

SCSC Data Safety Initiative – WG Meeting 53

1st July 2020, Teams/Teleconf

Minutes

Attendees

Mike Parsons (MP) – CGI, Paul Hampton (PH) – CGI, Tim Rowe (TR) – GTR Safety Management Ltd, Mark Templeton (MT) – Arcade Experts (Part), Brent Kimberley (BK) – Region of Durham, Dave Banham (DB) – Blackberry., Alastair Faulkner (AF) – Abbeymeade, Andy Williams (AW) – independent consultant, Jim Mateer (JM) – SQEP, Louise Harney (LH) – Leonardo, Paolo Giuliani (PG) – EDF (Part), Paul Mayo (PM) – SQEP.

Apologies

None.

Agenda

- COVID-19
- Data-Centric Safety Book
- SafetyTech (from Lloyds Register)
- Response to Thor’s Questions
- SSS’21 Abstracts
- Content of Guidance v4.0
- Sales/Downloads Update
- Future Events
- Minutes and action status
- AOB, etc.
- Data Safety in the News
- Next Meeting

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 [1].

1. COVID-19

MP talked about the Covid-19 response and in particular a recent news article: “*Coronavirus: Local testing data to be shared with councils*” [2]. In the article there was acknowledgement of the

importance of the timeliness of data, which he said was clearly a data safety concern. Another recent article stated that a lack of local Covid-19 testing data is hindering the UK's outbreak response [3].

MT also commented on the quality of data that has been available, especially around the information being provided by other countries. For example, he thought data from China (for whatever reasons: cultural, political or the collection methodology) has been inaccurate and the models on which the data is based were incorrect e.g. the rate of infection being roughly half of what it should have been. He raised the question on how allowances could be made for the possibility of misinformation in a data supply chain.

It was also noted that the UK data for Covid-19 rates has looked worse than other countries, but other countries were not collecting data as diligently or collecting at different rates, which has skewed the results.

It was also felt that the public and the media have not really understood the data and its significance.

AF asked how good the models were, and whether they had undergone any assurance. PH said there was an article on this subject in the May edition of the SCSC Newsletter by Harold Thimbleby, <https://scsc.uk/r158.1:1>. Harold concluded that the quality is poor and not assured to any quality standard. Indeed, the article stated that some of the code is so poor written and documented, it is difficult to verify whether it is correct. The group was extremely surprised by this, given the level of criticality of decision making the models were being used for.

BK said that in Canada, they are researching the collection of sewage samples to help model and anticipate a second wave of Covid-19 [4].

2. Data-Centric Safety Book

AF has now published a book: *"Data-Centric Safety – Challenges, Approaches and Incident Investigation"* [5] based on ideas on data safety he first had on the topic in 2013. He said the book supports the guidance and amongst other topics, summarises the issues with data-centric systems and asks the bigger questions such as how data diversity can be used in the same way it is done for software to help reduce systematic errors. He noted that the largest section is on data-safety accident investigation, as data related accidents present some unique challenges.

3. SafetyTech (from Lloyds Register)

MP introduced an intelligent safety wear report from Lloyd Register, which he thought discussed technologies that would have data-safety significance [6]. This was thought to be most applicable to the industrial, energy and utilities sectors. His main concern was that there would be a long safety chain, including analytics phases, so it would be hard to be confident that everything is working.

BK, who works in these industries, shared his experiences in this area:

- There is an issue with calibration of sensors and not taking sensors offline while calibrating;
- Often only expert judgment is used to assess if data is correct;
- Need to also assess meta data errors in the stack that accompanies the raw data;
- Timely traceability is important from an operational perspective; the value of the data may be questionable if it is not possible to trace back to the sensor or virtual sensor in under 5 minutes.

With regard to non-destructive testing, BK said that it is a good idea to set sensor uncertainty and bias expectations. To that end, for physical systems, from a control 'quantum in time' perspective, you can determine how frequently your system violates the laws of classical physics, and use that as a starting point to evaluate uncertainty, bias, etc.

MT said it may be something the group needs to cover in the next version of the guidance.

4. Response to Thor's Questions

PH noted that Thor Myklebust had provided some feedback and questions on the Guidance in an email on 10th June 2020. DB had provided an initial response to some of the questions. Other questions were discussed and it was agreed to address the comments in the next version of the guidance.

5. SSS'21 Abstracts

MP said he has received a good number of abstracts for SSS'21 and some with data-safety significance were discussed:

- The use of digital twins: these require vast amounts of data and complex modelling and so are heavily reliant on data;
- Misuse of statistics in software reliability;
- Safety Indicators;
- Using Bayesian Modelling to predict software incidents.

6. Content of Guidance v4.0

MP said the Guidance needs to be updated to cover things emerging from the Covid-19 pandemic such as (large scale) modification/filtering or falsification of data and cross validation. It was also agreed to add some content on digital twins.

