DSIWG

SCSC Data Safety Initiative – WG Meeting 42

28th September 2018, NATS Brettenham House, London

Minutes and Actions

Attendees

Paul McKernan (PM) – DSTL, Nick Hales (NH) – Consultant, Dave Banham (DB) - Rolls–Royce PLC, Louise Harney (LH) – Leonardo, Mike Parsons (MP) – NATS, Bill Blackburn (BB) – Process Renewal Group, Gordon Hurwitz (GH) – Thales, Mark Templeton (MT) – QinetiQ, Martin Atkins (MCA) – Mission Critical Applications, Divya Atkins (DA) - Mission Critical Applications

Apologies

Ali Hessami (AH) – Vega, Mike Ainsworth (MA) – Ricardo, Paul Mukherjee (PM) - Astellas, Paul Clugston (PC) – Consultant, Phil Williams (PW) – Engineer For Safety, Graham Meaden (GM) – Kipstor, Robert Green (RG) – NATS, Amira Hamilton (AH) – CGI, Duncan Dowling (DD) – DARD, Robert Oates (RO) - Rolls Royce PLC, Paul Hampton (PH) – CGI, Phil Wright (PW) – DSTL, Fan Ye (FY) – ESC, Chris Hartgroves (CH) – Leonardo, Shaun Cowles (SC) – EDF, Paul Dart (PD) – NCC Group, Maria Kelly (MK) – Leonardo, Andrew Eaton (AE) – CAA, Ashley Price (AP) – Raytheon, Bernard Twomey (BT) – Rolls-Royce PLC, Peter Smith (PS) – Highways England, John Bragg (JEB) – MBDA

Agenda

1. Update on Tools / Lloyds Register Foundation (MA, DA)
2. Data Safety & Autonomous / Machine Learning Update (All)
3. Data Strategy Session (All)
4. Training Course Development (MT)
5. New Health Trifold (PH/MP)
6. SSS'19 Update (MP)
7. Overleaf status (MT)
8. Move to KDP and Sales/Downloads Update (MP)
9. Formal Modelling Activity Update (DB)
10. Dissemination update (All)
11. Standards update (All)
12. Future Events (MP)
13. Minutes and actions status (All)
14. AOB (All)
15. Data Safety in the News (All)

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].*

Slides used in the meeting are available in [1].

1. Update on Tooling / Lloyds Register Foundation

DA and MCA provided an update on the activities to support data safety tooling. A lot of dissemination has been done through this activity, which is beneficial in itself. In parallel with the development of a proof-of-concept demonstrator, DA/MCA have discussed a student project with a lecturer in the Computer Science Department at Bath University, to explore an alternative way of modelling the Data Safety Guidance (DSG) for tool support. More details to follow if the project is selected by a UG/MSc student this year.

Feedback from the Alan Turing Institute was that from their perspective 'data services' and 'data sciences' are different domains (although related), so this would not fit into their data-centric engineering funding. They referred the enquiry back to Lloyds Register Foundation (LRF).

LRF have given encouraging feedback, but would like to see letters of interest from potential industrial collaborators to justify a grant to support development of the tool. They also proposed that a steering group direct this work - members of the DSIWG would be obvious candidates. DA can supply information on the time investment required from a member; a rough estimate would be around 10-12 days, so a business case needs to be presented to each participating company.

The proposal to LRF was to develop an Open Source tool, with no licencing fees, but with fees for customisation and ongoing support. Dual-licensing with a second commercial license was also discussed, since some organisations might have difficulty using Open Source software directly. Either license would allow licensees to see, use and modify the source code. The need to integrate with existing requirements and safety tools was discussed, and it was agreed that the Data Safety tool will need to have open interfaces so it can be integrated with existing tools.

Example data sets would be needed to aid development of the tool, public datasets were suggested, and the collaborators would be asked if they have suitable material.

MP raised the question how will the tool be baselined against the version of the guidance it has been developed against to maintain traceability? MCA replied that the tool would support several versions of the guidance simultaneously.

Action 42.1 – MP to send an email to all members of the DSIWG requesting expressions of interest to support the tool development

Any volunteers to join the steering group or to collaborate on development of the tool should contact DA / MCA.

