The Safety-Critical Systems Club Newsletter

Safety Systems

Vol 28 No. 3 - Oct 2020

DARK DATA
Why what you don't know matters

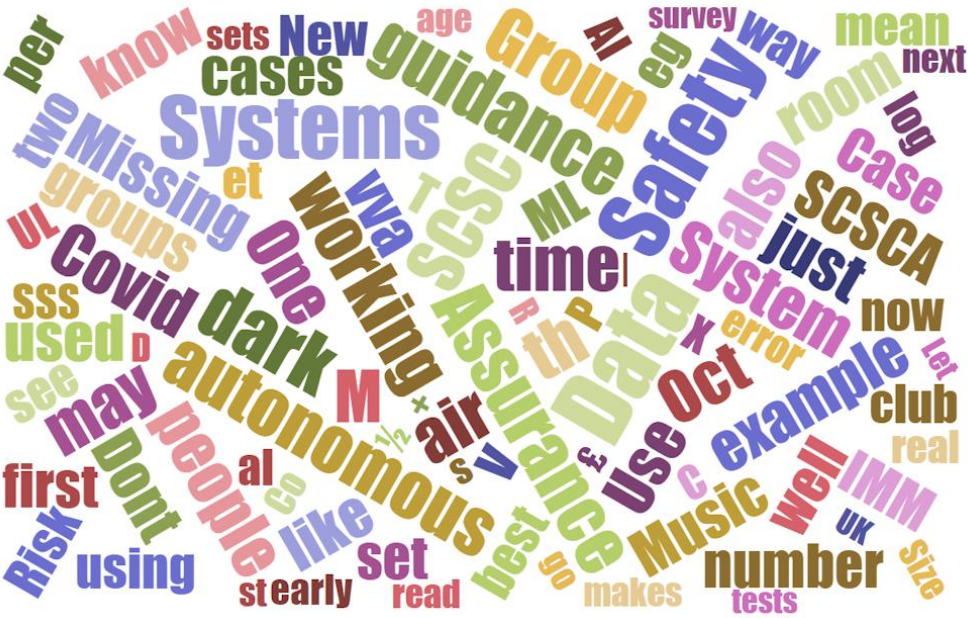
AUTONOMOUS STATE
How to assure autonomous systems

SOMETHING IN THE AIR
Has COVID-19 guidance gone far enough?

For everyone working in Systems Safety

SCSC

thescsc.org



SCSC Publication Number: SCSC-160

Contents

SCSC

FOR EVERYONE WORKING IN SYSTEM SAFETY

www.scsc.uk/sss

SSS'21

Safety-Critical Systems Club Annual Symposium

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

www.scsc.uk

THE SAFETY-CRITICAL SYSTEMS CLUB

29th Safety-Critical Systems Symposium

9-11th February 2021, Online

The Safety-Critical Systems Symposium in 2021 (SSS'21) will be held online. The event comprises three afternoons of live presented papers, including keynote presentations and submitted papers. There will also be pre-recorded 3-minute 'pitches' and poster talks.

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

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

The symposium features are:

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

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

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

For **further information, exhibition** and **booking queries** please contact: Alex King, Dept of Computer Science, University of York, Deramore Lane, York, YO10 5GH. Tel: 01904 325402; Email: alex.king@scsc.uk

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

Editorial

Our lives continue to be dominated by the Covid-19 pandemic and there seems little prospect of any immediate relief from its grip. Its impacts still remain far-reaching, and it is now unfortunately starting to affect the SCSC itself, with all future SCSC seminars and the annual symposium in February 2021 now being hosted online for the first time in its history.

On a positive note however, these events are now free to SCSC members so I would encourage you to renew your membership if you haven't done so already, and perhaps encourage friends and colleagues to also take advantage of these excellent events. Further savings can also be made through corporate membership arrangements, student discounts and the new feature of multiyear subscriptions, so please visit the link on the back cover and page 45 for more information.

In the many months since the initial nationwide lockdown, one would have thought that our understanding and handling of the situation would have improved and it seems incredible to see critical data handling issues still arising, such as the loss of 16,000 Track and Trace records. There also seems to be confusion on how data is captured and its meaning – is the record of a death "from" Covid-19 or "with" Covid-19, and how is the data being extrapolated from the location, demographic and motivation of those that are tested to say something about the population as a whole? These issues are very much at the heart of our first article from Prof. David Hand on "Dark Data" where he discusses the perils of the data we *don't* have, which can be just as important, if not more important than the data we do have. In the subsequent article, Mike Parsons continues this Dark Data theme and provides outline guidance on how to manage Dark Data risk from a safety perspective.

In our third article, Peter Ladkin takes a rigorous approach to assessing the guidance for musicians gathering in bands and orchestras. He looks at the practicalities of the droplet-focussed guidelines and analyses the risks and mitigations for aerosol-based infection, which undoubtedly raises important considerations for any indoor gatherings.

John Spriggs continues the theme of mistakes and misconceptions in our fourth article, with a discussion on the common pitfalls in constructing assurance cases, and he asks whether lessons are being learnt when things go wrong.

We conclude with an article from the SCSC Safety of Autonomous Systems Working Group and the progress they have made in developing practical guidance for assuring systems involving Artificial Intelligence (AI) and Machine Learning (ML) – areas that are rife with the issues of Dark Data we encountered in the first two articles.

Our 60 second interview is with Prof. John McDermid.

Note that the **scsc.uk** domain will be transitioning over to **thescsc.org** to reflect its more international presence and applicability, so please look out for changes and further announcements in the coming months.

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

In Brief

Covid: Test error 'should never have happened'

The health secretary has said a technical glitch that saw nearly 16,000 Covid-19 cases go unreported in England "should never have happened". The error meant that although those who tested positive were told about their results, their close contacts were not traced. The technical error was caused by some Microsoft Excel data files exceeding the maximum size after they were sent from NHS Test and Trace to Public Health England. *bbc.co.uk*

First death reported following a ransomware attack on a German hospital

German authorities are investigating the death of a patient following a ransomware attack on a hospital in Düsseldorf. The patient, identified only as a woman who needed urgent medical care, died after being re-routed to a hospital in the city of Wuppertal, more than 30 km away from her initial intended destination, the Düsseldorf University Hospital. *zdnet.com*

Uber's self-driving operator charged over fatal crash

The back-up driver of an Uber self-driving car that killed a pedestrian has been charged with negligent homicide. Elaine Herzberg, aged 49, was hit by the car as she wheeled a bicycle across the road in Tempe, Arizona, in 2018. Investigators said the car's safety driver, Rafael Vasquez, had been streaming an episode of the television show The Voice at the time. *bbc.co.uk*

Ethernet failure on Swiss business jet prompted emergency descent

An Ethernet failure aboard popular Swiss-made business jets could prompt the aircraft to move into an emergency descent as flight systems entered a "degraded" mode, the European Aviation Safety Agency (EASA) has warned. *theregister.com*

Stonehaven train derailment: Crash investigators confirm train struck landslip

The driver, conductor and a passenger died when the 06:38 Aberdeen to Glasgow service crashed near Stonehaven. The RAIB said the train had turned back towards Aberdeen after reports of a landslip further down the track. The six-vehicle train had travelled more than a mile when it was derailed after hitting a separate landslip. *bbc.co.uk*

Shining a Light on Dark Data

Professor David Hand discusses the perils of "dark data" – where the data that you *don't* have can be just as important, if not *more* important than the data you do have. He provides examples of where dark data has let to catastrophic outcomes and provides insights into managing the risks of dark data, and even, how to start using dark data to your advantage.

When you first start to learn statistics, you learn from small complete data sets. You calculate the mean of ten numbers, their median, their standard deviation. Later you learn to use computers, carrying out more sophisticated statistical analyses like regression or analysis of variance. Again, initially, when you learn how to use these tools – the statistical methods and the software which implements them – you apply them to perfect data, to clean and complete data sets. This all makes perfect sense. If you are trying to learn about averages or regression analysis, you don't want to be sidetracked by having to deal with peripheral matters.

Except that, when you graduate from learning how to use the tools and begin to apply them to real data sets, you realise that you have been misled. You learn that real data are not as clean as those early ten-point data sets suggested, that the data for your regression aren't as balanced as in the text books. At first, you ignore the problems. You discount the fact that quite a few people you approached for your survey refused to take part so your data set does not include them, that you were unable to take accurate measurements of some of the chemical concentrations, that some of the numbers were so outlandish that you felt obliged to drop them from the analysis, that the data collected when the patients were admitted to the hospital isn't *exactly* what you need to answer your research question, that the instrument providing the telemetry feed never gives a value greater than 2.3.

But then you begin to think more closely about the issues. What if those who refused to respond to the survey were different from those who agreed? Would this mean my statements about the overall population were wrong? What if the telemetry instrument had a hard ceiling at 2.3, so that it could never read above this, even if the actual values were way larger? Would this mean I was being very misled about how the system was working?

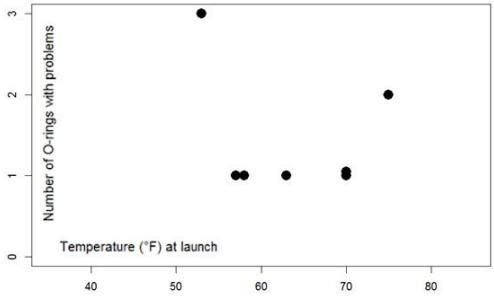
These problems, these data inadequacies, are examples of *dark data*. Dark data are, in short, data you don't have. They are the data you believe you have or hoped to have or would have had *if* something else had happened but you don't actually have. And the painful truth is that real life is riddled with *dark data* problems. Dark data are ubiquitous. It means that, as well as looking at the numbers we see in front of us, we need to ask *what am I not seeing? What is missing? What are these numbers concealing?* What is worse, these dark data can have very serious consequences. More than just meaning your analysis is wrong, it can result in lost fortunes and even deaths. Here are some examples.

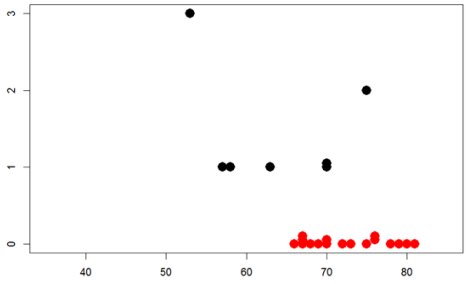
How it all went horribly wrong

You may well be familiar with the 1986 Challenger Space Shuttle disaster, in which the teleconference the night before studied the relationship between O-ring distress and air temperature at launch for those launches which had had O-ring problems. As the Rogers Commission report on the disaster says ([1], p146-7):

"there is nothing irregular in the distribution of O-ring 'distress' over the spectrum of joint temperatures at launch between 53 degrees Fahrenheit and 75 degrees Fahrenheit."

But what was revealing about the plot was not so much as what it showed, but as what it didn't show. In particular, no data was included for those launches which had had *no* O-ring problems.

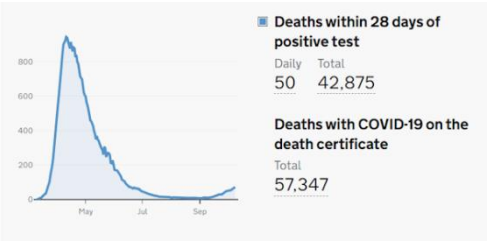




Why include data which appeared to show nothing about the problem being studied? But including those extra (dark data) points in the graph (red points in graph opposite) gives a completely different picture – so different that it is impossible to believe the launch would have gone ahead if the conference participants had seen it. The original data plot had just seven points. Seven astronauts died in the disaster.

A completely different kind of dark data arises from ambiguities or differences in definitions. The police recorded crime (PRC) statistics show that crime in the UK went up between 1997 and 2003 (from 4.6 million to 5.5 million), but the Crime Survey for England and Wales (CSE&W) show that crime *dropped* over the same period (from 16.5 to 12.4 million). If you had studied just one of these data sets you might be very misled – especially if you were really interested in the question answered by the other. The results are different, of course, because the two sources used different definitions of what a crime was. The CSE&W is a

victim survey, asking people what crimes they have experienced. That means it omits some – like possession of drugs. On the other hand it will include crimes that people didn't report to the police (perhaps because they felt no action would be taken). In contrast, the PRC omits those not reported. And there is a complicated potential feedback effect: the number of people arrested for drug possession will depend on the level of police activity, and that will depend on the perceived need for police activity.



