Science and Technology Facilities Council

Supporting Infrastructure Research Communities on DAFNI

Kyle Stevenson (on behalf of the DAFNI group)
UKRI (STFC)
20th March 2023

Abstract

DAFNI (a Data and Analytics Facility for National Infrastructure) is a major national facility, based at the UK Science and Technology Facilities Council (STFC), which supports cross-disciplinary collaborative research in infrastructure engineering and environmental science. DAFNI holds users' data and allows them to run models together in a system-of-systems approach. Thus, the platform provides a common place for data and models to be shared and used together, to explore new scenarios without needing direct access to the code, and then kept as a legacy space, which can be accessed by stakeholders after the end of research projects. This allows DAFNI to support programs that aim to solve complex problems in national infrastructure, and to support and enhance decision making.

The DAFNI platform is currently supporting a diverse array of infrastructure projects, including projects aimed at evaluating the impact of climate change on transport and energy networks, and informing how best to mitigate these impacts. We are also leading a drive to implement digital twins, a concept which will combine real time data collection and infrastructure modelling, in order to inform policy, and provide feedback into decision making loops in a scientifically informed manner. The preservation of historical data and models is critical to this process and is central to the DAFNI project.

In this paper we will discuss the infrastructure underpinning DAFNI, and the results from ongoing work. We will present work from several projects, including, but not limited to**,** the Climate Resilience Demonstrator (CReDo), emissions modelling for global shipping (a UCL based project), and OpenCLIM (The Open Climate Impact Framework). After reviewing our current work, we will cover future challenges and objectives for the DAFNI project.

Introduction

National infrastructure is of vital importance to the economic and social welfare of a country, and its interplay with the environment is increasingly of interest. DAFNI (Matthews, Hall, & al., 2023) is a national platform for supporting infrastructure research in the UK, with a focus on the security and resilience of systems. DAFNI is also concerned with developing effective responses to threats – such as climate change – to society and the economy.

DAFNI provides a shared space for infrastructure researchers, so distributed teams can work collaboratively on developing models and sharing data. Support is given for models written in a variety of languages and different operating systems, and we also support an approach where models can be used flexibly together within workflows. Maintaining a legacy environment is considered to be of critical importance to DAFNI, and we are aiming to provide a service where the provenance of data and models is considered central to operations and persists beyond the original lifetime of individual projects.

Workflows on DAFNI are run on a dedicated HTC cluster which, currently, is comprised of 792 CPUs, with a working memory of 16.8 TB. Approximately 1.3 PB of mixed type data storage is available to DAFNI users. The cluster is managed using Kubernetes, which orchestrates cluster management, scaling, and the deployment of software. Kubernetes has full native support for Docker, which is used to containerise models in DAFNI. This containerisation allows a diverse array of models to be uploaded and run on DAFNI.

The Architecture and Design of DAFNI

The DAFNI architecture is designed to allow models, often developed in different computing environments, using a variety of computing languages, to be run seamlessly within DAFNI. Models are often developed by different teams, and differences in library dependencies, languages and operating systems can form a significant impediment to creating multi-model systems – DAFNI intrinsically avoids these issues through the adoption of containerisation. Models are typically dockerised and provided to DAFNI by users, so they can be integrated into workflows that will run on DAFNI.

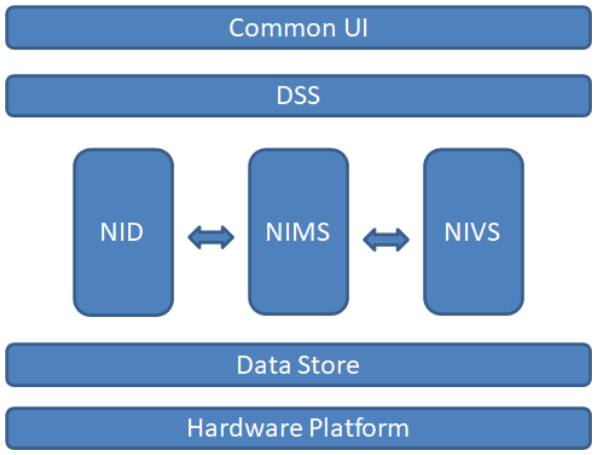

Figure 1. The main components of the DAFNI architecture.

In order to achieve this the DAFNI architecture is split into three main components - the NID (National Infrastructure Database), NIMS (National Infrastructure Modelling Service), and NIVS (National Infrastructure Visualisation Service), which are all accessible to users using a common web interface.

