The Safety-Critical Systems Club Newsletter

Safety Systems

Vol 28 No. 1 - February 2020

FUTURE PROOF
Assurance challenges in a changing world

RISKY BUSINESS
Harmonising the language of risk

For everyone working in Systems Safety

SCSC

scsc.uk

SCSC Publication Number: SCSC-157

Contents

THE SAFETY-CRITICAL SYSTEMS CLUB, Seminar:

Safe Use of Multi-Core and Manycore Processors

Thursday 30 April, 2020 - London, UK

This seminar is relevant to all those involved with Multi-core and Manycore processing systems, including software and hardware developers, certification bodies and sector regulators.

Safe Use of Multiple Core Systems in Critical Applications

This seminar will consider how to use processors with multiple cores in a way that safety can be assured, and such that the resulting system can be certified against industry standards and guidelines.

(A multi-core processor is typically made of several independent processor cores on the same chip, connected through an on-chip bus. Manycore processors are specialist multi-core processors designed for a high degree of parallel processing, containing a large number of simpler, independent cores (from a few tens of cores to thousands or more). Manycore processors are used extensively in embedded and high-performance computing.)

Critical systems, such as those used in avionics, are moving from single core processor to multiple core (multi-core) processor architectures. This enables a reduction in size, weight and power and the use of common processing platforms, reducing costs and allowing common spares. Software certification policies and guidance are currently evolving as experience is gained with creating certification evidence for multi-core processor architectures.

There are some unique challenges for using multi-core processors in certified platforms and these will be highlighted and discussed, including the investigation of multi-core interference channels.

This seminar will be held in central London at the Radisson Blu Edwardian Grafton, 130 Tottenham Court Rd, Bloomsbury, London W1T 5AY

Speakers include:

Iain Bate, University of York - "Multi-core architectures and timing analysis: Their influence on the scheduling of certifiable real-time systems"

TBC, Lynx Software Technologies - TBC

Guillem Bernat, Rapita Systems - "Independently Verifying the effectiveness of RTOS Hypervisors at reducing Multicore Interference"

Olivier Charrier, Wind River - TBC

Mark Hadley and Mike Standish, Dstl - "A Practical Assurance Approach for Multi-Cores (MCs) Within Safety-Critical Software Applications"

WWW.SCSC.UK

scsc.uk/e638

Editorial

As we turn the corner into a new year, and new decade, we are finding ourselves emerging into a dynamic and changing world; a world full of many novel challenges in the field of systems safety, with no clear direction on how these are going to be resolved. There are now many "disruptors" that challenge the traditional methods of safety management: Artificial Intelligence (AI), Machine Learning (ML), autonomous systems, highly data-centric systems (DCS) and remotely piloted commercial drones, to name a few.

The world seems eager to embrace these technologies, and momentum seems to be gathering as the value of these technologies is already being realised. We are already seeing successful application of AI in a wide range of sectors such as agriculture, transport and healthcare. For example, one report suggests AI is now better at predicting some cancers than trained clinicians.

The fundamental question is therefore: how will we assure systems using these disruptive technologies? How do we prove to ourselves that a future full of such systems will be sufficiently safe, especially when it may be difficult to apply traditional methods of safety management?

A number of articles in this edition of the newsletter provide insight into the challenges of assurance in this changing world.

The first article from Shakir Laher and Mark Sujan, discusses the challenges of assuring health IT systems that use Artificial Intelligence and Machine Learning technologies in delivering patient care, such as the autonomous use of an infusion pump to deliver medication to a patient.

Mike Standish and Dr Mark Hadley then discuss the challenges of developing non process-based assurance arguments using diverse evidence, to accommodate aviation regulatory bodies who are creating new paths to developing software safety assurance cases.

Alastair Faulkner then describes the assurance challenges with highly data-defined and data-driven systems in a world prevalent with ML and autonomy.

Our final article from Dave Banham, provides an introduction to an exciting area of work that is formalising the language we use to talk about risk. This could potentially allow the risks arising from system safety and security concerns to be described unambiguously using a common language, thus harmonising collaboration between the two domains.

In this edition we have SCSC event reports covering the Senior Leadership Forum and two seminars on Data Safety Evolution and Creating and Maintaining an Effective Safety Culture.

I have also introduced a new '60 second interview' feature, and Tim Kelly, the previous SCSC Director, has kindly agreed to be my first interviewee.

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

In Brief

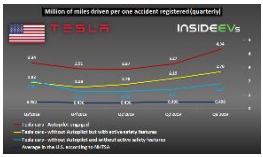
Lion Air 737 MAX Final accident report cites AOA sensor, MCAS among multitude of contributing factors

Indonesia's National Transportation Safety Committee (KNKT) published its final 322-page accident investigation report on the October 2018 Lion Air Boeing 737 MAX flight JT610 crash, finding that it was caused by a combination of an improperly aligned angle of attack (AOA) sensor, lack of pilot reporting and training as well as a breakdown in safety oversight of certification and design flaws shared between Boeing and the FAA within the aircraft's manoeuvring characteristics augmentation system (MCAS) system. *aviationtoday.com*

Clutha crash inquiry: "Pilot took risk with fatal consequences"

An inquiry into the helicopter crash that claimed the life of the pilot and nine other people, has found that the pilot ignored low fuel warnings. The Fatal Accident Inquiry report into the crash at the Clutha bar in Glasgow city centre in November 2013 was released yesterday and said that the captain took a chance by ignoring the safety signals. *greenocktelegraph.co.uk*

Tesla car safety increased in Q3 2019 to record level



According to the latest Tesla safety report, during the third quarter of 2019, Tesla cars set a new all-time record of avoiding an accident. On average, the company registered one accident for every 4.34 million miles driven when Autopilot was engaged. That's almost 10-times better than the U.S. average. Even those without using Autopilot or other safety features, Tesla was way above the industry average. *insideevs.com*

Self-driving Uber car that hit and killed woman did not recognise that pedestrians jaywalk

A self-driving Uber car that struck and killed an Arizona woman wasn't able to recognize that pedestrians jaywalk, federal safety investigators revealed in documents released earlier this week. Elaine Herzberg, 49, died after she was hit in March 2018 by a Volvo SUV, which had an operator in the driver's seat and was traveling at about 40 mph in autonomous mode at night.

The fatal accident came as a result of the automated Uber's not having "the capability to classify an object as a pedestrian unless that object was near a crosswalk." *nbcnews.com*

Assurance Challenges for Artificial Intelligence and Machine Learning in Healthcare

Photo Credit: Dr Nick Reynolds, Royal Derby Hospital

Shakir Laher and Mark Sujan discuss the challenges of assuring health IT systems that use Artificial Intelligence and Machine Learning technologies in delivering patient care.

The rate at which Artificial Intelligence (AI) and Machine Learning (ML) technologies are being developed in healthcare does not show any signs of passing by as just the latest fad. NHSX, recently set up by the government to drive NHS digital transformation and lead policy, published a report entitled "Artificial Intelligence: How to get it right". In it they surveyed 112 developers of AI products, finding 24 were very likely to deploy at scale in a year, with up to 82 in five years [1]. There will be many development challenges to overcome, which becomes the primary focus. However, consideration of potential patient harm resulting from this form of technology and the assurance requirements of the technology have never been more crucial.

"Medication errors present a significant and stubborn cause of patient harm, with as many as 237 million medication errors occurring in the NHS in England every year"

This article focuses on the potential assurance challenges faced by a specific instance of AI/ML technology in the form of delivering insulin autonomously to Intensive Care Unit (ICU) patients. Medication errors present a significant and persistent cause of patient harm, with as many as 237 million medication errors occurring in the NHS in England every year, causing around 700 deaths and contributing to around 1700 patient deaths [2]. Infusion errors are particularly error-prone and high-risk due to their complexity and the acuity of the patient. The findings have been extrapolated from the Safety Assurance of Autonomous Intravenous Medication Management Systems (SAM) project [3], which is part of the wider Assuring Autonomy International Program (AAIP). The project was tasked with identifying stakeholders' safety assurance requirements in the context of an autonomous intravenous medication management systems. In addition, NHS Digital's contribution was to conduct a case study at Royal Derby Hospital (RDH) to develop a preliminary hazard analysis and explore possible safety assurance strategies for different levels of an automated and autonomous intravenous medication management system. It is with this focus the article presents its findings of assurance challenges faced by autonomous healthcare systems.

The Regulatory Space

Technology that is driven by software and used in the health domain is generally classified into 1 of 2 types: Medical Device (MD) or Health Information Technology (HIT). If the technology is defined as a MD, it will be subject to a regulatory framework established under European law. Adhering to the framework establishes an acceptable level of safety (from a legal perspective) of the product to be placed on the market. If the technology is classed as HIT, it falls out of scope of European legislation. The NHS does however provide standards established under the Health and Social Care Act that need to be followed to assure technology that is placed in or accesses NHS IT infrastructure.

NHS Digital's clinical safety team utilises two standards to assure the safety of HIT in the NHS. DCB0129 [4] is used by manufacturers of HIT to evidence the clinical safety of their products. DCB0160 [5] is in place for healthcare organisations who deploy and use the HIT.

Safety Modelling, Assurance and Reporting Toolset (SMART)

Since the introduction of the standards, the clinical safety team have observed that organisations struggle with key concepts around hazard analysis, risk assessment and development of safety cases [6]. There is a variety of reasons for these difficulties, which are covered in the study at [6] and not reproduced here. However, if there were a software tool that guided organisations through the implementation of the standards, the safety cases produced would be of much richer quality. This is where SMART takes centre stage.

SMART is a software tool which has been developed in-house by NHS Digital's clinical safety team. It guides users through entering information that forms a safety argument based on the principles of a goal structuring notation (GSN) argument [7]. This technique allows users to concentrate on the intellectual work involved in the clinical risk management process and removes the burden of generating the documentation as the tool exports the documentation in the form of a clinical safety case report (CSCR).

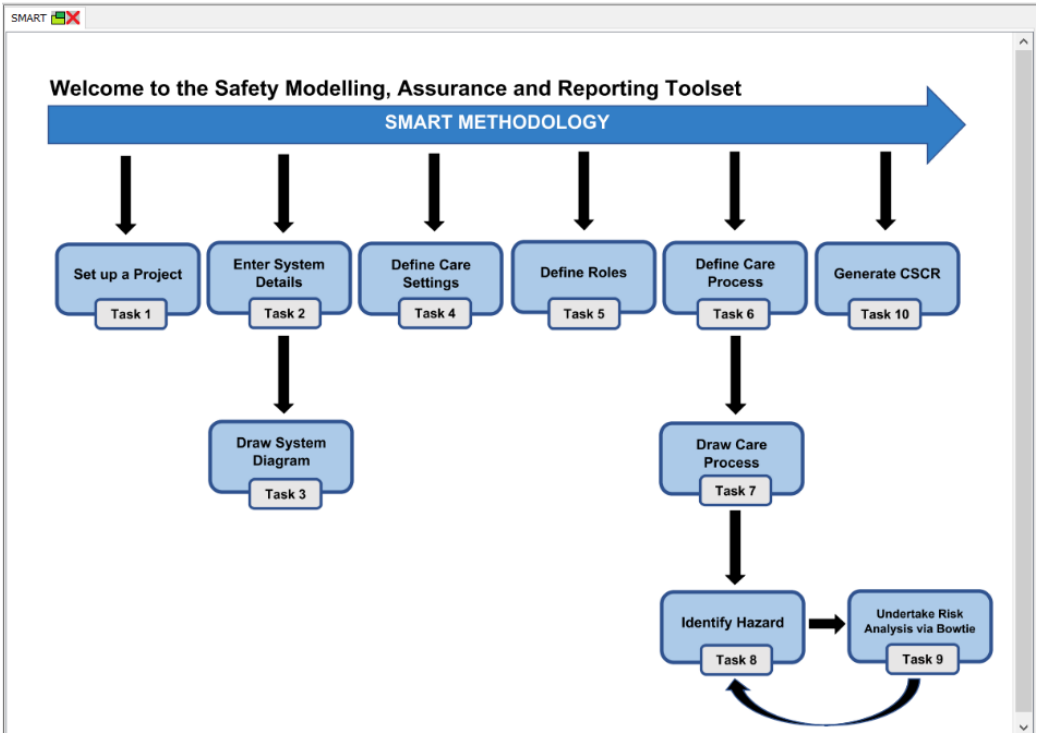
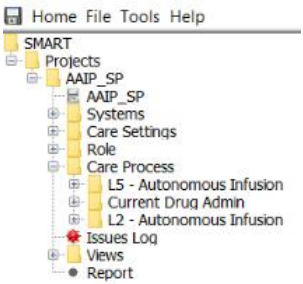


Figure 1: SMART Dashboard

Figure 1 graphically demonstrates the sequence of tasks the user needs to follow to complete the safety argument. The content can be navigated through the tree in the left-hand panel as shown in the figure opposite. The clinical risk management process for the case study was completed using SMART. It is envisaged SMART could become a standardised tool for clinical risk management.



Automated to Autonomous

To understand how the technology within the SAM project could be implemented, an approach was taken to describe the level of autonomy through levels similar to the automotive domain. Figure 2 illustrates different levels of capabilities of the technology, moving from automated to an autonomous implementation. This aided the preliminary hazard analysis and safety assurance argument as the technology took greater control of the process.

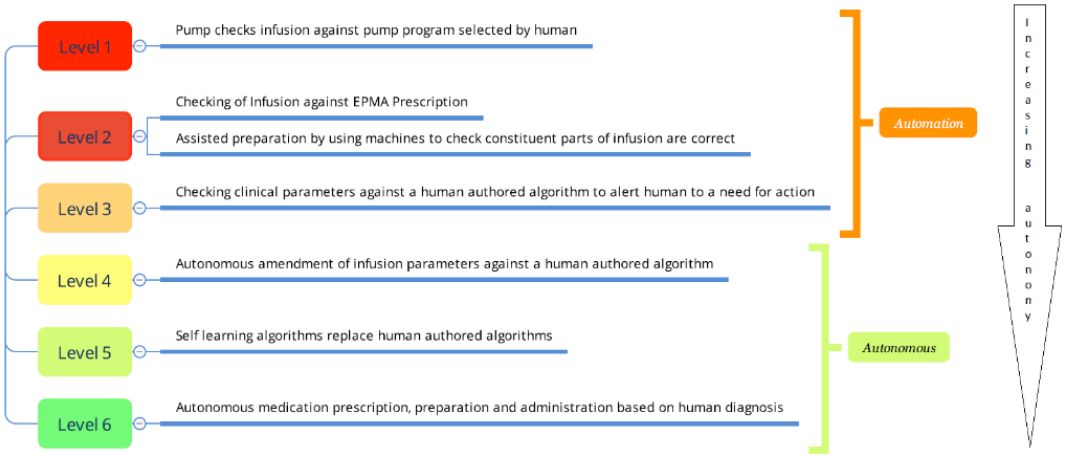


Figure 2: Automated to Autonomous

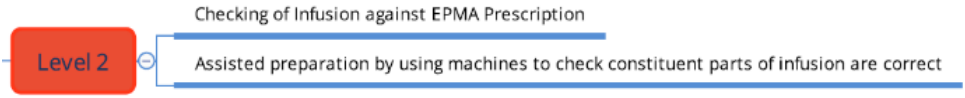
Using SMART's graphical process editor, three process maps were drawn to fully understand the current drug-administration process compared to the future processes at Level 2 (L2) and Level 5 (L5).

Care Processes: Through the Levels

The component parts of the drug administration process are:

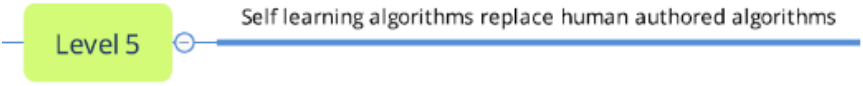
- Prescriber – Fills in the type, amount and dose of the medication
- ICU Nurse – Gathers, administers and monitors medication
- Second checker – Second checks at set intervals
- HIT – The ICU ward's computer system. Houses patient data
- Infusion pump – Delivers insulin into the patient's body

Current drug administration The current process is driven by a paper-based prescription chart that is passed around the staff. The consultant initiates the process by prescribing the medication on a prescription chart. The nurse assumes the responsibility of identifying the patient, gathering the medication, administering and monitoring infusion. The use of the pump in this scenario is limited to administering the medication with no decision-making functionality.



