

The PUNCH4NFDI Consortium

Newsletter Number 2

February 2022



Contents

- 0. News***
- 1. Highlights***
- 2. NFDI and related topics***
- 3. Formal, legal and financial topics***
- 4. Status of the consortium***
- 5. Communication and collaborative tools***
- 6. Events and upcoming excitements***
- 7. Recent talks, results, and publications***

0. News

PUNCH4NFDI is embedded in a very actively evolving field: More and more consortia of the second NFDI round get going (and PUNCH4NFDI is in good contact with a few of them – see below), entering the discussions with the older first-round consortia. The NFDI e.V. is forming an increasing number of discussion circles and “sections” – cross-sectional working groups on a selection of important topics. It is of high importance for PUNCH4NFDI to be involved in and contribute to the NFDI-wide efforts and to obtain visibility at this NFDI level!

Our consortium itself is picking up speed at a considerable pace. The task areas and work packages are staffed with more and more new people (we introduce a few of the below), and the consolidated work programmes are beginning to be executed. A few first steps have been taken towards achieving our first deliverables – e.g. a Storage4PUNCH demonstrator (see below). You are warmly invited to join the task area meetings most relevant to your research¹.

At the same time, new challenges arise: The German Research Foundation DFG has issued a call for base service consortia, and applications need to be handed in on a very short timescale². We are now actively discussing our participation in such an endeavour – stay tuned! We are also approaching our first quarterly report – there is a bit of work coming towards the TA / WP leads and the representatives of co-applicant and participant institutions. But considering the necessity to document our achievements and to be able to convince, in a not too far future, expert reviewers of the importance to further fund PUNCH4NFDI, this effort seems well invested.

In case you have news you want to share with PUNCH4NFDI, don't hesitate to send a mail to info@punch4nfdi.de. You can access all newsletters on our web site³.

1. Highlights

- One of the first deliverables of Task area 2 is a prototype for distributed storage, also known as Storage4PUNCH, leveraging "data lake" technologies that are also developed e.g. in the ESCAPE project (deliverable D-TA2-WP1-1 “Prototype data lake setup @ DESY & GSI”). A first test system has been configured at DESY, where most effort was spent to enable the access via the PUNCH-AAI and the documentation of the required clients. The system will soon be complemented with one or more storage instances at different locations. Documentation is available in the intranet⁴.
- In early February, the gammapy project⁵ was awarded the Jury Award of the Open Science Award for open-source for research software of the French Ministry of Higher Education, Research and Innovation⁶. Gammapy is an open-source demonstrator from task area 6, work package 4, and we are very pleased that efforts to which PUNCH colleagues contributed significantly are acknowledged in this formal way.

¹ <https://indico.desy.de/category/741/>

² https://www.dfg.de/foerderung/info_wissenschaft/2022/info_wissenschaft_22_08/index.html

³ https://www.punch4nfdi.de/news_amp_events/newsletter/

⁴ <https://intra.punch4nfdi.de/?md=/docs/TA2/WP1/StoragePrototyping.md>

⁵ <https://gammapy.org/>, <https://github.com/gammapy>

⁶ <https://www.ouvrirelascience.fr/open-science-free-software-award-ceremony/>

2. NFDI and related topics

- A driving theme of the entire NFDI and beyond is the German Research Foundation (DFG) call for base service consortia mentioned above⁷. PUNCH4NFDI is deeply involved in the discussions on how to best structure the field of base services. These discussions prove to be rather complicated
 - due to the various possibilities for the realisation of base service consortia,
 - because of a lack of a clear definition of base services, and
 - because of the many players involved (NFDI consortia, external institutions like the DFN, Helmholtz, ...) and the necessity to stay compatible with developments e.g. on the European level (EOSC etc.).

PUNCH4NFDI has a clear interest to contribute to the development and provision of base services like AAI, file transfer systems, software repositories and others, and we aim at realising these mainly in the context of the physics-related consortia in the NFDI. Stay tuned for interesting developments in this field.

