Seminar: New Developments in Rail Safety Systems



The "New Developments in Rail Safety Systems" SCSC Seminar was held at the Eurostars Book Hotel in Munich, Germany, on 4th December 2025 and was hosted by Carmen Carlan and Mike Parsons. Although the event had 'rail' in the title, many of the talks addressed the role of Artificial Intelligence (AI) and Machine Learning (ML). These are topics that have applicability in many other sectors. Roger Rivett provides highlights of the day.

This was the second seminar held in Munich and a follow up to the successful seminar held in November 2024. Carmen Carlan was again instrumental in organising the event, arranging the speakers and helping select a suitable venue. She also co-hosted with Mike Parsons, who, along with Alex King, Brian Jepson and myself, made the journey via train from the UK.

It was again decided to travel by train and carry the high-value AV equipment. Fortunately, the trains to Germany were on time and all connections were made without the difficulties encountered during the 2024 trip. After leaving St Pancras at 11:04 we arrived at the hotel at the relatively civilised time of 21:15. However, the plans for the journey back suffered an immediate set back when the train from Munich to Frankfurt was two hours late. We did all eventually get home in the early hours of Saturday, except for Brian who had to stay over-night in London and take a train in the morning; he arrived home 24 hours after he started!

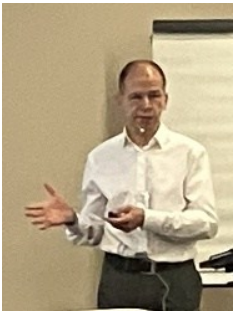
The book-themed hotel features the title of books, and the names of authors, as decoration and if you had trouble sleeping you could always read the quote thoughtfully written on the ceiling above the head of the bed. The event was once again well-served with a large room for the event and good catering, including a 3-course meal for lunch.

Mike Parsons opened the event with an introduction to the SCSC and the reason for organizing a seminar focused on rail safety systems, and the role of AI and ML in improving train operations. Mike then handed over to Carmen Carlan who introduced each of the speakers in turn.

The seven speakers presented talks discussing AI-based obstacle detection systems, machine learning graph compilers, autonomously driven trains, the operational design domain and the use of mobile mapping technology for railway infrastructure surveying. All topics were focussed on contributing to the safety and reliability of railway systems.

- **Lennart Kilian**, Siemens Mobility GmbH - safe.trAIn: AI based obstacle detection for trains - risk analysis, functional scalability, safety on system level and lessons learned
- **Jelena Frtunikj**, NVIDIA - Hardening Machine Learning (ML) Inference for Safety-Critical Applications: From Static Libraries to ML Graph Compilers
- **Carsten Thomas and Philipp Jaß**, HTW Berlin - N-Version Architectures for Railway Track Segmentation with Deep Neural Networks
- **Arnold Bitner**, Modelwise - In automation we trust - also for FMEDAs?
- **Gereon Weiß**, Automation Systems at Fraunhofer IKS - Safe Autonomy in Rail – The Role of Operational Design Domains
- **Stephen Bull**, Ebeni - Assuring Changes to a Complex System – Lessons From Experience
- **Volker Wegener**, DATADEV - Railway geodesy - 5 mm at 80 km/h: A modern method for efficient and precise surveying and digitalisation of railway infrastructure without any safety risk

Lennart Kilian



Lennart reported on work performed as part of the EU Safe.trAIn project https://safetrain-projekt.de/en/, which is contributing to the goal of 'Fully Automated Train Operation'. The benefits of this are that it addresses the shortage of drivers and provides opportunities to improve efficiency and delivery of a better service. The latter will also contribute to reducing emissions if it leads to a shift from personal to public transport.

This project focused on the detection of objects, based on the assumption that this can only be done using ML. A number of challenges were noted including that the application falls under the European AI ACT as a high-risk application and the fact that there is no safety standard for AI-based perception systems in the rail domain. There are also no clear requirements for interpretability and robustness that have to be met and no established tools and processes.

