

SSS'24 Event Report



The Safety-Critical Systems Club's annual Symposium returned in February 2024 for another fully in-person event hosted at the Bristol Royal Marriott. Paul Hampton, the SCSC Newsletter Editor, provides highlights of the three-day event.

Mike Parsons introduced the symposium and noted that this year's event was very well attended with 176 delegates. Mike said one of the key themes of the symposium was managing the risks around Artificial Intelligence (AI), and so amongst other AI-related talks, a special panel session was going to be held on the second day allowing delegates to pose AI-related questions to the panel of experts.

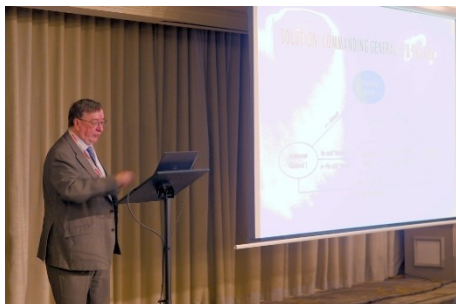
New for this symposium was the SCSC timeline (see page 43), and other features introduced last year such as the crossword, roving video interviews and a light-hearted take on last year's "Room 101" where delegates could, this time, vote on the most compelling narrative from a panel providing stories of events that influenced or even change their system safety world.

Byzantine Generals Attack an Airbus A320

In the first key-note speech of the Symposium, Dewi Daniels expanded on a talk he had previously given in 2022, which discussed an Airbus A320 incident involving the failure of all three primary flight control computers on touch-down.

Airbus claimed from in-service data that no similar fault had occurred in over 44.3 million flight hours, but Dewi said that another incident did occur relatively soon after his previous talk. Dewi said that both incidents suffered from the same problem – “The Byzantine Generals Problem”, first described by Leslie Lamport in his relatively unknown 1978 paper: “The implementation of Reliable Distributed Multiprocess Systems”.

This was reimagined as “The Byzantine Generals Problem” in 1982, which presented the ideas more in a story-telling narrative, and this has now had over 10,000 citation showing the power of storytelling in communicating ideas and concepts.



Dewi reprised the problem Byzantine General Problem:

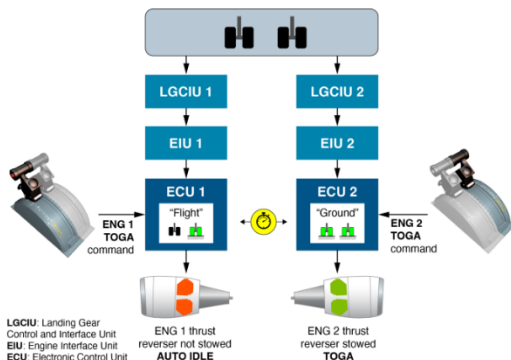
- A General in charge of several Lieutenants must get agreement on common actions amongst the lieutenants for the overall campaign to be successful
- However, one of the commanders is unreliable and so may send misleading messages to the other commanders to sabotage the campaign

Dewi said this is effectively a form of 2/3 voting system. Lamport pointed out that if input is faulty then it could confuse the voting systems. Faults may also arise if the processors read the input at different times.

Dewi showed how this mapped to the Airbus Flight control architecture where monitor system were employed that would switch off a command lane if errors were detected. In the original Airbus incident, cross winds and bouncing on landing caused the 'weight on wheels' sensor readings to fluctuate resulting in different values being read by the monitors. Dewi noted that there were no system faults here, it was simply that the readings were not in synchronisation. In less than a second, all three primary computers had shutdown and the pilots were only able to stop aircraft through manual braking.

Dewi then discussed the 2022 incident where an A320 was landing at Copenhagen during a cross wind and he shared the Airbus article that described the incident. Shortly after touch-down, having deployed reverse thrusters, a pilot initiated a go-around but the aircraft veered to the left and did not climb as expected as only the right engine had engaged – the left engine remained idle. This caused the aircraft to veer to the left and the left wing only cleared the ground by 1ft before being able to land safely on another runway.

