FAIR@HPC - Improving HPC usage in ESS by FAIR data and compute services (2024/25 update)

Interest Group (IG) "High-performance Computing in Earth System Sciences" of the National Research Data Infrastructure (NFDI) Consortium Earth System Sciences (NFDI4Earth)

Jan 2025 Status Report and Addendum to the Concept Paper of May 2022
– reflecting the status of recommendations, as for fulfilment or adjustments, by the NFDI4Earth Plenary May 2024 –

1 Introduction

In our concept paper[1] from May 2022, we addressed the continuously increasing demand for using High-Performance-Computing (HPC) infrastructure in the Earth System Sciences (ESS), and the challenges arising from this for the NFDI.

From a description of typical use cases, we derived challenges in FAIR[2] (Findable, Accessible, Interoperable, Reusable) Research Data Management (RDM) and computational workflow/data-flow management, as a motivation for driving NFDI4Earth developments on the HPC-centric side. We issued recommendations that relate to these possible developments, and scheduled a recommendation status review after 12-24 months.

This addendum discusses current and future topics with increasing significance for RDM in HPC in the ESS – together with a set of new recommendations related to these topics (Section 2). It then reviews the fulfilment status of our previous recommendations and lays out necessary adjustments (Section 3).

2 New Challenges and Recommendations

New challenges in our field are constantly arising from the changes in the HPC, FAIR RDM and Big Data landscape related to ESS activities. High Performance Computing (HPC) has traditionally played a critical role in ESS by enabling researchers to process vast amounts of data and perform complex simulations. Nowadays, HPC centres and systems enable large-scale generative or discriminative Artificial Intelligence (AI) and Machine Learning (ML) applications. As we lay out below, we see a particular need in this context to adapt our recommendation catalogue to cover developments concerning *i)* data lakes for data-intensive science, *ii)* indexing and data catalogues, *iii)* increasing usage of AI and ML methods in the ESS, and *iv)* the development of data-project concepts and storage grant schemes at supercomputing sites.

[1] https://doi.org/10.5281/zenodo.6565405
[2] Wilkinson et al., https://doi.org/10.1038/sdata.2016.18

Data Lakes

Context/Challenge

Given the extremely data-intensive nature of typical computing workflows in ESS, one key area that has attracted increasing interest in recent years is the integration of HPC with Data Lakes. Data Lakes offer a promising solution for storing and analysing vast and diverse datasets in their raw form, providing a repository for data-intensive applications. Depending on the architecture, Data Lakes can be organised in a centralised or federated way. The concept of Data Spaces, built according to the Dataspace Protocol[3] of the International Data Space Association (IDSA) and supported by the EU[4], is meant to play a crucial role in the federation of Data Lakes on a European scale in the next years. Beyond the architecture of Data Lakes, the integration of HPC and Data Lakes poses a number of challenges that need to be addressed in order to both maximise the use of HPC resources in conjunction with Data Lakes and to answer important questions in ESS. This integration aspect in particular needs further attention by our IG.

The prime challenge we see is the efficient utilization of compute resources within the HPC environment when dealing with Data Lakes. The sheer volume and variety of data stored in Data Lakes require sophisticated data management techniques to ensure that HPC systems can effectively access and process the data. These include efficient data partitioning or chunking, parallel processing, and data movement methods to streamline data workflows and prevent bottlenecks that impact performance.

Producing data for Data Lakes introduces additional challenges. Ensuring data quality and consistency becomes paramount for the usefulness of a Data Lake with complex data. The diverse sources and formats of data stored in Data Lakes can hinder re-usage, where a common bad-practice scenario is the creation of data silos with format compatibility and data quality issues. All this can impact the efficiency, accuracy and reliability of scientific computations performed on HPC resources. Data governance frameworks and implementing data cleaning and normalization procedures are thus becoming essential components in this complex ecosystem.

Moreover, the increasing adoption of cloud storage technologies, such as Amazon S3, poses technological challenges when integrating them with HPC systems. Solutions such as S3 introduce complexity aspects not found in traditional HPC storage architectures, possibly provoking issues with transfer speeds, latency and interfacing. Ensuring seamless integration between HPC and cloud storage platforms necessitates addressing compatibility issues, optimising data transfer protocols, and managing the scalability of storage solutions to accommodate the large datasets processed by HPC systems. Organizations must also consider data access controls and other practices that align with both HPC and cloud storage environments to mitigate potential security and compliance risks associated with utilising these diverse technologies in tandem. The direct utilisation of data from federated, remote Data Lakes or from Data Spaces in HPC poses potentially problematic challenges related to overheads from metadata- and data transfer over wide-area networks. In the best case, good data-indexing strategies (cf. next section) and appropriate metadata catalogues help to avoid performance problems, although this is not guaranteed under all circumstances.