And for a third, and again very different type of dark data, consider what summary statistics conceal. You may learn that a particular country has 20 cases of Covid-19 per day per 100,000 in the population. But what this doesn't tell you is that Covid-19 outbreaks tend to be clustered. A look at a map of where the cases occur will show you hotspots, where the rate is much higher than 20 per 100,000, and the rest of the country where it is much lower. In fact, the coronavirus pandemic is a rich source of examples of dark data. Early reported infection rates were wildly misleading because they reported only people with symptoms, and only those that came forward at that. Death rates can be misleading, according to whether they report people who died *of* Covid or *with* Covid, and how effective the strategy is for identifying such people. Tests are not perfect, and always have a non-zero rate of false positives and false negatives. If critical risk factors like age and deprivation are not measured we will not learn how important they are: what else have we missed? All of these, and other causes, serve to conceal what is really going on and have the potential to mislead any analysis.

"The coronavirus pandemic is a rich source of examples of dark data"

A taxonomy of dark data

These examples show that dark data can occur (or perhaps that should be 'not occur') in all sorts of ways. In my book *Dark Data: Why What You Don't Know Matters* [2] (and see https://darkdata.website/) I give a taxonomy of 15 types of dark data. These are not mutually exclusive – just because your data suffers from one of them does not mean it will not also suffer from others. These 15 types are:

1. *Data we know are missing*
2. *Data we don't know are missing*
3. *Choosing just some cases*
4. *Self-selection*
5. *Missing What Matters*
6. *Data Which Might Have Been*
7. *Changes with Time*
8. *Definitions of Data*
9. *Summaries of Data*
10. *Measurement Error and Uncertainty*
11. *Feedback and Gaming*
12. *Information Asymmetry*
13. *Intentionally Darkened Data*
14. *Fabricated and Synthetic Data*
15. *Extrapolating beyond Your Data*

The first two will be familiar to you, because they are Donald Rumsfeld's famous 'known unknowns' and 'unknown unknowns'. A familiar example of the first would be a missing value in a table giving ages of people. If a value is missing, we still know that that person had an age; we just don't know what it is. An example of the second is given by the Challenger disaster: the participants in the teleconference did not recognise that they were missing (crucial) data. Another example of this second type arose as a consequence of a problem with a NASA satellite, which suggested that the extent of fires in August 2020 in Brazil's section of the Amazon rainforest was 5% lower than the previous August [3]. It was only later that extra data showed this was not true, and in fact the number of fires had increased over the year [4].

The third and fourth type of dark data above, choosing just some cases and self-selection, are particularly common and pernicious types of dark data. If people are allowed to decide whether or not to be included in a database (to share their medical records, for example) then there is huge potential for the database to be non-representative of the population, with clear risk of subsequent analysis being misleading. If all the people who begin to feel better drop out of a clinical trial, so that you don't know that they are recovering, the implications for the conclusions are obvious. The Gartner organisation focuses on type 3 dark data (also called 'cold data' by some), ignoring data which have been collected and stored [5] but not analysed – perhaps they were collected for reasons of regulatory compliance.

And so it goes on. Dark data are ubiquitous, cropping up in all sorts of unexpected ways, possibly misleading you with potentially serious consequences.

"I have to chuckle whenever I read yet another description of American frontier log cabins as having been well crafted or sturdily or beautifully built. The much more likely truth is that 99% of frontier log cabins were horribly built—it's just that all of those fell down. The few that have survived intact were the ones that were well made. That doesn't mean all of them were." Mike Johnston at The Online Photographer [7]

What can we do about dark data?

So, that's the problem. But what can we do about it?

An obvious strategy which can sometimes be applied (at least if you know you have dark data) is to get more data. Larger samples reduce measurement inaccuracy, a sample of applicants you would normally have rejected gives you more information on overall population risk characteristics (so, for example, you can build better credit scoring models), follow up those who don't respond to surveys, and so on.

These are all sensible things to do. But there's an old saying that all complex problems have simple, straightforward, and easy-to-understand wrong answers. The same applies to strategies for coping with dark data. While basic statistics textbooks don't deal with the problem at all, implicitly encouraging people to ignore it, some statistical software packages include such simple strategies as default solutions, encouraging people to think they are reasonable approaches. One such strategy is *complete case analysis*, where you drop incomplete records. That's fine, but on the one hand you might discover it leaves you with very little data left to work with, or even none at all, while on the other hand how do you know that the complete cases are properly representative of the population? You could be being seriously misled.

Anthropologist Stephen Jay Gould wrote: "In July 1982, I learned that I was suffering from abdominal mesothelioma … mesothelioma is incurable, with a median mortality of only eight months after discovery." But what Gould realised was that that average mixed together all sorts of different people. He went on to say "half the people will live longer; now what are my chances of being in that half. I read for a furious and nervous hour and concluded, with relief: damned good. I possessed every one of the characteristics conferring a probability of longer life: I was young; my disease had been recognized in a relatively early stage; I would receive the nation's best medical treatment; I had the world to live for; I knew how to read the data properly and not despair." Gould lived for 20 years after his diagnosis, eventually dying of an unrelated cancer. [8]

Another simple strategy is to use all the data you have. Calculate the average height from those who gave their height, and the average weight from those who gave their weight. That's also fine – except what if the shorter people and the fatter people were more likely not to answer. Your two averages could give the impression of a population of very tall very thin people. Worse, correlations calculated from pairs of variables in this way can lead to contradictions: the sample correlation between x and y, and between y and z, contradicts that between x and z.

A third simple strategy is imputation: replace all the missing age values by the average of the ages recorded, for example. Except that this will underestimate the variance of age in the data. Having said that, modern sophisticated methods for coping with dark data – like the so-called *EM-algorithm* – are based on elaborations of this idea.

"And that, perhaps, is the most important message … be suspicious, recognise that you might be missing something"

To cope with dark data effectively, it is necessary to have an understanding of why it is dark. That means, in the first place that you have to know it is lurking out there. And that, perhaps, is the most important message of this article: *be suspicious, recognise that you might be missing something*. Having done that, you can try to detect it. While detection is very context dependent, varying from situation to situation and data set to data set, there are general strategies, like using expected properties of data – like the Benford distribution [6], for example.

Data can be dark for a huge variety of reasons, but a useful categorisation is into data which are unobserved for reasons unrelated to the study (e.g. sickness preventing students from sitting an exam), data which are unobserved for reasons related to data you do have (e.g. exam results preferentially missing for students who previously did badly on tests), and data which are unobserved because of the values they would have had (e.g. a student realising he had revised for the wrong exam, and so skipping it). I call these *Not Data Dependent*, *Seen Data Dependent*, and *Unseen Data Dependent* respectively in my book.

Using dark data to your advantage

My message is that dark data are ubiquitous and dangerous. But there is more. In fact, you can take advantage of dark data, in what I call the strategic application of ignorance. I do not mean hiding information from people, as in fraud, but something much more ethical. In fact, you already do this in other contexts, such as when you use a computer password (itself dark data) to keep data dark from others. And statisticians do it all the time, in all sorts of other ways. To spot this, it is necessary to look at statistical analyses from the other perspective, the dark data perspective.

So, for example, sample surveys study just a (carefully randomly selected) portion of a population – they keep the remainder dark – while randomised controlled trials conceal the true treatment each person receives.

But there's even more. Some of the most sophisticated tools use unobserved *data which might have been*. For example:

- in simulation, where data *which could have arisen from a model* are generated
- in regularisation, equivalent to slightly randomly perturbing the observed data
- in bootstrap methods, which take multiple subsamples from the observed data
- in boosting methods which replicate misclassified cases and force models to focus attention on them
- in Bayesian methods, which combine the data you have got with imprecisely specified extra data in the form of priors.

While it is possible to take advantage of dark data ideas, my main aim in this article has been to draw attention to the risks. These risks have been with us since we began collecting data, and they have not gone away with computers and big data. Computers, while enabling us to do amazing things, necessarily stand between us and the data. They enable us to see projections – summaries, plots, and so on – of the data. And big data do not erase the dark data challenges. Rather, they amplify the risks: you cannot examine your billion point data set point-by-point (there are only 32 million seconds in a year) but have to rely on computer summaries. Big data have all the problems of small data, and extra problems of their own.

It is always necessary to ask: *what might we be missing*?

References

[1] Rogers Commission: Report to the President by the Presidential Commission on the Space Shuttle Challenger Accident, June 6th 1986, NASA, Washington DC

[2] Hand, D.J. (2020) Dark Data: Why What You Don't Know Matters. Princeton University Press.

[3] www.reuters.com/article/us-brazil-environment-fires-exclusive/exclusive-brazil-amazon-fires-likely-worst-in-10-years-august-data-incomplete-government-researcher-says-idUSKBN25T349

[4] unearthed.greenpeace.org/2020/08/13/amazon-rainforest-fires-brazil-protected-areas

[5] www.gartner.com/en/information-technology/glossary/dark-data

[6] Berger A and Hill T (2015) An introduction to Benford's law. Princeton University Press, Princeton

[7] theonlinephotographer.typepad.com/the_online_photographer/2017/02/i-find-this-a-particularly-poignant-picture-its-preserved-in-the-george-grantham-bain-collection-at-the-library-of-congres.html

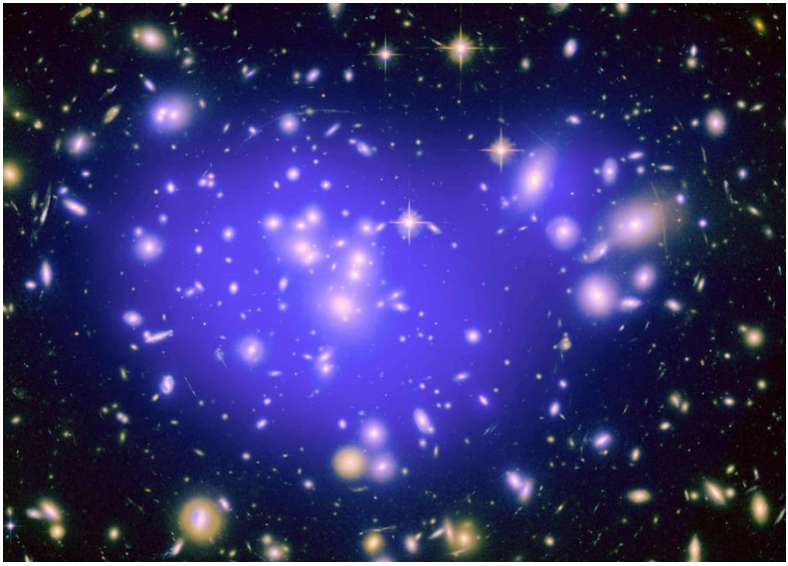
[8] csn.cancer.org/node/213889

Reference websites accessed Oct 2020.
Image Attributions
Lead article: 191433897 © Angkana Kittayachaweng | Dreamstime.com
Covid-19 graph (from coronavirus.data.gov.uk/deaths) licensed under Open Government Licence v3.0

Professor David Hand, Imperial College London

David Hand is Emeritus Professor of Mathematics and a Senior Research Investigator at Imperial College, London. He is a Fellow of the British Academy and a former President of the Royal Statistical Society. His 31 books include *Principles of Data Mining*, *Measurement Theory and Practice*, *The Improbability Principle*, *The Wellbeing of Nations*, and *Dark Data*.

Dark Data and Safety



Mike Parsons discusses the results of a focussed analysis by the Data Safety Initiative Working Group on Dark Data. The assessment looked at the safety implications arising from the various types of Dark Data and the plans to provide more specific guidance to address Dark Data safety risks.

At the last SCSC Data Safety Initiative Working Group (DSIWG) meeting (DSIWG#55, [1, 2]) the group had some initial thoughts on the implications of David Hand's work on Dark Data (see the previous article) for the Data Safety Guidance (DSG) [3]. It was suggested that we need to go through all the data categories in the current guidance and look for common "dark" examples, for example:

- DSG Category 3: **Requirements data** may not be formally written down
- DSG Category 13: **Staff and training data** may be missing or may be falsified
- DSG Category 23: **Justification data** may be missing e.g. for a COTS component

In the meantime, the meeting examined the previously identified categories of Dark Data:

Dark Data Categories and Safety Examples

1. **Data We Know Are Missing: "Known unknowns"**

This case is very common in safety justifications where assurance information may be withheld for commercial reasons or does not exist, but we know, or are informed, that it isn't available. Common examples include:

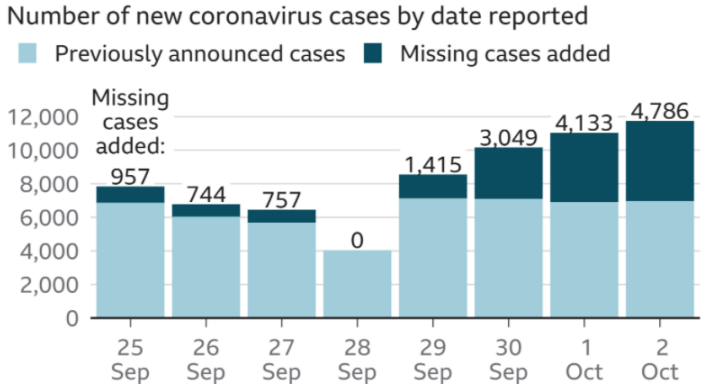
- Assurance information for COTS components
- Technical information about legacy systems

If this is the case we can mitigate in several ways, including use of warnings, training, restrictions of use, etc. We can also substitute with other more indirect assurance e.g. established organisational track-record in the sector, or audit reports.

