A Schematic Plan for Train Position Identification Using Digital Twin and Positioning Sensors

Albert Lau, Hongchao Fan, and Hailun Yan(✉)

Department of Civil and Environmental Engineering, Norwegian University of Science and Technology, 7491 Trondheim, Norway
{albert.lau,hailun.yan}@ntnu.no

Abstract. This paper presents a schematic plan outlining a methodology for determining train positions using digital twin technology and commercially available positioning sensors. The plan involves an offline step and an online step. The task in the offline step includes: building a library of 3D models for objects on railway tracks, and constructing a real digital twin environment. Then in the online step, train positions can be accurately determined by automatically generating real-time semantic point clouds from images and matching them with the 3D objects in the digital twin. Along with these tasks, a City Geography Markup Language (CityGML) Application Domain Extension (ADE) for modeling railway tracks in 3D will be developed and suggested as Open Geospatial Consortium (OGC) standard for similar applications. All these procedures will be conducted in a controlled environment using remote-controlled cars before being applied to an actual railway track. It is hoped that this localization methodology can enhance traditional positioning methods, ultimately leading to improved network operation and maintenance.

Keywords: railway Operations · trains · train positioning · digital twin · railway maintenance

1 Introduction

Train positioning, while crucial, was not a primary focus in the earlier days of railway operations. The static traffic block system was sufficient to ensure the safety of trains, with margins typically being adequate. However, with the introduction of the European Rail Traffic Management System (ERTMS), the emphasis on train positioning has grown [1, 2]. ERTMS has been instrumental in modernizing railway operations, emphasizing interoperability, and enhancing safety across European rail networks. Additionally, there's a shifting focus towards railway track and infrastructure condition monitoring using in-service vehicles. This shift underscores the importance of accurate train positioning, especially in a regional context, as efforts to revitalize regional rail intensify and the need to monitor track conditions using onboard sensing devices becomes paramount [3].

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 212–218, 2026.
https://doi.org/10.1007/978-3-032-06763-0_31

Despite the emphasis on accurate train positioning, numerous challenges persist. Trains often operate in harsh railway environments, where they encounter interference, insufficient satellite availability, and signal blockages. These blockages are especially prevalent in specific areas like tunnels, valleys, or due to meteorological conditions. Such challenges have been identified and discussed by several researchers [4, 5]. The quest for precise train positioning often presents a trade-off: opting for high accuracy necessitates expensive equipment and infrastructure, such as advanced communication systems. Conversely, a moderate infrastructure and equipment setup might result in less accurate positioning.

However, technological advancements, particularly in the realms of digital twins and digital mapping, combined with the emergence of cost-effective and efficient positioning sensors, are paving the way for more affordable and efficient train positioning solutions. Vision-based methods, for instance, are being developed for line-side switch rail condition monitoring and inspection, ensuring safer railway operations [6]. Building on this foundation, this paper presents a schematic plan that leverages digital twin technology and onboard train positioning sensors to accurately determine train positions, offering a comprehensive solution to the challenges faced in the realm of train positioning.

2 Methodology

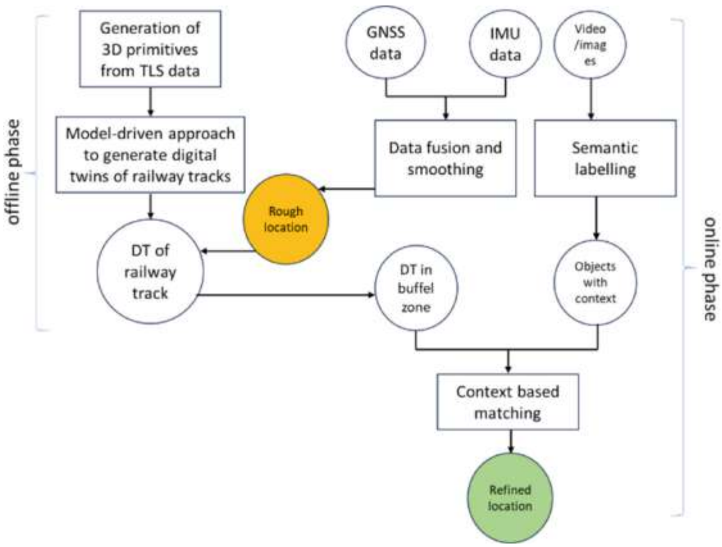

Fig. 1. The framework of the proposed approach for fine positioning of trains

As depicted in Fig. 1., the proposed method consists of two phases: an offline phase, which involves generating a digital twin of the railway track environment using a model driven approach with Terrestrial Laser Scanning (TLS); an online phase, for which data captured from on-board sensors such as GNSS sensors, an IMU (Inertial Measurement

Unit), and cameras, will be combined and matched with the offline digital twin model to obtain the train's real-time highly accurate locations.

2.1 Offline Step

There are two tasks in the offline phase: defining CityGML schema for railway track and generating digital twins with semantic information for railway tracks in the test area. So far, there exists research work about integrating railway infrastructure into CityGML [7]. However, in their concept, railway is modelled as Level of Detail 1 (LoD1) -like object without further decomposition in semantical and geometric domains. In this work, we will conduct a comprehensive inventory of objects that could appear on railway tracks to collect information about their geometrical, semantic, and attributive characteristics. Then knowledge graphs will be used to describe the semantic and topological relations of objects on railway tracks. On this base, code lists will be defined for all possible objects on railway tracks but following the sequence of from leaves to roots on the knowledge graphs. Finally, UML diagrams will be generated to describe feature classes and their hierarchical relations of objects on railway tracks.

For each object on the leaf node of the knowledge graph, a 3D digital model will be created to build up a library of 3D models for railway track modeling. In our project, TLS will be used to acquire 3D point clouds with high density, full coverage and high accuracy in positioning and geometries. In the next step, semi-automated method will be used to extract objects from 3D point clouds. At the same time. Line segments, circle-formed objects and planes will be detected from point clouds of the objects. Then outlines of the components of an object can be extracted, and further regularized in terms of shapes and topology. Finally, these primitives of objects can be modelled in CityGML. It should be noted that we will follow the mechanisms similar to city furniture modelling in CityGML. In other words, they will be modelled in a local coordinate system with the centroid of the objects as the original of local coordinate system.

After the library of the primitives of 3D models are made available, a digital twin for railway track in the test area could be generated by using model-driven approach from TLS point clouds [8]. The reconstructed 3D models of railway tracks will be stored in CityGML format in the online phase.

2.2 Online Step

In the online step, the train will be equipped with GNSS sensors, an IMU, and at least four cameras at the front, back, and two sides of the train. Data captured from all the on-board sensors will be combined and matched with the offline digital twin model to obtain the train's real-time highly accurate locations.

Typically, standard GPS systems offer an accuracy of about 5 to 10 m under open skies. However, this accuracy is often not sufficient for most railway operations. To improve accuracy, many railway systems employ Differential GPS (DGPS) or other augmentation systems like RTK (Real-Time Kinematic) GPS. DGPS can improve accuracy to less than a meter, and RTK GPS can provide centimeter-level accuracy. Nevertheless, the adoption of these technologies is associated with significant financial implications.

The proposed method avoids the adoption of such expensive sensors while maintaining high positioning accuracy.

First, the GNSS signal will be employed to determine an approximate position of the train. In addition, the train's IMU data will also be utilized, which provides the train's orientation and velocity. This IMU data, based on the train's latest known position, can further refine the current estimated location of the train. Such refinement is crucial for enhancing the accuracy of GNSS data, particularly in scenarios where GNSS signals are compromised, such as in tunnels or valleys, or when signals are completely obstructed due to extreme weather conditions. Utilizing the approximate position derived from both GNSS and IMU inputs, a buffer zone is then created around this position within the offline 3D digital twin.

Next, four cameras, ideally positioned at the train's front, rear, will capture images of the surrounding environment. Railway track and trackside objects will be extracted from the multi-view images using image detection and segmentation neural networks, and their neighborhood contexts will be established. These objects will be matched with the objects in the buffer zone of the digital twin. The buffer zone can largely reduce the searching area and increase the positioning efficiency.

2.3 Control Environment Testing

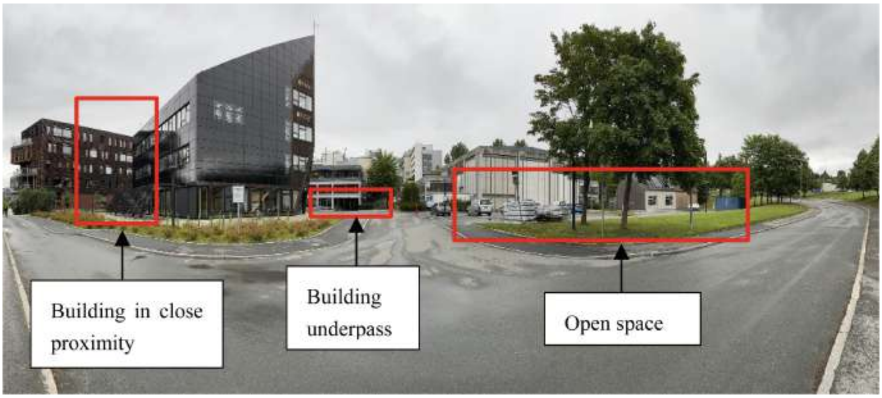

Fig. 2. Test Environment at NTNU

Before implementing this methodology on an actual train, it's prudent to test it in a controlled environment, primarily due to safety and security concerns. From a safety perspective, the installation of equipment, especially cameras, necessitates strict regulations and approvals from train operators and owners. On the security front, the train's position, especially if derived from on-board data, may not be readily accessible to external or unauthorized individuals. Thus, it's advisable to initially test the methodology in a broader setting and subsequently adapt it to the railway domain.

Given the generic nature of the approach, the strategy is to conduct tests in an open environment without compromising safety. A suitable location for this would be the campus of the Norwegian University of Science and Technology (NTNU), as depicted in Fig. 2.

The campus environment offers a variety of features that can simulate railway conditions. Buildings in close proximity can mimic valleys, building underpass can represent tunnels or enclosed areas, and open spaces can serve as typical open terrains. Specific landmarks and objects within this environment will be positioned in the digital twin for accuracy.

To simulate the train, a remote-controlled car will be employed. This car will be equipped with the same sensors as those on the train (i.e., GNSS sensors, cameras, and an IMU). First, an offline 3D digital twin of the test site will be generated using the methods described in Sect. 2.1. Next, the car will traverse the campus, while the data captured by all its sensors will be combined and matched with the digital twin to obtain the car's real-time locations. The algorithm will be continuously tested and revised until the positional accuracies achieve satisfactory.

3 Potential Challenges

Several challenges can be anticipated at this stage:

Monotonous Objects: While railway tracks are lined with numerous objects, many of these are repetitive in nature. For instance, the shapes of rails and sleepers remain consistent over long stretches. Relying solely on images for object identification can be problematic when environmental differences are minimal. Hence, it's crucial to incorporate other data sources, such as velocity, train position, and high-frequency trackside objects.

Weather Conditions: Conditions like heavy rain, snowfall, fog, or a snow-covered environment can obstruct or degrade camera visibility, compromising image quality for identification. In such scenarios, alternative data sources or methods should be considered to address these challenges.

Difficult Locations: Locations like tunnels and valleys are notorious for obstructing GPS signals. Additionally, these areas are often dimly lit, further complicating positioning efforts.

Lighting: Proper lighting is essential for accurate camera image identification. In low-light conditions, the accuracy of train location identification may be compromised. Potential solutions could include using cameras capable of functioning in low light or installing additional lighting on the vehicle.

Odometer Data: The resolution of odometer data logging can vary, presenting challenges if there's significant deviation. Different odometers have varying logging frequencies, which can affect the accuracy of positioning.

Communication System: Transmitting real-time position data to a central system can be challenging, especially when commercial communication infrastructures are unavailable, particularly in regional contexts. It's essential to process and compress position data to ensure it's transferable, even though systems like GSM-R.

System Integration: Once developed, system compatibility might pose challenges. The on-board system could differ from the testing version. If the system is to be implemented on-board, seamless integration with existing systems is vital to avoid substantial investments. Ideally, a plug-and-play solution would be most beneficial in this context.

4 Expected Outcomes

Through this methodology, the ambition is to pinpoint train positions with a precision of ± 25 cm. This level of accuracy is a marked improvement over certain maintenance vehicles, which currently have positioning accuracies more than ± 15 m. Another anticipated outcome is the potential for this methodology to be seamlessly implemented across all trains in Europe. The ideal scenario would be a plug-and-play solution, subject, of course, to rigorous security checks and full compliance with established standards.

Acknowledgements. This research project is funded by Norwegian Railway Directorate (NRD) and Europe's Rail Joint Undertaking (grant number 101101962-FP6-FutuRe-HORIZON-ER-JU-2022-01).

References

1. Steuer, M., Krzykawski, M., Simiński, D., Chema, W., Burdzik, R.: Train detection methods as the foundation of positioning systems of railroad traffic control. Diagnostyka (2023) 2023305. https://doi.org/10.29354/diag/168655
2. Fikejz, J., Kavička, A.: RegioRail—GNSS train-positioning system for automatic indications of crisis traffic situations on regional rail lines. Appl. Sci. **12**(12), 5797 (2022). https://doi.org/10.3390/app12125797
3. Tsunashima, H., Ono, H., Takata, T., Ogata, S.: Development and operation of track condition monitoring system using in-service train. Appl. Sci. **13**(6), 3835 (2023). https://doi.org/10.3390/app13063835
4. Loidolt, M., Marschnig, S.: Evaluating short-wave effects in railway track using the rail surface signal. Appl. Sci. **12**(5), 2529 (2022). https://doi.org/10.3390/app12052529
5. Wang, M., Xiao, Y., Li, W., Zhao, H., Hua, W., Jiang, Y.: Characterizing particle-scale acceleration of mud-pumping ballast bed of heavy-haul railway subjected to maintenance operations. Sensors. **22**(16), 6177 (2022). https://doi.org/10.3390/s22166177
6. Ye, J., Stewart, E., Chen, Q., Chen, L., Roberts, C.: A vision-based method for line-side switch rail condition monitoring and inspection. Proc. Inst. Mech. Eng. Part F: J. Rail Rapid Transit. **236**(8), 986–996 (2022). https://doi.org/10.1177/09544097211059303
7. Kumar, K., Labetski, A., Ohori, K., Ledoux, H., Stoter, J.: Harmonising the OGC standards for the built environment: a CityGML extension for LandInfra. ISPRS Int. J. Geo-Inf. **8**, 246 (2019). https://doi.org/10.3390/ijgi8060246
8. Zhang, Y., Weng, Q.: Model-driven reconstruction of 3D building using LiDAR data. In: IEEE Geoscience and Remote Sensing letters, Vol. 12, No.7 (2015)

Designing the Procurement Logistics Processes of a Smart Factory Based on Virtual Building Models

Bennet Zander[1](✉), Kerstin Lange[1], Chris Gieling[2], and Christian Struck[2]

[1] Jade University of Applied Sciences, 26931 Elsfleth, Germany
bennet-zander@gmx.de

[2] Saxion University of Applied Sciences, 7513 AB Enschede, The Netherlands

Abstract. Smart factory research has primarily concentrated on the manufacturing sector, without taking the building sector into closer consideration. Combining these two ventures can lead to a variety of synergies to minimize the CO_2 emissions of existing buildings. By producing facade and roof panels on an industrial scale for energetic refurbishments, Construction 4.0 and prefabrication can make a significant contribution to achieving climate targets. In this context, the purpose of this paper is to identify, how data of a virtual building can be used in procurement logistics of a smart factory. To achieve this, this paper follows a case study analysis. At the beginning, the procedure of developing a virtual building model is described. This is followed by the design of the inbound workflows of a smart factory in the construction industry. The findings demonstrate how the information from virtual building models can be leveraged to control procurement processes of a smart factory.

Keywords: Virtual Building Models · Smart Factory · Procurement Logistics

1 Introduction

The expected effects of Industry 4.0 on the economy are remarkable. The Organization for Economic Cooperation and Development predicts that the global gross domestic product will almost triple by 2060 [1]. The implementation of smart factories alone could add at least €1.35 trillion in value to the global economy. Currently, most studies focus on the six main industries – machinery and equipment, electrical and electronic equipment, automotive, chemicals, agriculture, and information technology – which are expected to have a value creation potential of €78 billion by 2025 [2].

In comparison, the potential of Industry 4.0 for the construction industry has received relatively little attention. The existing building stock contributes to 40% of the total energy use and 36% of the CO_2 emissions. The majority of European residential buildings, of which approx. 51% have been built before 1969, are in high need of renovation. The North Sea region (NSR) alone contains 22 million dwellings which cause 79 Mton CO_2 emissions annually. To deal with the climate crisis, climate and energy objectives were defined. By 2050, CO_2 emissions should be cut to at least 80% compared to 1990

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 219–225, 2026.
https://doi.org/10.1007/978-3-032-06763-0_32

and buildings should be energy-neutral [3]. However, the building renovation industry applies mainly manual on-site techniques, resulting in a low renovation pace (ca. 2% p.a.), high labor costs and a long duration [4]. Thus, the ambition to achieve the climate targets is challenging for the sector.

The integration of smart factories could enable growth in the construction industry, as the manufacturing of housing units could take place on an industrial scale. Automated processes create efficient manufacturing approaches in residential renovation to reduce CO_2 emissions to a necessary limit [5]. In this context, this paper aims to identify how data generated by 3D models of residential buildings can be used for managing the procurement processes of a smart factory in the construction industry.

2 Related Work

2.1 Impacts of Digitalization on the Construction Industry

In recent years, efforts have been intensified to invest in the digital transformation of the construction industry. Initiatives are being subsumed under the term "Construction 4.0". However, similar to Industry 4.0, there is no consensus on what Construction 4.0 means. Some authors describe it as a manifestation of Industry 4.0 in the construction industry and the use of modern technologies such as the IoT, Big Data, AI or advanced manufacturing, as this could enable a decentralized, cost-effective production in smart factories [6]. A second group considers Construction 4.0 as an opportunity to find complementarities between technological approaches to address the current challenges facing the construction industry [7].

This paper follows the view of the first group of authors that industrialization in construction can lead to the implementation of smart factories. Accordingly, this requires a redesign of supply chains, as building material is now processed in factories before being delivered to the construction site, instead of being delivered directly.

2.2 Smart Factories in the Construction Industry

A smart factory is defined as the integration of artificial intelligence (AI), mechanical engineering, and information technologies. It can be understood as a flexible production facility of Industry 4.0, based on intelligent units, where machines coordinate production processes independently and autonomous transport vehicles carry out logistical tasks. The involved objects, such as machines, products and/or vehicles, exchange information autonomously in real-time via the internet [8]. The Internet of Things (IoT) is used to integrate resources, allowing the manufacturing system to have sensing, connecting and data integration capabilities. In this context, data is the most important basis for a smart factory. It is used in all areas, from the initial design of the products, through procurement to manufacturing [**Error! Reference source not found.**]. This enables a dynamic production structure where processes are no longer statically predetermined, but adjusted based on the individual customer order. A flexible manufacturing system is created, which utilizes a continuous stream of data from connected systems to learn and adapt to new requirements [9].

In regard to the future of the construction industry, entire house units can also be autonomously prefabricated in factories. These Smart Construction Factories aim for more standardized, cost-effective, and faster production compared to on-site construction processes. The labor costs of the construction sites are consequently converted into capital costs. In addition to new buildings, the market for refurbishment can also be addressed, as these factories can produce serial renovation packages. They consist of facade and roof panels, solar collectors and heating, ventilation, air conditioning (HVAC) systems.

2.3 Virtual Building Models

The building sector adopted in the past innovations for design such as 3D drawing and Building Information Models (BIM), which allow to seamlessly integrate different engineering disciplines, to have a closer look on construction details and material use as well as to realize a higher construction quality. To further automate the procurement and production processes, more information is needed and thus more detailed models of the products are required. Therefore, high-resolution models can be used. The data level of detail of these models extends BIM. High-resolution models and BIM both contain the same data such as 3D geometrics, material or components use and installations. However, high-resolution models also contain additional data of (hygro-)thermal bridges, air flow paths, air leakage and operational schedules for e.g. lighting and occupancy, HVAC and domestic hot water systems as well as embedded renewable energy systems [10]. Using this data, a comprehensive virtual model of the building and its performance is generated, paving the way for subsequent mass production.

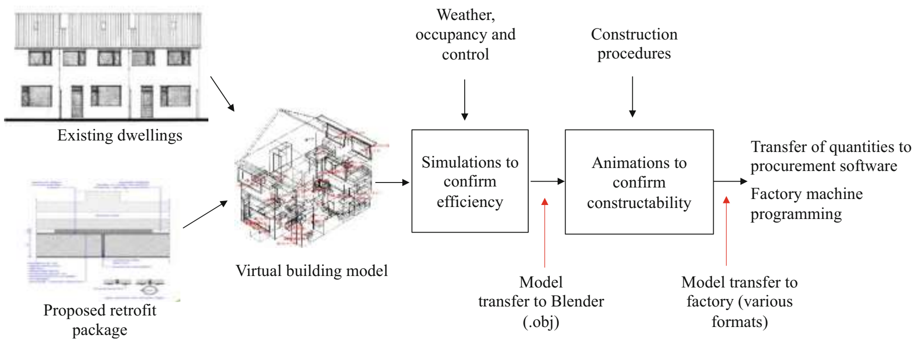

Fig. 1. Construction processes based on high-resolution building models.

High-resolution models can help to simulate all properties of the building, such as the condition following an energic refurbishment in which the building is renovated with panels from the Smart Construction Factory. The virtual model consists of the data of the existing building and the new facade and roof panels. Within the model, energy calculations and assembly simulations can be performed (in Blender) to verify the retrofit achieves the desired standard. If it passes, the.obj-file formats can be forwarded to the smart factory and be transformed to other machine language protocols to initiate production and logistics processes, such as procurement logistics (see Fig. 1) [11].

3 Methodology

This contribution focuses on describing and analyzing a case-study. Case study-based research distinguishes between a descriptive analysis, in which correlations are described, and an explicative analysis, in which correlations are additionally examined with respect to their variability [12].

As this contribution emphases on the identification of the influence of virtual building models on procurement processes of a smart construction factory, a descriptive case study methodology according to [13] is applied. Here, five components of a research design are important:

1. The research question,
2. Its propositions,
3. The unit of analysis,
4. The logic linking of the data to the propositions,
5. The criteria for interpreting the finding.

Following the five components, this study focuses on the research question, how building data can be used to control procurement processes of a smart factory (1). The assumption is, that the use of virtual building models can increase the efficiency of the procurement processes (2). The case study of this paper is the scientific project INDU-ZERO (3). A virtual building model can increase the costs and the effort involved in the design stage (4). At last, the study needs to verify that the new automated processes can cover the additional effort of creating a high-resolution model (5).

4 Application of Virtual Building Models in Procurement Logistics

To tackle the need for energetic renovations, six countries of the NSR collaborate in the Interreg project INDU-ZERO "Industrialization of house renovations toward energy-neutral" with the ambition to upscale the current renovation process and to create the necessary facilities. INDU-ZERO aims to design a smart factory blueprint to produce 15,000 fully integrated prefabricated renovation packages per year per factory suitable for all NSR countries at a reduced cost of 50% [4, 14].

In the INDU-ZERO project, a high-resolution model was generated as a 3D representation of one building to be refurbished after it was scanned. Based on the measurement points and using specifications, it was developed into a digital design that includes a structured hierarchy of product information and associated naming criteria, tabular metadata and a common organization via databases.

Procurement logistics comprises the connection between the supplier's distribution and the manufacturer's production. Its objects are raw materials, supplies and merchandises that have to be made available to the manufacturer. Correspondingly, it is the function of procurement to provide, maintain and develop delivery capacities.

To provide coherent instructions for the procurement processes of the Smart Construction Factory, the manufacturing data describing the product parameters had to be converted into readable data sets for the product lifecycle management (PLM) system. Material requirements such as material quantities, design and insulation specification were generated in the PLM system. For each construction order, the combined design and PLM data was in turn queried by the Enterprise Resource Planning (ERP) system and ordered from suppliers according to the pull principle (see Fig. 2).

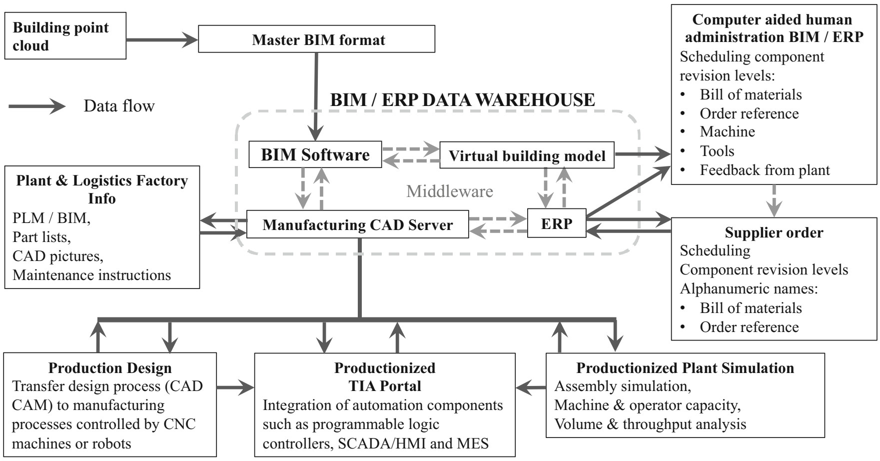

Fig. 2. Design of the information flows in a Smart Construction Factory.

In the implementation phase of the case study, an enterprise service bus (ESB) was used as a middleware for the data transfer between the systems. It was able to forward the data from the high-resolution model to the PLM and ERP system. After receiving the data of the materials, the suppliers prepared the transportation. Regarding the purchase, suppliers referred to predefined agreements so that individual orders do not have to be renegotiated. With this scenario, analogue administrative processes between the procurement, production and sales departments were avoided.

Furthermore, it was not only possible to benefit the procurement logistics processes from the virtual building model, but also to use its data to design the product and simulate the production processes of the smart factory and thus test their success.

5 Discussion and Results

The case study shows how high-resolution building models can be used for the procurement logistics of a smart factory in the construction industry. This applies in particular to the supplier ordering process of materials by transferring the BIM data to the PLM system and then to the ERP system with the help of an ESB middleware.

From a logistical point of view, the process changes that resulted in the case study have the most impact on the accuracy of purchase orders, the rate of emergency purchases and purchasing time. Automated purchase orders and predefined master agreements not only shorten order times, but also allow the immediate release of order quantities from actual materials used. Supplier availability, fulfillment rates, and invoice accuracy from contracted details are also met, ensuring that there should be no bottlenecks as suppliers commit to fixed quantities at specific prices and times.

From a manufacturing point of view, the most valuable components of the new information flows between the systems are algorithmic and parametric product designs. They describe the geometry of all individual parts that compose the bill of materials for the facade and roof panels, which can then be used by the PLM and ERP system to comprehend the design context and automate production and logistics processes.

Thus, the assumption made in the methodology that efficiency gains are possible is confirmed. By creating uniform interfaces in the software system landscape in advance, procurement and production processes can be automatically adapted to the building structures in the future without increasing the workload during operation.

References

1. OECD: https://data.oecd.org/gdp/gdp-long-term-forecast.htm. Accessed 20 May 2023
2. Capgemini Service SAS: https://www.capgemini.com/insights/research-library/smart-factories-at-scale/. Accessed 24 July 2023
3. European Commission: https://ec.europa.eu/energy/topics/energy-efficiency/energy-efficient-buildings_en. Accessed 21 Apr 2023
4. Decorte, Y., et al.: Upscaling the housing renovation market through far-reaching industrialization. IOP Conf. Ser. Earth Environ. Sci. **588**, 032041 (2020)
5. Klinc, R., Turk, Ž.: Construction 4.0 – digital transformation of one of the oldest industries. Econ. Bus. Rev. **21**(3), 393–410 (2019)
6. Craveiro, F., Duarte, J., Bartolo, H.G., Bartolo, P.J.: Additive manufacturing as an enabling technology for digital construction: a perspective on Construction 4.0. Autom. Constr. **103**, 251–267 (2019)
7. Boton, C., Rivest, L., Ghnaya, O., Chouchen, M.: What is at the root of construction 4.0: a systematic review of the recent research effort. Arch. Comput. Methods Eng. **28**, 2331–2350 (2021)
8. Jacobi, H.F.: Computer Integrated Manufacturing (CIM), in Westkämper, E., Spath, D., Constantinescu, C. Lentes, J. (eds.), Digitale Produktion, Springer Vieweg, Stuttgart, pp. 51–92 (2013)
9. Wang, S., Wan, J., Zhang, D., Li, D., Zhang, C.: Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination. Comput. Netw. **101**, 588–600 (2016)
10. Bongsug, C.: Developing key performance indicators for supply chain: an industry perspective. Supply Chain Manag. **14**(6), 422–428 (2009)
11. Tao, F., Zhang, M., Nee, A.Y.C.: Background and concept of digital twin. In: Tao, F., Zhang, M., Nee, A.Y.C. (eds.) Digital Twin Driven Smart Manufacturing, pp. 3–28. Academic Press, Cambridge (2019)
12. Baškarada, S.: Qualitative case study guidelines. Qual. Rep. **19**(40), 1–25 (2014)
13. Yin, R.K.: Case study research: design and methods. 5th edn. Sage Publications, Thousand Oaks (2014)
14. INDU-ZERO: https://www.induzeroblueprint.eu/cover/. Accessed 22 July 2023

A Framework for Compliance Verification Based on Trusted Data and Blockchain Identity

Nikita Karandikar(✉) and Knut Erik Knutsen

Den Norske Veritas, Veritasveien 1, 1363 Høvik, Norway
nikita.karandikar@dnv.com

Abstract. Ships generate huge amounts of sensor data, which is largely underutilized due to lack of standardization. The ad-hoc data collection and transmission procedures and proprietary nature of sensor naming conventions make it challenging to implement and scale digital solutions based on this data. Ensuring data integrity can present an additional challenge as data is generated by multiple sensors onboard a ship with constrained connectivity to shore. Sensor as well as stakeholder identities are significant here, in order to reflect ownership structures, maintain access control and link actions with the entity that performed it. In this paper we present Vidameco, an integrated hardware and software framework for standardized data collection and tamperproofing. Dockerized microservices are developed for data collection, hashing, publication of hashes on the blockchain and verification by data consumers. Further, we present an architecture to integrate this tamperproofed data with sensor and stakeholder identity in order to create an underpinning for a digitalized value chain. Sensors are named in a hierarchical structure using the ISO19848 standard and this structure is interfaced with digital identities of stakeholders on the blockchain such as from the European Blockchain Services Infrastructure (EBSI). Finally, taking EU's Monitoring Recording and Verification (MRV) regulations as a case study we conceptually demonstrate this framework. MRV reporting is a complex value chain involving multiple stakeholders and extensive documentation and forms the basis for EU's Emissions Trading System (ETS), which is set to include shipping from 2024. This case study presents and evaluates an end-to-end digitalized ETS ecosystem based on tamperproofed data using MRV as the Oracle in smart contracts for ETS trading.

Keywords: Blockchain · tamperproofing · MRV · ETS · identity · standardization

1 Introduction

Maritime transport is the backbone of the global supply chain, carrying over 90% of the world's merchandise trade, an estimated 11 billion tons of cargo each

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 226–232, 2026.
https://doi.org/10.1007/978-3-032-06763-0_33

year [1]. The importance of maritime transport came into sharp focus during COVID-19 crisis as shipping dependent supply chains were key not only for business continuity and keeping the economies functioning, but also to critical supply lines, delivering essentials and keeping food on the table. The pandemic has also shown the vulnerabilities of the sector and the need to increase resilience in the face of challenges. A recent World Bank-IAPH report [1] advocates for a broader and better coordinated integration of digitalization into the supply chains, particularly creation of a digital ecosystem which can unlock significant efficiency gains with lower emissions and more resilient supply chains.

However, today, data streams generated on board vessels are transmitted to data consumers on shore by ad-hoc and disparate means and thus requires keeping track of several different and non integrated data pipelines. Such a complex system is not easily scalable as adding a new data resource or supporting multiple services using several separately collected data points becomes challenging. Also challenging, is ensuring data integrity between processes in resource constrained environments where data transfers are not immediate. The integrity of the data is vital, as possible undetected tampering can severely undermine trust in the data and services relying on it. Further, investing into these improvements can translate into the bottom line. A study by BCG [2], estimated that digitalization can reduce OPEX in shipping by 15%, significant in such a price sensitive sector. Regulatory actors such as IMO have also pushed for digitalization through their Facilitation Convention, making electronic data exchange mandatory since 2019 with guidelines for integrity, authentication and confidentiality of information exchanges in the maritime sector.

EU's MRV and IMO's DCS regulations as well as the inclusion of shipping in the ETS program show the growing importance of traceability of emissions in shipping. ETS has a market value of €751 billion in 2023 [3] and the projections for ETS costs for shipping alone are €3.1 billion (2024), €5.7 billion (2025) and €8.4 billion (2026) [4], factoring in the 3 year phase in. With the introduction of market based measures comes the incentive to adjust emissions reported in order to reduce costs. For instance, some years ago, Europol found that carbon credit fraud in the ETS system cost a loss of tax revenues of €5 Billion [5]. More recently, in 2023, an instance of a multi-million euro ETS fraud was found in Bulgaria [6]. With the current granularity and accuracy of emissions as well as verification procedures, this may not be easy to discover. The estimated daily ETS cost of €10k for a container ship would translate to north of a million per annum per vessel [7] assuming more than 100 sailing days. With the current acceptable margin for error in MRV (5%) this would mean a difference of €50k per vessel per year. As such, the proposed approach if implemented can improve granularity and accuracy and reduce attempts at cheating so that the EU and IMO regulations will have greater effect and contribute to a fair distribution of costs.

2 Architecture and Conceptual Implementation

2.1 Standardization and Tamperproofing

Our solution addresses standardization and edge data integrity using two salient building blocks- Dockerized microservices and Blockchain respectively. Docker is a platform that simplifies the delivery of application code by bundling it with operating system (OS) libraries and dependencies creating containers—self-contained and standardized execution units that can run in any environment. Docker containers can be run in a Swarm mode to ensure high availability and scalability. Further, Docker containers can be used to deploy microservices. Microservice architecture granularizes and modularizes functionality into separate standalone components that can collaborate. This architecture creates interchangeable, scalable components, plug and play by design.

Blockchain, also known as a Distributed Ledger Technology (DLT) is the technology that underpins Bitcoin, and other cryptocurrencies but the applications of this technology have now expanded to enterprise use cases such as supply chain management, personal identity, IoT asset administration and trading renewable energy. Blockchain is so named as it consists of blocks of transactions chained together though cryptographic hashes. It is a shared ledger of immutable records, maintained by peers in a network using consensus algorithms such as Proof-of-work, Proof-of-stake or Proof-of-Authority to determine the transactions and the correct order of transactions in a block. This immutability makes it possible to use it for tamperproofing of data.

Figure 1 shows the architecture of the Vidameco solution. Microservices are developed to collect edge data batches. These batches are hashed, and the hash is published to the blockchain. Later the data consumer can use our DApp to access a microservice for verification of data integrity based on the published hash. These microservices are deployed as Docker containers making a modular solution with plug and play components. MQTT broker is used for inter-process communication. Vidameco can be used to enable standardization and data integrity for a variety of use cases. In more traditional sectors such as shipping this can drive digitalization. Focusing on EU's MRV regulations, implementing the Vidameco approach, can reduce redundancy in data collection and quick verification of both, the data used as input to the value chain and the certification produced at the end, ensuring traceability from the former to the latter. Tying the stakeholder activities and data to their identity will reduce complexity and avenues for mistakes and fraud.

2.2 Oracles and Federated Identity

Blockchain based solutions often need to take in data from the outside world. Oracles enable a blockchain ecosystem to access data as well as advanced computations. For instance if a smart contract encodes a bet between two individuals based on the outcome of a race, this information must be delivered into the smart contract. However, blockchain oracle mechanisms using a centralized

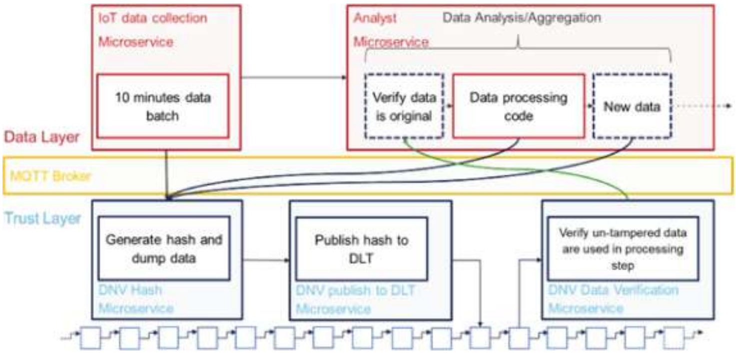

Fig. 1. Vidameco architecture

entity to inject data to a smart contract create a single point of failure, defeating the entire purpose of a decentralized blockchain application. Further, identity on the blockchain is important as it is in the real world. The EU has created EBSI with the ambition to provide a digital identity to every person and legal entity in the EU. However, identity can be used not only to authenticate parties, but also as an intuitive way to structure data by identifying components. The DNV sensor naming rule Vessel Information Structure (VIS) [8] based on the ISO 19848 standard provides a hierarchical naming scheme for sensors and components based on function. As seen in Fig. 2, we propose an hierarchical and integrated identity to ensure traceability of data at the granularity level of sensors and linking this with the EBSI identity of stakeholders. Further, such an identity created by interoperability of VIS and EBSI can be issued by a trusted issuer and be used to tamperproof the data streams. The identity, certificates and tokenized carbon assets as well as coins can be stored in a blockchain wallet. The source data can be stored in an offchain storage solution. This enables the creation of hybrid smart contracts combining on-chain and off-chain code, data and infrastructure. Further, as the oracle inputs are based on tamperproofed data, this creates an evidence based system where proof can be produced on demand. Thus, an Oracle for trusted data, middleware for blockchain interoperability between VIS and EBSI, interface between on and offchain storage, access control as well as a wallet will be needed. To provide all these capabilities, a Decentralized App will be developed for end to end traceability covering this entire value chain and providing a user friendly experience for all stakeholders.

3 MRV

MRV will be used for ETS reporting when maritime transport is included in the ETS system in 2024. The MRV reporting value chain is complex, with many stakeholders, each with their own systems of record and data. Currently, annual emissions calculations are not based on sensor measurements on board a ship.

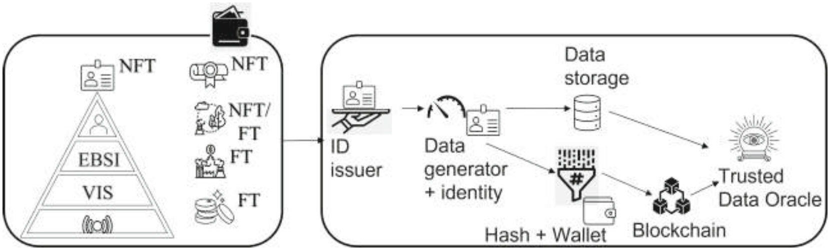

Fig. 2. Identity and trusted data

These are estimated based on the fuel consumption data and operational data reported by the ship in daily reports called the noon reports. The fuel consumption reported in the noon reports is substantiated by invoices on fuel sale and by sample studies done in laboratories on samples of fuel from a batch. In order to comply with reporting requirements, all this data must be submitted to an accredited verifier like DNV, for consistency check and verification before being submitted to the EU Commission. The documents are delivered to the verifier by email from each stakeholder in an unstructured fashion, as seen on the left hand side of Fig. 3. Sorting and structuring these documents, requesting missing documents, as well as checking for class certifications and validations is an extensively paper based and document heavy process. Additionally, as regulation is now moving towards demanding more granularity and transparency in emissions reporting, such aggregated annualized reporting may not be sufficient in the near future. Finally, as regulation continues to evolve, this unstructured data makes it more challenging to keep up. Data that is granular, tamperproofed and structured by identity will ease combining data sets in different ways, reducing rework and producing evidence based reports.

The right hand side of Fig. 3 shows our proposed solution. Here, a fuel supplier issues a bunkering invoice and tamperproofs it using our solution so that the hash of the bunkering invoice is published on the blockchain against the identity of the fuel supplier. This hash is meaningless without the source data and cannot be used to get any information about the source data including size of data. Similarly, the laboratory produces and tamperproofs their reports. The ship starts the voyage and collects and hashes data in 10 min batches using the hierarchical identity and publishes the hashes to blockchain. The source data for all three stakeholders is stored in an offchain repository. Later, when the MRV verifier is given access to this off chain database, they are able to retrieve the data stored, structured by stakeholder as well as sensor identity and verify its integrity against the hash stored on the blockchain.

4 Discussion and Conclusions

We now compare and contrast our approach with the current document heavy approach. **Robustness** and resilience to failure is an important characteristic for

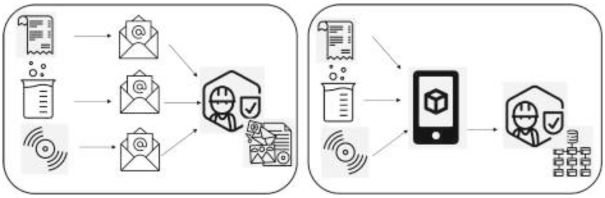

Fig. 3. MRV data pipeline current(left) and proposed(right)

such a system. Here, the Docker based data collection infrastructure in Swarm mode ensures availability so that when one Docker container fails, another one is automatically spun up to take its place. Further, the tamperproofing as well as the identity infrastructure uses blockchain which is maintained by a peer network meaning that each peer maintains a copy of the network, ensuring availability. In the current approach, there are many single points of failure that can be as simple as missing an email. **Scalability** is one of the main benefits of such a solution as it is scalable by design. The Dockerized microservices can be replicated as needed, and the standardized and structured fused data and identity eases adding in a new data source or combining multiple data sources, making it much more scalable than the current approach. **Verifiability** of presented reports will be essential as the regulations evolve, enforcing more traceability. Currently, the values presented in the reports are checked for consistency based on a model but integrating granular sensor data as proposed will improve Verifiability at each step of this value chain. Further, compared to a centralized Oracle, that defeats the purpose of blockchain, evidence is essential to ensure accountability.

The proposed approach will digitize the maritime value chain, beyond compliance reporting like MRV, improving robustness and ensuring end to end traceability.

Acknowledgement. This work was partially funded by the Research Council of Norway.

References

1. World Bank, IAPH: Accelerating Digitalization: Critical Actions to Strengthen the Resilience of the Maritime Supply Chain. World Bank (2020)
2. KPMG: Digital transformation in the shipping industry is here (2021)
3. Reuters: Global Carbon Markets Value hit record $909 bln last year (2023)
4. Hellenic Shipping News: Shipping set for €3 billion-plus 2024 EU ETS bill (2023)
5. Europol: Carbon Credit fraud causes more than 5 billion euros damage for European Taxpayer (2010)
6. EPPO: Bulgaria: EPPO probes into multi-million euro fraud regarding greenhouse gas emissions (2023)
7. Biermans M., van der Klip, M.: Prow Capital BV: The Financial Consequences of bringing Shipping under ETS: a preliminary review (2023)

8. Lag, S., Vindoy, V., Ramsrud, K.: A standardized sensor naming method to support digital twins and enabling new data driven applications in the maritime industry. In: 13th Symposium on High-Performance Marine Vehicles, HIPER (2021)

MOTIONAL: Advancing the Future of European Railway Systems Through Digitalization and Integration

Marco Ferreira[1](✉), Magnus Wahlborg[2], Lars Deiterding[1], and Anders Johnson[2]

[1] Hacon Ingenieurgesellschaft mbH, Lister Street 15, 30163 Hannover, Germany
marco.ferreira.smo@hacon.de
[2] Trafikverket, Röda Vägen 1, 781 70 Borlänge, Sweden

Abstract. The Flagship Project "MOTIONAL", supported by Europe's Rail initiative, paves the way for implementation of the future European Capacity Planning and Traffic Management System, built on digitalisation, automation, connectivity and multimodal integration.

Lead jointly by Hacon and Trafikverket "MOTIONAL" with its 89 partners aim to achieve a significant advancement in the state-of-the-art railway systems by pursuing the way to digitalization and interoperability across Europe. To achieve the expected results activities are carried out in two different working domains. The first domain covers three focus areas; planning, operational activities - which includes managing future interactively coupled timetable planning; and finally operational traffic management systems - which also includes integration activities for door-to-door mobility. The second domain refers to the delivery of a set of digital enablers for all European rail destinations, i.e., crosscutting, to support the development of destination-specific digitalisation solutions. The main outputs of the project will be significant improvements to the railway systems in Europe through the development and implementation of a range of technical enablers. The results of FP1 MOTIONAL will enhance the strategic and tactical capacity planning of the network by enabling a seamless cross border planning and integrated yard and station capacity planning and using modern optimisation and simulation technologies.

Keywords: Mobility · Interoperability · Rail Traffic Management Systems · Simulation · Rail Operations · Cross border

1 EU-Rail Flagship Project 1: MOTIONAL

The EU-Rail Flagship Project "MOTIONAL" aims to improve the planning and operational management of rail services through digitalization, automation, connectivity, and multimodal integration. This project, supported by Europe's Rail initiative, will enable the future European railway system to be interoperable, resilient, and integrated efficiently with different services, including last-mile operations. The project methodology is based on use cases, functional requirements and associated specifications coordinated

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 233–238, 2026.
https://doi.org/10.1007/978-3-032-06763-0_34

with involved stakeholder groups and close dialogue with the EU-Rail System Pillar. Key advancements include the improvement of planning and operational management of rail services, the development of digital enablers like digital twin technology, and the integration of rail with other transport modes. The goal is to make rail the backbone of a multimodal transport system for passengers and freight, with the perspective of achieving a Single European Rail Area. The "MOTIONAL" research project aim to achieve a significant advancement in the state-of-the-art railway systems by pursuing the way to digitalization and interoperability across Europe.

2 Planning

The main objective for Planning is to improve the strategic and tactical planning of the network with optimization and simulation functionality short term (adjustments) and long-term planning (longer than one year). This is done by specification, development and demonstrations.

Today new processes are under development to enable European wide capacity allocation and to connect planning and operational traffic. There is an ongoing technical and digitalization development in CMS (Capacity Management Systems)/TMS (Traffic Management Systems). Planning activities are focused in three areas:

- In cross border planning we improve planning between borders nodes/network areas within nation and between nations.
- In Advanced algorithms we develop and demonstrate decision support for timetable planners/Capacity management system operators.
- In Simulation and operational feedback, we develop simulation methods for improved capacity planning and simulating capacity effects for new DATO technologies (Digital automated train operations) and European Train Control System (ETCS).

Cross Border Planning and Integrated Capacity Planning

One of the objectives is to deliver future railway processes and solutions with seamless cross border planning and integrated yard and station capacity planning. Seven demonstrations are planned to achieve expected results in three technical enablers. These are: European cross-border scheduling with international train path planning; Improved capacity allocation using rolling planning and TTR; Integration of planning systems and TMS with yard capacity planning and station capacity planning.

Advanced Algorithms and Decision Support Tools

Another objective is to deliver optimized timetables by enhancing the processes, as well as by using knowledge from actual operations and simulations. Development of advanced algorithms for the generation and adjustment of timetables and rolling stock planning. The algorithms are demonstrated in decision support tools based on defined use cases and problems/solutions/effects. Four demonstrations are planned to achieve expected results in three technical enablers. These are: Decision support for short term planning; Train path and schedule optimization methods and strategies for capacity efficiency, punctuality and energy saving for different parts of the network and different traffic situations (level of punctuality); Integration of planning systems and TMS with yard capacity planning and station capacity planning.

Simulation and Operational Feedback
To simulate operational processes and to develop feedback loops between operations and planning as well as to develop simulation methods for TMS – C-DAS/ATO and simulate/study effects. Development of simulation methods for evaluating C-DAS/ATO. Development of simulation methods for evaluating European Train Control System (ETCS), such as optimal braking behavior (new generation braking coming from Destination 2) and ETCS Hybrid level 3. Two demonstrations are planned: Improved simulation models; Simulation of new technologies which includes ETCS HL3 and ATO to estimate capacity effects.

3 Operations

Today, rail traffic is managed by IMs (Infrastructure Managers) on national or regional level, usually supported by legacy systems with a poor level of digitalisation and weak integration with systems of other domains or actors participating in the overall traffic planning and management process.

Taking advantage of the potential of digitalisation, operational and technological solutions are developed for different business areas, paving the way towards the implementation of the future European Rail Traffic Management System to make rail the backbone of a multimodal transport system for passengers and freight.

Integration of TMSs and Processes
Within use-cases, processes and interfaces are developed for achieving a much higher integration level of functions and decision processes including the increase of the precision of the traffic prediction as part of TMSs. In focus are also the alignment between different TMS areas including cross-border and the integration of TMS with yard or station functions, Electric Traction Systems as well as RU (Railway undertakings) related data sources for consideration of e.g., actual rolling stock constraints. The availability of appropriate interfaces between the different clients and stakeholders and applications will help to support an aligned re-planning and management of capacity on railway lines and station or yard tracks including graphical visualisation, conflict detection and resolution. The new possibilities will increase the interests and the level of engagement of RUs in traffic operation.

Improved Resilience and Efficiency of Disruption Management
We develop modules for a cooperative multi-actor optimisation and decision support supporting the management of incidents and disruption. Through an advanced HMI the system resilience and efficiency of disruption management using future TMSs will be increased. TMS operators' workload will be reduced by task automation where possible to enable human focussing on critical situations.

More efficient traffic management is also achieved by better consideration of human factors being enabled by new HMI design and technology as well as extended decision support featuring “what-if?” scenario evaluation with a specific focus on disruption and incident management. Automated feedback operational observations to planning allow for improved robustness of future plans.

Linking TMS to ATO/C-DAS for Optimized Operations
Aims on testing integration between TMSs and ATO/C-DAS systems or components providing significant added value by increasing network capacity, timetable robustness, energy efficiency and punctuality. A seamless integration of TMS and ATO will be demonstrated by testing both together in an emulated “real-world” environment by human-in-the-loop demonstration and test bench. The testing environment will cover interaction of train-drivers, TMS operators and signallers, each using their own (emulated) workplace working and communicating together in simulated traffic states. The activities include human-factors research to achieve optimum and realistic results considering behaviour of loco drivers, signallers and TMS operators.

Automated Decisions and Decision Support for Traffic Management Optimization
Focuses on development of algorithms and modules including verification of their suitability for future TMSs providing decision support and whenever possible automatic decisions for traffic management based on optimisation and real-time conflict detection & resolution. Automation is demonstrated e.g., for requesting real-time Movement Authorities in ETCS L3 hybrid and L3 full moving block-based operation.

Different algorithmic or technology approaches including AI are used to demonstrate their application railway networks, lines and local environments such as yards, terminals, depots or stations. Different solution options are evaluated in accordance with “What-If?” and impact analysis paradigms prior to triggering the implementation of the final decision in the operational plan. Thus, disruption and operational imbalances can be dealt with providing a higher level of confidence.

4 Multimodal Integration

In terms of operational outcomes, the integration with other transport modes to deliver door-to-door mobility is a major driver for the development of rail attractiveness. The very objective of MOTIONAL in this area is to make Rail the backbone of Multimodality. This is supported by three technical and business pillars described in the text below.

Inclusive Mobility
Commuting between transport modes, especially for travelers with disabilities has always been complex and time consuming leading to a reluctance to use rail transport. MOTIONAL answer to that consists in a set of services focusing on the environment In railway hubs in order to ease customer journeys, we foresee: Cutting hedge technologies delivering Hands-Free User Experience; In Door Navigation with UWB (Ultra-Wide Band), BLE (Blue Tooth Low Energy) and Biometrics; Platform based guidance, as well as Interactive digital travel assistance, and finally In station PRM services including dedicated kiosks and communication with staff and escort, and/or Interactive assistance during transfer.

B2B Integration
Multimodal integration is not achievable without the cooperation between Mobility Providers. In order to favor this cooperation, MOTIONAL proposes the improvement and development of B2B platforms and services in the areas of Sales, Distribution, Financial services and Traffic information. Key technical features associated with these business objectives includes Cross-operator and cross-mobility support and the Compatibility with SaaS deployment, as well as the usage and enhancement (when applicable) of standards including: OJP, OSDM, SIRI, NETEX.

Anticipate Demand
Anticipating passenger demand and adjusting the service accordingly is key to rise travelers' satisfaction and to optimize mobility providers' operation. MOTIONAL ambition in this area is to design solutions and services integrating multiple transport modes, allowing hence a thigh and dynamic cooperation between rail and other transport modes. Features in scope includes Long-term demand forecast calculation, and short-term demand forecast calculation in real-time. It also includes simulation of demand and associated reaction of mobility networks by using Digital Twins, as well as Disruption management, and finally processing demand data across mobility modes.

5 Digital Enablers for the European Rail System

Digitalization is a major and fundamental transformational process of the European Rail System that pervades all domains: from delivering passenger services to maintenance, from engineering, construction and building to energy management, from traffic management and train operations to financials and accounting, and others. While the range of specialized advanced digital applications needed to cover such a large scope is very large and diversified, a set of underlying common enablers has been identified by the EU-Rail partnership to provide fundamental capabilities that can be leveraged by each specialized field of application.