Action 53.1 [MT] To consider what we need to do in the guidance for digital twins.

DB said there has been some progress on the ontology, but the data ontology is on hold until the more general risk model is matured.

7. Sales/Downloads Update

MP said there were 324 downloads of the electronic Data Safety Guidance Version 3.2 to date.

PH said there were 4 purchases of the Data Safety Guidance v3.2 hard copy in June.

8. Future Events

The New Safety Analysis Techniques seminar will be hosted now in November, and could be in a remote and more spacious location such as Hartham Park in Corsham (where Ebeni have an office).

It is still hoped to hold the Multicore seminar in London on 24th September 2020, using both local and remote technologies. [Now moved to April 2021]

The December event is still planned to be held in Thales premises in Reading.

For SSS'21, certain presentations, such as keynotes, may be streamed live, but the plan is to record everything else for access later. MP noted that 70 attendees are required to make the event financially viable.

9. Minutes and action status

See table at end.

For action 44.2, LH said Wikipedia have been difficult and have now deleted the article that was previously submitted. The action has been reassigned to MP to discuss with AK to come up with an alternative form of text.

It was discussed whether the guidance should be extended to look at security related properties like confidentiality, and hence cover security informed safety aspects. For example, the guidance could look at security related ways in which properties could be compromised.

BK suggested that the group look at 'principle of virtual work' – least action to compromise a given system from a security perspective – to look at exploitability. The term derives from the Lagrange–d'Alembert principle used to help solve mechanical engineering problems.

10. AOB, etc.

None.

11. Data Safety in the News

See articles [2], [3] and [4] referenced in meeting.

12. Next Meeting

A Microsoft Teams meeting in August was proposed [Now September 2020].

13. Thanks

Thanks to all for taking part. Thanks to PH for minutes and MP for chairing the meeting.

Summary of Open Actions

Rows have been greyed-out to indicate that the actions were closed during this meeting. Those entries will be deleted from future versions of the action log.

| Ref | Owner | Description | Target Guidance Version |
|-------|-------|--|-------------------------|
| 42.6 | PH | Define the process to publish a document developed in Overleaf via Amazon | 4.0 |
| 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 |
| 44.1 | MT | Review last 12 months of DSIWG minutes and put any actions referring to v4.0 into Appendix O. | 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 |
| 49.6 | MT | Review Overleaf briefing material and aim to hold a briefing Webex (between 16 th and 20 th December 2019) in the use of Overleaf in the production of the guidance. | N/A |
| 49.11 | DA/MP | Prepare an introductory email to send out to the DSIWG group inviting people to join the tooling subgroup. | N/A |
| 50.4 | PG | Arrange meeting at EDF to demonstrate the tool | N/A |
| 50.5 | DA | Start up the sub-group and initiate first teleconf | N/A |
| 51.1 | MA | To ask Brian Jepson to set up a Kanban board for the tooling | N/A |
| 51.2 | DA | provide access to the tool for all registered users. | N/A |
| 52.1 | All | Contact MP or PH if you would like to contribute to the LRF proposal or work on activities | N/A |
| 52.2 | All | Contact MP or MT with any news accident stories | N/A |
| 52.3 | MP | Contact RO of Rolls-Royce to seek further input on some of the new areas for data safety | N/A |
| 52.4 | All | If interested in taking part in some of the TSG activities, please contact DA directly | N/A |
| 52.5 | MP | Ask MT / PH to see which of the previously identified updates can be done for DSG3.3 | 3.3 |
| 53.1 | MP | To consider what we need to do in the guidance for digital twin. | 3.3 |

14. References

| Ref | Title | Location |
|-----|--|---|
| [1] | Meeting slides | https://scsc.uk/file/gd/53rd_DSIWG_Slides_V2-784.pptx |
| [2] | COVID-19 and data | https://www.bbc.co.uk/news/uk-53244626 |
| [3] | Lack of Testing Data | https://www.ft.com/content/301c847c-a317-4950-a75b-8e66933d423a?fbclid=IwAR16KRvf_CKgM-tMdASeWQFGyXBuvw8l0-SlQOu78LkFEwjzproiy-YJOOQ |
| [4] | What you flush away could help uncover the arrival of a second Covid 19 wave | https://news.ontariotechu.ca/archives/2020/06/what-you-flush-away-could-help-uncover-the-arrival-of-a-second-covid-19-wave.php |
| [5] | Data Centric Safety | https://www.elsevier.com/books/data-centric-safety/faulkner/978-0-12-820790-1 |
| [6] | Lloyds Register Safetytech Market Report | https://info.lrfoundation.org.uk/l/12702/3461-LRF-SafetyTech-report-pdf/97r81j |