Dewi explained what happened around the logic of each engine's Electronic Control Unit (ECU), which exhibited different behaviour as each engine acts independently. Due to differences in each engine's sensor reading for 'wheels on ground', the left engine considering it was still in flight did not stow the reverse thruster, while the right engine sensing that both wheels were on the ground did allow the reverse thrusters to be stowed.



Dewi said this was example of the Byzantine General's problem with the wheel sensor being the 'traitor'. He noted that neither reports mention the Byzantine general solution but Airbus has since performed a software update.

Dewi said the paper is well known but not applied by certification authorities, perhaps due to the age of the paper, but he noted the number of citations are increasing especially in the last 10 years mainly coming from blockchain. Dewi said he is not claiming that systems need to be redesigned to Lamport's solution but it seems to have been forgotten that the problem exists. Dewi concluded by asking how we can raise awareness of this problem and suggested it could be taught at symposiums, in academia and could be explored during design review. However, he said it left him with a wider concern about what other problems may exist with solutions that we don't know about – the unknown unknowns – and what steps should be taken to uncover them.

Safety in unmanned aviation – how much have we got and much do we need?

Steve Wright from Wright Airborne Computing introduced himself as previously working for Rolls-Royce, Airbus and Universities as a lecturer and said he has now moved into Uncrewed Aviation. He previously worked on large aircraft but is now working on uncrewed aircraft "crazy" vehicles.

Steve showed a video of 'drones through the internet' where drone data was being transmitted between France and the UK in real-time.



Steve also showed some of the other unusual projects he has been working on:

- Airborne robotics – a drone that unfolds itself and disguises itself as a brick
- Medical package delivery for the military
- Counterdrone aircraft

There have however been setbacks and Steve showed videos of drone crashes and battery fires and noted that the time imperative is strong and so failures are inevitable when 'pushing'.

Steve noted that reliability is never 100% especially during development and it can be fun building systems in 'wild west' environments, but there is a serious note and we are ultimately trying to preserve life. Steve covered differences between his old and new lives:

- **Old:** Meticulous design and tested, beautifully made, expensive and reliable
- **New:** Drone aero tech – multiple technologies that have evolved, borne out of mobile phone development. Software that should be £1M but is open source on the internet.

Steve then gave a live demonstration of a small toy drone and said it was an extraordinary piece of kit as it crams in all the usual flight systems you would expect in a conventional aircraft.

Steve then talked about aerospace reliability and showed that for conventional aerospace accidents the tolerable failure rate leading to catastrophic events is 10^{-9} per flight hour. This is compared with toy and commodity drones that may be as low as 10^{-1} for the one he had just flown. Steve illustrated the concept of the *gap of the six zeroes* showing a graph comparing reliability of vehicles from toy drone through to aircraft such as F16 and Predator drones and highlighted the gap between drones and conventional aircraft.



Steve made the distinction between availability – “does what we want” and integrity “doesn’t do what we don’t want” and noted the latter can be a counter-intuitive way of thinking. He said that the approaches to regulation are different between crewed and uncrewed aircraft with the latter focussing more on containment of accidents (via regulation CAP 722 and Open-A3-1), so the vehicle is expected to fail.

The environment is critical as well as aircraft performance and some companies like Manna are exploring these with urban deliveries on the west coast of Ireland – meeting the two considerations half way by boosting aircraft reliability and modifying the mission to reduce risk.

Steve then presented his personal opinions:

- It is not whether you crash but where and how. Availability is very low and integrity is more important
- “Flyaway” is a more important risk to avoid than crashing
- Safety personnel are disconnected from operations, so revert to no crash mindset.

Steve’s final thoughts were:

- Crewed aviation: it’s all about the vehicle
- Uncrewed: it’s about mission plus environment
- Crewed/Uncrewed want same thing (want to operate safely) – but the debate is how we get there.

Steve finished his talk by describing his ground-breaking work with Aquark Technologies that involves dropping a scientific package in free fall (containing supercooled Rubidium atoms) to determine the performance of the apparatus in a zero-g environment.

