



DARE UK Scientific and Technical Advisory Group (STAG) Meeting

Date: Tuesday 17 May 2022

Time: 10.30am-12.00pm

Location: Zoom

Attendees

STAG Members

- Professor David Ford (Chair) Professor of Health Informatics, Swansea University
- Dr Claire Bloomfield, Deputy Director Value of Data Centre for Improving Data Collaboration, NHSX
- Professor David De Roure, Professor of e-Research, University of Oxford
- Professor Ben Goldacre, Director, Bennett Institute for Applied Data Science
- Alison Kennedy Director, STFC Hartree Centre
- Phil Kershaw, Technical Manager, Centre for Environmental Data Analysis, Harwell Camp
- Maisie McKenzie, Public contributor
- Chris Monk, Public contributor
- Professor Tom Rodden, Chief Scientific Adviser, UK Department for Digital, Culture, Media, and Sport (DCMS)
- Professor Elena Simperl, Professor of Computer Science, King's College London
- Peter Stokes, Director of the Integrated Data Programme, Office for National Statistics

UKRI and DARE UK representatives (observers)

- Hans-Erik G. Aronson, Director, DARE UK Phase 1
- Professor Gerry Reilly, Design Authority, DARE UK Phase 1
- Fergus McDonald, Senior Programme Manager, DARE UK Phase 1
- Michelle Amugi, Programme Manager, DARE UK Phase 1 (Secretariat)
- Dr Ekaterini Blaveri, Interim Head of Digital Research Infrastructure, UKRI Medical Research Council (MRC) (Secretariat)

DARE UK Programme Board members

- John Marsh, Public contributor
- Professor Felix Ritchie, UWE Bristol

Apologies

- Professor Søren Holm, Professor of Bioethics, University of Manchester
- Professor Máire O'Neil, Director of Electronics, Communications & IT, Queen's University Belfast
- Professor Mark Parsons, Director of Edinburgh Parallel Computing Centre and Associate Dean for e-Research, University of Edinburgh

1. Welcome and minutes of the previous meeting

David Ford, STAG chair, welcomed members and he shared apologies from Professor Søren Holm, Professor Máire O'Neil, and Professor Mark Parsons.

STAG members' attention was drawn to the minutes from the previous meeting on March 30th which had been circulated ahead of the meeting (paper [DAREUK_STAG-05-22-2_March 2022 minutes](#)). Members were asked whether they felt that the minutes were an accurate reflection of the discussions which took place in the meeting. In response, a member reflected that they did not adequately reflect the rationale given on the need for a deeper landscape review.

In terms of actions from the March meeting, the following was shared (tabled at the end of [DARE UK_STAG-05-22-2_March 2022 minutes](#) paper):

- Action 1 will be addressed by UKRI once the information is available
- Action 2 will be discussed in the next STAG meeting on June 29th
- Action 3 will be discussed in the next STAG meeting on June 29th
- Action 4 is closed
- Action 5 is ongoing
- Action 6 is closed
- Action 7 is closed
- Action 8 to be discussed in this meeting under item 3
- Action 9 is closed
- Action 10 is closed

Action	Owner(s)	Deadline/ status
Reword minutes from March STAG minutes to better reflect the rationale given for a deeper landscape review to be conducted.	Delivery team	Closed

2. Brief update from the Programme Board meeting

Next, David Ford gave a brief summary of the key points and actions which were identified in the March Programme Board meeting:

- Taking into consideration the STAG's advice, the PB requested the Delivery Team (liaising with UKRI/MRC) to prepare a scoping paper outlining the objective, scope approach, and timeline for a comprehensive landscape review of the UK data infrastructure, which is designed to support sensitive data
- PB recognised the challenge of the task, members agreed that it was important to ensure that the whole activity around digital research infrastructure for sensitive data across UKRI was captured during DARE UK Phase 1 including, identifying tractable research use cases that cannot be addressed by current investments.
- PB advised the Delivery team to consider organising another design-thinking workshop with a different audience to capture the data producer user journey

No comments or questions were raised from STAG members regarding this item.

3. Deeper landscape review discussion

In response to the STAG's calls in the March meeting for the DARE UK programme to conduct a deeper review of the sensitive data landscape, the delivery team prepared and shared a scoping paper ([DAREUK_STAG-05-22-3_Deeper Landscape review](#)) outlining an approach to this piece of work.

