10 Video and sensor streaming for remote train operation

Katharina Hartmann, Oliver Röwer;
DLR, Institute of Transportation Systems, Braunschweig, Germany

10.1 Introduction

Work package 2 of the EU-funded TAURO project (Technologies for the Autonomous Rail Operation) aims to provide a solution for remote driving and remote command. For remote train operation (RTO) video and sensor streaming plays a vital role. Based on considerations in the projects Lucy Train Lab and 5G Living Lab and an additional literature review this paper gives an overview of existing approaches and architectural solutions for video and sensor streaming in a safety-critical environment. Furthermore, we propose an architecture that can be utilized for video and sensor streaming for remote train operations and point out how these activities could be integrated into the operational and technical workflow.

10.2 Technological background

This section presents relevant technologies for streaming videos from one computer to another. First, the terms aspect ratio, video resolution, and bitrate are introduced. After that, different techniques for video encoding are presented followed by protocols for video transmission.

10.2.1 Video streaming

When dealing with the topic of video streaming there are some aspects to be considered. The first item is the **aspect ratio** which describes the ratio between the width and height of a video and/or screen. Nowadays, common aspect ratios are 16:9, 21:9, or 4:3. The **video resolution** explains how many pixels each frame in the video has. Typical resolutions are high definition (HD) with 1280x720 pixels and full HD with 1920x1080 pixels. The third factor, **video bitrate**, describes the amount of data that is processed per unit of time. The video bitrate is usually measured in bits per second, common bitrates for HD videos vary between 20 and 30 Megabits per second (Mbps). The higher the bit rate, the higher the level of detail. Thus, it is common for a high resolution to also use a correspondingly high bit rate.

10.2.2 Video Encoding

When mapping the situation, the camera first generates a stream of raw data. Depending on the resolution, this can already exceed the entire available upload bandwidth of the 5G connection. Therefore, it is state of the art for video streaming and many other video-based services to use encoding and decoding algorithms, so-called video codecs, to compress the data to be transmitted. However, not all video codecs are suitable for all purposes. While the video codec H.264 is used for many purposes like video conferences and HD television, its successor H.265 became established for ultra-high definition supporting resolutions up to 8192x4320 pixels. The codec MPEG4 was, similar to H.264, developed for audio and video encoding in multiple multimedia applications. Another open-source codec is VP8. In [1] the video codecs

H.264, MPEG4, and VP8 are compared regarding the quality of the video stream and end-to-end latency. The results show that MPEG4 outperforms VP8 and H.264 in terms of quality while H.264 has the best end-to-end delay.

10.2.3 Video transmission

Regarding the transmission of the video and sensor data from the onboard to the trackside the following aspects have to be considered:

- protocols in the application layer (RTP, RTCP, RTSP, WebRTC, HTTP(S))
- protocols in the transport layer (UDP, TCP)
- access / multiplex technologies (NOMA, TDMA, FDMA, OFDM)

In communication and entertainment systems that use streaming media, the Real-time Transport Protocol (RTP) is commonly used to provide audio and video over IP networks. The real-time Control Protocol (RTCP) is used in combination with RTP to track transmission statistics and quality of service (QoS) [2]. Media sessions between endpoints are created and managed using RTSP [3]. Play, record, and pause are just a few of the commands that make real-time control easier.

All three protocols, RTP, RTCP, and RTSP, must be used for a video streaming application to work properly.

WebRTC is an Application Programming Interface (API) that enables real-time audio and video communication between web browsers and mobile applications [4, 5]. The video codecs H.264, VP8, and VP9 are supported by WebRTC across all browsers and once more utilize RTP. When the endpoints do not have public IPs, that is, they are hidden behind a NAT or firewall device, WebRTC enables media session establishment by utilizing Session Traversal Utilities for NAT (STUN) and Traversal Using Relays around NAT (TURN) servers to commence the connection.

