



# JASMINx expansion project: use case report findings

Jisc on behalf of UKRI/NERC February and March 2022

Victoria Moody, Jisc

# JASMINx expansion project: use case consultancy

## Background:

- [JASMIN](#) is a large data-intensive computing facility built to support the [NERC](#) environmental science community designed for potential expansion to service other communities within UKRI.
- The [Centre for Environmental Data Analysis](#) (CEDA) was established in 2005 to incorporate the British Atmospheric Data Centre and the NERC Earth Observation Data Centre.
- Since April 2018, CEDA has been a component part of the NERC Environmental Data Service, which brings together the five NERC data centres into a single service commissioned by NERC as a National Capability.
- JASMIN provides the infrastructure upon which the CEDA data, archives and services are delivered.
- Designed, built, and managed by STFC's Scientific Computing Department (SCD) for NERC, JASMIN is part supercomputer and part datacentre and provides a globally unique computational environment.

# JASMINx expansion project: use case consultancy

## Background:

- JASMIN increasingly provides flexible data analysis capabilities to a growing community, which benefits from high performance compute and a private cloud, co-located with peta-scale data storage.
- Environmental data are unique because the exact environment from which the data were derived cannot be reproduced. Environmental data will continue to grow, and projected demand for storage continue to increase.
- The CEDA Archive provides access to thousands of atmospheric, climate change, and earth observation data, accessible via a file system from the shared science machines on JASMIN.
- CEDA is a trusted repository under the Core Trust Seal. Priority is given to data generated as a result of funding by NERC however, data generated through other funding sources is considered for deposit at the discretion of the CEDA Archive.

# JASMINx expansion project: use case consultancy

## Project background:

- A new programme of work **JASMINx**, has been proposed to take the existing system through the next decade. This programme was costed to deliver against the known requirements of the environmental science community **but also designed for potential expansion to service other communities within UKRI.**
- The **next step for a wider JASMIN remit**, under the auspices of JASMINx is to examine the appetite in the community and amongst funders.
- **UKRI/NERC** contracted Jisc in February 2022 to support this evaluation project.

# JASMINx expansion project: use case consultancy

## Objectives and scope:

- Explore the potential for offering an expansion of the JASMIN platform to additional individuals, groups and communities who wish to share and analyse data from a wider range of disciplines.
- Enable JASMIN to understand how the infrastructure could be adapted to meet the requirements of researchers from a wider range of research domains, including for environments that support trusted research.

## In scope

- The JASMIN platform technology, data, user base, outputs and processes and exploration of the potential for adaptation for support of a range of additional disciplinary use cases.
- Other infrastructures' platform technology, data, user base, outputs and processes and exploration of the potential for JASMIN adapting to meet these use cases.

## Out of scope:

- Exploration of other infrastructure investments' potential for adaptation for support of a range of additional or multi-disciplinary use cases.

# JASMINx expansion project: use case consultancy

## The project considered:

- **Research Needs:** What requirements flow from research needs to infrastructure? Who is involved?
- **Infrastructure Requirements:** How should those needs be delivered?
- **Support Requirements:** How are users supported, by whom?
- **Finance and Governance:** What models for ownership and governance might best suit a diverse constituency?

# JASMINx expansion project: use case consultancy

24 representatives from among the following five user groups:

- Key individuals involved in **scoping and defining a joined up UKRI digital research infrastructure**
- UKRI staff **responsible for data-intensive programmes** within their research council
- **Users** of existing data intensive facilities.
- **Providers** of existing data intensive facilities
- **Research staff in universities** and UKRI research centres with data-intensive computational requirements

# JASMINx expansion project: use case consultancy

## Overall response:

- Significant enthusiasm
- A high level of need for infrastructure across the disciplinary areas consulted
- Clear gaps in provision, clear potential for meeting objectives of JASMINx
- A strong willingness to work in partnership across funders research collaborations and sector agencies

# JASMINx expansion project: use case consultancy

## Overall response:

Where they had knowledge of JASMIN interviewees felt it was a good model with strong cross-disciplinary possibilities.

Many were also eager to see the results of this work and willing to be involved in future activity to assess the potential for expanded provision of research infrastructure.