Senior Programme Manager, Fergus McDonald, gave a brief summary of the proposed 3 steps outlined in this paper before inviting discussion from the STAG on this proposal:

Step 1:

Establishing a broad baseline overview of sensitive data research digital infrastructures across research domains

Step 2:

Engagement with UKRI research council heads of digital infrastructure to validate the output of step 1, give a sense of priority to the digital infrastructures in terms of their value proposition, and provide additional context from a funder's perspective

Step 3:

Direct engagement with digital infrastructure teams to understand their capabilities and extract lessons learnt from past approaches

STAG members were invited to give their views. Members began by suggesting that the scoping paper should begin with an explicit articulation of the purpose and target audience of this work rather than the approach. Another member added that it is particularly important to give clarity on the purpose of the deeper landscape review, given the resource implications highlighted in the document. It was stressed that the main motivation of conducting this piece of work should be to establish what and where to invest in by gaining an understanding of the existing landscape and where the gaps are. This will allow the DARE UK programme to learn from what is or isn't working and why/why not and focus its investments on key known challenges.

In order to successfully do this, members stated, it would be imperative to ensure that the breadth of this deeper landscape review extends past UKRI investments as, without gaining insight into the capabilities of the UK sensitive data landscape as a whole, it will be difficult to ascertain exactly what gaps and challenges DARE UK needs to focus its investments on.

Additionally, it was highlighted that the ability to identify which initiatives will be within the scope of this exercise, is dependent on what the definition of sensitive data in the context of DARE UK is. Thus, reinforcing the urgency for this definition to be finalised.

A member also stated that this deeper landscape review should not only look at present work but be forward looking, allowing infrastructures to identify what their future plans to develop their sensitive data, digital research infrastructures are, in order to effectively inform where DARE UK should direct their efforts.

Another member ventured the suggestion that it could be useful to focus on a smaller representative sample group and test the questions out on them before then looking to extend the survey to a wider range of people. Citing that this approach would be beneficial as it will be more manageable and allow for things to

DARE UK

be corrected and reframed along the way. The chair responded to say that it might be difficult to create a sample when it isn't clear what the population is like. Nevertheless, he acknowledged the merit in the point being made.

Next, it was suggested that to alleviate the concerns raised in previous meetings that some disciplines and infrastructures have been underrepresented in DARE UK's engagements, a conscious effort to identify and communicate with infrastructures and disciplines which have been underrepresented, should be made to balance this bias. The member suggested that this could be done by focussing the deeper landscape review specifically on infrastructures and disciplines which have not yet been considered. It was suggested that to do this, the DARE UK delivery team could explore the research that the Infrastructure Portal has done by, understanding and exploring the methodology that the infrastructure portal used to develop their case studies as this could offer valuable insight and this approach would have much lower research implications than having to look at all infrastructures which store or process sensitive data.

To supplement this discussion and in response to the delivery team's suggestion that the UKRI infrastructure portal will provide a good starting point for this deeper landscape review, Ekaterini Blaveri, STAG secretariat, shared insights into the current discussions which have been had with the UKRI infrastructure portal team on the challenges to accessing data via this portal:

- The Data was gathered through the UKRI research infrastructure road map work and it's not limited to infrastructures supporting sensitive data.
- Aggregated survey response dataset from landscape analysis exercise cannot be shared due to permissions obtained during collection.
- Sub-set of information from the landscape is analysis exercise available on the infrastructure portal, with additional infrastructures added. However, the only information available via this data set, is what is online.

A member suggested that another useful approach would be to create an open library through a simple web infrastructure with the capability of providing self-reporting templates for DRI providers that ask DARE UK specific questions which can be structured to optimise data capture and intelligent querying of the data such that synthesising outputs and results is efficient. The Office of Science, High Energy Physics, Exascale requirements review, was suggested as a useful point of reference to exemplify this approach as they employed a co-authorship methodology to conduct their review.

The proposal to create an open library via a self-reporting model was welcomed by the DARE UK delivery team and they invited STAG members to collaborate on the development of this. In prompt response, a member wrote up and shared suggestions on a series of questions which could be asked. This included a useful suggestion on how to surface the concerns many raised in the Goldacre Review around programmes which have been funded in the past with little known yield: a separate page where people can refer other programmes/initiations in order to allow lessons to be learnt from only marginally if it all successful investments. It was suggested that this could be achieved via a mechanism whereby programmes can be identified and sent a report to complete based on this identification and, should they decline, this non-response can be noted somewhere for information. This community approach to information gathering would also help overcome gaps in the programme's potentially incomplete knowledge of all relevant sensitive data programmes/initiatives/infrastructures which exist within the UK.

