SCSC Data Safety Initiative – WG Meeting 80

2nd November 2023, Zoom

Minutes

Attendees

Mike Parsons (MP) – Ebeni, Roland Rosier (RR) – TomTom, Martin Atkins (MA) – MCA, Dave Banham (DB) – Blackberry, Paul Hampton (PH) – CGI, Paolo Giuliani (PG) – EDF, Divya Atkins (DA) – MCA, Tim Rowe (TR) – Consultant, Arch McKinley (AM) – NGA, Nick Hales (NH) - Consultant.

Apologies

Graham Sutherland (GS) – Consultant, Paul Butcher (PB) – AdaCore, Oscar Slotosch (OS) – Validas, Brent Kimberley (BK) – Durham, Daniel Clegg (DC) – BAE Systems, Dave Murray (DM) - BAE Systems.

Agenda

1. **Welcome**
2. **Google Maps Error**
3. **Collapse of a GE Vernova turbine**
4. **Russian Failed Moon Landing**
5. **Definition of Corrupt and Incorrect Data**
6. **Poster for SSS'24**
7. **Oldest Data Safety Error Paper**
8. **NATS recent outage – blame game**
9. **Data Affecting Data Update**
10. **New Guidance Version**
11. **Spread of Accidents**
12. **SITN (Data Safety in the News)**
13. **Update on Tooling**
14. **Actions**
15. **Next meeting**
16. **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/80th_DSIWG_Slides_v2-1633.pptx

1. Welcome

MP opened the meeting and welcomed those attending.

2. Google Maps Error

MP noted the recent Google Maps error which navigated a driver over a collapsed bridge where physical warnings had been removed; the driver unfortunately died:

https://www.bbc.co.uk/news/world-us-canada-66873982

It will be interesting to see the outcome of investigations and legal cases on this. RR said that drivers become lazy and just use Sat Nav to direct them and trust it too much. We all know that the navigation system shouldn't be trusted, but the general public will.
RR was concerned about a new EU Regulation 'Intelligent Speed Assist' which could lead to increased driver loss of attention:

- https://road-safety-charter.ec.europa.eu/resources-knowledge/media-and-press/intelligent-speed-assistance-isa-set-become-mandatory-across
- https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=PI_COM:Ares(2021)2243084&rid=1

There was a general consensus that all the new vehicle technology was leading to a situation where the driver was no longer in complete control, and relying on the technology too much. The line between driver control and driver assist was becoming blurred. There was also concern that the T&C's and disclaimers for such systems were not easy to read or understand. The conclusion was that there will inevitably be lots of cases such as this Google Maps one before a clear legal precedent is set.

3. Collapse of a GE Vernova turbine

MP showed a report of a collapsed GE Vernova turbine at a wind farm in Lithuania which had been completely destroyed:

https://renews.biz/88958/enefit-says-sensor-malfunction-caused-turbine-collapse/

This was a massive turbine and, had anybody been near, the consequences would have been severe. The report into the failure states:

> *"…detailed root cause analysis has led to the conclusion that a malfunctioning sensor sent incorrect information to the turbine controller which led to an excessive load on the tower structure and resulted in the collapse of the turbine"*

MP thought that the sensor failure should not have caused this, as the control system should be been resilient to faulty data from sensors. Instead of blaming one sensor a more general, data-based analysis should have been done. [There are possible analogies with the Boeing 737 Max crashes where initially one AOA sensor was blamed.]

4. Russian Failed Moon Landing

MA explained how data was involved in the recent Russian Space Probe crash into the moon:
https://youtu.be/TJ_a4NCIlmk?si=2E91UW1DFmaotEg9

First and foremost this was a priority miscalculation (or maybe inversion?) on the communications bus of the spacecraft, but communicating "no data" as zero is a data error!

5. Definition of Corrupt and Incorrect Data

MP highlighted recent communcations with Richard Garrett of SQEP about formal definitions of 'Corrupt Data' and 'Incorrect Data'. His suggestions were:

> **Corrupt Data** - Unwanted, unexpected and potentially hazardous changes to data (or Loss of integrity of data), where the changes may be hard to detect
>
> **Incorrect Data** - similar, but not as strong, with the implication that the changes can be more easily detected and therefore rejected.