To conclude, the integration of HPC with Data Lakes holds a great potential for advancing ESS research, but several challenges need to be addressed to realise its full benefits. Effective data management strategies, data quality assurance and data governance measures, and the adoption of cloud storage best practices including robust security measures are essential for an optimal interoperation of Data Lakes with HPC ecosystems.

[3] https://docs.internationaldataspaces.org/ids-knowledgebase/dataspace-protocol

[4] https://digital-strategy.ec.europa.eu/en/policies/data-spaces

Recommendations

- Data Management Workflows: Develop efficient data management workflows tailored to integrate HPC with Data Lakes. Implement data partitioning, replication, and movement strategies to streamline data processes.
- Data Quality Assurance and Governance: Encourage the establishment of robust and transparent standards for data formatting, data quality assessment, data validation and HPC-compatible data policies in the context of ESS-related data lakes. Criteria for certification of data lakes shall be considered, also in a context of international interoperability (e.g. via the Dataspace Protocol of the IDSA).
- Cloud Storage Best Practices integrated with HPC: Implement best practices for integrating cloud storage technologies such as Amazon S3 with HPC systems to optimise data storage, management, scalability, and data exchange. Security and regulatory requirements must be met for integration with HPC.

Catalogues, Indexing and Data Space Metadata

Context/Challenge

The exponential growth of data generated in ESS necessitates efficient access strategies. Data cataloguing and indexing offer a structured approach to discover, describe, organise and explore vast datasets. By providing metadata-driven access, these techniques streamline and hopefully accelerate the retrieval of appropriate data from data lakes as well as other data stores, and foster collaboration among researchers.

To ensure interoperability and efficient data sharing across multiple research institutions, standardization of data catalogues, indexes and access techniques is paramount. Frameworks such as STAC[5] and kerchunk[6] have emerged as valuable tools in this regard. STAC, a community-driven standard, provides a JSON-based schema for describing geospatial assets, facilitating the creation of comprehensive catalogues. Kerchunk, a Python library, offers a unified interface for representing various chunked and compressed data formats, enabling efficient data access from both traditional file systems and cloud storage.

On a cross-discipline interoperability side beyond STAC, metadata from data indexes and catalogues can be exported, for example via an OAI-PMH[7] server or a Data Space Connector according to the Dataspace Protocol (cf. above). General-purpose metadata standards such as Dublin Core[8], DCAT[9] and DataCite[10] are important in this context.

As a crucial aspect, system architectures will – where needed – have to distinguish rich library-type indexes or catalogues (for cold or warm data) from simple indexes facilitating high-performance hot data access. When correctly implemented, these indexes will address data-discovery and retrieval problems, reduce latencies, and simplify scientific workflows. Intelligent chunking and subsetting in data retrieval (cf. e.g. Kerchunk feature catalogue) is particularly important in the ESS for processing and analysing huge amounts of 2- and 3-D simulation and satellite data.

Recommendation

- Review emerging technologies and standards for data cataloguing and indexing in ESS, with focus on STAC and kerchunk but also alternatives and complements. Potential inputs to this review can be usability surveys or case studies.

[5] https://stacspec.org
[6] https://fsspec.github.io/kerchunk/
[7] https://www.openarchives.org/pmh/
[8] https://www.dublincore.org/
[9] https://www.w3.org/TR/vocab-dcat/
[10] https://schema.datacite.org

Requirements from ML/AI Applications

Context/Challenge

HPC Workflows involving methods from Artificial Intelligence and Machine Learning often require huge datasets as input, which may be generated from models and/or observations. The so trained ML-model is then used in applications. Innovative models will re-train with continuously incoming data streams during long term execution, as it is common practice, e.g., in continuously running Digital Twins. For this kind of "incremental-update" application cases it may turn out to be infeasible or even not useful to preserve input data streams for longer, owing to storage-volume requirements and/or lack of compute capacity for reprocessing old inputs. But also the output that users extract from this kind of models is probably difficult to continuously record. Making such applications FAIR thus poses a fundamental challenge, as it clearly requires recording a continuous stream of metadata and the registration of persistent identifiers on the in- and outputs as well. First concepts in this direction have recently been published[11].