A secure communication protocol for computer networks is HTTPS. Dynamic Adaptive Streaming over HTTP (DASH), commonly known as MPEG-DASH, is a high-quality adaptive bitrate streaming solution for media supplied via traditional HTTP web servers over the Internet. On the one hand, the statelessness of DASH offers the benefit of preventing streaming failures during handover procedures. Contrarily, every video chunk is supplied in response to an HTTP GET request, which adds at least 1 second of latency even when a chunk's length is reduced. [6]

The most used transport layer protocols are User Datagram Protocol (UDP) and Transmission Control Protocol (TCP). Since TCP is connection-oriented, a connection must first be made between the client and server in order to send data. Additionally, TCP offers applications on an IP network reliable, orderly, and error-checked delivery of a stream of data. TCP is considerably sluggish because of the guaranteed delivery of data, which is done by retransmitting data when a packet is defective or lost. As a connectionless protocol, UDP cannot ensure that sent data will arrive in the correct sequence or be protected from duplicate transmissions. UDP is substantially faster than TCP since there aren't any handshaking conversations or retransmissions like there are with TCP. UDP is, thus, often employed in time-sensitive applications.

Within the access technologies, we consider Frequency-division Multiple Access (FDMA), Time Division Multiple Access (TDMA), Orthogonal Frequency Division Multiplex (OFDM), and Non-Orthogonal Multiple Access (NOMA).

In the Time Division Multiple Access (TDMA) and the Frequency Division Multiple Access (FDMA) multiple users will be allocated different time and frequency resources in an orthogonal manner. NOMA allows users to share the same time and frequency resources which leads to higher spectral efficiency [7]. Furthermore, according to [8] NOMA achieves better performance compared to other orthogonal multiple access techniques.

10.3 Related Work

In recent years there has been a lot of work on the topic of remote driving and, thus, video and sensor streaming from vehicles to remote operator's places. As there is only little research on video streaming for RTO, work in the context of remotely driven cars is also considered.

Bećirbašić et al. proposed 2017 a video-processing platform for semi-autonomous cars over 5G [9]. To prevent blind spots the test vehicle was equipped with 10 cameras with overlapping fields of view with a resolution of 1280*720. For remotely driving a car the authors stated there has to be a maximal latency of 50 ms, which means that 4G networks are ineligible. In addition to the cameras, multiple sensors were integrated into the vehicle. In [10] they enhanced the considerations to a virtual cockpit application with three screens and proposed a solution for a maximum target speed of 50 km/h. To comply with a maximal latency of 70 ms (so the car would not have moved more than 1 m) the H.264 video codec, a bitrate of 10-15 Mbps, and a lower video resolution of 720*480 for two of the three video streams were used.

In 2018 Kang et al. tested various communication protocols, bitrates, and resolutions in the context of remotely driven cars [11]. A single raw video was compressed with H.264 or MPEG-4 video codec which resulted in a reduction of 5-15 % per frame. The researchers transmitted the video via UPD using LTE- or WiFi-based communication. In addition, multiple bitrates (0.5, 1, 4 Mbps) and different resolutions (320x240, 640x480, 1280x960) were considered. To record the relevant parameters, they used an app for measuring e.g. the latency. There was a median latency of 100 ms for LTE-based connections and 50 ms for transmissions via WiFi. The frame loss rate varied from 0.5% to 2% for different bitrates and resolutions. Even though the paper described various combinations the researchers did not recommend a resolution, bitrate, or video quality based on their results.

Yu and Lee proposed in their paper 2022 a UDP-based video transmission for remote driving systems [12]. While the transmission of commands requires control mechanisms to prevent data loss or faulty information, the main aim of the video transmission is low latency. As UDP has a limit regarding packet size the frames have to be split. The proposed solution was tested with model cars at a speed of ca. 65 km/h.

RTO is in combination with Automatic Train Operation (ATO) a growing market in the railway sector since 2019.