Rail Federated Dataspace
A cybersecure, reliable, scalable, distributed, data sovereignty-preserving and interoperable mechanism for data sharing and communication across and open ecosystem of autonomous business entities and heterogeneous systems is a recurring problem in the Rail System, particularly as the progress of digitalization extends the requirement across multiple domains. The MOTIONAL project is developing the "Rail Dataspace", a natively distributed system for content-independent, cybersecure, reliable, scalable, interoperable, data sovereignty-preserving data exchange across a federation of participants providing by design of digitalized trust, and governance policy enforcement and digitalized data provision contract negotiation. The Rail dataspace will be designed as the underlying technological infrastructure of the European Data Strategy and supported by actual available open-source software.

Conceptual Data Model
While the Rail Dataspace must provide an essential mechanism for data sharing a communication that is independent of the data contents, it must be complemented with a set

of machine-readable and processable description of this data contents likewise independent of how it is shared. Conceptualization of the digital representation of Rail concepts, objects and phenomena will allow for heterogeneous and autonomous entities participating in the exchange supported by the Rail dataspace to communicate not just the data, but the meaning and intent of the exchange. Only the combination of common conceptual data models and common data sharing and communication mechanisms independent of each other can provide full interoperability of heterogeneous systems and autonomous organizations.

Digital Twins Development and Execution
Finally, while the adoption of digital twin technology able to model structure and dynamic behavior of cyber-physical systems accelerates in the Rail System, the need is recognized to provide developers with a common set of architectural guidelines, modeling languages and tooling to enable the composition of complex simulation applications involving interactions between multiple cyber-physical systems. Specialized digital twins developed independently by multiple organizations must therefore also be able to interoperate. To this end, the MOTIONAL project will deliver a common digital twin development and runtime environment for the Rail System as an enabler of the development of natively interoperable digital twins.

Acknowledgements. Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Europe's Rail Joint Undertaking. Neither the European Union nor the granting authority can be held responsible for them. The project "FP1-MOTIONAL" is supported by the Europe's Rail Joint Undertaking and its members.

References

1. Europe's Rail Multi Annual Work Plan – Europe's Rail Webpage. https://rail-research.europa.eu/wp-content/uploads/2022/03/EURAIL_MAWP_final.pdf. published on 2022/03/01
2. Interview with FP1 Motional Coordinator – Europe's Rail Webpage. https://rail-research.europa.eu/news/europes-rail-fp1-motional-shares-ambitions-for-future-european-rail-traffic-management/. published on 2023/05/31

Intelligent Access, e-waybills and eFTI Developments in Estonia

Taavi Tõnts[1] (✉), Marko Jürimaa[1], Eva Killar[2], and Ulrika Hurt[3]

[1] Transport Administration, Valge 4, 11413 Tallinn, Estonia
taavi.tonts@transpordiamet.ee
[2] Ministry of Economic Affairs and Communications, Suur-Ameerika 1, 10122 Tallinn, Estonia
[3] Digital Logistics Centre of Excellence, Teaduspargi 6/1, 12618 Tallinn, Estonia

Abstract. The research objective is to contribute to more effective and greener road freight transport with HV's, without damaging aging road infra below. One of the best solutions in 21st century is to use all kinds of digital data (temperature; GNSS; OBW etc.) and road maps to control the logistics in the most optimum way, depending on the used vehicle's load type (Intelligent Access IA). It's a very cost-effective system. Since the HVTT15 & HVTT16 conferences, much work has been done in Estonia in the logistics digitalization area. In Transport Administration the e-waybill has been piloted to become in 2024 as a compulsory measure in building contracts for the greener infra buildings. Estonian Ministry of Climate has developed the architecture and started to pilot eCMR and eFTI in wider scale and also more widely all over Europe.

Keywords: Intelligent Access (IA) · Estonian VELUB System · OBW Automated Mass Control Integration · CEDR Road Freight Transport (RFT) workgroup · Estonian eCMR and e-waybill development · EU electronic freight transport (eFTI) regulation

1 Introduction

Since 2010 Estonian Transport Agency (ETA) has been developing Intelligent Access (IA). In the beginning, it was meant only for the wintertime 52t timber transport, if the pavements were frozen at the min 0,5m depth. For that reason, the digital vertical temperature sensors (max 2,5m deep) were installed over Estonia [1]. At the same time, the strong infra corridors were analyzed and mapped [2] (in the SmartRoad – the green color digital road corridors with daily update time at 4 pm). In the VELUB system, it is possible to apply for the special vehicles permit for up to 1 year [3].

As the climate change proceeded, the wintertime allowed corridor window decreased quickly, we started in ETA to develop the year around 52t (the violet corridors) IA system [4] with the legislation [5]. IA has been used since 2015 by ca350 truck/y, mostly by timber trucks so far, but also increasing others.

ETA signed on 9th April 2020 the memorandum to develop with 9 other stakeholders the e-waybill for the bulk material transport in road maintenance sector. After that we

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 239–245, 2026.
https://doi.org/10.1007/978-3-032-06763-0_35

have had many pilot contracts and are going to use it as mandatory since year 2024 in our road building.

To be ready for the application of the provisions of the Regulation of the electronic freight transport information (eFTI), Estonian Ministry of Economic Affairs and Communications has developed and tested during the years 2020–2023 a cross-border eCMR indexing prototype together with the partners from the Nordic Baltic region.

Also, Estonia as a lead partner, together with other EU Member States have just recently in spring 2023 started a new, 3-year long project eFTI4EU co-funded by Connecting Europe Facility (CEF) with a big part of it being piloting the eFTI Gate solution.

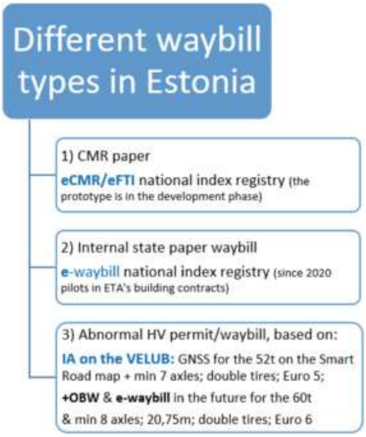

Fig. 1. – Different paper & e-waybill types in Estonia

In the following paragraphs, a very short overview will be given about Estonian's different e-waybills developments and research (see Fig. 1).

2 VELUB Intelligent Access

Since the HVTT15 paper [6], much work has been done in Estonia in the logistics digitalization area which has been also described shortly below to understand in the better way, the different cloud-based logistic systems, and their common ground (see also the HVTT16 paper [7] and the EU Aeroflex project workshop [8]).

For example, if the (i.e. abnormal) HV is in the wrong road corridor or has a total OBW [9] mass over the limited value etc., the cloud-based eCMR can change the cell red and send the notice to controlling organizations (and also on the Smart Road map can blink as the red dot, etc.). It's possible in the longer future to make automated direct enforcements as well – kind of like the speed cameras does now.

In CEDR report Intelligent Access (IA): current NRA practices [10], there has given good overview about IA ongoing developments. Also, in the Revision of the Weights and Dimensions Directive 96/53/EC [11] there is already mentioned eFTI and IAP using possibilities. The proposal is also establishment of **technical and operational standards for information exchange on the transport of indivisible loads** and the development of Intelligent Access Policies (**IAP**).

3 eCMR Pilots and Ongoing Process

EU Regulation 2020/1056 Electronic Freight Transport Information (eFTI) demands EU member states to accept the digital transport documents at the level of competent authorities starting from the first half of 2026. In order to be ready for eFTI Estonian Ministry of Economic Affairs and Communications (as of July 2023 Ministry of Climate) has developed and tested during the years 2020–2023 a cross-border eCMR indexing prototype between Estonia, Latvia, Lithuania, and Poland during the DIGINNO-Proto, DINNOCAP and The NDPTL Goes Real-Time economy: eCMR projects. During the pilot projects the technical development took place, and the concept followed a logic of not uploading any eCMR documents to the central database but linking them via national indexing scheme. Central point or connecting layer for secure data exchange between economic operators and competent authorities such as Police or Customs was created and called as index registry (see Fig. 2) [12].

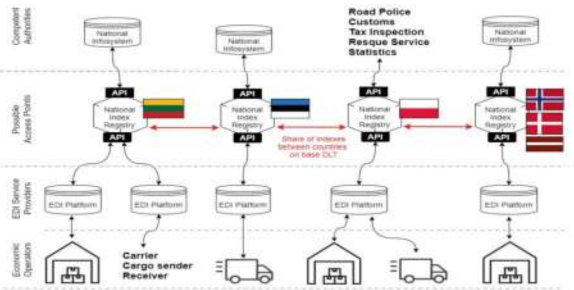

Fig. 2. Cross-Border eCMR indexing ecosystem

The process of establishing implementing acts for eFTI is ongoing at European Commission and it is an honour to note that the architecture that was tested during these above-mentioned pilot projects proved to be also the preferred option for eFTI technical architecture (chosen from the 8 different options).

The findings from the pilot projects confirmed the understanding that a successful distributed eCMR and eFTI architecture needs a middle layer - a National Access Point at every member state (see Fig. 3) in order the cross-border data exchange to function.

In 2022 Estonia conducted an analysis on Estonian National Access Point [13] which provided a comprehensive and exhaustive mapping of the different options, solutions, risks, and threats for the development, ownership and location of eFTI NAP (now called eFTI Gate), taking into account the legal and economic specificities of the different economic sectors.

During the same years, and intensively starting from 2019, Estonian experts from the abovementioned projects have provided support to European Commission and its

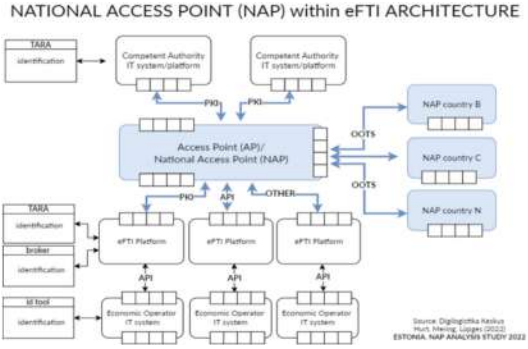

Fig. 3. Estonian NAP Analysis Study 2022 [13]

expert group Digital Transport and Logistics Forum, gathering feedback and reflecting the architecture principles for eFTI and the structure of indexing and identifiers.

In 2023, Estonia as a lead partner together with project partners from Finland, Lithuania, Germany, France, Italy, Portugal, Austria and Belgium and observers from Spain, Ireland and the Netherlands got funding from CEF for project called eFTI4EU. The goal of the eFTI4EU Project envisages to develop the eFTI Gate technical requirements to the fullest and establish harmonised rules for trusted networks of eFTI Gate platforms and its components. Project involves in total 22 different partners (public and private) from the abovementioned countries. The total budget is 28.3 million euros from which 50% is funded by CEF, and the project will last three years – from 2023 to 2026.The work is carried out in parallel with the setting up of the requirements of the eFTI Platforms and Service Providers – the ones responsible for managing and maintaining the transport data for economic operators.

4 Internal State e-waybill Using Experience

In 2020, ETA carried out three procurements, where it was mandatory to use an e-waybill of lading for the transport of the bulk materials for the road building. The special requirements for documenting, the works applied to both: the contractor and the owner's supervision. All consignment notes for the bulk materials and the summary tables, compiled based on them, had to be prepared in an electronic data exchange platform [14].

The pilot projects assumed the use of either the Waybiller (a service provider) environment developed in Estonia or an analogous electronic data exchange platform (at the moment 3 different service providers). The procurement required that the digital platform allow the creation of separate objects and GNSS location-based tracking of each load. The vehicle and/or trailer number information had to be generated automatically from the traffic register database and had to reach the electronic environment on the e-waybill. If the truck had a special cargo permit of 48 or 52 tons, its data, permit number, and the validity period had to be included. Supervisors and subscribers needed to have access to the environment so that they could control the information through the cloud.

The e-waybill had to contain at least the number of the truck and/or trailer, the number of axles, the permissible weight/load capacity of the truck; the mass and name of the material, the name of the driver, the hauler, the owner of the load and the quarry. If the transport was on public roads, an e-waybill had to be created for the transport from the intermediate warehouses to the site (except for the intermediate warehouses immediately adjacent to the site). An indication had to be made on the e-waybill, whether the material came from a quarry or an intermediate warehouse. Building contract annex (2 pages) has been prepared for contractors, which contains detailed minimum technical requirements for the e-waybill. This will make it easier for the new service providers to bring data exchange platforms to market and will be better traceable to all parties [15].

As mentioned, the owners' supervision was also obliged to use the electronic environment of the e-waybill. For example, he had to check the e-waybills for bulk materials provided by the contractor and confirm receipt of the load in the digital environment. Also, during the asphalting works, the engineer used a data exchange platform to validate the e-waybill of asphalt loads arriving at the site. In total, more than 5000 paper waybills were avoided by using e-waybills in the 3 smaller pilot building contracts.

In 2022–2023 over 30 pilots were made to prepare for the full transition to e-waybills in 2024, for the transport of the bulk materials.

5 Conclusions

E-waybill development in Estonian state road building, abnormal transport IA, and other areas have given us already a lot of savings in CO_2 emissions, making transport more efficient at the same time - with the single data entry to the cloud, etc. Defiantly there is still a lot to do and develop in coming years, to achieve EU's climate targets in the transport sector, using digitalization as a tool for helping it.

Estonia together with neighbouring countries, and other EU states, are working together to achieve the EU digitalisation goals and to make the freight transport between the states, and inside the state, more efficient.

ETA's bulk material e-waybill piloting and usage is just a small part of internal transport (ca 5%), but it is a huge step to digitize all internal waybills (over 2 mln/y) together with Climate Ministry, and together with international eCMR, based on the eFTI regulation.

In 2024 ETA is planning to demand e-waybills in all building contracts. There are already now many interested IT companies with big interest to develop e-waybills in Estonia, as it's not so complicated in 21 century anymore.

Digitized cloud-based transport and road information systems are allowing much further IA development - to protect road constructions and increase traffic safety. At the same time, transport transparency increases. Hopefully standardized IA data fields (vehicle data, see p.2) will be compatible with the eFTI data fields (freight data, see p.3), so the EU road owners can have good tool in the future to protect the aging infra and at the same time to allow more efficient and greener transport. Standardized fields allow the service provider to offer the standardized service all over EU and compete each other on the free EU market with the increasingly better service – it opens the doors to the better and greener HV transport system in EU.

References

1. Study of the relationship between the weather conditions of embankment freezing and the bearing capacity of different roads 2010, AS Teede Tehnokeskus. https://transpordiamet.ee/media/3236/download
2. Mapping of possible routes for overweight (52 tons) cargo on national roads, National Road Administration 2010 (2010). https://transpordiamet.ee/media/3239/download
3. E-service of the Estonian transport administration for applying the special permits (VELUB). https://eteenindus.mnt.ee/main.jsf?lang=en
4. Estonian ETA's SmartRoad digital year-round max 52t map (the violet corridors). https://tarktee.mnt.ee/#/en/link/GIQJOMqQE61p
5. Estonian abnormal transport, valid legislation. https://www.riigiteataja.ee/akt/109092015002?leiaKehtiv
6. Estonian VELUB system for more efficient and greener transport VELUB System is a small part of the much bigger, theoretical, MVMC System (En dimensional movement controlling system), T.Tonts, 2018 HVTT15 (2018). https://hvttforum.org/wp-content/uploads/2019/11/Taavi-Tonts-ESTONIAN-VELUB-SYSTEM-FOR-MORE-EFFICIENT-AND-GREENER-TRANSPORT.pdf
7. HV intelligent access and e-waybills development in Estonia, T.Tõnts, M.Jürimaa, i. Nosach. https://hvttforum.org/wp-content/uploads/2021/10/Tonts-Hv-Intelligent-Access-And-E-Waybills-Development-In-Estonia.pdf
8. Experiences with the stakeholder process from Estonia, T.Tonts (2019). Aeroflex/i4DF, Special Focus Workshop, Intelligent Access Part II, Paris. Aeroflex_i4DS_Stakeholders_ERA_TT_300919_TT.pdf
9. OBW tests with HV's (Tänapäeva ja tuleviku rakendused sõidukite kaalumisel, sõidukite masside jälgimissüsteemi loomine ja jälgimine reaalajas), Maanteeamet, ETA, AS Teede Tehnokeskus, Fleet Complete Eesti OÜ, Tallinn 2020y. https://transpordiamet.ee/media/3112/download
10. Conditions for efficient road transport in Europe, CEDR (2017). https://www.cedr.eu/download/Publications/2017/2017-5-Conditions-for-efficient-road-transport-report.pdf
11. Revision of the weights and dimensions directive 96/53/EC. https://www.eesc.europa.eu/en/our-work/opinions-information-reports/opinions/revision-weights-and-dimensions-directive-9653ec
12. Final report eCMR index registry prototype (2020). Estonian ministry of economic affairs and communications. https://www.diginnobsr.eu/ecmr-testing-results
13. NAP analysis. Study on the operational model for eFTI NAP. Final report. Hurt et al. (Estonia 2022).pdf (realtimeeconomy-bsr.eu)
14. E-waybill future possibilities in road owners' point of view, Roadpaper (Teeleht), pp. 29–33 nr 100, 2020y, Taavi Tõnts, ETA. https://www.transpordiamet.ee/media/1435/download
15. EU sees "YES" for eCMR (Euroopa ütles e-veoselehele "jah"), Roadpaper (Teeleht), pp. 19–21 nr 102, 2021y, Taavi Tõnts, Inna Nosach, Ulrika Hurt, ETA. https://www.transpordiamet.ee/media/1433/download

Hybrid Approach for Estimation of Traffic Hazards: Fusion of ML and Pure Statistical Model

Natasa Mojic[1](✉), Vijay Mudunuri[1], Radim Slovák[2], Thomas Mariacher[1], and Peter Hrassnig[1]

[1] ASFINAG Maut Service GmbH, Lakeside B03, 9020 Klagenfurt, Austria
natasa.mojic@asfinag.at
[2] TietoEVRY, Tieto Austria GmbH, Handelskai 94–96, 1200 Vienna, Austria

Abstract. Precise travel and traffic information are becoming increasingly important to road users. This is backed up by a growing number of users who are planning their trips before they start their journey. People primarily aim to avoid or minimize the travel delays by knowing alternative routes or changing the departure time beforehand. Furthermore, they are increasingly customizing their travels to maximize efficiency and convenience. From years of experience, and considering the Traffic telematics report [1], ASFINAG, as the national operator of Austrian motorways and highways, has recognized this need and has focused on bringing the accuracy of our traffic information to real traffic conditions. In this paper we present our hybrid approach that combines two methodologies to provide traffic information that is closer to the real traffic conditions, with a special focus on the traffic jams and delays. Our main goal is to improve the accuracy of the travel delays and traffic jams both in terms of time and location. The first methodology, which is a fusion of data from various data sources, is based on pure statistical model, whereas the second methodology is a machine learning (ML) model. This hybrid approach is not a comprehensive one but rather restricted to selected road sections.

Keywords: Traffic state · Traffic jams · Machine Learning · Data Science · Data Fusion · C-ITS

1 Overview

ASFINAG plans, funds, builds, maintains, operates, and collects toll along almost 2,249 km of Austrian motorways and expressways. In the last years, we have been extending our sensor infrastructure where most of the sensors support current wireless technologies such as the CEN Dedicated Short-Range Communication (Toll stations) [2], Bluetooth (BT), WLAN, ITS-G5 (C-ITS) [3], and TLS (Technische Lieferbedingungen für Streckenstationen) also called as speed detectors in this paper.

Based on our analysis of the traffic data, it's clear that relying on a single statistical model to estimate traffic jam and travel delays across the entire road network is not the

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 246–253, 2026.
https://doi.org/10.1007/978-3-032-06763-0_36

most effective approach. Apart from the traffic volume in urban and rural areas, traffic jams and travel delays are impacted by traffic events, and road and weather conditions. Keeping these additional parameters in mind, we came up with idea of developing a hybrid approach for traffic jam and travel time delay estimation. Here we combine the statistical model with the ML model to make our estimation more context aware. We primarily focus on the cases or regions where the pure statistical methods are not giving us satisfactory results. In this paper, we present this approach as well as the comparison and the evaluation results of non-ML (pure statistical model) model and the ML model. We accomplish this as a part of project and system "ASFINAG Traveltimes & Trafficstate Management System" (ARMS) [4].

2 Non-ML Approach

In the non-ML (pure statistical) model, we employ fuzzy logic to determine weights for individual data sources: Bluetooth and WLAN detectors, Toll transactions from the trucks, and speed detectors. The raw data from these data sources is post processed to calculate the single vehicle travel times for predefined section. The only exception is the data from the speed detectors. The data from the speed detectors are resampled in one minute intervals. The result is an average speed of the vehicles in a 60 s time window. To harmonize the traffic data from different sensor types, we have divided our road network into 200-m segments, which are referred to cells. The measurements obtained for a particular cell are interpolated to the nearest cell. Each nearest cell is assigned a weight, which is determined by factors such as measurement age, variance, and the proximity to the measurement location. Linear interpolation is the method we utilize for this purpose.

$$c_k = \beta_k \cdot d(mcell, icell) + \beta_{max} \tag{1}$$

$$\beta_k = (\beta_{max} - \beta_{min}) \cdot |\alpha|_T + \beta_{min} \tag{2}$$

$$|\alpha|_T = \frac{var_T}{var_{max}} \tag{3}$$

$$var_{max} = \sum\nolimits_{t=0}^{N} \lambda(t) \tag{4}$$

$$var_T = \sum_{t=0}^{N} \lambda(t) \cdot |v_t - v_{T-t}| \tag{5}$$

Equation 1 calculates the confidence level (c_k) of the received measurement based on the distance between the measured cell (*mcell*) and the interpolated cell (*icell* – cell to which we interpolate the measurement) and slope coefficient β_k. Equation 2 calculates the slope coefficient based on the standardized slope (weighted variance / max weighted variance). Weights are calculated by using a time function $\lambda(t)$ (the older the measurement lesser the weight). Equations 4 and 5 calculate the weighted and maximal variance values.

Once the data from each data source is assigned and interpolated to every cell in the network, our ARMS system performs a weighted mean calculation to determine a traffic state for each individual cell. The traffic state is quantified on a numeric scale ranging from 0 to 100. We refer this value as "availability". A value of 100 signifies that the current traffic state is indicative of free-flow conditions, whereas a value 0 indicates stationary traffic. Furthermore, we have employed an aggregation technique to assess the extent of a traffic jam. In this approach, we spatially and temporally aggregate the neighboring cells based on the similarity of the calculated traffic conditions.

3 ML Approach

In this section we explain our ML model, which is implemented to predict future travel times. Like the non-ML, the ML-model also uses the single vehicle travel times from the Bluetooth, WLAN, and Toll detectors, and the average speeds sampled over one minute from the speed sensors. Likewise, the data is harmonized to bring to a homogeneous data structure. In addition to these data, the model incorporates contextual information such as weather data, planned and unplanned events, and calendar data. To train our model, we have considered the last 6 months of data from above mentioned data sources. Although the 6 months of data appears to be relatively small, it has proven to be sufficient for our needs. It is detailed and consistent, which enabled the model to reach our expected accuracy in travel time prediction.

The data from Bluetooth, WLAN, and Toll detectors are preprocessed with Hampel outlier detection method [5]. The Hampel's approach is designed to be robust when dealing with extreme values in the data set. This is especially relevant in the travel times data set as some vehicles could stop in a parking area for a significant amount of time, resulting in extreme travel times. Concerning to the data from the speed detectors, it is important to note that the data used for our model has already resampled at 1-min intervals. This ensures the data is free from outliers. After the outlier removal, the travel times data are resampled at N-minutes time window (where N = 15, 20, 30), followed by feature engineering. In our model this involves assigning time features, holidays, unusual travel time patterns, etc. A detail explanation of feature engineering is skipped, as the focus is on the hybrid model.

The modeling is done using XGBoost [6] regression model, which is further finetuned by hyperparameter random grid search and 5-fold split cross-validation. XGBoost is an advanced machine learning technique, which operates by constructing an ensemble of decision trees to predict continuous numeric values. Each decision tree learns from the errors of its predecessors iteratively, allowing XGBoost to progressively refine its predictions. During training, the algorithm assigns initial predictions and calculates the associated residuals. Subsequent trees are then built to minimize these residuals. To prevent overfitting, regularization terms are introduced, and trees are pruned based on their depth and leaf weight. The trees predictions are combined, with each tree contributing to the final result, proportionally to its performance. XGBoost also uses a technique called gradient boosting, which updates the model's parameters by minimizing the gradient of the loss function. This iterative approach optimizes the model's fit to the data, resulting in a robust regression model capable of handling complex relationships and noisy datasets.

The assessment of model performance typically relies on metrics such as root mean square error (RMSE) and mean absolute error (MAE). These metrics were computed on a test set, resulting in an RMSE of 11.1 s and an MAE of 2.2 s. These values, while relatively small, align with expectations for scenarios without traffic congestion. Estimating travel time under normal, free-flowing conditions is generally straightforward.

However, during traffic jams, accuracy in travel time estimation is challenged by unpredictable events, such as vehicle accidents. Although the XGBoost model successfully identifies all traffic jams, it struggles to provide precise travel time predictions in such scenarios. Consequently, the error metrics are notably higher, with an RMSE of 126.3 s and an MAE of 107.6 s. That also explains why RMSE is higher than MAE because it is more sensitive to outliers.

4 Advanced Data Fusion – Hybrid Approach

As explained in Sect. 3, our non-ML approach relies on a pure statistical model to estimate the traffic state and the extent of traffic jam. The travel time and travel delay calculations are averages of travel time samples collected over specific time window. One notable drawback of these methods is their lack of predictive intelligence for anticipating traffic jams and delays. At times, they report travel time delays, which significantly deviates from the real-time conditions. As a result, traffic jams are identified late, and queue lengths are not measured with expected accuracy. Figure 1 illustrates a situation where our statistical method did not yield precise travel-time delay that match the actual conditions. This resulted in discrepancies of up to 30 min when compared to the real time delays experienced. Therefore, to narrow down these discrepancies, we have developed an ensemble strategy that combines our pure statistical model with the outcomes of our ML model. We refer this strategy as the "hybrid approach".

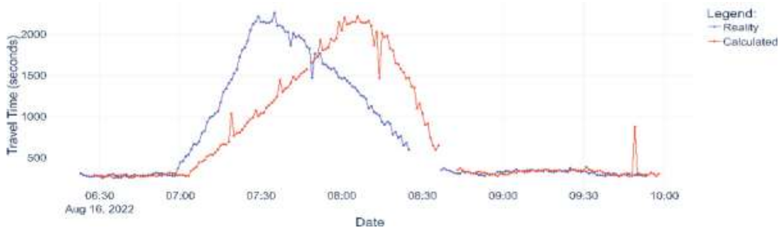

Fig. 1. Delay in the estimation of travel time delays

It is important to understand that this approach does not replace our conventional approach (non-ML), instead, it serves as a complementary strategy. Our goals are as follows: 1) Accurately estimate travel time delays, approaching real-time accuracy, 2) Identify the traffic jams at an earlier stage and 3) Enhance the precision of estimating the starting and ending points of traffic congestions.

Reference [7] describes various statistical and ensemble methods to integrate the results from multiple models. In our hybrid approach, we have chosen Bayesian statistics [8] to calculate individual weights for each data source, and thereby include or exclude a specific data source for traffic jam prediction based on their weightage. The data sources here also include the outcomes of the machine learning.

From the historic traffic data, we have pre-selected the traffic hotspots (sections and time periods). For each of these hotspots we have estimated prior probabilities for the data sources based on the fact whether they have identified the traffic jams within a threshold time or not. Although the data sources are independent, they have a strong correlation due to the similarity of the data that they produce. Therefore, we not only can calculate individual probabilities but also a joint probability of the likelihood of predicting a jam.

4.1 The Bayesian's Model for Data Source Selection and Aggregation

In this model, we consider our main data sources Bluetooth and WLAN data (B), toll data (T), and speed sensors (S). To find out the likelihood of each of these data sources detecting a jam, we compute the conditional probabilities of traffic jam (J) given average speeds from speed sensors P(J|S), toll detectors P(J|T) and Bluetooth detectors P(J|B) using Bayes rule. Subsequently, we also calculate the conditional probability of observing measurements from the data sources S, T, and B given the presence of traffic jam J by using the below formula:

$$P(J|S, T, B) = \frac{P(S, T, B|J)P(J)}{P(S, T, B)} \tag{6}$$

where P(S,T,B) represents the joint probability of observing the measurements from the data sources S,T,B, regardless of the traffic jam conditions.

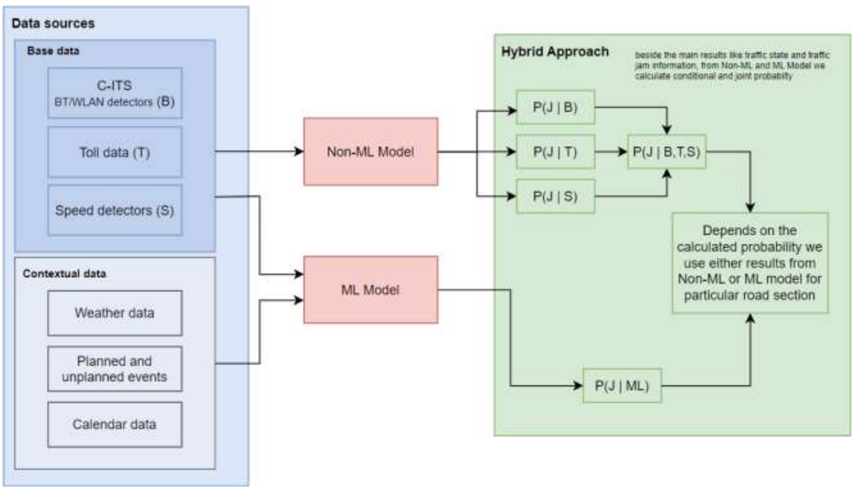

Fig. 2. Data Flow Visualization with focus on Bayesian's model for data source selection and aggregation

We first apply the joint probability, to check if it is greater than 0.7, indicating that there is strong likelihood of predicting the jam. In these cases, we assume that the data sources are capable of estimating the jam with less discrepancies to the reality without the help from the ML prediction. However, if the joint probability is less than 0.7 then:

1. We evaluate the conditional probabilities of each data sources to remove those with a conditional probability less than 0.7
2. We calculate the conditional probability of the prediction from ML model $P(J|ML)$
3. Finally, we perform a weighted average of the remaining data sources and the ML predictions.

This way we can dynamically assess and select the most reliable data sources for predicting the traffic jams by considering the outcomes from the ML model when needed.

5 Evaluation of Non-ML and ML Model

In this section an in-depth comparison between the outcomes of non-ML and ML models is presented. For this we also used RMSE and MAE metrics. Travel times obtained from non-ML model is rescaled by a constant to compensate a different length between start and end sensors position of ML model. Data from Google maps are utilized for more detailed, accurate comparison to non-ML and ML model.

Table 1 presents the overall metric results, and Fig. 2 illustrates the comparison between the ML model and non-ML approaches. Both approaches predicted the traffic jams and travel times in free flow cases correctly. However, the ML model predicted them earlier. In Fig. 2, we can see that the ML model is the first source to predict the change in the traffic trend. The difference between ML model and non-ML model data could be explained by the fact that for ML model features from BT and speed detectors data are more significant than from Toll transactions and thus is not able to predict travel time specially for trucks but instead predicts travel time for all vehicle classes. On the other hand, non-ML is calculating travel time separately for cars and lorries. The ML model predicted the change in the travel time pattern (free flow to queuing traffic) at least 5–10 min faster than our non-ML model and Google data. Likewise, it also detected the second pattern (queuing to free flow traffic) 15 – 20 min faster than the others (Fig. 3).

Table 1. Comparison between non-ML and ML model travel time RMSE and MAE metrics results together with comparison to non-ML and Google data.

	RMSE [s]	MAE [s]
ML to Google data	27.2	9.6
non-ML car to Google data	52.9	17.5
non-ML lorry to Google data	52.9	18.7
ML to non-ML car	26.2	7.0
ML to non-ML lorry	24.7	8.7

Fig. 3. Comparison of the results from ML and non-ML models together with Google data for road section on A23 during selected traffic jam.

6 Conclusion

Due to the emergence of various detection technologies, we have an increased diversity of data sources which help us in improving the accuracy of the traffic information. However, to fully leverage the strengths of these diverse data, we need an advanced approach that can effectively ensemble data from all the available data sources. Our paper focused on this important topic, where it presented a hybrid approach to ensemble data from multiple data sources, and furthermore ensembles the outcomes of the ML model when needed. At present we have employed this approach on selected road section. As part of the future work, we shall implement this proposed approach on the entire ASFINAG road network. Furthermore, we shall also research and experiment with other ML and AI models to conclude if they fit better for our use cases. We shall specifically focus on Graph Neural Network to model the whole high-level road network and compare the results with our current approach.

References

1. Federal ministry Austria. https://www.bmk.gv.at/themen/mobilitaet/alternative_verkehrskonzepte/telematik_ivs/publikationen/berichte/bericht2023.html, Accessed 25 Sep 2023
2. DSRC. https://www.itsstandards.eu/25-2/cen-dsrc/. Accessed 25 Sep 2023
3. C-ITS. https://www.car-2-car.org/about-c-its. Accessed 7 Jan 2024
4. Mariacher T., Bretis K., Rainer B., Hrassnig P., Pletzer F.: ARMS – Asfinag traveltime management system. In: Proceedings of 7th Transport Research Arena TRA 2018
5. Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.A.: Robust Statistics: The Approach Based on Influence Functions. Wiley, New Jersey (1986)
6. Chen, T., Guestrin, C.: XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794. ACM, New York (2016)
7. Mienye D., Sun Y.: A survey of ensemble learning: concepts, algorithms, applications, and prospects. IEEE Access **10**, 99129–99149 (2022)
8. Chipman, H.A., George, E.I., McCulloch, R.E.: Bayesian ensemble learning. In: 20th Annual Conference on Neural Information Processing Systems, NIPS, San Diego (2006)

Future Mobility Campus Ireland: A Testbed for Advanced and Autonomous Vehicles

Diarmuid Ó Conchubhair(✉), Russell Vickers, Wassim Derguech, Danijel Benjak, and Brían ó Cualáin

Future Mobility Campus Ireland, Shannon Free Zone, Co. Clare, Ireland
diarmuid@fmci.ie
https://www.futuremobilityireland.ie/

Abstract. The Future Mobility Campus Ireland (FMCI) was founded with the sole purpose of creating and delivering future mobility testbed facilities for stimulating research, development and innovation in the area of Autonomous Connected Electric Shared Vehicles (ACES), including Connected and Autonomous Vehicles (CAV) and Advanced Air Mobility (AAM), in Ireland. FMCI focuses on comprehensive mobility technologies that span both ground (autonomous driving, micro-mobility, smart cities, V2X communications) and air (unmanned drones, eVTOL, AAM, UTM) uses.

FMCI supports a range of parties, from individual researchers to multi-national corporations, as well as start-ups and government entities. FMCI assists organisations to conceive, develop, trial, and deploy societally transformative transport solutions both nationally and internationally.

The installation of state-of-the-art connectivity technologies at the FMCI testbed is providing an environment for the automotive industry, component suppliers, telecommunications companies and research institutions with the opportunity to adopt new approaches to develop innovation for future technologies. FMCI has been operational for over a year and is already facilitating various companies and higher education institutes (HEIs) to further their innovation and/or technologies.

Keywords: Testbed · Transport · Mobility

1 Background

Mobility has a major impact on our economy and our society. Travel of the future will focus on the needs of the individual rather than the operators. It will be multi-modal involving a mixture of cars, buses, trains and planes [1]. Future mobility must be more environmentally friendly, more efficient and overall, much smarter. It will be driven by data and enabled by the latest technological advances. FMCI [2] seeks to facilitate the technological solutions which will enable Ireland to benefit socially and economically from more sustainable and efficient transport systems.

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 254–259, 2026.
https://doi.org/10.1007/978-3-032-06763-0_37

In 2017 the CAV (Connected & Autonomous Vehicles) Ireland [3] forum was formed with members from IDA, EI, SFI, Lero, Insight, and Intelligent Transport Systems (ITS) working with industry partners from across the island of Ireland. The group focussed on identifying a value proposition for CAV in Ireland and helping to identify policy and regulatory barriers to our adoption of CAV, as well as the need for infrastructure to support the emerging ecosystem of companies involved in the development, manufacturing and testing of products and services related to CAV. The group concluded that an entity such as FMCI would support growth and expansion in these companies while providing opportunities to attract new inward Foreign Direct Investment (FDI) and create new start-up opportunities for Irish researchers and entrepreneurs (Fig. 1).

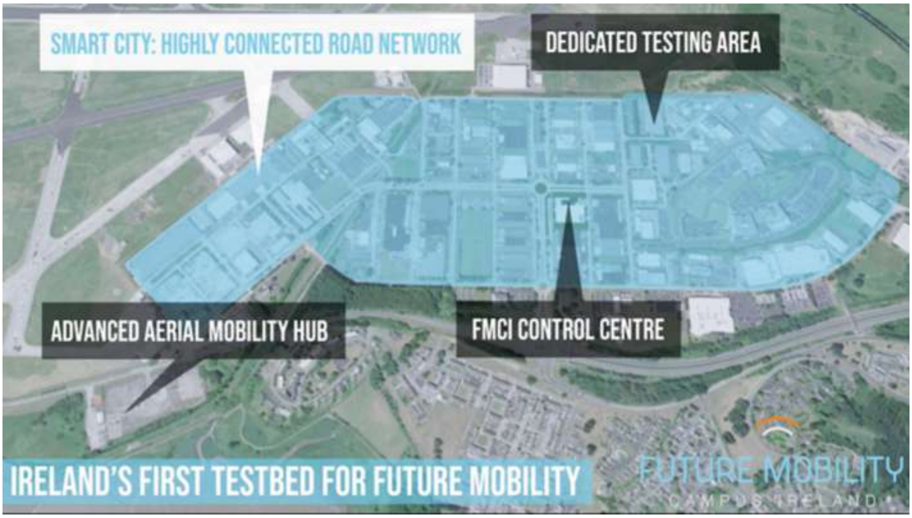

Fig. 1. Illustrating FMCI facilities at the Shannon free Zone (Image reproduced with permission from Future Mobility Campus Ireland (FMCI), copyright FMCI, 2024).

FMCI is delivering on the provision of a CAV test facility located in real-world settings, providing technology companies and researchers the ability to test and enhance their innovative solutions. The test facility is a non-for-profit entity and consists of a road network which is retrofitted with interconnected state of the art sensing and telecommunication technologies, all of which are fully accessible and controlled via a dedicated control centre located within the FMCI Headquarters (HQ). Within its first year of operations FMCI has expanded its offering which was initially focused on the provision of CAV text facilities, to also include aerial mobility R&D facilities.

FMCI facilities are available to its users/customers via a rental model with target users/customers of the facility including national and international automotive and technology providers that are interested in testing/validating their technology in real-world settings before the product release phase.

2 FMCI - An Automotive Testbed

FMCI's open testbed spans over 4 km of public road network within the Shannon Industrial Estate. The testbed is highly connected with fibre network to our dedicated data centre. It is equipped with multiple sensors, CCTV, a set of traffic lights along with high accuracy location systems, as well as a data management and control centre. FMCI provides its customers access to an advanced control room capabilities enabling insightful and data-driven monitoring as well as visual analytics of mobility data. The control room is fully equipped with an audio-visual system for real-time monitoring of the testbed (Fig. 2).

Fig. 2. (above left) and 3 (above right). Demonstrating a selection of FMCIs test vehicles (Images above reproduced with permission from Future Mobility Campus Ireland (FMCI), copyright FMCI, 2024).

At the Future Mobility Campus Ireland headquarters, customers can avail of private and secure workshop bays with all relevant services including high-speed broadband, a private virtual network, and a secure facial recognition access control. Customers are also offered access to a private or a common laboratory for any hardware works, etc. FMCI's fleet includes a Jaguar I-Pace and two Land Rover Defenders. All vehicles are retrofitted with a wide range of sensing and data processing technology which are secured to a flexible and adaptable roof rack, connected to a data logger equipped with large data storage capacity, and fully integrated to our in-house data centre. Recently FMCI has also added a mobile control centre which allows for additional monitoring and data analysis in offsite areas which may not be covered by our fibre connection.

2.1 Stimulating Innovation

New technologies and communications systems now embedded in vehicles are resulting in connectivity that opens up new possibilities, redefining mobility and changing public and private transport options [4]. While this is disrupting traditional business models it is also delivering opportunities for new business growth. FMCI facilitates innovation by providing start-ups; SMEs, MNCs as well as researchers with the infrastructure to trial, test, demonstrate and validate solutions for systems and technologies in a "laboratory with real-life conditions."

2.2 FMCI Customer Use Case: Provizio

Provizio [5] is an Irish SME who have been based at FMCI for the past 12 months. Provizio's technology fuses radar technology and on-board vehicle camera sensors with AI-on-the-edge processing and SLAM into one affordable, software-enhanced platform which can identify potential hazards with a high level of accuracy and speed. Each sensor is fully software-defined and integrated on an edge GPU - where proven AI-powered imaging techniques further improve their point cloud by more than 10x and in five dimensions. Provizio has been an anchor tenant for FMCI since its inception and utilise all aspects of FMCI infrastructure including, office rental for staff, test vehicle for technology integration, workshop space for hardware testing and vehicle integration as well as the FMCI data centre.

2.3 Diversity of R&D Projects

In its first 12 months of operation, FMCI has hosted a plethora of diverse R&D projects across a variety of organisations and research preforming entities, some of which include:

- Jaguar Land Rover: JLR Shannon [6] is at the cutting edge of Jaguar Land Rover's vision for the future and play an important role in developing new technologies to support electrification and self-driving features.
  - FMCI partner and several ongoing technological projects in conjunction with FMCI including V2X using FMCI fleet vehicles and testbed infrastructure.
- Valeo [7]: Vision systems and automated parking, producing physical hardware and generating software intellectual property with artificial intelligence.
  - FMCI partner as well as using FMCI facilities to further develop their technologies.

3 FMCI – Facilitating Advanced Aerial Mobility

FMCI Air targets a high growth potential market that is Advanced Aerial Mobility (AAM). From an industry perspective there are good synergies between air mobility and road mobility. Many of the same industry and research partners who are developing products for the future of road transportation will also utilise these products for air mobility. Users of the FMCI 'Drone port' can also avail of the facilities and infrastructure as listed above including the office space, laboratories, workshops, data centre and control room. The FMCI 'Drone Port' is designed to be scalable and mobile. To support testing and operations, the drone port offers a flight operations centre and a mobile unit to allow flexibility for off-site testing. The flight operations centre is based on a modular container equipped with the necessary radio and communications equipment to ensure flight safety.

4 FMCI – Leveraging National Funding

FMCI has also targeted European funding to further develop and demonstrate research projects in a real-world environment while also expanding its portfolio of clients throughout Europe. By leveraging nationally funded infrastructure, FMCI has been successful in a European proposal entitled 'ÉALÚ-AER [8]' which is a three-year project involving Irish and European partners. ÉALÚ-AER is one of the first European Digital Sky Demonstrators to focus on U-space and Urban Air Mobility and aims at accelerating the delivery of an inclusive, resilient, and sustainable Digital European Skies through research and innovation. Over the past 6 months FMCI has also been requested to join 6 European consortia with regard to submitted R&D proposals under the Horizon Europe funding programme, all of which have either an automotive or ariel research component.

References

1. Spickermann, A., Grienitz, V., von der Gracht, H.A.: Heading towards a multi-modal city of the future? Technol. Forecast. Soc. Chang. **89**, 201–221 (2014). https://doi.org/10.1016/j.techfore.2013.08.036
2. FMCI homepage. https://futuremobilityireland.ie/. Accessed 10 Oct 2023
3. CAV Ireland homepage. https://www.cavireland.com/. Accessed 10 Oct 2023
4. Alessandrini, A., Campagna, A., Site, P.D., Filippi, F., Persia, L.: Automated vehicles and the rethinking of mobility and cities. Transport. Res. Procedia, vol. 5, pp.145–160 (2015). https://doi.org/10.1016/j.trpro.2015.01.002
5. Provizio homepage. https://provizio.ai/. Accessed 10 Oct 2023
6. JLR Shannon homepage. https://www.jaguarlandrovercareers.com/content/Ireland/?locale=en_GB. Accessed 10 Oct 2023
7. Valeo Ireland homepage. https://www.valeo.com/en/ireland/. Accessed 10 Oct 2023
8. ÉALÚ-AER homepage. Accessed 10 Oct 2023

The More You Know: Digital Twins of Travelers

David Käthner(✉), Klas Ihme, and Erik Grunewald

German Aerospace Center, Lilienthalplatz 7, 38108 Braunschweig, Germany
david.kaethner@dlr.de

Abstract. Despite increasing digitalization, rail operators and other mobility service providers often need more real-time information regarding the occupancy of trains and stations. Worse, only general assumptions regarding passengers' specific destinations, individual needs, and current and past experiences can be made. Digital Twins of Travelers (DTT) address the issue by holding real-time digital representations of passengers throughout their journey from start to destination. Being informed about a passenger's route instead of isolated sections enables customized real-time forecasts, accessibility information for each transfer location, and veridical updates about how much time may be spent there. For example, a traveler would benefit directly when parts of a journey must be rescheduled due to delays on the original route. When continuing with means of transportation different from those chosen initially, forecasts and other information can be immediately adapted to the new route. In the same scenario, mobility service providers would benefit by being informed about occupancy changes, allowing them to prioritize scheduled connections if necessary. DTT consist of a real-world twin, its digital counterpart, and an interchange component linking them. The digital twin represents how travelers experience current events, holds information regarding their individual itineraries' progress, and can exchange information with mobility service providers. DDT's purposes range from supporting passengers in real-time by providing information relevant to their journey milestones, analyzing the effect of events such as delays on the travel experience, to helping to align transport services with the existing demand.

Keywords: Human Digital Twins · Public Transport · Travel Experience · Demand Forecast · Knowledge Management

1 Human Digital Twin Systems

Digital twins were initially conceived in the context of product lifecycle management (PLM), intended as a structured approach to the design, manufacturing, deployment, and maintenance of physical products. [8, 9] argued that successful PLM requires detailed knowledge of the product state at all times and that a continually updated digital representation of the product can provide this knowledge. In this conception, knowing the state of a digital twin equals knowing the state of the represented object. The digital twin allows for complete predictability of its physical counterpart because of a mechanism linking the digital twin's virtual space with the real-world twin's real space. The

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 260–266, 2026.
https://doi.org/10.1007/978-3-032-06763-0_38

mechanism continuously updates the digital twin, enabling data-generating processes like forecasts and analytics whose results may flow back toward the real space.

The core idea of coupled virtual representations of physical objects has received much attention, particularly in domains where product-related processes generate detailed yet isolated data [12]. However, the purposes of digital twins have been explored beyond their original conceptualization, precluding uniform definitions of what such twins may encompass. [14] suggested a purpose-related classification with two dimensions. The first dimension differentiates between an intended use for exploratory purposes, where no information flows from the virtual to the real space, and an intended use for decision-taking, where such information flows are characteristic. The second dimension differentiates between three kinds of entities the digital twin can represent. The first kind of representation concerns types instead of specific instances, for example, to enable virtual testing. The second kind represents individual instances, such as those required to track the life of a specific resource. The third kind of representation addresses more general system properties like resource consumption.

In recent years, the application of the digital twin concept to humans and human-related processes has been discussed. In their review of human digital twins, [13] discussed use cases in medicine, ergonomics, and sports performance. Notably, the authors suggested an ontology of Human Digital Twin Systems (HDTS) comprising a real-world twin, its digital twin counterpart, and an interchange component linking both. Additional components, such as a visualization engine, can significantly enhance the system's effectiveness, particularly for exploratory purposes such as during the prototyping stages of system development.

Minimally, the real-world twin comprises one human agent, one sensor to collect information regarding the current environment, and one process, such as a manufacturing workflow in which human behavior occurs. In the virtual world, the digital twin contains models of the human agent, the mission or task to be accomplished by the human agent, and performance goals to track how well the task is executed. Further, a prediction engine allows forecasting how changing properties within the virtual world would influence human or system performance, particularly regarding performance goals. Finally, the HDTS interchange component allows for bidirectional communication between the real and the virtual world and data-related functions such as persistent storage, cleaning, and fusion.

A critical difference between the concept of HDTS and PLM-related systems is the need to consider the enormous variety inherent to human experiences, human behavior, and the environments in which they occur. Further, there is an essential difference between measuring physical attributes and human-related constructs. In the case of physical products, many relevant attributes can be represented by measuring physical properties such as temperature. Precluding measurement errors, a user of the digital twin system can safely infer the real-world twin's state based on these measurements. In the case of human-related attributes, that is not necessarily the case. For example, based on an indicator such as a low heart rate, a person's stress level might also be classified as being low. However, such a classification can be substantially misleading without knowing that person's base rate. The relation between the level of indicator variables and the assumed level of human-related attributes is far less certain than for physical products. As such, mapping the real-world twin's state onto the digital twin's state often amounts to a profound measuring problem. Interpreting human-related measurements often requires extensive context information, primarily if few constraints govern the expected human behavior, like in many mobility systems.

We argue that HDTS can substantially help addressing problems of measuring and interpreting human behavior in open and complex environments. Designing, maintaining, and improving human-centric mobility services requires knowledge of many human-related attributes. Such knowledge does not have to be instantaneous or all-encompassing and should be restricted to the relevant information regarding the represented person. Still, a core tenet in Human Factors Engineering is "to ensure that something goes well, it is necessary to understand what goes on" [10, p. 358]. Achieving such an understanding requires integrating many aspects of travelers' experience during their journey, and what is relevant for an individual traveler may not always be evident from the outside. For mobility providers, travelers are only visible when entering their area of responsibility. For an individual traveler, however, the entire journey is essential, regardless of who is responsible for ensuring the journey goes well.

As such, representing the way travelers take through the various spaces of their journey poses similar challenges as representing physical products in the context of PLM. As a domain-specific application of HDTS, we argue that the concept of Digital Twins of Travelers (DTT) is a promising approach to represent humans as they journey through the traffic system and to integrate multiple data sources indicating their experiences during this time. Crucially, the approach provides a conceptual basis to represent the mobility-related context of the travelers' experiences, holding information about chosen routes or the journey trajectory concerning reached travel milestones [15]. A twin system can help focus on a subset of human-related phenomena and the specific situations in which they occur, for example, to aid in understanding complex issues such as travel mode choice based on real-world data. In short, a DTT provides the framework to implement conceptualizations of mobility as a travelers' journey through time and space [1].

Implementing the DTT concept requires (a) a clear theoretical and empirical basis for the construct of travel experience to know which human attributes can be represented and (b) use cases for applying the DTT concept to know what should be represented. Both aspects will be discussed in the following sections. Finally, open questions regarding data privacy and data ownership will be discussed.

2 Travel Experience

As an analogy to the construct of user experience that relates human affective experiences to technical system usage [20], travel experience refers to the collective subjective impressions and evaluations a traveler has in the context of a journey [19]. Travel experience is not restricted to using vehicles such as shuttles, trains, or cars but also addresses travel-relevant infrastructure such as stations, hubs, and technical systems such as ticket vending machines or apps [2, 3, 5, 11]. Further, relevant experiences may include the feeling of pleasantness when looking out of the window in a calm high-speed train, the frustration when learning that a bus has been canceled, or the discomfort caused by standing on a crowded platform.

Unfortunately, when using trains and other public transportation, many passengers regularly experience delays, missed connections, a lack of journey-relevant information, or long waiting times. If the negative experiences outweigh the positive experiences over a sufficient period, the likelihood of using individual modes of transportation, such as private cars, can rise considerably. However, we argue that properties of the public transport system, such as the state of the infrastructure or the number of delays, are only one factor influencing travel experiences and may not always be the most important one. As an affective-emotional state, journey-related experiences are not one-dimensional and do not perfectly correlate with any single measurement [18].

Instead, travel experiences are heavily influenced by a person's goals, resources, and cognitive and physical state. Depending on those factors, a situation such as a train's delayed arrival at the next station could lead to various cognitive and emotional responses. Whereas some passengers may be indifferent because the station is their final destination, others could experience significant stress for fear of missing a connecting train with drastic consequences for the remainder of their journey. Particularly in case of missed connections, travelers with access to relevant information may be able to reschedule their journey quickly and consequently experience much less uncertainty than those relying on outside help to adapt their itinerary.

Ideally, the travelers' digital twin can represent a broad spectrum of emotional travel experiences such as frustration, uncertainty, anger, fear, discomfort, flow, or pleasantness. However, deriving such a wide variety of states from sensors such as fitness trackers, which can plausibly accessed during the travelers' journey, is a considerable challenge. Instead, a lean but robust approach mapping indicators of arousal to positive or negative valence could serve as a more practical first step [16]. Recent work has employed this approach when assessing travel experience, recording valence or similar constructs based on one or two self-report items and arousal using physiological indicators [2–4]. Although this approach provides only a low resolution of the emotional spectrum, it operationalizes the construct traveler experience as an immediately applicable method that yields valuable results.

