SAFETY SYSTEMS

The SCSC Newsletter

Winter 2018-2019 Volume 27 Number 1

2019 marks the centenary of three long-forgotten but significant aviation accidents: the Airship N.S.11 explosion (15th July 1919) during a mine-hunting patrol over the North Sea; the crash of the Tarrant Tabor triplane bomber on its maiden flight near Farnborough (26th May 1919); and the structural collapse of the Caproni Ca.48 airliner near Verona (August 2nd 1919). All three aircraft were pushing at technological boundaries: the Tarrant Tabor was the largest aircraft in the world, the Caproni Ca.48 accident was one of the earliest heavier-than-air aircraft crashes, and the N.S.11 held the world distance record for non-rigid airship flight. It is worth pausing for a moment to reflect on these pioneering craft, and their fates, and on the current state of our discipline.

2019 is a time of great anticipation, and also a time when we are faced with considerable challenges. We are privileged to live at a time characterised by the promise of game-changing technologies. For example, the potential applications of autonomy, across many domains, promise vast benefits for society, including, we hope, improved safety. Advances in medical science and engineering offer us hope of effective, personally-tailored treatments, mitigating the effects of cancer and other diseases.

As safety engineers, though, 2019 brings us new challenges, and an enhanced awareness of long-standing challenges which we find it hard to address. Incidents like the Grenfell Tower fire, for example, remind us that engineering projects exist in a complex context of non-engineering issues, which also contribute to the maintenance or reduction of safety. At a time of rapid technological advances, where economics and politics is ever more reliant on engineers to deliver, it is important that we ensure that the concerns of safety engineers are heard, respected and acted upon. Let's keep talking!

Katrina Attwood
newsletter@scsc.uk

In this Issue

p.2 Nick Hales considers some of the "troubling questions" around autonomy

p.8 Paul Hampton and Mark Thomas discuss the application of the data safety guidance to clinical risk management

p.15 John Ridgway ponders the safest way to kill

p.19 Ken Nathan and his colleagues look at the future of MBSE

p.26 Ludmila Musalova explores connections between employee well-being and safety

p.32 Events Diary

Autonomy's Troubling Questions

by Nick Hales

I was asked at a recent SCSC seminar whether I thought standards should be maintained in the face of the pressure by manufacturers to reduce standards so that they could compete, in the small aircraft business, with developing economies. My response at the time was, "Oh yes! It's the duty of every civilisation to raise others to their standards". This implies that we must persuade developing economies to adopt our rigid safety standards, but I realised later the statement is also fraught with the need for caveats. To use an orthogonal example, an environmental activist such as myself may say, "but I should add that in the case of primitive tribes such as the Andaman island dwellers or remote tribes in the Amazon, the standards we raise them to should be simple, such as providing free medical care where requested, leaving them to live as close to harmony with their wilderness homes as they wish". It is often true that we should make allowances and see the downsides of spreading a civilisation, or in the case of this essay, an ideology - autonomy - before we promote ourselves or impose.

This essay examines the ethics of autonomy as a developing technology. We must decide what is best of our safety-oriented behaviours to impose on the future of that technology without interfering in its purpose, i.e. to relieve mankind of the burden of labour. The particular focus of my discussion is lethal robots, so for the purpose of the argument the underlying principle is that it is morally permissible to resist aggression via force to achieve peace, the Falklands war being considered by the author an ethically reasonable example.

An Unbreakable Principle of Autonomy

In their joint work 'Formal Verification of Ethical Choices in an Autonomous System' [1], Dennis et al state what I will call the *Unbreakable Principle of Autonomy*: that "All participants in society are required to follow specific regulations and laws. An autonomous system cannot be an exception." Of course, many societies exist on earth and they follow diverse regulations. One can lose one's head for adultery in some countries but just one's partner in the UK. It is therefore essential to tie rules on autonomy to internationally-recognised standards such as the UN's *Universal Declaration of Human Rights*, thus protecting people from autonomous vehicles which may be used to try to deny those rights.

Among the Oxford Martin School's recommendations in the policy paper, 'Robo-Wars: The Regulation of Robotic Weapons' [2] is the following: "Ensure compliance with legal and ethical frameworks, guaranteeing that they are taken into account at every stage of the design process." That illustrates again that mainstream thinking considers current frameworks, such as the Declaration of Human Rights, as essential to create a code for autonomy, the framework, rules and regulations or laws for which are not yet clear.

One's initial thoughts on autonomy may be to consider some form of Formal Logic to control behaviour, as Formal Mathematics is used in the development of the design of safety critical software systems. However, one may also have to consider whether developments in Artificial Intelligence will enable future machines to act independently as moral entities. For some people, this brings to mind the frightening prospect of the *Terminator* films, where an interconnected system

of systems becomes 'self-aware' and then programs, via a system akin to the already emerging internet of things, all other connected machines to destroy mankind as the only threat it faces. While such may sound fanciful, the question arises, "As far as we can estimate it, what is the likelihood of this type of scenario?"

To illustrate how incremental increases in capability can create unexpected outcomes, it is only necessary to consider evolution. Meteorites have been found with amino acids preserved inside, even after falling through the earth's atmosphere and experiencing extremely high temperatures on their surface. Over a century ago, if one asked a group of people, "Do you believe that the fundamental building blocks of life can exist in rocks in space, and if so, are your notions so fanciful that you believe a rock could break up on impact and those building blocks (amino acids as we now know them), falling into a warm ocean or pool of water, could develop into a life-form as complex and capable of making ethical or non-ethical choices as yourself?", all or nearly all would have said "No". But that looks like a possible route for the beginning of our evolution. Therefore, one could conclude that once the reasoning of a computer connected to the Internet of Things becomes so fast and innovative that we cannot fully comprehend at any one instant of time how it arrived at its current state, or whether it is developing its own sub-routines to perhaps defend itself from viruses, worms and Trojan horses etc., then there must be a perhaps significant probability that such a similarly inanimate building block could develop consciousness. We cannot escape this by believing that humans are differentiated from the rest of nature – or indeed of matter – by having 'souls' or something equal, since observation of animal behaviour since the 1960s leads us to the conclusion that, in some cases, they too have some form of ethics, reasons for conflict over resources, resolution of differences capacity, friendship, identification of enemies etc. The possibility that apparently inanimate rocks from outer space contain amino acids which are responsible for life on earth, illustrates how things can develop from next to nothing, biologically speaking.

Ethical Decisions By Machines

Continuing with the point made in the paragraph above, self-protection by an autonomous vehicle could be permitted until a decision is to be made which will affect a human life, or upon which a human life may depend. At that point, the decision the machine makes must be identical to that which a consensus of humanity would make in that position but excluding the legal principle of 'capacity' since the machine cannot be claimed to have capacity. The principle of capacity enables, for instance, a decision to be made about whether the life of a mother is more important than the life of a child she bears, by adults with capacity but not the unborn child. For instance, abortion is often justified as being within the woman's capacity, so she takes (moral) precedence, but in Roman Law, as a last resort, the child could be removed from a dead or dying mother: the woman of course would then die but the child live - the child took precedence. (The birth of Caesar this way is almost certainly a myth since his mother is reported to have lived long after his birth, an impossibility then). However, in the argument here, the survival of the machine is irrelevant, since it is non-human and therefore has no capacity. Only the non-combatants can be counted as having capacity which the machine should consider. Since the mission obviously has a high priority and may involve the death of opposition combatants, the machine would not be required to consider the capacity of those combatants, since their involvement in a

combat scenario effectively removes from them the capacity they would have in normal everyday life. The precedence rule therefore applies only to those who would be affected by a bad decision and are not opposition combatants. It effectively preserves the sanctity of life and possibly the environment if one takes the principle of 'Ecocide' (Higgins [3]) on board. The machine may sacrifice itself or terminate its mission to preserve the life of its masters, allies and non-combatants.

The question could then become, "Is there an intention to consider non-combatants as collateral damage?" If there is, then the machine, it is suggested, would have to be fully aware of the complete environment to test whether or not a 'shot' could be taken a few minutes later when there are fewer non-combatants in danger. It would also need to be able to calculate the numbers of non-combatants likely to be affected as collateral damage, since it will clearly be more likely to be acceptable to the public if four non-combatants got caught up in an explosion than if a crowded market place is hit and 200 innocents die in the explosion and any subsequent conflagration which rages through the market. This will be so difficult that it is considered unlikely ever to be permitted. The difficulty is that the decision as to whether to attack when the death of non-combatants is inevitable is, in current circumstances, a command decision passed to a combatant, (a classic example being the decision of Commander Dick of the London Metropolitan Police to permit termination of what proved to be the life of the innocent Jean Charles de Menezes on the London Underground in 2005). Thus, if a machine was given the power to make that decision then the machine – rather than the front-line soldier - would effectively become the Commander, and by the primary rule above a machine must not be allowed to become an exception to the principle that "all participants in society are required to follow specific regulations and laws". By becoming a front-line soldier with high command control it would be indeed doing that.

The need would be for the unmanned system to have recognition systems which allow it to perform various judgemental functions without human intervention, i.e. to know the difference between combatants and non-combatants, to calculate the numbers involved and to be able to predict whether or not a more opportune moment is likely to arise. At the beginning of the James Bond movie, *Skyfall*, M orders an agent to take a shot, knowing James Bond could be hit, simply because it is the best chance, as she sees it, to kill the enemy agent Bond is in close combat with. It is argued here that society cannot ask a robot to play the role of M in similar scenarios. In any case, it could be argued that M made the wrong decision, in which case programmers could get blamed for acts which commanders would have been responsible for, freeing commanders from guilt but making system designers culpable. 'Friendly Fire' casualties due to poor robot engineering will be unacceptable to the public.

Ethics, Cyber and Stealth

An ethical autonomous system would need to be able to discern what choices it has and make the best decision. If no option is available which fulfils its programmed 'ethical' criteria - if indeed a program can be designed to answer the questions posed by this complex and context specific problem - it is thought by some that it may then examine options in order of least "unethicalness" and select the best. But how that could be achieved is open to question. How can safety be maintained if security is not so powerful that guarantees that the system cannot be penetrated can be provably factual? Option

programmes may be the first point of attack by cyber agents to turn the machine against its masters. Also, what would happen if given the above, the least unethical options available is something completely unacceptable to humanity, despite other options being worse? The machine must clearly also be capable of knowing what the public would accept, given the risk that its priority target stack could be altered by an enemy.

A large international company involved in the development of an advanced robotic weapons has already made it clear that, with regard to unmanned aircraft systems, they believe and design to the principle that there will always be a need for a human in the loop, in particular regarding any use of weapons, both now and in the future. It remains to be clarified at what stage of the loop exactly the operator is involved. This was also the opinion of an international group of safety-aware engineers meeting in 2015 near Washington which the author attended. However, in case politicians change their minds, it is certain that many companies' future-proofing of unmanned systems technology will involve the inclusion of options for complete autonomy. In flight, for example for evading missile attacks, this is perfectly acceptable, since it would be impossible for a remote human operator to manoeuvre the vehicle in a sufficiently agile manner to achieve such objective-related mission critical activities. However, in a dog fight, could there be certainty the UAS would not attack its own manned platforms in the confusion? It would therefore appear that stealth to evade such combat problems would be a preferred option to autonomous combat. In-between options exist, where a human operator may be actively involved in particular steps of the targeting process while other steps are left to the robot, once it has been programmed accordingly. But if control and communication are lost, the intended intervention may not take place.

