The Safety-Critical Systems Club Newsletter

Safety Systems

Vol 29 No. 1 - Feb 2021



ANGLE OF ATTACK

Boeing 737 MAX is back – but is it safe?

IDENTICAL TWINS?

How to assure digital twins

THE BRAINS TRUST

Establishing trust in Autonomous Systems

For everyone working in Systems Safety

SCSC

thescsc.org



SCSC Publication Number: SCSC-162

While the authors and the publishers have used reasonable endeavours to ensure that the information and guidance given in this work is correct, all parties must rely on their own skill and judgement when making use of this work and obtain professional or specialist advice before taking, or refraining from, any action on the basis of the content of this work. Neither the authors nor the publishers make any representations or warranties of any kind, express or implied, about the completeness, accuracy, reliability, suitability or availability with respect to such information and guidance for any purpose, and they will not be liable for any loss or damage including without limitation, indirect or consequential loss or damage, or any loss or damage whatsoever (including as a result of negligence) arising out of, or in connection with, the use of this work. The views and opinions expressed in this publication are those of the authors and do not necessarily reflect those of their employers, the SCSC or other organisations.

Except where explicitly stated that licensed use of this work is otherwise restricted, this work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/ or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA. You are free to share the material in any form and adapt the material for any purpose providing you attribute the material to the Safety-Critical Systems Club (SCSC) newsletter, reference the source material, include the licence details above, and indicate if any changes were made. See the license for full details.

Cover image: 200281294 © Khairul Marasani | Dreamstime.com
Back image: 177930975 © Pairoj Ponglerdnadda | Dreamstime.com

Contents

WELCOME

Editorial 3

Opening words from the SCSC Newsletter Editor.

In Brief 4

Recent system safety news items from around the world.

FEATURES

Boeing 737 MAX – Safe to Fly? 5

Paul Hampton and Dewi Daniels provide an update on the 737 MAX as it now returns to service, and they ask if it really is now safe to fly.

Assurance of Digital Twins for Safety-Critical Systems 17

Kevin King introduces some principles to support the assurance of Digital Twins for safety-critical systems.

Safer Autonomous Systems Training Network 29

Davy Pissoort discusses a new training initiative to help establish people's trust in autonomous systems.

The Safety Futures Initiative 37

Zoe Garstang introduces a new initiative to help encourage and nurture the next generation of skilled safety engineers.

Seminar Reports 40

Richard Trebilcock reports on the Safety Analysis Techniques Seminar.

Andy Williams reports on the Management and Oversight of Complex Systems Seminar.

60 Seconds with … Harold Thimbleby 55

Prof. Harold Thimbleby answers some quick-fire questions about system safety and life!

GROUPS

Working Groups 51

Details of the current SCSC Working Groups.

SCSC Steering Group 58

Contact details for members of the SCSC Steering group.

EVENTS

Calendar 60

Events Diary 61

SCSC

FOR EVERYONE WORKING IN SYSTEM SAFETY

www.scsc.uk/sss

SSS'21

Details of this and other events at: www.scsc.uk

Safety-Critical Systems Club Annual Symposium

www.scsc.uk

THE SAFETY-CRITICAL SYSTEMS CLUB

29th Safety-Critical Systems Symposium

9-11th February 2021, Online

The Safety-Critical Systems Symposium in 2021 (SSS'21) will be held online as a virtual conference. This event comprises three afternoons of live presented papers, including keynote presentations and submitted papers. There will also be an evening session with additional talks. This event will be run through the SCSC website and the Whova application, with messaging between delegates and virtual rooms for exhibitors. Details and brochure can be found at: www.scsc.uk/sss

The Symposium is for all of those in the field of systems safety, including engineers, managers, consultants, students, researchers and regulators. It offers wide-ranging coverage of current safety topics, focussed on industrial experience.

It includes recent developments in the field and progress reports from the SCSC Working Groups. It covers all safety-related sectors including aerospace, defence, healthcare, highways, marine, nuclear and rail. There will be a special session on the systems and services related to Covid-19.

The symposium features are:

- Six keynote presentations and submitted papers;
- Updates from the SCSC Working Groups;
- Audience participation and submitted questions;
- Full proceedings volume with hardcopy available on Amazon and free PDF download;

The symposium is a regular event and fosters a community spirit in the field: it is a great place to network, learn about the latest practice in safety and develop new business contacts.

To **book your place** please see: www.scsc.uk/sss

For **further information, exhibition** and **booking queries** please contact: Alex King, Dept of Computer Science, University of York, Deramore Lane, York, YO10 5GH. Email: alex.king@scsc.uk

For **technical aspects** and queries on **talks, speakers, abstracts, papers** or **posters,** please contact: Mike Parsons, mike.parsons@scsc.uk

Editorial

As we approach the anniversary of the first occurrences of Covid-19 related lockdowns, we are still left speculating on when there will be a return to what we might have considered "normal". The rapid rollout of several different vaccines is encouraging, but the emergence of new, potentially more virulent and dangerous variants, brings further uncertainty and there are doubts as to whether things will ever be the same again.

A year on, and many people and organisations have found through necessity, the need to adapt and evolve to working within governmental constraints as a kind of "new normal". The SCSC is no exception, and has now run two successful online seminars covering new safety analysis techniques and the management and oversight of complex systems respectively. You can find reports of these seminars in this edition on pages 40 and 45, and a further Seminar on developments in Safety Culture is planned for April this year (see page 28). The stage, albeit a virtual one, is also set for the first SCSC Symposium (SSS'21) to be run online, and is entirely free to SCSC members (see opposite for details).

The SCSC Working Groups have also been active in assessing the impact of Covid-19 and as well as a dedicated Covid-19 Working Group now under Peter Ladkin's lead, other groups such as Data Safety and Service Assurance have been revising their guidance material to address the novel challenges that the Coronavirus presents in these areas.

It's been a difficult year for airlines, not least Boeing who have been dealing with the grounding of their 737 MAX fleet after fatal accidents in Indonesia and Ethiopia as reported by Dewi Daniels at SSS'20. The 737 MAX fleet is now back in service in some locations, so our first article provides an update to Dewi's previous work, and asks whether the aircraft really is now safe to fly. It has been a great privilege to work with Dewi in producing this update for the newsletter.

In our second article, Kevin King focuses on the challenges of assuring Digital Twins from a safety perspective. He presents a set of principles, developed as part of his MSc thesis, to support the assurance of Digital Twins in their role in through-life safety assurance.

Pace continues to gather on the technological development of autonomous systems, but are people ready to trust these systems? In our third article, Davy Pissoort discusses a new initiative to develop a training network with the goal of establishing people's trust in autonomous systems, by making these systems demonstrably safer.

The safety community, and engineers in general, are approaching a retirement 'cliff' with a large proportion expected to retire in the next 5 years. The Covid-19 enforced hiatus is, perhaps, only accelerating matters. To address this, we must encourage and nurture the next generation of skilled safety engineers and in our fourth article, Zoe Garstang introduces an exciting new development to do exactly that – the SCSC Safety Futures Initiative (SFI).

Our 60 second interview is with Harold Thimbleby.

Paul Hampton
SCSC Newsletter Editor
paul.hampton@scsc.uk



In Brief



Why flight safety faces new dangers during downturn



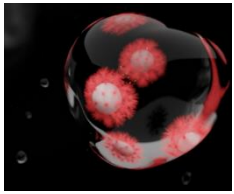
One uncomfortable aspect of the extraordinary crisis inflicted on air transport operations is the extent to which the pandemic might, at some point, be considered a contributing factor to aircraft accidents, despite – or perhaps because of – the reduction in flight activity. *flightglobal.com*



Police probes compromised after computer records deleted

Police investigations have been compromised by an error that led to hundreds of thousands of records being deleted from UK-wide databases, according to a letter seen by the BBC. The National Police Chiefs' Council said 213,000 records were deleted - more than the 150,000 first reported. This resulted in a couple of "near misses" for serious crimes when trying to identify an offender. *bbc.co.uk*

The ten worst Covid data failures



Throughout the pandemic, the government and its scientific advisers have made constant predictions, projections and illustrations regarding the behaviour of Covid-19. Their figures are never revisited as the Covid narrative unfolds, which means we are not given an idea of the error margin. A look back at the figures issued shows that the track record, eventually validated against the facts, is abysmal. *spectator.co.uk*



Sriwijaya Air crash places Indonesia's aviation safety under fresh spotlight Indonesia's poor air safety record is again in the spotlight after a Sriwijaya Air jet carrying 62 people crashed into the Java Sea minutes after take-off, marking the country's third major airline crash in just over six years. *reuters.com*

Boeing 737 Max flights resume



The American flight follows the Federal Aviation Administration's decision last month to clear the beleaguered aircraft's return to the sky. As part of this decision the agency has ordered Boeing and airlines to make repairs to a flight control system blamed for both crashes and increase pilot training. *cnet.com*

Boeing 737 MAX
Safe to Fly?



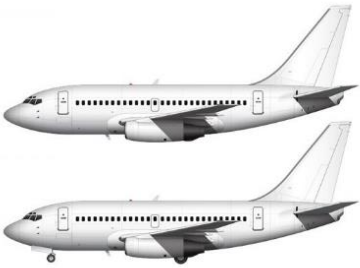
In March 2019, the Boeing 737 MAX aircraft, a newly designed aircraft from a highly respected manufacturer, was grounded after being implicated in two fatal accidents. After 20 months, the aircraft is now approved by the US regulator for commercial operations and has seen its first commercial flights. Paul Hampton and Dewi Daniels provides an update on events based on recent developments and ask if the aircraft is really safe to fly again.

Like the 2016 films 'Sully: Miracle on the Hudson' and 'Deepwater Horizon', it is possible that we may in the future see a film dramatizing the events surrounding the Boeing 737 MAX accidents – and what a tale it will be: a tragic story of human loss caused by failures in design, maintenance, operation and the regulation of aircraft, leading to not one, but two accidents; of lax safety practices, corporate cover ups, exit of the Boeing CEO and billion dollar fines. As the adage goes, you couldn't make it up.

Dewi presented an analysis of the accidents and causes at SSS'20 in Feb 2020 based on the interim reports and information that was available at the time. Since then, there have been a number of developments, such as the publication of updates to the accident reports [1, 2], the final report from the US House Committee for Transportation & Infrastructure [3] and details of the actions Boeing have taken to get the aircraft airborne again [4, 5]. We therefore present an update of that analysis to reflect the current situation and our story starts at the beginning – with the aircraft itself.

Boeing 737 MAX Evolution

The prototype Boeing 737-100 made its first flight in April 1967 and has undergone continuous evolution over the past 50 years. The Boeing 737 MAX is the latest variant, superseding the Boeing 737-700, -800 and -900.



To stay competitive in the market, the Boeing 737 MAX uses CFM Leap 1B engines that are more fuel efficient, but have larger fan diameters, and so need to be mounted further forward than on previous variants of the aircraft to provide sufficient ground clearance.



During wind tunnel testing however, Boeing found that the Boeing 737 MAX tended to pitch up during high angles of attack, which violated federal aviation regulations that demanded that the gradient of the stick force versus manoeuvring load factor must lie within satisfactory limits.

Boeing tried to fix the problem by changing the physical design of the aircraft, but no effective aerodynamic solution was found. They therefore decided to solve the problem in software by introducing a new system called the Maneuvering Characteristics Augmentation System (MCAS) developed by Rockwell Collins (now Collins Aerospace).



MCAS was designed to push the aircraft's nose down at a high Angle of Attack (AOA) to compensate for the tendency to pitch up, by changing, or 'trimming' the angle of the aircraft's horizontal stabiliser – an aerodynamic control surface on the aircraft's tail.

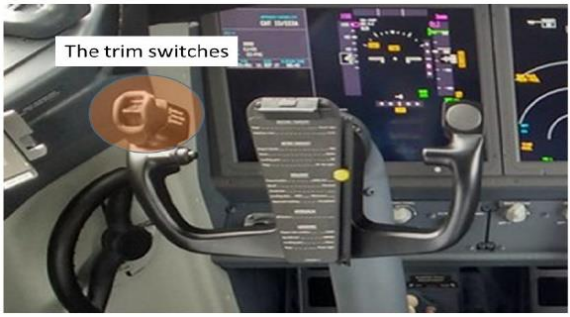
The Angle of Attack reading is taken from one of two independent AOA sensors mounted on the left- and right-hand sides of the aircraft. A physical vane protrudes from the device and changes as the flow of air passes over the vane. The MCAS system would take the AOA reading and other data, such as aircraft speed, and if the conditions warranted (flaps up, autopilot off, high speed, high AOA) would activate the horizontal stabiliser.



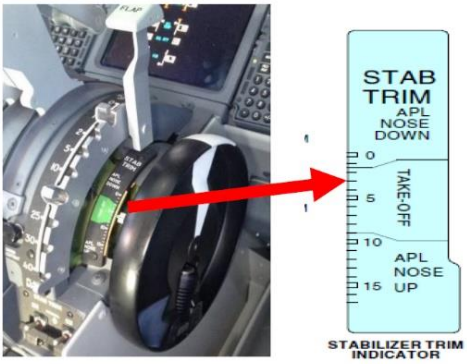
MCAS was added to the existing Speed Trim System, which in a similar fashion, helped the aircraft maintain the speed set by the pilot, usually during critical flight phases such as take-off, by trimming the stabiliser to force the aircraft's nose up and nose down as appropriate.

Boeing conducted a safety assessment of the MCAS system and concluded that failures in the system would lead to only "Major" hazards, as the maximum command authority was constrained to 0.6 degrees at high speeds, and failures in the system would manifest as a stabiliser 'runaway'.

Also, such a runaway, so it was assumed, would be readily identified by the pilots and the crew's non-normal checklists for such events would quickly instruct the pilots to use the electronic trim switch on the control column to counteract the behaviour, or if this didn't work, to switch off the electronic trimming function altogether and trim the stabiliser manually.



This assumption of prompt pilot reaction was critical – early in 2012 it took one test pilot over 10 seconds to respond to an uncommanded activation of MCAS in a flight simulator, a condition the pilot found to be catastrophic due to the inability to arrest the airplane overspeed [3].



Manual trimming is carried out by turning the black wheels next to the stabiliser trim (STAB TRIM) indicator. 15 turns are required to move the STAB TRIM by 1 unit.

A STAB TRIM value of 0 indicates full nose down and a level trim value for level flight, is at around 5 units.

A trim deflection of 0.6 degrees was applied in a simulator and the "Major" hazard assessment confirmed as it would not cause any significant handling difficulties.

A "Major" assessment required the failure condition to be shown to have a probability of 1 x 10^{-5} or less per flight hour and for the item to be developed to Development Assurance Level (DAL) C. This probability target is usually achievable by a single device – higher targets usually require extensive use of redundant components. Neither did DAL C assurance demand more detailed analysis of the impact of individual component failures (FMEA/FTA). This meant that failure of the AOA sensor, and the potential impact on other dependent aircraft systems was not considered or tested; and crucially, neither was the scenario where MCAS repeatedly activated.

Flight simulator testing however revealed that MCAS was not sufficient at low speeds and so, the decision was taken to increase the trimming authority to 2.5 degrees at these low speeds – 4 times the original – and incrementally decrease the authority as speeds increase to around 0.6 degrees. At low speeds, the hazards were not considered to be as serious as at high speeds and so no further hazard-based simulator testing was conducted.

The Road to Certification

The US regulator, the FAA, does not have enough resource to be closely involved in the review of all aircraft safety related activities and so can delegate this authority, in some cases, to trusted manufacturers by issuing an Organization Designation Authorization (ODA). Boeing's ODA meant they could effectively self-certify lower risk activities and bring only the more safety-critical aspects of any change to the FAA's attention. This is not an unusual arrangement and similar practices occur under the European regulator EASA.

FAA Order 8110.49A [6] specifies how to determine the level of certification authority involvement in a software project. Even under the previous regulatory regime before Boeing was given an ODA, the FAA would not have participated directly in Stage of Involvement (SOI) audits for a DAL C software project by an established supplier such as Rockwell Collins. In this instance, Boeing ODA would have performed the independent oversight.

Boeing's certification basis was to treat the Boeing 737 MAX as a change to existing model's Type Certification (TC). In doing so, an application for a new TC would be avoided, which would otherwise require the entire aircraft to be resubmitted for review. However, this is only possible if the FAA can be convinced that changes are not significant and Boeing reassured the FAA that this was the case.

The final committee report [3] however, is critical of Boeing's actions during this period, criticising their "opaque description of MCAS to some FAA officials and its limited description to others". The report also revealed that some Boeing staff did raise safety concerns about the MCAS design internally, but "these concerns were not investigated thoroughly enough and, in some cases, dismissed by other Boeing colleagues".