3 Use Cases for Digital Twins of Travelers

The first use case concerns exploring factors influencing passengers' travel experience. Representing attributes of specific persons with known itineraries allows a much better understanding of how they solve the task of traveling to their destination, react in case of unexpected events, and evaluate certain situations along their journey. Crucially, understanding factors influencing the travel experience can be improved iteratively by analyzing the data, formulating hypotheses based on insights gained, and testing those hypotheses by making predictions regarding novel situations. If specific context data is required to understand travel experience, it is possible to add such data post-hoc by querying relevant sources.

A second use case addresses the support of passengers directly during their journey. For example, by improving the digital twin's process model to represent the task-relevant goals of travelers during their journey, providing the real-world twin with relevant information to achieve their next goal becomes considerably more straightforward. Whereas the construct of travel experience addresses what a journey feels like, a DTT must also represent what needs to be done to complete the journey, i.e., hold a passenger's itinerary. In this view, the digital twin of a traveler is a construct for mapping, updating, and forecasting all relevant individual customer requests regarding their transportation. Specifically, itineraries can be modeled as the essential points or milestones of the service trajectory of the journey. Milestones include the journey's starting point and destination, the means of transportation on each leg of the journey, and, importantly, required transfers and associated waiting times between the legs of the journey [6]. During the journey, the DTT updates arrival and departure times for the upcoming milestones depending on the actual traffic situation and adjusts the transfer times accordingly. Such a real-time itinerary would enable communication of a passenger's progress on their specific travel chain, allowing them to manage delays and disruptions of transport services with much better consideration of individual needs and preferences [17].

As a third use case, a DTT could be a valuable data source for better adjusting supplied services to the actual demands for transportation. Whereas travelers often know in advance the specific relations they will use, mobility service providers do not have access to that information. Thus, decisions regarding offering transport services must be based on forecasts, potentially causing providers to miss business opportunities for relations where an unknown demand exists. In local public transport, ticket options such as transit passes place few constrictions on how customers use existing transport services, making the capacity planning for specific routes challenging. Even if a service has been booked well in advance, it is not certain that the customer will use it. Obtaining insight into current and planned journeys and perhaps individual preferences may allow mobility service providers to provide more targeted services than today, even in day-to-day operations. For example, if passengers miss a connection due to a service failure, the impact will vary depending on the trip's purpose. Taking the aggregated impact of service failures into account could help allocate resources in a more targeted way.

Finally, the fourth use case concerns the predictions enabled by the DTT's process model. Reducing complex journeys to essential milestones discretizes otherwise continuous processes, making simulations of complex relations in the traffic system much more accessible. For example, the approach allows for predicting the effect of travelers' preferences on the usage of buses or trains. Preferences refer to variables such as the time travelers are willing to wait on average for a connection, the price they are willing to pay, or the maximum vehicle utilization they will tolerate. When considering changes to existing transport services, the effect of those changes can be simulated together with varying preference distributions. It may be challenging to validate the veracity of the simulation's outcome. However, we argue that even considering travel experience when planning transport services is a crucial first step to aligning offered services with existing demand.

4 Data Privacy and Data Ownership

In a nutshell, the concept of DTT suggests combining data indicating travelers' emotional state with relevant context information underpinned by a powerful abstraction of the traveling task. We have argued that such an approach offers many attractive features supporting a better understanding of human behavior and experiences in complex environments and solving mobility-related practical issues. However, we also believe that implementing the concept requires further discussions regarding data privacy and data ownership. To be useful, the digital twin must represent its real-world counterpart's highly sensitive personal data, such as physiological data, self-reports, timetables, and traveler trajectories. As described in the use cases, leveraging the DTT's full potential requires sharing parts of the collected data with third parties, like mobility service providers. As such, effective precautions to eliminate, minimize, and mitigate threats to data privacy are indispensable. On the other hand, the EU's General Data Protection Regulation (GDPR) already regulates many aspects of collecting and processing private data. Any implementation and operation of DTT would need to adhere to the GDPR and other data protection regulations.

In any case, we suggest thoroughly addressing data privacy and data ownership already on the architectural level. First, travelers using a digital twin must have complete control over the collected data, especially regarding sharing such data with third parties. Second, if users are unwilling to share some data, they need to be able to use the subset of the DTT's functionality, which is still available. Third, data must be anonymized or at least pseudonymized immediately. For example, aggregated data is likely sufficient if a mobility service provider is interested in how service failures impact traveler experience. Fourth, data ownership at each processing step must be clear. As such, roles and responsibilities ensuring compliance with data privacy regulations must be considered as part of the DTT design phase.

Funding Source. This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No: 101101973. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

References

1. Al Maghraoui, O., Vallet, F., Puchinger, J., et al.: Modeling traveler experience for designing urban mobility systems. Des. Sci. **5**, e7 (2019)
2. Barría, C., Guevara, C.A., Jimenez-Molina, A., et al.: Relating emotions, psychophysiological indicators and context in public transport trips: Case study and a joint framework for data collection and analysis. Transport. Res. F: Traffic Psychol. Behav. **95**, 418–431 (2023)
3. Bosch, E., Luther, R., Ihme, K.: Travel experience in public transport: a geospatial analysis by experience mapping. In: Stolze, M., Loch, F., Baldauf, M., et al. (eds.) Mensch und Computer 2023, pp. 417–421. ACM, New York (2023)
4. Brandebusemeyer, C., Ihme, K., Bosch, E.: Travelers' information need in automated vehicles - a psychophysiological analysis. In: 2022 Human-Centered Cognitive Systems (HCCS). IEEE, pp. 1–6 (2022)
5. Castro, M., Guevara, C.A., Jimenez-Molina, A.: A methodological framework to incorporate psychophysiological indicators into transportation modeling. Transport. Res. Part C: Emerg. Technol. **118**, 102712 (2020)

6. Engler, E., Gewies, S., Banyś, P., et al.: Trajectory-based multimodal transport management for resilient transportation. Transp. Probl. **13**, 81–96 (2018)
7. European Parliament: Council of the European Union Regulation (EU) 2016/679 of the European Parliament and of the Council
8. Grieves, M.W.: Product lifecycle management: the new paradigm for enterprises. Int. J. Prod. Dev. **2**, 71–84 (2005)
9. Grieves, M., Vickers, J.: Digital twin: mitigating unpredictable, undesirable emergent behavior in complex systems. In: Kahlen, F.-J., Flumerfelt, S., Alves, A. (eds.) Transdisciplinary Perspectives on Complex Systems, pp. 85–113. Springer International Publishing, Cham (2017). https://doi.org/10.1007/978-3-319-38756-7_4
10. Hollnagel, E.: The changing nature of task analysis. In: Salvendy, G., Karwowski, W. (eds.) Handbook of Human Factors and Ergonomics, pp. 358–367. Wiley, Hoboken (2021)
11. Ihme, K., Christ, T., Walocha, F., et al.: Digitizing travel experience: assessing, modeling and visualizing the experiences of travelers in shared mobility services. In: European Transport Conference 2023 (2023)
12. Liu, M., Fang, S., Dong, H., et al.: Review of digital twin about concepts, technologies, and industrial applications. J. Manuf. Syst. **58**, 346–361 (2021)
13. Miller, M.E., Spatz, E.: A unified view of a human digital twin. Hum.-Intell. Syst. Integr. **4**, 23–33 (2022)
14. Sjarov, M., Lechler, T., Fuchs, J., et al.: The Digital twin concept in industry – a review and systematization. In: 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). IEEE (2020)
15. Rudolph, F., Grunewald, E., Tremer, D.: Dividing journeys in milestones – a passenger (agent) trajectory reference. Transport. Res. Procedia **72**, 2046–2053 (2023)
16. Russell, J.A.: A circumplex model of affect. J. Pers. Soc. Psychol. **39**, 1161–1178 (1980)
17. Scheelhaase, J., Scheier, B., Grunewald, E., et al.: How to manage delays and disruptions at intermodal transportation hubs in a better way? Eurasian Bus. Econ. Perspect. **26**, 135–145 (2023)
18. Scherer, K.R.: What are emotions? And how can they be measured? Soc. Sci. Inf. **44**, 695–729 (2005)
19. Schiefelbusch, M.: Analyzing and assessing the experience of traveling by public transport. JPT **18**, 46–72 (2015)
20. Schrepp, M., Hinderks, A., Thomaschewski, J.: Design and evaluation of a short version of the user experience questionnaire (UEQ-S). IJIMAI **4**, 103 (2017)

Different Facets of Artificial Intelligence-Based Predictive Maintenance for Electric Powertrains

Ali Serdar Atalay[1], Ahu Ece Hartavi Karcı[2], İbrahim Arif[4], Salih Ergün[3,4], and Alper Kanak[3,4](✉)

[1] AI4SEC OÜ, Harju Maakond, Kesklinna Linnaosa, Tornimäe tn 7-5, Talinn, Estonia
ali.atalay@ai4sec.eu

[2] Center of Automotive Engineering, University of Surrey, Guildford GU2 7XH, UK
a.hartavikarci@surrey.ac.uk

[3] Ergünler R&D Co. Ltd., Pirimehmet Mh. 113. Cad. No. 75, Isparta, Türkiye
{salih.ergun,alper.kanak}@erarge.com.tr

[4] Ergtech Sp.Z.O.O., ul. Ryżowa 49, 02-495 Warszawa, Poland
{ibrahim.arif,salih.ergun,alper.kanak}@ergtech.eu

Abstract. Maintenance, traditionally perceived as a reactive cost and a hindrance, poses challenges to efficiency when components succumb to unforeseen breakdowns. In addition to the financial implications, the repair process also incurs substantial time wastage. To overcome these obstacles and achieve enhanced efficiency and cost savings within the manufacturing sector, this paper presents a conceptual study of a technologically advanced predictive maintenance (PdM) approach, particularly in the realm of artificial intelligence-powered digital twins. The effectiveness of these solutions hinges on their data-driven nature, technical feasibility, and acceptance by industry stakeholders.

Keywords: Predictive Maintenance · Artificial Intelligence · Electric Powertrains · Performance Optimization · Operational Efficiency · Sustainability

1 Introduction

The multifaceted nature of predictive maintenance for electric powertrains entails comprehensive approaches to diagnose the internal physical phenomena of various components within electric vehicle (EV) powertrains. By integrating advanced sensing technologies, data analytics, and machine learning algorithms, predictive maintenance enables real-time monitoring and analysis of key performance indicators, facilitating early detection of faults or abnormalities. This proactive approach not only optimizes the performance, longevity, safety, and reliability of electric powertrains but also mitigates unplanned downtime, lowers maintenance costs, and enhances overall operational efficiency. Through the utilization of multidimensional predictive maintenance strategies, stakeholders in the electric vehicle industry can achieve significant advancements in powertrain maintenance and contribute to the development of sustainable transportation solutions.

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 267–273, 2026.
https://doi.org/10.1007/978-3-032-06763-0_39

This study aims to investigate the facets and benefits of implementing predictive maintenance strategies, specifically focusing on artificial intelligence-based approaches, in the context of electric powertrains. The European Union-funded ESCALATE and OPEVA approach aims to assure the electric vehicle user reaches its current destination to boost customer satisfaction, as it offers not only a safe journey but also promises no more breakdowns before the destination is reached. By examining the multifaceted aspects of predictive maintenance, including advanced sensing technologies, data analytics, and machine learning algorithms, this research aims to explore how these building blocks are brought together to enhance the performance, longevity, safety, and reliability of electric powertrains. Furthermore, the study seeks to evaluate the impact of advanced predictive maintenance on reducing unplanned downtime, optimizing maintenance costs, and improving overall operational efficiency. Through data-driven insights, this study aims to provide valuable insights and recommendations for stakeholders in the electric vehicle industry, contributing to the development of sustainable transportation solutions.

2 State-of-the-Art

Concurrently, extensive research is underway to enhance BMS performance, particularly in terms of predicting battery ageing and ensuring safety. These techniques also aim to predict battery lifespan and detect faults. To address these challenges, James C. Chen et al. [1] propose empirical mode decomposition (EMD), grey relational analysis (GRA), and deep recurrent neural networks (RNN), whilst Y. Che et al. [2] propose optimized health indicators and online model correction with transfer learning for the RUL prediction of lithium-ion batteries.

The practical integration of AI into power electronic applications encompasses deterministic and stochastic environments, data collection and analytics methods, forecasting techniques, and cost-effective algorithms and hardware implementations [3]. In recent years, physics-informed PdM methods aligned seamlessly with the strategy of making AI a frontrunner in e-mobility applications. Among these, Li et al. [4] propose a model to calculate the cumulative fatigue damage of IGBT modules in EVs to evaluate the reliability. Rao et al. [5] propose a novel approach to PdM of power electronics by integrating optical and quantum-enhanced AI techniques.

The electric motor, a pivotal component within the electrical powertrain of electric vehicles powered by batteries, operates under demanding conditions characterized by significant temperature fluctuations, pronounced vibrations and voltage stress stemming from inverter power supply [6]. Wotawa et al. [7] employ model-based, simulation-based and machine learning-based real-time PdM approach for e-motors. Rjabtšikov et al. [8] and Dettinger et.al [9] use a Digital Twin-based approach when dealing with the predictive maintenance of e-motors.

Recent developments [10] employ AIoT-based preventive diagnostic by enabling recommendation and notification prediction for Run-to-Failure maintenance, planned preventive maintenance, and predictive maintenance for EV thermal management system. Yi et al. [11] utilize Long Short-Term Memory (LSTM) based Digital Twin to achieve real-time temperature forecasting and analyze degradation models for lithium-ion batteries. Dettinger et al. [9] introduce a machine learning-based model for fault detection in EV powertrains, employing a Digital Twin. In parallel, Bhatti et al. [12] advocate the application of Digital Twin technology in EVs through their systematic review.

3 Digital-Twin-Powered PdM Services Architecture

Integrating technologies like IoFog, Arrowhead, and Mimosa into predictive maintenance (PdM) for Electric Vehicle (EV) powertrains can significantly enhance real-time data processing, communication, and overall management of maintenance activities by leveraging the capabilities of edge computing, interoperable automation, and standards for enterprise integration respectively. The PdM services are designed to be used as online services, in the form of PdM-as-a-Service, within a digital twin framework as depicted in Fig. 1. The framework has a data acquisition and management module which is built on a secure multi-agent system. This module is implemented as a heterogeneous data flow backbone based on an Arrowhead-ioFog-Mimosa architecture (See Fig. 2). This module is strengthened with cyber security solutions enabling very fast agent data encryption and person/node authentication. The multi-agent system enables effective data acquisition from EVs, grid infrastructure (i.e., charging infrastructure) and environmental data.

The acquired data can be classified as JSON files and stored in a cloud database, such as MariaDB, as proposed in the digital twin framework. There exist auxiliary backend services such as data and query management services, cloud services management, security and privacy management and Identity Management System (IDMS).

A two-way Kafka-based messaging system is designed to feed the Smart Services Modules which are utilized as Docker containers at the righthand side. The Docker containers receive the classified data in JSON format or other messages or requests from the secure multi-agent system and backend services. When the Digital Twin services including AI-powered PdM solutions produce an output, these outputs are also delivered to the visualization and user interfaces layer through a results gateway, or as feedback to the Secure Multi-Agent System (for instance, in case of an anomaly).

The framework is realized through a layered architecture (Fig. 2) that utilizes the combination of Arrowhead [13], Eclipse ioFog [14] and MIMOSA [15] and is composed of i) Edge layer (ioFog), ii) Communication & Automation Layer (Arrowhead), iii) Enterprise integration layer (MIMOSA), and iv) Execution layer.

The **Edge Layer** facilitates decentralized data processing, enabling real-time analysis of powertrain data (such as battery voltage, current, and temperature) directly within the EV or local edge nodes. This layer implements LSTM or other predictive models on the edge for immediate anomaly detection and alerting for powertrain issues. Here ioFog can manage data transmission over the cloud, ensuring only relevant or summarized data is sent, which is vital for efficient bandwidth usage and managing large fleets of EVs. This layer delivers an Edge-AI-based processing that enables localized decision-making, and reduces the need to transmit all raw data to a central server, thus saving bandwidth and enhancing data privacy. Instantaneous actions, such as emergency alerts or minor adjustment suggestions, could be generated directly from the AI models deployed at the edge.

The **Communication & Automation Layer** provides a framework for ensuring interoperable communication between various systems involved in a PdM solution space, such as between EVs, charging stations, and maintenance facilities. This layer is implemented by deploying an Arrowhead framework which is an open-source technology that establishes a standardized and secure interoperability platform for seamless communication and interaction between diverse services and devices. The Arrowhead-based solution facilitates automated workflows that enhance PdM, like scheduling maintenance activities, based on predictive insights while considering logistic and operational constraints.

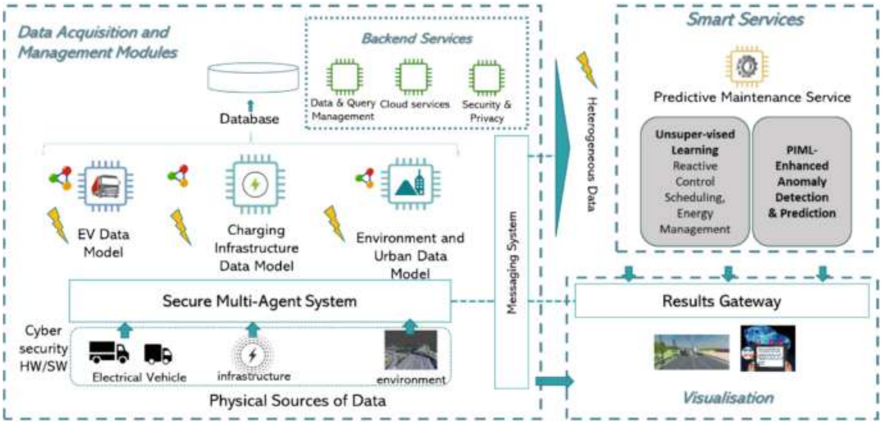

Fig. 1. Proposed Digital Twin Framework (Reproduced with permission from AI4SEC OÜ, copyright AI4SEC OÜ, 2024).

The **Enterprise Integration Layer** is implemented by utilizing the MIMOSA framework which is based on the Open System Architecture for Enterprise Application Integration (OSA-EAI) standard. Implementing MIMOSA standards ensures consistent data handling and interpretation across various entities involved in PdM systems, such as service providers, component manufacturers, and operators. Utilizing MIMOSA's OSA-EAI standards for enhanced data sharing and integration between enterprise systems and field-level entities ensures that PdM insights and actions are cohesively managed and executed across the organization (e.g. fleet operator, vehicle manufacturer).

The **Execution Layer** is the highest level of applications that are actively used by the end users such as maintenance service providers, vehicle operators, etc. The bi-directional arrows between this layer and the Enterprise Integration Layer indicate the cooperative work of service maintenance scheduling and coordination.

The four-layer architecture enables an effective data and action flow between layers. Data flow starts with raw data acquired from edge nodes like edge devices, gateways, smart grids, and/or onboard systems. Processed data and initial insights are transferred between the edge and communication layers. The standardized and structured data are propagated to the enterprise layer along with global insights and alerts. On the other hand, action flows can be instructions for actual maintenance activities that are streamed from the enterprise layer to the execution layer. Feedback, status updates and any corrective actions are exchanged between the execution and enterprise layer.

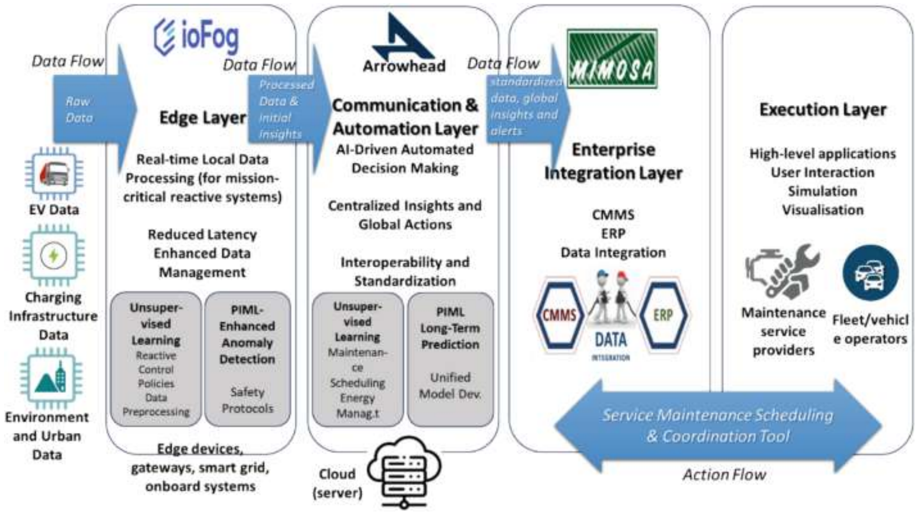

Fig. 2. The layered Digital Twin Architecture (Reproduced with permission from AI4SEC OÜ, copyright AI4SEC OÜ, 2024)

In both edge and communication & automation layers, unsupervised learning (UL) algorithms can be deployed for reactive control and data preprocessing as well as maintenance scheduling and energy management. The proposed architecture at the edge layer implements UL algorithms to develop local control policies that react to real-time data, adjusting operating parameters to enhance performance or mitigate anomalies. The edge layer also enables optimal energy management and creating policies that consider both localized powertrain health and broader energy demands, like grid stability or energy prices. Physics-Informed Machine Learning (PIML) for more effective PdM can also be tacked both at the edge and global level. ioFog can be used to implement PIML for real-time analysis, using physics-based principles to enhance anomaly detection in powertrain components, ensuring that the alerts are not just data-driven but also physically plausible. Here, safety protocols leverage PIML to establish safe operating boundaries for the powertrain, using physics principles to prevent settings or actions that might risk component health or safety. At the automation layer, PIML models can be used for analysing long-term physical degradation patterns in powertrain components, ensuring that PdM strategies are aligned with anticipated wear and tear. A unified model can be developed to integrate physics-based principles with data-driven insights, presenting comprehensive models that ensure predictive insights and prescribed actions that are scientifically substantiated.

4 Conclusion

This paper presents a digital twin framework based on a 4-layered architecture that integrates Edge (ioFog), Communication & Automation (Arrowhead), Enterprise integration (MIMOSA), and Execution layers. IoFog, Arrowhead, and Mimosa are utilized for orchestrating physics-informed PdM and UL for EV powertrains so that organizations can create a robust, efficient, and scalable system. This integration not only enhances real-time data processing and communication but also ensures that the maintenance activities are streamlined, standardized, and optimized for the best possible performance and longevity of EV powertrains.

Acknowledgement. This work has been supported by the European Union's Horizon Europe ESCALATE and OPEVA projects under grant agreements No. 101096598 and 101097267, respectively.

References

1. Chen, J.C., Chen, T.L., Liu, W.J., Cheng, C.C., Li, M.G.: Combining empirical mode decomposition and deep recurrent neural networks for predictive maintenance of lithium-ion battery. Adv. Eng. Inf. **50**, 101405 (2021)
2. Che, Y., Deng, Z., Lin, X., Hu, L., Hu, X.: Predictive battery health management with transfer learning and online model correction. IEEE Trans. Veh. Technol. **70**(2), 1269–1277 (2021)
3. Zhao, S., Blaabjerg, F., Wang, H.: An overview of artificial intelligence applications for power electronics. IEEE Trans. Power Electron. **36**(4), 4633–4658 (2020)
4. Li, L.-L., et al.: Predicting the power module cumulative damage degree in new energy vehicle: improved Manson model. J. Cleaner Prod. **374**, 133945 (2022)
5. Rao, P.S., et al.: Integrated artificial intelligence and predictive maintenance of electric vehicle components with optical and quantum enhancements. Opt. Quant. Electron. **55**(10), 1–19 (2023). https://doi.org/10.1007/s11082-023-05135-7
6. Galea, M., et al.: Reliability-oriented design of electrical machines: the design process for machines' insulation systems MUST evolve, IEEE Ind. Electron. Mag. **14**, 20–28 (2020)
7. Wotawa, F., et al.: Real-time predictive maintenance–model-based, simulation-based and machine learning based diagnosis. AI for Digitising Industry–App. River Publishers, pp. 63–81 (2022)
8. Rjabtšikov, V., et al.: Digital twin service unit development for an EV induction motor fault detection. In: International Electric Machines & Drives Conference (IEMDC). IEEE (2023)
9. Dettinger, F., et al.: Machine-learning-based fault detection in electric vehicle powertrains using a digital twin. No. 2023-01-1214. SAE Technical Paper (2023)
10. Kumar, L., et al.: Electric Vehicle (EV) preventive diagnostic system: solution for thermal management of battery packs using AIOT. In: 13th Annual Computing and Communication Workshop and Conference (CCWC). IEEE (2023)
11. Yi, Y., et al.: Digital twin-long short-term memory (LSTM) neural network based real-time temperature prediction and degradation model analysis for lithium-ion battery. J. Energy Storage **64,** 107203 (2023)
12. Bhatti, G., Mohan, H., Singh, R.R.: Towards the future of smart electric vehicles: Digit. twin technology. Renew. Sustain. Energy Rev. **141**, 110801 (2021)
13. Delsing, J. (ed.). IoT Automation: Arrowhead framework. CRC Press (2017)
14. Desbiens, F.: Building enterprise IoT solutions with eclipse IoT technologies: an open-source approach to edge computing, pp. 67–101. Apress (2023)
15. MIMOSA: open standards for physical asset management. https://www.mimosa.org/

Analyzing the Enabling Factors to Implement MaaS in Asian, African and Latin American Cities

María Rosa Muñoz B.[1](✉), George Panagakos[2], Shritu Shrestha[1], Emilie Martin[1], Marc Hasselwander[4], Samuel Bonsu[3], Grace López Realpe[3], Fabio Bachetti[5], and Michael Bruhn Barfod[2]

[1] Wuppertal Institute for Climate, Environment and Energy GmbH, Döppersberg 19, 42103 Wuppertal, Germany
maria.munoz@wupperinst.org
[2] Technical University of Denmark, Building 358, Akademivej, 2800 Kongens Lyngby, Denmark
[3] Urban Electric Mobility Initiative, Kopenhagener Straße 47, 10437 Berlin, Germany
[4] Institute of Transport Research, German Aerospace Center (DLR), 12489 Berlin, Germany
[5] Pluservice s.r.l, S.S. Adriatica Sud 228/d, 60019 Senigallia, (AN), Italy

Abstract. The EU-funded SOLUTIONSplus project, which aims at kick starting urban e-mobility in developing countries, included the provision of a white label app customizable to the needs of the cities as part of its offer to its 7 demonstration cities in Asia, Africa and Latin America. Despite having the possibility of testing the customized app free of charge for the duration of the project, only 2 out of 7 cities, Quito (Ecuador) and Kigali (Rwanda), started and continued the process. Yet, only Quito was able to test the customized app in real operations. Thus, the paper analyzes the MaaS level and implementation barriers of the 7 cities and conducts an in-depth expert assessment of the technology, organization and environment enabling factors (TOE) to implement the MaaS concept in Quito, Kathmandu and Kigali. The results show that despite some progress towards an intelligent and integrated transport system in the analyzed cities, an important number of conditions that are a given in the Global North (e.g.: formality and integration of PT system), still need to be met in cities in developing countries before MaaS could be realized.

Keywords: Mobility as a Service (MaaS) · Enabling Factors · Developing Countries

1 Introduction

In an increasingly urbanized world, where more than half of the global population already lives in cities (2/3 by 2050), digital technologies will have a pivotal role in advancing sustainable urban development and formulating effective policies and measures for climate change mitigation. In parallel, the emergence of new mobility concepts and the

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 274–281, 2026.
https://doi.org/10.1007/978-3-032-06763-0_40

widespread use of smartphones worldwide has brought along the emergence of a significant number of mobility applications aimed at improving the efficiency of urban mobility services both for freight and passenger transport. Local and international mobility (and micromobility) apps such as Uber (Eats), Ola, Gojek, SafeBoda, Glovo, PedidosYA, tembici, Moveo, etc. already operate in Africa, Asia and Latin America, showing the great penetration and potential that app-based mobility services already have in the Global South. The integration of such services under the Mobility as a Service (MaaS) concept could contribute to modal shift and reduced energy consumption [2, 3].

MaaS is a new transport and mobility concept that integrates existing and new mobility services (NMS) into a single digital platform, providing customized mobility options and offering personalized trip planning and digital payment possibilities [1]. MaaS has the potential to improve the travelling experience, reduce travelers' costs, efficiently manage travel demand, and improve environmental and social outcomes [2–4]. MaaS has been gaining popularity in Europe / the Global North and has been implemented in several cities with positive results related to sustainable and low-carbon mobility [2, 3]. However, in developing countries, in contexts of fragmented and often informal public transport systems, lack of data, cash economies, low digitalization levels, the implementation of such a concept comes with a lot of challenges [5]. While important scholarly contributions have been made to understand the barriers and enablers of MaaS [5–8], these studies primarily rely on literature reviews and expert interviews. In this article, we contribute to this body of knowledge–which has mainly focused on the European context– by validating these findings with data from real-world MaaS schemes in the global South.

The EU-funded SOLUTIONSplus project, which aims at kick starting urban e-mobility in developing countries, included the provision of a white label app customizable to the needs of the cities to enable access to the e-mobility services for the end-users as part of its offer to its 7 demonstration cities in Asia, Africa and Latin America [9]. In a MaaS approach, the app offered the integration of all services, payment and information useful for the end-users allowing the possibility of integrating e-bikes, e-3-wheelers, e-taxis, etc. to the existing public transport (PT) system.

Thus, the purpose of this paper is to define and analyze the enabling conditions needed to adopt the MaaS concept in developing countries. The analysis starts by assessing the MaaS level and the implementation process in the 7 SOLUTIONSplus demonstration cities in Asia, Africa and Latin America (Sect. 2.1). The analysis of the enabling factors in Sect. 2.2 focuses on Kigali, Kathmandu and Quito, the 3 cities in which the feasibility of implementation of a MaaS app in the context of the project was further explored.

2 Analysis

2.1 MaaS Level and Pilot Implementation in SOLUTIONSplus Cities

In an attempt to create a comparison tool for the level of penetration of the MaaS concept in different cities, Sochor et al. (2018) propose a topology that consists of 5 levels based on the degree of integration [10], i.e.: 0 - no integration; 1 - integration of information; 2 - integration of booking and payment; 3 - integration of the service offer; 4 - integration of societal goals. As it can be seen in Table 1, the 7 SOLUTIONSplus cities fall between the Levels 0 (in transition to 1) and 2. This means that most cities already have access to some sort of route planning app and the more advanced ones are already in the process of integrating the in-app payment functionality.

Regarding the level of implementation under the SOLUTIONSplus project, it is shown that despite having the possibility of testing the customized app free of charge for the duration of the project, only 2 out of 7 cities, Quito (Ecuador) and Kigali (Rwanda), started and continued the process. Yet, only Quito was able to test the customized app in real operations. Nevertheless, it is worth noting that other demonstration cities such as Hanoi, Dar Es Salaam and Montevideo have taken their own path to develop mobility apps for specific needs.

Table 1. MaaS level and implementation in SOLUTIONSplus cities

City	MaaS level	Level of MaaS implementation under SOLUTIONSplus
Africa		
Kigali	L1	Discussions between the SOLUTIONSplus team and the representatives of the City of Kigali to develop a customized MaaS app took place between July 2021 and May 2023. However, a joint decision not to pursue it due to high constraints and limited time was taken.
Dar Es Salaam	L0 in transition to 1	Dar Es Salaam is developing its own planning and ticketing app for which the SOLUTIONSplus team had an advisory role. The conditions to implement a MaaS app between BRT buses and paratransit modes are not in place yet.
Asia		
Kathmandu	L0 in transition to 1	After a prefeasibility assessment conducted by DTU [11], it was determined that the conditions to implement a MaaS app were not yet in place in Kathmandu. In particular, related to the lack of an intelligent transport system, a business model and the regulations to incentivize transport operators to take part in a MaaS platform.
Hanoi	L2	It was decided to put all efforts in the development of a booking app for the e-2-wheelers being implemented in the pilot.
Pasig	L1	The governance of PT in Pasig would have required the involvement of a larger number of stakeholders, including the national government. Due to the complexity, after the assessment it was decided not to pursue it.
Latin America		

(*continued*)

Table 1. (*continued*)

City	MaaS level	Level of MaaS implementation under SOLUTIONSplus
Montevideo	L1 in transition to L2	The Municipality of Montevideo declined the offer arguing that they already have their own municipal app called "Como ir", which at the moment only allows trip planning, but is supposed to integrate in-app payment once the ticket validators in the buses are replaced
Quito	L1	A customized MaaS app including a trip planner, in-app payment and e-ticketing was developed in close collaboration with the municipal PTOs (BRTs and subway). In order to address the specific needs of the city, two complements were added: 1) a web app to top up the e-wallet with cash in the ticket booth and 2) A mobile app to validate the e-tickets until the automatic turnstiles are procured and installed in all stations. The 3 apps were piloted in Q4 2022 with 50 university students for a period of 4 weeks in 1 BRT station. The circumstances that led to the pilot implementation in Quito were: 1) the imminent launch of the subway line, by which the city is in the process of modernizing and integrating the PT system, which encompasses the Integrated Payment System (SIR), the Data Exploitation System (SAE) and the User Information System (SIU). The Municipality was reluctant to continue and scale-up the pilot due to the lack of knowledge of the new authorities about the MaaS concept, the regulatory framework for its implementation and the linkage to and benefits for the SIR. The main concerns were related to the business model and costs after the project end, as well as the ownership of the data and the application.

2.2 The Enabling Environment in Quito, Kathmandu and Kigali

A literature review was conducted to investigate the main requirements, but also the barriers that have been defined / identified both in the Global North and South for the implementation of the MaaS concept [5–7]. Based on that and following the methodology applied by Hasselwander et al. (2022), the technology, organization and environment

(TOE) framework was used to classify what from now on will be called enabling factors [5]. A total of 20 number of variables have been identified as enabling factors for MaaS implementation. The elements considered under Technology refer to the level of digitalization in the selected cities. In terms of Organization enablers, policies, plans and the governance structure related to digitalization of transport and intelligent transport systems were analyzed. Under Environment, the enabling factors are related to the integration level of the public transport system and other transport services.

The results of the analysis using the TOE framework are summarized in Figs. 1, 2 and 3. As it can be seen, only very few factors are fully met in the 3 cities. Nevertheless, in most factors the Quito, Kathmandu and Kigali comply partially, which means there is already some level of advancement.

City	Quito	Kathmandu	Kigali	In summary
Enabling factor	Expert Assessment			
Access to smartphones and mobile internet	PARTLY	PARTLY	PARTLY	Access to a smartphone doesn't seem to be a barrier, at least not in the age group expected to use the app, and not in urban areas. Mobile internet access has risen in these cities, but the affordability of data plans still represents a restriction for an important percentage of the population.
Availability of payment gateways (e.g.: PayPal)	PARTLY	PARTLY	YES	MTN Mobile Money is widely used for transport services in Kigali. In Kathmandu and Quito there some payment gateways, such us e-Sewa and DeUna, but their use for transport services is limited or non-existent.
Penetration of mobility apps (e.g.: Uber) 1) Availability and use 2) In-app payment (credit card usage, payment gateways, etc.)	YES	YES	YES	There are, however, several mobility apps, local and international, operating in the 3 cities. However, the possibility to pay in-app with credit / debit cards or payment gateways is still limited in most cases.
Availability of journey planner (e.g.: Google maps, Waze, municipal apps, etc.)	PARTLY	PARTLY	YES	Global apps such as Google maps work in the 3 cities. However, with different functionalities. While in Quito and Kigali it is possible to see the bus routes and the estimated times, in Kathmandu it only shows the walking route.
Users' acceptance and profile	PARTLY	PARTLY	PARTLY	There positive experiences of the widespread use of mobile money and e-wallets in Kigali and Kathmandu. It is expected, as in other contexts, that the primary target group is young people. However, depending on the user friendliness, affordability and benefits of using the app, a larger and more diverse number of users could be attracted.
Data collection and standards 1) Standardization and aggregation of data, open data or willingness to share data 2) Real-time data collected 3) GTFS files available	PARTLY	NO	PARTLY	The use of ITS by PTOs is still quite limited / fragmented. (Real-time) data collection, sharing and ownership is still an unresolved issue in the 3 cities.The only city from the 3 with a GTFS available was Quito.
Privacy and security of user data is not a limitation for users	YES	YES	PARTLY	Concerns regarding data privacy / security in the general population of these cities is low in comparison with European countries / cities. In the case of Kigali, however, the privacy and security of user data are quite relevant for the national authorities.

Fig. 1. Technology enabling factors in Quito, Kathmandu and Kigali

City	Quito	Kathmandu	Kigali	In summary
Enabling factor	Expert Assessment			
Transport digitalisation policy objectives	PARTLY	PARTLY	PARTLY	The 3 cities analyzed have some type of plan, framework or strategy either at the national or local level, where the use of ICT to improve public transport is mentioned. The MaaS concept, however, is only explicitly mentioned in Quito's SUMP, yet without further detail.
General transport authority / degree of centralization / governance	PARTLY	NO	NO	There's not a unique authority. Transport governance is pretty disperse. Quito is trying to create a general transport authority, but the process will take time.
Knowledge & capacities 1) Experience / Clarity regarding role and responsibilities / expected benefits 2) ICT expertise	PARTLY	NO	NO	There's lack of knowledge / clarity about the MaaS concept, its application, operationalisation and benefits. In the case of Quito, however, we counted with the support of a public official that had clarity about the concept and benefits. It was one individual, at the institutional level it's still an unknown concept.
Transport authorities structures, fast decision processes, innovative strategies and integrated planning approaches	NO	NO	NO	The complexity of transport governance, including the relationship between public and private actors, and the low speed of processes, makes the introduction of MaaS difficult.
Political will and institutionalization 1) Top management support 2) Public sector engagement / Political support 3) Necessity of a window of (political) opportunity / political cycle	PARTLY	YES	YES	In general terms, these municipalities were engaged. The processes in these 3 cities were able to move forward only because of the top management support provided. However, the system and the governance is not mature enough for its full implementation. In the case of Quito, the political cycle played a huge role, as the new authorities were not as supportive of the project as the previous ones.
Financial resources and funding	PARTLY	NO	NO	The municipalities didn't have their own resources because MaaS is not yet in their priorities / radar. In Quito, there are resources as the Integrated Payment System (SIR) is being implemented as part of the PT modernization.
Clarity about the business model / ownership of app / data collection	NO	NO	NO	Business model and data ownership are among the main concerns when deciding to move forward. There was very little knowledge in the counterparts about how to implement the MaaS concept in the given environment.
Enabling regulatory framework	PARTLY	PARTLY	PARTLY	At the moment, the regulations in the analyzed countries are either at a very early stage or very restrictive. In particular, we encountered some regulations that make the collaboration with international companies for the development of MaaS or other mobility applications difficult. In Rwanda, for instance, cloud storage can't happen outside of the country. In Ecuador, on the other hand, the fare collection of PT has to be in the corresponding bank account of the PT authority within 24hours.

Fig. 2. Organizational enabling factors in Quito, Kathmandu and Kigali

City	Quito	Kathmandu	Kigali	In summary
Enabling factor	Expert Assessment			
Integrated and regulated operator landscape 1) Form of Route Service contracts between Authorities and transport operators 2) Formal, multimodal and integrated PT system (physical and tariff) 3) Low informality rate in transport services	PARTLY	NO	PARTLY	In the 3 cities, there is some form of route service contracts that formalises and regulate transport. In the recent years, the 3 cities have tried to improve the processes. However, the PTO landscape is still very fragmented in Kathmandu, and partially fragmented in Quito. Paratransit / informal transport is only relevant in Kathmandu. Thus, the level of integration of the system varies, but is still not ideal in any of the 3 cities.
Private sector engagement / Support from transport groups and associations	PARTLY	PARTLY	YES	The development has not reach the level of opposition. However, the level of integration that the proper functioning of the MaaS concept requires will need important negotiations with all PTOs. Thus, in the case of Quito and Kathmandu, where PT integration is not completed, it can lead to some level of opposition.
Integrated payment system for PT	PARTLY	NO	PARTLY	Kigali has a smart card that can be used in all PT operators. Quito is currently working on the introduction of an Integrated Payment System, which will include the proper infrastructure, the IT systems and fund reconcialiations. In Kathmandu, there have been some failed attempts. In the 2 latter, cash is still the mean of payment in PT, while in Kigali mobile money is used.
Existance and integration with other transport services (bike sharing, e-scooter sharing, etc.)	PARTLY	NO	PARTLY	In the 3 cities there are transport services, be it bike sharing, e-scooters or moto-taxi services. Some of them linked to mobility / booking apps. However, there's no proper integration between other services and PT. In Kigali and Quito there are some located close to PT stations, but there's not a proper multimodality strategy.
Successful adoptions (in the country/region) serving as blue print	PARTLY	PARTLY	PARTLY	There are some mobility apps in the 3 region trying to start a MaaS service with many limitations. However, so far, in most cities they only have trip planning functionalities.

Fig. 3. Environment enabling factors in Quito, Kathmandu and Kigali

3 Conclusions

As it has been shown throughout the paper, the implementation of the MaaS concept require the fulfillment of a set of technology, organization and environment conditions. The MaaS level analysis shows that despite the fact that some of the analyzed cities are still in level 0, they are already transitioning to level 1. Most cities are in level 1, some transitioning to level 2. Thus, the cities analyzed, despite their differences, reveal a slow, but steady progress towards the adoption of the features of an intelligent and integrated transport system that will enable MaaS. There are, however, still a series of barriers that need to be overcome related to the digitalization level, as well as in the transport system and governance, before the MaaS concept implementation is feasible in developing countries. A step-by-step approach could be desirable, starting by the gradual integration of all PTOs into the system in one digital platform and then the addition of other mobility service providers.

References

1. Jittrapirom, P., Caiati, V., Feneri, A.M., Ebrahimigharehbaghi, S., Alonso-González, M.J., Narayan, J.: Mobility as a service: a critical review of definitions, assessments of schemes, and key challenges. Urban Plann. **2**(2), 13–25 (2017)
2. Hensher, D.A., Ho, C.Q., Reck, D.J.: Mobility as a service and private car use: Evidence from the Sydney MaaS trial. Transport. Res. Part A: Policy Pract. **145**, 17–33 (2021)
3. Sochor, J., Strömberg, H., Karlsson, I.M.: Implementing mobility as a service: challenges in integrating user, commercial, and societal perspectives. Transp. Res. Rec. **2536**(1), 1–9 (2015)
4. Durand, A., Harms, L., Hoogendoorn-Lanser, S., Zijlstra, T.: Mobility-as-a-Service and changes in travel preferences and travel behaviour: a literature review (2018)
5. Hasselwander, M., Bigotte, J.F.: Mobility as a Service (MaaS) in the global south: research findings, gaps, and directions. Eur. Transp. Res. Rev. **15**, 1–16 (2023)
6. Butler, L., Yigitcanlar, T., Paz, A.: Barriers and risks of Mobility-as-a-Service (MaaS) adoption in cities: a systematic review of the literature. Cities **109**, 103036 (2021). https://doi.org/10.1016/j.cities.2020.103036
7. Karlsson, I.C.M., et al.: Development and implementation of Mobility-as-a-Service – A qualitative study of barriers and enabling factors. Transport. Res. Part A: Policy Pract. **131**, 283–295 (2020). https://doi.org/10.1016/j.tra.2019.09.028
8. Matyas, M., Kamargianni, M.: Investigating heterogeneity in preferences for Mobility-as-a-Service plans through a latent class choice model. Travel Behav. Soc. **23**, 143–156 (2021). https://doi.org/10.1016/j.tbs.2020.12.002
9. Panagakos, G., et al.: E-mobility solutions for urban transportation: user needs across four continents. Transportation Research Procedia, in press (2023)
10. Sochor, J., Arby, H., Karlsson, I.C.M., Sarasini, S.: A topological approach to mobility as a service: a proposed tool for understanding requirements and effects, and for aiding the integration of societal goals. Res. Transp. Bus. Manag. **27**, 3–14 (2018). https://doi.org/10.1016/j.rtbm.2018.12.003
11. Ortving, A., Brodthagen, M.: An analysis of Kathmandu's public transport system and viable improvements through MAAS IT solutions; a pre-feasibility study. Technical university of Denmark, DTU Management (2022)

Designing Mobility Systems-of-Systems

Pontus Svenson[1(✉)], Jakob Axelsson[1], Charlotta Glantzberg[2], and Robert Nilsson[3]

[1] RISE, George Town, Cayman Islands
{svenson.pontus,axelsson.jakob}@ri.se
[2] AFRY, Solna, Sweden
charlotta.glantzberg@afry.com
[3] Volvo Cars, Gothenburg, Sweden
robert.nilsson2@volvocars.com

Abstract. In this paper we describe the results of a study on the design of open and co-opetitive systems of systems for mobility. A system of systems (SoS) is a set of independent systems (CS – Constituent Systems) that interact to create capabilities that none of the constituent systems can accomplish on their own. A CS can be simultaneously part of several SoS. The independence of the constituent systems is an important and central ingredient in an SoS, and is often divided into managerial and operational independence.

We argue that the mobility SoS of the future can benefit from being open and co-opetitive (i.e., simultaneously collaborating and competing) to enable combined mobility (MaaS) for real. If competition between CS is not allowed, new entrants will choose to compete with the SoS instead of becoming part of it. If a new entrant develops much better vehicles or better ways to find transport solutions for users, that entrant must be able to become part of the SoS and take market share from other actors.

Keywords: Mobility · transport · system of systems · architecture · sustainability · MaaS

1 Introduction

Mobility as a service (MaaS) has emerged as a core component of smart city development. There are a number of definitions circulating around the smart city of the future, but basically it is about having a holistic perspective in the construction of a city that is economically, socially and ecologically sustainable with the help of technology.

An important part of the smart city will be smart mobility and transport solutions. In recent years, a vast array of new mobility services have been introduced. For instance, cities have been flooded with electric scooters and a large number of different car-sharing services have been introduced. Despite this, there are no city-wide services that allow residents to easily book a trip from A to B with a combination of several services. Rather, what we see on our streets is

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 282–287, 2026.
https://doi.org/10.1007/978-3-032-06763-0_41

a hodgepodge of services that often do not survive for very long, while the interconnected system of transport services that would constitute real system innovation is missing. One reason for this is the lack of language with which to formulate a systems of system mobility solution. Although there are many initiatives around smart cities worldwide that deal with transportation, and many research activities in the area, they are usually carried out as isolated projects.

MaaS is also a popular research topic that involves the integration of various mobility services [1,2]. There is a need to take account of different regulatory frameworks, infrastructures, and user preferences when implementing such systems [3]. There are many different actors involved in MaaS, such as vehicle operators, brokers that connect users to vehicle operators, users, and municipal and regional authorities.

Current approaches to MaaS are based in transport systems engineering and mostly consider them as integrated systems, i.e., they envision a system that is tightly integrated and controlled by a central orchestrator. This means that the technical architecture for a MaaS solution must integrate the different transport services tightly into a single digital platform [4]. There are obstacles in the form of data sharing and ownership issues as well as how to avoid silo solutions that lock-in the user to specific companies. From the perspective of companies involved in MaaS solutions, there is an understandable reluctance to cooperate too closely with competitors. But in order for the MaaS to be useful and gain enough market share to make a difference, there must be such cooperation.

Prompted by these possibilities to combine SoS and MaaS, we undertook a project (Maus) that investigated a *systems of systems* (SoS) [5] approach to mobility as a service by considering a scenario in a semi-rural area near a major Swedish metropolitan. In this paper, we describe some of the results of this project. Main results include a design method for SoS [6], a business model evaluation method, SoS architecture patterns [6–9], and processes and methods for governance and management of the SoS, as well as a survey of existing SoS design patterns [10].

The main scientific contribution of this paper is to take a first step towards applying SoS concepts to MaaS.

This paper is outlined as follows. Section 2 introduces key SoS concepts, both standard ones and results of the Maus project. Section 3 illustrates the mobility SoS. We conclude with a summary and discussion in Sect. 4.

2 Mobility Systems of Systems

An SoS is a set of independent systems (CS – Constituent Systems) that interact to create capabilities that none of the constituent systems can achieve on their own. A CS can be part of several SoS at the same time. The independence of the constituent systems is a central ingredient in an SoS. An important question when studying and designing SoS is what entities should be considered as the CS. For a mobility SoS, this boils down to the deciding if the CS should be the individual vehicles that carry out the transports or the transport companies that own the vehicles. For mobility SoS, it is usually the vehicles that should be considered CS, but this can vary depending on which aspects of the SoS that are studied. In addition to the vehicles, a mobility SoS will also need to have support services of various kinds, such as travel planners and payment intermediaries. Such support services are CS of a special kind called mediators. Figure 1 shows this general structure of an SoS. A mediator can be seen as a special case of CS and is something that provides services necessary to organize the collaboration, but does not have the same degree of independence as a CS. In a mobility SoS, someone who manages an intangible service, for example for information sharing, payment transfer or developing proposals for constellations whose combined abilities solve the transport assignments, is a journey planner. An example of a mediator could be to ensure that the trip can be booked and payment can be made.

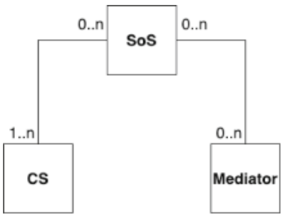

Fig. 1. An SoS contains one or more CS and optionally some mediators.

In earlier work, different types of SoS have been defined [5,11], for example depending on how much central control there is in the SoS. In a Directed or Acknowledged SoS, each CS is subordinated to a central control, while in Collaborative SoS the CS retain their control and choose when to collaborate. Here we will focus on Collaborative SoS that are also open and co-opetitive. An open SoS is an SoS which new actors can join. A co-opetitive SoS is an SoS whose constituents simultaneously both collaborate and compete.

3 Applying the Mobility SoS to a Semi-Rural Example

What might an SoS solution for mobility look like? In the Maus movie (available at sos-4-mobility.se or https://www.youtube.com/watch?v=e-DdHS5o6yk) we show this for parts of the scenario described here. The Maus project used the area around Harestad, north of Gothenburg, as a concrete case to work with. At the start of the project, a list of specific user stories and use cases was developed.

How would the people solve their transport needs today? Many of them would routinely choose their own car, because "all other ways are so hard". Some would check out the possibilities of using public transport. Someone would call different taxi companies to get prices. Some would ride bikes. What these solutions have in common is that the people focus on the means of transport instead of the

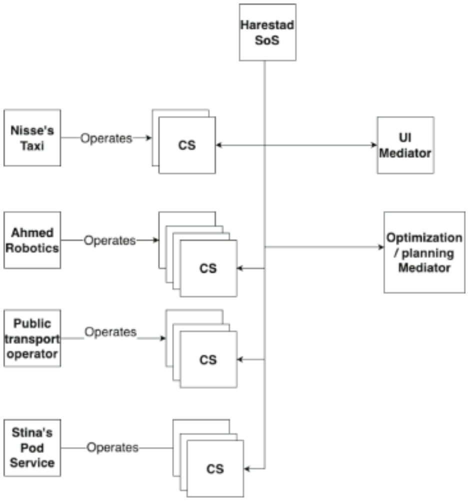

Fig. 2. Schematic view of the mobility SoS that could solve the travel needs in Harestad. The SoS has access to a large number of CS from four different companies, and also needs mediators for interfacing to the users, travel planning/optimization, and payment distribution.

transport itself. If you want to make an optimal choice of transport, you need to contact several different transport companies and compare their offers. Many people will then choose to travel with their own car because it is the easiest. The purpose of the Maus project was to show how you can choose transport instead of choosing a means of transport. In the next section, we show what such a solution could look like.

What is required for an SoS solution to be realized? There is a need for a robust payment service that enables the user to pay only once, and then distributes the money to all involved parties in a fair manner. This is one of the necessary mediators in the mobility SoS. Agreements are needed that regulate service levels and services that collect data on the quality of travel and goal achievement [9]. It is also important that all actors who are part of the SoS trust each other and that agreements are in place that regulate, for example, compensation levels and the consequences of delays and other problems.

Solutions are also needed to enable the users of the SoS to formulate their travel needs and for the SoS to match these with possible solutions. Here, it is important to specify what are acceptable waiting times and investigate the balance between cost and goal achievement. The major optimization problem posed by matching can not be solved exactly, which makes it important to study which trade-offs are acceptable. These services are best divided into two different mediators. One is the travel planner service. This mediator takes requests for transports and matches them to the available transport capabilities of the SoS. We have earlier [8] described how this could work. The other needed mediator is a user interface mediator. This service provides the interface between the users and the SoS.

Figure 2 displays the SoS structure that could be implemented. While we here include only one of each type of mediator, it is important to note that a SoS should be open for innovation and competition also for its mediators. If a new company develops a better travel planning service, it should be free to join the SoS and provide its service as a mediator that competes with the existing mediators.

4 Conclusion

In conclusion, we discussed some of the results of the Maus project and how SoS concepts could be applied to mobility and MaaS applications. We introduced some standard SoS concepts as well as new concepts developed within Maus.

