DSIWG

SCSC Data Safety Initiative – WG Meeting 68

30th March 2022, Zoom

Minutes

Attendees

Oscar Slotosch (OS) – Validas, Paul Hampton (PH) – CGI, Mike Parsons (MP) – AAIP, Andy Williams (AW) – Consultant, Dave Murray (DM) – BAE, Nick Hales (NH) – Consultant, Martin Atkins (MA) – MCA, Divya Atkins (DA) - MCA.

Apologies

Richard Garrett (RG) – SQEP, Mark Templeton – Qinetiq, Michael Green (MG) – Ecomergy, Mark Nicholson (MN) – University of York, Paul McKernan (PMcK) - Dstl, Fan Ye (FY) – ESC, Tim Rowe (TR) – Consultant, Paolo Giuliani (PG) – EDF, Sam Robinson (SR) – EDF

Agenda

1. Welcome
2. Reception of DSG V3.4
3. DSTL 'Crumbs'
4. Bosch Australia Interest
5. Dragon Kings, Black Swans and Data
6. Retrospective: Aims of DSIWG#1
7. Migrating, Porting and Importing Data
8. New BSI Guidance
9. AOB
10. Actions
11. 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: https://scsc.uk/file/gd/68th_DSIWG_Slides_v1-1365.pptx

1. Welcome

MP opened the meeting and welcomed those attending.

2. Reception of DSG V3.4

MP said that V3.4 of the Data Safety Guidance had been well received and noted that production had gone well with the editorial team of OS, MT, PH and MP working hard over Christmas. Many additions and improvements were made.

Download: https://scsc.uk/scsc-127G [179 downloads since SSS'22]
Amazon: https://www.amazon.co.uk/Data-Safety-Guidance-v3-4-Templeton/dp/B09RG3K1Z7

PH suggested that the SCSC could run a 'masterclass' course on data safety. MP said this is definitely worth considering, especially in sectors where there is a requirement to address data safety.

3. DSTL 'Crumbs'

MP noted that the DSTL booklet "DSTL: Crumbs! Understanding Data" was available at SSS'22. This features the DSG quite a bit. PDF version at:
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1023385/20210901-Crumbs_biscuit_book_DIGITAL.pdf.

4. Bosch Australia Interest

MP had an email exchange with Malcolm Watts, Robert Bosch (Australia) about Data Safety training courses / materials. He has responded with what is available on the web site but anyone wishing to contribute more should contact MP initially.

5. Dragon Kings, Black Swans and Data

MP and PH presented some of the work they have been doing on Black Swan and Dragon King events related to data (see slides):

- A Black Swan (BS) is an event that comes as a surprise, has a major effect, and can change our world view after the event. It is often rationalized[1] after the fact with the benefit of hindsight.
- A Dragon King (DK) is an event that is both extremely large in size or impact and born of unique origins relative to other events from the same system. DK events tend to occur in non-linear and complex systems, and can amplify to extreme levels.

There followed some interesting discussion. MP mentioned that there are really two types of *data-related black swans*: (i) where data problems eventually lead to the BS event, and (ii) where the BS is within the data itself (i.e. something wholly unexpected within the data).

PH described the cases he had found which were discussed by the group. PH mentioned some of the 'pathological data' cases he had come across and MP noted the 'bad flight plan' incident. There was thought to be some relationship with 'corner cases' in testing which are rare and can cause severe problems.

It was agreed that a BS is from a particular perspective / stakeholder, i.e. some people may aware of the problem (but have not communicated it to others) and so it is not a BS to them.

[1] Sometimes inappropriately

MA mentioned that can also have positive BS, e.g. the smartphone has totally changed society and not many people saw it coming[2,3].

There was a discussion of what is a DK and what is a BS. Where there is engineering to address a 'once in a hundred years' problem then that is more likely to be a DK, as it has been anticipated. A BS is unlikely to have been considered and will likely not be in the hazard log.

MA mentioned the issue of CRC32 checksums in very large TCP/IP network flows (e.g. in data centres), where, even with the additional TCP checksum, people are reporting several undetected corruptions per month due to the huge quantity of data flowing.

There was noted to be a risk in suppressing rare cases or even 'fake news' information as some of this may suggest new BS cases. The relationship to Dark Data was also noted.