# JASMINx expansion project: use case consultancy

## Focus of interview questions:

- The capacity to build and run own infrastructure, procurement of components and tools, platform and software, connectivity and compute capacity, transfer of data
- Diversity of disciplines and interdisciplinary approaches, associated complexity,
- Potential to address duplication, opportunities for efficiency and economies of scale
- Governance requirements, collaboration protocols, access and security requirements
- Data management and linkage, archiving, storage and preservation, volumes and complexity of data collected, generated or expected
- Licensing, IP and commercialisation needs
- Environmental considerations including strategy and targets in carbon footprint management and net zero
- Funder requirements in terms of interdisciplinary research and data from other research domains.

# JASMINx expansion project: use case consultancy

## Headline findings:

- New approaches to data intensive research are being developed across a range of research domains, with increasingly complex workflows.
- Decisions by researchers about which facility to use or infrastructure to implement were not always based on a full understanding of the best fit.
- There is a lack of comprehensive information available to researchers about the options that are available and most appropriate for them.
- There may be potential for assessment of extending an existing infrastructure such as JASMIN to mitigate against the potential for new technical infrastructure duplication or proliferation.
- Extending or expanding one investment to cover the full range of possible research use cases would not be feasible.
- Multi-disciplinary research increasingly requires the creation of and access to linked data with a concomitant risk of disclosure of personally, commercially or otherwise sensitive data.

# JASMINx expansion project: use case consultancy

## Headline findings:

- Use of more powerful compute facilities for increasingly complex problems may need new approaches to and skills in data governance, management, and support, and to extend computational capabilities.
- The development of modular approaches to governance, management and support across a more interconnected research data infrastructure, taking account of disciplinary distinctiveness where necessary, has significant potential.
- There is an opportunity to test and iterate the scope for deploying elements of the JASMIN technical, archival and cataloguing infrastructure across specific interdisciplinary research domain use cases to optimise understanding of the potential.
- Obtaining a comprehensive understanding of cases where other research domains using or oriented towards environmental data could potentially utilise elements of other facilities, for example JASMIN, would be of benefit.

# JASMINx expansion project: use case consultancy

## Recommendations across the following areas:

- Undertake awareness-raising of JASMIN with other disciplines.
- Expand the evidence base for enquiry in the form of a wide-ranging survey.
- Understand the potential routes to utility in real research settings, in terms of end-to-end research lifecycle requirements.
  - Define the technical characteristics of JASMIN as discipline agnostic for others to map their research workflows to, to test and iterate the scope for technical deployment utility across wider disciplinary and cross-disciplinary research domains.
- Enable understanding of potential alignment with disciplinary and cross-disciplinary research domains that are particularly suited to JASMIN scale and capability.
  - Consider a focus on disciplines that do or will collaborate, or have data linkage, with NERC disciplinary areas; broadly defined and appropriate to JASMIN scale and capacity specifically, rather than aiming to serve all potential use cases.

# JASMINx expansion project: use case consultancy

## Recommendations across the following areas:

- Focus on skills needs and opportunities from expansion of technical facilities to multi-disciplinary research and additional research domains.
- Explore modular approaches to governance, management and support which may be applied as technology agnostic towards a range of infrastructure technology settings, including JASMIN.
- Understand what is envisaged by and the subsequent potential for, 'federated' approaches in the context of the JASMIN technology and data architecture/archival model for other disciplines.
- Develop joint pilot projects to develop roadmaps and pathways to ascertain the financial, technical, governance, security, management, methodological, impact and research integrity benefit and an understanding of any risks.
- Develop an understanding of potential costs incurred relative to potential savings.
- Engage the full range of stakeholders including research councils and data infrastructure investments as well as researchers and sector bodies such as Jisc to understand, define and test the potential across the research ecosystem.

# JASMINx expansion project: use case consultancy

## Potential pilot projects for further exploratory activity:

1. Extend the research domains consulted in the form of wide-ranging survey and interviews with a larger representative group of research domains.
2. Engage an emerging, related / cross-disciplinary investment to directly test the potential for the JASMINx programme at the infrastructure development phase (an example would be the NERC Digital Solutions Programme led by the University of Manchester: <http://www.digital-solutions.uk/>.)
3. Engage the AHRC and its community to assess the potential for JASMINx to support the scoping of an AHRC research compute capability taking account of the media diversity and dimensional complexity of data in arts and humanities workflows.
4. Engage the ESRC funded HDRUK and ADRUK and UK Data Service programmes along with ONS and Research Data Scotland to understand technical, governance, management and support needs and opportunities in terms of trusted research environments.
5. Engage wider UKRI to develop a focus on technical innovation in modularity of provision for research governance, management and support which may be applied as technology agnostic in a range of infrastructure settings to support a wide range of research domains.



**Thank you**