SCSC Data Safety Initiative – WG Meeting 77

22nd June 2023, Zoom

Minutes

Attendees

Mike Parsons (MP) – Ebeni, Tim Rowe (TR) – Consultant, Paul Hampton (PH) – CGI, Nick Hales (NH) – Consultant, Mike Standish (MS) – Dstl, Oscar Slotosch (OS) – Validas, Jennifer Kracht (JK) – TomTom, Brent Kimberley (BK) – Durham, Daniel Clegg (DC) – BAE Systems, Roland Rosier (RR) – TomTom.

Apologies

Divya Atkins (DA) – MCA, Martin Atkins (MA) – MCA, Mark Nicholson (MN) – University of York, Graham Sutherland (GS) – Consultant, Fan Ye (FY) – ESC, Mark Templeton (MT) – Qinetiq, Michael Green (MG) – Ecomergy, Andy Williams (AW) – Consultant.

Agenda

1. Welcome
2. Indian Rail Crash
3. Moon Landings & Data
4. Calculus of DSALs
5. Oldest Data Safety Error
6. Timeliness and Dynamic Data
7. AI / ML and security
8. DSITN (Data Safety in the News)
9. Update on Tooling
10. Actions
11. Next meeting
12. AOB

NOTE: All comments or opinions in these notes are attributed only to individual attendees of the meeting, not to their respective organisations.

*[Note that actions are presented in the form **N.Mx** where **N** is the meeting number, **M** a reference number for the action raised in that meeting and **x** is an optional letter that differentiates related actions arising from the same discussion point].*

The meeting slides are available at: https://scsc.uk/file/gd/77th_DSIWG_Slides_v1-1574.pptx

1. Welcome

MP opened the meeting and welcomed those attending.

2. Indian Rail Crash

MP mentioned the recent Indian rail crash with the tragic loss of 288 lives. This would appear to have been caused by a mismatch between the signalling and the points. It is thought that work was being undertaken on the signals at the time, causing the driver of an express train to continue at high speed, when the points were set to move to another line[1]. The signal error could well have a data contribution.

A report on the accident is awaited; note that the UK RSSB will also release a report.

E.g. https://news.sky.com/story/india-train-crash-at-least-233-people-dead-and-hundreds-injured-in-collision-in-odisha-12895199
There is also a good account here with track diagrams:
https://www.nytimes.com/interactive/2023/06/04/world/asia/india-train-crash-cause.html

3. Failed Moon Landing

MA and MS supplied links regarding the Japanese failed space mission, the Hakuto-R Mission 1 moon lander. According to the incident report, this ran out of fuel at height and crash-landed on the surface. It lost accurate altitude information when moving over high crater rims, causing a sensor to go offline as its data was no longer considered credible.

https://science.slashdot.org/story/23/05/28/2012238/a-japanese-made-moon-lander-crashed-because-a-crater-confused-its-software?utm_source=feedly1.0mainlinkanon&utm_medium=feed

https://abcnews.go.com/Business/wireStory/crash-private-japanese-moon-lander-blamed-software-minute-99630046

https://youtu.be/2JlUnOAiMm4

Note that many news reports blamed the 'software' whereas the group considered this to be more a requirements or testing failure. There was a discussion as to whether there were redundant or backup altitude sensors. It was thought the software should not necessarily have stopped using the sensor and instead used e.g. last known good values for a time. Further simulations might also have helped.

There was also quite a discussion on LinkedIn on the loss: https://www.linkedin.com/posts/philip-koopman-0631a4116_japanese-moon-lander-crashed-because-it-was-activity-7068202209099821057-lxOV?utm_source=share&utm_medium=member_desktop

ACTION 77.1 (OS) – Investigate this moon landing accident further.

There was also a discussion about the missing sub taking passengers to the Titanic wreck. Data involvement could be to do with search and rescue sensing technologies, communications and beacons, etc. [Post-meeting note: this sub is now reported lost due to structural implosion.]