2. **Data We Don't Know Are Missing: "Unknown unknowns"**

This case is hopefully less common than (1), but we have to acknowledge that it happens. Some examples are:

- The recent Covid-19 Track and Trace data loss [4], where the organisation handling the data was unaware that rows were missing from a spreadsheet for some time



- Somebody knows a problem with a system (i.e. has counter-evidence regarding the safety claim) but does not communicate this to the person creating the safety position

In many cases the loss may be discovered after some period, and it is incumbent on the organisation involved to analyse the impact of the missing data over the time period, including subsequent decisions and actions. Its effects should not be underestimated. This case can fundamentally change the safety picture and is probably in the highest safety risk category.

The DSIWG discussed the issue of data that is discovered to be missing but then subsequently found. It may still be available so what should be done with the "rediscovered data"? Four options were identified: (1) Apply the missing data; (2) Ignore the missing data (3) Mention that the data was missing but not take it into account and (4) perform an impact analysis on the missing data and then act according to the results.

3. **Choosing Just Some Cases**

This is where something or somebody has been selective. Examples might be:

- Selection of test runs that succeeded (ignoring failed runs and their diagnostics)
- Selective sampling from sensors, or where the sampling intervals are chosen badly
- Incorrect filtering of the data, leaving out more cases than intended

Note that with complex or informal criteria the effects could be as case (2), i.e. you don't know what has been left out.

4. **Self-Selection**

This was considered to be similar to case (3), but could be even more informal or ambiguous.

5. **Missing What Matters**

This was considered similar to case (2) in impact terms. Examples might be:

- Measuring the wrong things, e.g. poor safety metrics / indicators
- Too much data to deal with or analyse, so some is ignored
- Too much filtering or processing, so losing information along the way
- Being too close to the data, i.e. the "wood for the trees". This is when the detail masks the overall issue with data, e.g. a slow trend or bias masked by peaks.

6. **Data Which Might Have Been**

This was considered an interesting case, an example might be:

- Inappropriate system architecture e.g. single data channel when a multiple channel approach should have been used, such as in the Boeing 737 MAX accidents [5]

7. **Changes With Time**

This was recognised as a common problem in a safety context. Data in safety systems often becomes obsolete or out-of-date and may still be mistakenly used. Some examples are:

- System configuration data not kept up to date as software or hardware changes
- Medical drug interaction databases
- Software patches require updates to configuration or system data – not always done.

8. **Definitions of Data**

This was considered a common case for systems with databases or those exchanging data with external systems, e.g.

- Data schemas in medical record systems often evolve over time. These can render old data obsolete / subject to misinterpretation, and possibly needing migration/translation.

9. **Summaries of Data**

Often seen in data about safety systems and projects:

- Safety metrics or indicators where data is aggregated to create a composite value
- Could be misleading or cause "boundary reactions", e.g. Red-Amber boundary
- Data fusion across multiple sensors

10. **Measurement Error and Uncertainty**

Sensors can degrade and fail over time, especially in harsh environments such as automotive, marine or aviation:

- Sampling techniques can cause artefacts, interval polling interval can be incorrect, etc.
- Data fusion again

11. **Feedback and Gaming**

This can happen in safety justifications or test case production:

- Early production of a safety argument could lead to only those artefacts which support the claim being generated
- Confirmation bias in safety justifications

12. Information Asymmetry

This is common where there are multiple stores or sources of the same data:

- Multiple / backup databases where they are not kept in sync

13. Intentionally Darkened Data

This can and does happen, e.g.

- Defence, security and government sectors where data is purposefully hidden or destroyed
- This could apply to records deleted after an accident to make things 'look better'

14. Fabricated and Synthetic Data

Fabrication is surprisingly common, e.g. medical, policing and maritime sectors. Synthetic data is often used where there are difficulties in producing enough real data with the right characteristics:

- Data is retrospectively entered / patched to make a "clean" record
- Synthetic autonomous vehicle training databases can have issues with artificial data if not realistic

15. Extrapolating Beyond Your Data

Machine learning systems have to cope with extrapolation outside of their training data, but the outcomes may be unexpected:

- Machine learning data, especially real or recorded data that may not contain edge/corner cases

A decision was taken by the meeting to create a standalone appendix on Dark Data in the next version of the data safety guidance (3.3) to be issued in Feb 2021 at SSS'21 [6].

References

[1] DSIWG, "Minutes of Data Safety Initiative Meeting #55", SCSC, scsc.uk/file/gd-main/SCSC-Data-Safety-Initiative-Meeting-55-FINAL.pdf, Oct 2020

[2] DSIWG, "Slides for Data Safety Initiative Meeting #55", SCSC, scsc.uk/file/gd/55th_DSIWG_Slides_v1-899.pptx, Oct 2020

[3] DSIWG, "Data Safety Guidance (Version 3.2)" SCSC, February 2020, SCSC-127E, ISBN-13: 9798601577359, scsc.uk/scsc-127E

[4] Leo Kelion, "Excel: Why using Microsoft's tool caused Covid-19 results to be lost", BBC, accessed 5th Oct 2020, www.bbc.co.uk/news/technology-54423988

[5] Boeing 737 MAX, en.wikipedia.org/wiki/Boeing_737_MAX_groundings, accessed 22 Oct 2020

[6] SSS'21, "Safety-Critical Systems Symposium (SSS'21)", SCSC, February 2021, https://scsc.uk/e683

Image Attributions
Top image: galaxy cluster Abell 1689, with the mass distribution of the dark matter in the gravitational lens overlaid (in purple). NASA, ESA, E. Jullo (JPL/LAM), P. Natarajan (Yale) and J-P. Kneib (LAM). Licensed under the Creative Commons Attribution 3.0 Unported license.
Covid-19 graph: Gov.uk dashboard, PHE. Contains public sector information licensed under the Open Government Licence v3.0

Mike Parsons, SCSC DSIWG Chair

Mike is the SCSC Director and Events Coordinator. He also leads the SCSC Service Assurance, Data Safety Initiative and Covid-19 Working Groups. He is currently a safety engineer at CGI UK working on various healthcare projects. He has been in the business of safety since 1989.

If Music be the Food of Love, Sing On, Sing On, But How?

Guidance on tackling Covid-19 places restrictions on group gatherings, so how can bands and orchestras continue to play music safely together? Peter Ladkin analyses the prevailing droplet-focussed guidelines for musicians and questions whether aerosol dispersion risks have been properly addressed.

Three weeks ago, at the farm store, I bumped into a neighbour who plays bass clarinet with the Bielefeld Philharmonic orchestra ("Bielefelder Philharmoniker" in German: Bielefelders who love harmony. We do). We often talk about breathing technique and mental attitude, her expertise, but now we talk about my speciality, safety, too. Playing musical instruments, especially wind instruments, or singing, together with others in an enclosed space, is as of early 2020 no longer safe.

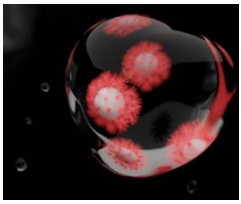
Communal music making is one of the most rewarding experiences many people have. Besides the professionals, we have in my city of 330,000 over a dozen amateur choirs and orchestras, as well as large numbers of bands such as mine. What to do? Make it safe again. This interdisciplinary endeavour involves studies of human anatomy and physiological behaviour, fluid dynamics, ventilation and germicidal engineering. Here is a guide to answering the title question.

When the music stopped

All public events were cancelled in Bielefeld from 15th March 2020. The city theatre company encompasses opera, musical theatre and spoken theatre, as well as the Philharmoniker. There is a YouTube video series on their channel Theater Bielefeld called Gemeinsame Sache, "[making] common cause". The series features solo performances – readings, dance, music – in various Bielefeld cafes and shops, which were at that time also all closed. Here is my neighbour, Margarete Fiedler, playing a folk tune in classical style on this lovely instrument [1], in the Kachelhaus café – a building covered in white and light-green glazed tiles, called "Kachel".

Margarete told me at the farm store that the Philharmoniker had just commenced again with performances, with a reduced contingent and of course a widely-spaced audience. The orchestra pit in the city theatre has been approved to accommodate 10 musicians. The engineer in me wonders how they got to that number. The theatre was renovated a few years ago to the design of a company called theapro, and there are some photographs in a diorama on their website [2] as well as a geometric floorplan (published without a scale, but some dimensions are given). The Philharmoniker and I are aware that the Institute for Music Medicine (IMM) at the University of Music in Freiburg has published indicators for distancing and other parameters of dealing with Covid-19 for musicians, based on the existing literature as well as some local experiments.

Distancing guidance – for droplets



I read the first version of the IMM guidance (in German and English) when it was first published in April 2020 [3]. It is now available in seven languages, including English [4]. It concerns the Covid-19 risk assessment of the activities of playing and teaching music and singing. I'll refer to the authors and report collectively as IMM. The assessment was updated four times until 17th July 2020, so there have been five versions. Other documents on guidance for musical activities refer to it, such as that of Public Health Ontario [5] and Swalje and Hoffman [6].

The flute family consists of transverse flutes, as well as recorders and other fipple flutes, and there is a Japanese flute which is blown in the same configuration as a fipple flute, that is, held parallel to the air stream, but with a transverse edge without a fipple, but it is not used (much) in Western music. The IMM regards flutes as particularly problematic, because much of the air expelled in playing is not necessarily contained within the instrument (a fair amount is, in experienced players, but less so with beginners). Based on preliminary experiments with symphonic players from Bamberg, as well as considerations derived from the vocal-fluid-dynamic literature, the April edition of the IMM guidance was recommending distancing of 3-5m in the blowing direction for flute players and 1.5m transversely. The document updated on 17th July 2020 relates some experiments in which disturbance of the air around wind-instrument players was measured. Essentially no disturbance was found at 2m distance from flautists and singers. IMM thus revised its advice to 2m radial minimum distancing for all musicians (including singers). This advice addresses the issue of droplets which fall to the ground under gravity.

Something in the air

The issue of aerosols is different, however, and must also be addressed. The distinction is made qualitatively, droplets being those particles which fall to the ground under gravity, and aerosols those smaller particles which remain suspended in the air. For convenience, the distinction is set at a reference level, arbitrary but realistic, of 5μm (droplets greater than, aerosols smaller). While droplets were generally the main concern as aerial vehicles of Covid-19 transmission in March and April 2020, largely it turns out for reasons of tradition, possible Covid-19 transmission through aerosols has been a scientific concern of late.

Aerosols can remain suspended in air in an enclosed space for a number of hours. Mürbe et al say three hours viability for SARS-CoV-2 [7, p4], citing van Doremalen et al [8]. Wendy Barclay, a British virologist and member of SAGE, suggested a practical one-hour viability in a BBC television interview in July 2020 [9].

"Aerosols can remain suspended in air in an enclosed space for a number of hours"

If we assume aerosols containing viable virus accumulate at a fixed rate, then as a first approximation the amount of aerosol in the air at a specific time follows a linear progression to a stable state, as follows. Assume viable virions in aerosol (vva) accumulate at a rate of V per unit time. Start at Time 0 with 0 vva. At 1 time period, there are V vva; at 2 time periods there are $(V + V)$ vva; ….. at n periods there are $V \times n$ vva. At some point those virions become unviable; let us assume after T time units. Call T the vva persistence. At 0 time, there are no aerosols; at T time units, the persistence limit of the initial V units expelled between time 0 and 1 has been reached, and they will all become unviable up to time $(T + 1)$, but V new vva will have been added also between T and $T + 1$, so the vva present at $T + 1$ is the same number as were present at T. A stable vva count is therefore reached at time T, and that will be $V \times T$. Obviously, this is a discrete approximation of a continuous process. Does it yield an accurate result?

A better approximation might be to take 1 time unit equal to the half-life of vva accumulations. Van Doremalen et al found this to be around an hour [8, Figure 1]. Let us take V to be the number of expelled vva that survive at 1 time period. As before, at time 0 there are 0 vva; at time 1, V vva; at time 2, $(V + ½ V)$, since only half the vva from the preceding time period are still viable, and there are V new vva.
At time 3, there are $V + ½ (V + ½ V)$; and so on. We have the recurrence relation:

- $a_0 = 0$;
- $a_n = ½ a_{n-1} + V$ for $n > 0$

which has the general solution $a_n = V \times (2^n - 1)/2^{n-1}$ which obviously converges to $2V$ in the limit, from below. The maximum concentration attained in the room is therefore a little below twice the amount of vva accumulated during a half-life period. By twice the half-life period it is only one and a half times V, but by three or four times the half-life period it is pretty close to twice V.