Transitioning to Live Operation

Boeing and the FAA engaged in extensive discussion about the appropriate content of Boeing 737 MAX training and manuals for a period of several years prior to Boeing 737 MAX certification. Ensuring that no simulator training was required was a "design objective" for the 737 MAX programme.

Boeing argued that MCAS was an automatic system that need not be mentioned in the crew's operation manual, which was finally agreed by the FAA. Exposure of MCAS to pilots would mean additional training and familiarisation for pilots thus making the aircraft transitional training more involved, with the associated time and cost impacts. In fact, Boeing's transition training was based on only a few hours of computer-based training with no simulator time [7], and the 13-page guide to the 737 MAX did not mention MCAS at all.

Had Boeing not been able to convince the FAA that non-simulator based training was sufficient, it could have cost the company $400 million with one airline alone due to contractual guarantees it had made [3].

"Boeing argued that MCAS was an automatic system that need not be mentioned in the crew's operation manual, which was finally agreed by the FAA.

The Fatal Flaws

It was of course understood that failure of the AOA sensor could trigger erroneous behaviour of MCAS. However, data from the AOA sensor is also used to feed into other aircraft systems, specifically those involved in calculating airspeed and altitude (ADIRU), and another for managing stall warning (SMYD).

These interdependencies meant that AOA sensor errors could lead to, amongst others, activation of the control column stick shaker and clacker (to warn of a stall when one was not actually present), and alert lights for indicated airspeed, altitude and feel differential pressure values (feel differential pressure is used to simulate aerodynamic force effects on the control column).



One might expect that a mismatch in AOA readings would also be alerted to the crew, and indeed this had been a standard feature on the previous model, the 737 NG. However, on the 737 MAX, the AOA Disagree alert was only annunciated if the airline had ordered the optional AOA Indicator display. This option had not been taken up by the accident airlines and so no AOA alert was presented to the pilots in both accident flights.

This dependency on the display option was not the intended functionality; it had been inadvertently introduced by the supplier when fixing another related issue, and was not detected by Boeing's testing of the fix. In fact, the issue affected more than 80% of the 737 Max fleet, a fact that Boeing did not publicly acknowledge until after the first accident [3].

Lion Air Flight 610 – Accident Flight

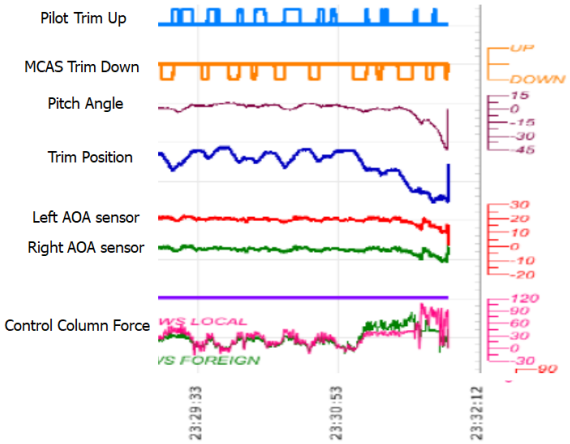


Lion Air Flight 610 was a scheduled domestic flight from Jakarta to Pangkal Pinang on 29 October 2018. This Boeing 737 MAX aircraft registered PK-LQP was a new aircraft whose Certificate of Airworthiness had been issued on 15th August 2018.

After earlier problems with a faulty AOA sensor, the aircraft had been refitted with a refurbished sensor, was most likely miscalibrated by the maintenance organisation used by the airline resulting in a 21 degree bias on the left-hand side AOA sensor.

Shortly after take-off, the left-hand stick shaker activated and continued for most of the

remaining flight. Airspeed and altitude mismatch alerts were also reported to the crew. The erroneous left-hand AOA sensor led to repeated attempts by MCAS to command the aircraft nose down. This was repeatedly countered by the pilot using electric nose-up trim and applying more nose-up aft force to the control column, while the First Officer tried to follow the non-normal checklist for the air-speed unreliable alerts. Increased workload and distractions from ATC communications meant that the checklist was never completed. Later, the Captain handed over control to the First Officer. As the aircraft accelerated, the resulting aerodynamic forces that built up made it impossible for the pilot to pull back on the control column; and a minute later, the aircraft was lost with 189 fatalities.



The figure opposite, adapted from the flight recorder data, shows the final minutes of the flight. The manual trimming conducted by the crew is shown at the top in blue, with MCAS trimming shown in orange. The pitch angle and trim positions show how the crew finally lost their trimming battle with MCAS.

Pilot performance during these events was by no means perfect, but not being aware of MCAS behaviour, absence of AOA disagree alerting and multiple concurrent alerts and stick shaker activation, meant the crew were ill-equipped to diagnose and mitigate the problems and did not consider cutting off electric trimming and trimming the aircraft manually, as was the anticipated behaviour in the original safety assessment.

Response to the Accident

Both Boeing and the FAA were quick to blame the pilots of Lion Air flight 610 for the crash and were slow to truly and fully understand and acknowledge how dangerous the design decisions and operational authority granted to the MCAS was to the flight crew and safety of the airplane [3].

On 6th November 2018, Boeing issued Operations Manual Bulletin (OMB) TBC-19 to emphasize the existing procedures provided in the runaway stabilizer Non Normal Checklist (NNC) and the next day the Federal Aviation Administration (FAA) issued Emergency Airworthiness Directive (AD) 2018-23-51 reiterating this procedure. However, Boeing and the FAA did not immediately alert MAX pilots of the existence of MCAS, and the FAA removed reference to MCAS from its draft AD [3].

"Both Boeing and the FAA were quick to blame the pilots of Lion Air flight 610"

MCAS however, eventually started to be named publicly in messages to operators and concerns were being raised. For example, the Pilot Union President called for accountability from Boeing and a reassessment of the FAA certification process [8].

7 weeks after the accident, in a presentation to the FAA, Boeing deflected blame and continued to assert that appropriate crew action would save the aircraft [9].

Ethiopian Airlines Flight 302- Accident Flight



Ethiopian Airlines Flight 302 was a scheduled international flight from Addis Ababa, Ethiopia to Nairobi, Kenya on 10th March 2019. The Boeing 737 MAX aircraft registered ET-AVJ was another new aircraft, which had been delivered to Ethiopian Airlines on 15th November 2018.

Again, shortly after take-off, the AOA sensors readings began to deviate with a difference being greater than 59 degrees for the rest of the flight. As with the previous accident, the stick shaker activated and airspeed and altitude values began to deviate. When the autopilot was disconnected and the flaps fully retracted, MCAS activated for 9 seconds putting the stabiliser at 2.1 units with the pilot requiring a column force of about 90lbs to keep the aircraft nose up.

A second MCAS activation was interrupted briefly by the crew applying electric trim and at 2.3 units the crew decided, as per the directives, to switch off the electric trimming function. The crew, even working together, were having difficulty pulling back on the control columns and were applying an average of 94 lbs. The First Officer tried to manually trim the aircraft stabilisers using the manual trimming wheel, but he reported that it wasn't working. This was because at 2.3 units of trim and an airspeed in excess of 300 kts, aerodynamic forces make the wheel impossible to turn manually.

A YouTube video [10] shows the control forces required to counter nose down trim at high airspeed. Subsequent tests show that even at 220 kts, and a more favourable trim position, the wheel is barely movable and something like 45 turns would be required to restore the trim to level position.

The crew appear to have decided to reactivate the electric trimming control, which allowed MCAS to activate a final time setting the stabiliser to 1 unit, from which the crew could not recover, and the aircraft was lost 23 seconds later with 157 fatalities.

Aftermath

Ethiopian Airlines suspended 737 MAX flights the next day and a cascade of groundings around the world took place with the FAA amongst the last after claiming there was no reason. Boeing's ODA to issue airworthiness certificates for the 737 MAX aircraft was also eventually revoked on 26th November 2019.

Congressman Sam Graves in a hearing of the Subcommittee on Aviation on 15th May 2019 [11] still maintained the position that the pilots were to blame and claimed that *"pilots trained in the US would have been successful"* in handing the emergencies on both jets, but there are significant doubts about this:

- On 10th May 2019, Aviation Week reported that a US-based Boeing 737 MAX crew tried to replicate the Ethiopian accident in a simulator [12]. The crew found that keeping the aircraft level required significant aft column pressure by the Captain, while aerodynamic forces prevented the First Officer from moving the trim wheel a full turn. The crew were eventually able to recover the aircraft by using a technique known as the 'roller coaster' procedure, which is not described in the Boeing 737 MAX flight manual. The pilot said he had only learned of the roller coaster procedure from excerpts of a Boeing 737-200 manual posted in an online pilot forum following the two Boeing 737 MAX accidents. Aviation Week concluded that *"the Ethiopian crew faced a near-impossible task of getting their 737 MAX 8 back under control".*
- Capt. Chesley "Sully" Sullenberger III, of "Miracle on the Hudson" fame, spent time in a simulator to experience the same conditions as the pilots and commented in an article to the New York Times Magazine, *"I know firsthand the challenges the pilots on the doomed accident flights faced, and how wrong it is to blame them for not being able to compensate for such a pernicious and deadly design. These emergencies did not present as a classic runaway stabilizer problem, but initially as ambiguous unreliable airspeed and altitude situations, masking MCAS. The MCAS design should never have been approved, not by Boeing, and not by the Federal Aviation Administration (FAA)."* [13]

On 30th October 2019, the Boeing CEO Dennis Muilenburg testified before US Congress committees, defending the company's safety culture, but damning internal correspondence reveals a very different environment: one of cynicism, open criticism of the 737 MAX design and its minimal transition training plan. Mulienburg resigned 2 months later.

On the 16th September 2020, the US House of Representatives released its concluding report placing the blame squarely on Boeing and the FAA for lapses in the design, construction and certification [3].

On 7thJanuary 2021, the US Department of Justice [14] fined Boeing $243.6 million in resolution of its criminal charges related to a conspiracy to defraud the FAA through its *"misleading statements, half-truths and omissions"* and states:

"Boeing's employees chose the path of profit over candor by concealing material information from the FAA concerning the operation of its 737 Max airplane and engaging in an effort to cover up their deception."

The settlement also includes Boeing's agreement for a passenger compensation fund of $500 million and $1.77 billion to compensate its airline customers [15].

Return to Service

The FAA in the US and EASA in Europe mandated the necessary changes (AD 2020-24-02 and AD 20-184 respectively) to allow the aircraft to return to service. These are summarised as follows [5]:

Changes to MCAS Software

- To use inputs from both AOA sensors and disable MCAS if outputs differ by more than 5.5 degrees
- To use algorithms to cater for unusual sensor behaviour such as sinusoidal variations
- To limit how much MCAS can move the horizontal stabiliser so it can always be counteracted by use of the control column alone
- To modify the activation schedule so MCAS can only operate once per high AOA event not multiple times

Changes to Documentations

- To revise many of the aircraft's Flight Manual and Quick Reference Handbooks to revise the procedures.
- To update the Flight Crew Operating Manual to described MCAS
- To enhance the maintenance documentation including the Airplane Maintenance Manual and Fault Isolation Manual

Changes to the Minimum Equipment List

To make changes to the minimum list of operational equipment that must be available before a flight can be safety dispatched. This includes, amongst others, ensuring both control wheel trim switch systems are operative.

Changes of the on board flight displays

- Making the AOA Disagree Alert standard and offering the AOA display as a free option

Conversion Training

Pilot aircraft type conversion training to cover speed trim and MCAS, existing crew procedures and related software changes. Pilots with 737 NG or MAX qualifications will require:

- 5 hours of transition training
- 3 hours of simulator briefings and live scenarios
- 2 hours of computer-based course work

Other Jurisdictions

Note that EASA and Transport Canada also requested additional features:

- To introduce a 3rd AOA sensor. As there are only 2 AOA probes, the aircraft will not know which one to trust in the event of a failure, and so, a 3rd sensor cross-check would allow this to be determined. A synthetic AOA sensor that computes data from a variety of other existing sources is suggested (similar to the synthetic airspeed concept on the Boeing 787)
- To introduce a means by which pilots can disable the stick shaker (eg. by pulling a

labelled circuit breaker) to stop the stick vibration and clacker to reduce noise and stress levels in the flightdeck, thereby making the situation easier to handle.

These additional features are not mandatory for returning the aircraft to service but may become so in a few years. Note also that the FAA has rejected the second request and the debate continues.

EASA have also gathered factual evidence that a single AOA failure can result in loss of flight guidance in certain approach scenarios and so will prohibit 737 MAX family aircraft from flying those types of approaches [16].

Regulatory Changes

The FAA is to reform how it certifies new airplanes in line with legislation passed by the US Congress [17]. The new law:

- Requires an independent review of Boeings safety culture
- Repeals rules allowing FAA employees to receive bonuses or other financial incentive based on meeting manufacturer-driven certification schedules or quotas
- Authorises civil penalties against aviation manufacturer supervisors who interfere with employees acting on behalf of the FAA
- Authorises new resources for the FAA to add key technical staff and requires it to review pilot-training.

Is it Safe?

The fundamental question is whether the 737 MAX is now safe, and one that the FAA and EASA firmly believe to be true. The head of the FAA, Steve Dickson, a former pilot, stated [18] that he was *"100% confident"* in the aircraft's safety and they have done *"everything humanly possible to make sure"*. This latter, of course, seems a little at odds with the omission of the 3rd sensor.

"The fundamental question is whether the 737 MAX is now safe, and one that the FAA and EASA firmly believe to be true"

The head of EASA, Patrick Ky, has stated that he is *"certain"* the aircraft is safe to fly and his organisation had *"left no stone unturned"* and they *"are very confident that it is now a very safe aircraft"* [19].

This is supported by the focus not being just on MCAS; the FAA had asked Boeing to audit other critical systems based on new assumptions about how long pilots may take to respond to emergencies [20]. This revealed two bundles of critical wiring used to control the tail, which were too close together and at risk of short circuit. The FAA also found other issues with the aircraft and have been fining them for other breaches, such as certifying a sensor, without it being tested or approved, as being compatible with the Head-Up Guidance System [3]. However, others are not so convinced – a recent report by former Boeing Manager, Ed Pierson, [21] raises questions around the build quality at the Boeing factory in Renton, Washington, which he claimed to be a "dangerously unstable production environment".

With FAA approval, the aircraft is nevertheless back in service in the US and Brazil with its first commercial flight taking passengers from São Paulo to Porto Alegre, along Brazil's eastern coast in December 2020.

However, there are concerns that persist. There is no doubt that the Boeing 737 MAX, including MCAS, has now had comprehensive and intense scrutiny, but modern airliners are hugely complex systems of systems. How do we know that MCAS-like problems are not lurking in other aircraft designs? Aircraft are also likely to become more complex with a greater reliance on automation.

How do we know that MCAS-like problems are not lurking in other aircraft designs?

Although the widespread changes to the processes and practises for the manufacturer, airlines, maintainers and regulator all look positive, this needs to be framed in an environment where airworthiness regulations are often criticised for being too onerous, placing an undue burden on airframe manufacturers and equipment suppliers resulting in unnecessary expense. Yet the 737 MAX issues still slipped through, because of drop-offs from those very processes and practices, which is a very human problem – not one fixable in software.

Modern airliners are exceptionally safe; Boeing's own Statistical Summary of Commercial Jet Airplane Accidents [22] reports that the 10-year average fatal accident rate for scheduled commercial passenger operations is 0.16 fatal accidents per million departures. Until now, both Boeing and Airbus airliners have had very low accident rates. However, the prolonged absence of accidents, especially those caused by software, can lead to a gradual degradation of safety margins as small deviations go unpunished until they accumulate to a much larger issue.

Is it Safe? Probably, for now…

References

[1] Aircraft accident investigation report, PT. Lion Mentari Airlines, Boeing 737-8 (MAX), PK-LQP. Published October 2019. http://knkt.dephub.go.id/knkt/ntsc_aviation/baru/2018%20-%20035%20-%20PK-LQP%20Final%20Report.pdf. Accessed 13 Jan 2021.

[2] Aircraft accident investigation preliminary report, Ethiopian Airlines Group, B737-8 (MAX) registered ET-AVJ, Report No. AI-01/19. Published March 2020.

[3] The design, development & Certification of the Boeing 737 MAX, September 2020, The House Committee on Transportation & Infrastructure. https://transportation.house.gov/committee-activity/boeing-737-max-investigation, accessed 22 Jan 2021.

[4] https://www.boeing.com/737-max-updates, Boeing, accessed 13 Jan 2021.

[5] http://www.b737.org.uk/mcas.htm#fix, Chris Brady, accessed 13 Jan 2021.

[6] Order 9110.49A Software Approval Guidelines. https://www.faa.gov/documentLibrary/media/Order/FAA_Order_8110.49A.pdf. Accessed 14 Jan 2021.

