



NFDI4Earth

IG High-Performance Computing in Earth System Sciences (HPC in ESS)

NFDI4Earth Kick-Off Dresden 9/10 June 2022

For the IG: Stephan Hachinger, Leibniz Supercomputing Centre (LRZ, Garching b.M.)

IG High-Performance Computing in ESS: Motivation

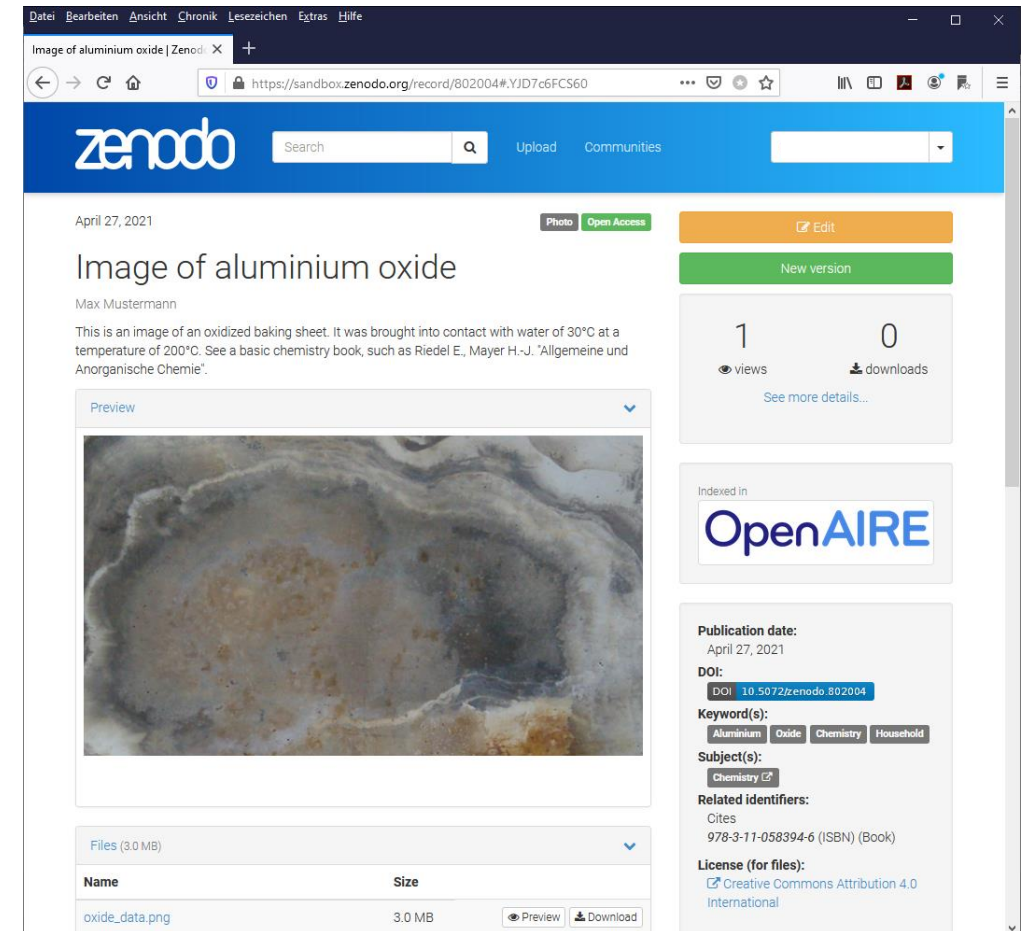
In HPC, FAIR methodology is a challenge!

Consider the role of repositories:



- offer space to deposit data and descriptive (& other) metadata
- make data findable, get you DOIs
- fulfill requirements of funding agencies

Problem: maybe 50 GB of storage is not quite enough for your multi-Terabyte HPC data.

→ Standard repositories are out!



The screenshot shows a Zenodo record page for a file named "Image of aluminium oxide". The page includes a search bar, navigation links for "Upload" and "Communities", and a "Photo" button. The record is dated April 27, 2021, and has 1 view and 0 downloads. It is indexed in OpenAIRE. The publication date is April 27, 2021, and the DOI is 10.5072/zenodo.802004. The keyword(s) are Aluminium, Oxide, Chemistry, and Household. The subject(s) is Chemistry. The related identifiers include Cites and 978-3-11-058394-6 (ISBN) (Book). The license for files is Creative Commons Attribution 4.0 International. A table at the bottom shows the file "oxide_data.png" with a size of 3.0 MB and options for preview and download.

Name	Size	Preview	Download
oxide_data.png	3.0 MB		

IG HPC in ESS: Motivation

FAIR „vs.“ HPC in ESS: Lots of challenges!

- **Technical:** Datasets produced too large for repositories – satellite data & global/large-scale high-resolution datasets
- **Methodical:** Reproducibility – simulation/analytics output depends on compiler switches, libraries, etc.
“Wild” community software usage: netCDF, eccodes, ...
- **Administrative:** Computing needs grants and accounting, but researchers want to quickly work where their data are.
- **Practical:** “Everyone talks while I have 1PB to handle!”

We have collected challenges on a „living poster“ on this conference!
Tell us what is important to you & give your opinions!

IG High-Performance Computing in ESS: Context

Scientists + Computing/Data Centres: NHR, Gauss-Allianz, GCS, ...