A special class of machine learning models that appeared quite recently, the Foundation Models, may be considered as a prime example posing such challenges to a FAIR RDM. On the one hand, Foundation Models may improve iteratively, and on the other hand so-called down-stream tasks, i.e. applications, need a task-specific further training before they can be solved. As in the case of large language models (LLMs), having permanently accessible instances of the model for exploitation by a range of user communities poses a large-scale RDM problem and a challenge in the field of non-commercial IT infrastructure provision.

The recently formed RDA interest group FAIR4ML[12] stated "The FAIR principles also can apply to machine learning tools and models, though a direct application is not always possible as machine learning combines aspects of data, software and computational workflows." First orientation has been given in a poster[13] by Katz et. al (2020) at the RDA Virtual Plenary 16, where it was stated that ML model data have characteristics of software and data at the same time.

We conclude that the challenges for making AI-related HPC-data FAIR are threefold:

1) keeping the pre-processed training data available for ML-model updates, (possibly also for new applications, requiring documentation and quality managed metadata);

2) preserving the machine learning model training results for re-use in other contexts. For this case, the storage format may be software framework dependent, i.e. not easily translatable to new versions or different kinds of frameworks and programming languages; and

3) running AI-models permanently as a kind of cloud service, or, at least, offering software and data containers for execution on local resources, with the aim of enabling a wide range of applications and developments, possibly from more than a single scientific discipline.

Recommendations

- Establish collaboration with FAIR4ML interest group in RDA.
- Work towards concept for long-term storage, accessibility and FAIRness of AI and ML data, considering first approaches available.
- Data projects for re-use/exploitation of ML-models (see also below).

---

[11] e.g. Sherpa et al., https://datascience.codata.org/articles/10.5334/dsj-2024-055

[12] https://www.rd-alliance.org/groups/fair-machine-learning-fair4ml-ig

[13] Katz et al., https://doi.org/10.5281/zenodo.4271996

Data Projects and Cross-Centre/Cloud Usage of Data

Context/Challenge

Current models of storage-grant management at HPC centres are oriented at primary dataset production, usually coupling storage grants to the compute-time grant. Thus, the storage granted traditionally has to be emptied after the end of the HPC calculations and some grace period. After utilisation in "lighthouse" data-analysis projects, data are then often moved to slow tape archives or even discarded. Recently, awareness on the value of data has led to some more flexible schemes in data keeping, e.g. to longer grace periods.

We argue that a stronger and consistent paradigm-shift towards data-centric projects (which we will call "data projects" in the following) may be helpful. We would see a first line of data projects concentrating on data sets being stored for re-use by fellow researchers that are given access. Another category or funding line would support "community data projects", hosting data that is reused by a larger community, where data-transfer volumes may be higher and consistent curation can be important. For both categories, storage could be granted on the basis of a scientific proposal (as it is done for computing time on large HPC clusters). These would include a plan or discussion outlining the value of a dataset for the field when analysed further by the proposing group or the general community. The proposal would also suggest storage and/or archival periods. Computing centres would respond to justified requests by granting storage and ideally also compute time on systems for post-processing and visualisation. These systems or services are not necessarily bare-metal clusters; they may follow cloud-computing concepts in the form of IaaS (infrastructure-as-a-service), PaaS (platform-as-a-service), or even SaaS or FaaS (software-as-a-service, function-as-a-service).

When re-usage by a broader community is really intended, supporting the basic ideas behind FAIR and open data, processing power would also have to be made available to general community users of the respective datasets. This would allow them, for example, to perform necessary basic data-reduction operations for downloading only parts of an extremely huge dataset.

Finding the financing and billing mode for computing-facility access with a data project as sketched above, in particular if it involves the scientific-community users of a dataset – who have not written the data-project proposal – is certainly a more complex challenge. However, it may be a crucial step to incentivise re-usage of data, including data validation by people other than the original authors. The re-financing of such an access and usage, which has to comply with administrative spending rules, is currently completely unclear. We see an opportunity and potential for NFDI4Earth and NFDI activities here. From a technical point of view, we would guess that relatively small amounts of computing resources would be sufficient for a large usability gain; user authentication to such resources remains a challenge, in particular as federated Single-Sign-On (SSO) solutions would be highly desirable.