The main contributions of this paper are to start the discussion on how to use SoS concepts to overcome some of the barriers currently preventing MaaS solutions and to present a first example of how a MaaS based on SoS concepts could be designed. The focus on the different perspectives that different roles in an SoS entail can contribute to better understanding of the roles and responsibilities within MaaS. By abandoning the idea of MaaS as a single system and embracing the co-opetitive SoS concept, it becomes easier to design solutions that both satisfy the needs of users and municipalities while still allowing the competition between actors that drives innovation.

There are ample opportunities for further research in this area. There is a need to further explore the usefulness of SoS concepts in transport and mobility research and implementing the ideas in demonstrations and pilots. Before this, however, it would be useful to perform large-scale simulations based on realistic travel patterns, to be able to quantify both the costs and the benefits of SoS mobility solutions.

Acknowledgements. This research was funded by Vinnova, the Swedish Innovation Agency, under grant no. 2019-05100. The research results d not necessarily agree entirely with the researchers employers views and standpoints in all details.

References

1. Liimatainen, H., Mladenović, M.N.: Developing mobility as a service - user, operator and governance perspectives. Eur. Transp. Res. Rev. **13**, 37 (2021)
2. Musolino, G., Rindone, C., Vitetta, A.: Models for supporting mobility as a service (maas) design. Smart Cities **5**(1), 206–222 (2022)
3. Kriswardhana, W., Esztergár-Kiss, D.: A systematic literature review of mobility as a service: examining the socio-technical factors in maas adoption and bundling packages. Travel Behav. Soc. **31**, 232–243 (2023)

4. Reyes García, J.R., Lenz, G, Haveman, S.P., Bonnema, G.M.: State of the art of mobility as a service (maas) ecosystems and architectures—an overview of, and a definition, ecosystem and system architecture for electric mobility as a service (emaas). World Electric Vehicle J. **11**(1) (2020)
5. Maier, M.W.: Architecting principles for systems-of-systems. In: INCOSE International Symposium 6(1), pp. 565–573 (1996)
6. Svenson, P., Reichenberg, F., Axelsson, J.: A design method for collaborative systems of systems applied to metropolitan multi-mode transport system. In: 2021 16th International Conference of System of Systems Engineering (SoSE), pp. 13–18 (2021)
7. Axelsson, J., Svenson, P.: On the concepts of capability and constituent system independence in systems-of-systems. In: 2022 17th Annual System of Systems Engineering Conference (SOSE), pp. 247–252. IEEE Press (2022)
8. Svenson, P., Axelsson, J.: Should i stay or should i go? how constituent systems decide to join or leave constellations in collaborative sos. In: 2021 16th International Conference of System of Systems Engineering (SoSE), pp. 179–184. IEEE Press (2021)
9. Svenson, P., Olsson, T., Axelsson, J.: Constituent systems quality requirements engineering in co-opetitive systems of systems. In: 2022 17th Annual System of Systems Engineering Conference (SOSE), pp. 347–352 (2022)
10. Axelsson, J.: Systems-of-systems design patterns: a systematic literature review and synthesis. In: 2022 17th Annual System of Systems Engineering Conference (SOSE), pp. 171–176. IEEE Press (2022)
11. Dahmann, J.S., Baldwin, K.J.: Understanding the current state of us defense systems of systems and the implications for systems engineering. In: 2008 2nd Annual IEEE Systems Conference, pp. 1–7 (2008)

Transport Data Sharing

Towards Integrated Traffic Management for All of Austria: Realtime Traffic Information and Multimodal Journey Planning Beyond Administrative Borders

Tobias Schleser(✉) and Martin Nemec

Asfinag Maut Service GmbH, Alpenstraße 99, 5020 Salzburg, Austria
tobias.schleser@asfinag.at

Abstract. All major players in road operation and traffic management of Austria joined forces in the EVIS.AT platform and developed a harmonized real time traffic information network. It provides a countrywide, comprehensive, and authorized real-time data basis for traffic information, traffic management and traffic analysis. The platform delivers traffic messages for planned and unplanned events as well as for traffic regulations and access restrictions and traffic flow data such as current and predicted speeds and LOS. This is done for all major street levels (highways to rural roads) for all of Austria and is based on sensors, probe data and journalistic data of the authorities. The involved parties committed to long-term operation and provision of real-time data through a public-public cooperation. Thus, the platform EVIS.AT serves as strong basis for road-related traffic data in the context of integrated traffic management as well as comprehensive traffic information. The authorized real-time road data of EVIS.AT is combined with public transport data of all operators, including timetable, real-time and message data in the intermodal routing platform VAO.AT. The platform delivers high-quality intermodal routing and traffic information services, based on the comprehensive data sets of Austria's various public authorities. It is used by partners and b2b customers to create powerful end user applications as well as various tools for administration.

The paper reports on objectives and status for these platforms, what makes them unique in the European data and service landscape and how these data and services are being used in national and international applications.

Keywords: Multi modal journey planner · real time traffic information · travel times · standardization · vehicle probe data · incident management

1 Introduction

The European Commission passed the Delegated Regulation (EU) 2017/1926 on 31 May 2017 with regard to the provision of EU-wide multimodal travel information services (the revision will be published soon). Even long before this event Austria took the approach to develop a platform called VAO („Traffic information Austria"): a nationwide intermodal traffic information platform that integrates approved traffic information of all modes

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 291–298, 2026.
https://doi.org/10.1007/978-3-032-06763-0_42

of transport provided by the major traffic infrastructure and traffic service providers. The system includes public transport timetables, stations and other connecting points, a detailed intermodal graph called GIP [1], traffic messages and Level-of-Service data for all major motorways, all addresses in Austria, park + ride and other parking facilities as well as additional data such as points of interest, bike rental stations and short term parking zones. VAO's various services rely on data covering all of Austria, making it a comprehensive toolset for traffic information and end user services. Intermodal routing services are available through a well-defined interface. [2] Currently, VAO is including shared mobility services to its routing and maps database.

EVIS.AT (Real time traffic information on Roads in Austria) on the other hand, aims at providing real time traffic information for roads and besides ASFINAG as Austria's motorway operator, brought aboard the 9 Austrian federal states, ITS organizations, major cities as well as the Ministry of Internal Affairs (police) and the automotive and mobility club ÖAMTC. They joined forces to improve the real time traffic data landscape of Austria significantly and to harmonize data and processes nationally. The project has been an ITS emphasis in the past few years. It ended in 2022 and started a permanent operation of data provision and quality processes to ensure long term support. [3] Subsequent sections of this paper report on objectives, technical and organizational construction as well as provided data and services and use of that data.

2 Objectives of the EVIS.AT Platform

EVIS.AT is a nationwide, harmonized and unified platform for all kinds of traffic information for motorized traffic, built by the experts who already developed and operate their existing solutions. The mission of the platform is to harmonize and unify traffic data for motorized traffic, with the goal of providing a better real time traffic data basis for VAO through creation of a data platform as well as processes for all major infrastructure and traffic information [5, 6]. EVIS.AT became the main content platform for dynamic aggregated road traffic information in Austria.

The main objectives of the EVIS.AT platform are:

- As public actors, **create and collect own real-time traffic data** as strong basis for administrative tasks and end user services but also when engaging with the large private mobility service providers (i.e. routing engines, navigation companies,…)
- **Harmonize all data of all public actors** (technically, structural, content-wise) to create a meaningful, national data set. Agree upon standards for implementation formats as well as delivery processes.
- **Collect and provide all data at a single point of contact** (EVIS.AT data platform) to make the data accessible; utilize this for data dissemination within the partners but also towards third parties, i.e. media, routing and navigation providers.

3 The EVIS.AT Platform

3.1 Development

The EVIS.AT platform (Echtzeit Verkehrsinformation Straße; real-time road traffic information) was established in an implementation project in three phases (planning, technical implementation, operation) and focused content-wise on the two key areas: incident management and real time traffic information (i.e. LOS on the roads).

Launching a project with broad participation of road and traffic authorities was based on the following considerations: *Use local expertise*. (collection of traffic information with respect to local road layouts and topographies but also knowledge of the specific workflows in provinces and federal states). *Therefore, get everybody together*. (EVIS is led by Austria's motorway operator ASFINAG and supervised by BMK. Moreover, most of the administrative stakeholders for traffic in Austria are project partners). *Develop common standards*. (When using local expertise and even varying methods and means to collect traffic information it is a must to agree upon common standards and the quality of service along with data types and formats. This is meant to ensure that a national data aggregation and logical decentralized unit can produce harmonized traffic information and foster the exchange of data to efficiently use them for traffic information and network management. Basis for this are European standards like DATEX2.) [4].

On the technical level, the main approach in developing the platform was to utilize existent well working solutions of the partners (be it systems, data formats, processes,…), select the fittest and deploy them nationally.

The project targeted the whole lifecycle of traffic information: from fundamentals like extending the stationary traffic sensor network throughout Austria, extending the traffic network graph GIP.GV.AT and tendering floating car data and mapping them on the graph. Concerning standards, the partners agreed upon data formats, definitions, message types and minimum standards for message creation. The definition of processes focused on clear responsibilities, local know-how and individual implementation. The partners implemented several tools for common usage (see next section) and in the area of governance, they set up a public-public partnership and contracts.

3.2 Technical Implementation

The EVIS.AT platform is structured as a distributed system. On the one hand, there are several local data aggregation and provision services. Following political and administrative conditions, partners from different areas in Austria are responsible for data generation and provision for their respective region. However, they all agreed upon common formats, standards and quality. To bring the data together in a harmonized way, these so called "central services" provide services to all users but are only implemented once. Therefore, each member can use all services implemented by itself or other parties. The following table gives an overview of the services (Table 1).

Table 1. Central services of the EVIS.AT platform

Central service	Tasks
Data hub	The **Data Hub** (data collection and distribution service) receives and distributes data between partners and 3rd parties like Verkehrsauskunft Österreich VAO and the free market
Clearing service	The **Clearing Service** is responsible for the collection of incident data and for the conflict free distribution in terms of location and content
Geo-Services	**Geo-Services** translate different versions of the above mentioned GIP format into the newest. They also translate the GIP system into other location referencing systems like WGS84.
FCD Services	The **Central FCD Service** collects and tenders nationwide probe data, processes them from raw data to speeds and travel times and makes the data available to partners
Traffic Message Acquisition Tool	The **Traffic Message Acquisition Tool** provides a tailored web front end for the input of traffic messages for authorities and supports the GIP as well as DATEX2 output and is integrated in a national identity provision and authorization framework and administration processes

3.3 Legal and Organizational Implementation

Legally, the EVIS.AT platform is constructed as a public-public cooperation of public sector stakeholders based on respective agreements. This allows for a cooperation agreement, where each partner provides services, data and licenses to the consortium while avoiding to found a corporation or directly providing cash.

The main parts of the cooperation agreement include:

- Governance and steering of the operating platform;
- Central Services and responsibilities of the respective responsible party and operator within the distributed system EVIS.AT, for details see Sect. 3.2;
- Provision of real-time traffic data of the respective responsible authority;
- Consummation and rights of use of real-time traffic data of other parties;
- Technical sales and distribution of data services to third parties;
- Public relations and external communication of the platform EVIS.AT.

Goods and services that can not be provided from within the consortium are tendered on the market – where one or several of the partners tender the respective services for the whole platform. As of summer 2023, there have been two major tenders:

- Clearing services: the objective of the central service "Clearing" is to receive, process, quality check and provide all traffic messages of more than 60 regional message providers (local public authorities); most of these checks are automated but manual checks are performed as well.

- Floating car data: dynamic procurement system established by Salzburg Research and the federal states of Salzburg, Tirol and Vorarlberg where tenderers bid to demonstrate their qualification and eligibility based on legal and commercial terms as well as data specific KPIs [7] to enter the procurement pool and deliveries are requested based on available funds from the pool [8].

The technical but also legal and organizational implementation of the platform has been operating for over a year now and is working well.

4 EVIS.AT Data and Service Offering

Through the EVIS.AT platform, Austrian road authorities can provide their respective real-time traffic data to third parties at a one-stop-shop, harmonized and comprehensive for all of Austria: motorways but also the lower-level network and several cities.

ASFINAG is the distributor of all EVIS.AT real-time traffic information data, i.e. not only for motorways but also for the lower-level network. This is based upon respective agreements with all 9 federal states of Austria, several cities, the BMI (Ministry of the Interior/Police), the BMK (Ministry of climate and traffic) and ÖAMTC (Automotive Club) and several cities via the operating agreement.

The data streams included are listed in Table 2.

Table 2. Data streams of EVIS.AT

Data stream	Data description	Data format
Traffic messages of planned events	i.e. constructions, closures, events, winter equipment,…	DATEX2
Traffic messages of unplanned events	congestion, accidents, bad weather,…	DATEX2
Seasonal conditional access restrictions	closed off-ramps of highways (for through-traffic in Tirol and Salzburg)	DATEX2
Traffic State	level of service, current speeds	JSON
Traffic State Prognosis	level of service, speed prognosis	JSON
Yearly average speeds	Yearly average speeds	JSON

All data is accessible via a secure https REST interface. Authentication and authorization are implemented via OpenID Connect.

To access the data, interested organizations need to sign a license agreement with the distributor ASFINAG. A license fee applies that reflects on the cost of providing the real-time data interfaces, clearing services and SLA. The data itself is free of charge.

All data is listed on Austria's National Access Point mobilitaetsdaten.gv.at as requested by the ITS directive [9].

5 Application of EVIS.AT Traffic Data

EVIS.AT traffic data is being used in a variety of end-user applications but also for internal tools in traffic management.

Verkehrsauskunft Österreich (VAO) is providing white-label multimodal end-user applications as described in the introduction and implemented all EVIS.AT data as reference service. This means, that through the EVIS.AT and VAO eco system, public authorities are now able to push traffic messages and measures in about 30 apps and websites within minutes.

Internal tools of the various partners that implement EVIS.AT data and utilize it to optimize operation and traffic management include:

- **GIS and maps tools**: they are used for strategic planning but also for traffic management to get an overview of adjacent networks and traffic situations; see Fig. 1 for an example of ASFINAG's Maps application.
- **Winter service**: EVIS.AT data is used in winter service applications. In return, they feedback traffic information from winter services to EVIS directly, i.e. snow closures, avalanche warnings and the like.
- **Abnormal load management**: abnormal loads (width, height, weight) need to be admitted by road authorities who take traffic, road and weather conditions into consideration. These authorities now also use EVIS.AT data to better evaluate and assess the suitability of certain routes (i.e. constructions)
- **Operation planning and routing for emergency vehicles**: certain traffic conditions or regulations (i.e. closures) are taken into account for route planning

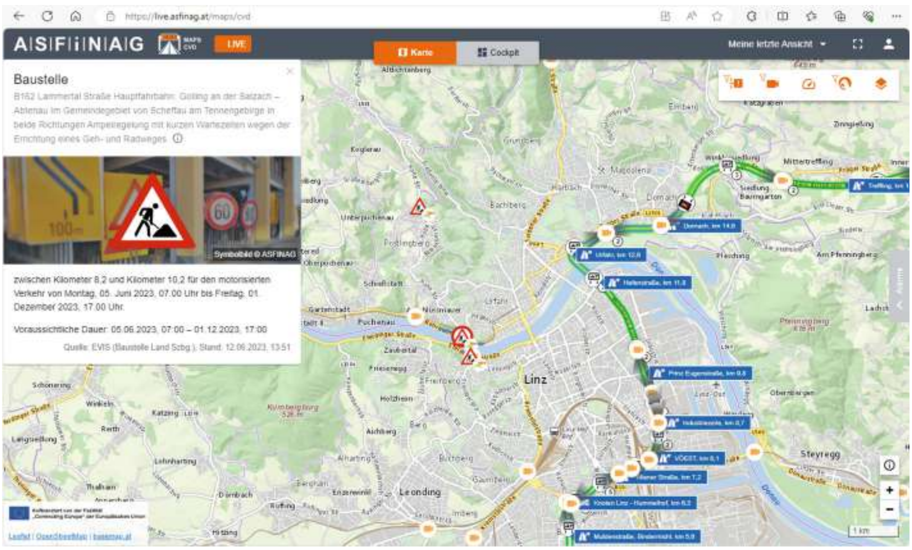

Fig. 1. Live.asfinag.at – ASFINAG's internal maps application for strategic traffic management and operations. EVIS.AT data of partners is implemented so that operators have an overview of the traffic information on adjacent networks. Reproduced with permission from ASFINAG, copyright ASFINAG, 2024.

Currently, there are also major efforts in place, to move towards Integrated Traffic Management in Austria to allow for policy making. Integrated Traffic Management is more than just integrating different data sources into a data hub. Main requirements are precise strategies and traffic management goals of politics and authorities as well as the legal delegation of traffic management duties to a capable platform that oversees the different networks and modes of transport. Additionally, the inclusion of several modes of transport, i.e. public transport and new mobility services will be included. In fall 2023, several projects on this topic were started by the major infrastructure and traffic information providers. The projects builds on EVIS.AT, VAO and other Austrian ITS platforms and datapools. Their main objective is to lay out a set up plan to introducing integrated traffic management in Austria.

The platform EVIS.AT will take on several roles in an "Integrated Traffic ´Management Austria": it will provide harmonized base data for KPIs and road traffic overview to be gathered by traffic operation centers; the consortium will serve as a platform for policymaking and strategic decisions of the road operators; EVIS.AT will also serve as data distribution hub to disseminate the policies and traffic regulations of operative ad hoc traffic management to various end user services.

6 Conclusion

We showed that EVIS.AT established a platform and legal framework that includes the motorway operator, federal states, cities, automobile club and police, addressing much-needed harmonization issues with respect to real-time traffic data for a whole country.

The novelty of the approach lies in the strong and legally binding cooperation of all major stakeholders beyond administrative borders and harmonization for a whole country. While most European countries have "national access points" in place, that allow data owners and administrations to list or publish their data, these data sets are hardly harmonized and come in an abundance of data formats, dialects, extensions, semantics, update characteristics and bundling. The data sets are also not complete for larger regions and therefor very hard to use in wider services.

The EVIS.AT partners agreed upon data formats, definitions, message types and minimum standards for message creation. The definition of processes focused on clear responsibilities, local know-how and individual implementation. The cooperation agreements ensure complete coverage for all of Austria.

EVIS.AT has become an important and established part auf Austria's ITS landscape, enabling administrations to publish data and generate end user and internal applications as well as third parties to use the real time traffic data in their services. The next step will be to utilize the platform for policymaking in the context of integrated traffic management – as data basis for road traffic as well as dissemination platform.

References

1. GIP, Project web site. http://gip.gv.at. Accessed 21 Sept 2023

2. Hintenaus, M.: VAO – Traffic Information Austria. PIARC Case Study (2018). https://www.piarc.org/
3. Schleser, N.: EVIS.at – Realtime Traffic Information Austria. PIARC Case Study (2018). https://www.piarc.org/
4. Schleser, T., Nemec, M.: Towards integrated traffic management for all of Austria: real-time traffic information and multimodal journey planning beyond administrative borders. In: Proceedings of the 24th ITS European Congress, Toulouse, 30 Mai–1 June 2022 (2022)
5. EVIS consortium. Project application at Climate and Energy fund: EVIS AT: Echtzeit VerkehrsInformation Straße – Österreich (project application number KR14IM4S12342) (2015)
6. EVIS consortium. Rolloutplan for project: EVIS AT: Echtzeit VerkehrsInformation Straße – Österreich. (project application number KR14IM4S12342) (2017)
7. Rehrl, K., Henneberger, S., Leitinger, L., Wagner, A., Wimmer, m.: Towards a National Floating Car Data Platform for Austria. Proceedings of the 25th ITS World Congress, Copenhagen. 17–21 September 2018
8. Salzburg Research Forschungsgesellschaft m.b.H.: EVIS.AT - Floating Car Data. In: ANKÖ Bekanntmachung öffentlicher Ausschreibungen. https://gv.vergabeportal.at/Detail/132118. Accessed 25 Sept 2023
9. EVIS.AT consortium: listing of EVIS.AT data streams on Austria's national access point. https://mobilitaetsdaten.gv.at/daten?search=EVIS.AT. Accessed 25 Sept 2023

Enhancing Accessibility and Interoperability of Mobility Data: MobilityDCAT-AP, a Metadata Specification for Mobility Data Portals

Peter Lubrich[1](✉), Marco Comerio[2], Petr Bureš[3], and Eva Thelisson[4]

[1] Federal Highway Research Institute (BASt), Brüderstr. 54, 51427 Bergisch Gladbach, Germany
lubrich@bast.de
[2] Cefriel, Milan, Italy
[3] Czech Technical University in Prague, Prague, Czech Republic
[4] Swiss Federal Roads Office (FEDRO), Bern, Switzerland

Abstract. This paper introduces a formal metadata specification for mobility data portals as an extension of DCAT-AP, called mobilityDCAT-AP. It addresses a scenario in which mobility data is offered on a data portal, and is intended to be found, assessed and reused by data users. Unlike in other domains, a structured and community-based metadata for the wider mobility domain has not been established yet. With such specification, an agreed usage of metadata among different portals; easier access to mobility data; improved interoperability in the mobility data eco-system; and the leveraging of semantic technologies are envisioned. In addition, the Resource Description Framework (RDF) as a de-facto standard for metadata, is applied to model the metadata vocabulary. The paper elaborates on the overall goals, previous works on metadata specification and harmonisation, the working process, concrete deliverables and future prospects of mobilityDCAT-AP.

Keywords: Mobility Data · Data Portals · Metadata · National Access Points · Data Harmonisation · Data Interoperability

1 Introduction

1.1 Context and Motivation

How to make different data assets from different actors in the mobility system discoverable and accessible? This is the role of internet portals for mobility data which have been developed all over the world in recent years. These portals often have specific spatial or thematic coverage, e.g., providing public-transport timetable data for a specific region. In addition, legal obligations, such as those established through the European Union's National Access Points (NAPs) [1], mandate the creation and population of such portals.

This work aims at supporting potential data users during the discovery and assessment of data offerings in NAPs and other mobility-related data portals. Metadata can

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 299–304, 2026.
https://doi.org/10.1007/978-3-032-06763-0_43

be particularly helpful in understanding the features and options of corresponding data, especially in complex and ever-changing ITS environments [2]. Thus, this work enables to describe data offerings via metadata, answering questions such as what data are offered, by whom they are published, what access conditions apply, etc. More specifically, the work structures and harmonises such descriptions in the form of a formal metadata specification called mobilityDCAT-AP. This specification provides precise and unambiguous metadata designations for any mobility related data offerings, facilitating harmonised, platform-independent metadata descriptions both in human-readable and machine-readable formats.

An important driver for such metadata specification is found in the context of interoperability. The mobility data ecosystem consists of different elements that complement and relate to each other. For example, a single data set may be bound to a single use case or a single region. To get a complete picture, multiple data sources from multiple portals may need to be combined. However, handling multiple data sources requires that the corresponding metadata are interoperable in terms of semantics and syntax. Interoperability depends on domain-specific metadata specifications or standards [3]. They aim to provide an agreed and uniform way of describing metadata, regardless of the metadata origin or user. They precise formal structures, vocabularies and agreements on their use. Another driver is the need for efficient ways of storing and sharing metadata, giving a key role to Resource Description Framework (RDF), which is also a prerequisite for Linked Data and Semantic Web concepts [4]. As a result, many of the established metadata standards and specifications are based on RDF.

This work has been carried out in the context of NAPCORE, an EU-cofunded Programme Support Action under the GRANT AGREEMENT No MOVE/B4/SUB/2020-123/SI2.852232 [5].

1.2 Previous Works

One well-known framework for defining metadata and specifications is the Data Catalog Vocabulary (DCAT), as an RDF vocabulary designed to describe data catalogues [6]. DCAT Application Profile (DCAT-AP) specifies DCAT as a basic profile for data portals in Europe to facilitate the aggregation, exchange, search and auto-mated processing of metadata [7]. DCAT-AP scope is cross-border and cross-domain, and thus further specified in different, domain-specific extensions like GeoDCAT-AP for spatial data [8] and StatDCAT-AP for statistical datasets [9]. In the field of mobility, some extensions such as TransportDCAT-AP in the domain of public transport, as created in the OASIS project [10], and spsDCAT-AP in the domain of Smart Parking Systems are based on DCAT-AP [11].

The Shift2Rail IP4 SPRINT project provided an automated solution for ingesting and harmonising metadata from selected NAPs to enable cross-border multimodal journey planning [12]. Similar work was carried out in the LOD-RoadTran18 project [13], which includes a specific metadata model to represent road-traffic information datasets. Both projects again rely on the DCAT-AP vocabulary.

Besides such DCAT-AP-based approaches, there are some proprietary metadata specifications, i.e., approaches based on proprietary metadata models and ontologies. The Coordinated Metadata Catalogue of the EU EIP project was a first common blueprint for metadata structures for NAPs in Europe [14]. It defines a common minimum set of metadata, including descriptions, formatting constraints and obligation levels. However, these definitions are only available in a proprietary, human-readable format.

This review identifies some previous approaches to metadata specification and harmonisation. However, many of these have not been developed for the generic domain of mobility. Moreover, some seem to be at an experimental level, not ready for community-wide and long-term application. Finally, some did not fully address interoperability and conformance with semantic technologies via RDF. Therefore, the presented work is a first approach towards a domain-wide, application-ready and RDF-compliant metadata specification in the domain of mobility data. It adds the mobility domain to the DCAT-AP extension family, with an own extension called mobilityDCAT-AP.

2 Development of a Metadata Specification

2.1 Preparatory Works

As a starting point, the NAPCORE Metadata Working Group [5] defined a roadmap for the design, implementation and publication of the envisaged metadata specification. A first action was a detailed requirement analysis involving experts and mobility-data stakeholders [15]. Based on this, a conceptual model was developed, which translates the before-formulated requirements into a technology-independent data structure. We focused on essential information about a mobility data offering on a data portal, and analysed common practice in NAPs and other mobility portals, harmonisation exercises, and feedback from experts. Figure 1 shows an aggregated view of the essential aspects identified.

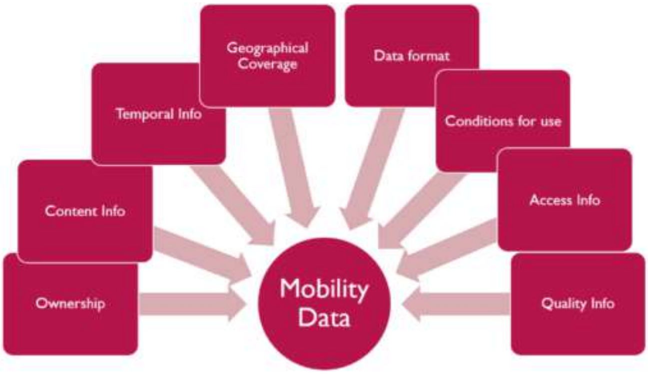

Fig. 1. Essential aspects of mobility data – Basis for the conceptual model of mobilityDCAT-AP

Such essential aspects were then disaggregated in much more detail, also considering mobility-related characteristics. For example, the aspect "Content info" was specified by the transport mode, the part of the transport network and other details covered by the data offering. The final conceptual model was developed iteratively in a tabular format, listing essential metadata elements including definitions, obligation levels, and usage notes.

2.2 RDF Vocabulary

The elements of the conceptual model were then mapped to corresponding RDF elements. RDF data representations are based on a data model architecture that includes classes and properties. As mobilityDCAT-AP is planned as an extension to DCAT-AP [7], the first step was to check if and how existing elements from the DCAT-AP vocabulary correspond to the elements of the conceptual model. In fact, many existing elements could be reused, although some definitions and usage notes had to be rewritten, and some obligation levels had to be changed. Further, additional classes and properties from other vocabularies were incorporated. For example, the quality description of a data offer, being an essential aspect to be provided via metadata, was considered by reusing the established Data Quality Vocabulary [16], which is another RDF vocabulary.

In cases where neither DCAT-AP nor any other known RDF vocabulary provided applicable elements, mobilityDCAT-AP added its own vocabulary elements. These additions aim at capturing some specific characteristics and features of mobility data. This is done by creating a separate namespace “mobilitydcatap:” and individual class/property declarations. Among other things, a property “mobilitydcatap:transportMode” has been introduced this way. MobilityDCAT-AP has also re-moved some optional properties from DCAT-AP. The reason for this is the rather broad scope of DCAT-AP, and the aim to keep the vocabulary size of mobilityDCAT-AP as compact as possible.

The final data model, comprising all RDF elements representing the conceptual model, was visualized as a Unified Modelling Language (UML) diagram. Figure 2 shows central classes from this diagram, and selected properties, as embedded within such classes.

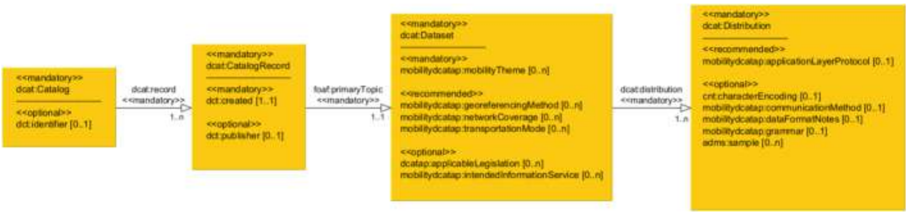

Fig. 2. UML Diagram of mobilityDCAT-AP

Accordingly, four central classes represent a hierarchical concept when describing metadata via mobilityDCAT-AP. Firstly, the Catalogue as such is described, i.e., being the metadata register in a data portal. Secondly, there is the Catalogue Record, which describes the metadata entry, including its publication date. Thirdly, the Dataset is described. In fact, most metadata elements are covered here, including the content theme; the spatial and temporal context; quality information and others. Finally, the distribution describes a technical and organisational way to access the Dataset. In addition to the data format (e.g., a machine-readable syntax standard), the licensing terms are described here. Some additional classes (not shown above) act to support these general classes.

Some of the vocabulary elements are restricted in such a way that the possible expressions of the vocabulary are bound to predefined value lists. Such value lists are stored under Controlled Vocabularies. For example, the Controlled Vocabulary of the above-mentioned property “mobilitydcatap:transportMode” contains the values “car, truck, bike, pedestrian, etc.”. Such predefinitions aim at the unambiguous use of frequently used terms. In addition, Controlled Vocabularies are also expressed as RDF vocabularies and allow interoperable processing. For example, a Controlled Vocabulary could be provided in multi-lingual versions.

2.3 Documentation and User Support

There is online documentation of the mobilityDCAT-AP specification [17], including a human-readable vocabulary description; a set of usage notes; a set of machine-readable serialisation formats to allow IT systems process the vocabulary; and schemas for automated conformance validation. In addition, user support is provided via a dedicated GitHub repository where users can interact with the editors, for example by raising questions and issues.

Finally, there is an Accompanying Guideline that serves as a practical orientation for users of mobilityDCAT-AP. This includes additional explanations and examples for specific vocabulary elements; recommendations for metadata handling and exposure on individual IT systems; and advice on metadata quality and validation processes.

3 Outlook

Following its release in autumn 2023, this specification is promoted for wide-spread use in mobility data portals. Some NAPs are interested in implementing mobilityDCAT-AP in their systems as early adopters. The NAPCORE Metadata Group supports such adopters through direct support and feedback. In addition, the development of a cross-border metadata directory using mobilityDCAT-AP will be demonstrated as a proof-of-concept.

Lastly, a maintenance and governance organisation will take care of future versions of mobilityDCAT-AP. These versions will consider, among others: a refinement of class and property definitions; further precision of controlled vocabs; alignment with other DCAT-AP extensions; and the adoption of progress in the DCAT(-AP) super-specifications. Such maintenance and governance structures will ensure the long-term acceptance and sustainability of mobilityDCAT-AP, and ensure safeguards are in place for a responsible use of metadata in the mobility field. Lastly, efforts will be made to care about privacy to prevent and mitigate the risk of misuse, related to metadata exchange [18].

References

1. NAPCORE information about National Access Points (NAPs). https://napcore.eu/description-naps/. Accessed 01 Sept 2023
2. Wigan, M., Grieco, M., Hine, J.: Enabling and managing greater access to transport data through metadata. Transp. Res. Rec. **1804**(1), 48–55 (2002)

3. Duval, E.: Metadata standards: what, who & why. J. Univ. Comput. Sci. **7**(7) (2001)
4. Riley, J.: Understanding metadata: what is metadata, and what is it for?: a primer. NISO Primer Series (2017)
5. NAPCORE Sub-Working Group on Metadata. https://napcore.eu/metadata/. Accessed 01 Sept 2023
6. Data Catalog Vocabulary (DCAT) - Version 2, https://www.w3.org/TR/vocab-dcat-2/. Accessed 01 Sept 2023
7. DCAT-AP 3.0, https://semiceu.github.io/DCAT-AP/releases/3.0.0/. Accessed 01 Sept 2023
8. GeoDCAT-AP - Version 2.0.0. https://semiceu.github.io/GeoDCAT-AP/releases/2.0.0/. Accessed 01 Sept 2023
9. StatDCAT Application Profile for data portals in Europe. https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/statdcat-application-profile-data-portals-europe/release/100. Accessed 01 Sept 2023
10. Chaves, D., Rojas, J.A., Colpaert, P., Chamorro, J., Corcho, O.: A common metadata model for representing public transport at the European level, TransportDCAT-AP and Controlled Vocs 1.2. Deliverable of the OASIS project (2015)
11. Lubrich, P.: Smart parking systems: a data-oriented taxonomy and a metadata model. Transport. Res. Rec. **2675**(12), 1015–1029 (2021)
12. Carenini, A., Fiano, A., Scrocca, M., Comerio, M., Celino, I.: Enabling cross-border travel offers through national access point federation via metadata harmonisation. In: Proceedings of Sem4Tra 2021: Third International Workshop on Semantics and the Web For Transport (2021)
13. Gutiérrez Moret, J., Samper-Zapater, J.J., de Siqueira Rocha, J.M., Martínez-Durá, J.J.: Implementation and deployment of an Lod enhanced tis. Expert Syst. Appl. (not yet published)
14. Lubrich, P., Jorna, R., Hendriks, L.: Harmonised metadata for transportation data portals. In: Proceedings of the 27th ITS World Congress, Hamburg, Germany (2021)
15. Scrocca, M, Azzini, A., Bureš, P., Comerio, M., Lubrich, P.: Towards napDCAT-AP: roadmap and requirements for a transportation metadata specification. In: Proceedings of SEMANTICS 2022 EU: 18th International Conference on Semantic Systems, Vienna, Austria (2022)
16. Data on the Web Best Practices: Data Quality Vocabulary, https://www.w3.org/TR/vocab-dqv/. Accessed 01 Sept 2023
17. mobilityDCAT-AP - Version 1.0.0, https://mobilitydcat-ap.github.io/mobilityDCAT-AP/releases/index.html/. Accessed 08 Aug 2023
18. De Montjoye, Y.-A., Hidalgo, C.A., Verleysen, M., Blondel, V.D.: Unique in the crowd: the privacy bounds of human mobility. Sci. Rep. **3**, 1376 (2013)

The PRESORT Project: Improving the Use of Third-Party Data by National Road Authorities

Dave Cowell[1](✉), Candida Spillard[1], Anastasia Tsapi[2], Andy Graham[3], Giovanni Huisken[4], Tomi Laine[5], Marwane Avaida[6], and Scott Stephenson[1]

[1] AECOM, London, UK
dave.cowell@aecom.com
[2] Royal Haskoning DHV, Amersfoort, The Netherlands
[3] White Willow Consulting, Redhill, UK
[4] MAPtm, Utrecht, The Netherlands
[5] Traficon, Helsinki, Finland
[6] UPHF University, Valenciennes, France

Abstract. PRESORT is a research project funded by the Conference of European Directors of Roads (CEDR) through the CEDR Transnational Road Research Programme (Call 2022). The aim is to improve the use of third-party data by National Road Authorities (NRAs).

NRAs are seeing increasing levels of digitisation and are open to the potential opportunities through the utilisation of third-party data – that is, data aggregated from multiple sources – to support them in the delivery of their core business services in traffic management, asset management, and construction.

PRESORT will eventually deliver practical, implementable, easy to use online guide to support NRAs to make better decisions regarding HOW and WHEN to acquire and use third-party data, which will assist their core business activities.

The initial phase of this project involved capturing the current state of third-party data use by NRAs. Part of this phase explored the challenges and barriers NRAs face regarding use of third-party data through a literature review and engagement with data providers and crucially, the NRA end-users. This paper reports on the results of this engagement.

Keywords: Third-party data · Cooperative ITS · Road safety · Road user charging

1 Introduction and Development of Questionnaire About NRA Use of Third-Party Data

Third-party data is aggregated or collated by a provider not involved in the original collection of the data. The data is then provided to a National Road Authorities (NRA).

PRESORT completed a broad literature review to understand current use of third-party data by NRAs. Given the nature of the subject, the search was restricted to identifying evidence from recent years – with over 280 separate articles discovered. This

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 305–311, 2026.
https://doi.org/10.1007/978-3-032-06763-0_44

literature review provided the background to an online questionnaire designed to engage key individuals from NRAs and third-party data providers.

Prior reporting highlighted data accessibility, data quality, data interoperability, legal matters such as GDPR considerations, and a general lack of the skills required to address these issues. There appears to be limited or no data flow between three data eco-systems in the transport system – cooperative ITS, data for road safety, and road user charging.

2 Questionnaire Respondents

A questionnaire was sent out to 79 targeted individuals working in NRAs or data providers, as well as to stakeholder representatives and independent experts. Participants were also encouraged to forward the questionnaire to others whom they deemed appropriate. A total of 34 individuals completed the questionnaire.

The questionnaire contained 24 questions related to the current and future use of third-party data by NRAs, the risks and challenges of using such data, and the preparedness of NRAs and their workers. This paper provides a report about the challenges and barriers to third-party data use by NRAs.

Questionnaire respondents reported they worked across 16 different countries or regions, as shown in Fig. 1, and many worked globally or across multiple countries. The majority (22) work either for or on behalf of an NRA, and most (30) reported their organisation used third-party data.

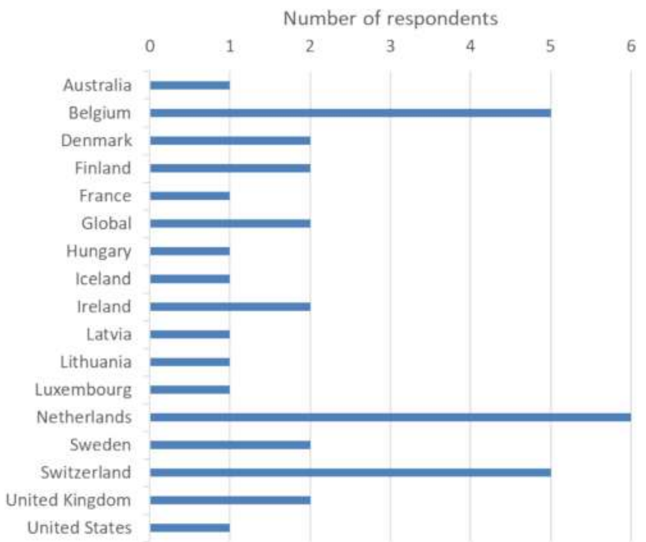

Fig. 1. Countries from which the questionnaire respondents operate.

3 Use of Third-Party Data by Questionnaire Respondents

Questionnaire respondents were also asked what type of third-party data they use. Their responses are summarised in Fig. 2, although it is worth noting this cannot be used as representative of the entire industry and only helps to qualify their other responses. However, it is a reflection of the types of third-party data that are available from data providers and partly a reflection of the previous demand for such data.

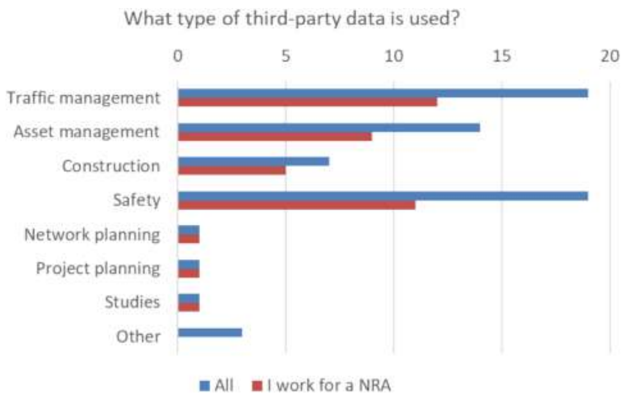

Fig. 2. The types of third-party data used by the questionnaire respondents.

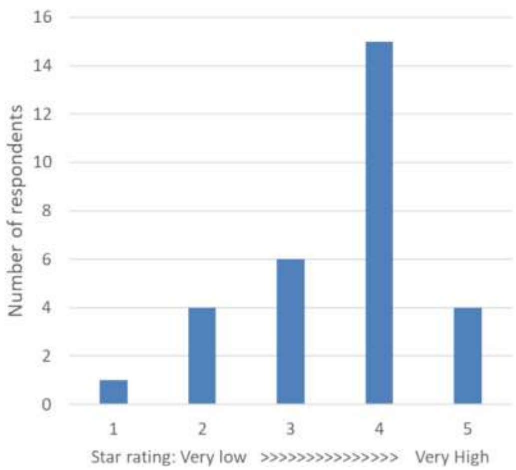

Fig. 3. Questionnaire respondents' estimation of their organisation's maturity level in using third party data.

The questionnaire also asked respondents to rate their organisation's maturity level in using third-party data. Generally, the respondents believed their organisation was mature (as shown in Fig. 3, their data management systems were ready, and the people in their organisations have the skills to use third-party data.

4 Challenges and Barriers to Third-Party Data Use by NRAs

Participants were asked what challenges there are to using third-party data. They were given several options to tick as well as the option to identify additional issues. Table 1 summarises the challenges and barriers identified by participants and a description of the issues.

Table 1. Challenges and barriers to third-party data use by NRAs identified by questionnaire respondents.

Challenge	Description
Privacy and security	GDPR (and similar) apply if the data has not been anonymised. Anonymisation is reversible for very small datasets or sections
Data sharing	Parties may be reluctant to share data or have little incentive to share
Practical applications	A common issue is the existence of seemingly useful third-party data without a specific application. A solution in search of a problem
Incompatible data formats	This may result either from the data type, or from the method used to collect it, for example different timeslots, leading to the need for interpolation
Data maintenance	Use of third-party data relies on the third-party, with no control over its consistency or indeed continued availability
Trust	Many respondents identified trust in third-party data's integrity and continuity as possible issues
User skills and training	Extra training will take time and resources away from the main functions of NRA. Extra recruitment may be needed
Business models	Building the business case for the use of third-party data and investment needed to enable the use of third-party data
NRA culture	Uptake of new ideas can depend upon the culture within the organisation. The need for extra data may not be obvious to the relevant decision-makers
Data use restrictions	NRAs may want to re-use or re-purpose the data for other needs. Often there are restrictions on such activities

As can be seen in Fig. 4, *Privacy or security* was the most selected issue by respondents, although the small questionnaire sample size and characteristics of the respondents does not make this a quantifiable finding. Notably, the results of this question highlight the broad range of challenges to the use of third-party data for NRAs. Participants took the opportunity to highlight issues related to business models, NRA culture, and data use restrictions.

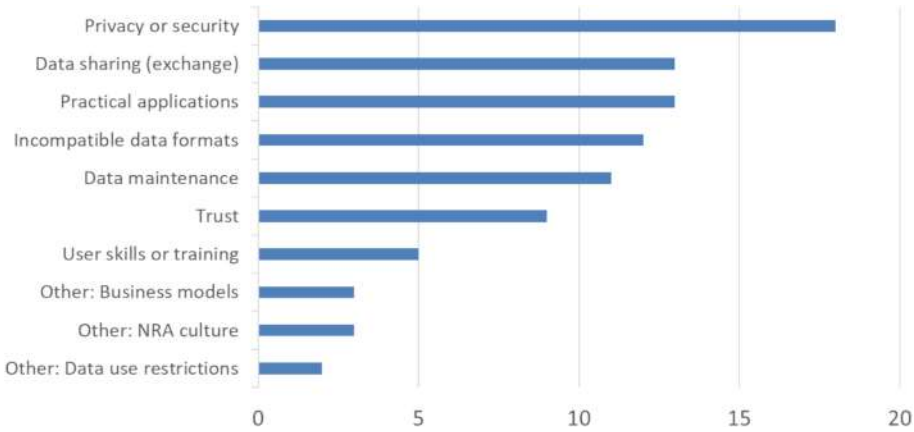

Fig. 4. Histogram showing the challenges to third-party data use selected by questionnaire respondents.

5 Scale of NRA Third-Party Data Usage

A wide variety of national and international bodies have drawn up strategies and roadmaps for the coordination, promotion, and implementation of digital roads including the use of third-party data. For instance, the EU Data Strategy [1] includes the creation of common European data spaces in key sectors, including mobility. Data spaces bring together the governance and infrastructure to facilitate the pooling and sharing of data in a controlled and secure way. The first stage involves gathering and coordinating all data already available. The creation of NAPs (National Access Points) is a key part of the strategy. More aspects of this strategy, including the relevant actions, are detailed on the site *EU Transport – Road* web page [2].

Some notable strategies include:

– The national vehicle authority in the Netherlands, RDW, have a three-pillar strategy [3]: innovation based on continuity; customised service, and collaboration in networks. This includes regard of ICT and connected and autonomous vehicles.
– In Germany die Bundersregierung Strategy for Automated and Connected Driving [4] has objectives specific to connectivity. These are to remain a lead provider, to become a lead market, and to implement connectivity on the road network.
– Belgium has separate strategies for Flanders and Wallonia.
– In the UK, National Highways' Digital Roads Strategy [5], involving construction and asset management and operation of the Strategic Roads Network (SRN) was announced in 2020. This includes the delivery of *Digital Roads 2025*, the relevant sections of which are *Single view of the network*, *Every customer has an option (ECHO)*, and *Enabling a connected and autonomous future*.

– EU-wide, the Network-Wide Road Safety Assessment Methodology [6] sets out the method to ascertain road network rating and ranking for all strategic roads throughout, specifying what data needs to be collected.
– The EU-wide initiative Data For Road Safety [7] (DFRS) provides the dissemination of live information throughout the European Union and the UK.

There are also cross-border initiatives such as the C-Roads Platform [8] that may involve the use or third-party data. The C-Roads Platform brings authorities and operators together to harmonise the deployment of C-ITS across Europe. Objectives include the effective exchange of data.

6 Conclusion

Individuals working in relevant positions have highlighted a broad range of challenges to the use of third-party data by NRAs, which warrant further detailed exploration and documentation.

PRESORT will carry out consultations and workshops with stakeholders (NRAs and third-party suppliers) to explore data needs and use cases, as well as generating interest in data sharing between the C-ITS, data for road safety, and the road user charging data ecosystems. It will also identify the key technical, commercial, legal, and organisational challenges of NRAs and third-party data service providers with regards to acquisition, use, and maintenance of third-party data.

References

1. Unlocking the potential of mobility data, European Commission. https://digital-strategy.ec.europa.eu/en/policies/mobility-data. Accessed Mar 2024
2. Intelligent transport systems: Road, European Commission. https://transport.ec.europa.eu/transport-themes/intelligent-transport-systems/road_en. Accessed Mar 2024
3. Safe and reliable on the road: The mission, vision and strategy of the DRW, Rijkswaterstaat. https://www.rdw.nl/-/media/rdwnl/overrdw/documenten/brochures-en-folders/rdw_strategy_en_def.pdf. Accessed Mar 2024
4. Strategy for Automated and Connected Driving, Federal Ministry for Digital and Transport. https://bmdv.bund.de/SharedDocs/EN/publications/strategy-for-automated-and-connected-driving.pdf. Accessed Mar 2024
5. Digital roads, National Highways. https://nationalhighways.co.uk/our-work/digital-data-and-technology/digital-roads/. Accessed Mar 2024
6. Network Wide Road Safety Assessment, European Commission (2023). https://road-safety.transport.ec.europa.eu/eu-road-safety-policy/priorities/infrastructure/road-infrastructure-guidelines_en. Accessed Mar 2024
7. Data for road safety. https://www.dataforroadsafety.eu/. Accessed Mar 2024
8. Harmonisation of C-ITS related reployments throughout Europe, C-Roads. https://www.c-roads.eu/platform/about/about. Accessed Mar 2024

Cross-Border Data Exchange to Improve Traffic Management and End-User Information Services in Central Europe

Wolfgang Kernstock(✉)

ASFINAG Maut Service GmbH, Schnirchgasse 17, 1030 Vienna, Austria
wolfgang.kernstock@asfinag.at

Abstract. In the European vision of a free movement of people and goods, it is obvious that traffic management strategies must not be limited to national borders. Due to the fact that Austria is a small country within an important transit area it is necessary to have a close cooperation with the neighbouring countries. Data is an important basis for modern traffic management, and it became apparent that motorway operators already have many relevant management strategies available. Therefore, ASFINAG has ever since pursued the coordination and harmonisation of Member-State specific approaches, on at technical and organisational level, by aligning topics such as traffic information and management at cross-border level.

The ASFINAG solutions are in principle quite easy with mainly list views beside map based information delivered. Therefore similar operators might have no problems in establishing a similar service at low costs as well as in short term. However the content delivering backend systems need to be at hand in order to fuel the customer service.

Keyword: Cross-Border · Data exchange · End-user services

1 Motivation

Traffic and transport across main corridor routes hold great potential for substantial improvements in efficiency and fluency. With that however come risks of disturbances, either on regular (e.g. holiday traffic congestions) or exceptional basis (e.g. heavy weather conditions).

Especially during the last years, additional draw backs have been experienced due to increased cross-border waiting times, related to control procedures in conjunction with applicable COVID-19-regulations. This is a possible risk for commuters and truck drivers, who have to cross borders on a regular basis – often several times a day. Beyond that, situations such as increased traffic during holiday seasons or harsh weather events, although more predictable, can quickly lead to major delays, accidents and frustrated road-users.

In the past, activities from the involved ministries, police and operators were often focussed on quickly restoring traffic flow. This was merely about combating the symptoms but not properly addressing the core of the problem, which was a lack in cross-border coordination and appropriate access to data and information. As transport is not

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 312–317, 2026.
https://doi.org/10.1007/978-3-032-06763-0_45

ending at organisational borders, it is all the more important to draw the links to adjacent urban areas here. While acknowledging the efforts of the previous years, the need for further development of common initiatives reaching beyond borders is still present in order to include real-time events and information that could reach drivers, commuters, HGV drivers or tourists equally.

2 Objectives

As for tourist travelling seasons, there are lot of travellers going from the Northern part of Europe to the seaside in Italy, Croatia and Slovenia during the summer period. To coordinate the mass of traffic going through Austria it is necessary to have a close cooperation between all affected road operators. First TMPs (traffic management plans) for the summer season were already agreed in past years between Italy, Croatia, Austria, Hungary and Slovenia. But it is important to continue this cooperation and to develop additional cross-border TMPs to handle the traffic during the summer period.

The most pressing issues are ongoing coordination and harmonisation on technical and organisational level in line with European legislation. In this way operators and service providers can draw from improved access to data and offer the best possible services to end users. Due to different stakeholders (public/private motorway operators, public authorities, other infrastructure/logistics operators) being responsible for their respective infrastructure segments, a special focus must be on addressing and harmonising the link between high-level corridors and adjacent infrastructure (urban areas, ports, etc.).

3 Data Collection

The cross-border data exchange focus on provision of EU-wide real-time traffic information related to priority action b), on provision of traffic data related to priority actions c), on the provision of information services for safe and secure parking places for trucks and commercial vehicles related to priority action e) and on the provision of multi-modal traveller information related to priority action a) of the ITS Directive (2010/40/EU). The partners are heavily committed to a corridor-based approach and the coordination necessary to ensure long-term relevance and high quality.

Due to the efforts in terms of data exchange and DATEX II profiling, a solid basis has already existed. All the more important it was to focus on the details, e.g. parameters to consider when describing a detected event, the process of detour selection as well as exact and harmonised location referencing.