In September 2019, Lucy Train Lab performed remote driving via 5G for the first time [13]. Lucy Train Lab is a modified train with integrated cameras to monitor what's happening on and around the train. One use case is a remotely operated trip from a storage area to a platform. For this, the camera images must be transmitted to the remote operator without any noticeable delay and in good quality. The remote operator's workstation has a display that shows the information of the Human Machine Interface (HMI) of the train and also an operating lever to properly steer the train. The train is equipped with an intermittent automatic train running control (PZB, short for German "Punktförmige Zugbeeinflussung"), Continuous Train Control (LZB, short for German "Linienzugbeeinflussung"), an ETCS-capable on-board unit (OBU), LiDAR sensors, thermal cameras as well as GNSS and cellular antennas. Furthermore, the driver's cab has an inertial measurement unit (IMU), and stereo and infrared cameras with a high-speed ethernet interface. A high-definition industrial camera (Nano-C1930 Teledyne DALSA) with 80 frames per second (fps) was chosen. This camera has advantages over common surveillance cameras as there is no internal compression and, thus, no additional latency. The individual frames were compressed via Motion JPEG. The video was streamed from the onboard system via VPN to the remote operator's workstation at 25 fps and approximately 85 kbytes/frame in 1200 x 1000 pixel format. In addition to transmitting camera images, an application on the remote operator's workstation displays the train's platform-specific geographic location. Work on Lucy will continue in follow-up projects as part of the 5G Living Lab.

In 2020 the results of the TC-Rail project, a partnership formed by SNCF, Thales, Actia Telecom, CNES, and Railenium, on safe remote driving without a driver in the train cabin were presented [7]. The project was the first proof of concept for remotely driving a train without European Railway Management System (ERTMS) infrastructure at a maximum target speed of 100 km/h. The paper compares LTE and 5G systems and concludes that 5G is in all situations superior to LTE. Additionally, the case of two trains in the same cell was considered. The authors proposed using NOMA for the allocation of resources in the channel to guarantee good performances for both trains.

10.4 Proposed architecture for video streaming

Based on the literature review and experience from projects like Lucy Train Lab we propose a general architectural solution for video streaming in the context of RTO. The proposed architecture for streaming video data is shown in Figure 10-1 and is based on the high-level functional architecture for remote driving that was developed within the project TAURO.