L2 This process moves the reliance away from the paper-based prescribing approach to a fully technological solution. The prescribing here is done using an Electronic Prescribing and Medicines Administration (EPMA) system, which is a component of the ward's HIT. The nurse is still responsible for gathering and administering the medication. However, the identification of the patient, viewing prescriptions and second checking the right syringe is placed in

the correctly programmed pump are all done by cross-checking against the HIT. The HIT prints a prescription barcode that has patient and prescription specific information. This barcode can be scanned by any device which can then communicate with the HIT to confirm any details. For example, scanning the barcode into an infusion pump will cross-reference whether it has been programmed correctly. At this level, much of the prescription second checking is now done by communicating with the HIT. This removes some of the duplication and double-checking tasks, from the nurse, which are error-prone and can lead to drug-administration errors.



L5 This level extends level 2 significantly by making use of AI & ML. The HIT is monitoring the patient's blood sugar levels with sensors. This gives it the ability to make the same decisions a nurse would make in terms of smaller adjustments to the drug administration based on established clinical guidelines. The technology could also assume greater responsibility and go beyond the initial clinical guideline by learning from the patient's physiological response to develop a personalised drug administration protocol. To maintain a level of human supervision, the HIT has certain hard limits (e.g. upper limit on drug dose and rate of administration, which are known to have fatal outcomes). If it decides to act outside of guidelines, authorisation is required from a healthcare professional. It is within these bounds that the technology is considered to be autonomous.

Hazard Analysis

Hazard analyses were conducted for each level by firstly identifying the hazards. The causes and subsequent harm events were added. Where possible, mitigations were included. This was documented using bowtie diagrams.

Assurance Challenges

A fundamental component of this work was to assess if the clinical risk management process outlined in DCB0160 would be applicable when assessing autonomous technology. The process did not present significant difficulties and appeared relevant and applicable. The analysis considered the different levels of automation. The major hazards remained the same across the levels of automation (e.g. patient harm resulting from overdose / underdose), but the hazard scenarios and contributory factors changed, with a shift away from people towards the technology. This seems a logical shift as the technology assumes more responsibility. However, it is also worth considering that novel human factors challenges are introduced, such as automation bias and handover between the technology and people [8]

"Hazard scenarios and contributory factors changed, with a shift away from people towards the technology"

The role of the human, mainly the nurse, in the baseline process is of someone who does the actual tasks of double-checking and drug administration. In L5, the technology undertakes these tasks and the role of the human becomes one

of providing supervision and expert oversight.

It is anticipated that this approach would lead to fewer drug administration errors. However, what assurances are needed to have confidence that the technology is functioning as it should be?

The technology at some point will fail or need to hand over to a clinician due to encountering a situation where it cannot continue. A protocol for handover will need to be established. Current nursing skills do not encompass handover from technology to humans, especially when there is an unplanned handover at L5. To assure safe operation of the service, the staff would need additional and new forms of training. It may be the case that a new specialised role might have to be established, such as an AI Operative Nurse who provides expert oversight and who would be present to give assurance that an appropriately trained human is in place.

"There would need to be a fundamental rethink of how the service is provided to the patient"

It is not anticipated that introducing AI/ML technology into a ward setting, while keeping all the other processes as they are, would be a wise approach. There would need to be a fundamental rethink of how the service is provided to the patient. It may be the case that the staff will need to work around the AI/ML product and fundamentally change their working practice. From an assurance perspective, AI/ML solutions will require certain staff and working practices to be in place.

Regulation will play a pivotal role in how AI/ML technology develops in this area. Based on the experiences within the SAM project, it appears that existing standards are reasonably well-suited, with some modifications, to address the technical challenges of assuring the safety of AI and autonomous systems in healthcare. However, further consideration needs to be given to the human factors aspects of using AI/ML in healthcare safely, and to the organisational and institutional context. For example, at present there is no communication between the current HIT and infusion pump.

At L5, as they begin to work in tandem, one of the two, or even both would be classed as a MD. This raises questions for the manufacturers, such as, will they want to take on the additional regulatory compliance and investment? The regulatory process under MD Regulation is far more complicated to navigate as opposed to going down the route of NHS standardisation. This will ultimately impact on the architectural design of the system, which in turn, will lead towards the types of assurance requirements that will be sought.

Conclusion

AI/ML technologies being developed in health domain show no signs of dissipating. As more of these technologies make their way on to the market, assuring them is, and will be a significant challenge. As observed from the case study, existing approaches to regulation and assurance still apply. However, deeper analysis is required in how humans and AI/ML technology collaborate to provide a service to the patient without harming them.

References

[1] Joshi, I., Morley, J.,(eds). Artificial Intelligence: How to get it right. Putting policy into practice for safe data-driven innovation in health and care. 2019. London, United Kingdom: NHSX

[2] Elliott, R A, Camacho, E, Campbell, F, Jankovic, D, St James, M M, Kaltenthaler, E, Wong, R,Sculpher, M J, and Faria, R, "Prevalence and economic burden of medication errors in the NHS in England". 2018. Sheffield: Policy Research Unit in Economic Evaluation of Health & Care Interventions.

[3] Sujan, M, Furniss, D, Embrey, D, et al. "Critical barriers to safety assurance and regulation of autonomous medical systems". 22–26 Sept 2019. In: Proceedings of the 29th European safety and reliability conference (ESREL 2019) (ed Beer, M, Zio, E), Hannover.

[4] NHSDigital, "DCB0129: Clinical Risk Management: its Application in the Manufacture of Health IT Systems – Specification NPFIT-FNT-TO-TOCLNSA-1792.06, V4.2, 02.05.2018", NHS, Jun 2018 [Online] Available: https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/dcb0129-clinical-risk-management-its-application-in-the-manufacture-of-health-it-systems [Accessed: 18.11.2019]

[5] NHSDigital, "DCB0160: Clinical Risk Management: its Application in the Deployment and Use of Health IT Systems – Specification NPFIT-FNT-TO-TOCLNSA-1793.05, V3.2, 02.05.2018 ", NHS, Jun 2018 [Online] Available https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/dcb0160-clinical-risk-management-its-application-in-the-deployment-and-use-of-health-it-systems [Accessed: 18.11.2019]

[6] Habli, I., White, S., Sujan, M., Harrison, S. and Ugarte, M., "What is the safety case for health IT? A study of assurance practices in England". 2018. Safety Science, 110, pp.324-335

[7] , SCSC Assurance Case Working Group (ACWG), "Goal Structuring Notation Community Standard Ver 2", 2018, https://scsc.uk/SCSC-141B

[8] Sujan M, Furniss D, Grundy K, et al, "Human factors challenges for the safe use of artificial intelligence in patient care". BMJ Health & Care Informatics 2019;26:e100081. doi: 10.1136/bmjhci-2019-100081

Shakir Laher, Safety Engineer, NHS Digital

Shakir Laher is a Safety Engineer currently working at NHS Digital. His role involves developing a software tool that generates safety cases, clinical risk management activities that support the safe development/deployment and use of Health Information Technology (HIT) within the NHS and writing academic research papers that explore the assurance requirements of HIT.

Mark Sujan, Senior Consultant, Human Reliability Associates

Mark Sujan, PhD, is Senior Consultant at Human Reliability Associates located in the North West of England, and Associate Professor of Patient Safety at the University of Warwick. He has more than 20 years of experience developing, applying and teaching human factors and safety engineering methods in both safety-critical industries and healthcare.

THE SAFETY-CRITICAL SYSTEMS CLUB, Seminar:

New Safety Analysis Techniques

Thursday 11 June, 2020 - London, UK

This seminar is relevant to safety engineers and safety consultants who have to perform analysis of systems. It will also be useful for safety auditors and assessors who may have to interpret or review analyses in these new methods.

Safety Analysis Techniques of the Future

This seminar will look at emerging, novel and recently established techniques for analysing aspects of safety systems: their overall properties, their architecture and interactions, their software, hardware and their data.

Safety systems require analysis for potential failures that can lead to hazards. Traditional techniques tend to have limited applicability in today's world of highly complex, interconnected, continually updated systems. Learning systems bring new analysis problems as the faults may be contained in the training data rather than the system itself.

Techniques such as STAMP/STPA will be covered as well as emerging methods for analysing hazards in context (Environmental Hazard Analysis). The uses and abuses of Bow-Ties will be covered. The Functional Resonance Analysis Method (FRAM) will be considered. New techniques for analysing autonomous and machine learning systems will be discussed. The issues of analysing systems of systems will be covered. Tools and techniques for analysing service aspects (e.g. based on Swimlane Diagrams and Business Process Model and Notation, BPMN) and data safety will also be covered (e.g. use of data FMEA).

There will be a range of speakers covering different techniques. A wrap up session at the end of the day will discuss the most promising contenders.

Speakers include:

Christina Goddard, Frazer-Nash Consultancy -"TBC"

Chris Harper, Bristol Robotics Laboratory - "Environmental Survey Hazard Analysis"

Mike Parsons, SCSC - "Data FMECAs"

WWW.SCSC.UK

scsc.uk/e654

Adapting to Changes in a Software Safety Assurance Approach

Mike Standish and Dr Mark Hadley discuss the challenges of developing non process-based assurance arguments using diverse evidence.

An aircraft (like many systems) comprises many interconnected elements, of varying complexity, which provide the fundamental functionality for the system's operation. Many of these elements will perform mission, security, or safety critical roles and the system's functionality is commonly underpinned by software. If there is a *failure* in the software then there can be a *failure* of the system to perform its function. To deploy a safety-critical system it is imperative to have *confidence* in the system's underpinning software and this is gained by performing software safety assurance. The term *confidence* is defined as providing "trust in a thing" and to "show [a level of] certainty" [1]. If there is an insufficient level of confidence in the software then there is an insufficient level of confidence in the system.

This article provides a number of thoughts on the wider technical and procedural aspects that may need to be considered if there are changes to how process-based software safety assurance is currently conducted. The thoughts relate to vendors, system integrators, equipment procurers, and regulators.

Judging Software Integrity

A traditional way to gain confidence in the software is to develop it to a pre-defined *process*. Within the civil aviation domain, a commonly applied guideline is DO-178C [2]. The guideline can also be adopted within the military aviation domain as an acceptable means of compliance to provide design assurance of airborne safety-related software (SRS).

There are also other approaches cited as being suitable to develop safety-critical software within wider domains (e.g. SafeScrum [3]). However, regulatory bodies are creating new paths to reflect revised software safety assurance stances (e.g. non-prescriptive regulatory oversight). The Federal Aviation Authority (FAA), for example, is aiming to adopt a set of three Overarching Principles (OPs) (*intent, correctness,* and *innocuity*).

The aim of the OPs is that they will *not* supplant the existing approaches (e.g. DO-178C) but to act as *another route* to gain certification approval [4]. The OPs provide an alternative means to justify the belief in the suitability of the software for inclusion into an aircraft. The revised, non-prescriptive, stances should allow (in theory) wider forms of evidence to be included within software assurance cases.

"Regulatory bodies are creating new paths to reflect revised software safety assurance stances"

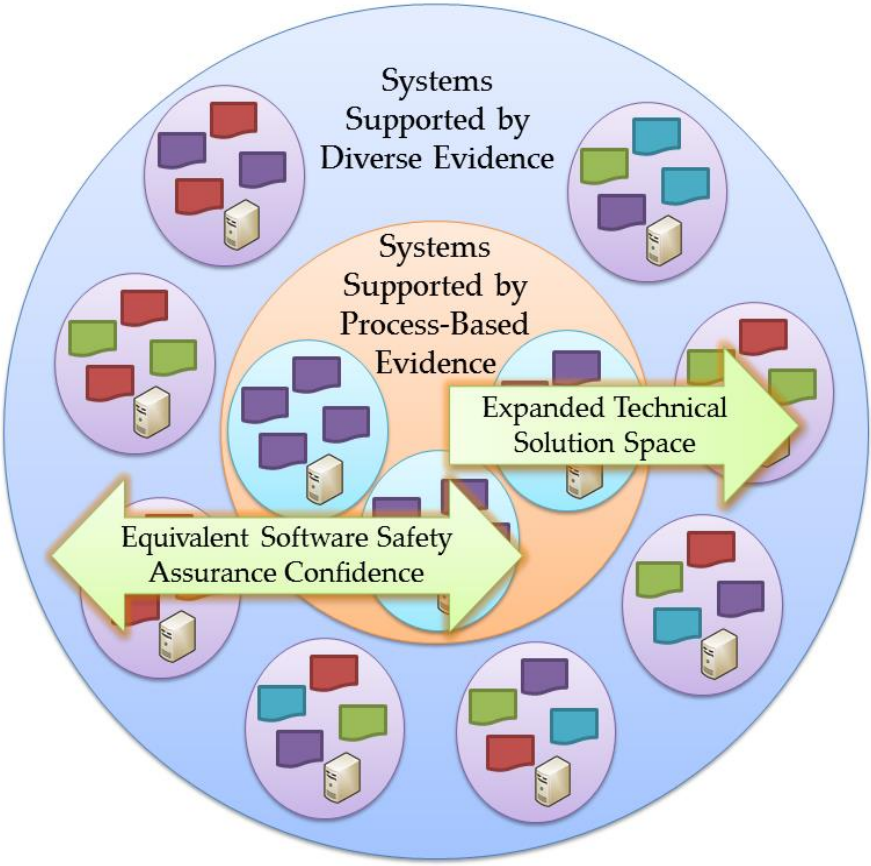
These would not necessarily be solely focussed on process-based developments and artefacts. However, we believe that adopting such approaches requires a shift in how evidence is currently generated, gathered, and judged as the traditional objective-by-objective activities (e.g. DO-178C) may not feature within a revised software assurance approach.

Use of Diverse Evidence to Expand the "Pool of Solutions"

Having the ability to adopt varying forms of evidence can support, at least, two aims:

- it can potentially assist with fulfilling the aims of the OPs; this indicates a benefit in *implementing a safety assurance argument*; and
- it can have benefits in articulating the evidence and the supporting arguments for *systems which do not conform* to the traditional processes, for example "DO-178C compliancy".

Accepting only systems which are software process compliant over other systems (which are supported by other evidence strands such as in-service data) *reduces* the pool of systems which could provide effective capability solutions, e.g. improved Size, Weight, and Power (SWaP) characteristics. It should be noted that any system considered within a wider technical solution must still have the evidence available to demonstrate that the safety constraints can be met to achieve an equivalent level of safety. This expanded solution space is illustrated simplistically in the figure below.



Managing the Change in a Software Safety Assurance Approach

To adopt a non-process software assurance approach there may need to be a fundamental shift in *how* software safety assurance is conducted within a number of safety-critical domains. The following points detail a number of considerations which may be necessary to implement (in-full or in-part) when evidence is accumulated within a revised software safety assurance approach. The considerations (stated at a high-level in the figure below) range from stakeholder skillsets to how safety-critical software is procured.

Knowledge and Skills of Stakeholders

Suitably Qualified and Experienced Personnel (SQEP) requirements and skillsets need to alter

The ability to review and judge a wider range of diverse evidence requires a revised skillset compared to a pure process-based approach. These include, but are not limited to: the procurement environment, system and platform interaction, and effective industry/customer engagement.

Principles of evidence theory need to be restated

Any revised approach requires evidence to be holistically assessed with an understanding gained into *how* evidence can be structured and supported. There is a need to (re-)educate stakeholders in the construction of diverse evidence arguments, *how* to structure evidence, the steps to improve existing data, and also the *theory of evidence*.

Through-Life Appraisal

Judgements should be captured for the life of the software/system

There are distinct phases of any system such as the development and in-service periods. These systems may also be subject to regular updates which require frequent judgements to be made on the software evidence. Standards and guidelines evolve, albeit slowly, but the attributes associated with any evidence can remain constant. An example of an evidence attribute is the *contribution* it makes to the wider evidence it supports. Therefore, a mechanism could be created to capture judgements through-life; acting as "snap-shots" to reflect the evidence characteristics. This is particularly relevant as a system and its software ceases to be developed and is maintained as legacy equipment.

Capture the confidence being built rather than only ongoing problems
In-service evidence should be a method to *build* confidence rather than to only question it. Traditionally, any in-service data for a system is used to gain feedback on problem reports to query the belief in the process evidence (indeed, counter-evidence is an important property to capture). However, to actively validate the prior-belief in any process, evidence could be used to maintain, or *improve*, confidence.

> ***"In-service evidence should be a method to build confidence rather than to only question it"***

Actively gather metrics to inform diverse evidence
There is a need to actively seek metrics and evidence *throughout the complete life-cycle* of the software. This includes during the development stage. There is a myriad of metrics that can be obtained to support or refute the confidence in a system, e.g. technical suggestions raised in Technical Interchange Meetings (TIMs), which lead to *actual* implemented changes. Opportunities to gather such data should be encouraged.