There was broad consensus amongst members that a multiplicity of methods would be the best approach, including the infrastructure portal and the direct open self-reporting proposals. However, a clearer definition of the aims and objectives and a considered definition of sensitive data was needed before deciding on the approach that should be taken to deliver this deeper landscape review.

Action	Owner(s)	Deadline/ status
Clearly set out the aims of the deeper landscape review before considering the methodology	Delivery team	Closed- amended scoping paper shared with PB
Clarify the definition of sensitive data in the context of DARE UK	Delivery team	Closed- will share in next meeting
In addition to conducting a survey, consider producing an open library where all infrastructures using sensitive data can self-report	Delivery team	May 27 th – to be discussed in this meeting
Refine deeper landscape review scoping document to include the above suggestion for consideration	Delivery team	May 27 th - shared with PB

4. DARE UK Public Dialogue findings

The delivery team then went on to give an overview of the insights derived from the DARE UK Public dialogue, having already shared the full report “Building a trustworthy national data research infrastructure: A UK wide public dialogue” as a paper (DAREUK_STAG-05-22-4_ Public Dialogue report).

3 key messages were highlighted:

- Participants wanted all types of researchers – including commercial organisations- to have access to their data when they are thoroughly vetted and it is in the public benefit
- Participants felt left in the dark about data research- they wanted more proactive transparency about how and why data is used; and meaningful and inclusive public involvement
- Participants were surprised at how long and hard it can be to access data and wanted processes improved so the public benefits can be realised more quickly

Following this, a brief walk through of the six recommendations that emerged from the dialogue was given:

Recommendation 1: Proactive transparency should be practised by those handling and using sensitive data for research.

Recommendation 2: Public involvement and engagement should be inclusive and meaningful.

Recommendation 3: Efforts should be made to raise awareness of security processes to protect data, and make sure those processes remain fit for purpose.

Recommendation 4: The processes and systems supporting data research across the UK should be unified in their approaches where possible.

Recommendation 5: Where feasible, processes enabling access to sensitive data for research should be standardised and centralised.

Recommendation 6: Sensitive data should be made available for research when it is in the public benefit.

STAG members were invited to respond to these findings. The work was commended with a member noting that the dialogue addressed areas that absolutely needed to be addressed. It was positively noted that the findings aligned with those of similar public dialogues on the topic. Members were pleased to see that concerns around the definition of public good had been highlighted and that participants in the public dialogue had suggested that public good should be assessed by an independent panel. However, it was argued that there is still a need for an even more granular approach.

Caution was advised with the use of the phrase public benefit “main-driver” being used to assess whether a project was in the public’s interest as the word “main” could allow for projects with ulterior, less beneficial, secondary motives to be approved.

A request was made for the workshop materials to be shared directly with the STAG so that they could be aware of the input from which participants responses were prompted.

Next, the delivery team were asked whether the focus of the public dialogue had been on health-related sensitive data or on all types of sensitive data as whilst a lot of work has been done with the public on the use of sensitive data in health, not enough has been done on other types of sensitive data. In response, the delivery team reassured members that this dialogue had focussed on all types of sensitive data and not just health.

A final point on this item was the advice that the delivery team need to have a clear strategy on how to address initial negative reactions from the public to what DARE UK is trying to achieve.

Action	Owner(s)	Deadline/ status
Share the DARE UK Public Dialogue workshop materials directly with the STAG	Delivery team	Closed

5. Mid-Sprint exemplar day reflections

Gerry Reilly shared that the Mid-Sprint exemplar day took place on April 21st and was an opportunity for the nine sprint projects, funded as part of DARE UK Phase 1, to share their progress to date.

He also extended an invite for STAG members to attend the Final Sprint Exemplar Day due to take place on 27 July 2022.

He noted that it may be useful to test some of the overall concepts explored in the projects further in subsequent phases of DARE UK but stressed that the funding of any future projects would be selected via open competition and that all current sprint projects would conclude, as planned, on August 31st.

DARE UK

David Ford raised that it would be useful to create and publish a summation of the total learnings from all of the projects.