Things That Changed My World

At the end of the first day, Tom Anderson chaired an informal session with Phil Koopman, Yvonne Oakshott and Harold Thimbleby acting as panellists. Each panellist had three opportunities to tell stories that were challenging, important or entertaining and that influence their system-safety world. The audience then voted on their favourite story.

Round 1 was won by Phil (47% votes) who told a story about an 8-reactor nuclear aircraft carrier that he had worked on previously.

Yvonne won the second round (44% votes) with her moving story around the crash of an AW609 tilt rotor aircraft during development.

The final round was won by Phil again (72% of votes) with a story about the challenges he faced when he was tasked with investigating an Uncommanded Acceleration (UA) event by inspecting source code.

Events

The morning of the second day started with a number of special events:

- Walking Tour of Bristol
- Bristol Cathedral Tour
- Bristol Museum Tour

These were all well-attended with all events being fully subscribed.



System Safety for Complex Projects – The Crossrail Approach



The second day's keynote was given by Hayat Zerkani discussing the establishment of Crossrail and the building of the Elizabeth Line.

She said the line is mostly deployed in tunnels, with 10 new stations (9 subsurface) with portals and shafts for evacuation and emergencies, platform screen doors and step-free access throughout with a focus on a Central Operating Centre.

All this has been built by different joint ventures and needed to link with existing infrastructure.

The system is highly complex with many challenges:

- Noise reducing rails, non-traction power, UPS for critical systems
- Forced ventilation for cooling, control station and fire control
- Communications-based train control (CBTC)
- New trains
- European Train Control System (ETCS)
- Automatic Warning Systems (AWS)
- Signalling systems can change between different sections
- Rail milling trains to smooth tracks and other types of maintenance vehicle
- Regulations from Interoperability Directive 2008/57/EC, Safety Directive 2004/49/EC and the Common Safety Method for risk ((EU) No 402/2013, CSM-RA)

To deal with this complexity, a more progressive handover was seen as more practicable rather than one as originally thought. Three stages were established:

- trial running
- trial operation
- entry into passenger service

Not all systems would need to be in place during non-passenger trial running.

Also a number of activities to ensure consistency were undertaken, such as developing a project-wide hazard record (PWHR), which was a key management tool to record hazards and safety justifications. This was implemented as a live database in DOORS¹ with a user friendly front end. This was maintained throughout the project and used as evidence for compliance with CSM-RA Regulations.

The safety justification had several parts: railway, contract and element level and it needed to bring these elements together including the station, shaft, portal and railway systems. It also essential to show integration with Network Rail so a Safety Justification was needed for the interface with railway companies and Crossrail.

This resulted in about 90 documents between contractors and Crossrail that needed to be maintained. Key success factors were:

- Progressive Assurance – deferring of some functionality to later configurations
- Safety Justification with dependencies, dependencies agreed with stakeholders and agreement of mitigations
- Ensuring consistency across all contracts was essential
- There were 600 dependencies so robust controls were required

Lessons that were learnt:

- reduce things into smaller more manageable components to help ensure consistency
- make use of tools
- focus on delivery
- be open with safety
- keep stakeholders engaged

¹ IBM® Engineering Requirements Management DOORS

Panel Discussion “Can we make safe AI?”



The day concluded with Alan Simpson hosting a panel discussion with six experts in AI answering questions from the audience.

The following questions were covered:

- Q1: Can we make AI safe?
- Q2: How do we make AI safe?
- Q3: How do we improve the competency of those developing AI systems?
- Q4: What is the degree of Functional Safety in AI regulation?
- Q5: Can you provide an example of the most complex AI system that has a safety case now?
- Q6: Should individuals/the community be more proactive in highlighting the dangers of AI?
- Q7: How do we prove AI systems are working correctly in law?
- Q8: What might limit the arguably unbridled expansion of AI systems?
- Q9: Would you be happy being driven by an Autonomous Vehicle?
- Q10: Why do you trust Autonomous Vehicles less than high risk human drivers?
- Q11: Can use of analysis techniques prove the absence of defects in a formal AI model?
- Q12: What could ensure ethical use of ML and would that achieve safety as a consequence?
- Q13: What advice would you give to the Safe AI Working Group?