There was discussion about this. PG thought that external causes were important in corruption cases, involving things such as failures in communications or multi-tasking. RR asked if data translation (or migration) errors would count as corruption and it was thought they could. AM thought that the definition could include aspects of undesirability or emergent changes. Data drift could also be counted as corruption. It was thought the root causes of corruption should be considered. AM also notes [1]in the chat. NH thought website was useful, NordVPN.com - Home page "corruption".

6. Poster for SSS'24

MP said that the DSIWG will present a poster at SSS'24 in February next year, https://scsc.uk/e1007 and that he thought the previous poster might be updated.

He showed a couple of examples from last year which he thought worked well (from SAWG https://scsc.uk/re898.6:1 and SASEWG https://scsc.uk/re898.5:1) because they had strong central image or graphical theme.

There was discussion about this. It was thought an annotated map with metadata as a central image would be really good. RR suggested that TomTom might be able to help. MP suggested the recent NATS outage (see below) would be good to include somewhere.

ACTION 80.1 (RR) – Ask Jennifer Kracht at TomTom if she could assist with the new DSIWG poster

[1] From Arch Mckinlay:

- need to add "undesired" to cover nuisance issues that are assumed to be non-safety and the PM accepts the risk with no further changes. These include emergent as well as undesired. Such as new geo features added in the data that cause the mapping function to throw a nuisance fault any human would probably ignore
- data drift mimics this error
- unwanted are anti-requirements
- unexpected are uncertainty which is a huge issue
- uncertainty and complex data systems is causing misinterpretation as non-determinism
- FAA is requiring uncertainty analysis as usual but we used to talk about these as controls on corruption among others
- root cause of corruption is system-level during operations or if during development then is the data steward's and data assurance/security responsibility. If data is enriched without controls then it mimics corruption.

7. Oldest Data Safety Error Paper

MP said that there had only been limited contributions (from PMcK and MP) so far and so he thought it best to use this work in an SCSC Newsletter article rather than at SSS'24.

[Update we now have 8 suggestions including Amelia Earhart's Last Flight, the Thermopylae battle, The Scilly Isles loss of ships 'Association', 'Eagle', 'Romney' and 'Firebrand', but could still do with more.]

8. NATS recent outage – blame game

MP noted the recent press reports from airlines related to the NATS Air Traffic Management problem in the UK[2]. A blame game appears to be developing with the airlines, particularly RyanAir disputing the NATS figures for the number of delays and cancellations:

- https://www.ftnonline.co.uk/2023/10/13/the-nats-blame-game-commences/
- https://www.independent.co.uk/travel/news-and-advice/air-traffic-control-failure-ryanair-nats-b2431567.html

This was discussed by the group. There was surprise at why the flight plan was not simply rejected. TR said there were thought to be similarities with an earlier outage NATS suffered some years ago. AM suggested some organizational factors come into play with these sorts of systems where contractors are very reluctant to update or upgrade systems.

9. Data Affecting Data Update

MP reworked the list of how one data item might affect another:

1. Calculations
2. Configuration parameter
3. Tuning/filter/selection data settings
4. Tags or metadata are incorrect or lost
5. Noisy or bad data (e.g. timestamps)
6. Sentinel values are incorrect
7. Bad CRCs cause data items to be rejected
8. Incorrect indices in arrays or tables
9. Hashing
10. Block chain?
11. Data chaining, where value used in next stage

MP will continue to work on this and report back. AM offered to assist.