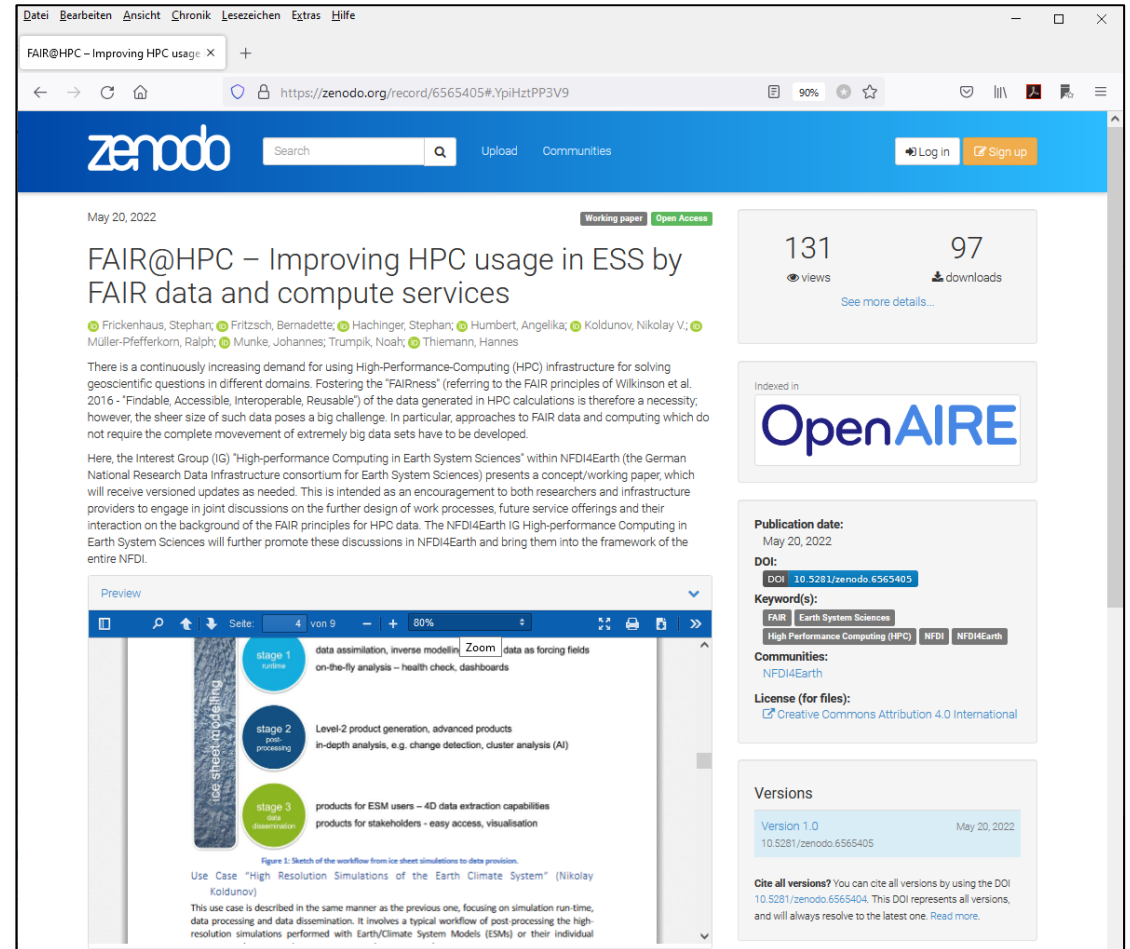


IG HPC in ESS: Format / Aims

- Every 2nd Monday in month: meet 9:00 am: <https://meet.lrz.de/ig-hpc-nfdi4earth>
- Some goals:
 - semi-automatically enrich results from HPC systems with metadata
 - produce (meta-)data according to NFDI4Earth standards
 - steps towards federated 'in place' data-analysis facilities
 - foster and simplify re-use of HPC data
- Current „projects“
 - concept paper “Use Cases → Challenges → Recommendations for Development”
 - workshop with computing centres (NHR, Gauss-Allianz, GCS)
 - contact with NFDI sections (Metadata, Common Infra)
 - tech research projects: e.g. on latencies in distributed data MGMT

IG HPC in ESS: First Result – Concept Paper

- IG activities – two „pillars“:
 - (Meta-)data formats, interoperability and reproducibility (→ reusability)
 - Federated access, findability and accessibility
- Two use cases & their needs examined
 - Ice sheet simulations
 - High-resolution climate simulations
- Challenges identified
 - In HPC/system usage
 - In FAIRness, sharing, re-use
- Recommendations (monitor every 1.5 yrs)
 - Infrastructure registry & unified access
 - Homogeneous computing environments
 - Standard formats (data/metadata)
 - Standardised (meta)data access / portal solutions



The screenshot shows a Zenodo record page for a concept paper. The title is "FAIR@HPC – Improving HPC usage in ESS by FAIR data and compute services". The page includes a search bar, navigation links, and a list of authors. The abstract discusses the demand for HPC infrastructure and the challenges of data management. The page also features a preview of a workflow diagram, a list of keywords, and a version history section.

Concept paper:
<https://doi.org/10.5281/zenodo.6565405>



Thank You !



<https://www.nfdi4earth.de/>

Want to get active with us ? – Join by mail to hachinger [at] Irz.de !