One important factor is that some extreme politicians may, simply to gain popular support, insist on ethical controls being lifted. This would imply over-ruling all engineering and corporate advice and demanding fully autonomous systems, saying damn to the risks and the potential for unwelcome tit-for-tat responses by less scrupulous regimes, instead of pursuing international agreement through the auspices of organisations like the UN.

Three Principles

Robo-Wars [2] cites three principles for developers, users and regulators: "Recommendations for States","Recommendations for manufacturers and the military" and "Recommendations for States and the military".

The principle for regulators or government has been dealt with partly in the discussion on ethical frameworks above. For the users or military/ police forces, the essence of the recommendation



is that requirements should include allocated responsibility and the enabling of over-ride by operators. However, as discussed above, a chicken-and-egg situation arises here, since the requirement must be bound by regulators no matter what the military wants. Beyond these Oxford Martin recommendations, there is also a case for an automatic "return to base" program to be triggered if communication is lost (this principle might also be useful for hijacked planes, by the way, where the pilot immediately on realising a suicide, bomb or hijack attack is underway, presses one single button which initiates a programme that takes the plane to the nearest safe airport automatically and which cannot be over-ridden).

The last of the three, "Recommendations for states and the military" for reasons discussed, must be given priority because otherwise manufacturers may produce whatever anyone requests, even if this lies outside the envelope of developing ethics. It can be summarised thus: "At an international level, develop standards for the deployment of robotic weapons, specifying and delimiting the contexts and purposes for which robotic weapons can be used, preventing use outside legal paradigms provided by international, humanitarian and human rights law", the former of which requires distinction and proportionality in war, the latter of which concerns law enforcement or police activities.

If the three principles can be summed up in one sentence it is this: "Autonomy is used to enhance human decision-making, not substitute for it".

The Lack of Chivalry

Most people will have thought at some point about uprisings, as weapons become more sophisticated, "They may be the losers and I would not want them to win, but one cannot deny their bravery". The problem of autonomous systems is that, as they increasingly become able to operate without involving the personnel of the military force that uses them, the former codes of chivalry (reliant on ideas of loyalty, fair-play and honour) no longer have any meaning. Armed forces may be seen as disabled by their use of robots in the eyes of those who oppose them in asymmetric conflicts. These insurgents and terrorists can capitalise on this model, attracting sympathy where otherwise their cause would be despised. Military doctrine in the west already holds high principles, so though it has been suggested in [2] that "responsibility and integrity" could be virtues ascribed to non-combatant Remotely Piloted Air System operators, where chivalry and bravery cannot. This is unlikely to have much traction, and rotation through the role by former front-line combatants who have already earned their stripes and medals is more likely.

Conclusion

The future can look bleak or rosy, depending on one's state of mind and the information one has received. A rosy future may imply a shared sense of global "civilisation", but would you want it policed by robotic systems, with all the questions of ethics that then arise and the potential for unsafe operation or penetration of the system by cyber-attacks? A bleak future may entail endless conflict against asymmetric forces or even ludicrous virtual total destruction initiated by advanced civilisations, armed with unthinkable power, engaging in relatively trivial disagreements or, worse still, due to cyber-attacks by terrorists on nations otherwise at peace. In any of these events, forethought by engineers in producing safe robotic systems operating under moral paradigms of behaviour, could be the difference between the possibility of peace or total destruction.

I will close by offering one of the most moving and concerning statements given by a scientist recently. Professor Brian Cox, talking about life elsewhere in the Universe, made a comment to the effect that, because we have developed from cavemen to advanced armed civilisations so relatively quickly in astronomical terms, our intellectual power has not caught up with our power to destroy. We are like children operating machine guns. If this has always been the case, wherever intelligence evolves in our universe, it may be that every advanced life form vanishes in an orgy of self-destructive violence soon after evolving and that therefore, despite the billions of possibilities, we are alone in the universe. Earth-based life forms may be doomed never to communicate with intelligent life on other planets because we cannot make the necessary ethical decisions to enable long term survival. This is very sad. If we are to move safely forward to the future we must maintain our humanitarian standards and encourage others to adopt them. Developing new codes of practice for autonomy is therefore essential: future generations of technology, and of humans, depend on them.

References

[1] L. Dennis, M. Fisher, M. Slavkovik and M. Webster 'Formal Verification of Ethical Choices in Autonomous Systems' , Robotics and Autonomous Systems vol 77 (2015), pp. 1-14 http://www.sciencedirect.com/science/article/pii/S0921889015003000

[2] A. Leveringhaus and G. Giacca, *Robo-Wars: The Regulation of Robotic Weapons*. Oxford Martin School, University of Oxford, 2014. http://www.oxfordmartin.ox.ac.uk/news/201412_Robo-Wars

[3] P Higgins http://eradicatingecocide.com/

Until his recent retirement, ***Nick Hales*** *was a Senior Engineer at DE&S, specialising in critical avionics systems.*

PLEASE NOTE: The views expressed in this article represent the fruits of Nick's own thinking, and should not be ascribed in any way to his previous employer, or to UK MOD.

We are pleased to announce a new service for SCSC members and member organisations: SCSC Announcements. From March 2019, we will be happy to take suitably-worded announcements of interest to members (for example, of forthcoming events, publications, calls for tender or vacancies) and advertise them on the SCSC website. The link will be provided in *Safety Systems* and will also be clear from the home page. For further details, please contact Alex King: *alex.king@scsc.uk*



Data Safety Guidance for Clinical Risk Management

by Paul Hampton and Mark Thomas

Health organisations are increasingly using Health IT Systems in the provision of patient care. As with other sectors, the use of IT brings system safety risk and this in turn can have implications on patient safety. For example, if a system records patient allergies incorrectly, this could have dangerous if not fatal outcomes for patients if relied upon when administering medication. These risks have long been recognised by the industry and standards such as SCCI0129 and SCCI0160 have been established to inform the safety management systems of the manufacturers of Health IT System and those health organisations deploying such systems. As with other standards, however, risks associated with hardware and software is the main focus and there has been little or no explicit guidance on the management of safety data.

NHS Digital, in conjunction with the Safety-Critical Systems Club (SCSC) and the Data Safety Initiative Working Group (DSIWG), have therefore now published guidance - *Clinical Risk Management: Data Safety* - specifically to address the management of data in the clinical



healthcare context. This guidance is based on the work of the DSIWG, but has been tailored for the Healthcare context and specifically adapted to augment and complement the existing SCCI standards.

The guidance document is intended to be used in conjunction with the SCCI standards and aligns well to these, as the Data Safety Guidance (on which it is based) and the SCCI standard all follow a similar structure based around the established practices for the management of risk. The document is split into 3 broad sections:

- A general introduction to data safety;
- More detailed guidance on data safety aligned with the requirements of the SCCI standards;
- A structured methodology for assessing the risks.

Part 1: General Introduction to Data Safety

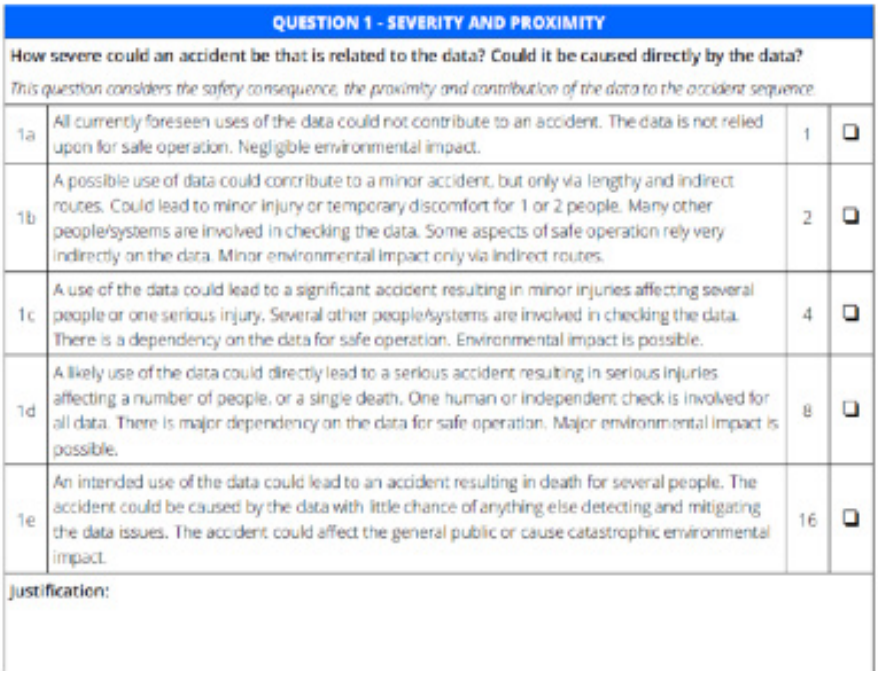
These sections provide an introduction to the subject, discuss the types and properties of safety data, the typical issues that can arise in healthcare settings, and provide a method for assessing the high-level applicability of data safety risks in an organisation.

These sections form an important part of the document as, irrespective of the remainder of the document, they play a significant role in raising awareness of the types of issue that data can give rise to, and as such, can help inform and educate those involved in the clinical risk management processes.

One early step in the risk management process is to establish the data safety context and to assess, at a high level, how much data safety is an issue for an organisation and the overall level of risk they

may be exposed to. This early assessment could then, for example, be used to help determine the scale of effort to invest in managing perceived data safety risks.

This assessment is conducted using an Operational Data Risk (ODR) assessment form. This is a tool developed by the DSIWG, and is structured as a questionnaire. Figure 1 shows an example screenshot from the tool. There are 8 questions in total, each covering different aspects of data safety that might typically contribute to risk assessment, for example: *How severe could an accident that is related to the data be? Could it be caused directly by the data?"* and *"What would be the impact on the organisation, client or public if an accident related to the data occurred?"* Answers are provided as a sliding scale of applicability with each option carrying a score. A total is then generated for all of the responses, and an overall ODR assessment level determined for the entire enterprise.



QUESTION 1 - SEVERITY AND PROXIMITY			
How severe could an accident be that is related to the data? Could it be caused directly by the data? *This question considers the safety consequence, the proximity and contribution of the data to the accident sequence*			
1a	All currently foreseen uses of the data could not contribute to an accident. The data is not relied upon for safe operation. Negligible environmental impact.	1	❑
1b	A possible use of data could contribute to a minor accident, but only via lengthy and indirect routes. Could lead to minor injury or temporary discomfort for 1 or 2 people. Many other people/systems are involved in checking the data. Some aspects of safe operation rely very indirectly on the data. Minor environmental impact only via indirect routes.	2	❑
1c	A use of the data could lead to a significant accident resulting in minor injuries affecting several people or one serious injury. Several other people/systems are involved in checking the data. There is a dependency on the data for safe operation. Environmental impact is possible.	4	❑
1d	A likely use of the data could directly lead to a serious accident resulting in serious injuries affecting a number of people, or a single death. One human or independent check is involved for all data. There is major dependency on the data for safe operation. Major environmental impact is possible.	8	❑
1e	An intended use of the data could lead to an accident resulting in death for several people. The accident could be caused by the data with little chance of anything else detecting and mitigating the data issues. The accident could affect the general public or cause catastrophic environmental impact.	16	❑
Justification:			

Figure 1: Organisational Data Risk Assessment Form

Part 2: Detailed Guidance on Data Safety Aligned with SCCI Standards

The SCCI standards are structured by defining requirements for each stage of the risk management process. The guidance sections provide, for each set of the standards' requirements, an interpretation of the requirement as it applies to data and the typical activities to consider when manging risks associated with data. These sections are described through the use of a worked example that describes a typical scenario where a manufacturer is providing a Health IT System to a health organisation who will be deploying it.