Action	Owner(s)	Deadline/ status
Compile and publish a summative overview capturing the overall key learnings from the DARE UK sprint projects	Delivery team	October 2022

6. Persona and user journeys workshop findings

Gerry Reilly, DARE UK's Design Authority, gave an overview of the Persona and User journeys workshop which took place in May. Detailing the following:

Aim: Short, focused activity to establish the key persona and user journeys for a future pan-UKRI council digital research infrastructure for research on sensitive data.

Format: Half-day workshop at STFC Hartree Centre supported by design thinking professional(s), participants were split into 5 groups with 5 to 7 members and a mix of backgrounds in each.

Attendance: there were 30 participants (incl. delivery team of 6) from across research domains, with 11 participants from industry.

Findings:

A range of persona types were identified. In most cases, the personae are not mutually exclusive and often intersect.

Examples of key classes of personas identified were citizens, domain researchers, data custodians, data owners, skilled data practitioners and business collaborators

Next steps:

The findings will now be synthesised to identify key personas. These will be shared with PB and STAG members once available

This exercise is by no means concluded. Half a day was not enough time to exhaust the topic. However, it provided a good starting point and a varied group of people who we will be able to draw upon for further iterations.

A series of follow up workshops will need to be conducted for example with domain specific groups.

In response, a member raised a point: that the DARE UK Delivery team need to carefully consider and come up with an action plan to surface the perceived conflicts of interests between DARE UK as a convenor and HDR UK and ADR UK as delivery partners. For example, concern was expressed over the view that the persona and user journeys workshop had been conducted by a group who uses things that they have built in this space. This led the member to ask who is checking the homework? In response, Hans Erik Aronson, stated the workshop was purposely conducted by STFC colleagues and not by the DARE UK delivery team. The member noted that regardless the issue of conflicts of interest between DARE UK as commissioner and delivery partner, was something that needed to be surfaced.

It was noted that this was a concern that had been raised by some members of the research community and was particularly pertinent as this is a space which has seen substantial investment over the years with little progress. The opinion ventured was therefore that this issue of vested interests needs to be addressed. With a question being raised as to "who is checking the homework" of DARE UK? The chair responded by stating

DARE UK

that the STAG exists to provide some counter action to the potential for bias and as such, the group need to remain vigilant in their advice to the programme.

To close this item David Ford noted that it will be useful for the personas identified to be shared with the STAG once available, to allow them to advise on whether any personas have been left out or need developing further.

Action	Owner(s)	Deadline/ status
Share final personas created in the 27 th April Persona and User journeys workshop	Delivery team	Working on it- will share once completed
Consider how to manage perceived vested interests in DARE UK	Delivery team	Ongoing

AOB

Before closing, a member again emphasised the need for the issue of perceived vested interests between DARE UK and other initiatives to be considered and addressed.

Finally, the planned date for the next STAG meeting was shared: Wednesday, June 29th from 10:30-12:30. Members were asked to inform the secretariat should they not yet have received an invitation.

Hans-Erik Aronson, DARE UK Phase 1 Programme Director, closed by offering thanks to the STAG for their contributions, noting that their feedback is invaluable.

Appendix – Summary of the actions from the 17th May 2022 UKRI DARE UK Scientific and Technical Advisory Group meeting

#	Actions	Owner	Deadline/status
1	Reword minutes from March STAG minutes to better reflect the rationale given for a deeper landscape review to be conducted	Delivery team	Closed
2	Clearly set out the aims of the deeper landscape review before considering the methodology	Delivery team	Closed- amended scoping paper shared with Programme Board (PB)
3	Clarify the definition of sensitive data in the context of DARE UK	Delivery team	Closed- will share in next meeting
4	In addition to conducting a survey, consider producing an open library where all infrastructures using sensitive data can self-report	Delivery team	Closed
5	Refine deeper landscape review scoping document to include the above suggestion for consideration	Delivery team	Closed - shared with PB on May 27 th
6	Share the DARE UK Public Dialogue workshop materials directly with the STAG	Delivery team	Closed
7	Compile and publish a summative overview capturing the overall key learnings from the DARE UK sprint projects	Delivery team	Sept 31 st
8	Share final personas created in the 27 th April Persona and User journeys workshop	Delivery team	Working on it- will share once completed
9	Consider how to manage perceived vested interests in DARE UK	Delivery team	Ongoing