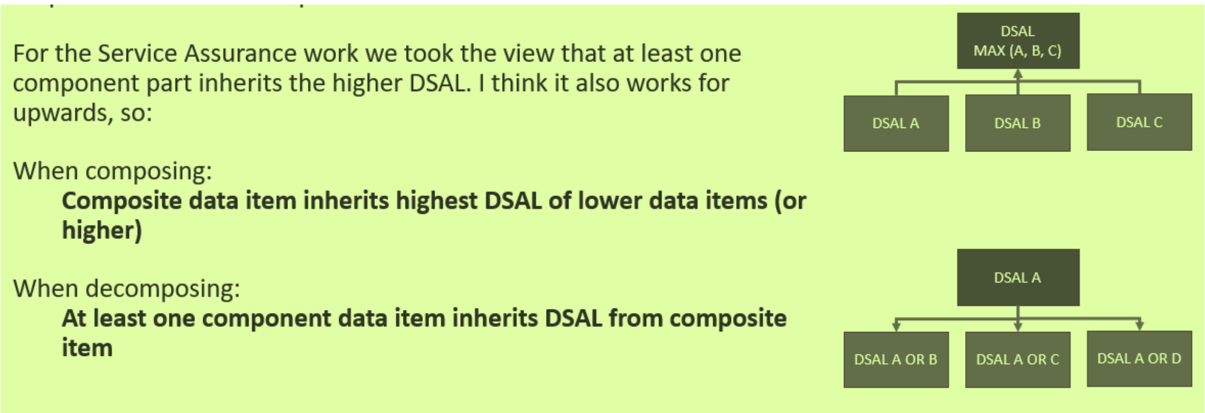
[1] TR mentioned that the India derailment has some echoes of the 2017 Waterloo derailment - signals did not correctly reflect the points setting because of work being done on the system.

4. Calculus of DSALs

MA had raised the issue of data fusion where data from multiple sensors is combined to give an overall view. This is not covered in the existing guidance as DSALs may need to be combined[2].

MP presented his recent thinking about DSALs across data composition / decomposition as a possible way forward. The sorts of things considered were merging of databases of different DSALs or using parts of medical records (e.g. the address part may be less critical than the blood type part).

It was noted that in the SCSC service assurance guidance a simple approach is taken for Levels of Service Assurance (LSA) whereby on composition the highest LSA of the sub-elements is used at the next-higher level and on decomposition at least one sub-elements inherits the LSA from the higher level, so:



For the Service Assurance work we took the view that at least one component part inherits the higher DSAL. I think it also works for upwards, so:

When composing:

Composite data item inherits highest DSAL of lower data items (or higher)

When decomposing:

At least one component data item inherits DSAL from composite item

It was thought that for the upward composition this approach may make sense for data, but for downwards decomposition the situation is more complex, i.e. how to assign DSALs to lower-level components is not straightforward.

The situation of multiple sparse or diverse sets of data leading to more assured data was discussed (i.e. filling in 'holes' in data sets makes a better overall set). Redundancy of data was considered and how that might lead to higher assurance.

RR said that in the ISO 26262 world the highest ASIL is used on aggregation[3]. PH said that in aviation, the ARPs (4754, 4761) also have a methodology for combining levels[4].

There was discussion about dependent / independent data sets and how that would affect the picture.

MP expressed his hope that a simple methodology might be produced. PH said that approaches are likely to be different to those for existing software and systems standards.

[2] JK asked if this will mean the Data Safety assessment needs to be done on component level because TomTom tried to do it on product level in the past...

[3] RR mentioned that ISO 26262-9:2018 section 5.4.9 has some diagrams on decomposing a higher ASIL into multiple lower ASILs. Section 5.4.3 also mentions that the elements which are part of the decomposition should be sufficiently independent.

[4] MS said section 5 onwards of ARP4754A includes information on assurance assignment for FDALs (e.g. system-level) and IDALs (e.g. software-level). This is for a combination of errors and propagation.

5. Oldest Data Safety Error

MP presented his proposed abstract for SSS'24 and this was updated in the meeting:

"Data-intensive systems are now around us everywhere and so the opportunities for data to cause accidents have increased. However such data safety errors are not new and research by members of the Data Safety Initiative Working Group has identified some important cases throughout history where data-related errors have had a major impact: from the recent Covid-19 spreadsheet silent loss of rows causing an estimated 1,500 deaths, the Gemini space mission where the landing site was miscalculated due to an error in the value used for the Earth's rotation rate, to the loss of sailing ships caused by longitude errors due to the lack of accurate time fixes, to early civilisations when pyramid dimensions were likely incorrect. This paper offers a timeline of data safety problems in systems and shows that they have been with humankind for millennia."

6. Timeliness and Dynamic Data

PH mentioned that he had found a problem in the current guidance document: If you look at the techniques tables from 6.4.2.3 to 6.4.2.10 there is no technique suggested for Dynamic Data Category ("D"), Data Property "M" (for timeliness[5]) and DSAL1. He showed the problem using data within the RADISH tool which helped identify the issue. He has fed back some points to MCA Ltd who are developing the tool.

MP had some suggestions for some methods where the technique should be added: SD.08, SD.09, DD.03, DC.04, DC.05. He said that we probably also need some new methods for this case, e.g.

