SCSC Data Safety Initiative – WG Meeting 58

12th January 2021, Zoom

Minutes

Attendees

Mike Parsons (MP) – University of York, Divya Atkins (DA) – MCA, Martin Atkins (MA) – MCA, Paul Hampton (PH) – CGI, Dave Banham (DB) – Blackberry, Dale Caldicott (DC) - BAE, Michael Green (MG) - Ecomergy, Paolo Guiliani – EDF, Paul McKernan (PMK) – DSTL, Jim Mateer (JP) - SQEP, Emma Taylor (ET) - SaRS, Carl Tipton (CT) – Johnson Matthey, Mark Templeton (MT) - Qinetiq, Andy Williams (AW) – NewTechNo.

Apologies

Louise Harney (LH) – Leonardo

Agenda

1. Intro and Welcome
2. Covid-19 Bad Data Pandemic
3. Update to V3.3
4. SSS' 21
5. SCSC Events
6. Minutes and action status
7. Next Meeting
8. AOB, etc.

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/58th_DSIWG_Slides_v1-991.pptx

1. Introduction and Welcome

MP welcomed the group and thanked everyone for attending, noting that Zoom appeared to working for us.

2. Covid-19 Bad Data Pandemic

MP noted that NH has produced some information on what has gone wrong with data handling during COVID 19 pandemic. MP summarised some of the issues from this, with the items in bold (10-13) added in the meeting:

1.	In predictive models where many assumptions may be wrong[1], or the algorithms inappropriate
2.	Presentation of statistics in a selective or misleading way
3.	Methods of data collection - selective or incomplete. E.g. reporting on the number of positive test cases is misleading - many people asymptomatic with the virus so never get tested.[2]
4.	Calculations and filtering
5.	Allowances for delays in collection or processing
6.	Intentional and unintentional bias
7.	Use of averaging (e.g. moving averages) and smoothing of plots hiding sudden increases
8.	Loss of data (e.g. in the recent UK Test and Trace system due to old versions of Excel)
9.	Misinformation e.g. on vaccination risks
10.	**Largely unauditable. Poor rigour. Amateur vs. Engineer view.**
11.	**Need better 'Data Hygiene'**
12.	**Underlying issues communication, education and awareness of data risk**
13.	**Need better software (Noting Harold Thimbleby's upcoming SSS'21 talk)**

There was a lengthy discussion on aspects of these issues. Also discussed were Covid 19 "Data Reliant" Systems where poor data can produce poor results. MP also discussed with ET her thoughts on the communications issues associated with Covid data and ET noted that she will include a section on this in her 10th Feb SSS'21 presentation.

3. Update to V3.3

Updates to V3.3 of the guidance document were discussed including:

- Dark data
- Addressing Thor Myklebust's comments
- Adding PH's mapping tables between data issues and loss of data properties
- Updates to the Organisational Data Risk (ODR) form
- Covid 19 section
- Reference to Alistair's book

There was a discussion about how the data war story cases could be renumbered such that adding new examples did not mean that all references needed to be updated. A solution was proposed based on using the date.

---

[1] PMK. There has also been a challenge with population behaviour with significant numbers refusing to follow the guidance and thus skewing the predictions

[2] PMK. One of the challenges or the modelling and simulation is that they all rely on an infection rate and we have been unable (indeed it may be impossible) to obtain an accurate infection rate thus the basic data is incomplete at the very start…

4. SSS' 21

MP provided an update on the plans for SSS' 21, https://scsc.uk/e683 . The speaker for the evening 'after dinner' session on the Wednesday will be Patrick Hudson, https://flightsafety.org/wp-content/uploads/2017/02/Patrick-Hudson.pdf . ET will also be presenting on communication issues. Two weeks after event there will be a follow up session https://scsc.uk/e791 to address questions that could not be answered at the time, because some of the presentations will be pre-recorded.

5. SCSC Events

SSS'21 and follow-up, https://scsc.uk/e683. There will also be a safety culture event in early April 2021.

6. Minutes and action status

See table at end.

7. Next Meeting

Next meeting by Zoom 23rd February 2021. MP to arrange. [See: https://scsc.uk/e794]

8. AOB

Post meeting notes:

1. At the previous meeting, RR had presented on their ethical framework for AI. This framework has now been made generally available. https://www.rolls-royce.com/sustainability/ethics-and-compliance/the-aletheia-framework.aspx

2. MA's proposed text for a new war story in the guidance document v3.3 related to the loss of Excel track and trace data: https://scsc.uk/file/gd/CovidSpreadsheet_-_War_Story-v1-994.docx

9. Thanks

Thanks to PMK for taking the minutes and MP for chairing.

Summary of Open Actions

Actions 50.5, 51.2, 55.1 and 57.2 closed.

Ref	Owner	Description	Target Guidance Version
42.6	PH	Define the process to publish a document developed in Overleaf via Amazon	3.3
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.	3.3
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 before end of March 2021 in the use of Overleaf in the production of the guidance. #58 MT to provide updated slide set and on-line briefing	N/A
53.1	MP	To talk to Kevin King about what we need to do in the guidance for digital twin.	3.3
55.2	MP/MT/PH	Create appendix on Dark Data for version 3.3 of the guidance	3.3
56.1	PH	Consider linkage of Data Safety Properties to Dark Data types	3.3
56.2	DA	Consider impact of Dark Data on the Data Safety Tool	N/A
56.5	PH	Address Thor's comments on the guidance	3.3
56.6	PH	Consider suggested updates to the cover of V3.3, e.g. elephant with mask(!)	3.3
57.1	ALL	Suggest updates to NH contribution on Covid-19 and data	3.3