- In order to best address commonalities with other NFDI consortia especially in the physical sciences and in related fields, PUNCH4NFDI and DAPHNE4NFDI organise a workshop in April 2022⁸. The goal of the workshop is to foster the exchange on existing competences and solutions and, based on this fuller picture, to discuss knowledge exchange and the development / setup of common tools and services. Main topics to be addressed will be:
 - Analysis workflows and distributed computing
 - Software repositories for common software solutions
 - Metadata schemas
 - AAI and other collaborative tools

Please register on the INDICO page if you are interested in participating or contributing to the workshop.

- So far, four NFDI sections have been established, and a number of NFDI task forces have started working, and PUNCH colleagues are involved in all of them. In future newsletters, highlights from these groups will be sketched. For the time being, the task force on “Monitoring” might be the most important one, since it aims at defining criteria for the evaluation of consortia in view of their prolongation.

3. Formal, legal and financial topics

As previously discussed, the financial and legal arrangements of PUNCH4NFDI are settled. All contracts are signed, first funds have been distributed, and the first round of “Verwendungsnachweise” was successfully mastered. A great “Thanks” goes to Sonja Kempe at DESY and to all administrative contacts at all partner institutions for making this work out so smoothly!

Irrespective of the smoothly running formal processes, PUNCH4NFDI – like all other consortia – is suffering from a “underspending” problem – see figure 1. The NFDI is discussing mechanisms to mitigate this issue. So far, the only identified mechanism is to ask for a postponing of funds from one year into the next year. This worked for the

⁷ https://www.dfg.de/foerderung/info_wissenschaft/2022/info_wissenschaft_22_08/index.html

⁸ <https://indico.desy.de/event/33410/>

transition from 2021 to 2022, but there is no guarantee that the DFG will answer future requests favourably. All partners are therefore encouraged to spend the allocated resources without delay.

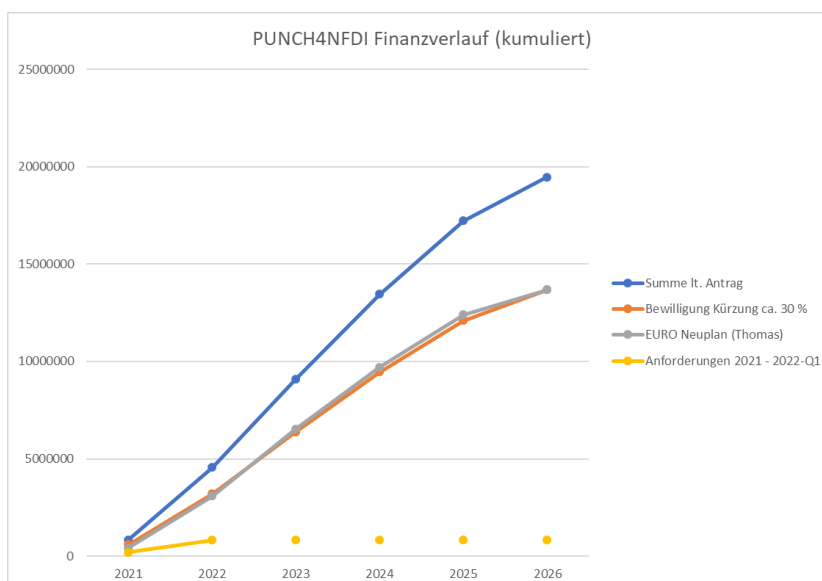


Figure 1: Cumulated spending curves. Shown are the sums of the proposal (blue) and the sums granted by the DFG (grey / orange) together with the current spending situation (yellow).

4. Status of the consortium

Work programme and reporting

The reduced and revised work programme of the consortium has been finalised and is available on the intranet⁹. Substantial changes to this work programme need to be discussed and agreed upon in the Management Board.

