Science and Technology Facilities Council

Scientific Computing

Science and Technology Facilities Council

Ada Lovelace Centre

# Catering to the bespoke data analysis needs of scientific facilities using the Ada platform

As scientific experiments have grown larger and more complicated, so has the kind of data that they generate, and the need to store and manage this data has become increasingly more complex.

The DAaaS (Data Analysis as a Service) team is tackling this issue through its new 'Ada' platform, named for the famous Ada Lovelace who wrote the world's first computer program. The Ada platform allows scientists at STFC's and other large-scale facilities, such as the Diamond Light Source's ePSIC (electron Physical Science Imaging Centre), to access vital computing hardware and software through the STFC Cloud, using an easy and convenient user-friendly interface. The Ada platform forms part of a wider Ada Lovelace Centre (ALC) programme to solve bottlenecks to delivering science at STFC's large facilities.

## The Challenge

STFC is home to a variety of facilities, each with its own unique data analysis needs, from electron microscopy software to the ability to process vast amounts of satellite data from RAL Space.

**"Analysing data is a collaborative effort between STFC facilities, users of the facilities and other collaborators. To perform this analysis, you need time to analyse the data and provide the flexibility for everyone involved to access the data, share what they're doing, and access computing resources for calculations."**

***Jonah Loughlin**, Graduate Student, STFC Scientific Computing.*

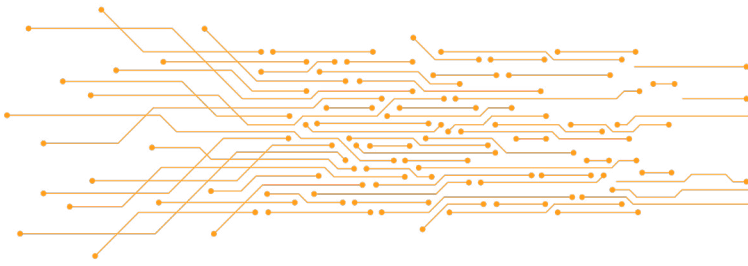
Accessing data in collaborations can often hit challenges with sharing, arranging access and permissions to different systems. The challenge is to make the data easily accessible. While we want to share data with our collaborators, privacy and ownership of scientific experiments and their results also need to be considered and are of the utmost importance within research. We take care to address the threats of cyberattacks and malware to mitigate the risk of data, including unfinished research data, being leaked, as this can pose huge risks for the cybersecurity and reputations of organisations.

The Ada platform provides a self-service style of functionality, through which scientists can access bespoke workspaces with preinstalled software. This provides them with a remote desktop environment that runs very much like their own laptop.

## Our Approach

By maximising the potential of the STFC Cloud, Ada matches pre-existing compute resources to the needs of the user, all from the convenience of a web browser.

Scientists can upload data straight from their instruments into specialised cloud storage, minimising the time between data collection and data analysis. This storage is what enables Diamond users to access their data on the Ada platform securely, and allows us to create collaborative environments, paving the way for further research groups and users to analyse our facilities' experiment data in the future.



**“The cloud storage allows large amounts of data to be stored within a centralised location that the Ada platform can then access, and control access to. The data can be stored, along with processed results and other information, allowing greater flexibility in how scientists approach analysing their data. There is also a possibility to review and re-analyse from different perspectives on high performance and high-spec computers, long after the experiment has closed.”**

***Jonah Loughlin**, Graduate Student, STFC Scientific Computing.*

**The Ada Lovelace Centre (ALC) is a key programme within STFC Scientific Computing, set up in recognition of the critical role of scientific software and data analysis in realising scientific impact from experiments.**

This project was undertaken by graduate software engineer Jonah Loughlin, with specific input relating to the data sharing functionality from fellow Scientific Computing graduate Aaron Larkins, who works on the DAFNI platform (Data and Analytics Facility for National Infrastructure).

## The Benefits

Keeping up with this ever-changing landscape is a complex issue with business-critical implications: if data cannot be analysed, science cannot be performed. As experimental research pushes the science to new frontiers, STFC and its associated research communities require well-designed technological solutions that can make otherwise inaccessible compute resources available to scientists with ease. Ada operates with a fully-realised appreciation of this, embedded into its core design.

The evolution of Ada as a platform is indicative of the ALC programme capacity and drive to meet challenges accessing data and compute. It is also facilitating new ways of working to meet the data analysis needs of scientists across the organisation and users of the national facilities.

Diamond Light Source, Harwell

**If you would like to find out how the DAaaS team can help with your research data, please contact supportanalysis@stfc.ac.uk.**

Written by Cora Jamieson and Jonah Loughlin.
Designed by Cora Jamieson, Communications and Impact, STFC Scientific Computing.