PUNCH4NFDI

B-1 Progress Report Part 1 for the consortium

Particles, Universe, NuClei and Hadrons for the NFDI

September 2024

Table of Contents

1. General Information

Name of the consortium:

PUNCH4NFDI - Particles, Universe, NuClei and Hadrons for the NFDI

PUNCH4NFDI is the consortium of particle, astro, astroparticle, hadron, and nuclear physics. The goal of PUNCH research is to identify the elementary building blocks of matter, the laws governing their interactions, and their impact on the development of the universe and the formation of structures in it. Thus, PUNCH4NFDI addresses the research domains: 3.24 Particles, Nuclei and Fields; 3.25 Astrophysics and Astronomy; 3.22 Statistical Physics, Nonlinear Dynamics, Complex Systems, Soft and Fluid Matter, Biological Physics; 3.23 Optics, Quantum Optics and Physics of Atoms, Molecules and Plasmas.

To a large extent, PUNCH research is carried out at large and international facilities: accelerators and detectors of particle and hadron & nuclear physics, observatories, and telescopes. Observations are accompanied by high-precision theory predictions, model calculations, and phenomenological simulations.

Spokespersons: Thomas Schörner (BLV), Andreas Haungs and Christiane Schneide (NFDI)

Links of the consortium's website and repositories used for publishing output:

Name	**Link**
PUNCH4NFDI web site	https://www.punch4nfdi.de
PUNCH4NFDI results page	https://results.punch4nfdi.de
Zenodo community	https://zenodo.org/communities/punch4nfdi/
Public GitLab	https://gitlab-p4n.aip.de/punch_public

2. Summary

PUNCH4NFDI, the NFDI consortium of particle, astro, astroparticle, hadron, and nuclear physics, is providing more and more services and infrastructure for a FAIR[1] and open data management in the PUNCH sciences and beyond. Building upon long experience in data management and dissemination, the goal is to support the FAIR handling and publication of research outcomes, such as data or software, and to facilitate easy and seamless access to and handling of data. To this end, the consortium has collected existing

[1] https://doi.org/10.1038/sdata.2016.18

initiatives and concepts used in its communities, upon which – in a second step – the goals and detailed work schedule were defined.

The core of PUNCH4NFDI is the Science Data Platform, which will allow via single sign-on access to federated computing and storage resources (Compute4PUNCH and Storage4PUNCH), as well as to a registry for research products (RPR). Hereby, the access to research outcomes, which are registered as digital research product (DRP), and their reuse by providing associated computing and storage resources is simplified. **First pilot implementations of several elements of this ambitious plan have already been made accessible to users.** These include

- the PUNCH-AAI for single-sign on, which is operated as part of the Helmholtz-ID,
- federated computing resources (Compute4PUNCH),
- federated storage resources (Storage4PUNCH),
- a container registry and a registry for digital research products (DPR),
- a workflow engine (REANA), which can be used to submit workflows to Compute4PUNCH,
- a metadata and file catalogue service, and
- several use cases, i.e. examples of workflows, to demonstrate the usage of the individual elements for scientific analyses (e.g. analysis of ATLAS Open Data[2], or plotting astronomical data from Gaia data release 2[3]).

The management of **federated heterogeneous resources** has a long tradition in the PUNCH communities. Still, major developments were necessary in order to allow token-based access. The onboarding of further sites to the distributed compute and storage infrastructures is still ongoing. The workflow engine was extended to accept Compute4PUNCH as backend, and this update will be made publicly available with the next REANA release. Specialised containers were developed in order to enable user interaction with running workflows, e.g. via JupyterLab or VS Code.

As a **result of these efforts**, a user can now

- submit a complex analysis as workflow to REANA, which will execute each step of the workflow under the specified computing environment,
- access data at Storage4PUNCH,
- submit jobs to the specified backend (e.g. Compute4PUNCH or Kubernetes), and

[2] https://gitlab-p4n.aip.de/compute4punch/tutorials/atlas_open_data_example_1
[3] https://gitlab-p4n.aip.de/punch_public/reana/tutorials/-/tree/main/advanced/example1?ref_type=heads

- store results of the analysis and capture all relevant metadata of the workflow for reproducibility.