A first reporting cycle will be started in mid-March, with the aim of having a first quarterly report ready by the middle of April. A template for the report – to be filled in by task area and work package leaders as well as by institute representatives has been shown at the general meeting on 16 November 2021 in Thomas' presentation¹⁰.

New people

New people who contribute to the work programme of PUNCH4NFDI are our most valuable asset! In this category of the newsletter, we try to give short introductions of newcomers to the consortium – welcome to everybody!

⁹ <https://intra.punch4nfdi.de/?md=/docs/Consortium/Documents/Documents.md>

¹⁰

<https://indico.desy.de/event/32298/contributions/113001/attachments/69833/88648/20211116.GeneralMeeting.pdf>



Credit: Jasper Dohse

David Ohse (Bonn): *"I studied physics and astrophysics at the University of Bonn. My bachelor thesis at the Max Planck Institute for Radio Astronomy was dedicated to pulsar observations, whereas my master thesis used theoretical approaches to investigate millisecond pulsars. Therefore, I'm trained in observational radio astronomy as well as in theoretical stellar astrophysics. Besides studying, I've been involved in diverse science communication activities with the University of Bonn Physics Show and the young German*

Physical Society (jDPG). For PUNCH, I'm coordinating public outreach activities within the task area 7 "training, education, outreach & citizen science"."



Lisa-Marie Stein (DESY): *"I work in the library and documentation group at DESY in Hamburg, engaged mainly in publication issues. As a studied linguist and media scientist (M.A.) I like focusing on details, playing with different concepts of meaning and supporting the interchange of different communities. As a newcomer to PUNCH4NFDI in the field of research data management, I intend to assist the participating communities to get together and find a common language in the shape of metadata structures. Contributing to Work Package 2 within Task Area 4 ("Data Portal") I pursue the task to conceptually analyze and design FAIR metadata."*

Nicola Malavasi (LMU Munich, nicola.malavasi@physik.lmu.de): *"I am a postdoc in astrophysics studying galaxy evolution, the structures of the cosmic web, and the relation between the two. I obtained my Ph.D. at University of Bologna, then worked as a postdoc in the U.S.A. and France. I am now working as a postdoc at the University Observatory Munich (USM-LMU). I am part of TA3-WP4 and WP1, where I am trying to implement a workflow to detect the cosmic web from galaxy surveys using specific algorithms in a way which is reproducible and accessible."*



Jörn Künsemöller (Bielefeld, jkuensem@physik.uni-bielefeld.de): *"I am a software developer in radio astronomy at the University of Bielefeld with a background of IT in the natural sciences and a Ph.D. on cloud infrastructure. I bring several years of DevOps experience from working on control software and long-term archiving for the International LOFAR Telescope. Within PUNCH, I contribute primarily to TA2, where I work on the storage infrastructure and on connecting radio-astronomical resources to the emerging NFDI."*

David Clarke (Bielefeld, dclarke@physik.uni-bielefeld.de): “I work at Bielefeld University in the area of lattice field theory and the understanding of the QCD phase diagram. I did my Ph.D. at Florida State University. In PUNCH, I am involved mainly in TAs 3 and 4, helping to plan metadata structures for lattice data. In particular, we work on expanding the existing scheme, QCDml, and on bridging this with the overall PUNCH4NFDI metadata structure.”



5. Communication and collaborative tools

For more details on PUNCH4NFDI AAI and collaborative tools efforts, please refer also to the intranet and to the last newsletter.

Since end of 2021 the Indico system at DESY supports login via the Helmholtz-AAI. This enables most users to use the credentials of their home institution to login. Users should register in the PUNCH-AAI, which is actually hosted within the Helmholtz-AAI (details are linked here: https://www.punch4nfdi.de/services/collaborative_tools/). If a user has already a local Indico account, the system usually detects this and joins the AAI login to existing account. In case that procedure fails, one should contact Indico admins via indico@desy.de. In the future PUNCH4NFDI internal agenda pages will only be accessible after login and the present password protection will no longer grant access.