4 ASFINAG B2C Solutions Towards Requirements of ITS Directive

4.1 ASFINAG Smartphone App

In 08/2011 the first version of the ASFINAG Smartphone Application "Unterwegs" was introduced in the Apple App Store and the Android Play Store (by 2021 a Huawei version was established). Since then a professional ITSM/ ITIL based process management was established to manage the lifecycle of the product leading to an ongoing improvement of the product and the provided information. Because it is now more than ten years old, a general upgrade of this smartphone app was necessary. The new ASFINAG app contains up-to-date information on traffic and road conditions on the Austrian motorway network according the requirements of the ITS Directive. Following main features are available with the new smartphone application (Fig. 1):

- **Personalised homescreen** - a personalised homescreen by adding personal destinations, routes, or webcams.
- **Webcams** - Access to over 1800 live webcam images along with the motorway network including the webcams of the neighbouring countries.
- **Traffic information** - overview of all events and road works along with the motorway network including a push-message function for personalised routes
- **Rest areas** - Overview of all rest areas along with the motorway network including the rest areas of the neighbouring countries.
- **E-Charging station** – Overview of charging stations for electric cars in Austria
- **All about tolls** - purchase of digital vignette or digital route toll via the ASFINAG toll shop, Austria-wide overview of all vignette and GO-Box sales points.
- **Planning & Navigating** - Intermodal Route Planner (Central Europe-wide) with the help of Traffic Information Austria
- **Languages** - ASFINAG App is available in 12 languages

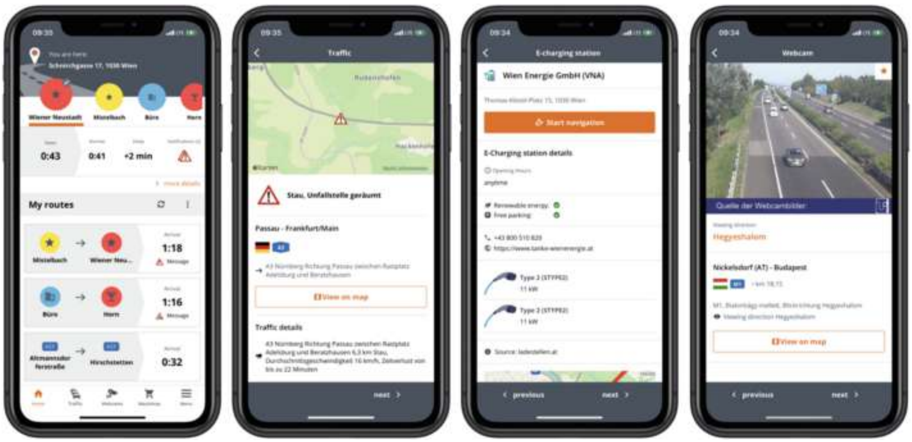

Fig. 1. ASFINAG Smartphone App (https://www.asfinag.at/asfinag-app/)

4.2 ASFINAG Website

The ASFINAG website provides information by means of maps widgets. The focus is on the needs towards requirements of ITS Directive (incl. E-Charging points) and information from neighbouring countries (e.g. Webcams, rest areas, traffic messages…). The update cycle takes place in four regular updates per year and is regularly extended to include data from neighbouring countries (Fig. 2).

Fig. 2. Rest Areas (https://www.asfinag.at/parken-rasten/rastanlagensuche/)

4.3 Common Cross-Border TMP Tool

Together with DATEX II and the exchange of event information, that is relevant for road conditions in neighbouring countries, the digitalisation of traffic management plans (TMPs) has become one essential achievement in terms of harmonised cross-border traffic management pursued in Austria.

Although a mature status on TMPs on national level has already existed, a connection on transnational level was needed. Besides the obvious language barriers, missing automatisms and a unified way of displaying information (both internally and to drivers) were major obstacles when updating existing TMPs or adding new ones.

Therefore, operators from Austria, Croatia, Hungary, Italy and Slovenia have established a common process for this coordination of TMPs. In order to properly present the information from those neighbouring motorway operators, Slovenian motorway operator DARS conceived an interactive road map and integrated it into its central system for displaying and entering traffic information throughout the entire Slovenian road network. The system was built in a centralised way, allowing to enter and review information for all Slovenian (and neighbouring countries') roads from a single tool.

The TMP application consist of an interactive map, where users can see TMPs and active events for the CROCODILE corridor region.

Events can be added if the section is covered by a TMP and then add according information, e.g. type of event and duration. For already existing events, all involved parties will receive a notification and are also able to select TMPs and rerouting strategies. For every event only one strategy can be selected. Parties not involved in the certain plan or strategy can view them, but not select or edit (Fig. 3).

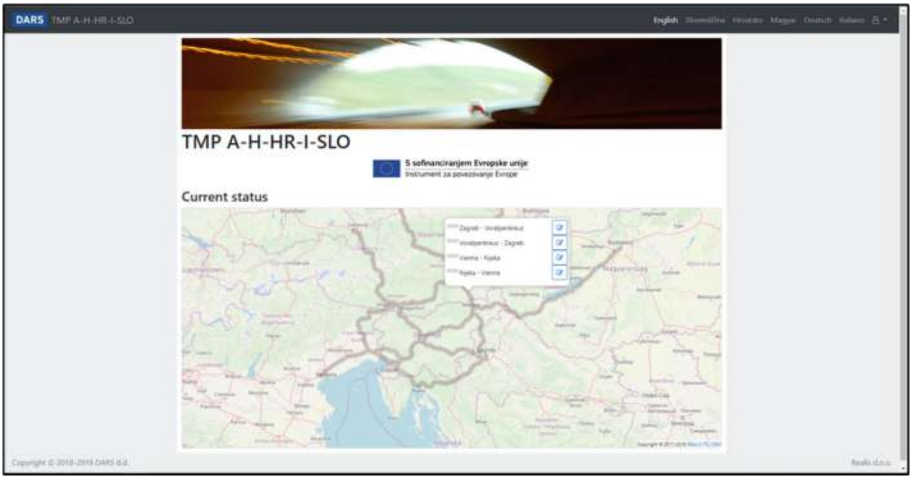

Fig. 3. Common TMP tool (https://tmp.dars.si/tmp/app/default)

5 Results and Evaluation

As Austria is situated in the heart of Europe, special challenges arise to the transportation network as:

- High percentage of foreign drivers (especially during holiday seasons)
- High percentage of the Austrian motorway classified as critical infrastructure in terms of road safety. The total length of the network is per 2020: 2.249 km; 400 km of the network is situated in tunnels (166 tunnels). In addition 5.818 bridges are located on the Austrian motorway network.
- Due to topographical reasons in some areas are only limited rerouting possibilities.

The ASFINAG solutions are in principle quite easy with mainly list views beside map based information delivered. Therefore similar operators might have no problems in establishing a similar service at low costs as well as in short term. However the content delivering backend systems (in Austria’s case e.g. webcam system, roadworks management tool, event database or the Austrian wide route planner) need to be at hand in order to fuel the customer service.

Major findings of the provision of ASFINAG end user services (since beginning in 2011):

- Transparency: delivering real time information pushes the level of data transparency (of course not necessarily bad). With that increase, customers will raise data and service issues, additional ideas and wishes as well as difficulties they experience. Service providers need to be prepared for a high volume of customer contacts in order to satisfy the customers.
- Multilingualism: the provision of the ASFINAG App (now in 12 languages) resulted in reduced flexibility when adding new features as a translation agency has to be consulted for each change request. Additional costs - beside initial costs – for each change are also o be calculated. But the overall feedback show positive response by customers using the service in their preferred language.

6 Conclusions

The past years have shown that cooperation is important to find common solutions. Especially since traffic does not end at organisational borders.

While acknowledging the efforts, the need for further collaboration beyond borders is ever more present. The most pressing issues in this regard are the ongoing coordination and harmonisation at technical and organisational level in accordance with European legislation. In this way operators and service providers can draw from improved access to data and offer the best possible services to end users. Due to different stakeholders (public/private motorway operators, public authorities, other infrastructure) being responsible for their respective infrastructure segments, a special focus must be put on addressing and harmonising the link between high-level corridors and the adjacent transport network (urban areas, secondary road network, multi-modal nodes etc.).

Other issues include the harmonised implementation of the ITS Directive and its supplementing Delegated Regulations. While basic principles are set out as mandatory provisions, the technical and organisational details often remain unclear and subject to Member-State-specific approaches. To ensure a common European and corridor-based strategy, it is essential to further foster the alignment of the previously mentioned topics such as DATEX II profiles as well as traffic information and management at cross-border level. This approach has been successful in the past and should be pursued further for sustainable and long-term corridor benefits.

References

1. CROCODILE 3 corridor project https://www.its-platform.eu/its-corridors/crocodile/. Milestone 29 Report: Recommendations based on the findings of the CROCODILE 3 Working Groups
2. Traffic Management Plan tool. https://tmp.dars.si/tmp/app/default
3. ASFINAG Website. https://www.asfinag.at/

AI-Assisted Services for Content Acquisition and Creation Interlinking in Transport Research

Xenophon Kitsios and Afroditi Anagnostopoulou(✉)

Centre for Research and Technology Hellas/Hellenic Institute of Transport, Kasimati 1, 18531 Piraeus, Greece
a.anagnostopoulou@certh.gr

Abstract. This paper presents the gateway of the domain-specific Knowledge Graph (KG), which was built based on related data sources and by exploiting the OpenAIRE ecosystem and EOSC services. These enabled it to offer services for integrated KG smart browsing based on impact and reproducibility using AI by also serving several categories of stakeholders. Following the current trends and stakeholders' needs, the areas of highest interest were identified and gaps in data and knowledge were also detected. On top of that knowledge space, transport research-inspired information retrieval scenarios were implemented by tuning the use of individual AI-assisted services or combinations of them. More specifically, scenarios were built around different use case directions including reproducibility/reusability reports for publications and/or datasets in the transport research sector which were automatically identified and they included information about the ease to reproduce/reuse them and the extent to which the work has already been reproduced by meta-analysis studies.

Keywords: Knowledge Graph · OpenAIRE · Reproducibility · Data sets · Software PoCs · Open Science

1 Introduction

Transportation research sector involves different modes of transport (i.e., road, rail, water, air, cross-modal) and different types of mobility (i.e., passenger and freight) producing innovation uptakes that could enable users and societies to be better informed and make safer, more coordinated and 'smarter' use of transport systems. More specifically, information processing is combined with engineering studies, technologies and human sciences producing research data span multiple domains (e.g., algorithms, software Proof of Concepts (PoCs), methods for technological impact, user acceptance, inputs to draft standards and regulations of newly introduced technologies, etc.). In addition, several domain-specific knowledge bases (i.e., TRID, ERTRAC, ERTICO, ECTRI, IRU, WEGEMT, EASA, ERRAC) are exploited to enhance scientific quality, identify remarkable knowledge and share research activities across national boundaries.

In this context, a domain-specific Knowledge Graph (KG) developed by relying on related data sources and exploiting the OpenAIRE ecosystem and EOSC services. These

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 318–324, 2026.
https://doi.org/10.1007/978-3-032-06763-0_46

enabled it to offer services for integrated KG smart browsing based on impact and reproducibility using AI by also serving several categories of stakeholders (i.e., researchers, developers, regulators). The areas of highest interest were identified and gaps in data and knowledge were also detected. On top of that knowledge space, transport research-inspired information retrieval scenarios were implemented by tuning the use of individual AI-assisted services or combinations of them. More specifically, scenarios were built around different use case directions including reproducibility/reusability re-ports for publications and/or datasets in the transport research sector which were automatically identified, and they included information about the ease to reproduce/reuse them and the extent to which the work has already been reproduced by meta-analysis studies.

The remainder of this paper is organized as follows: Sect. 2 provides an overview of the key developments and trends in the field of knowledge graphs up to the present day. The followed domain-specific Knowledge Graph in the context of transport research and the derived results of its implementation are presented in Sect. 3. Finally, in Sect. 4 conclusions are drawn and pointers for future research are provided.

2 Literature Review

Knowledge graphs have emerged as a fundamental tool for organizing and representing structured knowledge in a graph-based format (Hogan et al., 2021). They have gained significant attention in recent years due to their wide-ranging applications in various domains, including natural language processing, data integration, semantic web, recommendation systems, and artificial intelligence (Chen et al., 2020). The concept of knowledge graphs can be traced back to early AI research, but it gained substantial traction with the advent of the Semantic Web initiative (Chaudhri et al., 2022). Researchers recognized the need for a standardized, machine-readable representation of knowledge to enable machines to understand and reason about information on the web.

The key components of Knowledge Graphs (KG) involve nodes and edges with semantics underlying, and they could be represented as Resource Description Framework (RDF) or Web Ontology Language (OWL). The KGs consist of nodes representing entities (e.g., people, places, concepts) and edges representing relationships between these entities (Hamilton et al., 2018). This graph-based structure allows for the representation of complex relationships and hierarchical information. In addition, the RDF and the OWL (Gutiérrez and Sequeda, 2021) are commonly used formalisms for representing knowledge in a structured and machine-readable way. RDF provides a simple, flexible format for expressing subject-predicate-object, while OWL allows for more complex ontology modeling.

The application of KGs varies as they could be used in (i) semantic search (Li, 2017) allowing users to retrieve relevant information and answer complex queries, (ii) in recommendation systems (Dongliang et al., 2022) to help in understanding user preferences and item attributes, (iii) in question answering (Chakraborty et al., 2021) for extracting relevant information from vast knowledge bases to provide concise answers to user queries, and (iv) in natural language processing (He et al., 2022) for entity recognition, sentiment analysis, and text summarization by incorporating semantic context.

In recent years, several techniques have been developed to learn continuous vector representations of entities and relationships in knowledge graphs, enabling better reasoning and link prediction (Ali et al., 2021). Moreover, large-scale KGs (such as Wikidata and Google Graph) created and offer vast, interconnected knowledge bases covering a wide range of domains. KGs also integrated with Graph Neural Networks (GNNs) to improve the performance in various applications, such as entity linking, node classification, and recommendation (Park et al., 2019). Finally, KGs are used to provide explanations for AI decision-making processes, increasing transparency and trust in AI systems (Lecue, 2020).

3 Knowledge Graph in Transport Research

3.1 Knowledge Graph Structure

The KG consist of five entities, i.e., 1) Results that represent the outcomes of research activities, 2) Data Sources that are the sources from which the metadata of graph objects are collected, 3) Organizations that correspond to companies, research institutions, and public authorities involved in projects, 4) Projects that are research project, and 5) Communities that are groups of people with a common research intent, such as the Transport Research Community. Furthermore, the OpenAIRE research graph is utilized to develop various services offered to three distinct categories of organizations within the Transport Research Community. These categories are: 1) Content Providers, 2) Research Organizations, and 3) Public Authorities.

Regarding Content Providers, OpenAIRE offers validator services to automatically verify compliance of exported records with set guidelines. Additionally, full-text mining algorithms utilize the full-texts of Open Access publications to enhance metadata records by linking them to projects, publications, datasets, software, organizations, research infrastructures, and terms from standard classification schemes. For Research Organizations, the services aim to support the adoption of Open Science publishing practices and monitor the implementation of Open Science practices by their researchers. In cases where an official Open Access repository is unavailable, any organization can easily create a dedicated community and encourage affiliated researchers to deposit their research outputs there. Within the user interface, users can access content reports and project lists. The content report provides information about the research outcomes associated with the organizations, while the project list outlines the projects the organization is involved in and their respective research outcomes. Lastly, Public Authorities, such as national and international research funding organizations, can access services to monitor the impact of their funding within the Transport research community.

The gateway for the Transport Community is provided in the https://beopen.openaire.eu/ focusing on promoting territorial and cross border cooperation and contributing to the optimization of open science in transport research. Integrating the described services and leveraging the OpenAIRE Graph, the Transportation Research community benefits from enhanced publication deposits and tools for in-depth exploration.

3.2 Results

The studied scenarios of transport research sector following the aforementioned gateway of the OpenAIRE, and the derived results are presented below.

(i) The first scenario studies the use case of researchers that are interested in replicating experiments and verify algorithm's performances on benchmark datasets utilizing open access algorithms/software. Therefore, the developed KG offers a **benchmark reproducibility report of algorithms/software** (Fig. 1) that addresses complex logistics problems by specifying "open access" in the Access category and using as keywords "optimization", "algorithms" and "logistics". Further results with closed access are also found and could be presented by altering our research.

(ii) The second scenario studies the use case of datasets that used for optimization purposes in the logistics sector. The developed KG offers a **benchmark datasets report** (Fig. 2) that used to assess algorithms performance by utilizing the keyword "optimization". This keyword is used by the KG to deliver relevant research data sets as it is specified in the Type category selecting "research data". The results are referring only to the open ones as "open access" was selected in the Access category.

(iii) The third scenario studies the use case of generalization, retrieving information about research publications that study "automation" in different transportation domains. The developed KG offers a **generalization report of publications** (Fig. 3) including information on studies and experiments, such as how an algorithm's performance varies with problem conditions or complexity. The keyword "automation" is utilized for searching the KG and open results are presented as "open access" is selected in the Access category.

Figure 1 reproduced from the https://beopen.openaire.eu/, *developed within the BE OPEN and SCILAKE projects.*

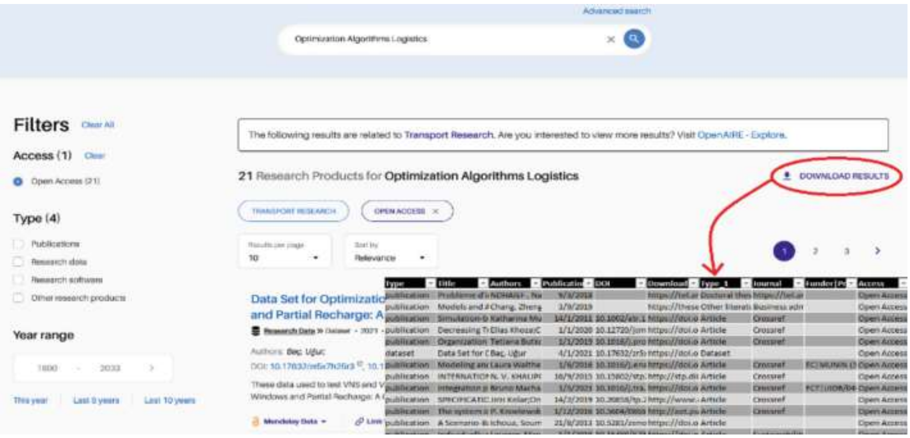

Fig. 1. Benchmark reproducibility report.

Figure 2 reproduced from the https://beopen.openaire.eu/, *developed within the BE OPEN and SCILAKE projects.*

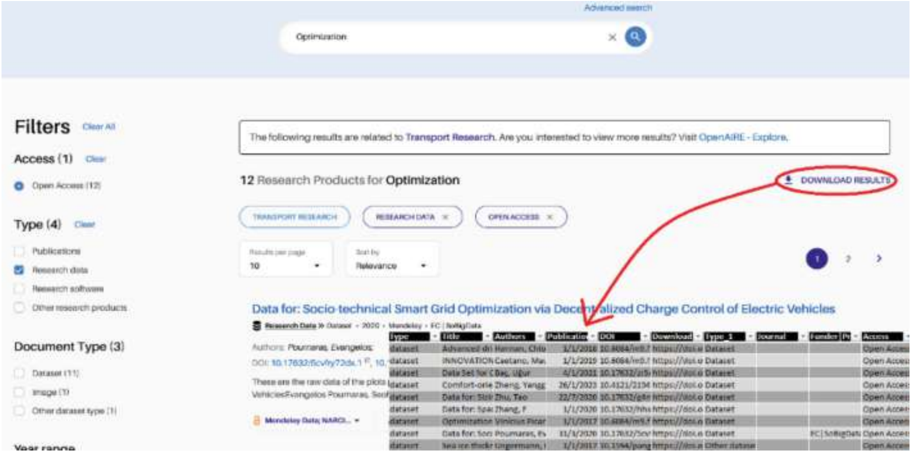

Fig. 2. Benchmark datasets report.

Figure 3 reproduced from the https://beopen.openaire.eu/, developed within the BE OPEN and SCILAKE projects.

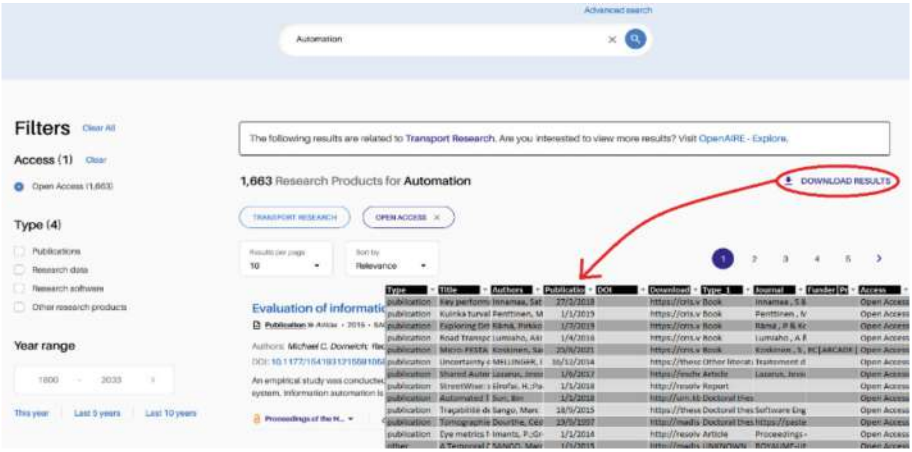

Fig. 3. Generalization report of publications.

3.3 Next Steps

The developed AI-assisted services could be further enhanced by allowing algorithms comparisons and reproducibility to demonstrate improvements of new algorithms. This should include the code for the baseline algorithms, configuration settings, and clear instructions for conducting fair comparisons in order to enable others to verify the reported performance gains. In addition, parameter tuning is another future step for improvement. Optimization algorithms often have hyperparameters that significantly impact their performance. Thus, reproducibility reports that can outline the process of hyperparameter tuning, including the range of values tested, the criteria for selecting the best configuration, and the scripts used for this purpose could contribute to the advancement of knowledge and the development of more robust optimization algorithms.

4 Conclusions and Future Research

AI-assisted services in the transport research sector introduced to promote science and foster collaboration among research community and other relevant stakeholders. Towards this direction, this paper presents and analyses the services of the gateway for the Transport Community that exploits OpenAIRE Graph. According to the derived results, an open and sustainable scholarly communication is established monitoring and providing all research outcomes in the transport research sector. The presented scenarios and future steps reflect a commitment to advancing research in the sector and optimizing algorithm performance, with a strong emphasis on open-access resources and reproducibility. Consideration is also given on future research and directions are drawn. Ethical and legal considerations should be examined related to AI-assisted services and open access initiatives by exploring issues such as data privacy, intellectual property, and responsible AI deployment in scholarly communication.

Acknowledgements. This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101058573. The results in this paper reflect only the authors' view.

References

Hogan, A., et al.: Knowledge graphs. ACM Comput. Surv. **54**(4), 1–37 (2021)

Chaudhri, V., et al.: Knowledge graphs: introduction, history and perspectives. AI Mag. **43**(1), 17–29 (2022)

Chen, Z., Wang, Y., Zhao, B., Cheng, J., Zhao, X., Duan, Z.: Knowledge graph completion: a review. IEEE Access **8**, 192435–192456 (2020)

Hamilton, W., Bajaj, P., Zitnik, M., Jurafsky, D., Leskovec, J.: Embedding logical queries on knowledge graphs. Adv. Neural Inf. Process. Syst. **31** (2018)

Gutiérrez, C., Sequeda, J.F.: Knowledge graphs. Commun. ACM **64**(3), 96–104 (2021)

Li, Y.: Research and analysis of semantic search technology based on knowledge graph. In IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), vol. 1, pp. 887–890. IEEE (2017)

Dongliang, Z., Yi, W., Zichen, W.: Review of recommendation systems based on knowledge graph. Data Anal. Knowl. Disc. **5**(12), 1–13 (2022)

Chakraborty, N., Lukovnikov, D., Maheshwari, G., Trivedi, P., Lehmann, J., Fischer, A.: Introduction to neural network-based question answering over knowledge graphs. Wiley Interdisc. Rev. Data Min. Knowl. Disc. **11**(3), e1389 (2021)

He, Q., Yang, J., Shi, B.: Constructing knowledge graph for social networks in a deep and holistic way. In: Companion Proceedings of the Web Conference 2020, pp. 307–308, IEEE (2022)

Ali, M., et al.: PyKEEN 1.0: a python library for training and evaluating knowledge graph embeddings. J. Mach. Learn. Res. **22**(1), 3723–3728 (2021)

Park, N., Kan, A., Dong, X. L., Zhao, T., Faloutsos, C.: Estimating node importance in knowledge graphs using graph neural networks. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 596–606 (2019)
Lecue, F.: On the role of knowledge graphs in explainable AI. Semant. Web **11**(1), 41–51 (2020)

MBSE Approach for Railway Digital Continuity

Sana Debbech[1](✉), Simon Collart-Dutilleul[2], and Philippe Bon[2]

[1] SNCF Réseau, 93210 Saint Denis, France
sana.debbech@gmail.com
[2] COSYS/ESTAS, Université- Gustave Eiffel, 59650 Villeneuve d'Ascq, France
simon.collart-dutilleul@univ-eiffel.fr,
philippe.bon@univ-eiffel.fr.com

Abstract. Railway systems are known as Safety Critical Systems (SCSs). In this kind of system, safety measures derived from the dysfunctional analysis used to be expressed in an informal way. This latter has several gaps in the context of the one going numeric transition: in the early phase of SCSs design, there is a need to link these safety measures to main safety goals. A first step provides a knowledge structure, where the considered knowledge is composed by a set of data and a set of engineering rules. These rules, including safety measures, correspond to a knowhow built through information sharing between actors during previous industrial system life-cycle. From this structured knowledge, models using main concepts can be designed. As concepts come from ontology, the system models are naturally high-level ones and directly linked to the source needs. Indeed, source needs are expressed on the basis of the structuring concepts of the ontology. Obviously, obtained models are abstractions of the real systems. Model based system engineering (MBSE) allows a systematic reasoning and tooled conformance checking and it is possible to assign a meaning to measured data during the whole life cycle of the railway system. A fundamental assumption is the validity of models used during this life cycle. As an abstraction is a partial point of view, the relevance of this partiality must be monitored during the system life cycle in order to avoid ambiguous interpretations. In this paper, the semantic interoperability is tackled to avoid ambiguities and to ensure the railway digital continuity.

Keywords: MBSE · Digital Continuity · Railway · Ontologies · Safety

1 Introduction

Railway system's design, operation and maintenance require the involvement of several stakeholders such as system engineers, safety engineers, railway experts etc. Indeed, in the system design phases, actors and needs are raised and requirement engineers identify functional requirements, their allocation to stakeholders, flows and interfaces between physical objects. Each phase of the system development involves multi-disciplinary experts that share heterogeneous data and collaborate without an unambiguous terminological guide about exchanged concepts. Actual practices are mainly based on sharing specification and documents in textual form in order to communicate about system

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 325–329, 2026.
https://doi.org/10.1007/978-3-032-06763-0_47

related choices. Furthermore, models are built independently and refers to different levels of Model-Driven Architecture (MDA), such as Computational Independent Model (CIM), Platform Independent Model (PIM) and Platform Specific Model (PSM) [1]. In the railway domain, different tools are used for each phase of the development process from the industrial needs specification to the prototype implementation. These tools does not generally allow a digital continuity in the information, model or data life-cycle. In order to ensure models traceability and to facilitate the communication between stakeholders, Model-Based System Engineering (MBSE) approaches are adopted in critical systems development, like railway systems [2]. These approaches allow a robust documentation of models and improve their reuse in different contexts. Nevertheless, MBSE does not allow to deal with semantic heterogeneity and to share a common terminology during the whole system development process. From this context, the following Research Question (RQ) is identified: How to ensure semantic interoperability in such multidisciplinary ecosystem, like the railway domain?

In this paper, we introduce a MBSE approach which is able to help decision making and to have an efficient collaboration between actors. This approach is based on the integration of ontology as a structured and shared representation which is maintainable during the system life-cycle [3]. This knowledge representation has been widely used in order to disambiguate different domains. In the next section, the articulation between knowledge engineering and system engineering is introduced with the aim to meet railway issues.

2 The Proposed Approach

The proposed approach is illustrated by Fig. 1. It is based on three domains of interest:

1. Knowledge engineering,
2. System engineering and
3. Deployment.

This approach is structured in a top-down way from the abstract level which represents the need's specification until the implementation level. Knowledge structuring using ontologies is proposed as a crucial step in the system development, especially in complex operational activities related to system performance evaluation. Indeed, an ontology is defined as "a formal, explicit specification of a shared conceptualisation" [4]. It is considered as the basis of the whole process since it disambiguates the domain terminology and provides a common and shared representation between all stakeholders.

A domain ontology is formalised using the Web Ontology Language (OWL) and represents CIM level which is crucial to represent the domain view. In order to fulfil the critical needs of railway systems and to federate the railway data models representing heterogeneous subsystems of the railway infrastructure, an upper ontology, Unified Foundational Ontology (UFO) [5], is used. Upper ontologies provide foundational concepts and relations and their definition, such as space, process, situation, etc. [6]. Domain ontologies that are grounding in upper ontologies are more reusable since this practice facilitates the semantic alignment between heterogeneous knowledge domains. Conceptual models are considered as lightweight ontologies represented in Unified Modeling

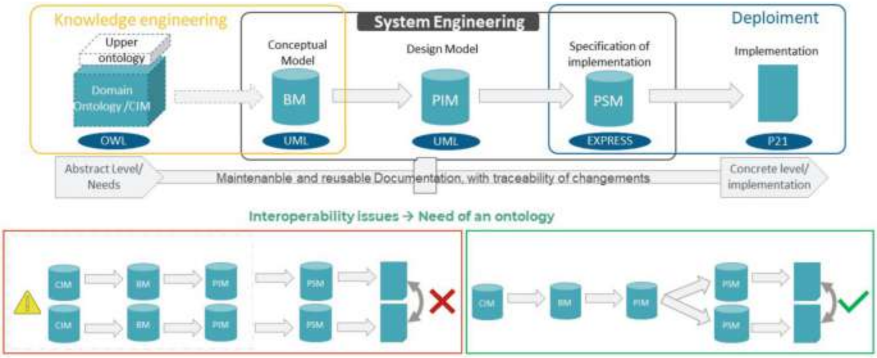

Fig. 1. The proposed MBSE approach for railway digital continuity

Language (UML) [7] and are a common part of both knowledge and system engineering. They are also used to be Business Models (BM) which represent the most pertinent concepts of the domain.

Using MBSE foundations, model-driven approaches and related standards, the conceptual model is then extended and refined in design model (PIM) to represent system features and attributes. Design models include several system properties as design constraints in the system architecture models for example. Therefore, these models are transformed into PSM to provide models in a specific language/format. At this level, the implementation/serialisation is performed to have a concrete prototype.

The whole process is supported by a documentation process which is required to maintain and manage changes traceability during the system design. Furthermore, the key aspect of this approach is the digital continuity between all system design phases since the ontology is the essence of the common semantic interpretation. This methodology is as generic as possible to ensure semantic interoperability in the design of the railway system assets. In the next section, the validation of this MBSE approach in the railway domain is presented.

3 The Approach Application and Validation

In previous work, an ontological approach was proposed with the aim to integrate dysfunctional analysis into the system design process of railway systems [8]. Then, this approach is illustrated and validated by real railway accidents. The key contribution of this approach is to ensure a semantic link between the safety measures and safety requirements in the railway system. This approach helps the safety-decisions making process and enhances the collaboration and communication.

Several initiatives have been developed as railway reference models that represent the infrastructure, such as Rail System Model (RSM) [9], IFC Rail[1] and EULYNX[2] that are built independently in UML. However, there is a digital discontinuity and a lack of

[1] https://www.buildingsmart.org/standards/rooms/railway/ifc-rail-project/.
[2] https://www.eulynx.eu/.

interoperability between these standards. The proposed MBSE approach is successfully used to deal with this semantic heterogeneity and to ensure digital continuity between the railway ontology and the system architecture. The developed railway domain ontology aims to avoid semantic ambiguities of railway concepts, such as Track, Signaling, Energy and Telecom and relations between them [10]. Then, this railway ontology is used to help decisions making in maintenance activities.

In order to ensure the digital continuity in the development of new systems with high performance, the proposed approach is validated and applied in the autonomous train development, namely the Autonomous Train Map (ATM) subsystem [11]. An ontology was developed in order to represent the ATM concepts and relations. Then, conceptual and design models were built, transformed in PSM and then serialised in code to be integrated in the final system.

4 Conclusions and Perspectives

The proposed top-down approach has powerful and generic capabilities to be applied in different fields and use-cases for the whole life-cycle phases from the design to operation and maintenance of systems. In the railway domain, this approach is used to ensure digital continuity. Indeed, the ontology layer is integrated as a knowledge basis to establish the semantic interoperability between railway actors. This MBSE approach improves collaboration and helps decisions-making process during the system life-cycle. It has been widely used and validated in the autonomous train development process and to ensure digital continuity in the railway domain.

In future work, we intend to adapt this approach for the development of railway digital twins (DT) with the integration of a digital representation layer, such as Building Information Modelling (BIM) [12]. Therefore, the integration of ontologies, Machine Learning (ML) algorithms and BIM in the DT development for the railway domain should meet scientific and industrial challenges related to data interoperability and integrity regarding the multiplicity of data sources. Finally, we plan to validate it by real use cases of predictive maintenance of track components using its BIM models.

References

1. Castellanos, C., P´erez, B., Correal, D., Varela, C.A.: A model driven architectural design method for big data analytics applications. In: 2020 IEEE International Conference on Software Architecture Companion (ICSA-C), p. 8994. IEEE (2020)
2. Henderson, K., Salado, A.: Value and benefits of model-based systems engineering (MBSE): evidence from the literature. Syst. Eng. **24**(1), 5166 (2021)
3. Hagedorn, T.J., Smith, B., Krishnamurty, S., Grosse, I.: Interoperability of disparate engineering domain ontologies using basic formal ontology. J. Eng. Des. **30**(10–12), 625–654 (2019)
4. Studer, R., Benjamins, V.R., Fensel, D.: Knowledge engineering: principles and methods. Data Knowl. Eng. **25**(1–2), 161–197 (1998)
5. Guizzardi, G., Botti Benevides, A., Fonseca, C.M., Porello, D., Almeida, J.P.A., Prince Sales, T.: UFO: Unified foundational ontology. Appl. Ontol. **17**(1), 167–210 (2022)

6. Borgo, S., Galton, A., Kutz, O.: Foundational ontologies in action. Appl. Ontol. **17**(1), 1–16 (2022)
7. Koc, H., Erdogan, A.M., Barjakly, Y., Peker, S.: UML diagrams in software engineering research: a systematic literature review. In: Proceedings, vol. 74, no. 1, p. 13. MDPI (2021)
8. Debbech, S., Dutilleul, S.C., Bon, P.: An ontological approach to support dysfunctional analysis for railway systems design. J. Univers. Comput. Sci. **26**(5), 549–582 (2020)
9. Ciszewski, T., Nowakowski, W., Chrzan, M.: RailTopoModel and RailML–data exchange standards in railway sector. Arch. Transp. Syst. Telematics **10** (2017)
10. Rahmig, C., Debbech, S., Errandonea, I., Arrizabalaga, S.: Ontologybased Conceptual model development for the railway domain: a maintenance case study. In: 4th SmartRaCon Workshop, San Sebastian (2022)
11. Chouchani, N., Debbech, S., Perin, M.: Model-based safety engineering for autonomous train map. J. Syst. Softw. **183**, 111082 (2022)
12. Heaton, J., Parlikad, A.K., Schooling, J.: Design and development of BIM models to support operations and maintenance. Comput. Ind. **111**, 172–186 (2019)

Effect of Different Weather Elements on the Delay Prediction of Trains

Pranjal Mandhaniya(✉), Nils O. E. Olsson, Anders S. Larsen, and Caroline Skjøren

Norwegian University of Science and Technology, 7034 Trondheim, Norway
pranjalmandhaniya@gmail.com, pranjal.mandhaniya@ntnu.no

Abstract. Estimation of train delays is crucial for customer information. One cause of the train delays can be easily blamed on the weather. The effect of weather on signaling and dispatching can be indirectly articulated from the arrival and departure delays at the stations. The study uses Norwegian delay data from 2021, 2022 and parts of 2023. This data contains scheduled and actual departure and arrival times of trains on the Dovre line between Oslo and Trondheim. This article talks about acquiring freely available weather data using APIs at the stopping station and checking the effect of weather elements on the departure delay. Weather elements correlated with the departure delays were rainfall (precipitation) and temperature. This study attempts to articulate the quantitative nature of the effect of these weather elements on the departure delays of the trains. The delay prediction model uses different neural network algorithms. The prediction results from different algorithms are compared to provide a deeper insight into the effect of weather on the delay characteristics.

Keywords: Delay prediction · railway operations · neural networks · weather effects · predictive modeling · transportation management

1 Introduction

Due to the unpredictability of weather conditions, transportation networks are experiencing unprecedented challenges. The railway system is thought to be the least weather-sensitive of the transport modes, often expected to run when other forms of transport are disrupted [1]. However, weather conditions can influence the railway traffic and cause delays. Additionally, extreme weather can affect the train operator's performance directly or indirectly through disturbances in the railway infrastructure.

It is hard to identify the best implementation options to lessen the negative effects of disruptions. However, it is widely acknowledged that transportation planning and management organizations must increase system resilience to minimize losses and disruptions [2]. This is partly caused by an inadequate knowledge of how various weather phenomena and specific extreme weather occurrences affect the functionality of transportation systems. Uncertainty exists over

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 330–337, 2026.
https://doi.org/10.1007/978-3-032-06763-0_48

how effects vary across various event types and geographically and temporally for various means of transportation within the same event. Therefore, providing a comprehensive picture of their effects is crucial, mapping all essential steps toward increased network robustness and resilience and finding chances for mode substitution [3].

As stated by [4], weather and climate are frequently treated as synonyms even though they differ. Weather is the short-term state of the atmosphere at a specific time and place. On the other hand, climate is the long-term manifestations of weather and other atmospheric conditions in a given area during a period long enough to ensure that representative values are obtained. Weather refers to the daily variation in the atmosphere, including the temperature, relative humidity, cloud cover, precipitation, and wind. In contrast, climate refers to the general conditions over a long period in a given location. More studies on climate change and its impact on the transport sector have been published in the 21st century, hence the railway system [5].

There are several ways to examine the impact of the weather conditions on the transport mode, according to [2]. The differences in performance can compare transport systems between regions with very different weather conditions. According to [4], the weather influences individual travel demand. Travelers have availability to information on the exact weather conditions and advanced weather forecasts, which seems to have an impact on their choice of transportation mode.

Each railway system has its particular challenges related to the local weather. While the cold weather is unfamiliar in the southern nations, it isn't easy in the Northern countries. The weather-related issues can be solved in a variety of ways. The British railway system has its own *Leaf Fall Timetable Changes* [6]. Thousands of tonnes of leaves fall across the railway network during the autumn. The leaves on the track are compressed by passing trains, producing a thin, black coating that affects train braking and acceleration. Timetables are modified throughout the autumn to prevent unforeseen delays to journeys.

There has been some recent research on how weather conditions affect disturbance in the railway industry. However, the examination of published literature has a heavier emphasis on other modes of transportation [7,8]. Yet, as [9] pointed out, the literature has begun to pay more and more attention to the need to comprehend how weather and climate change affect railway delays and punctuality.

Including weather variables in the train delay prediction models has been shown to improve the overall accuracy of the models. [10] developed a dynamic train delay prediction system that uses state-of-the-art tools and techniques to combine heterogeneous data sources and interact with dynamically changing systems to provide meaningful information to traffic management and dispatching operations. These models are further enhanced by including exogenous meteorological data [10,11]. [12] attempted to comprehend the trends in weather and railway delays. According to the study, train delays in severe weather are most influenced by poor weather. Still, delays in fair weather are primarily influenced by the duration and frequency of previous train delays. Further, [12] found that infrequently occurring weather, such as snow in southern cities, has a stronger influence and causes longer train delays.

The impacts of weather phenomena, including wind, temperature, and precipitation, on railway operators' performance of passenger train services are estimated by [1]. The study concluded that 4–8% of all train disturbances were caused by bad weather conditions. [13] discovered that trains frequently display the same delay patterns as in the past and that adverse weather frequently results in longer delays. [14] found that 60 % of late arrivals in Finland between 2008 and 2010 were connected to winter weather by modeling the co-variation between harsh weather and freight train delays. The weather factors that influence the punctuality of trains on the Norwegian railway Nordland Line were examined by [8]. The study demonstrates that extreme winter cold is a significant influencing factor that leads to delays and poor punctuality. Moreover, snow depth is the meteorological factor that most explains changes in passenger train punctuality daily and weekly. [15] investigated the impact of weather on railway punctuality. As the temperature drops below 0°, punctuality reduces exponentially; at –5°, it drops by 7.5%, and at –30°, it drops by 50%. The temperature variation was highly correlated with the travel distance. [16] introduced a train delay prediction model that included a fully connected neural network with two LSTM components. The study highlighted the importance of considering the interactions between trains, stations and weather-related factors regarding prediction accuracy.

The present study tries to provide data-driven insights into delay modeling by including weather elements. This research specifically targets the existing delays and the effect of different weather elements on the departure delays of the trains.

This research is a part of an EU-funded project "FutuRe". The flagship project FP6 – FutuRe (GA 101101962) [17] under Europe's Rail Joint Undertaking) aims at providing new innovative technical requirements, methods, solutions, developments and services based on the latest leading-edge technologies to make regional rail cost-efficient while meeting safety standards and improving the reliability, availability and capacity of the railway system. The work presented here is related to the FutuRe project area Regional Rail Customer Services, focusing on customer service and aiming to develop highly accurate multimodal passenger information on-board and/or at stations for passenger and freight management.

2 Methodology

The data was acquired from Bane NOR for the arrival and departure delays of long-distance trains running on the Oslo-Trondheim line from 1 January 2021 to 28 February 2023 at the stopping and non-stopping stations. The data was acquired from different sources into TIOS database as shown in Fig. 1. The data includes trains starting from Oslo and trains starting from Trondheim. Machine learning model, specifically the long short-term memory (LSTM) algorithm of the neural network domain, was used for prediction modeling. Three weather features (temperature and precipitation) were acquired from the Norwegian Meteorological Institute for the corresponding dates. Correlations were established between the effects of different aspects of weather. These results were compared to the predictions without considering the weather effect. The data was cleaned and analyzed using Python modules.

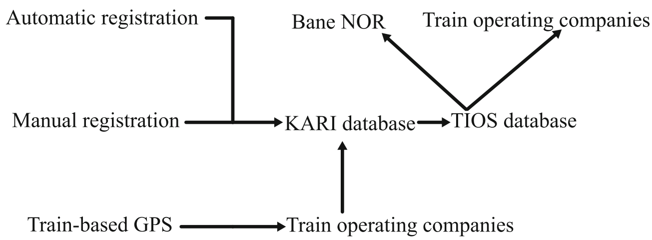

Fig. 1. Illustration of the data flow of the train record information (reproduced from [18])

The delay data was selected between the interquartile range (IQR), a statistical measure of the spread of a data set, as shown in Eq. 1. It is calculated as the difference between the third (Q3) and the first quartile (Q1) of the data distribution, which is the range between the 25th and 75th percentiles of the data. The IQR method for identifying outliers is used by setting up a "fence" outside of Q1 and Q3, or minimum and maximum values. Values outside of these boundaries are considered outliers.

$$lower\_boundary = Q1 - 1.5 \cdot IQR; upper\_boundary = Q3 + 1.5 \cdot IQR \tag{1}$$

The delays, temperature, and rainfall values (parameters) were graded using a formula in Eq. 2 to derive additional features for the prediction model. The value of variable *grad* ranges from 1 to 4 while the values v1, v2 and v3 are changed as per the parameter.

$$Parameter = \begin{cases} < v1 & \text{grad} = 1 \\ \geq v1 < v2 & \text{grad} = 2 \\ \geq v2 \leq v3 & \text{grad} = 3 \\ > v3 & \text{grad} = 4 \end{cases} \tag{2}$$

There were negative delays at some non-boarding stations in the dataset, which were ignored in this analysis. The Entur Geocoder API was used for getting the latitude and longitude of the stations to connect them with the nearest weather stations. The FrostMET API was used to get historical data on temperature and rainfall from weather stations nearest to the train stations.

Entur Geocoder API is free to use and requires no credentials. However, bulk requests are blocked by the server. The application of FrostMET API requires duly formed credentials, which can be overcome by web scraping. However, there are some unresolved issues regarding bulk requests and multiple weather elements. These issues will be resolved in the future research.

3 Analysis and Results

The effect of gradation as a modeling feature is shown in the form of Pearson coefficient values (R^2), mean absolute error (MAE, in minutes) and root mean square error (RMSE, in minutes) of the LSTM prediction model (as per Fig. 2) in Table 1.

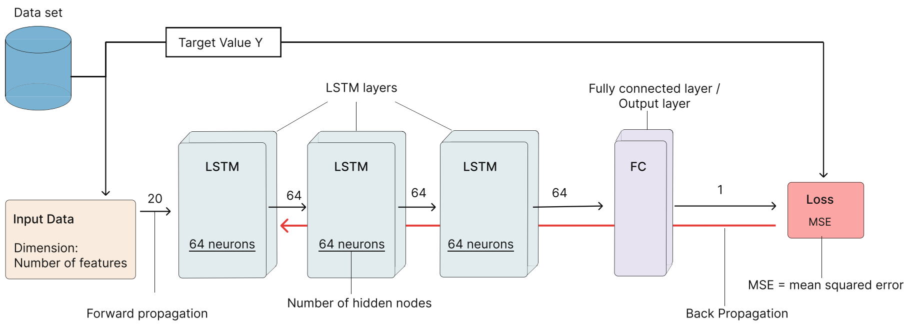

Fig. 2. LSTM model architecture (adopted from [18])

The general gradation of delay values was present in all the datasets by default unless stated in the table. It can be seen that gradation increases the R^2 values while keeping the error values in check. It can be seen that without gradation, the weather data correlation becomes much weaker [19].

Table 1. Statistical inferences from different prediction models

Models	MAE	RMSE	R^2
Weather without gradation (delay gradation absent)	2.0231	2.5375	0.0109
temperature and rainfall data without gradation	0.9358	1.4162	0.6919
temperature and rainfall data with gradation	0.9813	1.2976	0.7414
Only temperature data with gradation	0.9528	1.2730	0.7500
No weather data at all	0.9712	1.3205	0.7312

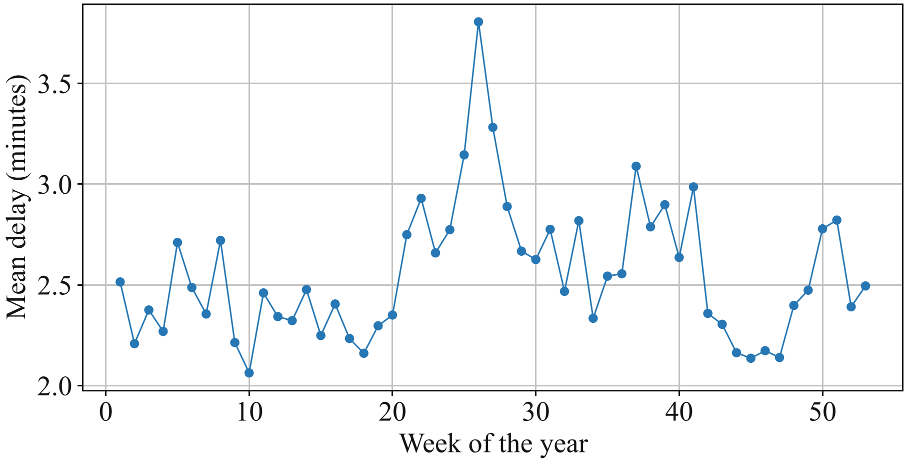

Fig. 3. Delay averaged over the weeks of the year (1 is the first week and 53 is the last week on the x-axis)

The delay averaged over the weeks of the year is shown in Fig. 3. The delays can be related to two factors, i.e., congestion-based delays due to increased demand and weather-based delays. The peak delay is in the middle of the year, specifically the 26th week. This coincides with the peak of the holiday season in Norway, and people travel during this time. The second-highest peak is at 37th week, around the arrival of the fall season in Norway. This requires more maintenance at the railway tracks, thus causing delays on single-line tracks.

4 Conclusions

The major conclusion that can be drawn from the study and the literature studied in the present study are:

1. The application of weather elements to the prediction model surely increases the quality of the results. However, clustering and gradation as a feature are more helpful in bolstering their application.
2. The peak delays can be related to two major factors: congestion due to increased demand and maintenance due to extreme weather.

Acknowledgements. We are delighted to acknowledge the help of Bane NOR for providing the delay data for this study.

Funding. This research activity is part of the FP6 FutuRe project, partially funded by the European Commission through Europe's Rail Joint Undertaking under the Horizon Europe Programme with the grant agreement number 101101962.

Data Availability Statement. The data used in the present research is made available on https://github.com/pranjalm/Future_Europes_rail.

References

1. Xia, Y., Ommeren, J.N., Rietveld, P., Verhagen, W.: Railway infrastructure disturbances and train operator performance: the role of weather. Transport. Res. Part D: Transp. Environ. **18**, 97–102 (2013). https://doi.org/10.1016/j.trd.2012.09.008
2. Koetse, M.J., Rietveld, P.: The impact of climate change and weather on transport: an overview of empirical findings. Transport. Res. Part D: Transp. Environ. **14**(3), 205–221 (2009). https://doi.org/10.1016/j.trd.2008.12.004
3. Stamos, I., Mitsakis, E., Salanova, J.M., Aifadopoulou, G.: Impact assessment of extreme weather events on transport networks: a data-driven approach. Transport. Res. Part D: Transp. Environ. **34**, 168–178 (2015). https://doi.org/10.1016/j.trd.2014.11.002
4. Sabir, M.: Weather and travel behaviour. PhD thesis, Vrije Universiteit Amsterdam (2011). https://research.vu.nl/ws/portalfiles/portal/42203407/complete+dissertation.pdf
5. Baker, C.J., Chapman, L., Quinn, A., Dobney, K.: Climate change and the railway industry: a review. Proc. Inst. Mech. Engineers Part C: J. Mech. Eng. Sci. **224**(3), 519–528 (2010). https://doi.org/10.1243/09544062JMES1558
6. National Rail Enquiries. Leaf fall timetable changes - autumn (2022). https://www.nationalrail.co.uk/255905.aspx
7. Nagy, E., Csiszar, C.: Analysis of delay causes in railway passenger transportation. Period. Polytech. Transp. Eng. **43**(2), 73–80 (2015). https://doi.org/10.3311/PPtr.7539
8. Zakeri, G., Olsson, N.O.E.: Investigating the effect of weather on punctuality of Norwegian railways: a case study of the nordland line. J. Mod. Transport. **26**(4), 255–267 (2018). https://doi.org/10.1007/s40534-018-0169-7
9. Palmqvist, C.W.: Delays and timetabling for passenger trains. PhD thesis, Lund University (2019). https://portal.research.lu.se/files/70626078/Carl_William_Palmqvist_web.pdf
10. Oneto, L., et al.: Advanced analytics for train delay prediction systems by including exogenous weather data. In: 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), pp. 458–467 (2016). https://doi.org/10.1109/DSAA.2016.57
11. Oneto, L., et al.: Dynamic delay predictions for large-scale railway networks: deep and shallow extreme learning machines tuned via thresholdout. IEEE Trans. Syst. Man Cybern. Syst. **47**(10), 2754–2767 (2017). https://doi.org/10.1109/TSMC.2017.2693209
12. Wang, P., Zhang, Q.: Train delay analysis and prediction based on big data fusion. Transport. Saf. Environ. **1**(1), 79–88 (2019). https://doi.org/10.1093/tse/tdy001
13. Ling, X., Peng, Y., Sun, S., Li, P., Wang, P.: Uncovering correlation between train delay and train exposure to bad weather. Physica A: Stat. Mech. Appl. **512**, 1152–1159 (2018). https://doi.org/10.1016/j.physa.2018.07.057
14. Ludvigsen, J., Klæboe, R.: Extreme weather impacts on freight railways in Europe. Nat. Hazards **70**(1), 767–787 (2013). https://doi.org/10.1007/s11069-013-0851-3
15. Palmqvist, C.W., Olsson, N.O.E., Hiselius, L.W.: Some influencing factors for passenger train punctuality in Sweden. Int. J. Prognost. Health Manag. **8**(3), 1–14 (2017). https://doi.org/10.36001/ijphm.2017.v8i3.2649
16. Huang, P., et al.: Modeling train operation as sequences: a study of delay prediction with operation and weather data. Transport. Res. Part E: Logist. Transport. Rev. **141**, 102022 (2020). https://doi.org/10.1016/j.tre.2020.102022

17. Europe's Rail. Fp6 future homepage (2022). https://projects.rail-research.europa.eu/eurail-fp6/
18. Skjøren, C., Larsen, A.S.: Predictions for railway traveller information. Master's thesis, Norwegian University of Science and Technology (2023). https://hdl.handle.net/11250/3088477
19. Mandhaniya, P., Olsson, N.O.E., Larsen, A.S., Skjøren, C.: How complex systems sometimes follow murphy's law: train delay prediction at a station using delays at previous stops as the features. Procedia Comput. Sci. **239**, 1778–1783 (2024). https://doi.org/10.1016/j.procs.2024.06.357

An Analysis of Dock-Less Bike Sharing Service in Dublin, Ireland

Mounisai Siddartha Middela[2], Laura Bennett[1], Vikram Pakrashi[1], and Bidisha Ghosh[2](✉)

[1] University College Dublin, Dublin, Ireland
[2] Trinity College Dublin, Dublin, Ireland
bghosh@tcd.ie

Abstract. Cycling promotes a healthier lifestyle and alleviates traffic congestion and air pollution. The present study focuses on the dockless bike sharing system in Dublin, Ireland and identifies the socio-economic and built environment factors that affect the dwell time. The origin- destination and timestamp databases obtained from Bleeper Bike are used to estimate the dwell time. The data is analysed using Uber H3 hexagonal zones, and non-spatial and spatial regression models (both spatial lag and spatial error models) are developed. The spatial error model was found to provide a better fit to dwell time. The study identified that key factors such as the presence of public transport stations and car ownership impact dwell time significantly.

