

Questions and answers from the UKRI Trusted and Connected Data and Analytics Research Environments - Collaborative Design and Dialogue event

28 July 2021

Most questions were answered in the webinar which you can [watch here](#). This document answers the remaining questions that were not answered in the webinar.

1. The 'public' includes a very wide range of individual values, views & opinions. How will you ensure that you involve a representative cross-section (& avoid the tendency for a vocal, confident, demographic of the 'usual suspects with axes to grind' to fill the places)? PPI in clinical medical research has had challenges with this.

We agree, having a diverse range of views and opinions is crucial when it comes to involving the public. As well as providing different ways for people to get involved so it may appeal to a wider range of people, we are continuously building new relationships with existing public networks. This enables us to reach their audiences and ensure we are working with a more diverse range of people. We will also work with those who do get involved to build a robust plan that fully supports public contributors to take on those roles.

2. Development and building of TREs are mostly driven by very large players (UK Research Councils and Industry) - how can we ensure mid and small players (Start-ups, charities, the ecosystem supported by Innovate UK) are not left out or forced to use platforms which are not optimal for their size (e.g., too expensive to implement) and be part of the ecosystem?

During this initial design and dialogue phase we will be seeking ideas from a wide range of organisations of all types and sizes. We will be looking to build a set of recommendations that support the development of a network of trusted research environments suitable for different styles of research and different types of data. It is definitely important to us that any future federated network is suitable for a diverse mix of research needs and based on open standards and so accessible to the full range of organisation sizes and capabilities.

3. Excellent initiative from UKRI/HDR! To what extent is stakeholder engagement expected during Sprint Exemplar Projects? Would potential applicants be informed on the outcome of the proposed stakeholder engagement outcomes from UKRI/HDR?

We will be openly sharing the outcomes from our stakeholder engagement as we go through this first phase. The early engagement will also be used to refine the design of the Sprint Exemplar Projects and we would expect potential applicants to also use this in refining their proposals. In the launch of the call for the Sprint Exemplar Projects, we will share more details and also provide an opportunity for questions and discussions. We also welcome informal calls with potential applicants prior to the call opening.

4. Re the funding for sprint exemplars: £100-400k is a lot to spend in 8 months with only 3 months warning. Is the typical use cloud burn or hardware? (Because something like this was run by HDR UK before...) Even if you have a project in mind, it feels too short to realistically hire staff to develop / flesh out a project.

We realise this is a short period of time. However, we want to be able to use these sprints to inform the design options for Phase 2 and therefore projects need to be completed in this timeframe. Based on past experience with the Digital Innovation Hub sprints, we saw this approach was successful. Our expectation is that the typical spend will be on staffing time and not hardware or cloud services. Given the structure of sprint projects we would anticipate that organisations would be using existing staff and therefore be able to move quickly, and our expectation would be that if a project required access to data, all approvals would already be in place.

We would be happy to have early pre-competitive calls with potential applicants to discuss project ideas.

5. Any good sprint proposal will be dependent on data access - and presumably linkage: how will you possibly resolve the access negotiations within those project timescales?

We would expect any sprint applications that require access to data to be based on existing data access agreements. We do not anticipate projects collecting data or requiring a new data access requests to be approved. If a new data access request is required, this will be expected to be negotiated prior to the submission of the application. We would also welcome applications that make use of well-defined synthetic datasets that are available to the project team.

6. It's fantastic that UKRI is embarking on this initiative. However, the programme does seem to be very focused on the technology and I would argue that the technology barriers will be overcome relatively easily. Covid research accelerated at pace due to a change in the legislation (COPI) and due to public support in using identifiable data at scale to develop new treatments and vaccines. However, this may not be the case for use of other data, for example household income, socio-economic data etc, especially if linked to health data. How will the programme approach these challenges?

Data governance and rules for data access, ethics, public engagement and use cases will be key parts of this programme, as well as technology, to form a trustworthy ecosystem. We would also welcome application for Sprint Exemplar Projects that explore novel approaches to data governance, ethics and public engagement.

7. Is there any link with the UK Data Service? Some potential users will already have taken their trusted researcher training and will have been certified.

This programme is being delivered jointly by HDR UK and ADR UK. We are working closely with Economic and Social Research Council and this will include stakeholders from the UK Data Service. In this phase of the programme, we will be looking at requirements around researcher training and certification, but any changes in approach would only be implemented in subsequent phases. We

will also be taking into account the collaboration on researcher training that has been part of the COVID-19 National Core Studies work.

8. Apologies if I am being naive, but what will the final output look like? Will it be bringing together data from TREs, if so, doesn't that go against the function of TREs? There is also a danger of us developing too many TREs.

This is genuinely a listening phase; we do not know what the final output will look like. HDR UK, ADR UK and others have been developing approaches to trusted research environments to ensure safe and secure research and to minimise data travel. This remains fundamental and as Sir Ian Diamond mentioned on the call, the intent is not to move towards large data lakes, but to support analytics across a federated network of trusted research environments.

9. Can you explain how this initiative correlates with the ONS Integrated Data Programme?

The ONS Integrated Data Programme, which is awaiting final approval by HM Treasury, is entirely complementary to the investment ADR UK is making into the ONS Secure Research Service. If funded, the Integrated Data Plan would provide enhanced technology and significant additional capacity to support internal, civil service-led analysis of linked datasets. The upgraded service will also provide a new pipeline of administrative datasets, that can be curated for external, ADR UK-funded, research use.

10. Are there plans to use the data for education purposes, as in, train the next generation of data scientists through, for instance, student projects?

The use of data for education purposes is outside the scope of this programme, however, stakeholders involved or interested in educating the next generation of data scientists may contribute to this design and dialogue phase. Whilst the focus of the phase is on sensitive data, we would expect the final infrastructure to also support access to open and synthetic datasets that may be more appropriate for training.