[7] https://qz.com/1574878/pilots-trained-for-boeing-737-max-with-one-hour-ipad-lesson/, accessed 14 Jan 2021

[8] https://www.alliedpilots.org/News/ID/7030/Pilot-Union-President-to-Call-for-Accountability-from-Boeing-and-Reassessment-of-FAA-Certification-Process, accessed 13 Jan 2021.

[9] https://www.seattletimes.com/business/boeing-aerospace/after-lion-air-crash-boeing-doubled-down-on-faulty-737-max-assumptions/, accessed 13 Jan 2021.

[10] Boeing 737 Unable to Trim!! Cockpit video (Full flight sim). https://www.youtube.com/watch?v=aoNOVlxJmow, from 10:06 to 13:47. Accessed 12 August 2019.

[11] Ranking Members Sam Graves & Garret Graves Statements from Hearing on Status of the Boeing 737 MAX, 15 May 2019. https://republicans-transportation.house.gov/news/documentsingle.aspx?DocumentID=404188. Accessed 9 August 2019.
[12] Ethiopian MAX crash simulator scenario stuns pilots. https://aviationweek.com/air-transport/ethiopian-max-crash-simulator-scenario-stuns-pilots, Accessed 14 Jan 2021.
[13] Sullenberger letter to the New York Times Magazine http://www.sullysullenberger.com/my-letter-to-the-editor-of-new-york-times-magazine/
[14] https://www.justice.gov/opa/pr/boeing-charged-737-max-fraud-conspiracy-and-agrees-pay-over-25-billion, accessed 13/01/2021.
[15] https://www.theverge.com/2021/1/7/22219370/boeing-737-max-fraud-conspiracy-criminal-charges-fine-crashes, accessed 13/01/2021
[16] https://aviationweek.com/air-transport/safety-ops-regulation/easas-max-return-proposal-adds-nuisance-stick-shaker-disabling, accessed 14 Jan 2021.
[17] FAA to reform new airplane safety approvals after 737 MAX crashes | Reuters, accessed 13 Jan 2021.
[18] https://www.reuters.com/article/boeing-737max-faa-int/faa-chief-100-confident-of-737-max-safety-as-flights-to-resume-idUSKBN27Y1N1, accessed 13 Jan 2021.
[19] https://www.bbc.co.uk/news/55366320, accessed 13 Jan 2021.
[20] https://www.nytimes.com/2020/01/05/business/boeing-737-max.html, accessed 13 Jan 2021.
[21] Boeing 737 MAX – Still Not Fixed https://img1.wsimg.com/blobby/go/ec12e28d-4844-4df3-a140-ca706a04c0f7/downloads/737%20MAX%20-%20Still%20Not%20Fixed.pdf?ver=1611597399852, accessed 25 Jan 2021.
[22] Statistical summary of commercial jet airplane accidents, worldwide operations 1959-2017. https://www.boeing.com/resources/boeingdotcom/company/about_bca/pdf/statsum.pdf. Accessed 5 August 2019.

Image Attributions
Unless specified, images are from the Lion Air Aircraft accident investigation report [1]
Top image: ID 202388759 © Patrick Barron | Dreamstime.com (Test flight 6/11/2020)
Lion Air Flight 610: https://commons.wikimedia.org/wiki/File:Lion_Air_Boeing_737-MAX8;_@CGK_2018_(31333957778).jpg licensed under the Creative Commons Attribution-Share Alike 2.0 Generic license.
Ethiopian Airlines Flight 302: https://commons.wikimedia.org/wiki/File:Ethiopian_Airlines_ET-AVJ_takeoff_from_TLV_(46461974574).jpg licensed under the Creative Commons Attribution-Share Alike 2.0 Generic license.
Boeing 737 engine designs: https://www.norebbo.com/
Modified MCAS schematic licensed under CC BY-SA
AOA sensor licensed under CC-BY

Paul Hampton, SCSC Newsletter Editor

Paul is a Chartered Engineer with over 30 years' experience in IT. He has spent 15 of those designing and developing enterprise systems in many diverse sectors and the remainder involved in system safety in a variety of capacities including: safety engineering, safety management, independent auditing and corporate governance and assurance.

Dewi Daniels, Software Safety Limited

Dewi is a Chartered Engineer with nearly 40 years' experience in the development and verification of safety-critical software. Dewi was one of the authors of DO-178C/ED-12C. He is currently a member of the RTCA/EUROCAE Forum on Aeronautical Software and one of the UK experts on the IEC 61508 committee.

The authors retains copyright of this article.

Assurance of Digital Twins for Safety-Critical Systems



Digital Twins are seen as a formative technology for the Internet of Things (IoT), central to future systems development. When Safety-Critical Systems fall in scope of Digital Twins, and it is clear they will, then proportionate assurance will be necessary. From work in support of his MSc thesis, Kevin King presents Principles to support the assurance of Digital Twins in their role in through-life safety assurance.

The Problem of Static System Assurance

Historically, when a product or system enters into service, the manufacturer presents an argument that their system is acceptably safe for its intended purpose when measured against its design intent, generally known as a system safety case, or in some fields, a safety assurance case. However, the arguments it sets out are typically static at the point of delivery, only evolving to reflect the system in service when a baseline update occurs, which is typically slow and reactive. There are many examples where system safety-cases have not kept track with their design intent, notably the Nuclear accidents at Three-Mile Island and Chernobyl, the Chemical accident at Bhopal and the Space Shuttle Challenger disaster. In all these cases, changes to the design, operating procedures or environmental envelope were contributory factors to the accident.

Moving to a Dynamic Assurance Solution

These examples were back in the 20th century and in recent years systems have become significantly more complex, predominantly borne out of a comparable increase in the resources available to support design, implementation and operational service. Parallel to this complexity growth has been the evolution of smart technologies to enable resource management of these ever more complex systems in a world where consumers expect 'more for less'. Crucially, these proposed smart solutions consider the systems they underpin through their entire lifecycle; pointing to the need for a more dynamic approach to their assurance, an approach that may have mitigated the incidents mentioned and many others.

For some time, industry has been seeking to understand what this complex future really means; with a broadly held view that it will be based on the 21st century concept of the IoT. Consequently, systems are expected to become more dynamic, a step change from the static solutions of the past. The expansion of system complexity, real-time data explosion, and growth in novel technologies such as Cloud Computing, Big Data Analytics, Machine Learning and Artificial Intelligence will all play their part in these dynamic IoT-based systems.

One technology proposed to underpin this new, dynamic world is the Digital Twin. A term first advocated by Grieves and Vickers in 2002 [1] and seen as a catch-all for the burgeoning collection of 'smart' technology-based assets that could be seen to collectively mirror a physical system. Crucially, the Digital Twin is seen as remaining interconnected with what could be viewed as its 'Physical Twin', through its entire life.

From that earliest formulation, there has been a growing view, notably since 2017, that promotes Digital Twins as the answer to the complexity/resource conundrum presented by the IoT. Equally, the dynamic nature of the Digital Twin is seen by some as a means to maintain system safety through life, albeit theoretically.

Issues of Confidence with Digital Twins

Considering the complexity of a Safety-Critical System, the Digital Twin paradigm adds a further dimension, which the author believes raises fundamental questions around maintaining confidence that it will provide an acceptable level of assurance of the safety of its Physical Twin dynamically through its in-service life.

Digital Twins have for some time been seen by some as an immature construct, with Gartner positioning them at the 'peak of inflated expectation' in the August 2018 issue of their Hype Cycle with the 'Digital Twin of the Person' still climbing the innovation trigger in 2020. To gain confidence as the concept matures, if Digital Twins are to become a mainstream, productive technology, then it is appropriate to consider what issues might hamper their development.

There are lots of positives to support Digital Twins in a potential safety-assurance role in service, however, there are some significant issues to ponder, notably:

- **Environmental Similitude:** Can we be certain the virtual environment of the Digital Twin equates absolutely to the substantive environment of its Physical Twin?
- **Twin Divergence:** Can we be certain that the twins will maintain their sibling bond? Will one or other of them diverge?
- **Twin Utility:** A view that a Digital Twin emulating 100 percent of its Physical Twin is unrealistic. A Digital Twin will only be of true value if it associates data from the Physical

Twin with processes in the digital model seen as useful, in context of the physical systems operation.

- **Data Practicalities:** Vast amounts of data are predicted to be available to the Digital Twin. How that data is collected, filtered, analysed and put to use is seen as essential for the Digital Twin to be of value, with some researchers proposing the use of Big Data Analysis to address traditional frailties. Equally, concerns exist over protocol mismatches across the data interconnection.
- **Data Ownership:** Digital Twins' life-blood is the data it receives from its Physical Twin, but who owns that data? If fed from the operator of the physical system, how content will they be for their data to be fed back to the entity maintaining the Digital Twin? Is there scope for an operator to give their data not only to the manufacturer of a device but also a competitor, contracted to manage the Digital Twin?
- **Ethical and Societal Impact:** What if the data fed to or derived by the Digital Twin is in any way sensitive? This is of utmost relevance in the health care sector where concerns over data privacy, derived benefits and transparency of use are key. Will Digital Twins provide therapeutic solutions or advise on 'enhancements' and who will be able to afford their personal medical replicant?
- **Managing the Twins:** The point in the lifecycle of the physical system at which its Digital Twin is born is crucial to their ongoing relationship, as is how it endures through the physical system's life. Tools exist to evaluate the performance of physical systems; equivalents need to be considered to manage the Digital Twin to maintain the digital to physical bond.
- **Hyperbole:** Some researchers believe that by 2021 around half of large companies will be using Digital Twins. This links back to the hype mentioned earlier but the author believes is unrealistic.
- **Cost Effectiveness:** Creating Digital Twins for one-off complex, Safety-Critical Systems is seen as time consuming and costly when weighed against the benefits the Digital Twin provides. Significant upfront costs of Digital Twin development may not be viewed across the whole life of the physical system, especially in a keenly contested consumer world where 'cost to market' is an essential factor.
- **Incongruent Digital Twins:** Although generally seen as whole lifecycle assets, there is a view that in some sectors, Digital Twins will be developed focussing solely on characteristics of the design and build phase of a system. Once that system enters into service a fresh 'in-service' Digital Twin will be developed, potentially by a different organisation (service provider) or even a collection of them. There is scope therefore for the 'in-service' Digital Twin to be distinct from its design and build predecessor.

Many researchers are focussing on case studies where Digital Twins can be applied, in effect presuming the concept of Digital Twins is established and not considering sufficiently the key facets of the Digital Twin and notably, where targeted at Safety-Critical Systems, the role the Digital Twin may play in the assurance of the safety-related aspects of that physical system.

Key Attributes of a Digital Twin

Before proceeding to set out an approach to support the assurance of Digital Twins, it is appropriate to understand exactly what a Digital Twin is. Many definitions exist; breaking these down and considering across research in this area, Digital Twins are seen to have the following key attributes:

- The precise, ultra-high-fidelity replication of relevant facets of a physical system including the processes, people, interfaces and environment with which it interacts
- A unique portrayal of an instance of a physical system in a close one-to-one relationship
- The effective use of real-world data to iteratively maintain the sibling bond between the digital and physical twins through the whole product lifecycle
- A data interconnection that provides consistency and context between the digital and physical twins
- The ability to effectively simulate and predict events without impacting the physical twin
- The ability for the physical twin's performance to be enhanced, life to be extended and operating costs reduced through learning from its Digital Twin
- Components that exist wholly in the virtual world and can potentially be widely distributed
- Ability to provide quantifiable reliability and digital certification of the physical twin

Distilling these attributes, the author has formed a pertinent definition of a Digital Twin.

Digital Twin
A dynamic, comprehensive, wholly digital one-to-one representation of a physical system, encompassing its structure, components, properties, users, behaviours, interfaces, and target environments. The Digital Twin shall be born when the embryo of its physical sibling is conceived and should evolve dynamically, remaining intrinsically linked to it. The sibling bond shall be maintained through the life of the physical system through a data interconnection, providing real-world, near real-time data. The Digital Twin shall be able to effectively simulate and predict events without impacting the physical system, enabling enhanced performance, supporting potential life extension, and reducing operating costs.

When is the Digital Twin of most benefit?

If assurance on entry into service is a mature concept across industry, why do we still witness industrial incidents such as Deepwater Horizon? It would seem that systems assurance wanes over time. Theoretical research sees the Digital Twin as a core technology in systems management through life, but can we be certain once their Physical Twin enters into service that will include an ongoing assurance role? The critical issue is the symbiosis between the Digital and Physical Twins which is expected to endure through the systems life. The evolving life of systems in service would appear to be a fundamental factor influencing the role of Digital Twins. Thus, it makes sense to understand the best solutions to both assure Digital Twins and provide sufficient confidence in their role in assuring their Physical Twin.

Research and industry have focussed across a range of areas where Digital Twins are seen as beneficial, many of these being safety-critical, including autonomous vehicles [3], healthcare [4], aerospace [5], and the Rail Sector [6].

Of particular interest to the author is the defence industry which finds its systems no longer immune to the IoT explosion. The United Kingdom's Defence Science and Technology Laboratory (DSTL) have defined the Internet of Military Things (IoMT), which is expected to employ the technologies and concepts of the IoT. Although defence systems historically have run in isolation for reasons of safety and security, the evolution of web-based Maintenance Management Systems coupled with budgetary efficiencies brings defence systems into the scope of the Digital Twin. Linked to the Internet of Military Things, it would also seem realistic to expect a data explosion in defence systems employing secure cloud-based data management e.g. to interlink battlefield assets with command posts, a 21st century implementation of the datalink technology of the 1990s. Digital Twins are likely to play as important a role in the IoMT as they are in the IoT, making their capability for providing 'sentient-assurance' through the life of a system as important in that arena.

Sentient Assurance

It would seem logical that systems assurance will need to keep pace with the 'smart', dynamic solutions within the new IoT-based world and evolve beyond the fundamental shortcomings of the static, reactive approach to a more dynamic, proactive, 'sentient-assurance' methodology. As a 'living' concept, Digital Twins could play an intrinsic role in providing that 'sentient-assurance'; ensuring the Physical Twin they support maintains its design intent through its operational life. Where the system being twinned is safety-critical, there is an added requirement to consider if any certification or accreditation of the Digital Twin is required.

Digital Twins are gaining traction as a 'sentient-assurance' methodology but dependence on them could come at a cost. If they are relied upon without any plausible certification or accreditation, it is conceivable that they could provide a misleading view of confidence in the safety of their Physical Twin. It would appear essential that a credible methodology is needed against which Digital Twins can be established to provide the 'sentient-assurance' of Safety-Critical Systems through their entire lifecycle, notably when in service. This methodology would also support the certification and accreditation being sought.

The Digital Twin Assurance Gap

Commonly, approaches to assurance of Safety-Critical Systems need to confidently demonstrate that all risks to life from the system have been considered with all intolerable risks eliminated and all other risks, through mitigation, reduced As Low As Reasonably Practicable (ALARP) in line with relevant guidelines, e.g. The Office of Rail Regulation, Railway Safety Principles and Guidance [7]. In many domains that approach is further underpinned by some form of independent accreditation or certification, such as the Certificate of Airworthiness sought by Farnell et al. [8]. For Digital Twins to be seen as a key pillar in the assurance of Safety-Critical Systems then it is sensible to believe they will require equivalent objective endorsement, likely accreditation and certification.

Glaessgen and Stargel [9] believe Digital Twins will "revolutionise certification". However, the author's research across the fields of safety governance, accreditation, and certification was unable to find examples of Digital Twin certification that address the need identified. There is an assurance gap that needs to be addressed. That research covered the main sectors where Safety-Critical Systems predominate, and the related standards, e.g. SAE ARP 4761, ISO 14971, ISO 26262, Def Stan 00-056, ISO 61508.

As the NZ Government's Chief Digital Officer states [10] "A principles-based approach provides confidence in the delivery of outcomes without resulting in excessive levels of assurance". Therefore, to support the assurance 'gap' some generic principles would seem appropriate.

Several researchers have proposed what they term frameworks or models applicable to Digital Twins, some in a system safety context and some in related technologies such as Cyber Physical Systems. Importantly, none explicitly address the endorsement of Digital Twins for Safety-Critical Systems. Rather than dismiss these various works for not providing the whole solution to the gap, it seemed sensible to consider the benefits they present that could be valuable in considering assurance in the Digital Twin paradigm.

The Approach to Developing Digital Twin Assurance Principles

Rather than defining a framework or model, it was assessed that supporting principles would provide a more generic, pan-sector solution for the assurance gap. In developing these principles, it was necessary to understand what form they would take and what methodology should be applied to ensure the results were credible and repeatable.

There are two main styles of principles; they are either based on characteristics and behaviours as with the Gemini Principles [13], where artefacts exhibiting a sufficiency of the characteristics and behaviours can be considered compliant with the principles, or rule-based, as in the Software Safety Principles [14] where compliance with the rules means an artefact has met the principles. This research employed the former style given the rapidly changing theories in the Digital Twin paradigm and need to understand the characteristics exhibited by Digital Twins that fall in scope of assurance.