Digital research products can already be registered in the RPR. A facility to import and export workflow metadata from REANA by the RPR is under development. The consortium developed a schema for DRPs, based on DataCite, including required and optional metadata which is still being further evolved, building on expertise of various PUNCH communities. This will enhance the search for DRPs and thus improve their findability.

In order to **enable the community** to make their research outcomes FAIR and to make use of the tools and services provided by PUNCH4NFDI, the consortium has rolled out a training programme covering for example REANA, Storage4PUNCH and Compute4PUNCH, but also JupyterHub, software development with Git, and efficient use of HPC computing.

PUNCH4NFDI is actively **contributing to the NFDI sections as well as the Base4NFDI initiative and its projects**. The consortium is involved in IAM4NFDI, which is currently in the integration phase. Furthermore, PUNCH4NFDI participates in the recently approved Base4NFDI project nfdi.software which will start in November 2024, as well as in the MultiCloud proposal which is currently under evaluation.

On national and international level, PUNCH4NFDI is strongly connected with further initiatives relevant for the scientific community as well as infrastructure providers. These include, for example, the ErUM-Data Hub[4], as well as Helmholtz initiatives like HMC and HIFIS, as well as ESCAPE[5], the European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures.

While PUNCH4NFDI concentrated on developing prototypes of the core elements of the Science Data Platform and interfaces during the first part of this funding phase, the main focus during the next years is on improving the stability and user-friendly handling of the provided services, provide improved support for users via a helpdesk system, and in particular integrate various community data publishing sites.

3. Composition of the Consortium

Applicant institution:

Applicant institution	Location	Duration
Deutsches Elektronen-Synchrotron (DESY)	Hamburg	10/21 - today

[4] https://erumdatahub.de/en/
[5] https://projectescape.eu/

Spokesperson:

Spokesperson (ORCID)	Institution, location	Duration
PD Dr. Thomas Schörner (0000-0002-7213-0352)	Deutsches Elektronen-Synchrotron DESY, Hamburg	10/21 - today

Co-applicant institutions:

Co-applicant institution	Location	Duration
Forschungszentrum Jülich (FZJ)	Jülich	10/21 - today
Frankfurt Institute for Advanced Studies (FIAS)	Frankfurt / Main	10/21 - today
Georg-August-Universität Göttingen (GAU)	Göttingen	10/21 - today
GSI Helmholtz-Zentrum für Schwerionenphysik GmbH (GSI)	Darmstadt	10/21 - today
Hochschule für Technik und Wirtschaft (HTW)	Berlin	10/21 - today
Johannes Gutenberg-Universität (JGU)	Mainz	10/21 - today
Karlsruher Institut für Technologie (KIT)	Karlsruhe	10/21 - today
Leibniz-Institut für Astrophysik (AIP)	Potsdam	10/21 - today
Ludwig-Maximilians-Universität (LMU)	München	10/21 - today
Max-Planck-Institut für Radioastronomie (MPIfR)	Bonn	10/21 - today
Max-Planck-Institut für Kernphysik (MPIK)	Heidelberg	10/21 - today
Rheinische Friedrichs-Wilhelms-Universität (RFWU)	Bonn	10/21 - today
Ruprecht-Karls-Universität (RKU)	Heidelberg	10/21 - today
Thüringer Landessternwarte (TLS)	Tautenburg	10/21 - today
Technische Universität Dresden (TUDD)	Dresden	10/21 - today
Technische Universität Dortmund (TUDO)	Dortmund	10/21 - today
Universität Bielefeld (UB)	Bielefeld	10/21 - today
Universität Hamburg (UHH)	Hamburg	10/21 - today
Universität Regensburg (UR)	Regensburg	10/21 - today

Co-spokespersons:

Co-spokesperson (ORCID)	Institution, location	TA	Duration
Prof. Dr. Susanne Pfalzner (0000-0002-5003-4714)	FZJ, Jülich	3	10/21 - today
PD Dr. Andreas Redelbach (0000-0002-8102-9686)	FIAS, Frankfurt	5	10/21 - today
Prof. Dr. Arnulf Quadt (0000-0002-0098-384X)	GAU, Göttingen	7	10/21 - today
Dr. Kilian Schwarz (0000-0002-0800-2743)	GSI, Darmstadt	2, 6	10/21 -08/2022
Dr. Mohammad Al-Turany (0000-0002-8071-4497)	GSI, Darmstadt	6	08/2022 - today

Prof. Dr. Hermann Hessling (0000-0001-6809-6895)	HTW, Berlin	5, 6	10/21 - today
Prof. Dr. Volker Büscher	JGU, Mainz	5	10/21 - today
Dr. Andreas Haungs (0000-0002-9638-7574)	KIT, Karlsruhe	4, 7, 1	10/21 - today
Prof. Dr. Matthias Steinmetz (0000-0001-6516-7459)	AIP, Potsdam	4	10/21 - today
Prof. Dr. Joseph Mohr (0000-0002-6875-2087)	LMU, München	3	10/21 - today
Prof. Dr. Michael Kramer	MPIfR, Bonn	5, 7	10/21 - today
Prof. Dr. Jim Hinton	MPIK, Heidelberg	6	10/21 - today
PD Dr. Philip Bechtle (0000-0003-3479-2221)	RFWU, Bonn	4	10/21 - 06/2023
Prof. Dr. Frank Bertoldi	RFWU, Bonn	7	06/2023 - today
Prof. Dr. Stefan Wagner (0000-0002-7474-6062)	RKU, Heidelberg	6	10/21 - today
Dr. Matthias Hoeft (0000-0001-5571-1369)	TLS, Tautenburg	2	10/21 - today
Prof. Dr. Arno Straessner (0000-0003-2460-6659)	TUDD, Dresden	5	10/21 - today
Prof. Dr. Kevin Kröninger (0000-0001-9873-0228)	TUDO, Dortmund	7	10/21 - today
Prof. Dr. Dominik Schwarz (0000-0003-2413-0881)	UB, Bielefeld	2	10/21 - today
Prof. Dr. Gregor Kasieczka (0000-0003-3457-2755)	UHH, Hamburg	3	10/21 - today
PD Dr. Sara Collins	UR, Regensburg	3	10/21 - today

Participating institutions:

Participating institution	Location	Duration
Albert-Ludwigs-Universität	Freiburg	10/21 - today
Europäisches Kernforschungszentrum CERN	Geneva (CH)	10/21 - today
Deutsches Luft- und Raumfahrtzentrum (DLR)	Jena	10/21 - today
Deutsche Physikalische Gesellschaft (DPG)	Bad Honnef	10/21 - today
Johann Wolfgang Goethe-Universität (GU)	Frankfurt / Main	10/21 - today
Hochschule Darmstadt (HDA)	Darmstadt	10/21 - today
Humboldt-Universität (HU	Berlin	10/21 - today
Helmholtz-Zentrum Dresden-Rossendorf (HZDR)	Dresden	10/21 - today
Leibniz-Institut für Sonnenphysik (KIS)	Freiburg	10/21 - today
Leibniz-Rechenzentrum	Garching	10/21 - today
Max Planck Computing and Data Facility (MPCDF)	Garching	10/21 - today
Physikalisch-Technische Bundesanstalt	Braunschweig	10/21 - today

(PTB)		
Ruhr-Universität (RUB)	Bochum	10/21 - today
RWTH Aachen University (RWTH)	Aachen	10/21 - today
Technische Informationsbibliothek – Leibniz Information Centre for Science and Technology (TIB)	Hannover	10/21 - today
Technische Universität Darmstadt (TUDA)	Darmstadt	10/21 - today
Technische Universität München (TUM)	München	10/21 - today
Universität Potsdam (UP)	Potsdam	10/21 - today
Universität Siegen (USi)	Siegen	10/21 - today
Universität zu Köln (UzK)	Köln	10/21 - today
Verein für datenintensive Radioastronomie e.V. (VdR)	Jena	10/21 - today
Westfälische Wilhelms-Universität (WWU)	Münster	10/21 - today