The safety objectives are to protect both passengers in the train and persons on the track. This involves the detection of objects and persons on the track. The safety target is 'overall as good as a driver' and quantified as a Probability of Failure on Demand (PFD) of 1%, although the use of PFD is controversial in European Rail Groups.

The PFD is considered to be equivalent to 1–recall, where recall =TP/(TP+FN), TP is a true positive and FN is a false negative.

The 1% PFD target is derived by comparing event trees for the driver and the object detection system. The analysis is independent of the assumed initial event probability, which means that it's the ratio that is relevant, not the absolute scale. It was observed that expanding the sensor capabilities was more effective than decreasing the PFD and that increasing the PFD to 10% is significant while decreasing it to 0.1% has little effect.

Overall, it was considered that safe perception for driverless regional trains could be achieved and that, although posing many challenges, these were already being overcome in the automotive domain.

Jelena Frtunikj



Jelena described how, since 2012, market demand has driven an evolution of neural network architectures. This has seen an increase in core tasks and representative models, combined with model complexity.

To date, these architectures have been based on the use of pre-validated static kernels. In these, the safety subset of the kernels is compiled ahead of time and each kernel is validated according to ISO 26262. The compilers are qualified and the kernel executed on an ISO 26262 safety-certified runtime environment. This has the benefit of clear, deterministic execution. However, it cannot support new operations, it cannot fuse operations for efficiency and new hardware requires months of validation. Performance is also suboptimal because it is limited to what is in the library, and generic libraries cannot keep pace with modern deep learning.

The proposed way forward is the use of an ML compiler that translates a computation graph, representing the ML neural network, into a form that can be executed on the target platform. It consists of two main components, a graph compiler and a kernel compiler that performs complex across-layer optimizations that holistically evaluate and optimize the entire network. The graph compiler takes advantage of model-aware optimization, hardware-aware specialization, operation fusion, memory optimization and scheduling and concurrency.

While this new approach improves performance, it also has benefits for safety as it moves from being tool-centric ('Is the compiler qualified?') to being output-centric ('Is the generated compiler code correct?'). The old paradigm was based on compliance with ISO 26262 for the compiler, the safety runtime and the static libraries. The result was safe but limited. With the new approach, the output of the graph compiler is validated at build time and each generated kernel is subjected to a multi-layer verification (graph level, kernel level, application level) before being executed on a deterministic safety runtime developed according to ISO 26262. The result is safe and flexible.

A question was asked about the need for a correct semantic definition of the graph compiler, as for a code compiler. This is not possible as the semantics are application-dependent; instead, validation is performed at different levels. Kernal is developed as a Safety Element out of Context (SEooC) and the user is provided with a safety manual, a safety case and tools for specific contextual validation to identify any non-deterministic behaviour.

Another question concerned having a validated graph compiler. Again, this is considered not to be possible due to the rapid development of the technology and the lack of semantics, due to optimisations across multiple measures.

Carsten Thomas and Philipp Jaß



Carsten explained that while driverless metro systems had been in operation since the late 1960s, autonomous train operation on mainline railways is a lot more challenging. It is a less controllable environment, longer braking distances are required and the network is used by different types of trains.

In the conventional approach, the train driver must observe the track, the signals, the level crossings and the overhead power lines in order to avoid hazards such as obstacles on the track, unjustified emergency braking, and persons to persons on, or close to the track. In this way, harm is prevented from occurring to persons in the train and those outside it.

To avoid obstacles, it is necessary to determine which track, the ego-track, the train will be on after crossing points. An autonomous train cannot rely solely on GPS, because it is not accurate enough, or train schedules, because that are not reliable enough. The approach taken here is to use an AI-based perception system to identify all possible future tracks and combine this with switch and signal detection and localization.

The most critical failure of the AI-based perception system is not detecting an obstacle that is present (false negative). The acceptance criteria used was that these are not more frequent than for a human train driver; this equates to a PFD < 1% (which cannot be achieved by AI alone) and falls between SIL1 and SIL2. However, this criterion is still subject to debate.

