

# Safe Data Access Professionals Quarterly Meeting

12<sup>th</sup> June 2019

Cancer Research UK, London

## Minutes

### Attendees (in person)

Richard Welpton, *The Health Foundation* (Chair); Yannis Kotrotsios, *Cancer Research UK* (local organiser); Joshua Symons, *Imperial College*; Simon Parker, *Cancer Research UK*; Owen Maynard, *Office for National Statistics*; Anthony Macken, *Central Statistics Office Ireland*; James Dainty, *HMRC*; Cuna Ekmekcioglu, *University of Edinburgh*; Beate Lichtwardt, *UK Data Service*; Deborah Wiltshire, *UK Data Service*; Sarah Al-Adely, *University of Manchester*; Libby Bishop, *GEISIS*, Germany

### 1) Introduction

The Chair thanked everybody for attending this June meeting of the Safe Data Access Professionals, and welcomed new participants. The Chair also thanked Yannis Kotrotsios and Cancer Research UK for the kind and generous help making the local arrangements for the meeting.

### 2) Introduction from new participants

Sarah Al-Adely introduced himself and the work that she undertakes at University of Manchester. Sarah works with the Trusted Research Environment (TRE) as part of her role with the Manchester Foundation Trust. Extracts of secondary care data are created and made available for researchers. The data are centred around rheumatology, and include non-structured data from over 50 sources. The data will be housed in a data warehouse for researchers, who can submit research proposal to the team. Data will then be made available in the TRE.

### 3) Round-table updates (all services)

#### **HMRC Datalab**

- HMRC is scheduled to move from the transitional Canary Wharf site to the Stratford regional centre when it opens in late 2020. This is likely to affect the location of the Datalab.
- Data from PAYE systems are now available for researchers to analyse on high-spec PCs
- Will be running the Safe Researcher Training programme from August 2019

#### **UK Data Service Secure Lab**

- The service is looking at ways to streamline access to data application processes, and whether new applications, such as remote execution (DataShield for example) could be used
- New data arriving include clinical trial data, and data from the Whitehall Study
- New training for researchers accessing clinical trial data is being developed

### **Central Statistics Office, Ireland**

- Information on available datasets will be published on website soon, including metadata
  - This follows work by the office to create a register of datasets
- Research accreditation programme is under development, and CSO have been in touch with ONS about accreditation under the Digital Economy Act

### **Cancer Research UK – Cancer Intelligence**

- The new secure data environment is almost ready for use, and will be tested shortly
  - This will include the ability to securely whitelist web sources to help with e.g. managing R packages
- A recent change in policy on checking outputs for statistical disclosure control involves analysts reviewing the outputs they would like to have released with one of the data management staff. This appears to have led to a reduced number of issues arising when analysts request outputs to be released, as they are more aware of what is expected.

### **Edinburgh University Data Safe Haven**

- Now has a licence to enable up to 100 users to log into the environment
- Will apply for Scottish Government accreditation and the NHS Digital Data Security and Protection Toolkit following ISO27001 certification
- New projects established include one investigate AI
- Secure remote access is being considered
- Running courses on SDC and anonymisation
- Gatekeeper concept enabling researchers to check outputs created by other researchers for statistical disclosure. Discussed generally as a group
- Has adopted the SDAP SDC Handbook

### **Office for National Statistics**

- Accreditation under the Digital Economy Act is almost complete
- Output checking course is under development, and will be tested during July. Anybody interested in participating can contact Owen.
- Assessment to accompany the Safe Researcher Training will be rolled out soon

### **Update on IDAN: International Data Access Network**

- Beate Lichtwardt provided an update about work involving partner organisations from the UK, France, Germany and the Netherlands

## **GEISIS, Germany**

- Data access categories are in the process of being restructured
- Team are investigating platforms for providing secure remote access

## **Imperial College**

- Increasing computing environment capacity, but also looking to outsource
- Interest in BS1012 standard, for managing personal data; this would help with GDPR-compliance demonstration
- Joshua Symons reported that he'll soon be leaving Imperial College to take up a position at NHS Digital. The group wished Joshua all the best for the new post, and very much hopes that he'll continue to join these meetings.

## **The Health Foundation**

- Move of SDE to UK Cloud has been working without problems for the last 6 months
- ISO27001 6-month check about to be undertaken
- Gap without a dedicated SDE Manager
- UNECE Workshop on Statistical Confidentiality: two abstracts accepted:
  - SDC Handbook
  - 'Personalities and Data Access Decisions'

### 4) Other topics discussed

During the previous session, a number of topics arose which the group discussed, including:

- Charging for data access
- Data distribution
- Synthetic data
- SDC and managing expectations
- Training

### 5) Presentation by Yannis Kotrotsios on SDC, outputting checking and assessment

- Yannis demonstrated the assessment that Cancer Intelligence are asked to complete, to test their knowledge of how to apply SDC to statistical output
- Materials were taken from the SDAP SDC Handbook, and the assessment created on Microsoft Forms.
- There was general excitement about the assessment, and Yannis was asked to circulate a link to the assessment around the group

#### 6) SDC Handbook

- Richard updated the group by stating that a designer has been commissioned to professionally design the SDC. More information about this, and a workshop in October/November, to be circulated soon.

#### 7) Discussion on safe data professionals and career development

- The group agreed that Data Management needs to have more resources invested in staff
- HDR\_UK may be looking into data management training for secure data
- Travel and time can be a factor; short online modular courses could be useful
- Organisations such as CESSDA have some data management training available
- Online courses could be produced

#### 8) Discussion on secure archiving

- The group discussed this topic further:
  - Archiving and syntax used to create data would be a useful beginning
  - It is important to understand the concerns of data controllers, and there are many benefits to them which they might not realise
  - Curation aspects to be considered
- The group will continue working on this with a view of drafting a proposal during the summer

#### 9) Meeting closed

The meeting was closed at 16:00, and the Chair thanked everybody for their attendance and contributions.