6. Events and upcoming excitements

- 3rd PUNCH4NFDI general meeting: 12 April 2022, 14:00 hrs, <https://indico.desy.de/event/33451/>
- Next PUNCHLunch seminar: I. Kiesel (FIAS): Workflows for efficient pattern recognition under different computing architectures. 24 February 12:30 hrs, <https://indico.desy.de/event/33453/>
- A complete list of PUNCHLunches can be found in the intranet and in INDICO: <https://indico.desy.de/category/743/>). If you have suggestions for the seminar series – let us know at info@punch4nfdi.de.
- TA 5 workshop on concepts for metadata, 21 February 2022, 15:00 hrs, <https://indico.desy.de/event/33215/>
- Workshop on NFDI tools / services / synergies between physics-related consortia and others, 5 April 2022, <https://indico.desy.de/event/33410/>
- For a complete list of TA and other working meetings, see the INDICO category <https://indico.desy.de/category/741/>
- DPG workshop “Forschungsdaten im Physikstudium“, 20-22 April 2022 in Bad Honnef, <https://www.dpg-physik.de/veranstaltungen/2022/workshop-forschungsdaten-im-physikstudium>
- NFDI Infratalk on 2 May 2022: “Data management in the PUNCH sciences” (details coming soon)
- NFDI Tooltalk on 11 May 2022: AAI and collaborative tools in PUNCH4NFDI efforts so far (Harry Enke and Kilian Schwarz, more information coming soon)

7. Recent talks, results, and publications

To be found at least partly on the web page and in ZENODO (<https://zenodo.org> – just search for “PUNCH4NFDI” or – if you are interested in the broader scope – “NFDI”).

At the PUNCH4NFDI General Meeting 18 January 2022 (<https://indico.desy.de/event/32343/>), Jakob Nordin (DESY Zeuthen) and Laura Spitler (MPIfR) focused on the relevance of concepts of data irreversibility that will be further developed in task area 5.

- Jakob Nordin introduced basic workflows in multi-messenger astronomy. Combining a large number of data streams and selecting interesting signals in real-time constitutes a major challenge. In this context, the modular AMPEL platform has been introduced to systematically process a set of heterogeneous datasets. This framework is designed to select, analyze, update, combine, enrich and react to data and can be applied also in fields beyond experimental astrophysics.
- Laura addressed some of the future challenges related to the quasi real-time identification of fast radio bursts (FRBs). The enormous increase of data rates requires the processing of real-time search pipelines to extract interesting signals using classification via machine learning techniques. As an example, Laura pointed out that results from the convolutional neural network FETCH have shown a high efficiency in distinguishing FRBs from non-interesting radio frequency interference backgrounds. Laura: *“Our human experience of the night sky is that it goes through daily and annual rhythms but is otherwise largely static and predicable. In recent decades, astronomers have discovered a dynamic Universe of explosions and time variability on normal human time scales. Modern astronomical surveys are designed with these time scales in mind, which results in a massive increase in data rates, inevitably leading to the situation that only the interesting data can be saved long term. This data irreversibility manifests in several areas of a survey: having to commit to what signals are deemed interesting, designing algorithms that robustly find all of the astrophysical signals while reducing the number of false positives, and an understanding that we cannot re-search the full data looking for entirely new signals.”*

At the same meeting, Yori Fournier (AIP) presented a prototype of a research workflow packager (called kuFAIR as a working title). kuFAIR allows scientists to create, run, and explore a part of their research workflow in various contexts, either locally (on a single machine), or in a distributed environment. The idea of a research workflow package is to contain the instructions to setup the full operational environment necessary to view, access and interact with the scientific data. This allows scientists to share their research with collaborators in a format, that allows the latter to open a fully operational environment in their web browser, without the need to download the data or install the used software components.