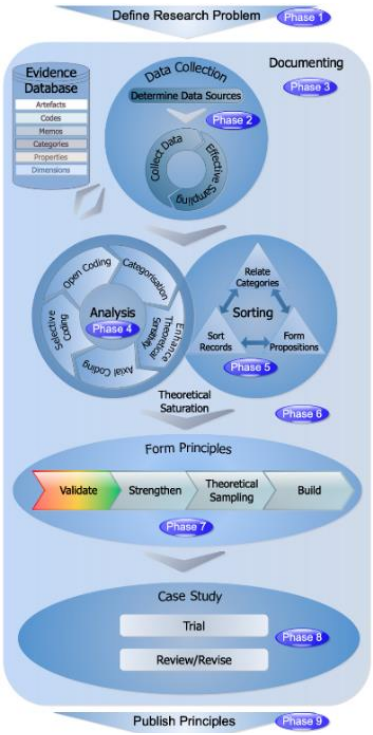


Figure 1 Refined Grounded Theory

The methodology chosen to support the development of assurance principles needed to be set out clearly to maintain credibility as the Digital Twin paradigm moves towards the 'Plateau of Productivity' [2]. Considering the scope of material available across industry and academia and the wide range of theories they present made it difficult to approach data analysis with any preconceptions. Practically, the approach chosen needed to reduce instability in the results by ensuring the design is repeatable within agreed parameters. Significant variability of outcomes from the methodology would cast doubt over its

veracity. Three aspects needed to be ensured. Firstly, the chosen technique needed to stand up to academic scrutiny. Secondly, the methodology needs to be sufficiently generic to hold value across all sectors of industry. Finally, the themes that are derived by the application of the methodology must be grounded in real-world data to be seen as useful. This directed to the use of a refined application of the social sciences-based technique of Grounded Theory Analysis (GTA) (Figure 1). As with all qualitative approaches, themes emerge from the analysis undertaken. Importantly by employing GTA, those themes are 'grounded' uniquely in the chosen datasets analysed. Furthermore, GTA is structured around a set of clearly defined activities, offering consistency, credibility and repeatability for the analysis undertaken.

Details of the refinement of the GTA and how it was applied are in the author's paper on Digital Twin Assurance presented at SSS '21 [12].

The Principles

Application of the refined GTA led to the emergence of 7 principles as set out below. Should a Digital Twin of a Safety-Critical System exhibit these principles then it could be considered to be providing an acceptable level of assurance of its Physical Twin.

Detailed objectives have also been proposed for each principle, but for brevity, these are not included in this article. Rather they can be found in the author's paper on Digital Twin Assurance presented at the Safety-Critical Systems Symposium 2021 [12].

In addition to the emergent principles, it is appropriate to consider the level of risk the Digital Twin presents to its Safety-Critical Physical Twin. Therefore, an additional +1 principle is proposed in line with the approach taken by Hawkins et al in defining the Software Safety Principles [14].

1 Conformance



Outline

All constituent elements of a Digital Twin must conform to the context and intended use of the Physical System with which it is being twinned. They must also conform to any applicable Digital Twin standards and regulations in place. Importantly such conformance must be maintained through the life of the Physical System which itself would be expected to conform to relevant standards for the sector in which it is operating.

Digital Twins will be expected to align to a specified level of data abstraction consistent with the equivalent levels in their Physical Twin.

Rationale

For a Digital Twin to be of any value it needs to be appropriately aligned to the Physical System it is to be twined with centred around its intended purpose. The extent of that veracity depends on:

1. The quality of the data that the twin will be reliant on
2. The fidelity of the algorithms within the twin's constituent models
3. The conformity of the source code within the digital representation
4. The quality of the twin's output
5. The level of safety criticality in the Physical Twin.

2 Awareness



Outline

Digital Twin processes of Safety-Critical Systems will need to be able to be prognostic/proactive and identify future unsafe events through perception of the Physical Twin's environment, comprehension of the current situation, and projection of future states. Such situational awareness and decision making should trigger mitigating controls in the Safety-Critical Physical Twin which positively influence the path it takes. Digital Twin awareness should also mitigate against divergence across instances.

Rationale

Consideration needs to be taken of the predictive nature of assurance that technologies such as Machine Learning and Artificial Intelligence require. For Digital Twins to be perceptive (i.e. have situational awareness and make appropriate decisions to ensure the Physical Twin remains acceptably safe), the whole asset lifecycle is an essential element of this paradigm.

3 Interoperability



Outline

Digital Twins of Safety-Critical Systems must be able to establish equivalence between their constituent representation models, the 2-way data interconnection and their Physical Twin. Open standardised data formats and sector independence must form the basis to enable such interoperability.

Rationale

Without a clear understanding of data formats, standards and hosting domains there is scope for Digital Twins to grow incongruent with their Physical Twin. Equally, for safety-critical applications it is essential that the Digital Twin is able to account for changes in standards and data formats through the service life of its Physical Twin.

4 Credibility



Outline

For Digital Twins to be credible, the relationship between them and their Physical Twin needs to be relevant, assured, secure and maintained at an appropriate level of integrity. All models that constitute the Digital Twin and all data that it interacts with need to consistently be of an acceptable quality. Evidence of the Digital Twin's credibility needs to be maintained in accordance with the *Curation* principle.

Rationale

Digital Twins need to maintain a level of similitude with their Physical Twin conforming to their design intent (which may have evolved), otherwise the assurance they provide cannot be considered credible.

The information interchange between Digital and Physical Twins is critical, the life-blood of their relationship. The integrity of that information is therefore essential.

5 Curation



Outline

For Digital Twins to provide precision in supporting the assurance of their Physical Twin, it makes sense that all artefacts within the Digital Twin paradigm are maintained in a configured, orderly way with assured records. Curation in this way provides an audit trail of changes to the Digital Twin to support the *Conformance* and *Evolution* principles.

Rationale

In Safety-Critical applications, to ensure a Digital Twin maintains conformance with its requirements, design and through-life operation, it is essential that consistency of performance is maintained. Notably, given the impact of Machine Learning and Artificial Intelligence this should include all artefacts in the Physical Twin and Data Interconnection.

6 Evolution



Outline

Digital Twins are expected to grow in parallel to the growth in their Physical Twin. For Safety-Critical Physical Twins that growth needs to be in a controlled manner facilitated by assured protocols and facets of the *curation* principle.

Rationale

Cyber Physical Systems that Digital Twins will be related to will evolve through their service life. They will likely involve Artificial Intelligence and Machine Learning technologies and be repurposed during their working life. The Digital Twin will need to evolve to maintain that relationship.

7 Clarity



Outline

For Digital Twins of Safety-Critical Systems it is appropriate to have clarity around the safety-related goals and priorities of their Physical Twin and the expectations of the owners/operators of it. This clarity needs to permeate through all artefacts in the virtual world of the Digital Twin.

Rationale

Aligned to the *Evolution* principle, it is sensible to develop a clear understanding of the functionality of all artefacts of the Digital Twin from cradle to grave of its Physical Twin. Clarity would also provide benefit in the structure and management of relationships throughout the Digital Twin Paradigm. Not just the inter-relationships betwixt Digital and Physical Twins but also the intra-relationships between the virtual components (models) in the Digital Twin.

+1 *Proportionality*



Outline

Across all 7 principles it is important that the effort applied is balanced against the level of risk the Digital Twin presents in the assurance of the safety properties of the Safety-Critical System it is supporting.

Rationale

In the context of Safety-Critical Systems, excessive effort expenditure to assure Digital Twins of low-risk Physical Systems is considered disproportionate to the management of that risk in line with the principles set out in the UK Health and Safety Executive Process "Reducing risks, protecting people". The application of these principles for Digital Twin Assurance is seen as requiring a similar approach to proportionality of application.

NOTE: This principle will need to adapt to account for changes in the required level of assurance where the Digital Twin is related to Machine Learning/Artificial Intelligence based Physical Systems.

Related Work

During this research, it has become evident that an ISO standard for Digital Twins (BS EN ISO 23247-1) is under development. Currently, that draft standard does not consider safety-assurance.

It is envisaged that parallel activities discussing issues of cyber-security in Digital Twins will ensue, so why not overlay this research onto that subject? The author is aware there are security into safety principles in the form of a new IET code of practice [16] and plans to engage in that arena to build on the Code of Practice.

The majority view appears to be that Digital Twins will by their nature be built on software. This raises a question over levels of Digital Twin integrity and whether BS EN 61508 and Safety Integrity Levels (SIL) need to be considered in the context of Digital Twins. Also, worth considering are Levels of Digital Twin Assurance supporting levels of data abstraction and safety criticality, akin to the Data Safety Assurance Levels [15].

Conclusion

This research has presented a set of 7+1 pan-sector assurance principles to provide confidence in the capability of a Digital Twin to safely support a Safety-Critical System.

Accepting that in reality the Digital Twin phenomenon is still in its infancy, if we are not careful, industries will run head-long into implementing them and trip over some substantial assurance hurdles they don't understand, and crucially, are not being required to by their independent regulator or relevant standards. Importantly, the timeframe to Digital Twin maturity is also uncertain.

There is a clear need for these principles, however, that must be caveated in scope of the 'hyped' domain of Digital Twins from which they evolved. Equally, it would be illogical to promote them as any form of panacea to the assurance need without a sufficiency of empirical evidence to support such promotion. This aligns with the view of Rae et al. [17] who believe that "safety science" needs to be firmly bedded in reality.

It seems sensible that, aligning with their manifesto, future practical work is enacted across industry to strengthen the foundations of the 7+1 principles before they can be considered substantively plausible and importantly are allowed to extend or evolve. It is naïve to expect the researcher to move the principles into that plausible space in isolation; the principles need exposure.

References

[1] M. Grieves and J. Vickers, 'Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems', In: Transdisciplinary Perspectives on Complex Systems 2017, pp. 85-113.

[2] K. Panetta, '5 Trends Drive the Gartner Hype Cycle for Emerging Technologies, 2020', Gartner, 18 August 2020 [Online]. Available: https://www.gartner.com/smarterwithgartner/5-trends-drive-the-gartner-hype-cycle-for-emerging-technologies-2020 [Accessed: 20-Oct-2020].

[3] M. Sharma and J. P. George, 'Digital Twin in the Automotive Industry: Driving Physical-Digital Convergence', White Paper, TATA Consultancy Services, 2018.

[4] K. Bruynseels, F. Santoni de Sio and J. van den Hoven, 'Digital Twins in Health Care: Ethical Implications of an Emerging Engineering Paradigm', Frontiers in Genetics, vol. 9, 13 Feb 2018.

[5] H. Aydemir, Z. Ugur and D. Umut, 'The Digital Twin Paradigm for Aircraft Review and Outlook,' 2020, doi: 10.2514/6.2020-0553.

[6] R. Tomar, J. Piesk, H. Sprengel, E. Isleyen, S. Duzgun and J. Rostami, 'Digital twin of tunnel construction for safety and efficiency', Proceedings of the World Tunnel Conference 2019 pp. 3224-3234, 2019.

[7] Office of Rail Regulation, Railway Safety Principles and Guidance, UK Health and Safety Executive, 2005.

[8] G. P. Farnell, A. J. Saddington and L. J. Lacey, 'A new systems engineering structured assurance methodology for complex systems', Reliability Engineering & System Safety, vol. 183, pp. 298 - 310, 24 Nov 2018.

[9] E. H. Glaessgen and D. S. Stargel, 'The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles', 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, vol. 9, 14 Jun 2012.

[10] NZ Government's Chief Digital Officer, 'Principles of good assurance', 2019 [Online]. Available: https://www.digital.govt.nz/standards-and-guidance/governance/system-assurance/all-of-government-portfolio-programme-and-project-assurance-framework/principles-of-good-assurance [Accessed: 20 Oct 2020]

[11] V. Bolbot, G. Theotokatos, L. M. Bujorianu, E. Boulougouris and D. Vassalos, 'Vulnerabilities and safety assurance methods in Cyber-Physical Systems: A comprehensive review', Reliability

Engineering & System Safety, vol. 182, pp. 179 - 193, 2019.
[12] K.C. King, 'Digital Twins for Safety-Critical Systems: A rational approach to forming principles to support their role in safety assurance', Feb 2021, Proceedings of the 29th Safety-Critical Systems Symposium, UK.
[13] Bolton A, Enzer M, Schooling J et al, 'The Gemini Principles: Principles to guide the development of the National Digital Twin,' ed. Centre for Digitally Built Britain: Centre for Digitally Built Britain, 2018.
[14] R. D. Hawkins, I. Habli and T. P. Kelly, 'The Principles of Software Safety Assurance', presented at the International System Safety Conference, Boston, USA, 01-Aug-2013.
[15] Data Safety Initiative Working Group (DSIWG), Data Safety Guidance, v3.3. SCSC-127F (https://scsc.uk/scsc-127F), SCSC, 2021.
[16] Williams, P. and St. John Green, M. 'IET Code of Practice for Cyber Security and Safety', The Institute of Engineering and Technology, v1.0 ISBN 978-1-83953-318-1, 2020.
[17] A. Rae, D. Provan, H. Aboelssaad, and R. Alexander, "A manifesto for Reality-based Safety Science," Safety science, vol. 126, 02/12 2020, doi: 10.1016/j.ssci.2020.104654.

Kevin King

Kevin is Product Safety Functional Manager for BAE Systems Maritime Services, based in Portsmouth. He has over 35 years' experience in the defence industry initially as a software engineer, but for the last 12 years working in Product Safety across a wide portfolio of products including torpedoes, radars, small boats and training systems. Kevin has recently completed an MSc in Safety-Critical Systems Engineering at the University of York. His thesis for that degree forming the basis for this article.

All opinions expressed in this article are entirely his own and do not reflect the views of his employer.

this article is © BAE Systems PLC.
All images are copyright the author unless stated below:
Digital Twin Header image: ID 128043723 © Leowolfert | Dreamstime.com
Icons: free resource from www.flaticon.com [Accessed 2nd July 2020].

SCSC
FOR EVERYONE WORKING IN SYSTEM SAFETY

Seminar: Safety Culture Developments

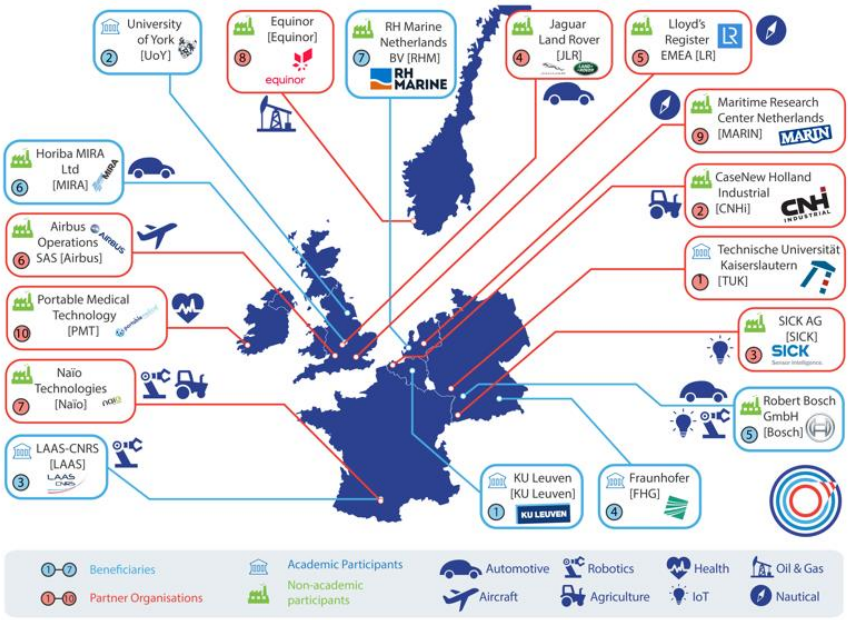
THE SAFETY-CRITICAL SYSTEMS CLUB, Seminar:

Safety Culture Developments

Tuesday 13 April, 2021 - Zoom

This online seminar, held from 14:00-16:30 UK time, will look at new developments in Safety Culture and will feature results from the SCSC SCWG Online Survey. Further details TBA.

Safer Autonomous Systems Training Network



Few people are really ready to trust autonomous machines and are much less prepared to tolerate a mistake made by them. Davy Pissoort, professor at KU Leuven Bruges Campus in Belgium, discusses the Safer Autonomous Systems Marie-Curie Innovative Training Network – a four-year initiative looking at ways to establish people's trust in autonomous systems by making these systems demonstrably safer.

Introduction

The coming of autonomous systems doesn't just mean self-driving cars. Advances in artificial intelligence will soon mean that we have drones that can deliver medicines, crew-less ships that can navigate safely through busy sea lanes, and all kinds of robots, from warehouse assistants, to search-and-rescue robots, down to machines that can disassemble complex devices like smartphones in order to recycle the critical raw materials they contain.

