard sensor systems and their potential for more accurate emissions. Other topics are the current industry standards for reporting and calculating transport-related emissions and the role of the European Union. These standards and policy documents will focus on the institutional domain. Three databases with distinct search terms were used to establish the scope and exact research questions. The literature will stay a knowledge input throughout this research in addition to expert interviews and industry experts, highlighting the research's stakeholder domain. The literature revealed the knowledge gap that onboard sensor systems could provide 100% accurate emissions for road logistic vehicles via primary data. However, very little is known about the exact data structures needed to support such systems for road logistics, especially considering the socio-technical environment and not just the technical domain. Thus, the main research question is:

"What onboard sensor systems architecture can enable road logistics operators to gather primary data from their fleet to accurately determine their vehicle emissions?"

The main research question aims to address the technical, stakeholder and institutional domains of the problem to capture the socio-technical considerations of the problem, as this is currently not present in the existing body of knowledge. Thus, the thesis aims to create an abstract system architecture that incorporates technical, institutional and stakeholder perspectives within the road logistics sector for the usage of primary data for GHG emissions calculations.

The research will utilise a design science approach, specifically the Design Science Research Methodology (DSRM) approach to answer the main research question by proposing a system design for using primary data in road logistics for the calculation of GHG emissions. The DSRM approach has different stages that correspond and can be used to understand the multi-domain character of the research. Each design stage will have a research sub-question to steer and answer the main research question. The first sub-question will evaluate the characteristics of the socio-technical environment and the stakeholder needs. The second will discuss the technical requirements. The third is focused on the design of a potential system, and the final fourth question will be used to evaluate/ validate the outcomes of each previous sub-question. The research and sub-questions will be facilitated and evaluated by interviews with experts from the industry that will be accessed via the collaborating parties of the research. The involved parties are explained in the next section.

1.3. Involved Parties

The research was conducted in collaboration with shipzero. Shipzero is a Hamburg-based startup operating a data-infrastructure platform which companies can use to manage and organise their logistical emissions data. It enables the automated calculation of GHG emission and time savings via primary logistical data [111]. They also offer reporting possibilities of emission factors and benchmark analyses. Shipzero first collects its customers' data from various sources, such as ERP systems and telematics providers. Then it cleans and checks the data for its accuracy and completeness. Based on that, they can calculate emissions via direct energy consumption and activity-based modelling. The data and insights are then connected and communicated to their partners and customers. They fully confirm with the international GLEC framework and are also a member.

The Smart Freight Centre is a non-profit organisation with the aim to decarbonise the logistical industry and net-zero logistics [113]. They are an industry-wide accepted organisation and have published multiple frameworks and position papers on how to accurately quantify GHG emissions in the logistical sector and recommendations on achieving a carbon-neutral logistical sector for various modes of transport. Their most famous framework is the "Global Logistics Emission Council Framework" (GLEC Framework), used throughout the industry and continuously updated to reflect state-of-the-art techniques [112]. They are an indirectly involved party as they are a close partner of shipzero and can provide inside information on the global freight system [113].

1.4. Structure of the Thesis

The remainder of the paper will first address a state-of-the-art literature analysis in chapter 2 followed by the research approach and methodology in chapter 3. Chapter 4 initiates the research with the first sub-question and a socio-technical analysis of the space. Chapter 5 will finalise the requirement analysis and creation. The design phase of the research starts in Chapter 6 with the creation of design options that address the established requirements and artefact design of the most important concepts. The evaluation of the artefact design will be done in chapter 7. Moreover, chapter 7 will elaborate on the limitations of the proposed design. Lastly, chapter 8 will conclude the research by providing high-level implications of the research and providing key recommendations to relevant stakeholders.

2
Literature Research

This chapter will present a state-of-the-art literature review of current practices for emissions tracking in road logistics and the derived knowledge gap with the corresponding main research question. Section 2.1 will outline the criteria, search strings for the article selection and the scoping of the articles, followed by a definition of core concepts in section 2.2 and the presentation of the results of the selected articles in section 2.3. There are three different parts of the literature review. The first part will cover the current limitations of emission factors (sections 2.3.1 and 2.3.2). The second part will discuss onboard sensor systems in combination with road logistics and emission factors (section 2.3.3), and the third part will elaborate on institutional consideration and GHG emissions quantification methodologies (section 2.3.4). The results of each literature review section will be summarised in Tables that cover the main findings and contributions. Finally, section 2.4 will conclude with the main research question of this study, based on the defined knowledge gap.

2.1. Article selection

The article selection will follow the PRISMA 2020 method, which is used to help researchers to communicate their selection criteria transparently [103]. The method involves a three-step procedure. The first step is the identification of databases and open libraries to access relevant literature. The second stage is the screening, which outlines selection criteria. Lastly, step three will summarise the used literature [103].

For the first step, three databases were used. The first and main one was Scopus, which selected relevant articles based on a ”Title, Abstract & Keyword” selection criteria (the used keywords can be found in Tables 2.1 and 2.2). The second database was Google Scholar, which ranks articles based on the number of citations they have received, the year it was published, the author of the article and if other articles have recently cited it [61]. Lastly, the platform Connected Papers was used. Connected Papers is a platform that can directly compare a paper with more than 50.000 other papers. For that, it will compare the references and citations of the input paper with all other relevant articles based on similarities [18]. Thus, two papers can be highly alike even though they have not cited each other. The results are presented in a weighted graph, where two highly similar articles are displayed as larger, and the colour indicates the year of the publication. Fig. 2.3 visualises an exemplary output of the graph for the earlier mentioned article by Bojan et al. (2014) [14], and Fig. 2.1 visualises the used search terms per database.

Knowledge was also added from external and internal sources, identified in talks and conversations with other researchers and experts in road logistics and emissions factors. They were a combination of policy documents and emissions calculation methodologies, academic literature, and company press releases about the impact of policy decisions on the logistical sector. These documents were not gathered via traditional literature research but added especially valuable information for the regulatory analysis of emission factors. Table E.3 visualises the documents that were gathered via external sources.

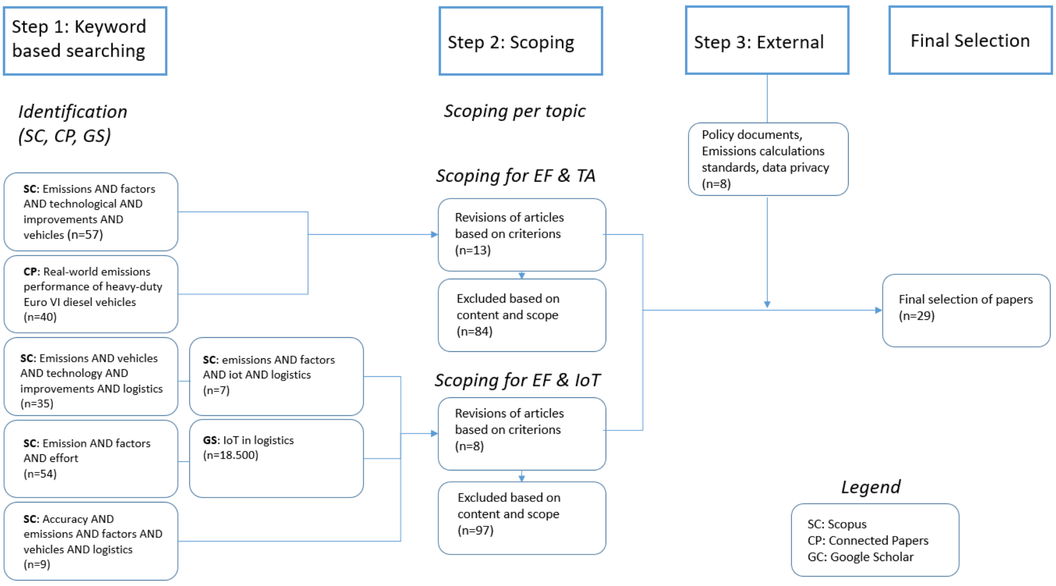

Figure 2.1: Detailed outline of the search terms used per topic.

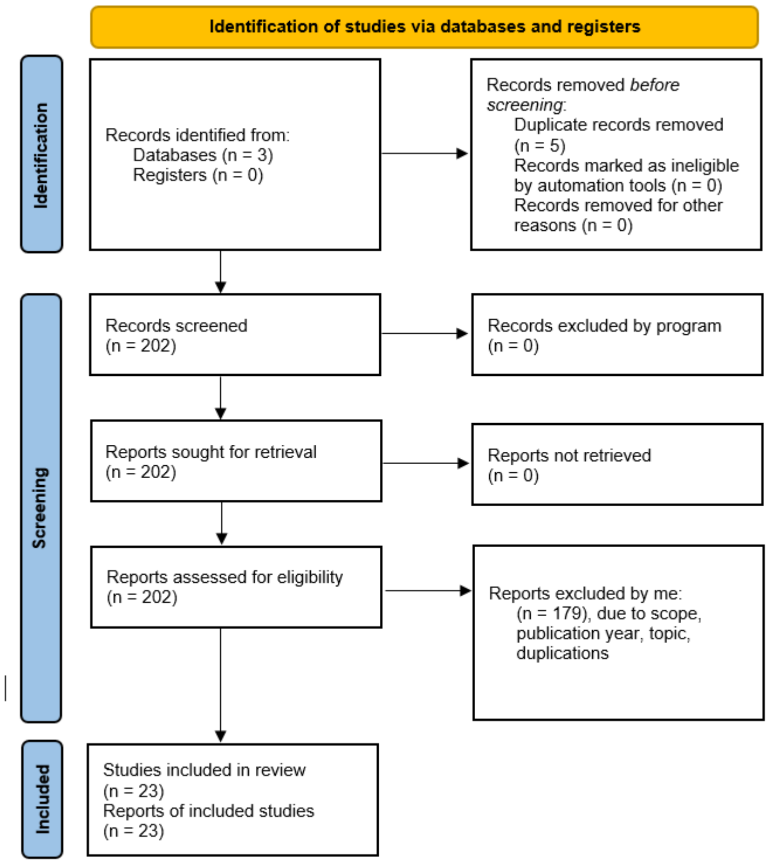

Figure 2.2: PRISMA 2020 flow diagram, which outlines the search strategy for the literature review.

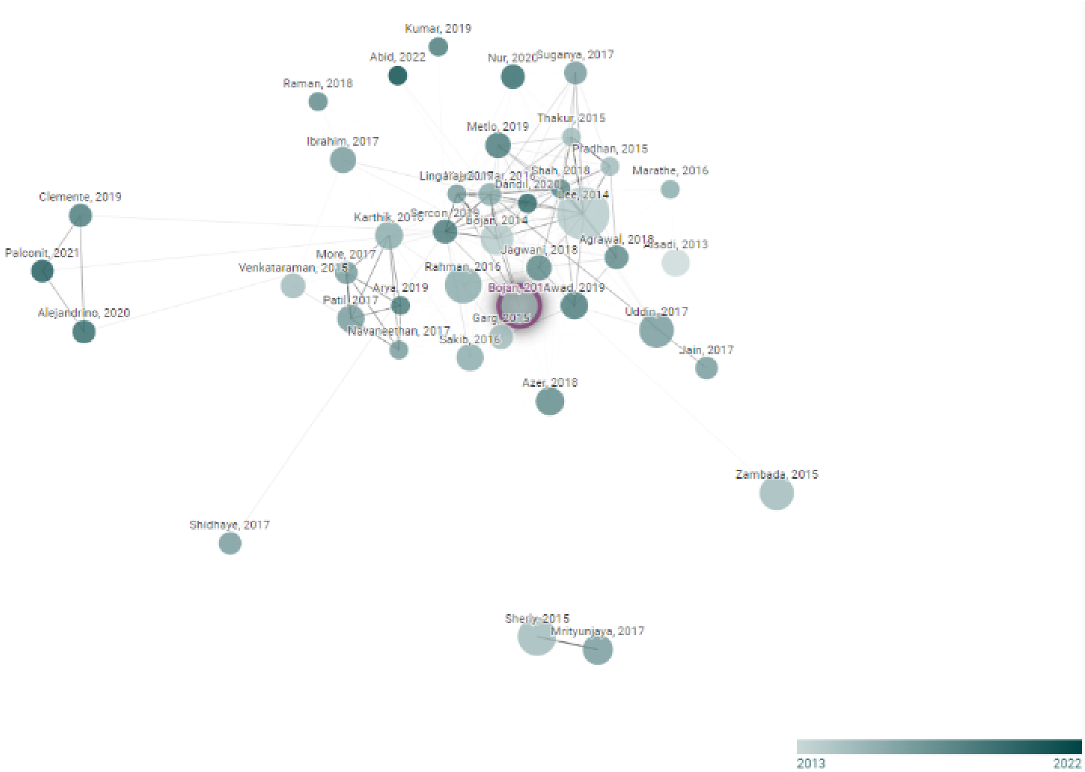

Figure 2.3: Exemplary output of the article identification platform "Connected Papers" for the article from Bojan et al. (2014) [14].

2.1.1. Search strings

The Tables 2.1 and 2.2 summarise the exact search efforts by mentioning the used database for the search (first column), the exact search string in the database (second column), the results of the specific search string (third column), and the fourth column visualises how many articles were selected. The next section will elaborate on the scoping and selection procedure of the articles.

Table 2.1: Overview of the used databases, research strings and results per search for articles with a focus on emission factors and technological improvements.

Database	**Search string / article**	**Results**	**Used** , (exclusion criteria)
Scopus	Emissions AND factors AND technological AND improvements AND vehicles	57	8 (Scope, topic and publication date)
Connected Papers	Real-world emissions performance of heavy-duty Euro VI diesel vehicles	40	5 (Scope, topic and publication date)

Table 2.2: Overview of the used databases, research strings and results per search for articles with a focus on IoT and emission factors.

Database	Search string / article	Results	Used, (exclusion criteria)
Scopus	Emissions AND vehicles AND technology AND improvements AND logistics	35	2 (Scope, topic, year, conference papers)
Scopus	Emission AND factors AND effort	54	1 (Scope, topic and publication date), many medicine-r elated articles
Scopus	emissions AND factors AND iot AND logistics	7	1 (Scope, topic and publication date)
Scopus	Accuracy AND emissions AND factors AND vehicles AND logistics	9	2 (Scope, topic and publication date)
Scopus	iot AND emissions AND logistics AND road	4	1 (Scope, topic, and publication date)
Google Scholar	IoT in logistics	18.500	1 (Scope, topic and publication date)

2.1.2. Scoping

Scoping is an important process of article selection and was done based on three criteria:

1. Topic/Scope of the paper
2. Publication date of the paper
3. Type of article

Thus, articles that fulfilled all three criteria would be used for the literature review.

Topic/Scope: One of the major reasons for articles to be excluded from the literature review was the content fit. For example, at the time of the literature review, one would find 35 results when using the search string *"Emissions AND vehicles AND technology AND improvements AND logistics"* on Scopus. However, not all articles apply to the topic at hand. Articles that covered those search terms but focused on different topics, such as electric vehicles and their emission factors, road energy analysis, or predictive fuel consumption models, were not included. Only articles that directly discussed and added to the topic of road vehicles emission factors, technological improvements and the issue of updating emission factors or onboard sensors for emission factors were used. For the search on Google Scholar, only articles directly connected to onboard sensors, emissions estimations and the logistical sector were used to cope with the large number of available papers[1].

Publication date of the article: All articles that were selected should not be older than five years to represent the state-of-the-art knowledge on that topic, thus articles older than that were excluded unless they had a highly impactful contribution to the topic and the mentioned content was not updated in other articles yet. Examples of that are the articles that covered the differences between real-life and laboratory emissions of vehicles. Their publication date ranged from 2017 to 2019, however, the knowledge was still correct today and underlined the issue of laboratory-measured emission factors of vehicles, which has been known longer.

Type of article: The literature review used independent articles, not conference or symposium papers.

The final selection of articles can be found in Table 2.3 and in Table 2.4. The Tables E.1, E.2, and E.3 showcase which articles were retrieved from which search string.

[1]The scope of the analysis of the articles was only the first page of article results, as Google Scholar portraits the most suited articles at the top of the page.

2.2. Definition of core concepts

Emission factors (EFs) are vehicle-dependent constants that aim to predict the number of emissions emitted per driven distance, speed, or fuel type [58]. They depend on many factors that impact driving and the resulting emissions, such as the engine's temperature, the mileage, the slope of the road, or the driving behaviour [58]. The quality and accuracy of EFs are, therefore, heavily dependent on how they were derived and where and under which conditions they are being used. The EFs constants for the various vehicle types and configurations are commonly derived based on experimental data in either controlled laboratory conditions or in real-world conditions [58]. Fig 2.4 visualises the potential problem with EFs and their high dependence on driving conditions and vehicle type. It summarises the emissions of 41 different driving cycles and sub-cycles of 15 different Diesel Euro 4 with a similar engine capacity for passenger cars[2]. Even though all vehicles have similar engines, their NO_x emission heavily fluctuates, especially for extremely low and high values.

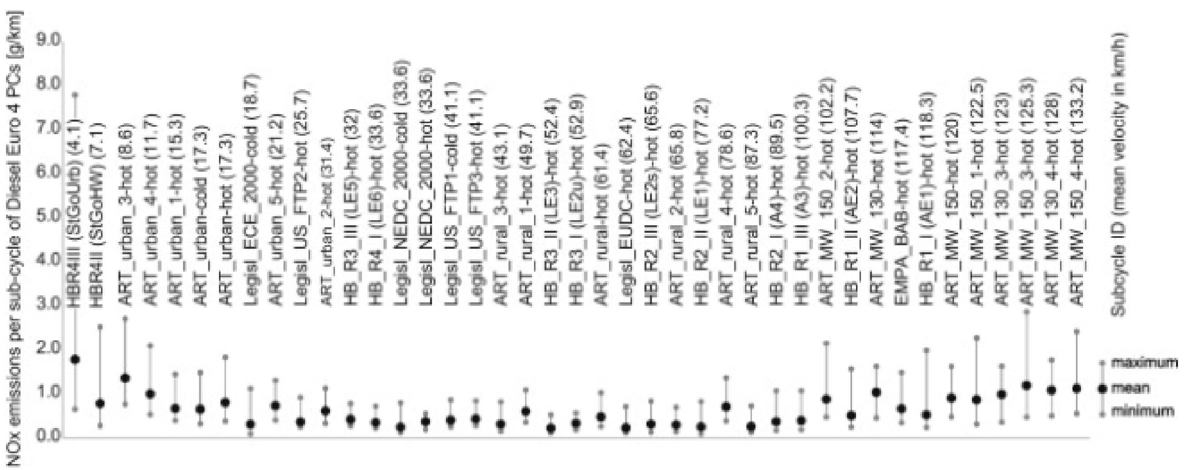

Figure 2.4: Overview of the different emissions test results of 15 different Diesel Euro 4 passenger cars. Retrieved from [58, p. 87].

The **Internet of Things (IoT)** is a concept that describes an interconnected logistics environment, where infrastructure, vehicles or physical objects, such as containers, are equipped with a sensor [121]. The sensors can collect and send real-time data, parameters and information to a central server. It can collect information about driving behaviour, speeds, duration of travel, or fuel consumption. IoT is the extension of telemetries data of vehicles, as it offers the possibility to share and **monitor it in real time**[3]. The collected and generated data can be used to create and calculate the actual emissions of vehicles. Thus, in a world where every car has onboard sensors, the general emission factors could become obsolete, as everyone calculates the emissions of their vehicles based on actual data rather than standard factors. IoT is part of the bigger category of **onboard sensor systems** and will be used interchangeably in this research. Onboard sensor systems can gather the actual fuel consumption data of the vehicle, and thus it is possible to calculate the vehicle's emissions accurately. Not all onboard sensory systems have the option of real-time data monitoring. The search strings and literature review will focus more on IoT systems, as these are currently a central research topic in the field and also mention and cover onboard sensors in their findings. However, the research aims at gaining knowledge on how primary data can improve the accuracy of emission factors for road logistic vehicles. Thus, IoT and onboard systems are applicable and can gather primary data on road logistics vehicles.

Transport is defined by the ISO 14083:2023 standard as:

> „movement of passengers and/or freight from one location to another performed by nodes of transport such as air, cable car, inland waterway, pipeline, road or sea“ [75, p. 4].

[2]Passenger cars have the same issues as HDV, as the driving behaviour of the driver impacts the total emissions of the vehicles, which are difficult to capture with standard emission factors accurately.

[3]Real-time monitoring depends on the definition. For some, it is instantaneous monitoring with latencies of milliseconds. For others, a delay of five to 30 minutes is still considered "real-time"

With the more strict emissions reporting rules, also activities that are not directly related to the movement of passengers or freight itself (thus supporting activities), like emissions caused to receive the fuels or energy costs of logistical sites & hubs will be incorporated within this sector.

2.3. Literature Results

2.3.1. Challenges of traditional emission factors

Emission factors (EFs) are one of the major techniques and constants to estimate and calculate emissions of both private and commercial road vehicles. They have been used for many years and have thus far done a sufficient job to estimate and calculate emissions caused by road vehicles. The various vehicle engines, power, age, and mileage yielded many vehicle-specific EFs constants. The magnitude of variables that impact the estimation of emission factors adds another layer of complexity when determining vehicle emission. Most organisations did not need 100% accurate emissions of their vehicles, thus, this limitation was not perceived as a significant disadvantage of using emission factors. However, the shift towards more regulations regarding the reporting and collection of accurate road logistics GHG emissions may push many logistical operators to gather more accurate emissions of their vehicles (a more in-depth regulatory analysis will be conducted in chapter 4). Soon not just logistical operators will need to report their emissions on a detailed level to comply with the mentioned regulation, such as the Corporate Sustainability Reporting Directive (CSRD), or to avoid high carbon princes via the carbon certificate trading program. Thus clearly indicating a need for identifying other options to monitor the emissions of vehicles [51, 119]. New advances in technology, such as the Internet of Things (IoT) and smaller and more precise onboard sensors, now open the door for other methods to calculate the emissions of vehicles more accurately. However, it is unclear whether introducing industry-wide applications of these sensors will benefit anyone and what requirements such a system must fulfil to be used. Before diving into the opportunities that the new technologies offer, it is important first to understand why current EFs are not able to give 100% accurate estimates and how the different vehicle-dependent constants influence this. All discussed articles and their main conclusions and implications can be found in Table 2.3.

Research of Ibañez-Acevedo et al. (2022) established that especially the motor performance and the fuel consumption of vehicles are the most influential factors when trying to estimate the emissions of vehicles [71]. However, the engine performance is highly dependent on the age and maintenance of the vehicle, which constantly changes over time [71]. The second most influential factor (after fuel consumption) also depends on other factors [71]. Posada-Henao et al. (2023) established that fuel consumption is, opposed to the dominant view, more dependent on the vehicle's weight and the road's slope than the vehicle's speed [105]. The road's slope and the vehicle's weight are two highly fluctuating constants, which are difficult to generalise. Thus, they concluded that the current models[4] are not sufficient to accurately calculate the emissions of vehicles, as they take the vehicle's speed as the dominant constant [105]. Krecl et al. (2017) also established this conclusion, identified in a case study in Sweden, that leading models for estimating vehicle emissions are unsuitable and should be updated to more sophisticated and complex models. Models that use speed as the main factor are still not capturing the complexity, as the EFs of vehicles differ for different velocities [84]. Grigoratos et al. (2019) discovered that each velocity segment of vehicles has significantly different EFs, which are difficult to capture by using a vehicle's average speed over time [64]. This leads to another issue of EFs, which is that real-world emissions of vehicles are up to 10% higher compared to the emissions measured in laboratories [64, 84]. Other studies and research from Mendoza-Villafuerte et al. (2017), Tan et al. (2019), Suarez-Bertoa & Astorga, Dixit et al. (2017), and Quiros et al. (2018) came to similar conclusions when comparing actual emissions of vehicles under real-world conditions with the estimates based on EFs for both heavy-duty logistical vehicles and passenger vehicles [90, 117, 116, 28, 106]. This discrepancy between real-world data and the estimates is also partly driven by the different factors influencing the emissions and the corresponding effort it would need to determine them for every vehicle.

[4] They used the HDM-4 model (highway development and management model) as an example

2.3.2. Emission factors and technological improvements

Another issue is the general technological improvements of vehicles regarding aerodynamics, engine efficiency, weight, fuel consumption, and road surfaces. All these factors play a significant role in vehicles' EFs, and capturing them accurately means that the EFs have to be updated frequently. For example, Martin, Bishop & Boies (2017) established that annual technological improvements decrease fuel consumption rates of light-duty vehicles (LDVs) by 4.3% on average, thus underlining the need to update EFs frequently [87]. Roeth (2020) added that the transportation sector is at the beginning of a drastic technological push facilitated by shorter development cycles in science and more disruptive technologies which will be deployed [109]. Thus, again highlighting that technological advancements increase the need for more frequent EFs updates. However, as Mosquim & Mady (2022) and Höltl et al. (2017) have pointed out, these technological improvements are not global, and to get an accurate picture of emissions, policymakers and companies would need the technological improvement factors at a country level rather than on a global scale (this is especially true for areas with less intense collaboration between countries like in Latin America or Africa, in comparison to the EU) [99, 68]. This would mean that the European Union or the automotive industry would need to establish a national technological improvement factor that could be multiplied with the current EFs of vehicles to reflect the increased efficiencies of vehicles and, thus, a lower emission rate. Based on Mosquim & Mady (2022) and Höltl et al. (2017), it is not sufficient to take any improvement factors from a different region, such as the USA or China, as the European factor could deviate significantly.

Thus, researchers seem to agree that EFs are rarely 100% accurate and do not represent real-life emissions of logistical road vehicles due to their inability to capture vehicle and road-specific constants. Additionally, current models are not complex and sophisticated enough to accurately predict emissions. However, due to their simplicity, these emission factors and models are still widely used for global and European policy decisions and for determining the emissions of both logistical and passenger vehicles. It must be said that the discrepancies between EF-based estimations and real-world data can be minimal. Not every organisation (currently) needs their emissions on such a detailed level that they want or have to abandon the traditional way of estimating emissions with EFs[5]. However, Höltl, Macharis & De Brucker (2017) discussed how current models and EFs are used for policies and white papers and as a foundation for policy decisions [68]. This can be problematic, as many have outlined how these models consistently underestimate vehicle emissions. The results of the discussed papers can be found in Table 2.3 summarises the findings per article.

Table 2.3: Overview of the results and contributions of the selected papers about EFs and technological improvements.

Author/ Year	Topic	Location	Issue with EFs	Additional	Method
Ibañez-Acevedo et al. (2022), [71]	Emission mitigating strategies via EFs (passenger).	Mexico	Motor performance & Fuel consumption is the most influential for vehicle estimations.	Motor performance is based on the age of the engine, hard to constantly estimate over time.	Dynamic modelling via Stella 9.0.3.
Mosquim & Mady (2022), [99]	Technological improvement in passenger fleets and their impacts on the fuel efficiency.	Brazil	Technological improvements are not always used for improving the emissions of vehicles. Increase in weight or speed of the vehicle.	Technological improvements are region/ country dependent. And often vary from the estimates.	Knittel and Mackenzy's method for estimating efficiency tradeoffs and technological trade-offs.

[5] Note that an industry-wide application of mandatory GHG emissions reporting for all companies would also force these companies to move beyond the traditional EFs.

Table 2.3 continued from previous page

Author/ Year	Topic	Location	Issue with EFs	Additional	Method
Posada-Henao et al. (2023), [105]	Real-life emission data of vehicles is more accurate and road factors play a more important role than assumed.	Colombia	Fuel consumption is more dependent on the weight and the road slope than on the speed of the vehicle.	HDM-4 model is not suitable to estimate fuel consumption data, must be updated at least every 10 years. Slope and weight are more important.	Linear & Non-Linear regression models. Minitab.
Roeth (2020), [109]	Technological advancements in transportation.	Global	The transportation the economic and regulatory factors at the beginning of drastic technological advancements push sector.	The increased development cycle and disruptive technologies might create the need for even more frequent updates of EF for vehicles.	Literature Review.
Grigoratos et al. (2019), [64]	Difference between real-world and laboratory emissions of HDV.	EU	Even though EURO 5 had better values than the previous models, there were still significant differences between the different types of HDV. Varying levels of emissions especially low speed is problematic. Existing EF misscalculated the CO2 emissions by 10%.	Every velocity segment has different emissions, thus very hard to estimate the emissions based on average speeds.	A 1-year experiment of real-road conditions with on-board sensors.
Höltl et al. (2017), [68]	How are EFs used for estimations and how accurate do you need them (HDV)?	EU	EF are often the basis for policy papers and scenario analysis, however, their accuracy is not enough. Would need detail on a country level.	Problematic to base the entire scenario analysis based on non-ideal EF data, which are then used for policy recommendations.	Backcasting approach.

Table 2.3 continued from previous page

Author/ Year	Topic	Location	Issue with EFs	Additional	Method
Martin et al. (2017), [87]	Technological advancements of LDV and the impact on emissions and fuel consumption.	UK	Annual technological improvements of 4.3% for fuel consumption.	Importance to have EF that represents the status quo of technology.	Data sets of 95% of all vehicle sales in the UK.
Krecl et al. (2017), [84]	Even leading EF models are unable to accurately determine EFs. More elaborate models are needed for that, which takes more effort.	Sweden	More sophisticated methods to determine EFs yield significantly higher EF results than the average laboratory ones.	Major leading models that are used to calculate EFs should be updated to get more accurate results.	Real-world measurement of EF via next -to-road measurements and the tracer method.
Dixit et al. (2017), [28]	Variation between real-world and laboratory emissions of vehicles (passenger and freight trucks)	USA	Stop and go in metropolitan streets heavily increases emissions compared to standard EF	Peoples in metropolitan areas have greater emissions exposure	Real-world measure-ments via Mobile Emissions Laboratory (MEL)
Tan et al. (2019), [117]	Real-world emissions of HDV for NOx	USA	Large difference between measured and certified NOx emissions	Low cost NOX sensors are a valid tool to measure NOx emissions	Selective Catalytic Reduction (SCR) system
Suarez-Bertoa & Astorga (2018), [116]	How temperature of the vehicles affect their emissions	Europe	Emissions of vehicles increased dis-proportioned at cold temperatures (-7 Celsius)	The winter season can increase emissions	In climatic cell to simulate temperature profiles
Mendoza-Villafuerte et al. (2017), [90]	Real-life emissions are higher in real-life scenarios	Europe	Different boundary conditions of the measurements resulted in different emission values	HDV may also create considerable N2O emissions	Measure-ments via Portable Emissions Measurement System (PEMS)
Quiros et al. (2018), [106]	Fuel-based emissions estimations of HDV	USA	Fuel-based emissions calculations result in more accurate and higher emissions of HDV	Measuring sensor are often malfunction and result in wrong results	Diesel particle filters (DPF)

2.3.3. Emission factors and the opportunities of onboard sensor system

Emissions reporting and accurate tracking are essential today and possibly even more than previously due to the many climate change mitigation agreements and mandatory reporting schemes on a European and global scale [37, 35]. Thus, many academics and logistics operators have looked for other possibilities and turned to IoT and sophisticated onboard sensory systems. Onboard sensor systems can, directly and indirectly, improve the sustainability factors of modern supply chains and the corresponding emissions factors. Sensors could bridge the gap to 100% accurately estimate emissions for organisations and studies that need it, where the current methods of EFs fall short of doing so. Hopkins & Hawking (2018), Yu et al. (2022) and Chen et al. (2022) pointed out that IoT and onboard sensors can track and report actual vehicle-specific factors, such as fuel consumption, speed, road slope, engine data, and vehicle-specific emissions [69, 126, 17]. It would remove the main issue of current EFs, as IoT and onboard sensors eliminate the need to estimate EFs based on similar vehicle categories, average velocities, and fuel consumption. Onboard sensor systems would be able to create accurate vehicle-specific emissions. They all utilised the ability to collect first-hand and vehicle-specific data to calculate emissions instead of using uniform emissions factors. Yavari et al. (2022) even created and used an IoT-based platform to monitor and track data of last-mile delivery parcels in real-time to create emissions generated per parcel based on its weight, the velocity of the vehicle, the surrounding temperature, and real-life driving conditions [125].

IoT and onboard sensors can also indirectly improve logistical operations and make them more sustainable. Xue et al. (2022) estimated that new technologies and digitalisation could reduce emissions between 14% and 19% by using onboard data to optimise routing and driving behaviour [124]. Mi et al. (2021) proposed to use Big Data analytics and the newly available freight data as input to create more sophisticated models for freight-specific emission models [91]. Thus, also improving current models that are based on traditional EFs. Hopkins & Hawking (2018) showcased that the combination of heavy-duty logistical truck telematics and truck IoT data reduced emissions in an Australian case study by 42% [69]. More surprisingly was that 32% of these 42% were emission reductions based on the improved driving behaviour of the drivers, which were achieved by analysing the braking, acceleration, velocity, and engine data of the vehicles [69].

However, data management is essential regardless of whether IoT and onboard sensors are being used indirectly or directly. Nearly all authors highlight the importance of data management structures when using IoT and sensor systems in logistics and for emissions tracking. Liljestrand, Christopher & Andersson (2015) pointed out the importance of having access to data in the first place but also highlighted the significance of only using the required ones not to get lost in the large amounts of data generated by onboard sensors and IoT [86]. They continue that IoT and other new technologies only have an added value if used correctly, as poor data management can also reduce the entire system's efficiency [86]. D'Amico et al. (2021) pointed out that data flow management is crucial when designing an IoT and onboard digital infrastructure [23]. Thus, the data flow structures of the generated data at the logistical vehicles must enable a smooth connection that enables the transmission of the data from the vehicles to the final destination of the data. Table 2.4 summarise the articles' findings.

Table 2.4: Overview of the results and contributions of the selected papers about IoT and logistics.

Author/ Year	Topic	Location	Sensors & Logistics	Additional	Methods
Xue et al. (2022), [124]	Decolonisation of the Chinese heavy-duty-trucks.	China	Logistical improvements can only reduce emissions between 14 to 19%.	Digitisation and new technologies are key to reducing emissions.	Fleet-based dynamic model.
Mi et al. (2021), [91]	Digitisation and Big Data applications within freight travel.	China	Current research is focused on mainly passenger travel but lacks to address the characteristics of freight travel.	Opportunities to create optimised freight models based on data and to leverage new technologies.	Literature review.

Table 2.4 continued from previous page

Author/ Year	Topic	Location	Sensors & Logistics	Additional	Methods
Hopkins & Hawking (2018), [69]	How can IoT & Big Data reduce the environmental impacts of logistics?	Australia	Truck telematics and IoT sensors can gather data about speed, distance, braking, location,and engine performance. This data can be used to change drivers driving behaviour and thus reduce emissions. The total emissions cut was 42% of which 32% was due to a change in driving.	IoT can help gather important data to calculate the exact emissions of vehicles, but can also be used to change the driving the behaviour of drivers to further reduce emissions.	Case study under real-world conditions.
Liljestrand et al. (2015), [86]	How can transport portfolio frameworks reduce the environmental impacts of logistics?	Sweden	Importance of data in logistics and how to properly manage this data. Availability, correctness, and strategic usage of data.	Important to not get lost in the amount of generated data.	Literature review, framework creation, tested on empirical data from a case study.
Chen et al. (2022), [17]	Creation of improved fuel consumption models under complex geographical conditions.	China	Data from the Internet of Vehicles (IoV) platform is used to create better fuel consumption models.	Data and IoT can have many direct and indirect emission reduction effects.	Back Propagation fuel model based on Cauchy Multi-Verse optimiser (CMV).
Yu et al. (2022), [126]	Driving behaviour that impacts emissions of light-duty gasoline vehicles.	China	Portable emission measurement systems (PEMS) were used to collect data about driving behaviour.	Saturated traffic drastically changes driving behaviour and increases emissions.	Logistic regression, Naive Bayes, decision tree, support vector machine & BP neural network.
D'Amico et al. (2021), [23]	Data-based approaches have become necessary to realise smart and sustainable logistics.	Global	New technologies such as IoT, sensors, Big Data analytics or AI are needed to enable data-based decisions and sustainable logistics.	Especially data flow management is important when designing such systems.	Literature Review.

Table 2.4 continued from previous page

Author/ Year	Topic	Location	Sensors & Logistics	Additional	Methods
Yavari et al. (2022), [125]	Development of an IoT platform to measure last-mile parcel delivery emissions.	Australia	Real-time analysis of IoT data to obtain emission data on a parcel level, considering, real-life driving conditions, temperature, weight, distance.	The insight on a parcel emissions level can be used to be extrapolated to a freight level.	Case study.

2.3.4. Regulatory standards of emission factors

One of the most influential standards for the freight sector is the GLEC framework, which the Smart Freight Centre developed in collaboration with industry players. The GLEC framework was one of the first and most recognised methodologies to consistently and coherently identify, calculate and report emissions caused by any transport activity [112]. The latest version (the GLEC 2.0 framework) was updated in 2022[6]. The standard addressed the industry's call for coherent and clear instructions for emissions counting beyond Scope 1 emissions. The GLEC Framework also includes a detailed description of the calculations of Scope 2 & 3 emissions that are caused by activities derived from logistical activities. Fig. 2.5 summarises the three scopes of emissions and defines the framework. The GLEC framework outlines best practices for emissions calculations and defines explicit scopes and boundaries that certified companies must comply with. Examples are the exact scopes of the fuel life cycle (well-to-wheel, WTW), best practices for distance calculations of the freight or vehicle-specific load factors and default factors [112]. Even though the methodology is not mandatory, many shippers and carriers seek compliance to prove their efforts to reduce emissions within their transport operations to their end consumers. The European Union has defined the goal to reduce all transport-related emissions by 55% by 2030 and already started policies such as the pricing of carbon emissions [119]. Thus, companies already aligned with the GLEC framework can reduce their logistics emissions, save money in the future, and use it as a competitive advantage.

To give the standard even more acceptance in the industry, the Smart Freight Centre asked for official guidance and regulations and collaborated with the International Organisation for Standardisation (ISO) to develop an industry wide-standard for all emissions accounting in the logistical sector, which was based on the GLEC Framework [75]. The newly released ISO 14083:2023 standard (March 2023) will be the new benchmark for the industry for companies to comply with. It clarifies and gives a more detailed description of real-life emissions tracking and allocation problems. It also introduces two new transport modes compared to the GLEC framework (cable car and pipeline) [75]. It also extends the scope of emissions related to transport activities, such as emissions caused by data centres for the process, maintenance and handling of transport operations [75]. Additionally, it gives a detailed description of what companies have to report when declaring their emissions.

[6] An updated GLEC 3.0 version is expected in September of 2023, however, it was not available when writing this thesis.

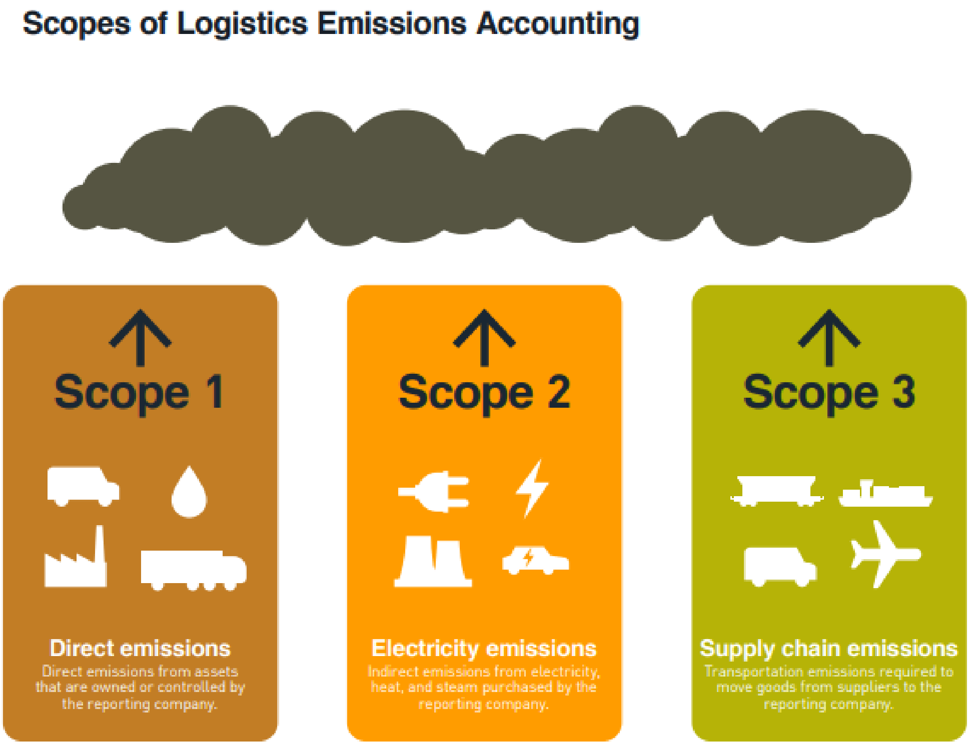

Figure 2.5: Emissions scopes defined by the GLEC framework specifically for the freight transport sector. Retrieved from [112, p. 16].

The GLEC framework and the new ISO 14083:2023 standard are non-binding documents that aim to harmonise and align standards in the industry. The General Data Protection Regulation (GDPR) is a mandatory document that also applies to data sharing in the logistical sector. Thus, any data sharing within the logistical chain must conform with the regulations outlined in the GDPR. This is even more the case for the customer parcel sector but does not exclude the commercial freight and road logistical sector [128], especially when onboard sensors collect the data on the driving behaviour of the drivers of the trucks. Data protection is a joint task from all participants in the supply chain to protect personal data, and if not done correctly, it can result in fines in the millions [27]. For logistics companies, that means that they must be more careful with whom they share their data (i.e. with external suppliers or service providers). They should have detailed documentation about their flow of personal data in the supply chain and account for how the data will be stored, processed and analysed [63]. Lastly, each participant must identify weaknesses in their own data security and take appropriate measures to eliminate or mitigate those [63, 53].

The shift towards a sustainable transport sector also comes with mandatory emissions reporting of large-scale companies and freight cargo shippers. Large non-transport companies[7] need to declare their non-financial impact on the environment and people, thus also their emitted emissions from 2024 onwards, via the Corporate Sustainability Reporting Directive (CSRD) [51]. Reporting of their emissions will also become mandatory for cargo shippers in 2024 [29]. Thus, frameworks such as the GLEC or the ISO14083:2023 will take a vital supporting role in the calculation of the emissions, which need to be reported. The European Union is even proposing that all companies that must report their emissions under these directives must use the new ISO 14083 standard, which highly prioritises primary vehicle data for it and, thus, not the traditional EFs [51].

Next to the regulatory, technical and policy considerations are also industry standards and interoperability considerations that need to be considered when looking at data sharing and sensory systems. A more detailed policy analysis will be conducted in chapter 4.

2.4. Knowledge Gap and Research Question

From the literature overview, it becomes clear that IoT and onboard sensors in road freight vehicles have the possibility, on the one hand, to measure 100% accurate emission on a vehicle level if that is needed and required by an organisation or government. On the other hand, they could collect vehicle-specific constants, which could be used indirectly to improve current EFs-based emission models.

[7] Companies with 500 or more employees

As outlined in the articles, detailed and highly accurate emissions were not required until now, but decarbonisation and detailed emissions tracking on a vehicle-level force fleet operators to explore new technologies.

To drive the required innovation step change, data flow management and the creation of appropriate system architectures are one of the most significant barriers to the industry-wide application of onboard sensory systems on freight vehicles, as identified by multiple articles. Their introduction would come with new challenges for fleet operators and require investment into their digital infrastructure to capture and successfully analyse the newly gathered data. What also became clear from the literature review was that none of the studied articles included any stakeholder or socio-technical considerations for such a system. All articles only considered the technical component and left out the other two important aspects, which is also considered an important finding/ knowledge gap of the literature review. Based on the available literature and the current gaps, this research will focus on the possibilities of onboard sensory systems to increase the accuracy of emissions reporting by investigating the data management structures and architectures to successfully implement sensory systems in logistical vehicles for emissions reporting to gather primary vehicle data.

Additionally, this research will consider the socio-technical, the stakeholder and the technical aspects of the system to derive a meaningful analysis. Based on the identified gaps and shortcomings of the available literature, the following main research question was formulated:

What onboard sensor systems architecture can enable road logistics operators to gather primary data from their fleet to accurately determine their vehicle emissions?

3

Research Approach & Methodology

This chapter will discuss the used research approach and methodology. The research approach will be discussed in section 3.1 followed by the presentation of the derived subquestion in section 3.2. Each subquestion will be examined in terms of the main knowledge and data inputs needed to answer the subquestion and the expected output. Section 3.3 will give an overview of the usage and creation of one of the most important knowledge inputs in the study, the expert interviews, by explaining how the interview partners were selected and accessed, how the interviews were conducted and how their insights were used. Finally, section 3.4 will visualise the usage of the interviews, and the research flow diagram is presented.

3.1. Approach

The main research question needs insights into the three different domains of technology, stakeholder and institutions to properly analyse, understand and develop an artefact to answer the main research question (see Fig. 1.4 for the problem decomposition). After understanding these three topics, developing a potential artefact to address the research question is possible. Thus, the research requires an approach that can capture the initial problem scoping and understanding before designing the artefact. One of the most common approaches for that is the Design Science Research Methodology (DSRM) proposed and developed by Peffers et al. (2007) [104]. The approach includes a framework with six steps that start with problem identification and motivation, followed by the definition of objectives, design and development, demonstration of the solution, evaluation and lastly, the communication with the stakeholders [104]. The different stages of the approach offer the researcher flexibility to enter the research cycle at different stages depending on the specifics of the case. The approach has advantages compared to other dominant research approaches, such as knowledge creation while designing the artefact, the four different entry points to the approach and the detailed six-step guide, which also helps to structure the research [15]. Lastly, the used approach is relevant to address socio-technical aspects of the system's design, as the stakeholders (here, participants within the logistical sector) and their needs and requirements for such a system are a main part of the first design stage.

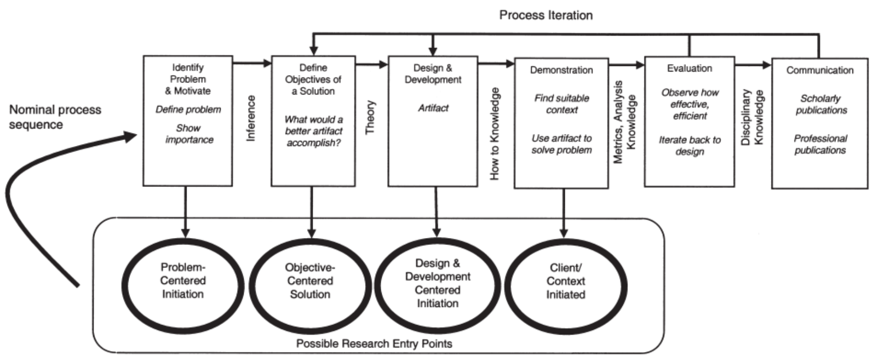

Figure 3.1: Design Science research methodology for information systems. Proposed and developed by Peffers et al. (2007) [104, p. 54 (p. 10 in PDF)].

3.2. Research Questions & Data Inputs

The main research question identified in section 2.4: ***"What onboard sensor systems architecture can enable road logistics operators to gather primary data from their fleet to accurately determine their vehicle emissions?"*** needs appropriate sub-questions that address the multi-domain aspects of the research question and the different DSRM stages. The main research question aims to create a design-focused approach for the onboard sensory systems architecture in road logistics. Four sub-questions were formulated based on the proposed research approach and the main research question.