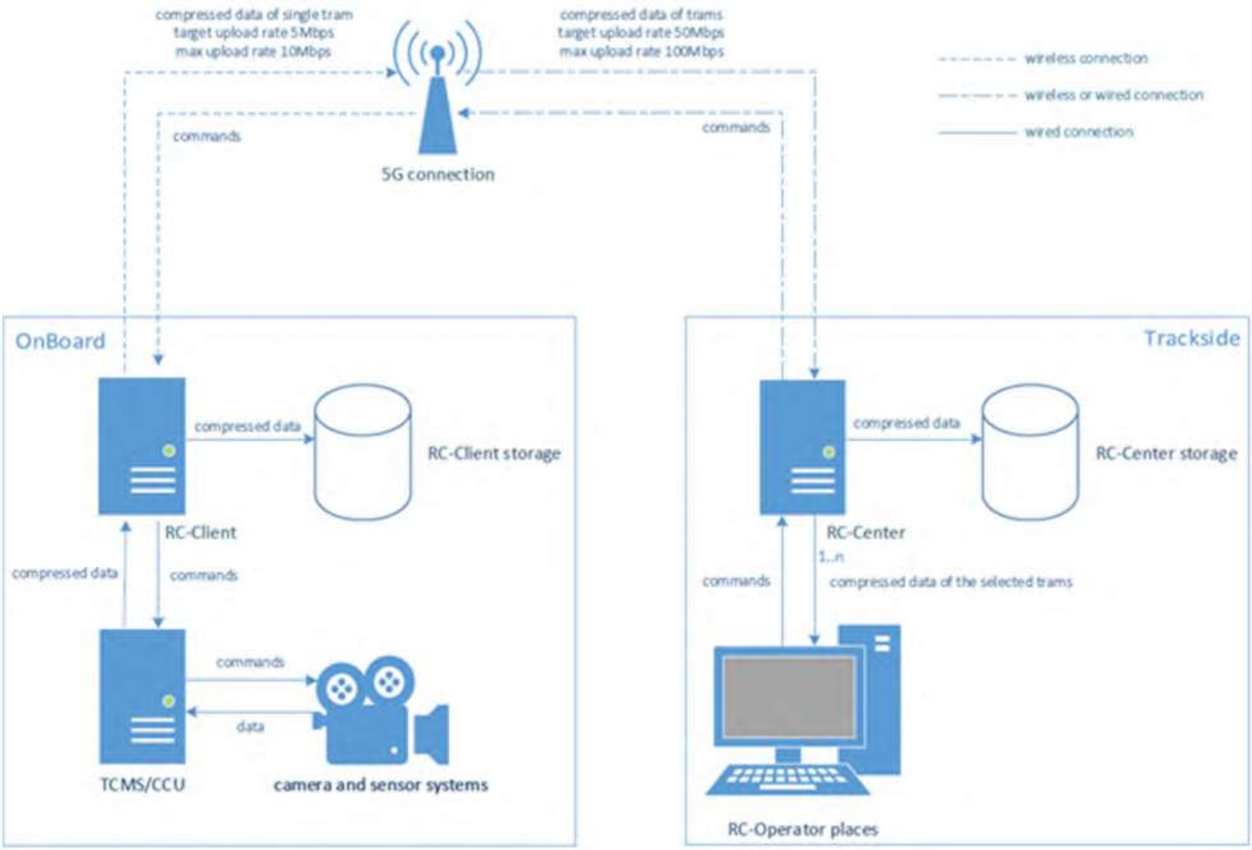

Figure 10-1: Proposed architecture for video and sensor streaming

The presented system consists of an onboard and trackside system with a Remote Control (RC) client and a RC center respectively. Additionally, the onboard system includes Train Control and Monitoring System (TCMS) to which the cameras and sensors are connected and a storage for recorded video and sensor data. We propose the use of industrial cameras without internal compression to avoid additional delay due to internal compression. Based on similar research and projects in [12], [11], and [9] we propose a video resolution from 640*480 px up to 1280*720 px. As the video has to be compressed before the transmission the video codec H.264 should be used to ensure low latency and suitable quality at once. The storage is crucial since it allows for reading out of what occurred during the absence of a corresponding transmission, even if a connection is lost. This could be particularly vital for the study of accidents.

The RC client transmits data to the RC center using a 5G connection to a linked radio tower at a targeted upload rate of minimum 5 Mbps which is considered a requirement for RTO within TAURO. For the transmission of video and sensor data the network protocol UDP and the RTP respective RSTP are frequently used. Thus, we propose to use UDP in the transport layer and RTP in the application layer. As no browser application is involved, WebRTC does not have to be used. A VPN connection is reasonable to decrease glass-to-glass latency and safeguard the data from outside cyber assaults. It is possible that the 5G tower is reached by a number of rail vehicles. Therefore, NOMA should be considered as an access technology to allow for multiple trains in the same cell. The tower transmits all incoming data from the various RC clients to the corresponding RC center. Here, the data is also stored in order to support, among other things, an accident analysis even if the data already stored on the rail vehicle is no longer accessible.

The video and sensor data are now decompressed and transmitted to the designated RC operator place. The aspect ratio depends on the preferred screen size at the remote workplace. The RC operator now has the possibility to process the data according to his activity and to transmit necessary control signals and instructions to the train.

10.4.4 Extension of the architecture for additional sensor information

In addition to the visual perception of the train's surroundings by the remote driver by means of using a camera, further sensors can be installed in the train. The remote driver needs knowledge of the train metadata in order to operate remotely. This incorporates data on the position obtained using GPS or GNSS receivers as well as speed, as determined, for instance, by an odometer or radar sensor. Additionally, while operating the train, train drivers would experience both audio and visual sensory experiences. From a psychological perspective, adding noises through the proper microphones can be a helpful complement to the video stream because it can improve the remote operator's situational awareness, especially in case of low video quality [14]. Additionally, the video data can be processed so that automatic object detection in the image aids in remote control operation or even starts automatic procedures, like emergency braking, to prevent a collision with an object on the track. Radar or lidar systems are examples of sensors that assist in the identification of potential threats on the track.

The data volumes of sensor systems in addition to the camera sensor technology also have an effect on the network load and, in particular, on the upload rate required per train. It can be assumed that the resulting data volumes are comparable to or even exceed those of the video stream. A remedy for this can be a digital map in which objects detected by the other sensor systems are noted and ultimately only these objects are transmitted.