As long as these autonomous systems stay out of sight, or out of reach, they are readily accepted by people. The rapid and powerful movements of assembly-line robots can be a little ominous, but while these machines are at a distance or inside protective cages we are at ease. However, in the near future we'll be interacting with "cobots" – robots intended to

assist humans in a shared workspace. For this to happen smoothly we need to ensure that the cobots will never accidently harm us.

This question of safety when interacting with humans is paramount. No one worries about a factory full of automated machines that are assembling cars. But if these cars are self-driving, then the question of their safety is raised immediately. People lack trust in autonomous machines and are much less prepared to tolerate a mistake made by one. So even though the widespread introduction of autonomous vehicles would almost eliminate the more-than 20,000 deaths on European roads each year, it will not happen until we can provide the assurance that these systems will be safe and perform as intended. And this is true for just about every autonomous system that brings humans and automated machines into contact.

Autonomous vehicles, indeed all autonomous systems, need to be made safe enough so that people trust them. The destination, therefore, is clear; the route, however, is a difficult one. The Safer Autonomous Systems ITN project is designed to get us to our destination, safely.

Until recently, safety assurance has been integrated into the design processes, based on safety standards and demonstrating compliance during the system's test phases. However, existing standards were developed primarily for human-in-the-loop systems, where a human can step in and take over at any time. They do not extend to autonomous systems, where behaviour is based on pre-defined responses to a particular situation. What's more, current assurance approaches generally assume that once the system is deployed, it will not learn or evolve. On the one hand, advances in machine learning mean that autonomous systems can be given the potential to learn from their mistakes, and the mistakes of all the systems they are connected to, making their abilities to operate safely infinitely better than previous generations. On the other hand, machine learning means more uncertainty about how the system will decide to react to a particular circumstance in the future, making safety assurance a hard task, which can only be accomplished by a highly-skilled, interdisciplinary workforce.

Are you ready yet, to take a seat on an autonomously controlled airplane? If you hesitate to say "yes", then you are tacitly acknowledging the need for a training and research programme such as the *Safer Autonomous Systems (SAS) Innovative Training Network (ITN).*

SAS Objectives

The main objective of the Safer Autonomous Systems (SAS) [1] project – which started on 1st Nov 2018 and runs for four years in total – is to identify ways that we can establish people's trust in autonomous systems by making these systems demonstrably safer. In order to achieve this objective, we identified three challenges to be addressed by the early-stage researchers (ESRs) in their 15 individual research projects. This simply-stated objective, and the interdisciplinary needs required for its realization, is of such complexity that we saw a large training network involving some of Europe's flagship companies – such as Bosch, Airbus and Jaguar Land Rover – together with leading European universities – like KU Leuven and the University of York – as the best way to tackle these challenges, which are briefly described as follows:

- Increased autonomy, by definition, means a significant reduction of the time during which a human is involved in the system's decision making, thereby reducing the residual control afforded to humans. Studies have shown that it may take minutes for a non-actively involved human operator (e.g. a passenger in a self-driving car) to take over control in case of an emergency. Moreover, just putting a self-driving car to a full stop on a busy highway by removing its power (so-called fail-stop behaviour) is definitely not a safe action. In contrast, an autonomous system should be fail-operational (perhaps with reduced functionality) under all circumstances, monitor its own safety and make its own decision about a sensible and safe reaction. The challenge therefore is to design autonomous systems in such a way that they remain safe under all conditions, even in the case of component failures.
- Testing is the most intuitive way to reveal unsafe behaviour. However, autonomous systems must operate in a near infinite range of situations. When we test autonomous systems, we must therefore systematically determine which range and diversity of situations should be simulated and tested. We need to test them on roads, in the rain, and with people in the way. We need to test them when they're in intermittent supervisor contact and when they've got an unbalanced wheel. And we need to test them in all the possible combinations of those cases. Testing autonomous systems in the field is clearly too costly and too time consuming and might even be harmful for the system or its environment. Hence, virtual model-based testing is the only viable option. However, breakthrough solutions are required to guarantee the rigour of our virtual testing and to optimize its overall coverage.
- More autonomy is possible only through new technologies, e.g., machine learning, for which no accepted safety-assurance strategies currently exist. Legacy experience as well as established standards and regulations are lacking. Implicitly or explicitly, current safety-assurance practices and safety standards assume that the behaviour of the system is known at the design stage and can be assessed for its safety prior to system deployment. As autonomous systems might learn and evolve over time, this is no longer possible. This means that meeting the current safety standards for autonomous systems is either impossible to do or completely insufficient to assure safety throughout the life time of the system.

To achieve the main Scientific/Technical (S&T) objective of trust in autonomous systems by overcoming the 3 challenges, we decided on three sub-objectives that are the aims of the project's 3 research Work Packages (WPs):

- **Objective 1:** To integrate guaranteed acceptably safe behaviour directly into the architecture/design of the autonomous system (WP1)
- **Objective 2:** To prove by model-based safety-analysis techniques that the behaviour of an autonomous system remains acceptably safe under all possible conditions (WP2)
- **Objective 3:** To ensure that the safety-assurance strategies that combine the architectural/design measures with the evidence allow us to have trust in the autonomous system, which is very likely to be learning and evolving (WP3).

Overview of the SAS consortium

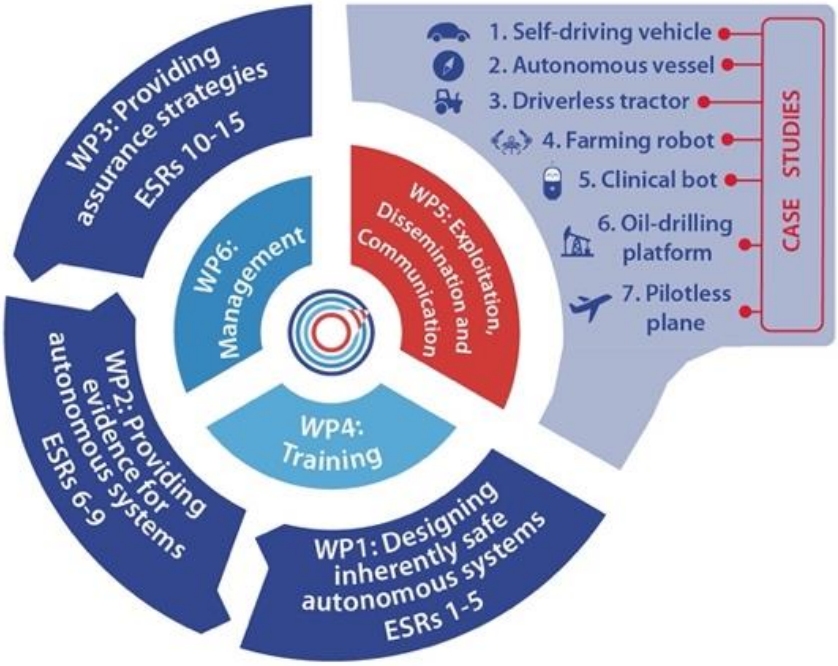
SAS provides specialised training to 15 early-stage researchers (ESRs), for 540 researcher months. The 15 ESR positions are distributed among 7 Beneficiaries in 5 countries. In addition, SAS has 9 non-academic Partner Organisations and one academic Partner Organisation.

The SAS consortium includes the 2 leading members (University of York and Lloyd's Register) of the Assuring Autonomy International Programme (AAIP) – which is the perfect partner for SAS. The AAIP is a prestigious initiative that targets the development of standards, open toolsets and training resources for assurance and the public acceptance of robotics and autonomous systems by collaborating with industry, regulators and research teams around the world.

The already well-established collaborations between the institutions involved will ensure that the network runs smoothly, while strengthening the interactions and the exchange of academic and non-academic resources. SAS aims to actively research the development of safer autonomous systems at multi-nationals like Bosch, medium-sized companies like MIRA and RH Marine, but also to stimulate the development of new safety designs, modelling and assurance techniques by involving the ESRs in SMEs and, potentially, their own start-ups.

SAS Work Programme

The SAS project is based on 6 Work Packages (WPs), three of which are S&T WPs (WP1–3), one for training (WP4), one for Exploitation, Dissemination and Communication (WP5) and one for Management (WP6).



The S&T WPs are organized along 3 research tracks covering the 3 main steps in the safety-assurance process:

- building safety and dynamic risk mitigation into the system by design
- gathering evidence that the behaviour of the system will actually be safe
- combining these into a clear strategy that allows us to put our trust in the system.

WP1: Designing inherently safe autonomous systems



WP1 involves 5 ESRs and tackles the actual safety-aware design, i.e., making safety inherently part of the design process of resilient autonomous systems.

ESR1 and ESR2 take up the challenge of *developing generic frameworks to monitor and handle the safety of autonomous systems during run-time*. The role of such a safety monitoring-and-handling framework is to observe the system and its environment and to trigger interventions that maintain the system's safety. For non-autonomous systems, a human operator takes a significant role in this fault-monitoring and, definitely, in the fault-handling. In contrast, versatile autonomous systems will have to deal with a much richer set of safety rules. Moreover, these safety rules have to take into account the wide application of machine learning in autonomous systems, causing them to evolve over time, and the a-priori largely unknown open-context in which autonomous systems will be applied. ESR1 and ESR2 will be working on related, but complementary projects, with ESR1 having the task to extend current safety-monitoring frameworks such that they cover the whole chain from safety-constraint definition to the actual autonomous reactions to avoid a possible hazard, while ESR2 starts from a MAPE-K cycle (i.e., Monitoring, Analyse, Plan and Execute based on Knowledge) to enable real-time adaptations of functionality, structure, and fault-tolerance mechanisms in order to assure the run-time resilience of autonomous systems.

ESR3 will run in parallel with ESR1 and ESR2 and go one step further and integrate *dynamic safety handling of autonomous systems-of-systems through run-time safety contracts* into the adaptive safety monitoring and handling framework. Driven by trends like ubiquitous computing and cyber-physical systems, new application domains for autonomous systems-of-systems have emerged. In such systems-of-systems, different devices are combined during run-time to fulfil higher-level emergent functionalities in a collaboration that cannot be provided by one of the involved systems on its own. Classic safety assurance relies heavily on a complete understanding of the structure and behaviour, which is not available at design-time for an autonomous system-of-systems. It is therefore more reasonable to use the idea of safety contracts between the different subsystems. Safety contracts are an effective way to conditionally describe the safety guarantees that a component should fulfil in order to make sure that the overall system-of-systems remains safe. Up until now, the use of safety contracts has mainly been limited to static, non-evolving systems. ESR3 will extend this approach to dynamic, modular safety contracts.

ESRs 4 and 5 will work on effective techniques and measures that assure by-design that even under fault conditions the autonomous system remains safe without any human intervention. When autonomy increases, so does the software complexity and thus the likelihood that it contains faults. Therefore, ESR4 focuses on *software design guidelines and testing specifications for non-functional requirements in safety-critical autonomous systems*. Future applications of autonomous systems will rely heavily on different communication technologies to connect and interact with other devices, infrastructure, the "cloud", etc. Although adding connectivity has its benefits, it also adds challenges, among which are most definitely its robustness and resilience. ESR5 focuses on more hardware-oriented design and testing specifications, which make *connectivity work reliably under a diverse range of environments*. This takes into account a combination of stresses, including electromagnetic interference, temperature and vibrations, aging, etc.

WP2: Designing inherently safe autonomous systems



WP2 targets novel methodologies that allow us to evaluate, validate and verify the safety-aware design (WP1), meaning that safety can be guaranteed given the complex environment and extremely varied use-case scenarios that autonomous systems will be subjected to. This challenge cannot be underestimated. As Michael Bolle, President of Bosch, Corporate Research once said in a speech: "*We have looked at what it takes to physically validate autonomous driving, and the time needed was estimated at 100,000 years. We need breakthrough solutions from the research community.*" As physical testing is too costly and too time consuming, we must turn to virtual, i.e., simulation- and model-based, testing.

ESRs 6 and 7 will collaborate to achieve a breakthrough with respect to the overall coverage of the model-based safety analysis. ESR6 will address the issue of the *virtual-worlds generation* and will apply this to autonomous robots. In other words, ESR6 answers the question "which operational situations and environments should be tested in the virtual world?" and starts from a criticality analysis. Once the most critical virtual worlds have been generated, ESR7 will *evaluate and maximize the situation coverage of each of the virtual worlds*. Exploiting techniques like Environmental Survey Hazard Analysis, ESR7 will answer the question about which faults and failure modes should be seeded into the model-based safety analysis to cover the most threatening challenges.

Whereas a classic model-based safety analysis often limits itself to failures of one or multiple components, the open-context nature of autonomous systems forces us to also consider the safety-applications of functional insufficiencies. A typical example being a camera in a self-driving car that should prevent a collision with a human being, but only detects a human correctly in 99.9% of the cases. Therefore, ESR8 is going to look at *model-based system-analysis techniques to determine propagation paths of functional insufficiencies in software-intensive systems* and will use probabilistic ways to model the uncertainties.

Complementary to ESR5 (WP1), ESR9 will also take up the challenge of the strong reliance of autonomous systems on wireless communication and will perform a *model-based system analysis of the robustness of autonomous systems against electromagnetic interference*. Combining efficient statistical electromagnetic modelling with behavioural modelling, the resulting behaviour of an autonomous system upon electromagnetic disturbances will be forecasted and evaluated.

WP3: Providing assurance strategies



The S&T WPs conclude with WP3, which pilots novel safety-assurance strategies, combining the previous 2 research WPs, thereby allowing us to put trust in the safe behaviour of autonomous systems. In total this WP involves 6 ESRs and, besides safety, also covers other design constraints such as security, reliability, availability and liability.

ESRs 10 and 11 both focus on dedicated assurance cases for autonomous systems. Existing standards, processes and practices place a great emphasis on how safety can be certified throughout the design and development stages. However, there is little guidance on how

safety assurance should be maintained throughout the system's operational life. Many assumptions about the environment and the system performance and use, particularly for complex and novel autonomous systems, that are made during the design and development stages might turn out to be incorrect during operation. From a safety point of view, this can threaten the validity of the safety case and weaken confidence in the actual safety of the system. Within SAS, two complementary approaches will be pursued to tackle this:

On the one hand, ESR10 aims at making the transition *from static assurance cases during design-time to executable assurance cases during run-time*. Here, a safety-assurance case is structured argumentation, supported by evidence (WP2), intended to justify that the system is designed (WP1) such that its behaviour is acceptably safe when being put into service. While for non-autonomous systems, the whole safety case is traditionally developed, documented and accepted prior to operation, the safety case for some autonomous systems may instead need to be posed with residual obligations that are only satisfied during run-time. For example, the vast number of possible inter-vehicle and infrastructure configurations that an autonomous vehicle may encounter may require run-time verification of safety properties (such as end-to-end response times, or the integrity of received data) to sustain the safety case. Therefore, ESR10 will establish a new way of working with an executable set of claims that will be sustained and maintained during run-time.

On the other hand, ESR11 will study *assurance-case structures for machine learning in the decision making of highly autonomous systems*. Currently, the use of machine learning or any other artificial intelligence technique, is not recommended for safety-critical tasks. However, many autonomous systems will rely on machine learning and ways to address this are urgently needed.

ESRs 12, 13, and 14 all start from a specific application scenario, i.e., autonomous vessels, clinical conversational bots and autonomous vehicles, respectively. In addition, they always consider other, possibly conflicting, design constraints, as is the case in industrial practice. ESR12 wants *to assure safe autonomous sailing from A to B while minimizing operational costs* by combining a cost-optimization algorithm, a collision-avoidance algorithm and situational awareness. ESR13 looks at the *safety assurance for clinical conversational bots* by combining safety engineering with typical clinical processes. ESR14 will cover the whole *dependability assurance for autonomous vehicles*, covering, besides safety, also reliability, availability and cyber-security.

Last, but certainly not least, ESR15 will take up the emerging challenge of the liability aspects of autonomous systems in safety-critical domains. ESR15 will propose a *liability allocation framework for safe autonomous systems* that explores new avenues for the allocation of liability for autonomous systems in a way that strikes a balance between the commercial interests of operators and manufacturers and the safety and fair compensation of the general public.

Summary

Autonomous systems offer humankind tremendous opportunities, like freeing us from mundane tasks, carrying out risky procedures and generally giving us more time to enjoy the things we like doing. However, we lack trust in many forms of autonomous systems: partly this is human nature, but primarily because these systems, such as self-driving cars, have not demonstrated their safety credentials yet. Only by making these systems safer can we expect their widespread acceptance. The Safer Autonomous Systems (SAS) ETN is about getting people to trust these systems by making the systems safer. In order to achieve this objective and to train a group of highly skilled, responsible, future innovators, we will bring together 15 early-stage researchers (ESRs) to investigate new forms of system-safety engineering, dependability engineering, fault-tolerant and failsafe hardware/software design, model-based safety analysis, safety-assurance case development, cyber-security, as well as legal and ethical aspects. Hence, the most important output of SAS will be 15 well qualified people who have been trained to tackle many of the problems now being faced by European industry.