SQ1 Problem Setting: "What is the social-technical environment that impacts onboard sensor systems for the gathering of primary data?" The first sub-question was formulated to establish an overview of the socio-technical environment in which the system will be implemented. Important sub-questions to this are: *Which stakeholders are involved?, What policy aspects influence the environment?, What are the industry standards?* It will act as a problem-setting and exploration phase to scope and identify the main areas for the research. The sub-question deliverables are twofold. First, a socio-technical analysis of the policies and institutional boundaries was conducted to identify the institutional requirements of a potential system architecture by using and analysing policy documents. Moreover, the socio-technical investigation was completed by assessing industry data sharing standards and analysing existing GHG emissions quantification methodologies and collaboration efforts in the space. Data inputs were official publications, expert interviews and news articles. The second part was the stakeholder analysis. The relevant stakeholders were identified to define their needs for an abstract system architecture. Data inputs were literature, white papers and expert interviews. The expert interviews were a vital part of not just this but also the other sub-questions. Section 3.3 will briefly outline how the interviews were used throughout the research, however, the concrete usage will be explained in the corresponding chapters of this research. The final deliverable was the first set of requirements based on the socio-technical environment, stakeholder needs, and a stakeholder overview. The sub-question was positioned at the first entry point of Peffers et al.'s (2007) design science research methodology [104]. This problem-centred project start guided the research and set the first focus points that needed to be considered throughout the design.

SQ2 Objectives & Requirements: "What are the technical requirements for onboard sensor systems in road freight logistics?" The second sub-question was formulated to establish an overview of the main tasks and the corresponding requirements of onboard sensory systems within the logistical road sector. The previous subquestion focused on the socio-technical and regulatory aspects of the problem. Thus, this subquestion will elaborate on the technical requirements but will also merge all requirements into a final list. It will act as a foundation for the rest of the research and as a knowledge input for the following research sub-questions. An overview of the relevant requirements will be important for creating a proposed system architecture.

The insights and the first list of the socio-technical and stakeholder requirements of SQ1 were the first knowledge inputs, especially to understand the legal and regulatory requirements that operators would have to follow. The initial literature review and expert interviews tackled the research approach's first step (problem identification). The interviews were also used to create the final requirements and design principle list for the second sub-question. The first data input was the available scientific literature and white papers about opportunities and tasks of onboard sensor systems in road freight logistics. White papers had many sources, from the European Union over research institutes to industry players. Scientific literature contributed knowledge on similar topics in different geographic locations and thus increasing the geographical scope. The second type of data input for the second sub-question was primary data that was accessed via expert interviews. The final output of this sub-question was a complete categorised requirements list for the system's design with corresponding design principles.

SQ3 Design & Development: "How does a potential onboard sensor system look like to gather primary data in road logistics?" SQ1 will establish the relevant socio-technical and stakeholder requirements of onboard sensor systems within the logistical sector, and SQ2 will highlight the technical and complete requirements for such a system. Subsequently, SQ3 will focus on the conceptual design of such a system for road logistics by considering the requirements derived from the previous sub-questions. Both outputs from SQ1 and SQ2 were direct inputs needed to complete this sub-question. The result was to gather and organise the insights and learnings of both previous sub-questions and apply them to the exact case of onboard sensor systems while considering the social-technical, technical, and regulatory aspects. The final deliverable was a list of design options that addressed and fulfilled the derived systems' requirements and a stakeholder-specific system architecture that embodied parts of the established design options for a specific case.

SQ4 Evaluation: "What is, according to industry experts, the quality of the designed artefacts and requirements?" This sub-question is aimed at evaluating the design artefacts of the earlier sub-questions with industry experts. Data inputs were the created stakeholder overview, the created system requirements, and the design efforts of SQ3. The inputs were evaluated with the help of industry expert interviews, literature that focused on the same topics, and informal conversation with industry experts. The designs and requirements lists were evaluated based on their completeness and correctness. The concrete interview evaluation structure with the corresponding questions can be found in the Appendix in section C and in the corresponding chapters. The deliverable of this subquestion was to understand the quality of the designed artefacts/requirements lists and improve them based on the insights of the literature and the expert interviews.

A general theme for the data inputs of this research was the expert interviews and the insight knowledge of the employees of shipzero. These expert interviews were conducted to get a general overview and an in-depth understanding via face-to-face (online) interviews.

Lastly, all sub-questions were directly connected to answer the main research question as described by Peffers et al. (2007) [104]. Fig 3.3 summarises the research flow diagram. SQ1 and SQ2 established the problem and objectives of the research and design, SQ3 the design and development phase, which was followed by the evaluation, which was based on the insights of the previous outcomes of the sub-questions.

3.3. Interviews

In order to answer the sub-questions, one knowledge source was expert interviews, and the other was informal conversations and speeches from and with industry experts. For that, six expert interviews were scheduled [1] with various stakeholders in the industry, varying in size and application area, and a total of 15 informal conversations and speeches were conducted or attended[2].

3.3.1. Access & Selection of experts

The interview partners were accessed with two different methods. The first one was through the external network. This included the participating company that contributed with two different interview partners (carrier and LSP) and one evaluation interview partner (External IT company), and one interview partner was accredited through a contact of a university professor (data market provider). The second access method was a first informal conversation at a logistical conference with a stakeholder (emissions factors and modelling expert), who then agreed to participate in the research and provided another expert interview from its network (Fuel consumption, sensor and IT emissions tracking expert). The interview partners were selected based on their industry role, size, their expertise and availability in the time frame for this research. Most of the knowledge gained from the informal talks happened during the logistical conference. Every participant had a connection to the decarbonisation of logistics. Thus, multiple stakeholders were approached, and related speeches were attended. Not all information gathered during these activities was used in the end. The two remaining informal conversations were accessed via personal connections.

3.3.2. Interview medium & Questions

Each interview was conducted online and recorded for later analysis. Each participant was contacted prior via email and received a consent form and a questions document. The consent form indicated the willingness to participate in the research and that the knowledge of the interviews could be used for the study. The provided question catalogue can be accessed in Appendix A.1 and provides an overview of the discussed topics. The interviews were conducted in a semi-structured manner to, on the one hand, provide a certain consistency of discussed topics and, on the other, to also allow for individual comments of the different interview partners. Each interview lasted about one hour. The informal conversations were conducted in person (face-to-face), and each participant received a brief summary of the topic and the research. Conversations lasted between 5 minutes to 15 minutes and did not follow a specific structure.

3.3.3. Processing & Usage

All interviews were recorded, and these recordings were used to create anonymised interview summaries, which highlighted the main insights (see Appendix A.3 til A.7 for the design interviews and Appendix C.2 til C.5 for the evaluation interviews). The summaries were sent again to the interview partners to ensure their consent about the provided statements. Each interview partner gave their consent, and no adaptations had to be made. The interview summaries were an important knowledge source and were referred to as official scientific literature throughout the study and were cited when their insights were used. Thus, each interview summary can also be found in the bibliography of this study with a link to the anonymised summary of the interview in the Appendix to communicate the findings derived from the interviews transparently. Next to citing them, they were also used for quotes that supported certain statements during the design and evaluation process. Table 7.2 visualises all conducted interviews and highlights their contribution to the study. Table A.1 provides an overview of the experience and position of each interview partner, thus validating their ability to give recommendations for the design and evaluation phase. The informal conversations were summarised after each talk, and their main knowledge contributions were merged in Table B. The knowledge gained from the conversations was not used as an official literature source but rather as a supportive tool via footnotes to specific statements.

[1] Six interviews were scheduled in total, of which two were used solely for the first two sub-questions, three were used for both design and evaluation purposes, and one solely for the evaluation. Table 3.1 summarises the contributions per interview

[2] Both informal conversation and speeches added up to 15. Table B summarises them

Table 3.1 summarises the exact contributions of each interview partner towards the research's design and evaluation phases. The first column mentions the interview partners (the names are consistent with the names of the interview summary sections in Appendix A and the evaluation interview summaries in Appendix C). The second column summarises to which degree the interview contributed to the design of the research. Thus, an interview that mentions *"Stakeholder"* contributed to the knowledge creation of important stakeholder considerations. The exact contribution can be found in the interview summary mentioned below the contribution[3]. The comment of *"None"* in the contribution row of the interview partner means that the interview was not used for the design phase of this research (i.e. like the final interview with the External IT company). The third column summarises the contributions of the interviews towards the evaluation of the created designs and requirement lists. It follows the same structure as the design interviews, where a *"None"* indicates that the interview partner was not used for the evaluation of the designs (like the first two interview partners) and if it was used for the evaluation the keyword indicates for which parts. The exact contribution of the evaluation can again be accessed through the link to the summary of the evaluation interview [4].

[3] The link to the interview summaries always start with A.x

[4] The links to the evaluation interviews start with C.x

Table 3.1: Overview of the contribution of the different interview partners for the design phase and the evaluation phase. Each contribution (for the design and evaluation phase) is noted with the corresponding citation (first) and with the summary of the interview (second). Find the description of each interview partner and their professional function in Table A.1.

Interview Partner	Design Interviews (contribution)	Evaluation Interviews (contribution)
Fuel consumption, sensor and IT emissions tracking expert	Stakeholder, Requirements, Institutional environment, Customer Onboarding, Data pulling, General Improvements, [11], A.3	None
Emissions factor and modelling expert	Stakeholder, Requirements, Institutional Documents, Limitations, General Improvements, [10], A.4	None
Data marketplace and IoT aggregator expert	Requirements, Standards, Data pulling, Stakeholder, Business relations, Policies, General comments, [7], A.5	System architecture, General comments, [6], C.2
Logistical actor (carrier)	Challenges, (Requirements), Current Practices General comments, [8], A.6	Stakeholder, Customer onboarding, Requirements, System Architecture, [3], C.3
Logistical actor (LSP)	Current practices, General comments, [9], A.7	Stakeholder, Requirements, Customer Onboarding, Business relations, System Architecture, General comments, [5], C.4
External IT company	None	Stakeholders, Requirements, Design Options, Multi-tenant architecture, Customer Onboarding, Business relations, Data Pulling, Architecture, [4], C.5

3.4. Design sequence & Research Flow Diagram

Fig. 7.6 visualises the design and evaluation process for the entire research. The exact interview composition and usage will be discussed in the corresponding chapters, however, this already presents how and where each interview was used to either create knowledge or evaluate outputs. The rectangles with round edges indicate analysis or design activities, thus knowledge creation via the interviews. If multiple designs were created, the number in brackets indicates the amount of created concepts or designs. The hexagons indicate an evaluation activity. The arrows indicate the knowledge flow with the possibility of external knowledge input via scientific articles or expert interviews. The red arrows indicate the implementation of the feedback from the evaluation interviews into the designs.

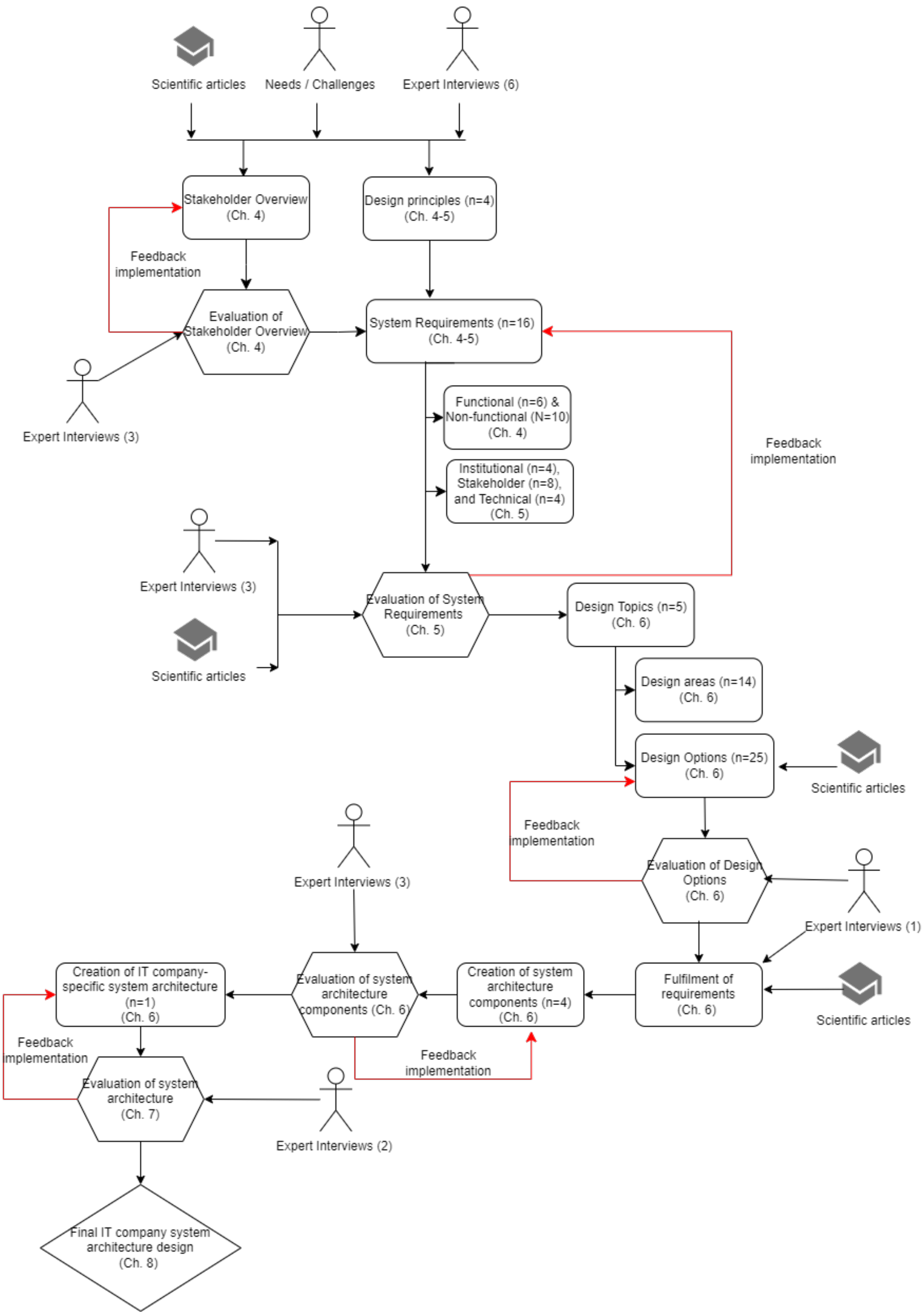

Figure 3.2: Sequence of the design and evaluation process with each intermediate product

The Research Flow Diagram summarises the knowledge flow from the four sub-questions by highlighting each subquestion's in and outputs[5].

[5]Note that only the green arrow and the knowledge flow directly contribute to answering the main subquestion and that the others are indirectly contributing to it.

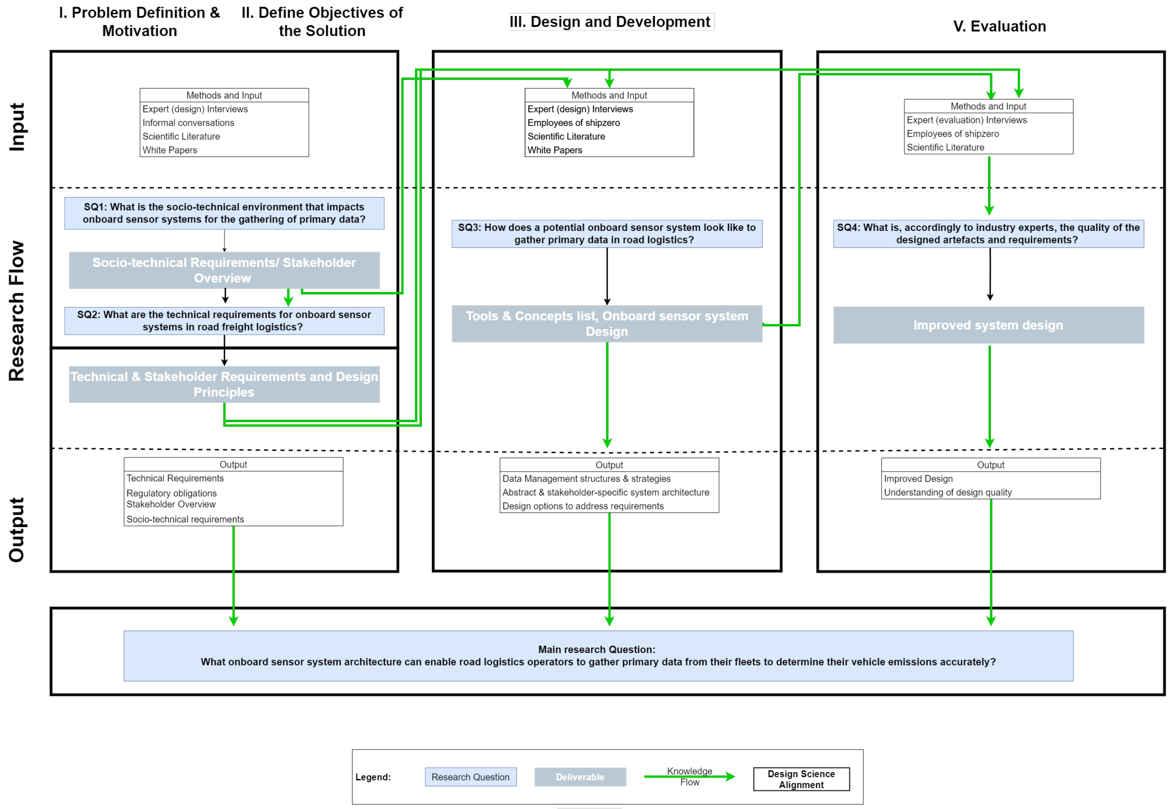

Figure 3.3: Research Flow Diagram for the proposed Master thesis with data and knowledge inputs and outputs and corresponding deliverable.

4
Socio-Technical System Analysis

This section will elaborate on the first sub-question *"What is the socio-technical environment that impacts onboard sensor systems for gathering primary data?"* by first elaborating on how the expert interviews were used for the analysis in section 4.1. It is followed by a broad overview of the system in section 4.2.1. The general overview will include an analysis of the institutional environment and policy documents that might affect the system in section 4.2.2. Section 4.2.3 elaborates on existing GHG emissions quantification methodologies and standards, especially for the data collection and sharing standards at the vehicle level and relevant collaboration efforts. The main influences of the policy and methodology domain will be translated into design principles and system requirements. Section 4.3 presents a stakeholder analysis and their relations. Their needs were also formulated into requirements. Throughout the chapter, insights from the interviews were already implemented and were also included in the final representation of the results for this sub-question. Section 4.4 elaborates on the gathered requirements and concludes the chapter by providing an overview of all derived requirements and the design principles. Section 4.5 addresses the fourth sub-question *"What is, according to industry experts, the quality of the designed artefacts and requirements?"* by presenting the conclusions of the expert evaluation process of the stakeholder overview (the requirements validation will be conducted in the next chapter). The section will conclude with changes to the overview based on the evaluation process and a general conclusion in section 4.6.

4.1. Interview Usage

The earlier mentioned experts' interviews were a vital knowledge input and source for the socio-technical analysis, the final stakeholder overview, and the first set of derived system requirements. Section 3.3 already elaborated on the access and selection of the interviews (see section 3.3.1), the types of questions (see section 3.3.2), and the processing and usage of the gained insight from the interviews (see section 3.3.3). This section will elaborate on how and which interviews were used for socio-technical analysis by highlighting their contribution to the creation of requirements and design principles as well as their contribution to the stakeholder overview in section 4.1.1.

4.1.1. Creation of Design Principles, Requirements & Stakeholder Overview

Design principles are building blocks in system design tasks that should be addressed and thus guide the design process. They formulate general design objectives, which the requirements should address (a more elaborate explanation of the design principles will be given in chapter 5). They are based on and developed around challenges or needs that the design must fulfil and identified via the interviews and the literature, and the analysis of the socio-technical environment. System requirements define what the system should meet.

Interviews were used to establish parts of the design principles and the requirements. For that, three expert interviews were conducted with a *Fuel consumption, sensor and IT emissions tracking expert* (see interview summary in A.3), with a *Emission factor and modelling expert* (see interview summary in A.4), and a *Data marketplace and IoT aggregator expert* (see interview summary in A.5). The interview partners received a questions list in advance of the interview (see the interview structure in Appendix A.1) to familiarise them with the questions and topics. The insights from the interviews were used throughout this and the coming chapter, and all insights and statements gathered from the interviews will be indicated. The interviews were vital to getting an industry perspective of the current limitations and challenges, as the stakeholder perspective was often incomplete or completely missing in the available literature. Also, the first design output of the research, the stakeholder overview, was partly established on the insights of the same three interviews. Here the interviews provided essential insights into the relations between the various stakeholders and also which stakeholder should be included in the overview in the first place.

4.2. Impacts & Environment

The environment of onboard sensory systems within the road logistics sector to gather primary data on the vehicle and its emissions is a multi-stakeholder/domain problem influenced by many. The magnitude of involved logistical stakeholders, infrastructure stakeholders (sensor systems creators), transport and sustainability policymaker, audit and verification companies, external modelling and calculations organisations, and external and independent NGOs within the transportation sector makes up most of the involved parties. Thus, many different actors are involved (see the definitions used and the overview of the mentioned stakeholders in Table 4.3 til Table 4.5). The logistical sector is highly fragmented, especially from the carriers, with thousands of small and larger businesses, due to the low entry and exit barriers into the sector [88].

4.2.1. Overview of the field

At the top of the regulatory environment that indirectly influences onboard sensors are the global and European desires and commitments to decrease CO_2 emissions on a general level and on a transport level. The European Green Deal set a clear target to reduce European net carbon emissions by 55% by 2030 [35] and introduced a carbon tax and, with it, a carbon certificate trading program per tonne of emitted CO_2 [119]. The price is marked-driven, thus changing dynamically dependent on the supply and demand for these certificates. The increased awareness and urgency let the price increase from 14 €/tonne in May 2018 to 90 €/tonne in May 2023 [119]. Another influencing factor driving the increased awareness and changes towards sustainable logistics, and with it, the need for more accurate emissions tracking, are consumers' expectations [122]. Consumers today demand more eco-friendly and sustainable products in nearly all sectors, from clothing to agriculture to everyday products. One significant aspect of that is the logistical services in the background, and many contracting companies (shippers) demand more details of the caused emissions due to their transportation and reductions from the carrier site[1]. Thus, the players in the industry are highly influenced by the desire and need to decarbonise all sectors in general, especially the transportation sector.

Higher customer expectations, more transparency within the logistical processes and the need for accurate carbon tracking all have a similar challenge. The current digitalisation status of the logistical sector is low, and the digitalisation opportunities that have already happened in other sectors went for a long time unnoticed. For a long time, the industry was still dependent on many manual processes, old and inefficient legacy-IT systems and many error-prone practices [66]. This led to limited innovation in the sector, and no authentic digital DNA was established. Transportation companies identified this as the biggest challenge in 2016 [118]. However, this lack of used potential has long been seen and has created a wave of digital, innovative and disruptive logistics start-ups to enter the market. In 2017 a new logistics start-up was founded every five days, and venture capitalists investments in the sector more than doubled from 2015 to 2016 [22, 66]. The Covid-19 pandemic made the need for more digitalised transport chains even more visible, and investments increased [19]. What becomes more evident is that especially the start-ups that can automate operations are the ones that still receive the most funding [19]. Thus, the low-digitalisation status of the transport industry faced several environmental challenges with the need for more transparency and more sustainable obligations. Digitalisation and automation are one of the most prominent solutions to those challenges, which were again facilitated by the increased capital investments and policy regulations, which will be discussed in more detail in the next section.

4.2.2. Policy regulations & Institutional documents

The transport sector is heavily embedded within European legislation due to its wide impact on daily life and the potential to harm people and the environment (all legislations that were relevant for the analysis can be found in Table 4.1). The main harm of the transportation sector is due to sound, air, and light pollution and the destruction of the environment for the infrastructure and indirectly by increasing carbon emissions in the atmosphere. Thus, the European Union has been trying to limit especially the sound and air pollutants of vehicles by introducing regulatory limits and standards, such as the Emission standards Euro 1 to Euro 6 starting in 1992, to reduce harmful effects of the needed transportation sector [97].

European regulations can have different angles and ambitions when addressing vehicle-based emissions. Classical regulations limit and reduce the emitted emission of vehicles, such as the Commission Regulation (EU) 2018/1832, which introduced the new Euro VI emissions standards for pollutants[2]. However, it also specifically introduced new requirements and guidelines for emissions measurement via onboard sensor systems by introducing the "Worldwide Harmonized Light Vehicle Test Procedure" (WLTP) [45]. The regulation outlines and specifies the entire boundary conditions for the testing and generation of the emissions factor by defining an exact cycle distance (23.25 km), the maximum speed (131 km/h), the average speed (46.5 km/h), the driving phases (52% urban roads and 48% nonurban-roads), the different gear shifts points or the test temperatures [45]. Thus, if an entity wants to measure and report vehicle emissions factors and values officially, it must comply with the Commis-

[1] see keynote speech at logistical converse I.3 in Appendix B

[2] such as the Euro 6d-TEMP for stricter NOx emissions for diesel-based vehicles or the Euro 6d for stricter restriction for particulate matter for diesel and gasoline-based vehicles

sion Regulation (EU) 2018/1832. These emission factors are also used to control if vehicle manufacturers comply with the emissions limits. Thus, the Commission Regulation (EU) 2018/1832 was one of many policy documents that established the system requirement of *"Conformity with European emission factor measurement legislation's"*. This was also confirmed by the emissions factor and modelling expert, who mentioned the importance of conformity with these requirements, especially with the testing boundary conditions of the exact cycle distance etc. (see expert interview summary in Appendix A.4) [10].

The Commission Regulation (EU) 2018/1832 was established to expand Directive 2007/46/EC of the European Parliament and of the Council and the Commission Regulations (EC) No 692/2008 and (EU) 2017/1151 by introducing the WLTP-principle mentioned above and the concept of Real Driving Emissions (RDE) for measuring pollutants for commercial vehicles and by aligning the technical requirements for measuring emissions and fuel consumption with the state-of-the-art industry standards [52, 41, 42].

Another highly influential regulation is the Commission Regulation (EU) 2019/1939, which firstly outlines new standards for Portable emissions measurement systems (PEMS) that are used and equipped into vehicles to measure their real emissions output under real-driving conditions and secondly mandates Member States of European Union to create and maintain national emissions factor databases for the monitoring of emissions [46, 10]. The regulation expanded the Commission Regulation (EU) No 582/2011 by stating that the WLTP concept should also be used for the determination of emission factors of HDV and not just for light-vehicle, thus aligning the methodologies of emissions factor determination and using state-of-the-art techniques to do so [50]. This regulation established first the general design principle of *"Conforming with European Legislations and Methodologies "*, as this need is valid for the entire system. Secondly, the regulation also created the system requirement of *„Conformity with European vehicle reporting and monitoring standards"* as it demands Member States to have national emission databases for vehicles and to monitor them.

The Commission Regulation (EU) 2021/392 introduced the concept of Real-Driving Emissions (RDE) for HDV, rather than testing them under laboratory conditions to ensure more accurate and realistic measurement of emissions [47]. Moreover, it expanded the requirements of the Onboard diagnostics systems (ODS), which should start to collect real-world data of HDV about emissions and fuel consumption starting from the 20th of May 2023, as part of the efforts to establish national emission databases. ODSs are not just used for collecting emissions and fuel data but were also designed to detect and report the malfunction of vehicles, which would lead to greater emissions [47]. The regulation was created to extend and update the Commission Regulation (EU) No 1014/2010, (EU) No 293/2012, (EU) 2017/1152 and (EU) 2017/1153 [48, 49, 43, 44]. Commissions Regulation (EU) No 1014/2010 was one of the first regulations that established a framework and format for the reporting and verification of CO_2 emissions in the transport sector [48], thus was updated by the addition of the RDE and OBS methodologies by the new regulation (EU) 2021/392. Commissions Regulation (EU) No 293/2012 made a strong contribution to emissions monitoring and verification by establishing aligned, working and harmonised OBS systems and introducing standardised tests and reporting mechanisms [49]. This allowed for comparisons between the different vehicle classes. Commissions Regulation (EU) No 2017/1152 and Commissions Regulation (EU) No 2017/1153 supported a robust and realistic emissions reporting scheme. Commissions Regulation (EU) No 2017/1152 added new test procedures and conformity factors for the RDE standard to increase the tests' robustness and account for uncertainties and noise in the measurements [43]. Moreover, it added minimum vehicle documentation and updated requirements of PEMS [43]. Commissions Regulation (EU) No 2017/1153 focused on aligning fuel consumption and emissions reporting and calculations to increase reliability and robustness [44].

Regulation (EU) 2018/956 is solely dedicated to the requirements and monitoring of CO_2 emissions and fuel consumption of vehicles in the categories M1, M2, N1, and N2[3], while the regulations and directives above had CO_2 monitoring as their main contribution [56]. This regulation requires Member States to collect and report 78 different parameters and data points of newly registered HDV in the EU (see Fig. 4.1 for some of the required parameters) [56]. This is mandatory for all HDV that are being registered after 2019. The Member State will collect and store the information and will then forward it to the Commission from the 28th of February of each year [56]. The Commission will keep a register of all registered HDV and will make this data openly available[4] and will also report an annual report indicating minimum and maximum performance and average fuel consumption [56]. Each Member State and manufacturer is responsible for the data quality, and the EU can also impose their own verification procedures for the provided data. If the data is inaccurate or transmitted too late, the EU can impose fines (maximum fine is 30 000 € per HDV) [56]. It will also be possible for the Commission to verify the emission values stated by vehicle manufacturers. Regulation (EU) 2018/956 is one of the most influential directives and created three implications for the system. First, it strengthened the earlier-mentioned design principle of *"Conforming with European Legislations and Methodologies"* and the requirement of *„Conformity with European vehicle reporting and monitoring standards"*, while also establishing a new requirement to *„Not track purely driver specific data"*, as it classifies that only vehicle-related parameters should be tracked and reported.

[3]a mass greater than 2610kg, thus Heavy duty vehicles

[4]except for vehicle identification numbers, name and address, and manufacture of the transmission, axel and tires

60	Average speed (km/h)	2.2.1	Vehicle driving performance (for each mission profile/load/fuel combination)
61	Minimum instantaneous speed (km/h)	2.2.2	
62	Maximum instantaneous speed (km/h)	2.2.3	
63	Maximum deceleration (m/s^2)	2.2.4	
64	Maximum acceleration (m/s^2)	2.2.5	
65	Full load percentage on driving time	2.2.6	
66	Total number of gear shifts	2.2.7	
67	Total driven distance (km)	2.2.8	
68	CO_2 emissions (expressed in g/km, g/t-km, g/p-km, g/m^3-km)	2.3.13-2.3.16	CO_2 emissions and fuel consumption (for each mission profile/load/fuel combination)
69	Fuel consumption (expressed in g/km, g/t-km, g/p-km, g/m^3-km, l/100km, l/t-km, l/p-km, l/m^3-km, MJ/km, MJ/t-km, MJ/p-km, MJ/m^3-km)	2.3.1-2.3.12	

Figure 4.1: Snapshot of the parameters that HDV manufacturers and Member States need to collect and report to the European Commission that are relevant for emission factors. Retrieved from the Regulation (EU) 2018/956 [56, pp. 11-12]

The General Data Protection Regulations (GDPR) and the European Green Deal are more general yet highly influential policy documents. The European Green Deal bundles the European ambitions to be the first climate-neutral continent by 2050, the reduction of 55% GHG emissions in 2030 (compared to 1990), and to plant three billion additional trees by 2030 [35]. It provides incentives and communicates European ambitions. Thus entities in the logistical sector have more clarity about what is expected from them. The European Green Deal directly or indirectly influences nearly all parts. The GDPR also indirectly impacts the environment of the transport sector, as all logistical actors have to comply with it. This is especially the case when it comes to collecting, storing and processing data of vehicle drivers that are collected via the onboard sensors. All drivers must be asked for their explicit consent when sensitive data from them is being collected [53]. Thus, any emissions tracking or onboard systems that collect the drivers' data must be aligned with the GDPR [7] (this was also echoed in the interviews; see Interview A.5). The GDPR created the second design principle of "Data Sovereignty " and, which was also captured in a system requirement of *„Conformity with European Data Privacy Legislation's"*. This critical requirement was also confirmed by the data marketplace expert, who clarified that every party that touches the data is responsible for it and must conform with European data legislation (see Interview A.5).

Finally, the Corporate Sustainability Reporting Directive (CSRD) will force large companies with more than 500 employees to report their environmental impact from 2024, including the transport emissions of their logistical operations [51]. Moreover, the European Commission just released a proposal in which they suggest making the new ISO14083 standard the default methodology for all GHG emissions quantification under their new "CountE-missionsEU" project [54]. Meaning that all companies that need to report their emissions under the CSRD must use the new ISO14083 (if the proposal is accepted). This resulted in the system requirement of *"The collected data is compliant with ISO10483 and GLEC 2.0"*.

Table 4.1 summarise the findings of all discussed policy documents and also again highlights which document resulted in the creation of design principles or system requirements.

Table 4.1: Overview of the important policy documents that influence the Socio-technical environment of emission factors and their determination.

Document	**Influence-Level**	**Content**	**Author/ Contributor**
Commission Regulation (EU) 2018/1832, [45]	Direct to light passenger cars and commercial vehicles (HDV). **Created requirement**.	The regulation sets out specific emission standards and limits with a special focus on HDV above 3.5 tons for pollutants such as NO_x, PM and CO_2 depending on the vehicle's weight, type and engine power. It introduced the new Euro VI emission standards and created new regulations in regard to testing and maintenance of the vehicles. Moreover, it introduces the WLTP, which creates strict guidelines for the measurement of emission factors of new vehicles, such as the distance (23.25 km), the maximum speed (131 km/h), the types and the mix of different roads (52% urban, 48% nonurban-roads). This was done to ensure accurate, reliable and harmonized emission factors for all vehicles in the EU.	Regulation of the European Union.
Directive 2007/46/EC, [52]	Indirectly via Commission Regulation (EU) 2018/1832	Framework directive, which harmonised other vehicles directives for the approval and monitoring of engines and their related emissions of NOx, CO_2 and PM. Great focus on emission standards for vehicles and the promotion of alternative fuels. Highlighted the importance of market surveillance and the need for vehicles to comply with legislation throughout their entire lifespan.	European Parliament and Council
Regulation (EC) N0 692/208, [41]	Indirectly via Commission Regulation (EU) 2018/1832	Addresses and limits emission of vehicles, while promoting accurate measurements of fuel consumption. It expands the limits for pollutants by adding values for CO and HC. Moreover, it demands from fuel producers to monitor the emissions caused by creating the fuel and allows authorities to monitor those.	Commission Regulation
Regulation (EU) 2017/1151, [42]	Indirectly via Commission Regulation (EU) 2018/1832	Creation of a system for the reporting of emission factors and includes provisions for data collection and reporting related to vehicle emissions, which authorities can monitor.	Commission Regulation
Commission Regulation (EU) 2019/1939, [46]	Directly **Created Design Principle and requirement**	Is focused on the monitoring and measuring of emissions of vehicles with new CO_2 standards. It requires Member States to create national databases for the monitoring of emissions. It adds sets out standards for PEMS to measure the emissions of vehicles.	Commission Regulation
Commission Regulation (EU) No 582/2011, [50]	Indirectly via Commission Regulation (EU) 2019/1939	Limits of pollutants for HDV based on types, weight, and engine capacity. Members states must establish systems for the monitoring of HDV emissions.	Commission Regulation

Table 4.1: Overview of the important policy documents that influence the Socio-technical environment of emission factors and their determination.

Document	Influence-Level	Content	Author/ Contributor
Commission Regulation (EU) 2021/392, [47]	Directly	Introduces emissions testing of HDV under RDE, rather than in laboratory conditions. Moreover, it introduced new requirements for OBD systems. As these are firstly needed for accurate emissions determination, but they are also designed to detect and report malfunctions in the vehicle to limit additional emissions because of that.	European Union
Regulation (EU) No 1014/2010, [48]	Indirectly via the Commission Regulation (EU) 2021/392	Sets out and established requirements for monitoring, calculation, reporting and verification of GHG emissions in the transport sector. This is especially aimed at the fleet operators.	Commission Regulation
Regulation (EU) No 293/2012, [49]	Indirectly via the Commission Regulation (EU) 2021/392	Sets and established monitoring and reporting standards for CO_2 emissions in the maritime sector with maximum thresholds for CO_2 consumption and fuel consumption.	Commission Regulation
Regulation (EU) 2017/1152, [43]	Indirectly via the Commission Regulation (EU) 2021/392	Mandates HDV manufacturers to monitor and report their CO_2 emissions and fuel consumption of their vehicles and introduces minimum documentation that they need to provide about the vehicles.	Commission Regulation
Regulation (EU) 2017/1153, [44]	Indirectly via the Commission Regulation (EU) 2021/392	Sets out new targets for passenger and commercial vehicles for CO_2 emissions with these limits getting stricter over the years.	Commission Regulation
Regulation (EU) 2018/956, [56]	Direct **Created Design Principle and 2 requirements**	Mandatory reporting of parameters for HDV, such as average speeds, fuel consumption, load values, and maximum acceleration. Possibility for the EU to verify emission standards and impose fines.	Commission Regulation
GDPR, [53]	Indirectly by storing and processing data of drivers and vehicles. **Created Design Principle and requirement**	Bounds fleet operators that store, monitor and process data of their vehicles and thus also from their drivers (via geolocation and driving behaviour data and driers cards) to comply with data regulations.	European Union
European Green Deal, [35]	Directly	Directly drives and impacts the decarbonization ambitions of the European Union within the transport sector, via carbon pricing, emission reduction targets, improved emission reporting, green fuels, improved emission factors and market incentives.	European Union
CSRD, [51]	Directly **Created requirement**	Large companies (more than 500 employees) and certain SMEs are required to report their impact on the environment from 2014. This will be further expanded in the future. The reporting also includes the GHG emissions caused by transport.	Commission Regulation

Table 4.1: Overview of the important policy documents that influence the Socio-technical environment of emission factors and their determination.

Document	Influence-Level	Content	Author/ Contributor
Count-EmissionsEU, [54]	Directly **Created requirement**	Proposal to make the ISO14083 the standard methodology for the reporting and quantification of GHG emissions caused by transport. This would require every company that is required to report their emissions using the ISO methodology.	Commission Regulation

4.2.3. Methodologies, Standards & Collaboration efforts

The following section will discuss the main methodologies, standards and collaboration efforts that are of importance. These will be grouped within the respective subsections. However, nearly all are a combination of at least two aspects. I.e. every methodology is always created with strong collaboration within the sector and can become an industry standard.

Methodologies

The two most dominant emissions quantification, calculation, and reporting methodologies within the European Union are the GLEC framework and the newly released ISO 14083:2023 [112, 75]. Section 2.3 already mentions their impact and implications for the transport sector. The insights and learnings will not be repeated here, but this section will rather focus on the implications of those methodologies for transport stakeholders. Section 4.2.1 already established the desire of end consumers to have sustainable and green products, which also includes the logistical operations. This is especially relevant for shippers as they present the last entity before the end customers. Shippers outsource their logistical operations to carriers, which then need to re-communicate the emissions they produced for transporting the shippers' goods back to the shipper so that the shippers can report those for their scope 3 emissions. Thus, shippers can directly reduce their emissions by selecting greener carriers (see interviews A.3 and A.4) [11, 10]. Carriers, on the other hand, are in the majority of cases unwilling to report their exact and detailed emissions to the shippers, as they could face competitive disadvantages from it (see interview A.3) and use third-party emissions calculations companies that track their total transport emissions proving to shippers their efforts and improvements. Thus, the methodology frameworks from the Smart Freight Centre (GLEC 2.0) and ISO (ISO14083:2023) are tools for both carriers and shippers first to get accredited and thus imply their sustainability ambitions but also show their regulatory compliance (i.e. with the European Green Deal). Also, publicly listed companies must disclose in their Annual Reports Financial and Non-Financial Statements. Both are audited in analogue standards such as EU Taxonomy / IFRS standards. Therefore, a requirement for many logistical stakeholders is their company accreditation with these leading methodologies [8, 9]. This led to the strengthening of already established requirement *„The collected data is compliant with ISO14083 and GLEC 2.0"*, as the EU proposed to make the ISO14083 the default GHG emission methodology for any reporting of transport-related emissions and if shippers want to be compliant with those regulations their collected data must be compliant with the ISO methodology. Third-party IT and emissions calculation companies have the unique role of a (neutral) middle company by guaranteeing the shippers the validity of their emissions caused by the carrier. In contrast, the carriers can show their GHG emissions reduction without giving other carriers and shippers detailed insight into their logistical operations.

Standards

Data sharing and communication standards play an essential role in determining total emissions and emission factors within the road logistical sector, as the pre-installed onboard sensory systems of the vehicle can already give insights into the fuel consumption, weight of the cargo or the travelled distance[5]. In the past, every vehicle manufacturer had different standards and protocols for their vehicles and even sometimes a different standard between different vehicle types of the same manufacturer (see interview summary A.3) [11]. This made it difficult for carriers, shippers and third-party emissions calculation organisations to access the data, as every vehicle manufacturer and, worst case, each vehicle type needed its own software. Thus, more resources and time needed to be invested or fewer vehicle types and manufacturers could be used and accessed for the calculations of emissions [11]. The Fleet Management System (FMS) was an answer to this issue, and eight of the largest European truck manufacturers [6][7] created a uniform standard within their vehicles, which recorded data about

[5] Without the installation of any third-party sensors

[6] Daimler AG, MAN Truck & Bus AG, Scania AB, Volovo Corporation, Renault Trucks, CNH Industrial, DAF Truck N.V., and VDL Bus & Coach B.V.

[7] Daimler AG and Volvo are also represented with their bus divisions, Scania with Scania CV

the fuel consumption, acceleration, idling, weight, or geolocations of the vehicles [65]. Reading this data via the already available onboard systems without additional equipment was possible.

The first FMS standard in 2002 offered shallow data quality in terms of reporting frequency and measurement accuracy (fuel-level reporting at the high and low ends, vibration and noise issues), which led many third-party sensor companies to install their own equipment with higher reporting frequencies and more accurate measurements (here almost real-time monitoring was possible) (see interview summary A.3) [11]. Thus, few companies collected large amounts of vehicle data of various vehicle types over multiple years on a very high granularity with a high frequency and accurate measurements. Today's FMS protocols have improved their accuracy and are sufficient for calculating the total emissions via the total consumed fuel. However, the reporting frequency is still low, with sample sizes of 15 minutes up to the entire transport leg. Companies that have detailed emissions and vehicle parameter databases can use the FMS data as input to then enrich the data with their own to achieve highly accurate data with high granularity and frequency to analyse minute-by-minute fuel consumption and vehicle speed data to derive driving behaviour data [11]. The importance of incorporating multiple data sources created the design principle of *"System Interoperability "*, as this is a key pre-requisite for combining different data sources, but also established the two requirements of *"The system can incorporate other data sources "* and *"The system supports interoperability between platforms "*. Both requirements are directly connected to the importance of data enrichment. This was also communicated and confirmed by the emissions modelling expert (see interview summary A.4), who mentioned that today's fuel emission factors are derived from multiple national databases and by the sensor and IT emissions tracking expert (see interview summary A.3), who uses various databases for the enriching. Moreover, the mentioned standards of ISO and GLEC also note the importance of external emissions data, thus supporting the two requirements.

This data smoothing, analysis and monitoring via the FMS and enriched data was only possible afterwards. Thus vehicle manufacturers developed a remote access FMS version (rFMS), which allowed fleet operators to purchase a subscription to the rFMS software and access the data of their vehicles already during their journey (not on a real-time basis) [108]. Fig. 4.2 is retrieved from the official rFMS documentation, visualises possible request names, the unit they must be reported in, if the information should be reported (thus mandatory) or if it is optional, and lastly, a brief description of the request. The detail and availability of each request are OEM-specific and might vary [108]. Thus, modern-day emission factor determination combines multiple datasets and data origins with sufficient granularity and quantity (a combination of FMS/rFMS data, third-party databases, European emissions databases, and additional sensor input). For that, it is vital to have common standards that increase database communication and alignment [11]. This led to the requirement *"The system can receive, store and process data"*.

Collaboration

Collaboration is highly influential within the logistical sector regarding data sharing, interoperability of data platforms and databases, national and international emission databases and (regional) emission default values. Examples of that are the Handbook of emission factors for road transportation (HBEFA), which started as a national collaboration between the countries of the DACH-region (Germany (D), Austria (A), and Switzerland (CH)) to develop and store emission values and factors for road vehicles. The collaboration expanded to Norway, France and Sweden [73]. Now the HBEFA and INFRAS databases are used for a European partnership with the European Research On Mobile Emission Sources (ERMES), which has compiled a database with national emissions factor values of each European Member State (again highlighting the importance of the design principle and system requirements of interoperability and integrating multiple data sources to calculate GHG emissions) [33]. Another example of collaboration focusing on interoperability between different data platforms and their communication is the FEDeRATED project, co-funded by the European Union [57]. The project aims to create a European standard which allows different databases to communicate almost naturally while adhering to security and data sovereignty principles. Industry experts see this project as highly influential in increasing the interoperability of platforms, which will facilitate data bundling and reduce the magnitude of software tools to access the various databases [11, 10, 7].

A summary of the mentioned methodologies, standards and collaborations is summarized in Table 4.2.

Name	Unit	M/O	Description
durationWheelbaseSpeedOverZero	Seconds	M	The time the vehicle speed has been over zero.
distanceCruiseControlActive	Meter	M	The distance the vehicle has been driven with cruise control active
durationCruiseControlActive	Seconds	M	The time the vehicle has been driven with cruise control active
fuelConsumptionDuringCruiseActive	MilliLitres	M	The fuel consumption the vehicle has consumed while driven with cruise control active
durationWheelbaseSpeedZero	Seconds	M	The time the vehicle speed has been equal to zero. Engine on (RPM>0) and no PTO active
fuelDuringWheelbaseSpeedZero	MilliLitres	M	The fuel consumption the vehicle has consumed while the vehicle speed has been equal to zero. Engine on (RPM>0) and no PTO active
fuelWheelbaseSpeedOverZero	MilliLitres	M	The fuel consumption the vehicle has consumed while the vehicle speed has been over zero. Engine on (RPM>0)
ptoActiveClass		M	2 Classes: wheelbased speed >0, wheelbased speed =0 at least one PTO on during wheelbased speed =0 counter in seconds and millilitres at least one PTO on during wheelbased speed>0 counter in seconds, millilitres (includes consumption of driving) and meter driven
brakePedalCounterSpeedOverZero		M	The total number of times the brake pedal has been used while the vehicle was driving.
distanceBrakePedalActiveSpeedOverZero	Meter	M	The total distance the vehicle has driven where the brake pedal has been used.
accelerationPedalPositionClass		M	In percent. Minimum 5 classes [0, 20[[20, 40[[40, 60[[60, 80[[80, 100]

(a) System responds of the rFMS protocol with the indication of mandatory (M) and optional (O) information. Retrieved from [108]

Name	Unit	M/O	Description
accelerationClass		M	In m/s^2 Minimum 13 classes.], -1.1]]-1.1, -0.9]]-0.9, -0.7]]-0.7, -0.5]]-0.5, -0.3]]-0.3, -0.1]]-0.1, 0.1[[0.1, 0.3[[0.3, 0.5[[0.5, 0.7[[0.7, 0.9[[0.9, 1.1[[1.1, [
retarderTorqueClass		M	In percent. Minimum 5 classes [0, 20[[20, 40[[40, 60[[60, 80[[80, 100]
HighAccelerationInClasses		O	In m/s2 Minimum 11 classes], -3.0]]-3.0, -2.5]]-2.5, -2.0]]-2.0, -1.5]]-1.5, -1.1]]-1.1, 1.1[[1.1, 1.5[[1.5, 2.0[[2.0, 2.5[[2.5, 3.0[[3.0, [
drivingWithoutTorqueClass		M	The total number of seconds/meters while driving without torque. With gear (clutch is engaged)
engineTorqueClass		M	In percent based on EEC1 value (Actual Engine-Percent Torque). Minimum 10 classes [0, 10[[10, 20[[20, 30[[30, 40[[40, 50[[50, 60[[60, 70[[70, 80[[80, 90[[90, 100]
engineTorqueAtCurrentSpeedClass		O	In percent based on EEC2 value (Engine Percent Load At Current Speed). Minimum 10 classes [0, 10[[10, 20[[20, 30[[30, 40[[40, 50[[50, 60[[60, 70[[70, 80[[80, 90[[90, 100]
vehicleSpeedClass		M	In km/h Minimum 40 classes. [0, 4[[4, 8[[8, 12[[12, 16[[16, 20[[20, 24[... [156, 160[[160, [Engine on (RPM>0)
engineSpeedClass		M	In RPM Minimum 10 classes [0, 400[[400, 800[[800, 1200[[1200, 1600[[1600, 2000[[2000, 2400[[2400, 2800[[2800, 3200[[3200, 3600[[3600, [

(b) System responds of the rFMS protocol with the indication of mandatory (M) and optional (O) information. Retrieved from [108]

Figure 4.2: rFMS protocol responds. Retrieved from [108, pp. 18-19]

Table 4.2: Overview of European standards, methodologies and collaborations connected to the measuring, reporting, or transmitting of data of emissions factors.