10. New Guidance Version

MP said that the only updates which are likely to make it into the next version of the guidance document to be issued at SSS'24 are:

1. Corrections and minor wording updates

[2] A flight plan with duplicate waypoint data outside UK airspace caused both primary and backup systems to fail, causing an outage in automatic flight plan processing

2. Reworking of the accident list (see item below)
3. Work some appendices into the body of the document
4. New appendix on tooling, mentioning that ‘Radish’ is available. Possibly also showing an example of application

ACTION 80.2 (TR) – See which appendices are candidates to work into the main body of the document
ACTION 80.3 (DA, MA) – Write appendix on tooling and RADISH

11. Spread of Accidents

MP said that he had done his action to look at the sector spread of accidents across sectors mentioned in the current guidance document. The results are:

Air (Civil): 9 H.7, H.15, H.17, H.18, H.21, H.28, H.30, H.31, H.35	Policing: 3 H.3, H.12, H.33
Air (Other): 1 H.10	Maritime (Civil): 4 H.19, H.24, H.27, H.29
Air (Military): 3 H.13, H.20, H.22	Maritime (Military): 1 H.14
Defence: 1 H.32	Medical: 6 H.3, H.5, H.6, H.16, H.23, H.26
Internet: 1 H.4	Rail: 3 H.9, H.25, H.36
Oil & Gas: 1 H.37	Space: 4 H.2, H.8, H.11, H.34

He said that some sectors such as: Drones, Power Generation/distribution, Highways, Automotive, Government, Nuclear, ATM, and Buildings were not covered. DA said she and MA were collecting all known data safety accidents into the data-safety.tech accident list, and that this would be available to the DSIWG. It was thought that this was useful and once all data had been entered it should be easier to extract a better sector spread of accidents for the guidance. [Note if this is used to hold historical accidents then this could be used for the Newsletter article on the oldest data safety error too.]

ACTION 80.4 (DA, MA) – Update the data-safety.tech web site with all known data safety accidents (from past DSIWG minutes and any other sources).

ACTION 80.5 (MT) – (i) Add a ref to the data-safety.tech website in the guidance and (ii) Use the updated data-safety.tech website to produce a better spread of accidents for the new version of the guidance.

12. DSITN (Data Safety In The News)

MA reported that, in the ongoing war in the middle east the remote control monitoring and weapons systems on the Israel wall were apparently disabled by drones attacking the mobile phone masts. It would be good to find a link on this.

TR reported on the current AI systems: they are only as good as the training data they are fed, and this can be maliciously or accidently biased or modified. Apparently safeguards within ChatGPT can be overcome by using rare language subsets or language tricks (e.g. phrasing things in the negative). False information ‘Hallucinations’ supplied by AI tools is a major issue as well.

13. Update on Tooling

DA and MA said that the RADISH tool would be updated to use the latest version of the guidance.

14. Actions

76.1 Add TR as actionee
77.2 Close
79.5 Closed
See table at end.

15. AOB

None.

16. Next Meeting

The next meeting will be held 6th December 2023 via Zoom, https://scsc.uk/gd

17. 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
69.3	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

Ref	Owner	Description	Target Guidance Version
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
76.1	MP, PH, MT, TR	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.	-
77.2	MP, PH, MT	Review proposed updates and add to list of changes to version 4.0 of the guidance	4.0
78.1	MP/MA	Investigate representation / format issues further. Consider adding another issue to the next version of the guidance on representation.	4.0
78.2	MP	Investigate data-data interactions further.	4.0
79.1	MT	Introduce EBA / EBTA with a short presentation at next meeting	
79.2	DA	Investigate the Chandryan 2 crash further and report back at next meeting	
79.3	PH	Report on use of the RADISH tool at next meeting	
79.4	PG	Re-visit the earlier nuclear data safety work using the RADISH tool	
79.5	MP	Review the DSG accident list in Appendix H and see if a better sector spread can be achieved	4.0
80.1	RR	Ask Jennifer Kracht at TomTom if she could assist with the new DSIWG poster	
80.2	TR	See which appendices are candidates to work into the main body of the document	4.0
80.3	DA, MA	Write appendix on tooling and RADISH	4.0
80.4	DA, MA	Update the data-safety.tech web site with all known data safety accidents (from past DSIWG minutes and any other sources).	
80.5	MT	(i) Add a ref to the data-safety.tech website in the guidance and (ii) Use the updated data-safety.tech website to produce a better spread of accidents for the new version of the guidance	4.0