

# Computing and Data

## Air Quality and STFC technology

STFC's computing capabilities have grown over the years to support our need to analyse large and complex datasets arising from our core science programmes and facilities. For example the Large Hadron Collider at CERN, in which we are a partner, generates tens of petabytes of data every year. These must be analysed to produce new breakthroughs in physics, such as the discovery of the Higgs Boson. STFC's computing capabilities across both the STFC Scientific Computing Department and the Hartree Centre can provide any type of organisation with the expertise and support they need to develop software, tame data mountains, discern patterns in complex datasets and harness computing to improve forecasting.

■ STFC's Scientific Computing Department has world-leading experts in computational chemistry, computational engineering, software engineering, petascale storage and many other disciplines. They support some of the UK's most advanced scientific facilities.

■ The Hartree Centre work with partners on collaborative research and development projects, either on a commercial basis, or as a partner in a funded research project, addressing a challenge within a business and delivering a bespoke digital solution. The team can also help you to apply for both STFC-based and external sources of funding to carry out collaborative projects.

■ The Centre for Environmental Data Analysis (CEDA) Archive is a long-term data archive primarily for the Earth observation and atmospheric science communities. The CEDA Archive, which hosts over 15 petabytes of data, is co-located with the JASMIN computing facility at the STFC Rutherford Appleton Laboratory, and both are run by STFC on behalf of the NERC. The partnership enables scientists to explore large environmental data collections in order to solve the most pressing environmental issues we face.

■ JASMIN consists of multi-Petabyte fast storage co-located with data analysis computing facilities. It provides a range of computing services (batch, interactive, community cloud) and supports a variety of data types in a scalable environment, as scientists bring their data to JASMIN. Users can access the Met Office NAME atmospheric dispersion model as a supported service, without needing to install software or transfer data locally. Other datasets relevant to air quality include the high-volume Sentinel 5 data, and in-situ measurements from the stations and mobile instruments of the NCAS Atmospheric Measurement and Observation Facility (AMOF) and the Facility for Airborne Atmospheric Measurements (FAAM).



### Case study: Data and Analytics for National Infrastructure (DAFNI)

There is an ever-increasing awareness that human actions have significant impacts on our planet, and so major infrastructure projects cannot be planned without also considering the implications for the environment. As populations increase, the demand for new and improved infrastructure - housing, roads, electricity, clean water etc. – rises with them. More than ever there is a need to analyse impacts on the environment associated with new infrastructure, including effects on air quality.

Computer modelling is a vital element of infrastructure design. However, at present it is extremely difficult to consider all aspects of infrastructure and how they impact on one another at scale. Until now there has been no central resource available for researchers to share knowledge and analysis expertise to build new analytics tools.

The DAFNI team (Data and Analytics for National Infrastructure), part of STFC's Scientific Computing Department, is bringing together researchers, organisations and authorities to enable a smarter, collaborative approach to data analysis for national infrastructure. Funded by the UK

Collaboratorium for Research on Infrastructure and Cities (UKCRIC), DAFNI is doing this by building a common platform for data storage and analysis. Organisations will be able to lodge their data on a secure platform, upload their models and link to a wide range of other information. The approach will enable large scale analysis and, crucially, provides the visualisation tools so that results can be shared in a meaningful way.

DAFNI is working with the Urban Observatories run by researchers in six major UK cities. The observatories collect a wide variety of city and environment-related data from sensors placed in strategic locations monitoring local air quality, thermal and visual imaging, weather conditions, traffic flows and energy use to build digital images of the cities. Computational models can then be used to investigate resilience, potential improvements and the societal and environmental impacts of any changes in infrastructure.

#### Access

Access to the CEDA archive and JASMIN is free at the point of use for academic researchers working in environmental science areas, and there is training available for those that need it. STFC's Scientific Computing Department and The Hartree Centre collaborate with organisations on a project-by-project basis and funding is usually required.



#### Key contacts

**CEDA** – Wendy.Garland@stfc.ac.uk

**JASMIN** – Poppy.Townsend@stfc.ac.uk

**Scientific Computing** - Marion.OSullivan@stfc.ac.uk

**The Hartree Centre** - hartree@stfc.ac.uk