Ensure the right metrics are informing the right judgements
There is a need to be mindful of the use of metrics, how they shape the safety arguments, and how they alter the confidence being gained. Measuring the "wrong" thing can have unintended consequences [5], for example, for future decisions taken or by the introduction of inherent weaknesses in the confidence argument. There are other dangers with metrics such as *gaming* a quantification-based approach (i.e. where the choice of metrics is deliberately interpreted so as to exploit favourable outcomes) and the difficulties with Subject Matter Experts (SMEs) agreeing metric values.

Managing Evidence Complexity

Management and stakeholder understanding of an increased range and depth of underpinning evidence
Process-based compliance that is benchmarked against standards/guidelines can result in sets of structured findings against defined objectives. However, a diverse evidence approach will make use of a depth and range of evidence with no pre-defined benchmarks. Therefore, the wider evidence needs to have the structure *managed/captured* in a *consistent* manner. There is also a greater reliance on the judgements made by SMEs.

Growth in the solution space requires assistance for decision makers
A correlation may exist between the *diversity* of the evidence and the *quantity* of evidence which is under review. Stakeholders can select a myriad of evidence to support an assurance argument and therefore the potential solution space grows significantly. This larger solution space may be complicated for stakeholders to assess and to determine suitable actions.

"Evidence and its confidence should be driven from the context of the system/software and not just necessarily from a process-based approach"

Communication of evidence to stakeholders is key
Allowing stakeholders to comprehend the evidence and the confidence which can be placed in it is obviously an important element. This allows judgements to be debated and understood. Not all stakeholders who have a vested interest in the outcomes need to have visibility of *all* of the evidence judgements. An abstraction of the information would be of use to gain buy-in and to assist with stakeholder discussions. An approach such as the *Wheel of Qualification* could assist with this [6].

Treatment of Evidence

Not all objectives are equal
Within any standard/guideline, particularly those which are focussed on process conformance, there are unstated degrees of *importance* to each of the objectives. At present, a number of the standards and guidelines that are adopted for software assurance do not provide information on the weighting of the objectives and therefore give no indication of the priorities or consequences of any shortfalls. Any assurance regime should allow for the priorities or weightings of the objectives to be stated within the extant standards/guidelines. An option is to allow weightings, e.g. *contribution*, to be formally established for objectives and to allow these to be managed and reasoned with.

"At present, a number of the standards and guidelines do not provide information on the weighting of the objectives and therefore give no indication of the priorities or consequences of any shortfalls"

Evidence should be collated to be judged rather than it occurring at distributed stages
Presently, a range of evidence is captured as part of the assurance process and evidence is captured via a *staged* review process rather than making holistic judgements on all relevant evidence. The range and depth of evidence could be significantly increased in any revised assurance approach.

What does "good" look like?
With the adoption of diverse evidence there are a number of standards/guidelines that can inform a view of what an ideal scenario would look like for any given evidence. An example is Capability Maturity Model Integration (CMMI) for process improvement. The metric for the success may not be *full* compliance with a standard/guideline but to allow a measure to be gained of the shortfalls and therefore, the confidence that can be assigned to such evidence. However, diverse and radical evidence that does not have precedence will lack a supporting framework. Such evidence will be significantly reliant on SME judgement.

Need to build evidence from the bottom-up
Due to the novelty of a diverse evidence approach, there is a need for SMEs to understand the *reasoning* for the choice of evidence and the place it has within the assurance argument.

The relevance and weight of the underpinning evidence needs to be ascertained. There is an argument that, in some contexts, the assurance confidence needs to be gained from the *available* evidence and this requires a bottom-up approach. Evidence and its confidence should be driven from the context of the system/software and not just necessarily from a process-based approach.

Diverse evidence associated with a system will be unique and should be treated as such

In essence, each system has unique characteristics; e.g. a level of process-based conformance, a level of in-service data, a level of third-party oversight, etc. This is a change from the pure process-based approach, which doesn't always account for the wider supporting evidence. Therefore, there is a requirement to treat systems as bespoke items with such judgements on the evidence being captured.

Elements of a system, e.g. software, will not be as easily labelled as being standard/guideline compliant

As the solution space opens up, the confidence that is assigned to systems will be derived from non-process evidence. Therefore, the systems and software may not receive a "label" that succinctly states a process-based standard compliance. This is due to there being a wider and diverse set of evidence providing the *confidence*. There should be consideration to remove the labels that are attributed to software, such as "DO-178C DAL B compliant". The concept of "assurance confidence" would be a valid approach to reflect the range of underpinning evidence supporting the target measurement.

Judgment of Evidence

Judgements could be captured with consistent attributes

An assurance approach that uses wider evidence could have the ability to allow *priorities and the purpose of evidence to be managed/captured*. Evidence attributes (e.g. *contribution*) allow the detail of any judgements to be captured and to maintain a consistent narrative to describe the evidence. If a *quantitative* approach is not adopted, the evidence attributes can still drive the judgements and act as prompts for any *qualitative* arguments. A consideration is to also capture the *overheads* (e.g. time, cost, or quality implications) associated with gathering/generating any evidence to achieve a target level of confidence.

A more dynamic evidence landscape requires a defensible position rather than a repeatable one

A move from process-based evidence assessment to one which uses a wider set of evidence, means there is a need to accept, and place value on, the subjective opinion of SMEs. This approach is, arguably, more difficult to measure. Regulatory assurance stances should take an approach to provide further credence to SME judgement. This would allow judgements to be made from a *defensible* position rather than one that is based upon known and accepted benchmarks, i.e. that which is *repeatable*. A *defensible* claim by a stakeholder would be one that is justified with an opinion which can be argued to be *good* [1] (akin to joint reports/submissions regrading differing expert witness perspectives within the criminal justice domain).

Debate and Consensus

Collective judgements need to be captured
An approach underpinned by judgement rather than process conformance may require altered methods for the evidence measurement by stakeholders. A number of subjective stakeholder opinions and their statements of evidence acceptance need to be captured. This information should be able to be debated and reasoned upon, via a suitable structure and method.

Move to a more consensus based approach for evidence judgement and acceptance
To remove any issues regarding single SME judgement, a more consensus driven approach with stakeholder cooperation and approval of the evidence and judgements is needed. As a result, there may be a *joint acceptance* by the stakeholders of the software safety assurance argument.

Procurement and Contracting

Procurement types and stages influence the diverse evidence adopted
There are a number of procurement options for delivering capability with each having merits and demerits. Which option is chosen has consequences in terms of the availability and type of evidence that will be received as part of any procurement. In addition, the stage of the lifecycle, e.g. design, will influence the evidence available and can act as an opportunity to influence any future evidence which is received.

Method to contract for software assurance evidence will need to reflect diverse evidence approach
If systems are treated as unique entities, then the method to gain evidence will also have to be bespoke. With a process-based approach, the evidence requested is very much artefact-driven, e.g. provision of a development plan. However, the *relevant* and *available* evidence (and its context) needs to be understood. Requests to suppliers for the provision of evidence will need to be informed via conversations and an understanding of the software. Suppliers will need to be engaged in a revised manner compared to traditional assurance approaches.

Closing Remarks

This short article aims to highlight a number of considerations if wider diverse software safety assurance evidence is adopted within any revised assurance approach. Any change in stance to how software safety assurance is conducted is certainly not solely reliant on the underpinning evidence. There is a need to look at the systems and processes that support the assurance activities, and this includes how any assurance cases will be articulated to stakeholders [6]. The end-to-end process for undertaking safety assurance needs to reflect *how* any evidence is generated, gathered, and judged.

The authors of this short article have adopted the use of wider diverse evidence within a number of projects and there are challenges with such an approach (e.g. stakeholder buy-in). It certainly isn't an easier route to software qualification; however, there are rewards as the approach, if implemented correctly, can allow systems to be assured efficiently and effectively.

Disclaimer

This article is an overview of UK Ministry of Defence (MOD) sponsored research and is released for informational purposes only. The contents of this article should not be interpreted as representing the views of the UK MOD, nor should it be assumed that they reflect any current or future UK MOD policy. The information contained in this presentation cannot supersede any statutory or contractual requirements or liabilities and is offered without prejudice or commitment.

References

[1] Collins, "English Dictionary and Thesaurus", 1995.

[2] RTCA, "DO-178C: Software Considerations in Airborne Systems and Equipment Certification", 2011.

[3] G. Hanssen, T. Stålhane, and T. Myklebust, "SafeScrum - Agile Development of Safety-Critical Software", 2018.

[4] C. M. Holloway, "Understanding the Overarching Properties", July 2019.

[5] R. Ashmore, M. Standish, "The Measurement of Software Maintenance and Sustainment: Positive Influences and Unintended Consequences", CrossTalk Journal, October 2016.

[6] M.J. Hadley, M. Standish, "Using Tiers of Assurance Evidence to Reduce the Tears! Adopting the 'Wheel of Qualification' for an Alternative Software Safety Assurance Approach", High Integrity Software (HIS) Conference, 5th November 2019.

Mike Standish CEng MBCS MINCOSE

Mike is a senior engineer in systems at the Defence Science and Technology Laboratory (Dstl). Mike has experience of all aspects of software and systems lifecycles, which has been gained in over 15 years within the defence sector. Mike is in the final stages of completing an Engineering Doctorate (EngD) in Systems at the University of Bristol.

Dr Mark Hadley CEng MIET

Mark has been involved in the airborne safety critical software domain for over 20 years with Dstl (and its predecessor organisations) working on a range of UK MOD airborne systems. Mark is a senior principal consultant in software and provides Independent Technical Evaluation (ITE) and SME advice to a host of MOD Project Teams.

SCSC

FOR EVERYONE WORKING IN SYSTEM SAFETY

This seminar is an opportunity to hear a range of talks by industry experts on the practical integration of safety and security. It is aimed at those involved in assurance activities in any sector. It will also be useful for those reviewing, auditing and regulating where both safety and security is involved.

Update on Safety and Security Integration

WWW.SCSC.UK

scsc.uk/e666

THE SAFETY-CRITICAL SYSTEMS CLUB, Tutorial:

Combining Safety With Security

Thursday 17 September, 2020 - London, UK

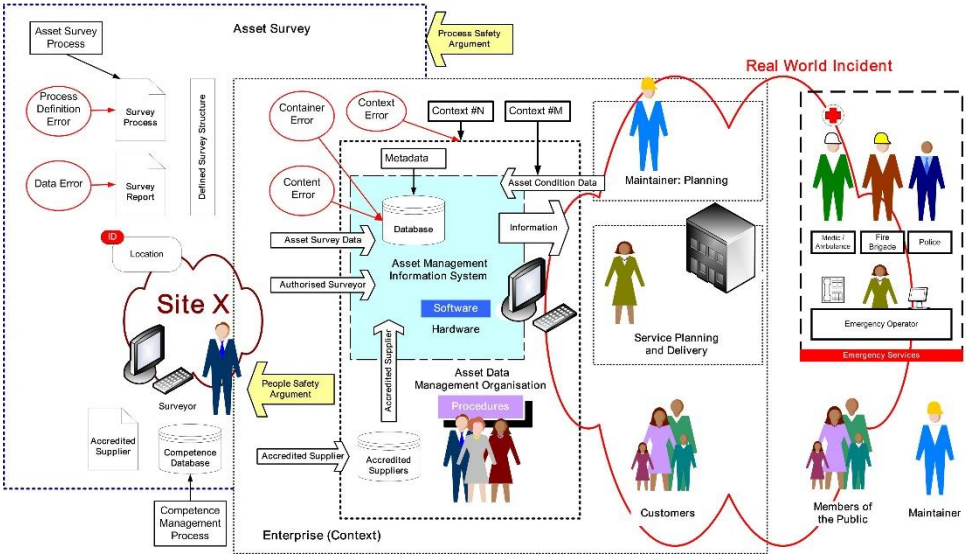
The SCSC Tutorial "Combining Safety With Security" will be held on 17th September 2020 in London.

This tutorial is an update on the latest developments in practical integration of Safety and Security covering: analyses, processes, management, assessments and assurance cases. There will be a variety of speakers from a range of sectors explaining how these two critically important aspects can be assessed and integrated for systems. The ways in which risks, analyses and justifications, including assurance cases, can be linked and combined will be discussed.

Real-life case studies and examples will be used to demonstrate the techniques, across Rail, Aviation, Automotive and other sectors.

There will be a wrap-up session at the end of the afternoon chaired by Mark Nicholson. It also gives the delegates a further opportunity to put further questions to the speakers.

Data, Data Everywhere ...



The easy availability of data has led to Data-Centric Systems (DCS) that are highly data-defined and data-driven. The nature of how the data is used, and depended on, is also radically changing with the advent of Machine Learning (ML) and Autonomous Agents (AA). Alastair Faulkner discusses the challenges of assuring Data-Centric Systems.

Data (in all its forms) is often unchallengeable, unverifiable, ubiquitous, unrecorded and invisible. Yet this data increasingly determines the behaviour of systems and through this behaviour our access to products (goods and services). Data may be internal or fed to systems with safety responsibility. As a result, data error or omission may go undetected with potentially hazardous or catastrophic consequences. There may also be consequent damage to assets. Failure of such systems may also contribute to harm indirectly through incorrect decisions made by actors (human or computer) who rely on, or trust, these systems and the data they supply. How should we reason about Data-Centric Systems (DCS) so that our reliance on, or trust in their correct operation can be justified?

The domain of safety-related DCS is immature, and as such, the safety community has yet to reach consensus on many aspects of architecture, design, implementation, operation and maintenance. Past debates over software best practice and guidelines serve to illustrate the difficulties in reaching consensus. The questions raised in this article require further consideration and I encourage the reader to participate in the debate.

"Swimming in sensors, drowning in data"

The widespread application of infrastructural technologies breakdown old barriers creating an age of connected systems. Seemingly innocuous sensors use highly capable computational platforms. These technologies have the potential to produce vast quantities of data. [1] This combination of production, communication and consumption shifts the focus away from hardware and software to data. This leads to self re-enforcing pressures to create evermore data reliant systems.

"Data, once scarce, is now superabundant"

Superabundance [2] recognises that our ability to produce data exceeds the resources available to transform it into information. This simple statement creates additional challenges beyond production and consumption. Which data should be retained (and why)? For how long? Who should have access? More importantly, what can data be relied on, and why is it good enough?

Before we explore these issues, it is worth a small thought experiment. Consider a defined system. A suitable and sufficient safety analysis identifies several safety functions and associated safety requirements. The apportionment, based on the system architecture, determines at least one of these safety functions is higher than SIL2. A solution constraint requires the use of only hardware and software with the limited use of data. A dual-channel 2oo2 architecture is selected using diverse implementation. This conventional approach fits within the confines of many safety standards and existing safety solutions.

At the last minute, an alternative solution is proposed. It consists of a database supported by generic hardware and software. It is claimed that failures of the generic hardware and software have no (safety) impact. That all the behaviour is described by and contained within the database. Therefore, the safety functions are implemented using data. The existing safety standards constrain safety assessment. For example, two questions arise; firstly, where is the required diversity for this SIL2 system; and secondly, how can this solution implement a dual channel 2oo2 architecture?

"High integrity safety systems require strong contexts and depend on strong-data"

An inescapable conclusion

The debate as to whether data is a separate system component is over. It requires an equitable treatment of data and raises awkward questions as to the safety of existing data reliant safety systems. Nevertheless, the case for data builds day by day; it is now inescapable.

Data Safety Challenges

The combination of data volume and its use to characterise, parametrise, configure and describe the system behaviour provide an overwhelming argument. Data **is** a separate system component and often the dominant element. Data has always been present, typically though not exclusively in lower volumes where its influence is constrained to specific functional areas and domains. Infrastructural technologies are the primary catalyst. It is essential to create foundations as constructs for data safety.

Data is only *valid* in a defined set of contexts

Context is a difficult concept to express. A context may be closed and separate, or open interacting with other different types and categories of systems and operating environments. Its definition includes any information used to characterise the situation of an entity [3]. An entity is a person, place, or object that is considered relevant to the interaction between an actor, an application, systems and their environment, including the actors and systems element themselves [4]. Therefore, defining a context concerns capturing the conceptual structures and frameworks used to construct the system, its boundaries, and its utility in the operational environment. Typically, complex systems contain many different types of actor. Each type of actor has a viewpoint that contains a subset of elements of the context [5]. Extensive datasets are often associated with data ecosystems that include collections of infrastructure, analytics and applications used to capture and analyse data [6].

The properties of a data context relate to its antecedence. For example, strong-data is finite in volume, very specific, such as a sensor or medical record [7]. In contrast, one source of weak-data is data derived from sources often vast in quantity, typically fuzzy and ambiguous, by data analytics [7]. High integrity safety systems require strong contexts and depend on strong-data.