References

[1] European Training Network for Safer Autonomous Systems, http://etn-sas.eu, accessed 31 Oct 2020

Davy Pissoort, Professor KU Leuven Bruges Campus & SAS Project Coordinator

Davy received his PhD degree in electrical engineering from Ghent University, Belgium in 2005 and worked as an R&D Engineer at Agilent Technologies (now Keysight Technologies) in Ghent. He is currently professor and head of the Mechatronics Group at KU Leuven Bruges Campus.

The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No 812.788 (MSCA-ETN SAS).

This article reflects only the authors' view, exempting the European Union from any liability. Project website: http://etn-sas.eu/

The Safety Futures Initiative



Zoe Garstang introduces an exciting new development – the SCSC Safety Futures Initiative (SFI) - to help encourage and nurture the next generation of skilled safety engineers. She introduces the SFI, describes its aims and activities and explains how to become a member.

Safety Demographics

As the safety community, and engineers in general, approach a retirement 'cliff' [1] it is now more important than ever to encourage and nurture the next generation of skilled safety engineers.

With approximately 19.5% of UK engineers due to retire by 2026 [2], there is an urgent need to diversify the demographic of safety engineers to ensure knowledge and skills are maintained.

There are a number of challenges contributing to the growing gap between the number of safety professionals required and those available. These include: an outdated view of the safety profession, a lack of understanding of what it involves, a lack of awareness of routes into the industry and an unclear career pathway once the individual has embarked on their journey.

Introducing the Initiative

The SCSC Safety Futures Initiative (SFI) seeks to address these by bringing together Young and Early-Career (YEC) individuals working, or interested, in safety. The Initiative provides a supportive environment within which to network, develop skills and share knowledge.

This Initiative is designed to encourage more people to consider a career in safety by showcasing what it has to offer. In doing this, the SFI can challenge the outdated perceptions about safety engineering and the routes into industry.

"The SCSC Safety Futures Initiative (SFI) seeks to address this challenge by bringing together Young and Early-Career (YEC) individuals working, or interested, in safety"

SFI Aims

The SFI has four main aims:

1. To establish a community of like-minded individuals to share ideas and experiences as they progress through their studies/careers.
2. To create a forum in which YECs can share technical knowledge and develop both personally and professionally.
3. To increase awareness and visibility of the safety-related career opportunities and pathways.
4. To influence SCSC activities in order to meet the needs of SFI Members.

For those considering a career in safety, the SFI increases visibility of different career paths (e.g. apprenticeships, study options, government initiatives, etc.) to ensure they are widely accessible. For those in the early stages of their career, the SFI provides professional support and an insight into potential next steps. In the SFI, promotion and awareness of safety careers is just as important as retaining existing skilled professionals.

"In the SFI, promotion and awareness of safety careers is just as important as retaining existing skilled professionals"

The SFI integrates with a number of existing SCSC activities including involvement in the Steering Group, Working Groups and opportunities within the annual Symposium (SSS). Members of the SFI will benefit from a supported transition from the SFI to the SCSC as they progress in their career.

SFI Activities

Following an initial Launch Event in early 2021, the SFI will be organising a variety of further events and workshops throughout the year. These will be in addition to, and sometimes alongside, existing SCSC events (such as the one-day seminars, web events and SSS). The range of activities aims to provide members with the opportunity to network, develop new and enhance existing technical skills, to learn more about the industry as a whole, and explore potential career pathways within safety.

The Initiative has also started to increase the social media outreach of the SFI and wider SCSC activities. Alongside the current SCSC Twitter and LinkedIn pages, an SCSC Facebook page has also been established. In the future, we will seek to create more SFI-specific social media outlets to provide a space for members to interact.



Finally, the SFI is establishing a mentoring network utilising the expertise within the SCSC population and sharing the benefits that a diverse group can bring.

The SFI will reinforce the idea that the work undertaken within safety engineering and safety assurance really matters, and it will help YEC safety professionals to be more responsible, professionally and ethically, in their work and wider society.

Membership

Membership to the SFI is currently open to the following individuals:

- Those in full-time education, or
- Those aged 26 or under, or
- Those embarking on a safety career path later in life (less than 3 years in safety)

Membership costs £35 per year and members will receive access to all online events hosted by both the SCSC and SFI. Please see: thescsc.org/membership

If you are interested in joining the SFI, getting involved in developing the Initiative, or would like further information please get in touch.

References

[1] T. Bowers: https://www.globalrailwayreview.com/article/103410/addressing-rail-skills-shortage-covid19-apsco/, accessed Dec 2020.

[2] J. Cole: https://www.jonlee.co.uk/blog/2020/01/whats-the-reason-behind-the-shortage-of-engineers-in-uk, accessed Dec 2020.

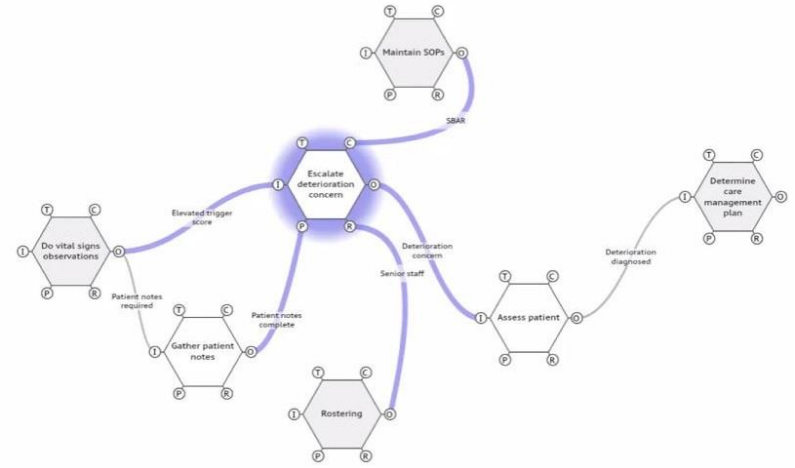
[3] E. Sharf: https://www.itproportal.com/2015/11/03/urgent-need-for-cyber-security-talent-in-the-uk/, accessed Dec 2020

Image Attributions
Top image: ID 54923434 © Rawpixelimages | Dreamstime.com
Word cloud (embed in Top image): ID 98430886 © Amir Zukanovic | Dreamstime.com
Pathways: 110448 © Jack Moreh | freerangestock.com. Licensed under Equalicense

Zoe Garstang, Airworthiness Engineer and SFI Lead

Zoe is a Flight Safety Analyst at BAE Systems, providing in-service support to the Typhoon aircraft. She previously undertook an Advanced Engineering Apprenticeship with the company before joining the Continued Airworthiness team.

New Safety Analysis Techniques Seminar Report



Richard Trebilcock reports on the SCSC Seminar on New Safety Analysis Techniques which, due to the COVID19 restrictions, was held virtually using Zoom. Richard summarises the four presentations discussing a range of new safety analysis techniques and their application.

This was an afternoon seminar hosted by Mike Parsons with approximately 20 delegates attending from a wide range of industries including Aviation, Defence, Healthcare, Rail and Space. Mike opened the seminar with an overview of the afternoon's proceedings and timetable for the Seminar presentations.

Managing the deteriorating patient – A FRAM analysis

Mark Sujan from Human Factors Everywhere Ltd., discussed the "Management of the Deterioration of Healthcare". In particular, the miss-management of the deterioration of cases, i.e. where patients are admitted into hospital, and their deterioration is not picked up by the medical staff. Retrospectively, it can be seen that there are obvious signals that patient deterioration is taking place but it has gone undetected or not acted upon.

National Early Warning Scores (NEWS) are used to indicate the requirement for early action, standardising the assessment of acute-illness severity in the NHS. These scores, based on an ascending scale in the previous 24 hours, trigger the initiation of a range of clinical responses. However, patients are still dying, and the early warning scores have not delivered as hoped for. Some hospitals are better at managing patient deterioration than others and there is a need to move from failures to interventions. This highlighted the need for a simple support tool, which is where the research started.

The approach was to turn the problem on its head, i.e. rather than look at why some hospitals fail, look at why some hospitals succeed. Also, there was a need to look at Safety II and create an environment where resilient behaviours are created.

Knowing what to look for, in particular, looking at how to improve Resilience Abilities.

The Functional Resonance Analysis Method (FRAM) prompts the analyst to consider a range of factors impacting on an Activity or Function. FRAM is an analysis method, and is not a model. The purpose of FRAM is for an analysis to make sense of an Activity or Function. The analyst builds a framework. Output of the analysis is a list of tables, which help to identify:

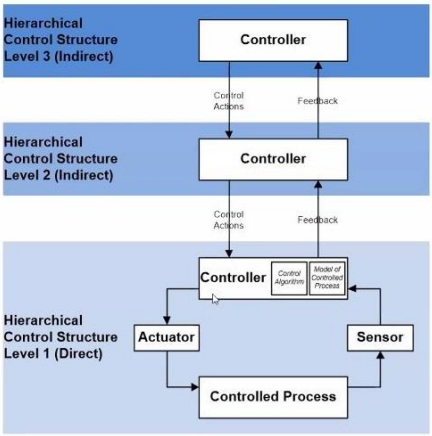
- What is the variability?
- What are the reasons for variability?
- Finally, linking variabilities to resilience abilities

Improvement in the pre-alerting for a patient should be achieved through better understanding of what is going on, on an everyday basis, and ensuring resilience is built into the systems and processes.

Overview of STAMP & STPA Hazard Analysis

Simon Whiteley from Whiteley Aerospace Safety Engineering & Management Limited, provided an overview of the Systems Theoretic Accident Model Processes (STAMP) & System-Theoretic Process Analysis (STPA) hazard analysis.

Simon started by introducing STAMP, which is an 'Accident Causality Model' based on Systems Thinking, Systems Theory and Control Theory Concepts.



STAMP-based processes comprise four model-based analysis processes:

- Systems Theoretic Early Concept Analysis (STECA);
- Systems Theoretic Process Analysis (STPA);
- Systems Theoretic Process Analysis-Security (STPA-Sec);
- Causal Analysis based on Systems Theory (CAST).

Simon spent time explaining where STAMP-based processes fit in the system lifecycle. He suggested that thinking in terms of control structures helps to organise analyst thinking and understanding. In addition, it provides a natural path to follow and a structure to work to, this is a key reason why STAMP-based assessments are so powerful. Points raised when discussing STAMP:

- Control Loops and control structures, the distance from control to actions.
- It was pointed out that we often have a constrained definition of hazard, which is reflected by our mental model of a process, and subsequently the hazard. We do not always take account of control and process lag times.

STPA is a relatively new hazard analysis technique, which builds on STAMP, a process and accident model using concepts of system and control theory. Moreover, it will give the safety analysts different views on the STPA hazard analysis process. The Ultimate Goal for STPA is to:

- Establish:
  - what the Control Structure is intended/required to do
  - how and why it allows, or contributes to, inadequate control, not just failures or "what went wrong" but "what went right"
- Create and Develop:
  - the changes necessary to enable adequate control
  - Safety Requirements / Constraints

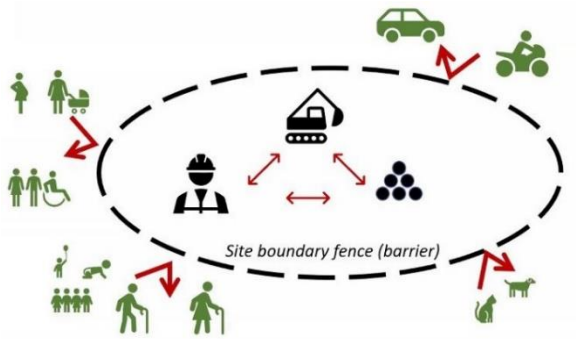
A process overview of STPA is as follows:

- **Foundation Phase**: Considering Losses, Hazard(s) and High Level Safety Constraints (HLSC)
- **Modelling Phase**: to Model Control Structures
- **Analysis Phase**: to identify Unsafe Control Actions, Causal Scenarios & define Safety Constraints
- **Development & Implementation**

When identifying causal scenarios it is necessary to identify the right level of abstraction for the control model. This is an acquired development skill, and for the novice there is the need for a coach.

Environmental Survey Hazard Analysis (ESHA)

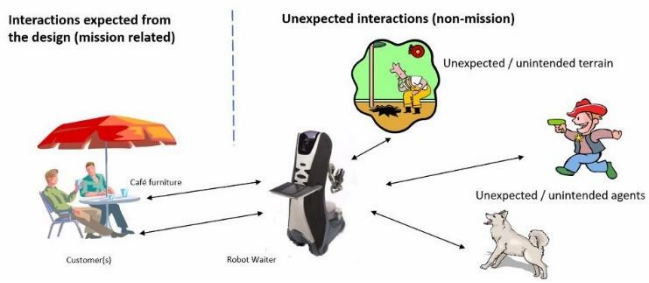
Chris Harper from Bristol Robotics Laboratory, University of the West of England, discussed robotics and how robotic Autonomous Systems (ASs) interact with their environment. He considered the problems presented by bounded and unbounded domains. The bounded domain reduces the number of interactions required by the system. For example, a barrier allows you to cut down the interactions e.g. site boundary fence blocks deflect interactions, reduces the complexity of the problem.



In the field of robotics, with Autonomous Systems moving towards interactions in a much more open domains alongside humans, where we can no longer assume the presence of barriers like steel cages, there are many more features to interact with. Many of the interactions are not eliminated by design. As such, more complex behaviour is required, with more advanced system technology required to achieve intelligent behaviour (i.e. Machine Learning/Artificial Intelligence). As a consequence, it is very much harder to show that a safety analysis is complete.

An Autonomous System (AS) needs to be able to perform Mission and Non-mission tasks/interactions:

- AS must survive in general (without human intervention) as well as perform their mission
- Hence, AS must perform "non-mission" tasks and interactions, as well as mission tasks
- Non-mission tasks are often not covered well by traditional system engineering methods.



In the context of ASs, ESHA Guide Words help designers to classify features of the environment. However, ESHA and Safety Arguments presented problems:

- Conventional systems offer service provision while ASs have to accommodate situation behaviour
- Typical arguments is "acceptable residual risk" which is difficult to define in ASs
- The problem of unbounded domains amplifies existing Safety-Critical Systems methodological problems
  - Expansion of states arising from an increase of AS scope to interactive processes causes 'combination explosion'
  - Coverage metrics become (even more) questionable
  - What is residual probability of failure if the domain is unbounded / infinite?

There is a need to look for alternative types/styles of argument; a So Far as is Reasonably Practicable (SFAIRP) argument may be more feasible.

There was a requirement for formal classification schemes for environmental features, i.e. an Ontology. These tend to be organised into 2-3 layers:

- **Foundational Ontologies**: 'Universal' Abstract Types e.g. continuant, occurrent, physical, abstract
- **Domain Ontologies**: Extended abstract type and higher-level types for domains e.g. system safety, autonomy, environmental domains (e.g. domestic, roadside)
- **Application Ontologies**: Terminology related to a particular application domain e.g. driverless vehicles, healthcare assistive robotics

The Unified Foundational Ontology (UFO) has been selected as the base formalism. The UFO has been modified to support ESHA. While there is still a large programme of work ahead, it was proposed that ESHA may, in the future, be able to support a Safety Case or safety argument.

Dialectic Argumentation

Yvonne Oakshott from Leonardo, presented on behalf of the Assurance Case Working Group (ACWG) and discussed the current interim status of the Dialectic Argument activities. Yvonne covered the following topics:

What is a Dialectic Argument?

Dialectic Argument provides a framework for creating, challenging and questioning assurance cases. A Dialectic Argument is more than just the consideration of counter-evidence. A 'Dialectic Model' is based on a set of moves or responses to each part of the argument.

Who would use a Dialectic Argument?

- Assurance Case Practitioners /Authors
- Reviewers e.g. ISA, Customers, Operators, Regulators
- Various domains e.g. air, sea, land, rail, medical

Why use a Dialectic Argument?

- Assurance arguments, for example Safety Cases:
  - Are by their nature subjective and ultimately require stakeholder acceptance of the subjective position
  - Common Criticisms:
    - May be defeated by facts that destroy the argument
    - Logical fallacies – deceptive appearance of being a good argument
    - Insufficient premise in support of the conclusion
    - Even if the evidence appears to well support the argument, the evidence may not be trustworthy
- Over emphasis on the high level claims can lead to bias and insufficient focus on the wider picture
- Explicit Dialectic argumentation can redress the balance but is little practised

Barriers to use

A number of issues have prevented the adoption of a wider approach to challenging arguments such as Awareness: many practitioners are unaware of dialectic argument and Acceptance: dialectic argument is not routinely used.