Standards (S), Methodology (M), Collaboration (C)	Influence-level	Content	Author
GLEC 2.0 framework (M) [112]	Direct **Created requirement**	A methodology/ framework that is designed to calculate and report GHG transport emissions of multi-modal transport chains. Covers excessive lists of default emission factors for various fuel and vehicle types.	Smart Freight Centre
ISO14083:2023 (M) [75]	Direct **Created requirement**	European standard for the quantification and reporting of GHG emissions in the transport sector. Replaced the old European standards EN16258 and is based on the GLEC 2.0 framework. Provides excessive fuel and vehicle emission factors.	ISO
HBEFA (C) [73]	Indirect	Handbook for emission factors for road transport for various vehicle classes. Initially DACH-focused, now as well expanded to Sweden, Norway and France and in collaboration with the JRC.	INFRAS
FMS standards (S), (C) [65]	Indirect **Created design principle**	Fleet Management system standards (FMS), which was developed by eight leading European truck manufacturers. It records and stores truck information about fuel consumption/ level, speed, idling, geolocations, engine hours, vehicle distances and many other parameters in a uniform manner. The information can be accessed by connecting to the system (local) and reading out the data.	ACEA
rFMS standard (S), (C) [108]	Indirect **Created design principle**	Remote Fleet Management System standard, which records and stores the same truck parameters as the FMS but offers the opportunity for remote access to it (if a connection is established via an API). The system is offered as a SaaS by the truck manufacturers.	ACEA
FEDeRATED (C) [57]	Indirect	EU Member States driven project to construct and validate federated platform communication to share data within the logistical sector by ensuring interoperability between different platforms.	European Commission
ERMES (C) [33]	Indirect	European project with the aim to coordinate and establish a comprehensive mobile emission factor database with data sets from all EU Member States. Used as emission factor default values in methodologies (i.e. GLEC and ISO) and for data enrichment and improved emission models.	ERMES, European Union
COPERT(C) [32]	Indirect	European emissions calculator for vehicles	EEA, JRC, LAT, Emisia

4.3. Stakeholder Analysis

The logistical sector is highly fragmented (especially on the carrier side; see Figs. 4.3, 4.4) and is directly and indirectly influenced by many stakeholders. Five main stakeholder clusters were identified: the logistical actors/-operators (directly involved in the logistical operations), policymakers (control the logistical actors via regulations and norms), framework and standards developer and owners (create GHG emissions reporting standards used by logistical actors), external third-party companies (support and interact with logistical actors along the chain), and the consumers (influence logistical actors, mainly shippers, by their purchase behaviour and expectations). Each stakeholder cluster and the individual stakeholder have their own needs and requirements for onboard sensor systems to determine emission factors. This makes the socio-technical environment increasingly complicated. Tables 4.3, 4.4, 4.5 summarise the mentioned stakeholders clusters, individual stakeholders, and provide a definition, their involvement & influence and needs of each stakeholder and the final stakeholder overview is visualised in Fig. 4.6.

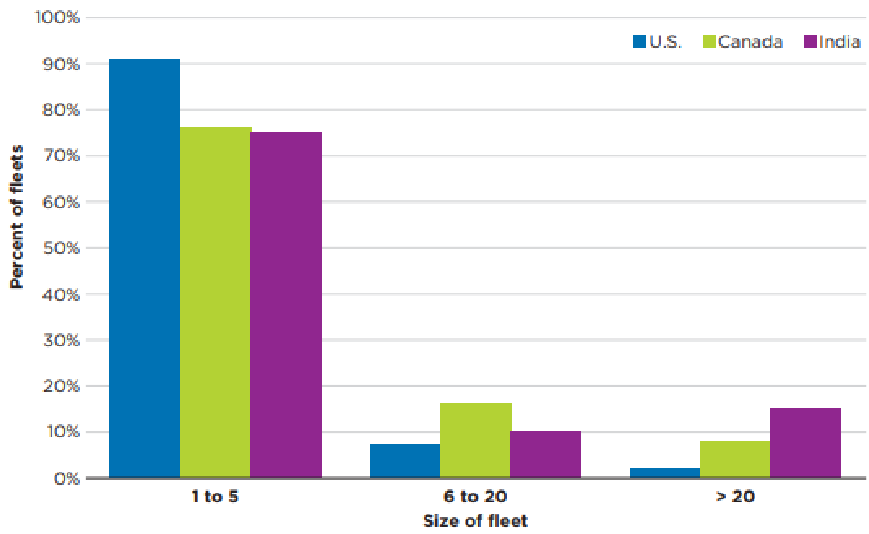

Figure 4.3: Average fleet size for US, Canada and India logistical operators. Most operators (<70%) operate a fleet with less than 5 vehicles. Retrieved from [82, p. 10].

Fahrzeuge	Anteil am Fuhrpark	Ø Anzahl pro Betrieb
Lastkraftwagen (normale Bauart)	16 %	17
Anhänger (normale Bauart)	9 %	13
Sattelzugmaschinen	19 %	21
Sattelauflieger (normale Bauart)	25 %	30
Motorwagen für Wechselbehälter	7 %	14
Anhänger für Wechselbehälter	10 %	17
Sonstige Fahrzeuge (Kühl-, Jumbo-, Silofahrzeuge, Pkw-Transporter etc.)	14 %	–

(a) The average fleet size of German carriers in 2010 per vehicle class. Retrieved from [30, p. 14]. German carriers had on average 17 HDV. (Translation: Fahrzeuge = vehicles, Anteil am Fuhrpark = Percentage of the fleet, Anzahl pro-Betrieb = Average number of vehicles per carrier, Lastkraftwagen = Heavy duty vehcile)

Beschäftigte pro Betrieb	Betriebe insgesamt	Einbetriebsunternehmen (ohne Niederlassungen)	Mehrbetriebs-unternehmen[1]
1 bis 10	17 %	26 %	7 %
11 bis 50	36 %	36 %	36 %
51 bis 100	20 %	23 %	17 %
101 bis 200	13 %	11 %	16 %
über 200	14 %	5 %	25 %
Gesamt	100 %	100 %	100 %

(b) The average employee's size of German carriers in 2015. Retrieved from [31, p. 22]. More than 50% of German carriers had, on average less than 50 employees. (Translation: Beschäftigte pro-Betrieb = Number of employees per carrier, bis = til, Betriebe insgesamt = Amount of carriers)

Figure 4.4: Fragmentation of German carriers. Left-site is the average fleet size of German carriers. On the right is the average employee size of German carriers.

Shippers start the logistical chain of the logistical actors by needing to transport a good by another external party (thus not by themselves). For that, they either directly contract a carrier for their transportation needs or hire a Logistical Service Provider (LSP) as a middle party that organises the transportation for them. In the second option, the LSP would then contract a final carrier for the transport. The carrier will then transport the goods from the shipper [8].

The entire logistical landscape is regulated by the policies and restrictions of the European Union (here the European Commission as an executive body), which in this case are Directives and Regulations that force shippers and carriers to report and disclose their GHG emissions of Scopes 1, 2 and 3 and desire national and European emission factor databases. Thus, shippers that outsource their transportation activities to carriers must retrieve the detailed emissions data for their shipments and Scope 3 reporting. However, carriers are highly reluctant to share emissions details on a transport leg basis with external organisations (mainly with shippers and other carriers), as it would be possible to conclude and understand their logistical operations directly (see I.9 & I.11 in Appendix B and interview A.3). Thus, they either report the aggregated total GHG emissions directly to the shippers or use the services of third-party IT and emissions calculation companies. Shippers can also use the services. These IT and emissions companies have dedicated contracts with both or one of the parties to clean, analyse and evaluate the emissions data and report them back to the shippers (see interview A.5)[9]. This intermediate step has two main advantages for the parties. Firstly, shippers do not need to invest time and resources in emissions determination (for some shippers, sustainability is one of their core pillars, thus they will also want to do these calculations in-house)[10]. Secondly, carriers can prove to shippers their ambition and progress in reducing GHG emissions without showing them detailed logistical data on a leg basis. The emissions reduction process is also getting more critical as shippers receive pressure from legislative bodies to reduce their emissions, but also end-consumers demand sustainable and green products nowadays. Thus, shippers are starting to select their carriers to a certain degree based on their emissions footprint (see interview A.3) [11].

The provided emissions and transport data of carriers to emissions modelling and supporting IT companies can also be used as model inputs for transport simulations, which then will be used for policy decisions and evaluations. Another entity involved in the data gathering is sensor manufacturers, which are often contracted by carriers to observe the fleet's fuel consumption due to shippers' requests for more detailed scope 3 data (see interview A.3) [11]. This additional primary data from the sensors can also be given to the emissions modelling and supporting IT companies (with the consent of the carriers) to enrich the data set and as an additional data source [11]. Both carriers and shippers need to report their GHG emissions, which is done based on emissions quantification methodologies provided by the framework and standard owners. The standards and methodologies are aligned with European regulations and aim to harmonise reporting standards. Optionally these reports can be audited by Verification, Validation and Assurance companies, which on the one hand, verify the used (historical) emissions data that was used to determine the total GHG emissions and, on the other hand, validate the used methodology and whether it is sufficient to make statements about future scenarios (see I.6 in Appendix B). Lastly, the auditor will either give a reasonable, limited or no verification for the report of GHG emissions. Especially for Scope 3 emissions, the industry standard is limited verification due to less available data[11]. Lastly, Venture Capital firms can also indirectly influence the landscape by financing innovative supporting IT companies, emissions modelling & calculation firms or sensor manufacturers. The same is the case for other financing and credit companies that provide funds to logistical operators (see evaluation interview C.4)[12] [5]. Fig. 4.5 visualises the stakeholder taxonomy, summarising the five main stakeholder clusters with the corresponding actors. All presented stakeholders were then integrated into the stakeholder overview in Fig. 4.6 with their related relationships and actions.

[8] this can be done by one large carrier or multiple smaller ones that each do a leg or part of the goods

[9] many IT companies are both emissions modelling and calculation firms*

[10] see I.14 in Appendix B

[11] see I.6 in Appendix B

[12] added after evaluation process* based on interview C.4

Table 4.3: Overview of relevant stakeholders directly or indirectly connected to European emissions factors. The first column represents the stakeholder cluster, the second column represents the individual stakeholder, the third column the definition of the stakeholders, the fourth column the involvement and influence of the stakeholder and the faiths column the need of the stakeholder.

Stakeholder Cluster	Stakeholder	Definition	Involvement and influence	Need
Logistical Actor		Actors that are directly involved in the movement, organization or planning of freight	Active entities of the transportation sector and obliged to reduce emissions.	
	Carrier	Logistical operators in charge of moving the good	Increased pressure from shippers to reduce emissions	Accepted GHG emissions calculation schemes and tools to prove their effort to reduce emissions without giving access to leg-specific data
	Shipper	The contracting party of the operation (normally the owner of the shipped good)	Increased pressure from end consumers & policymakers to reduce emissions	Higher detail on Scope 1 & 3 emissions Emissions reporting & Due-diligence Cost-savings from a carbon tax
	LSPs	Logistical service providers in charge of warehousing, planning, distributing or transportation of the goods.	Reduction of GHG emissions	Accepted GHG emissions calculation schemes and tools to prove their effort to reduce emissions without giving access to leg-specific data

Table 4.4: Overview of European standards, methodologies and collaborations connected to the measuring, reporting, or transmitting of data of emissions factors. The first column represents the stakeholder cluster, the second column represents the individual stakeholder, the third column the definition of the stakeholders, the fourth column the involvement and influence of the stakeholder and the faiths column the need of the stakeholder

Stakeholder cluster	Stakeholder	Definition	Involvement and influence	Need
Policymakers				
	European Commission	Part of the executive body of the European Union	- Green Deal - GDPR - Other policy documents	- Decarbonization of industries and transport activities - Adherence to standards - Data Privacy
Framework/ Standard owner				
	Smart Freight Centre	NPO with the aim to reduce GHG emissions that originate from freight activities	GLEC Framework	- Industry-wide standards and alignment of methodologies - Driving the decarbonization of logistics - Increased partnerships with stakeholders
	ISO	International Organization of Standardization	ISO14083:2023	- Clear instructions for GHG emissions calculations

Table 4.5: Overview the stakeholder clusters ”External Parties” & ”Consumers” with related European standards, methodologies and collaborations connected to the measuring, reporting, or transmitting of data of emissions factors. The first column represents the stakeholder cluster, the second column represents the individual stakeholder, the third column the definition of the stakeholders, the fourth column the involvement and influence of the stakeholder and the faiths column the need of the stakeholder

Stakeholder cluster	Stakeholder	Definition	Involvement and influence	Need
External third-parties				
	Verification, validations & assurance companies	- Verification = evaluation of historical data and the determination if the statements are materially correct - Validation = evaluation of used methodology, assumptions and limitations that support the a prediction of a future outcome - Assurance = Statement of an external auditor about the confidence of the used GHG inventory (verification) or emissions reduction program (validation)	Contracted by logistical actors or emissions calculation companies	Aligned methodologies
	Supporting IT-companies	Businesses that offer services to logistical companies such as emissions tracking or calculation tools	Contact along the entire logistical chain, depending on the business	Aligned methodologies and clear policy expectations
	Sensor manufacture	Producer/ assembler or maintainer of additional sensor equipment on road vehicles	Mainly in contact with carriers and IT companies to provide data of sensors	High-quality data and interoperable databases for data enrichment
	Emissions modelling & calculations firms	Emissions modelling specialists that create or use local or national emissions models to predict emissions	Results are the basis of policy decisions and validation of emissions reductions efforts	Alignment of methodologies and complete and detailed emissions factor databases
	Venture Capitalist firms	A private organization that funds other businesses with the goal to receive a return on their investment	Funding of disruptive start-ups in the logistical sector	
	Financing services*	A private or public organisation that supports businesses financially and can influence investments and strategies with it	Influence via their financing power. Can also ask for more sustainability	
Consumers	End customers	Buyers of end products (can be consumer or industry goods)	Indirectly influence the market with expectations and buying behaviour	Transparent and sustainable products

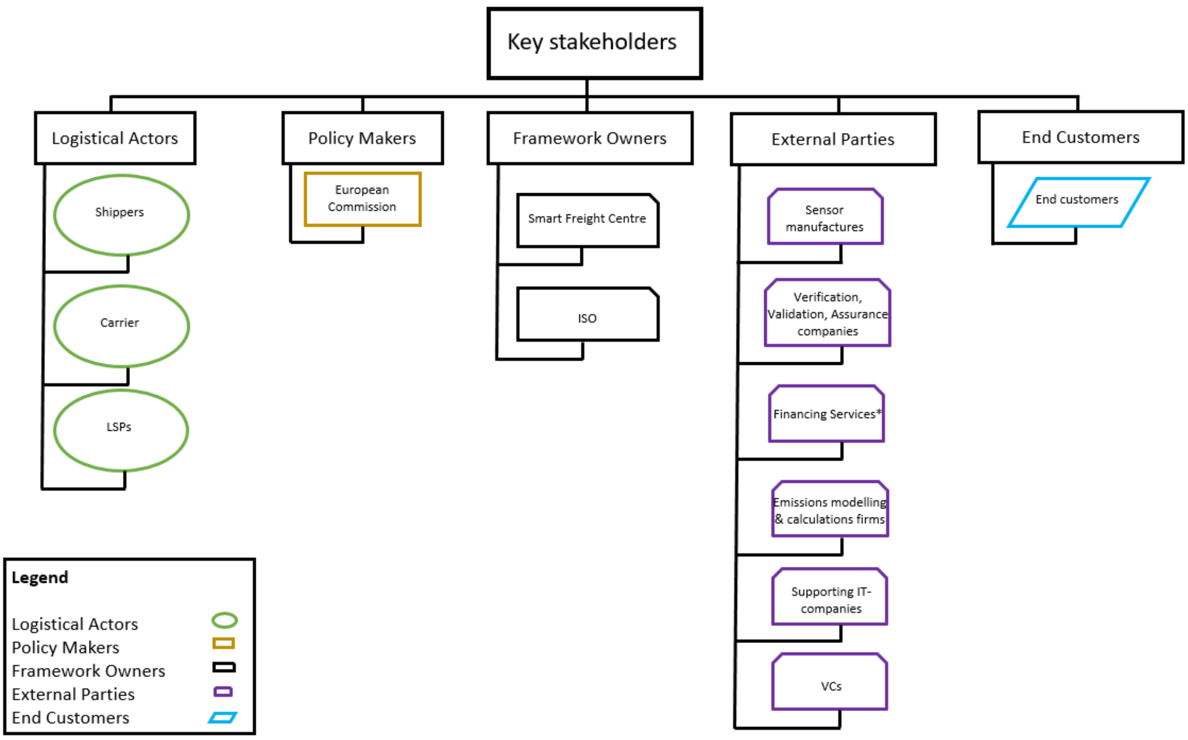

Figure 4.5: Stakeholder taxonomy with the five main stakeholder clusters and the corresponding stakeholders.

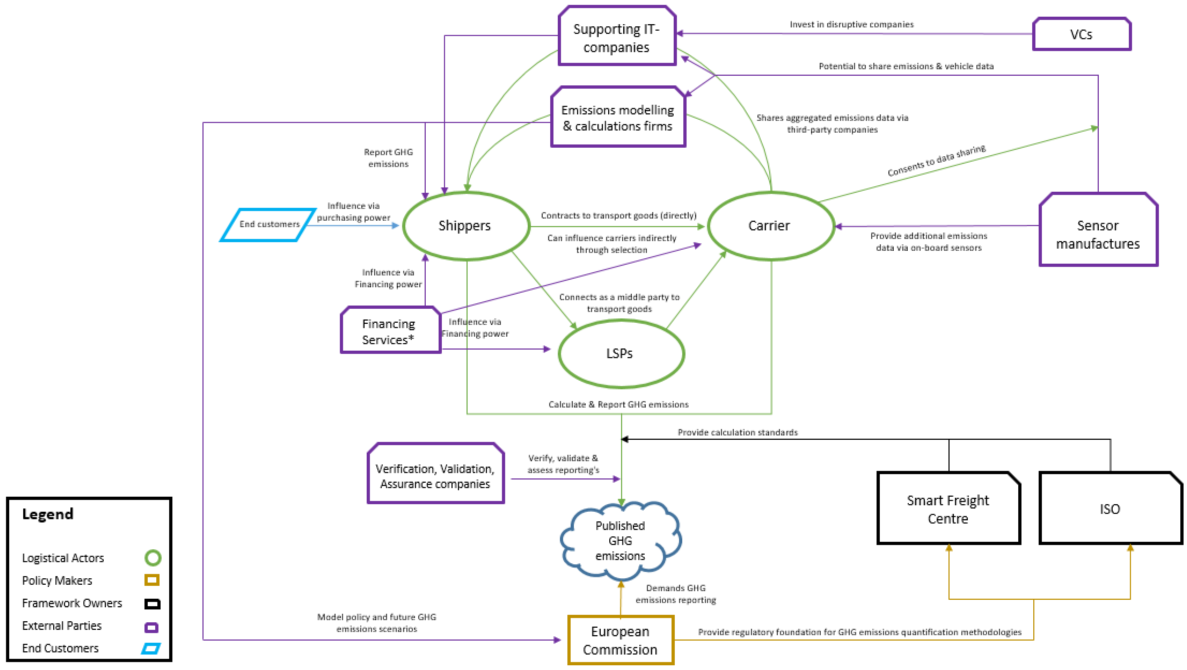

Figure 4.6: Most relevant stakeholders with main relationships between them, Logistical Actors and relations in green, Policymakers in brown, Framework and Standard owners in black, External third parties in purple, and Customers in light blue. A star indicates that the Stakeholder was added after the evaluation.

4.4. Results

4.4.1. Socio-Technical Analysis

Policy Analysis

The policies of the European Commission are the main regulatory influence on the system, which also result in design requirements and design principles of the system. Above all, the European Green Deal is a main driver, which sets emissions reduction targets in the logistical sector. The general trend in the sector to move from laboratory-based emission factors to real-world ones is facilitated by the Commission Regulation (EU) 2021/392, which demands vehicles to collect real-world data of their journey via their OBS system since March of 2023 and forces Member States to create and maintain a national emissions database. Moreover, Regulation (EU) 2017/1152 & 2017/1153 established a robust reporting scheme for emissions, which was further clarified by a list of vehicles' minimum reporting parameters to align standards. Regulation 2018/956 solely focused on reporting emissions and fuel consumption of vehicles from the classes M1, M2, N1 and N2 based on 78 parameters and made it mandatory for all HDV from 2019. Connected to it all is the GDPR with its data privacy concerns, which need to be adhered to while also adhering to the reporting and collection obligations. The upcoming mandatory reporting obligations for large companies in 2024 via the CSRD, in combination with the EmissionCountEU proposal of using ISO14083 as the new standard for GHG quantification, will further increase the need for not just logistical companies to familiarise them with its contents directly or indirectly via external GHG emissions quantification companies. Based on the institutional analysis, the following design principles and requirements were created as the first research contribution:

Table 4.6: Results of subquestion 1, Policy Analysis

Developed concepts	Concept	Source
Design Principle	Conformity with European legislations and methodologies	Policy Analysis, EU 2018/956, EU 2019/1939
Design Principle	Data Sovereignty	Policy Analysis, GDPR
Requirement	Conformity with European Data Privacy legislations	Policy Analysis, GDPR, [7]
Requirement	Conformity with European vehicle reporting & monitoring standards	Policy Analysis, EU 2018/956, EU 2019/1939
Requirement	Conformity with European emissions factor measurement legislations	Policy Analysis, EU 2018/1832, EU 2021/1392, [10]
Requirement	No tracking of purely driver-specific data	Policy Analysis, EU 2018/956
Requirement	The collected Data is compliant with ISO14083, GLEC 2.0	Methodology, Policy Analysis, CSRD, EmissionCountEU [13]

Methodologies, Industry Standards & Collaboration

The policy implications are supported by industry-wide GHG emissions calculation methodologies, such as the GLEC framework 2.0 and ISO 14083. They represent best practices in the reporting and quantification of GHG emissions, and their accreditation is sought after by many companies to showcase their sustainability efforts to their customers and creditors. Detailed data is needed for the GHG emissions quantification after ISO standards. However, many logistical operators are reluctant to provide their data, as it could result in competitive disadvantages. Thus, external IT companies have the unique role and opportunity to act as middle companies to connect shippers and carriers and to facilitate data sharing via their servers. The shared data can come from various sources, and the FMS standard was a first effort of industry partners (here OEMs) to create a harmonised reporting scheme for trucks in road logistics. The accuracy was poor. Thus many logistical operators hired external sensor companies to install third-party hardware, which led to many third-party emissions databases. The updated rFMS standards improved the accuracy of the reported data, but the third-party databases' importance stays for operations such as data enriching. Thus, modern emissions calculation requires consistently combining multiple data sources and sets with sufficient granularity and quality to be aligned with ISO and GLEC methodologies. Collaboration efforts, such as the ERMES project to create a European emission database with organisations of each Member State contributing to it or the FEDeRATED projects, are European answers to the need for more harmonisation and collaboration in the field. However, they fail to get all stakeholders equally onboard, as they are often heavily dominated by logistical operators, not data management experts. The derived requirements were as follows:

Table 4.7: Results of subquestion 1, Methodologies, Standards & Collaborations

Developed concepts	Concept	Source
Design Principle	System Interoperability	Standards, rFMS
Requirement	The system can receive, store and process data	Standards, ISO and GLEC 2.0
Requirement	The collected data is compliant with ISO 14083 & GLEC 2.0	Methodologies, Policy Analysis CSRD, EmissionCountEU[14]
Requirement	The system supports interoperability between platforms	Standards, Collaboration, ISO and GLEC [11, 10]
Requirement	The system can incorporate multiple data sources	Standards, ISO and GLEC, [11, 10, 7]

Stakeholders

One reason why it is so difficult to get all stakeholders involved in these joint collaboration projects is the sheer number of them. Thousands of small and micro transport companies move goods on the carrier site for the shipper, which increases the difficulty of reaching all of them. However, there are, in general, many different stakeholders involved in the system. Thus, they were clustered into five general groups: logistical actors/operators, policymakers, framework & standard owners, external third parties, and consumers. Each cluster has its own characteristics and needs when it comes to a system of onboard sensors to calculate GHG emissions. Thus, it might not be possible to create a one-for-all solution, but it might be necessary to create an abstract system architecture which groups all the general requirements for all stakeholders. The creation of a stakeholder overview with possible relations between the stakeholders and general working procedures supported the analysis. The main logistical actors start the chain with the movement and generation of data. The system is governed by the regulatory framework provided by the European Commission. Again the intermediate role of external IT companies in the data-sharing process can be highlighted to connect shippers and carriers. It benefits the shippers and carriers as they can outsource the activities and focus on their core service. At the same time, they can also prove their sustainability efforts to cut emissions with the results. A final realisation was that it would not be possible for a single stakeholder, yet a stakeholder cluster[15], to make lasting changes to the system. Each change towards an industry-wide system architecture to calculate GHG emissions via onboard sensor system must happen as a unit.

4.5. Evaluation of the Stakeholder Overview

4.5.1. Interview usage for the evaluation process

The evaluation process of the created stakeholder process was done by conducting evaluation interviews with industry experts, the available literature and with the knowledge insights of the informal conversations (see Table B). For that, three different evaluation interviews were done with a *Carrier* (see evaluation interview C.3), with an *LSP* (see evaluation interview C.4), and an External IT company (see evaluation interview C.5). Thus, the interview partner represented the main and diverse actors in the system to access the stakeholder overview. The evaluation interviews followed the same semi-structured structure as the design interviews, and participants received the stakeholder overview, the stakeholder taxonomy and the corresponding evaluation questions in advance. Participants were asked and evaluated the presented overview based on the completeness of relevant stakeholders and the correctness of the established relations between the various stakeholder. The exact evaluation structure with the corresponding questions can be found in Appendix C.1. Each participating evaluation expert only evaluated designs that they did not contribute to in the design phase to have an unbiased opinion. The only exception is the carrier that mentioned (unsolicited) requirements in the evaluation process. Due to the diversity of the topic, it was expected that not every stakeholder could evaluate every part of the research. Thus, only directly important remarks were used and added to the evaluation interview summaries. Fig. 4.7 represents the evaluation process and the medium for the evaluation while presenting the coming evaluation steps.

The evaluation process followed a formative evaluation, meaning that the stakeholder's feedback was already implemented into the designs and created concepts to improve them. However, parts that were altered because of it were highlighted[16], and the feedback was clearly communicated (see Appendix D.1 for the designs of the

[15] apart from the policy bodies

[16] all changes in the designs or requirements were indicated with a "*"

first iteration) [78]. The exact evaluation strategy followed an ex-ante and naturalistic approach. The artefact was a typical socio-technical artefact with many involved stakeholders. Thus, the naturalistic approach was chosen [78]. The artefact was not deployed yet; thus the ex-ante approach was chosen, and real users and stakeholders were used for the evaluation [78]. The evaluation methods (interviews, informal talks, and literature) also suit the combination of naturalistic and ex-ante evaluation strategy [78].

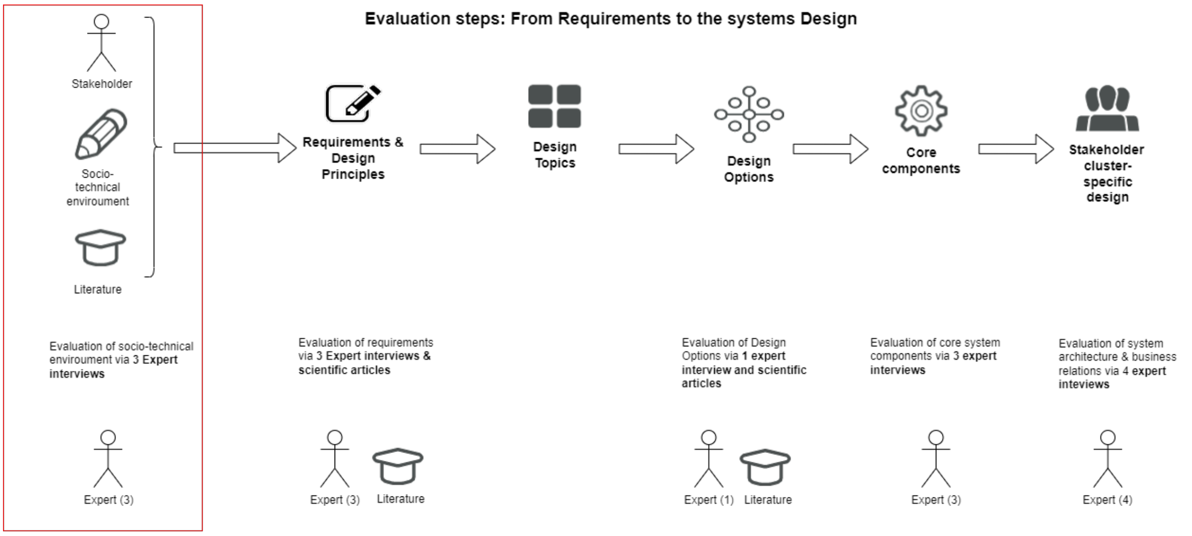

Figure 4.7: Evaluation procedure for the stakeholder overview via 3 expert interviews.

4.5.2. Stakeholder evaluation

The first validation was conducted on the stakeholder overview in Fig. 4.6. The general comments for the abstract representation of stakeholder relationships were highly positive, with all validation partners agreeing on the validity and correct representation of the socio-technical system on an abstract level. A carrier approved the idea of clustering different stakeholder groups together and commented that it was a *"Good idea to cluster the stakeholder"* (see evaluation interview C.3) [3]. However, what also became clear was that the evaluation of the stakeholder overview was scope-specific, and many experts had different opinions on where to make that system cut. While the carrier and the LSPs argued to include more stakeholders such as financing organisations (banks, credit firms), shareholders or fuel and OEM producers, did the IT expert argue that the current scope is sufficient and commented *"This level of detail makes the most sense to me if you would increase the scope [of the overview] until scope 3 emissions than the main relations could be lost in the complexity"* (see evaluation interview C.5) [4]. This was the opposite conclusion from the previously mentioned evaluation experts, who concluded the following *"Potential to add the fuel producer (especially for SAF and bio-fuels) and shareholders for reporting issues"* (see evaluation interview C.3) [3] or mentioned *"The first idea I would have are banks and credit companies, which evaluate the sustainability score of companies and would give credits and financing based on that. For us, that is a main driver to have also green transport to get positive scores from them and thus also financing."* (see evaluation interview C.4) [3] [5]. Thus, a correct scope seems to be dependent on the point of view of the stakeholder. Other points were that, especially for IT companies, a concrete declaration of either "Supporting IT company" or "Emissions modelling & calculation firm" was difficult as many offer both supporting and calculation services (see evaluation interview C.5) [4]. Lastly, it was mentioned that also other stakeholders create pressure towards other logistical actors to increase their sustainability efforts *"I will support the statement that the end customer creates pressure on the shipper to reduce emissions, but there are also companies such as Carrier and LSP, that create this pressure, however, the scope of this representation is clear and correct"* (see evaluation interview C.5) [4].

Based on the received feedback, the following steps were taken:

1. The addition of indirect pressure from other stakeholders (apart from end customers) to increase the sustainability of logistical processes
2. The addition of financing services for the logistical operators
3. The addition of fuel providers and OEM was not done, as these are only indirectly relevant for the logistical operators and would therefore increase the complexity. However, they will be mentioned in the text and the limitations of the overview. The same is the case for the possibility of IT companies having multiple roles.

4.6. Conclusion

The first sub-questions aimed to explore and identify the socio-technical environment's influences on road logistics sensor systems to gather primary data from the vehicles. The main conclusion of the first sub-question was the importance of the institutional analysis to the system's design, as it showcased the connection and influence of policymakers to the system. The European Commission acts as an active system designer by setting general design principles and requirements for the system. Thus, the European Union first defined a goal for the system (the design principle) very broadly, and it continues by what such a system should fulfil and how this should be achieved. One example of that is the Commission regulation (EU) 2019/1939 [46], which first outlines the importance of emissions monitoring of vehicles and creates a new CO_2 standard (the design principle) as a general goal of the system to create greater transparency in the monitoring of emission values of vehicles. It then defines that each Member State of the European Union should create and maintain a national emissions database of vehicles. Lastly, they established concrete instructions on how to test by setting the exact parameters that should be reported and then how to measure them, such as the vehicle not drive faster than 145km/h, it should drive at least 16km in rural areas, and not drive more than 43% of the total distance in those [98]. Another important insight was the significance of external IT companies and their possibility to connect the shipper with carriers and LSPs to combine their data without risking a competitive disadvantage for the carriers. This position would imply a neutral position for the IT company. Lastly, the stakeholder overview also showcased the magnitude of stakeholders with their different needs and requirements, which already indicated the issue of creating a list of requirements for the system that applies to all stakeholders.

5
Objectives & Requirements

This chapter covers the second sub-question *"What are the technical requirements for onboard sensor systems in road freight logistics?"* using the insights of the previous institutional and stakeholder analysis and the knowledge gained from the expert interviews. The expected deliverable is a list of all functional and non-function requirements for the system and their respective clustering (institutional, stakeholder, and technical), influence level to a specific stakeholder cluster, their source and the relation to a design principle. Section 5.1 elaborates on the requirement creation via the expert interviews, which is followed by the explanation of the design principles (see section 5.2) and the explanation of the requirement creation and categorisation in section 5.3. The requirements are derived in section 5.4 by first deriving the functional requirements (section 5.4.1) and then the non-functional requirements (section 5.4.2). Section 5.5 will present the list of the final requirements, followed by the evaluation of the requirements (per cluster) with the help of expert evaluation interviews and literature in section 5.6, thus answering the fourth sub-question *"What is, according to industry experts, the quality of the designed artefacts and requirements?"*. Finally, section 5.7 will conclude the chapter.

5.1. Interview Usage

Interviews with industry experts were used for the requirements analysis, creation, and finalisation. Three interviews with a *Fuel consumption, sensor and IT emissions tracking expert* (see interview A.3), an *Emissions factor and modelling expert* (see interview A.4) and a *Data marketplace and IoT aggregator expert* (see interview A.5) were conducted. A *Carrier* also mentioned some requirements during the interview, even though they were not asked for it. However, the insights are still mentioned here. The interview partners received a consent form[1] and the question catalogue in advance (see the questions catalogue in Appendix A.1). The asked questions were aimed at understanding what the experts value in a potential system that would track and monitor their GHG emissions. For that, the experts were asked about reporting frequencies of the system, user-specific usage of primary data at this moment, the importance of access control and data sovereignty or the importance of compliance with existing standards. The insights from the interviews were processed and used to finalise the design principle and requirements analysis in combination with the available literature on the topic of systems requirements and informal conversations.

5.2. Design Principles

The design principles are high-level building blocks in the system architecture that should be addressed and thus guide the design process. The previous socio-technical analysis in chapter 4 already established three design principles. The coming section will briefly recap the three design principles before introducing the final one.

DP.1 = Conformity with European legislation and methodologies

A system that does not conform with European legislation is considered useless to any stakeholder and thus not viable. The common industry standards for GHG emissions quantification and calculation methodologies are closely related to that. Many stakeholders mentioned the importance of showing to their customers or the European bodies their efforts to reduce their GHG emissions via the previously mentioned methodologies like the ISO14083 or the GLEC framework [11, 10]. Thus, the proposed design should also be aligned with the dominant calculation methodologies.

[1] To indicate their willingness to participate in the study

DP.2 = System interoperability

A second challenge, which echoed in nearly every interview, informal conversation and many articles, was the system interoperability (see section 4.2.3, and informal conversations I.1, I.4, I.9, I.11, and I.13 in Appendix B and interview summary A.5) [7]. Here the phrase system interoperability mainly refers to the ability of the design to incorporate multiple databases as data sources.

DP.3 = Data sovereignty

A third challenge, which is especially important for the directly involved logistical operators (shippers, carriers, LSPs), is sharing sensitive data about their operations and transport legs.

Lastly, the fourth challenge concerns the system's flexibility and capabilities. Interviews have revealed that many shippers use the data from their transport legs on an aggregated level to report the total GHG emissions (thus, they do not need very detailed ping rates, as long as they have the total fuel consumption and GHG emissions, in the end, [7]), however, if certain legs are continuously troublesome and delayed they might want to increase the ping rate and detail of this specific trips [7]. Thus, they need to ensure that the system offers some flexibility on an event basis for specific legs with different reporting frequencies. Thus, the final design principle is:

DP.4 = The system should allow event-specific reporting

5.3. Requirement categorisation

The proposed architecture had two requirements types that it needed to fulfil. Firstly, functional requirements describe the system's functional behaviour and the tasks the system should execute, while non-functional requirements describe how the system should work [79, 24]. The requirements were categorised into three clusters. The first cluster was the institutional requirements, which were mainly derived during the first sub-question and the analysis of the socio-technical environment, but also throughout the interviews with stakeholders. The second cluster was the stakeholder-specific requirements gathered throughout the interviews and the informal conversations. Lastly, the technical requirements were based on the technical aspects the system needed to fulfil. The requirement analysis was based on the ISO/ICE/IEEE 29148 standard, which provides guidance for requirements engineering in systems and software engineering [77]. The standard distinguishes between requirements derived from policy and regulatory conditions (see section 4.2.2), stakeholder-specific needs and thus their requirements (see Tables 4.3, 4.4, and 4.5 and section 4.3) and general technical requirements (see this chapter 5). Thus, the approach for the requirements creation was first to identify the relevant stakeholder in the system and their needs regarding such a system. Secondly, the interviews and literature provided additional insights into the general requirements of the system. The third step was the analysis of the general requirements and the categorisation into institutional, stakeholder and technical requirements. Fourth the clustered requirements were specified into concrete statements and grouped within the functional and non-functional requirements domain [77]. Each requirement received a requirement ID to identify the type and the cluster (see Table 5.1 for an explanation of the different requirement IDs).

Table 5.1: Explanation of different requirements nomenclature

Requirements ID	Definition
IFR.x	Institutional Functional Requirements number x
SFR.x	Stakeholder Functional Requirements number x
TFR.x	Technical Functional Requirements number x
INFR.x	Institutional Non-Functional Requirements number x
SNFR.x	Stakeholder Non-Functional Requirements number x
TNFR.x	Technical Non-Functional Requirements number x

5.4. Requirement Analysis

5.4.1. Functional Requirements

Table 5.2 presents the final identified functional requirements of the analysis. All institutional requirements were derived from the policy analysis in combination with the expert interviews. IRF.1 is relevant for all stakeholders in the system that deal with or handle data. It is directly derived from the GDPR and is essential for the entire data handling, storing and processing chain. Nearly every interview mentioned this legislation's direct or indirect importance, which shaped and determined their operations [11, 7]. One interview partner summarised the importance of the GDPR as follows:

„Our data sharing we hold very, very strictly to GDPR rights, so we set up the consent […] where we […] will speak with the carrier […] and ask them are we OK to share your data under these circumferences […]?“ (see interview A.5) [7], and when asked about their responsibility of using data from others (like OEMs) under their umbrella architecture (thus, they would not be the main party that would need to conform with the GDPR), they answered: *„We have a role as a data owner. Once you touch that data, you are responsible for that data disseminated“* (see interview A.5) [7]. Both quotes highlight the importance of the GDPR and all related data-handling practices.

The European Directives and regulations EU2018/956 and EU2019/1939 directly established IRF.2, as the monitoring and reporting of newly registered HDV are mandatory for vehicles after 2019. Thus, the system must be compliant with it. The same is true for IRF.3, as emissions measurement systems are directly regulated by EU2018/1832 & EU2021/1392, which was also emphasised during the interviews (see interview A.4) [10]. One stakeholder emphasised this by stating that they often work for the European Bodies: *We have tested more than 50 or 60 vehicles on behalf of the European Commission as part of the so-called […] Service Conformity Program of the European Commission. So […] when you measure emissions on the road, you do not just […] turn on the engine and you start driving. […] There are certain conditions that need to be met in order for the test to be considered valid“* (see interview A.4) [10]. And they continued by stating *„In the RDE legislation […] they have prescribed certain driving styles so there are certain limits for speed or acceleration“* (see interview A.4) [10]. Thus, again emphasising the strong influence of the legislation.

SFR.1 & SFR.2 were directly derived from the expert interview with a logistical operator[2] (see interview A.6) [8]. The carrier strongly argued for restricted access to the system by first highlighting the privacy and GDPR issues if the system had open access. Full access to the system and data from their customers would need additional security and privacy concerns, and secondly, it would also give too much insight into their system. SFR.2 is a requirement to ensure that the generated data can be stored, sent and processed on the vehicle itself to have complete and consistent data (see interview A.6) [8].

Finally, the technical and functional requirement TFR.1 was established by the interviews with experts (see interviews A.3, A.4, A.5) [11, 10, 7]. Everyone stressed the importance of having access to multiple datasets and bases to add additional value to the data, increase its quality or enrich and make predictions based on it. Some statements were as follows: *„We have different data sources. Not only […] our database, but also from other institutions, and other massive databases“* (see interview A.4) [10]. *„We collaborate with third parties for enriching the data to get it in compliance with our system“* (see interview A.5) [7]. *„We can predict, based on historical data, what the fuel consumption would be and we can provide […] a more accurate visualisation and not just visualisation but also insights for the customer“* (see interview A.3) [11]. All these statements indicate the importance of using different data sets and sources for the system.

Table 5.2: Overview of the derived functional requirements, sorted into each cluster (institutional (IFR), Stakeholder (SFR), and Technical (TFR)).

Requirements ID	Requirements
IFR.1	Conformity with European Data Privacy legislations
IFR.2	Conformity with European vehicle reporting & monitoring legislation
IFR.3	Conformity with European emissions factor measurement legislations
SFR.1	Access to the system is restricted
SFR.2	The system can receive, store, and process data
TFR.1	The system can incorporate other data sources

5.4.2. Non-Functional Requirements

Table 5.3 presents the final identified non-functional requirements for the system. It has has one institutional non-functional requirement related to data privacy. INFR.1 specifies that no purely driver-specific information can

[2] SFR.2 was also established through the standards analysis in section 4.2.3

be tracked to limit complications and GDPR violations[3]. Many stakeholders expressed the effort they had to go through, explaining to authorities which data they wanted to track (see interview A.5) [7]. Thus, only vehicle, GHG emissions and vehicle parameters should be tracked. Many vehicle parameters are also very closely related to drivers and their behaviour (like acceleration and driving behaviour), which complicates the issue.

The remaining stakeholder and technical-related non-functional requirements were derived from the interviews and the analysis in chapter 4. SNFR.1 is an important non-functional requirement for nearly all stakeholders, especially the main logistical actors, and is concerned about the system's security. Nearly all interviewed stakeholders mentioned that the sharing of operational data from the logistical actors is a sensitive topic, thus no data leaks or wrongful data access can happen. Experts summarised the dilemma of data sharing and the resulting requirement of security of the system as follows: *"The funny thing is, everyone talks about cooperation and data sharing and everyone is very keen to get their hands on data of others, but nobody wants to share their data "* (see interview A.5) [7]. Thus, the general security of the system to reduce the chance of a data leak or accidental data sharing between stakeholders is important. The requirement SNFR.2 again emphasises the root cause of many of the digitalisation and GHG emissions efforts of the shippers, LSPs and carriers. They need to reduce their emissions, or they risk high carbon prices and fines in the future. For them, the way to prove those ambitions are accepted and approved methodologies like the GLEC 2.0 and the ISO 14083 standard. Thus, for them, it is a requirement that they contract a third-party company for the handling, analysis or calculation of their fleet data to calculate or report GHG emissions to be in line with these standards. An LSP mentioned that the accreditation adds another level of authority to the GHG reports, which is desired by their customer *"An officially accredited system (ISO or GLEC) is very important to us for our customers"* (see interview C.4) [5].

SNFR.3 is closely related to TFR.1 and incorporates different data sources. The discussion above in subsection 5.4.1 already touched upon the importance of having multiple data sources to analyse, enrich and make predictions (see the provided quotes and the expert's interviews A.3, A.4, and A.5). Thus, a way to achieve the incorporation of different data sources would be to support interoperability between different (data sources) platforms via common data structures and formats. The fourth stakeholder requirement is the flexibility of the system towards user requests. What is meant by that is that the system would be able to adjust reporting frequencies or analysed parameters depending on the customer's needs. The expert of the data marketplace is often faced with this challenge, and every customer needs specific services, which sometimes have to be changed if problems occur (see interview A.5) [7]. An example is the following quote from the same interview: *"Large and complex supply chains tend to work on exceptions. 85-90% of it is going well, but there is that 15 to 10% that you know messes up every time. […] So it is a lot about increasing visibility in these stages. Thus, customers look for more events. What is happening?, What is going wrong […]? "*.

The SNFR.5 requirement concerns the usability of the system and structures. Many customers of the shipper and carrier have great IT knowledge for order fulfilment, customer support or their internal networks. However, the technical knowledge that is required for a system concerned about tracking vehicles and calculating GHG emissions needs different kinds of technical knowledge that spans over multiple organisations and on-top as well knowledge about GHG emission quantification standards [7]. *"Most of the companies IT are focused on customers facing websites and services and the data on logistics is not their expertise"* (see interview A.5) [7]. This is, of course, stakeholder-specific and might change in the future as companies start to need this expertise, however, the system's usability should be customer friendly. Lastly, SNFR.6 is about a feature that the system should have. The system should visualise the potential GHG emissions savings from alternative fuel types or energy carriers and then compare them to the traditional ones. With this feature, it would be possible to directly compare the impact of new eco-friendly energy carriers and use this information to justify investments [8]. "*Dashboards that compare fuel types and their savings, based on primary data, could be used to validate investment decisions*" (see interview A.6) [8].

The technical non-functional requirements were also derived from the stakeholder interviews and the analyses above. However, they were placed into the technical cluster, as they are mainly related to the system's technical specifications. The first requirement, TNFR.1, is about the system's scalability. Scalability is important to continuously add new data sources and clients (which could also be considered data sources) to the system. This is also facilitated by the institutional regulations making monitoring of newly registered HDV mandatory. Thus, many companies are looking for solutions. This requirement is especially important for IT platform providers. This was confirmed by existing literature from Yun et al. (2019) [127] and by comments of experts: *"We are always trying to stay ahead of it (their storage/ IT capacities). We want to give ourselves a lot of room to grow and not cut cost on […] data storage and things like that, because it is so cheap "* (see interview A.5) [7]. The second technical requirement, TNFR.2, is about the translatability of the various reporting formats and structures in which the raw data arrives at the IT or shipping company. The system must be able to accept raw input data that is not directly in line with their formats and convert it into correct ones (see informal conversation I.13 in B and interview A.5) [7]. This is also closely related to interoperability and having multiple data sources for the system. Finally, TNFR.3 is fundamental to the system. It should transmit consistent and complete data sets to limit the guessing factor along the transport leg.