You can see videos of the panel’s responses to these on the event’s Programme page on the SCSC website: <https://scsc.uk/e1007prog>

Poster Session



The days talks were followed by a poster session in the main exhibition area where individuals and members of the SCSC Working Groups presented posters summarising their work on large presentation boards. There were many posters this year, which meant spreading the stands across several rooms. Some posters were also interactive inviting delegates to contribute their own ideas on the particular poster subject.

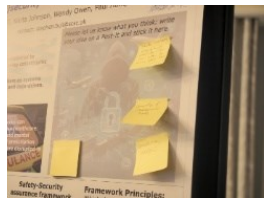
Posters were presented on the following topics:

- Dynamic Risk Assessment in Automated Vehicles: An elaborated Approach to HARA
- A Bow Tie Approach to Quantitative Risk Assessment and Control for Maritime Vessels
- Towards Trustworthy AI: Legislation, Techniques and Challenges
- SafeLLM: A Novel Framework for Safety Evaluation of Large Language Models: A Case Study of Offshore Wind Maintenance Planning
- Using Rust for Safety-Critical Systems
- Ensuring SOTIF for Off-Highway Autonomous Vehicles
- Adaptive Safety Measures: A Concept to Optimize Safety in Automated Driving Systems
- Safety and Complexity (Paper: How do we make the case for "Safe" software and AI systems? - the Horizon Example)
- IEC 63187: Systems engineering – System safety Complex systems and defence applications
- Safety and Complexity

And posters from the Working Groups:

- Safety Culture Working Group
- Assurance Case Working Group
- Data Safety Initiative Working Group
- Multi-Core Working Group
- Systems Approach to Safety of the Environment Working Group
- Safety of Autonomous Systems Working Group
- Service Assurance Working Group
- Safer Complex Systems Working Group
- Security Informed Safety Working Group
- Ontology Working Group

See page 51 for more details of most of the posters.



Safety Timeline

A new feature for this symposium was the Safety Timeline where delegates were invited to add safety events that are important to them to a timeline stretching from 1400 to the current year. This was for example, filled by notable accidents such as Chernobyl (1984) and the Challenger Disaster (1986). The intention is to present the captured information in an interactive form on the SCSC website so that the timeline can be explored in more detail and new events can be added.



Banquet

The end of the second day concluded with the traditional symposium banquet in the fine Palm Court Room at the hotel.

After an excellent meal, the attendees were entertained by Simon Davies from Vertical Aerospace who gave a talk and presentation on aviation and his role as test pilot for the new class of electric Vertical Take-Off and Landing (VTOL) aircraft currently under development by his company amongst others.



The 'Golden Thread' of Information: A Systems Approach to Construction

The first keynote speech on the final day was given by Mikela Chatzimichailidou, Professor in Design for Mobility, Health, People and Society at UCL.

She covered three main areas in her talk:

- Background: motivation and literature
- Applied research: scope and publication
- Findings: Benefits and Challenges

Background

Mikela started by saying that construction came into her life after the 2017 Grenfell tower block fire that she witnessed firsthand from another building.

She said that the resulting independent review of building regulations and fire safety (the Hackitt Report), concluded that there were four main problems:

- Ignorance, regulations are not always read and often misunderstood and misinterpreted
- Indifference and commercial imperatives to deliver quickly and cheaply
- Lack of clarity on roles and responsibilities
- Inadequate oversight throughout the lifecycle

The proposed solution is to start thinking about the building as a system (it has a life) and she introduced the report's "golden thread of information", this being a tool – intended to be a digital platform – to manage all buildings as systems and to allow people to use digital records (acting as a database and single source of truth) to safely and effectively design, construct and operate residential buildings.

Applied Research

Mikela then went on to say that there is much innovation in modernising methods of construction, such as modular construction (constructing components offsite and assembling), sustainable building and 3D printed houses.