The Principles of Dialectic Argument

Challenges can be applied at any level with the assurance case and may defeat parts of the assurance case:

- **Rebuttal**: Evidencing an argument which results in a counter to the conclusions of the assurance case (i.e. refutes, disproves, shows to be false)
- **Undercutting**: The introduction of additional facts (evidence) which challenge the reasoning within the assurance case
- **Undermining**: Defined in some existing work specially raising doubts about the validity of evidence.

Wigmore charts for dialectic arguments are used in law, so there is precedence for their use.

Further developments from SCSC ACWG

Unfortunately, Goal Structuring Notation (GSN) does not provide an accepted way to represent a dialectic discourse. However dialectic arguments can be represented in Structured Assurance Case Meta-model (SACM). The intention is to include dialectic support in version 3 of the GSN community standard. The Objective of the ACWG is to raise the profile of dialectic argumentation.

Management and Oversight of Complex Systems Seminar Report



The SCSC held this Seminar on Thursday 3rd December 2020. Unfortunately, given the ongoing COVID-19 pandemic, this event (like many others this year), had to be held over ZOOM rather than the physical face-to-face gatherings to which we were all previously accustomed. But despite the lack of cake and coffee, the presentation was very well attended online, with around 45 attendees drawn from a variety of industry sectors and of course academia.

Background

This seminar was aimed at those who have some involvement in the management, approval, regulation or operation of complex systems that have a safety aspect, for example Air Traffic Control systems, Nuclear Plant, Military Vessels or Power Transmission systems.

Complex systems have behaviours that are emergent properties of the interactions between the parts of the system, which are not apparent from knowledge of those parts and their interactions alone. By this definition, many modern safety-critical, socio-technical systems, for example healthcare and transport, are also complex. As well as emergent properties, complex systems have other defining characteristics such as dynamic and ill-defined boundaries (the scope of the system can change) and tipping points (the behaviour of a system can change suddenly). Complexity therefore provides challenges for safety engineering and management – for example, safety engineering normally starts by defining the system boundary.

The Seminar

SCSC Lead – Mike Parsons opened the seminar and presented the outline agenda. The Seminar itself comprised three short presentations with a follow-on wrap up session after each to address questions arising from the attendees. Mark Nicholson then presented a final wash-up session at the end of the afternoon.

The rationale for holding the seminar is that it is obvious that so many systems are becoming more complicated. Everything is much more connected and there is less isolation. Legacy systems that were originally designed to be standalone are now being interconnected.

There is often lots of software integrated into these systems that is continually being updated. Therefore, interaction between components is often difficult to predict with any certainty.

The situation is likely to become even more complicated by the development of autonomous vehicles whose system boundary may be difficult to clearly define – we cannot predict with certainty which roads will be taken down or what obstacles/incidents such vehicles will encounter.

The recent COVID-19 outbreak and the very complex logistical task of a vaccination roll-out programme probably demands application of systems thinking.

A Framework for Governance and Management of the Safety of Complex Systems

John McDermid, University of York presented a three-layer framework which has been developed as part of a study funded under the Lloyds Register Foundation/Royal Academy of Engineering Safer Complex Systems Programme. The framework considers drivers of complexity, systemic failures, exacerbating factors and controls to reduce the risks arising from complexity (as opposed to the normal focus of safety engineering on "failures").

John outlined the basis of the Framework which originates from a short study funded by the Royal Academy of Engineering known as 'Engineering X'. The study itself consisted of a project team (four people) with John responsible for delivery of the programme overall.

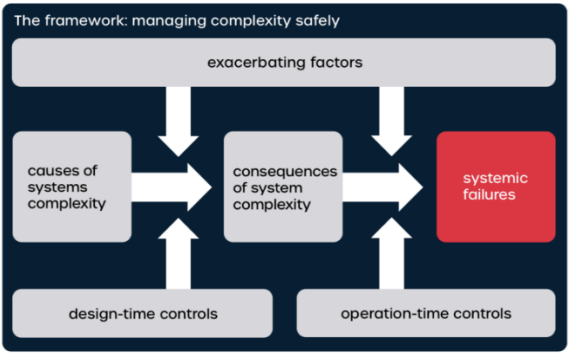
The project aim was to develop conceptual clarity around what is meant by 'Safer Complex Systems' by producing a framework that supported a common way to communicate about the safety of complex systems across sectors and between different levels of expertise.

John began by identifying the nine 'characteristics of complex systems' including Unity, Emergence, Semi-Permeable Boundary and Non-Linearity.

The essence of the Framework subsequently presented was the concept of a three-layer model comprising Governance, Management and Task & Technical.

- **Governance** being the cross-jurisdictional incentives and requirements for organisations to adhere to best practice through direction regulation
- **Management** being risk management and informed design trade-offs (supply chain dynamics and the sustainment of long-term institutional knowledge for long-lived and evolved systems)

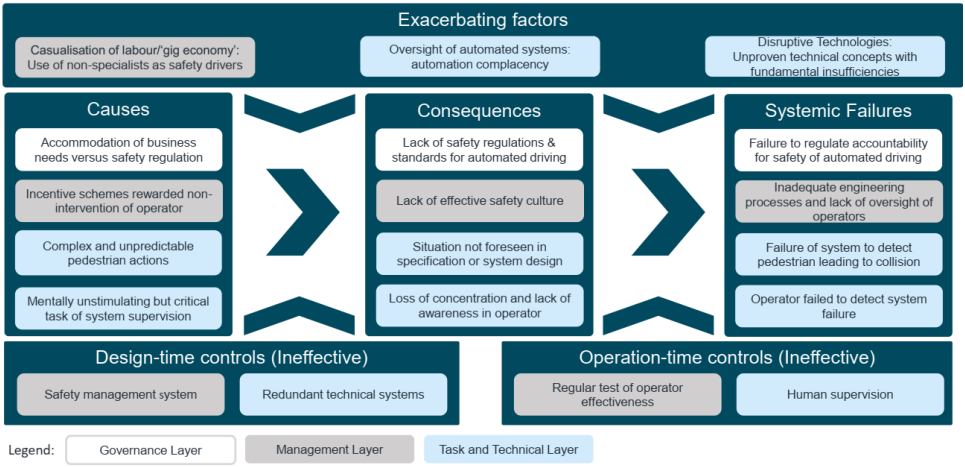
- **Task & Technical** being the technological components and tasks by the users, operators and stakeholders within a socio-technical context.



Given Safety comprises Causes, Consequences and Controls that culminate to produce 'systemic failures', John presented a series of slides which in-turn 'overlayed' this Framework on to each of these components and interactions to identify topics of concern.

Using the framework it is possible to identify potential 'causes of complexity' not just consider actual failures, and the resulting consequences.

John gave several examples of applying the framework to real-life incidents such as the Boeing 737 Max and a NATS outage in the Aerospace sector, and also in the Automotive Sector with the Uber Tempe incident analysis as follows:



John wrapped up the presentation by looking at the conclusions from the study.

It is clear the Framework has many new domain applications including Urban Air Mobility, Automated Lane Keeping Systems (ALKS) and integration of such systems with Smart Motorways etc.

"technology moves far faster than the regulators can"

The Framework produced is an initial output from the Safer Complex Systems programme. Largely descriptive but can help take a holistic view, and is rich enough to support expert workshops.

Areas for future work are as follows:

- Refinement of the framework and guidance in its use
- Extending to include existing analysis methods, STAMP/SPTA etc.
- Better reflection of interdependencies (true complexity)
- Better treatment of human-system interactions
- Needs more usage to refine and mature
- Explore more successes not just the failures – which controls are the most effective?

Type 23 Frigates

James Maltby from BAE Systems focused on the different challenges posed for the management of safety on complex warships such as the Type 23 Frigate. James is based with BAE Systems Maritime Services who operate out of two Naval Bases at HMNB Portsmouth and Devonport.

These facilities are physically huge with a large number of personnel supporting approximately 70% of the Royal Navy surface ships currently in-service including the Type 23 Frigate, QEC, River Class Batch 1 OPV, Hunt Class MCNV, River Class Batch 2 OPV and Type 45 Destroyer.

For the Type 23 Frigate, there are 13 ships in this Class. Support for the vessels is delivered as an Alliance split between the Portsmouth and Devonport sites. That Alliance includes BAE Systems, Babcock International and the UK Ministry of Defence (MoD).

Each vessel is an incredibly complex piece of engineering, with many sub-systems providing many different functions with equipment sourced and integrated from multiple suppliers.

The Type 23 Frigate incorporates a mix of both 'tangible' and 'intangible' products. Tangible products being defined as physical items of equipment and kit to be installed on to the platform; and intangible products being the provision of service for installing and maintaining that equipment.

The Service Provision identifies the agreed level of safety through identification of activities that could possibly cause harm. It also identifies responsibilities for safety delegation and the assurance processes. The focus is understanding how service outputs might deviate from the required condition or quality necessary to maintain safety. The company mitigates these through a rigorously defined business processes and assurance to ensure integrity is delivered.

Classical techniques such as implementation of a Risk Register and Hazard Log are used to manage safety through the agreed documented processes, with a strong emphasis on learning from experience to improve process. This is supplemented throughout the installation process by integral design reviews and safety assessments, inspections and trials.

For the Type 23 vessel, the original design of the ship dates back to the late 1980s, consequently there are subtle variations and deviations found in the build standard across the fleet. These variations make a traditional safety case more challenging to present, but the aim is to bring all vessels toward a common standard. The integration is often further complicated with legacy equipment that may not have a contemporary safety case in existence.

Managing Operational Safety in a Large, Complex Airline

Adam Johns from KeolisAmey Docklands (and ex-Cathay Pacific Airways), presented the challenges of managing operational safety in a large, complex airline, that has adopted concepts around complex systems, systems thinking and contemporary safety science to enhance their understanding of normal operations.

Adam's background is principally from the aerospace domain (Teledyne Controls, Virgin Atlantic, CAA) but mostly recently he has become involved with Operational Health & Safety on the KeolisAmey Docklands light railway in London.

Cathay Pacific Airways is a large operation with approx. 200 aircraft, 180,000 flights per year with over 100 worldwide destinations. Cathay Pacific partition their business operation into five distinct areas: Flight Operations, Cabin Operations, Engineering, Ground Operations and Cargo, all of which are overseen by a Corporate Safety group under the Executive and Operations leadership.

Adam presented an engaging animated slide illustrating three mechanical cogs to represent the core elements of their Safety Management System: hazard identification, risk analysis & assessment and mitigation of risk by management – all underpinned by a 'just reporting culture'.

The presentation summarised the differences between Safety-I and Safety-II concepts. Safety-I tends to focus on accidents, incidents and near-misses, compared to Safety-II, which focuses on analysing everyday actions and their outcomes – to identify those events whose outcomes ended well and to understand the reasons why failure was averted.

Cathay-Pacific adopted the Safety-II culture and began to socialise new ideas by hosting a series of regular short safety presentations on bite size chunks of topics to encourage staff throughout the organisation to embrace and discuss safety openness.

"Make everyday work your unit of analysis, not just adverse outcomes"

Adam outlined the differences between 'complicated' systems and 'complex' systems. Essentially, the definition given of a complicated system being one that is technical/mechanical, static, linear, predictable, well-defined boundary etc – a jet engine was given as the example. Complex systems on the other hand include the human interaction element, so by definition are socio-technical, dynamic, non-linear, unpredictable, infinite etc – jet engine maintenance was given as the example.

Safety becomes an emergent property of the complex system of engine maintenance; but it is not the only interaction of concern since a given 'worker' interfaces with many different elements within the organisation. This can highlight discrepancies between 'work-as-imagined', the 'work-as-prescribed' and the actual 'work-as-done' offering the opportunity to identify risks to mitigate.

As a result of the safety briefings, Cathay Pacific introduced the Operational Learning Review (OLR). OLRs take a systems approach to ensure learning and improvement are the sole objectives, not to identify blame. They are designed to help people identify where work went well and why, to highlight recommendations in working practices that can be fed back into the process.

Report by Andy Williams

Connect

The Newsletter

The newsletter is published three times annually, in February, May and October and sent to paid-up members of the Safety-Critical Systems Club.

Do you have a topic you'd like to share with the systems safety community? Perhaps an interesting area of research or project work you've been involved in, some new developments you'd like to share, or perhaps you would simply like to express your views and opinions of current issues and events. If you are interested in submitting an article, then get in touch with the Newsletter Editor to discuss ideas: paul.hampton@scsc.uk

The SCSC Website

Visit the Club's website thescsc.org for more details of the Safety-Critical Systems Club including past newsletters, details of how to get involved in working groups and joining information for the various forthcoming events.

Facebook

Follow the Safety-Critical Systems Club on its very own Facebook page.

www.facebook.com/SafetyClubUK

Twitter

Follow the Safety-Critical Systems Club's Twitter feed for brief updates on the club and events: @SafetyClubUK

LinkedIn

You can find the club on LinkedIn. Search for the Safety-Critical Systems Club or use the following link:

www.linkedin.com/groups/3752227

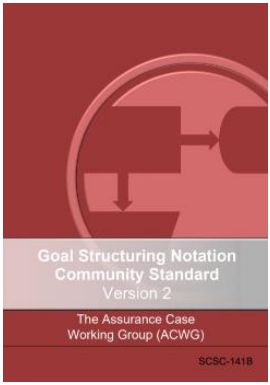
Advertising

Do you have a product, service or event you would like to advertise in the Newsletter? The SCSC Newsletter can reach out to over 1,000 members involved in Systems Safety and so is the perfect medium for engaging with the community. For prices and further details, please get in touch with the Newsletter Editor.

SCSC Working Groups

The Safety-Critical Systems Club is committed to supporting the activities of working groups for areas of special interest to club members. The purpose of these groups is to share industry best practice, establish suitable work and research programmes, develop industry guidance documents and influence the development of standards.

Assurance Cases



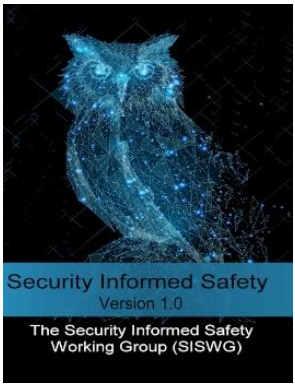
The Assurance Cases Working Group (ACWG) has been established to provide guidance on all aspects of assurance cases including construction, review and maintenance. The ACWG will:

- Be broader than safety, and will address interaction and conflict between related topics
- Address aspects such as proportionality, rationale behind the guidance, focus on risk, confidence and conformance
- Consider the role of the counter-argument and evidence and the treatment of potential bias in arguments

One of the working group's initial activities is to take on board the maintenance of the Goal Structuring Notation (GSN) Community standard.

Lead Phil Williams phil.williams@scsc.uk

Security Informed Safety



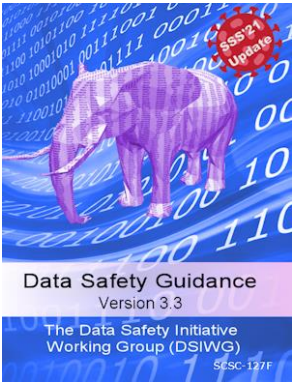
The Security Informed Safety Working Group (SISWG) aims to capture cross-domain best practice to help engineers find the 'wood through the trees' with all the different security standards, their implication and integration with safety design principles to aid the design and protection of secure safety-critical systems and systems with a safety implication.

The working group aims to produce clear and current guidance on methods to design and protect safety-related and safety-critical systems in a way that reflects prevailing and emerging best practice. The guidance will allow safety, security and other stakeholders to navigate the different security standards, understand their applicability and their integration with safety principles, and ultimately aid the design and protection of secure safety-related and safety-critical systems.

Lead Stephen Bull stephen.bull@scsc.uk

SCSC Working Groups

Data Safety Initiative



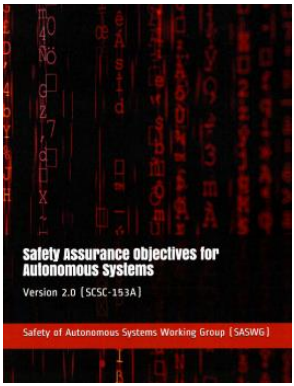
Data in safety related systems is not currently sufficiently addressed in current safety management practices and standards.

It is acknowledged that data has been a contributing factor in several incidents and accidents to date. There are clear business and societal benefits, in terms of reduced harm, reduced commercial liabilities and improved business efficiencies, in investigating and addressing outstanding challenges related to safety of data.

The Data Safety Initiative Working Group (DSIWG) aims to have clear guidance on how data (as distinct from the software and hardware) should be managed in a safety related context, which will reflect emerging best practice.