[3]However, there is always the threat of data breaches

Table 5.3: Overview of the derived non-functional requirements, sorted into each cluster (institutional (INFR), Stakeholder (SNFR), and Technical (TNFR)).

Requirements ID	Requirements
INFR.1	No tracking of purely driver-specific data
SNFR.1	The system and data storage is secure
SNFR.2	The collected data is compliant with ISO 14083 & GLEC 2.0
SNFR.3	The system supports interoperability between platforms
SNFR.4	Flexibility per user request
SNFR.5	Ease of use for user
SNFR.6	The system can portray GHG emissions savings and compare fuel types
TNFR.1	Scalability of the system
TNFR.2	Translatability of various data structures & languages
TNFR.3	The system transmits complete and consistent data

5.5. Results

The combination of literature sources and the conducted interviews with different stakeholders experts resulted in the establishment of four general design principles being:

1. Conformity with European legislation and methodologies
2. System interoperability
3. Data sovereignty
4. Event-specific reporting

All four design principles address challenges and hopes for the system, which will be influential throughout the entire design process. A total of six functional and ten non-functional requirements were determined. Table 5.4 summarises all requirements and relates them to the most influenced and important stakeholder of that requirement, provides the source of the requirement and relates it to the corresponding design principle[4]. Fig. 5.1 relates the functional and non-functional requirements to each design principle in a requirement tree structure.

[4] The abbreviation A.x indicates an interview summary as the source for the requirement

Table 5.4: Overview of the final functional and non-function requirements. The first column provides the requirement ID (IFR = Institutional Functional Requirement, INFR = Institutional Non-Functional Requirement, SFR = Stakeholder Functional Requirement, SNFR = Stakeholder Non-Functional Requirement, TFR = Technical Functional Requirement, TNFR = Technical Non-Functional Requirement), the second column specifies the requirement, the third column addresses which stakeholder(s) are mainly influenced by it, the fourth column provides the source of the requirement and the final column relates the requirement to a design principle.

Requirement ID	Requirement	Stakeholder	Source	Relation to the design principle
IFR.1	Conformity with European Data Privacy legislations	All	Policy Analysis, GDPR, A.5 , A.6	DP.1
IFR.2	Conformity with European vehicle reporting & monitoring legislations	Shippers, Carriers, LSPs	Policy Analysis, EU 2018/956, EU 2019/1939	DP. 1
IFR.3	Conformity with European emission factor measurement legislations	Shippers, Carriers, LSPs, Emissions measurement	Policy Analysis, EU 2018/1832, EU 2021/1392, A.4	DP. 1
INFR.1	No tracking of purely driver specific data	Shippers, Carriers, LSPs	Policy Analysis, EU 2018/956	DP.1
SFR.1	Access to the system is restricted	Shippers, Carriers, LSPs	A.6	DP.3
SFR.2	Can receive, store and process data	Shippers, Carriers, LSPs	A.6, Standards	DP.2, DP.3
SNFR.1	The system and data storage is secure	OEMs, carriers, shippers, LSP	A.3, [20]	DP.3
SNFR.2	The collected data is compliant with ISO14083 & GLEC 2.0	Shippers, Carriers, LSPs	A.3,A.4, Methodologies	DP.1
SNFR.3	The system supports interoperability between platforms	Shippers, Carriers, LSPs, Emissions measurement	A.3,A.4,Standards & Collaboration Analysis ISO and GLEC	DP.2
SNFR.4	Flexibility per user request	Shippers, Carriers, IT companies	A.5, L [127]	DP.4
SNFR.5	Ease of use for user	Shippers and Carriers	A.5, L [127]	DP.2
SNFR.6	The system can portray GHG emissions savings and compare fuel types	Shippers, LSPs and Carriers	A.6	DP.4
TFR. 1	The system can incorporate other data sources	OEMs, carriers, shippers, LSP	A.3,A.4, A.5, Standards Analysis, ISO and GLEC	DP.2
TNFR. 1	Scalability of system	All	A.5, L [127]	DP.2, DP.4
TNFR.2	Translatability of various structures and languages	IT-companies	A.5	DP.2
TNFR.3	The system transmits complete and consistent data	All	A.6	DP.1, DP.4

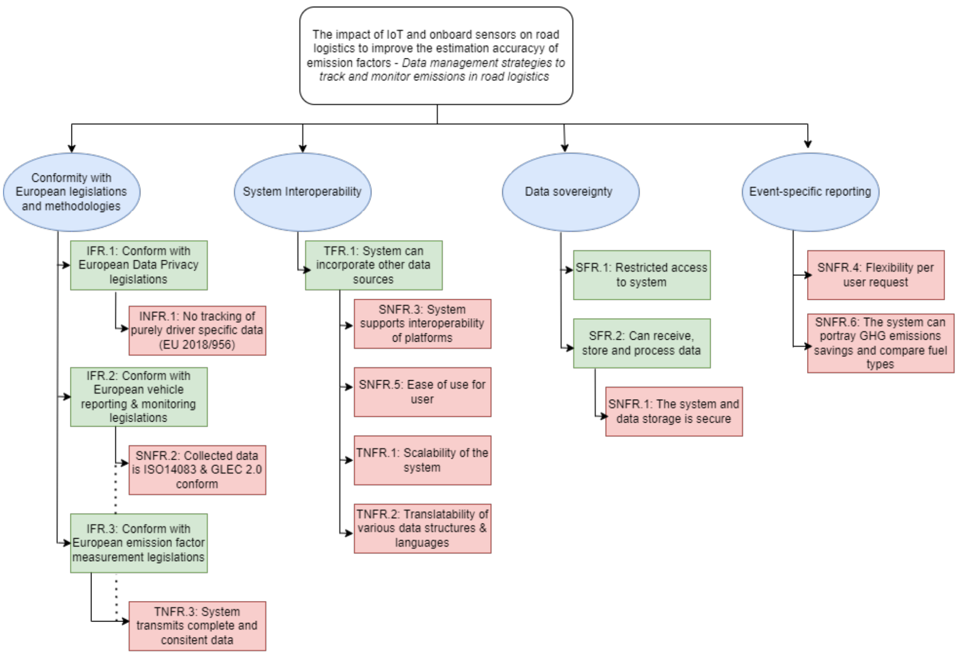

Figure 5.1: Requirements tree derived from literature, expert interviews and policy document.

5.6. Evaluation of requirements

5.6.1. Interview usage for the evaluation process

The evaluation process of the system requirements was conducted in the same manner as the stakeholder overview and followed the same semi-structured approach where the evaluation experts received the to-be-evaluated system requirements (they received Table 5.4 beforehand with a consent form). The system requirements were evaluated in three interviews with a *Carrier* (see evaluation interview C.3), with an *LSP* (see evaluation interview C.4), and with an *External IT company* (see evaluation interview C.5). The interview partners were asked to evaluate the completeness and correctness of the system requirements, and Appendix C.1 provides the exact structure of the interview. Thus, the evaluation criteria were first to evaluate if the gathered systems requirements were correct and secondly, if stakeholders thought the requirements were complete or if additional ones should be added. The insights of the evaluation process were again directly implemented into the results of this section and highlighted if an adaptation happened based on an evaluation interview. Some requirements were also evaluated with the help of literature. Table 5.5 summarises the system requirements evaluation and indicates the evaluation method (thus, via interviews or literature). Fig. 5.2 visualises the evaluation process of the requirements.

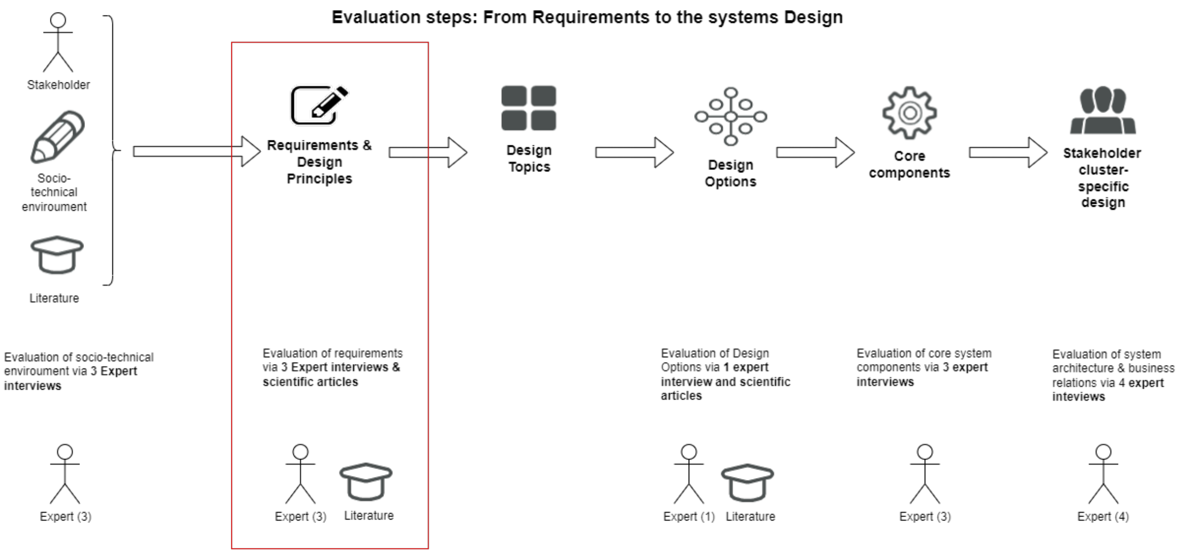

Figure 5.2: Evaluation process for the requirements.

5.6.2. Requirements evaluation

Table 5.5 provides an overview of all the previously mentioned and derived requirements. A checkmark indicates that the expert approved the requirement, a checkmark in brackets indicates partial approval (on a specific system or customer setup), and a cross indicates that the requirement was not approved. A citation of an evaluation interview indicates that the requirement was directly mentioned during the interview process (the evaluation interview citations are [6, 3, 5, 4]), while the capital letter ”L ” followed by a source indicates that the requirement was approved through a literature source. At first glance, it can be observed that the evaluation interview partners approved all requirements; some directly by explicitly mentioning them, some indirectly.

The first vital evaluation contribution was the assessment of the used requirements clustering into institutional, stakeholder-specific and technical requirements. The IT expert concluded that *”[...] it makes sense to group the various requirements into the different clusters to showcase the main application point and also clearly separate the technical requirements from the rest”* (see interview C.5) [4]. The expert continued that creating a clear boundary for the requirements is difficult, as many stakeholders might have different requirements for such a system (here, the requirement SNFR.4 was especially highlighted) [4]. Other critical contributions from multiple evaluation partners were that all presented requirements were correct and logical to them. However, they also highlighted that some requirements are more important based on the role of the company in the system. The LSP summarised it as follows: *”All requirements are correct and relevant, however, some are more relevant for us”* (see interview C.4) [5], while the IT expert supported the statement by concluding: *”None of the requirements seems to be completely off-target or inappropriate, and I see that they might have different importance per stakeholder”* (see interview C.5) [4]. Thus, it is important to understand that every stakeholder has their importance per requirement, depending on the technical knowledge inside the company or the specific needs.

5.6.3. Institutional requirements evaluation

One of the most important requirements is data privacy (IFR.1). Multiple experts highlighted this point. An LSP added that they would not access any primary data sources of their carriers directly to ensure compliance with the requirement and the GDPR *"We do not want to have a direct connection to the subcontractors to stay within data privacy regulations and to respect the sovereignty and business model of the carriers"*, thus directly validating the requirement (see interview C.4) [5]. The IT expert added *"We always need consent to access the data from the data owner for the shipper. Even if we have already worked with both the shipper and the carrier before on different projects, we again need to receive consent for this specific partnership"* (see interview C.5) [4]. Again, this highlights the importance of data privacy and the requirement for all activities connected to the data.

IFR.2 was not explicitly mentioned by any evaluator. Reasons for that could be that non of the evaluator was an OEM or vehicle producer. Thus, they would expect the European vehicle reporting legislation to be fulfilled by the vehicle providers. Requirement IFR.3 was indirectly supported by the LSP by stating that *"An officially accredited system (ISO or GLEC) is very important to us for our customers"* (see interview C.4) [5]. Thus, the system must also conform with the European emission factors legislation.

INFR.1 is especially significant for the actors that are processing the data, as they would be the first entity to recognise a potential data breach due to non-vehicle-related data points. This is often of control for them, as they have to work with the data they receive. *"We present the data to our customers the way we receive them, thus if a client sends us data with driver names in it, we would reach out to them and inform them of potential data breaches on their site. Normally they ask us to delete these data points [...]* (see interview C.5) [4]. Data leaks related to requirements INFR.1 must constantly be checked to comply with GDPR guidelines. The expert continued by stating that they also deleted this information immediately without the customer telling them to do so. Moreover, they have integrated structures in their system that would detect the potential data breach and delete the information once they have received the data. Thus, the sensitive information would never be uploaded to the database, and the potential risk is limited (see interview C.5) [4].

5.6.4. Stakeholder requirements evaluation

Both functional stakeholder requirements, SFR.1 & SFR.2, were approved by the evaluation partners. A carrier expressed their position towards requirement SFR.1 (restricted access to the system) as follows: *"We accept a direct connection to our database if it is clearly regulated which data will be accessed and what data will be made available to the shipper, however, only in a very small scope, as we do not see the added value of sharing data beyond the fuel consumption of our fleet."* (see interview C.3) [3]. The statement makes it clear that not every stakeholder is willing to share all their data without a clear reason. Thus, the system should have restricted access. SFR.2 was concerned that each vehicle must be able to store (temporary) data on the vehicle, thus guaranteeing multiple data sources. The IT expert highlighted especially that point: *"For me, the importance of multiple data sources and telematics system is very important for such a system from a developer point of view "* (see interview C.5) [4].

SNFR.1 was indirectly approved through IFR.1, as it is related to security and the GDPR. The compliance of the collected data with ISO and GLEC (SNFR.2) was already approved by the statement of the LSP about the importance of these accreditations for their customer (see interviews C.3 and C.4) [3, 5], and the comment of the IT expert about the importance of multiple data sources from various platform also directly confirmed requirement SNFR.3 (interoperability between platforms) [4]. The flexibility towards user requests (SNFR.4) is a highly stakeholder-specific requirement. The IT expert already mentioned this, and it was also apparent during the evaluation. During the interviews to establish the requirements, the fuel consumption specialist highlighted the importance of frequent and detailed data to create accurate results (thus a high ping rate) (see interview A.3) [11]. During the evaluation interviews, a carrier said the following: *"For us real-time monitoring of the data is not important, the same is the case for the system latency"* (see interview C.3) [3]. Thus, directly contradicting the statement. The LSP also mentioned that they need their GHG emissions reporting only quarterly and not more, but both agreed that this could be highly dependent on the role [5]. The usability of the system (SNFR.5) was a main point for logistical actors to select an external IT company with the calculations of their GHG emissions. The LSP approved the requirement with one simple statement *"We hired an external company to worry less about the IT, thus we want an easy system"* (see interview C.4) [5]. Finally, SNFR.6 needed some additional explanation for the LSP, who then agreed with the requirement.

5.6.5. Technical requirements evaluation

The technical requirements were closely related to each other, especially TFR.1 (the system can incorporate multiple data sources), TNFR.1 (scalability of the system), and TNFR.2 (translatability of various structures and languages). TFR.1 and TNFR.1 were directly approved through Oliveira et al. (2019) (thus literature), who identified the interoperability and scalability as general non-functional requirements for an IoT middleware model [102]. This was also supported by comments of the experts about interoperability and multiple data sources (see interview C.5) [4]. TNFR.2 is a direct prerequisite for TFR.1, thus also approved. Lastly, a carrier highlighted the importance of consistently transmitted data, as especially broken sensors can lead to incomplete datasets.

5.6.6. Results of the evaluation

Based on the evaluation, the following steps were taken and improved:

1. The derived requirements were kept. However, the scope and the stakeholder-specific characteristics for it must be emphasised

Table 5.5 summarises the results. The first column is the representative requirement-ID, the second column presents the requirement, and the third column indicates the evaluation status of the requirement. A checkmark indicates that the requirement was "approved", meaning that the requirement is correct. The citations [6, 3, 5, 4] after a checkmark indicate that the requirement was directly mentioned and approved by the experts and an "L" followed by a citation indicates that the requirement was evaluated and approved by a literature source.

Table 5.5: Expert requirement evaluation

Requirement ID	Requirement	Approved
IFR.1	Conformity with European Data Privacy legislations	✓, [3, 5, 4]
IFR.2	Conformity with European vehicle reporting & monitoring legislations	✓
IFR.3	Conformity with European emission factor measurement legislations	✓, [5]
INFR.1	No tracking of driver specific data	✓
SFR.1	Access to the system is restricted	✓, [3, 4]
SFR.2	Can receive, store and process data	✓
SNFR.1	The system and data storage is secure	✓,[4]
SNFR.2	The collected data is compliant with ISO14083 & GLEC 2.0	✓,[3, 5]
SNFR.3	The system supports interoperability between platforms	✓,[3, 5]
SNFR.4	Flexibility per user request	✓,[3]
SNFR.5	Ease of use for users	✓, [3, 5], L[102]
SNFR.6	The system can portray GHG emissions savings and compare fuel types	✓
TFR. 1	The system can incorporate other data sources	✓,[4], L[102]
TNFR.1	Scalability of system	✓, [6] , L[102]
TNFR.2	Translatability of various structures and languages	✓,[4]
TNFR.3	The system transmits complete and consistent data	✓, [3]

5.7. Conclusion

This chapter gives the answer to the second subquestion *"What are the technical requirements for onboard sensor systems in road freight logistics?"* The first conclusion was the approval of the clustering of the system requirements into institutional, stakeholder, and technical groups by the experts. Moreover, all system requirements were approved during the evaluation interviews with the experts and the available literature. However, one of the most significant insights was the stakeholder-specific character of the requirements. Different stakeholders in the system have different priorities and thus prioritize certain requirements higher compared to other stakeholders. One example of that is the importance of real-time monitoring and the reporting frequency of the sensors. While IT experts and fuel consumption specialists prefer more detailed reporting and data, logistical operators only require the data every quarter. This was an important insight, and it will impact the design of the system in the next chapter, as these general requirements will lead to a more general and conceptual system architecture. However, the next chapter will use all requirements for the system's design.

6

Design & Development

The chapter will address the third sub-question *"How does a potential onboard sensor system look like to gather primary data in road logistics?"* by proposing various design options to fulfil the created system requirements of the last two chapters. Section 6.1 will first outline how interviews were used for the process of designing the system architecture, which is followed by the explanation of the used approach to do that in section 6.2. Section 6.3 will provide an overview of general abstract systems architectures already adapted to logistics and transport. Section 6.4 will cover the design by first outlining the different design options to fulfil the requirements in subsections 6.4.1 to 6.4.9. Section 6.5 will summarise the results and visualise four main aspects of the final system architecture in subsections 6.5.1 to 6.5.4. Section 6.6 will address the fourth sub-question *"What is, according to industry experts, the quality of the designed artefacts and requirements?* and will explain how the evaluation process of the derived design options and designs was conducted and will then present the insights of the evaluation interviews in section 6.7 with a list of actions that followed the evaluation process. Based on the evaluated design components section 6.8 will present an abstract stakeholder-specific system architecture (the evaluation will follow in the next chapter). Finally, section 6.9 will conclude the chapter.

6.1. Interview Usage

The conducted interviews during the earlier stage of the design process (mainly in the socio-technical analysis in chapter 4 and during the finalisation of the requirements in chapter 5) were an indirect knowledge source for this final design process. No additional interviews were conducted for the sole purpose of the creation of the system architecture. However, the obtained knowledge during the earlier interviews was vital for the understanding, focus and design of the system, thus was a constant indirect knowledge input during this chapter and research sub-question.

6.2. Creation of design options and components

The final design process followed a five-step approach, of which the first step already happened in chapter 5 (the establishment of the design principles and requirements). Fig. 6.1 visualizes the process. The stakeholder analysis, the socio-technical environment and the conducted expert interviews led to the creation of six functional and ten non-functional requirements, categorized into institutional, stakeholder and technical requirements. Four design principles were established, and the various system requirements were allocated to them (see Fig. 5.1).

Based on the established design principles, five larger design topics were established to which the systems requirements were allocated (see the design topics and the related requirements to each design topic in Table 6.1). The requirements were allocated based on the closeness to the design topic, i.e. the design topic of *„Data Privacy“* had the requirements IFR.1 and INFR.1 allocated, as both covered the conformity with European data privacy legislations (IFR.1) and the tracking of driver-specific parameters (INFR.1).

For each design topic, various design options were researched that could fulfil the design topic and, thus, the corresponding system requirements. The design options represent the design space of the design and provide an overview of possible tools for the design. Multiple design options will be researched per design topic. Thus, they recommend viable tools that can be used in the design process.

After the viable tools have been established per design topic (and thus as well for the system requirement), core components for the design of the final system will be established. They are the building blocks of the complete system architecture and incorporate individual or multiple identified design options. The core components were based on the most important aspects and requirements for the system determined during the interview process. Finally, the stakeholder-specific system architecture that incorporates the core components and some design options will be created.

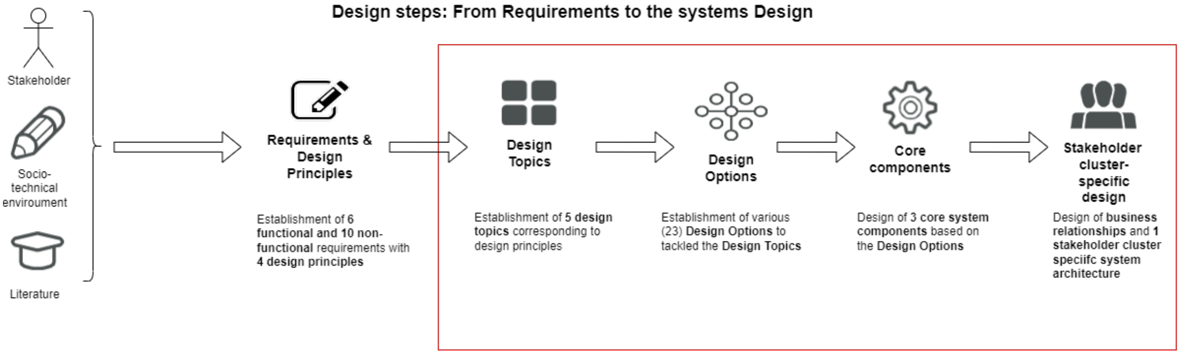

Figure 6.1: Design Process: from the initial requirements to the final stakeholder-cluster specific system architecture

6.3. IoT & Onboard system architecture overview

A basic onboard/ IoT sensor architecture at its core has three main layers that make up the infrastructure, which has been discussed and applied to many different use cases such as smart industries (see Fig. 6.2 for the basic structure of the system architecture) [107, 59], fleet management systems [81], or digital twins [2]. The lowest layer is the perception layer, which enables and supports data collection at the physical infrastructure via sensors and accumulators [59]. It is the backbone of the architecture, as it is the only connection between the physical and digital layers. At this layer, only minor data storage is present.

The second layer, often called the transportation or middleware layer, is connected to the perception layer via communication networks such as WiFi, 4G/5G or satellite internet and is responsible for access to the system [59]. The layer transmits the collected data from the physical infrastructure to the cloud or databases for other actors to access the collected data. Thus, the three main tasks are transmission, storage and the connection to the physical layer to access and use the collected data [107, 59]. Finally, the last layer is the application layer, which is visible to the final user of the system, and it provides all requests of the final customer of the system [59].

This three-layer structure was also followed for this artefact design. Fig. 6.2 visualises the basic system architecture, already with adaptation for the road logistics and onboard sensory systems. The perception layer starts with the physical layer at the vehicle level with various physical sensors that collect data (here, four types of sensors were selected: IoT sensors, Onboard diagnostics systems (OBDS), rFMS/FMS sensors, and telematics)[1]. These data points are stored in the vehicle and transmitted to the middleware layer via different communication networks (4G/5G, WiFi, or satellite internet). The main storing and processing happen in this layer. This storing and processing can happen at the fleet operator, OEM, or other external databases for road logistics. The middleware layer is the connection between the perception and application layer, at which the customer accesses the (processed) data via visualisations and dashboards. The following system design will follow this structure.

[1] this research will not cover advanced onboard processing and calculation methods, such as edge computing

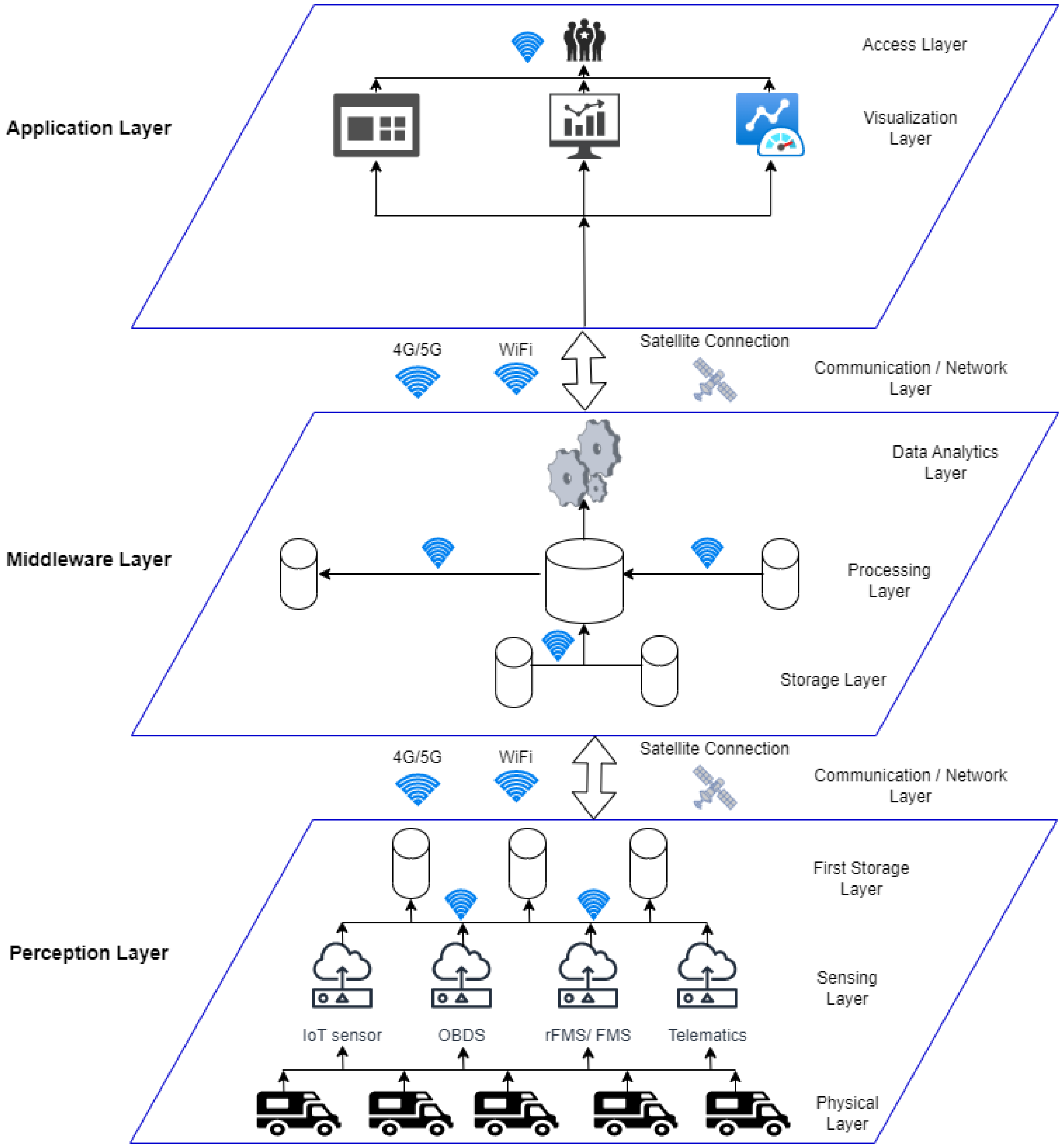

Figure 6.2: Basic system architecture for IoT and onboard sensor systems in road logistics. The structure was adopted from [107, p. 1], and visualisation and modifications for road logistics were done independently created.

6.4. Architecture Design

This section will portray the main system design in line with the design approach described in section 6.2 and Fig. 6.1. Each requirement was allocated to a larger design topic on which the design should focus. Five design topics were established based on the design principles: Data privacy, data parameters, access & security, user, and interoperability. Table 6.1 summarises the requirements allocation to each design topic. Each design topic was again split into smaller areas to facilitate the design process and cover all aspects of each larger design topic. Fig. 6.3 presents the separation per design topic. Data privacy was split into four areas that all contribute toward data privacy (encryption, restricted access, contractual agreements, and internal data privacy), data parameters into two (data sources and data enrichment), access & security into three (access and control, external security, and internal data privacy), user was split into three aspects (connections, tools, and system architecture) and finally interoperability was split into two parts (connections and processing). In the first design step, various design options per design area (i.e. data privacy via encryption) were researched as possible options to fulfil the targeted requirements for the system. Selected design options recommendations will be used and implemented in the exemplary component design in section 6.4 and in the system architecture for an IT company in section 6.8.1.

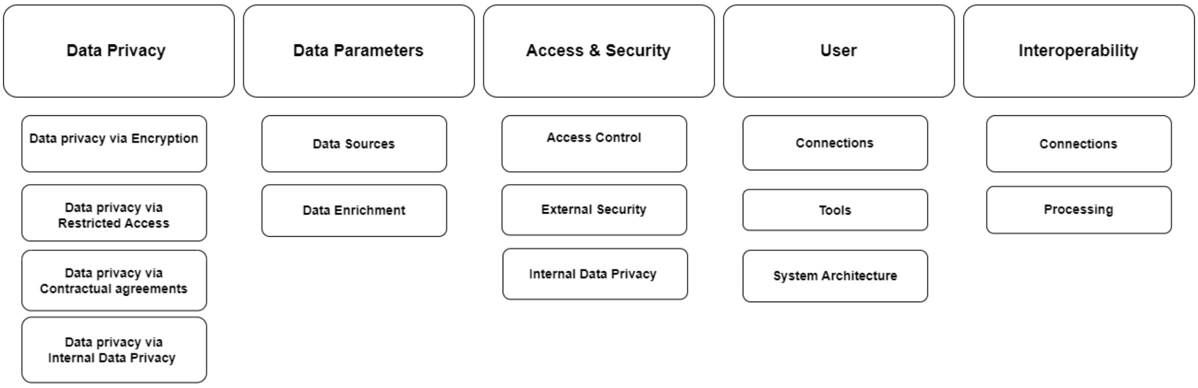

Figure 6.3: Identified Design topics (5), each with the sub-categories per design topic to facilitate the design process.

Table 6.1: Allocation of derived requirements to larger design topics

Design Topic	Corresponding functional requirements	Corresponding non-functional requirements
Data Privacy	IFR.1	INFR.1
Data Parameters	IFR.2 IFR.3	SNFR.2 TNFR.3
Access & Security	SFR.1	SNFR.1
User	SFR.2	SNFR.4 SNFR.5 SNFR.6 TNFR.1
Interoperability	TFR.1	SNFR.3 TNFR.2 TNFR.3

6.4.1. Data Privacy

Data privacy is a fundamental topic for the system architecture design and one of the main design requirements from, on the one hand, stakeholders and customers, but on the other as well, from a policy perspective. There are many design options to achieve this data privacy conformity. The design topic of data privacy was separated into four design areas: Data privacy via encryption, restricted access, contractual agreements, and company internal data privacy measures.

Data privacy via encryption

One of the most known and used data privacy measures is the usage of encryption for the data and all related communications. One possible option is TLS encryption (Transport Layer Security), which can be used to increase the security of all communication [101]. Other forms to ensure data privacy are the direct connection to databases via remote VPS or private tunnel database access to limit the exposure when accessing and transferring the data [123]. Lastly, an indirect tool to encrypt the data is via data aggregation. Thus, the aggregated data cannot be traced back to the raw data and, therefore, also protect the individual or the company behind the raw data (this is especially possible when presenting GHG emissions on an entire company level, thus no leg-specific GHG emissions can be derived from it).

Data privacy via restricted access

Data privacy is also heavily increased by proper access management to the system and to the data itself. Thus, the first step for data privacy is already happening while onboarding new data sources [7, 11]. A unique link is sent to the potential new data source, which allows the access/ registration option to only the intended party. The setup also requires the creation of login details for later authentification in the system and to track users' activities in the system [7]. This method is already used in the industry in combination with contractual agreements to establish data usage consent, as explained by a data market provider [7]. The access to the entire system can be further restricted by predefining access of that specific user to specific datasets or tools. The chosen general system architecture can support access management to increase data privacy. One possibility is Multi-tenant database architecture. Multi-tenant database architectures can separate different customer databases, tools, and services into unique sub-digital environments under the umbrella of the main digital infrastructure. Thus the various customers' (or tenants') architectures are logically separated [13]. Other advantages are the system's scalability (a more detailed discussion about the multi-tenant system architectures will follow in section 6.5.2).

Data privacy via contractual agreements

Another option for achieving data privacy is with contractual agreements between all parties about the exact actions that can be done with the data. These contracts can be highly customised to fit the needs of each stakeholder. They can elaborate on topics, such as with whom which data can be shared for which reasons, the duration of data storage or access restrictions for other parties [7]. Such agreements, of course, do not ensure data privacy from a technical standpoint but ensure the responsibility of each party, which can be enforced if breached. Closely related to general data contracts are consent forms. Here the explicit consent of the data owner to the data processor/ user is critical. Thus, only if consent is given to use, store, process or share the data actions are allowed. This means it is also possible to withdraw consent in hindsight (also in line with GDPR Art. 7 [53]). The contractual agreements are often part of the earlier mentioned onboarding process of new data sources.

Internal Data Privacy

Finally, each data handling company is obliged to follow internal "built-in data privacy". This can be accomplished via training and education sessions for all company employees to properly handle data and be aware of the most common security risks, such as fishing emails or data handling procedures (such as certificates of DataGuard [25]). Thus, it is possible to train the employees of an organisation to be sensitive about data privacy, as the human component is one of the main attack points for intruders. Another tool is to make the digital infrastructure more secure against attacks and potential data leaks by appointing Information System Security Manager (ISSM) [16]. An official internal data privacy standard is the ISO 27001 norm (Information security management systems), which offers companies of all sizes and industries guidance for creating safe systems [76]. Moreover, article 37 of the GDPR also mandates all companies that handle EU citizens' sensitive data to appoint a data protection officer [53]. Another option is to use external services compliant with existing data privacy norms (especially for databases and storage). Here cloud systems that are located in Europe (and thus have to be in line with European legislation, such as the GDPR) are a tool to achieve that. The data never leaves the EU, and there must be security the entire time. Options for that are Microsoft Azure cloud systems or the equivalent of Amazon Web Services (AWS)[2]. To further strengthen security, only transport-related data should be stored in the databases, thus no personal or financial information.

6.4.2. The Implication of the Findings for data privacy

Data privacy is affected by many design areas, which all provide different design options to address it. The analysis provided twelve different recommendations to ensure the fulfilment of the related system requirements and related system privacy concerns. The options were established from different literature sources and insights from the expert interviews. Especially the presented multi-tenant system architecture has great potential to increase the data privacy in the system, thus will be elaborated on in more detail in section 6.5.2.

6.4.3. Data parameters

Again, the requirements related to the larger design topic of data parameters were separated into two main design options: the data sources themselves and the potential data enrichment processes.

Data sources

The foundation of GHG emissions quantification and calculation methodologies is primary data of various vehicle parameters and their shipments [75, 112]. Therefore, multiple data sources and collection systems must be used and integrated to access the data to comply with existing methodologies, regulations, and requirements. Data sources that are compatible with existing emissions quantification methodologies are the aforementioned rFMS/FMS systems that are operated by the vehicle and truck OEMs, external IoT sensor companies that equip and manufacture different sensor systems for vehicles and can collect the needed data, OBDS (i.e. ondometer) and vehicle telematics systems which can also be used as the foundation for emissions calculations. Another advantage of these larger sensor systems is the common format [108]. Thus, reduced pre-processing is needed if a company receives only one kind of data (i.e. only primary data via rFMS or only IoT sensor data).

Data Enrichment

The possibility of data enrichment is highly important due to three main reasons. Firstly, it bridges the gap between various reporting standards (rFMS vs IoT sensor vs OBDG vs telematics) by harmonising all data formats to the

[2] a more detailed analysis of the European cloud servers will follow

same level of detail and granularity if multiple data sources and standards are used for a single calculation or reporting [11]. Having your own data resources is necessary to enrich the data or access other external larger primary databases. A second reason for the importance of data enrichment is the potential for malfunctions of the sensors. During an interview, an LSP expert mentioned that the malfunctions of sensors were main reasons for incomplete datasets and technical challenges that occasionally happen [9]. Thus, sensors would report only partly complete data or, in the worst-case no data for specific legs. Thus, data enrichment was needed to create consistency and counter potential sensor or system failures at the physical layer [8]. An extensive primary database (or at least access to one) can be a massive asset for external GHG emissions calculation companies to strengthen their services against possible data inconsistency treats [11]. Lastly, enriching data also enables more advanced data analysis and features such as digital twins of entire vehicles or driving behaviour analytics [1].

6.4.4. The implications of the Findings for data parameters

Data parameters are the data foundation of the system. Two main conclusions were made based on the analysis. First, the system must be able to incorporate many primary data sources into the system, and second, data enrichment is a design option that increases the system's robustness. The system should be able to incorporate external primary databases next to the primary data sources. The design option of data enrichment via a primary external database was visualised in section 6.5.3 in combination with the design topic of interoperability and data pulling.

6.4.5. Access & Security

The design topic of access & security was separated into three main design areas: access control, external, and internal security protocols.

Access control

Access control to the system ensures controlled and restricted access but also has advantages for the system's security. Subsection 6.4.1 already mentioned principles that are relevant for the data privacy of the system and security and access control. The explained onboarding procedure of new customers via unique links and authentification via customised login credentials are tools to ensure restricted access [7]. Multi-factor authentification can further strengthen access control to reduce the risk of unauthorised access. Another known concept is called "Know your customers" (KYC), which can be used for the identity management and protection of users [80]. The design option can be combined with the aforementioned onboard procedure, login credentials, and logging of user activities in the system to increase the visibility of which user accessed which files and data.

External security protocols

The environment of onboard and IoT sensor systems in road logistics is influenced by a variety of stakeholders, as shown in section 4.3. Thus, no company or stakeholder can operate in isolation and relies on external security measures and protocols. One example is vehicle OEMs' security and access control protocols for their databases and systems. Thus, selecting partner companies with sufficient security and access protocols is highly important. All companies that operate and sell their products in the European market are also bound by the same security and data compliance regulations, thus having to ensure compliance.

Internal security

Finally, internal security is also important for access and security and is again closely related to the design options mentioned in section 6.4.1 and data privacy. The design option of a muti-tenant system architecture would restrict the access possibilities of different customers to data [13], the built-in privacy and security measures through ISSM, ISMS, data protection officer, and educational training further strengthen the security and access threats [16, 76, 53, 25]. Internal security measures are especially important for data processing companies [11, 7]. European databases and centres ensure data compliance with European standards for external cloud and processing purposes.

6.4.6. The implications of the Findings for access & security

Security & access is, similarly to data privacy, a design topic with many different perspectives and therefore offers many design options to address the related requirements. The main finding of the section was that a structured and customised onboarding process to the system could result in significantly improved security. This, in combination with contractual agreements and consent forms (see the section on data privacy), could be a main building block of the proposed system architecture and is already being used in the industry [7]. Thus, this customised onboarding process, combined with contractual agreements and consent forms, will be visualised and elaborated on in section 6.5.1 as a system component for the final system architecture design.

6.4.7. User

The design topic of user requests was separated into three main design areas: connection to the users and from the users to the external database, the provided tools and features of the system and the system architecture.

Connection

Connection to different systems and databases is one of the most important topics within this environment. This section will not focus on the security and data privacy concerns of connections (this was done earlier) but rather on how connections can increase data quality by accessing multiple external databases and by increasing the flexibility of the system towards customer requests by creating pull-APIs, which can be integrated into different company systems. Pull-APIs enable efficient resource management, as the accessing company is not forced to build an entire system around it [7].

Tools

The provided services and tools are directly related to the requests and needs of customers. During the interview, the logistical operator expressed the desire to have benchmarking/ comparison tools to either show their CO_2 reduction efforts compared with competitors or analyse different investment options [8]. These tools could enable data-based decision-making by supporting modal shifts based on CO_2 reduction targets (i.e. comparing CO_2 emissions of rail and road) or highlighting CO_2 reduction possibilities of new fuel types.

System architecture

Finally, user requests can also influence the system architecture itself. A modular system with multi-tenant database management increases flexibility and guarantees sufficient scalability opportunities in combination with external European cloud systems. Tsai et al. [120] have shown how multi-tenant databases in combination with horizontal (tenant level) partitioning are more cost-efficient and increase the reliability and availability of the system while ensuring linear scalability possibilities (compared to the vertical partitioning).

6.4.8. The implications of the Findings on user requests

User requests for the system are highly stakeholder-specific and thus hard to generalise for an abstract system architecture. However, a common theme that was recognised throughout the interviews with logistical actors was their desire to store their data in European data centres [8, 9]. Thus, a more detailed analysis of the European cloud infrastructure on the example of Microsoft Azure was conducted in section 6.5.4.

6.4.9. Interoperability

The design topic of interoperability is separated into two main design options: the connection possibilities to different databases and the processing of raw data.

Connection

The connection capabilities are not just vital for user requests but also for the interoperability of the system. The usage of pull APIs can significantly increase the interoperability of the system by including various databases and sources outside the own system. They offer connecting possibilities to various databases without creating completely new system designs. However, just a pull API is insufficient to increase the systems' interoperability, as the pulled data could still be in different formats. Thus the combination of sufficient connection possibilities and related pre-processing capabilities at the company is fundamental for interoperability.

Processing

Processing means in this context, the conversion of various data formats and standards from external sources to a consistent and uniform format, which then can be used by the company for further analysis. This consistency can be achieved by pre-processing pulled data from various databases to create a common company standard. An elaborated version of pre-processing is a collaboration between different companies or Joint Ventures to facilitate further the data exchange between databases and create common standards. Collaborations and Joint Ventures also facilitate the connections possibilities and reduce the efforts for direct pull-APIs.

6.4.10. Implications of the Findings of interoperability

The core message of the design topic of interoperability was the ability of a company to connect to as many data sources and databases with the least effort for data alignment afterwards as possible. Pull APIs take a major role in this procedure. A Joint Venture is a special form of collaboration, as this offers (after the initial setup) the least effort for receiving and pulling data from the data source to the IT company [4]. Thus, the conclusion of the section was to elaborate on a possible Joint Venture setup of an IT company with an OEM for the final system architecture design in section 6.5.3 [7].

6.5. Results

The presented design options for the system architecture design have presented a wide range of design choices to ensure the feasibility of the design based on the gathered requirements. Fig. 6.4 presents the discussed design options per design topic and each sub-category. However, as described in section 4.3, the sector and its stakeholders have highly individual requests and needs. Thus, generalisations of design options are only possible until a certain degree of detail. Therefore, four main design options were discussed in more detail (see the implications of the findings per design topic). The first one relates to the access & security aspects and was presented via the customised onboarding process in section 6.5.1. Second, the data privacy design topic was elaborated on with a more detailed description of the multi-tenant system structure in section 6.5.2. Third, the interoperability and data parameters design topics were elaborated on in more detail by the example of data pulling in section 6.5.3. Lastly, section 6.5.4 concludes the section with an example of the European cloud infrastructure of Microsoft Azure.

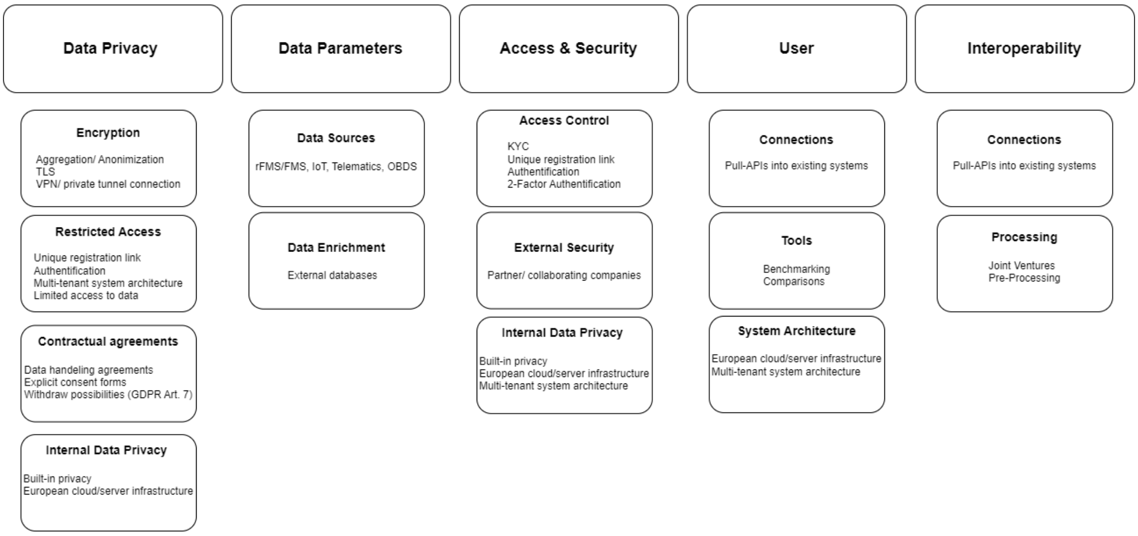

Figure 6.4: Five main design topics for the system architecture of onboard sensor systems in road logistics (top layer). Subcategories with proposed tools and concepts to ensure feasible system architecture based on requirements analysis.

6.5.1. Access & Security: The Registration/ Onboarding Process

The onboarding and registration process is a vital part of the proposed system as it sets the scene for data privacy and access & security concerns. Fig. 6.5 visualises the entire onboarding procedure with all relevant steps. This onboarding process is needed if new shippers, carriers, LSPs or subcarriers want to start a partnership or use the services of external companies for either data-sharing practices or GHG emissions calculations.

The onboarding process is stakeholder specific, however, during interviews, common steps were identified and were elaborated on [7]. The process includes five different stakeholders (shipper, external company, known carrier, new carrier, and OEM /3PL) who play an active or indirect role. The first step is the contact of the shipper to the external company to use their services. The shipper outsources their transport activities and thus does not have the primary data the external company needs. Thus, the shipper sends a list of used carriers and LSPs to the external company. The external company checks the list to identify if they already have contracts with the used carriers and LSPs. The onboarding process is quicker if all carriers and LSPs are known and have worked with the external company. The external company can directly reach out to the carriers/ LSPs to receive their consent to access the data for the shipper, and the data flow can start. Here it is company specific if a new consent form and a contract are needed or if the existing ones can be used. Some companies have internal and GDPR-related guidelines to have a new contract with each carrier for each shipper. If the carriers and LSPs are known, the onboarding and registration process can happen within hours and up to 2 days (depending on the response speed of the carrier) [7].

The process is more elaborate if the shipper uses new carriers or LSPs [7]. Now the external company cannot reach out to the new carriers and LSPs directly but has to create the first contact via the shipper (to conform with data privacy regulations and GDPR guidelines) [7]. Thus, they send a unique registration link to the shipper, who then forwards it to the new carrier. The registration link has two purposes. First, it expresses the desire of the shipper that the external company is accessing their data; second, it gives the external company the consent of the carrier or LSP to access and use their primary transportation data for the transport operations of that specific shipper (thus both design topics of access & security and data privacy were addressed). The new carrier confirms (or denies) the requests and consent to the shipper and uses the unique link. After the confirmation of the new carrier, the external company can directly contact them to determine if they use external transportation services or OEM data telematics services [7]. Thus, the new carrier sends a list of used carriers and OEMs to the external company, which the external company checks if they have worked with them before. If all OEMs, sub-carriers, or 3PLs are known, the onboarding process is done, and an API connection can be established for the data transfer. If the new carrier also uses unknown service providers, their consent must also be established before an API connection can be established. The onboarding process of new carriers and LSPs can take up to one week, depending on complexity and response frequency [7].