Estimating the steady-state vva in the space by the first method, with persistence at three hours, leads to steady-state accumulation of $3 \times V$. The second estimation with half-life of one hour leads to an accumulation approaching $2 \times V$ from below, which is already

somewhat better. We shall see below that IMM says that an air exchange of 6 times room volume per hour (or a 10m ceiling!) yields "sufficient" removal of aerosols, and, failing that, thorough room ventilation every 15 minutes. If V is the accumulated vva per hour, extrapolating linearly we estimate maximum vva accumulation to be $\frac{1}{6}V$, respectively $\frac{1}{4}V$, which is a twelfth, respectively an eighth, of the steady-state unventilated accumulation. But they are both still multiples of V. So if V is, say, more than 6 times an infectious dose, even following IMM air-exchange guidelines you still have an infectious accumulation in the room. Looking at it this simple way, having seven infectious people in the room might be problematic for others even if you ventilate often! But the length of time of exposure also seems to play a critical role in transmission of Covid-19; the 15-minute guideline was also used in the definition of "close contact" in initial general advice about infection opportunities in February and March 2020.

What is an infectious vva concentration? No one yet knows what it takes in viable-virion intake to become infected, as noted by Germany's "superstar" virologist, Christian Drosten, in a recent interview [10]. It is implausible that there would be just one "magic" number – people likely differ as to their susceptibility to infection, and it does depend on how the virus can enter into your body – but a reference value would be helpful. There isn't one.

A refreshing approach to decontamination

If one is refreshing, respectively decontaminating, air constantly, then stable levels of virion and thus persistence and accumulation do not matter as long as the refresh/decontamination rate is lower than the persistence as well as low enough to keep current aerosols at a non-infectious level.

Bahl et al reviewed the propagation of breath-expelled aerosols and droplets in general [11] (note that the paper was submitted on 26th Feb 2020, so relatively early on in the progression of Covid-19). It includes recent work of Lydia Bourouiba's MIT Laboratory [12]. Bahl et al noted that most of the more recent papers report droplet ranges of over 2m, depending on the size of the droplets – 16µm droplets were reported to travel up to 7m, and Bourouiba reported just short of 8m range for all droplet sizes. The scientific issue with all these experiments is the accurate detection of smaller droplets during experiments.

"What might not have been an infectious concentration at first could turn into one during accumulation"

Aerosols remaining in suspension in air in enclosed spaces used for musical activity thus pose a problem. They accumulate, to a limit, and what might not have been an infectious concentration at first could turn into one during accumulation. IMM recommends conducting musical activities outside, under distancing, where possible, since the threat from aerosols in the open air is thought to be negligible. When inside an enclosed space (room, hall) IMM says that the size of the room, the number of people present, and the length of time spent in the room all seem to play a role, and reference a 2006 review of aerosol transmission of influenza B virus [13].

A recent addition to the IMM document says "*from an air exchange rate of 6/h on, a sufficient removal of aerosols can be assumed*" [4, p16]. "6/h" means 6 times room volume per hour, which I shall write as 6 Rvol/hr. The authors mean a replacement with air from outside the room space, assumed to be pathogen-poor. I summarise the IMM general recommendations as follows:

- as much space as possible ("*cathedral situation*"); they speak of a ceiling height of 10m or more being comparable to an air exchange of 6 Rvol/hr [4, p16];
- when possible, air exchange of at least 6 Rvol/hr [4, p16]
- if this air exchange rate is not available, thorough indoor ventilation (without occupants) of smaller spaces every 15 minutes, or using rooms with an HVAC system (presumably there are some requirements for the HVAC system, but these are not specified); this amounts to 4 Rvol/hr, although such replacement will not be as perfect as this arithmetic implies
- assessment of CO_2 amounts, via a CO_2 "traffic light" based on the Pettenkofer number (1000ppm), since the quantity of CO_2 is dependent on outbreathed air, which is where aerosols (in whatever concentration) originate also. So, when the light goes red (>1000ppm), evacuate and ventilate.

There are two issues which I see arising from this advice.

First, 6 Rvol/hr is regarded as in any event sufficient, however "ventilation" every 15 minutes, which amounts to 4 Rvol/hr (if perfect), is regarded as adequate. Those are two different amounts. Is 4 Rvol/hr adequate or not? If so, the 6 Rvol/hr figure is superfluous. If not, the "ventilation" requirement should surely be raised to "every 10 minutes" to maintain consistency.

Second, it is not clear to me why it should be that the Pettenkofer number itself is key. Obviously it is key from the point of view of CO_2, and the authors are assuming that V is related linearly to CO_2 exhalation, which seems reasonable. But, as noted above, nobody knows what the not-infectious/infectious vva limit is. Why should it coincide with the Pettenkofer number?

The guidance in practice

What are the practical consequences of this guidance?

My music space is sparsely furnished, somewhat over 7m in length and 3.5m in width, with an average ceiling height of about 2.75m, which makes it just short of 80m^3 in volume. Using the previous IMM figure of 5m in front of flutes, we have played during rainy weather indoors with three flutes (one transverse, two fipple) in three of the four corners, with open windows to the outside near all players and a good draft coming through. That fits the original IMM guidance, but is difficult when the weather is cold.

Under the latest IMM guidelines of 2m radial distancing I could likely manage to configure 4-5 players geometrically, good enough for chamber music and small traditional-music sessions, but not for my 6-singer choral group (maybe one of us can sing in the corridor outside, like Harry Secombe in the Ying Tong song [14]). However, winter will make ventilation problematic. Breaking every 15 minutes for ventilation might work for professionals, but will likely put off amateurs having or trying to have fun. Which brings us to filtering, ventilation and disinfectant devices.

Germicidal irradiation

Let us start with ultraviolet germicidal irradiation, UVGI. There are various UVGI devices on the market, in the UV-C band, which is most effective in destroying RNA/DNA of microorganisms – as well as damaging human eyes. Such devices can be installed inside existing HVAC systems moving building air, but if you don't have one (most older houses,

like mine, do not) then you could consider a standalone or wall-mounted device. Such devices suck room air into a tube enclosing a UV-C lamp and evacuate it out the other end. No UV-C radiation makes it outside the tube.

For microbe disinfection, including viruses, the most effective wavelength is said to be 260-265nm (the DNA absorption peak). Mercury vapor lamps produce 95% of their energy at 253.7nm, which is regarded as effective germicide [15, p18, esp. Figure 2.1]. Mercury is of course highly poisonous, and thus requires careful handling (don't drop that lamp!). There are said to be LED variants but I haven't found any. I looked for such devices in stand-alone variants, considering only those for which relatively detailed performance parameters are publicly available. The least expensive sells for around €500, and achieves nominally 80m^3 per hour throughput. The most powerful standalone device I found achieves nominally 300m^3 per hour throughput, and costs over ten times as much at €5,200 wholesale (retail, add €1,000-2,000). So the inexpensive device achieves a circulation rate of 1 Rvol/hr in my music room, and the most expensive 3.75 Rvol/hr, both short of the rate of 6 Rvol/hr highlighted by IMM, but the higher-performance device does almost attain the "ventilation" requirement of 4 Rvol/hr. Since these devices only experience flow resistance internally, one could expect actual throughput performance to be close to nominal. The noise the devices make when operating is relevant, but the levels seem acceptable at around 30dB per device. So it is possible, although expensive, for me to achieve appropriate levels of clean air/replacement air through UVGI devices on the market.

Another option for non-ventilated rooms is an "upper room system", which UV-irradiates the upper part of a room, above the people, during use. The idea is that air containing aerosols emanates from people and is warmer than the ambient, thus rising into the upper part of a room, where it is then disinfected by the UV-C radiation. Additional circulation mechanisms might be required: "*[g]ood air mixing is necessary*" for effectiveness; "*[w]hen rooms lack adequate air movement, the addition of mixing fans is a satisfactory solution*" [15, Chapter 9]. Such systems have been in use "*for decades, and data that has accrued from installations in hospitals, schools, army barracks, and other indoor environments has shown them to be predictably effective provided the systems are properly designed and appropriate safety precautions are taken*." [15, Chapter 9].

One of the main safety issues with upper-room systems is the level of reflected UV-C radiation which arrives where the people are. UV-absorbing paints are available, which can help. Kowalski cautions about use of such systems where floor-to-ceiling heights are lower than 2.3m, because of the hazard of eye exposure. Mine is 2.70m, resp. 2.78m. It does seem as if significant room-preparation is likely to be necessary, as well as testing – "*[i]nvariably, photosensor measurements will be taken to establish that no hazardous levels of UV exist below the UV zone in the habitable areas*" [15, Chapter 9].

How much irradiation do aerosols need to ensure adequate SARS-CoV-2 virion destruction? One speaks of log 1 (90%), log 2 (99%), log 3 (99.9%) and log 4 (99.99%) levels of disinfection. Hessling et al surveyed current knowledge concerning UV-C and coronavirus inactivation in May 2020 [16]. The authors conclude "*To date, UVC radiation has been effective against all coronaviruses in all published investigations, although the absorption properties of the sample media reduced inactivation success. The calculated upper limit for the log-reduction median dose (in low-absorbance media) is 10.6 mJ/cm^2, but the probably more precise estimation is 3.7 mJ/cm^2.*"

Heilingloh et al achieved complete inactivation of SARS-CoV-2 at a concentration of 5×10^6 $TCID_{50}$/mL with a dose of 1048 mJ/cm^2 [17], about 100 times the log-reduction median dose reported by Hessling et al.

How much aerosol is emitted by singers or flute players? For this, we would turn to the results of the PERFORM project, led by the chemist Jonathan Reid and singer/surgeon Declan Costello [18]. The relative difference in emissions between speaking and singing is a factor of 1.5-3.4, which, while statistically significant, is not the real driver. That comes with the volume of the activity. Higher-decibel activity is associated with 20-30 times the emissions of quieter activity. As far as I know, results on instrument playing are not yet available.

Gregson et al [18] emphasise the relative quantity of emission. Since vva levels for infection are not known, it is not known how much virus in emissions from speaking/singing leads (with varying probability) to infection. It follows that the amount of irradiation to which room air should be exposed over what period of time, in order to eliminate an infectious concentration of aerosols, has not yet been established. Heilingloh et al's results concern exceptionally strong concentrations of virus in liquid preparation, which will attenuate the UV some, on a dish. Aerosols will afford nowhere near that concentration, and become more or less dry in quite a short time, so there will be no attenuation through liquid.

Filtration

Now to HEPA air filtration devices. They are commercially available, also stand-alone, filter out particles down to 3µm, but can be loud. It requires a lot of work to push-pull the air through the dense filter elements , and some of that work turns into noise – up to some 50+dB, which makes such devices impractical for medium-size music rooms [19], although some devices are claimed on their data sheets to be quieter. Filtering is effective down to 3µm, but a SARS-CoV-2 virion is some 20-30 times smaller than that [20], suggesting that an in-room recirculation device may not be effective prophylaxis against Covid-19.

Ventilation

So, on to ventilation. IMM seems fairly clear that effective air exchange is a solution to vva problems. Is it achievable in a music room retroactively? Household ventilation systems are widely available, and are widely installed in modern-build houses for a variety of reasons. Common models for entire houses shift nominally 350-600m^3 per hour, although I understand practical effectiveness is quite a bit less than that (some 60%?) inter alia because of air flow resistance in the tubing [19]. I understand they are generally aimed at a per-house rate of ½ House-Vol/hr [19]. There are also per-room installations, but they don't shift air at anything like the IMM-recommended rates. All these systems work to the same general principle: air is drawn from outside, prewarmed via a heat exchanger, filtered (at varying quality levels), circulated through the occupied space, filtered again, cooled in the heat exchanger, and exhausted to the outside. A key condition for such systems to work effectively is relatively good sealing of openings such as windows and doors, and there are standards and tests for such sealing capacity [21].

The ventilator itself is about 30dB loud, maybe louder at full power (my informal observation), but of course can be installed remotely. Such a system could be configured to ventilate just my music room. Just for this one room, I imagine airflow could be near-nominal: 450m^3 per hour yields 5.6+ Rvol/hr, and 350m^3 per hour yields 4.4- Rvol/hr.

However, as well as the structural installation of the circulation unit and pipes, room windows would almost certainly have to be replaced and the two doors modified to achieve the necessary sealing [21]. The devices themselves cost €3,000-4,000 plus tubing and installation; and then comes window replacement and door modification, so I would be looking likely at €8,000-€10,000. Whether to go with a higher performance device or a lower depends on whether I would consider the 6 Rvol/hr figure or the lower 4 Rvol/hr definitive, and we have seen that that is not decided. In any case, such an installation would be working to a different specification than the usual one (½ Rvol/hr), and there is no guarantee how that would work out.