Safety is a system property, so although there is no such thing as safe deep learning, what is required are safe cognitive systems. Possible safety measures include the use of dissimilar redundancy in sensing and implementation, plausibility checks using tradition algorithms and checking against external data. The work described here considered redundancy and dissimilarity in the implementation.

Philipp noted that a single neural network is not suitable for safety-critical systems before discussing n-version architectural dissimilarity, which is a well-known method from software safety standards for non-ML software. This uses redundancy to detect and potentially correct errors. An architecture was presented in which one input is processed by three ML models with a non-ML component producing the output.

The ML models were not pre-trained as the rail domain lacks a database of pre-existing datasets. Instead, the models were trained on a newly collated data set of 15,519 images collected from multiple sources. The predictions of the ML models were combined using two different algorithms, Maximum Confidence Voting (MCV) and Pixel Majority Voting (PMV). The combination of individual confidence scores was based on the similarity of individual predictions.

Overall, the project demonstrated that the use of N-version architectures may be a promising approach to ML safety. The diversity in the Deep Neural Network architectures improves

segmentation reliability while the combined confidence score correlates better with the actual prediction quality. Pixel-level combination enables error correction.

Arnold Bitner



Arnold presented on the integration of safety engineering processes with engineering processes. He identified 5 problems with the safety analysis process:

- Missing integration between requirements, safety and product development
- Months of manual work for Failure Mode, Effects, and Diagnostic Analysis (FMEDA) per design iteration
- Accuracy depending on scarce and expensive experts capable of work
- Budget and time constraints making full analysis impossible
- Suboptimal fit of design to requirements due to over, or under, engineering

These highlight the need for automation and data integration across the supply chain and the challenges of managing complex systems. There is a need to abstract and simplify data for human review while maintaining traceability. It is also important to connect stakeholders across the supply chain and share data securely.

His proposal was to take the engineering design, the safety requirements and the failure modes and rates and combine them in a single mode. Then, to use AI to perform automated reasoning on the failure behaviour and produce an FMEDA report that can be assessed; this reduces the human workload. Such a system should be scalable, efficient, easy to use and certifiable.

Gereon Weiß



Gereon outlined the problem he was addressing as being that autonomous systems in the real world must safely operate in an unpredictable and chaotic place and that understanding and defining the context in which they operate is key.

He then introduced the concept of the Operational Design Domain (ODD) as a means to help achieve this by confining the real world to that in which the autonomous system is intended to operate.

He mentioned two automotive standards (ISO 34503:2023-08 and ASAM OPENODD BASE STANDARD 1.0.0 SPECIFICATION V1.0.0:2025-04-03.) and two examples of automotive implementation (Waymo and Mercedes).

He introduced two definitions, one commonly used for Automated Driving Systems (ADS):

> *A set of operating conditions, including, but not limited to, environmental, geographical, and time-of-day restrictions, and/or the requisite presence or absence of certain infrastructure characteristics in which a given automated driving system, or feature thereof, is specifically designed to function*

and one for AI Systems:

> *A set of operating conditions, including all entities relevant for the AI system, in which it, or feature thereof, is specifically designed to function*

The ODD should be seen as a central artifact that provides traceability and consistency for many different aspects of the development process such as architecture, safety case, training, verification and validation and performance monitoring.

ODDs can vary in subject matter and detail depending on the context in which they are used. The ODD covering development aspects needs to be detailed whereas the ODD covering safety aspects can be more coarse-grained and abstract so as not to be over-specified.

There can be application-specific refinement of the ODD, for example starting with a more permissive one for describing the system environment, a refined one for system development and assurance and one specifically refined for ML-methods.

The level of complexity of the ODD decreases in the move from manual to fully automatous systems as manual systems can cope with a more complex ODD than fully automatous systems, which tend to be designed for narrow ODDs.