Data Definition

To support safety analysis, data should be defined. The data container may be structured, unstructured or semi-structured and represented in one or more data models. This analysis leads to consideration of the role of data and the reliance on the correctness of the data. Simple data may be a scalar, an array or a table. Data's role is not limited to characterisation, parameterisation and configuration. Data is also the basis of the definition and provision of function, flow and service. In high integrity safety systems, EN 61508 requires diversity as a different means of performing a required function [8]. Diversity is an architectural approach to addressing the issue of Single Point of Failure (SPoF) or Common Cause Failures (CCF) leading to a safety event. Standards typically identify two types of diversity: mechanistic and conceptual. One of the many challenges for Data Safety is the development of a consensus on data diversity.

Data Content

Much literature addresses data quality as a measure of the *content* stored in one or more *containers*. The IEC 25000 [9] series of standards provides a complete treatment of data quality. Data Safety cannot be assured using data quality measures alone.

Bottom-up versus Top-down

Real-world data plays an essential role in all safety systems and their management. As existing data-centric systems with safety responsibility are identified, post-implementation safety analysis is required. Experiential (bottom-up) techniques are necessary for the identification of hazards, their possible removal, mitigation and management. These are the core aspects of all Safety Management Systems (SMS). Experience Based Quantification (EBQ) is a collection of bottom-up techniques used to justify decisions based on data collected from the real world. Competencies derived from EBQ are often confined to those

managing operational process rather than informing and enriching corporate memory. Dis-associations induced by EBQ create a disconnect between corporate management and therefore reduce their ability to implement appropriate strategies.

Once system safety responsibilities are identified, the implementation of safety requirements and any remedial actions may prove difficult. For example, an examination of data errors in a criminal justice system database is only one part of the process, without the resources to correct those errors, they remain in place [10]. Efforts to correct existing data quality issues are bottom-up.

In contrast, top-down safety management follows a more traditional path from requirement, design, implementation, verification, installation operation and maintenance. This implies the use of development and safety lifecycle. On completion, the system is installed, and any changes are instantiated at baseline release. Inferred in this mechanism is that all hazards are known at the time of implementation.

The influence of data on the system extends beyond individual data elements or collections. One means of managing the safety risks associated with data require the consideration of context, container and content. Hawkins et al. [11] have developed 4Plus1 safety assurance principles for software as a pragmatic approach that considers both top-down and bottom-up. The 4Plus1 safety principles are reinterpreted by the author for Data Safety in Table 1.

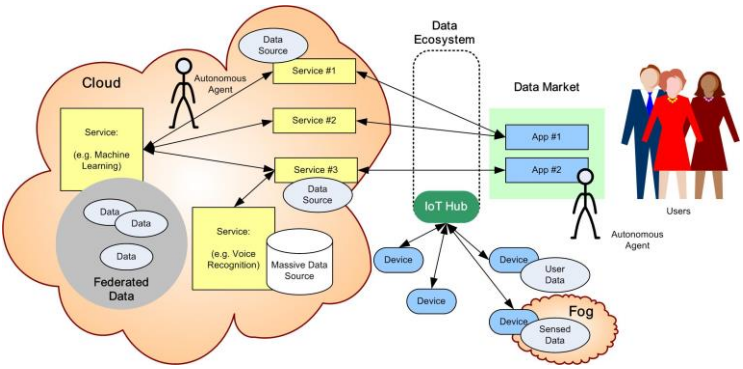
Table 1: 4Plus1 Data Safety Principles

Principle	Description
1	(Top Down) (Data) safety requirements shall be defined to address the data contribution to system hazards and associated risks.
2	(Top Down) The intent of (data) safety requirements shall be maintained throughout requirements decomposition (and apportionment to data components).
3	(Top Down) (Data) safety requirements shall be satisfied.
4	(Bottom Up) Hazardous behaviour has been identified and mitigated.
Plus1	The confidence established in addressing the (data) safety principles shall be commensurate to the contribution of data to system hazards and associated risks.

As the lifecycle progresses, the requirements and design are progressively decomposed. In this way, safety requirements relating to action response requirements are decomposed into knowledge requirements, then information requirements and finally, data requirements. A pragmatic approach requires consideration of the technological constraints, including assessing the requirements placed on the data elements such as sensors and their acceptable Operational Design Domain (ODD). A more detailed design is created. The intent of the requirements must be maintained as the safety requirements are decomposed. Data safety requirements must ensure that the safety intent is maintained as the data architecture and individual data elements emerge.

Data Safety Architectures, Design, Techniques and Measures

It is relatively easy to define an acceptable ODD for a sensor. This task becomes more complex when combing multiple sensors in an array. Sensor fusion addresses how different sensors are combined. Sensors are only one source of data.



Architecture

How should all these elements be combined into an appropriate set of safety architectures that provide robust defences against SPoF and CCF? This consideration illustrates the requirement for multi-channel systems and diversity. In the absence of safety community consensus, defining acceptable ODD for safety subsystems and systems will be challenging. Many of these systems will not be stand-alone but operate and co-operate with others in a hierarchy.

Design

Data Safety spans the spectrum of implementation from high-integrity protection systems to information systems. The increasing use of agile development creates challenges for safety analysis and safety acceptance due to the short iterative cycles and the currency of the documentation. Data-defined and data-directed systems often use generic hardware and software, their form and behaviours are wholly described by data.

As systems become more extensive in scale, scope and complexity, a system will likely consist of multiple instantiations. This will be the case for Autonomous Vehicles (AV). What design elements are required for version, configuration management to provide for maintenance, upgrade and incident investigation? One suggested method for addressing this is that an acceptable AV ODD should also include interaction with other AVs and their environment.

Techniques and Measures

A safety community consensus for the recommendation of techniques and measures [12] is essential to the process of assurance and safety acceptance. The rigour of their use provides evidence and sets the competence, training and experience requirement on the developers. This becomes a circular argument, as, without safety community consensus for data architecture, design and implementation of each system safety assessment and acceptance are highly dependent on the individual assessor. Therefore, one option is to include a risk of 'failure to gain acceptance' in the project risk register.

Data Consumption

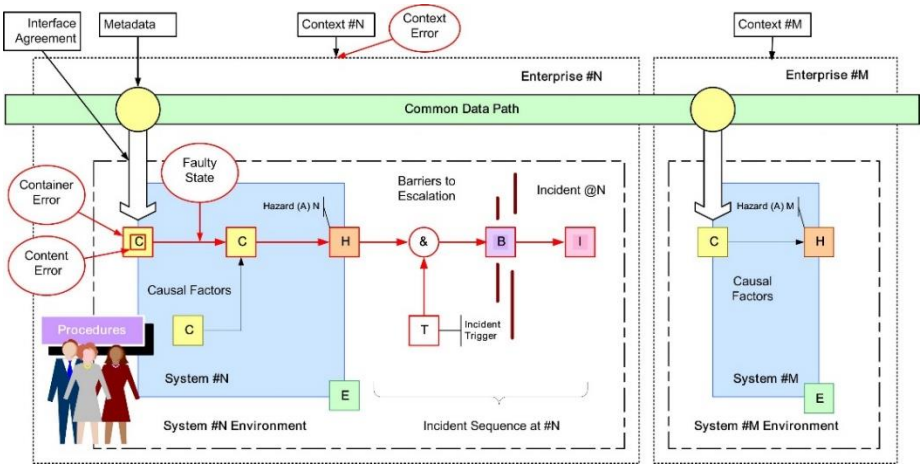
With a reliance on data, data integrity requirements are essential.

Given the contribution of data errors to the system behaviour safety analysis is likely to result in the apportionment of a high percentage of the safety risk to the data component. Data-Centric Systems (DCS) are unlikely to be stand-alone, their data integrity requirements are to be satisfied at each interface (and by implication the data path (data supply chain)).

Data Sources and Production

The ready availability of some data does not qualify it for use in a safety system. Solid foundations are essential to all engineered systems. This is also true for DCS. Data may originate from many sources; it may be highly processed and transformed. Its ownership may change many times. Who then will be liable for data error and consequent losses?

Is it reasonable to require an acceptable ODD for each element of the data path, or should the ODD be limited to the point of delivery (interface) of the consuming system? It depends on the use made of the data and the influence of data errors on the consuming system. A single data path might supply multiple systems. Only one of those systems might fail due to a specific data error.

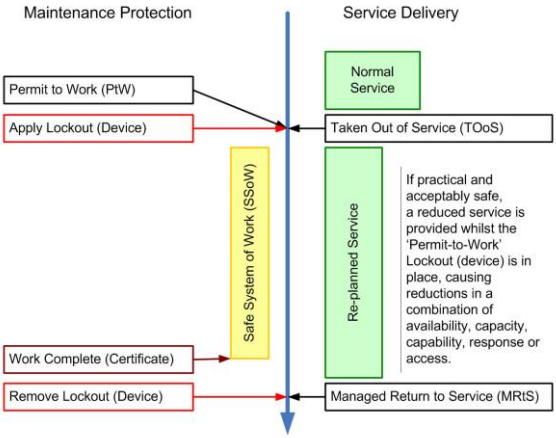


A DCS may require extensive data. Not all of this data will be needed by all the types of actors using the DCS. Classic safety engineering requires the identification of one or more boundaries, typically, although not exclusively interfaces. In this classical world, errors, faults and failures give rise to hazards at the system or component boundary. Where are these boundaries in extensive data? One suggested method for addressing this is to develop the concept of Interface Agreements (IA) as a means to describe both physical and virtual boundaries.

Interface Agreements

An interface is a point where two systems, system and environment, subjects, organisations, etc. meet and interact [13]. Conmy [14, 15] proposes Interface Contracts as a set of one or more [safety] constraints which must be upheld. These are typically used to describe high-integrity safety rely-guarantee (theorem) constraints. In a more generalised form, IA [16] are a set of one or more constraints which must be upheld by system components to meet one or more requirement(s).

Critical Control Points



Creating and installing a system also requires consideration of its maintenance. How should a system element be safely taken out of service? One suggested method develops the concept of a Critical Control Point: (CCP) as an IA used to provide data on the status, performance and behaviour of the system across the operational context. A CCP provides one means to implement Permit-to-Work (PtW).

Typically, a PtW is a management procedure where only persons with specific authority will sign a permit on which ostensibly the life of a worker might depend. To this end, responsibility for the PtW rests with the person in charge of the operation for which the permit is required [17].

Safety Acceptance

If safety acceptance to a safety standard such as IEC 61508, then, how should a DCS be assessed for compliance? Taking the assessors point of view, the overall architecture and hardware and software components are well-defined. The assessor is constrained by the standard, which says very little about data. What should the assessor do when a majority of the safety integrity requirements are apportioned to the data component?

Extensive scale, scope and complexity requires the assessor to recognise that elements of the system may be in different states at different times. It is proposed that acceptance and approval should consider:

- Staged start-ups
- Staggered start-ups
- Asynchronous start-ups (large distributed systems)
- Synchronous start-ups (aligning timing requirements and hand over [live to hot stand by])
- Localisation, Segregation and Quarantine (treatment – inoculation)
- Maintenance requirements and restrictions
- Operation without a Safe System State

Operation and Maintenance

Safety is a property of the operational system [18]. Where data describe the system and its behaviours, we should not assume that a system change will use the change procedures normally associated with the development lifecycle, change management, versions and baselines. Shortly after the widespread introduction of AVs, it is easy to imagine many vendors, each with many models and versions operating with conventional vehicles. In this AV system, there will be many contexts, containers and multiple (possibly duplicate) content.

We should not assume that the user is human. An actor may be an individual, entity, or combination of product, people and process. The role of the actor will increasingly be undertaken by one or more AAs. Where an AA is an entity operating on the owner's behalf as an actor without interference from the ownership entity. Typically, these are products that incorporate varying degrees of Machine Learning (ML) [19, 20].

It is proposed that the use of actors requires the definition and enforcement of identity management, such as the consideration of:

- Authentication
- Authorities
- Fraud Detection, Enforcement and Management
- (False) Identity [Agents, Users, Customers] (Joiners, Leavers and Renewal)
- (False) Identity (Components, Combinations of Components, and Systems)
- All identities (should) expire and need to be renewed? (over what period)
- Attacks and Malicious Event Response

Incident Investigation

Over the last five years, we have been fortunate to live through a period of relatively few major incident and accidents. One consequence has been a complacency and reduced focus on safety management activities. Many senior and experienced practitioners have reached or are approaching retirement.

In the absence of major incidents, existing safety margins in existing systems have proved sufficient. These technologies are now mid-life and approaching obsolescence; elements of these systems now require replacement. Fit-form-function replacements may be based on DCS.

Data is a decisive and disruptive technology; therefore, in the absence of a definition of safety-related DCS architectures or other desirable properties, the incident investigation method needs to be tailored to each context. Also, the use of data ecosystems requires characterisation for its potential influence on the incident. This characterisation is the basis of the formalised and documented approach to its investigation.

Data Safety

With the emergence of data as a separate system component arises the requirement for the consideration of Data Safety. In this sense, all that is argued for is equality with the other safety system components. The development of a safety community consensus on architecture, data structures, techniques and measures.

Data will enable and facilitate dynamic behaviours. This new dynamism finds form where data defines the system and its behaviours. Potentially, the system has escaped the controls normally associated with the development environment. In this sense, deployed products are 'unfinished'; their behaviours are modified by AA and ML through their operational experience. Data-Centric Incidents (DCI) will be caused by data component errors. Data ownership will become a contentious issue. Who will be liable for errors associated with data?

Conclusions

Hopefully, this article has provided some insight into the safety management challenges associated with DCS. One of the most profound changes is the shift from development based approaches to the creation of safety products to the operational arena. The ease with which data can be changed provides an irresistible economic factor for the proliferation of DCS.

As safety practitioners, we address engineered systems, created by a designer, to implement a design intent, developed, operated and maintained to that design intent. We commonly consider products as engineered artefacts. As a divisive technology, data will induce a paradigm shift. The underlying message of this article is the evolution of known and mature systems safety concepts, rather than revolution. DCS are only one aspect of change. Those organisations that use DCS will become increasingly reliant on them and in doing so become Data-Centric Organisations (DCO). One aspect of this evolution is to address increasingly open systems with external interactions across the system boundary. Increasingly open systems extend existing cyber-security risk and highlight issues of identity. It is proposed that that identity will be extended to ensure the unique labelling of attributes of the object (system resource) being accessed and of the actor requesting access in a given context.

In developing this article, it follows that there will be DCIs. How should they be investigated? Unlike physical incidents, DCIs may not leave physical witness marks such as tyre skid marks. Data error or omission may go undetected, and may also contribute to harm indirectly through incorrect decisions made by actors (human or computer) who rely on, or trust, these systems and the data they supply. Without proper design, DCS and DCO may become interconnected, interdependent, and as a result, will not be analysable.

Finally, the proposition that 'data is a separate system component' has widespread repercussions which are not simply limited to a system being comprised of hardware, software, actors, process and data. Data quickly becomes the dominant component. Without safety community consensus addressing context, container and content, it is unclear how the safety risks associated with data will be adequately managed.