Recommendations

- Keep discussing the idea of data projects and community data and the connection of these to the availability of computing systems to data producers or data users for post-processing in contexts such as Helmholtz, Nationales Hochleistungsrechnen (NHR) and Gauss Centre for Supercomputing (GCS).
- Activate NHR centres and beyond for piloting a conceptual framework for data projects and community data. Distil the minimum demands on such projects to be realizable under the current funding regulations. Initiate an NHR project on criteria and necessary development steps for implementing FAIR HPC Data projects.
- Monitor solutions at single centres or within centre associations, and help to devise feasible unified concepts.

3 Monitoring: Recommendations from 2022 vs. Reality 2024/25

In the tables below, our recommendations from our position paper of 2022 are reviewed group by group with respect to implementation status. We have found some recommendations which we will rather not follow up in the scope of our IG; in these cases the status discussion is typeset in **bold face**.

General Recommendations

Recommendation	Status
1. An infrastructure registry for NFDI4Earth to find matching computing resources	A living-handbook article has been prepared to begin with, giving an overview of resources and access modalities in the consortium.
2. Approaches to access computing resources through Single Sign On	Base4NFDI is establishing an AAI (project IAM4NFDI), but adoption at the various computing centres with their different administrative contexts is needed as a next step.
3. Seamless cross-centre and cross-system computing access for users including accounting	Within NFDI4Earth, the IG has launched the CAPICE pilot to implement cross-centre workflows involving DKRZ, DLR and LRZ. The NFDI Section Common Infrastructures is developing concepts for a federated Multi-Cloud. On EU scale, there are modern approaches as well, e.g. the LEXIS platform (www.lexis-project.eu). The development and consolidation of these concepts will need some years and will be monitored by the IG, making practical contributions as in CAPICE.
4. Interoperability & interconnection of computing systems	Network infrastructure is continuously enhanced, but cross-system interoperability is seldom addressed.
5. Robust computing environments for re-usability and reproducibility (e.g. containerisation)	Containers on HPC machines pose problems, and persistent individualities of each HPC system (e.g. special, site specific MPI[14] libraries also used in containers) might make reproducibility gains through a containerised software stack wishful thinking. SPACK[15] HPC package management makes software stack behaviour somewhat more predictable and may be recommended. Community framework usage for postprocessing is recommended to increase transparency and reproducibility of analysis steps. These steps can be usually executed well on virtualised infrastructure and packaged into containers. For smaller tasks, it is thus of great help to have virtualisation environments (e.g. Kubernetes clusters, computing clouds) connected to or integrated within HPC clusters at computing centres.

Specific Recommendations: Pillar "(Meta-)Data Formats, Interoperability and Reproducibility"

Recommendation	Status
6. Establishing (meta-)data formats in collaboration with standardisation bodies, RDA, etc.	– see below a)…d) –

[14] https://www.mpi-forum.org/

[15] https://spack.io/

6.a) Max. ~10 general storage formats	Up to now, in HPC we see certainly a usage of: HDF5, NetCDF, ZARR, and OGC-endorsed standards (e.g. GeoTIFF). Standards related to parallel-processing frameworks (Geospark, Hadoop, ...) should be monitored.
6.b) Max. ~10 metadata formats	Datacite is clearly endorsed for HPC data, and another NFDI4Earth IG is working on standards for Geochemistry. We are monitoring complementary developments also in other NFDI consortia; we expect these to be streamlined by the respective section in NFDI e.V. (https://www.nfdi.de/section-meta) and its Task Force Metadata. Interesting developments may come from the result evaluation of a workshop in Jan 2025 driven by this Task Force.
6.c) NFDI4Earth recommends these formats	We are considering including these recommendations in the NFDI4Earth label (cf. https://nfdi4earth.de/2interoperate/common-standards) process.
6.d) Support for these formats at all computing centres in NFDI4Earth	The Helmholtz, NHR and GCS sites involved in NFDI4Earth seem to have basic support for the (meta-)data formats mentioned above in place or in development. A more detailed monitoring will be undertaken later, as e.g. relevant developments within the InHPC-DE project of the GCS centres consolidate.
7. Standardised software environments at computing centres, documented within NFDI or NFDI4Earth	Here, no progress has been made yet. PUNCH4NFDI is making such an effort in the astrophysics sector, where we can monitor lessons learned. The standardisation of HPC software stack deployment techniques (e.g. via SPACK) may facilitate re-use and standardisation of software stacks to some degree.
8. Recommendations on RSE and software packaging	**This should be cut out of our recommendation catalogue and left to RSE groups such as the NFDI4Earth IG Research Software and https://de-rse.org.**
9. Recommendations on quality control	Data quality control is at the moment more or less in the scope of scientific work, as a generic method seems missing. As for metadata, quality relies on the choice of adequate standards and the avoidance of various potential problems in usage (e.g. partially filled fields). For filling in DataCite metadata correctly, there are guides which the researchers can use (e.g. https://doi.org/10.5281/zenodo.3559800) The group may try to generate momentum on quality control in the context of future data lakes (see above).
10. Political work towards viable solutions across state/national borders	We are working in international contexts (e.g. EuroHPC, EOSC via NFDI), and the cooperations are still intensifying. State-funded initiatives for international exchange (e.g. LRZ-IT4Innovations/CZ) complement this. Furthermore, associations of computing centres (NHR, Gauß-Allianz, GCS, …) will be essential for overcoming administrative obstacles across federal states.