Gereon than introduced a new document for the Specification of Operational Design Domain in Rail, DIN DKE SPEC 99004, which defines a taxonomy for the ODD for fully automated rail vehicles covering important operating conditions for railway-specific AI applications such as Automated Train Operation (ATO), e.g. environmental or track conditions, https://www.dinmedia.de/de/technische-regel/din-dke-spec-99004/390496147. This has a wide range of potential usage, including:

- Risk assessment for safety and security
- Reference System comparison
- System development
- ODD limit detection
- Scenario-based testing
- Conformity assessment and approval
- Intended purpose adequacy
- Reuse
- Increasing confidence
- Project planning

Gereon concluded his talk by giving three example ODDs for rail applications, each of which was described by a taxonomy, depicted graphically and also specified using YAML (a human-readable data serialisation language).

Stephen Bull



Stephen highlighted the fact that changes to a railway system had to take place while the railway was still operating and compared this to changing the engine of a plane while in flight.

A railway is a system of systems supporting the needs of many people including the train driver, the signaller, the maintainer and the customer. While being difficult to define what makes the railway complex, Stephen highlighted 15 characteristics that contributed to the notion of complexity.

He illustrated some of these with three examples from experience, the first being the high number of interfaces and legacy components with which the UK East Coast Digital Programme has to contend with. The second was the interaction between the need for very curved platforms at Tottenham Court Road and the clearance of the platform screen doors from the train on the London Elizabeth Line. The last one concerned radio communication capacity where the unexpected loss of communication resulted in the continuous attempt to regain a connection. This led to the system being swamped in what looked like a denial-of-service attack.

Stephen then considered a number of emerging challenges.

AI is both a risk and an opportunity. The key issues to be addressed for assurance being transparency, bias, traceability, cyber security and ethics.

Given the large quantity of data involved it is necessary to understand the different types and how it is used; reference was made to SCSC Data Safety Guidance document (see scsc.uk/dsiwg). An example was given where the data representing the gradient of the track has to be trustworthy because the train relies on this to determine when and how hard to break.

The link between security and safety was highlighted; the increasing attack surface is leading to an increasing number of attacks resulting in a higher risk of harm. Reference was made to the SCSC Security Informed Safety guidance (see scsc.uk/siswg).

The problem of standards becoming outdated was mentioned, along with the fact that they never reflect the latest technology by the time of publishing and that different ones often contradict each other.

Stephen summed up the key challenges as being:

- Understand the interactions and interfaces in your system
- Understand your assurance goals (e.g. for standards, AI, data)
- Know how you will integrate your system

His advice for addressing these was start early, agree a strategy, identify regulations and standards, decide how to demonstrate compliance, review regularly and plan to manage change.

Volker Wegener



Volker explained that the management of a railway requires knowledge of the infrastructure, for example the precise location of the rail, bridges and tunnels. It also needs up-to-date information concerning the state of the track and signals. There is also a need to create situation maps after floods, earthquakes and storms. Such knowledge has traditionally been collected using methods which were not efficient and disrupted the rail service. Therefore, there is a need for a more cost-effective method that does not disturb regular service and is also safe for customers and railway staff. Ideally, the information would be stored such that it can interface with existing IT systems.

To address this, the work had evaluated the use of a multi-sensor approach consisting of a Global Navigation Satellite System (GNSS), an inertial measurement unit (IMU), Light Detection and Ranging (LiDAR) and cameras mounted on a locomotive. These were used in conjunction with ground control points and a vehicle positioning systems with targets along the tracks. This was able to collect data at speeds ranging from 0 to 100 km/h and achieved a ±5 mm accuracy at 80 km/h.

The benefits of this approach include the fact that there are no persons in the track danger zone. The fast coverage of long track, with automatic detection of damage, cracks, markings and traffic signs leads to reduced inspection costs, better planning of renovations and it also minimizes downtime.



The data is collected in a format which can be used in combination with existing data sets providing a uniform database for all departments helping with improved decision-making through spatial analysis and forms the basis for smart city applications and digital twins.

The approach also helps with security and disaster management by providing mobile mapping in the event of floods, earthquakes and storms and leads to faster response times and improves deployment planning.

The overall conclusions were that multi-sensor systems are the most efficient railway surveying method providing high accuracy without operational impact and being fast, scalable and cost-efficient.

The presentations and videos for this event are available to SCSC Members, see scsc.uk/e5004 for details.