The DSS (DAFNI Security Service) manages the security of the platform giving users access to the data and models they have the correct permissions for. DAFNI will potentially hold sensitive data and models, so it is important that the security and integrity of data and models is maintained, while still allowing seamless access.

NID – National Infrastructure Database

The NID provides a one stop service point for data within the DAFNI framework. It covers all aspects of interest to users including a full data catalogue, flexible access to data and metadata, and a fully integrated data publication service. The NID is one of the fundamental building blocks of DAFNI, managing the uploading and publication of data.

NIMS – National Infrastructure Modelling Service

The NIMS provides a suite of accessible models and facilitates the development of multi-systems models. Uploading of models is language and OS agnostic within the DAFNI NIMS, as containerisation is fully supported and managed with docker. The NIMS has full support for searchable user defined metadata.

A model workflow system is integrated into the NIMS, and a corresponding searchable catalogue is also maintained for workflows. The workflow framework supports the use of multiple models, allowing users to build sophisticated system-of-systems simulations, and models can be run iteratively within workflows, with both parallel and sequential loops supported.

NIVS – National Infrastructure Visualisation Service

It is important to be able to both visualise and analyse output data from workflows and the NIVS addresses this need, providing basic visualisation support for users. Traditional graphing tools are provided by the NIVS, as well as support for Jupyter which provides a powerful tool for the analysis and visualisation of data in user developed notebooks.

Supporting Climate Change Response and Resilience Measures

One of DAFNI's core goals is to support the UK national effort on climate change, and to both host and support projects which aim to answer key questions about the resilience of the UK's infrastructure and support future infrastructure planning. Several projects are currently hosted on DAFNI, two of which we detail in this section (both these projects are focused on climate resilience).

CReDo – Climate Resilience Demonstrator

Currently concentrating on flood resilience, CReDo is a ground breaking project (CReDo, 2022) that aims to demonstrate how we can use modelling of systems to improve decision making, and mitigate the impact of climate change on national infrastructure sites.

In the first phase of CReDo a demonstrator was setup, using synthetic data, to show the potential of CReDo to inform and guide policy makers (the synthetic data was representative of a real network of infrastructure sites). This work was extended in phase 2 of the project which focussed on setting up secure data storage within DAFNI for clients – UK Power, Anglian Water and BT – which can be easily accessed, while remaining secure. Work on providing evidence-based guidance for decision making, with regards to flood mitigation, was also carried out during this phase of the project.

While the CReDo project is currently focussed on the impact and mitigation of flooding, it is expected that the project will expand to cover other aspects of climate change, such as changes in regional temperatures, and corresponding droughts.

OpenCLIM – Open Climate Impact Framework

The OpenCLIM project adopts a broad-spectrum approach aimed at accurately assessing the impact of climate change on the UK (Kennedy-Asser, Owen, Griffith, & al., 2022) and potential adaption measures (Jenkins, Ford, & Robson, 2022). This is achieved by linking physically based models in order to evaluate the effects of climate change on agriculture, temperatures, potential for flooding, and the interplay of these effects with future expansion in urban areas and infrastructure. DAFNI currently hosts a large suite of OpenCLIM models, data and workflows, which were used in this project.

OpenCLIM is currently considering two scenarios, which are considered representative of areas that will be considered in future studies. The first is an urban area – Glasgow and the Clyde region, while the second area will cover the Norfolk Broads and local environs.

Transport Infrastructure

Another core area of interest to the DAFNI project is supporting work on the modelling of transportation systems, investigating their efficiency, and their impact on the environment. How transportation systems can be improved in the future, and better adapted to serve society's needs is of great interest to DAFNI.

UCL Shipping Project

One of the most critical issues we face at the moment is the impact of global shipping on carbon emissions. In a business-as-usual scenario (Vandycke & Englert, 2017) it is expected that over the next 30 years emissions will increase anywhere between 50 and 250%, where shipping currently is responsible for 2-3% of total $C0^2$ emissions. The UCL shipping project aims to demonstrate that it may be possible, even within the limits of current technology, to significantly reduce emissions within the shipping industry.

Current shipping procedures have resulted in a practice known as SFTW (Steam fast, then wait) being incentivised, where cargo ships are encouraged to make best speed for a destination port, so as to declare a NOR (Notice of Readiness) in order to best satisfy maritime contract law, and also to be in a position to claim demurrage when such clauses exist in the contract. It is estimated that emissions could be reduced by 20-25% (as much as 100 million tons of CO^2 emissions) if this practice was modified and shipping moved towards a just-in-time model for port arrivals. This in turn would then allow for more fuel-efficient travel speeds.

