Commissioning a cloud secure data environment at CRUK

Simon Parker – Data Liaison Manager

CANCER RESEARCH UK

Together we will beat cancer

The first steps…

CANCER RESEARCH UK

Background

- Cancer Intelligence team works with patient-level data. This data is usually provided by Public Health England.
- In 2017, we launched a secure data environment built by our Technology team. This is a remote desktop that can be accessed by approved analysts, and provides access through a firewall where the data is stored.
- However, the environment is not computationally powerful enough for our needs, and maintenance is a low priority for the Tech team.
- We developed a business case, and began a tendering process in January 2019.

Our Security Model

- In Cancer Intelligence we use the 5 Safes Model as a framework for our data governance work.

  Safe Projects
  + Safe People
  + Safe Setting
  + Safe Outputs
  + Safe Data
  = Safe Use

- Links
  - Training: https://securedatagroup.org/training/
  - SDC: https://figshare.com/articles/SDC_Handbook/9958520

Requirements

- Our new secure data environment had to help us achieve Safe Use.
  - Safe Setting
    - Users shouldn’t be able to take files in or out.
    - Restricted Internet access.
    - 2-factor authentication.
  - Safe Outputs
    - A pipeline for administrators to transfer files out of the environment.
- Expandable and scalable.
- Possibility to have a HSCN connection.
- Supplier had to have relevant information security certifications.

Selecting, designing, and building…

CANCER RESEARCH UK

Selecting a provider

- A number of teams across CRUK were involved in deciding which supplier to go with.
  - Procurement – focussing on the cost of the contract.
  - Legal – to review the contracts.
  - Information Security – to check the accreditations were in place.

- We also had a large input in the process as we were recognised as the experts in this area within the charity.

- We felt that the bid from AIMES met our requirements at the price we had budgeted.

Designing the system

- We had a requirements gathering meeting at the beginning of April with AIMES where we discussed how we would like to use the environment.
  - This was more detailed than the tendering documents, with a particular focus on dataflows and system specifications.

- The design stage was an iterative process lasting around 4 weeks.
  - Selecting the software required.
  - Designing the folder structures and access permissions.
  - Deciding upon network access.

- We had to sign-off on the proposed design before production of the environment began.

- We had weekly meetings with AIMES during this period as well as having access to platforms such as IT Glue to share documents.

Building the Secure Data Environment

- The initial build took approximately 2 weeks, at which point we began testing.
- The initial testing was conducted by the administrators. We continued to have weekly meetings with AIMES during this time as fine-tuned the setup.
- The fine-tuning phase last around 3 weeks. We then began formal User Acceptance Testing.
- We continued to test the environment extensively before we gave a final sign-off and went ‘live’.
- In total the design and building phase lasted 10 weeks.

My first time in Liverpool…

CANCER RESEARCH UK

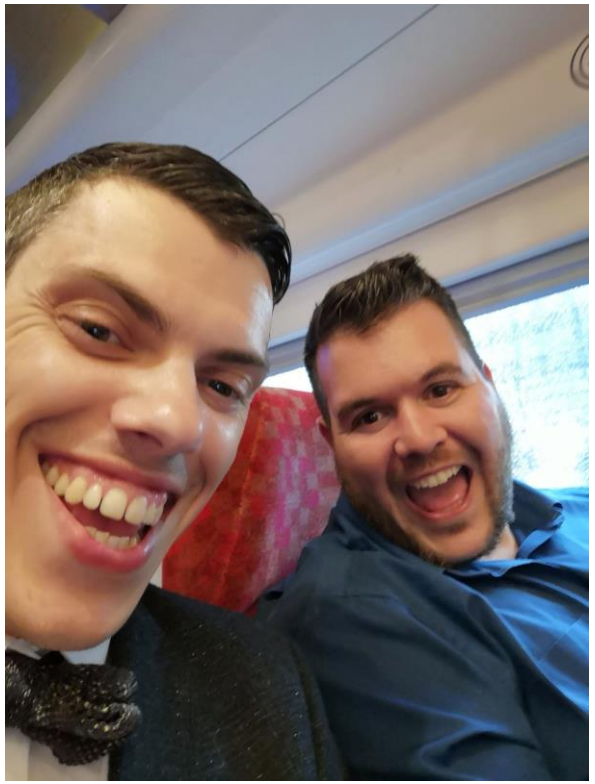
Speaking to patients

- We thought that patients might be concerned about us allowing patient-level data to be stored outside of CRUK.

- We presented the proposed project to our Patient Data Panel to get their feedback about the proposal and to discuss any concerns that they had.

- The Panel were very supportive of the proposal and encouraged us to be ambitious and to use the environment to support research across CRUK.