References

[1] Lt. Gen. David A. Deptula. Military 'Swimming In Sensors and Drowning in Data'. URL: http://www.nationaldefensemagazine.org/articles/2009/12/31/2010january-military-swimming-in-sensors-and-drowning-in-data

[2] Kenneth Cukier. Data, Data Everywhere. The Economist, 2010. URL: https://www.economist.com/special-report/2010/02/27/data-data-everywhere

[3] Anind K. Dey and Gregory D. Abowd. Towards a better understanding of context and context-awareness. Proceedings of the Workshop on the What, Who, Where, When and How of Context-Awareness, affiliated with the CHI 2000 Conf. on Human Factors in Computer Systems, 2000

[4] Robert Flood and Ewart Carson. Dealing with complexity: An Introduction to the Theory and Application of Systems Science. Volume 2nd ed. ISBN 978-0306442995. Plenum Press, New York, NY, USA, 1993

[5] BKCASE Governance and Editorial Board. Engineered System Context. 2017. URL: https://www.sebokwiki.org/wiki/Engineered_System_Context

[6] Marcelo Iury S. Oliveira and Bernadette Farias Lóscio. What is a Data Ecosystem? 978-1-4503-6526-0. ACM, 2018, 74:1–74:9

[7] Shomit Ghose. Engineered Influence: Weak Data, Machine Learning and Behavioral Economics. 2017 Sutardja Center for Entrepreneurship and Technology's annual journal AIR (Applied Innovation Review), 2017. URL: https://scet.berkeley.edu/engineered-influence-weak-data-machine-learning-behavioral-economics/

[8] IEC 61508: Functional safety of electrical/electronic/programmable electronic safety-related systems - Part 4: Definitions and abbreviations. International Electrotechnical Commission, 2010

[9] IEC 25000: Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – Guide to SQuaRE. International Electrotechnical Commission, 2014. URL:http://iso25000.com

[10] Bill Blackburn, Paul Hampton and Mike Parsons, Data in Police and Criminal Justice Systems: How data errors could lead to harm to innocent citizens (or, how the film "Brazil" was spot-on), in Evolution of System Safety, Proceedings of the Twenty-sixth Safety-critical Systems Symposium, York, UK, 6th-8th February 2018, https://scsc.uk/SCSC-140

[11] Richard Hawkins, Ibrahim Habli, and Tim Kelly. The Principles of Software Safety Assurance. International System Safety Conference (ISSC), Boston, 2013

[12] Data Safety Guidance. Version 3.1. ISBN-13: 9781793375766. Safety Critical Systems Club - Data Safety Initiative Working Group, Kindle Direct Publishing Platform, 2019, https://scsc.uk/SCSC-127D

[13] Interface - Definition. Oxford Living Dictionaries. URL: https://en.oxforddictionaries.com/definition/interface

[14] Phillipa Conmy, John McDermid, and Mark Nicholson. Safety assurance contracts for integrated modular avionics. Volume 33. SCS '03 Proceedings of the 8th Australian workshop on Safety critical systems and software, 2003, pages 69–78

[15] Phillipa Conmy. Safety Analysis of Computer Resource Management Software. University of York, PhD Thesis, 2005

[16] Alastair Faulkner and Mark Nicholson, Data-Centric Safety - Challenges, Approaches, and Incident Investigation, Elsevier, 2020, Version, ISBN 978-0-12-820790-1

[17] Institution of Engineering and Technology (IET). Safe Systems of Work (Health and Safety Briefing No. 32). 2015

[18] Alastair Faulkner: "Safety Arguments for Use with Data-driven Safety Systems", Proceedings of the fourteenth Safety-critical Systems Symposium, pp 263-276 ISBN: 1-84628-333-7, Bristol, UK 2006.

[19] Stan Franklin and Art Graesser. Is it an Agent, or just a Program?:A Taxonomy for Autonomous Agents. Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages, Springer-Verlag, 1996. URL: http://robotics.cs.tamu.edu/dshell/cs631/papers/franklingraesser96agents.pdf

[20] Jae-Gil Lee et al. "Can Autonomous Vehicles Be Safe and Trustworthy? Effects of Appearance and Autonomy of Unmanned Driving Systems". In: International Journal of Human–Computer Interaction 31.10 (2015), pages 682–691

Alastair Faulkner, Abbeymeade Limited

Alastair has over 40 years' experience in the application of systems, software and safety engineering. The last 20 years have been in safety consultancy. An early interest in data dependent systems led to an Engineering Doctorate at Warwick (*Data integrity: an often ignored aspect of safety systems*). Alastair is a highly competent practitioner, with comprehensive experience in senior development and management roles in systems, applications, and infrastructure environments. Alastair is a joint author of an Elsevier publication '*Data-centric Safety – Challenges, Opportunities and Incident Investigation*' will be available from March 2020 – ISBN 978-0-12-820790-1) with Mark Nicholson.

Email: alastair.faulkner@abbeymeade.co.uk

Formalising the Language of Risk

The use of natural language in engineering and specifically, engineering risk management, is often problematic due to assumed meanings and usage contexts of domain terms, with the result that misunderstandings can arise. Dave Banham provides an introduction to a formalised structure of words – an ontology – by which, at least, the risks arising from system safety and security concerns can be described unambiguously using a common language.

The Problem

The Data Safety Initiative Working Group (DSIWG) has been working on guidance for the management of data safety risks for several years and has now produced mature guidance material. The group has however, found that the language of safety and risk is contextually dependent with terms having multiple meanings, different terms being used with the same implied meaning, and, as is often the case in English, terms being used without qualification.

Add to this situation the risk terminology used by cyber-security professionals, because security informed data safety is a useful adjunct, and the result is a long glossary of terms with no self-consistency. It is therefore hard to produce guidance for data safety that is clear and accessible. Moreover, often this variability of meaning, introduces subtle misunderstandings that take conscious effort to detect and resolve.

To help resolve these issues the DSIWG established a sub-working group to establish a formalised structure of words – an ontology - to describe risk. The intention was to use the ontology to be more precise in the language used in the guidance, but it has been realised this ontology could have much wider applicability across the entire safety and security domains.

What do we mean by Risk?

In common usage, "risk" means an activity or situation that has a chance of a significantly unpleasant outcome in the worldview of the observer making the statement. Risk is generally used as the adjective "risky" to qualify the sense of uncertainty being expressed about the named activity: *parachute diving is risky; driving fast is risky; betting on slot machines is risky*, etc. Risk can be used as a noun when conceptualising it: *I accept the risks involved in free climbing; the risk of injury in rugby is high*. The two forms can be combined to yield sentences such as: *there is a risk of injury from risky driving*. However, note how easily the implied meaning of risk shifts from one of uncertainty in outcome (when used as adjective), to that of likelihood (when used as noun).

Moreover, the outcomes and likelihoods that are often inexplicitly stated as an assumed shared understanding, are significantly undesirable in the worldview of the person making the statement, but may be considered otherwise by somebody else. Free climbing is one person's *horror story*, but another's *pleasant sport*. The common language use of the word "risk" is completely inadequate for engineering where assumptions need to be eliminated in preference for precise terms, calculations, and a shared (and agreed) worldview.

Articulating Risk

Engineering makes use of standardised terms to help articulate risk. The international standard for risk management is ISO 31000 [1], with the compendium ISO vocabulary of risk terms, ISO Guide 73 [2].

ISO Guide 73 defines risk as the *effect of uncertainty on objectives of stakeholders*. Objectives are things that stakeholders seek or want to avoid. We don't want harm to arise from the use of our goods and services; conversely, we want to make money from selling our goods and services. Uncertainty exists in the fulfilment of these objectives due to phenomena such as natural processes, unforeseen circumstances, competition, etc. When things are certain (perhaps because they have already happened), there is no risk.

This leads to the idea of positive risk (*seeking a benefit*) and negative risk *(avoiding a harm*). The common use of "risk" is in the sense of negative risk; harmful (often physical) situations that need to be avoided. However, risk arises from the worldview, frame, or context that the stakeholder has since they own the objective. Consider theft. The owner of a valuable asset wants to protect that asset from, amongst other things, theft. Theft results in a harm that creates a loss, to the owner, of the stolen asset and is thus a negative risk concept to the stolen asset's owner. Whereas to the criminal, theft is the means by which value is gained (a benefit) and is thus a positive risk concept to them, notwithstanding the negative risk of being caught.

The language of safety is formalised around that of risk. Let us start by defining "harm". Harm is the *consequence of a failure* to meet stakeholder objectives when the consequential situation is undesirable to them. The converse is a "Benefit", when the *consequence is desirable* to them. A subset of the total set of possible harms is the set of safety-related harms. A safety-related harm is generally defined as a physical harm that impacts the health or life of a person or persons, or impacts the wellbeing of the natural environment. Although a stakeholder may include other impacts such as the loss of an asset, loss of reputation, etc. in their definition.

How can harms arise? Since harms are generally not certain, specific situations need to occur to allow them to arise as a *consequence*. These causal situations are referred to as *incidents*. An incident is a *dangerous event* (i.e. a moment in time) and is therefore a *danger source*. (That is, danger may lead to harm.)

A near miss is an incident that did not lead to harm, but had the potential to do so. An accident arises from an unintentional incident that leads to harm; that is, an accident arises from unintentional sources of danger.

Incidents can be intentionally created and sometimes maliciously; for example, by arsonists, thieves, or by misguided misuses of a system (i.e. by incompetent users). As such, the term "incident" is more useful than purely safety terms such as "accident", as it allows the safety analysis to consider a wider set of concerns that have traditionally been, for example, the reserve of security specialists.

One purpose of a system safety analysis is to theorise about what potential harms a system may cause and to identify the potential danger sources that may lead to them. An identified danger source is called a *hazard*; a hazard is a known danger source that may lead to an incident that causes harm.

A risk score is a metric arising from a function of a potential incident's likelihood and the desirability of the potential outcome. Hence, numerically:

risk = likelihood × desirability

where *likelihood* is a probability of occurrence, *desirability* is a positive score when the objective is sought and a negative score when it is to be avoided, and where × is a binary operator (i.e. a function) taking two parameters. Hence, a negative risk score indicates the risk of a harm, which conforms to the ISO 31000 framework. In the context of harms, desirability is often stated as a severity score and the equation of negative risk can be stated as:

risk = -(likelihood × severity)

To understand how harm may arise we either start with a harm and ask the deductive question of *how can it arise*, or we start with some other aspect of the system such as a system input, or a subsystem and ask the inductive question of *what would happen if*. Where a weakness or susceptibility to failure is found in a constituent part of a system (which also includes people when they form an active part of the system) then a *vulnerability* is said to exist.

A danger source is something that can exploit a vulnerability to create an incident. An incident that is a system failure can result in harm, although typically what happens is that the incident is a failure that is more localised to a constituent part of the system; that is, a part no longer completely fulfils its objectives. A localised failure manifests as a fault (failure condition) that can propagate through a system exploiting other vulnerabilities (that is, triggering other failures) until potentially the system fails with some harmful outcome.

A Model of Risk Terminology

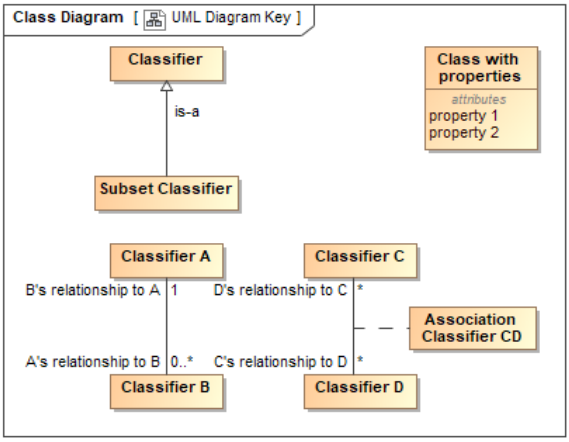
We can describe this terminology formally in an ontology and use UML class diagrams to represent aspects of that ontology through a series of diagrams. The ontology captures the ISO 31000 concept of desired and undesired stakeholder objectives through the consequences of benefit and harm. However, from a safety and security point of view, our main interest is in the risks associated with harms and, as a result, the ontology is significantly more refined in this area. Nevertheless, it is important to understand the opportunity and benefit cases that malicious threat actors may have towards a system.

The figure on the next page shows the graphical notation subset of UML 2 [3] class diagrams that are used to model the risk ontology. The rectangular shapes within the diagram frame represent classifiers, which are used to describe the language terms in the ontology.

The *is-a* relationship denoted by the hollow closed arrow headed edge (▷) describes a taxonomical relationship between classifiers where the specialised classifier is a subset concept of the more general classifier at the arrowhead end of the edge. The ontology optionally clarifies the specialisation with an annotation text shown with double angle quotes next to the generalisation edge as follows: «Classifies» and «Subsets»

A «Classifies» relationship denotes a set of specialisations that can be used in an additive fashion; that is, they are overlapping and additive concepts. Whereas a «Subsets» relationship denotes a set of specialisations that are distinct from each other; that is, they are non-overlapping concepts.

For example, a vehicle can be classified by its means of its source of power, its means of motion, and its colour (to name just a few). Each of these classifiers can be subset, so for example for power, we could have diesel engine, electric engine, gas turbine, etc., and for the subsets of means of motion we could have, wheels, wings, hull, etc., and for colour some set of colours.



From this set of classifiers and their subsets we can describe a vehicle as red, with diesel engine and wheels, or as red with gas turbine and wings. The classifiers are additive and individually describe an aspect of the thing (a vehicle in this example) they classify. They also provide discrimination by class, so in this example all the red vehicles can be identified, irrespective of their other classifiers.

A class can relate to other classes in non-taxonomical ways and these are denoted by edges with either no arrowheads, where the relationship is bidirectional (as shown in the figure for classifiers A and B), or with a single open arrowhead end (→) to denote a unidirectional relationship. The meaning of the relationship is denoted by the verb phrases at each end of the edge for bidirectional relationship, or just at the arrow headed end for unidirectional relationship.

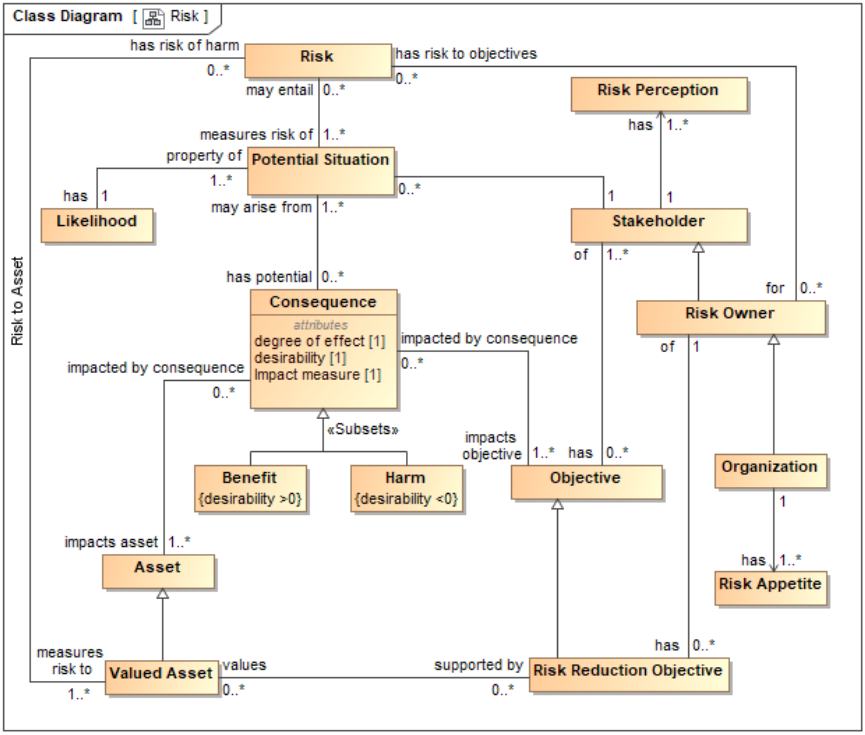
A relation end multiplicity quantifies the permissible number of relationships that may exist when the related terms are being used. The table below lists the typical multiplicities that have been used and their corresponding meaning. The combination of the multiplicity and the associated end verb phrase combine to provide a quantified relation from one classifier to the related classified.

Multiplicity designation	Meaning
1	One
0..1	May have (i.e. none or one)
1..*	Some (i.e. one or more)
*	May have some (i.e. none, one, or more)

An association can be further qualified by an association class (as shown in the figure for classifiers C and D with association class CD). The purpose of an association class is to provide a class based definition of the association; that is, an association class is a class with edges. One benefit this provides in ontology modelling is that it allows relationships to be defined terms by virtue of the association class name.

A Model for Risk

To show an example of the ontology, consider the following figure that shows the relationships associated with the entity "Risk".



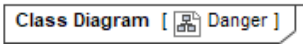
From the "Risk" Class Diagram on the previous page, we can derive the following narrative:

Risk is a measure of the uncertainty in attaining **stakeholder** objectives. **Objectives** are things that **stakeholders** want or do not want to happen, which are not certainties.

Risks are contextualised against the **asset** or assets that are impacted by the **objectives**, given the *likelihood* of the **potential situations** that may arise from them and the *desirability* of the **situation** that may arise as **consequence**. Such **consequences** can be *desired* or not *desired* and we call this a **benefit** and a **harm** respectively.

In terms of managing risk, there needs to be an identified **risk owner** that has **risk reduction objectives** that relate to the subset of **assets** that are considered to be of value (i.e. **valued assets**). This corresponds with the pragmatic view that the formulation of a risk treatment strategy needs to be targeted to be cost effective.

A further example of the ontology is given in "Danger" class diagram in the figure on the opposite page. In this figure we start to see how the ontology provides a language that is common between safety and security domains.



As with the previous figure, we can derive a narrative from the diagram opposite as follows:

Danger is described by the *possibility* that an **undesirable situation** may cause **harm**, as denoted by the relationship between these two terms. More specifically, we can state that **harm** arises from **incidents** that cause it; an **incident** is the cause and **harm** is the **consequence**.