SD.25	Data expiry detection
SD.26	Transmission failure detection
SD.27	Time window detection
DD.11	Timing models

ACTION 77.2 (MP, PH, MT) – Review proposed updates and add to list of changes to version 4.0 of the guidance

7. AI / ML and security

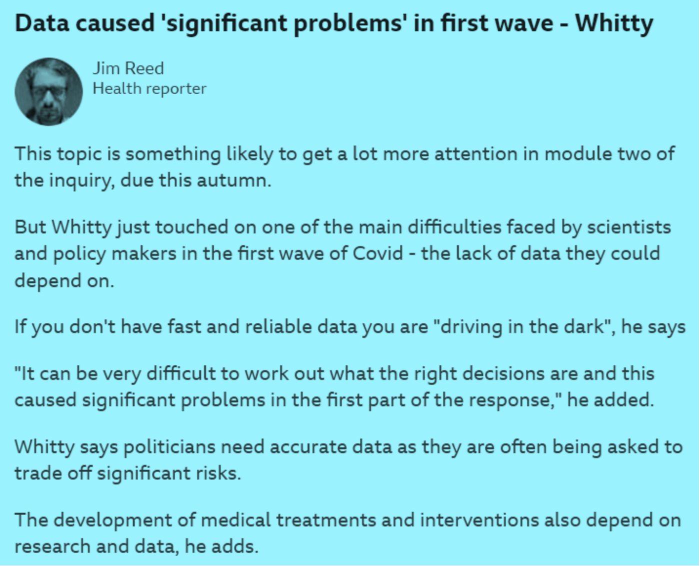
NH said that the document, "AI SECURITY CONCERNS IN A NUTSHELL", https://scsc.uk/file/gd/BSI__AI_Security_Concerns-1560.pdf contained some apparent anomalies e.g. use of multiple diverse systems as, although these might improve safety, they could create additional security vulnerabilities by creating more ways into a system. Also redundancy could create additional channels of weakness, especially if the duplicated systems are in different physical locations. However OS said that there is a trade-off to be assessed as diversity can help. RR mentioned that many road vehicles are going to be connected to the internet and this creates many

[5] BK noted that Fidelity - the degree of exactness with which something is copied or reproduced. Timeliness is a subset of fidelity.

more vulnerabilities, with hackers potentially having a direct route into the vehicle systems with live data feeds[6].

8. DSITN (Data Safety in the News)

MP mentioned the UK Covid-19 enquiry which is now underway. There has been an explicit mention of data already, and how this hampered efforts to control the virus, https://www.bbc.co.uk/news/live/uk-65967979 (excerpt below):



Data caused 'significant problems' in first wave - Whitty

Jim Reed
Health reporter

This topic is something likely to get a lot more attention in module two of the inquiry, due this autumn.

But Whitty just touched on one of the main difficulties faced by scientists and policy makers in the first wave of Covid - the lack of data they could depend on.

If you don't have fast and reliable data you are "driving in the dark", he says

"It can be very difficult to work out what the right decisions are and this caused significant problems in the first part of the response," he added.

Whitty says politicians need accurate data as they are often being asked to trade off significant risks.