Action 68.1 (MP/PH) – Develop the Black Swan / Dragon King Data work further and consider publishing as a newsletter article

6. Original Objectives Review

MP presented the original DSIWG objectives from 2013 (https://scsc.uk/file/gd-main/Data-Safety-Initiative-Meeting-1-KO-240113-FINAL.pdf), together with an approximate score:

Agreed Objectives of Group for 2013	Score?
1. Produce cross-sector guidance by end of the year including a clear statement on handling of data as a separate component within safety related systems	A*
2. Produce a high level strategic plan by the end of the year for fuller adoption e.g. into existing standards or a new standard	A
3. Influence standard updates currently in progress where we can	B
4. Actively promote and disseminate objectives and outcomes of the initiative to the wider community, professional bodies, etc.	A

It was thought that the objectives had been met, but that the objectives going forward should be regularly reviewed, e.g. every 24 months. They should also be reviewed at the next meeting.

7. Migrating, Porting and Importing Data

MP presented slides on migration of data addressing action 62.3 developed with MT. It was suggested that could be developed into material for the next issue of the guidance. It was thought the cases should be expanded to four:

1. Importing
2. Exporting
3. Porting

[2] MA: In retrospect, a simple application of the consequences of Moore's law "should" have made smartphones obvious!

[3] NH: I am not sure that it is relevant for data safety in systems. I'm thinking of surprise data in a software programme, could it ever be positive?

4. Migrating

At least 13 issues were identified, and these need to be reviewed, expanded and mapped to the four cases above:

1. Data may not map: different database formats, endian issues, etc
2. Data may not translate: no equivalent data/record/field type in new system, etc
3. Data may not be valid: less numeric range available in new system, etc
4. Data may have different meaning in new system context
5. Data may be correctly rejected (detection of error on import/export)
6. Data may be wrongly rejected as incompatible
7. Two or more data items may map to the same item in the new system – one may overwrite the other
8. Migration may be incomplete: loss of data
9. Migration may change values: corruption of data
10. Migration may be out of date: time taken to do migration causes stale data
11. Migration may create conflicts with some values: overwrite existing data, or both migrated separately
12. Use of cloud storage and hosting (reduction in visibility and control over data process, backup etc)
13. Configuration management of changes during migration

It was noted that there may be overlap with the existing migration table of mitigations in the existing DSG and this needs to be resolved.

Action 68.2 (MP) – Develop the migration work further and present at next meeting

8. New BSI Guidance

It was noted that some new guidance from BSI addresses issues of data explicitly in new standards: https://www.bsigroup.com/en-GB/standards/bsi-flex-236-v1.0-landing/

9. AOB

NH presented a slide on data and environmental issues:
https://scsc.uk/file/gd/Presentation_for_DSIWG_March_22-1366.pptx .
MP mentioned that a proposal to start a WG on environment is tabled for SCSC Steering Group next week. He will report back.

There was a discussion as to what is different in environmental safety and MP suggested:

1. Potentially longer time frames for harm to materialise
2. Likely wider scope
3. Many more interactions and relationships between elements (& stakeholders)
4. Less well-defined boundaries
5. Could be one or any combination of the following: commercial, governmental, national, NGO and individual interests
6. More unknowns and ‘Dark Consequences’, i.e. we don’t know what the effect will be
7. Currently no known formal assurance or use of safety cases, etc.

10.Actions, etc.

See table at end.

11.Next Meeting

Next meeting will be held end May 2022 in Bath UK and on Zoom. MA/DA to arrange.

12.Thanks

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

8. 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
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.	N/A
53.1	MP	To talk to Kevin King about what we need to do in the guidance for digital twins.	4.0
56.2	DA	Consider impact of Dark Data on the Data Safety Tool	N/A
61.2	AW	Research the relevance of digital currencies and report back to the group (with MA and MT)	4.0
62.3	MP/MT	Consider issues of porting and importation of legacy data (see 68.2)	4.0
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
64.1	MP	Contact Thor and establish the details of the guidance proposals in the paper.	4.0
65.1	MP	Contact Davy Pissoort and see if any interest in this funding route for Data Safety	-
65.2	DA/MA	See if access to the tool can be given to TomTom for evaluation	-
66.6	MT	Add these three properties ['Analysability', 'Explainability', 'Verifiability'] to the user-visible further work section. If time allows then develop into the guidance further.	3.4
66.9	ALL	Volunteers are needed for proof-reading the drafts of the new guidance document version. Please contact MP if you can help in January 2022.	3.4
68.1	MP/PH	Develop the Black Swan / Dragon King Data work further and consider publishing as a newsletter article	4.0
68.2	MP/MT	Develop the migration work further and present at next meeting	4.0