10.4.5 Integration of video and sensor streaming in the operational and technical workflow

To pave the way for RTO the video and sensor streams have to be integrated into the remote operator’s workflow.

The first aspect to be considered is the coherence between the sensor data and the video stream. To ensure that valid information is presented to the remote driver, the video data that the remote driver sees should have been generated at the same time as the simultaneously arriving sensor data. However, other shared variables like system status have to be synchronized as well. Therefore, the architecture has to ensure that outdated video and sensor data are discarded. This could be implemented with a time stamp that is created onboard and verified on the remote side. If the incoming data have a time stamp that is more than one second behind the data should be discarded.

Furthermore, there are a variety of possible causes that can lead to the impairment of the remote control. In the following the interference that directly affects video and sensor data transmission is explained.

Once a connection has been established and the remote driver is active, it is conceivable that the train will enter an area that is not sufficiently well covered by 5G. In this case, the available bandwidth and thus the potential upload volume drop. Thus, the high-resolution camera images may no longer be transmitted in the required quality. These deteriorated connection conditions can also lead to an interruption in communication. A comparison with the predecessor of 5G shows that, although the construction and expansion of the 4G network began in 2010, even today there are repeated connection drops during train journeys. Since remote driving on open tracks is also considered in TAURO, this problem may also occur with 5G. The project also considers shunting in shunting yards, and as coverage of such a delineated area is much easier to implement, the occurrence of this type of interference can be more easily prevented here.

A similar interference occurs when there is enough bandwidth available, but it has to be divided among too many participants. Particularly when multiple trains are in the shunting yard, the 5G network can become overloaded. In this case, it is necessary to consider whether one or more masts should be erected specifically for the shunting yard in order to counteract overload.

In addition to inadequate bandwidth, disruption can also occur due to hardware defects on the train itself. For example, the cameras, power or data lines, and computer hardware that processes the recorded video and sensor data could fail. As the video and sensor data are stored on the train, there

Interference with video and sensor data transmission can also result from the train having an insufficient amount of power to turn on or run the required hardware.

However, the disruptions that are specifically relevant for video and sensor streaming are poor connection quality, network congestion, or the disruption of corresponding hardware and software elements such as sensors, cameras, and systems that are then supposed to make this data available to the remote operator via the network.

If disruptions cause such a poor connection that remote control is not possible or even a connection abortion there have to be procedures so the drive is not interrupted. The question of which procedures should be applied in such a case and what aspects determine the triggering condition requires further research.

10.5 Conclusions

In summary, this paper gives an overview of necessary considerations regarding remote train operation. First, the theoretical background information was presented. Then, we discussed related work in the context of remote driving of cars as well as trains. Especially the experiences from the project Lucy train laboratory and the project 5G Living Lab have influenced this paper. Based on the literature review we proposed an architectural solution for video and sensor streaming for RTO. Especially in 5G networks, new access technologies, e.g. NOMA, should be considered to allow for multiple trains in one network cell. Furthermore, we considered possible disturbances in the technical and operational workflow and how the proposed architecture can prevent long-term interruptions.

10.6 References

[1] ROWSHANRAD, Shiva ; NAMVARASL, Sahar ; JAMASB, Behnaz ; KESHTGARY, Manijeh: Video Codec Standards Comparison for Video Streaming Over SDN 5 (2015)

[2] HENNING SCHULZRINNE ; STEPHEN L. CASNER ; RON FREDERICK ; VAN JACOBSON: RTP: A Transport Protocol for Real-Time Applications. RFC Editor, 2003. URL https://www.rfc-editor.org/info/rfc3550

[3] HENNING SCHULZRINNE ; ANUP RAO ; ROB LANPHIER ; MAGNUS WESTERLUND ; MARTIN STIEMERLING: Real-Time Streaming Protocol Version 2.0. RFC Editor, 2016. URL https://www.rfc-editor.org/info/rfc7826

[4] CULLEN JENNINGS (CISCO), HENRIK BOSTRÖM (GOOGLE), JAN-IVAR BRUAROEY: WebRTC 1.0: Real-Time Communication Between Browsers. URL https://www.w3.org/TR/webrtc/. – Aktualisierungsdatum: 2022-01-26 – Überprüfungsdatum 2022-05-17