For example, the following requirements are defined in the SCCI standards:

SCCCI0129 Section 4.4.1: For each hazard identified the Maunfacturer MUST estimate, using the critieria specified in the Clinical Risk Management Plan:
- the severity of the hazard
- the likelihood of the hazard
- the resulting clinical risk.

SCCI0160 Section 4.4.1: For each identified hazard the Health Organisation MUST estimate, using the criteia specified in the Clinical Risk Management Plan:
- the severity of the hazard
- the likelihood of the hazard
- the resulting clinical risk.

The guidance document then goes on to describe a method for conducting hazard identification from a data perspective. This, for example, introduces the following dimensions to be considered when thinking about data hazards:

- Proximity: how directly a data failure will lead to an accident;
- Dependency: how dependent the application is on the dataset;
- Detection: the likelihood of being able to detect a data failure prior to an accident;

- Prevention: the ability ofthe system's architect/ developer to guard against errors;
- Correction: the ability of the system to work around or correct errors.

These can then be used in conjunction with a likelihood assessment to determine the overall data contribution to risk, as shown in Table 1 below. The likelihood calculated in this way, along with the assessed severity of the resulting impact, can then be used to determine the categorisation of the risk using a Clinical Risk Matrix, which will have been established as part of the risk assessment of hardware and the software hazard identification process.

We include a worked example to help illustrate the management of data safety using the guidance. The case study details are as follows:

A Manufacturer is building a new integrated health and social care system to support holistic care for community health services. The system supports clinical workflows for aspects such as referrals, tracking clinical encounters, appointment scheduling, outcome measures through to letter and report generation. The system follows a typical development lifecycle as a series of phases: business modelling, requirements, analysis and design, implementation, test and deployment. The system is being targeted to meet the requirements of a Health Organisation who are procuring a solution to help their cilinicians maintain a high level of quality of care

Concern	Very High/High	Medium	Low/Very Low
Proximity	A known use of the data is highly likely to lead to an accident	A possible use of the data could lead to an accident	All currently foreseen uses of the data could lead to harm only via lengthy and indirect routes
Dependency	Data is completely relied on	Data is indirectly relied upon	Little reliance on data
Detection	Low or no chance of anything else detecting an error	Some other people/systems are involved in checking the data	Many other people/systems are involved in checking the data
Prevention	Difficult or impossible to guard/barrier against errors	Possible to guard/ barrier against errors	Easy to guard/barrier against error
Correction	Difficult or impossible to correct or work around er-rors	Possible to correct or work around er-rors	Easy to correct or work around errors

Table 1: Data Hazard Risk Likelihood

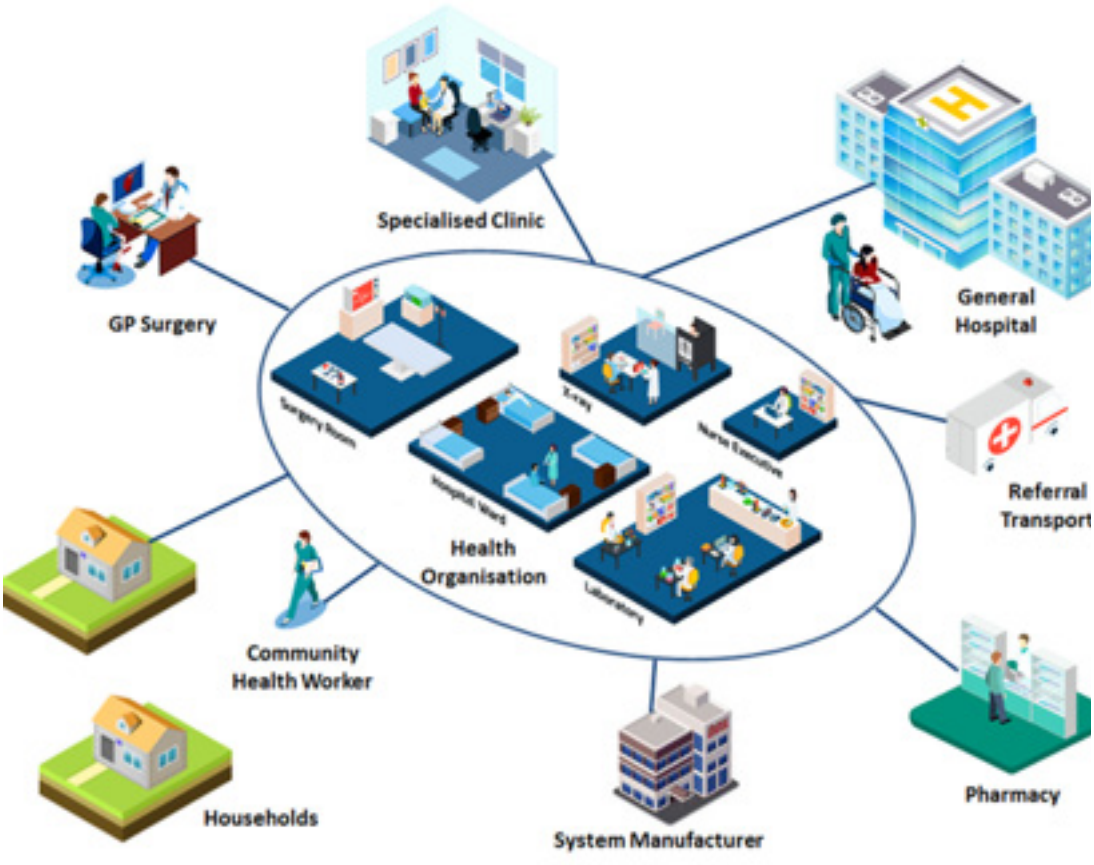


Figure 2: Case Study Scenario

in the face of increasing volumes of patients and pressure to reduce staff costs. The context for the integrated system is presented in Figure 2.

A Structured Methodology for Assessing Data Safety Risk

The final section in the guidance takes the form of an Appendix which provides a more formal, structured approach to the assessment of data safety risks and, inportantly, allows specific and appropriate 'good practice' risk management methods and techniques to be determined. We use the case study introduced above to illustrate the use of this methodology.

Use of the methodology is not mandated as part of the guidance, but it is highly recommended, since it allows linkage to be established between the idetified data safety risks and the good practice methods and techniques that are recommended to mitigate those risks.

The methodology uses the concept of Data Safety Assurance Levels (DSALS), which provide a more structured and systematic method of categorising risk. The DSALS are determined by extneding the risk assessment table already established in earlier sections, as indicated in Table 2.



Severity or Impact of data related accident	Very High / High	Medium	Low / Very Low
Catastrophic			
Multiple patient deaths. Permanent lifechanging incapacity and any condition for which the prognosis is death. Permanent lifechanging Incapacity for multiple patients. Severe injury or severe incapacity from which recovery is not expected in the short term for multiple patients.	DSAL4	DSAL4	DSAL3
Major			
A single death. Permanent lifechanging incapacity and any condition for which the prognosis is death for a single patient. Permanent lifechanging incapacity for a single patient. Severe injury or severe incapacity from which recovery is not expected in the short term for a single patient. Severe injury or severe incapacity from which recovery is expected in the short term for multiple patients. Severe psychological trauma for multiple patients.	DSAL4	DSAL3	DSAL2
Considerable			
Severe injury or severe incapacity from which recovery is expected in the short term for a single patient. Severe psychological trauma for a single patient. Minor injury or injuries from which recovery is not expected in the short term for several patients. Significant psychological trauma for multiple patients.	DSAL3	DSAL2	DSAL1
Significant			
Minor injury or injuries from which recovery is not expected in the short term for a single patient. Significant psychological trauma for a single patient. Minor injury from which recovery is expected in the short term for multiple patients. Minor psychological upset; inconvenience for multiple patients.	DSAL2	DSAL1	DSAL0
Minor			
Minor injury from which recovery is expected in the short term for a single patient. Minor psychological upset; inconvenience for a single patient.	DSAL1	DSAL0	DSAL0

Table 2: Data Safety Assurance Levels (DSALs)

Likelihood						
	Very High	DSAL1	DSAL2	DSAL3	DSAL4	DSAl4
	High	DSAL1	DSAL2	DSAL3	DSAL4	DSAL4
	Medium	DSAL0	DSAL1	DSAL2	DSAL3	DSAL4
	Low	DSAL0	DSAL0	DSAL1	DSAL2	DSAL3
	Very Low	DSAL0	DSAL0	DSAL1	DSAL2	DSAL3
		Minor	Significant	Considerable	Major	Catastrophic
		Severity				

Table 3: Clinical Risk Matrix

If a Clinical Risk Matrix is being used that is consistent with the matrix recommended in Section 4.4.3 of the Implementation Guidance to SCCI0129 and SCCI0160, then this table can be presented, as in Table 3 above.

Using this DSAL-oriented matrix simplifies the categorisation of data safety risk and allows recommended mitigations to be systematically determined.

Specific guidance on how data-related risks may be treated is presented in the Risk Treatment tables in Section 6.4 of the Data Safety Guidance, based on the data type and the DSAL of the data. These methods and techniques represent good practice in mitigating the risks associated with safety-related data. The Risk Treatment tables indicate where a particular method or approach is applicable to a given lifecycle data type and data property and, for each DSAL, whether the method/technique is recommended (R) or highly recommended (HR_. The lifecycle data types considered in the referenced version of the guidance are as follows:

- Verification (V);
- Infrastructure (I);
- Dynamic (D);
- Performance (P);
- Justification (J).

The applicable types of data property where the given technique could be effective in preserving that property are provided. These are coded to the 20 or so data properties identified, so for example 'I' indicates Integrity, 'C' Completeness, 'Y' Continuity etc..

Table 4 presents an example Risk Treatment table applicable to the case study example used in this paper.

It is expected that some interpretation may be required for these recommended techniques, depending on the nature of the system and the data, and indeed some may prove not to be applicable at all for certain applications. Nevertheless, it is recommended that each is given consideration, relevant interpretations are recorded and justification is provided where a technique is not adopted.

Conclusions

Health IT Systems are increasingly dependent on data and when issues arise with that data, there are risks to patient safety with potentially the same severity of impact as those normally associated with accidents caused by software or hardware. Management of data safety risks has not historically been addressed in safety standards; standards do exist, but these are primarily focussed on software and hardware. NHS Digital guidance has now been developed in conjunction with the Safety-Critical Systems Club and the Data Safety Initiative Working Group; this has been tailored to the Healthcare domain and supplements the existing Healthcare Standards. This guidance describes practices, methods and techniques that represents good practice in the

6.4.2.3 System Design

Technique	Data Types	DSAL 1	DSAL 2	DSAL 3	DSAL 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.....
Checksums / Cyclic Redundancy Checks (CRCs) / Hashes	..D..	-	R	HR	HR	Digests of data sets are produced, included with the data set and checked to provide confidence that the data is unaltered.	IC.....
Digital Signatures	..D..	-	R	HR	HR	For non-repudiation and integrity of data.	I.....T.....U.....
Sequence Numbers	..D..	R	R	HR	HR	Data bears sequence numbers so the integrity of a data stream can be checked (e.g., monotonic increase, duplicate detection).	ICN.....PQ.....
Automatic Repeat Request	..D..	R	R	HR	HR	Automatic Repeat-reQuest (ARQ) to repeat transmission of data which has not been received correctly.	IC.Y.....