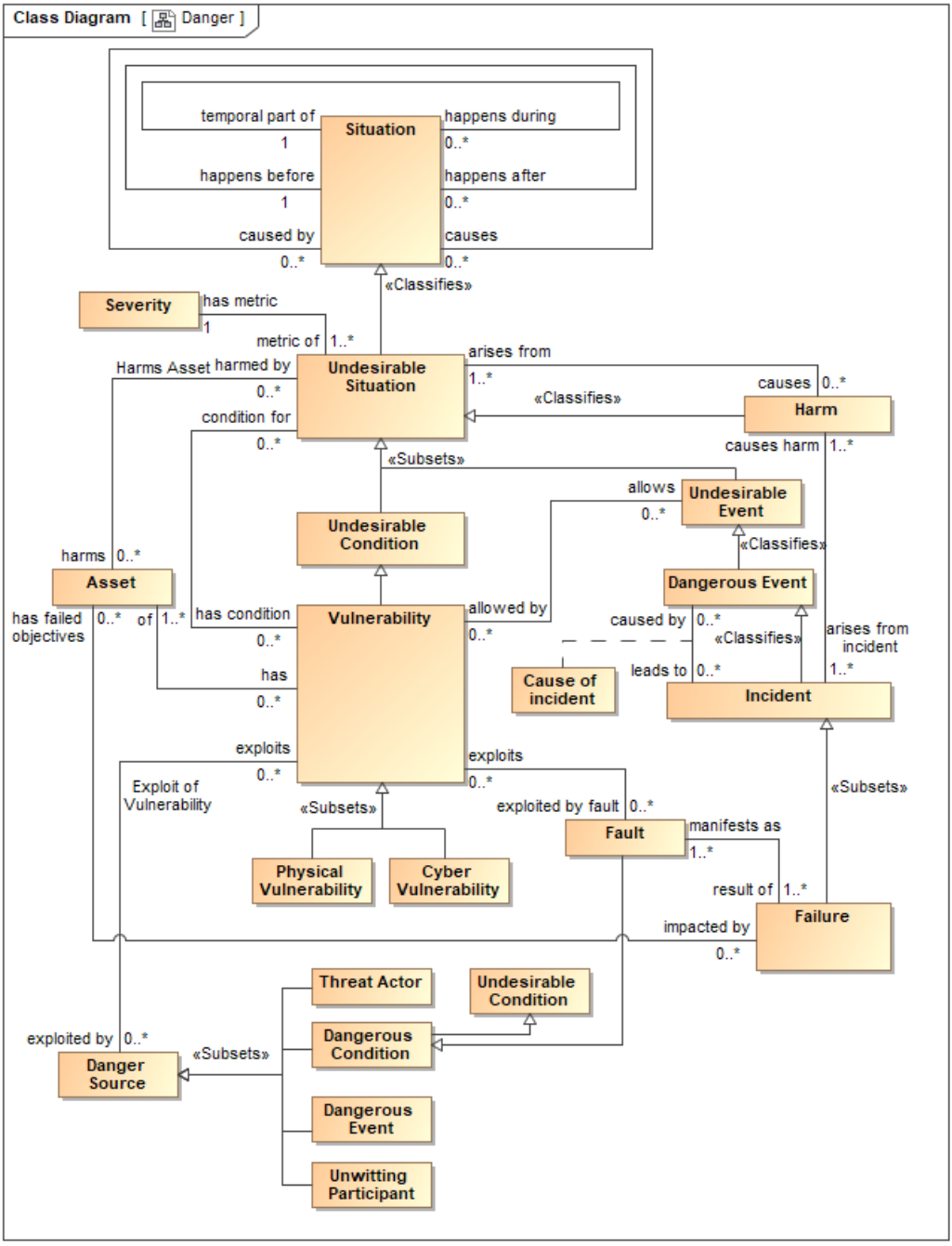
An **incident** is both a **danger source** (a **dangerous event**) and an **undesirable situation** (an **undesirable event**). Hence, the *possible* relationship between **undesirable situation** and **harm** becomes a substantiated one between **incident** and **harm**.

The degree of *danger* posed by an **undesirable situation** is captured by its **severity** property.

Assets can have **vulnerabilities**, where a **vulnerability** defines the conditions that *allow* an **undesirable event** to occur. A **danger source** is the generic term for something – natural, systematic, or intentional – that can *exploit* a **vulnerability** to create **an undesirable event**.

Since **undesirable events** can be classified as **dangerous events**, which is a **danger source**, a chain of events can be created whereby a series of **vulnerabilities** are exploited until an **incident** occurs and **harm** arises.

The term "exploit" is used here with both its "use" and "abuse" meanings. Physical things have **physical vulnerabilities** that are subject to the laws of physics and particularly the law of entropy; physical things break. They break through the *wear and tear* of natural use, and they break by being abused and misused. Complex systems have both physical vulnerabilities and vulnerabilities arising out of design limitations and design flaws (i.e. systematic defects). In computer-based systems, these design vulnerabilities are called **cyber vulnerabilities.**


Class Diagram [Danger]
temporal part of
1
Situation
happens during
0..*
happens before
1
happens after
0..*
caused by
0..*
causes
0..*
«Classifies»
Severity
has metric
1
metric of
1..*
Harms Asset
harmed by
0..*
Undesirable Situation
arises from
1..*
«Classifies»
causes
0..*
Harm
condition for
0..*
causes harm
1..*
«Subsets»
allows
0..*
Undesirable Event
Undesirable Condition
«Classifies»
harms
0..*
Asset
Dangerous Event
has failed objectives
0..*
of
1..*
has condition
0..*
Vulnerability
allowed by
0..*
caused by
0..*
«Classifies»
arises from incident
1..*
Cause of incident
leads to
0..*
Incident
has
0..*
exploits
0..*
exploits
0..*
Exploit of Vulnerability
«Subsets»
exploited by fault
0..*
Fault
manifests as
1..*
«Subsets»
Physical Vulnerability
Cyber Vulnerability
result of
1..*
Failure
impacted by
0..*
Threat Actor
Undesirable Condition
exploited by
0..*
Danger Source
«Subsets»
Dangerous Condition
Dangerous Event
Unwitting Participant

For this article, it has only been possible to provide a brief introduction to the ontology with only a small subset of the model being presented. For further information and a much more detailed description of the model, refer to the paper "Formalising the Language of Risk" published as part of the 2020 Safety-Critical Systems Club Symposium Proceedings [4].

Conclusion

This article has set out to introduce the Risk Ontology that the Data Safety Initiative Working Group has assembled. The language is self-consistent (by virtue of the ontology formalism), and provides the means for describing causal situations that can result in harm to assets. The power of expression in the language is aided by the ontology classification meta-language as it allows terms to inherit higher order concepts.

An example being that whilst harm is a consequence that results from some other situation, harm is also a situation. This allows causal modelling to show, for example, how harms can propagate. For example, a fire in a bin (a localised harm) spreads (due to lack of adequate containment and/or proximity to other combustible materials) to destroy the building (a larger scale harm). Moreover, the causality modelling afforded by the situation related terms can be used in incident investigation (i.e. after the fact) where evidence is being assessed to determine why and how something occurred.

The ontology attempts to find common ground between the safety and security risk analysis by using unified terms such as "incident" and "vulnerability". The language described may not cover all aspects of safety or security analysis, but it is hoped that it provides enough common language to enable greater productivity in achieving security informed system safety.

Acknowledgements

The author would like to acknowledge the significant work of the "Threat and Risk Community" (threatrisk.org) in creating an ontology that paved the way for the creation of the DSIWG's own Risk Ontology. In that respect, our ontology shares many of the same concepts and terms as the Threat & Risk Ontology, although ours is smaller, in part because it lacks some of the foundational terms that the Threat & Risk Ontology formally defines.

The author would also like to thank the following people for their contribution to this paper and to the Risk Ontology: Paul Hampton, Divya Atkins, Martin Atkins, and Mike Parsons.

[1] "Risk Management - Guidelines," ISO 31000, 2018.

[2] "Risk Management - Vocabulary," ISO Guide 73, 2009.

[3] "Information technology - Object Management Group Unified Modeling Language (OMG UML), Superstructure," ISO/IEC 19505-2, 2012.

[4] "Assuring Safe Autonomy: Proceedings of the 28th Safety-Critical Systems Symposium (SSS'20) York, UK, 11th-13th February 2020", SCSC, January 2020, https://scsc.uk/SCSC-154

Dave Banham, Functional Safety Specialist at BlackBerry QNX

Dave Banham is a Chartered Engineer specialising in dependable software intensive and cyber physical systems. His professional career spans 28+ years including time at GEC, Alstom, Areva, Rolls-Royce, and most recently he joined the functional safety team at BlackBerry QNX. He is an active member of the DSIWG, MISRA C, MISRA C++ , and OMG SysML 1.7 working groups.

SCSC Senior Leadership Forum: Safety Management Systems

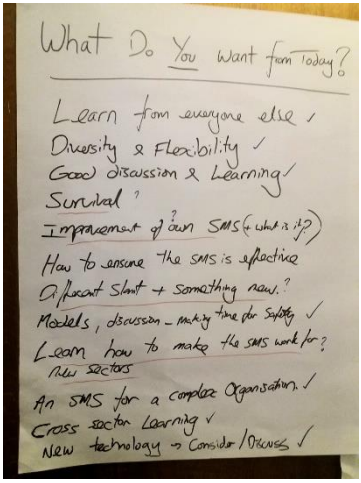
Senior leaders, drawn from a wide range of safety-related domains, met in York in October 2019 to discuss best practices in Safety Management Systems.

This was a one-day event held in the Grand Hotel in York and the forum was led by Silas Hays from THI Safety Management Systems. There were 16 delegates representing sectors such as Aviation, Railway, Healthcare, Energy and Academia. The main topic of the event was to discuss the group's experiences and views on the effectiveness of Safety Management Systems.

Safety Management System Effectiveness

Silas opened the session with a personal perspective of the Nimrod MR2 aircraft and his involvement in the subsequent investigation of the XV230 accident in 2006 that culminated in the Hadden-Cave report. In his account, he noted the challenges he experienced in the absence of a formally recognised Safety Management System.

Each individual in the group introduced themselves and described their own personal objectives for the day, which were then collated.



The group then shared their individual views and perspectives of how effective Safety Management Systems (SMSs) were in their own organisations and a number of themes arose:

- SMSs can place inordinate responsibility on the individual at the lowest level of the organisation, with Top Management being disconnected and out of touch with the actual organisational risks;
- SMSs can be overly prescriptive, and thus promote a tendency for staff to think only in terms of compliance to process, rather than applying creative thought in managing safety risks;
- SMSs can be well defined but the organisation may not have the expertise or budget to implement the processes effectively;
- SMSs can be hindered by the complexities of organisational relationships with other stakeholders especially around contractual and regulatory boundaries;
- SMSs can be inflexible to change, as organisations are reluctant to deviate from established practices even when there may be good reasons to challenge those processes – analogous to the "Five Monkeys" experiment.

"Five Monkeys find themselves in a room with a ladder leading to a bunch of bananas. Any monkey that tries to climb the ladder to reach the bananas is sent sprawling with a spray of cold water that soaks the entire company, and they all sit miserably wet, cold and hungry. An experimenter replaces one of the monkeys with a new monkey. When this new, naïve, monkey tries to reach the bananas, he is soon pulled back down from the ladder by the others, as they can't bear to be soaked again. Another wet monkey is replaced, and again, when this new monkey tries to reach the bananas she is duly pulled back by the group, including* the monkey that has no experience of being soaked. *One by one, the original wet monkeys are replaced and eventually, even though there are no monkeys left that have ever experienced the soaking, none are permitted to climb the ladder again."

Measuring Effectiveness

The means by which the effectiveness of an SMS could be measured was discussed by the group. A number of methods used in practices were proposed by the group, such as audits, reviews, performance metrics etc. and there are various tools to support the assessment such as the CAA SMS Evaluation Tool.

It was however, acknowledged that measuring more subjective aspects such as "Top Management engagement" was more challenging. One suggested method was to use metrics on attendance of management (planned versus actual) at safety meetings. Safety culture questionnaires are also another method to measure effectiveness as they represent an individual's perception of how well the SMS is understood and being followed.

The use of a "traffic light" dashboard across the SMS processes was also recommended:

- To provide an "at a glance" overview of the status of the entire SMS
- To highlight those areas (marked in red) that are underperforming and requiring attention
- To allow trend analysis for a service or engagement to see if practices are improving over time.

It was also noted that it is important to ensure that the measurements and records that are being captured are meaningful and can be used in an effective and timely manner. An analogy was given to the Grayrigg derailment in 2007 where, it is argued, that video footage from a measurement train could have been used to detect the faulty points, but there was too much data being captured to make this practicable.

Risk Matrices

The use of risk matrices was identified as a problematic area with a number of issues identified:

- There can be many different and conflicting matrices in play such as those for safety, security and financial aspects and it is challenging to get stakeholder agreement on priorities across all of these;
- There are multiple pathways through hazards so for example, should a "worst credible" assessment be favoured over considerations of those that are "most likely"?

It was suggested that a combined risk matrix covering all areas of the organisation's risk might be a way of reconciling the first point.

Managing Change

It was agreed that an effective Safety Management System not only has to support the management of current perceived risks, but also to facilitate and manage changes however this might occur. For example, new features and modifications of a supply need just as much safety assurance as the original build, but this assurance may be done by different staff in different roles as a project transitions from say, build to service.

Artificial Intelligence

The session concluded with a discussion on how Safety Management Systems could be used to manage risks in projects that have Artificial Intelligence and Machine Learning aspects. A number of challenges were identified:

- AI systems can adapt in real-time; so how can a safety management system be applied to a system that is changing itself?
- AI removes people from being in direct control so the way humans interact with the technology will need to change, possibly in novel ways;
- In systems of systems context, one system could learn "bad habits" from another;
- Innovation is so rapid at the moment that regulation is struggling to keep pace.

Ideas such as having a "fail fast" approach and treating data as a separate consideration were proposed, but the group agreed that this was a challenging area that would require further discussion.

Conclusion

Despite the diversity of sectors that the group represented, it is clear that many of the problems with establishing an effective Safety Management Systems are common. For example, issues with resourcing, management buy-in, effort prioritisation and apportionment of responsibilities, were all issues echoed throughout the group. There were however, useful examples of best practice that were shared, such as the use of an SMS dashboard for trend analysis, that the group could take back to their respective organisations for consideration.

It is apparent that the discipline is now moving into new challenging and unexplored areas with the advent of Artificial Intelligence and Machine Learning technologies. There is a clear desire for Safety Management Systems to be able to cater for these technologies but the evolutionary pathway to achieving this is uncertain.

Report by Paul Hampton, SCSC Newsletter Editor

SCSC Seminar
Data Safety Evolution

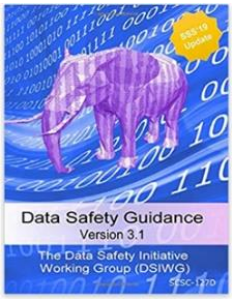
The SCSC held its third Seminar on Data Safety in London on the 14th November 2019. Presentations were given by a number of experts in the field discussing recent developments in the practices and methods in managing the safety risks association with data.

This was a one-day seminar held in the Westminster Park Plaza Hotel in London led by Mike Parsons with individual presentations being hosted by Dave Banham. Approximately 20 delegates attended from a wide range of industries including Aviation, Defence, Healthcare, and Rail. This was the third seminar on this topic, with previous events being held in 2012 (*'How to Stop Data Causing Harm'* scsc.uk/e209) and 2015 (*'How to Stop Data Causing Harm: What you need to know'* scsc.uk/e343).

Why Is Data a Safety Problem?

Mike Parsons open the seminar with a recap of why data safety is important and why it needs to be recognised separately from hardware and software considerations.

Mike presented some historical accidents across several sectors where data has been a contributing cause. This included the recent Boeing 737 MAX accidents where data from a faulty angle of attack sensor led to unintended actuation of the elevator trim that ultimately led to the crashes.



Mike then covered the work of the Data Safety Initiative Working Group (DSIWG) formed in 2013, which has culminated in the publication of a guidance document – Data Safety Guidance (see scsc.uk/scsc-127D and www.amazon.co.uk/gp/product/1793375763). He also covered other DSIWG initiatives currently underway such as the development of an ontology for risk led by David Banham, and the development of tooling to support the implementation of the Data Safety Guidance led by Divya and Martin Atkins (see below).

Introduction to the SCSC Data Safety Guidance

Divya Atkins then presented an introduction to the Data Safety Guidance for those not familiar with its structure and content. Divya described how the document is split into Normative, Informative and Discursive content: Normative covering the formal specification, Informative describing means of compliance and Discursive providing additional information such as list of accidents.