The pilot UCL shipping project (Manola, Varga, Owen, & Stevenson, 2023) uses AIS data for selected cargo ships, and then estimates potential fuel savings that could be made by adopting a lower, more efficient transit speed. This information can then be relayed using the ships communication system (typically ships will have access to the internet via INMARSAT, covering an area that lies approximately within 70 degrees of the equator). The UCL Shipping model has been implemented in DAFNI and is currently running on a 6-hr cycle, writing data and results to a PostgreSQL database. A prototype communication system, which sends results via e-mail, has been setup to form the initial steps necessary for a digital twin.

Sheffield Traffic Project

The Sheffield Traffic Project (Genes, 2021) is a pilot study which combined the use of traffic sensors (which monitor traffic flows within the urban area of Sheffield, UK), with AI based traffic models in order to predict future traffic flows. The monitoring data was based on 640 traffic sensors which are deployed at strategic positions in

Sheffield as part of the Sheffield Urban Observatory. The sensors report the traffic flow every five minutes and this data was subsequently accessed through an API call to the Urban Observatory servers. A machine learning regression model was trained to estimate the correlation between historical traffic data and recent data in order to predict future traffic flows. The model was setup on DAFNI to run using both real-time and historical data, and a visualisation was setup showing predicted traffic flows on a map in real-time.

Digital Twins and DAFNI

When using models to both describe the real world and to predict future behaviour, we must always be aware of the limitations of the particular models used. Models, and conglomerates of models, by their nature, are typically simplified representations of what are often highly complex systems in real life, and though they are often extremely useful tools they are by no means fully accurate. Therefore, it is very important to understand the limitations of our models, and this means comparing our computer models to the real world, which is at the core of what a digital twin is. The UCL Shipping Model and the Sheffield Traffic Study – which were both described in the previous section, can – as such - both be considered as pilot digital twins.

Conclusion

DAFNI is a fully realised facility able to support a myriad of multi-systems models, using a wide spectrum of data types, languages and computing platforms. Intrinsic support exists for collaborative sharing of models and data, providing for a distributed community of researchers and stakeholders. It is recognised that while large scale computing resources are invaluable in solving research and development challenges, these resources can be difficult to access and utilise for non-experts. DAFNI provides a user environment that aims to overcome these challenges and provide intuitive access. Containerisation has been highly effective in supporting DAFNI's goals, and we are currently looking at ways to improve, automate and further streamline this process to make it easier for users to start working within the DAFNI framework.

DAFNI has matured, since its inception in 2017, from a development project to a fully realised service platform for users, which can be flexibly scaled in the future. This enables DAFNI to grow in the future in order to meet operational needs and, is considered key to supporting future infrastructure goals in the UK, working together with stakeholders in government, academia and industry.

Acknowledgements

Dafni has been supported by the EPSRC Grants EP/R012202/1 *UKCRIC National Infrastructure Database, Modelling, Simulation and Visualisation Facilities* (2017-21) and EP/V054082/1 *Dafni-Research Only Strategic Equipment (Dafni-ROSE) (2021-23)*. The author would like to thank the members of the Dafni support and

development team in STFC, and the collaborators and champions involved in specifying requirements and developing pilot cases within the Dafni environment.

References

CReDo. (2022). CReDo Project. *https://digitaltwinhub.co.uk/credo/reports-resources*.

Genes, C. (2021). DAFNI as a Digital Twin Platform. *https://www.dafni.ac.uk/dafni-as-dt*.

Jenkins, K., Ford, A., & Robson, C. (2022). Identifying Adaption 'on the ground': Development of a UK adaption Inventory. *Climate Risk Management 36*, 100430.

Kennedy-Asser, A. T., Owen, G., Griffith, G. J., & al., e. (2022). Projected Risks Associated with Heat Stress in the UK Climate Projections (UKCP18). *Environmental Research Letters, 17* , 034024.

Manola, E., Varga, L., Owen, J., & Stevenson, K. (2023). Maritime Shipping Digital Twin toward Decarbonisation of Concrete. *Manuscript in preparation*.

Matthews, B., Hall, J., & al., e. (2023). DAFNI: A Computational Platform to Support Infrstructure Systems Research. *Manuscript submitted for publication*.

Vandycke, N., & Englert, N. (2017). Three Reasons why Maritime Transport must act on Climate Change. *World Bank*.