Table 4: Example Risk Treatment Table

management of safety risks associated with data.

Paul Hampson *is a Chartered Engineer with over 28 years' experience in IT. He spent 15 years designing and developing enterprise systems in sectors including energy & health. He has extensive and varied systems safety experience, including corporate governance and assurance, and healthcarse standards-based safety engineering.*

Mark Thomas *is a safety engineer with 30 years' experience in the defence, nuclear and health sectors. He holds full memberships with the IET and the IMA and represents the UK via the BSI on the international committee responsible for developing an update to ISO80001. Mark holds degrees in Mathematics from the OU and LMU..*

On the Safest Way to Kill

by John Ridgway

The fascination of hunting as a sport depends almost wholly on whether you are at the right or wrong end of the gun.

P. G. Wodehouse, *The Adventures of Sally*

I like to think that business is very much like hunting. Basically, people hotly pursue their goals, and there are those who will benefit from the pursuit and others who will… well, maybe not benefit so much. To capture this idea, organisations like to talk of differing stakeholders, which is just a lofty way of saying that not everyone shares the same goals and concerns.

Stakeholding can make the construction of a safety case very interesting (if not to say contentious) since one person's safety can often set the scene for another's downfall. Nowhere is this contrast as stark as it is for nuclear weaponry. When nuclear weapons work safely, people are still allowed to die in their millions. Yet an unsafe weapon might only kill a handful of people, albeit an unintended handful. For example, a maintenance team could be killed by faulty detonators. Worse still, your warhead may fail to go off altogether, thereby allowing the enemy a greater opportunity to kill millions. And, of course, we are talking about the 'wrong' millions here.

Military analysts refer to the concept of Always/Never, which means that a weapon must always fire when required but never fire when not required. Consequently, if safety devices are to be applied, they cannot be allowed to compromise the mission – the weapon bearer's stakeholding must be protected under all circumstances. This uneasy balance between the demands of both mission and safety is central to the development of the world's nuclear arsenal, and it is a good reason why you shouldn't sleep too easily tonight.

How Safe Do Nuclear Weapons Have to Be?

Before I question too deeply the logic of developing a safety case that permits mass destruction, I need to consider the level of risk that can be tolerated on behalf of the population you claim to be protecting.

The first attempt to answer this question was made by the United States Army in 1955. They calculated how many fatalities had resulted from floods, earthquakes and other natural disasters in the preceding fifty years and used this as the target mortality rate for nuclear accidents. From this they calculated that the acceptable odds for an atomic (fission) weapon to explode accidentally on U.S. soil should be 1 in 250 during the course of a year. The same target equated to a 1 in 100,000 chance for thermonuclear weapons. Remember, we are talking about accidental death here. Determining the acceptable number of communist deaths required quite a different calculation.

In 1957 an organisation known as the Armed Forces Special Weapons Project revisited the question and concluded that there should be no more than a one in five chance of an accidental thermonuclear detonation within a given decade. The same organisation condoned the virtual certainty of a fission weapon accidentally detonating within the same timescale. This was the military viewpoint. One can imagine that civilians would come up with a different answer.

Since those early days, nuclear weaponry has seen many changes. Furthermore, the social and political backdrops have

evolved. Consequently, the safety standards required for nuclear weapons have undergone significant reassessment. The latest statement on the subject (chapter 2 of US Defense Standard DoD 31 50.2-M, 1996) states that the probability that a nuclear warhead might accidentally detonate, whilst in normal storage and operational conditions, shall not exceed one in a billion per warhead lifetime. However, this target drops to one in a million for abnormal environments such as those resulting from fires. Keep in mind that this limit is per warhead, so a goodly stockpile will increase the risk significantly. It seems, even now, no one expects nuclear weapons to be entirely safe, and many expert authorities fear that an accidental or unauthorised detonation is only a matter of time.

So How Safe Are They?

Or, more to the point, should one really expect a straight answer to this question? For example, when it was proposed that Strategic Air Command bombers should carry live nuclear bombs as part of a permanent airborne readiness, President Eisenhower was advised that the chances that any aviation accident could result in a nuclear explosion were, "essentially zero". In truth, zero was the only probability that could be confidently ruled out. And given that B52 bombers were, at that time, crashing once in about every 20,000 flying hours, the real prospects of avoiding a very bad day at the Oval Office were not nearly as good as Mr President thought.

So, I won't insult your intelligence by plying you with 'official' facts and figures. Nevertheless, a survey of things that might go wrong can still be illuminating.[1] Firstly, I'll address questions of physical vulnerability, then I will take a quick look at failures of command and control. As I do so, I will discuss safeguards and the ever-present conflict between the pursuit of both safety and mission objectives.

Wooden Bombs

In the earliest days of atomic weapons, the risk that a bomb might accidentally detonate was minimised by delaying final assembly until the point of deployment.[2] Prior to that point, the fissile component would be stored separately from the high-explosive components used to detonate it. Whilst this arrangement improved the safety margins whilst in storage, it was bad news as far as operational readiness was concerned. Furthermore, the weapons were difficult to maintain and very bulky and heavy. Weapons development therefore focused on creating relatively lightweight[3], preassumbled bombs that could be safely stored at dispersed sites and then loaded and primed within minutes – so-called "wooden bombs", sat inertly on the shelf awaiting use. Since such weapons would have their fissile core confined within a solid metal casing, they were said to have a sealed pit.

One-Point Safety

The resulting weapons were smaller and more powerful but had the disadvantage of being ready to bang if subjected to physical trauma, as in a bullet strike or a fire. Consequently, they were required to be "one-point safe". This decreed that detonation at a single point within a bomb's high explosive system (used to implode the fissile core) shall not result in a nuclear explosion.[4] Furthermore, any non-nuclear explosion's yield should not exceed the equivalent of four pounds of TNT (in consideration of the damage that a non-nuclear explosion can still cause on a ship). Notwithstanding these arrangements, the behaviour of high explosives under abnormal conditions, such as in a fuel fire, is sufficiently uncertain as to warrant extreme caution when dealing with such conditions. Each make of bomb had a "time factor", typically only a few minutes, after which every

firefighter and their dog was advised to run to the hills.

More Things to Wory About

Irrespective of one-point safety, two serious concerns remained to preoccupy the safety conscious combatant:

Firstly, a non-nuclear explosion may still result in the dispersal of highly radioactive material. And if this happens to be plutonium, you are dealing with a very serious risk indeed, since plutonium-contaminated dust can be easily inhaled, to deadly effect. Worse still, plutonium has a half-life of about 24,000 years, so any contamination is likely to outstay its welcome.

Secondly, early sealed pit weapons ran the risk of accidental detonation caused by a rogue firing-signal. This could result from a maintenance accident, a lightning strike or a fire. To counter this risk, a trajectory sensing switch was developed, whereby the firing mechanism would only become operational following weapons launch or release. Prior to that point, it should be physically impossible for current to flow to the detonators. But Mother Nature is ingenious, so no mechanism is guaranteed to carry zero risk.[5] And the harder you try to reduce the risk of accidental fusing, the more likely you are to end up with a dud when you press the button. The military have good reason to worry about this. A safety device added to Polaris missiles was later found to be subject to corrosion, rendering the missile dud. In fact, a 1963 routine maintenance check discovered that at least 75% of the Polaris fleet's missiles were destined to disappoint.

If in Doubt, Build More Bombs

Perversely, the early development of safety mechanisms was to make the world less safe. Concerned that a good many of their weapons might now be of the "Never/Never" persuasion, the military compensated by increasing the nuclear stockpile, with the intention of hitting any given target with many and varied warheads. This clearly increases the risk that a full-scale nuclear conflict might be instigated in error, making that of a single, accidental detonation appear quite attractive in comparison. So this is probably a good time to discuss how a nuclear contretemps could happen and how it may be avoided.

The Importance of Self-Control

First and foremost, protection[6] from nuclear attack requires that one



possesses an early warning system that is both effective and trustworthy. Consequently, any indication of an incoming attack must be properly evaluated before a response may be decided. However, the checks and balances required to avoid responding to a false alarm inevitably take up valuable time, and the "use it or lose it" principle is bound to influence deliberations. Clearly, therefore, the risk of hastily responding to a false alarm has to be suitably balanced with that of failing to respond in a timely fashion. The good news is that the military are able to practise their response using computer-simulated attack and so, presumably, have tested their procedures to remove any inefficiencies. The

not so good news is that in 1979 someone in the North American Aerospace Defense Command loaded a training tape onto a computer without anyone realising and, in so doing, very nearly precipitated World War Three. Catastrophe was averted only because surveillance satellites failed to provide corroborating evidence of an attack. Other close calls have resulted from mistaking a scientific rocket for an ICBM, a faulty computer chip, sunlight reflected from high-altitude clouds, and mistaking the moon for an incoming salvo of ICBMs.

As a further concern, one of your nuclear weapons may fall into the wrong hands, and there is always the chance that one of your own personnel may not be on message. Consequently, controls and arrangements that help avoid an unauthorised launch[7] are essential. These include: Permissive Action Links (PALs), i.e. devices that require a coded input to enable a device's firing mechanism; centrally-controlled code management; dual key arrangements, with mutual monitoring of potentially rogue behaviour; secure and reliable communications for authorisation and confirmatory dialogues; effective security protection for arsenals and armed delivery systems; and appropriate psychological profiling and selection of personnel. I could go on, but I doubt if I could convince you that the arrangements amount to a fool-proof system. As long as battle readiness remains a priority, the Devil will always find a way.[8]

From Stakeholder Management to Holocaust

The development of nuclear weaponry provides a prime example of the efforts required to meet the Always/Never challenge. However, war is what happens when stakeholder management gets out of hand, so it is no surprise that, in the battlefield, Always/Never makes sense only when seen from the appropriate stakeholder perspective. Not only are safety and mission efficiency arrangements often at odds with each other, a successful balance between them is likely to maximise the death toll, once the slaughter of all stakeholders has been taken into account. Safety arguments can look very odd when reliability equates to lethality, since the concepts of the right people to kill and the right circumstances in which to kill have a very shaky ethical foundation.

Finally, it's not as if stakeholder conflict is unique to military systems. Even in civilian applications, your stakeholder management obligations may lead you to make some morally dubious decisions. For example, what price public safety if it requires that the lives of the emergency services be placed at risk? The bottom line is that if you ever find yourself constructing a safety argument and asking yourself, "What would Jesus do?", you might want to consider getting another job.

Notes:

[1] In this article I do not attempt a quantitative assessment of the safety risk. However, those who are interested may wish to consult On the Risk of an Accidental or Unauthorised Nuclear Detonation – RM-2251 U.S. Air Force Project Rand (1958), in which Bayesian techniques are used to assess probabilities based upon "accident opportunities" and their incident rates.

[2] So, for example, an airborne weapon's nuclear core would only be inserted after take-off. To further reduce the risk, nuclear missions were not to be flown from U.S. bases. Instead, expendable airfields such as those in Norfolk, England were to be used. As far as the U.S. government was concerned, the good people of Norwich were the wrong kind of stakeholders.

[3] The weight and size reduction was principally achieved by using solid electrolytes to provide the power to activate the warhead, rather than relatively heavy, liquid electrolytes. In addition, a tritium and deuterium gas was used to boost the weapon's yield, whilst enabling the reduction of fissile material. Note that the fission boost is provided by extra neutrons produced following fusion of the lithium and deuterium. However, the energy released by the fusion

does not itself significantly contribute directly to the yield, so the weapons are still basically fission devices.