What becomes clear with this example is that data privacy and consent are vital parts of these external companies. This elaborate process is firstly directly increasing the data privacy of all parties (as the shipper also has no direct access to the carrier's data, but only to the final processed ones) and secondly, that access to the system is only possible via a unique link which increases the security. The derived onboarding sequence will be an important part of the system architecture.

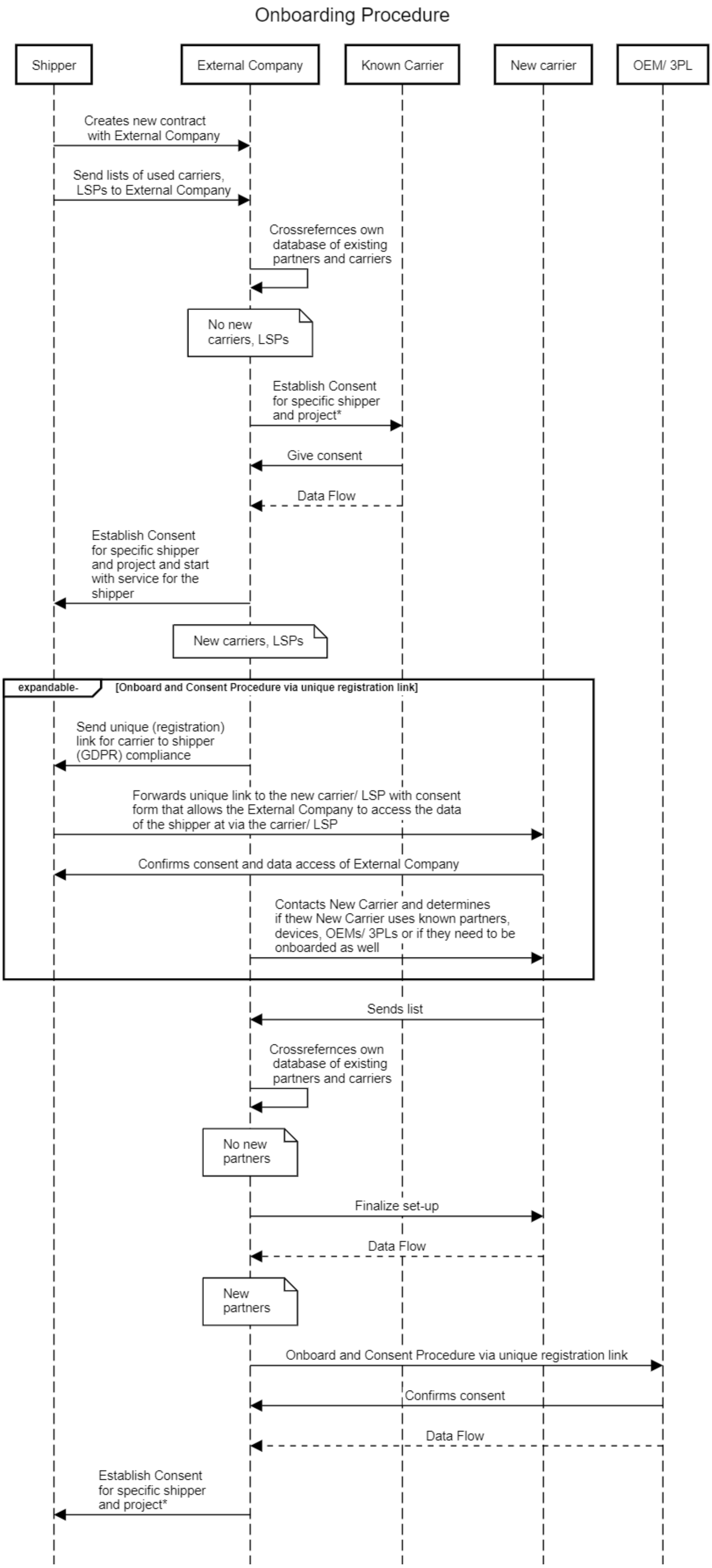

Figure 6.5: Abstract sequence diagram of the onboarding process of known or new carriers/ LSPs with an external IT company.

6.5.2. Data Privacy: The Multi-tenant system architecture

The idea behind multi-tenant system architectures is the separation and isolation of data from various users and the restriction of their access to the data of the other users. Fig. 6.6 displays what such a system could look like. Every tenant has only access to their data, and the risk of unwanted data exposure and leaks is limited [120]. It is commonly used for Customer Relationship Management (CRM) and has next to access restrictions the benefits of improved machine utilisation via load balancing [120]. Table 6.2 compares different multi-tenant database architectures with regard to scalability, isolation, costs, performance, development complexity and operational complexity. It is important to remember the related system requirements to decide which system architecture structure fits best for the envisioned design. The architecture should support data privacy and manage access to the system (and data) while ensuring the scalability and flexibility of the system (see Table 5.4). The standalone system architecture (the second column in Table 6.2) excels in privacy (tenant isolation) but lacks scalability possibilities and is expensive. The shared multi-tenant system architecture (the fourth column in Table 6.2) offers the best scalability possibilities at the lowest costs but offers low tenant isolation and security. Lastly, the database per-tenant system architecture (the third column in Table 6.2) combines sufficient scalability options at low costs with high privacy, thus combining the best of both words. Therefore, it was decided to continue with the database per-tenant system architecture.

Table 6.2: Comparison of various database configurations, retrieved from [95]

Measurement	**Standalone app**	**Database-per-tenant**	**Sharded multi-tenant**
Scale	Medium 1-100s	Very high 1-100,000s	Unlimited 1-1,000,000s
Tenant isolation	Very high	High	Low; except for any single tenant (that is alone in an MT db).
Database cost per tenant	High; is sized for peaks.	Low; pools used.	Lowest, for small tenants in MT DBs.
Performance monitoring and management	Per-tenant only	Aggregate + per-tenant	Aggregate; although is per-tenant only for singles.
Development complexity	Low	Low	Medium; due to sharding.
Operational complexity	Low-High. Individually simple, complex at scale.	Low-Medium. Patterns address complexity at scale.	Low-High. Individual tenant management is complex.

Fig. 6.6 visualises a potential representation of such a system with three different tenants. Each tenant accesses the system via the same interface (app). The individual login of the tenant only grants access to the individual database with the corresponding visualisations and dashboards (thus every tenant has its own database). They do not have any access to the interfaces and data of the other tenants and can only access their own. This builds on the prerequisite that the system host correctly sets up the appropriate access rights (here, user and access management is key, and mistakes can be major challenges. See informal conversation I.15 in appendix B).

The flexibility and advantages of the proposed multi-tenant system architecture with databases per tenant yielded to use it as a general system architecture for the later system design.

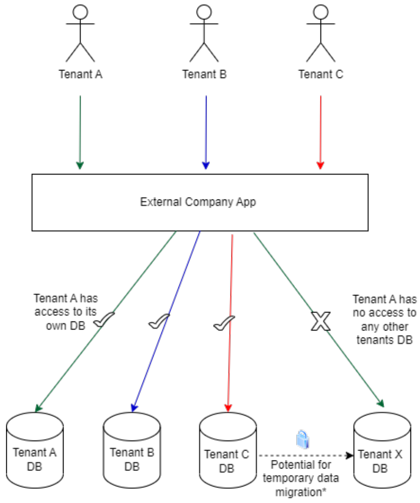

Figure 6.6: Abstract representation of a multi-tenant system architecture with databases per tenant, allowing each tenant access to only the own database. Access to other databases is restricted.

6.5.3. Interoperability: Data Pulling via Pull-APIs

The data-pulling processes via APIs are vital for external companies to receive the data of their customers (or their carriers) and specific user requests. Fig. 6.7 provides a graphical overview of the data-pulling process of external company with an OEM after establishing a Joint Venture. The onboarding & registration process, described in section 6.5.1, and with it, the consent forms is a prerequisite for the data-pulling via APIs. It starts with the shipper, who outsources their transportation activities to a carrier directly or via an LSP, which can outsource the transportation service to a fleet operator [7]. The fleet operator owns a fleet of vehicles purchased from an OEM and subscribed to the OEM-specific rFMS system, providing them with various data points and parameters. The shipper also contracts an external company with the GHG emissions calculations, which requires the primary data [7]. Thus, the external company reaches out (via the shipper) to both the fleet operator and the OEM to access the data for the specific shipper (same procedure as explained in section 6.5.1).

Once all administrative tasks have been established, the actual data-pulling can start. The vehicles of the fleet operator constantly collects data on the journey via the rFMS system and stores it in the vehicle[3]. The collected data is batched and stored until it is pulled or pushed to the OEM and their database (the frequency of the pulling is OEM-specific but usually is between 5 to 30 minutes) [7]. The data arrives nearly instantly at the OEM and their database. Here it is being processed, and parameters such as fuel consumption rates or emitted CO_2 are calculated for each vehicle. This processing happens in minutes [7]. Once the data is processed, it's stored in the OEM database. The external company directly accesses the OEM database for the vehicle-specific primary data (vehicle identification numbers or licence plates can be searched per vehicle) [7]. This can happen as the external company has the aforenoted consent of both the shipper and the fleet operator to access the data. The access happens via a pull API and creates a connection between the OEM database and the database of the external company. Once the external company has accessed the OEM data, it can further process the data (depending on the shipper's requests) and use it for its own visualisations and tools. The services of the external company determine the processing time, which can vary from milliseconds to days [4] [7].

The described procedure showcases how different data sources and bases can increase interoperability via direct API connections in combination with pre-processing. It also increases the system's flexibility as the external company can react to specific customer requests or data sources for their calculations. The direct pull APIs are a resource-efficient tool for external companies, as they require fewer development resources than complete new structures (see interview A.5). The derived connection procedure will be implemented into the system architecture design.

[3]also directly related to requirement SFR.2
[4]Normally, more complicated processes are done, which increases the time considerably

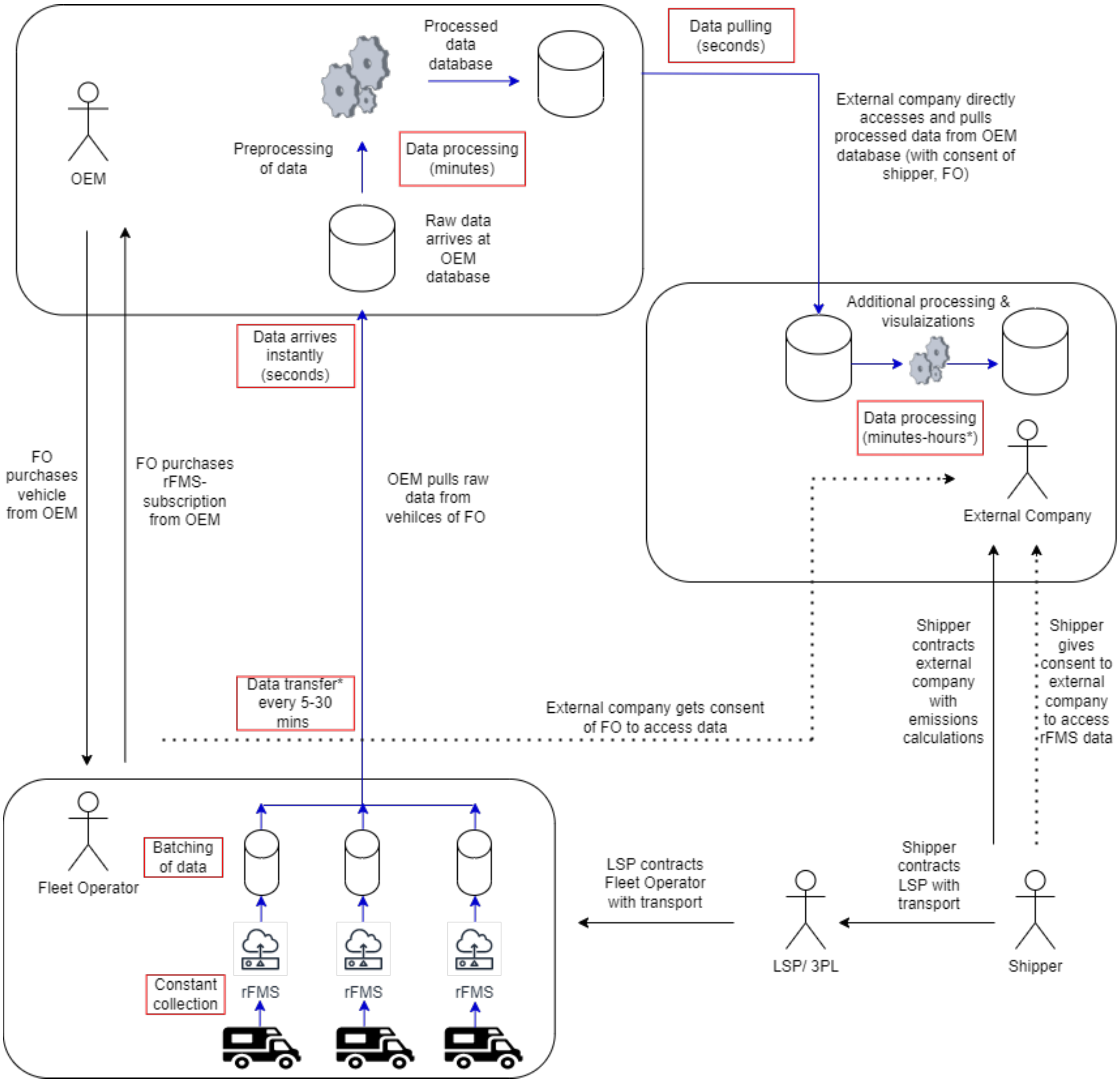

Figure 6.7: Abstract representation of data pulling of an external company with OEM and fleet operator via APIs

6.5.4. User: The European Cloud Infrastructure

External cloud and server infrastructure are vital for many IT-based companies, as they offer cheap and easy scalability possibilities based on subscription models without the need for capital investment by the company itself. Large companies such as Microsoft, with their Azure cloud system or Amazon, with their AWS product, offer many server locations worldwide with many locations in Europe (see Fig. 6.8). These server centres must comply with European legislation and country-specific legislation depending on the server's location (i.e., the server in Ireland has different regional compliance measures than the one in the Netherlands, see Table 6.3). This is useful for the companies, as they can trust that their stored data is sufficiently protected under global standards such as the ISO 20000 (standards for IT service management), but also European standards such as the GDPR or regional standards such as the "BIR 2012" (Dutch-specific standard, which outlines that companies that operate in the Dutch government sector must demonstrate compliance with the Baseline Informatiebeveiliging Rijksdienst standard (BIR 2012) [92]) or even industry-specific standards such as the NEN7510 (Dutch industry standard about data privacy in the healthcare industry). Customers can specifically select the option to solely use the European data centre to eliminate the risk that the data is taken outside of Europe and the corresponding data privacy legislation [93]. Thus, these systems are highly adaptable to the different needs of external companies and are an important point for many shippers, as they want their data not to leave the EU and its regulations [9]. The design option of European cloud centres is selected for the system architecture design.

Figure 6.8: Locations of data centres of Microsoft Azure globally, retrieved from [94]

Table 6.3: Data compliance and data storage description of Microsoft Azures data centres in Ireland (left) and the Netherlands (right), retrieved from [93]

Regions	North Europe Start for free >	West Europe Start for free >
LOCATION	Ireland	Netherlands
YEAR OPENED	2009	2010
AVAILABILITY ZONES PRESENCE	Available with 3 zones	Available with 3 zones
Compliance ^	**Global Compliance** CIS Benchmark, CSA STAR Attestation, CSA STAR Certification, CSA STAR Self-Assessment, ISO 20000, ISO 22301, ISO 27001, ISO 27017, ISO 27018, ISO 27701, ISO 9001, SOC 1, SOC 2, SOC 3, WCAG 2.0 **Regional/Country Compliance** EN 301 549, ENISA IAF, Standard Contractual Clauses, GDPR **Industry Compliance** EBA, CDSA, GxP, PCI DSS, Shared Assessments, TruSight Azure compliance offerings	**Global Compliance** CIS Benchmark, CSA STAR Attestation, CSA STAR Certification, CSA STAR Self-Assessment, ISO 20000, ISO 22301, ISO 27001, ISO 27017, ISO 27018, ISO 27701, ISO 9001, SOC 1, SOC 2, SOC 3, WCAG 2.0 **Regional/Country Compliance** BIR 2012, EN 301 549, ENISA IAF, Standard Contractual Clauses, GDPR **Industry Compliance** AFM/DNB, NEN 7510, EBA, CDSA, GxP, PCI DSS, Shared Assessments, TruSight Azure compliance offerings
DATA RESIDENCY	Stored at rest in Europe	Stored at rest in Europe

6.6. Interview Usage for the evaluation

Another important aspect of the design is the evaluation of made assumptions, the design options and the different core components. For that, expert evaluations were conducted in a semi-structured format. The participating experts received the to-be-evaluated results and a consent form in advance and were asked to evaluate the results based on correctness and completeness (see evaluation interview structure and asked questions to the experts in Appendix C.1). The design options and their potential to fulfil the design topics (and thus the requirements) were evaluated in an interview with a *Carrier* (see interview summary in section C.3) [3], an *LSP* (see interview summary in section C.4) [5], and an *External IT company* (see interview summary in section C.5) [4]. The created core components that are based on the design options were evaluated based on the same correctness and completeness principles and experts (see evaluation interview structure and asked questions to the experts in Appendix C.1).

6.7. Evaluation of Design Option & Results

The third evaluation round was conducted for the proposed design options, followed by the evaluation of the system architecture components. Section 6.7.1 covers the design options' evaluation, done with expert interviews and literature. The final result of the evaluation round was a list of all derived design requirements and the proposed design options contributing to their fulfilment. Table 6.4 summarises the evaluation efforts. Afterwards, sections 6.7.2 til section 6.7.4 contain the evaluation of the components of the developed system derived from the design options. The system components were solely evaluated via expert interviews, and the feedback was directly implemented. Each component evaluation has a detailed list of changes that were done to it based on the evaluation interviews. The evaluation of the created artefacts will contribute directly to the fourth sub-question by assessing the design's current quality and, secondly, improving the derived components and design options for the system architecture. Fig. 6.9 summarises the evaluation steps of the design efforts.

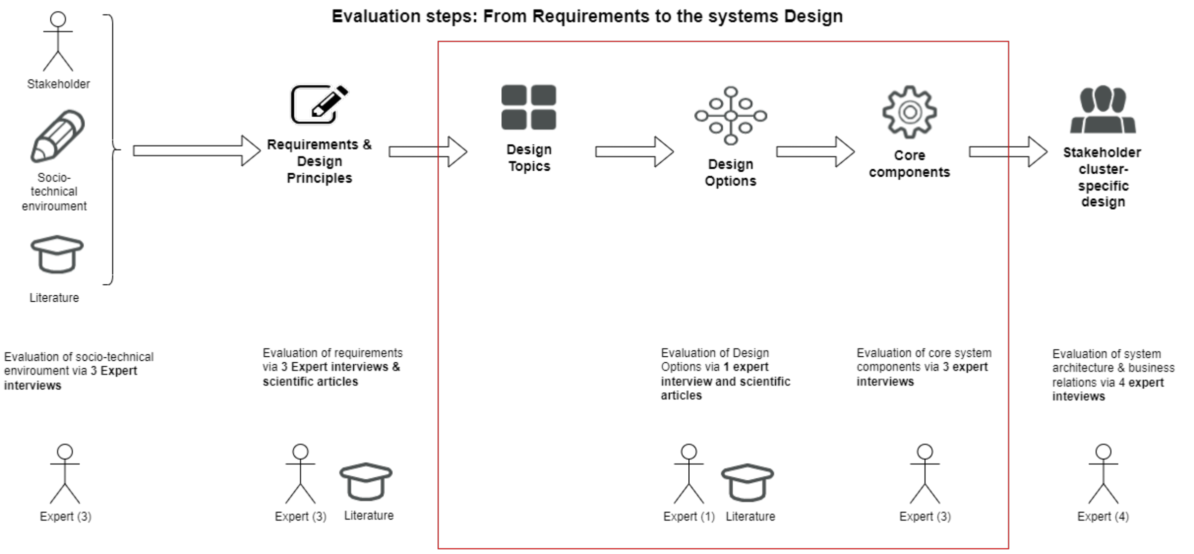

Figure 6.9: Evaluation process of the design options and core components with all the intermediate steps.

6.7.1. Evaluation of Design Options

The result of the evaluation of the recommended design options is presented in Table 6.4. The system requirements themselves were already approved earlier. Thus this evaluation is solely about if the proposed design options can fulfil the design topics and, thus the system requirements. The evaluation was done with an IT expert and literature. The first important comment of the IT expert was the clarification that: *„In my opinion, it would be good to explain that these presented tools and concepts are a way to reach the goal of the design topic, but not all tools and concepts have to be used for it. It might be sufficient to only use one“* (see evaluation interview C.5) [4], which was then added to make it clear. Secondly, the IT expert already had a clear perspective on the usage of multiple data sources and how these data sources, on the one hand, allow ISO compliance (SNFR.2) but also make sure to conform with the European vehicle reporting and emission standards (IFR.2 and IFR.3) *For me, the importance of multiple data sources and telematics system is very important for such a system from a developer point of view* (see evaluation interview C.5) [4]. Multiple data sources also ensure consistency of the data (TNFR.3), as gaps can easier be filled.

When asked about the importance of the European cloud centres, the IT expert commented: *„The location of the servers has cost considerations (non-European are cheaper), but the European Server has a faster response time compared to servers in i.e. China. For us, the topic is important but not a daily discussion point“* (see evaluation interview C.5) [4]. Thus, the design option of European data centres is a valid option for data security (IRF.1) and access & security (SNFR.1), however, from a developer standpoint, it also contributes to a quicker but more expensive system. When asked about data privacy and aggregation & anonymisation, the expert said the following *"We present the data to our customers the way we receive them, thus if a client sends us data with driver names in it, we would reach out to them and inform them of potential data breaches on their site. Normally they ask us to delete these data points, however, the concepts of Aggregation & Anonymisation are used the least when talking about data privacy, for that we use tools, such as the mentioned TLS“* (see evaluation interview C.5)

[4]. The statement resulted in two conclusions. First, data breaches can sometimes be out of the control of the IT company. Thus the requirement INFR.1 can not be entirely fulfilled by the proposed design options (and will be elaborated on in the following evaluation chapter). Secondly, the security and isolated databases per tenant are even more critical to ensure the damage is limited if a data breach happens. Therefore the IT expert directly continued by stating that all their customer have their own isolated environment, which ensures proper security and restriction to the system (SFR.1) (see evaluation interview C.5) [4].

The design topic of interoperability and user requests was mainly addressed via pull API calls. The expert confirmed this and mentioned how these increase the flexibility for them and, thus, indirectly for the user (thus validating the requirements SNFR.3, SNFR.4) [4]. Moreover, pull APIs ensure that various data sources (TFR.1) can be incorporated.

The design options of an advanced form of collaboration (Joint Ventures) and 2-factor authentification were exciting to the IT expert, and both were seen as valid options to fulfil SFR.1 and SFR.2, but currently not used by them (see evaluation interview C.5) [4]. Joint Ventures also allow easier translatability of different data formats (TNFR.2), as they can directly access the data and can make agreements with the other parties about common standards [4].

Finally, both the ease of use for the customer (SNFR.5) and the potential features of the system (SNFR.6) are highly stakeholder and IT company-specific. Thus, it should be evaluated on a case-by-case basis and will be elaborated on in the next chapter.

Based on the evaluation, the following steps were taken:

1. The list of recommended design options was kept
2. Addition that not all design options are needed to fulfil the requirement
3. INFR.1 can not be fully fulfilled as it is often out of control [5]
4. SNFR.5 and SNFR.6 are stakeholder-specific and must be evaluated on a case-by-case basis [6]

Table 6.4 summarises the evaluation process of the design options. The first column presents the requirement ID, and the second column the corresponding requirement. The third column presents all proposed design options to fulfil the specific requirement. The last column indicates if the requirement can be fulfilled based on the option of the evaluation experts or literature. A checkmark indicates the fulfilment of the requirement with the proposed design option, and a checkmark in brackets indicates the partial fulfilment of the requirement. The reasons for the particle fulfilment can be the stakeholder specific-character of the requirement, which makes it not possible to generalise the fulfilment or that the requirement is out of control. However, all proposed design options are, in theory, able to fulfil the requirement. The provided sources after the check marks indicate the expert interviews which approved the fulfilment of the requirement and an ”L” followed by a source indicates the evaluation with a literature source.

[5] Implications will be discussed in chapter 7
[6] Implications will be discussed in chapter 7

Table 6.4: Result of the evaluation of the design options for the fulfilment of the design requirements

Requirement ID	Requirement	Proposed concepts	Fulfilment
IFR.1	Conformity with European Data Privacy legislation's	Data handling agreements Explicit consent Customer onboarding process European Cloud Infrastructure Built-in-privacy	✓, [4]
IFR.2	Conformity with European vehicle reporting & monitoring legislation's	rFMS/FMS, Telematics, OBDS, IoT External databases	✓
IFR.3	Conformity with European emission factor measurement legislation's	rFMS/FMS, Telematics, OBDS, IoT External databases	✓, [5]
INFR.1	No tracking of driver specific data	Aggregation & Annomization Built-in-privacy	(✓), [4]
SFR.1	Access to the system is restricted	Onboarding process KYC Authentication (2-Factor)	✓, [4]
SFR.2	Can receive, store and process data	Joint Ventures Collaborations	✓, [4]
SNFR.1	The system and data storage is secure	Multi-tenant system architecture, European Infrastructure, Built-in-security, Onboarding process & Authentication, TLS, VPN &private tunnel connections	✓, [4]
SNFR.2	The collected data is compliant with ISO14083 & GLEC 2.0	rFMS/FMS, Telematics, OBDS, IoT External databases, Collaborations	✓, [3, 5]
SNFR.3	The system supports interoperability between platforms	Pull-APIs, Joint Ventures, Pre-processing, External data sources	✓
SNFR.4	Flexibility per user request	Multi-tenant system architecture, Pull-APIs, Tools	✓, L [21]
SNFR.5	Ease of use for users	KYC, Multi-tenant system architecture	(✓), L [21]
SNFR.6	The system can portray GHG emissions savings and compare fuel types	Benchmarking, Comparison	(✓)
TFR. 1	The system can incorporate other data sources	Pull-APIs, Joint Ventures, Pre-processing, External data sources	✓, [4] , L [21]
TNFR.1	Scalability of system	Multi-tenant system architecture, European Cloud Infrastructure	✓
TNFR.2	Translatability of various structures and languages	Pull-APIs, Joint Ventures, Pre-processing, External data sources	✓, [4]
TNFR.3	The system transmits complete and consistent data	rFMS/FMS, Telematics, OBDS, IoT External databases, Collaboration	✓

6.7.2. Evaluation of Accessibility & Security

Customer onboarding (see Fig. 6.5) is a crucial point of the system, as it impacts many areas and requirements, such as IRF.1, SFR.1, SNFR.1 and the main design topics of data privacy and security. At the core of the process are the multiple contractual agreements for data access and usage between the external IT company, the shipper and the carrier. Both the carrier and the LSP agreed with the correctness of the process (see evaluation interviews C.3 and C.4) [3, 5]. The LSP highlighted their role as the first contact person to the carrier to ensure them that no additional costs will occur to them, and they always have the option to withdraw their consent at any time (see evaluation interview C.4) [5]. The external IT company also verified the process and emphasised the importance of getting consent from each involved party for every single collaboration (see evaluation interview C.5) [4]. Thus, they would still need it again even if they received the consent form from all parties in earlier projects.

The following improvement was made based on the evaluation:

1. Indicate that each new project needs new consent (currently, this is only the case for completely new carriers or shippers)

6.7.3. Evaluation of Data Privacy

The Multi-tenant system architecture is a tool which is commonly used for access and security topics and for data privacy by separating the various databases of each user (related requirements SNFR.1, SNFR.4, SNFR.5, and TNFR.1). The IT expert validated both the usage and effectiveness of this method for proper security and data privacy (see evaluation interview C.5) [4]. The IT expert emphasised the security of the multi-tenant architecture with the following example *"If data migration from the database of the carrier to the shipper happens, this data will only be stored temporarily, and neither party would have access to both data sets [...]"* (see evaluation interview C.5) [4]. Thus, such a system structure would be a valid tool for fulfilling the corresponding requirements.

The following improvement was made based on the evaluation:

1. Indicate that there is the option of temporary and secure data migration between the different databases to match the order with transport data

6.7.4. Evaluation of Interoperability

Data pulling via pull API connections are a fundamental tool for the enrichment and data exchange of the various players in the system (related requirements are: SNFR.3, SNFR.4, TFR.1, TNFR.2) and are also indirectly needed for all data transfers of the various sensors and standards from the vehicle level to the next layer of storage (related requirements: SNFR.2, TNFR.3). Nearly all involved stakeholder mentioned the usage and the benefits of data pulling in some way. However, they are crucial for IT companies. The expert concluded that the presented design options (see Fig. 6.7) are a realistic representation. However, two comments were made. First, it is also possible that the vehicle itself pushes the data to the OEM or other sensor companies, thus not just a pull from the OEM (see evaluation interview C.5) [4]. This has the benefit of bridging connection issues. Thus, if the vehicle has no connection at a certain moment and the OEM tries to pull the data at that time, problems and backlogs could occur. Thus, a push from the vehicle could solve it [4]. Second, the presented times for the processing at the OEM and IT company are highly dependent on the specific service, and it is more likely that the processing times at the IT company are in the minutes to hours (see evaluation interview C.5) [4].

The following improvements were made based on the evaluation:

1. Add the data pushing of the vehicles to the OEM by adding the term "data transfer"
2. Adapt the processing times at the level of the IT company

6.8. Stakeholder-specific Architectures

After evaluating the design options and proposed system architecture components, a system architecture architecture was created. Fig. 6.11 presents the derived system architecture. It was built upon the previous results and was created in collaboration and through the knowledge obtained from expert interviews within the sector of the previous chapter [11, 7]. Every stakeholder has particular needs and demands for their architecture, thus an IT company was chosen as an example, as they have one of the most complex architectures due to their intermediate position between shippers, carriers and LSPs and the various requirements for data compliance and sharing regulations. Additionally, not every IT company will have the same needs, thus deviations from the presented architecture are possible. First, the relationships and business connections between the main logistical actors, the OEMs and the external company were discussed and presented (see Fig. 6.10), which was followed by the architecture. Both designs were evaluated in the next chapter to check their correctness and completeness.

6.8.1. External IT-companies

The proposed business relationships were developed based on the expert interview (see A.5) and represent an ideal multi-stakeholder collaboration between four stakeholders[7] (shippers, external IT companies, fleet operators, and OEMs). The shipper pays a fleet operator (this can also happen indirectly through an LSPs) to move its goods. The fleet operator receives compensation for the movement of goods. The shipper also contracts an external IT company to calculate GHG emissions and gets validated scope 3 emissions values of their transport in return. The fleet operator is an asset owner and has vehicles from one OEM[8], with the extension of data tracking via the rFMS

[7] the overview only includes business relations and not compliance relation with the European Union or other institutions

[8] in reality, fleet operators typically have vehicles from multiple OEMs, this scenario will be presented in the final system architecture representation

system. The rFMS system collects data about the vehicle and the journey and sends it back to the OEM (see the detailed description of the pulling process in section 6.5.3). The external company has a contractual agreement with the shipper, the OEM and the fleet operator to access the shipper's data at the OEM from the carrier (see the explicit onboarding & consent procedure in the onboarding process in section 6.5.1).

The OEMs are currently only monetising around 7% of the rFMS subscription, however, they have created a Joint Venture with the external IT company to access the cleaned and processed data (see interview A.5) [7]. The Joint Venture also resulted in easier access for the external IT company. They created a "no-log-in solution" in which the external IT company can access the OEM database directly and search specific vehicle parameters via the vehicle identification number and the licence plate (the fleet operator provides this information). Thus, the shipper contracts and pays the external IT company and the fleet operator to calculate emissions and movement of goods, respectively. The fleet operator pays the OEM for the vehicle and potentially the rFMS system. The external IT company pays the OEM for the data access, which again then offers the data in combination with the data handling processes to the shipper (the OEM also pays its suppliers etc. However, this is out of the scope in this system). The actors created an environment where each stakeholder profits from the services and is dependent on the other actors. There is also the possibility of a connection between the external IT company, the fleet operator, and the shipper with the OEM[9]. The process was visualised in Fig. 6.10.

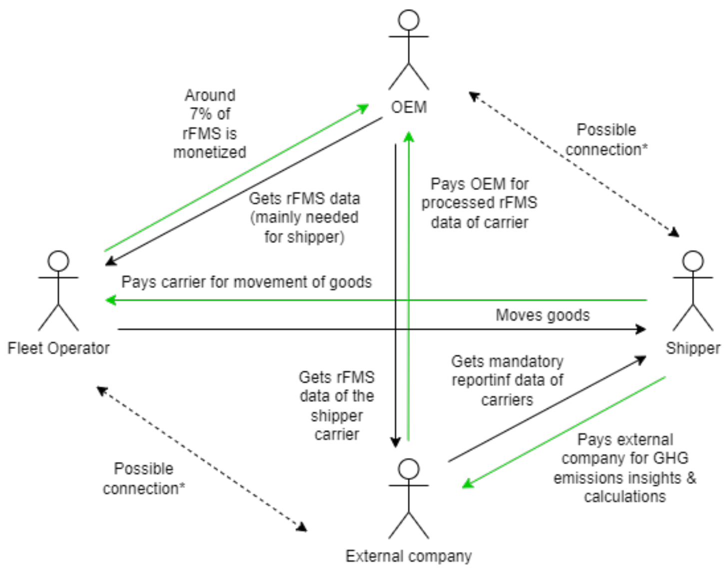

Figure 6.10: Business relationship between shipper, OEM, fleet operator and external company for the data collection, creation, processing and usage

The final system architecture design for an external IT company was created based on the expert interviews, the derived requirements for the system and available literature on IoT and onboard sensor systems architectures [11, 10, 1]. The aforementioned design options, such as the onboarding process, data pulling, encryption or the multi-tenant system architecture, were included in the design and will not be mentioned again explicitly. Again, this is a general representation of the system architecture, and every company has their own requirements and needs. However, the representation visualises a general stakeholder-cluster-specific design.

The design starts with the previous business relationships (see Fig. 6.10), in which the shipper contracts a fleet operator through an LSP to move its goods and pays them a fee. Moreover, they contract an external IT company to validate and calculate their GHG emissions (scope 3) due to corporate values and reporting obligations [9]. The fleet operator has a heterogeneous fleet of multiple vehicles of various brands and different methods to collect the data of the vehicles. A part of that fleet is from an OEM with the rFMS extension, to which the external IT company has a direct link to access the shipper's data at the OEM from the fleet operator. The other three data sources and onboard sensor systems are the OBDS system, the default telematics system and another external IoT sensor system to which the external fleet operator has access. The external IT company created detailed and contractual agreements with all relevant stakeholders (OEM, shipper, fleet operator, LSP, and potentially with the external IoT sensor company) to access the shipper's data at each level without sharing the raw data with the shipper. Thus they act as an independent middle platform for data sharing and processing without the direct sharing between the fleet operator and the shipper.

Each vehicle collects data via its respective onboard system, which is temporarily stored at the vehicle (indicated by the small cylinder at the vehicle level) and transmitted to either the OEM, the IoT sensor provider or the

[9] added after the expert evaluation

database of the fleet operator. The collected parameters and the reporting frequency depend on the service of the sensor provider, the OEM and the age of the vehicle and can be different per vehicle. The connections from the vehicle to the respective database one layer higher (i.e., to the OEM or sensor manufacturer) are protected and not openly accessible.

Depending on the contracts of the external IT company, they have access to the data at different times. With the OEM, the IT company can either directly access the data after the collection process on the vehicle or once the data has been processed and cleaned by the OEM (as described in the data pulling process, see section 6.5.3, the same is the case for the IoT sensor provider) (see interview A.6) [8]. The various contractual agreements and consent forms also enable the IT company to directly access the database of the fleet operator (or they receive a batch of data every pre-defined time) to have complete access to all transport and emissions data. There is also always the possibility that the fleet operator already shares parts of their data with other stakeholders (can be other shippers with whom they cooperate or regulatory parties such as other Member States (MS) or the European Commission (EC))[10].

Once all data is present at the external IT company, they start pre-processing to create consistent formats and frequencies. This can also happen via the enrichment of the data with other data sources, external services, or default emission values (the enrichment of the data is only necessary if the provided primary data is not complete or detailed enough). Suppose the data cleaning and processing is complete. In that case, the external IT company will create visualisations and dashboards for the shipper, based on their requirements, in an isolated digital environment, only for the shipper to access (the raw primary data is not accessible to the shipper). This process is visualised in Fig. 6.11 and represents one possible solution. One key takeaway is the intermediate position of the IT company to bring together the three main logistical players by acting as a neutral party in the middle, with whom everyone is sharing data, as they do not fear a competitive disadvantage (as they are not a logistical service provider).

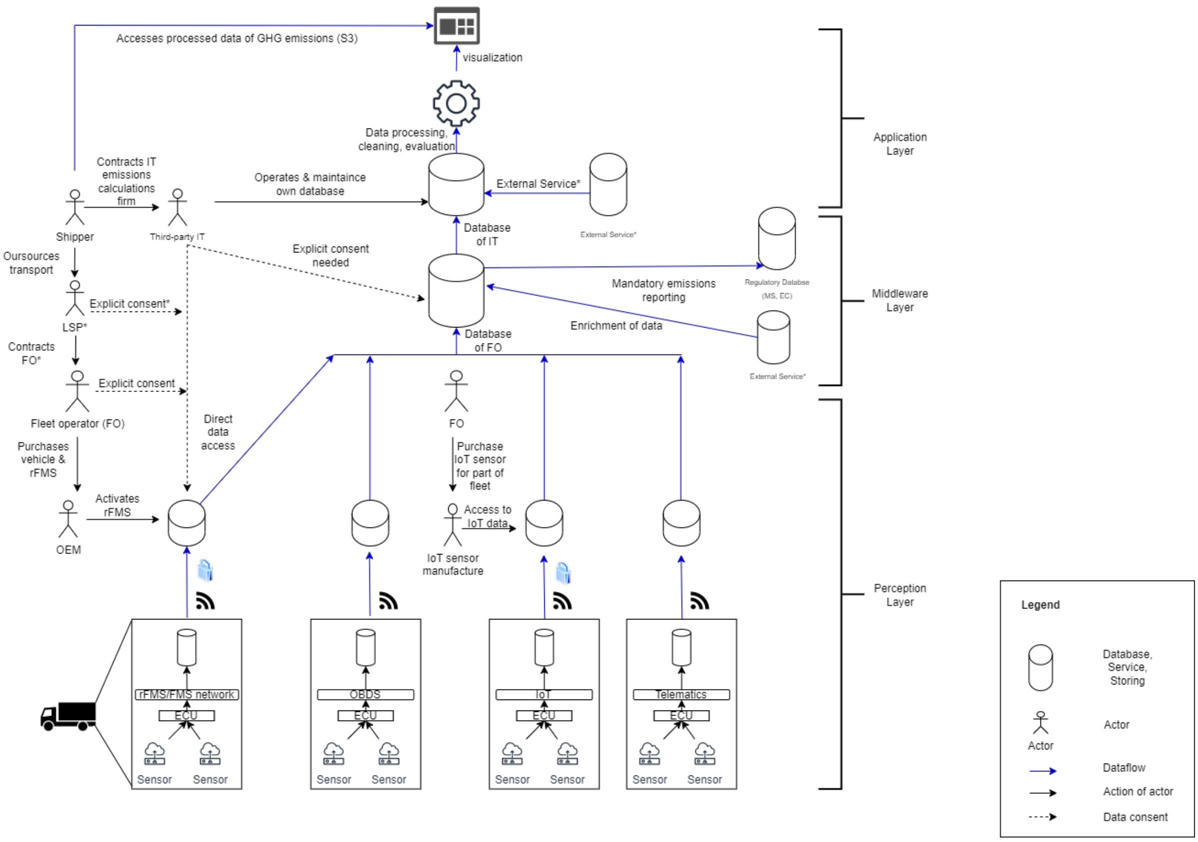

Figure 6.11: Abstract system architecture for onboard sensor system in road logistics for an external IT company.

6.9. Conclusion

The third subquestion *"How does a potential onboard sensor system look like to gather primary data in road logistics?"* was the design aspect of this research and built upon the design principles and system requirements

[10] The connection to institutions were not presented in Fig. 6.10, as these were compliance relations

that both SQ1 & SQ2 had established earlier. The design process had two parts. The first one was to group all gathered requirements into five broad design topics (Data Privacy, Data Parameters, Access & Security, User, and Interoperability), which were also closely related to the design principles. The design topics were very broad; thus, sub-topics of each design topic were established. Fig. 6.4 provided an overview with the splitting of each design topic and the established design options per category. Recurring design options were a unique onboarding process for new customers to IT data platforms, which firstly contributed to data privacy and GDPR-compliant data handling agreements and secondly to the access restrictions and security of the system. API connections from the IT companies to the various logistical actors and other external data sources ensured compliance with leading methodologies and made sure that enough data sources could be accessed while also facilitating the interoperability of the system.

The second part of the design phase was the creation of abstract and stakeholder-specific process components that take a central point in the system architecture. Visualised concepts were the unique onboarding process, the multi-tenant system architecture, the data pulling of external databases via pull APIs from the site of IT companies and the business relationships between four stakeholders for the data pulling. Evaluations of the designs happened after every step via expert interviews and scientific literature. Finally, an abstract system architecture from an IT perspective was created that incorporated the presented tools and concepts. Fig. 6.11 presents the final design output with dataflows and important stakeholder relations and actions. The main conclusions were the potential for various stakeholder-specific system architectures resulting from the different and stakeholder-specific requirements impacting the design process. The second conclusion was again the potential market position of external IT companies that can act as a bridge and data consolidation point between shippers, LSPs, carriers and OEMs. This important role of the IT companies was already hinted at in the stakeholder overview of the socio-technical analysis and also proved to be relevant in the design. The next chapter will continue with the evaluation process of the final system architecture and with a more elaborate discussion about possible limitations of the design and the impact of the only partly fulfilled system requirements INFR.1, SNFR.5, and SNFR.6.

7 Evaluation

This chapter covers the final evaluation round and the fourth sub-question *"What is, according to industry experts, the quality of the designed artefacts and requirements?"* by evaluating the stakeholder-specific business relationships and the system architectures. The previous chapters already evaluated the stakeholder overview in chapter 4, the derived system requirements in chapter 5, and the design options and derived system architecture components in chapter 6. The evaluation followed the mentioned evaluation process of Johannesson & Perjons (2021) [78]. Section 7.1 will summarise the evaluation strategies and steps for the final evaluation round, followed by the explanation of how the interview contributed to it in section 7.2. Section 7.3 will present the final evaluation of the system architecture and the business relationships. Section 7.4 will expand the limitations of the designs and the evaluation process. Moreover, potential path-dependencies of the system will be discussed in section 7.5 and the design's success factors and complexity considerations will be elaborated on in section 7.6. Finally, section 7.7 will conclude the section by providing a conclusion.

7.1. General evaluation strategies and approach

The evaluation process, which can be seen as the validation of the design and requirement outputs, followed evaluation steps outlined in Johannesson & Perjons (2021) [78] and utilised three different methods to do so. The methods were semi-structured evaluation interviews with different stakeholders (see Appendix C for the exact question structure), informal conversations with various stakeholders to parts or components of the design (see Appendix B for the complete list of informal talks, that influenced the research), and lastly literature. The expert evaluation interview with a data scientist and engineer specialising in emissions calculations with primary data in road logistics, data fetching, efficient processing systems architectures and databases needs to be highlighted for its contribution to all created requirements and stakeholder overviews, as well as all created design artefacts.

Due to the diversity of the topic, it was expected that not every stakeholder could evaluate every part of the research. Thus, a shipper would have most of the knowledge in the field of stakeholder management and potentially the important requirements for their business rather than the exact details of the entire system architecture. Therefore only directly important remarks were used.

The evaluation process followed a formative evaluation, meaning that the stakeholder's feedback was already implemented into the designs and created concepts to improve them. However, parts that were altered because of it were highlighted[1], and the feedback was clearly communicated (see Appendix D.1 for the designs of the first iteration) [78]. The exact evaluation strategy followed an ex-ante and naturalistic approach. The artefact was a typical socio-technical artefact with many involved stakeholders. Thus, the naturalistic approach was chosen [78]. The artefact was not deployed yet; thus the ex-ante approach was chosen, and real users and stakeholders were used for the evaluation [78]. The evaluation methods (interviews, informal talks, and literature) also suit the combination of naturalistic and ex-ante evaluation strategy [78].

7.2. Interview Usage for the evaluation

The structure of the evaluation interviews can be found in Appendix C.1, and the corresponding evaluation summaries are in Appendix C.2 til Appendix C.5. The final evaluation process for the business relations and the stakeholder-specific system architecture was evaluated by all experts. Again the interview evaluation partners

[1] all changes in the designs or requirements were indicated with a "*"

received the designs upfront with the consent form and were asked to evaluate the design based on correctness and completeness.

Four interview were conducted with a *Data marketplace expert* (see evaluation interview in section C.2) [6], a *Carrier* (see evaluation interview in section C.3) [3], an *LSP* (see evaluation interview in section C.4) [5], and an *External IT company* (see evaluation interview in section C.5) [4]. The business relations were only evaluated by the *External IT company* and the *LSP*, as they had detailed knowledge. Fig. 7.1 visualises the evaluation process with the inputs.

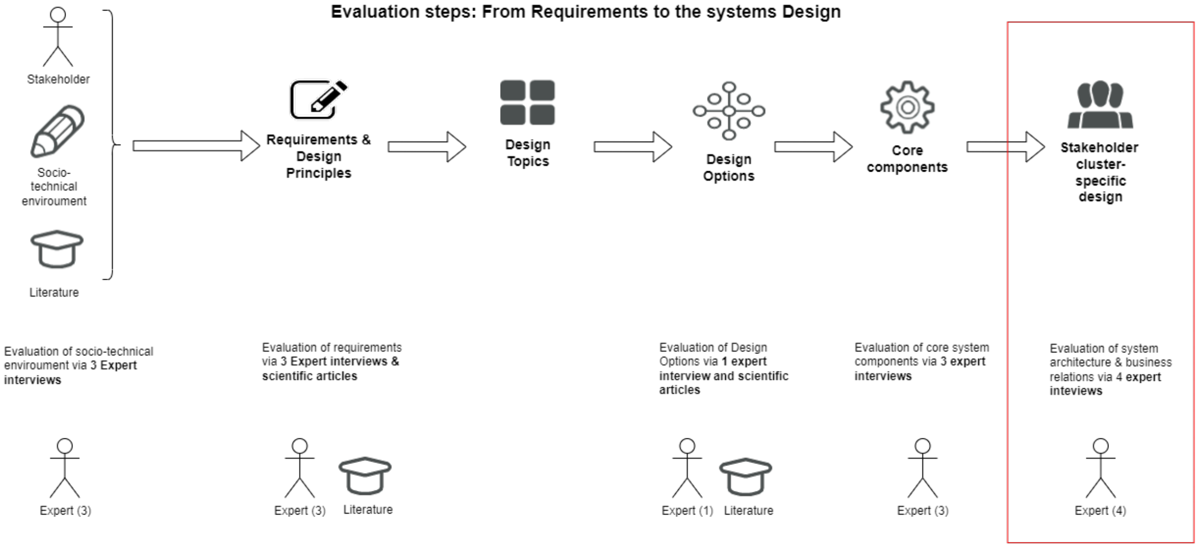

Figure 7.1: Evaluation of the final stakeholder-cluster specific system architecture.