The legal position

My final theme is law. Musical activity in public spaces (whether rooms or open air) is now part of law in my state of residence, North-Rhine-Westfalia in Germany, and most amateur choirs or music ensembles use public space in this sense to practice, and of course for performance.

Current law [22, §8(5)] regulates the use of public spaces (as above) for music and singing, and is being revised monthly. For 1st Oct 2020 to 31st Oct 2020, it simply requires adherence to the accompanying hygiene and infection-protection standards. Those standards [23, Section XII], also time-limited, are relatively detailed, encompassing how wind instruments should be cleaned during use, washing hands afterwards; how players should be situated (flutes in the front row of the orchestra), transparent material separating wind instruments from others, and so on. 2m radial distancing for everybody; 4m from musicians to audience; no audience of any sort in practice rooms. Apparently in the previous version (for September) it was required that each musician should occupy $7m^2$ of floor space, but in the current version that condition is no longer present [24, comment on differences].

I cannot say I am content with this guidance.

It is good on distancing, and on cleaning instruments. For example, dispose of your cloth after pulling through (There is no way I am going to throw my favourite chamois away, but I should obviously think how to clean differently if/when I play in ensembles outside my house). When you open the condensation-release valve on your trumpet, don't just shake out into the environment – catch what comes out, dispose of it appropriately, and disinfect. And then wash your hands. All good stuff.

But all to do with droplets and fomites. There is nothing on adequate ventilation/disinfection of the room, concerning aerosols. Which is what this article has mostly been about.

References

[1] YouTube channel Theater Bielefeld, GemeinsameSache – Margarete Fiedler im Kachelhaus, at www.youtube.com/watch?v=j8gDom2jJvk, accessed 15th Oct 2020.

[2] theapro GmbH, Projekt Stadttheater Bielefeld, at theapro.de/projekt/bielefeld-sanierung-stadttheater, accessed 15th Oct 2020.

[3] Hochschule für Musik Freiburg, Risikoeinschätzung einer Coronavirus-Infektion im Bereich Musik – viertes Update vom 17. Juli 2020 (Risk Assessment of a Coronavirus Infection in the Field of Music), links to the report in seven languages, at www.mh-freiburg.de/hochschule/covid-19-corona/risikoeinschaetzung.

[4] C. Spahn & B. Richter, with collaborators, Risk Assessment of a Coronavirus Infection in the Field of Music, Hochschule für Musik Freiburg, Freiburger Institut für Musikermedizin, fourth update of 17th July 2020 https://www.mh-freiburg.de/fileadmin/Downloads/Allgemeines/RisikoabschaetzungCorona-MusikSpahnRichter17.7.2020Englisch.pdf, accessed 15th Oct 2020.

[5] Public Health Ontario, Santé publique Ontario, Synopsis: COVID-19 Transmission Risks from Singing and Playing Wind Instruments – What We Know So Far (dated 7th Sept 2020). Available at www.publichealthontario.ca/-/media/documents/ncov/covid-wwksf/2020/07/what-we-know-transmission-risks-singing-wind-instruments.pdf, accessed 15th Oct 2020.

[6] A. T. Schwalje & H. T. Hoffman, Wind Musicians' Risk Assessment in the Time of Covid-19, University of Iowa Health Care, Iowa Head and Neck Protocols, 14th August 2020. medicine.uiowa.edu/iowaprotocols/wind-instrument-aerosol-covid-era-covid-19-and-horns-trumpets-trombones-euphoniums-tubas-recorders, accessed 15th Oct 2020.

[7] D. Mürbe, P. Bischoff, M. Fleischer & P. Gastmeier, Beurteilung der Ansteckungsgefahr mit SARS-CoV-2-Viren beim Singen (in German), Charité, Klinik für Audiologie und Phoniatrie/ Institut für Hygiene und Umweltmedizin, 4th May 2020. audiologie-phoniatrie.charite.de/fileadmin/user_upload/microsites/m_cc16/audiologie/Allgemein/Singen_und_SARS-CoV-2_Prof._Mürbe_et_al._04052020.pdf, accessed 15th Oct 2020.

[8] N. van Doremalen, T. Bushmaker, D.H. Morris et al, Aerosol and Surface Stability of SARS-CoV-2 as Compared with SARS-CoV-1, Correspondence, New England Journal of Medicine 2020; 382:1564-1567 DOI: 10.1056/NEJMc2004973 , 16th April 2020. www.nejm.org/doi/full/10.1056/NEJMc2004973, accessed 15th Oct 2020.

[9] W. Barclay, Interview on the Andrew Marr show, BBC Television, 12th July 2020. www.youtube.com/watch?v=4kG87HpWpc0, accessed 15th Oct 2020.

[10] C. Drosten, It's Up To Us, interview with F. Schumann and J. Simmank, Die Zeit, 8th Oct 2020. www.zeit.de/wissen/2020-10/christian-drosten-coronavirus-infection-winter-virologist/komplettansicht, accessed 15th Oct 2020.

[11] P. Bahl, C. Doolan, C. de Silva, A.A. Chughtai, L. Bourouiba, & C.R. MacIntyre, Airborne or Droplet Precautions for Health Workers Treating Coronavirus Disease 2019?, *The Journal of Infectious Diseases*, 16th April 2020, academic.oup.com/jid/advance-article/doi/10.1093/infdis/jiaa189/5820886, accessed 15th Oct 2020.

[12] The Bourouiba Group, Homepage, lbourouiba.mit.edu/home, accessed 15th Oct 2020.

[13] R. Tellier, Review of aerosol transmission of influenza A virus, Emerg Infect Dis. 2006 Nov;12(11):1657-62

[14] The Goons, The Ying Tong Song, 1950's. www.youtube.com/watch?v=Nebe1zuEtbc, accessed 15th Oct 2020.

[15] W. Kowalski, Ultraviolet Germicidal Irradiation Handbook, Springer-Verlag, 2009.

[16] M. Hessling, K. Hönes, P. Vatter & C. Lingenfelder, Ultraviolet irradiation doses for coronavirus inactivation -review and analysis of coronavirus photoinactivation studies, GMS Hygiene and Infection Control 2020,Vol. 15, ISSN 2196-5226. Available at www.researchgate.net/publication/341371696_Ultraviolet_irradiation_doses_for_coronavirus_inactivation_-review_and_analysis_of_coronavirus_photoinactivation_studies_Ultraviolette_Bestrahlungsdosen_fur_die_Inaktivierung_von_Coronaviren_-Review_u, accessed 15th Oct 2020.

[17] C.S. Heilingloh, U.W. Aufderhorst, L. Schipper, U. Dittmer, O. Witzke, Yang D., Zheng X. K. Sutter, M. Trilling, M. Alt, E.Steinmann, & A. Krawczyk, Susceptibility of SARS-CoV-2 to UV irradiation, American Journal of Infection Control 48(10):1273-1275, Oct 2020. www.sciencedirect.com/science/article/pii/S0196655320307562, accessed 15th Oct 2020.

[18] F.K.A. Gregson, N.A.Watson, C.M. Orton, A.E. Haddrell, L.P. McCarthy, T.J.R. Finnie, N. Gent, G.C.Donaldson, P.L. Shah, J.D. Calder, B. R.Bzdek, D. Costello & J.P. Reid, Comparing the Respirable Aerosol Concentrations and Particle Size Distributions Generated by Singing, Speaking and Breathing, preprint. chemrxiv.org/articles/preprint/Comparing_the_Respirable_Aerosol_Concentrations_and_Particle_Size_Distributions_Generated_by_Singing_Speaking_and_Breathing/12789221, accessed 15th Oct 2020.

[19] M. Schmidt, Schmidt HLS GmbH & Co. KG, personal communication 13th/14th Oct 2020

[20] Y.M. Bar-On, A. Flamholz, R. Phillips & R. Milo, SARS-CoV-2 (Covid 19) by the Numbers, eLife 2020: 9; e57309. www.ncbi.nlm.nih.gov/pmc/articles/PMC7224694, accessed 16th Oct 2020.

[21] J. Niederhommert, Architekturbüro Hammesfahr + Niederhommert, personal communication 16th Oct 2020.

[22] State of North-Rhine-Westfalia, Coronaschutzverordnung (CoronaSchVO), Corona-Protection-Regulation, version of 30th Sept 2020 (valid 1st Oct 2020 until 31st Oct 2020). www.land.nrw/sites/default/files/asset/document/2020-09-30_coronaschvo_ab_01.10.2020_lesefassung_0.pdf, accessed 15th Oct 2020.

[23] State of North-Rhine-Westfalia, Anlage "Hygiene- und Infektionsschutzstandards" zur CoronaSchVO NRW (Appendix "Hygiene and Infection-Protection Standards" to the CoronaSchVO of North-Rhine Westfalia). www.land.nrw/sites/default/files/asset/document/2020-09-30_anlage_zur_coronaschvo_ab_01.10.2020_lesefassung.pdf, accessed 15th Oct 2020.

[24] Chorverband Nordrhein-Westfalen e.V., Homepage variant for "Aktive" (active Singers), available at www.cvnrw.de/index.php?id=start, accessed 15th Oct 2020.

Prof. Dr. Peter Bernard Ladkin

Peter Bernard Ladkin works in system safety and software-based system dependability. He is retired Professor at Bielefeld University, and Managing Director resp. CEO of British and German companies, both called Causalis, providing services in engineered-system RAMSS. His method Why-Because Analysis (WBA) is used worldwide by some 11,000 engineers. His flute playing is improving.

Mistakes & Misconceptions

John Spriggs discusses how mistakes and misconceptions can unwittingly creep into our world and, in particular, when presenting assurance cases. He explains the perils of failing to understand the context or grasp the bigger picture and asks if lessons are being learnt from past failures.

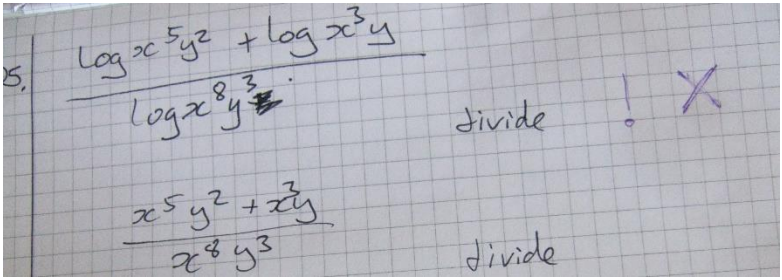
Talking with some school maths teachers recently, I found that one of their key skills is the ability to anticipate misconceptions in a topic, and plan to address them in the lessons before they arise. Similarly, over a period of teaching a topic, they identify common mistakes, and so can put more emphasis on those parts of the topic that cause problems. This is like identifying a vulnerability and implementing appropriate controls (or mitigations for hazards). What mistakes and misconceptions could make your assurance project go pear-shaped?

Mistakes

Some mistakes arise due to a change of context. For example, a student may be perceived as proficient with logarithms but, when a more complicated logarithm problem is posed, they make a basic error. For example, they are told to simplify an expression like (log a + log b)/log c where a, b, and c are scary-looking expressions, like x^5y^2. "This is obviously an algebraic manipulation problem", so the student unthinkingly divides through by log, as if it were a quantity. Such a switch in perceived context led to an assurance mistake I once saw, where a system supplier assured the regulator of their new customer that all was OK, because they complied with Standard X. The assurance was rejected; more work was required. The supplier was surprised, because they had sold an equivalent system into another

regulated industry recently, and compliance with Standard X was all that was required of them then.

What the supplier did not know was that the earlier regulator had their own explicit requirements, which they had placed on the customer, and within those was a note saying that demonstrating compliance with Standard X is an acceptable means of compliance to these requirements. The customer had taken the easy route and just specified compliance with that standard, rather than passing the detailed requirements on. The regulator in the new market had a different set of requirements, and did not identify use of Standard X as an acceptable means of compliance. In this scenario, it is up to the assurer to justify their claims that meeting Standard X also satisfies that regulator's requirements. The mistakes made by the maths student and the supplier are both understandable given the changed context.