[4] One way to ensure this is to design the bomb's core so that, notwithstanding its implosion, reaching criticality still requires the boost neutrons that are released following fusion of the surrounding lithium and deuterium gas. The mechanism that activates this fusion is independent of that which detonates the high explosives. Consequently, an accidental activation of the high explosive implosion system will always be insufficient, on its own, to instigate a nuclear reaction.

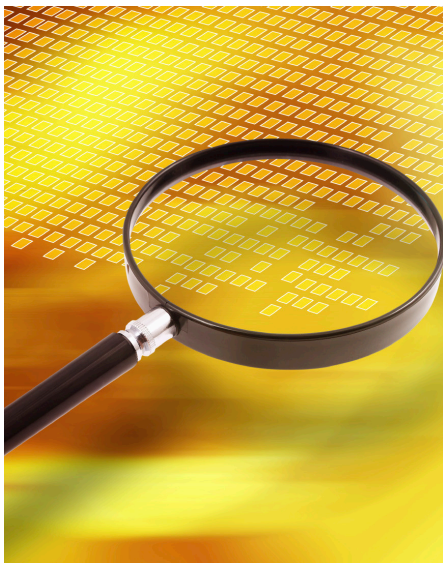
[5] For example, short circuits can result from melted solder, charred plastic or from arcing between newly proximate wiring. Errant currents might then energise the firing mechanism. Such risks are minimised by isolating the firing mechanism within a physical barrier and then using so-called 'strong links' to connect to external components. 'Weak links' within the isolated section are then designed to predictably fail 'open' in abnormal environmental conditions, long before the strong links have chance to fail in a potentially unpredictable fashion.

[6] By 'protection', I mean ensuring that your nation's imminent destruction shall not go unnoticed.

[7]The U.S. terminology for nuclear mishaps is provided by DoD Directive 5230.16. Apparently, an unauthorised launch of an ICBM, resulting in detonation, is an example of a so-called 'NUCFLASH'. I wonder how long it took them to come up with that one.

[8] To expedite launch procedures, the security code required to activate the Minuteman missile was set to '00000000' in all silos.

John Ridgway *is a retired analyst who spent most of his career working for Serco Transportation Systems. Having failed to change the world, he now lives in quiet retirement with his wife and two dogs. He can be contacted at j.ridgway4@ntlworld.com*



Requirements in the Loop: The Future of Model-Based System Engineering

by Yves Génevaux, Ken Nathan, Fabien Gaucher & Bruno Joffret

In this article, we consider the level of maturity of Systems Engineering as compared with other phrases of embedded systems development. Although much progress has been made over the years with the introduction of Requirements Management tools, Traceability tools and support for organising the elaboration of an architecture with Model-Based System Engineering (MBSE), essentially through various SysML flavours, there still remains a vast area for improvement. It is known that something of the order of 70% of the errors introduced during the various activities of an embedded system development project are actually introduced during the specification phase, whereas only some 4% of the errors are discovered during this phase. Figure 1 illustrates the percentage of errors typically introduced and detected during the development process. Given that the general consensus is that the earlier you find the errors, the more efficient and cost-effective it is, it is clear that the processes and tools used during the specification phase need to be improved.

One blocking factor to improvement is that most of the information relating to the expected behaviour of the system and of its components remains in a number of requirements written in natural language. Imagine dozens, or hundreds, of requirements that are to be considered in parallel, where any of them can be 'activated' at any

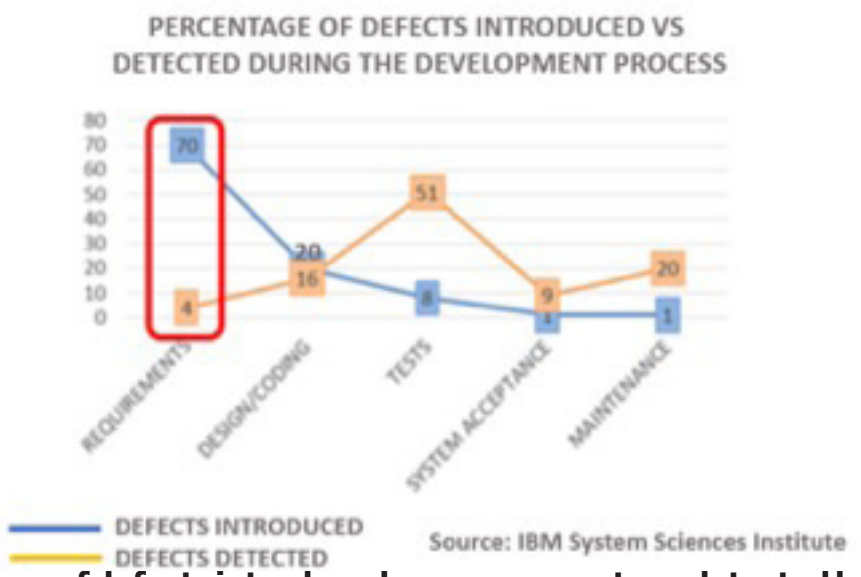


Figure 1: Percentage of defects introduced versus percentage detected by development process phase

moment. The human brain is simply not designed to handle such parallelism. Conflicts, missing requirements, ambiguities and semantics that do not actually reflect the designers' intentions are numerous. Even if MBSE is applied rigorously, a very large amount of critical functional information is simply out of the model, radically reducing the ambition one can have for the accuracy and correctness of the model. Figure 2 captures this mismatch and its implications. How can we improve this situation? What fundamental principles should we apply in order to make the specification less error-prone, more effective and supportable by easy-to-use tools?

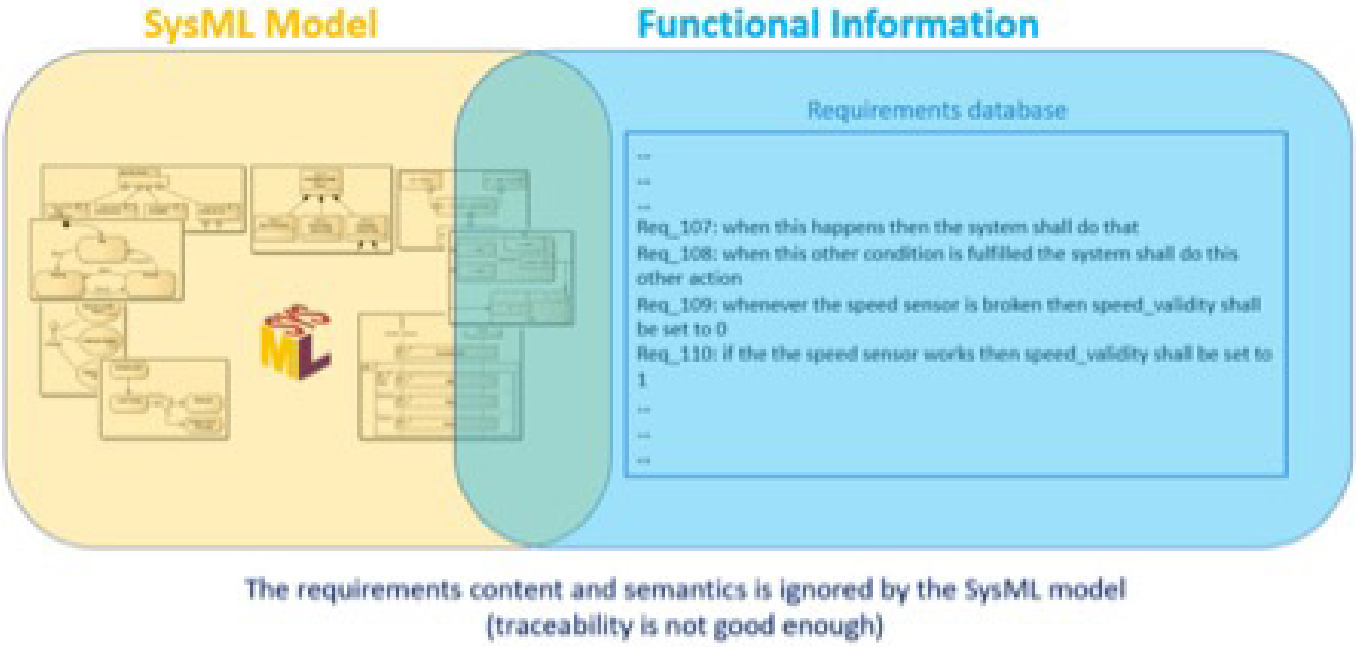


Figure 2: Model-Information mismatch and its implications for requirements

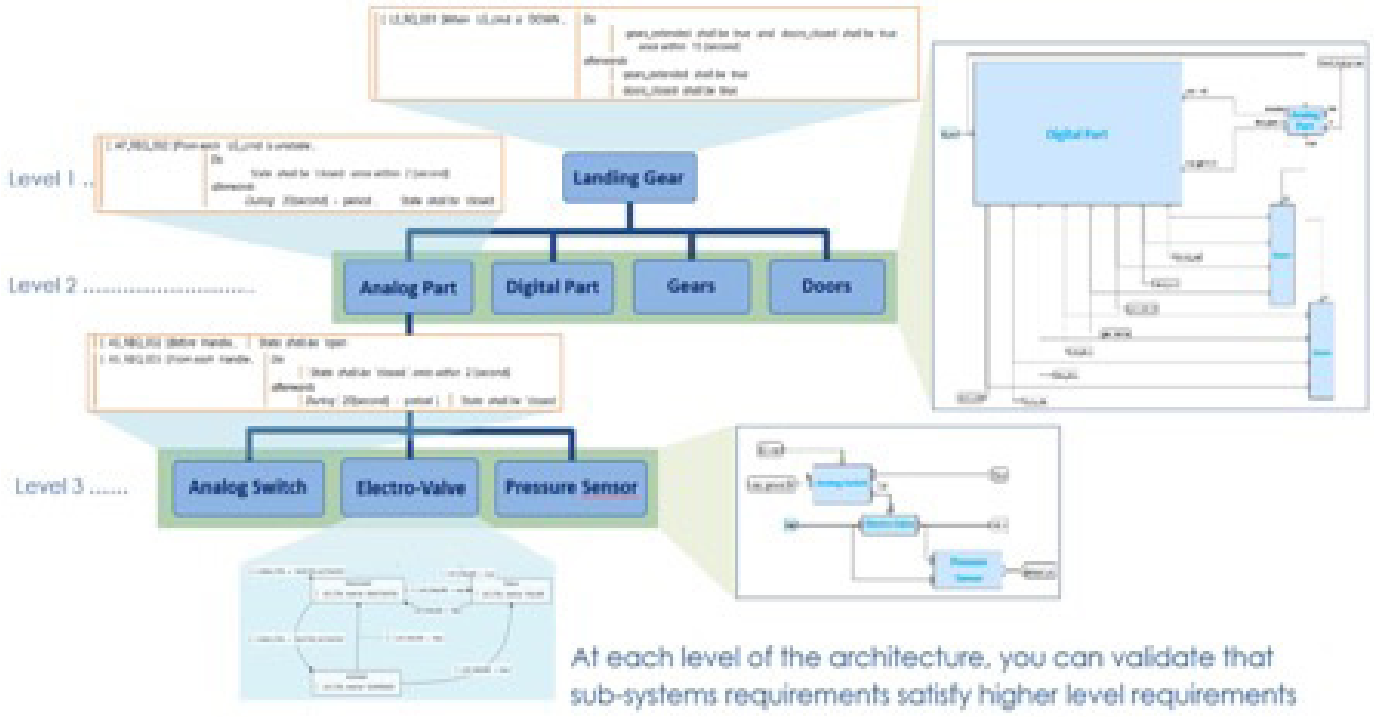


Figure 3: STIMULUS prcess model

We believe that three key points should be taken into account, in designing a better method:

1. The fact that SysML-based specification models ignore the behaviour of the requirements is a severe limitation. The requirements are not part of a model and therefore any attempt to propose a more effective process should ensure that any model includes the requirements.
2. The requirements can no longer be based on free text, or informal templates. This permits the introduction of too many mistakes. However, any solution should remain very close to natural language, and should also allow customisation so that it is possible for a company to implement company-specific variations according to its particular needs.
3. Simulation is a must. In the world of embedded systems, behaviour matters.