Keywords: Bike sharing · dwell time · spatial regression models

1 Introduction

Cycling is an environmentally conscious form of transport that promotes a healthier lifestyle and helps alleviate traffic congestion and air pollution [1, 2]. In the past two decades, bike-sharing systems, which allow for flexible and environmentally friendly travel, have been gaining popularity in many cities worldwide. Recently, dockless bike sharing systems that allow users to park their bike anywhere is growing across the world. Bleeper bikes [3], is the first to introduce dockless bike sharing systems in Dublin, Ireland.

Bike sharing research areas include identifying factors that influence the demand, analysing the use of bike sharing schemes to solve the last mile problem, and repositioning of bikes [4, 5]. Most earlier studies focused on docked bike sharing systems to identify the factors influencing their demand [6, 7]. These studies have found that socio-economic characteristics, built environment factors, and weather affect the demand. Recently, dockless bike sharing systems have been introduced all over the world. The analysis of these systems and identification of factors affecting their demand is still in the nascent stage, and more studies are essential.

Earlier studies have employed various regression models. These range from Bayesian structural time series model, geographically weighted regression, autoregressive model

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 338–344, 2026.
https://doi.org/10.1007/978-3-032-06763-0_49

with exogenous variables, spatial lag model, graph regularisation technique, negative binomial regression, machine learning models, and semi-parametric methods [8]. All the above studies have modelled bike sharing trip counts. None of the studies have analysed the effect of built environment factors on dwell time, particularly in Ireland.

In this study, we have developed regression models for dwell time by spatial zones. To evaluate the presence of and consider spatial dependence in dwell time, we developed spatial lag and error models and compared them with the non- spatial model. The present study contributes to the literature by modelling dwell time, in contrast to earlier studies modelling bike trip counts. The study findings reveal various factors that affect dwell time and provide initial evidence of the need for repositioning. The findings provide valuable information that could be used to understand and improve the bike sharing systems.

The rest of the paper is organised as follows. Section 2 presents the data and descriptive statistics. Sections 3 and 4 provide details on the methodology employed and the study results. Finally, Sect. 5 presents the study conclusions.

2 Data and Descriptive Statistics

The data utilized in this study has been sourced from Bleeper, which is the first dockless bike-sharing system established in Dublin, Ireland. This system operates within a defined service area known as the "purple zone", which is 100 km^2. Users can both access and return bikes only within this zone. The origin and destination (OD) database and the timestamp database are used in this study. The timestamp database contains the location data of the bike every 5 minutes. Since these dockless bikes can be parked anywhere, for the purpose of the analysis, the study area was divided into zones using Uber H3 hexagons [9] (Fig. 1). The resolution used is 9, implying each hexagon has an area of 0.10533 km^2.

For each zone, dwell time was calculated as the time between the destination of journey A and the time of the origin of journey B, where journeys A and B are consecutive trips made by the same bike. If the time at destination A and the time at origin B occur on different days, the dwell time was readjusted to account for the five hours (12.00 am to 5.00 am) that the bikes cannot be unlocked. In the present study, dwell times exhibited considerable variability, ranging from 0.26 to 46,234 h, with a mean value of 4,132 hours and a standard deviation of 6,248 h for the H3 zones. This is because of the variation in bike supply and usage by H3 zone. The number of bike trip origins from each zone ranged from 0 to 290. Figure 1 compares bike trips and dwell time. A few H3 zones with more bike trips have shorter dwell times and vice versa.

We obtained the data on the socio-demographic characteristics and built environment from multiple sources. The socio-demographic characteristics such as the number of households, population by age and ethnicity, employment by sector, and car ownership were obtained from the Central Statistics Office (CSO) public datasets [10]. This information is available at a resolution of 20 m^2 as small area population statistics and was remapped to estimate different variables for the H3 zones. The built environment variables were counted as the number of bus stops, bike stations, bars and restaurants, and others using their geolocations obtained from OpenStreetMaps.

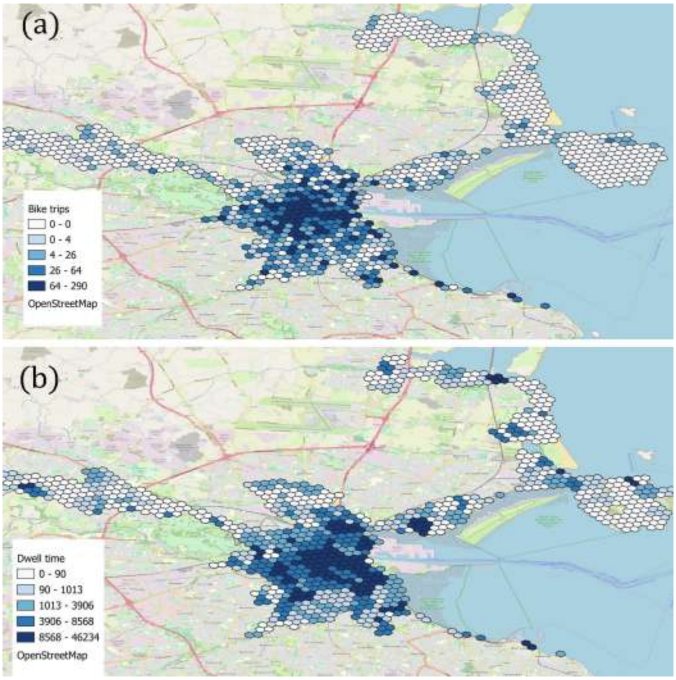

Fig. 1. Study area showing the (a) bike trips and (b) dwell time (hr) variation with resolution 9 Uber hexagons. (Basemap source: Google Maps™).

3 Methods

In the present study, non-spatial linear regression models (NS), spatial lag models (SLM) and spatial error models (SEM) are developed for dwell time. The NS model is estimated using the ordinary least squares estimator. Spatial models are considered to capture spatial dependence and understand the neighbour inter- actions, utilising a spatial weights matrix. The spatial weights matrix chosen is the 6-nearest neighbour matrix, similar to [11]. This matrix differs from the contiguity weights matrix only in the case of boundary zones. SLM and SEM are estimated using the maximum likelihood estimator (Fig. 2).

3.1 Spatial Lag Models

Spatial Lag Models (SLM) capture the spatial correlation in each cross-sectional unit's dependent variable as a spatially weighted average of dependent variables of other cross-sectional units. The weighted average is the spatial lag ($\boldsymbol{W}\,\boldsymbol{Y}$), with its parameter known as the spatial lag parameter ρ. While $\boldsymbol{X}$ is a matrix of independent variables, $\boldsymbol{\beta}$ is a vector of parameters to be estimated. The corresponding reduced-form version of Eq. (1), given by Eq. (2), reveals a global range of spillovers, with $\boldsymbol{A^{-1} = (I - \rho W)^{-1}}$ being the spatial multiplier or the Leontief inverse [12].

$$Y = \rho WY + X\beta + \varepsilon \tag{1}$$

Variable	Min	Max	Mean	SD
Dwell time (hr)	0.26	46234	4132	6248
Bike origins (#)	0	290	25	43
Socio-demographic characteristics				
# households	51	1560	590	275
Population by age - proportion				
Age < 10	0.00	0.28	0.11	0.04
Age < 20	0.01	0.57	0.09	0.05
Age 20-25	0.02	0.52	0.08	0.05
Age 25-30	0.03	0.28	0.10	0.05
Age 30-35	0.01	0.25	0.10	0.05
Age 35-40	0.01	0.18	0.08	0.03
Age 40-45	0.00	0.16	0.07	0.02
Age 45-50	0.00	0.13	0.06	0.01
Age 50-55	0.00	0.14	0.06	0.02
Age 55-60	0.00	0.12	0.06	0.02
Age 60-65	0.00	0.14	0.05	0.02
Age > 65	0.00	0.55	0.15	0.08
Age > 50	0.00	0.71	0.32	0.11
Population by ethnicity				
White Irish	32	2771	1090.71	484.82
WI Traveller	0	191	3.53	12.63
Other whites	7	1336	216.32	196.62
Black Irish	0	229	23.07	31.09
Asian Irish	0	626	65.3	77.17
Other race	0	412	45.15	46.49

Variable	Min	Max	Mean	SD
Number of people employed by sector				
AFFT	0	12	1.5	1.7
BCT	0	85	21.6	13.8
MIT	0	144	41.7	22.4
CTT	1	755	231.9	129.1
TCT	0	486	103.5	70.0
PAT	1	166	44.0	25.8
PST	1	513	180.4	93.9
Others	2	1574	197.6	205.0
Car ownership proportion in total households				
0 cars	0.01	0.8	0.25	0.19
1 car	0.08	0.65	0.39	0.09
2 cars	0.01	0.58	0.24	0.13
3 cars	0	0.14	0.04	0.03
>3 cars	0	0.27	0.06	0.04
Built environment				
# Bus stops	0	31	2.05	2.87
# Bike stations	0	19	1.01	2.18
# Bars and restaurants	0	94	2.4	6.73
# Entertainment and cinema	0.00	5	0.16	0.58
# Educational institutions	0.00	5	0.16	0.53
# Financial institutions	0	9	0.29	0.92
# Health centres	0	8	0.47	1.1
# Government institutions	0	8	0.61	1.16
# Luas stops	0	1	0.03	0.17
# Parks	0	4	0.05	0.3
# Rail stations	0	1	0.03	0.18
# Schools	0	2	0.02	0.15
# PT stops	0	31	2.11	2.93

Note: AFFT - Agriculture, forestry and fishing; BCT - Building and construction; MIT - Manufacturing industries; CTT - Commerce and trade; TCT - Transport and communications; PAT - Public administration; PST - Professional services

Fig. 2. Descriptive statistics of the dependent and independent variables.

$$Y = A^{-1}X\beta + A^{-1}\varepsilon \tag{2}$$

where, ***Y*** is the dependent variable. ***I*** is an identity matrix. ***W*** is the spatial weights matrix. ε is normally distributed with mean 0 and variance σ^2.

$$Y = X\beta + u \tag{3}$$

where,

$$u = \lambda Wu + \varepsilon \tag{4}$$

4 Results

The effect of neighbourhood characteristics on dwell time are presented in this section. Table 1 presents the different models developed in this study. The R2 and AIC of the models are also shown in the table. The best model is the SEM model in terms of AIC. The spatial dependence is present in both the dependent variable (SLM) and the residuals (SEM). In both the spatial and non-spatial models, the socio-demographic characteristics such as population by age and ethnicity, employment by sector, and car ownership are found to affect the dwell time. Similarly, built environment variables also affect the dwell time.

A few variables were found to be insignificant in each of the models. Dwell time did not differ for the populations aged 25–30 and 35–40 in the SLM model. Further, the proportion of white Irish employed in the building and construction, commerce and trade sectors and the proportion of households with either 1 or 2 cars are found to be not significant. The number of Luas stations also did not affect the dwell time of a zone. Similar results are also observed in the case of SEM, except for the number of households with either 1 or 2 cars, which are found to influence the dwell time.

The SEM model is only interpreted here for the sake of brevity. The results show that zones with a higher proportion of households owning either 1 or 2 cars have shorter dwell times. This is perhaps because these zones have higher usage of bikes mixed with a lower supply. Further, zones with more bus stops and rail stations tend to experience longer dwell times, perhaps because most bike trips in these zones are for park-and-ride trips that result in higher dwell times. Additionally, zones with more bike stations, bars and restaurants, and financial institutions have higher dwell time. On the other hand, zones with more educational institutions have shorter dwell times. Further, zones with a higher proportion of black Irish population aged 55–60 have shorter dwell times. This is perhaps because the younger population and black Irish tend to rely more on bikes for their mobility than other population groups.

Table 1. Non-spatial and spatial regression model results for dwell time

Variable	NS	SLM	SEM
(Intercept) *Population by age* Age 25–30	9881.00 (6.35) 1183.91 (1.74) 6249.44 (5.20) −15460.00 (2.30)		
Age 35–40	−15400.00 (1.62)		
Age 55–60	−33400 (3.36) −26637.39 (2.96) −28209.33 (2.62)		
Population by ethnicity WI	−2.41 (3.80)		
BI *Employment by sector* BCT	−32.88 (4.58) 30.26 (1.50)	−23.76 (4.55)	−27.90 (4.46)
CTT	8.29 (2.56)		

(*continued*)

Table 1. (*continued*)

Variable	NS	SLM	SEM
TCT *Car ownership* No cars	18.26 (3.63)	15.95 (6.40) 2187.65 (1.88)	20.59 (7.25)
1 car	−4906.00 (2.21)		−7101.25 (3.36)
2 cars *Built environment* Bus stops	−8817.00 (4.49) 207.20 (3.25)	178.92 (2.90)	−4591.75 (2.68) 165.70 (2.63)
Bike stations	764.20 (6.75)	787.58 (7.19)	844.99 (7.69)
Bars and restaurants	240.40 (6.50)	213.53 (5.92)	247.14 (6.73)
Educational institutions	−797.90 (2.23)	−891.32 (2.58)	−996.18 (2.94)
Financial institutions	400.70 (1.71)	581.10 (2.60)	485.74 (2.21)
Luas stations	1239.00 (1.40)		
Parks	1227.00 (2.51)	1049.71 (2.21)	944.02 (2.00)
Rail stations	1615.00 (2.01)	1270.81 (1.63)	1148.50 (1.51)
Spatial parameters ρ	NA	0.23 (5.75)	NA
λ	NA	NA	0.29 (6.09)
No of observations *Model fit*		704	
R^2	0.64	NA	NA

5 Conclusions

The present study considers the spatial dependence in the bike dwell times by hexagonal zones. Our findings show the presence of spatial interactions in both the dependent variable and the residuals, and considering spatial effects produces a better model. The spatial error model provides the best fit. Notably, zones with a higher proportion of car ownership have a high bike usage conditional on the supply. Bike supply can be increased in these zones. Zones with more public transport stops experience higher dwell times due to a predominant share of park-and-ride trips. Further, the analysis shows specific locations exhibiting longer dwell times, indicating potential oversupply issues or a need for bike repositioning.

References

1. Fan, Y., Zheng, S.: Dockless bike sharing alleviates road congestion by com- plementing subway travel: evidence from Beijing. Cities **107**(102), 895 (2020)
2. Ma, X., Yuan, Y., Van Oort, N., Hoogendoorn, S.: Bike-sharing systems' impact on modal shift: a case study in delft, the netherlands. J. Clean. Prod. **259**(120), 846 (2020)
3. Bleeper. Bleeper Bike (2023). https://www.bleeperactive.com/
4. Zhang, Y., Thomas, T., Brussel, M., Van Maarseveen, M.: Exploring the impact of built environment factors on the use of public bikes at bike stations: case study in Zhongshan, China. J. Transp. Geogr. **58**, 59–70 (2017)
5. Usama, M., Shen, Y., Zahoor, O.: A free-floating bike repositioning problem with faulty bikes. Procedia Comput. Sci. **151**, 155–162 (2019)
6. El-Assi, W., Salah Mahmoud, M., Nurul Habib, K.: Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto. Transportation **44**, 589–613 (2017)
7. Ravensbergen, L., Buliung, R., Mendonca, M., Garg, N.: Biking to ride: Investigating the challenges and barriers of integrating cycling with regional rail transit. Transp. Res. Rec. **2672**(8), 374–383 (2018)
8. Venkadavarahan, M., Joji, M.S., Marisamynathan, S.: Development of spatial econometric models for estimating the bicycle sharing trip activity. Sustain. Cities Soc. **98**(104), 861 (2023)
9. Brodsky, I.: H3: Uber's hexagonal hierarchical spatial index (2019). https://eng.uber.com/h3/
10. Central Statistics Office. Census 2016 Small Area Population Statistics (2023). https://www.cso.ie/en/census/census2016reports/census2016smallareapopulationstatistics/
11. Middela, M.S., Ramadurai, G.: Incorporating spatial interactions in zero- inflated negative binomial models for freight trip generation. Transportation 1–22 (2021)
12. LeSage, J., Pace, R.K.: Introduction to Spatial Econometrics. Chapman and Hall/CRC, Boca Raton (2009)

Mobility Data Harmonisation: The TANGENT Solution

Marco Comerio(✉), Andrea Fiano, Marco Grassi, and Mario Scrocca

Cefriel, Milano, Italy
{marco.comerio,andrea.fiano,marco.grassi,
mario.scrocca}@cefriel.com

Abstract. Data interoperability is a challenging objective to enable different stakeholders to communicate and exchange information effectively and with unambiguous meaning. Indeed, stakeholders adopt different (legacy) systems for managing and exchanging data that cannot be directly integrated. The H2020 TANGENT project needs to face data interoperability issues to enable the development of innovative tools for multimodal transport network management and for optimising traffic operations. This paper describes the design and the development of the TANGENT solution for data harmonisation and fusion, realised to face heterogeneous interoperability issues related to data discovery, access, harmonisation, integration, and extraction.

Keywords: Mobility Data · Data Interoperability · Data Harmonisation

1 Introduction

The TANGENT project[1], funded by the European Commission under the Horizon 2020 Programme, is developing new complementary ICT tools for optimising traffic operations in a coordinated and dynamic way from a multimodal perspective and considering automated/non-automated vehicles, passengers, and freight transport.

Data coming from intermodal mobility (transport schedules, transport network, traffic flows, etc.), captured by sensors deployed to monitor the transport network (trains, road traffic, ferries, etc.) or generated by users and vehicles will feed the TANGENT tools to determine the traffic conditions and to deliver warning and recommendation services. One of the project's main objectives is gathering data sources related to the TANGENT case study cities (i.e., Athens, Lisbon, Greater Manchester, and Rennes Metropole) and enabling their usage in the TANGENT tools. The definition of a proper solution is not straightforward and should face the following challenges: (i) **discover** the data (*Which data are available and where?*), (ii) **access** the data: (*How to obtain access to the data?*), (iii) **harmonise** the data: (*How to convert the data according to a reference data model?*), (iv) **integrate** the data: (*How to merge data from different sources?*), *and* (v) **extract** the data: (*How to consume harmonised/integrated data?*).

[1] https://tangent-h2020.eu/.

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 345–350, 2026.
https://doi.org/10.1007/978-3-032-06763-0_50

The analysis of the proposed challenges considering the TANGENT case studies has highlighted several metadata and data interoperability issues to be addressed. As shown in Fig. 1, data are available in different data catalogues, i.e., web portals that describe and often host a set of data sources (datasets and data services). Considering different data catalogues, each of them may contain:

- datasets in different formats and/or using different data models (A, B, C);
- data services relying on heterogeneous specifications and technologies (I, II);
- data sources described according to different metadata specifications (d,e).

Identifying a single solution to the abovementioned challenges is impossible since a single interoperability problem cannot be formulated. Indeed, the identified data interoperability issues are widely heterogeneous and pose various requirements that can be possibly faced only by considering a set of tools appropriately configured [4].

Fig. 1. Metadata and data interoperability issues

The following sections describe the design and the TANGENT solution for data harmonisation and fusion, realised as a set of tools to address the various data interoperability requirements elicited by the project.

2 TANGENT Solution for Data Harmonisation and Fusion

This section describes the components of the TANGENT solution identified to face each of the mentioned challenges: discover, access, harmonise, integrate, and extract.

The European Commission is also addressing the first two challenges (discover and access) by implementing National Access Points (NAP) for transportation data. The concept of NAP leverages the one of Data Catalogue, i.e., a digital platform to facilitate the sharing of data sources and their findability by other stakeholders. However, even in the case of NAPs, each Member State adopted different approaches to implementing their NAP, thus creating interoperability issues at the European level. For this reason, the NAPCORE[2] project is currently working on coordinating and harmonising such platforms around Europe. One crucial objective is supporting the findability of data in each mobility data platform and defining a uniform mechanism to access the data sources. Adopting structured metadata descriptors according to well-known vocabularies (e.g., DCAT-AP [2]) is extremely important to facilitate the implementation of searching functionalities within one or multiple data catalogues. Moreover, proper data governance must be defined to regulate the usage of the catalogue between the different involved

[2] https://napcore.eu/.

stakeholders. Finally, data catalogues should support the harmonisation of technological modes to access the data sources and/or provide detailed documentation to assist the data user.

For these reasons, the first two components of the TANGENT solution for data harmonisation and fusion are:

- The **TANGENT Data Catalogue**: a digital platform (i) describing uniformly all the relevant data sources collected for the TANGENT case studies according to the TangentDCAT-AP metadata specification, and (ii) implementing governance mechanisms according to the TANGENT Data-Sharing Governance model. A complete description of the metadata specification and the governance model is available in [1].
- The **TANGENT Data API**: a uniform API to access data sources and define integrations with the overall TANGENT solution. The basic idea of the TANGENT Data API is to make it easier for the final user to access the data without locating and accessing several data portals. The user can access a data source by knowing the endpoint at which the TANGENT Data API is located and by retrieving the identifier of the data source to be accessed from the TANGENT Data Catalogue. The TANGENT Data API handles authorisation mechanisms for the different data sources. Moreover, in cases where a data source that is published online should be filtered according to the specificities of a case study, the TANGENT Data API is already configured for providing access only to the relevant data (e.g., adding parameters to filter data according to a defined temporal/geographical scope).

The remaining three data interoperability challenges (harmonise, integrate, and extract) require a flexible solution that can be configured to address heterogeneous requirements in terms of: (i) *schema and data transformation* to obtain syntactic (structural) and semantic interoperability and (ii) *integration with existing information systems* as data sources (i.e., components generating or storing the data) and/or data sinks (i.e., components consuming or archiving the data).

Different approaches can be exploited and implemented, spanning from ad-hoc solutions targeting a specific scenario to more general and scalable solutions supporting multiple stakeholders and data representations. Into the last approach falls the any-to-one mapping, which reduces the number of mappings, i.e., translations from one representation to another, needed to implement interoperability by different stakeholders [1, 4]. This approach selects a reference model for the domain of interest (e.g., the Transmodel[3]), and each stakeholder is responsible for defining mappings from their own data representation to the reference model and vice versa.

A specific application of any-to-one mappings relies on using Semantic Web technologies and defining the central model as a reference ontology. The main advantage of this approach is the definition of an interoperable and integrated Knowledge Graph, modelled according to a formal reference ontology, as an additional valuable product of the conversion procedure [4].

The implementation of the semantic any-to-one mapping approach requires the definition of two procedures: (i) a ***lifting*** procedure from the source data representation to

[3] https://www.transmodel-cen.eu/.

the reference ontology and (ii) a ***lowering*** procedure from the reference ontology to the target data representation.

Different approaches for lifting and lowering can be suitable considering a specific scenario. Moreover, the harmonisation, integration and extraction process may require the definition of custom pre- and post-processing to manage interoperability issues. Therefore, composing and configuring different components should be possible considering the specific requirements for integrating certain data sources.

Semantic-based ETL (Extract, Transform and Load) tools are proposed in the literature to address the problem of defining composable procedures considering Semantic Web technologies. Chimera[4] [2] is an open-source solution based on Apache Camel[5], enabling the definition of semantic data transformation pipelines by composing different components for knowledge graph construction, transformation, validation, and exploitation. The advantage of Chimera is that the integration with Apache Camel provides a wide set of production-ready components to integrate pipelines with heterogeneous systems. Moreover, custom components can be easily defined and integrated within a pipeline (e.g., to implement custom extraction/filtering approaches). For these reasons, Chimera has been selected to implement the semantic any-to-one mapping approach, considering the integration requirements of the TANGENT project.

Accordingly, two additional elements are identified to support the TANGENT solution for data harmonisation and fusion:

- The **TANGENT Reference Conceptual Model**: a model defining a reference ontology for the transportation-related data handled within TANGENT.
- The **Semantic Harmonisation and Fusion Pipelines**: a set of workflows defined with Chimera to fulfil the integration requirements associated with the TANGENT data sources.

3 A Prototype of the TANGENT Solution for Data Harmonisation and Fusion

This section discusses a set of illustrative semantic conversion pipelines developed within the project, considering real data sources from case studies and enabling the harmonisation of traffic data according to a uniform output format.

The following data sources from two different case study cities (i.e., Greater Manchester and Rennes Metropole) are considered:

- MANCH_DATA_01: a data service exposing real-time journey times data for the Greater Manchester case study area in a custom JSON format.
- MANCH_DATA_02: a data service exposing the position of sensors measuring the journey times in MANCH_DATA_01 in a custom JSON format.
- RENN_DATA_01: a data service exposing real-time road traffic data for the Rennes Metropole case study area in a custom CSV format.

[4] https://github.com/cefriel/chimera.

[5] https://camel.apache.org/.

The three custom formats are different but represent similar data associated with real-time traffic measurements. The three data sources contain travel times data, while only REN_DATA_01 includes additional information regarding average speeds and traffic status.

The harmonisation requirement identified is to represent the information of the three data sources using a uniform JSON format based on DATEX II[6] semantics that can be interpreted by other TANGENT services. The resulting data flow should contain the real-time measurements and the position of the sensor that carried out the measurements.

The three data sources are described in the TANGENT Data Catalogue, using a uniform metadata descriptor, and exposed through the TANGENT Data API that filters the incoming data according to the case study areas. The two data sources from Greater Manchester should be merged to obtain the needed data, while the Rennes Metropole data source already integrates both measurements and the position of the sensors. The data in the selected data sources should be lifted to their RDF representation (according to the TANGENT Reference Conceptual Model) by defining a set of declarative mappings. A single lowering mapping is needed to extract data from the produced knowledge graph and lower the information to the common DATEX II-based JSON format. This example shows the advantage of the any-to-one mapping approach: only the lowering mapping should be modified if a modification is introduced in the expected target format since the lifting operation is completely decoupled. Moreover, the intermediate representation through an RDF graph facilitates the possibility of merging different data sources and querying an integrated dataset.

The implementation of the designed pipelines was done following these steps:

- Analysis of the TANGENT Reference Conceptual Model to identify the relevant class and properties to be used for the mappings of concepts specified in the input data sources.
- Encoding the lifting mappings from the input JSON format to their RDF representation according to the identified class and properties.
- Definition of the lowering mappings from the RDF representation to the output JSON format relying on DATEX II-based semantics. The harmonised and possibly fused data are accessed from the knowledge graph and lowering mappings are defined according to the expected output format.
- The semantic harmonisation and fusion pipelines are defined and configured using the Chimera components and Apache Camel.

The pipelines can then be executed and exposed as a data service listed in the TANGENT Data Catalogue and integrated with the TANGENT Data API.

4 Conclusions

This paper has described the design of the TANGENT solution for data harmonisation and fusion to address heterogeneous interoperability issues related to data discovery, access, harmonisation, integration, and extraction. The solution has been realised as a set

[6] https://www.datex2.eu/.

of tools (i.e., the TANGENT Data Catalogue, the TANGENT Data API, the TANGENT Reference Conceptual Model, and the Semantic Harmonisation and Fusion Pipelines) to address the various data interoperability requirements elicited by the project. Considering the state-of-the-art and best practices, the solution has been designed by applying the any-to-one centralised approach for semantic interoperability solutions, enabling data conversions and exchange with unambiguous and shared meaning.

The solution has been released, and its integration with the overall TANGENT solution is ongoing.

Acknowledgements. The presented research has been supported by the TANGENT project (Grant Agreement no. 955273), co-funded by the European Commission under the Horizon 2020 Framework Programme.

References

1. Comerio, M., Azzini, A., Scrocca, M., Fiano, A., Metta, S.: D2.2 Data-sharing governance model (2022). https://tangent-h2020.eu/wp-content/uploads/2023/01/TANGENT_D2.2_Data-Sharing_Governance_Model.pdf
2. European Commission. DCAT Application profile for data portals in Europe (DCAT-AP). Version 3.0.0 (2023). https://semiceu.github.io//DCAT-AP/releases/3.0.0
3. Grassi, M., Scrocca, M., Carenini, A., Comerio, M., Celino, I.: Composable semantic data transformation pipelines with chimera. In: Proceedings of the 4th, International Workshop on Knowledge Graph Construction. CEUR Workshop Proceedings (2023)
4. Sadeghi, M., et al.: SPRINT: Semantics for PerfoRmant and scalable INteroperability of multimodal transport. In: 8th Transport Research Arena TRA 2020, pp. 1–10 (2020)
5. Scrocca, M., Comerio, M., Carenini, A., Celino, I.: Turning transport data to comply with eu standards while enabling a multimodal transport knowledge graph. In: Proceedings of the 19th International Semantic Web Conference, pp. 411–429. Springer, Heidelberg (2020)
6. Vetere, G., Lenzerini, M.: Models for semantic interoperability in service-oriented architectures. IBM Syst. J. **44**(4), 887–903 (2005)

Towards Privacy-Preserving Connected Vehicles: A Blockchain Approach for Vehicle Identity Management and Data Sharing

Dermot O'Brien(✉), Vasileios Christaras, Ioannis Kounelis, Igor Nai Fovino, and Georgios Fontaras

Joint Research Centre, European Commission, Ispra, Italy
{georgios.fontaras,georgios.fontaras}@ec.europa.eu

Abstract. As vehicular connectivity and digitisation surge, escalating data transmission to Original Equipment Manufacturers (OEMs) and public authorities is vital for the digital transition, for applications from legal compliance to traffic management. Amidst burgeoning data-sharing ecosystems, ensuring secure, private data transmission and General Data Protection Regulation (GDPR) compliant user control over vehicle identity and data-sharing permissions becomes pivotal, barring legal and enforcement exceptions. This research explores employing blockchain technology to safeguard privacy and security within a Vehicle Identity Management system, using CO_2 emissions monitoring as an exemplar. Utilising an emulation-based environment, replicating vehicle interactions with European authorities and the European Commission (EC), the study demonstrates that blockchain systems, specifically for CO_2 emissions monitoring, can meet transaction rate and latency demands of large-scale transport applications, accommodating 280 million vehicles reporting annually. This inquiry not only amplifies understanding of blockchain's applicability in connected transportation systems and secure data exchange among vehicles, authorities, and stakeholders but also lays groundwork for future advancements in trustful, efficient, and secure data interchange, potentially benefiting authorities, industry, and end-users alike.

Keywords: Sustainable Transport · Digital Identity · Data Integrity · Security and Privacy · Vehicle Connectivity · Regulatory Monitoring · Blockchain

1 Introduction

[1]European policy increasingly emphasises the importance of connectivity and interoperability of various systems in the road transport sector. Hundreds of millions of vehicles are projected to interact on a daily basis with people, automated systems and each other. The deployment of new mobility solutions that will improve the sustainability, effectiveness and energy efficiency of the transport sector will rely on connectivity and secure data exchange between entities.

[1] The views expressed in this document are purely those of the authors and shall not be considered as an official position of the European Commission under any circumstance.

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 351–357, 2026.
https://doi.org/10.1007/978-3-032-06763-0_51

One technology suitable for this purpose are distributed ledgers, and particularly blockchain (BC) systems. At their core, blockchain systems are a type of distributed ledger technology that enables secure and transparent documentation of digital transactions. They are comprised of a series of blocks, each containing a group of validated transactions, which are cryptographically linked to the previous block, thereby creating an immutable and tamper-proof record.

Blockchains could revolutionise how real-world data and assets are connected to their digital counterparts in several ways:

- **Smart Contracts**: Automate transactions for predefined conditions.
- **Identity Management**: generate secure and easily verifiable digital identities for entities, that may be used to control access to data.
- **Data Sharing and Interoperability**: secure data sharing, fostering collaboration and interoperability.

This paper builds upon previous research [1], which investigated the application of BC technology for secure and reliable data communication in the road transport sector. This research changes the network structure to more accurately reflect the real-world. The key research questions addressed are the plausibility of using blockchain in a wide scale application covering large portions of the fleet, and what would be a possible system architecture for doing so. In this paper we investigate the communication of vehicle emissions data to European authorities. Such data retrieval directly from the vehicles is foreseen by existing EU regulations for CO_2 emissions [2–4] and have been proposed for the monitoring of regulated CO_2 and pollutant emissions.

In the context of emissions monitoring of vehicles, the crucial element for preserving privacy is ensuring the Vehicle Identification Number (VIN) is hidden while verifying that it is both a valid and unique VIN registered at one of the EU27 Vehicle Registration Authorities. Centralised access management of identities raises concerns, such as security risks due to a single point of failure, performance issues from technical problems or Denial of Service (DoS) attacks, and privacy breaches following database leaks [5–7].

2 Methodology

This research explores the application of BC technology for securely managing emission data from vehicles, targeting a throughput capable of handling the EU27 fleet of 280 million vehicles. A comprehensive description of the hardware and software utilized is documented in the JRC Science for Policy Report [1].

Infrastructure and Experimentation: Utilising JRC's EPIC cluster, we emulated a realistic blockchain (BC) network for simulation purposes, involving the deployment of 80 EPIC nodes connected via a 1 Gbps line, each facilitated with 6 GB of RAM and 40 GB of storage [8]. Hyperledger Fabric (HLF) 2.2 [9] was employed for its modular and private architecture on Ubuntu 18.04, and experiments were executed within an orchestrated environment utilizing Kubernetes atop EPIC to efficiently handle network layers, storage allocation, and scaling.

Blockchain Network and Data Privacy: The Hyperledger Fabric employs Private Data Collections (PDCs) to assure data privacy by ensuring only authorized entities access sensitive data, enabling GDPR compliance via irreversible data deletion, and offering adjustable access control to safeguard data while promoting secure sharing among organizations on the channel. A detailed breakdown of PDCs and Hyperledger terminology is available in referenced sources [9–12].

2.1 Experimental Setup

Changes were made to the endorsement policy from previous experiments [1] to:

1. require approval only from the Vehicle Registration Authority (RA) and the EC endorser, thereby reducing traffic and conserving storage, and
2. transition to a realistic model, where clients connect, transmit transactions, and disconnect, more closely mimicking real-world operational circumstances.

Vehicles, upon registration with the relevant RA, are assigned an X.509 certificate and a VIN recorded in the RA's global state database, enabling them to transmit emissions data to the pertinent EC organization. Vehicle data is stored on the peers of each RA and the EC, accessible only to these entities, while others receive a hash of the emissions data.

In a network emulating real-world scenarios, 100,000 VINs per RA were registered and then used to generate transaction processes involving connect-discover-authenticate-disconnect procedures, ensuring an extensive engagement of the entire system. A script enabled five connections, each representing a vehicle transmitting data, to mimic a genuine scenario involving multiple vehicle connections to the network. The GNU parallel tool was utilized to manage a substantial number of connections and execute multiple script instances concurrently [1].

3 Results

The successful simulation of a comprehensive network, including 27 Member State organisations and the EC, is depicted in Fig. 1 and Fig. 2. The system managed to support 200 clients per Member State (5400 total clients connected simultaneously), handling 2.7 million transactions while maintaining stability and reasonable performance. A saturation point of around 250 Transaction Per Second (TPS) was consistently reached, regardless of the number of organisations, indicating the maximum throughput achievable with the current configuration. Figure 1 shows that when there are only a few Member State organisations with a limited number of clients, the system does not reach saturation. For both 3 and 5 Member State configurations, saturation is achieved with 200 clients per Member State, as shown in Fig. 1 and Fig. 2. With 100 clients per Member State, the network reaches its peak performance at around 280 TPS and a latency of approximately 0.5 s. However, increasing the number to 200 clients per Member State results in a lower TPS of 245. The constant latency, as seen in Fig. 2, indicates that "this reduction is due to an increase in CPU utilisation caused by the larger number of clients" [1].

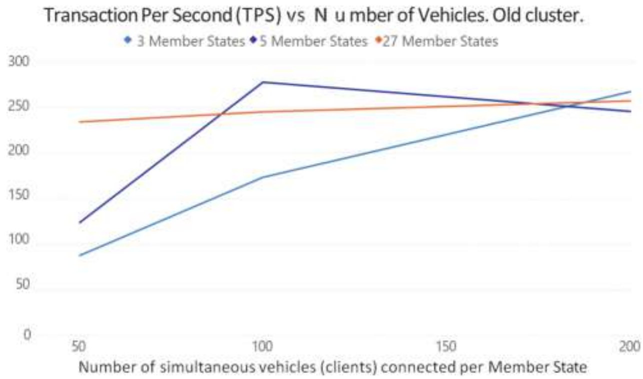

Fig. 1. Experiment 2: "*Transactions per second for 3, 5, 27 Member States + EC with varying numbers of simulated vehicles*". Source: JRC, 2022 [1].

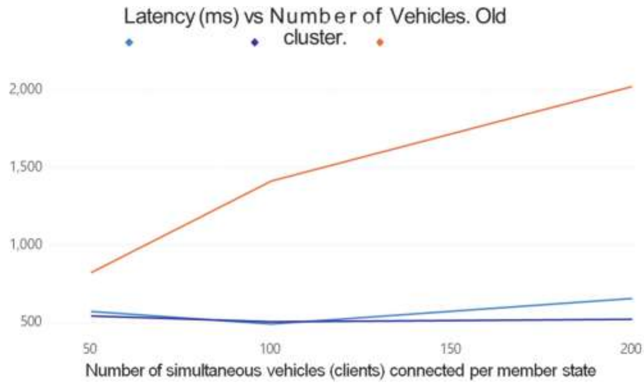

Fig. 2. Experiment 2: "*Latency in seconds for 3, 5, 27 Member States + EC with varying numbers of simulated vehicles*". Source: JRC, 2022 [1].

Despite the BC network performance remaining stable, full utilisation each of the virtual CPUs was observed for all endorsing peers. The system with 27 Member States reached saturation even with a small number of clients. When 50 clients were connected simultaneously, "*latency was significantly higher than in the 3 and 5 Member State configurations, increasing to almost 2 s with 200 clients. It is important to note that the system remained stable and no transaction failures or timeouts were observed on the client side*" [1]. However, trials with 300 simultaneous clients, which are not shown in the figures, revealed increased instability across all Member State configurations, mainly due to timeouts and connection failures.

This research builds upon the initial experimental setup findings described in [1, 13], incorporating a modified consensus approach and client configuration to more closely resemble real-world scenarios. By mandating signatures solely from the endorsing peer of the RA in which the vehicle is registered and the endorsing peer of the EC, the new consensus approach addresses the concerns of maintaining data integrity and audit ability, reducing storage footprints and lowering network traffic.

The updated client configuration, which incorporates 100,000 VINs per RA and simulates a more intricate connect-discover-authenticate-disconnect process, provides a more accurate and comprehensive representation of real-world use cases. This enhanced approach captures the complex aspects of vehicular connections within the blockchain network, offering valuable insights into the network's performance under realistic conditions.

In this experiment, the system was able to support a significant number of clients (200 per Member State, totalling 5400 simultaneous clients) while maintaining stability and reasonable performance. The saturation point of approximately 250 TPS was consistently reached, regardless of the number of organisations involved. This result indicates that the current configuration has a maximum achievable throughput, which may need further optimisation to accommodate higher transaction rates.

However, the experiment also revealed some limitations, such as increased latency and resource constraints with higher numbers of clients. These findings indicate that further optimisation of the network design, consensus approach, and hardware resources may be necessary to improve performance and accommodate larger numbers of clients.

The primary limitation of the system was its ability to accommodate a large amount of clients. "*The system could not handle more than 200 simulated vehicles for each Member State organisation. When the number of vehicles in the simulation*" [1] exceeded this limit, the system experienced difficulties. There was a dramatic increase in the number of transactions that timed out and clients that were disconnected from the system. This suggests that the system has a limited capacity to manage a high volume of clients at the same time.

In conclusion, this research demonstrates the potential of the modified consensus approach and client configuration for handling real-world use cases involving vehicular connections to the blockchain network. The results highlight the scalability and robustness of the proposed architecture while also identifying areas for further optimisation and refinement. Future work should focus on addressing these limitations and further improving the network design to ensure its suitability for large-scale emissions monitoring and other related applications.

4 Conclusion

This research aimed to evaluate the technical feasibility of sharing vehicle data through a blockchain-based system, utilising current technology. "*Real-world fuel consumption monitoring served as a representative use case, involving a fleet of 280 million vehicles*" [1], where each Member State is legally obligated to report CO_2 emissions from road vehicles to the EC annually. Optimisations in the experimental setup enabled the system to successfully support 5400 concurrent clients (vehicles) distributed across 27 Member States. Each vehicle completed a full life cycle, taking approximately three seconds, achieving a maximum throughput of 257 TPS while maintaining system stability and responsiveness.

Ideally, processing the 280 million required transactions per year would take approximately 12.6 days for all EU vehicles. Although this result may not reflect real-world conditions with non-optimal timing and variable demand, it does indicate that the system can "*meet the legal requirement of one transaction per vehicle per year*" and even has the potential to reach the experimental goal of one transaction per vehicle per month. However, this work also identified significant HLF limitations, such as considerable "*transaction overhead in terms of disk usage for each endorsement*" [1] and managing numerous open client connections simultaneously. By addressing these limitations, the project team observed substantial performance improvements in the new experimental setup detailed in this research.

References

1. O'Brien, D., Christaras, V., Kounelis, I., Fovino, I.N., Fontaras, G.: Final report of the exploratory research project, blockchain for transport (bc4t). Scientific analysis or review KJ-NA-31161-EN-N (online), KJ-NA-31161-EN-C (print), Luxembourg (Luxembourg) (2022). https://doi.org/10.2760/309745(online),10.2760/491939(print)
2. EC. Commission regulation (eu) 2018/1832 of 5 November 2018 amending directive 2007/46/ec (2018). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32018R1832
3. EC. Regulation (eu) 2019/631 of the European Parliament and of the council of 17 April 2019 (2019). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32019R0631
4. E. Parliament. Vote to amend regulation (eu) 2019/631 on co2 emission standards for cars and vans; last accessed July of 2022 (2022). https://oeil.secure.europarl.europa.eu/oeil/popups/summary.do?id=1706841&t=d&l=en
5. S.E. Haddouti, M.D.E.C.E. Kettani, arXiv preprint arXiv:1902.11184 (2019)
6. Cao, Y., Yang, L.: In 2010 IEEE International Conference on Information Theory and Information Security (IEEE, 2010), pp. 287–293 (2010)
7. Dabrowski, M., Pacyna, P.: In 2008 Second International Conference on Emerging Security Information, Systems and Technologies, pp. 232–237. IEEE (2008)
8. Siaterlis, C., Genge, B., Hohenadel, M.: IEEE Trans. Emerg. Top. Comput. **1**(2), 319 (2013)
9. Hyperledger. Hlf docs v2.2; last accessed April of 2022 (2020). https://hyperledger-fabric.readthedocs.io/en/release-2.2/ledger/ledger.html
10. HLF. hyperledger-fabricdocs documentation release master v1.2; last accessed April of 2022 (2020). https://buildmedia.readthedocs.org/media/pdf/hyperledger-fabric/release-1.2/hyperledger-fabric.pdf
11. HLF. Glossary — hlf docs; last accessed April of 2022 (2020). https://hyperledger-fabric.readthedocs.io/en/latest/glossary.html
12. HLF. hyperledger-fabricdocs documentation release master v1.4; last accessed April of 2022 (2021). https://readthedocs.org/projects/hyperledger-fabric/downloads/pdf/release-1.4/
13. O'Brien, D., Christaras, V., Kounelis, I., Fovino, I.N., Fontaras, G.: In 2022 14th International Conference on Knowledge and Smart Technology (KST), pp. 1–6 (2022). https://doi.org/10.1109/KST53302.2022.9729059

A Methodology for Planning and Executing Mobility Data Labs: Fostering Collaboration, Data Sharing, and Innovation

Anna Kontini[1], Alexandros Papacharalampous[1](✉), Thierry Chevallier[2], and Johannes Lauer[3]

[1] AETHON Engineering, Emmanouil Benaki 25, 10678 Athens, Greece
{a.kontini,a.papacharalampous}@aethon.gr
[2] Akkodis Research, 7 Boulevard Henri Ziegler, 31700 Blagnac, France
thierry.chevallier@akkodis.com
[3] HERE Technologies, Am Kronberger Hang 8, 65824 Schwalbach Am Taunus, Germany
johannes.lauer@here.com

Abstract. Data plays a pivotal role in modern mobility research, and data labs have emerged as powerful platforms for promoting collaboration, innovation, and data sharing in the mobility sector. This paper presents a comprehensive methodology designed to plan and execute data labs in the mobility sector, along with its successful implementation within the EU-funded MobiDataLab project context. Central to the methodology is the identification of critical stakeholders, the precise delineation of objectives and challenges, ensuring alignment with genuine mobility issues. It staunchly adheres to a user-centric approach, actively engaging end-users, entrepreneurs, developers, researchers, start-ups, and SMEs with data expertise through a systematic, replicable strategy. Additionally, this paper seeks to introduce the technical infrastructure that was designed and deployed to facilitate these endeavors. The methodology emphasizes the establishment of robust data management systems, protocols, and privacy safeguards. Simultaneously, the methodology advocates for the promotion of interoperability standards and open data formats to facilitate seamless access to diverse data sources. Advanced open data catalogues and data enrichment processors, with anonymization features, further enhance data collaboration and privacy. Capacity-building initiatives enhance stakeholder skills, supported by an Open Knowledge Base for sharing best practices. The methodology's efficacy is finally illustrated through case studies which underscore the concrete benefits of data labs in advancing formal innovation and collaboration within the mobility sector.

Keywords: Mobility Data Sharing · Urban Transport · Data Science · Data Labs · Living Labs

1 Introduction

In the rapidly evolving landscape of mobility solutions, harnessing the power of data has become the cornerstone for driving innovation, enhancing user experiences, and making transport systems more efficient and sustainable. Data labs have emerged as

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 358–363, 2026.
https://doi.org/10.1007/978-3-032-06763-0_52

dynamic hubs where stakeholders from various domains converge to co-create, experiment, and leverage data-driven insights for mobility solutions. This paper outlines a robust methodology for orchestrating the creation and operation of mobility data labs, fostering collaboration among diverse stakeholders, facilitating data sharing, and igniting innovation.

In the upcoming sections, we will delve into the core elements of our methodology. Within the methodology section, we emphasize the significance of a user-centric approach and delve into how we seamlessly integrate the quadruple helix model and the Plan-Do-Check-Act (PDCA) tool. As we progress, the implementation in the MobiDataLab project section provides concrete evidence of our methodology's success. Here, we explore the ambitious goals of the MobiDataLab project, its holistic design principles, and the indispensable role played by stakeholders in propelling innovation forward. Concluding our exploration, we delve into the crucial role of FAIR principles in enhancing the efficacy of mobility services, detailing how the MobiDataLab project has successfully integrated these principles into its framework. Finally, through reflection and iterative refinement, we have forged a methodology that paves the way for progress in the realm of mobility data labs.

2 Methodology

To embark on the journey of establishing and operating mobility data labs, it is paramount to establish a user-center methodology that not only leverages the expertise of diverse stakeholders, but also prioritizes the real-world needs and challenges of the mobility sector. This paper, therefore, seeks to provide real-world evidence and insights into the establishment and operation of mobility labs based on the findings of the EU MobiDataLab research project.

2.1 Living Labs and the Quadruple Helix Model

The Living Lab approach, along with its adaptations like the mobility labs, operates on the concept of the quadruple helix model of innovation collaboration. In this model, government, industry, academia and the general public join forces, collaboratively generating innovative solutions [1]. This approach aligns seamlessly with the mobility data lab framework, as it actively involves these key stakeholders, promoting a holistic approach to mobility innovation. Schaffers et al. (2007) underscores the significance and promise of engaging (end) users actively and consistently throughout the development process, through a collaborative co-creation approach with the developers [2]. The shift in living labs towards emphasizing early-stage innovation like user research and co-creation holds promise for enhancing user participation in ICT innovation and unlocking their innovative potential [3].

2.2 Beyond the State of the Art

As a notable advancement beyond conventional practices, the MobiDataLab project has forged a groundbreaking methodology for the inception and execution of mobility data labs. This innovative approach seamlessly integrates the well-established Plan Do Check Act (PDCA) tool, bolstering it with invaluable insights distilled from the iterative experiences of MobiDataLab living labs. These lessons learned from real-world iterations have profoundly enriched our methodology, elevating it to a new standard of effectiveness and practicality.

The PDCA cycle involves planning, executing, evaluating, and refining solutions iteratively. It ensures that innovations are not only ideated but also rigorously tested, refined, and aligned with real-world challenges. This methodology, rooted in the living lab concept, takes innovation a step further by actively involving a diverse range of stakeholders, including end-users, in each phase of the PDCA cycle (Fig. 1).

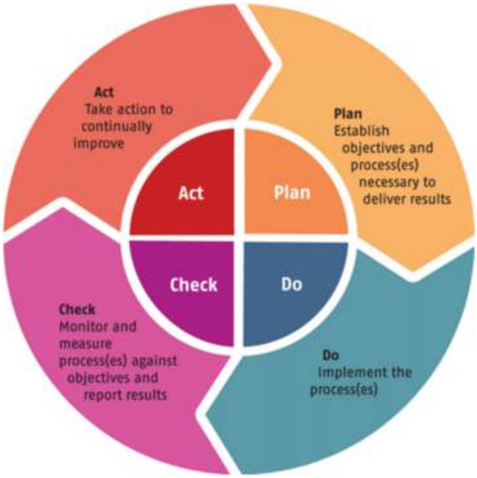

Fig. 1. The Plan Do Check Act Cycle [4]

3 Implementation in the MobiDataLab Project

The MobiDataLab uses the PDCA approach to guide the project stakeholders: project partners (10 project partners from industry, research, academia, consultancy and governance sectors, located in 7 countries), the reference group, the living labs attendees and external participants to pave the way for successful mobility data sharing.

The execution of the approach did start with an overview on the state-of-the-art, the implementation of the transport cloud, the execution together with the stakeholder group within the three living labs (the Datathon, the Hackathon and the Codagon) and the review of domain expert along the path to reflect and integrate the feedback.

The efficacy of our methodology is rooted in its ability to actively involve stakeholders across the mobility spectrum. Entrepreneurs, developers, researchers, start-ups, and SMEs, armed with data knowledge, become essential participants, equipped to tackle the challenges posed by data providers. Figure 2 shows the five stakeholder groups with their motives and needs as a result of a comprehensive stakeholders' analysis within the mobility sector, coupled with an extensive literature review of successful implementations of mobility living labs.

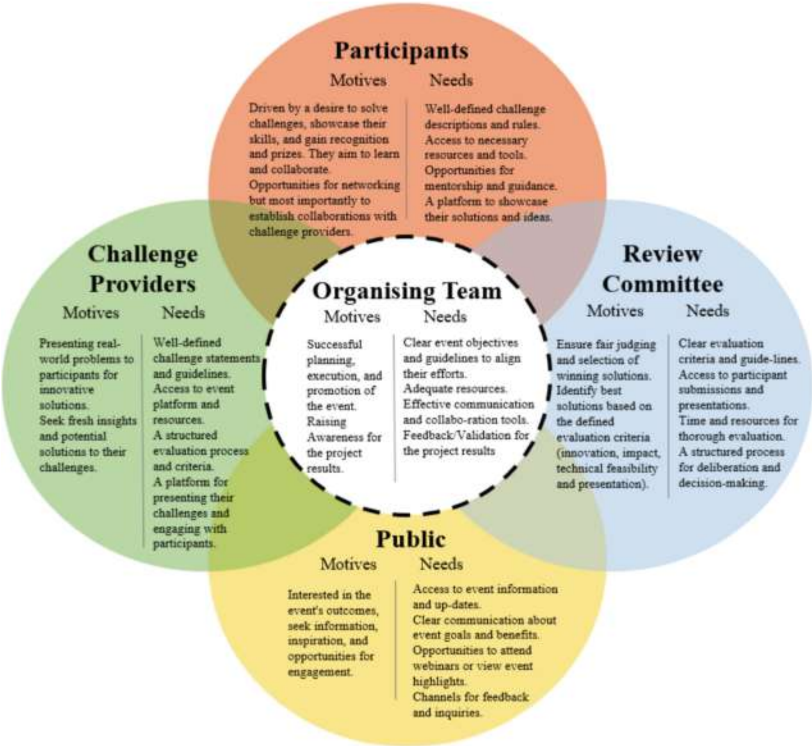

Fig. 2. MobiDataLab Living Labs Stakeholders Analysis

Organising Team: Project partners involved in the planning, execution and evaluation of the MobiDataLab living labs.

Challenge Providers: Municipalities of Leuven, Eindhoven, Milan and Paris.

Review Committee: Consisting of project members and external domain experts.

Participants: Innovators (SMEs, start-ups, researchers, developers etc.).

Public: End-users who use mobility services (e.g. commuters, residents, and citizens) and indirectly affected population groups.

The Virtual Lab platform, a significant component developed within the MobiDataLab project, plays a pivotal role in the realisation of the project's mobility labs. Virtual Lab is a digital counterpart to the traditional living lab concept and contains various functionalities that support discussion and promotion of solutions throughout their inception life cycle, from challenge to idea to a prototype [5].

4 FAIR Data and Services

Effective open innovation hinges on the accessibility and availability of open data. As a result of the project and improved by the living labs, the Virtual Lab platform is following the FAIR principles (findable, accessible, interoperable and reusable) [6] and provides access to mobility metadata catalogues and a service catalogue. The practical application of these catalogues is exemplified through the various challenges tackled during the MobiDataLab Living Labs, demonstrating the platform's ability to address real-world mobility issues and catalyze innovative solutions[1].