7.3. Evaluation of System architecture

7.3.1. Business relations evaluation

Before evaluating the final stakeholder-specific system architecture, the business relationships that govern the final system architecture were evaluated. The possible business relationships between the different actors were presented to both an LSP and an external IT company. Both agreed that it represents a realistic visualization of a possible business relationship between the multiple parties (see evaluation interviews in Appendix C.4 and C.5) [5, 4]. The external IT company added that there are possible connections between the external fleet operator and between the shipper and the OEM depending on the nature of collaboration [4].

The following improvement was made based on the evaluation:

1. Addition of possible connections between shippers & OEM and external IT company & fleet operator.

7.3.2. System architecture of an IT company evaluation

The stakeholder-specific system architecture represents the final design artefact, which combines and incorporates all components of the onboarding process, the multi-tenant system architecture, the data pulling and possible business relationships. All interview partners approved the design and concluded that it represents a realistic and abstract design of such a system with correct data flows and relationships between the different stakeholders [6, 3, 5, 4]. An LSP mentioned that there is also the possibility of another layer of complexity in the contracting chain (see evaluation interview in Appendix C.4) [5], which was also supported by the data marketplace provider, who added that there is usually an LSP between the shipper and the fleet operator (see evaluation interview in Appendix C.2) [6]. Lastly, the external IT company mentioned that enrichment via external databases could also happen with external services. Thus, they would never access an external database for the data but would also purchase an external service to enrich and support their GHG emissions calculations (see evaluation interview in Appendix C.5) [4].

The following improvements were made based on the evaluation:

1. Addition of LSP to the order chain
2. Addition of external services for the data enrichment and not just external databases

7.4. Limitations of the Design & Evaluation

Limitations to the designs are an important factor in this research due to their highly stakeholder-specific character. The goal was to design and investigate general and abstract systems and data architectures to investigate the potential of onboard sensor systems to improve the estimation of GHG emissions in road logistics while using a multi-domain and socio-technical perspective for the analysis. Each stakeholder within the sector has different needs; some are uniform, such as the data privacy and security aspects. However, some might be different or even contradict the desires of others. One example of that is the reporting frequency. Here the shippers and LSPs require less frequent reporting, while IT companies need as much detail and data as possible to improve their calculations and predictions.

What also became apparent in the evaluation phase was that three non-functional requirements were not completely fulfilled with the proposed design options:

1. INFR.1 = No tracking of driver-specific data
   (a) Design Option 1: Aggregation & Anonymization
   (b) Design Option 2: Built-in-privacy
2. SNFR.5 = Ease of use for users
   (a) Design Option 1: KYC
   (b) Design Option 2: Multi-tenant architecture
3. SNFR.6 = The system can portray GHG emissions savings and compare fuel types
   (a) Design Option 1: Benchmarking
   (b) Design Option 2: Comparison

The first non-functional requirement (INFR.1) is especially important due to its regulatory character and potential legal consequences. Under the GDPR, a data breach is defined as:

"a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed" [53, p. 34].

And personal data is defined as:

"personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person [53, p. 33].

In the case of such a data breach of personal data, the controller[2] is responsible to report the incident within 72 hours to the corresponding authorities [53]. Thus, the given example that led to the non-fulfilment of the requirement was a typical data breach as a driver's name was communicated. However, the design option recommendation of built-in privacy addressed the issue and immediately made the other parties aware. Additional structures were added to the system, that would immediately detect the data breach, delete the information autonomously and thus, would prevent the uploading of sensitive data to the database. The problem with data breaches is that the contracted IT company cannot eliminate them by themselves if they get the data transmitted like that. However, there are even more design options that would be able to eliminate them, such as a direct connection from the IT company to the database of the data source. The IT expert mentioned the following solution *"When clients send us data, they must understand what type of data we need, in which formats and what times. This issue can be bridged with a Joint Venture or direct access via pull APIs"* (see evaluation interview C.5) [4]. The internal procedures were able to reduce the exposure of the information. Thus, the requirement was partly fulfilled.

The other two partly-fulfilled requirements, SNFR.5 & SNFR.6, were not fully fulfilled due to their case-by-case characteristic. However, it is believed that the proposed design options would fulfil them. Various comparison tools are already available, and it needs to be evaluated with the customer to determine if these solutions are suitable. This is also in line with the comment of the data marketplace operator, who mentioned that a large part of their work is the expectation setting with customers about possibilities (see design interview A.5) [7]. Lastly, the usability of the system must be evaluated with the help of an existing system. Even if a customer company has lower technical skills, it is still possible to increase usability by offering training on the tools and system, which would fulfil the requirement.

The evaluation rounds also continuously demonstrated the issue of scoping for the designs and where to set the system boundaries. While a narrow scope might lead to the exclusion of important stakeholder and their requirements, a broad scope increases the complexity and stakeholder and needs that are not essential to the

[2] the natural party that is alone or jointly responsible for the purpose and means of processing the data

system. Thus, a balance needs to be found. This issue was mainly present in the stakeholder overview, where different interview partners had different opinions about what should be included. The system architecture design focused on the data flows and sources. Thus it was decided not to include OEMs and fuel providers directly in the stakeholder overview. However, the OEMs were included in the architecture overview as an external data source via the data pulling. This compromise is believed to still portray the system's essential actors and procedures without unnecessarily increasing the complexity. This was also proven in the evaluation interviews, where all evaluation partners agreed on the correctness of the abstract design.

Another limitation is the sensitivity of this system to external influences, especially from a policy and institutional aspect. Data privacy regulation changes or data classification adaptations would significantly influence the presented results. Stricter data privacy concerns could alter the onboarding process of new customers in all directions (easier or more difficult, depending on the change), a change in the GDPR of the vehicle and driver data classification could change the entire consent process, while European-founded collaboration projects and regional databases could eliminate the need for Joint Ventures or data sharing agreements. Another point was indicated by the emissions factor and modelling expert, who mentioned that the current discussions of GHG emission are mainly focused on the direct impacts and fuels. However, non-emission-related toxins, such as small particles of tires or breaks of the vehicle, or damage to the road surface, are not considered (see design interview A.4) [10]. These non-emission-related toxins are also not included in this system design or in the state-of-the-art GHG quantification methodologies.

Limitations are also connected to the evaluation procedure. This topic's complexity and multi-domain characteristics make it more challenging to find experts who can evaluate the entire design space of the artefact. Very few experts have expertise in the required technical domain of data management, knowledge about the details of GHG emissions calculations, and policy aspects. Thus, individual parts had to be evaluated by specific experts that know this single domain. Direct logistical actors were necessary for the stakeholder and requirements development, while more technical stakeholders were incredibly knowledgeable in data management and design options to fulfil the requirements. The expertise of the collaborating company was an immense help, as they combined the expertise of all three domains and acted as an overarching evaluation partner for all aspects of the design in addition to the other evaluation partners[3].

The experts themselves are another potential limitation, as they are bound by their desires and beliefs and their roles in this system. This is, of course, not intentional, however, it must be considered. The interview experts had various roles, thus reducing potential biases and single-perspective views in the evaluation and creation process. Lastly, the semi-structured interviews with predefined questions could also lead to the steering of the interview partners towards specific conclusions or opinions. Thus, each stakeholder was asked if they had additional comments or areas to add to the interviews and evaluation which had not been addressed yet.

7.5. Potential path dependencies of selected design options

Another potential limitation of the design choices and options is path dependencies that result directly from them. Path dependencies can arise in the system as past decisions impact the design options in the present. This can be due to high switching costs to different design options or customer lock-in [110]. Chapter 4.2.1 already touched upon this topic by mentioning the problem of legacy-IT systems in the logistics sector, which slowed down the digitalisation process and facilitated error-prone practises [66]. Thus, reflecting on potential path dependencies for the proposed designs is essential. Four different path dependencies were identified that could occur at different points in the design stage with different impacts on the system (see Table 7.1). The four types of path dependencies were separated into three different clusters. Physical path dependencies happen due to design choices at the perception layer at either the physical layer (the vehicles themselves) or the sensing layer (the sensors). Digital path dependencies occur mainly in the middleware layer. They mostly happen in the communication and processing layer and, in the provided example, are concerned with accessing and connecting various data sources and services (i.e., enriching the data). Lastly, institutional path dependencies are concerned with a potential lock-in due to design choices that occurred during the selection of the methodology used to calculate GHG emissions (here, the path dependencies impact the earlier decisions).

7.5.1. Physical path dependencies

A physical path dependency is based on a design choice for a physical component of the system, which influences parts or the entire design of the system. These design decisions happen in the first perception layer of the system. The first example is situated at the lowest layer of the potential design, the physical layer (see Fig. 6.2), and is concerned with the selection of vehicles (see the visualisation of the potential impact of the path dependency in Fig. 7.2). Different vehicles offer different OBDS/ telematics/rFMS possibilities to collect data, thus already predefining the next steps, such as storage, access or potential data enrichment.
Suppose an asset owner decides to purchase new vehicles that only offer standards OBDS and telematics sys-

[3] Of course, this overarching evaluation role can also be a limitation

tems and not rFMS (due to the age or brand of the vehicle). The influence of this design decision is upward to the middleware and parts of the application layer and impacts design topics such as the data parameters (via the vehicle selection itself and, with it, the available sensor systems), the interoperability (via possible API connections to external databases and services or to the database of the OEM), or access & security (via the potential onboarding of more data sources for the enrichment process if not enough data is available originating from the vehicle). If an asset owner opts for an rFMS vehicle, the data gets transmitted to the OEM database. Thus the first communication layer is to the OEM, and then a direct API connection could be used to re-access the data. This option requires no additional data enrichment (if granularity is sufficient), which would be required with less sophisticated telematics systems (which would need another connection to an external source). The design paths would combine again at the data analytics layer. Thus, the impact of the design options (data parameters) to opt for a certain vehicle starts in the perception layer and has an influence until the end of the middleware layer.

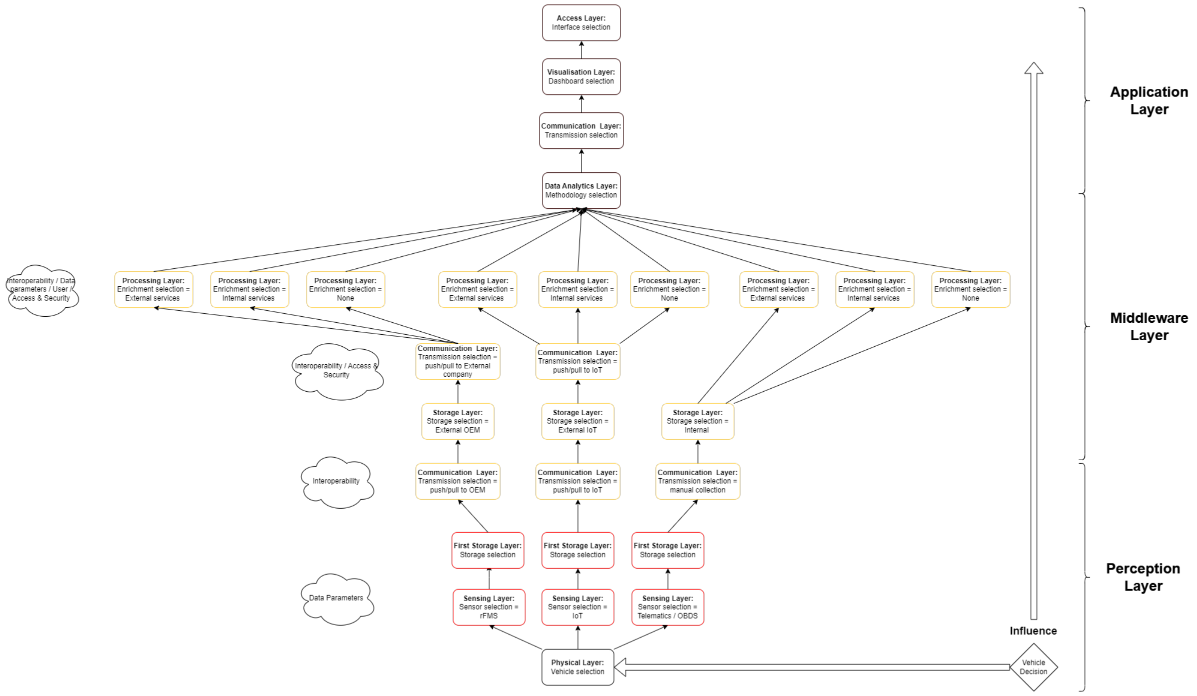

Figure 7.2: Potential path dependencies in the design of the system architecture, due to design option choices (vehicle selection, perception layer)

Similar path dependencies are also present when selecting sensors (see Fig. 7.3). Thus, the vehicles are already present, and the asset owner wants to make a data parameter decision to purchase additional sensors. Depending on the design options, various predefined paths can occur based on the decision for sensors. As described above for the vehicle selection, these can happen in the design topics of interoperability (if and how to connect to external databases to access the primary data), user requests and potential data enrichment procedures, and access and security considerations that are related to the access of other data sources. The influence of that design option is upward again.

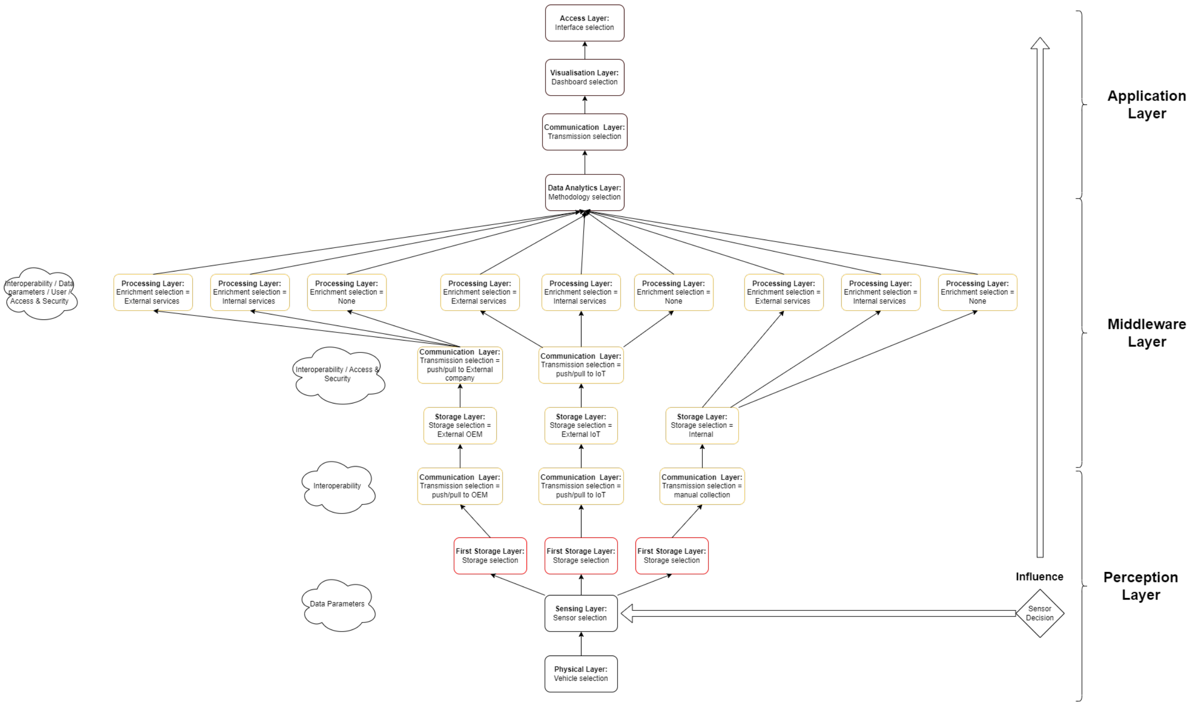

Figure 7.3: Potential path dependencies in the design of the system architecture, due to design option choices (sensor selection, perception layer)

7.5.2. Digital path dependencies

A digital path dependency caused by the design option can be related to the processing layer and potential connection possibilities needed for data enrichment via the different methods to connect to external services or databases. The design options of connecting via push/pull APIs to an existing service/data source or creating their own systems to which the data sources connect can result in different visualisation options. The flexible push/pull API option is easier to scale, as the company can directly access the already existing digital infrastructures at the data source in contracts to create its own platform to upload the data resulting in more effort for the customer [7]. The interoperability can also be negatively influenced when creating a platform, as APIs need to be more adaptable to various systems and formats. The flexibility of APIs can thus indirectly influence the visualisation layer by offering more data to portray.

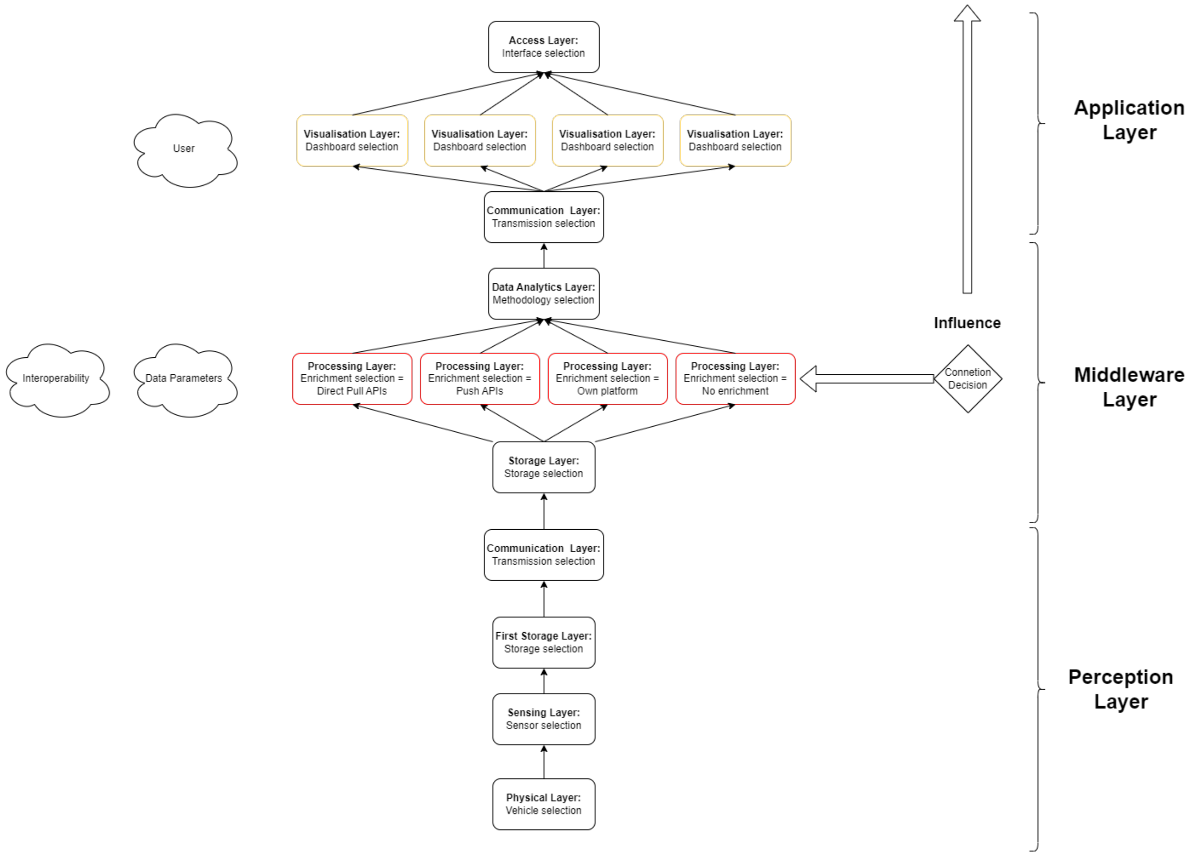

Figure 7.4: Potential path dependencies in the design of the system architecture, due to design option choices (connection selection, middleware layer)

7.5.3. Institutional path dependencies

Lastly, an institutional path dependency is related to governing principles in the system, such as the used GHG emissions quantification methodology and the different requirements (see Fig. 7.5). This is a user-specific design option, and in contrast to the other, the influence is up and downwards. Choosing one methodology over the other first influences the perception and sensor layer downwards. Different methodologies require different input primary data. Thus possible changes or upgrades can occur (design topic data parameters). The impacts of path dependencies when selecting sensors were already visualised in Fig. 7.3. Therefore, the implications of that will not be repeated here and were not added to Fig. 7.5. The selected methodology also has direct implications for data enriching (thus interoperability), as different methodologies can have different requirements for using default emission factors. Lastly, the design option of a methodology also has upwards influences in the application layer. Here the "user" design topic is influenced, as different methodologies have different reporting standards [75, 112]. These reporting standards can predefine the visualisations and dashboards the shipper (or customer) would need from the IT company. However, it is also logical that changes in the methodological approach influence all directions of the system, with already predefining specific design options such as the data sources and the used sensors.

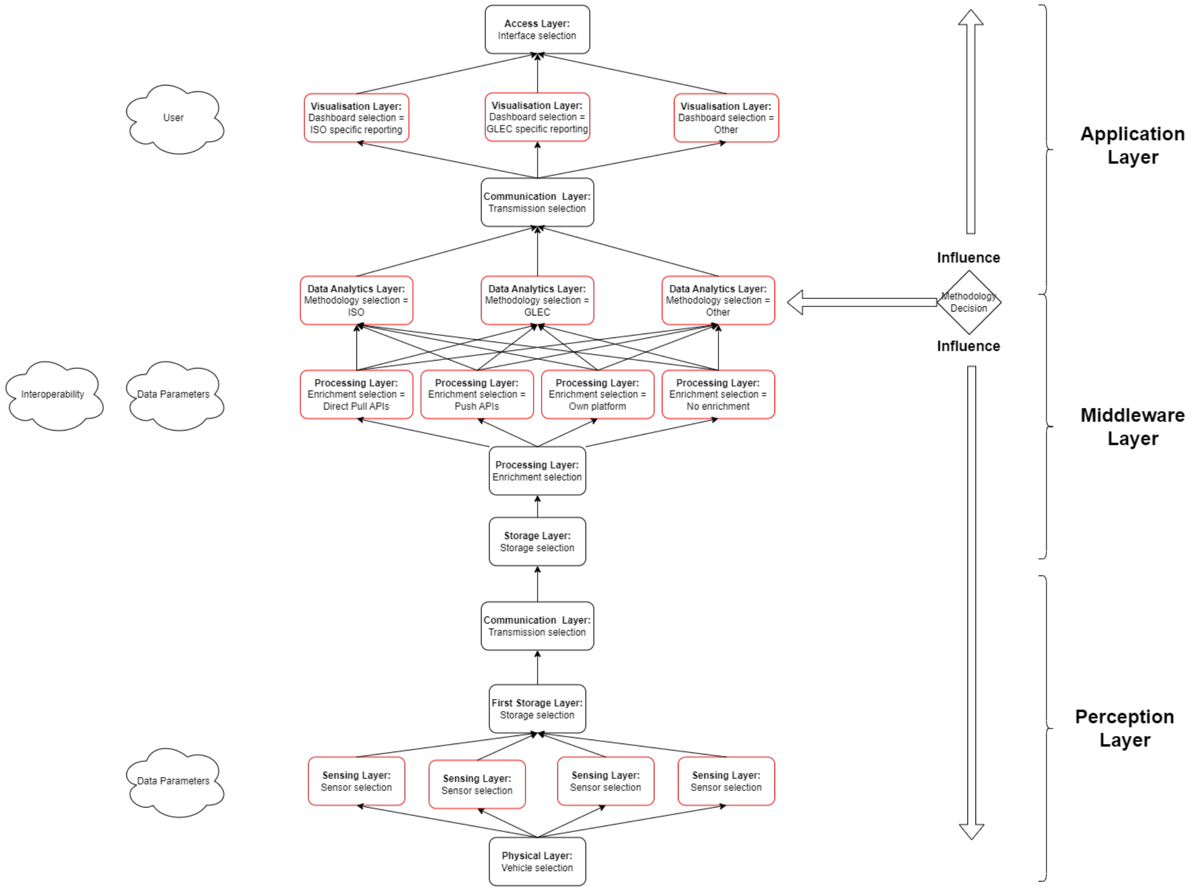

Figure 7.5: Potential path dependencies in the design of the system architecture, due to design option choices (methodology selection, middleware layer)

Table 7.1: Overview of potential path dependencies caused by the design options

Decision/ layer	Impact on other layers	Worst case scenario	Type of path dependency
Vehicle brand (perception layer/ physical layer)	All	Fewer technological improvements	Physical
Selection of sensor (perception layer/ physical layer)	All	Fewer technological improvements	Physical
Connection to a data source (Middleware, communication)	Middleware and application	Inflexibility/ non-interoperability	Digital
Methodology (Middleware/ Analytics)	All	Inaccurate emissions calculations, Old approaches	Institutional

7.6. Success criteria & complexity of the proposed design

The presented system portrays multiple complex and entangled relationships between stakeholders from various domains. Thus this section will briefly mention success factors that the system would need in order to work properly and will also elaborate on the complexity of the design with its various stakeholders. Three institutional-related and three private success factors were established.

7.6.1. Needed success criteria

The system's first and most important success criteria is the willingness of shippers, carriers and LSPs to adopt the current state-of-the-art GHG quantification methodologies for reporting and tracing their logistical emissions. Reporting GHG emissions is not a straightforward task and requires resources and time to be set up properly. Thus companies must be willing to invest their time into it. Otherwise, the system will not be successful, as its main value is currently a compliance matter. This compliance aspect fulfils this success factor, as it forces companies to invest the needed time and resources into the topic via legislative directives such as the CSRD (Corporate Sustainability Reporting Directive), which makes the reporting of business activities impact on the environment mandatory for many large companies in the EU from 2024. European proposals, such as the "CountEmissionsEU "project, further facilitated this, which will suggest using the new ISO14083 as the tool for all further GHG emissions reporting under the CSRD and future GHG-emissions-related policies.

This also includes that the European Union will take a regulatory and systems design role rather than an active stakeholder role. If the EU would completely take over all aspects of the system, thus providing its own system architecture, processing facilities and IT solutions, external and described roles and needed connections in the system would not be necessary. However, the EU's proposal and directives' current development indicates another trajectory towards the EU's regulatory and system-boundary designer-focused role. This means that the EU creates the frameworks and boundary conditions for private and semi-private companies to fill the space with their market-driven solutions. This will act as another success factor for the system, as companies can emerge in this newly created space, which is further facilitated by mandatory reporting schemes from 2024, accelerating companies' adoption rate.

A final institutional success factor will be the need for clear and aligned reporting standards. The system can only be successful if clear and uniform reporting standards exist. The topic of GHG emissions reporting is complex and requires expert knowledge, which very few companies currently have. Many logistical operators are currently overwhelmed by these reporting methodologies' sheer volume of required detail and requirements and thus seek help from external companies. However, the legislative bodies must be careful that the policies are not too complicated and have clear and understandable boundaries. Therefore, the commitment to one or two main reporting methodologies will be key in aligning these reporting standards. Currently, the EU is going towards the right direction by adopting clear methodological standards for reporting. However, these methodologies still have a high level of complexity and leave room for interpretation in the reporting, such as using and reporting bio-fuels or potential off-and-insetting options for CO_2 reduction.

This success factor is highly dependent on the role of external IT companies. More neutral and independent IT companies can act as an intermediate facility that consolidates the data of multiple sites, thus creating this value for all stakeholders. However, this is only the case if all sites (thus carriers and LSP on the one site and shippers on the other) trust these IT companies with their data and do not fear a competitive disadvantage. Currently, different trajectories are visible in the market. On the one hand, truly neutral IT companies that connect the two or more sites and thus increase the trust between all parties. On the other, business-driven companies that offer sourcing and a spot freight market, based on the available data from all sites, thus more as a marketplace. It is not clear which business model will be dominant or if they will co-exist. However, what is needed for both business models is trust.

This trust and the willingness of the logistical operators is a pre-requisite for the final private success factor, the need for the system for long-term value-added services. The system and, with it especially, the IT companies have to transform in the long-term from a compliance-driven business model to a truly value-adding service. The current legislative mandatory reporting requirements are a great start for the system to start off, as companies need to do it. However, the true added value will emerge if these companies develop beyond this reporting towards optimization, decarbonization and insight-driven activities. Compliance reporting must become a more convenient and supportive product, which is possible due to the available transport data. The currently needed complexity of the system is facilitated and supported by the regulatory needs and requirements that companies have to follow in the short term if they do not possess any competencies in the domain and have to outsource it. Thus, these companies create short-term success through regulatory needs rather than created value. However, in the long term, true value can be generated by combining the available logistical data (fuel, distance, truck types, etc.) with business-related data such as order and loading data. Thus, the system must evolve from a compliance-driven service to a supporting and optimization service, which supports companies in optimizing their logistical operations through state-of-the-art data-driven analysis and investment decisions. A carrier especially mentioned these companies' potential for validating investment decisions. An example was a potential new bio-fuel tank station, which could be validated through the cost savings of the new fuel in the long-term benchmarked against traditional fuels. Here the insight into cost savings due to lower CO_2 emissions and thus lower CO_2 costs could be visualized in the tools to showcase their potential.

7.6.2. Complexity of design

The presented complexity of the system is a potential barrier for the success of the system, but also very much needed to incorporate all important decision-makers and stakeholders. The logistical sector always connects at least two sides with each other, which already results in a higher number of stakeholders. In this system, in particular, the policy domain was also considered, which increased the complexity and added more stakeholders while also taking into account the multiple supporting and GHG emissions companies. Thus, the minimal amount of complexity is already, by default, high, making it, on the one hand, such an interesting system to analyze and design, but on the other, it significantly increases complexity and dependencies. However, it is believed that this complexity is needed for the system to work and create value, as it was already described in the success factors that only one type of data (i.e., the rFMS data) is insufficient for proper insights or value-added activities. Thus, a minimum complexity must be present in the system in order for the system to deliver what it was designed for in the first place; accurate GHG emissions reporting. This is further strengthened by the different systems' requirements that the various stakeholders impose, which also increase the complexity, such as the data-privacy concerns.

7.7. Conclusion

The fourth subquestion *"What is, according to industry experts, the quality of the designed artefacts and requirements?"* is the evaluation step in the DSRM and was done via evaluation interviews with industry experts and stakeholders. The evaluation process happened throughout the entire research, and the feedback was directly implemented. Fig. 7.6 summarises the approach, and Table 7.2 summarises the main evaluation comments made by the experts, which were also implemented into the designs.

Table 7.2: Main evaluation comments for the artefact design

Design Artefact	Comments	Comment from Experts
Customer Onboarding	Each new partnership needs consent, even if worked together beforehand	IT expert [4]
Multi-tenant system architecture	Indicate the option of secure and temporary data migration.	IT expert [4]
Data Pulling	Updated processing times and added the option of data pushing from vehicles	IT expert [4]
Business relations	Addition of possible connections between shipper & OEM and IT company & fleet operator	IT expert [4], LSP [5]
System architecture	Addition of LSP in the order chain and expansion of external services instead of just databases	Data marketplace expert [6], IT expert [4]

The second part of the chapter focused on the limitations and potential challenges of the system. The main insights were that the design and evaluation process via industry experts facilities potential biases in the design and evaluation due to the specific stakeholder perspective, which was tried to be mitigated by including a wide range of industry experts and additional scientific literature. Moreover, potential design-based path dependencies were identified on different levels, which showcased that the selection of one GHG emissions reporting methodology has the biggest impact on the system design and, thus, the design-based path dependencies in the system. Lastly, the design was evaluated on the required success factors that must be true for the system to work and the complexity of the system. Here the institutional success factors of the regulatory need, aligned reporting standards, and the willingness of the logistical operators to adopt those standards are given. From a private perspective, the trust of the various logistical operators to share their data with intermediate IT-companies is essential and currently successful in the space. However, the trend for more information sharing will be needed to unlock the much-needed value-adding optimization and decarbonization analytics of those IT companies, which is currently lacking.

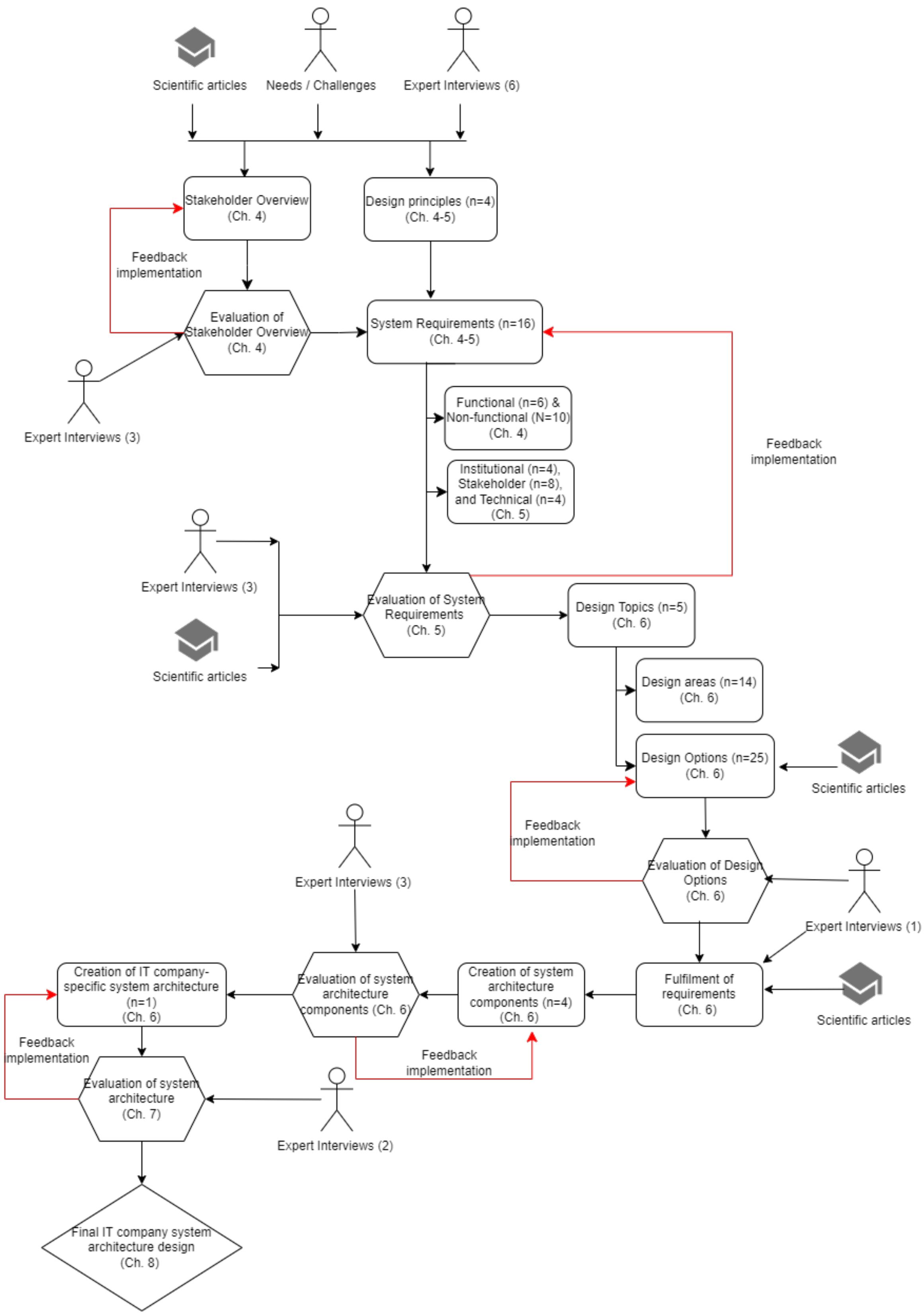

Figure 7.6: Sequence of the design and evaluation process with each intermediate product

8 Conclusion & Future Work

The research aimed to investigate how primary data can improve the accuracy of emission factors in road logistics by addressing the system from a multi-perspective and socio-technical point of view. Additionally, the research presents a final system architecture and outlines the systems' main stakeholders and data flows.

The main question addressed:

What onboard sensor systems architecture can enable road logistics operators to gather primary data from their fleet to accurately determine their vehicle emissions?

The research adopted the Design Science Research Methodology of Peffers et al. (2007) and divided the research into four parts. The research utilised industry expert interviews, literature, policy documents, and informal conversations as knowledge inputs.

The first part investigated the socio-technical system by conducting an institutional, methodologies and industry standards analysis resulting in a stakeholder overview and the first system requirements and design principles.

Second, the remaining design principles and systems requirements were established via additional expert interviews and literature and categorised into institutional, stakeholder and technical requirements in line with ISO29148, which is commonly used in systems requirement analysis and development.

In the third part, the insights from the previous chapters were used to design a stakeholder-specific system architecture.

The final part evaluated all outputs of each part with the help of industry experts and the literature. Past research into the topic had a single-perspective approach to the problem at hand, focusing primarily on the technical requirements of a potential system, thus neglecting the complicated socio-technical environment in which the system is embedded.

The following sections highlight and summarise the contribution of the work and recommendations for key stakeholders. Section 8.1 expands on the contribution of the research by highlighting the importance of the logistical sector and the methodological approach. Section 8.2 summarises the implications and the derived recommendations for different stakeholders. Section 8.3 provides a future outlook and potential future research topics, and section 8.4 will conclude the research by relating it to the degree programme of Complex Systems Engineering & Management.

8.1. Contribution of the research

The decarbonization of the logistical sector until 2050 is an ambitious goal of the European Union, which requires the alignment and readiness of multiple domains. CO_2 reporting based on primary logistical data is a core enabler to reach the goal, reduces the potential of green-washing, achieves end-to-end traceability, and enables private companies to quantify the monetary benefits of CO_2 reductions in their logistical operations. If the public and private domains fail to act now, two scenarios are realistic:

1. The European CO_2 reduction targets are not reached. This leads to further acceleration of the climate-induced negative aspects and would send a fatal signal to all public or private organisations that seriously drive the climate-friendly transformation.
2. The European Union enforces already agreed-upon climate mitigation agreements and reduction policies, such as the carbon pricing rules, while companies are not ready, leading to a significant competitive disadvantage.

To avoid these scenarios, the research contributed directly to the practical design process of an abstract system architecture in section 8.1.1 and to the theoretical research approach and methodology for future system designer in section 8.1.2.

8.1.1. Contribution to system architectures in road logistics

The main contribution of this research was the adoption of a wider systems design boundary, resulting in the possibility of including institutional, stakeholder and technical considerations in the systems development and the following seven contributions/ recommendations:

1. The inclusion of a multi-domain designer perspective when designing an abstract system architecture in the logistical sector - should be mandatory in the scoping of any project.
2. The identification of relevant stakeholder clusters and an abstract stakeholder analysis to be considered when designing in the socio-technical environment.
3. The categorization of systems requirements into institutional, stakeholder and technical requirements to represent the multi-domain character of the system.
4. The need for financial quantification tools of the CO_2 reduction for logistical operators to validate their investment decisions.
5. Compliance with leading European GHG emissions quantification methodologies and data regulations - should also be implemented in leading EU taxonomies.
6. A stakeholder cluster-specific system architecture, which incorporated the derived requirements and outlined business relationships and data flows in the system, evaluated and approved by industry experts.
7. The issue of the stakeholder-specific requirements for the system leading to multiple co-existing system architecture specifications.

Moreover, this research can be a conversation starter for discussions of how primary data can be implemented in state-of-the-art GHG emissions reporting schemes. The proposed system architecture can act as a blue-print of how the primary data is collected at the vehicle level and transmitted via the relevant stakeholders to a final processing site. This is not just the case for road logistics but also for other areas that involve a magnitude of stakeholders and a similar socio-technical environment.

8.1.2. Contribution to methodological approaches

The research contributed in two ways to the current systems design methodology. First, it created the stakeholder system requirements in line with the ISO 29140 norm for the specific area of onboard sensor systems in road logistics while utilizing the Pfeffers et al. (2007) Design Science Research Methodology. Thus, using the two theoretical systems design approaches and applying them to a practical example and ensuring that important industry and expert considerations could be integrated into the design process. Second, it further strengthens the results of Iacovidou, Hahladakis & Purnell (2021) [70] and Govindan and Bouzon (2018) [62], who both propose a system of system thinking designing approach within the logistical sector to consider multiple domains. Their proposal of a multi-domain design perspective in sustainable logistics was extended to the domain of stakeholders. Thus, further researchers are encouraged to use Iacovidou, Hahladakis & Purnell (2021) proposed system of a systems thinking approach with the included domains of governance, technologies, business, natural resources and patterns of behaviour and on top the identified stakeholder perspective.

8.2. Implications and recommendations for Stakeholders

Using predominantly primary data for reporting GHG emissions, which is transmitted via system architectures, will result in different implications and recommendations per stakeholder cluster, which were derived and gathered through the research. The coming section will outline the key points that stakeholders need to consider to deal with the more strict reporting schemes.

8.2.1. External IT-companies

The legislative push towards primary data-driven GHG emission reporting schemes is a great opportunity for many IT companies. However, their success will not be happening by default, as already outlined in section 7.6. Thus, IT companies will need to consider the following aspects: **Focus on Scalability and Flexibility.** The mandatory reporting deadline of the CSRD is getting closer, and many logistic operators who have not yet considered its implication will need a quick and knowledgeable supporting company to start their GHG emissions reporting.

The first recommendation for IT companies is, therefore, to design scalable and robust products while focusing on core product competencies that allow for high flexibility for specific user requests and for continuous adjustments and legal / reporting requirements.

The second recommendation for IT companies is to find the **right balance between short-term** needs – creating true business value for the logistical operator by solving today's real-life problems by either providing the full GHG emissions compliance reporting program or offering optimisation, freight brokering and benchmark analysis to improve internal logical operations by reducing, i.e. the empty running of trucks – and the **long-term business value** – enabling system adaptations and historical data transferability (see contribution number 4 in 8.1.1). This will create a competitive advantage and secure long-term business revenues.

8.2.2. Logistical Operators

Logistical Service Providers (LSPs)

LSPs and carriers will need to make significant investments in their digital infrastructure to track and monitor the primary data of their fleets. Reduced risks for these large investments can be made by proposing **Joint Ventures** with OEMs or shippers to not carry the financial burden alone or by negotiating **long-term contracts** with their customers to increase their financial safety when making these investments.

Shippers

Shippers must decide how they want to tackle GHG emissions reporting. They can either do it in-house by themselves or outsource the activities.

Nonetheless, the recommendation to shippers is, **to act now** and implement decarbonization strategies and carbon reduction measures (see contribution number 5 in 8.1.1). This also entails that they already now start to communicate to their **carriers that they need to have detailed primary data in the future**, as it will also take time on their site to set up appropriate systems and connect them to either the shipper or external companies. The additional efforts of capturing the GHG emissions based on the ISO 14083 on a transport operation chain, thus on high granularity, however, reporting them on an aggregated transport chain level will lead to additional efforts for them. But the gained insights can be used as a **business advantage** to create carbon footprints per product to target ecologically conscious consumers while complaining with legislation.

8.2.3. Policy makers

Policymakers play a crucial and active role throughout the entire system design, implementation and monitoring. The diversity of stakeholders makes it crucial that policymakers lay out the foundation of **minimal reporting standards for all logistical operators and a common reporting standard** (see contributions number 1, 2, and 5 in 8.1.1). Next to that, it is crucial that they make **supporting infrastructure investments** into vehicle charging, provide quick and easily accessible subsidies and continue to have active communication channels with the industry to react to industry demands. It will also be crucial that they keep an open attitude regarding **technology openness** not to limit the possible solutions for relevant stakeholders, while on the other hand, giving certainty to investors to ensure that their investments will pay off - otherwise, shippers and other logistical operators will take on an attitude of "wait and see", instead of innovating (see contribution number 7 in 8.1.1). Lastly, it will be critical that they allow companies to leverage their sustainability efforts via financial reporting to create incentives for companies to be at the forefront of GHG reductions.

8.2.4. Framework owners

Framework owners and developers should see the adoption of their GHG quantification standards as European standards as a validation for the great work and continue with it. Core tasks for them will be to find the right balance between appropriate complexity and state-of-the-art methodologies. Currently, especially non-GHG-related toxins and indirect land change uses (LCU) are not implemented in the fuel emission factors, which will be the next step to even more accurate emission reporting. Framework owners have hereby the responsibility to add the needs as a broadening to existing and not as a completely new framework. Lastly, it is also their responsibility to formulate the frameworks in such a conscience manner that **green-washing will be impossible** or multiple interpretations are possible for reporting GHG emissions.

8.3. Outlook to the field & Future Research

Current Developments

The topic of GHG emissions reporting is currently one of the most important topics, leading to the magnitude of policy documents and proposals that were and are published during the entire duration of this research. Just at the beginning of the research, the new ISO 14083 was published and approved, while during the research, the new CountEmissionsEU proposal was published that recommended to use the ISO 14083 as the common GHG quantification methodology for all future GHG reportings of the Corporate Sustainability Reporting Directive (CSRD). Especially the CountEmissionsEU proposal validates the proposed complexity of the system architecture, which was based on the ISO14083 recommendation to use primary data for emissions reporting. While writing this, the new European Data Act is about to be enforced, which would make the availability of vehicle data easier, would limit the OEMs to only use the data themselves, and make it more affordable to purchase the data for third parties [55]. Thus, anticipations about the future of the field are heavily dependent on new and upcoming legislation.

Short-Term
The logistics sector will face massive challenges in the short term, and not every company is ready for the upcoming legislative changes. Many company executives and CEOs are worried about the new regulatory constraints and are planning already now to devote more time to supply chain-related activities [12]. It will be a tough and complicated journey for many companies, and first movers will be able to create a real competitive advantage from it. Many companies will risk potential fines or must accept high carbon prices, or in the worst case, being washed away from the business market. The presented scenario in section 7.6, which summarised compliance GHG reporting as a site service, seems unrealistic in the short term. Many IT companies and system architectures will still have a clear focus on this core product and service, as this will still be the dominant driving force for many companies to seek their services. The presented system architecture is expected to still be adapted and further improved after the insights of the first mass implementations. Multiple companies in the market will also offer similar or identical products and services as no company consolidation has happened yet, and logistic companies will try to spread their data reporting to multiple companies to spread their risk. Scalable and robust services and companies will have the biggest potential to maintain in the market by innovating their services and offering ”freemium” products for additional insights into the logistic processes for their customers.
Long-Term
In the long term, it is assumed that the services will be more and more centred around combining multiple data types - e.g. fusing multiple types of non-financial reports into **one reporting**. The proposed system architecture can be one tool to gather the primary data to be combined with business and order data to generate more advanced insights for optimization purposes. Other technologies, such as blockchain, are also highly interesting for the purpose of data sovereignty and data sharing between different entities. It is also assumed that the emissions and carbon reporting sector will develop into something similar to the financial service, with end-to-end traceability of carbon and emissions for every product, similar emissions and carbon audits and detailed insights for every customer about the carbon footprint of every consumer product. This might be in the future also used for environmental fees (similar to VAT) and with that be reflected in the prices of the products, which have the price for the emitted CO_2 from transportation, extraction and production integrated into the final price that the consumer will pay. The proposed system architecture can contribute here by delivering the detailed emissions caused by the transportation of the good. Sustainability will become a core aspect of every product and thus will become the data related to the sustainability score of a product, an important commodity. The European Union is already starting to set up this environment with initiatives such as the digital product passport, which could also be a medium to save every product’s final transport-related carbon footprint.