In order to meet these challenges, Argoism proposes a disruptive approach which is supported by a software tool, STIMULUS.

1. As illustrated in Figure 3, the STIMULUS model provides for the description of the system architecture and the interfaces between architecture components with a hierarchy of Bloack Diagrams. These can be linked to textual requirements and hierarchical state machines, which are used to describe the expected behaviour of these components. Structuring the architecture in this way means that, at each level of the architecture, it is possible to validate that sub-system requirements satisfy requirements at a higher level.
2. As illustrated in Figure

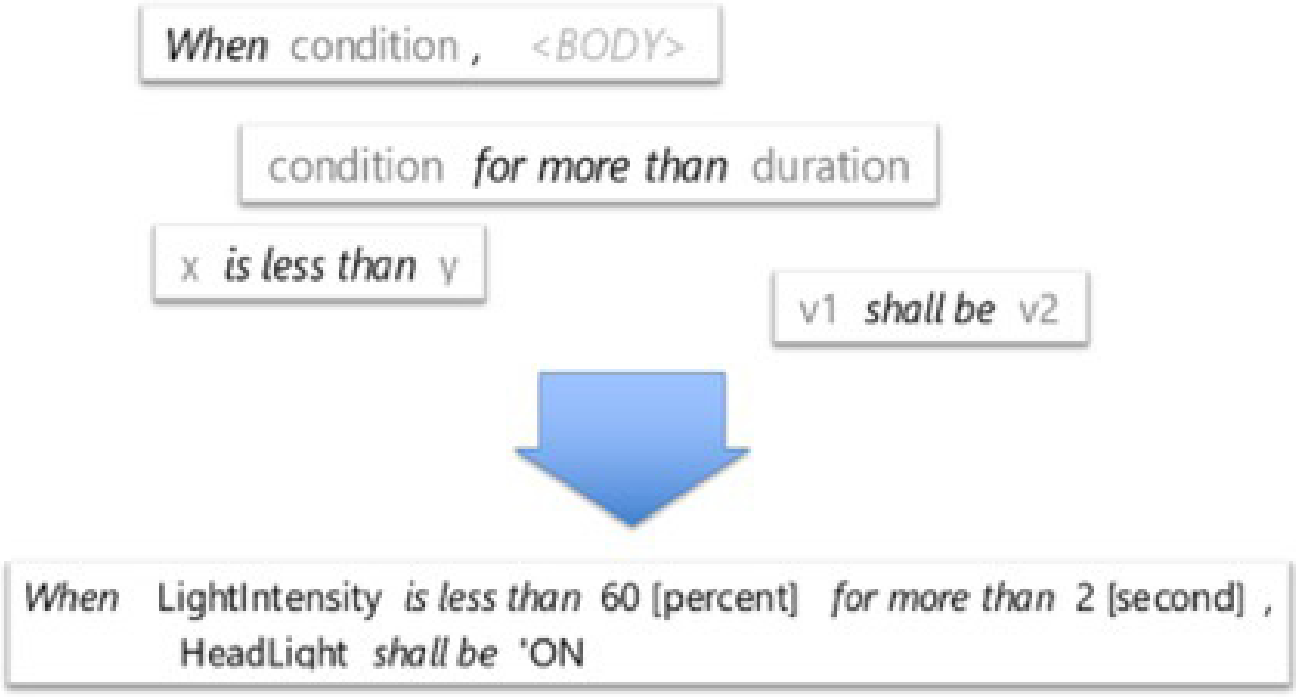


Figure 4: STIMULUS requirements templates

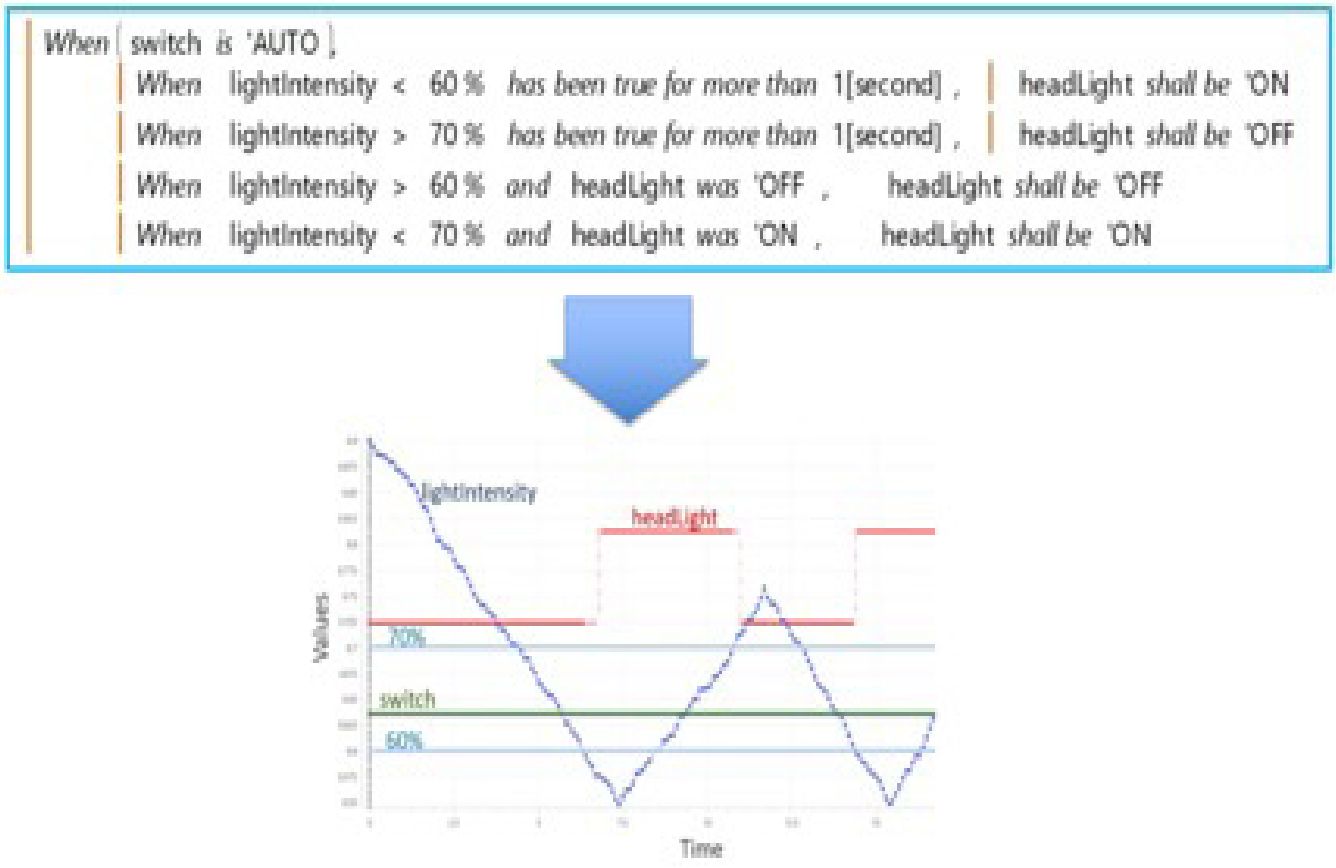


Figure 5: Debugging requirements

4, textual requirements are written using a series of predefined templates, called "operators". The library of operators is customisable and extendable by the user.

3. The state machines and textual requirements have executable semantics: simulation is therefore possible. You can now debug a specification in the same way that you would debug a piece of code. More easily, in fact. (Figure 5)

With these principles, the requirements are not only clear and executable; they are an intrinsic part of the specification model, which is what System Architects have been lacking so far. They are *"In the Loop"*, in exactly the same way as other *"In the Loop"* concepts adopted by industry to improve verification processes. Figure 6 illustrates these concepts with reference to a standard V lifecycle. Starting from the right-hand end of the V:

HIL - Hardware in the Loop. The objective is to test the system as close as possible to the production context.
SIL - Software in the Loop. Various execution and emulation environments help in executing and debugging the code.
MIL - Model in the Loop. During thr design phase, the models can be activated using simulation, so that future system behaviour can be verified as early as possible.
That "as early as possible" point has now shifted further up the left-hand side of the V:



Figure 6: Requirements in the Loop

RIL - Requirements in the Loop. During the specification phase, the requirements are part of an executable model, which allows the limits of the approaches described above to be overcome.

How does RIL work in STIMULUS?

Following recent research in the field of synchronous languages and verification methods, STIMULUS applies two basic principles:

1. Requirements are considered to be constraints on specified system behaviour. A constraints solver is therefore able to compute the space of possible system behaviours at a given execution cycle.
2. Use Cases are defined as a set of constraints on and between inputs. The constraints solver defines "acceptable" inputs for a given Use Case, and can generate any number of test vectors, between one and several thousand. This will allow an automatic exploration of the envelope of possible system behaviours. STIMULUS then produces classical simulation plots displaying the evolution of the inputs and outputs during the execution of the model.

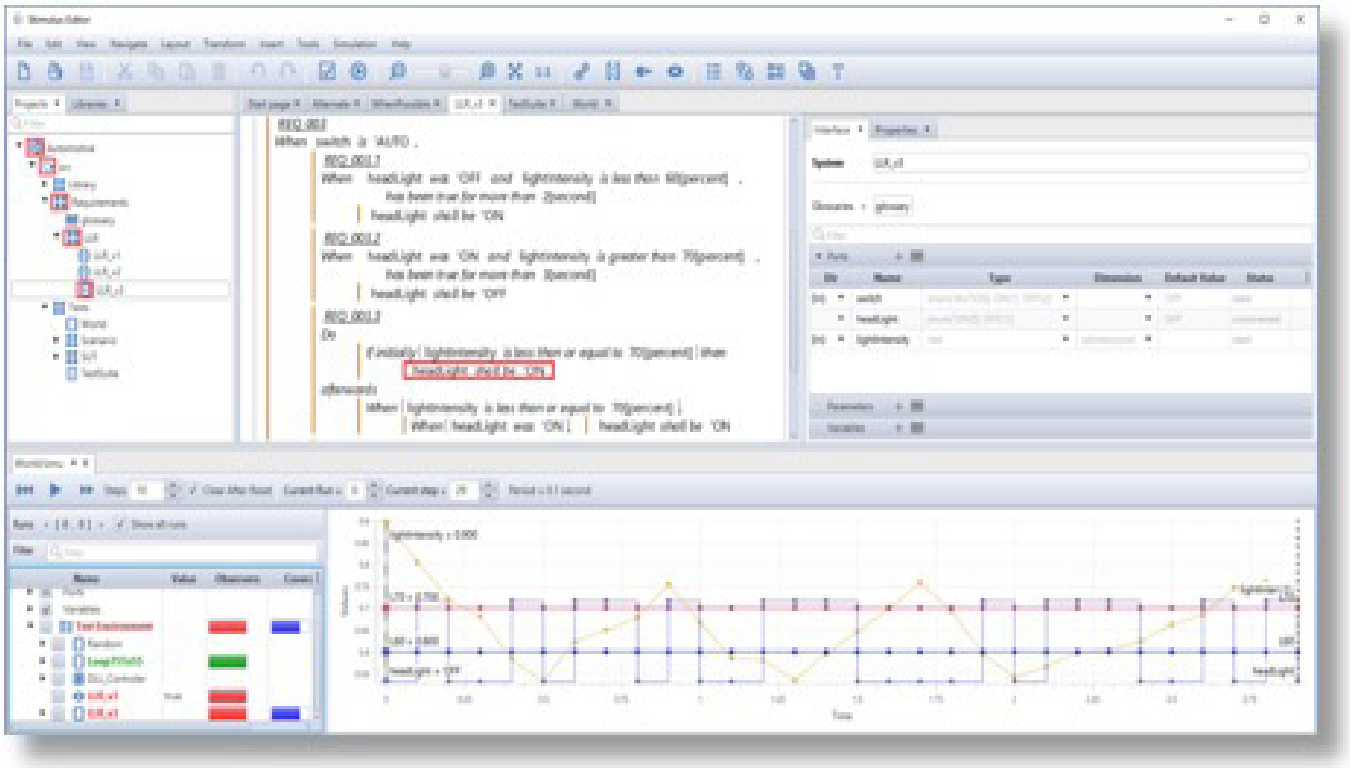


Figure 7: STIMULUS screenshot

The RIL approach has many benefits:

- The requirements elicitation process is supported by a tool. Since one requirement is enough to start simulating, an iterative process can be used to define the requirements incrementally.
- Missing and conflicting requirements are identified automatically.
- You can verify that a specification complies with safety requirements, putting the latter in Observers whilst the specification is simulated by STIMULUS. If the specification violates the safety requirements, the Observer will report it.
- You can verify the compliance of code with its requirements automatically. Again, using an Observer that contains the specification whilst the code (compiled in a DLL, or an FMU generated by a third-party tool) is executed by STIMULUS. You then have a completely automatic, functional validation environment.
- Different versions of the same specification can be compared. One is executed while the other one is put into an Observer. It becomes possible to verify different hypothesis, variants of the same specification, or the impact of a change.
- The functional coverage is directly measured by STIMULUS without having to use traceability links, which are always subject to errors. STIMULUS will indicate the requirements that have not been tested during a test campaign so you can design new use cases that will increase the functional coverage.
- The consistency between the requirements of different levels of the system can be verified, which has been extremely difficult to do so far. You can verify if the refinement of the requirements, from the top level down to lower levels, is correct.

In the present industrial context for embedded systems, projects face a stark choice: either explore

new ways of working, or carry on as before and continue to face the consequences of injecting multiple errors into the specification. The consequences of the latter choice are potentially very costly: there is a risk that an organisation will continue to develop the wrong system, and/or incorrect test cases, until realising, towards the end of the project, that project deadlines and budgets cannot be met, because of the lengthy correction loops that are required. The consequences of an error making it through into the final project are potentially very severe. The Requirements in the Loop philosophy, supported by the STIMULUS tool, offers system architects, requirements engineers and test engineers a means of finding errors when they are made, in the specification phase.

For more details of the STIMULUS tool, watch the videos at https://www.youtube.com/watch?v=aiu5pQRgm7o and https://argosim.com/product-videos/

Ken Nathan *graduated from Exeter University with a degree in Physics and Geophysics. He has worked in Sales and Marketing in embedded software and telecommunications for 30 years with such organisations as Vodafone, Nortel and PolySpace. He has run sales teams across EMEA, US and Australia and is currently the Country Manager for UK and Nordics at Argosim.*

Yves Génevaux *graduated from the Ecole des Arts et Métiers. He has worked in the fields of system and software engineering for embedded and critical systems for more than 25 years which led him to collaborate with key accounts in the aeronautical, railway, automobile, energy, telecommunication and defence industries, both in France and worldwide. After several years working as a development engineer, he developed several start-ups involved in the publishing of development and validation tools (Verilog, PolySpace Technologies) for the French and international markets (Europe, Asia, United-States). He has also worked with large US-based companies such as The Mathworks. He is currently the Sales Director at Argosim.*

Bruno JOFRET *graduated from EPITA with CSAC major (Cognitive Science and Advanced Computing). He has been working for more than 10 years around numerical simulation. Starting at INRIA, as a software architect, on Scilab Open source numerical computation software, he provided simulation/system expertise through Scilab Enterprises a company which he co-founded in 2010. He is currently Senior Consultant at Argosim.*

Fabien Gaucher *graduated from the Grenoble Alpes University and from the Institut National Polytechnique of Grenoble. He holds a PhD in Computer Sciences, carried out at the Verimag laboratory where he studied real-time system validation techniques. He then worked for 10 years at Dassault Systèmes, where he developed a simulation and code generation workshop for embedded systems. He is currently CEO of Argosim, which he co-founded in 2013 in order to create STIMULUS, the first simulation tool to debug and validate embedded system requirements.*



The Relationship Between Employee Well-Being and Safety-Related Outcomes

by Ludmila Musalova

Workplace safety, otherwise known as Occupational Health and Safety, focuses on the physical and psychological well-being of employees and has attracted considerable attention from government bodies, organisations, employers and safety researchers due to the far-reaching implications, both economically and socially. However, comparatively little attention has been given to the impact of stress on safety, despite research evidence demonstrating a clear link between these two factors.

What is work-related stress, and why does it matter?

Work-related stress is often referred to as Occupational or Job-related stress, and describes stress that originates in the work environment. The Health and Safety Executive (HSE) defines work-related stress as "the harmful reaction that people have to undue pressures and demands placed on them at work" [1]. In comparison to other types of stressors, organisations play a role in moderating stress levels experienced by their workforce, and possess the ability to mitogate factors which contribute to stress, through, for example, job design, managerial support and workload management.

Intial research has focussed on investigating stress amongst "working professionals" - i.e. white-collar employees. In the past two decades, there has been a shift in attitudes towards stress in high-hazard industries, bringing a recognition that more attention needs to be devoted to this topic. The initial research driving the change focused on the transport and driving, oil and gas (offshore drilling) and nuclear sectors.

Despite organisations' best efforts to raise awareness of stress, there is still a stigma attached to mental health, especially amongst operational levels of employees. Workers within high-hazard industries are often exposed to competing objectives, such as tight deadlines, budgets, and customer demands, relative to their personal safety and wellbeing. Without clear leadership and defined organisational priorities, workers can feel left to make a personal judgement on what and how to prioritise their work, which can contribute to or exacerbate feelings of stress.

Organisational failure to recognise and mitigate workplace stressors, often leads to impaired employee well-being, manifested by physical and emotional depletion, reduced efficiency and performance. Therefore, the effects of work-related stress will ultimately have a negative impact upon safety and safety related outcomes of workers, as well as the productivity and quality of output.

Stress has been shown to have negative implications for safety by reducing individual's lack of focus and concentration, impairing cognitive processes and psychomotor control, increasing proneness to fatigue, reducing effortful coping and creating task disengagement. This is also supported by research, which has shown the following findings:

- Stress is a contributory factor in the occurrence of incidents - there is a strong correlation between stress and involvement in accidents [2];
- High levels of stress are associated with self-reporting of errors, lapses and violations [3];

- Stress significantly elevates at-fault crash risk [4];
- Occupational stress is significantly predictative of self-reported, fatigue-related near misses [5];
- Job stress negatively affects safety behaviour in terms of safety compliance [6] and safety participation [7].

Stress Prevention and Stress Management Techniques

In order to improve employees' well-being and ensure a safe working environment, organisations need to make mental health of their workforce a priority. There are three levels of interventions aimed at stress prevention, namely primary, secondary and tertiary level.

- Primary-level interventions aim to eliminate the stress hazards by creating a healthy workplace and control stressors to ensure that the work is a) interesting, b) rewarding and c) within the person's capabilities;
- Secondary-level interventions aim to isolate the hazard by ensuring that only adequately trained and experienced personnel attempts to manage the causes of stress. This could be done by improving the fit between the person and the job by staff selection, on the job training, performance feedback and monitoring of problems. In addition to employee participation/ engagement in well-being programmes to enhance the prior points.
- Tertiary-level interventions aim to minimise the stress hazard enabling employee's return to work. The interventions are designed to help the person experiencing stress or deal with the consequences and aftermath of stress.

Examples of specific interventions applied at each

Type of intervention	Aim	Strategies
Primary	Dealing with stressors	Review of well-being and stress management policies. Work contingency planning strategies for all staffing levels. Organizational Stress Audits in the form of well-being questionnaires and consultation with employees.
Secondary	Helping people to cope	Management and supervisory roles training in mindfulness, resilience and stress management coaching. All staff training on how to recognize the signs of stress. Stress evaluation toolkits specific to high hazard industries. Identifying mental health champions to reduce stigma and aid support. Improving the fit between the person and the job. Healthy lifestyle programs encompassing mental and physical health.
Tertiary	Putting remedial measures in place.	Employee Assistance Programs (EAP). Access to counselling. Diagnosis of depression and care management. Planned Return to Work (RTW) and ongoing support in stress management.

Table 1: Primary, secondary and tertiary stress-management interventions

level are provided in Table 1. It is important to note, that stress management interventions and related training, encompassing all three levels should be executed by suitably qualified practitioners, e.g. psychologists.

Beacons of excellence in stress prevention and stress management

In recent years, there has been an increased effort from high-hazard industries (rail, aviation, nuclear and energy sector) to include the mental health of their workforce within their health and safety improvement programmes. However, these efforts are being hindered by poorly-implemented strategies (e.g. strategies focusing mainly on primary stress management interventions or exhibiting a reactive approach to stress management; strategies targeting a specific segment of employees only or being implemented by staff without suitable qualifications and experience). In order to truly initiate change, the stress management interventions should:

a) Design organisation-specific and worker-related strategies, involving managers, supervisors, employees and employee representatives in the continuous risk assessment and evaluation process.

b) Use management 'top-down' approach to cascade commitment and support for mental health. This will help to reduce the stigma of mental health and provide support and encouragement for staff to recognise the signs of stress.

c) Use employee 'bottom-up' approach to increase employee participation in stress management initiatives, recognition of stress and willingness and confidence to seek early help if needed.

d) Engage in an open communication about mental health and stress to normalise mental health issues.

The effectiveness of interventions is also determined by a holistic approach to stress management strategies accounting for the psychological, emotional and physical aspects of individual well-being, support for which is offered by Lazarus's Multi-Transactional Model of stress adopted by Palmer and Dryden [8]. The model recognises the effect of internal and external pressures and coping resources on the appraisal of situation which will ultimately affect the response to stress manifested in different modalities. The model is based on the premise that unless all seven modalities (behavioural, affective, sensory, imagery, cognitive, interpersonal and drug/biological elements) are considered, the interventions may not be successful, as significant issues may be overlooked. The seven modalities, collectively known as BASIC-ID are used to distinguish and classify different elements in the response to stress, as documented in Figure 1 below.

Stress management interventions and their effectiveness

Organisational interventions to improve employee health focus on physical and psychological well-being, encompassing variety of strategies. This can take form of a simple massage chair, aiming to aid relaxation to more elaborative strategies, such as wellbeing road-shows, training management and supervisory roles in mindfulness, resilience and stress alleviating thinking, wellbeing and suicide prevention toolkits, and workplace wellbeing indexes. In high hazard industries, these interventions are still in the infancy stages and greater efforts should be exercised to decrease employees' stress and improve their overall wellbeing.