In the context of the data labs, it is necessary to make mobility data easier to find, discover and reuse. This is achieved using data catalogues, which enable a more efficient search on the key aspects of each dataset, i.e. descriptive metadata. In the open data ecosystem, a number of software systems can serve this purpose, for example CKAN, GeoNetwork, OpenDataSoft, Socrata [7]. The first difficulty is choosing a solution. In MobiDataLab, the CKAN solution has been elicited due to its modularity, the number of available extensions, and their variety. In particular, the CKAN harvesting extension functionality was implemented, allowing to add datasets on demand into the MobiDataLab data catalogue, from data portals corresponding to the local area of referring municipalities. A future challenge for the MobiDataLab platform is overcoming organisational mistrust, rooted in fears of granting competitors an unintended advantage. Usability by data consumers across the European union remains a challenge since European citizens speak different languages. Harvesting metadata in the local language and translating them in real time into the consumer or most spoken language requested the implementation of dedicated extensions, resulting in MobiDataLab's contribution to open-source software. Not only the MobiDataLab platform improves findability, but also reusability, offering anonymization techniques to data providers, such as microaggregation for anonymising large mobility datasets [8].

Today's mobility services are very diverse and distributed. The FAIR principle is less established and implemented. Catalogues to register services and make them findable are becoming more requested but are not yet stablished. Also, standards for interoperable interfaces are not available or still in discussion. Within the MobiDataLab project, a catalogue with an easy interface to register OpenAPI (https://www.openapis.org/) described services is part of the Transport Cloud ecosystem. The catalogue allows the registration of REST APIs and makes them findable. Further demands are easy access, interoperable interfaces (e.g. aligned and agreed by service providers) to fulfill the reusability as the fourth step to make mobility services FAIR.

5 Conclusion

The methodology we present here is a blueprint for harnessing the full potential of data labs in the mobility sector. By bringing diverse stakeholders to the table, setting clear objectives, and fostering user-centric innovation, we enable the co-creation of solutions that address the pressing mobility challenges of our time. The PDCA approach implements the agile principle and aligns and supports all stakeholders on the way to a FAIR approach for mobility data sharing. The provisioning of data and services and connection of data producers and data and service consumers in the context of the Living and Virtual Labs improves the alignment and enables the connectivity between stakeholders through communication and collaborative solution enablement.

[1] Details of the challenges can be accessed at: https://labs.mobidatalab.eu/challenges/.

The MobiDataLab project stands as a testament to the effectiveness of this methodology, demonstrating that by working together, we can create a future of more efficient, sustainable, and inclusive mobility solutions.

References

1. Voytenko, Y., McCormick, K., Evans, J., Schliwa, G.: Urban living labs for sustainability and low carbon cities in Europe: towards a research agenda. J. Clean. Prod. (2015). https://doi.org/10.1016/j.jclepro.2015.08.053, retrieved on 23.08.2017
2. Schaffers et al.: Exploring business models for open innovation in rural living labs. 13th International Conference on Concurrent Enterprising, Sophia-Antipolis, France (2007)
3. Følstad, A.: Living labs for innovation and development of information and communication technology: a literature review. eJOV: The Electronic Journal for Virtual ldots. 10. 99–131 (2008)
4. PDCA Cycle. ASQ. [En ligne] https://asq.org/quality-resources/pdca-cycle
5. Virtual Lab. (n.d.). https://labs.mobidatalab.eu/
6. Wilkinson et al.: The FAIR guiding principles for scientific data management and stewardship, Scientific data 3 (2016) 1–9.7
7. https://ckan.org/. https://geonetwork-opensource.org/. https://www.opendatasoft.com/. https://open-source.socrata.com/
8. https://github.com/MobiDataLab

Data Sharing at the Edge of the Network: A Disturbance Resilient Multi-modal ITS

Igor Mikolasek[1(✉)], Saeedeh Ghanadbashi[2], Nima Afraz[2], and Fatemeh Golpayegani[2]

[1] Transport Research Centre CDV, Brno, Czech Republic
igor.mikolasek@cdv.cz

[2] School of Computer Science, University College Dublin, Dublin, Ireland

Abstract. Mobility-as-a-Service (MaaS) is a paradigm that encourages the shift from private cars to more sustainable alternative mobility services. MaaS provides services that enhances and enables multiple modes of transport to operate seamlessly and bringing Multimodal Intelligent Transport Systems (M-ITS) closer to reality. This requires sharing and integration of data collected from multiple sources including modes of transports, sensors, and end-users' devices to allow a seamless and integrated services especially during unprecedented disturbances. This paper discusses the interactions among transportation modes, networks, potential disturbance scenarios, and adaptation strategies to mitigate their impact on MaaS. We particularly discuss the need to share data between the modes of transport and relevant entities that are at the vicinity of each other, taking advantage of edge computing technology to avoid any latency due to communication to the cloud and privacy concerns. However, when sharing at the edge, bandwidth, storage, and computational limitations must be considered.

Keywords: Shared mobility · Mobility-as-a-Service · Decentralized Systems · Data sharing · Edge Computing

1 Introduction

The concept of Mobility-as-a-Service (MaaS) is gaining interest due to increased awareness of sustainable mobility and the adoption of Intelligent Transportation Systems (ITS) [1,2]. MaaS aims to shift people from cars to more sustainable modes of transportation, relying on ITS for efficient infrastructure use, real-time data provision, and user-guided routing. Presently, public transport (PT) systems commonly use ITS for real-time data on arrival times and vehicle locations. Similarly, bike/scooter-sharing services provide device location data, and traffic signals can optimize intersection throughput based on traffic conditions. These ITS systems often lack cooperation among different transportation modes, limiting their effectiveness particularly at the times of disturbances. Multimodal ITS (M-ITS) [5] can enhance MaaS, where different modes can support or substitute

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 364–370, 2026.
https://doi.org/10.1007/978-3-032-06763-0_53

one another [3,4,7]. This paper provides an overview of transportation modes, networks, and their interactions, focusing on disturbances and how M-ITS can mitigate them. M-ITS concept can incorporate edge computing paradigm [6] to place the processing of data closer to its source. This will address privacy concerns, reduce data processing delays, enable decentralize computation, and eliminate costly centralized systems. Using edge computing paradigm, decisions are made in a decentralized manner, relying on data from the local transport network, onboard vehicle sensors, and information from nearby vehicles and infrastructure. Identifying data-sharing needs, and the data flow among systems and vehicles during disturbances is essential. This paper outlines such possible disturbances and their impact on different modes and suggests coordinated impact-mitigation strategies for the M-ITS-based MaaS system.

2 Transportation Modes and Networks

To facilitate data sharing between multiple modes of transport, we first need to understand their dependencies and how they interact with one another. This paper focuses on transport in urban areas, although most modes can operate beyond. Expanding M-ITS and MaaS to intercity or interstate levels would involve considering additional modes, especially air transport that are out of scope for this paper. To do so, first, we model the dependencies between transportation modes and their infrastructure networks in urban areas.[1] Transport modes can operate on various networks, some shared and others exclusive. Networks often share infrastructure, for example roads are shared between cars, buses, and cyclists. They can be modeled collectively, with parameters defining mode usage. Multimodal modes indicate possible mode transitions. Figure 1 presents an overview of the modes noted as M_i and the networks they use, noted as N_i.

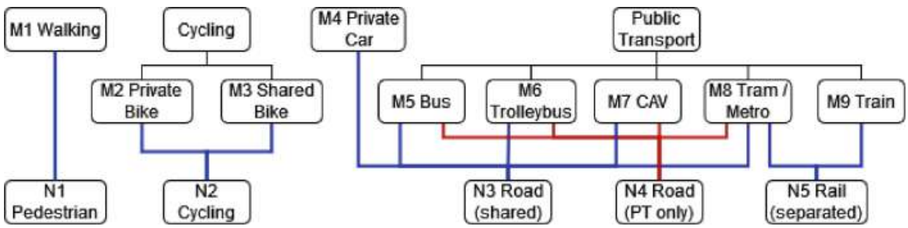

Fig. 1. Overview of fundamental transportation modes and networks employed. The indices M_i and N_i are employed for mode and network respectively and specifying the pairings between the two in Fig. 2.

[1] The commonly available transportation modes provided is not exhaustive – each category can represent multiple modes of similar nature.

3 Service Disturbances to Mode-Network Combinations and M-ITS Adaptation Strategies

Unplanned events can disrupt normal traffic operations, impacting system performance and requiring adaptations to limit such impacts. Data sharing and collective intelligence can inform about and address the challenges related to such disturbances. Five main causes of disturbances are identified, impacting various mode-network combinations (see Fig. 2). Here, only cases of possible direct impacts (need to re-route) are considered. Secondary impacts due to possible relocation of transport demand are considered later for individual disturbance scenarios. In Table 1, we present a comprehensive overview of individual disturbances, providing further insights into potential data sources for detection, strategies for adaptation, and special cases that might necessitate specific responses.

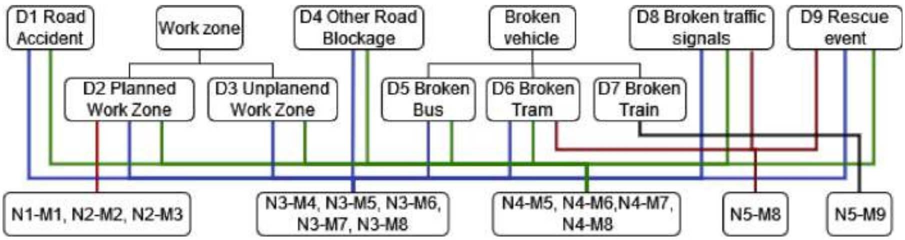

Fig. 2. Effects of different sources of disturbance on different combinations of modes and networks.

Below we provide some examples of disturbances and scenarios to handle them. For example, "Planned Work Zone" along the pedestrian network (N1-M1) might result in disturbance for users with wheelchairs. Such work zones are not commonly recorded in a machine-readable format. This is especially true if they are unplanned, hence this case is omitted in the schema in Fig. 2. Pedestrians and cyclists have the advantage of being very agile and can bypass most obstacles on the network with acceptable delay and they are not connected to any ITS service, except shared bikes. Therefore, except for planned work zones, they are not considered primary stakeholders. These modes are considered here mostly as transition modes and as an adaptation strategy in case of disturbance.

Additional causes of disturbance include major sport or cultural events, and gatherings, that significantly increase (or decrease) traffic demand within certain areas of the networks. Therefore, a registry of such events, including information about the event, time, location, affected area, and possibly expected number of visitors should be included. Some disturbances, especially the external ones mentioned above, can be observed and identified centrally. Then only the warnings with relevant information can be shared with relevant individual edge devices to be evaluated individually. This includes the processing of information about

Table 1. List of detection options, adaptation strategies, and specific cases for the basic identified disturbances.

D	Detection Options	Adaptation Strategies	Specific Cases
D1 - Road accident D4 - Another Road blockage	• User feedback apps (Waze) • C-ITS data from vehicles (V2I) • Traffic information centers (API, C-ITS DENM, RDS TMC, ...) • Integrated rescue system dispatching • Video detection (AI image processing) • Traffic flow (TF) data from sensors	• Warn road users nearby and with routes via the disturbance location • Info + guidance at affected stops to MaaS alternatives • Busses temporarily adjust routes if taxi/CAV can serve bypassed stops, if favorable • Intelligent traffic signs, VMS and other devices adjust to changes in traffic – warnings, guidance, change signal plan	• Other networks affected • Reserved road lanes are affected • Reaction varies based on road blockage • Treat D4 as D2 when extended
D2 – Planned Work zone	• Digitized registry of planned work zones with temporary traffic management measures • Intelligent "first/last cone" or contractor updates*	• Temporarily adjust networks for all affected modes and react aptly	• Adjust duration through regular detection means if not updated dynamically
D3 – Unplanned Work zone	• Integrated rescue system** • User feedback apps (Waze) • C-ITS data from vehicles • Intelligent "cone"* • Traffic information centers • TF data from sensors	• Mostly the same as D1 • For long-lasting work zones, treat as D2 once details are available	• Mostly the same as D1 • Response varies depending on expected duration
D5-D7 – Broken vehicle Rail block	• C-ITS data from vehicles (V2I) • Railway or public transport (PT) operator dispatching • Otherwise treat as D1/D4	• Inform passengers about alternatives within the MaaS • Inform M-ITS about the obstacle and react accordingly (mix of D1 and D3) • For D6, D7, replace operation with busses/CAVs	• Certain PT routes (e.g. airport line) have higher priority • Follow train/ metro safety standards
D8 - Broken Traffic signals	• C-ITS (I2I) from the device • Device operator • TF data from sensors	• Inform police to regulate traffic at the intersection • Warn relevant road users • Optimize nearby traffic signals for changed traffic • Adjust routing for MaaS	• Warn trams, if affected, and react suitably
D9 - Rescue event	• C-ITS data from vehicles • Integrated rescue system dispatching	• Warn nearby road users • Re-route vehicles or adjust signal plans for smooth rescue	• Consider operation size and duration

*C-ITS-equipped "traffic cone" or mobile app to update WZ start/end time.
**Police should know first, communication with M-ITS may vary among countries.

detected disturbances from the edge devices and transport users in cases where the impacts of the disturbance exceed the communication range of the local edge devices.

Other Impact Mitigation Suggestions are as Follows: –For damaged CAVs, inform its users about nearby replacements and send warnings if blocking the road. –For broken shared bikes and scooters, notify users of alternatives; other modes are not affected. –Temporarily adjust service limits for space-limited services when intermodal nodes near the borders are out of service. –Warn users at affected PT stops and guide them to nearby alternatives or other MaaS services. –Empty bike-sharing nests provide info about nearby nests with available bikes and PT alternatives. –Notify taxis and CAVs of disturbances and areas with increased demand. –Provide (multimodal) route change suggestions for users with planned routes. –General control measures for improved performance: Integrate (multimodal) navigation apps into M-ITS; Increase C-ITS systems and OBU penetration; Establish disturbance detection information systems; Implement VMS and variable LED signage; Use intelligent traffic lights, including cooperation, PT prioritization, and multimodal; Create incentives (cost evaluation, tolling, gamification); Implement multimodal ticket schemes; Consider multimodal performance in traffic control systems; Design new data standards for interoperability across the EU (identify best practices for data management; define minimal operational standards for data requirements).

4 Data Sharing at the Network Edge

The disturbance-mode-network model defines combinations of disturbances and affected transportation modes/networks, but not all users or vehicles of the affected mode-network combination must be warned. Only relevant edge devices should be informed with specific details. Given the small size of the basic warning data package, this can be spread quite broadly with more details given upon request, if available. Disturbance warnings sent to edge devices should include the following information, –Type of disturbance, –Location on the network affected modes, and class of the network segment for each mode (e.g., critical, major, inferior, minor), –Expected impact severity (one or more measures as applicable): Capacity reduction; Lanes affected; Severity index; Displaced traffic volume, –Estimated time of duration, and any other case-specific information.

To efficiently distribute warnings for disturbances, it is crucial to consider various parameters like the affected area, specific modes, networks, and user classes. The goal is to notify only relevant users and devices without overwhelming them. This involves a combination of general heuristic rules for area-based warnings, C-ITS traces for approaching vehicles, data from edge devices regarding their planned routes (including non-scheduled traffic), adaptation strategies for dealing with the disturbance, and machine learning algorithms for predicting trajectories. There are two main aspects to consider: (i) information about the disturbance, which is part of the warning itself (type, location, severity, duration), and (2) information about the recipient of the warning. For recipients, it is essential to account for their expected or planned trajectory, any expected TF changes along that trajectory, and whether they are part of the adaptation strategy.

Edge devices in this environment encompass a wide range of components, from sensors and processors within vehicles, through users' mobile phones, to infrastructure elements like traffic signal controllers and communication nodes. Edge computing can greatly enhance the efficiency of distributing warnings for disturbances based on the following considerations: **Efficient Data Processing:** Edge computing processes data closer to the source, enabling quick assessment of disturbances and reducing response times. For example, a CAV detects a fallen tree on the road and immediately sends an alert to nearby vehicles. **Localized Decision-Making:** Edge devices can make localized decisions based on the available information. They can analyze their planned routes, expected TF changes, and other relevant factors to determine whether a particular disturbance warning is relevant to them and address it accordingly. **Adaptation Strategies:** Edge devices can autonomously implement the pre-defined adaptation strategies and cooperate with other devices to optimize overall performance. **Reduced Communication Network Congestion:** By processing and filtering data at the edge, the system can significantly reduce congestion of the communication network. This is crucial during disturbances when there may be an increased volume of data traffic due to widespread alerts.

5 Conclusions

This paper has explored the critical role of M-ITS in enhancing MaaS efficiency, especially during disturbances. Establishing relationships between transportation modes, networks, and disturbances, has highlighted the importance of data sharing on the network edge. The selective distribution of disturbance warnings to relevant users and devices is crucial for effective disturbance mitigation. This framework offers valuable insights for building resilient and adaptive urban transportation systems that prioritize sustainability and user needs.

Acknowledgements. This project has received funding from the RE-ROUTE Project, the European Union's Horizon Europe research and innovation program under the Marie Skłodowska-Curie grant agreement No. 101086343.

References

1. Esztergár-Kiss, D., Kerényi, T., Mátrai, T., Aba, A.: Exploring the maas market with systematic analysis. Eur. Transp. Res. Rev. **12**(1), 1–16 (2020)
2. Golpayegani, F., et al.: Intelligent shared mobility systems: a survey on whole system design requirements, challenges and future direction. IEEE Access **10**, 35302–35320 (2022)
3. Jevinger, Å., Persson, J.A.: Potentials of context-aware travel support during unplanned public transport disturbances. Sustainability **11**(6), 1649 (2019)

4. Mnif, S., Darmoul, S., Elkosantini, S., Ben Said, L.: An immune multiagent system to monitor and control public bus transportation systems. Comput. Intell. **34**(4), 1245–1276 (2018)
5. Qureshi, K.N., Abdullah, A.H.: A survey on intelligent transportation systems. Middle-East J. Sci. Res. **15**(5), 629–642 (2013)
6. Safavifar, Z., Ghanadbashi, S., Golpayegani, F.: Adaptive workload orchestration in pure edge computing: a reinforcement-learning model. In: 2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI), pp. 856–860. IEEE (2021)
7. Yang, G., Wang, J., Zhang, F., Zhang, S., Gong, C.: A real-time timetable rescheduling method for metro system energy optimization under dwell-time disturbances. J. Adv. Transp. **2019**, 1–11 (2019)

nuMIDAS: The New Mobility Data and Solutions Toolkit

Sven Maerivoet[1(✉)], Steven Boerma[2], Rick Overvoorde[2], Chrysostomos Mylonas[3], Dimitris Tzanis[3], Carola Vega[4], Eglantina Dani[4], Magdalena Hyksova[5], Andre Maia Pereira[5], Valerio Mazzeschi[6], and Valerio Paruscio[6]

[1] Transport and Mobility Leuven (TML), Leuven, Belgium
sven.maerivoet@tmleuven.be
[2] MAPtm, Amsterdam, The Netherlands
[3] Centre for Research and Technology (CERTH-HIT), Maroussi, Greece
[4] Factual Consulting, Sant Cugat del Vallès, Spain
[5] Czech Technical University in Prague (CTU), Prague, Czech Republic
[6] Poliedra, Milan, Italy

Abstract. The mobility ecosystem is rapidly evolving, with the rise of new stake-holders and services, accompanied by new ways for the generation, collection, and storing of data. In the nuMIDAS (New Mobility Data and Solutions Toolkit) project, we provided insights into what methodological tools, databases, and models are required, and how existing ones need to be adapted with new data. We started from insights obtained through (market) research and stake-holders, as well as quantitative modelling. A wider applicability of the project's results across the whole EU was guaranteed as all the research was validated within a selection of case studies in pilot cities, with varying characteristics, thereby giving more credibility to these results. Finally, through an iterative approach, nuMIDAS created a tangible and readily available toolkit that can be deployed elsewhere, including a set of transferability guidelines, thus thereby contributing to the further adoption and exploitation of the project's results.

Keywords: Intelligent traffic and mobility management · Data management

1 Introduction

The mobility ecosystem is rapidly evolving, whereby we see the rise of new stakeholders and services. Examples of these are the presence of connected and automated vehicles, a large group of organisations that rally to establish various forms of shared mobility, with the pinnacle being all of these incorporated into a large MaaS ecosystem. As these new forms of mobility offerings start to appear within cities, so do new ways in which data are being generated, collected, and stored. Analysing this (Big) data with suitable (artificial intelligence) techniques becomes more paramount, as it leads to insights in the performance of certain mobility solutions, and is able to highlight (mobility) needs of citizens in a broader context, in addition to a rise in new risks and various socio-economic impacts. Successfully integrating all these disruptive technologies and solutions with the

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 371–379, 2026.
https://doi.org/10.1007/978-3-032-06763-0_54

designs of policy makers remains a challenge at current. Let alone being able to analyse, monitor, and assess mobility solutions and their potential socio-economic impacts.

nuMIDAS, the New Mobility Data and Solutions Toolkit, bridged this (knowledge) gap, by providing insights into what methodological tools, databases, and models are required, and how existing ones need to be adapted or augmented with new data. To this end, it started from insights obtained through (market) research and stakeholders, as well as quantitative modelling.

A wider applicability of the project's results across the whole EU was guaranteed as all the research was validated within a selection of case studies in pilot cities, with varying characteristics, thereby giving more credibility to these results. Finally, through an iterative approach, nuMIDAS created a tangible and readily available toolkit that can be deployed elsewhere, including a set of transferability guidelines, thus thereby contributing to the further adoption and exploitation of the project's results.

nuMIDAS, the New Mobility Data and Solutions Toolkit, ran during 2021 and 2022 under the Horizon 2020 programme and was developed by a European Consortium, composed of 9 partners from 6 countries: Belgium, Czech Republic, Greece, Italy, The Netherlands, and Spain.

The project built on a distributed selection of case studies in four pilot cities to provide a geographic coverage of the EU: Barcelona (Spain), Milan (Italy), Leuven (Belgium), and Thessaloniki (Greece). All public deliverables can be readily downloaded from our website (https://numidas.eu/). The novelty of our project lies in the fact that we now have a platform that is more generally and broadly applicable, whereby new use cases can be more quickly added, thereby leveraging the existing functionalities of the platform.

In the following paragraphs we will elaborate more extensively on the various aspects of our completed research. First, we describe how we selected the different use cases tackled in our project. Then we show our software architecture by explaining the construction of our back-end and front-end engines. Next, we elaborate on a selection of the use cases, how they were modelled, implemented, and applied. Finally, we provide some conclusions and recommendations.

Fig. 1. Geographical overview (figure created by Factual using QGIS).

2 Selection of Use Cases

In order to achieve the project's outcomes, we took the following steps:

- Perform an assessment of the current trends
- Define advanced methods and tools
- Execute case studies in pilot cities

- Create a consolidated toolkit for stakeholders

Some of these steps ran in parallel, with feedback mechanisms involved in order to capture and process the latest developments and insights during the creation of the nuMIDAS toolkit.

2.1 Process for Scoping the Use Cases

To successfully support transportation planners, researchers, and policy makers, emphasis was put on the harmonisation of the development of the toolkit. With this objective, a series of parameters was defined: extensibility, interoperability, minimisation of development and operational costs, reliability, and usefulness and user-friendliness of the toolkit. nuMIDAS provided high-level descriptions of the use cases, along with descriptions of problems which need to be addressed in the specific city(ies) and the technical approach to be adopted. In addition, we also provided UML mock-ups giving the visual representation of each use case approach, and capturing the entire process of a tool in combination with the interrelations with the involved actors. In order to make the scope of each use case more understandable and tangible, they set the boundaries of the tool, the steps to be followed each time, and the requirements needed to move forward (see also Fig. 1). Various use cases were implemented in the toolkit, some fully, others partially, thus providing researchers, planners, and policy makers with visualisations of the results derived from the methods and tools used in the case studies (Fig. 2).

Fig. 2. Overview of the processes followed for scoping the use cases in nuMIDAS. Figure created using Microsoft PowerPoint; icons created by Factual of the developed methodology.

2.2 Final Selection of Use Cases

Thanks to the requirements that emerged from the discussion and the interaction with the project pilot cities and relevant stakeholders, we identified the following 6 use cases (UC) to address the problems and provide the technical support to solve them by creating specific methods and tools: **Use case 1: Pre-planning of shared mobility services (Milan and Leuven)**: This use case designed a tool for optimising the fleet size of shared mobility services, like bikes and e-scooters, by considering various factors to minimise both the generalised trip costs and the system's operating costs. **Use case 2: Operative areas analysis (Milan)**: This use case sought to allocate shared mobility service fleets across metropolitan sub-areas to minimise service providers' economic losses and maintain equitable, high-quality service for citizens. **Use case 3: Air quality and vehicle emissions analysis based on multi-source data (Barcelona and Leuven)**: This use case involved integrating diverse data sources, such as traffic, emissions, and weather, to evaluate traffic's impact on air quality and support the forecasting and policy-making process for air quality improvement. **Use case 4: Planning for parking (Leuven)**: This use case analysed the impact of decreasing on-street parking in urban centers on traffic flow, assisting in crafting and applying policies that address longer parking search times and the shift of parking demand to adjacent areas. **Use case 5: Inflows and outflows in a metropolitan area (Barcelona)**: This use case focused on using data from ANPR-systems and census information to estimate vehicle movements between metropolitan districts, aiding in the enhancement and enforcement of environmental policies like low-emission zones. **Use case 6: Assessment of traffic management scenarios (Thessaloniki and Leuven)**: This use case evaluated traffic management scenarios through simulation and data analytics, integrating data from connected vehicles and traditional counting systems to support decision-making.

We implemented all of these use case, some more as a proof-of-concept, others more as prototype that could (and was) already be put to use by the cities. As such, our nuMIDAS toolkit now supports answering the following questions from policy makers:

- Milan: How many shared (micro)mobility services need to be provided, and where?
- Barcelona: How large are the vehicle emissions (CO2, NOx, and PM)? How much vehicles drive from A to B?
- Leuven: What is the impact of removing on-street parking spaces?
- Thessaloniki: What is the impact of traffic management measures (traffic lights and speed limits)?

3 Deploying a Suitable Software Architecture

In order to make sense, and to have a common and agreed upon language of terminology, the nuMIDAS project adopted a service-oriented architecture. Advanced methods and tools are required to analyse, assess, and monitor new mobility solutions and policies including new data management techniques. The toolkit (in the form of a dashboard) incorporated these methods and tools providing researchers, planners, and policy makers a visualisation of the results through a GUI. For each of the use cases, we devised new algorithms that took the current state of the art into account and add new elements to it. The algorithms were initially coded in Python and first tested under stand-alone lab conditions. Then they were ported to a back-end engine (forming part of the toolkit) that runs an instance of Amazon Web Services (AWS) with appropriate GIS-databases and input/output file handling capabilities. To perform smooth operations, the entire endeavour operates using private and public APIs, with the front-end system containing access control regulations and the visual part of the different use cases (e.g., graphs for geographical interpretations, monitoring of key performance indicators, etc.), as well as scenario management where suitable.

4 Implementation of the Specific Use Cases

4.1 Example Implementation of a Use Case: Shared Mobility Planning

As an example, in the city of Milan, the policy maker required a tool for the preplanning of shared mobility services. The nuMIDAS team designed and developed a tool that processes available data to produce the output required by Milan city mobility planners. The aim was to produce a tool replicable and scalable in different cities/regions in Europe. In the case of station-based car-sharing, as an example, the dashboard suggests the proper location of new car-sharing stations and a well-balanced car fleet. This outcome is produced by back-end algorithms that elaborate data provided by the municipality. New car-sharing services, or extensions of the existing ones, can then be planned based on the outcomes of the nuMIDAS tool. The dashboard allows tuning parameters of the algorithms to compare different scenarios, depending on the weight given to conflicting objectives. Examples of the latter are the increase of shared mobility options to citizens in low-demand areas, and the economic sustainability of car-sharing operators. A policy maker can use the dashboard to link data sources and input parameters, and to visualise the results of the computation in a user-friendly environment. This specific example was tested in the case of Milan, but can be extended to be used by any European city willing to make use of a tool that supports car-sharing planning. For Milan, we modelled and simulated various scenarios through the dashboard, using a suitable set of input parameters: scenario 1: high demand/low fleet; scenario 2: low demand/high fleet; scenario 3: weight factor significantly in favour of the society; scenario 4: weight significantly in favour of the service provider. For example, scenario 1's outcome showed the optimal fleet to be 84 free-floating bicycles. In addition, demand was about 97% covered, outlining a very good coverage of the service for the user segment. The average walking time was within an acceptable value for which the user may choose to use the bicycle for his or her commute. Scenario 3 on the other hand, considered a net imbalance for society. In fact, looking at the results, the optimal fleet value of 93 was deemed very close to the optimal fleet value for the end user (which was 94), indicating how the algorithm took into account that the preponderant objective for this simulation is to satisfy demand at the expense of the operator's profit. As a result, compared to scenario 1, demand coverage was greater while profits were slightly lower (the same reasoning applies to profits as for scenario 1).

4.2 Example Implementation of a Use Case: Planning for Parking

The main objective was to assess the potential impact of on-street parking restriction policies in the city of Leuven. The city is especially concerned with possible spill-over effects of hyperlocal policy changes, such as substantively reducing the available parking spaces in one street or neighbourhood. Operationally, the algorithm, relies on the discretisation of a road network using a grid comprised of cells of appropriate size and the use of parameters, such as the demand for parking and parking capacity (number of parking places) corresponding to each cell. By that means, the tool will enable the simulation of enforcing a parking restriction policy in one or more grid cells and the assessment of its impacts on the remaining cells. Through the simulation of enforcing a parking restriction policy in one or more roads of the network, both the policy maker as well as the transport planner are able to evaluate the effects, on the basis of the KPIs provided by the tool. Specifically, the evaluation of an on-street parking restriction policy includes the calculation of parking pressure, the calculation of searching time, and the examination of the extent of the effect. For the city of Leuven, we used the tool to evaluate a number of scenario's, varying in complexity. Here, several parking places were suspended, leading to a redistribution of the parking demand. Some areas indicated a drop in demand (as expected), with some redistribution to the neighbouring areas, mostly to the southern region. This redistribution had an effect on the parking searching times (Fig. 3).

Fig. 3. Example of the shared mobility planning (left) and planning for parking (right) use cases in the nuMIDAS toolkit's. Figures created within the nuMIDAS toolkit.

5 Conclusions and Further Recommendations

5.1 General Conclusions

Our nuMIDAS project provided insights into what methodological tools, databases, and models are required, and how existing ones need to be adapted or augmented with new data. To this end, we started from insights obtained through (market) research and stakeholders, as well as quantitative modelling. A wider applicability of the project's results across the whole EU was guaranteed as all the research was validated within a selection of case studies in pilot cities, with varying characteristics. Finally, through an iterative approach, nuMIDAS created a tangible and readily available toolkit that can now be deployed elsewhere, including a set of transferability guidelines, thus thereby contributing to the further adoption and exploitation of the project's results.

nuMIDAS's project results represented a reference for any other future scientific work in Europe and provided to the whole scientific community more advanced bases for the theoretical elaboration on mobility. This outcome of nuMIDAS can be quite beneficial for city administrations and local policy makers, as they not always have a clear view of the real role and potential of data in cities. This work can help them in making more informed and evidence-based decisions, and in getting the mobility system "ready" through land use planning and zoning, transport system design, information services and processes for bringing together the different interested actors (city departments, public transport and sharing-mobility operators, developers, etc.).

5.2 Upscaling and Transferability to Other Cities

In order to maximise the exploitation of project results, and in particular a widespread use of our new mobility toolkit, we drew up transferability guidelines. These describe how new cities can implement new instances of the tools within the nuMIDAS architecture framework, and use them for new mobility solutions analysis, assessing, and monitoring. From an operational point of view, due to the city-specific and location specific parameters and constraints (air quality, origin-destination matrix, etc.), the guidelines transferred not the models themselves, but rather provided a useful methodology to adapt the model to each city.

5.3 Further Exploitation via Suitable Business Models

The state-of-the-art analysis provided early on in our project the basis for the work on business models for particular services. As the mobility field is rapidly changing due to, among others, digitalisation, as well as the uprise of small and large electric vehicles, business models and stakeholder roles change. Furthermore, more and more services combine many disciplines, which often means that services require multiple types of stakeholders to become effective. For this reason, a stakeholder analysis was performed in the context of nuMIDAS, in combination with a business model analysis. The analysis was performed using a Service Dominant Business Model Radar (SDBM/R). The SDBM/R gives insight into which stakeholders can be found within the business model, as well as the way these stakeholders add value to this service. In addition to the SDBM/R, the analysis contains descriptions of the change in business models and stakeholder roles within the business models over time. All SDBM/Rs were created per service, based on the achievement of shared goals and value co-creation by a group of actors (businesses, firms, and costumers) which interact to achieve that shared goal. This activity required to identify the values-in-use as the focal points of the services, and further the value propositions (part of the central values contributed by the particular actors), coproduction activities (i.e. those that actors perform to achieve the co-creation of the values), and costs and benefits for particular actors. The radars were complemented with the description of the corresponding business model scenarios, descriptions of actors and their roles and activities, and the discussions of the transitions of business models influenced by new methods and tools. The summary resulted in a short list of the most common stakeholders within the SDBM/Rs. These are: municipalities, travellers, (mobility) service providers, intermediaries, and MaaS providers. The first three stakeholders within this list are most important and most common. These three stakeholders respectively provide policies, demand, and supply for transportation. The last two of these stakeholder types, intermediaries and MaaS providers, are not yet so common in the actual urban mobility field. These are categorised as future partners and are expected to become more prevalent. As transport systems become more and more integrated, while mobility service providers are competing, intermediaries stay independent and are able to handle data of competing partners with care to provide services which would otherwise not have been possible due to said competition. When the apps of MaaS providers finally really take-off, many services are expected to change, as mobility will be even more accessible from within our pockets. As such, our nuMIDAS toolkit is excellently positioned in light of further commercial exploitation.

Acknowledgements. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101007153. nu-MIDAS, the New Mobility Data and Solutions Toolkit, ran during 2021 and 2022 under the Horizon 2020 programme and was developed by a European Consortium, composed of 9 partners from 6 countries: Belgium, Czech Republic, Greece, Italy, The Netherlands, and Spain. The project built on a distributed selection of case studies in four pilot cities to provide a geographic coverage of the EU: Barcelona (Spain), Milan (Italy), Leuven (Belgium), and Thessaloniki (Greece). All public deliverables can be readily downloaded from our website (https://www.numidas.eu/).

References

1. nuMIDAS Deliverable 2.1: State-of-the-art assessment (2021). https://numidas.eu/wp-content/uploads/2021/11/nuMIDAS_Deliverable_2.1_v1.0.pdf
2. Mitsakis, E., Mylonas, C., et al.: nuMIDAS Deliverable 2.2: Use cases definition – UML model (2021). https://numidas.eu/wp-content/uploads/2021/11/nuMIDAS_Deliverable_2.1_v1.0.pdf
3. Shchuryk, O., Vega, C., Figuls, M.: nuMIDAS: Deliverable 3.1: Report on the orientation of advanced methods and tools and risk assessment (2021). https://numidas.eu/wp-content/uploads/2022/01/nuMIDAS_Deliverable_3.1_v1.0.pdf
4. See https://numidas.eu/index.php/project-deliverables/ for a complete set of deliverables that were used as references, in particular: D3.4, D4.1, D4.2, D4.3, D5.3, and D5.4

Integrated Evaluation of and Vision on Truck Parking in Flanders, Belgium

Sven Maerivoet[1](✉), Bart Ons[1], Sven Vlassenroot[2], and Gwynne Vankaauwen[3]

[1] Transport & Mobility Leuven (TML), Leuven, Belgium
sven.maerivoet@tmleuven.be
[2] Imec, Leuven, Belgium
[3] Tractebel-Engie, Brussels, Belgium

Abstract. The Flemish government launched two related projects in Belgium between 2020 and 2023. First, they rolled out an Intelligent Truck Parking Service (ITPS) along a part of the E17 motorway. Then we supported them to develop a vision on truck parking. The ITPS consisted of an app for truck drivers, combined with information on variable message signs. We evaluated the ITPS so as to provide answers to the following research questions: (i) which technology best measures parking occupancy? (ii) what is the impact of providing parking occupancy information to the truck driver? (iii) how do users deal with the information (user experience)? In order to evaluate the ITPS, we performed a technical analysis, an im-pact analysis, a user acceptance analysis, and performed interviews with stakeholders regarding the eco-system. In addition, the analysis of GPS measurements provided insights into used parking locations and occupancies for developing the vision.

Keywords: Intelligent traffic and mobility management · User behaviour and acceptance

1 Background and Research Questions

The Flemish government launched two related projects in Belgium between 2020 and 2023. First, they rolled out an Intelligent Truck Parking Service (ITPS) testing ground along the E17 motorway corridor between the community of Kalken and the French border (see also Fig. 1). Then we supported them to develop a vision on truck parking. After some quality improvements, the ITPS was commissioned with the launch of an app, a dynamic DATEX II flow, a web interface, and the visualisation on VMS boards. The goal of the evaluation of the ITPS was to provide answers to the following three research questions:

- RQ1: What is the best technology to measure parking occupancy?
- RQ2: What is the impact of providing information to the truck driver on the occupancy of the parking?
- RQ3: How do users deal with the information (user experience)?

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 380–386, 2026.
https://doi.org/10.1007/978-3-032-06763-0_55

During our evaluation, we performed a technical analysis, an impact analysis, a user acceptance analysis, and performed interviews with stakeholders regarding the ecosystem. Finally, we analysed GPS measurements to obtain insights into used parking locations and occupancies for developing a vision on truck parking. We will provide details on each of these aspects in the following sections.

Fig. 1. Geographical overview of the truck parkings along the E17 motorway within the pilot project. (created using QGIS and Microsoft PowerPoint)

2 Analyses

2.1 Technical Evaluation

A number of detection methods (not the same everywhere) have been made available at the various rest areas and service zones, including (i) loop detectors, (ii) barriers and ticket system, (iii) parking sensors, (iv) DSRC readers, (v) traffic sensors, and (vi) truck OBU data. In order to assess the accuracy of the measurement systems, manual counts were performed (which were also initially used to (re)calibrate the systems). During the baseline measurement, one measurement was performed per car park every day, at a specific time, for two weeks between September 15, 2020 and September 28, 2020. This means that a total of 14 measurements were performed per car park over the measurement period. Four measurement moments are possible per day: morning (between 6 am and 12 pm), afternoon (between 12 pm and 6 pm), evening (between 6 pm and midnight) and night (between midnight and 6 am). For each manual count, we determined the absolute and the percentage error, and the mean absolute percent error (MAPE) in relation to the different measurement methods (in the morning when there are fewer trucks, the absolute error is decisive, at night the percentage error is decisive, and in the afternoon both are important). For each type of count, we also looked at the statistical distribution of the errors, and whether outliers occur.

2.2 Impact Analysis

To measure the impact of the different incentives on the occupancy rate, we organised three successive phases:

- Phase 1: this is the baseline measurement (there is no app and the VMS boards are not active)
- Phase 2: an app is available to the truck drivers (Truckmeister, on Android and iOS)
- Phase 3: the app is available and the VMS boards are active.

For each of these phases, we performed an analysis of the parking occupancies, the impact of the VMS, and the results from the OBU data (Fig. 2).

		W38	W39	W40	W41	W42	W43	W44	W45	W46	W47
Fase	Info	14/sep	21/sep	28/sep	5/okt	12/okt	19/okt	26/okt	2/nov	9/nov	16/nov
Fase 1	Nulmeting										
Fase 2	App										
Fase 3	App+VMS										

(1/10) *(22/10)*

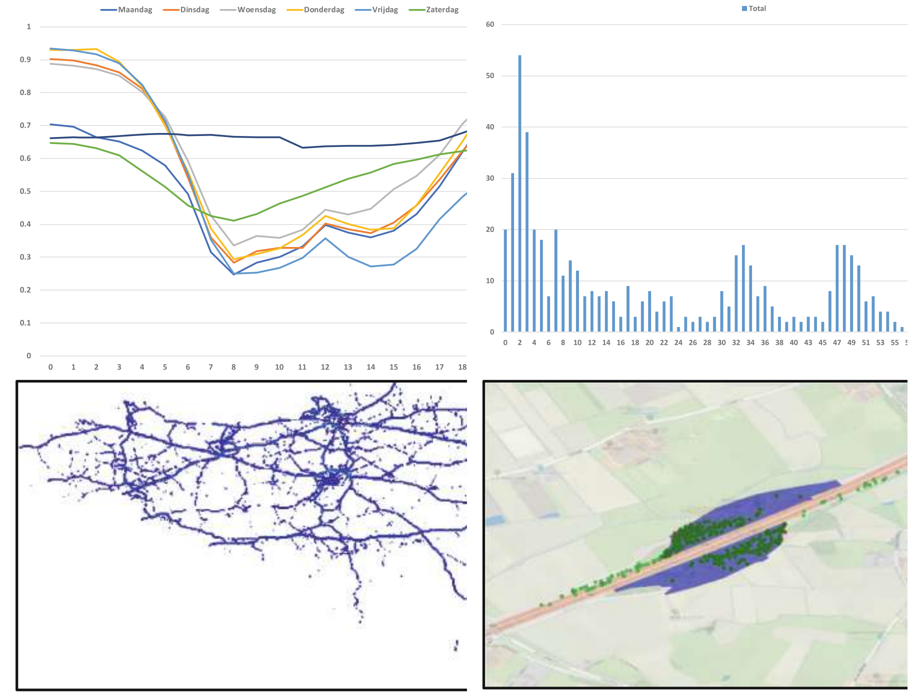

Fig. 2. Example data sources used for the analysis of the research questions: parking occupancies per day of the week (upper-left), distribution of parking durations in minutes (upper-right), truck OBU data from trucks (bottom-left), truck OBU data matched on motorway parking terrains (bottom-right). Figures created using Microsoft Excel, Java, and QGIS.

2.3 Analysis of User Acceptance

2.3.1 Surveying Truck Drivers

We rolled out a survey after the end of phase 3, containing an extensive questionnaire. The survey could not be conducted in the car parks themselves due to the COVID-19 pandemic and the associated measures. As such, we created it digitally. Because the truck drivers had to be reached on the route, the concession holders of the car parks were approached for this, advertisements were posted in specific Facebook groups of truck drivers (both those of the government and others), via LinkedIn posts, and finally it was also distributed among the truck drivers of a specific local distribution company. In addition, QR codes were also made available, which referred to the survey. The survey was available in Dutch, French, English, and German. The survey was closely monitored week after week, with appropriate actions taken to receive statistically significant feedback and a high response rate. In total we had 256 complete (42%) and 349 (58%) incomplete answers for a total of 605 together.

2.3.2 Surveying Truck Drivers

As a final step in our evaluation, we held interviews with nine relevant stakeholders from the ecosystem: transport sector organisations, private concession holders, public stakeholders, service providers, and stakeholders from policy. Through the interview we wanted to gauge the use, the possibilities, and the findings about services that provide information about the occupancy rate. Each interview included the following three main parts that were surveyed from all stakeholders:

- In the first part, we asked about more general aspects of service zones and the extent to which the interviewee or the organisation has knowledge about the services to communicate occupancy rates to truck drivers or others.
- In the second part, we mainly asked how the organisation feels about this service.
- In a final part, we discussed the specific aspects of this service: to what extent is it effective? What impact can it have on the organisation? How do you see the costs and benefits? Etc.

3 Conclusions

3.1 Research Question 1: Which Technology Best Measures Parking Occupancy?

After analysing parking occupancy measurements, the answer to this question appears to be nuanced, whereby the traffic sensors appear to score less well, barriers score quite well, as do parking sensors (if there is no overcrowding). This answer is based on a detailed technical evaluation of measuring methods for the dynamic occupancy rates of truck parkings.

Analyses of the OBU data showed that most trucks that were already in the car park leave before 7 am (these have mainly Poland, Belgium, Lithuania, the Netherlands, and Romania as countries of registration). Conversely, most trucks stop (for the rest of the day) more at Nazareth and Marke than the other car parks. The bulk of trucks routinely arrive at a car park throughout the day, with a slight dip around 8 am. When it comes to regular car park arrivals, Kalken, Nazareth, and Marke have the highest volumes. In these, the parking peak is earlier in the morning/early afternoon. Finally, we also analysed the lorries' dwell times in the car parks. There is a large spread here with local maxima. Nevertheless, the majority of trucks appear to be in the car park for only a very short time (less than 10 min), followed by two groups around 35 and 50 min. These two local peaks occur around 10h00 -11h00.

3.2 Research Question 2: What is the Impact of Providing Parking Occupancy Information to the Truck Driver?

To measure the impact of the different incentives on the occupancy rate, we organised three successive phases: a baseline measurement, an app made available to the truck drivers, and both an app and active VMS boards. Compared to the baseline measurement in phase 1, making the app available in phase 2 had a negligible to nonexistent effect on the parking behaviour of truck drivers (due to the very low penetration rate).

It was not possible to uncover a direct relationship between the messages on the VMS signs and parking occupancies. Nevertheless, the VMS signs were the only variable across the three phases, excluding other factors such as seasonal effects, for example, and we could not observe any other measures taken or behavioural changes. Consequently, the activation of the VMS signs in phase 3 probably did have a positive effect, with the indication of the current occupancy rate presumably leading to a better distribution in the car parks. Note that the VMS signs mainly provided information for the car parks at Nazareth, Kruishoutem, and Marke. Only the VMS sign near Kortrijk also gave information for Rekkem.

3.3 Research Question 3: How Do Users Deal with the Information (User Experience)?

We rolled out our own survey after phase 3, which contained an extensive questionnaire. Almost all respondents (239 of 256, 93%) drove on the A10 / E17 towards France, which gave relevant results. There were 164 respondents (69%) who used the information from the VMS. Almost all (89%) found the information offered useful to very useful. Most of the respondents (76%) thought that the information on the VMS was usually correct, and if this was not the case, it turned out that there was no space left while the signs indicated the opposite (just like with the app). In the future, a large proportion of truck drivers (80%) would find an app useful to very useful, and a smaller proportion (20%) rather useless.

4 Recommendations

The issue of truck parking in Flanders has been studied and plays out in three areas:

1. Truck parking in motorway car parks, especially in service areas during night and weekend rest. This problem occurs on all axes of freight transport on motorways in Flanders. The problem is most pronounced in the Ghent-Brussels-Antwerp triangle and extends along the E17 towards the French border;

2. Truck parking along regional roads and near residential areas. This often involves truck parking by employees who take the truck home and park it locally. It is more of a local problem where parked trucks take up public space and sometimes cause danger;
3. Truck parking for loading and unloading where trucks occupy public domain or wait on the street until it is their turn or their time slot to load or unload. The problem here is that trucks sometimes block public roads.

A total of over 5,300 long-term parking events per day are measured in Flanders. Parking on the secondary road network is often linked to the function of the location. Areas attracting freight traffic (industrial zones, airports, and ports) have more trucks parked along the road. Of the total number of trucks parked, about one-fifth are parked in unofficial car parks.

Our study had a very broad approach, in which a number of aspects were highlighted in great detail. To answer the three research questions, various analyses were performed, experiments were set up, and groups of users were questioned. There are consequently a number of comments with regard to the implementation of the study and the interpretation of the results.

- Before such a study like this can be started, the technical systems to be considered must already be operating under good conditions. This was not always the case, which meant that certain analyses had to be repeated several times.
- Various problems occurred during the study, including detection methods that were set over time with different parameters, problems with technical installations, missing measurements, manual counts that were partly incorrect, etc.
- Further research could be useful, given that a significant number of manual counts are available. This would then make it possible to determine whether data is best fused or not, or recalibrated regularly, and to what extent raw data is suitable for achieving a good fusion. A larger amount of data and a greater mutual comparability of the various locations are strong pluses. Another caveat to the impact analyses is that more data would in any case lead to more stable results.

A comment on the impact analyses is that more data would lead to more stable results anyway. Initially, a longer baseline measurement is recommended (at least four weeks). On top of that, the app penetration rate needs to be significantly higher if we want to measure any impact from that. Finally, it is also true that the spatial scope of the study is limited on two levels: on the one hand, we do not know what happens beyond the border with France (after Rekkem), and on the other hand, parking could also occur on breakdown lanes, and possibly even on (car parks on) the underlying road network.

It is not clear whether VMS or the app can guarantee a better service: VMS appears to be necessary if one gets closer to the car parks. An app allows for more options and can, if necessary, be more complete. The reluctance to an app is more about the use of a mobile phone and road safety. If this service is integrated into the truck itself, the reluctance is less. Providing the service is mainly placed in the hands of the government by some stakeholders, one expects that the government will outline the guidelines and assume responsibility for both the detection facilities and the provision of data.

The problem of overcrowded car parks and the need for a better distribution of use is almost unanimously endorsed. Service provision, as in the pilot project, will certainly make its contribution but should be seen as part of a more global approach: better legislative framework, enforcement, more parking facilities including (private) ones not along motorways, etc. A parking policy/vision therefore urges both regional and European. The main stakeholders seen are the government, parking operators (concessions/private/public), service providers and truck manufacturers. The provision of the service is mainly put in the government's shoes. People expect the government to set the outline and take on both the detection provisions and the provision of data.

There is certainly the will of organisations to contribute to tackle the problem. How far one wishes to go as an organisation to facilitate this service is closely related to the business potential of the service and what role one should assume as a stakeholder (facilitator, data provider, etc.). Minimum information that should be provided is indicating which car park is full or free and the travel time to the relevant (free) car park. Preference is given to using figures/icons and as little text as possible.

Acknowledgements. The projects referenced in this paper have received funding from the European Union's Connected European Facilities (CEF) programme under grant agreement No 2014-BE-TM-0694-S, the Flemish government, Administration of Roads and Traffic, contract No TDV.064.EVCITS_009, and the Flemish Department of Mobility and Public Works.

References

1. Maerivoet, S., Ons, B., Vlassenroot, S., Berghmans, R., Bourgeois, D.: Milestone 11 'Pilot Evaluation Report' of Activity 6 'Truck Parking: Project Evaluation' of the CEF project 'Safe and Secure Infrastructure in Flanders' (2021)
2. Final report 'Developing a vision for truck parking in Flanders', Department of Mobility and Public Works (2023)

The TANGENT Governance Model for Mobility Data Sharing

Antonia Azzini, Marco Comerio(✉), Sabino Metta, and Mario Scrocca

Cefriel, Milan, Italy
azzini@cefriel.com

Abstract. Guaranteed access to mobility for all is one of the conditions for the participation of all individuals in an inclusive society. On the one hand, there is a significant increase in the availability of multimodal mobility services for people. On the other hand, the increase in the movement of people using different mobility services makes traffic management an even more difficult challenge to overcome. Traffic management must move from the traditional management of traffic volumes to managing the different types of available vehicles. For this purpose, the EU-funded TANGENT project is developing new complementary services for optimising multimodal traffic operations. Data from heterogeneous data sources must be appropriately shared between the stakeholders involved in the project to enable the development and testing of these innovative services. This paper presents the methodology adopted by the TANGENT project for formalising the TANGENT Data Sharing Governance model and adequately managing and facilitating data-sharing tasks.

Keywords: Mobility Data · Data Governance · Data Sharing

1 Introduction

The European Strategy for Data [1] defines data-sharing as the process of making the same data source accessible to numerous users or applications. The strategy includes the creation of a common European mobility data space (EMDS) 2] that should facilitate the access, pooling and sharing of data from existing and future transport and mobility data sources. The goal is to overcome technical and legal barriers and to ensure trusted and secure data-sharing by combining data governance frameworks and technical infrastructures. With this objective, the Data Governance Act [3] has been stipulated to increase all the benefits from data for European citizens and businesses by fostering secure data handling and more transparent data governance.

The TANGENT project[1] is developing new complementary tools for optimising traffic operations, elaborating data related to multimodal transport services and considering automated/non-automated vehicles, passengers, and freight transport.

A structured and effective data governance model is crucial for managing and facilitating data-sharing between the different stakeholders involved in the project. This paper

[1] https://tangent-h2020.eu/.

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 387–392, 2026.
https://doi.org/10.1007/978-3-032-06763-0_56

presents the methodology adopted by the TANGENT project for formalising the TANGENT Data Sharing Governance model based on an extensive analysis of the literature and the adoption of best practices.

2 Methodology

Defining a governance model is crucial to regulate the data-sharing between data providers and the organisations involved in the TANGENT project. The methodology adopted by the project to define the data-sharing governance model includes the following propaedeutic activities: (i) extraction of a Data Governance Reference Schema from relevant papers in data governance literature representing the main aspects to consider to handle data governance properly; (ii) usage of the Data Governance Reference Schema to guide the analysis, comparison, and discussion of literature related to data governance in the transport domain and to identify best practices to be adopted in the TANGENT project; (iii) discussion of the best practices and collection of additional information about data governance aspects in dedicated workshops with relevant stakeholders of the TANGENT case study cities (i.e., Athens, Lisbon, Greater Manchester, and Rennes).

2.1 Data Governance Reference Schema

The analysis of relevant papers in data governance literature (e.g., [4–6]) has guided the definition of the *Data Governance Reference Schema* shown in Fig. 1.