An outline business case containing the following information would be needed to support other company's involvement:

- Start / end dates and duration
- Effort (i.e. total number of hours to be spent)
- Expenses and material costs
- Objectives

- Description (i.e. what is the tool and why do we need it; also needs to specify how this will work with IT/security procedures)
- Relevance to the business (e.g. benefit of an external tool / COTS tool)
- Why is this innovative?

Action 42.2 – DA/MCA to provide an outline business case which would be needed to promote the tool, stating what the tool does and what its benefits would be

Defence Standard 00-970 *'Design and airworthiness requirements for service aircraft'* includes a requirement for data safety arguments to be created, so there is a risk to the business of being unable to do this. Managing that business risk may include purchasing / acquiring a COTS tool, rather than trying to understand all the guidance available. This would support the business case.

2. Data Safety & Autonomous / Machine Learning Update

Action 42.3 – PM to talk to colleagues at DSTL working on autonomy/machine learning/artificial intelligence about data implications and what guidance is needed on how this data is managed.

During the meeting PM completed this action by emailing colleagues. It was agreed to contact Rob Ashmore (DSTL) to confirm that tasks are not overlapping between working groups. Support will be provided by DSTL, but someone else will need to lead specific development of the guidance.

Test coverage for machine learning and artificial intelligence data sets was questioned. For neural networks, rather than nodes and connections it is the output of the network which is of interest.

3. Data Strategy Session

An update was provided on each of the data strategy sessions.

More work is needed on cloud implementations with respect to data, in particular the use of software-defined virtual environments and the unclear mapping to hardware.

A consultation has begun in relation to NIST IR 8228, Considerations for Managing Internet of Things (IoT) Cybersecurity and Privacy Risks: https://csrc.nist.gov/publications/detail/nistir/8228/draft
It would be useful for members of the working group to review this document.

Action 42.4 – LH to get an update from BO on cyber-data properties mapping, which has been published.

It was agreed to add multicore and FPGAs to the list of data strategy topic areas
[the full list is now: **Internet of Things and OT/IT Convergence; Cryptography; Big Data; Cloud Computing; Cyber-Physical Systems; Machine Learning and Autonomy; Wireless Communications; Multicore; FPGA**]. FPGAs are already analysed already from a safety perspective by some organisations, but not necessarily from the data parameters perspective, so more work is needed.

4. Training Course Development

MT provided an update on the data safety training course development [2] and the 4 runs of the course already undertaken. Some benefits have already been observed, including changes in ways of managing databases. One weakness was that the guidance stops at 'here are some interesting techniques'. Safety requirements were not well understood by people who do not derive these on a day-to-day basis.

The course is still being improved from a visual perspective, so is not yet released. However, some feedback had already been received:

- Alignment of data types and data categories is required [the meeting agreed that the 'Data Types' should be renamed to 'Data Categories' throughout, and that the consistency of the second set of data types in the guidance (supposed to be a strict subset of the first set) should be checked.]
- There should be a role for someone responsible for data in the system
- MT explained how the data available to the ATCOs during the Uberlingen accident contributed, in combination with training, to the accident. Found this helped participants understand the causes of the accident when talking about the data properties, so the accident examples are a useful inclusion in the guidance.
- Someone needs to review the worked example in the guidance to confirm that it covers all of the process, and actually works with Version 3.0 of the guidance.

Action 42.5 – MT to make a PDF of the guidance document version 3.0.1 available for SSS'19 to support reporting of progress

5. Update of the Guidance Document

It is intended to release the next update of the guidance in February 2020, so it is necessary to begin work on the draft update.

Action 42.6 – PH to define the process to publish a document developed in Overleaf via Amazon

Feedback was provided from MT and MC after using the guidance:

- The guidance does not state where we go at the end of the process, such as how to mitigate the issues identified and define appropriate safety requirements
- The guidance seems to suggest that the lowest DSAL identified should be applied rather than the highest DSAL. This does not seem correct, so needs to be confirmed
- Within 'High-level mitigations', Editing limitations (e.g. encapsulation of data / access limitations) was not understood
- The guidance does not clearly separate between flows of data and items of data
- Some good outcomes have been found, such as possible improvements in homogeneous redundancy.

Action 42.7 – MP to ask John Spriggs (NATS) to write a template data safety argument for the next guidance update.