Divya then explained the role of the overarching 4+1 Data Safety Principles that mirror those in software assurance domains and how the sections of the guidance fulfil these principles.

Principle 1 Data Safety Requirements shall be defined to address the data contribution to system hazards

Principle 2 The intent of the Data Safety Requirements shall be maintained throughout requirements decomposition

Principle 3 Data Safety Requirements shall be satisfied

Principle 4 Hazardous system behaviour arising from the system's use of data shall be identified and mitigated

Principle 4+1 The confidence established in addressing the Data Safety Assurance Principles shall be commensurate to the contribution of the data to system risk

She then covered the actual process steps, which are structured around ISO31000 - an established risk management standard. In this process, data safety risks are managed by:

- establishing the context
- identifying data safety risks
- analysing and categorising risks by a Data Safety Assurance Level (DSAL)
- evaluating the acceptability of those risk and applying risk control measures.

The first step can be facilitated through the use of an Organisational Data Risk (ODR) Safety Assessment Form. This is a short questionnaire style form that allows an organisation to quickly establish their overall data safety risk exposure by answering 8 questions. The scores for the individual questions are then summed to give an overall ODR value ranging from ODR0 to ODR4 indicating the overall risk exposure from very low to high.

ORGANISATIONAL DATA RISK LEVEL	
Record the total score and use it to determine the ODR level based on the ranges given below. If the first 3 questions' scores sum up to 6 or less then disregard the scores for the remaining questions.	
Score 14 or less	ODR0
Score 15 to 21	ODR1
Score 22 to 37	ODR2
Score 38 to 47	ODR3
Score 48 and above	ODR4
Total Score for this scenario/context:	
ODR Level for this scenario/context:	

DSALs are a new concept created to support the Data Safety Guidance. The guidance uses a traditional risk matrix to categorise data safety risks into DSALs. The DSALs apply to a Data Artefact, which is the item or collections of items of data under concern.

Severity	Likelihood		
	High	Medium	Low
Minor	DSAL1	DSAL0	DSAL0
Moderate	DSAL2	DSAL1	DSAL0
Significant	DSAL3	DSAL2	DSAL1
Major	DSAL4	DSAL3	DSAL2
Catastrophic	DSAL4	DSAL4	DSAL3

Technique	Data Types	DSAL 1	2	3	4	Notes	Data Property
Built-in-Test / Built-in-Test Equipment (BIT / BITE)	..D..	-	R	HR	HR	Application tests the data (e.g., at start-up or when requested by an operator).	IC.....V.....
Cyclic / Continuous BIT	..D..	-	-	R	HR	Application applies tests to the data it is processing continuously (e.g., for a live data stream) or periodically (e.g., every nth message, every hour).	IC.Y.....VL.....
Backward recovery	..D..	R	R	HR	HR	If a fault in data has been detected, the system resets to an earlier internal data set, which has been proven consistent.	IC.....
Parity Checks	..D..	R	R	HR	HR	Within data, e.g., Hamming codes, Reed-Solomon, Hagelbarger.	I.....
Automatic Error Correction	..D..	R	R	HR	HR	Detected errors are corrected automatically.	IC.....

...

A Data Artefact will have a Data Type (Dynamic, Verification, etc.) and the DSAL will relate to the loss of a particular property of that Artefact, such as loss Integrity, Accuracy, Timeliness, Continuity, etc.

The combination of these 3 attributes: Data Type, Property and DSAL are then used to key into several tables of recommended methods and techniques to mitigate those data safety risks.

The process is illustrated through a worked example in the guidance document, relating to a Healthcare case study. The presentation concluded with Divya discussing up-coming areas of work, especially the Data Safety Tooling development, which is supported by a grant from the Lloyds Register Foundation and covered by Martin Atkins later in the seminar.

Data safety - doing it for real

Mark Templeton then presented his own experiences of managing data safety within the military aviation sector. Mark first described his earliest encounters with the data safety problem; he gave an example of SIL4 safety critical systems that communicated over a databus, but, at the time, there were no satisfactory established methods or guidance on how to assure the data exchanges between the systems.

He then went on to describe his first experiences of using the Data Safety Guidance to support an airworthiness case for an Unmanned Aerial Vehicle (UAV). He noted that the process was time-consuming and took a lot of hard work, but the discipline of logical examination of data flows, review of mitigations and criticality did lead to a demonstration that

controls were adequate.

Mark said he had then written a Data Safety training course and given this to around 40 people in 4 sessions and this has generated useful feedback. This feedback was used to improve the guidance process definitions and a revalidation of the previous UAV proved to be much easier with this additional clarity. The course feedback and the exercise have now been fed into the latest guidance (v3.1).

Mark then went on to cover another case study – a battle-space scenario with both air and ground actors. In this scenario, almost all of the guidance was applied with about 2 man-weeks of effort. Although some aspects proved useful, such as the data HAZOP guidance, he found that the derived requirements were overwhelming, although this was thought to be a symptom of working at the wrong level.

Mark concluded the presentation with the following observations:

- Techniques within the Guidance are effective
- Data safety HAZOP is particularly good
- The process led to unexpected issues

However:

- The "method" works, but the methods and techniques tables need extending for specific domains
- Manual use of the tables on non-trivial examples can be onerous
- Now needs wider usage and feedback to DSIWG

Applying the SCSC Data Safety Guidance: Practical Considerations

Paul Hampton discussed some practical considerations in applying the guidance in two case studies. The first was to elaborate the Healthcare example that is already in the Data Safety Guidance, and to apply the guidance to derive actual data safety requirements. Paul showed a summary of his working as he followed the guidance to arrive at the final requirement set. Paul noted that, as there were many interconnected systems, Data Artefacts were chosen to align with the individual data-flows leading to in excess of 20 Data Artefacts to assess.

His overall conclusions from the work were:

- Overall the process arrived at a reasonable set of 17 requirements
- The process was only manageable by simplifying the data safety properties under consideration
- Even then, there were many Data Artefacts to consider
- Tooling is essential

In his second case study, he assessed the command and control (C2) link for a commercial Remotely Piloted Aircraft System (RPAS) being used to conduct linear inspections of infrastructure such as electricity pylons.

(Image © Alpha Unmanned Systems)

The analysis was from the perspective of the C2 link provider alone, and so when selecting the Data Artefact, he concluded that there was only one - this being the link itself, as the link provider as no knowledge of the data actually being transmitted across the link.

His overall conclusions from this case study were:

- The process was relatively quick, although there was only one Data Artefact
- Effort spent on copying/pasting from guidance - Tooling is essential
- 60% of recommendations not applicable for this Use Case
- Final derived requirements were a subset of those actually implemented but were not exhaustive - suffers from lack of context
- There is no requirement in the guidance to dictate the level of rigour to be applied in meeting a recommendation

He concluded with some open questions on whether the methods/technique are too data-repository centric and whether more method/techniques for these cases are required.

After the 2 case studies, Paul went on to discuss the role of Organisational Data Risk (ODR) assessments. Although the form was originally intended at the bidding stage, to give a high-level assessment of the data safety risk, it is now seen as being useful in providing more than just raising awareness. Paul then explored its use in helping an organisation tailor its approach to managing data safety risks and how it can complement processes, even in regulated environments such as Aviation. He then concluded with some open questions on whether the ODR should be used to define the level of assurance rigour to apply, and its relationship to existing standards that have assurance levels (SILs/DALs/ASILs etc).

Paul finally looked at the correlation between DSALs and Standards Assurance Levels. In this section he assessed the role of DSALs in Healthcare, where there are standards but no explicit assurance levels, and Aviation where there are standards with Assurance Levels (e.g. DALs). Paul concluded that:

- DSAL severities and likelihoods can be aligned with existing standards
- The DSALs can inform on recommended techniques and methods
- Where a standard has Assurance Levels then DSALs can inform the risk assessment but is independent of the Level
- Overall Data Safety Guidance seems to augment and inform existing standards

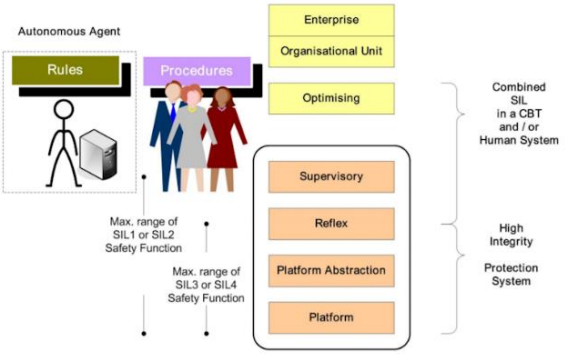
Data guidance in defining development of hydrographic data

Dale Callicott presented his view of safety-related data issues that can occur with hydrographic data; this is the data that underpins the maps and navigation systems used by seafaring vessels. He noted that while paper charts could be verified by the Hydrographic Office once produced, this is no longer the case where hydrographic data is being provided to 3rd party suppliers of navigation equipment with electronic map displays.

Dale gave several examples of accidents where hazards such as islands and underwater wrecks were not displayed on the navigation equipment due to issues such as scaling and overlapping icons.

Exploring the Data Safety Model (DSM) as part of the Assurance Solution

Alastair Faulkner discussed the use of a Data Safety Model (DSM) as part of an assurance solution. Alastair described the architecture of an Information System and then discussed some of the data safety challenges, introducing the concept of strong-data (from sensors) and weak-data (derived from other sources such as data analytics). Alastair then outlined how the DSM could be applied using a case study involving the autonomous flight of a drone taking off from an airport and transiting different airspaces.



He described the key steps of the process as the identification of:

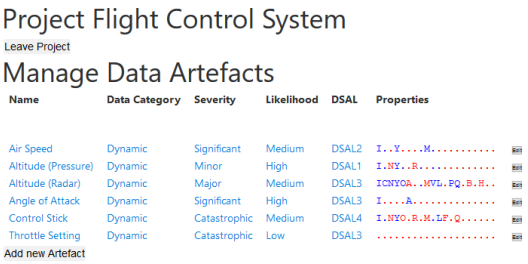
- the context, its hierarchy and the actors in the system
- enterprises and organisations and their respective boundaries within the context
- constituent systems (document in the respective system definitions)
- Interface Agreements (including their description and documentation)
- actors, their identities, authentication and authorities
- all information flows within the context and its hierarchy

Throughout, Alastair, highlighted important areas of concern with what he called "Scary Monsters"; these indicated significant and potentially sizable areas of uncertainty where there are no immediately obvious solutions. For example, the methods by which an organisation would investigate a data-safety accident – how can an investigator prove or repudiate a data supply chain stakeholder's contribution to an accident?

The DSM will be covered in more detail in Alastair's forthcoming book, which he is writing in collaboration with Mark Nicholson of the University of York: *"Data-Centric Safety: Challenges, Opportunities and Incident Investigation"*

Data Safety Tooling

Martin Atkins concluded the event by presenting the progress to date in the development of a Data Safety Tool to support the implementation of the Data Safety Guidance. The Tool prototyping phases are funded by a grant from Lloyds Register Foundation. The tool is web based, platform agnostic and supports multiple users.



Martin demonstrated a prototype of the system and ran through real-life examples of how the tool could be used, such as the failed angle of attack sensor in the recent Boeing 737 MAX accident.

He also illustrated the lifecycle and processing context in which the tool would be used.

Report by Paul Hampton, SCSC Newsletter Editor

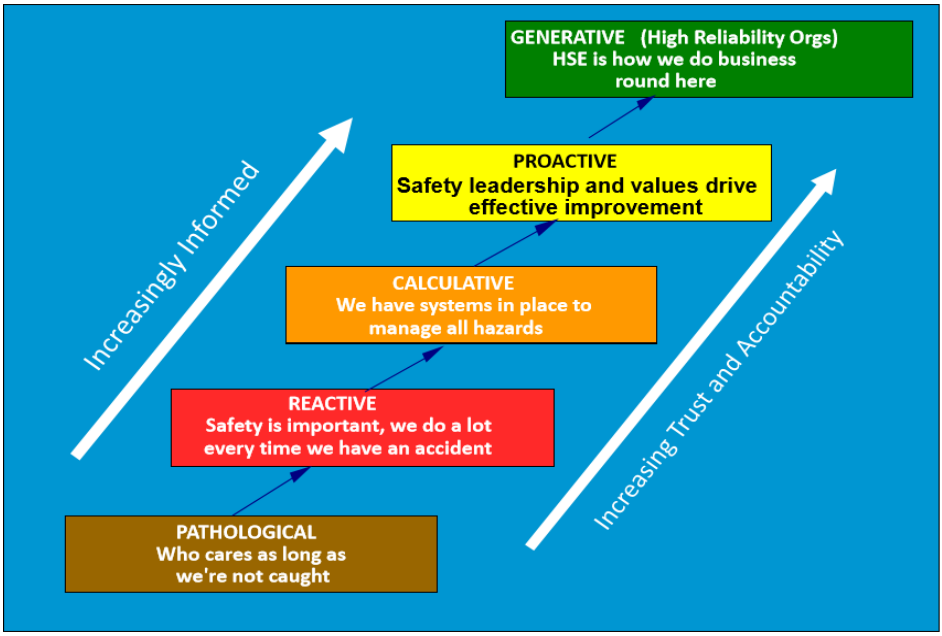
SCSC Seminar
Creating and Maintaining Effective Safety Culture

The SCSC held a Seminar on Creating and Maintaining an Effective Safety Culture on 5th December 2019. The event featured seven speakers from diverse sectors explaining how to create, assess, develop and maintain a strong and effective organisational safety culture — with the recognition that different industries may require different approaches depending on the safety risks they face, the regulatory environment and the staff they employ.

This was a one day seminar held at the DoubleTree by Hilton Hotel, London West End. There were approximately 40 delegates attending from a range of industries including Aviation, Defence, Nuclear, Healthcare, Rail and Utilities.

Snakes and Ladders:
Climbing Up the Culture Ladder Without Falling Off

Patrick Hudson, Professor of Human Factors in Safety at the Delft University of Technology, gave an overview of how safety culture has evolved in organisations and identified some examples of organisational causes to high profile safety incidents across several industries, such as reliance on past success and barriers to communication.

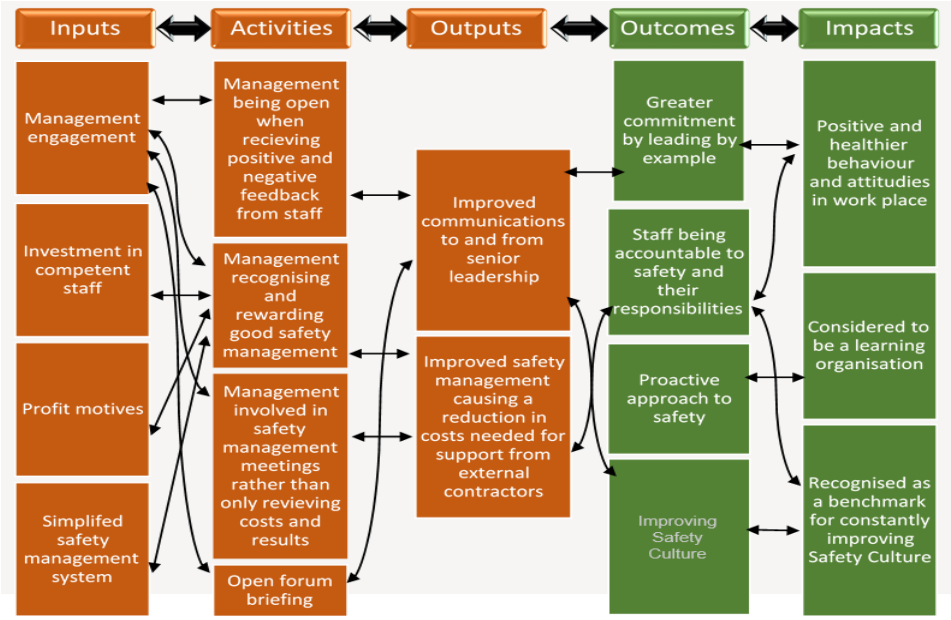


He reviewed conventional approaches involved in the measurement and development of safety culture, such as surveys, interviews and incident analyses. He also gave some examples of questions that might typically be asked in a culture survey, how the responses might be scored and then mapped to a ladder of development levels or safety culture indicators, progressing from Pathological to Generative. He explained how surveys of this nature are often phrased in terms of Generative indicators and are subject to measurement bias.

The talk considered an alternative approach where mappings to the development levels are based on rich descriptions of distinct observable characteristics, rather than quantitative scoring of survey responses. This approach may be preferable as it can be simpler to infer values, beliefs and attitudes based on observed behaviour, rather than predict behaviour based on known values, beliefs and attitudes. The talk concluded by proposing that organisations implement activities, processes and systems to move up the ladder, using lessons learnt from organisations that have fallen down the ladder.