[5] ADAM ROACH: WebRTC Video Processing and Codec Requirements. RFC Editor, 2016. URL https://www.rfc-editor.org/info/rfc7742

[6] MPEG, I.: Information technology-dynamic adaptive streaming over http (dash)-part 1: Media presentation description and segment formats. In: ISO/IEC MPEG, Tech. Rep (2012)

[7] ALSABA, Yamen ; BERBINEAU, Marion ; DAYOUB, Iyad ; MASSON, Emilie ; ADELL, Gemma Morral ; ROBERT, Eric: 5G for Remote Driving of Trains, Bd. 12574. In: KRIEF, Francine; ANISS, Hasnaâ.; MENDIBOURE, Léo; CHAUMETTE, Serge; BERBINEAU, Marion (Hrsg.): Communication Technologies for Vehicles : 15th International Workshop, Nets4Cars/Nets4Trains/Nets4Aircraft 2020, Bordeaux, France, November 16–17, 2020, Proceedings. 1st ed. 2020. Cham : Springer International Publishing; Imprint Springer, 2020 (Springer eBook Collection, 12574), S. 137–147

[8] MA, Zheng ; ZHANG, ZhengQuan ; DING, ZhiGuo ; FAN, PingZhi ; LI, HengChao: Key techniques for 5G wireless communications: network architecture, physical layer, and MAC layer perspectives. In: Science China Information Sciences 58 (2015), Nr. 4, S. 1–20

[9] BECIRBASIC, Damir ; MOLNAR, Marijana ; LUKAC, Zeljko ; SAMARDZIJA, Dragan: Video-processing platform for semi-autonomous driving over 5G networks. In: 2017 IEEE 7th International Conference on Consumer Electronics - Berlin (ICCE-Berlin) : IEEE, 2017, S. 42–46

[10] MOLNAR, Marijana ; BECIRBASIC, Damir ; LUKAC, Zeljko ; SAMARDZIJA, Dragan: Video processing for virtual cockpit over 5G networks. In: 2017 IEEE 7th International Conference on Consumer Electronics - Berlin (ICCE-Berlin) : IEEE, 2017, S. 170–172

[11] KANG, Lei ; ZHAO, Wei ; QI, Bozhao ; BANERJEE, Suman: Augmenting Self-Driving with Remote Control. In: KIM, Minkyong (Hrsg.): Proceedings of the 19th International

Workshop on Mobile Computing Systems & Applications. New York, NY : ACM, 2018 (ACM Conferences), S. 19–24

[12] YU, Yang ; LEE, Sanghwan: Measurement-Based Evaluation of Video Streaming Method for Remote Driving Systems. In: 2022 International Conference on Information Networking (ICOIN) : IEEE, 2022, S. 136–139

[13] TOBIAS MELZER: Konzeption neuer Technologien für autonomes Fahren. In: DER EISENBAHNINGENIEUR (2020), Nr. 02

[14] GRIPPENKOVEN, Jan ; MEIRICH, Christian ; ROTH, Michael ; CASPAR, Mirko ; HUNGAR, Hardi: Teleoperierte Triebfahrzeugführung als Rückfallebene der Hochautomation / Teleoperation as a fallback solution for highly automated rail traffic. In: SIGNAL + DRAHT (2020)

10.7 Acknoledgement

The project TAURO has received funding from the Shift2Rail Joint Undertaking (JU) under grant agreement No 101014984. The JU receives support from the European Union's Horizon 2020 research and innovation programme and the Shift2Rail JU members other than the Union.

Disclaimer: This dissemination of results reflects only the authors' view and the Shift2Rail Joint Undertaking is not responsible for any use that may be made of the information it contains.