Specific Recommendations: Pillar “Federated Access, Findability and Accessibility”

Recommendation	Status
11. Issue NFDI4Earth recommendations on efficient data access and transfer for Big Data	The results of several workshops and surveying the HPC (NHR, GCS, ...) landscape in Germany are being condensed into the NFDI4Earth Synthesis Architecture.
12. Removal of administrative barriers regarding federated data/compute access	– see below a)...c) –
12.a) Establish SSO following NFDI e.V. recommendations	Will presumably happen via IAM4NFDI – awaiting results and then discussing adoption possibilities. We aim at monitoring and fostering fruitful developments also in the context of existing federations between HPC Centres (NHR, Gauß-Allianz, GCS, ...) and across research institutes (Helmholtz/HIFIS, DFN, ...). Low-barrier data retrieval and on-site evaluation (e.g. on cloud infrastructures next to HPC) with SSO can play a major role in implementing FAIR and boosting science.
12.b) Establish portals giving free access to Open Data without moving data from original storage	There are examples where this can already be achieved, using e.g. EUDAT-B2SHARE (b2share.eudat.eu) or the LRZ FAIR Data Portal for posting metadata with links to the actual data remaining on-site. For file access and transfers, object storage (S3) is probably an accepted standard for storage and data transfer besides transfer protocols such as GLOBUS, UFTP or GridFTP. The availability and usability of respective middleware frameworks is improving constantly.
12.c) Survey to find users with inadequate data/compute access (every ~18 months).	**This has not been followed up yet; instead of a fixed frequency such surveys shall be conducted on demand.**
13. Data-portal and storage solutions with spatio-temporal and layer-/field-based data selection & download	The IG has not worked on this yet; solutions like PAVICS etc. will have to be explored.
14. Establish standardised (meta-)data interface to institute- or subdiscipline-specific interfaces	– see below a)...b) –
14.a) Metadata interfaces, starting with OAI-PMH	OAI-PMH and OGC standards distributing metadata are certainly endorsed**. This topic is, however, not HPC-specific and the recommendation should not be followed up by the IG.**
14.b) Advance usage of standard data interfaces, e.g. S3	S3, UFTP, GridFTP, GLOBUS are on the list of promising developments and the IG keeps discussing them on various occasions, encouraging use.
15. Clarification and NFDI4Earth index of LTA solutions	There is a new LTA Interest Group in NFDI4Earth; **thus, the topic should not be followed up by the IG.**
16. Interactive HPC including checkpointing	Jupyter-based prototypes are established within, e.g. GCS and subject of a new Base4NFDI proposal. Individual HPC sites have been implementing Jupyter for data exploration, processing and AI development. It is important to mention that these services should be connectable also to external storage such as S3.
17. Concepts for services reacting on data events (e.g. alerting)	The new version of ESGF (https://esgf.llnl.gov) with CMIP has a message queue where one can subscribe to new-data-arrival messages. This may be seen as an example implementation which others could follow.

Acknowledgements

This work has in large parts been funded by the German Research Foundation (DFG) through the project NFDI4Earth (DFG project no. 460036893) within the German National Research Data Infrastructure (NFDI). We are thankful for contributions from discussions and workshops involving the IG and its members during the preparation of this paper. In particular we thank our colleagues H. Nolte and J. Kunkel from GWDG/NHR for pointing us at current activities in data project realisation.