There is however debate as to whether Modern Method of Construction (MMC) building is safe and so Mikela along with an industrial research student undertook research (including a review of 74 relevant papers) to examine the link between Building Information Model/Modelling (BIM) systems, safety management and MMC. Mikela described BIM as a process with the output being the digital twin – the virtual representation of the system to be operated.

A BIM system can enable the duty holders to ensure accuracy and quality of design and construction, which are crucial for building-in safety up front and gives stakeholders access to a database from which to draw evidence.

Key findings from the research were categorised into three areas: training & warning, safety inspection and crane management.



Training: BIM can enhance training by supporting a virtual, interactive learning environment.

Safety Inspection: From manufacturing to construction, BIM-based dynamic automatic detection systems can effectively identify possible safety risks in modular construction.

Crane Management: operators and managers can more intuitively and accurately analyse potential conflict factors and find unreasonable lifting plans, thereby avoiding casualties caused by crane collision.

BIM therefore provides an opportunity to link worker safety and building safety under system safety to achieve the “golden thread of truth”.

Mikela provided some conclusions:

- Buildings are complex
- Safety is a system's property, an emergent property
- Manage safety through requirements that are traceable throughout the building lifecycle
- BIM can be used to achieve the golden thread
- The golden thread brings together worker safety and building safety under system safety

There are challenges however, and Mikela provided eight challenges such as data protection, data assurance and skills shortage.

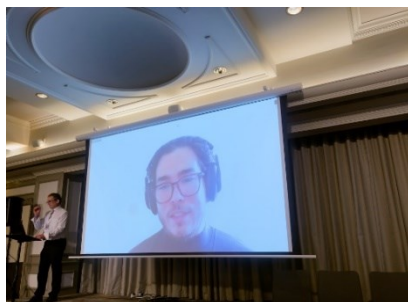
Mikela concluded with perspectives that are important not to forget:

- Digital tools are not here to replace human thought, critical thinking and evaluation still need to go into safety management; we should treat these as decision support tools
- Need to be confident about data assurance, its provenance, who has access and for how long and data quality
- Use of tools should be accompanied by contract commitments, for example, by putting use of BIM into contracts
- Seeing is believing – target and share successful case studies but be transparent
- It's a team effort, policy and practice need to go hand in hand
- One size doesn't fit all – be mindful of project specifics and nuances
- Need to remember the contribution being made in social/public values

Safe by Design – Autonomous Vehicle Systems

Adiac Aguilar from Volvo Cars works on electronics and software platforms for Autonomous Vehicles (AV) and he gave his keynote talk on the 'Safety by Design' methodology. Adiac presented remotely having been unable to attend the symposium in person.

Adiac started by introducing the foundations of Safe by Design. This is based on a safety blueprint of the methodology's process and principles than involves safety culture, Continuous Improvement (CI) and human/machine/environment considerations.



The process starts with Hazard/Threat Analysis and covers System Safety and Cybersecurity streams.

Safety (Assurance & Risk) Management Systems are required to ensure a Safety Management System (SMS) is established covering Safety Policy, Safety Processes, Safety Risk Management, Safety Assurance and Safety Product Development.

Architecture Design for Safety

Adiac showed a high-level block diagram for an AV Architecture comprising:

- Communication
- Decision Making
- Perception
- Power Management and Control

He stressed the importance of designing robust and fault-tolerant architectures to mitigate safety risk.

V&V for Safety Assurance: Adiac also stressed the importance of rigorous testing and validation to verify the effectiveness and reliability of safety measures. Integration of V&V throughout the development lifecycle is also important.

Safety Culture and Organisational Practices

It is important to foster a strong safety culture to promote safe behaviours and attitudes. Integration of organisation practices and policies will support and reinforce safety culture. Adiac discussed the Safety Culture Framework covering:

Leadership Commitment, Employee Engagement, Communication, Safety Training, Continuous Improvement, Organisational Practices, Incident Reporting and Safety Committee.

