

# The PUNCH4NFDI Consortium

## Newsletter Number 4

October 2022



### Contents

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

## 1. News and Highlights

With our **4<sup>th</sup> general PUNCH4NFDI meeting** in Göttingen on 28-29 September we held our first in-person meeting. The schedule included keynote speeches by NFDI colleagues (Regine Stein, SUB Göttingen and member of Text+ and NFDI4Culture, and Prof. Julian Kunkel, University of Göttingen and GWDG), a general overview of the status of PUNCH4NFDI, a poster competition<sup>1</sup>, a PUNCHLunch, and contributions by all task areas through highlight talks and hands-on sessions. For further information see the Indico page<sup>2</sup>.



Figure 1: Picture of all participants at the General Meeting in Göttingen (left) and of the contributed posters (right).

We congratulate the winners of our PUNCH poster competition, Achim Geiser (DESY), Victoria Tokareva (KIT), and the Lattice QCD group (Uni Bielefeld, Uni Bonn, DESY, FZ Jülich and Uni Regensburg). Snapshots of the winning posters are shown in Figure 2.

Attached to the general meeting, the second **Women4PUNCH** meeting took place on 28 September, with more than a dozen female participants. One outcome is the proposal to set up PUNCH guidelines for personnel selection processes in order to promote unbiased hiring situations.

Currently, PUNCH4NFDI members conduct four **user surveys**, and you are kindly invited to participate:

- Task area 3 is performing a survey on the usage of simulation codes in astrophysics<sup>3</sup>;
- task area 6 is performing a survey on data access and data analysis tools<sup>4</sup>;
- task area 7 is performing two surveys: The first on available material and recurrent events for education in data literacy<sup>5</sup>, and the second on organisations

<sup>1</sup><https://indico.desy.de/event/34710/sessions/13822/#20220929>

<sup>2</sup><https://indico.desy.de/event/34710/>

<sup>3</sup><https://go.fzj.de/astro-survey>

<sup>4</sup>[https://docs.google.com/forms/d/e/1FAIpQLSfucyXwVNdXKGGJ9\\_Zk1V7xrV0aDJr-zhDm1ujPvcY6tGgToA/viewform](https://docs.google.com/forms/d/e/1FAIpQLSfucyXwVNdXKGGJ9_Zk1V7xrV0aDJr-zhDm1ujPvcY6tGgToA/viewform)

<sup>5</sup> <https://umfrage.dpg-physik.de/index.php?r=survey/index&sid=744358&lang=de>

or projects providing resources for the education in data literacy<sup>6</sup>. Both of these surveys are in German only.

## CMS + ATLAS Higgs->4Lepton demonstrator project for PUNCH platform

A. Geiser, DESY, 29.9.22, for the Higgs->4Lepton PUNCH demonstrator project team

**Main purpose: demonstrate practical feasibility of a PUNCH use case on the PUNCH Science Data Platform going significantly beyond what is already available outside PUNCH** (i.e. not just an import of things already available elsewhere), using PUNCH resources already now wherever possible.

PUNCH goal stated last year: **"to set up a working prototype within the first year"**

**Transformation of data from different projects/sources to common analysis data format, TA4/WP3.**  
(Current prototype still limited to HEP as starting point, extension to other PUNCH4NFDI communities conceptually started).


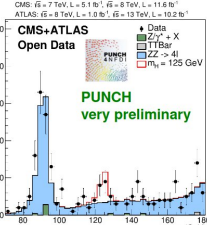
**Representation of the corresponding metadata in a common PUNCH scheme based on XML and datatec, TA4/WP2.**  
(Current prototype: very preliminary practical starting point, including visualization. Details to be discussed further).

**Actual usage of storage4punch resources, TA2/WP1,** (including test of the corresponding access procedures, TA4/WP3).

**Usage of compute4punch resources, TA2/WP2** (soon) (including test of the corresponding access procedures, TA4/WP3).

**Usage of portal resources for scripts and documentation, TA4/WP4** (including e.g. "automatic" transformation Twikis -> Markdown, TA4/WP3).

**Setup of corresponding Research Products, TA4/WP1** (soon) for PUNCH derived data sets and entire example (on the platform, TA4/WP4).

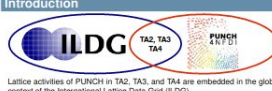
29.09.22

A. Geiser, Poster PUNCH general

## Lattice QCD Data in PUNCH

G. Ball, B. Bhattacharya, S. Gupta, D. Clark, S. Collins, O. Kaczmarek, F. Karsch, S. Kniehl, K. Meyer, C. Schmidt, J. Simeth, K. Simma, C. Urbach, T. Weing, U. Bielefeld, Uni Bonn, DESY, FZ Jülich, Uni Regensburg


### Introduction



Lattice activities of PUNCH in TA2, TA3, and TA4 are embedded in the global context of the International Lattice Data Grid (ILDG).

### What is ILDG?

ILDG is a federation of autonomous Regional Grids (RG) with agreed-upon standards for data and metadata.