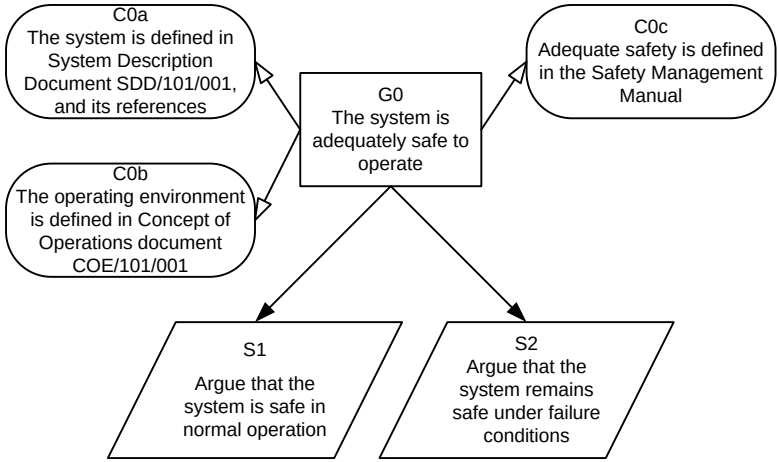
Misconceptions

What about misconceptions? This is where the mental model of a concept does not accord with the real 'mechanism'. This can often arise from a misunderstanding of what a teacher said or, in many cases, due to ambiguity in the way the concept was presented by the teacher. For example, a young student may have written 3.6 X 10 = 3.60 and challenged it being marked wrong, saying, "But you said to multiply a number by 10, just write 0 next to that number, like 5 times 10 equals 50"…

In the assurance world, we have the Goal Structuring Notation as one way to present an argument. It transpires that some people see this name and have the misconception that the notation is all about setting goals and achieving them. From this viewpoint, presenting assurance that "the top goal is met", means that all the sub-goals are "met", and consequently all the required evidence is available to support the lowest level goals. Presenting assurance under this misconception is therefore just accounting of evidence items as if they were all equally important. This view may have been valid for the origins of the notation, which was developed in support of engineering requirements decomposition (and verification composition). In an extreme example (where the argument was prepared by someone else, maybe as a template), as the goals are all "met", it is assumed that we do not need to see them clearly, and so the whole argument structure is presented in a single unreadable diagram, or not presented at all.

The Goal symbol in the Goal Structuring Notation … is just a proposition; a statement that may be either True or False. Truth is in the eye of the beholder.

The Goal symbol in the Goal Structuring Notation, when used to present an assurance argument, is just a proposition; a statement that may be either True or False. Truth is in the eye of the beholder in this context. Those who are to be assured need to see the argument to assess its validity, and how it appeals to the evidence. They can then judge whether the evidence presented is necessary and sufficient for their needs. Just claiming that, "The top goal is met, because the evidence is available for audit" should be insufficient to assure anyone.



Seeing the whole picture

The student's misconception arose from being exposed to only part of the picture. To multiply by ten, you must bounce the decimal point one place to the right, and you append zero if there is then an unprotected decimal point on the end; if there is no decimal point just append zero. We can be exposed to only part of the picture in engineering too. I was once told by a supplier that I would receive the "complete assurance package" the following week, which was a surprise because I was producing the assurance for that project. Another surprise was that it actually arrived midweek; "next week" usually means after five o'clock on the Friday. When it arrived, it was a set of test reports and explanations of how those tests demonstrated satisfaction of the requirements. This was actually what I was expecting from the supplier, as set out in the Contract Schedule. They had been contracted to supply various items, and the main body of the contract, to which I did not have access, specified them under sub-headings, like Training Materials, Maintenance Manuals, and Assurance Documentation.

Surely, there is no problem with a completely understandable mis-naming like that. But, what if that supplier bids for your next project, and the non-specialist negotiators accept their offer of a "complete assurance package", because it is cheaper than your proposed work package? Not a likely scenario, but I did experience something similar many years ago, where a bidder was accepted because they promised full compliance with a standard that had a number similar to the one actually specified…

There is a school of thought that specifying safe behaviour in a set of requirements, and then demonstrating that those requirements are "met", is sufficient assurance for the equipment part of a system. This too is a misconception. I also need to be persuaded, inter alia, that the set of requirements is correct and complete, and also that there are no unwanted interactions. Each requirement has been demonstrated separately, but what if the behaviour arising from one requirement adversely impacts on correct operation of a safety function (mitigation) specified by another? Similarly, satisfactions of the people and procedure requirements may not be sufficient assurance either. Such requirements often include ones of the form, "There shall be a procedure for performing periodic maintenance", for example; satisfaction is then demonstrated by pointing to the procedure in the management system. Assurance will also be required that the right people have been trained to use this procedure, and that records exist of its use with the system.

There is a school of thought that specifying safe behaviour in a set of requirements, and then demonstrating that those requirements are "met", is sufficient assurance for the equipment part of a system. This too is a misconception.

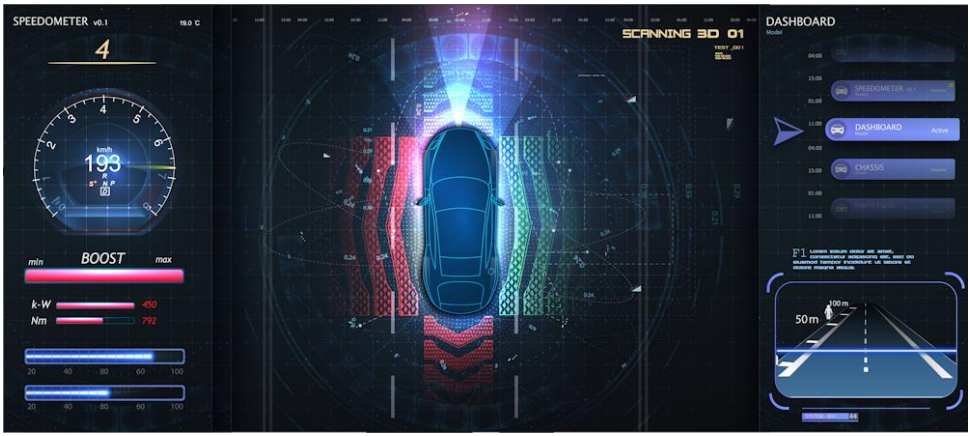
Lessons Learnt?

Can we improve our assurance processes by anticipating misconceptions and preventing common mistakes? Many companies have procedures for lesson learning; most concentrate on impacts to time and cost, but the principles can be extended to extracting lessons learned from assurance failures too. Make sure that the lessons are actually learned, so that the same problems can be avoided in future. It is easy to record the lesson and then forget it. Someone once told me that his employer has a formal meeting at the end of each project intended for lesson learning, but the minutes of the meeting are just archived with the project. There is no requirement to raise actions for process change. I think that is a mistake.

John Spriggs, Independent Author and Presenter

John Spriggs was a designer of avionics systems in the 1970s and '80s. In the 1990s, he became an assurer of equipment for Air Traffic Management and Airports. From 2000, he was the Safety Assurance Consultant for an innovations factory, and later provided assurance for air navigation services providers. John is now an independent author and presenter; he is author of the Goal Structuring Notation textbook.

The Safety of Autonomous Systems: An SCSC Working Group



The SCSC Safety of Autonomous Systems Working Group discusses the history and motivation of the group and presents a summary of the significant progress it has made in developing guidance in the uniquely challenging domain of safety assurance of autonomous systems and technologies.

The Group

In January 2016, the Data Safety Initiative Working Group (DSIWG) released version 1.3 of "*Data Safety Guidance*". Motivated by the continued success of the DSIWG, the Safety Critical Systems Club (SCSC) considered the potential for other working groups. Then, as now, rapid developments were being made in Artificial Intelligence (AI) using Machine Learning (ML) techniques. Autonomous systems, enabled by a combination of AI and ML, were becoming more important. Consequently, the Safety of Autonomous Systems Working Group (SASWG) was formed.

The group held its first meeting in January 2017. As is typical for SCSC working groups, this involved a collaboration between industry, academia and government. This balanced input has been critical to the group's success. It combines innovative research with practical concerns.

Following the DSIWG's example, the SASWG coordinates its main outputs with the Safety-critical Systems Symposium (SSS). In January 2018, a paper outlining "*Safety-Related Challenges for Autonomous Systems*", SCSC-143, was released. Developing this paper helped solidify the group's thinking.

Maintaining the annual drumbeat, the next publication, in January 2019 (for release at SSS'19), was "*Safety Assurance Objectives for Autonomous Systems*", SCSC-153 [1]. This established a structure for the group's guidance material. Whilst, in this edition, large portions of the structure were simply placeholder sections, initial guidance relating to the computational aspect of autonomous systems safety was included. Positive feedback was received.

An updated version of "*Safety Assurance Objectives for Autonomous Systems*", SCSC-153A [2], was produced in January 2020 and released at SSS'20. This contains a complete list of 45 objectives, along with guidance on how they may (at least partially) be met. As such, it provides useful guidance for those wanting to develop a safety (or assurance) argument for an autonomous system. Reaching this point just three years after the group's first meeting is a significant achievement.

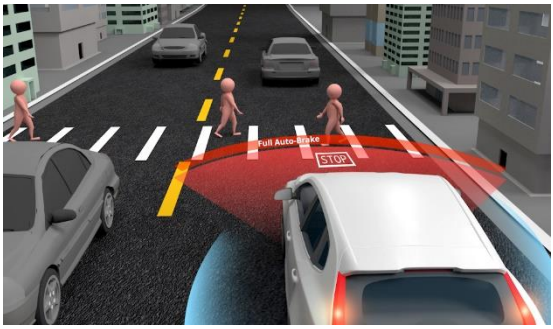
It is a great credit to the SASWG Chair, the volunteers' employers and, most especially, the volunteers themselves that the SASWG has functioned so well and achieved so much.

Like other SCSC working groups, the SASWG is entirely a volunteer effort. The group relies on the generosity of its members for everything, including: access to physical rooms (or virtual spaces) to host meetings; knowledge and insights to support discussions; drafts of potential guidance text; review and proofreading of guidance documents; and funding hardcopy distribution. It is a great credit to the SASWG Chair, the volunteers' employers and, most especially, the volunteers themselves that the SASWG has functioned so well and achieved so much.

The Challenge

Many aspects of systems engineering are equally applicable to autonomous systems and traditional (or "non-autonomous") systems. The same holds for, on the one hand, conventional software engineering and, on the other, AI and ML. Despite these similarities, autonomous systems, and the associated enabling technologies, present unique safety assurance challenges.

The requirements placed on autonomous systems can be difficult to verify. Top-level requirements can be stated, for example, "Do not run over pedestrians". However, since autonomous systems are used in open, dynamic environments, these requirements cannot be decomposed to a simple set of behaviours in a small number of situations. If that were possible then there would be no need to invoke the complexity associated with autonomous systems.



A similar discussion applies at the computation level. One of the key principles of software safety assurance assumes a traceable, hierarchical decomposition of requirements that encapsulate the safety concerns [3].

Typically, ML approaches do not include this type of decomposition. All the requirements are captured abstractly, for example, via the training data, model (e.g. neural network) structure and training approach (e.g. loss function).

Another challenge in providing safety assurance for autonomous systems is the pace of technological development. Whilst this challenge is not unique to autonomous systems, the significant amount of investment in this area, combined with the vast quantity of academic research that is being published (especially in relation to ML), makes it difficult to identify Recognised Good Practice (RGP). Another consequence of this rapid development is that the best performing technologies are somewhat immature, especially from a safety perspective. This means that rare, but important, failure cases may not have been identified. This also applies to security vulnerabilities, which could have significant safety implications.

The Scope

SCSC-153A aims "*to provide clear, practical, pan-domain guidance on the safety assurance of autonomous systems*". This is clarified in statements describing the document's scope. For example:

- There is a deliberate focus on aspects directly related to autonomy, and enabling technologies such as AI and ML, rather than more general safety engineering or system engineering, where it is assumed that relevant general standards, guidelines and best practice will be applied. The intent is to avoid duplicating existing guidance relating to these general topics.
- There is a deliberate focus on AS [Autonomous Systems] that use AI developed using ML. Although it is possible to envisage AS that do not use these technologies, AI and ML are considered to represent the greatest assurance challenges; they are also expected to be widely used.

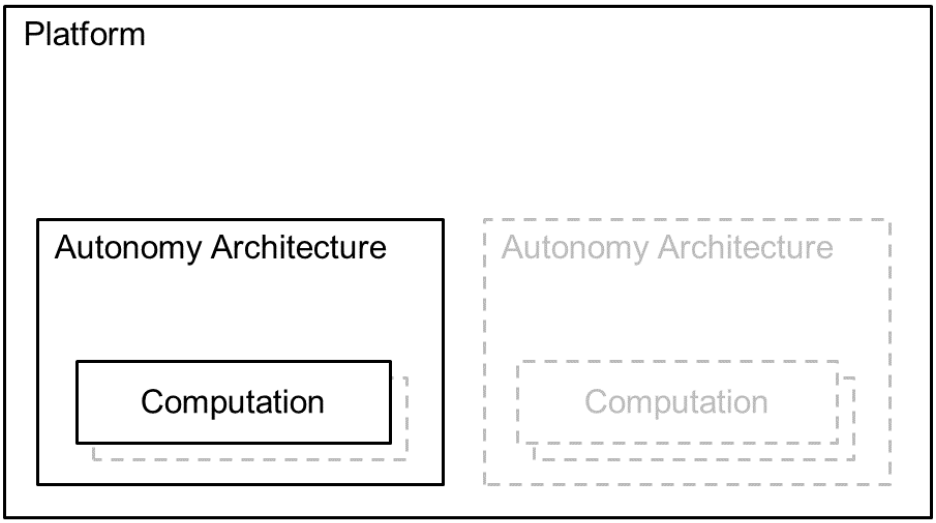
These reflect the SASWG's intent to focus on things specifically related to autonomy. This has allowed helpful guidance to be produced in an efficient manner. It does mean that SCSC-153A should be used alongside other guidance and standards.