Ethical Considerations

Adiac concluded by covering ethical considerations in three sections:

- Ethical Dilemmas: trolley problem, allocation of responsibility between human and Artificial Intelligence (transition between level 3 and 4)
- Ethical Principles: safety transparency, accountability, fairness, privacy and social responsibility
- Ethical Frameworks: IEEE ethically aligned design framework, ISO/other standards, NHTSA Federal Automated Vehicle Policies

Breaking the Tyranny of Net Risk Metrics for Automated Vehicle Safety

The last keynote speech² was from Phil Koopman from Carnegie Mellon University.



Phil noted that all Autonomous Vehicle (AV) manufacturers advertise as having safety as their first priority and the messaging that he is seeing is that these companies are saving lives, but the reality in his view is that a lot more miles will be required to prove it with sufficient confidence.

Robotaxi, for example, claimed in 2023 that humans are 'terrible drivers' citing 92% fewer collisions but the statistic only included 'at fault' accidents and hence, for example, did not count left turn faults so the statistics are potentially skewed.

There are however counterexamples where, for example, a Cruise robotaxi rear-ended an articulated bus seemingly ignoring the back half of the bus despite having LiDAR sensors. Phil said that "Robots won't drive drunk but will drive stupid"; humans will make same errors but for other reasons.

Further occurrences took place in August 2023 when an AV drove into cement and a crash into a fire engine resulted in the company being told to halve their fleet. The messaging that 'robots were good and humans terrible' has set companies up for a fall.

In San Francisco, Waymo have a better record but are not blame free. Phil noted that they are not required to report things where people were not injured but there have been safety-related incidents where incapacitated vehicles stop access to emergency vehicles.

² Phil's talk is based on a paper in the Safety Critical Systems eJournal, "Breaking the Tyranny of Net Risk Metrics for Automated Vehicle Safety" by Philip Koopman and William H. Widen <https://scsc.uk/scsc-191>

Nobody knows when AVs will be safer than humans but Phil believes companies are overstating the position and in fact possibly distorting the story. Phil estimates that 300+millions miles are required but they have only 5-8 at the moment.

A Cruise vehicle in Aug 2023 also drove into a road closed area and Phil asked if that was safe; there is more to it even when no harm is done.

The drive to demonstrate that AVs are safer than humans is called Positive Risk Balance but Phil said there is a lot more to it and there are other factors that need to be considered:

#1 Positive Risk Balance (PRB): means safer than a human driver, but which: (under influence, speeders, distracted, drowsy, older drivers, state and infrastructure) – national average are meaningless. What kind of vehicle – those with auto braking are safer rather than older vehicles.

#2 Avoid risk transfer: are children more at risk, or cyclists? There is a problem if some specific demographics will lose out so we shouldn't increase a group's risk and so adds additional constraints on safety from PRB.

#3 Avoid Negligent Driving: establish a duty of care to other road users. There is no precedent for an AV. Product liability different from tort and liability. Should be similar to a human driver being negligent. Miles driven is no protection from culpability.

#4 Standards and expectation on safety: only one of the standards are currently required.

#5 Fine-grain risk & regulators: want to avoid regulatory recalls. Recalls always relate to specific behaviour such as rolling through stop signs; saving lives is not enough to avoid recalls. Regulators are struggling to predict safety outcomes in advance.

#6 Ethical & Equity Concerns: Ride Hail made promises with disappointing results. Labour issues, cheap taxi undermines safety, ethical & related concerns eg: testing risk imposed upon vulnerable people and no independent technical oversight is required.

#7 Sustainable trust: trust-degrading rhetoric. Unreasonable claims and relentless blame on humans. Lobbying for municipal pre-emption. Need more trust and need to talk with and not "at" with more transparency on incidents and corrective actions. Need to state release criteria and performance metrics.

Phil concluded by saying that if you want to claim an AV system is safe then all these points need to be addressed.

Crossword Competition

The SCSC Safety System crossword was first published in the Oct 2023 edition of the newsletter and was also given to delegates to complete during the symposium. There were 13 entries in total, and the winner was drawn from the correct entries in the 'hat' by Mitali Atkins with the winner receiving a £250 gift voucher.

The answers to the crossword are given below.