Lead Mike Parsons mike.parsons@scsc.uk

Safety of Autonomous Systems



The specific safety challenges of autonomous systems and the technologies that enable autonomy are not adequately addressed by current safety management practices and standards.

It is clear that autonomous systems can introduce many new paths to accidents, and that autonomous system technologies may not be practical to analyse adequately using accepted current practice. Whilst there are differences in detail, and standards, between domains many of the underlying challenges appear similar and it is likely that common approaches to core problems will prove possible.

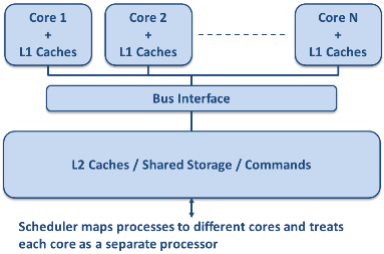
The Safety of Autonomous Systems Working Group (SASWG) aims to produce clear guidance on how autonomous systems and autonomy technologies should be managed in a safety related context, in a way that reflects emerging best practice.

Lead Rob Alexander rob.alexander@scsc.uk

SCSC Working Groups

Multi- and Manycore Safety

It is becoming harder and harder to source single-core devices and there is a growing need for increased processing capability with a smaller physical footprint in all applications. Devices with multiple cores can perform many processes at once, meaning it is difficult to establish (with sufficient evidence) whether or not these processes can be relied upon for safety-related purposes.

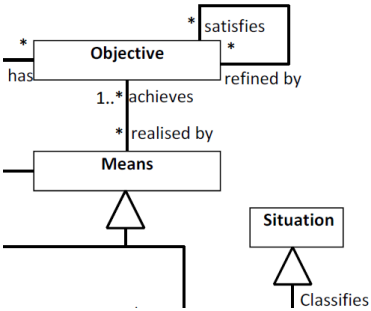


Parallel processes need to access the same shared resources, including memory, cache and external interfaces, so they may contend for the same resources. Resource contention is a source of interference which can prevent or disrupt completion of the processes, meaning it is difficult to know with a defined uncertainty the maximum time each process will take to complete (Worst Case Execution Time, WCET) or whether the data stored in shared memory has been altered by other processes.

The Multi- and Manycore Safety Working Group (MCWG) has been established to explore the future ways of assuring the safety of multi- and manycore implementations.

Lead **Louise Harney** louise.harney@scsc.uk

Ontology



The Ontology Working Group (OWG) develops ontologies that will form the basis of SCSC guidance, as well as having wider industrial and academic applications.

The OWG is currently working on the definition of an ontology of risk for application in guidance for risk-based decision making – notably safety and security – and for which ISO 31000 Risk Management principles are to be applied.

The Data Safety Working Group (DSIWG) developed the core aspects of the Risk Ontology, which has been migrated to this working group. The Risk Ontology will form the upper ontology to the Data Safety Ontology that the DSIWG will continue to develop.

Lead **Dave Banham** ontology@scsc.uk

SCSC Working Groups

Covid-19



The Covid-19 Working Group is involved with discussion, analysis and assistance related to the Coronavirus. The group meets remotely to see what a systems and assurance view of the situation brings. The group has compiled an extensive range of Covid-19 related material and made this available on the working group's website pages along with ongoing developments in the thoughts and ideas of the group.

Members are all experienced engineers, used to making reasoned arguments about safety. The aim is to apply the groups considerable technical expertise to the problem and find and assure appropriate solutions.

Lead **Peter Ladkin** ladkin@causalis.com

Service Assurance

Risks presented by safety-related services are rarely explicitly recognised or addressed in current safety management practices, guidelines and standards. It is likely that service (as distinct from system) failures have led to safety incidents and accidents, but this has not always been recognised. The Service Assurance Working Group (SAWG) has been set up to produce clear and practical guidance on how services should be managed in a safety related context, to reflect emerging best practice.

Lead **Mike Parsons** mike.parsons@scsc.uk



SCSC Safety Culture Working Group (SCWG)

The Safety Culture Working Group (SCWG) has been established to provide guidance on creating and maintaining an effective safety culture. The group seeks to improve safety culture within teams working with safety-critical systems. This is a well-researched area with a lot of guidance already in place, so the initial activities are considering what is working well, what is not effective, and where guidance is required. Meetings provide an opportunity to discuss any particular aspects attendees are interested in taking forward, and to help set future directions for the group.

Please contact michael.wright@greenstreet.co.uk if you are interested in joining this group.

60 Seconds with… Harold Thimbleby



Prof Harold Thimbleby is See Change Fellow in Digital Health, based at Swansea University, Wales. He is advisor to many organisations such as the Royal College of Physicians and the UK Medicines & Healthcare products Regulatory Agency (MHRA). Harold is a popular speaker, and has been invited to talk in over 30 countries.

He has been an expert witness in NHS criminal cases; his work exposing problems in digital healthcare has stopped nurses going to prison. Harold won the British Computer Society's Wilkes Medal, and his book, *Press On: Principles of Interaction Programming* (MIT Press), won several international awards.

What first attracted you to working in the field of System Safety?

In 2000 one of my students had an accident and ended up in intensive care. I noticed that a syringe driver (a device that uses a motor to operate a syringe to deliver drugs automatically to patients) being used in a hospital had a wide range of bugs. I told the MHRA and they responded that nobody else had mentioned any problems. I then set out to explore why, to try and understand how what I considered should have been routine programming, safety critical programming, could be of such poor quality.

What aspect of your career are you most proud of?

My students. I think it's a great privilege to be a professor and give a lecture course to a hundred students, or to have research students who get PhDs who I see go on to have wonderful careers.

What advice would you give to yourself age 12?

At 12 I was totally engrossed in electronics. Being advised to learn to speak real languages would have been really useful. The best advice I'd give myself at 12: ask for and listen to advice.

What worries you the most about the future of System Safety?

We have millions of programmers who think they are good, and they are often very productive. But being good is not the same as being safe. So I think it's going to be very hard to have a safety qualification that isn't swamped by the 'democratic' wishes that everybody wants to be qualified. Then there's the related more formal problem that any large program contains a mix of safety critical and non-critical code: sorting out the relations between the modules and the required qualifications of individual programmers doing different development jobs in large teams will be hard.

What's your most favourite quote or motto?

I've recently come across Desmond Tutu's insight: "There comes a point where we need to stop just pulling people out of the river. We need to go upstream and find out why they're falling in." In other words, we keep seeing failures in IT, harms and deaths caused by bad software: we've got to go upstream and change regulations and further upstream, change the law. The current desperately sad Horizon case, where over 550 postmasters and mistresses were accused of fraud, but which was caused by bugs and failure to understand computer evidence, is a case in point.

"There comes a point where we need to stop just pulling people out of the river. We need to go upstream and find out why they're falling in."

If you could learn to do anything, what would it be?

All my skills are pretty solitary and some are noisy: I can program, weld, fix a washing machine, or carve a bowl. So I'd like to learn some social skills, play an instrument, sing, dance, tell stories round a camp fire.

If you could be any fictional character, who would you choose?

How about Ponder Stibbons, Head of Inadvisably Applied Magic at the Unseen University in Discworld. He invented the computer Hex, which is made of glass tubes with ants running around inside. Naturally, Hex gleefully satirises real computers and inevitably has a sticker on it: "Anthill Inside."

What's the best piece of advice you've ever been given?

Tony Hoare told me to seek advice from my friends, because even if you don't like it, it's much better than criticism from your enemies after it's too late.

Which book title best sums up your experiences with Covid-19?

My own book is called *Fix IT: How to solve the problems of digital healthcare* (OUP 2021). We need to fix IT! My book has a chapter on COVID, the insights and the new problems. I'm giving a paper at the SCSC conference in 2021 on the software safety problems the pandemic has exposed – and my thoughts on some of their solutions. But a better book title that sums up my experiences with Covid would be Laura Spinney's engrossing *Pale Rider: The Spanish Flu of 1918* and how it changed the world (Vintage, 2017). They didn't know what they were doing then (they didn't even know what a virus was), so very much like today.

More details, including offprints of hundreds of Harold's articles and videos, are available from his website www.harold.thimbleby.net and his twitter handle is @haroldthimbleby

SCSC Membership

The SCSC provides a range of services to the System Safety community including seminars, tutorials, leadership events, specialist topic working groups, the annual symposium and a comprehensive body of publications. Membership brings many valuable benefits such as free access to online events, the SCSC Newsletter and access to presentations and other resources from events.

Individual Membership

To become an individual member of the SCSC please register on the SCSC website using the icon at the top right of any page and select "Register". Complete and save your account registration and then verify your email address. Once registered and logged in click the link "why not join the SCSC..." inviting you to become a member at the top right of the page or select "Pay membership" from the icon.

Individual membership can be paid online using a credit/debit card through our secure payment partner Realex Global Payments or contact Alex King for other payment methods. For student or retired member rates please contact Alex King to get your account status changed.

Corporate Membership

Your company contact with the SCSC should arrange the membership and any renewals for your organisation. To join as a member covered by a corporate membership, register as per the instructions for an individual member and then contact Alex King to confirm your affiliation.

Renewing Membership

You should be notified by email when your membership is almost expired or shortly after it has expired. These notifications will contain a link to the online renewal page or you will be able to renew when logging onto the website through the 'click to renew' link.

Membership Fees

The following fees are applicable for new and renewing members:

- 1 year Individual Membership: £125
- 2 year Membership: 20% discount: £200
- 3 year Membership: 33% discount: £250 (3 years for the price of 2)
- 1 year Student Membership: £35
- 1 year Membership, retired member rate: £35
- For Corporate Membership discounts contact Alex King

Contact Alex King using office@scsc.uk

The SCSC Steering Group



Tom Anderson

Honorary member



Robin Bloomfield

Honorary member



Stephen Bull

stephen.bull@scsc.uk



Dewi Daniels

dewi.daniels@scsc.uk



Jane Fenn

jane.fenn@scsc.uk



Zoe Garstang

zoe.garstang@scsc.uk



Paul Hampton

paul.hampton@scsc.uk



Louise Harney

louise.harney@scsc.uk



James Inge

james.inge@scsc.uk



Brian Jepson

brian.jepson@scsc.uk



Nikita Johnson

nikita.johnson@scsc.uk



Graham Jolliffe

Honorary member



Tim Kelly

Honorary member



Alex King

alex.king@scsc.uk



Mark Nicholson

mark.nicholson@scsc.uk



Mike Parsons

mike.parsons@scsc.uk



Felix Redmill

Honorary member



Roger Rivett

roger.rivett@scsc.uk



Emma Taylor

emma.taylor@scsc.uk



Phil Williams

phil.williams@scsc.uk



Sean White

sean.white@scsc.uk

Club Positions

The current and previous (marked in italics) holders of club positions are as follows:

Managing Director

Mike Parsons 2019-

Tim Kelly 2016-2019

Tom Anderson 1991-2016

Steering Group Chair

Roger Rivett 2019-

Graham Jolliffe 2014-2019

Brian Jepson 2007-2014

Bob Malcolm 1991-2007

Programme & Events Coordinator

Mike Parsons 2014-

Chris Dale 2008-2014

Felix Redmill 1991-2008

Manager

Alex King 2019-

Newsletter Editor

Paul Hampton 2019-

Katrina Attwood 2016-2019

Felix Redmill 1991-2016

University of York Coordinator

Mark Nicholson 2019-

Website Editor

Brian Jepson 2004-

Administrator

Alex King 2016-

Joan Atkinson 1991-2016

Safety Futures Initiative Lead

Zoe Garstang & Nikita Johnson 2019-

Calendar

January

M	T	W	T	F	S	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

February

M	T	W	T	F	S	S
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28

March

M	T	W	T	F	S	S
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

April

M	T	W	T	F	S	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30		

May

M	T	W	T	F	S	S
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31						

June

M	T	W	T	F	S	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				

July

M	T	W	T	F	S	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

August

M	T	W	T	F	S	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

September

M	T	W	T	F	S	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			

October

M	T	W	T	F	S	S
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

November

M	T	W	T	F	S	S
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30					

December

M	T	W	T	F	S	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

Events Diary



9-11 February 2021
SCSC Symposium

29th Safety-Critical Systems Symposium (SSS'21)

Online

scsc.uk/e683

24 February 2021
SCSC SSS'21 Follow Up

Follow up Q&A session to the 29th Safety-Critical Systems Symposium (SSS'21)

Online

scsc.uk/e791

25 March 2021
Conference

Nuclear engineering for safety, control and security

Online

events2.theiet.org/nuclear

13 April 2021
SCSC Seminar

Safety Culture Developments

Online

scsc.uk/e793

7-10 Sept 2021
SafComp 2021
Conference

40th International Conference on Computer Safety, Reliability and Security

York, UK

safecomp2021.hosted.york.ac.uk

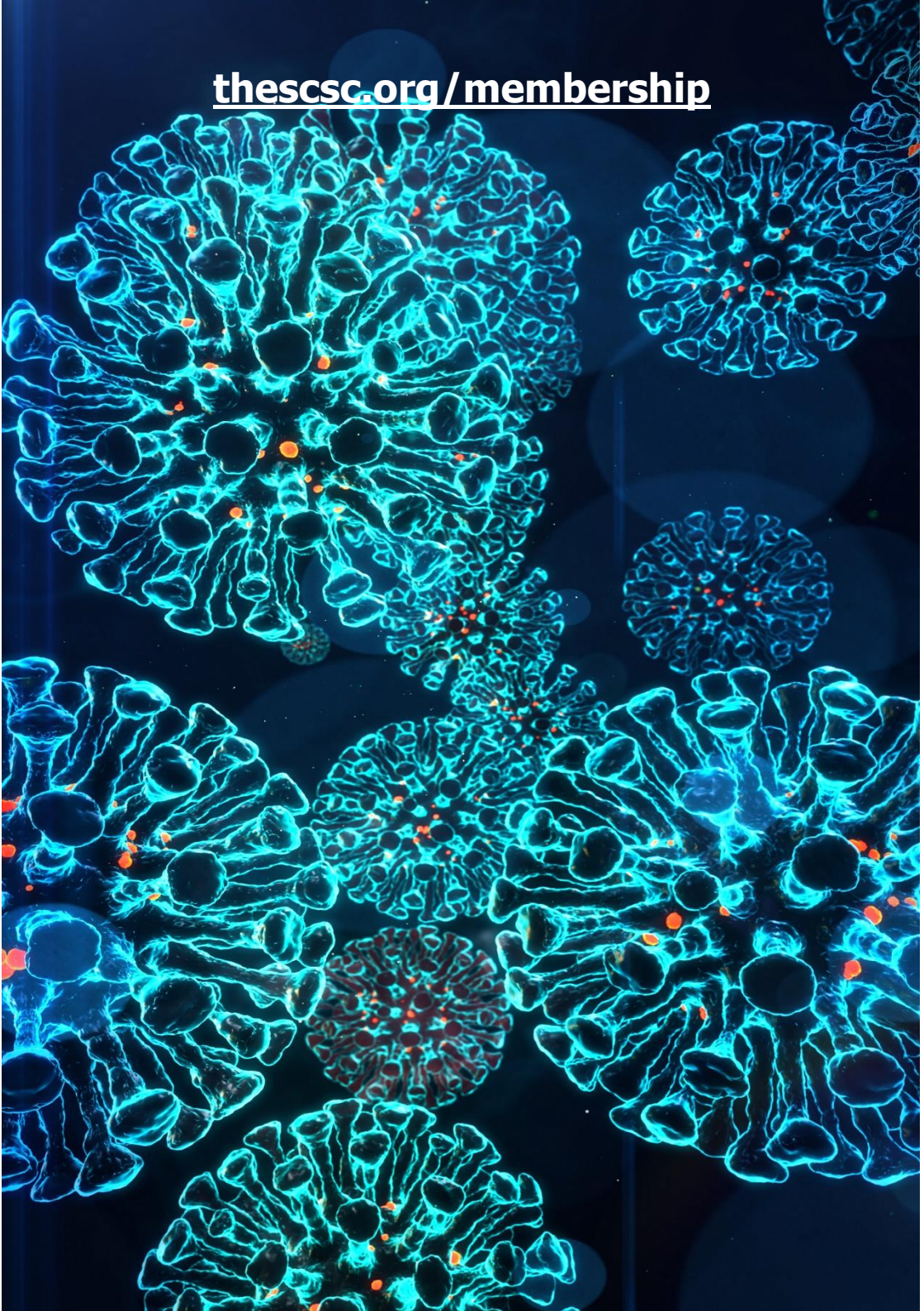
8-10 February 2022
SCSC Symposium

30th Safety-Critical Systems Symposium (SSS'22)

Online

scsc.uk/e797

NB: all events are subject to change due to the Covid-19 situation. Please check the SCSC website for up-to-date information: scsc.uk/events


thescsc.org/membership