8.3.1. Future Research
The research results have identified interesting knowledge gaps that should be investigated in future research activities. First, the stakeholder-specific character of the system requirements opens the door to many tailored design possibilities. Future research should identify if it is possible to create a detailed system architecture for onboard sensor systems to calculate GHG emissions on a stakeholder-cluster level or if that is only possible on the individual stakeholder level. The specific system architecture already indicates that it might be possible to create already specific architectures on a stakeholder-cluster level, but this is not fully validated.

The topic of GHG emissions calculations with detailed reporting requirements and methodologies is still relatively new for many industry players. Shippers, LSPs, and Carriers are just now starting to investigate and explore this field and are still very much at the beginning of it. Early adaptors and first movers in this area, like in every industry, use their head-start as a competitive advantage, but the general mass of companies is not ready yet. Not all vehicles are equipped with sufficiently accurate sensors, especially older ones, or are connected to a database. Thus, it could be interesting to investigate how the field developed by itself, what the main institutional and regulatory differences between the current and future state of the system might be, and how they influence the design of such a system.

Moreover, the research could be expanded to other modalities such as rail, sea or air to investigate whether these modalities have the same or completely different requirements for such a system architecture. There is a potential that the other modalities could use the main learning of this research and related ones to increase the development speed.

Data classification under the GDPR is a main influence factor for the system. Especially logistical operators and IT companies have many obligations in handling the data and have difficulties explaining to regulators how and why they want to use it [6, 4]. Currently, the data of the vehicles are still considered personal data of people, as the driver is operating it. However, the companies are not interested in the driver’s data rather than the vehicle’s data for their calculations. Thus, separating these transport data into human- and vehicle-centred would significantly improve the sector. Stakeholders have mentioned that they were planning to create direct Joint Venture API connections to other database providers (i.e. OEMs); however, the increased complexity of data privacy concerns and the GDPR is stopping them from doing so or significantly reducing the implementation [6, 4]. Thus, it could be a significant contribution to the field to research how a potential data classification scheme, which separates vehicle and human data, would look like and how this would influence the stakeholders.

Another point that this research did not focus on was other forms of data sharing apart from direct API connections and contractual agreements. Two efforts that could significantly impact the future of data sharing, also for the logistical sector and the design of system architectures for GHG emissions calculations, should be highlighted. First, the concept of data spaces, such as the International Data Space (IDS) [72]. Data spaces aim to create a digital environment for the future data economy, where data sharing is secure, governed by the data sovereignty principle to access the full data value/ potential [72]. Another form of secure data sharing is via blockchain applications. Here the European Blockchain Service Infrastructure (EBSI) could be an option for secure data sharing for all the stakeholders [34]. Both projects are not fully deployed yet, thus were not considered for the research. However, their potential is enormous and could significantly impact the design so that even all presented stakeholders connect to the services they provide for data sharing. The influence should be investigated for road logistics.

Lastly, the coming GDPR update will have potentially large implications for the proposed designs. Any change to the regulation can dramatically change the design of the onboard sensor system architectures in road logistics at unknown magnitudes. It is currently not possible to quantify these impacts, but the European Union is addressing problems with the GDPR, which they aim to release in the second quarter of 2023 [36]. Main updates will be about enforcing GDPR data breaches and more transparent rules for the Big-Tech companies on taxes, market entries or rights of customers [89]. There are currently no comments specifically about the transport sector or re-categorisation of transport-related data. The impacts could be enormous if these would appear or if other changes are made that indirectly affect the industry and the design process. Thus, it is vital to investigate the implications of the GDPR update in future research.

8.4. Link of the study to the CoSEM degree

The problem and research at hand summarise a typical Complex Systems Engineering & Management topic with a clear focus on the logistical and ICT sector. It is based on the foundation that the big questions and problems are a combination of multi-disciplinary parts combined into a larger system, which is heavily influenced by the people that act within the system, the available technologies and the influence of governmental organisations via legislation or incentives. The technical component is at the core of the research, as only this would allow the collection and monitoring of emissions of logistic road vehicles; however, the technology by itself is not able to achieve a change as only the combined efforts of policymakers, involved stakeholders and technology providers can change the status quo. The research connects the socio-technical analysis with establishing system requirements and design principles and evaluates it with industry actors to assess the functionality of the technical solutions with the stakeholder domain. Fig. 8.1 visualises the multi-domain and CoSEM-related aspects of the research.

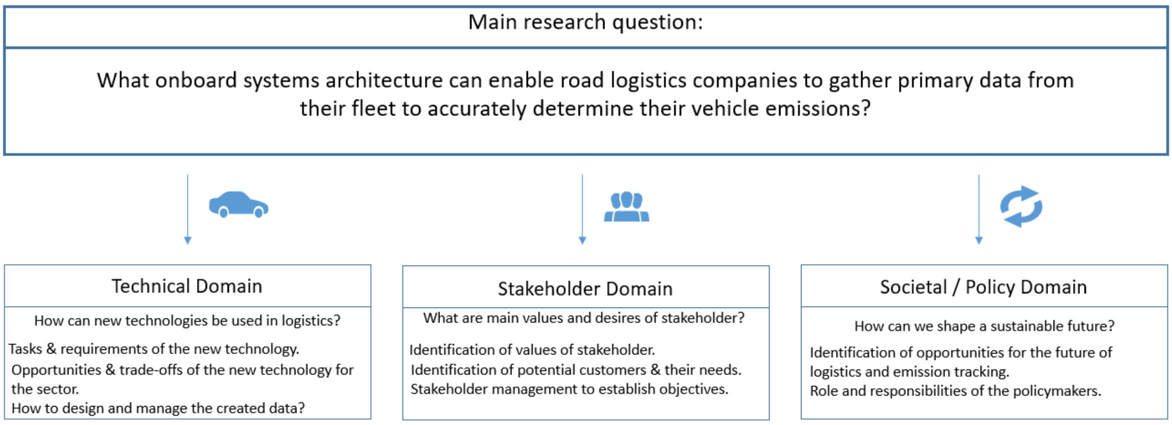

Figure 8.1: Problem decomposition into the three main components of technical, stakeholder, and societal aspects

References

[1] Kyriakos Agavanakis et al. "Driving Sustainability in Logistics Value Chains - A Telematics Data Hub Implementation for Accurate Carbon Footprint Assessment and Reporting Using Global Standards-based Tools". In: Tmrees, Technologies and Materials for Renewable Energy, Environment and Sustainability, 2023.

[2] Abdul-Rahman Al-Ali et al. "Digital twin conceptual model within the context of internet of things". In: *Future Internet* 12.10 (2020), p. 163.

[3] Anonymous. *Evaluation with a carrier, see Appendix C.3*. 2023.

[4] Anonymous. *Evaluation with a data scientist at an external IT company, see Appendix C.5*. 2023.

[5] Anonymous. *Evaluation with an LSP, see Appendix C.4*. 2023.

[6] Anonymous. *Evaluation with marketplace operator, see Appendix C.2*. 2023.

[7] Anonymous. *Expert Interview with a data marketplace and IoT aggregator expert, see Appendix A.5*. 2023.

[8] Anonymous. *Expert Interview with a logistical actor (carrier), see Appendix A.6*. 2023.

[9] Anonymous. *Expert Interview with a logistical actor (LSP), see Appendix A.7*. 2023.

[10] Anonymous. *Expert Interview with emissions measurement and modelling expert, see Appendix A.4*. 2023.

[11] Anonymous. *Expert Interview with sensor and IT emissions tracking expert, see Appendix A.3*. 2023.

[12] Ashcroft, Sean. *Supply chain ESG laws causing mass CEO anxiety, says Proxima*. 2023. URL: https://supplychaindigital.com/digital-supply-chain/supply-chain-esg-laws-causing-mass-ceo-anxiety-says-proxima.

[13] Cor-Paul Bezemer and Andy Zaidman. "Multi-tenant SaaS applications: maintenance dream or nightmare?" In: *Proceedings of the joint ercim workshop on software evolution (evol) and international workshop on principles of software evolution (iwpse)*. 2010, pp. 88–92.

[14] Thiyagarajan Manihatty Bojan, Umamaheswaran Raman Kumar, and Viswanathan Manihatty Bojan. "An internet of things based intelligent transportation system". In: *2014 IEEE international conference on vehicular electronics and safety*. IEEE. 2014, pp. 174–179.

[15] Jan vom Brocke, Alan Hevner, and Alexander Maedche. "Introduction to design science research". In: *Design science research. Cases* (2020), pp. 1–13.

[16] CDSE. *Information System Security Manager (ISSM) Toolkit*. 2023. URL: https://www.cdse.edu/Training/Toolkits/Information-System-Security-Manager-Toolkit/.

[17] Chenxi Chen et al. "Research on vehicle fuel consumption prediction model based on Cauchy mutation multiverse algorithm". In: *2022 9th International Forum on Electrical Engineering and Automation (IFEEA)*. IEEE. 2022, pp. 1115–1119.

[18] Connected Papers. *About Connected Papers*. 2023. URL: https://www.connectedpapers.com/about.

[19] Cosgrove, Emma. *VCs are abandoning supply chain tech — except for startups that can cut out the human*. 2022. URL: https://www.businessinsider.com/vc-investment-supply-chain-startups-tanked-third-quarter-2022-12?international=true&r=US&IR=T.

[20] Mauro A. A. da Cruz et al. "A Reference Model for Internet of Things Middleware". In: *IEEE Internet of Things Journal* 5.2 (2018), pp. 871–883. DOI: 10.1109/JIOT.2018.2796561.

[21] Mauro AA da Cruz et al. "A reference model for internet of things middleware". In: *IEEE Internet of Things Journal* 5.2 (2018), pp. 871–883.

[22] D'Inca, Joris and Borreck, Max-Alexander. *Digital Logistics Startups Are Both Challenge And Opportunity For Industry Incumbents*. 2017. URL: https://www.forbes.com/sites/oliverwyman/2017/07/28/digital-logistics-startups-are-both-challenge-and-opportunity-for-industry-incumbents/?sh=17cce4781589.

[23] Gaspare D'Amico et al. "Smart and sustainable logistics of Port cities: A framework for comprehending enabling factors, domains and goals". In: *Sustainable Cities and Society* 69 (2021), p. 102801.

[24] Mohammad Dabbagh and Sai Peck Lee. "An approach for integrating the prioritization of functional and nonfunctional requirements". In: *The Scientific World Journal* 2014 (2014).

[25] DataGuard. *Build trust and operationalise compliance*. 2023. URL: https://www.dataguard.com/.

[26] SNC Dellaert and R Dröge. "Uncertainty of the NOx, SOx, NH3, PM10, PM2. 5, EC2. 5 and NMVOC emissions from transport". In: *TNO, Utrecht* (2017).

[27] DHL. *Data protection and its legal implications in logistics*. 2019. URL: https://lot.dhl.com/data-protection-and-its-legal-implications-on-logistics/.

[28] Poornima Dixit et al. "Differences between emissions measured in urban driving and certification testing of heavy-duty diesel engines". In: *Atmospheric Environment* 166 (2017), pp. 276–285.

[29] DNV. *EU ETS: Preliminary agreement to include shipping in the EU's Emission Trading System from 2024*. 2023. URL: https://www.dnv.com/news/eu-ets-preliminary-agreement-to-include-shipping-in-the-eu-s-emission-trading-system-from-2024-238068.

[30] DSLV. *Zahlen - Daten - Fakten aus Spedition und Logistik*. 2010. URL: https://www.dslv.org/fileadmin/Redaktion/PDFs/10_Pressemitteilungen/Kompaktversion_deutsch.pdf.

[31] DSLV. *Zahlen - Daten - Fakten aus Spedition und Logistik*. 2015. URL: http://www.appen-gruppe.de/DSLVZahlen.pdf.

[32] EEA, JRC, LAT, Emisia. *copert - The industry standard emissions calculator*. 2016. URL: https://www.emisia.com/utilities/copert/.

[33] ERMES. *Organizations - Our assocciates*. n.d. URL: https://ermes-group.eu/organizations.

[34] European Co missions. *Use cases - EBSI*. 2023. URL: https://ec.europa.eu/digital-building-blocks/wikis/display/EBSI/What+is+ebsi#what-ebsi-used-for.

[35] European Commission. *A European Green Deal*. 2019. URL: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en.

[36] European Commission. *Further specifying procedural rules relating to the enforcement of the General Data Protection Regulation*. 2023. URL: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13745-Further-specifying-procedural-rules-relating-to-the-enforcement-of-the-General-Data-Protection-Regulation_en.

[37] European Commission. *Paris agreement*. 2016. URL: https://climate.ec.europa.eu/eu-action/international-action-climate-change/climate-negotiations/paris-agreement_en.

[38] European Commission. *Transport sector economic analysis*. 2023. URL: https://joint-research-centre.ec.europa.eu/scientific-activities-z/transport-sector-economic-analysis_en.

[39] European Environment Agency. *Greenhouse gas emissions from transport in Europe*. 2021. URL: https://www.eea.europa.eu/data-and-maps/indicators/transport-emissions-of-greenhouse-gases/transport-emissions-of-greenhouse-gases-12.

[40] European Environment Agency. *Greenhouse gas emissions from transport in Europe*. 2022. URL: https://www.eea.europa.eu/ims/greenhouse-gas-emissions-from-transport.

[41] European Union. *Commission Regulation (EC) No 692/2008*. 2008. URL: https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%5C%3A32008R0692.

[42] European Union. *Commission Regulation (EU) 2017/1151*. 2017. URL: https://eur-lex.europa.eu/eli/reg/2017/1151/oj/eng.

[43] European Union. *Commission Regulation (EU) 2017/1152*. 2017. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%5C%3A32017R1152.

[44] European Union. *Commission Regulation (EU) 2017/1153*. 2017. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%5C%3A32017R1153.

[45] European Union. *COMMISSION REGULATION (EU) 2018/1832*. 2018. URL: https://eur-lex.europa.eu/legal-content/GA/TXT/?uri=CELEX:32018R1832.

[46] European Union. *Commission Regulation (EU) 2019/1939*. 2019. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019R1939&from=EN.

[47] European Union. *Commission Regulation (EU) 2021/392*. 2021. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021R0392&rid=1.

[48] European Union. *Commission Regulation (EU) No 1014/2010*. 2010. URL: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%5C%3A32010R1014.

[49] European Union. *Commission Regulation (EU) No 293/2012*. 2012. URL: https://eur-lex.europa.eu/eli/reg_impl/2012/293/oj.

[50] European Union. *Commission Regulation (EU) No 582/2011*. 2011. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%5C%3A32011R0582.

[51] European Union. *DIRECTIVE (EU) 2022/2464 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 14 December 2022 amending Regulation (EU) No 537/2014, Directive 2004/109/EC, Directive 2006/43/EC and Directive 2013/34/EU, as regards corporate sustainability reporting*. 2022. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32022L2464.

[52] European Union. *Directive 2007/46/EC of the European Parliament and of the Council*. 2007. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02007L0046-20180331.

[53] European Union. *General Data Protection Regulation- GDPR*. 2018. URL: https://gdpr-info.eu/.

[54] European Union. *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the accounting of greenhouse gas emissions of transport services*. 2023. URL: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13217-Count-your-transport-emissions-%E2%80%98CountEmissions-EU%E2%80%99_en.

[55] European Union. *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the accounting of greenhouse gas emissions of transport services*. 2023. URL: https://ec.europa.eu/commission/presscorner/detail/en/ip_23_3491.

[56] European Union. *REGULATION (EU) 2018/956 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 28 June 2018 on the monitoring and reporting of CO2 emissions from and fuel consumption of new heavy-duty vehicles*. 2018. URL: https://eur-lex.europa.eu/eli/reg/2018/956/oj.

[57] FEDeRATED. *FEDeRATED - Network of Platforms*. n.d. URL: http://www.federatedplatforms.eu/.

[58] Vicente Franco et al. “Road vehicle emission factors development: A review”. In: *Atmospheric Environment* 70 (2013), pp. 84–97.

[59] Mario Frustaci et al. “Evaluating critical security issues of the IoT world: Present and future challenges”. In: *IEEE Internet of things journal* 5.4 (2017), pp. 2483–2495.

[60] Gerben Geilenkirchen et al. “Methods for calculating the emissions of transport in the Netherlands”. In: *PBL Netherlands Environmental Assessment Agency: The Hague, The Netherlands* (2020).

[61] Google Scholar. *About Google Scholar*. 2023. URL: https://scholar.google.com/intl/en/scholar/about.html.

[62] Kannan Govindan and Marina Bouzon. "From a literature review to a multi-perspective framework for reverse logistics barriers and drivers". In: *Journal of cleaner production* 187 (2018), pp. 318–337.

[63] Gravity. *How GDPR Affects The Logistics Industry And Why You Need To Be Prepared Now*. 2018. URL: https://gravitysupplychain.com/gdpr-checklist/.

[64] Theodoros Grigoratos et al. "Real world emissions performance of heavy-duty Euro VI diesel vehicles". In: *Atmospheric environment* 201 (2019), pp. 348–359.

[65] HDEI / BCEI Task Force. *FMS-Standard description, Version 04, 13.10.2017*. 2017. URL: http://www.fms-standard.com/Truck/down_load/fms%5C%20document_v_04_vers.13.10.2017.pdf.

[66] Heilemann, Fabian. *Why the logistics industry is ripe for disruption*. 2017. URL: https://medium.com/birds-view/why-the-logistics-industry-is-ripe-for-disruption-73358a655b04.

[67] David Martin Herold and Ki-Hoon Lee. "Carbon management in the logistics and transportation sector: An overview and new research directions". In: *Carbon Management* 8.1 (2017), pp. 79–97.

[68] Arne Höltl, Cathy Macharis, and Klaas De Brucker. "Pathways to decarbonise the European car fleet: A scenario analysis using the backcasting approach". In: *Energies* 11.1 (2017), p. 20.

[69] John Hopkins and Paul Hawking. "Big Data Analytics and IoT in logistics: a case study". In: *The International Journal of Logistics Management* (2018).

[70] Eleni Iacovidou, John N Hahladakis, and Phil Purnell. "A systems thinking approach to understanding the challenges of achieving the circular economy". In: *Environmental Science and Pollution Research* 28 (2021), pp. 24785–24806.

[71] Yidanes Alejandra Ibañez-Acevedo et al. "Greenhouse Gas Emission Scenarios and Vehicle Engine Performance in a Main Urban Road in Northwestern Mexico". In: *Applied Sciences* 12.23 (2022), p. 12502.

[72] IDSA. *International Data Space - The future of the data economy is here*. 2023. URL: https://internationaldataspaces.org/.

[73] INFRAS. *About HBEFA 4.2 (Handbook Emission Factors for Road Transport)*. n.d. URL: https://www.hbefa.net/Tools/EN/MainSite.asp.

[74] International Transport Forum. "Decarbonising Transport in Europe The Way Forward". In: (2021). URL: https://www.itf-oecd.org/sites/default/files/docs/decarbonising-transport-europe-way-forward.pdf.

[75] ISO. *ISO 14083:2023. Greenhouse gases — Quantification and reporting of greenhouse gas emissions arising from transport chain operations*. 2023. URL: https://www.iso.org/standard/78864.html.

[76] ISO. *ISO/IEC 27001; Information security management systems*. 2022. URL: https://www.iso.org/standard/27001.

[77] "ISO/IEC/IEEE International Standard - Systems and software engineering – Life cycle processes – Requirements engineering". In: *ISO/IEC/IEEE 29148:2018(E)* (2018), pp. 1–104. DOI: 10.1109/IEEESTD.2018.8559686.

[78] Paul Johannesson and Erik Perjons. *An introduction to design science*. Vol. 10. Springer.

[79] Paul Johannesson et al. "Knowledge types and forms". In: *An introduction to design science* (2014), pp. 21–38.

[80] Fadoua Khanboubi, Azedine Boulmakoul, and Mohamed Tabaa. "Impact of digital trends using IoT on banking processes". In: *Procedia Computer Science* 151 (2019), pp. 77–84.

[81] Patrick Killeen et al. "IoT-based predictive maintenance for fleet management". In: *Procedia Computer Science* 151 (2019), pp. 607–613.

[82] Irem Kok et al. *Accelerating ZEV adoption in fleets to decarbonize road transportation*. 2023. URL: %7Bhttps://theicct.org/wp-content/uploads/2023/05/ZEVTC_fleets-briefing_final.pdf%7D.

[83] Marina Kousoulidou et al. “Road-transport emission projections to 2020 in European urban environments”. In: *Atmospheric Environment* 42.32 (2008), pp. 7465–7475.

[84] Patricia Krecl et al. “Trends in black carbon and size-resolved particle number concentrations and vehicle emission factors under real-world conditions”. In: *Atmospheric Environment* 165 (2017), pp. 155–168.

[85] NE Ligterink, PS van Zyl, and VAM Heijne. *Dutch CO2 emission factors for road vehicles*. Delft: TNO, 2016.

[86] Kristina Liljestrand, Martin Christopher, and Dan Andersson. “Using a transport portfolio framework to reduce carbon footprint”. In: *The International Journal of Logistics Management* (2015).

[87] NPD Martin, JDK Bishop, and AM Boies. “Emissions, performance, and design of UK passenger vehicles”. In: *International Journal of Sustainable Transportation* 11.3 (2017), pp. 230–236.

[88] Alan McKinnon. “European freight transport statistics: limitations, misinterpretations and aspirations”. In: *5th ACEAa Scientific advisory group meeting. ACEA. Bruxelles*. 2010.

[89] McKinsey & Company. *The EU digital strategy: The impact of data privacy on global business*. 2023. URL: `https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/the-eu-digital-strategy-the-impact-of-data-privacy-on-global-business`.

[90] Pablo Mendoza-Villafuerte et al. “NOx, NH3, N2O and PN real driving emissions from a Euro VI heavy-duty vehicle. Impact of regulatory on-road test conditions on emissions”. In: *Science of The Total Environment* 609 (2017), pp. 546–555.

[91] GAN Mi et al. “Review on application of truck trajectory data in highway freight system”. In: *Journal of Transportation Systems Engineering and Information Technology* 21.5 (2021), p. 91.

[92] Microsoft Azure. *Baseline Informatiebeveiliging Rijksdienst standard (BIR 2012)*. 2023. URL: `https://learn.microsoft.com/en-us/compliance/regulatory/offering-bir-2012-netherlands`.

[93] Microsoft Azure. *Find the Azure geography that meets your needs*. 2023. URL: `https://azure.microsoft.com/en-gb/explore/global-infrastructure/geographies/#overview`.

[94] Microsoft Azure. *Locations of Microsoft Azure data centres*. 2023. URL: `https://datacenterlocations.com/wp-content/uploads/2020/08/Microsoft-Azure-Data-Center-Locations.jpg`.

[95] Microsoft Azure. *Multi-tenant SaaS database tenancy patterns*. 2023. URL: `https://learn.microsoft.com/en-us/azure/azure-sql/database/saas-tenancy-app-design-patterns?view=azuresql`.

[96] Dominika Crnjac Milić, Ivana Hartmann Tolić, and Marina Peko. “Internet of Things (IoT) solutions in smart transportation management”. In: *Business Logistics in Modern Management* (2020).

[97] Milieuzones in Nederland. *Tables Vehicle’s Euro emissions standard*. n.d. URL: `https://www.milieuzones.nl/tables-vehicles-euro-emissions-standard`.

[98] Vı́ctor Valverde Morales and Pierre Bonnel. *On-road testing with Portable Emissions Measurement Systems (PEMS)*. 2017.

[99] Rafael Fernandes Mosquim and Carlos Eduardo Keutenedjian Mady. “Performance and Efficiency Trade-Offs in Brazilian Passenger Vehicle Fleet”. In: *Energies* 15.15 (2022), p. 5416.

[100] Jagadish Nayak. “Round the clock vehicle emission monitoring using IoT for smart cities”. In: *International Journal of Advanced Computer Science and Applications* 9.11 (2018).

[101] Chaeyeon Oh, Joonseo Ha, and Heejun Roh. “A survey on TLS-encrypted malware network traffic analysis applicable to security operations centers”. In: *Applied Sciences* 12.1 (2021), p. 155.

[102] Artur Oliveira et al. “Comparing IoT Platforms under Middleware Requirements in an IoT Perspective”. In: ().

[103] Matthew J Page et al. “The PRISMA 2020 statement: an updated guideline for reporting systematic reviews”. In: *International journal of surgery* 88 (2021), p. 105906.

[104] Ken Peffers et al. “A design science research methodology for information systems research”. In: *Journal of management information systems* 24.3 (2007), pp. 45–77.

[105] John Jairo Posada-Henao, Iván Sarmiento-Ordosgoitia, and Alexánder A Correa-Espinal. “Effects of Road Slope and Vehicle Weight on Truck Fuel Consumption”. In: *Sustainability* 15.1 (2023), p. 724.

[106] David C Quiros et al. “Deriving fuel-based emission factor thresholds to interpret heavy-duty vehicle roadside plume measurements”. In: *Journal of the Air & Waste Management Association* 68.9 (2018), pp. 969–987.

[107] Hafiz Wahab Raza et al. “A review of middleware platforms in internet of things: a non–functional requirements approach”. In: *Journal of Independent Studies and Research Computing* 10 (2020), p. 18.

[108] rFMS version 2.0. *rFMS version 2 - API documentation*. 2016. URL: `http://www.fms-standard.com/Truck/down_load/Technical_Specification_rFMS_V2.0.0_21.09.2016.pdf`.

[109] Michael Roeth. “Transformational technologies reshaping transportation-an industry perspective”. In: *SAE International Journal of Advances and Current Practices in Mobility* 3.2020-01-1945 (2020), pp. 5–48.

[110] Knut-H Rolland and Ole Hanseth. “Managing path dependency in digital transformation processes: a longitudinal case study of an enterprise document management platform”. In: *Procedia Computer Science* 181 (2021), pp. 765–774.

[111] shipzero.com. *Decarbonize your supply chain Measure emissions accurately on a global scale and move towards net-zero logistics*. 2023. URL: `https://shipzero.com/`.

[112] Smart Freight Centre. *Global Logistics Emissions Council Framework for Logistics Emissions Accounting and Reporting. Version 2.0*. 2022. URL: `internal%20access`.

[113] Smart Freight Centre. *Towards efficient and zero emissions global freight and logistics*. 2023. URL: `https://www.smartfreightcentre.org/en/`.

[114] Solutions. *Mandatory Emissions Reporting Around the Globe*. 2020. URL: `https://www.ul.com/news/mandatory-emissions-reporting-around-globe`.

[115] Shruti G Hegde Soumyalatha. “Study of IoT: understanding IoT architecture, applications, issues and challenges”. In: *1st International Conference on Innovations in Computing & Net-working (ICICN16), CSE, RRCE. International Journal of Advanced Networking & Applications*. Vol. 478. 2016.

[116] Ricardo Suarez-Bertoa and Covadonga Astorga. “Impact of cold temperature on Euro 6 passenger car emissions”. In: *Environmental pollution* 234 (2018), pp. 318–329.

[117] Yi Tan et al. “On-board sensor-based NO x emissions from heavy-duty diesel vehicles”. In: *Environmental science & technology* 53.9 (2019), pp. 5504–5511.

[118] Tipping, Andrew and Kauschke, Peter. *Shifting patterns - The future of the logistics industry*. 2016. URL: `https://www.pwc.com/gx/en/transportation-logistics/pdf/the-future-of-the-logistics-industry.pdf`.

[119] Trading Economics. *EU Carbon Permits*. 2023. URL: `https://tradingeconomics.com/commodity/carbon`.

[120] Wei-Tek Tsai et al. “Towards a scalable and robust multi-tenancy SaaS”. In: *Proceedings of the Second Asia-Pacific Symposium on Internetware*. 2010, pp. 1–15.

[121] Mengru Tu. “An exploratory study of Internet of Things (IoT) adoption intention in logistics and supply chain management: A mixed research approach”. In: *The International Journal of Logistics Management* (2018).

[122] Katherine White, Rishad Habib, and David J Hardisty. “How to SHIFT consumer behaviors to be more sustainable: A literature review and guiding framework”. In: *Journal of Marketing* 83.3 (2019), pp. 22–49.

[123] Zhiwei Xu and Jie Ni. “Research on network security of VPN technology”. In: *2020 International Conference on Information Science and Education (ICISE-IE)*. IEEE. 2020, pp. 539–542.

[124] Xingyu Xue et al. “Assessing decarbonization pathways of China’s heavy-duty trucks in a well-to-wheels perspective”. In: *The International Journal of Life Cycle Assessment* (2022), pp. 1–15.

[125] Ali Yavari et al. “ParcEMon: IoT Platform for Real-Time Parcel Level Last-Mile Delivery Greenhouse Gas Emissions Reporting and Management”. In: *Sensors* 22.19 (2022), p. 7380.

[126] Q. Yu et al. “Driving behavior impact on emissions of light-duty gasoline vehicle based on portable emission measurement system”. In: *Jiangsu Daxue Xuebao (Ziran Kexue Ban)/Journal of Jiangsu University (Natural Science Edition)* 43.3 (2022), pp. 270–276.

[127] Jaeseok Yun et al. “Implementation of sensing and actuation capabilities for IoT devices using oneM2M platforms”. In: *Sensors* 19.20 (2019), p. 4567.

[128] R van Zenden. “The identity mangement solution that improves data sharing in logistics”. MA thesis. University of Twente, 2023.

A

Anonymised interviews summaries with experts in the field

A.1. General structure of Experts Interviews

Each interview was conducted with a stakeholder and based on their knowledge of the given field. Every interview partner was asked for consent to be part of this research and informed about their rights to withdraw the consent. The participants agreed to be recorded and signed a consent form finalising it. The questions were the first starting point of the conversation and were formulated in an open manner to facilitate a conversation.

The general interview structure was as follows:

General Questions:

1. Does your company use any primary data gathered via onboard sensory systems to calculate GHG emissions for your road fleet?
   (a) If yes, which data parameters are you using for that?
2. Do you have an internal data management system that is able to integrate multiple data sources for the calculation of GHG emissions?
   (a) If you use external systems, which ones are you using and at what point are you accessing the data?
3. To which degree are you sharing your logistical data with other companies or organisations and why are you doing that?
   (a) If yes, how important are data sovereignty and anonymisation for you?

Technical Questions:

1. What are the main requirements for your company for a general system that is able to track, collect, store and process primary data of road vehicles for the calculations of GHG emission?
   (a) Granularity of the reporting
   (b) Accuracy of the data
   (c) Real-time monitoring
   (d) Compliance with existing standards
   (e) Data security
   (f) Tools
2. What are the main problems with the collection and transmission of primary vehicle data?
3. How are you transmitting your primary data from the vehicle to the database‘?
   (a) If you use sub-carriers, how are you managing their data transmission to you or other companies?
4. Which requirements do your customer have towards you when it comes to the calculations of GHG emissions, data transmission and data security?
5. Are you the data owner?
6. How are you evaluating the success of such a data management system?

Additional Questions:

1. Are you, in principle, interested in advanced data analytics based on primary vehicle data, such as driving behaviour analysis or emission-based decision-making?

2. How important is your collaboration with other stakeholders when it comes to data exchange and industry standards?
3. Have you recognised a general shift in the industry towards more sustainable logistical practices, which also influences the contract tendering process?
4. Do you have any other additional points that you would like to mention that were not sufficiently addressed here?

A.2. Interview partner summary

Table A.1: Interview partner overview

Expertise	Industry/ Function	Interview type/ date	Contribution (design)	Contribution (evaluation)
25 years in onboard software, control system, and real-time data acquisition	Sensor manufacture, CTO	Online semi-structured interview, 04.05.2023	Stakeholder, Requirements, Institutional Environment, Customer Onboarding, Data Pulling, General Improvements	None
15 years in emissions modelling, emissions factor creation	Emissions Modelling, Managing Director	Online semi-structured interview, 09.05.2023	Stakeholder, Requirements, Institutional Environment, Limitations, General Improvements	None
12 years in the logistical sector, recently logistical data	Data marketplace, CEO	Online semi-structured interview, 16.05.2023	Requirements, Standards, Data Pulling, Stakeholders, Business relations, Policies, General comments	System architecture, General comments
8 years in methodology development, logistical processes, independent consultant, carrier operations	Carrier, Senior Manager	Online semi-structured interview, 31.05.2023	Challenges, (Requirements), Current Practices, General comments	Stakeholder, Customer Onboarding, Requirements, System Architecture
4 years in sustainability, accounting and reporting of CO2 emissions	LSP, Project Manager	Online semi-structured interview, 06.06.2023	Current Practices, General comments	Stakeholder, Requirements, Customer Onboarding, Business relations, System Architecture, General comments
1 year in data management structures, optimisation, CO2 emissions, logistical processes	IT company, Data scientists & engineer	Online semi-structured interview, 09.06.2023	None	Stakeholder, Requirements, Tools and Concepts, Multi-tenant system architecture, Customer Onboarding, Data Pulling, System Architecture

A.3. Fuel consumption sensors and IT emissions tracking expert

Summary of the interview with an expert in the field of emissions measurement sensor systems and IT carbon tracking platforms.

Monitoring of fuel consumption data has multiple use cases and benefits apart from detailed and accurate emissions tracking, such as fuel theft prevention (this needs real-time monitoring and specialized equipment), optimization of driving behaviour or prevention of long idle times of the engine without movement of the truck. This reduces the amount of wasted fuel and thus saves not just emissions but also money for the logistical operator.

Fuel monitoring of logistical vehicles can happen in different ways. Older and less modern trucks need the installation of specialised equipment in their fuel tanks, while modern trucks are obliged to have dedicated software and equipment onboard systems that are able to monitor and store their fuel consumption under the FMS standard (The European Materials Handling Federation). The FMS standard closed the gap and the issue of the past where every manufacturer had their own standards for reporting emissions and recording of their fuel consumption and geo-locations. In extreme cases, the reporting standards and data frames were even different per vehicle type of the same manufactures. This made creating an automatic program to read out the data extremely time-consuming and inefficient. The new FMS standards grouped eight of the largest truck manufacturers and created uniform standards, which are able to record the data consistently and uniformly. It is, however, not possible to communicate with the vehicle or to give it commands, but only to receive data.

The FMS standard provides information about various parameters of the vehicle such as weight, fuel consumption, cruise control/ speed, or when the next service is needed. There are three ways to get the fuel level: via the FMS (that was highly inaccurate 10 years ago), via their own equipment inside the tank (results in fantastic results, but needs time and money for installation) or via their already sensory and systems on the vehicles, called rFMS (remote FMS). The system will transmit the data when there is a connection to a network, thus storing it onboard.

The rFMS does not provide real-time and not high-frequency data, but rather processed data throughout the day (for each leg or every 15 minutes). Real-time monitoring is not needed when you want to analyse the fuel consumption or driving behaviour, also possible to do it afterwards with the same results. However, even though there is now an agreed-upon standard (the rFMS), not every truck manufacturer is following it.
The new rFMS standard now allows companies to access the data of the truck in satisfactory quantity and frequency without the need to install any additional equipment. Companies that have already a detailed and large database of high-quality data in high frequencies and various vehicle classes can now smoothen and enrich the data of the rFMS systems and even add more third-party telemetric sources to it. Thus they can overcome the limitation of the rFMS data and get data of the highest quality to monitor and capture the fuel consumption, the speed or the idling of the vehicle and even the operator's driving behaviour.

The owner of the vehicle or truck needs to pay a subscription to the vehicle manufacturer to access the vehicle's rFMS system. The truck manufacturers are in charge of all security and privacy concerns/ protocols of the data sharing of the rFMS data and only provide it to third parties with written consent. Thus, third-party IT companies do not need to worry about it.

Tachigrapf (similar to the black box of an airplane) records all data from the driver of the vehicle, which is needed to prove to the authorities that the driver confirms with legislation about the breaks of the driver, geolocation or the maximal speed. This detailed information in combination with the driving behaviour, idle times and fuel efficiency can be used to train drivers that perform worst or even allocate a bonus to highly efficient drivers. This must be in alignment with the GDPR and is regulated via contracts between the carriers with their drivers (they have to consent to the data tracking) and via contracts between the carriers and the third-party IT companies.

The different stakeholders have different needs when it comes to the sensor systems. Fleet owners that are faced with a lot of fuel theft opt for additional equipment on top of the FMS/rFMS system to get alarmed when a theft is happening. Additional equipment enables real-time monitoring of the fuel-level of the vehicle and is able to send an alarm if fuel is removed without the movement of the vehicle. This is also the case for a long period of idling of the vehicle, here the FMS/rFMS can only show the waste of fuel and idling afterwards, while the additional equipment can show it now and prevent it. Shippers, on the other hand, only care about the fuel consumption data, thus are satisfied with the rFMS and the enriched data.

Shippers are starting to select carriers based on their consumption and sustainability scores, thus indirectly, carriers start to now heavily invest in tools and ways to be more sustainable. Thus, it is becoming a competitive advantage for carriers to have detailed information on a high frequency of their fuel consumption, idling or fuel theft on a representative sample of the fleet to prove this to shippers. Thus, they also equip part of their fleet with additional sensors on top of the FMS/rFMS system. Based on these detailed records, it is possible to create a digital twin of the vehicles and then predict the consumption and other vehicle parameters of other vehicles with no additional sensors on a very high level. Thus, current developments are striving for fleets with mixed-equipped sensors on a carrier level.

Collaboration and consistency in concepts and ontologies are fundamental for the work and reporting of emissions. Another important project is FEDeRATED which should reduce the time for new concepts to be adopted and accepted in the sector. The project is hoped to connect multiple platforms and let them talk naturally because it is currently a financial constraint to develop and connect to all different platforms.

Old sensors were very susceptible to the vehicle's noise of movement or vibration, which needed to be accounted for. Modern devices can account for it but are expensive. However, an investment on the site of the truck manufacturers could be a big boost to the sector to get better data. Another improvement could be an increase in the computation power of the systems of vehicle systems to not transmit processed data on low frequencies but rather nearly real-time data. Other mentioned challenges in the sector are:

1. The Interoperability of different platforms. Here the FEDeRATED project could help
2. The computing power of systems on the vehicles
3. The early status of the sector, thus there are not too many investments yet from carriers
4. The lack of detailed representation of fuel levels at all times to understand why more or less fuel was consumed. This is not relevant for the total emissions generated but rather for analysing and optimizing driving behaviour
5. Assumptions of vehicle classes based on consumption to then use an emission factor to determine the GHG emissions, thus limited information about vehicles (especially old ones)

6. Carriers do not want to transmit detailed data of their operations to shippers (competitive reasons). Here third-party IT platforms can be a way to have a neutral party in between, but it also leads to aggregated GHG emissions calculations rather than to transport leg-specific emissions
7. Slow-moving corporate structures that hinder the quick implementation of the sensor systems, the people need training and need to acquire knowledge about the system, but also at the first place that they have a problem. Here carbon pricing can help

A.4. Emissions factors and modelling expert

Summary of the interview with an expert in the field of emissions measurements and road transport simulations of various vehicle classes.

They developed an online tool to estimate emissions caused by road traffic, which also covers heavy-duty vehicles (HDV) and also were one of the main developers of the Copert emissions calculator. Their expertise resulted in a long collaboration with the European Commission and other European bodies and with a magnitude of clients ranging from logistical companies to shippers to automotive companies for various transport modes.

All real-life emissions testing is governed by precise rules and standards laid out in the European Commission regulation 2018/1832 amending the Directive 2007/46EC of the European Parliament and of the Council, and the Regulations (EC) No 692/2008 and (EU) 2017/1151). It specifies the street types of the testing, the compositions, the speeds, the measurement system PEMS (Portable emissions measurement system) itself, the time of the testing and all other related system boundaries to eliminate external factors such as driving behaviour. The testing is part of the „European Conformity Program“ of the European Commission with a very high level of granularity of the emissions on a seconds-level, thus very high reporting frequency.

Currently, they are following the standards of the GLEC framework and maintain and operate their simulations accordingly to these standards. Standards and collaboration, which align the methodologies of emissions tracking, reporting and measuring are highly important for them and the collaboration between industry, academia and regulatory bodies and policy-makers is one of the main drivers of that. This collaboration has also led to many of the default emission factors of known reference standards such as the GLEC framework or the new ISO 14083:2023 to be a joint product of multiple entities and emissions measurement experts. Another form of collaboration is happening on a data level for emissions factors under the umbrella of the „European research on mobile emissions sources (ERMES)“, which created a large database of emissions factors of various vehicle classes. Each participating organisation provides its measured (mostly national) emissions data to this database and can in return use the data of all other contributing parties for their research, as a model input or other activities. The raw data is only accessible to contributing partners and not open to the public. The European Commission is very cautious in making this raw information available, as special knowledge is needed to understand and process this data. One threat why they restrict access is the attempt to green-washing emissions data or interpret the data in a biased manner. However, the processed data is available to the public large.

New regulations, such as the OBFCM (On-board fuel consumption meter) mandate car manufacturers to track and report their emissions and used fuel to the European Commission. These new efforts will lead to an enrichment of the emissions data and databases, which will eliminate the need for dedicated emissions-measuring companies for common pollutants such as CO2. The main reason behind this and other similar regulations is to monitor the efforts and commitments of car manufacturers and logistical companies of reducing their emissions. A side benefit is the improved accuracy of models and predictions that are then based on real-world data and driving conditions.

The status quo of emissions prediction models for larger geographical areas is highly complex and requires more effort than determining an emissions factor for one vehicle class. Dedicated models such as the COPERT emissions calculation model were designed to exactly determine emissions on a national level, however, the current development of obtaining real-world emissions data of modern trucks could eliminate the need for such models in the future if all central organisations would group, collect and calculate the caused emissions for each country directly based on the primary truck data. However, this is only the case for modern vehicles that have the ability to report their data and older trucks do not have that.

One current limitation is the number of pollutants, especially non-regulated pollutants. There is much information available on common pollutants such as carbon dioxide, however, other like ammonia or hydrocarbons, dioxides have less information available. Another part that is currently lacking is non-emissions related toxins, such as used rubber of tires, breaks and road surface factors. More research and effort are needed to also complete the database for these factors.

The main drivers for their customers to invest time and resources into emissions tracking are carbon accounting and the desire to decarbonise their operations, which is directly influenced by upcoming regulations and mandatory emissions reporting. The developed programs and models are a tool to firstly track and report their emissions, but also to show their efforts and progress in the decarbonization attempts. Nowadays many end customers also desire and demand sustainable products, thus shippers select their carriers based on their carbon footprint and emissions. Thus, carriers directly reduce the emissions of companies' scope 3 emissions.

They assume that the new ISO14083 will not critically impact their work and current models, as it is based on and aligned with the GLEC standard. However, they see the general need for the ISO norm, as the ISO brand name makes it more official and will help to increase the credibility. One major requirement is the alignment of all existing methodologies without contradictions.

A.5. Data marketplace and IoT aggregator expert

Summary of the interview with an expert in the field of data market platforms and emissions calculations.

The interviewed company is a data marketplace and IoT sensor aggregator that has connections and partnerships with hundreds of third-party telematics and onboard sensor providers and also has direct partnerships with multiple OEMs that provide them with various data. They work and operate mainly through API integration's, where they facilitate and enable data collection and sharing in companies to give them insights into their emitted emissions in their supply chain.

They categorize their customers as Data Consumers and work with all relevant logistical actors such as carriers, shippers, OEMs, research groups, insurance companies, fast payment companies, and third & fourth-party (3PL & 4PL) logistical actors. They offer their service to all actors that need increased transparency in their operations. But the logistical (main) actors are their main customers.

The different actors have different requirements and needs when it comes to the data that they want to use. However, every customer has a similar minimum requirement of data that they need (GPS position, ignition data), but the advanced requirements differ. I.e. shippers and actors that operate cross-docks are really interested in the percentage of empty or half-empty trips and loading factors compared to others, while insurance companies mainly care about risks and where the shipment was at what times.

Shippers and large corporations work on an "event basis" with their shipments. Meaning that they care about their shipments in specific legs/ corridors where problems occur more frequently. Thus, they might only want one-data point for a specific leg, but then still have the possibility to increase the frequency for the legs that are continuously troubling or delayed (however they would still ask for a 1-minute ping rate with great latency). In the end, they want to report their emissions on an aggregated level for a leg (preventive vs. hindsight problem-solving). Large companies see data as a tool to not be caught off-guard when shipments are delayed and not as a tool to optimise and restructure logistics (they own many semi-independent logistical sites & vehicles and do not see it as a large cohesive network).

The interview partner categorises their data provider into different categories. Their data provider used to be mainly third-party telematics providers (8 to 10 dominate the market with a long tail of smaller ones that increase the complexity of a common standard), now some OEMs provide much data through their FMS & rFMS standards.

The data of the telematics provider is received by pulling it (with the consent of the asset owner), the OEM create their own TMS/ FMS and they only have access to their vehicles. Thus, the interviewed company needs to integrate/combine multiple OEMs and vehicle-specific platforms. They also buy and sell the data of truck OEMs to trailer OEMs and vice versa, thus that everyone is able to create and compile their own database/ ecosystem (no interoperability).

They have two different strategies to deal with these various standards and formats of their data providers: First, an important aspect of their work is also the expectation setting with their customers about what is currently possible. The OEM data gets them the best data quality, ping rate and accuracy, however, the majority of the data is from telematics providers with 15 to 30-minute ping rates with two to three (maximal five) data points per trip. Second, the enrichment of the data by increasing the frequency and detail of it by using their knowledge and insights of what happens in between two data points (they use an external company for that). However, many data consumers are not technically inclined (problems here could be data backlogs), thus every customer needs a specific treatment for the details.

Their offer is the connection with one API, which guarantees a more usable dataset with common standards to their customers. The data that they receive via rFMS/ OBDS or telematics are collected on the vehicle and batched and sent every i.e. 5 minutes to the corresponding OEM and their database. The data engineers at the OEM then process/clean the data to prepare the final parameters of the vehicles (fuel consumption, speed, load), which then can be bought and directly pulled by third-party companies. Thus, they (the interviewed company) created a joint venture with the OEMs. The asset owner (carrier) is legally the data owner, however, they normally never use/ get the data, as it goes directly through the third-party company (with consent from the asset owner), who then also becomes the data owner. Currently, the OEMs are only monitoring around 7% of the data that they collect from their rFMS. The asset owner pays the third-party company for their service via a subscription, the third-party company is paying the OEMs for their rFMS data and then monitoring it for the asset owners and other carriers.

They have 4 different data sources: telematics, OBDS, rFMS/FMS, and IoT (IoT is for them interchangeable with third-party telemetrics data and the carriers pay for it). For carriers, the advantage of a third-party IoT sensor compared to the rFMS system is the common standard. The rFMS system might have different standards per OEM, however, the data quality of the IoT is not good (low ping rate and few transmitted parameters). Around 80% of the rFMS data format is the same for all OEMs and it is possible to get hundreds of various data points every minute. rFMS gives more insight into real-time fuel consumption and carbon emissions, load factors, road slopes etc.

The time it takes for the data to be created at the vehicle to arrive at the third party is minutes and the internal system architecture is real-time (21 ms, the industry average is around one minute). It takes the most time for the data to be batched at the vehicle and sent to the OEM. Currently, the real-time availability of the data is less important, but they see it as a USP in 3 to 5 years. Quick access is more relevant for highway control systems.

Their data-sharing is directly related to the GDPR and they have explicit contracts with each carrier and shipper of which data is allowed to be shared with whom in which situations. They mainly present aggregated and processed data and not raw data, thus anonymization is less important to them. The carriers don't want to share the data as well and the sub-contracting is also a challenge as those are considered secret to many.

The interviewed company sees everyone talking about data sharing and interoperability. Everyone wants to have data from others, but no one wants to share theirs, which is slowing down the entire sector (as there is also no clear market-dominated party).

For the onboarding process, they have two different methods: First, they use push APIs (as these are quicker to develop), which are connected to the client's database. Then they will cross-reference all the carriers that the company uses with their own database to identify which carriers need new contracts and which ones have already an agreement with them. The agreement will happen through a unique personalized link, which the shipper must send to their carrier to see their data through the interviewed company (due to GDPR compliance). Then the interviewed company will get in contact with the carrier and identify which sensors and vehicles they use. If they use third-party sensor technology that the interviewed company does not have an agreement with they need to again ask for consent to then reach out to the carrier again to tell them everything is set up. The set-up for new customers/ carriers/ third-party companies takes between 2 days to a week. Existing/ known carriers and technologies will happen within 48 hours.