The Structure

When SCSC-153A was being developed, there was no accepted way of structuring considerations (safety or otherwise) related to autonomous systems. Consequently, the group developed its own approach. This is based on three "frameworks". As SCSC-153A states:

- The computation-level framework addresses implementation at the software and computational hardware levels. It focuses on mapping an input to an output. Activities associated with this level typically relate to fault prevention. This is the lowest conceptual level considered.
- The autonomy architecture-level framework addresses how computations can be integrated into a system, or platform. Activities at this level typically relate to fault tolerance.
- The platform-level framework addresses what the final autonomous entity should do and what effects it should have on its environment. In essence, the focus is on requirements. This is the highest conceptual level considered.

As illustrated below, a platform may contain multiple "autonomy architectures". Likewise, an autonomy architecture may contain multiple computations.



To illustrate these concepts, SCSC-153A includes two examples. For ease of reference, these are reproduced below.

	Illustration One	Illustration Two
Example Platform	Self-Driving Car	Medical Diagnosis Application
Example Autonomy Architecture Components	Sensor Health Checks, Sanity Checks on Generated Route	Integrity Checks on Supplied Image, Using Multiple Classifiers
Example Computation	Route Planning	Image Classification (Benign / Malignant)

Each framework has an associated set of projections. These were used to help identify specific objectives. Whilst they proved useful, they also caused some confusion. In essence, "*projections provide different perspectives; they provide different ways of viewing each framework*". Hence, they share some characteristics with top-down and side-on projections of the same physical object. Like those projections, the projections in SCSC-153A "*are not intended to be strictly independent, distinct or non-overlapping*". For completeness, the projections are outlined in the following table.

Framework	Projection	Outline
Computation	Experience	Focused on the data that is available to train (or develop) the algorithm
	Task	Focused on the performance of the implemented algorithm; emphasizes requirements
	Algorithm	Focused on the type of algorithm that is used; emphasizes implementation
	Software	Focused on the software used to develop the algorithm and, separately, support its operational use
	Hardware	Focused on the computational hardware that is used, both for development and for operational use
Autonomy Architecture	Tolerance	Focused on tolerating faults and failures in computation-related items
	Information Provision	Focused on recording and maintaining information for subsequent use
	Adaptation	Focused on how updates to the algorithm (and associated software and hardware) are managed
Platform	Behavioural Specification	Focused on platform-level specification; includes defining the scope of autonomous aspects
	Interacting Items	Focused on things intended or required to interact with the platform, not directly owned by the platform developer or operator
	People	Focused on how the platform interacts with people; adopts a whole lifecycle perspective
	Environment	Focused on things in the operational environment that are outside the control of the developer or operator

The Objectives

There are three parts to each objective in SCSC-153A:

- The objective's text;
- A discussion illustrating how the objective contributes to autonomous system safety;
- Examples of approaches that could be taken to satisfy, or partially satisfy, the objective.

The third of these is, perhaps, the most interesting. In developing SCSC-153A, the SASWG was clear that the document should not contain objectives that were impossible to achieve. There was also a desire to avoid being overly prescriptive and an acknowledgement that the "state of the art" was subject to rapid change. The result is a collection of examples that draw heavily on academic papers: SCSC-153A includes 88 references. Moreover, SCSC-153A clearly states: "*These examples are not intended to be prescriptive; there may be other ways of satisfying an objective. Likewise, the examples do not necessarily represent a preferred way of satisfying an objective.*"

It is hoped that the separation of discussion (why the objective is important) from examples (how the objective might be satisfied) will help address the twin difficulties of, firstly, guidance material lagging excessively behind technological developments and, secondly, guidance material changing regularly [4]. In particular, an ambition is for the majority of future changes to SCSC-153A to be located in the examples. In total, SCSC-153A includes 45 objectives. The allocation of these to different framework levels is shown below.

Framework Level	Number of Objectives
Computation	19
Autonomy Architecture	11
Platform	15
Total	**45**

The Justifications

The SASWG devoted considerable effort to justifying the contents of SCSC-153A. Detail is provided in a series of appendixes. More specifically:

- The projections used in the computation framework were compared with a number of broadly similar items, including ones produced by the safety community [5] and ones developed from an ML perspective [6]. These projections were also compared against typical activities associated with, firstly, software development and, secondly, ML.
- The objectives listed in the computation framework were compared with a historical paper on neural network safety [7]: this provides an enduring perspective, rather than one dominated by recent technology trends. They were also compared against an analysis of gaps identified by viewing an automotive standard from an ML perspective [8].
- The projections used in the platform framework were compared with other similar categorisations, including ones developed by US authorities [9], academia [10] and industry [11].
- The complete set of SCSC-153A objectives was compared with the structure of the Assuring Autonomy International Programme (AAIP) Body Of Knowledge (BOK) [12]. The objectives were also compared with the sections in "*The Standard for Safety for the Evaluation of Autonomous Products*", UL4600 [13].

This collection of activities provides confidence that the SCSC-153A objectives are necessary and sufficient. It is, however, important to remember that autonomous systems, AI and ML are areas of rapid technological development. That, inevitably, constrains the amount of confidence that can be placed in any guidance document relating to this area.

The Relations

As noted above, since the SASWG began its work, two related items have been developed, specifically the AAIP BOK and UL4600.

The AAIP is funded by Lloyd's Register Foundation and the University of York. It is addressing global, pan-domain challenges associated with the assurance and regulation of Robotics and Autonomous Systems (RAS). The AAIP BOK is intended to become a key reference source for information related to these topics. It consists of a structured set of assurance objectives, each of which could be required for an assurance case.

UL4600 has been developed by Underwriters Laboratories (UL) and Edge Case Research (ECR). Despite the general terms in its title, UL4600 is targeted at the automotive domain. It concentrates on ensuring that a valid safety case is created. This would be expected to comprise goals, argumentation, and evidence.

The following table summarises the main properties of the AAIP BOK, UL4600 and SCSC-153A.

	AAIP BOK	UL4600	SCSC-153A
Domain	Pan-domain	Automotive focus	Pan-domain
Scope	System-level, based on a Sense-Understand-Decide-Act-like architecture	Whole system view, including systems engineering	Tight focus on autonomy-enabling technologies
Structure	Decomposition-based	Standard, with explicit clauses to produce arguments	Frameworks, projections and objectives
Approx. Size	67 assurance objectives	~190 clauses	45 objectives

All three items are concerned with autonomous system safety. However, they differ in their scope and utility. For example, the AAIP BOK's use of information from practical demonstrators and its non-linear nature may mean it offers most value by providing specific guidance for particular challenges. Conversely, the detailed nature of UL4600 may mean it is well suited to delivery of a mature product, intended for general use, most notably in the automotive domain. Alternatively, the smaller, more accessible nature of SCSC-153A may mean it is a valuable resource for early-stage or specialist products. In addition, the "autonomy architecture" framework provides a neat separation between platform and autonomy-enabled sub-systems, allowing SCSC-153A to be used at the sub-system level.

The smaller, more accessible nature of SCSC-153A may mean it is a valuable resource for early-stage or specialist products.

Given this discussion, it is suggested that the AAIP BOK, UL4600 and SCSC-153A be viewed as complementary (rather than as in competition).

The Future

After two years of rapid change and growth, there are no plans for a new edition of "*Safety Assurance Objectives for Autonomous Systems*" to coincide with SSS'21. Instead, SCSC-153A is being applied to a small number of specific cases. Experience gained from these applications will be used to refine the document, as necessary. Likewise, in due course, the document's content will be revised to account for developments in technology and safety standards.

SCSC-153A is a useful document and the SASWG has already accomplished a lot. However, much work remains to be done. Support in this endeavour would be greatly appreciated, whether that comes in the form of joining the SASWG, providing comments on SCSC-153A, or in any other way.

References

[1] SASWG, Safety Assurance Objectives for Autonomous Systems, v1.0, Feb 2019, scsc.uk/scsc-153

[2] SASWG, Safety Assurance Objectives for Autonomous Systems, v2.0, Feb 2020, scsc.uk/scsc-153A

[3] R. Hawkins, I. Habli, and T. Kelly, The principles of software safety assurance, 31st International System Safety Conference, Boston, Massachusetts USA, (2013).

[4] C. Johnson, Role of regulators in safeguarding the interface between autonomous systems and the general public, (2016).

[5] M. Douthwaite and T. Kelly, Safety-critical software and safety-critical artificial intelligence: Integrating new practices and new safety concerns for AI systems, in Evolution of System Safety, Proceedings of the Twenty-sixth Safety-Critical Systems Symposium, Safety-Critical Systems Club, 2018. ISBN 978-1979733618.

[6] E. Breck, S. Cai, E. Nielsen, M. Salib, and D. Sculley, What's your ML test score? A rubric for ML production systems, in NIPS Workshop on Reliable Machine Learning in the Wild, 2016.

[7] D. Bedford, G. Morgan, and J. Austin, Requirements for a standard certifying the use of artificial neural networks in safety critical applications, in Proceedings of the international conference on artificial neural networks, 1996.

[8] R. Salay and K. Czarnecki, Using machine learning safely in automotive software: An assessment and adaption of software process requirements in iso 26262, arXiv, 1808.01614 (2018).

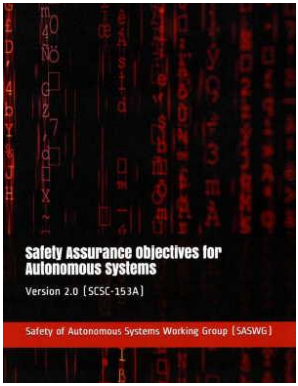
[9] Automated Driving Systems 2.0: A Vision for Safety, published by the US National Highway Traffic Safety Administration (NHTSA).

[10] AI Safety Landscape: https://www.ai-safety.org/landscape

[11] Uber Advanced Technologies Group, Safety Case Framework: https://www.uber.com/us/en/atg/safety/safety-case-framework

[12] AAIP Body Of Knowledge: https://www.york.ac.uk/assuring-autonomy/body-of-knowledge

[13] Underwriters Laboratories. UL 4600: Standard for Safety for the Evaluation of Autonomous Products.



SCSC Safety of Autonomous Systems Working Group (SASWG)

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.

Comments or suggestions for the Safety of Autonomous Systems Working Group (SASWG) and the Guidance Document are welcome at:

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

Errata

The Oct 2019 edition of Safety Systems (Vol 27 Nos. 2 [SCSC-155]) contains a type-setting error on page 19 at the end of the section titled "It is impossible to test all digital states even once". The numbers 2^{144} and 2^{96} were incorrectly shown as 2144 and 296 respectively.

SCSC

FOR EVERYONE WORKING IN SYSTEM SAFETY

THE SAFETY-CRITICAL SYSTEMS CLUB, 156th Seminar:

New Safety Analysis Techniques

Thursday 12 November, 2020 - Online, Free to Members

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 now take place in an online format in the afternoon of 12th November, from 14:00 to 16:30 UK time.

It will be a FREE EVENT but you have to be a full SCSC member to attend.

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 environment and their justification.

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 Survey Hazard Analysis). The Functional Resonance Analysis Method (FRAM) will be explained. The final talk of the day will look at Dialetic Arguments in safety cases (TBC).

A wrap up session at the end of the day will discuss the most promising techniques for specific areas.