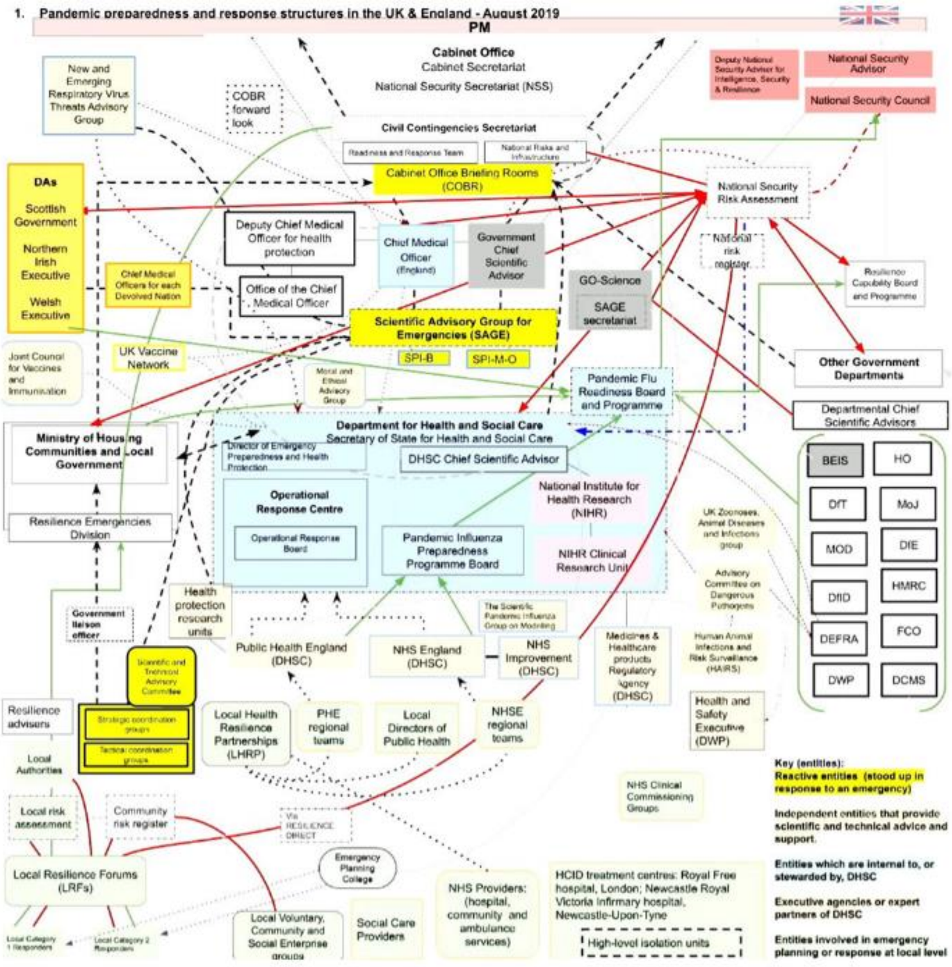
The development of medical treatments and interventions also depend on research and data, he adds.

.

The diagram of organisational relationships has also raised questions due to its complexity and number of information flows, https://www.theguardian.com/uk-news/2023/jun/13/a-bowl-of-spaghetti-covid-inquiry-opens-with-flowchart-on-uks-pandemic-planning Was incorrect, late or missing data a cause of poor decision making in this overall system?[7]

[6] BK said that in North America, manufacturers work through dealers to limit liability. Will live data streams impact manufacturer liability?

[7] BK said that in Canada, according to local media, we may have had trouble controlling COVID datasets. Likewise, there are still questions in the media WRT the origins of COVID.



1. Pandemic preparedness and response structures in the UK & England - August 2019

PM
Cabinet Office
Cabinet Secretariat
National Security Secretariat (NSS)
Civil Contingencies Secretariat
Readiness and Response Team
National Risks and Infrastructure
Cabinet Office Briefing Rooms (COBR)
COBR forward look
New and Emerging Respiratory Virus Threats Advisory Group
Deputy National Security Advisor for Intelligence, Security & Resilience
National Security Advisor
National Security Council
National Security Risk Assessment
National risk register
Resilience Capability Board and Programme
DAs
Scottish Government
Northern Irish Executive
Welsh Executive
Chief Medical Officers for each Devolved Nation
Deputy Chief Medical Officer for health protection
Office of the Chief Medical Officer
Chief Medical Officer (England)
Government Chief Scientific Advisor
GO-Science
SAGE secretariat
Scientific Advisory Group for Emergencies (SAGE)
SPI-B
SPI-M-O
Joint Council for Vaccines and Immunisation
UK Vaccine Network
Moral and Ethical Advisory Group
Pandemic Flu Readiness Board and Programme
Other Government Departments
Departmental Chief Scientific Advisors
BEIS
HO
DfT
MoJ
MOD
DfE
DfID
HMRC
DEFRA
FCO
DWP
DCMS
Ministry of Housing Communities and Local Government
Resilience Emergencies Division
Department for Health and Social Care
Secretary of State for Health and Social Care
Director of Emergency Preparedness and Health Protection
DHSC Chief Scientific Advisor
Operational Response Centre
Operational Response Board
Pandemic Influenza Preparedness Programme Board
National Institute for Health Research (NIHR)
NIHR Clinical Research Unit
UK Zoonoses, Animal Diseases and Infections group
Advisory Committee on Dangerous Pathogens
Human Animal Infections and Risk Surveillance (HAIRS)
Health protection research units
Government liaison officer
Scientific and Technical Advisory Committee
Strategic coordination groups
Tactical coordination groups
The Scientific Pandemic Influenza Group on Modelling
Public Health England (DHSC)
NHS England (DHSC)
NHS Improvement (DHSC)
Medicines & Healthcare products Regulatory Agency (DHSC)
Health and Safety Executive (DWP)
Resilience advisers
Local Authorities
Local Health Resilience Partnerships (LHRP)
PHE regional teams
Local Directors of Public Health
NHSE regional teams
NHS Clinical Commissioning Groups
Local risk assessment
Community risk register
Via RESILIENCE DIRECT
Local Resilience Forums (LRFs)
Emergency Planning College
Local Category 1 Responders
Local Category 2 Responders
Local Voluntary, Community and Social Enterprise groups
Social Care Providers
NHS Providers: (hospital, community and ambulance services)
HCID treatment centres: Royal Free hospital, London; Newcastle Royal Victoria Infirmary hospital, Newcastle-Upon-Tyne
High-level isolation units
Key (entities):
Reactive entities (stood up in response to an emergency)
Independent entities that provide scientific and technical advice and support.
Entities which are internal to, or stewarded by, DHSC
Executive agencies or expert partners of DHSC
Entities involved in emergency planning or response at local level