Second, they have created a no-login solution with an OEM, thus once the consent is set up they can pull the data from their database (with licence plate information) within 5 to 10 minutes (here the issue is that the OEMs are sometimes not updating their database).

The shipper will most of the time only access the processed data via third-party companies. The interviewed company also just started looking at the mandatory reporting to European bodies for their clients. The interviewed company has achieved savings of 16% in emissions for their customers via the GLEC framework and then managed to reduce the carbon emissions by another 20% by adding more data points to it like gross axel weight etc. They want to show companies the potential to reduce emissions and the money that can be saved there as a benchmark analysis. They also offer free driving behaviour analysis of the vehicle operators, however, the shippers are not really concerned about it (this is more important for the carriers). However, the services are meant as an incentive to the companies so that they share their data among others like quick payments.

For them, the interoperability projects (like FEDeRATED) are just the beginning for interoperability, common languages and spaces. However, many of those projects are run by logistical supervisors with little information about data and data structures. They are not sure in which final form these projects will end up, but they hope it will lead to increased interoperability.

Third parties can bridge the gap between logistical operators by acting as a neutral party in the middle with which the other logistical operators are less reluctant to share their data. This also means that the OEM/accessed company is mainly responsible for the security (when they directly access the database), however, everyone that touches and works with the data has to be responsible for it. However, the fragmentation and segregation of all companies and standards is the main challenge.

For logistical actors, the main incentive next to the mandatory emissions reporting is the potential costs that can be cut with improved operations based on data insights. Many companies that operate in these industries have a lot of secrets, thus are less likely to get a lot of data that would increase transparency. Another point for the slow implementation of their systems are the slow-moving corporate structures and that the logistical expenses are just a fraction of their total costs (compared to R&D etc.), as the systems architectures would need new IT skill and would need a lot of work to first align all their existing systems in the various semi-independent warehouses.

For the interviewed company their API connection into existing systems is a real USP. Old competitors were pushing platforms, which resulted in a magnitude of them. Five to six years ago was a VC boom, which lead to the founding and creation of many platform-based third-party IT companies. Now many large shippers are opting for their own in-house systems architecture solutions, which creates an ecosystem around their supply chain. Thus, their API connection to existing systems is still very useful. In the end, shippers want the cheapest and quickest system for their operations.

Data storage and transmission is not the main issue in their opinion, as much of it is automized and cheap (as a reference they pay between 5 to 15 cents per vehicle per moths for the supporting data structures in the back). A main requirement for their system is always scalability (via cloud storage AWS and their queue system).

General improvements in the space are firstly alignment of regulations that make the standards and requirements of each Member State equal. Secondly, it would be important for them that regulators would separate the vehicle data and human-based data and distinguishes them under the GDPR. Currently, all vehicle data is still protected under GDPR (because a human is in the car), however, they are not interested in the data of the human, but they are interested in the data of the vehicle. Thus, they want clear instructions on which type of vehicle data can be easier shared and used under the GDPR, which has nothing to do with the human operator. On a macro level, they want to see more consolidation. Lastly, they see the sub-contracting eco-system as a problem, which creates a lot of waste and inefficiencies (also for the data gathering and processing). Another idea to use data without interfering with the GDPR is to equip part of the fleet with highly detailed sensors to then aggregate the data to the entire fleet level to then use the data for modelling and optimisation proposes, thus not using data of humans and therefore no/ less GDPR restrictions.

A.6. Logistical actor (carrier)

Summary of the interview with a logistical operator (carrier).
The carrier company uses external IoT sensors for all their vehicles and the entire fleet. They use for that a third-party company that produces the sensors and also does preliminary data cleaning and processing. They access the data via the third-party company and use it on an operational level for the geo-tracking and scheduling of their vehicles but also for regulatory reporting such as driving times of drivers and emissions reporting. For that, they use the total fuel consumption of the vehicle and the consumption values for specific legs and corridors. The data, which is used for reporting is not accessed by them, but by another third-party emissions calculation specialist, who can directly access the data at the third-party IoT manufacturer (with their consent). They first had their own structure and architecture for the calculations but decided to outsource it. The reason for that was that they wanted to focus more on the reduction of emissions and not on the calculations of them (thus resource management).

They started to work with the company as the system was more agile than the traditional OBDS and rFMS tools and the third-party company offered dashboards and monitoring opportunities, which the other did not have.

Even though the data collection and management are managed by external experts there are still problems that occur. These problems mainly occur on two levels: first the technical implementation and second data quality, consistency and what to then do with the available data.

On a technical level, the connection and the equipping of their vehicles have to be perfect. Here common issues are maintenance issues or issues at the sensor manufacturers' sites, but also the connection and transmission of data or a defective sensor itself can be challenges. Another issue could also be that data is only partly transmitted and is not complete.

For them, it is always a consideration between effort vs benefits. The provided data might not be 100% accurate, but the deviations are still acceptable. Thus, here it is a challenge to not get lost in the details.

On a data quality level, enriching the data with multiple and external data sets is a challenge. This is because many of the different parameters have not been combined before (i.e. fuel consumption per weight kilometre within a specific transport leg). Thus there are only a few reference points and benchmarks for it, which makes it time intensive. Another point is the insights from the data, which will not have immediate benefits but will be more important in the future rather than now.

In the end, the challenges are a combination of resource issues and know-how issues. The resource issue is about setting up appropriate API connections and maintaining it. This has been done before but takes time to properly set up, which is often inefficient for companies. The know-how issue is about the proper GHG emissions methodologies and calculation steps, which only experts know in detail. The company is a logistical carrier for only other logistical actors (only for other carriers and LSPs) and they proactively offer the sharing to their customers of primary data, however, the desire to get those primary data points is currently very little. What they would like to see is that companies that are customers of the same external GHG emissions IT company could automatically share and access the data (with consent), thus no additional work would be needed. This could be interesting for bench-marking and reduction targets and alternative and eco-friendly fuels.

Here the usage and calculations of primary data of the exact fuel consumption per shipment with eco-friendly fuels could be used for investments. Meaning that dashboards and GHG emissions calculation tools could plot their savings (based on empirical primary and actual data) with these eco-friendly fuels to justify investments. This could be used for insetting and off-setting of GHG emissions and other reporting measures (these tools and monitoring do currently not exists).

Another advantage of these external data platforms is the validation. Thus, the interviewed company could refer their customers to the external and certified GHG emissions calculations company to show them the exact and validated savings. Here authorized access to the platform is very important. However, they also stressed that they do not like sharing their raw and rFMS data with their customers, as it would give too much insight into their operations but would also mean additional data protection and GDPR considerations.

Their main requirements for such a system would first be data security and compliance with EU regulations. Here especially the creation of driver profiles based on the tracked data is an issue, as the vehicle data can be traced back to an individual driver. The latency of the system (thus real-time monitoring) is less important for them, but what is important is the quality and the completeness of the data. However, in the end, they are a logistical service provider and not a data specialist. Thus, they care very much about that their third-party companies have the right data, but they do not want to expand their business into data management. The data reporting frequency is less important, as long as they have and can determine fuel consumption on a leg/route basis (currently they have every 5 minutes).

The data transfer from the vehicle to their external IoT sensor manufacturers is similar to the transmission of rFMS. The data is batched and collected at the vehicle and then sent to the IoT provider. There the data is processed and can be accessed (with consent) by other third-party companies and them.

A.7. Logistical actor (LSP)

Summary of the interview with a Logistical Service Provider (LSP)

They have structured their business into different segments, that are different in terms of further sub-contracting logistical services. The first segment is the classical „carrier“ segment, which they completely outsource to other sub-contractors and they do not own a fleet for this segment. The second segment is for equipment and intermediate parts for the industry, for which they operate around 50% of the orders themselves and the rest is again outsourced with some sub-carriers being officially employed by them while some receive the order via the spot market (like a tendering process for large orders).

They calculate their emissions for their own fleet via the already equipped telematics systems (they use the rFMS system for that) on the vehicles to calculate specific and vehicle and shipment-specific emissions values and emissions intensity values. Their approach is very closely connected to the new characteristics approach of ISO 14083, which categorizes the transport chains based on Transport operation characteristics (TOCs). They are able to retrieve all direct emissions and loading factors, but also indirect emissions (emissions caused for temperature-controlled transport or empty runnings etc.).

For their sub-carriers, they do not have a direct telemetrics connection. Now they work together with an external IT company that has it, but in the past, they would receive the data about emissions from their subcontractors on a yearly bases via an Excel sheet, that they provided. They tried to not use the default values of known GHG emissions methodologies, such as the GLEC framework, as they would not have any improvements in their emissions, only if they reduce the total transport amount.

Thus, far they do not have any automation software that calculates or manages their data. They used to calculate their emissions yearly for reporting. However, they contracted an external IT company to improve their data systems and reporting frequency and to have the IT company directly linked to the subcontractor. Thus, far they calculated the exact emissions of the subcontractor by multiplying the reported emissions with their own factor to account for the age and conditions of the vehicles (thus they aggregated the reported data from subcontractors based on their own emissions and expertise).

They on purpose did not request any direct connections to the raw data of the subcontractors to ensure data privacy and compliance, but also because they know that the subcontractor makes their profits with the fuels (as this is about one-third of their costs) and optimized planning. The data that they receive from the contractors are only vehicle data (age, mileage, weight, etc.) and are only used for the GHG emissions calculation because they are available. If they report emissions for customers, they will never show specific values of shipments and report them on an aggregated and abstract level to not violate data privacy.

For them a monthly reporting of emissions values is sufficient and they do not need live monitoring of the data on a weekly basis. However, want they strive for is quarterly standardised reporting to identify possible outliers.

Current issues with the data collection are occasional malfunctions of the sensors themselves, which stops the data tracking. Their main reason to start collaborating with external IT companies was the desire to move away from default emission values towards primary data, without them having direct access to their subcontractor's raw data. They were looking for a neutral middleman that connects both sites. These connections would also be (theoretically) possible from them, however, this would use a lot of resources from their site. Another point is the official accreditation of the external company, which increases the validity of their reports. Lastly, the automation process and frequency of reporting are a bonus for them (on a quarterly level).

Currently, around 60% of their emissions from subcontractors are done based on primary data, but they want to increase this further. They will not be able to connect every subcontractor with primary data, as some orders are still sold to their subcontractors via the spot market.

Their intentions and KPIs for onboard sensory systems to calculate GHG emissions are to generally reduce their emissions, however, the data system is just a tool for that. A goal is of them is to have around 80% of their contractor's calculations (the value is based on the experts' opinion), based on primary data, however, they never formulated explicit KPIs for such a system (they see it as a binary tool, you can either calculate the data on primary data or you cannot). They also see the benefit of the cost savings based on GHG emissions reductions, however, this is not a driver and they also do not want to create price pressure towards their contractors, by saying you optimise your operations and drove less (thus also fewer emissions) and now we pay you less, as you in theory had less effort. They rather want to work together with their subcontractors on collaborative projects to reduce general emissions.

They also see a general shift from the shippers towards more detailed estimations of GHG emissions already in the tendering process of contracts, however, not as a final decision criteria yet, as this is still too vague for that. Shippers are still very keen on working with the carriers to find ways to reduce emissions and the expert thinks that the new ISO and other standardised and automated GHG emissions calculations will help the industry to make the reporting of emissions more lenient. They are currently not asking their subcontractors for such emissions approximations, since the transport market is highly competitive and transport services are scarce, thus it is often not possible to not select a carrier if they are unsustainable. The pressure from an operative point of view is still too high, but they are also creating sustainable joint projects with their carriers to improve sustainability and they believe that the pressure for sustainable transport will increase, thus that they can also ask for it in the future.

They also have multiple collaborations with other industry players within the logistical sector to create common standards, however, the complexity of the transport industry due to the magnitude of players decreases the speed. It is a joint effort and no single stakeholder can make a difference individually.

They are already using advanced analytics of driving behaviour to evaluate their own drivers, and if a driver scores below a threshold they have to do extra classes to improve their efficiencies, but they do not do it for their subcontractors. If they would do it for their subcontractors, the granularity would be on a fleet level and not on a driver level. The usage of subcarriers and other sub subcarriers is daily business in the logistical sector, but also reduces the clarity of reporting and increases the complexity of the data transmission.

B

Informal Discussions / Logbook

Table B.1: List of informal conversations that resulted in a knowledge gain

Stakeholder (cluster)	Date and Place	Knowledge gained	Type of conversation	Abbreviation
Verification, Validation and assurance company (VVA)	18.04.2023, Amsterdam	Data foundation of accounting Evidence of data generation is needed Challenges: - Multiple data sources - Multimodal transport specifics - Access to Scope 3 data - Unit management	Keynote speech	I.1
Verification, Validation and assurance company (VVA)	18.04.2023, Amsterdam	Elimination of double counting via leg-specific calculating Needed data transparency Risk mitigation via detailed reporting	Keynote speech	I.2
Shipper, Carrier, VVA, company IT	19.04.2023, Amsterdam	Sustainability needs to become a core value and requires training Execution gap (strategies vs actions)	Keynote speech	I.3
VVA	19.04.2023, Amsterdam	Which primary data to use Problem of sub-sub carriers and Scope 3	Keynote speech	I.4
Sensor manufactures	19.04.2023, Amsterdam	Advanced sensors with high reporting frequencies need data management strategies	Informal conversation	I.5
VVA	19.04.2023, Amsterdam	Scope 3 GHG emissions normally do not get reasonable verification, but rather limited verification Industry-standard has less accurate S3 GHG emissions reporting because they do not know them Use more EF databases instead of primary data	Informal conversation	I.6
LSP	19.04.2023, Amsterdam	Large LSPs leverage the power that they have over carriers for net-zero projects to reduce S3 emissions	Keynote speech/ Informal conversation	I.7
Shipper	19.04.2023, Amsterdam	Currently do not use primary data and do not see the need to use it in future The carriers and primary data are not connected Use external services for emissions calculations	Informal conversation	I.8
Carrier	19.04.2023, Amsterdam	Leg-specific data sharing between shippers and carriers main issue Issues with sub-sub carriers	Informal conversation	I.9
Carrier	19.04.2023, Amsterdam	Shippers request more detailed emissions data (for S3)	Informal conversation	I.10
Shipper	19.04.2023, Amsterdam	Leg-specific data sharing between shippers' and carriers' main issue	Informal conversation	I.11
Carriers, Shippers, LSP	19.04.2023, Amsterdam	Sustainability is promoted by collaboration and long-term contracts (gives assurance for capital investments) Reward asset owners and add flexibility to contracts for the optimization of transport chains	Keynote speech	I.12
IT company	10.04.2023, Hamburg	Data alignment main issue	Informal conversation	I.13
Shipper	05.05.2023, Hamburg	If sustainability is a core value, then potentially company policies for in-house emissions determination	Informal conversation	I.14
IT Specialist	08.06.2023, Online	A main challenge in Multi-tenancy system architectures is the correct access and authority determination of users	Informal conversation	I.15

C
Evaluation Interviews

C.1. General structure of evaluation interviews

Each evaluation interview was conducted with a stakeholder and based on their knowledge of the given field. Every interview partner was asked for consent to be part of this research and informed about their rights to withdraw the consent. The participants agreed to be recorded and signed a consent form finalising it. Each evaluation step started with a brief explanation of the presented concepts and the context. It was asked if they have additional questions about it and if that was not the case the evaluation started. The evaluation of specific artefacts or concepts was removed from the process if an interview partner could not/ was not allowed to or did not feel confident about the topic or artefact.

What do I want to evaluate:
Stakeholder:

1. Do you understand the overview or do you have any questions about it?
2. Is the stakeholder analysis and overview complete from your point of view or are you missing specific stakeholders or relations between them?
3. Do you agree with the selected scope of the stakeholder analysis?
4. Does this abstract representation display a realistic overview of the status quo?
5. Presented image is 4.6 (Stakeholder overview)

Requirements:

1. Do you understand the overview or do you have any questions about it?
2. Do you agree with the three clusters of requirements?
3. Are the presented requirements relevant to you and your organisation?
4. If not, which ones are less and more relevant to you?
5. Are you missing any important requirements?
6. Presented Requirements table 5.4 (Requirements table)

Designs: For each of the different graphics:

1. Is the shown graphic clear to you or do you have any questions about it?
2. Does the image represent a realistic (abstract) view of reality from your expertise?
3. Are any parts/ aspects missing or is the representation complete?
4. Designs evaluated:
   (a) Presented design options overview 6.4
   (b) Multi-tenant system architecture 6.6
   (c) Customer onboarding 6.5
   (d) Business relations 6.10
   (e) Data Pulling 6.7
   (f) System architecture 6.11

Per design artefact

1. Presented Fig. 6.4 (System design options)
   (a) Are the presented design options able to address the challenges realistically and sufficiently?
   (b) If not, which ones would you add?

Per design artefact

1. Presented Fig. 6.7 (Data Pulling)
   (a) Are the times for the data pulling realistic?

Per design artefact

1. Presented Fig. 6.10 (Business relations)
   (a) Are the shown relations accurate and realistic?
   (b) Would you add stakeholders?

Per design artefact

1. Presented Fig. 6.11 (Abstract system architecture)
   (a) Is this a realistic (abstract) representation of data and consent flows?

Additional comments:

1. Would you like to add comments to either category, which currently was not discussed yet?
2. Do you agree with the evaluation criteria and the selected domains? If not why?

C.2. Evaluation with a data marketplace expert

Requirements:

1. "The scalability of the system is highly relevant and we do not want to cut costs there (TNFR.1)

Architecture:

1. "Exactly, I mean that really is the chain. The only thing worth putting in there is a lot of times the freight operator that the shipper contracts is not an asset-based company. So there's like a, [..] 3PL essentially that will organize that transport with a freight operator and actual company. So you get one more layer of essentially the shipper giving this order to a 3PL, the 3PL contracting the fleet operator."

Additional comments:

1. "It is important to create common standards with all relevant industry players, that are consistent throughout all European Member States"
2. Important next steps from a regulatory point of view is the separation of vehicle data parameters and human data points, that operate the vehicle. This will make the processing under the GDPR more lenient and easier.
3. "Independent middle companies could be the solution to more data sharing in the logistical sector and could bridge shippers and their LSPs"

C.3. Evaluation with a carrier

Stakeholder:

1. Very good visualisation
2. Good idea to cluster them
3. Potential to add the fuel producer (especially for SAF and bio-fuels) and shareholders for reporting issues

Requirements:

1. The most important requirements for them are:
   (a) The third-party accreditation after ISO or GLEC standards (SNFR.2)
   (b) Data Privacy concerns, also the server location (IFR.1)
   (c) The data consistency (TNFR.3)
   (d) And the restricted access to the data from all other parties (SFR.1)
      i. "We accept a direct connection to our database if it is clearly regulated which data will be accessed and what data will be made available to the shipper, however, only in a very small scope, as we do not see the added value of sharing data beyond the fuel consumption of our fleet."
2. "For us real-time monitoring of the data is not important, the same is the case for the system latency"

Customer Onboarding:

1. "We were onboarded like that"

Architecture:

1. "Yes, it looks correct and consistent, like an idealised version of the data flow in the system."

C.4. Evaluation with LSP

Evaluation with a project manager in sustainability and energy management at a logistical actor. The focus of the project manager is the accounting and reporting of CO2 emissions, that were caused by their transport activities for the entire organisation. They also create individual reporting for customers and recently increased their reporting scope to scope 3 emissions.

Stakeholder

1. “The first ideas I would have are banks and credit companies, which evaluate the sustainability score of companies and would give credits and financing based on that. For us, that is a main driver to have also green transport to get positive scores from them and thus also financing.”
2. “Another point is shareholders, which can also be municipalities, which also expect sustainability reports from companies that they invest in”.
3. ”But in general very realistic representation”

Requirements

1. A first explanation for the requirements was needed, however, afterwards, they were very clear
   (a) SNFR.6
   (b) SNR.1
2. All requirements are correct and relevant, however, some are more relevant for them, such as:
   (a) SNFR.2 / IFR.3 = “An officially accredited system (ISO or GLEC) is very important to us for our customers”
   (b) IFR.1 = „We do not want to have a direct connection to the subcontractors to stay within data privacy regulations and to respect the sovereignty and business model of the carriers“.
   (c) SNFR.5 = “We hire an external company to worry less about the IT, thus we want an easy system”

Architecture

1. “The system and the architecture look realistic and logical, however, it is possible that there is another complexity layer (in the form of another stakeholder) in the system, but we outsource these activities to focus more on our main logistical services”

Customer Onboarding

1. ”When we onboarded new carriers to external IT companies, we tried to have a conversation with everyone to express our intention behind it and to ensure them that they keep the data sovereignty, that the consent can be withdrawn at any point and that no additional costs will occur, after that the external company send a unique verification link to establish the collaboration and to receive the consent, thus as portrayed in this sequence”

Business Relationships

1. Is realistic to me, however, I did not know that some external IT companies have direct connections to the OEMs and pay them for part of the data”

General comments

1. „We were looking for someone that could help us to move from default emission values to primary data without us getting direct access to the raw data of our carriers. We were looking for a middle company that is able to get data from both sites, clean them, match them and then report the final emissions back to us.“

C.5. Evaluation with external IT company

Summary of the evaluation interview with a data expert in calculations of GHG emissions in road logistics from an external IT company.

Stakeholder:

1. „Many external IT companies take on multiple roles in this representation and could be identified as both a „Supporting IT company“ and as well an „Emissions modelling & calculation firm“
2. „I will support the statement that the end customer creates pressure on the shipper to reduce emissions, but there are also companies such as carrier and LSP, that create this pressure, however, the scope of this representation is clear and correct“
3. „The scope represents a good level of detail to the system and would not require the addition of OEM or fuel providers to be included in this stakeholder overview“
4. „This level of detail makes the most sense to me, if you would increase the scope until scope 3 emissions then the main relations could be lost in the complexity“

5. „It is sufficient for the calculation of emissions factors“
6. „The image gives a good and realistic abstract representation of relevant stakeholders in the area of GHG emissions calculation in road logistics“

Requirements

1. „I think it makes sense to group the various requirements into the different clusters to showcase the main application point and to also clearly separate the technical requirements from the rest“
2. „Where and how did you make cuts (in terms of the scope) for the requirements?“
3. „The used scope is very useful for such a general system, however, if you would want to make such a system for a specific stakeholder you could think about using a more detailed scope (i.e. specify the exact reporting frequency for SNFR.4)“
4. „For me, the importance of multiple data sources and telematics system is very important for such a system from a developer point of view“
5. „I would need a more detailed description for SNFR.4 to fully understand it“ Explanation from within the text was given, which then made it clear
6. „The requirement SFR.2 is very broad and general, and should be made very clear what is meant by it“
7. „None of the requirements seems to be completely off-target or inappropriate and I see that they might have different importance per stakeholder“

Design options:

1. When asked about Aggregation & Anonymisation
    (a) „We present the data to our customers the way we receive them, thus if a client sends us data with driver names in it we would reach out to them and inform them of potential data breaches on their site. Normally they ask us to delete these data points, however, the concepts of Aggregation & Anonymisation are used the least when talking about data privacy, for that we use tools, such as the mentioned TLS“
    (b) i. Addition to the above statement: We would also immediately delete this information (i.e. the names of drivers) without the customer telling us. We also built structures in our system that directly delete these kinds of information when we receive the data from our customers, thus that they never are uploaded to the database.
2. „In my opinion, it would be good to explain, that these presented design options are a way to reach the goal of the design topic, but not all design options have to be used for it, it might be sufficient to only use one“
3. When asked about multi-tenant-system architecture:
    (a) „All our customers have a unique login with different roles and rights, but no customer has ever access to any other database of a different client as you said here. We have a general database for the first log-in with hashed passwords, but once the user is log-in all actions only happen on the own database. Even if a shipper has two additional carriers with us, these databases are all separated“
4. When asked about Access control:
    (a) „We are currently not using a multiple factor authentication, as it requires additional infrastructure (even over known providers such as Google Authentication) and all customers would need to do additional verification steps, however, it is a good and valid idea and might become more important in the future if we continue to grow. It might also become mandatory from a policy site in the future, thus a valid tool“.
5. When asked about interoperability:
    (a) „We like APIs the most, as these make it very easy and flexible for us. Other methods for data exchange, such as SFTP, always require the action of the data providers, which increases the risk of inconsistent data transfer or mistakes (such as the transmission of data that is not meant for them) and increases the efforts for them“.
    (b) „Joint ventures are a very interesting concept and we might create something similar in the future to get direct data access. Joint Venture have the possibility to increase the speed of the data transfer, especially with the possibility to also create queries for the data.“
    (c) „When clients send us data, they must understand what type of data we need, in which formats and what times. This issue can be bridged with a Joint Venture or direct access via pull APIs.“
    (d) „It is more efficient, but of course also increases the complexity in regards to data privacy and GDPR“
6. When asked about internal data privacy:
    (a) „The location of the servers has cost considerations (non-European are cheaper), but the European Server has a faster response time compared to servers in i.e. China. For us, the topic is important but not a daily discussion point“.

(b) „We use the format TLS 1.2 for our encryption“

Multi-tenant system architecture:

1. „Each customer has their own database and no access to others“
2. „If we have to migrate data from one database (i.e. the carrier) to another (i.e. to the corresponding shipper) to calculate or match order and transport data, then the data is only stored temporarily in a complex matter (the shipper would still not have access to that data, only we as a company can see it) and removed again, once the calculations are complete.“
3. „The data and databases are never directly linked and we use scripts to do this automatically and securely“

Customer Onboarding:

1. „We always need consent to access the data from the data owner (can be carrier or LSP) for the shipper. Even if we have already worked with both the shipper and the carrier before in different projects we again need to receive consent for this specific partnership”
2. „If a carrier contracts other subcarriers we need consent from each layer“
3. „The collaboration with the main shipper can already start in advance, but only with the data, once all consents forms are signed.“

Business Relationships:

1. „Image looks very plausible, possible to have also a connection between the external company and the fleet operator or between the shipper and the OEM“

Data Pulling:

1. „We use the terminology of RESTAPI instead of pull-API“
2. „Realistic representation.“
3. ”Also possible that the vehicles themselves push the data to the OEM because it could be possible that the vehicle does not have a connection at the moment of the pull request of the OEM.“
4. „The pull is always happening at the last layer (from OEM to external IT company) beforehand it could also be another technique. You could use the terminology data transmitting to include all processes.“
5. „The processing time at the external IT company is dependent on the service, a process can also take hours or days, depending on API connection etc.“
6. „Make clearer that each vehicle has a storage solution on board, but of course no traditional large database.“

Architecture:

1. „Also the possibility to have external services and not databases for the calculations. They have their own database, but we only use their output and not the data.“
2. „Very plausible abstract representation of such a process“

General comments

1. No additional comments or questions

D
Additional Documents

D.1. First Designs

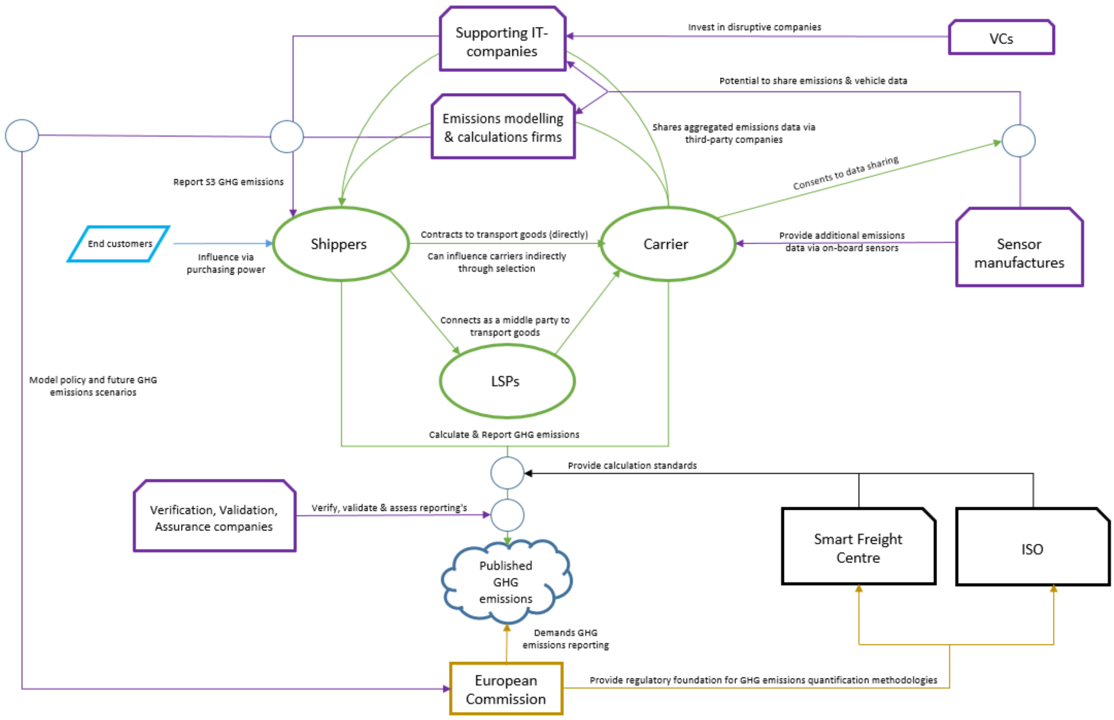

Figure D.1: Most relevant stakeholders with main relationships between them, Logistical Actors and relations in green, Policymakers in brown, Framework and Standard owners in black, External third parties in purple, and Customers in light blue.

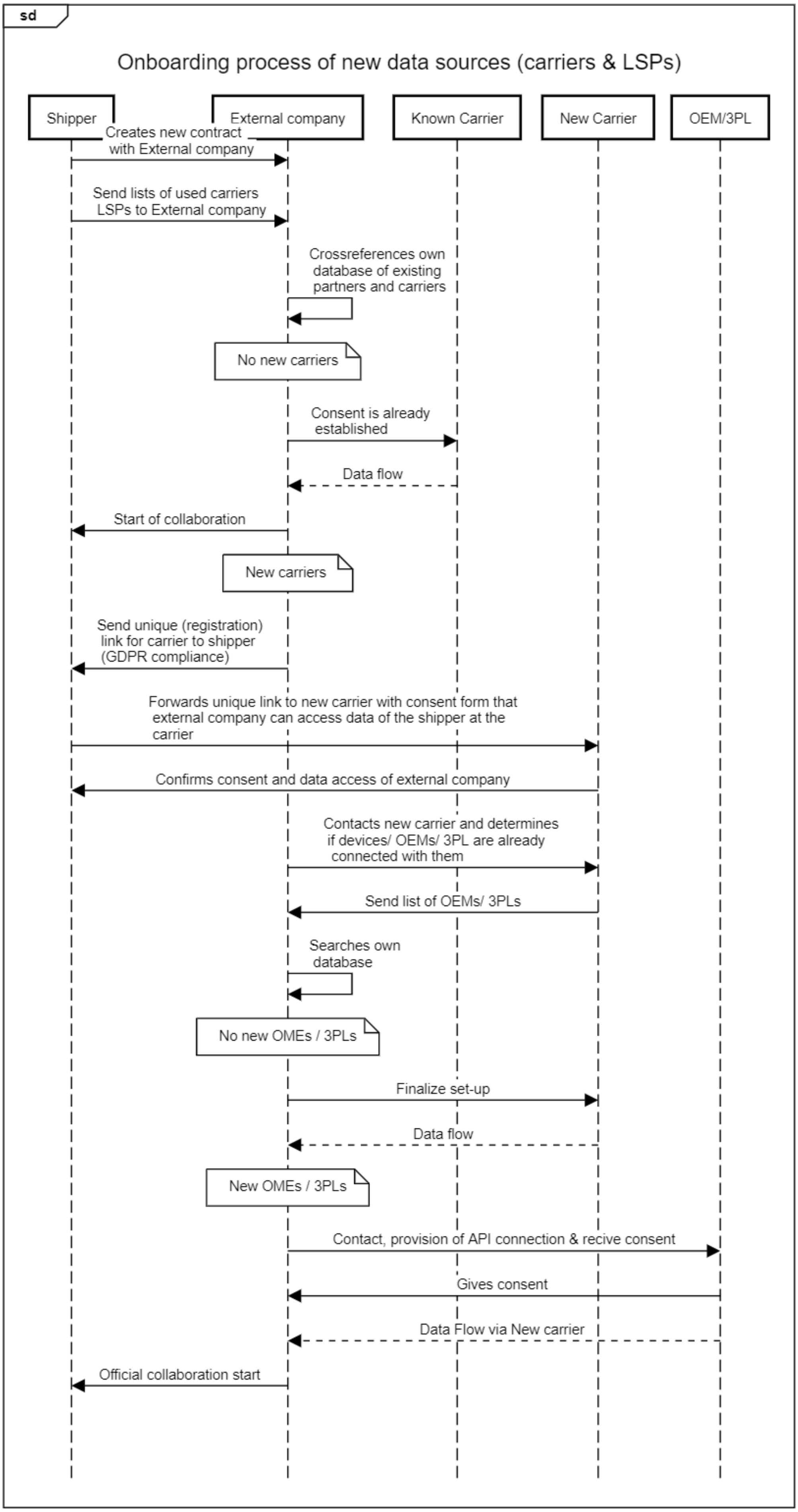

Figure D.2: Abstract sequence diagram of the onboarding process of known or new carriers/ LSPs with an external IT company.

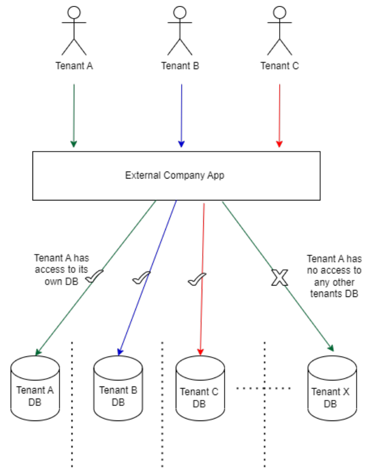

Figure D.3: Abstract representation f a multi-tenant system architecture, which allows access for each tenant to only the own database. The access to other databases is restricted.

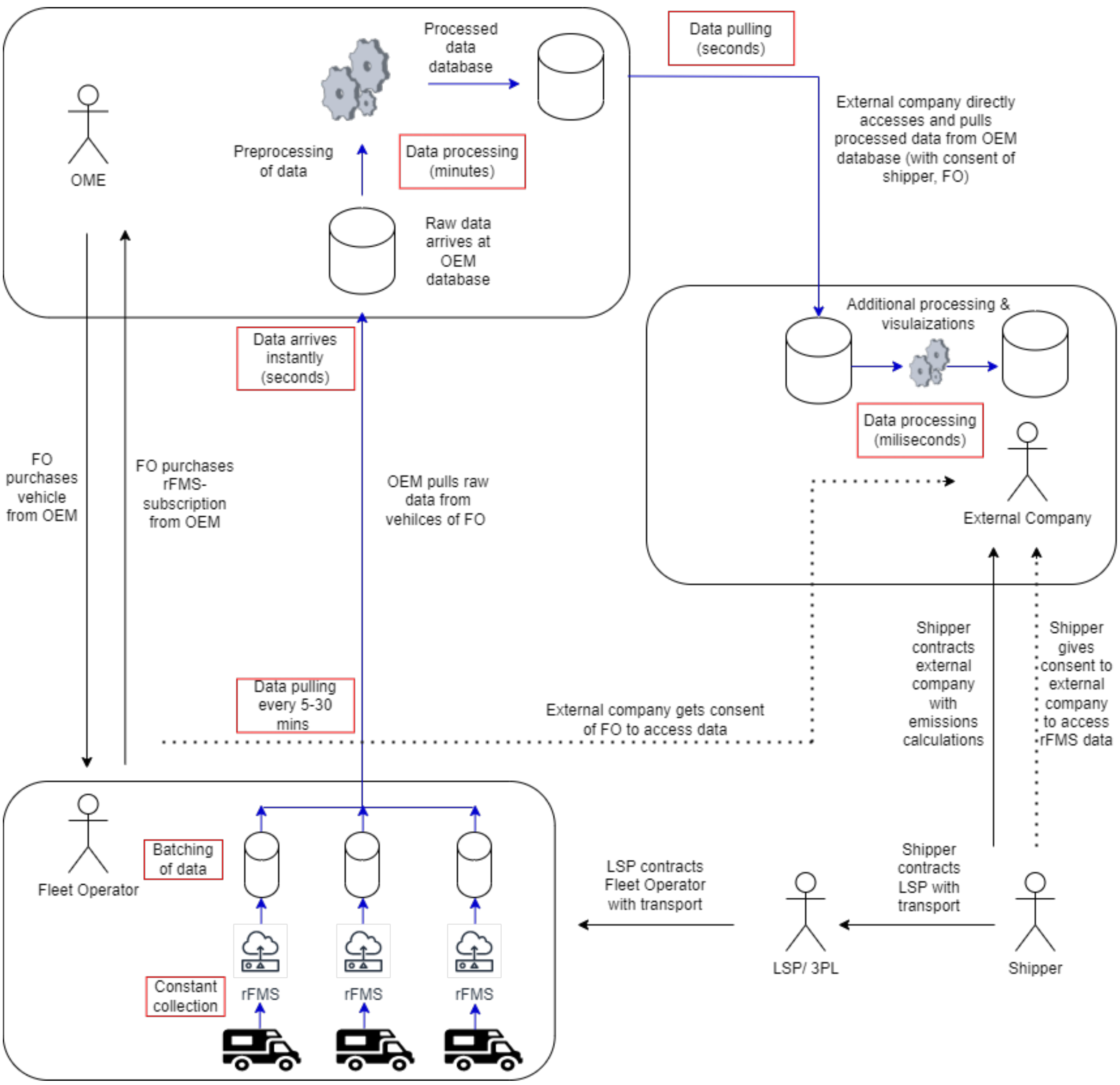

Figure D.4: Abstract representation of data pulling of an external company with OEM and fleet operator via APIs

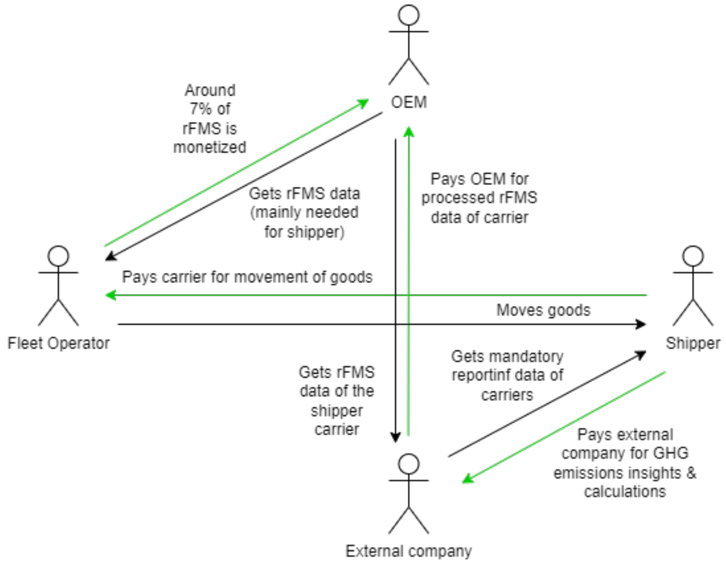

Figure D.5: Business relationship between shipper, OEM, fleet operator and external company for the data collection, creation, processing and usage

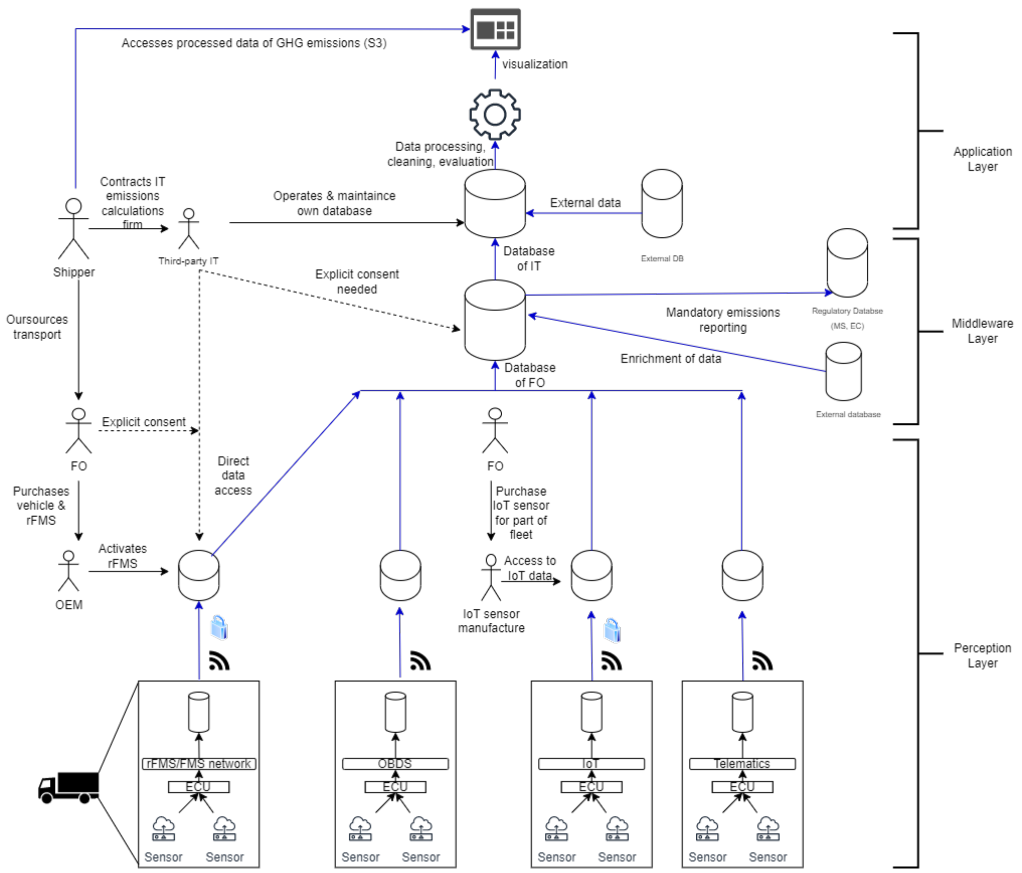

Figure D.6: Abstract system architecture for onboard sensor system in road logistics for a marketplace operator.

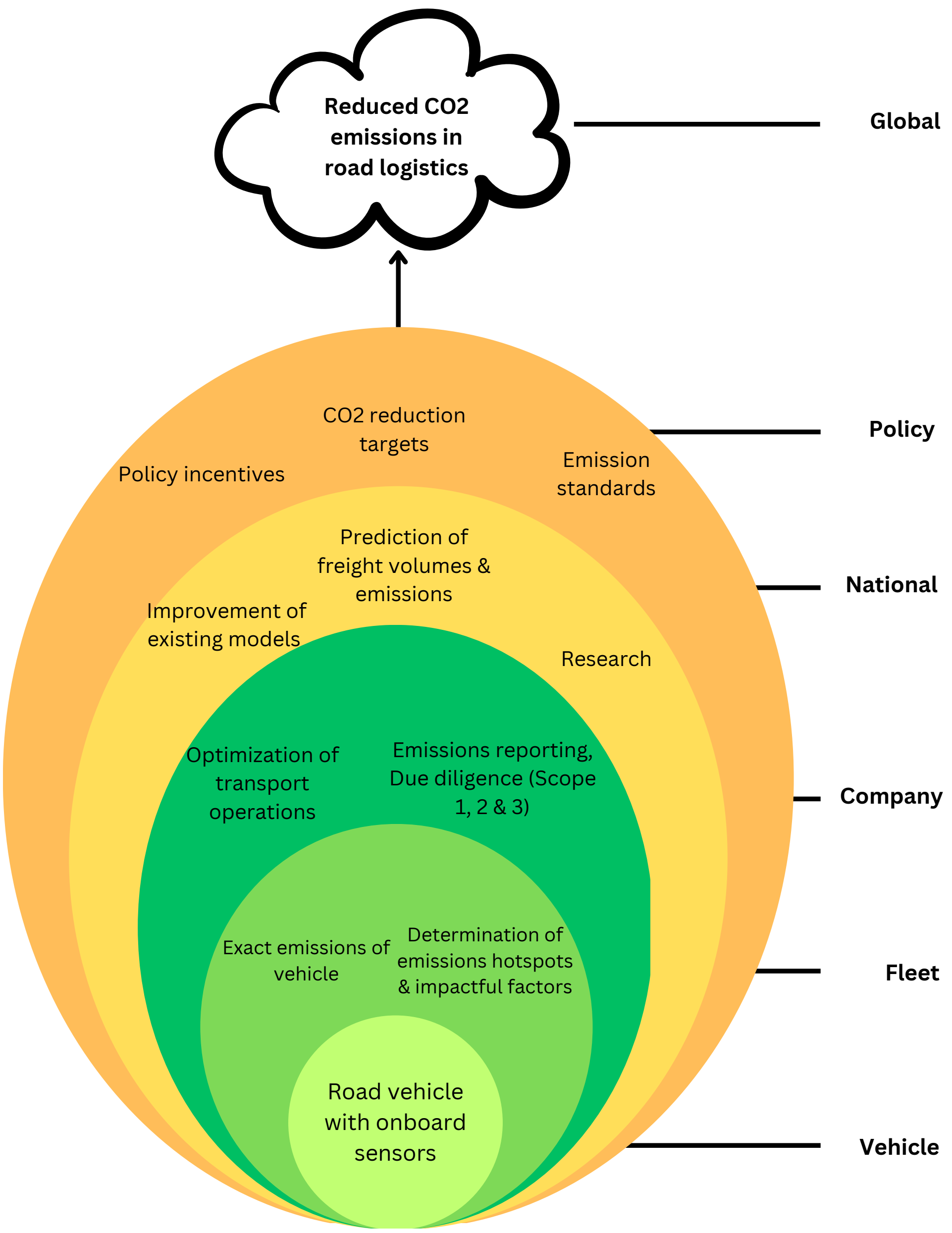

Figure D.7: Direct and indirect usage of primary data per operational level

E

Literature Review Tables

Table E.1: Articles selected per search term.

Search Term	Articles selected
emissions AND vehicles AND technology AND improvements AND logistics	Hopkins & Hawking [69], Xue et al. (2022) [124]
emission AND factors AND effort AND logistics	Liljestrand et al. (2015)[86]
emissions AND factors AND iot AND logistics	D’Amico et al. (2021) [23]
accuracy AND emission AND factors AND vehicles AND logistics	Chen et al. (2022) [17], Yu et al. (2022) [126]
IoT in Logistcis	Mi et al. (2021) [91]
iot AND emissions AND logistics AND road	Yavari et al. (2022) [125]

Table E.2: Articles selected per search term

Search term	Selected articles
Emissions AND factors AND technological AND improvements AND vehicles	Krecl et al. (2017) [84], Martin et al. (2018) [87], Höltl et al. (2017) [68], Roeth (2020) [109], Grigoratos et al. (2019) [64], Misquim & Mady (2022) [99], Ibanez Acevedo et al. (2022) [71], Posada-Henao et al. (2023) [105]
Connected Papers	Mendoza-Villafuerte et al. (2017) [90], Tan et al. (2019) [117], Suarez-Bertoa & Astorga (2018) [116], Dixit et al. (2017) [28], Quiros et al. (2018) [106]

Table E.3: External documents used as a knowledge input.

Type of document	Document	Identified
Policy Document	GDPR [53]	Expert talks
Methodology for emissions	ISO14083 [75], GLEC 2.0 [112]	Emission experts
Company press releases	DHL [27], DNV [29]	IT experts
Academic articles	Van Zenden (2023) [128]	Research experts
News articles	Trading Economics [119], Gravity [63], Solutions & Independent [114]	Independent Research