Corn Flakes and Safety Culture

Elizabeth Jacob of SNC-Lavalin Atkins presented a process methodology used for assessing and developing safety culture, which includes the use of a logic flow analysis tool for depicting planned work (in terms of business inputs, activities and outputs) and intended results (in terms of expected organisational outcomes and impacts):



The process methodology is adapted from a change management system developed by the Kellogg Foundation and consists of four steps:

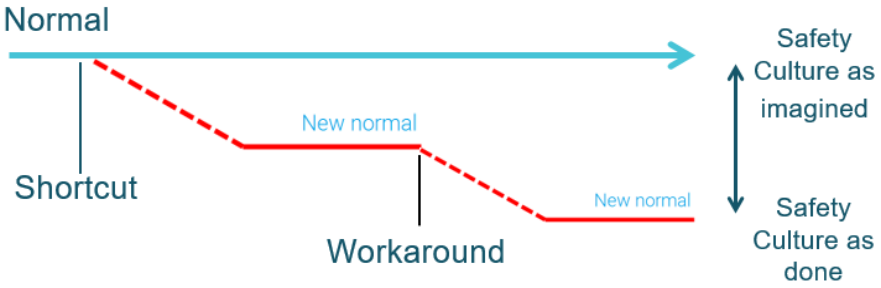
- Understand the processes which lead to the current safety culture outcomes
- Develop a set of desired safety culture outcomes which meet the organisational aim, and needs of staff and service users
- Construct and implement a sequence of events connecting actions to the outcomes desired
- Respond to learning from experience and external factors to continually improve

This approach is simple to implement and can increase clarity around the contributors to safety culture, required business resources, and potential business improvements.

The Wrong Culture Kills

Nick Flynn of National Air Traffic Services (NATS) stated that safety culture is embodied in the attitudes, actions and behaviours of staff and has to be owned and maintained by the organisation. Shortcuts and workarounds in normal working practices (e.g. caused by indifference, complacence and poor attitudes) can lead to a degraded organisational safety culture, which can increase the likelihood of accidents. Degradations in culture can build up over time and must be identified and managed.

Accidents are not typically caused by a single broken component. Employees deviate from normal working practices because they operate in complex and pressurised work environments with limited resources and conflicting business goals. To illustrate this point, the skills and competency requirements for healthcare practitioners were reviewed and correlations identified between the behaviour of surgeons and patient fatality rates. Nick emphasised the need to focus on actual staff behaviour over process in developing an effective organisational safety culture.



Fostering a Safe Culture – Approaches from GB Rail

Paul Leach of the Rail Safety and Standards Board (RSSB) gave an overview of Britain's railway system, its safety record, organisations involved in rail safety leadership and the benefits of developing a "fair culture". The RSSB's safety culture model identifies four main elements linked to organisational factors:

Main Element	Organisational Factors
Effective and appropriate safety management systems	• Barriers & Influences • Training • Communications
Demonstrable management commitment to Health and Safety (senior & line management)	• Organisational Commitment • Management Commitment • Supervisor's role
Participation, involvement and workforce attitudes to Health and Safety	• Personal role • Workmates influence • Risk taking behaviours • Employee Participation
Organisational learning and continuous improvement	• Organisational learning

The Rail Accident Investigation Branch (RAIB) considers the contribution of organisational factors to accidents. Causal factors can include issues such as staff competence, individual and group behaviours, local workforce factors, leadership, the safety management system (i.e. system mitigations), and the wider organisational culture.

The RSSB aims to promote a culture which accepts that staff will make mistakes commensurate with their experience and training, that incidents may be caused by a combination of human performance issues and system failures, and where staff feel able to report unsafe behaviours without fear of punishment. The talk also outlined non-technical skill requirements for railway staff and a method for assessing their behaviour.

Creating and Maintaining an Effective Safety Culture

Wood Nuclear designs, builds and maintains nuclear assets for its energy clients. Chris Gazard and Paul Gaynon gave an overview of Wood's safety culture expectations, health and safety management systems, and examples of their implementation. Wood has introduced a Safety Shield system to encourage preparation, engagement and intervention from staff. It has also identified Safety Essentials to raise awareness of the behaviours required of staff to prevent incidents. A Stop Work Authority may be authorised to ensure intervention and stop any activities considered unsafe.

Key to Wood's corporate safety management system is the oneCLICK Dashboard which provides staff with access to tools, reports and training via a single screen.

Wood's Independent Assurance group provides specialist advice and assessments independent from designers, operators and Safety Case authors. This group comprises Suitably Qualified and Experienced Personnel specialising in nuclear safety and other regulated industries.

Safety Culture Assessments and Improvement Strategy: Lessons from Defence and Elsewhere

Greenstreet Berman provides safety culture and Human Factors consultancy services to its clients. Its Director, Michael Wright, described a toolkit of tried and tested methods for assessing safety culture in Defence and other organisations.

The toolkit comprises a range of techniques including: surveys and workshops, observation and audits, and behavioural indicators. These are directed at assessing Elements of Safety Culture and the outputs may be triangulated and weighted so, for example, the results are not solely determined by subjective survey alone.

The talk identified benefits associated with corporate assessments including: benchmarking to a qualitative standard, comparable results and sharing examples of good practice.

It was recommended that assessment be owned at an appropriate organisational level and integrated with other business improvement plans.

It is also important to recognise that different safety cultures may exist across: Acquisition Safety, Product Safety, Occupational Health and Safety, and Operational Safety. The talk concluded by outlining steps towards an evolving model of safety culture based on "psychological safety". That is one whose elements include: a sense of trust, being part of a team, inclusive leadership, purposeful engagement, shared goals, improving the system of work for a common purpose, and no blame.

Safety Culture: Some Lessons from Healthcare

Patrick Waterson, a Reader in Human Factors and Complex Systems at Loughborough University, gave a talk on the importance of safety culture in healthcare, how it is measured and some lessons from high profile incidents involving NHS Trusts. Since 2004 there have been over 300 studies of Patient Safety Culture. It was suggested that many of these studies are unreliable and supported by weak evidence – e.g. low sample sizes, exclusive use of surveys rather than workshops and interviews, responses limited to a specific staff group.

Hospital Survey on Patient Safety Culture (Original Version)

Overall perceptions of safety "Patient safety is never sacrificed to get the work done"	**Frequency of error reporting** "When a mistake is made, but is caught and corrected before affecting the patient, how often is this reported?"	**Supervisor/manager expectations** "My supervisor/manager overlooks patient safety problems that happen over and over"
Organisational learning "We are actively doing things to improve patient safety"	**Teamwork within units** "People support one another in this unit"	**Communication openness** "Staff will freely speak up if they see something that may negatively affect patient safety"
Feedback/Communication "We are give feedback about changes put into place based on event reports"	**Nonpunitive responses to error** "Staff feel like their mistakes are held against them"	**Staffing** "We have enough staff to handle the workload"
Management support "The actions of hospital management show that patient safety is a top priority"	**Teamwork across units** "There is good cooperation among hospital units that need to work together"	**Handoffs and transitions** "Things "fall between the cracks" when transferring patients from one unit to another"

It was suggested that the assessment of safety culture in healthcare in general has a great deal to learn from other industries – e.g. use of team based reflection and education, self-assessments and safety maturity models. Some possible tools were outlined: Hearts and Minds toolkit used in Oil and Gas, Eurocontrol's Safety Discussion Cards used in Air Traffic Management, and the Oxford NOTECHS tool used for evaluating the non-technical skills of operating theatre staff.

Panel Session

A wrap-up session led by Mark Nicholson of York University gave delegates an opportunity to put further questions to the speakers.

Some key messages from the event were summarised:

- To beware of the accuracy of quantitative measurements of safety culture
- To avoid big leaps and instead, use small steps to achieve the desired level of safety culture
- To recognise that safety culture models are evolving and the five development levels referenced by many of the speakers are also likely to evolve
- To recognise that since an organisation's business areas, staff groups, partners, and operations could all have different safety cultures, the full "culture stack" should be considered

Report by Rick Vinter, Principal Safety Consultant, CGI IT UK

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 scsc.uk 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.

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

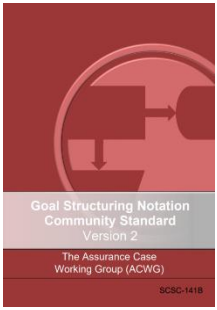
Advertising

Do you have a product, service, event or job vacancy 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 specialist 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:

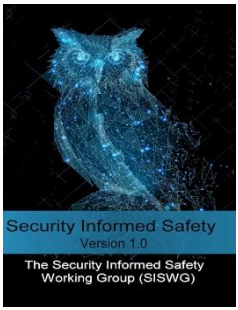
- Be broader than safety, and will address interaction and conflict between related topics
- Address aspects such as proportionality, rationale behind the guidance, focus on risk, confidence and conformance
- Consider the role of the counter-argument and evidence and the treatment of potential bias in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

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

The next meeting is on the 18th May 2020, in London.

Lead Phil Williams phil.williams@scsc.uk

Security Informed Safety



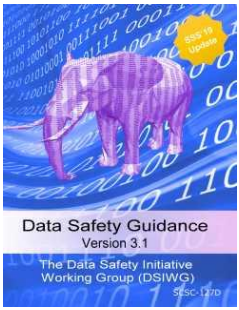
The Security Informed Safety Working Group (SISWG) aims to capture cross-domain best practice to help engineers find the 'wood through the trees' with all the different security standards, their implication and integration with safety design principles to aid the design and protection of secure safety-critical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

SCSC Working Groups

Data Safety Initiative

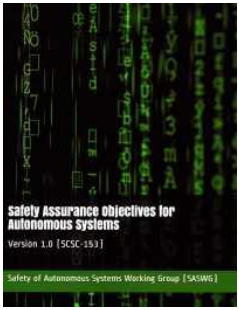


Data in safety related systems is not currently sufficiently addressed in current safety management practices and standards. It is acknowledged that data has been a contributing factor in several incidents to date. There are clear business and societal benefits, in terms of reduced harm, reduced commercial 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. The next meeting is on the 17th March 2020, in Bath.

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. The next meeting is on the 11th March 2020, in Bristol.

Lead Rob Alexander rob.alexander@scsc.uk

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. The next meeting is on the 24th March 2020, in London.

Lead Mike Parsons mike.parsons@scsc.uk

60 Seconds with…Tim Kelly

Tim has worked for over 25 years in the domain of high-integrity and safety-critical systems engineering. As full-time Professor in High Integrity Systems at the University of York, he has been at the forefront of the development of best practices in the field, and he has published over 150 papers in international journals and conferences. In 2019, following his ordination in the Church of England, he became Assistant Curate of Beverley Minster.

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

I started working (as a sponsored student) in a Rolls-Royce centre in Derby called the High Integrity Systems and Software Centre back in 1991. It was the software engineering angle that had attracted me at first. However, when I started to see the applications of software to safety-critical systems such as the aero-engine FADECs (Full Authority Digital Engine Controllers) I became more interested in the need to make sure that we 'got it right' and the engineering methods required to help develop and assure these systems.

What aspect of your career are you most proud of?

I am proud of my role in developing the Goal Structuring Notation (GSN) and that it has been so widely used by engineers developing safety cases across the world in so many different industries. When used correctly, it has helped many people to be clearer about the nature of the safety arguments that they are making, and the strengths and weaknesses of those arguments.

What advice would you give to yourself age 12?

Stick with the computers Tim. You'll have a lot of fun!

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

In the same vein as Haddon-Cave's concerns in the Nimrod Report, I worry about the commoditisation of system safety, i.e. it becoming something we pay others to do for us rather than it involving the significant intellectual input of design and operating authorities.

"Stick with the computers Tim. You'll have a lot of fun!"

What's your most favourite quote or motto?

"The Devil is in the Detail"

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

Play the guitar!

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

Harry Palmer (from the John le Carré novels).

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

When doing my PhD - "You've got to start with an idea that seems simple, because it probably isn't!"

The SCSC Steering Group

Tom Anderson

Honorary member

Stephen Bull

stephen.bull@scsc.uk

Jane Fenn

jane.fenn@scsc.uk

Louise Harney

louise.harney@scsc.uk

James Inge

james.inge@scsc.uk

Nikita Johnson

nikita.johnson@scsc.uk

Tim Kelly

Honorary member

Mark Nicholson

mark.nicholson@scsc.uk

Felix Redmill

Honorary member

Phil Williams

phil.williams@scsc.uk

Robin Bloomfield

Honorary member

Dewi Daniels

dewi.daniels@scsc.uk

Paul Hampton

paul.hampton@scsc.uk

Stuart Harrison

stuart.harrison@scsc.uk

Brian Jepson

brian.jepson@scsc.uk

Graham Jolliffe

Honorary member

Alex King

alex.king@scsc.uk

Mike Parsons

mike.parsons@scsc.uk

Roger Rivett

roger.rivett@scsc.uk

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	29	

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

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

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

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

11 March 2020
SCSC Group Meeting

Safety of Autonomous Systems Working Group Meeting

Bristol, UK

scsc.uk/e670

17 March 2020
SCSC Group Meeting

Data Safety Initiative Working Group meeting #51

Bath, UK

scsc.uk/e673

16-20 March 2020
Training Course

Introduction to System Safety Engineering and Management 2020

Brugge, Belgium

iiw.kuleuven.be/brugge/m-group/Events/ssc2020

23-26 March 2020
PSASS Workshop

STAMP Workshop

Massachusetts, USA

psas.scripts.mit.edu/home/stamp-workshops

24 March 2020
SCSC Group Meeting

Service Assurance Working Group meeting #21

London, UK

scsc.uk/e674

26 March 2020
IET Conference

Nuclear engineering for safety, control and security

Bristol, UK

events2.theiet.org/nuclear

30 April 2020
SCSC Seminar

Safe Use of Multi-Core and Manycore Processors

London, UK

scsc.uk/e638

18 May 2020
SCSC Group Meeting

Assurance Cases Working Group Meeting #12

London, UK

scsc.uk/e672

11 June 2020
SCSC Seminar

New Safety Analysis Techniques

London, UK

scsc.uk/e654

16-18 June 2020
VDA Conference

Quality, safety and security for automotive software-based systems

Potsdam, Germany

vda-qmc.de/en/software-processes/vda-automotive-sys

17 September 2020
SCSC Tutorial

Combining Safety With Security

London, UK

scsc.uk/e666

3 December 2020
SCSC Seminar

Management and Oversight of Complex Systems

London, UK

scsc.uk/e661

Call for Participation

SCSC Safety Futures Initiative (SFI)

The SFI aims to bring together young and early-career (YEC) safety professionals to gain experience, share knowledge and to build a community around the specialised interest area of safety engineering and assurance.

Young technologists are often highly trained and very knowledgeable about engineering topics such as programming, systems development, testing, etc. – however, there are very limited opportunities to learn about system safety, which requires a very particular set of skills and a particular mindset in addition to technical aptitude.

The SCSC SFI aims to build a community of like-minded individuals to share ideas, experiences and develop together. Support will be provided to understand the possibilities for a career in safety which is, perhaps, much more of a vocation than careers in other engineering disciplines.

The SCSC SFI will reinforce the idea that the work we are doing really matters, and will help YEC safety professionals to be more responsible, professionally and ethically, in their work and to wider society. In essence, the goal is to nurture the hearts and minds of the next generation of safety professionals.

The target demographic is engineers, students and safety professionals involved in safety systems with less than 3 years' experience. If you are interested, or know of anyone who might be, please contact Nikita Johnson nikita.johnson@scsc.uk

SCSC Ontology Working Group (OWG)

The OWG has been working on a common language for risk management, and has developed a model – an ontology – of terms so that safety and security professionals can share ideas and concepts on risk (see page 33). While initially focussed on data, the OWG is now looking for wider participation from safety and security communities to help advance this exciting area of work. Contact paul.hampton@scsc.uk for further information on the group and how you can get involved.

SCSC Safety Culture Working Group (SCWG)

A new working group is being established to provide guidance on creating and maintaining an effective safety culture (see page 52). Please contact mike.parsons@scsc.uk if you are interested in joining this group.

Data Safety Tooling (DSIWG-TSG)

The Data Safety Tooling Team is seeking support from user organisations to influence the future direction and evolution of the tool currently being prototyped to support the Data Safety Guidance (see page 51), under funding from the Lloyds Register Foundation. Several levels of participation are possible, such as: workshop attendance, provision of test data/scenarios, beta testing, through to membership of the project advisory board. For further details, contact Dr Divya Atkins divya@mca-ltd.com