10.8 References

[1] I. Errandonea, S. Beltrán, and S. Arrizabalaga, "Digital Twin for maintenance: A literature review"; Comput. Ind., vol. 123, 2020, doi: 10.1016/j.compind.2020.103316

[2] SIA Project official website: https://siaproject.eu/

[3] I. Errandonea, J. Goya, U. Alvarado, S. Beltrán, and S. Arrizabalaga, "IoT Approach for Intelligent Data Acquisition for Enabling Digital Twins in the Railway Sector"; in 2021 International Symposium on Computer Science and Intelligent Controls (ISCSIC), Nov. 2021, pp. 164–168, doi: 10.1109/ISCSIC54682.2021.00039.

[4] LinX4Rail project overview; https://projects.shift2rail.org/s2r_ipx_n.aspx?p=LINX4RAIL; last access: 02.03.2022

[5] In2Rail: Innovative Intelligent Rail. http://www.in2rail.eu/; last access: 20.05.2020

[6] Karlsson, M. and Wegele, S. (2017): Digitalisation - Toward seamless connectivitiy; in: Global Railway Review, Vol. 23, issue 05.

[7] LinX4Rail: MS6 report - Preliminary input on CDM for use case developments in WP4; 29.04.2021

[8] O. Betari, S. Filali, A. Azzaoui, M. A. Boubnad, "Applying a Model Driven Architecture Approach: Transforming CIM to PIM Using UML"; International Journal of Online and Biomedical Engineering (iJOE), Vol. 14, No. 9, 2018.

[9] Elmhadhbi, L., Mohamed-Hedi Karray, and Bernard A.: Toward the use of upper-level ontologies for semantically interoperable systems: An emergency management use case. Enterprise Interoperability VIII. Springer, Cham, 2019. 131-140.

[10] Debbech, Sana, Simon Collart Dutilleul, and Philippe Bon. An Ontological Approach to Support Dysfunctional Analysis for Railway Systems Design. J. Univers. Comput. Sci. 26.5 (2020): 549-582.

[11] Nilsson, Jacob, and Fredrik Sandin. Semantic interoperability in industry 4.0: Survey of recent developments and outlook. 2018 IEEE 16th international conference on industrial informatics (INDIN). IEEE, 2018.

[12] L. Obrst. Ontologies for semantically interoperable systems . In: Proceedings of the twelfth international conference on Information and knowledge management - CIKM '03. New York, USA: ACM Press, Nov. 2003, p. 366

[13] Gruber, T. R. (1993). A translation approach to portable ontology specifications. Knowledge acquisition, 5(2):199–220.

[14] http ://www.integrail.eu/

[15] https://ontorail.org/

[16] Hoinaru, O., Mariano, G. et Gransart, C. (2013). Ontology for complex railway systems application to ERTMS/ETCS system. In FM-RAIL-BOK Workshop in SEFM 2013, 11th International Conference on Software Engineering and Formal Methods, page 6p

[17] Antunes, G., Caetano, A., Bakhshandeh, M., Mayer, R., and Borbinha, J. (2013). Using ontologies to integrate multiple enterprise architecture domains. In International Conference on Business Information Systems, pages 61–72. Springer.

[18] In2Rail Project Deliverable D9.1: "Asset status representation"; October 2016.

[19] UIC: RailTopoModel (RTM) version 1.1 release webpage: https://uic.org/rail-system/railtopomodel

[20] railML.org: Official railML website. https://www.railml.org/en/;

[21] SIA Project Deliverable D5.3: "Modified XSD and matching RailTopoModel class diagram"; September 2020

[22] Linx4Rail Project Deliverable D4.1: "Demostration indicator and use cases description"; August 2021.

10.9 Authors

Katharina Hartmann has been a research associate at the DLR Institute of Transportation Systems since 2020. Her work focuses on testing components for rail and automotive communications as well as project management. Katharina holds a degree in computer science and is currently deepening her knowledge in technology-oriented management with a focus on service design for digital products.

E-Mail: Katharina.Hartmann@dlr.de

Oliver Röwer has been working at the RailSiTe® test laboratory of the DLR Institute of Transportation Systems since 2013 as a computer scientist. His research area is test methods and approval methodologies in the railway domain for command, control and signaling technologies. This includes in particular the construction and further development of test laboratories and simulation systems in the area of Automatic and Remote Train Operation.

E-Mail: Oliver.Roewer@dlr.de