6. New Health Trifold

PH has updated the healthcare flyer [3], and it would be useful to have a similar flyer format for other events / purposes. For example, autonomy-focussed flyers would be useful for High Integrity Software (HIS) 2018.

Action 42.8 – MP to get Data Safety flyers printed for HIS 2018 via Alex King.

7. SSS'19 Update

MP showed the draft programme for SSS’19, which has a varied and interesting set of talks.

8. Overleaf Status

MT has been waiting for Overleaf to stabilise following a recent update, so no update available.

9. Move to KDP and Sales/Downloads Update

Amazon have moved all publications from Amazon Createspace to the Kindle Direct Publishing (KDP) platform and this now has to be used for all SCSC publications. This is causing some learning issues. It is still possible to use print on demand however. Metrics are not as good on KDP and the same historical metrics are not available. It is not yet clear how different the publishing process using KDP is.

MP provided an update on the sales since moving to KDP, and the downloads of various documents from the SCSC website. Downloads of the Data Safety Guidance v3.0 have been particularly good.

10.Formal Modelling Activity Update

The modelling is facing some challenges due to applicability of data properties to the various data categories. This would be improved if there was a way to work out which properties need to be considered.

Action 42.9 – MP to work out a matrix of data categories (previously ‘types’) and data properties (based on DB discussion)

DB explained that he needed additional feedback on his current model [4].

11.Dissemination update

No further update was available.

12.Standards update

The ongoing work with IEC 61508 appears to have stalled. Some work has been performed, but no further update was available. It is necessary to have someone deeply involved with that group.

13.Future Events

MP showed the latest status from the SCSC website event page, specifically including:

- COTS, Legacy and Reuse seminar on 6 December 2018 will be useful and is closely related to data safety.
- SSS’19 has already been discussed.
- Evolution of Assurance Case Practice seminar will be held in April 2019.
- There will be a further seminar later in 2019 relating to accident responses.

14.Minutes and actions status

The following action statuses were updated:

- Action 40.1 – MP has identified the correct person to ask and will invite him to the next meeting; still ongoing.
- Action 40.4 – It was agreed to keep this action open and to contact BO for an update.
- Action 41.2 – This action is now closed.
- Action 41.3 – This action is delayed until the next meeting.
- Action 41.6 – This action is now closed.
- Action 41.7 – This action was closed and replaced with a new action:

Action 42.10 – MP to contact Phil Williams and request he shares the ontological model with a wider community.

- Action 41.8 – This action is now closed.
- Action 41.10 – This action is now closed.
- Action 41.13 – This action is now closed.

During meeting 41, an action was closed which is believed to still be relevant, so a new action was raised as follows:

Action 42.11 – MP to ask BJ how to ensure a search for the correct key words such as data safety finds the material created by the DSIWG on the first page of search results.

15.AOB

BB provided an update on the SCL conference he attended. It included a session on AI in the UK relating to a Select Committee on Artificial Intelligence paper '*AI in the UK: ready, willing and able?*'. This report includes many mentions of data. This document is publicly available here: https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf. It would be beneficial to get involved in this study and similar studies happening in future; many universities were included but not the University of York.

The Turing Institute has a research area relating to Data-Driven Design Assurance, which seems to be closely related to the work of the DSIWG. To gain support, it may be necessary to identify where the DSIWG strategically aligns with the Turing Institute research already ongoing.

16.Data Safety in News

MP showed an article relating to someone wrongly arrested after applying for a driving licence due to data sharing between government agencies. Note that this use of data may make the data being shared safety-related. https://www.buzzfeed.com/emilydugan/man-arrested-naked-home-office-raid

MP also showed an article listing current automotive software defects, many of which relate to data. https://betterembsw.blogspot.com/2018/09/potentially-deadly-automotive-software.html

17.Next Meeting

The next meeting date is provisionally planned for 21 November 2018, to be held at Thales, in central London.

18. Thanks

Thanks to LH for taking the minutes.
Thanks to MP for chairing the meeting.
Thanks to NATS for hosting.