Speakers include:

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

Mark Sujan, Human Factors Everywhere Ltd. - "FRAM in Healthcare"

Simon Whiteley, Whiteley Aerospace - "STAMP/STPA"

Yvonne Oakshott, Leonardo - "Dialectic Arguments" (TBC)

Membership details can be found here: https://scsc.uk/membership

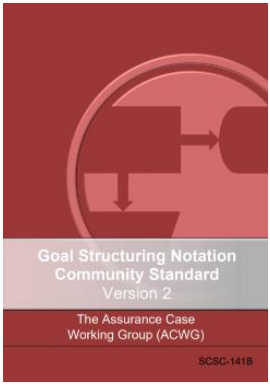
https://SCSC.uk/e654

WWW.SCSC.UK

SCSC Working Groups

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

Assurance Cases



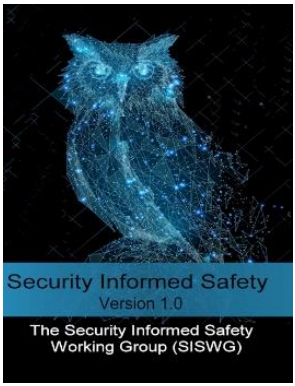
The Assurance Cases Working Group (ACWG) has been established to provide guidance on all aspects of assurance cases including construction, review and maintenance. The ACWG will:

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

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

Lead Phil Williams phil.williams@scsc.uk

Security Informed Safety



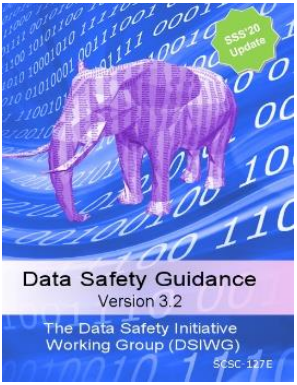
The Security Informed Safety Working Group (SISWG) aims to capture cross-domain best practice to help engineers find the 'wood through the trees' with all the different security standards, their implication and integration with safety design principles to aid the design and protection of secure safety-critical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

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

Lead Stephen Bull stephen.bull@scsc.uk

SCSC Working Groups

Data Safety Initiative



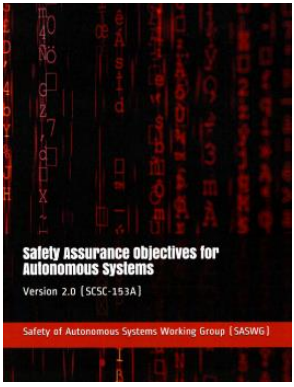
Data in safety related systems is not currently sufficiently addressed in current safety management practices and standards.

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

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

Lead Mike Parsons mike.parsons@scsc.uk

Safety of Autonomous Systems



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

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

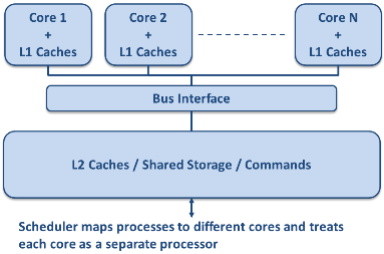
The Safety of Autonomous Systems Working Group (SASWG) aims to produce clear guidance on how autonomous systems and autonomy technologies should be managed in a safety related context, in a way that reflects emerging best practice.

Lead Rob Alexander rob.alexander@scsc.uk

SCSC Working Groups

Multi- and Manycore Safety

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

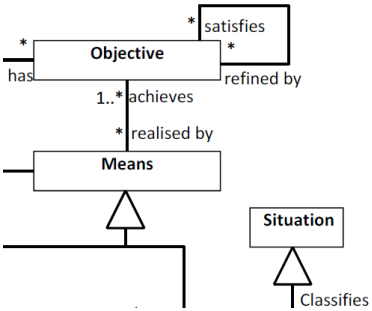


Parallel processes need to access the same shared resources, including memory, cache and external interfaces, so they may contend for the same resources. Resource contention is a source of interference which can prevent or disrupt completion of the processes, meaning it is difficult to know with a defined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

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

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

Ontology



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

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

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

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

SCSC Working Groups

Covid-19

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

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

Lead Mike Parsons mike.parsons@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.

Lead Mike Parsons mike.parsons@scsc.uk

SCSC Safety Culture Working Group (SCWG)

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

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

60s with … John McDermid

John is Professor of Software Engineering and previous Head of the Department of Computer Science at the University of York. John has published many books and about 440 papers in high integrity computer systems, especially in the areas of safety and security. He currently directs the Assuring Autonomy International Programme funded by the Lloyd's Register Foundation. His work has influenced industrial practice both directly and via standards and he has taught extensively at post-graduate level, including on safety courses for industry.

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

Two things at more or less the same time – investigating the failure of a syringe pump which killed two people, and doing static code analysis on some safety-critical software.

What aspect of your career are you most proud of?

Helping others to achieve their best and to overcome obstacles – I hope some of my former PhD students will recognise this …

What advice would you give to yourself age 12?

Trust your own judgement.

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

Possible failure to "grapple" with AI/ML – to be effective we need a cultural shift where the AI/ML and safety people work effectively together, and I don't see this happening although I hope the work I am leading through the Assuring Autonomy International Programme will help …

What's your most favourite quote or motto?

The University of York's motto is "in limine sapientiae" – on the threshold of wisdom – so I live in hope …

"in limine sapientiae" – on the threshold of wisdom – so I live in hope …"

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

Play a musical instrument – it would probably have to be the drums, as I have a sense of rhythm but no sense of pitch …

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

It would have to be a detective – maybe Inspector Chen Cao from Qiu Xiaolong's series of books, who is sometimes referred to as the "Morse of the Far East", and this would mean I could also write poetry, so it gets me a second answer to the previous question!

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

"Don't attribute to malice what can be explained adequately by incompetence" – from a former boss and this has helped me preserve my equanimity on many occasions.

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

1Q84 by Haruki Murakami – the book is set in 1984 but strange things happen, so it doesn't seem like 1984 anymore, hence the Q in the title – so 2Q20 now – oh, and it's a very, very long book …

THE SAFETY-CRITICAL SYSTEMS CLUB, Seminar:

Management and Oversight of Complex Systems

Thursday 3 December, 2020 - Online

This will now be an online event from 14:00 to 16:30 UK time.

It will be FREE TO ATTEND for all SCSC members.

This seminar is aimed at those who have to manage, approve, regulate and operate complex systems which have a safety aspect, e.g. Air Traffic Control systems, Nuclear Plant, Aircraft Carriers or National Power Transmission systems.

It will cover aspects such as design, operation and manning of control centres, and the use of dashboards and metrics used for monitoring. It will consider the key performance indicators that can be used to measure safety performance.

It will also cover the tools and skills that are required for management and understanding of such systems.

Speakers include: John McDermid, University of York

Further details TBC.

SCSC
FOR EVERYONE WORKING IN SYSTEM SAFETY

This online seminar will be useful for all those involved in running a complex operation that involves safety. It is aimed at Managers, Operators, Regulators and Assurance staff. If you operate a management or operations centre for your organisation then this seminar is for you.

How to gain an overview of complex, safety-related Systems

https://SCSC.uk/e661

WWW.SCSC.UK

SCSC Membership

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

Individual Membership

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

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

Corporate Membership

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

Renewing Membership

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

Membership Fees

The following fees are applicable for new and renewing members:

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

Contact Alex King using office@scsc.uk

The SCSC Steering Group

Tom Anderson

Honorary member

Stephen Bull

stephen.bull@scsc.uk

Jane Fenn

jane.fenn@scsc.uk

Paul Hampton

paul.hampton@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

Emma Taylor

emma.taylor@scsc.uk

Sean White

sean.white@scsc.uk

Robin Bloomfield

Honorary member

Dewi Daniels

dewi.daniels@scsc.uk

Zoe Garstang

zoe.garstang@scsc.uk

Louise Harney

louise.harney@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

Phil Williams

phil.williams@scsc.uk

Club Positions

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

Director

Mike Parsons 2019-

Tim Kelly 2016-2019

Tom Anderson 1991-2016

Steering Group Chair

Roger Rivett 2019-

Graham Jolliffe 2014-2019

Brian Jepson 2007-2014

Bob Malcolm 1991-2007

Programme & Events Coordinator

Mike Parsons 2014-

Chris Dale 2008-2014

Felix Redmill 1991-2008

Manager

Alex King 2019-

Newsletter Editor

Paul Hampton 2019-

Katrina Attwood 2016-2019

Felix Redmill 1991-2016

University of York Coordinator

Mark Nicholson 2019-

Website Editor

Brian Jepson 2004-

Administrator

Alex King 2016-

Joan Atkinson 1991-2016

Safety Futures Initiative Lead

Nikita Johnson 2019-

Calendar

October '20

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 '20

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 '20

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			

January '21

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 '21

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

March '21

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				

30

April '21

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 '21

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

June '21

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 '21

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 '21

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

September '21

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			

Events Diary

12 November 2020
SCSC Seminar

New Safety Analysis Techniques

Online

scsc.uk/e654

POSTPONED
SCSC Seminar

Safe Use of Multi-Core and Manycore Processors

scsc.uk/e638

3 December 2020
SCSC Seminar

Management and Oversight of Complex Systems

Online

scsc.uk/e661

9-11 February 2021
SCSC Symposium

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

scsc.uk/e683

25 March 2021
Conference

Nuclear engineering for safety, control and security

Bristol, UK

events2.theiet.org/nuclear

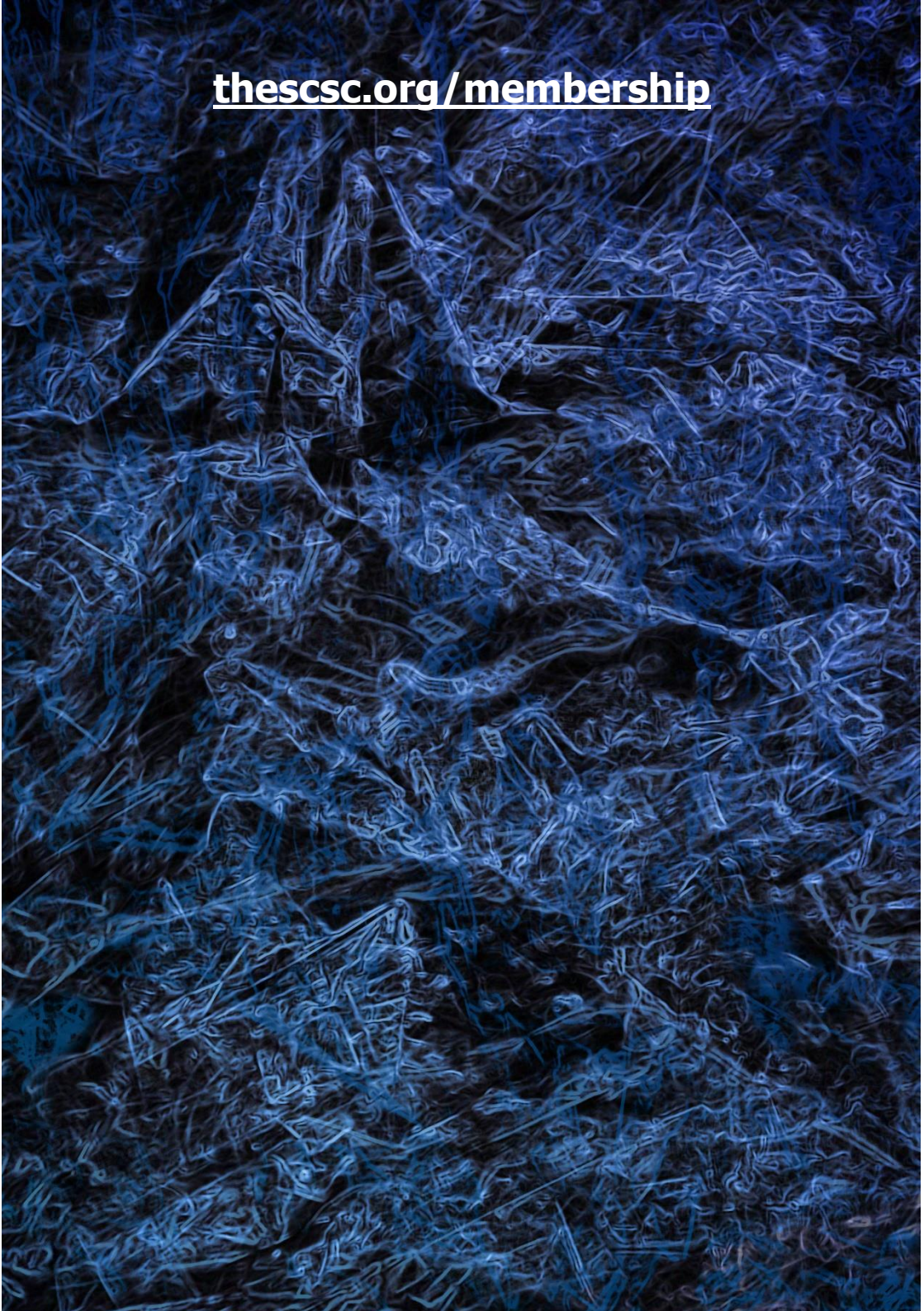
7-10 Sept2021
SafComp 2021
Conference

40th International Conference on Computer Safety, Reliability and Security

York, UK

safecomp2021.hosted.york.ac.uk

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


thescsc.org/membership