9. Update on Tooling

No update other than PH was currently using the RADISH tool[8].

10. Actions

See table at end.

11. AOB

MS asked if there were any existing talks on Data Safety that could be used to introduce some defence staff to the topic. It was agreed to send MS the slides from the 'Data Safety Evolution' seminar held in 2019.

ACTION 77.3 (MP) – Send MS the slides from the Data Safety Evolution seminar

12. Next Meeting

The next meeting will be held 3-5pm 26th July via Zoom:

[8] BK asked how does RADISH compare with MOIMS? https://cwe.ccsds.org/moims/default.aspx PH replied: Not familiar with MOIMS but RADISH is really an implementation of the Data Safety Guidance: https://dst.mca-ltd.com/ Note RADISH is "Risk Assessor for Data Integrity and Safety Hazards"

https://us02web.zoom.us/j/81685207668?pwd=RjJRQVFwUFE4dDRieEJ5aVNWSFNtUT09

13. Thanks

Thanks to all who provided contributions.
Thanks to MP for chairing and taking minutes.

Summary of Open Actions

Actions greyed out are considered closed and will be removed from the list at next issue.

Ref	Owner	Description	Target Guidance Version
42.9	MP	Work out a matrix of data categories (previously 'types') and data properties (as per DB discussion)	N/A
43.4	MP	Write up a data focussed FMEA approach.	4.0
44.2	MP	To discuss with AK on how to get the Wikipedia article published	N/A
46.1	MP	Review the application of DSALs to higher level forms of aggregation	N/A
53.1	MP	To talk to Kevin King about what we need to do in the guidance for digital twins.	4.0
61.2	AW	Research the relevance of digital currencies and report back to the group (with MA and MT)	N/A
63.1	CT	Look at both Dark Data and Dazzle Data for sensors (e.g. when a sensor is saturated, in noisy environment or when readings are below the detection level floor)	4.0
69.2	RR	Explore the issue of data / software compatibility issues and to what extent data can impose requirements on software	4.0
69.3	PMcK	Develop a scoping diagram that shows how the DSG fits into the overall lifecycle process and other standards	4.0
69.4	MA	Write a short note on the issues of aggregation	4.0
69.6	MA/DA	Update the data safety tool to use the latest version of the guidance document	-
71.3	PH/DA/RO	Develop security properties thinking further for next DSIWG	4.0
71.4	PH/DA/RO	Present security properties work to next SISWG meeting	-
71.5	AM	(i) Establish if any of this can be published within the DSIWG and (ii) Consider a structuring similar to that used in security standards or ISO26262	-
71.7	MP/CT	Consider impact of FAIR data on the guidance	4.0
73.1	JK, RR	Consider production of a short note which could be used as an appendix to the guidance on lessons learnt using the guidance at TomTom	4.0
73.2	MT	Consider how the guidance fits with different lifecycles considering 'V', Continuous Service, Agile and 'Data Pipeline'	4.0
73.4	DA	See whether any of the DSITN entries from previous minutes and slides could be used to enhance the list of accidents	-
75.3	MT	Fix the minor typos that have been reported with v3.5 and prepare updates for both the online version and also the KDP/Amazon hardcopy.	3.5 Update

Ref	Owner	Description	Target Guidance Version
76.1	MP, PH, MT	Consider whether Guidance Annexes can be integrated into the main body of the document.	4.0
76.2	TR	Look further at IEC 25012 to consider whether a new Annex to the Guidance would be beneficial.	4.0
76.3	MT	Provide slides from last October's SQEPtember presentation (an introduction to data safety) to MP.	-
77.1	OS	Investigate this moon landing accident further	
77.2	MP, PH, MT	Review proposed updates and add to list of changes to version 4.0 of the guidance	4.0
77.3	MP	Send MS the slides from the Data Safety Evolution seminar	