Visiting the data centre

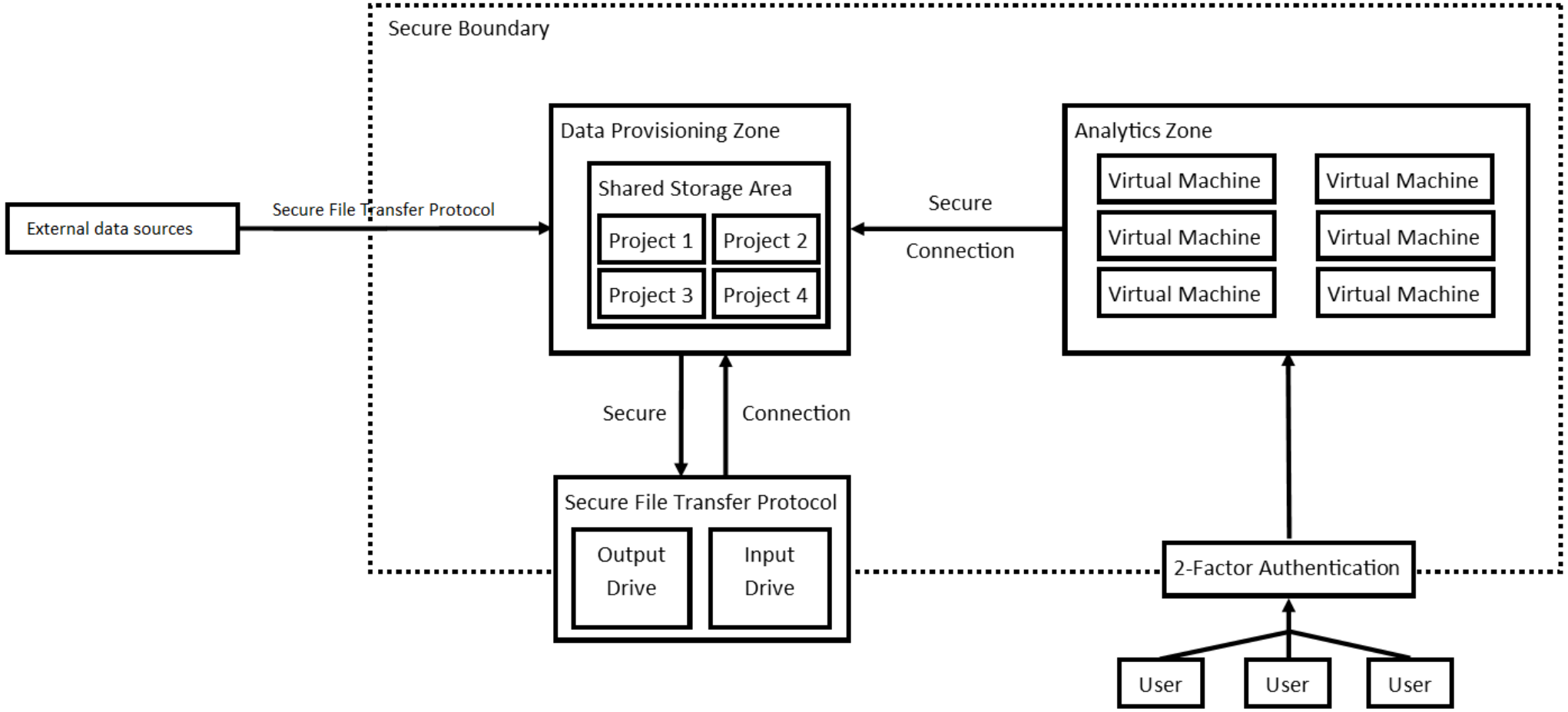


- The Patient Data Panel were pleased that our Information Security team had been part of the due diligence process. They suggested that we also visited the site to see where the data would be stored, so that we could see first-hand the security in place.
- A few members of the Patient Data Panel came with us.
- We were rather excited to go.

Using the Secure Data Environment

CANCER RESEARCH UK

Plan of the Secure Data Environment



Using the Secure Data Environment

- Each project has its own project folder. We use security groups to manage who can access which project.
- Users are able to access R packages from a specific CRAN repository, no other Internet access is enabled.
- Maintenance of the Secure Data Environment is the responsibility of AIMES.
- We are the administrators and must approve adding/removing users, changing permissions, expanding the provision.

Data Security and Protection Toolkit

- The use of data through our Secure Data Environment is covered by our DSPT accreditation.

- Section 10 relates to Suppliers.
  - Evidence that adequate due diligence has taken place for any supplier that will be handling personal information.
  - Evidence that the supplier has equivalent accreditation.
  - Supplier’s business continuity plan.

- We keep logs of the use of the Secure Data Environment for auditing purposes. These include a record of log in attempts and receipts for the transfer of files through the SFTP.

Thank you for listening

simon.parker@cancer.org.uk

CANCER RESEARCH UK

Together we will beat cancer