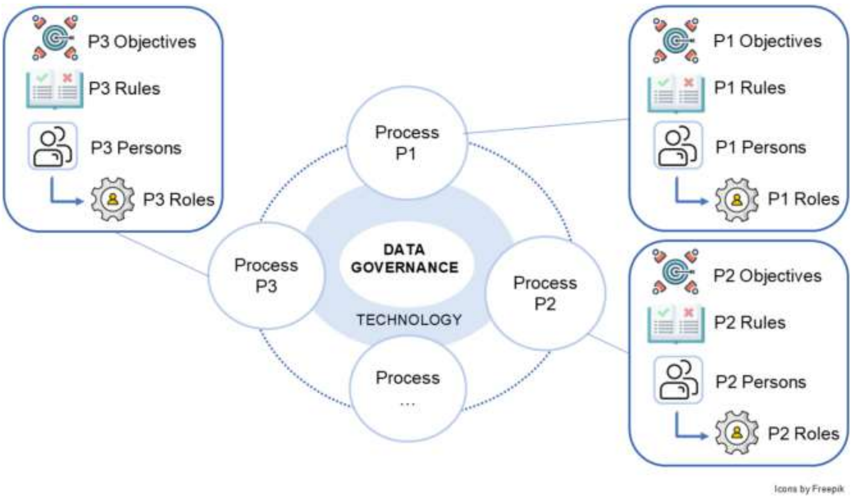

Fig. 1. The Data Governance Reference Schema

Data governance (DG) involves different *processes* (e.g., data collection, sharing, usage) characterised by several *objectives* (e.g., data quality, data accessibility, compliance with European directives). The *rules* that enable each process to achieve the objectives are defined through the DG model. An example is represented by the rules that must be applied and observed to allow for data accessibility in the data-sharing process, such as the acquisition of metadata, the definition of a descriptor, its validation, and so on.

The stakeholders involved in executing the different processes represent all the significant *actors* to be considered in the definition of the DG model. Each actor may assume one or more *roles* (e.g., data publisher, data controller, data owner). A committee or an authority body may be defined to cover the role of the central DG authority. Other roles often considered by a DG model are the data steering committee, the DG officer, the data owner, the data steward (i.e., who oversees aspects related in general to data management, e.g., format, accessibility, description and quality), and the technology steward (i.e., who is responsible for regulating the use of the different technological solutions and tools needed to perform a specific process).

2.2 Mobility Data Governance: Best Practices

The Data Governance Reference Schema has guided the analysis, comparison, and discussion of DG models defined by European research projects (e.g., Transforming Transport[2], HARMONY[3], MOMENTUM[4], NOESIS[5], LEMO[6]) in the transport domain. The analysis of the proposed DG models has highlighted the following best practices:

- definition of rules and establishment of a Technical Management Board for facilitating the processes of data collection, publication, and sharing;
- definition of a metadata catalogue for supporting the data-sharing;
- definition of rules to enforce compliance with European regulations on data (e.g., [7]) and to ensure interoperability and reuse of data and metadata;
- definition of rules to properly govern data manipulation (e.g., data harmonisation and fusion);
- definition of rules to properly govern data reuse.

2.3 Workshops with Stakeholders

The objective of data-sharing in the TANGENT project is to enable the testing of the TANGENT solutions for optimising traffic operations in each case study. Data-sharing must be performed following a structured and effective Data Governance model.

The project has organised multi-actor cooperation workshops with the stakeholders involved in the TANGENT case studies. Each workshop had a dedicated section on data governance to collect additional information about directives, regulations, and DG aspects to be considered for managing data related to the TANGENT case studies.

The following data governance processes were presented and discussed in the first part of each DG section: (i) dataset identification and description, (ii) data access, (iii)

[2] https://transformingtransport.eu/.
[3] https://harmony-h2020.eu/.
[4] https://h2020-momentum.eu/.
[5] http://noesis-project.eu/.
[6] https://lemo-h2020.eu/.

data harmonisation and fusion, and (iv) data usage. Then, in the second part, the participants were invited to actively join a working session aimed at collecting (i) feedback on rules proposed by TANGENT to regulate data-sharing within the project partners, (ii) information on active regulations and directives at the city or national level that TANGENT must be aware of, and (iii) information about stakeholders potentially involved in the process.

The analysis of the collected information has shown that all the participants agreed on the importance of:

- adopting a metadata profile for describing datasets;
- signing an NDA between TANGENT partners and each external data provider to enable and regulate access to and the use of private datasets;
- assuring a high level of data quality of the dataset;
- ensuring compliance with existing licences for handling datasets' access, manipulation and use.

3 TANGENT Data Sharing Governance Model

Based on the results of the propedeutic activities described in Sect. 2, the TANGENT project considered the following elements as crucial aspects for the definition of the TANGENT Data Sharing Governance Model:

- data governance is characterised by different *objectives* (e.g., enhance findability and reuse of data);
- data governance should consider *different data source types*;
- data governance should consider *tools and technologies* adopted for data-sharing;
- data governance refers to different *processes* (e.g., data publication, data access);
- different *actors with different roles* are involved in the governance of each process;
- different *rules* (both external and defined by the project) should be applied to establish effective data governance.

The TANGENT Data Sharing Governance Model focuses on five different processes identified according to the project objectives:

- *Data publication*: the process of describing a data source in a shared catalogue according to a metadata profile;
- *Data quality*: the process of assessing the quality of a data source and its metadata published in the shared catalogue;
- *Data access*: the process of accessing data contained in or exposed from a data source;
- *Data storage*: the process of storing (a portion of) data provided from a data source;
- *Data usage*: the process of using and/or manipulating data from one or more data sources.

This section focuses on the data publication process and describes the main data governance rules regulating it in the TANGENT Data Sharing Governance Model (listed below). These rules require the usage of the *TANGENT Data Catalogue* (i.e., a data catalogue implementing a single solution for the findability and accessibility of all the data sources shared within the TANGENT project) and the *TANGENT Metadata Profile* (i.e., a specification defining metadata fields for the description of the data sources in the TANGENT Data Catalogue). Moreover, the rules must be applied by stakeholders covering two different roles: the *TANGENT data publisher* (i.e., the person responsible for publishing and describing a data source within the TANGENT Data Catalogue) and the *TANGENT Data Catalogue Management Board* (i.e., a group of persons responsible for the management and control of a (set of) data source(s) within the TANGENT Data Catalogue).

- Rule "*Sharing of external data sources for the case study*": The external data sources (both open and proprietary) required for each case study should be published in the TANGENT Data Catalogue. The TANGENT case study leader assumes the role of TANGENT data publisher for all the external data sources related to a specific case study.
- Rule "*Metadata description of data sources*": The TANGENT data publisher should describe a data source in the TANGENT Data Catalogue according to the TANGENT Metadata Profile. The licence selected for metadata in the TANGENT Data Catalogue is Creative Commons CC-BY. Metadata requiring a stricter licence should not be added to the TANGENT Data Catalogue.
- Rule "*Approval of data source publication*": The TANGENT Data Catalogue Management Board validates the publication of data sources shared and described by the TANGENT data publisher. In particular, the TANGENT Data Catalogue Management Board should ensure the publication of data sources following signed NDAs and data licences.

4 Conclusions

Based on an extensive analysis of the literature related to data governance, this paper has provided a Data Governance Reference Schema describing the proper structures and processes to handle data governance properly. The reference schema has been used to guide a more tailored analysis of literature related to data governance in the transport domain. The comparison led to identifying best practices for the TANGENT project.

These propaedeutic activities and the inputs collected from stakeholders during the multi-actor cooperation workshops have been used to define the TANGENT Data Sharing Governance model. The model identifies crucial processes (data publication, data quality, data access, data storage, data usage) to be addressed by TANGENT stakeholders with different roles and following a detailed list of rules.

The TANGENT Data Sharing Governance model represents a best practice that can be easily adapted for the governance of a mobility data space, guiding data sharing between data owners and consumers properly.

Effective data governance, such as the one defined by TANGENT, can facilitate the sharing and reuse of data and enable the creation of new tools for optimising traffic management operations. Through simple rules and agreements on data-sharing, it is, in fact, possible to improve real-time traffic management and mobility planning, reducing road congestion and pollution and increasing our security.

Acknowledgements. The presented research has been supported by the TANGENT project (Grant Agreement no. 955273), co-funded by the European Commission under the Horizon 2020 Framework Programme.

References

1. European Commission (2020): European Strategy for Data. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX%3A52020DC0066. Accessed 03 Jan 2024
2. European Commission: Creation of a common European mobility data space (2023). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52023DC0751. Accessed 03 Jan 2024
3. European Commission: Data Governance Act (2020). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52020PC0767. Accessed 14 Jan 2024
4. Panian, Z.: Some practical experiences in data governance. Int. J. World Acad. Sci. Eng. Technol., 939–946 (2010)
5. Brous, P., Janssen, M., Vilminko-Heikkinen, R: Coordinating decision-making in data management activities: a systematic review of data governance principles. In: Int. Conf. on Electronic Government - EGOV 2016, pp. 115–125. Springer (2016)
6. Weber, K., Otto, B.: One size does not fit all - a contingency approach to data governance. J. Data Inf. Quality **1**(1), 1–27 (2009)
7. European Commission: Commission Delegated Regulation (EU) 2017/1926 supplementing Directive 2010/40/EU with regard to the provision of EU-wide multimodal travel information services (2017). https://eur-lex.europa.eu/eli/reg_del/2017/1926/oj. Accessed 03 Jan 2024

An Integration Framework to Support Data Sharing and Process Tracking in Intelligent Asset Management Systems

Mario Scrocca(✉), Ilaria Baroni, Alessio Carenini, Marco Comerio, and Irene Celino

Cefriel, Milan, Italy
{mario.scrocca,ilaria.baroni,alessio.carenini,marco.comerio,
irene.celino}@cefriel.com

Abstract. The advancement of Intelligent Asset Management Systems (IAMS) in the railway sector can be fostered by integrating data and trustworthy artificial/human intelligence. However, the implementation of solutions for maintenance prescription and optimized intervention plans requires the integration of multiple digital artifacts and the involvement of different stakeholders. This paper discusses an IAMS Support Integration Framework, facilitating the integration of diverse digital artifacts for the implementation of intelligent maintenance scenarios in a multi-stakeholder environment. To support the integration, the framework offers functionalities for enhancing data sharing and guaranteeing process tracking within an IAMS. The paper outlines the framework's requirements and architecture, demonstrates its implementation in practical scenarios from the DAYDREAMS project and presents the preliminary evaluation performed with relevant stakeholders.

Keywords: Intelligent Asset Management System · Data Sharing · Reproducibility · Data Catalogue · Process Tracking

1 Introduction

DAYDREAMS, which completed its activities in May 2023, is a project within Shift2Rail's 3rd Innovation Programme (IP3). The project's overall objective was to advance the integration and use of data and artificial/human trustworthy intelligence for Intelligent Asset Management Systems (IAMS) in the railway domain. The adoption of an integrated Intelligent Asset Management System is highly relevant in the digitalisation process of railway companies, and different solutions can be considered according to specific needs [2]. The DAYDREAMS project focused on a set of maintenance scenarios proposed by three railway maintenance stakeholders to implement intelligent asset management solutions such as (i) prescription based on asset status forecasting (prescriptive analytics, PA), (ii) optimisation of maintenance intervention plans considering different metrics and constraints (multi-objective optimisation, MOO), (iii) contextdriven interfaces to enhance human decision-making (human-machine interface, HMI).

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 393–399, 2026.
https://doi.org/10.1007/978-3-032-06763-0_57

This paper reports on the design and demonstration of an IAMS Support Integration Framework to enable the integration of different digital artifacts for an intelligent maintenance scenario in a multi-stakeholder environment. In this direction, the complementary objectives of the framework are: (i) to enhance data sharing to support better the reuse of solutions developed by different stakeholders, e.g., a prediction model for a specific type of asset may be reused, and (ii) to introduce a trusted way of tracking and auditing the adoption of digital solutions within maintenance process, e.g., to identify users and software components that failed to predict a breakdown. A prototype, implementing the IAMS Support Integration Framework for the three DAYDREAMS scenarios, is discussed and evaluated to demonstrate the proposed solution.

2 Challenges and Related Works

The SHIFT2RAIL[1] Innovation Program 3 promoted the development of solutions for intelligent maintenance in the context of railway asset management [9]. In this framework, DAYDREAMS tries to make a step forward (leveraging the knowledge and results of IN2RAIL, IN2DREAMS, IN2SMART and IN2SMART2 projects[2]), proposing a holistic approach where the information on the current and future status of the assets (e.g., actual faults to repair and probability of news faults) together with the estimated information on the maintenance (e.g., costs and duration), are exploited to deliver options to decisionmakers on how to schedule, improve, and reduce the impact of the maintenance with the purpose to improve the infrastructure maintenance management.

The implementation of such a solution requires the collection of data from different data sources and the implementation of specific software components for data analytics (PA, MOO) and visualisation (HMI). Moreover, an effective deployment in an integrated system is needed for users to adopt it within maintenance operations. A significant challenge is associated with multiple stakeholders involved in maintenance processes, e.g., asset manufacturers, maintenance supervisors, and maintenance teams. Moreover, the definition of digital solutions supporting asset management introduces a set of additional stakeholders to be considered, e.g., sensor manufacturers, vendors of data collection solutions, and data scientists developing models for prediction and prescription. Each stakeholder may contribute or use different digital artifacts and should be made accountable for its involvement in the overall maintenance process. Finally, other actors often have similar needs for implementing intelligent maintenance solutions but adopt different or non-interoperable software solutions, thus slowing down the digitalisation process.

To streamline and enhance the integration, two fundamental sets of guidelines were considered in the design of the proposed framework. On the one hand, the FAIR principles (Findability, Accessibility, Interoperability, Reusability) defined in the context of research data management [10], can be adapted to consider different digital artifacts, from software components [6] to machine learning models [4]. The principles emphasize: (i) the importance of structured metadata for facilitating the findability of digital

[1] https://rail-research.europa.eu/research-development/.

[2] https://projects.shift2rail.org/s2r_ip.aspx?ip=3.

artifacts, (i) the publication of digital artifacts online for accessibility, (iii) the adoption of best practices and standards to facilitate the interoperability of the solutions, and (iv) the need for precise documentation and legal terms of access and usage to support reusability. On the other hand, the Machine Learning Operations (MLOps) guidelines [7] were considered to promote automation in the development and deployment cycle of software components for prediction, prescription and optimisation.

Adopting a shared catalogue of digital artifacts [5] allows actors belonging to different roles to publish and manage descriptions of web services, datasets, and software components promoting their findability and reusability. In the DAYDREAMS project, we focused on how to effectively implement a shared catalogue for intelligent asset maintenance by identifying the relevant digital artifacts and associated requirements. Moreover, while many works in the literature focus on platforms for the development of machine learning models [1], we focused on how to enable the deployment and integrated execution of such models and related software components to support the end users. Finally, considering previous results from the IN2DREAMS project on the usage of blockchain technologies for auditability purposes in the railway domain [8], we define a seamless solution to keep track of operations associated with an intelligent maintenance scenario.

3 IAMS Integration Support Framework

To design a solution supporting the implementation of Intelligent Asset Management Systems, we focused first on the elicitation of the technical requirements associated with it. Integration means first of all the identification or development of the relevant digital artifacts (e.g., datasets, machine learning models, software components) and the availability of a runtime environment for the deployment. Then, it means allowing proper communication between components, which have been developed independently, and guaranteeing their integrated execution according to well-defined processes, i.e., sequence of operations. Finally, the execution of processes must be monitored to ensure that all the stakeholders involved in developing and using the components acted appropriately.

The architecture designed for the IAMS Integration Support Framework, shown in Fig. 1, is based on three macro-functionalities defined to support the elicited requirements: (i) a shared catalogue to enhance the findability and reusability of digital artifacts (IAMS Shared Catalogue); (ii) a runtime area for the integrated deployment and execution of artifacts associated with intelligent maintenance scenarios (IAMS Runtime Area); and (iii) a blockchain-based system to handle the monitoring and tracking of artifacts, processes and operations performed by users (IAMS Process Tracking).

Stakeholders can store and describe in the shared catalogue their digital artifacts, as well as Intelligent Maintenance Packages (IAM Package) defined to group artifacts for integrated execution. The catalogue supports the configuration of metadata profiles for the consistent description of artifacts, and lifecycle processes defining their governance (e.g., process to request/grant access to an artifact). Additionally, it implements programmatic access to the artifacts, allowing stakeholders to develop advanced functionalities (e.g., creation or update of artifacts at runtime). We identified four basic digital artifact types: Dataset, Software Components, Machine Learning Model, and Maintenance Plan. We defined a metadata schema for each of them relying on the generic ADMS-AP specification [3] and on more specific controlled vocabularies targeting each artifact type. The runtime environment enables stakeholders to deploy software components needed for the integrated execution of IAM Packages in a secure and integrated environment. Software components should be containerised[3] and expose well-documented interfaces (e.g., HTTP API documented through OpenAPI[4]) to facilitate their deployment and integration with other components. Finally, the tracking functionality allows stakeholders to monitor the lifecycle of digital artifacts, tracking it together with the runtime operations performed by users. Each lifecycle and runtime interaction is transparently registered as an operation, and a blockchain-based infrastructure is leveraged to provide additional trust among the involved stakeholders. This information can also be saved within the shared catalogue as an Intelligent Maintenance Session (IAM Session) artifact and exported for auditing and reproducibility purposes, ensuring transparency and accountability in managing digital artifacts.

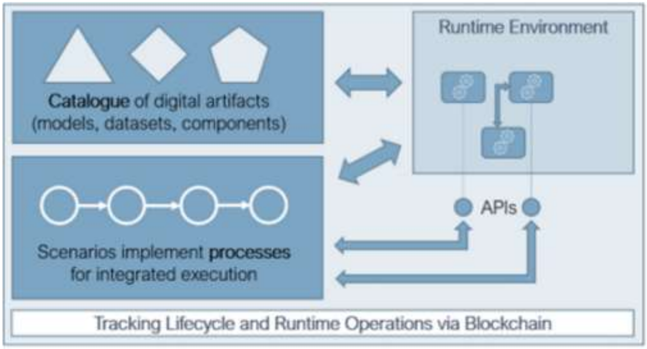

Fig. 1. Overview of the IAMS Integration Support Framework

4 IAMS Integration Support Prototype

The IAMS prototype, developed by DAYDREAMS and shown in Fig. 2, demonstrates the implementation of the IAMS Integration Support Framework and its application considering three intelligent maintenance scenarios proposed by railway stakeholders. The digital artifacts are described in the IAMS Shared Catalogue by different users, and an IAM Package is defined for each scenario grouping together the relevant artifacts. Each IAM Package comprises real datasets provided by railway stakeholders to address a proposed maintenance scenario, innovative software components and machine-learning models for PA e MOO developed by the project, and a dedicated HMI designed by UI experts. The users can access the catalogue through a dedicated web interface and

[3] https://opencontainers.org/.
[4] https://www.openapis.org/.

access all the metadata of the digital artifacts. All the lifecycle operations associated with digital artifacts are tracked and can be visualised through the catalogue. The containerised software components are automatically deployed in the IAMS Runtime Area in isolated sub-networks for each IAM Package. The IAMS Prototype functionalities are implemented transparently by a set of dedicated components, i.e., without requiring a modification of the integrated PA and MOO components. The IAMS Gateway handles the authentication and redirects the requests made by users interacting with the HMI. Additional functionalities enabling the tracking of runtime operations and the implementation of advanced integration with the shared catalogue are implemented by dedicated Package Handler components. In the IAMS prototype, we demonstrated the automatic update of the metadata of a machine learning model every time a training operation is performed, as well as the usage of the shared catalogue to store and retrieve maintenance plans generated and modified by different stakeholders. The IAMS Process Tracking component receives all the tracked operations and registers them through the blockchain-based infrastructure to ensure trust among various stakeholders interacting with the system. The user can retrieve the list of operations performed in each IAM Session and trace back all steps taken during the runtime interaction with the digital artifacts in an IAM Package. All the operations returned to the user are validated through the blockchain, and an alert is produced if the validation is unsuccessful.

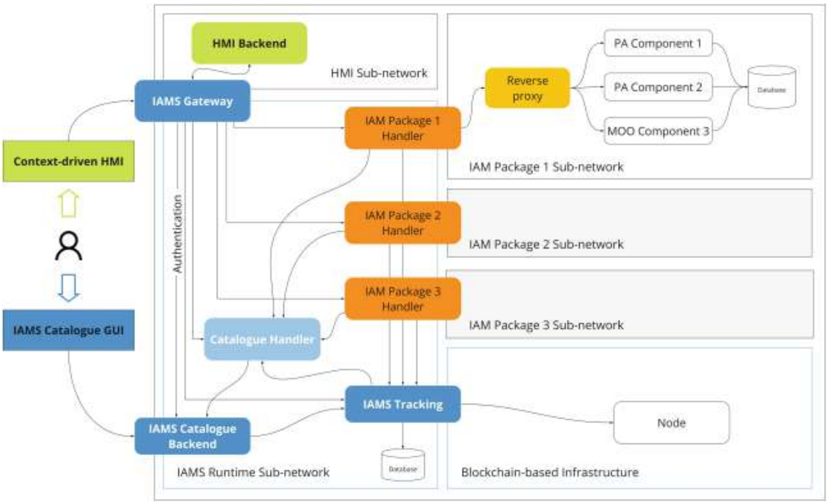

Fig. 2. Overview of components integrated within the DAYDREAMS IAMS prototype to demonstrate the IAMS Integration Support Framework.

The implemented prototype successfully supported the demonstration of the final project results within the three practical use cases for railway maintenance considered. It demonstrated the sharing of digital artifacts through structured descriptors and a common catalogue, the integrated deployment of the software components for end-users to visually inspect and interact with maintenance predictions and prescriptions, and the adoption of a blockchain-based solution to track and audit operations in a trusted way.

Moreover, we performed an initial assessment of the prototype by showcasing its capabilities to six stakeholders from three distinct railway-related companies specialized in infrastructure management and intelligent maintenance solutions. This assessment aimed to gather their insights through interviews regarding the technical feasibility and business potential of implementing this solution within their organizations. The participating stakeholders recognized the benefits of integrating similar functionalities into their operations to enhance the reusability of digital artifacts and foster trust among involved parties. All stakeholders indicated their interest in potentially adopting a similar solution in the near future, with an average rating of 3.5 out of 4. Notably, one stakeholder highlighted the importance of applying these principles to external entities (e.g., external service/product providers) and various units and geographical divisions within large corporations. Regarding potential enhancements, the stakeholders stressed the need to ensure the tool's user-friendliness for individuals with diverse technical backgrounds and to meet stringent security requirements for handling proprietary and confidential digital artifacts.

5 Conclusions

The paper presented an IAMS Integration Support Framework to foster the sharing and reuse of digital artifacts for different intelligent maintenance scenarios, and to enable process tracking mechanisms guaranteeing trust in a multistakeholder environment. The discussion of the IAMS Prototype, developed within the DAYDREAMS project, demonstrated the application of the proposed framework to a set of practical intelligent maintenance scenarios involving different digital artifacts and processes. The implemented prototype and its evaluation by stakeholders highlight how the proposed framework can effectively support the integration of digital artifacts for a generic intelligent maintenance scenario and the potential interest for adoption by companies.

Acknowledgements. This project has received funding from the Shift2Rail Joint Undertaking (JU) under grant agreement No 101008913. The JU receives support from the European Union's Horizon 2020 research and innovation programme and the Shift2Rail JU members other than the Union. This publication reflects only the author's view and the JU is not responsible for any use that may be made of the information it contains.

References

1. Agrawal, P., et al.: Data platform for machine learning. In: Proceedings of the 2019 International Conference on Management of Data, SIGMOD '19, pp. 1803–1816. Association for Computing Machinery, New York (2019). https://doi.org/10.1145/3299869.3314050
2. Crespo Márquez, A., Gómez Fernández, J.F., Martínez-Galan Fernández, P., Guillén López, A.J.: Maintenance management through intelligent asset management platforms (IAMP). In: Emerging Factors, Key Impact Areas and Data Models (2020). https://doi.org/10.3390/en13153762, https://idus.us.es/handle/11441/107312

3. ISA Programme of the European Commission: Asset Description Metadata Schema Application Profile (ADMS-AP) Version 2.0 (2016). https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/asset-description-metadata-schema-adms/release/20
4. Katz, D.S., Psomopoulos, F.E., Castro, L.J.: Working towards understanding the role of FAIR for machine learning. In: Proceedings of the 2nd Workshop on Data and Research Objects Management for Linked Open Science (DaMaLOS) Co-located at the ISWC2021. PUBLISSO-Fachrepositorium (2021). https://doi.org/10.4126/FRL01-006429415
5. Kuzmina, D., et al.: SPRINT project Deliverable D2.5 – Recommendation to IP4 – final version (2021). http://sprint-transport.eu/
6. Lamprecht, A.L., et al.: Towards FAIR principles for research software. Data Sci. **3**(1), 37–59 (2020). https://doi.org/10.3233/DS-190026. https://content.iospress.com/articles/data-science/ds190026
7. Salama, K., Kazmierczak, J., Schut, D.: Practitioners guide to mlops: a framework for continuous delivery and automation of machine learning. Google Could White paper (2021)
8. Spigolon, R., et al.: Improving railway maintenance actions with big data and distributed ledger technologies. In: Oneto, L., Navarin, N., Sperduti, A., Anguita, D. (eds.) Recent Advances in Big Data and Deep Learning, pp. 120–125. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-168414_12
9. Thaduri, A., Galar, D., Kumar, U.: Railway assets: a potential domain for big data analytics. Procedia Comput. Sci. **53**, 457– 467 (2015). https://doi.org/10.1016/j.procs.2015.07.323, https://www.sciencedirect.com/science/article/pii/S1877050915018268
10. Wilkinson, M.D., et al.: The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data **3**(1), 1–9 (2016)

Implementation of a Novel Concept to Unlock Data Value in Multimodal Systems

Ismini Stroumpou[1](✉), Slavica Dožić[2], Danica Babić[2], Josep Lluis Larriba Pey[3], and Milica Kalić[2]

[1] Sparsity Technologies, Carrer del Comte de Güell, 08028 Barcelona, Spain
ismini@sparsity-technologies.com
[2] University of Belgrade - Faculty of Transport and Traffic Engineering, 305 Vojvode Stepe, Belgrade 11000, Republic of Serbia
{s.dozic,d.babic,m.kalic}@sf.bg.ac.rs
[3] Universitat Politècnica de Catalunya, C. Jordi Girona 1-3, 08034 Barcelona, Spain
larri@ac.upc.edu

Abstract. This paper aims to present the basic elements of the Smart Contract Framework (SCF) and the value of data sharing in a multimodal system. SCF is a business process that defines data exchange rules among Transport Service Providers (TSPs) that share the common goal of getting the passenger to his/her destination through a multimodal trip chain. It provides a centralized hub for the generation and management of contracts (i.e., data sharing agreements and smart contracts) via a web platform that allows TSPs to create, negotiate and continuously monitoring and making use of signed contracts. Most importantly, the storage of TSPs' data and all the data sharing processes are executed outside this platform. Moreover, the SCF will provide a data driven environment where exchanged data can be further used for other purposes (e.g., analytics and business relationships' analysis). Additionally, data sharing has the potential to create benefits for TSPs in the form of greater transparency, reduced costs, increased revenue, strengthened business relationships, etc.

Keywords: Multimodality · Transport Service Providers (TSPs) · Seamless door-to-door (D2D) transport · Data Sharing Agreement (DSA) · Travel Companions (TC) · Single Ticketing

1 Introduction and Literature Review

The stakeholders of air transport are integrated into the air transport system. They permanently exchange their data to provide safe, secure, and expeditious transport. However, connection with the other modes of transport has not reached the same level, since there are no regulations and a common channel of communication between stakeholders of different modes of transport. The European Air Traffic Management (ATM) Master plan [1] envisions the integration of air transport into the transport system with all other modes of transport.

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 400–406, 2026.
https://doi.org/10.1007/978-3-032-06763-0_58

Multimodality as a concept implies full coordination and integration of different modes of transport including a single ticket for the whole journey, coordinated timetables, the possibility of remote check-in, etc., [2]. To provide multimodal service and to include air transport in multimodal chains, stakeholders from different modes of transport must collaborate. To ensure successful collaboration, some preconditions need to be satisfied, i.e., stakeholders involved in the multimodal chains need to have data in digital form and a standard format. Moreover, the stakeholders must recognize the benefits of multimodal service and must be willing to cooperate and share their data.

To fulfill a selected goal (e.g., single ticketing, synchronization of timetables etc.), data which need to be shared has to be precisely defined. More data is not mandatory but less can be an issue. Therefore, depending on the selected objective, data sets that need to be shared are defined [3]. The participants involved in providing the multimodal service should trust each other and should show a willingness to start the collaboration and further sign a data sharing agreement (DSA) and/or smart contract (SC) as well. SCs are proposed by [4] representing computerized transaction protocol that executes the contractual terms of an agreement. Although there is no uniform definition regarding the SC, it can be defined from two main perspectives, from the lawyers and from the IT perspective.

Regarding SCs, computer scientists refer to the code, while lawyers are more focused on the legal relationship of obligations [5]. Therefore, we can say that a smart contract is an e-contract composed of a legal part and a software part (i.e., external smarter contract model). Contractual clauses embedded in smart contracts enable automatic enforcement in the case that a defined condition is met [6]. They pointed out the advantages of SC over conventional contracts emphasizing that they reduced risks, cut down administration and service costs, and improved the efficiency of business processes. Data sharing agreement (DSA) is an e-contract that defines the data that should be shared between the two Transport Service Providers (TSPs) who signed the contract. More precisely, a TSP who is a provider should make data available to another TSP who has the role of data consumer. These contracts also set out the specific terms and conditions for the lawful use of such data. In our research there are predefined objectives of data sharing agreements that can be selected by TSPs who want to sign the DSA. Therefore, the objective of a DSA can be single ticketing, synchronization of timetables, optimization of resources, increased ridesharing, etc. This objective determines the required datasets. On the other hand, the SC is an operational contract that details the objective of the DSA and defines specific Triggers (conditions) and Actions (contractual obligation).

Collaborative transport requires that all participants share relevant data, but individual TSPs may hesitate to share (parts of) their information due to reasons of security and competition. However, the precise value of realistic and complete information sharing can be very high. This value can be seen in about 20–30% reduced travel costs, pollution, and congestion when different TSPs cooperate by combining transport services [7]. If TSPs cooperate in larger coalitions research shows that individual TSP can improve profit up to 800% [8]. In [9] can be found a comprehensive review on the modelling of information sharing in collaborative transport and the estimated value of sharing different levels of TSP data.

The objective of this paper is to propose a new business model aligned with EU policies, namely Smart Contract Framework (SCF) which sets common goals for data sharing between TSPs and points out the value of data sharing in the multimodal transport system. After a brief introduction and the literature review, the new concept is described. It is followed by the need and value of sharing the data in the multimodal transport chain. The paper finishes with some concluding remarks.

2 Smart Contract Framework

SCF, proposed in this paper and in the SYN + AIR[1] project, is a process designed to manage (i.e., negotiate, create, modify, terminate and monitor) DSAs and SCs between the two TSPs. To facilitate this process and to materialize the SYN + AIR's findings, a web platform will be developed by the SIGN-AIR[2] project. SC will detail different modalities of the operational collaboration between TSPs, depending on the selected multimodal transport objective(s). To provide a safe and secure data sharing process, the TSPs willing to collaborate should sign the DSA, which will regulate all actions. A DSA defines the "who" and "what" should be shared. Such an agreement could be complemented with a SC [10] which defines the rules concerning "how" and "when".

Fig. 1. Smart Contract Framework phases.

The platform's overall objective is to simplify and standardize the process of data sharing between two TSPs. It facilitates their mutual obligations in the field of multimodal transport. What is the motivation for TSPs to use SCF and sign SC? The usage of SCF will bring certain benefits to the TSPs that can motivate them to join the platform. For example, TSPs will get familiar with a standardized process for data sharing, and they will optimize negotiation time from a legal perspective using technological aids. Moreover, signing a DSA will clarify the minimum data sets required and the objective that should be shared. Finally, by signing a SC, the operational obligations of TSPs will be unambiguously defined, as well as the time when they need to be executed. Most importantly, by using the TSPs platform could identify potential collaborations and to unlock the value of their datasets (Fig. 1).

[1] Synergies between transport modes and air transportation, SYN + AIR, funded by SESAR Joint Undertaking, GA No 894116.

[2] Implemented Synergies, data sharing contracts and goals between transport modes and air transportation, funded by SESAR Joint Undertaking, GA No 101114845.

Contract preparation. It allows the TSPs to create their identity and find potential collaborations, and the platform to prepare for them the relevant contracts' templates. *Registration.* At the beginning of the process the Terms & Conditions (rules of platform provision and framework for DSAs & SCs) should be signed by a person who is authorized and entitled to act on behalf of the legal entity. This person will become platform user and provide access to certain colleagues to sign the User Agreement. The latter is used to identify the authorized end users of the TSP, namely the signing representative, administrative (negotiator), and technical (data/transport expert). *Discovery –Definition.* By the registration at the platform a TSP has three different pathways using a graph database running at the backend (Fig. 2). A TSP based on the reason that it is registered (e.g., needs, ambitions etc.) can: 1) Select a specific goal of DSA that applies to multimodality (e.g., Single ticketing, Synchronization of timetables, Optimization of resources, Increase ridesharing, etc.) The platform proposes a list of potential collaborators and the required data to be provided and consumed by the TSP to be able to fulfill the selected goal. 2) Select to provide the data that TSP wants to offer (create its catalogue). Then based on that data, the platform can indicate a list of goals and specific collaborators (TSPs). 3) Choose a specific TSP, then the platform provides the type of collaboration (goal) that they can fulfill and the data that are required to be shared between them.

Fig. 2. Triangle of collaboration vision.

Therefore, based on the lists/options of potential collaboration the TSPs can make a decision on how to proceed. The outcome is the same in all the pathways, the TSP will have defined the basic elements of the DSA creation which are the Goal of the Agreement, the data that are required to be provided and consumed and also the counterparty of the agreement. Based on these elements the platform provides the adequate DSA and SC template.

Contract Creation. It allows TSPs to proceed in negotiations with one or more of the identified collaborators and create DSAs and SCs. Furthermore, to test and sign their contracts. *Negotiation.* The two TSPs will use the provided template of a DSA. To be more precise on the one hand, a DSA: 1) Sets the rules between Data Provider TSP and Data Consumer TSP; 2) Terms and conditions for the allowed use of data; 3) Set out a Multimodal Transport Goal for the data sharing; 4) This goal determines the required datasets and the Triggers and Actions of the Smart Contract and the issues of Revenue

Sharing and Responsibility Sharing; 5) Legal principles already included in Terms & Conditions (incorporated); 6) Specific obligations of DSA will prevail (negotiation and drafting); 7) Mandatory law, such as GDPR, must be taken into consideration.

On the other hand, a SC is an operational contract that details the Goal of the DSA that defines specific Triggers (a certain condition, event, choice) and Actions (obligation that must be fulfilled in case a Trigger is realized). It also contains terms of Revenue Sharing (how the revenue or benefit of Actions are shared between TSPs) and Responsibility Sharing (how liability is shared (e.g., who will pay delay/cancellation compensation to a traveler in case of certain Triggers).

Settlement enables the use of the aforementioned created e-contracts for the data sharing agreements and the "external smarter contract model" (composed of a legal part and a software part) for the SC.

Post Settlement and Monitoring. It sets all the necessary mechanisms to support an effective data sharing process after the signature of the DSA. *Monitoring* enables smooth data sharing processes based on the contracts signed by the implementation of the TSP's Dashboard and 3rd parties' services such as Travel Companion(s) (TCs). A TC application is an integral part of the platform ecosystem, and it interfaces with the traveler and enables monitoring the contracts signed and passengers' consent (if it necessary). Additionally, in this phase the TSPs are able to select added value services that are offered by the platform such as analytics, optimization algorithms etc. *Modification and Termination.* The TSPs are able to perform modifications and terminate their contracts in accordance with the relevant clauses.

3 Data Sharing and Data Value

We live in a world full of data. However, most transport data remain in the storage of their owner with little or no chance to be reused by other TSPs who may extract more information from them. In each transport system there are a large number of actors but with limited accessibility to data that exist on that market. One of the possible solutions is raising awareness that data has more value when shared. The examples of good practice related to data sharing in the transport sector are Airport Collaborative Decision Making (A-CDM), Global Distribution Systems (GDS) and Mobility-as-a-Service (MaaS). These proven business models have a common approach in their strategy: place data sharing at the center and use data in the best way. It enabled stakeholders to use and analyze the shared data, which produced a new insight into a wider system and resulted in certain actions that have outcomes affected wider society (e.g., enriching the type of service or through the development of additional services, i.e., hub control center, information apps, journey planners, TCs, etc.).

Regarding multimodal service, there are number of opportunities for TSPs that arose from offering the new public transport infrastructure [2], and the one enabled by data sharing are following: increasing interconnection/connectivity, cost efficiency, performing complex analysis of data to guide decision-making processes (in terms of capacity, disruption management, infrastructure usage, etc.), creating personalized services, etc.

Besides the benefits that data sharing can unlock for the TSPs, it can also support developing transport systems towards more sustainable mobility systems in less developing countries (better infrastructure usage, driving innovations, etc.).

4 Conclusion

In this paper the theoretical foundations of a platform for generating, managing and monitoring DSA and SC between TSPs. The starting point are TSPs, who understand the value of data sharing and the multitude of ways to collaborate, but mostly willing to make progress towards seamless multimodal travel. Starting with this assumption, we developed a detailed framework that sets up all the processes required by two willing-to-collaborate TSPs, to write and sign both, DSA and SC. The SCF is a bottom-up concept aiming to facilitate TSPs to unlock their data value.

If well implemented, data sharing has the potential to create benefits for TSPs in the form of greater transparency, reduced costs (either on planning or operational level), valuable customer information provision, increased revenue, strengthened business relationships, etc. It can also unlock benefits for the environment and to society as a whole, such as improved quality of life and increased safety.

References

1. SESAR Joint Undertaking: European ATM Master Plan: Digitalising Europe's Aviation Infrastructure, (2020)
2. Babić, D., Kalić, M., Janić, M., Dožić, S., Kukić, K.: Integrated Door-to-Door Transport Services for Air Passengers: From Intermodality to Multimodality. Sustainability **14**, 6503 (2022)
3. D4.3: Report on Data Flow Model, SYN+AIR, (2021)
4. Szabo, N.: The idea of smart contracts, in: Nick Szabo's Papers and Concise Tutorials (1997)
5. IDSA: Whitepaper Smart Contracts and Distributed Ledger-A Legal Perspective (2017)
6. Zheng, Z., et al.: An overview on smart contracts: Challenges, advances and platforms. Futur. Gener. Comput. Syst. **105**, 475–491 (2020)
7. Gansterer, M., Hartl, R.F.: Collaborative vehicle routing: A survey. Eur. J. Oper. Res. **268**, 1–12 (2018)
8. Schulte, F., Lalla-Ruiz, E., Schwarze, S., González-Ramırez, R. G., Voß, S.: Scalable core and Shapley value allocation methods for collaborative transportation. In: International Conference on Computational Logistics, p. 3. (2019)
9. Los, J., Schulte, F., Spaan, M.T.J., Negenborn, R.R.: The value of information sharing for platform-based collaborative vehicle routing. Transportation Research Part E: Logistics and Transportation Review **141**, 102011 (2020)
10. D5.1: Report on the Smart Contracts Framework, SYN+AIR, (2022)

Knowledgeable Comprehensive and Fully Integrated Smart Solution for Resilient, Sustainable, and Optimized Transport Operations

Giulia Renzi[1](✉), Piergiuseppe Di Gregorio[1], Zoi Petrakou[2], Alexandros Papacharalamous[2], Alexandros Georgakopoulos[2], Camille Leotta[3], Andrea Mutti[3], and Fabrizio Borgogna[4]

[1] Unimore, Via Amendola 2, 42122 Reggio Emilia, Italy
giulia.renzi@unimore.it
[2] Aethon Engineering Single Member P.C., Alexandras Avenue 9, 11473 Athens, Greece
[3] T Bridge Spa, Vía Garibaldi 7-10, 16124 Genova, Italy
[4] Gruber Logistics, Vía Nazionale 75, 39040 Ora, Bolzano, Italy

Abstract. The overreaching goal of this paper is to present how KEYSTONE - a project funded under the Horizon Europe program – aims to develop a knowledgeable, comprehensive, and fully integrated smart solution for resilient, sustainable, and optimized transport operations. Treasuring previous projects that have worked on standardization and learning from the challenges and strength of current European platforms, KEYSTONE has the goal to support the development of a sustainable, efficient, and safe transport system, allowing enforcement authorities to access data for the purpose of checking compliance with rules applied in the transport of goods and passengers. The aim is to tailor standardized digital solutions that can be used from several realities to standardize the transport system. In this paper, the authors present the on-going activities and the methodology that will be followed in order to develop new solutions and demonstrate their validity through an app and through its test with two highly diverse pilots.

Keywords: Logistics · Digital Transport · Digital Transition

1 Introduction

According to EU objectives for a resilient, sustainable, and more efficient transport system, the improvement of smart enforcement concept constitutes a fundamental aspect. The objective is to apply harmonized communication and procedural standards, which create a more comprehensive and harmonized environment for transport enforcement. Such solutions should be viable, effective, and feasible under economical and technical aspects, allowing future integration with newly developed systems and devices in order to exchange relevant information [1]. In 2022, the European Commission launched the new European Standardisation strategy, outlining the importance of increasing the EU's

C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 407–413, 2026.
https://doi.org/10.1007/978-3-032-06763-0_59

competitiveness to enable a resilient, green, and digital economy [2]. This includes the necessity to improve and standardize technological applications, such as plug & play systems, connectivity technologies, and communication standards across Member States.

The following paragraphs are aimed to show the key contributions that will provide the EU-funded KEYSTONE project to achieve these objectives between June 2023 and May 2026. More specifically, they will explain in detail the scopes and the methodologies adopted in the KEYSTONE project, as well as other related EU projects aimed at achieving a more comprehensive, sustainable, and efficient transport system across Europe.

2 KEYSTONE Project Description

KEYSTONE is an acronym and stands for "knowledgeable comprehensive and fully integrated smart solution for resilient, sustainable and optimized transport operations". It is an EU-funded project under the Horizon Europe programme (Grant Agreement number 101103740), which started in June 2023 and ends in May 2026 that aims to develop an ecosystem for the integration and exchange of data and information of an interoperable, intermodal, and seamless transport and logistics system.

The KEYSTONE ecosystem will be established in eight phases which are characterized by a sequential approach. The first three phases are dedicated to the identification of the users' needs, to the requirements of the KEYSTONE ecosystem and to the development of a project vision. The identified solutions create a federation of platforms and standardised data and information exchange, which is then developed and tested within two pilot settings. Finally, a set of guidelines and recommendations are created to replicate and extend at European level the proposed ecosystem. Within these phases and through the development of the project, KEYSTONE aims to have a strong impact on the environment, on the society, on the economy, on the industry sector and on the science and knowledge field. In particular, KEYSTONE has the ambition to improve the ability to use transport data to find impactful actions to decrease the CO2 and NOX emissions, monitor the progress of the biosphere and to intervene on specific problems within the context of future R&D projects or initiatives. This will increase the ability to acquire knowledge and find concrete solutions which improve the quality of the goods and services offered, thus improving the quality of life within society.

3 Methods

KEYSTONE is an ambitious project that aims at revolutionising the European transport system, it stands at the forefront of innovation and efficiency. With its overarching goal of fostering sustainability, efficiency, and safety in the transportation of goods and passengers, KEYSTONE is poised to build upon the ground-breaking work of the Digital Transport and Logistics Forum's (DTLF) Subgroup 2 [3]. This subgroup, focused on "Corridor Freight Information Systems", has paved the way for a common framework for information sharing in multimodal transport and logistics chains. KEYSTONE's mission is to take this foundation and implement the principles of "plug & play" [4], enabling seamless data exchange among diverse stakeholders.

Additionally, KEYSTONE draws inspiration from previous EU-funded projects that acted as support projects to DTLF, FEDeRATED and FENIX, harnessing their knowledge, technologies, and semantic models to create a standardised API reference model. This collaborative effort will not only enhance connectivity but also advance automation, ultimately transforming the European transport landscape towards a more sustainable, efficient, and interconnected future, while also ensuring compliance with the rules and legislation that apply in the logistics sector.

3.1 FEDeRATED Project

FEDeRATED is part of the broader DTLF initiative chaired by the European Commission, Directorate for movement of persons and freight (EC DG Move). Its primary aim is to develop a federated network of platforms ensuring interoperability among existing logistics platforms and solutions. This network allows logistics enterprises and authorities to connect to a single solution, facilitating digital business collaboration [5]. The project's four core principles are: i) plug and play, ii) federation, iii) independent technology services, and iv) safety and security, all aligning with the European Interoperability Framework (EIF) [6]. KEYSTONE is keenly interested in FEDeRATED's semantic model and aims to use the knowledge and model developed. While FENIX focuses on platform connectivity, FEDeRATED enables business interoperability with a chosen solution, aiming for an open and neutral data-sharing infrastructure. This infrastructure can aid KEYSTONE in developing standardized solutions for seamless data exchange and collaboration across logistics platforms [7]. FEDeRATED's semantic model, defining shared data as "linked data" with consistent semantics and syntax, is invaluable for KEYSTONE's standardization efforts.

FEDeRATED aligns with the eFTI Regulation [8] and eCMR Convention [9], aiding KEYSTONE in ensuring regulatory compliance. Insights from FEDeRATED on data sovereignty, where data remains under its source's control, can enhance KEYSTONE's standardized solutions. Understanding stakeholder roles within the logistics ecosystem from FEDeRATED will help KEYSTONE tailor solutions to specific needs. Additionally, FEDeRATED's deployment approach with "Basic Data Infrastructure" (BDI) nodes serves as a practical reference for implementing standardized solutions in real logistics scenarios.

3.2 FENIX Project

FENIX is pioneering a European federated architecture for data sharing in the logistics community, aiming to establish seamless interoperability among various logistics platforms for Business-to-Administration (B2A) and Business-to-Business (B2B) data exchange [10]. Inspired by the European Commission's Digital Transport and Logistics Forum (DTLF), FENIX promotes efficient data sharing, enabling stakeholders to connect, collaborate, and comply with legislation while ensuring trust and security.

KEYSTONE finds FENIX highly relevant due to its guidance on standards in logistics and data sharing, which can inform KEYSTONE's standardization efforts and ensure alignment with industry norms. FENIX's commitment to European Commission policies offers insights for KEYSTONE to support compliance with European regulations effectively. Additionally, FENIX's achievements in interoperability among logistics platforms provide valuable knowledge for KEYSTONE's goal of creating a seamless digital transport ecosystem [11, 12]. KEYSTONE can leverage FENIX's federated architecture as a reference to develop an evolved transport system infrastructure, reducing implementation challenges and risks. FENIX's focus on semantic interoperability and experiences from pilot sites offer valuable lessons for structuring and standardizing data, crucial for developing standard APIs. KEYSTONE will build upon FENIX's governance model to create a framework suitable for its digital transport ecosystem, including roles, procedures, and certification processes.

3.3 Digital Transport and Logistics Forum (DTLF)

The Digital Transport and Logistics Forum (DTLF), established by the European Commission, brings together stakeholders from the transport and logistics sectors to promote digitization. This analysis covers DTLF's history, mandate, structure, and impact on European transport. DTLF addresses the need for digitization in the transport industry, aiming to create a common vision and roadmap for digitalization. It includes stakeholders from public and private sectors to coordinate efforts and support EU-level measures.

The DTLF operates through two subgroups: one focuses on electronic transport documents, and the other on optimizing goods flow through freight information systems. A significant milestone was the EFTI Regulation, which established standardization in the sector. The DTLF later expanded its working groups to address regulatory aspects, reflecting the sector's evolving needs. Subgroup 2, focusing on data sharing, is crucial for "plug and play" activities, aiming to integrate DTLF principles into market dynamics.

Despite challenges in practical implementation, the DTLF's work remains vital. It shifted from unique digital platforms to an "ecosystem of platforms," involving a wide range of participants to address transport and logistics issues comprehensively. While ambitious, the DTLF continues to push for a fully digitized and cooperative European transport sector.

3.4 Overview of the Main European Platforms

The KEYSTONE digital solution will consider the main existing European platforms today (RESPER, ERRU, IMI, TACHONET) used by authorities for the exchange of information in the logistics and road transport sector, in order to develop an innovative application that allow to directly access all relevant data to verify compliancy to social regulations, market rules as well as safety legislation. In terms of safety legislation, RESPER [13], (Réseau permis de conduire) is the EU network of driving licenses. It guarantees effective freedom of movement for license holders issued in the Member States, enhancing road safety, reducing the possibility of document fraud, and avoiding the issuance of multiple licenses in compliance with the EU Directive 2018/645.

The compliance with social road transport rules such as Regulation 165/2014 and 68/2016 on driving times and resting period, is verified by authorities through TACHONET [14], the telematics network for the exchange of information concerning the issuing of a tachograph cart, i.e., tool which records driving time, period of work, other periods of availability, breaks of work and daily rest periods, as well as vehicle speed and distance of the journey.

On the side of market rules, ERRU [15] (European Registers of Road Transport Undertakings) allows Member States to exchange information on road transport companies in line with the requirements of the Regulation 1071/2009 on the access to the profession of road transport undertakings. ERRU provides a means to interconnect the national registries through the exchange of structured (XML) messages to a central hub, so that the competent authorities can mutually exchange information contained in their respective databases.

Moreover, the system IMI [16] (Internal Market Information) allows authorities of each Member State to acquire information about companies that post workers to EU countries and about posted workers, necessary for the definition of inspections in the matter of transnational posting under the competence of the requesting authority, in accordance with directives 2005/36, 2006/123 and 2014/67.

4 Results and Discussion

KEYSTONE will face several key challenges. Establishing seamless access to real-time information for authorities and private stakeholders is crucial for ensuring compliance with transport rules and optimizing logistics. This requires breaking down data silos and promoting interoperability. Additionally, implementing "compliance by design" and "compliance by default" principles is essential to protect users' privacy while enabling efficient data sharing. Standardizing digital information and data exchange among logistics stakeholders is necessary to fully leverage digital technologies and improve freight transport services. Enhancing road safety through standard safety standards and innovative technologies, as well as improving social conditions for transport workers and reducing bureaucratic burdens, will support economic growth in logistics. Lastly, adopting an integrated approach to sustainability, covering economic, social, and environmental aspects, is vital to ensure that solutions align with sustainability goals.

At the end of its three-year project, KEYSTONE aims to create a sustainable, efficient, and safe transport solution through a web app designed to enhance data exchange between enforcement authorities and logistics operators. From defining the state-of-the-art, developing and testing the KEYSTONE solution, to evaluating and replicating it, stakeholders play a strategic role.

The web app will demonstrate the API standard and reference model proposed by the consortium. It will be an open-source service for small transport operators, accessible via eID without additional registration, and implement a plug-and-play concept. The app will provide complete documentation for border controls and highlight data inconsistencies for verification. Using business intelligence, it will extract data from users and fill any information gaps. The two pilot projects will refine the software and model to ensure maximum utility and acceptance.

Declarations. KEYSTONE (Knowledgable comprEhensive and fullY integrated SmarT sOlutioNs for rEsilient, sustainable, and optimised transport operations) has received funding from the Horizon Europe research and innovation programme under grant agreement No 101103740.

Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them.

References

1. Elkarmi, F. (ed.): Power System Planning Technologies and Applications: Concepts, Solutions and Management. IGI Global (2012)
2. European Commission, ICT and standardisation, retrieved from web on 28 June 2024
3. Digital Transport and Logistics Forum (DTLF): Subgroup 2 – Corridor Information Systems. Workprogramme, Version 24092019 (2019)
4. Fanti, M.F., Ukovich, W., Giansante, C., Berardi, A., Di Pierro, B.: Common requirements for the federated architecture of platforms. In: D2.2.2 Fenix Network (2020)
5. Hofman, W., Mulder, H., Hemeleers, R., Lind, K., Lind, M., Out, M.: Federated Library, Proposition, Architecture, Semantics. Accessed 18 Sept 2023
6. European Commission, DIGIT: National Interoperability Framework Observatory (NIFO). Accessed 18 Sept 2023
7. Hofman, W., Bouter, C., Burghoorn, M., Boertjes, E., De Graaf, E., D'Auria, A.: Towards a mobility data space: data sharing via linked semantic data, an example for eFTI. In: Transport Research Arena (TRA) Conference Proceedings (2022)
8. European Parliament: Regulation (EU) 2020/1056 on electronic freight transport information. In: Official Journal of the European Union (2020)
9. Buhl, H., et al.: Recommendation for the use of standards for corridor information systems and T&L services. In: D6.5.1 Fenix Network (2020)
10. Fanti, M.P., Ukovich, W., Berardi, A., Di Pierro, B., Giansante, C.: Existing platforms and technical requirements for the architecture of the federated ecosystem. In: D2.2.1 Fenix Network (2020)
11. Polytechnic University of Bari (POLIBA): Governance structure. In: D2.5 Fenix Network (2021)
12. European Commission, Directorate Energy and Transport: RESPER Project details
13. European Commission: TACHOnet website. Accessed 18 Sept 2023
14. European Commission: European Register of Road Transport Undertakings (ERRU) website (retrieved from web on 18 September 2023)
15. European Commission: Internal Market Information System (IMI) website. Accessed 18 Sept 2023
16. European Parliament: Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). In: Official Journal of the European Union (2016)