19.Summary of Open Actions

Ref	Owner	Description	Target Guidance Version
33.6	LH	Add everyone on the DSIWG distribution list to the LinkedIn page.	N/A
36.4	MC	Coordinate the production of training material (based on v3.0).	N/A
37.3	DB	Ensure the model is included in the next version of the document	4.0
37.5	MP	To coordinate with BJ on the close down old docuwiki, remove content, and refer future authors to MP	N/A
39.9	PH	To feed comments back to Paolo on the worked example.	3.1
39.11	MAs	To look at what funding routes may be available from Nuclear Sector for tooling.	N/A
40.1	MP	To get an expert from NATS to brief the WG on the configuration data challenges of cloud.	N/A
40.2	VM	To better define 'Cloud' in the context of data safety and cloud. What level (scope) should this be defined at?	N/A
40.3	All	Specific domain owners to identify experts to brief the DSIWG on their domain so as to identify the hazards and issues	N/A
40.4	PM	To contact RO and/or DB about real world examples of big data applied to data safety.	N/A
40.9	MP	To check with Audrey Canning on access to IEC 61508/update cycle. May be better to have SCSC as a permanent representative on 61508 committee	N/A
40.10	MT	To ask if his Qinetiq course can be released to DSIWG and be presented at SSS'19	N/A
41.1	ALL	Post on LinkedIn about new topics discussed under Data Safety Strategy	N/A
41.2	MP	Email around members asking for any input on the topics discussed under Data Safety Strategy	N/A
41.3	ALL	Review ontological model in own domain and provide review comments for DSIWG Meeting #42	N/A
41.4	TBD	Review Data Safety Guidance and identify where ontological model has highlighted inconsistencies, for discussion at the DSIWG Meetings *Note: owner TBD dependent on time available.*	4.0
41.5	BT	Find out who updates each of the standards the ontological model is based on and provide information to MP / LH	N/A
41.6	LH	Speak to TK about the possible impact on standards through the SCSC SG, with the view to raise a standing SCSC SG agenda item	N/A
41.7	MP	Think of how to better to broadcast the ontological model within the SCSC and wider	N/A
41.8	MP / BJ	Update Data Safety Guidance PDFs with ligature problems fixed	N/A
41.9	JB	Follow up with MT on how change tracking works in Overleaf and whether guidance can be added directly to Overleaf	N/A
41.10	PH	To update the healthcare trifold to include the NHS logo, if approved, and arrange distribution of the trifold at key healthcare events	N/A
41.11	BT	Provide key event dates at which data safety dissemination may be useful / welcomed	N/A
41.12	ALL	DSIWG members to identify any influential conferences in their domain for discussion at DSIWG meeting #42	N/A
41.13	MP	Agree date and location for next DSIWG meeting	N/A
42.1	MP	Send an email to all members of the DSIWG requesting expressions of interest to support the tool development	N/A

Ref	Owner	Description	Target Guidance Version
42.2	DA/MA	Provide an outline business case would be needed to get any traction going forwards stating what the tool does and what the benefit of it is	N/A
42.3	PM	Talk to colleagues at DSTL working on autonomy/machine learning/artificial intelligence about data implications and what guidance is needed on how this data is managed.	N/A
42.4	LH	Get an update from BO on cyber-data properties mapping, which has been published.	N/A
42.5	MT	Make a PDF of the guidance document version 3.0.1 available for SSS'19 to support reporting of progress	N/A
42.6	PH	Define the process to publish a document developed in Overleaf via Amazon	N/A
42.7	MP	Ask John Spriggs to write a template data safety argument for the next guidance update	N/A
42.8	MP	Get Data Safety flyers printed for HIS 2018 (via Alex)	N/A
42.9	MP	Work out a matrix of data categories (previously 'types') and data properties (as per DB discussion)	N/A
42.10	MP	Contact Phil Williams and request he shares the ontological model with a wider community	N/A
42.11	MP	Ask BJ how to ensure a search for the correct key words such as data safety finds the material created by the DSIWG on the first page of search results.	N/A

20.References

Ref	Title	Location
[1]	Slides from the meeting	https://scsc.uk/file/gd/42nd_DSIWG_MP_Slides-485.pptx
[2]	Mark Templeton update	https://scsc.uk/file/gd/Mark_Templeton_input-486.docx
[3]	Health Trifold Update	https://scsc.uk/file/gd/Working_Group_Trifold_Healthcare_Example_v4-487.docx
[4]	Ontology model paper	https://scsc.uk/file/gd/Proposal_for_a_model_of_data_risk_management_(full_report)_vB.1-464.pdf