Organisations which exercised conscious efforts to improve employees' wellbeing, such as EDF

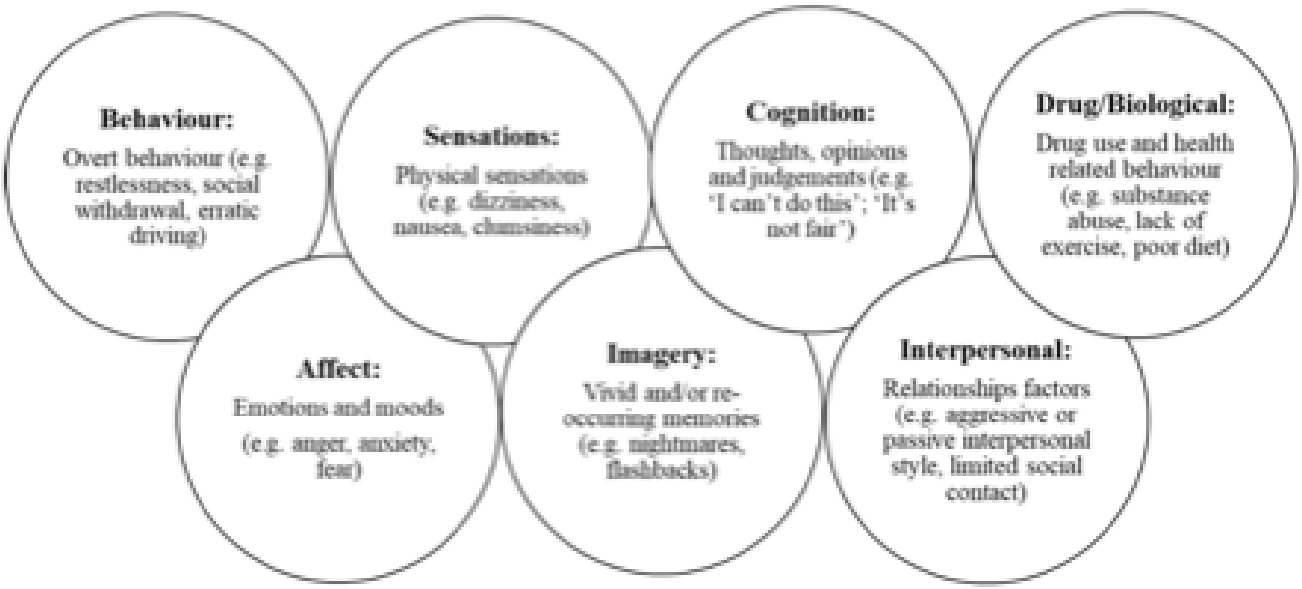


Figure 1: BASIC-ID distinguishing different elements in response to stress

(cognitive therapy and stress-recognition programmes) [9], Virgin Trains (education on mental health) [10] and BAE Systems (wellbeing awareness-raising programmes) [11], have noted a positive impact of these interventions, such as reduced stigma and increased reporting of mental ill health, improved working relationships and performance effectiveness, and overall job satisfaction. This also corroborates the documented benefits with academic research. For example, a meta-analytical study has shown that stress management training, namely cognitive behavioural training, relaxation techniques and multiple stress management training leads to improve psychological health and job satisfaction [12]. Positive effect was also demonstrated for comprehensible worksite stress management programme consisting of self-management and a stressor reduction process. The results have shown that this programme has increased work-units' productivity and decreased absenteeism [13]. There is also evidence that stress management interventions do improve safety related behaviour. For example, a study of nuclear plant control room operators testing the effect of mindfulness has confirmed that both safety compliance and safety participation behaviours are positively influenced by dispositional mindfulness (i.e. conscious awareness and attention to our thoughts and feelings in the present moment) [14].

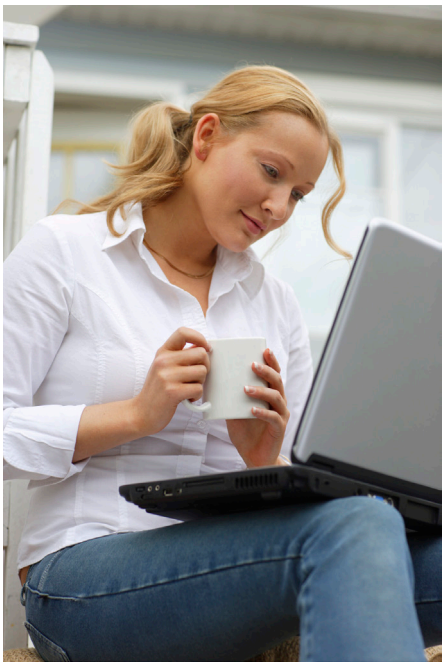
The benefits of stress-management interventions are undeniable, as they contribute not only to individuals' positive wellbeing, but also affect organisational outcomes, such as performance, absenteeism, safety and compliance. As a result, high-hazard industries need to place greater focus on stress prevention and stress management amongst their workforces, as individuals' mental

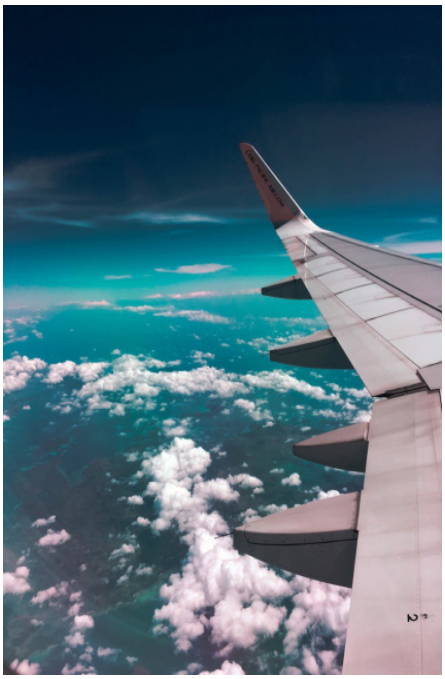
health is reflective of their safety-related behaviour, which can have detrimental effects not only for the individuals but for organisations.

References:

1. Author unknown. Work related stress and how to tackle it. Health and Safety Executive. 2015 [cited 13th June 2018]. Available from: http://www.hse.gov.uk/stress/what-to-do.htm
2. Sutherland VJ, Cooper CL. Personality, stress and accident involvement in the offshore oil and gas industry. Personality and Individual Differences. 1991 Jan 1;12(2):195-204.
3. Westerman SJ, Haigney D. Individual differences in driver stress, error and violation. Personality and Individual Differences. 2000 Nov 1;29(5):981-98.
4. Legree PJ, Heffner TS, Psotka J, Martin DE, Medsker GJ. Traffic crash involvement: Experiential driving knowledge and stressful contextual antecedents. Journal of applied psychology. 2003 Feb;88(1):15.
5. Strahan C, Watson B, Lennonb A. Can organisational safety climate and occupational stress predict work-related driver fatigue?. Transportation Research Part F: Traffic Psychology and Behaviour. 2008 Nov 1;11(6):418-26.
6. Lu CS, Kuo SY. The effect of job stress on self-reported safety behaviour in container terminal operations: The moderating role of emotional intelligence. Transportation research part F: traffic psychology and behaviour. 2016 Feb 1;37:10-26.
7. Wang D, Wang X, Xia N. How safety-related stress affects workers' safety behavior: The moderating role of psychological capital. Safety Science. 2018 Mar 31;103:247-59.
8. Palmer S, Dryden W. Counselling for stress problems. London: Sage; 1994.
9. Hampson E, Soneji U. At a tipping point? Workplace mental health and wellbeing. Delloite. 2017 [cited 14th June 2018]. Available from: https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/public-sector/deloitte-uk-workplace-mental-health-n-wellbeing.pdf
10. Clarkson. N. Virgin trains commit to tackling mental health stigma. Virgin. 2017. [cited 14th June 2018] Available from: https://www.virgin.com/careers/virgin-trains-commits-tackling-mental-health-stigma
11. Donaldson Feilder E, Podro S. ACAS Future of workplace relations -Discussion paper series: The future of health and well-being in the workplace. ACAS. 2012. [cited 13th June 2018] Available from: http://m.acas.org.uk/media/pdf/7/k/The_future_of_health_and_wellbeing_in_the_workplace.pdf
12. Kröll C, Doebler P, Nüesch S. Meta-analytic evidence of the effectiveness of stress management at work. European Journal of Work and Organizational Psychology. 2017 Sep 3;26(5):677-93.
13. Munz DC, Kohler JM, Greenberg CI. Effectiveness of a comprehensive worksite stress management program: Combining organizational and individual interventions. International Journal of Stress Management. 2001 Jan 1;8(1):49-62.
14. Zhang J, Wu C. The influence of dispositional mindfulness on safety behaviors: A dual process perspective. Accident Analysis & Prevention. 2014 Sep 1;70:24-32.

Dr Ludmila Musalova *is a Senior Consultant at Greenstreet Berman Ltd. She is a Chartered Psychologist and a certified Stress Management Coach. Her core areas of work include employee wellbeing and stress-management, safety culture assessment, non-technical skills training competence assessment, organisational change and research (analytics). She has worked in academia and with corporate clients across industries, in the UK and internationally.*











Spring 2019 Events Diary

SCSC Events

Safety Critical Systems Symposium 2019 (SSS'19)
Bristol Marriott Royal Hotel, Bristol UK,
Tuesday 5th - Thursday 7th February 2019
For further details, contact alex.king@scsc.uk

SCSC Workshop: Safety Assurance of Autonomy and AI
Bristol UK
Monday 4th February 2019
For further details, contact alex.king@scsc.uk.

Safety of Autonomous Systems Working Group Meeting No 11 (SASWG 11)
London UK
21st March 2019.
For further details, contact amjk@adelard.com

Assurance Case Working Group Meeting No 8 (SCSC ACWG 8)
Lincoln UK
March 27th 2019
For further details, contact phil.williams@engineerforsafety.co.uk.

SCSC Seminar: Evolution of Assurance Case Practice
London UK
4th April 2019.
For further details, contact alex.king@scsc.uk.

SCSC Seminar: Learning from Accident Investigations
London UK
13th June 2019
For further details, contact alex.king@scsc.uk

Other Safety and Dependability Events

ISO26262 Improving Safety of Advanced Mobility Conference
Munich Germany
25th - 28th March 2019
For further details, see https://iso26262-conference.iqpc.de

Simulation and Training for Resilience and Safety Symposium (STRS 2019)
London UK
26th - 27th March 2019.
For further details, see https://www.halldale.com/strs-call-for-papers/

Masterclass: Systems Approach to Safety and Cybersecurity
Manchester UK
8th - 11th April 2019
For further details, see https://stamp-consulting.com/training-classes/stamp-uk-masterclass/

38th International Conference on Computer Safety, Reliability and Security (SAFECOMP2019)
Turku Finland
10th - 13th September 2019
For further details, see https://conferences.ncl.ac.uk/safecomp2019/

7th Scandinavian Conference on System and Software Safety, Stockholm Sweden, 22nd - 23rd October 2019. For further details, see http://safety.addalot.se/2019

INCOSE UK Systems Engineering Conference 2019
Leeds UK
20th - 21st November 2019
For further details, see https://incoseonline.org.uk