Services operated by ILDG:

- Membership registration of Virtual Organization (VOMS) groups and roles of ILDG users for authentication to services, e.g. storage
- Website: specification of standards and conventions, URLs of services of each RG

Services operated by each RG:

- Metadata Catalogue
- File Catalogue
- Storage Elements
- Website (RG-specific information)

Organization: Board, Metadata Working Group, Middleware Working Group

### The FAIR Principles

FAIR (Findable, Accessible, Interoperable, Reusable) provides guiding principles for data management.

**Findable:**

- F1: globally unique and persistent ID assigned to IM/D
- F2: data described with rich MD
- F3: MD includes data ID of data
- F4: MD registered or indexed in a searchable resource

**Accessible:**

- A1: (MD) retrievable by ID using standardized protocols
- A1.1: protocol is open, free, and universally implementable
- A1.2: protocol allows AA procedure where necessary
- A2: MD accessible even if data is no longer available

**Interoperable:**

- I1: (MD) use a formal, accessible, shared, and broadly applicable language
- I2: (MD) use vocabularies that follow FAIR principles
- I3: (MD) include qualified references to other (MD)

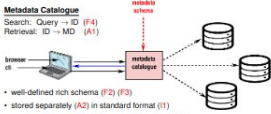
**Reusable:**

- R1: (MD) richly described with plurality of accurate and relevant attributes
- R1.1: (MD) released with clear and accessible data usage license
- R1.2: (MD) associated with detailed provenance
- R1.3: (MD) meet domain-relevant community standards

### ILDG and PUNCH go FAIR

**Metadata Catalogue**

Search: Query -> ID (F4)  
Retrieval: ID -> MD (A1)



**Well-defined rich schema (F2) (F3)**

- stored separately (A2) in standard format (I1)
- searchable in central catalogue with standard API

**Unique Identifiers (F1)**

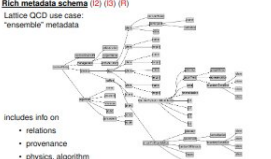
entity ID metadata includes information on relations content data storage

data objects	1:1	unique	content	data storage
data sets	1:1	unique	content	data storage
data publication	1:1	set of MD	content	data storage

Lattice QCD use case: data object = "configuration", data set = "ensemble"  
I1 = logical file name, MD = "Markov Chain UPT"

**Rich metadata schema (I2) (I3) (F1)**

Lattice QCD use case: "ensemble" metadata



includes info on:

- relations
- provenance
- physics, algorithm

### Plans for ILDG 2.0

Building on basic concepts of ILDG 1.0 (well-defined metadata schema, interoperable services, ...)

- make ILDG fully FAIR-compliant by supporting also data publishing process
  - registration of persistent identifier (DOI)
  - landing page (hosting and automatic generation)
    - harvesting of metadata (DOI-PURL)
- upgrade to modern technologies, in particular:
  - new metadata catalogue implementation with
    - support for multiple generic MD schemas
    - fine-grained token-based access control
    - REST API and user-friendly client tools
  - easy deployment for different PUNCH applications

## Metadata curation use cases in astroparticle physics\*

Victoria Tokareva\*\* for PUNCH4NFDI Consortium

Karlsruhe Institute of Technology PUNCH4NFDI in-person annual meeting & 4th PUNCH4NFDI General Meeting | Göttingen | 28-29 Sep 2022

Defining metadata for different physics or data science use cases is a key challenge in the era of big data. This poster presents a use case for the PUNCH4NFDI consortium in astroparticle physics. The use case is based on the PUNCH4NFDI consortium's data and metadata management strategy. The use case is based on the PUNCH4NFDI consortium's data and metadata management strategy. The use case is based on the PUNCH4NFDI consortium's data and metadata management strategy.

**Use case studies for NFDI**

- Commonly used in astroparticle physics
- Analysis of large distributed datasets
- Complex metadata structures
- Complex metadata structures

**Storage architectures**

GCRC, GRAALCI, etc.

**Comparative analysis of the usecases**

Use Case	Key Features	Challenges	Technologies
GCRC	Position of the tree, uniform, platform-agnostic, etc.	Data analysis, visualization, etc.	GCRC, GRAALCI, etc.
GRAALCI	Open access, platform-agnostic, etc.	Data analysis, visualization, etc.	GRAALCI, etc.

\* This work was partially supported by DFG fund "NFDI 381" for the PUNCH4NFDI consortium.  
\*\* Karlsruhe Institute of Technology, Institut für Astrobleichphysik (IAP) <https://www.iap.kit.edu>

Figure 2: Snapshots of the winning posters of the PUNCH poster competition. Top left: Achim Geiser (DESY), Bottom left: Victoria Tokareva (KIT). Right: Lattice QCD group (Uni Bielefeld, Uni Bonn, DESY, FZ Jülich, Uni Regensburg).

The PUNCH4NFDI consortium recently published its first flyer which is shown in Figure 3.

In case you have news that you want to share with PUNCH4NFDI, don't hesitate to send a mail to [info@punch4nfdi.de](mailto:info@punch4nfdi.de). You can access all newsletters on our web site<sup>7</sup>.

<sup>6</sup> <https://umfrage.dpg-physik.de/index.php?r=survey/index&sid=994277&lang=de>  
<sup>7</sup> [https://www.punch4nfdi.de/news\\_amp\\_events/newsletter/](https://www.punch4nfdi.de/news_amp_events/newsletter/)

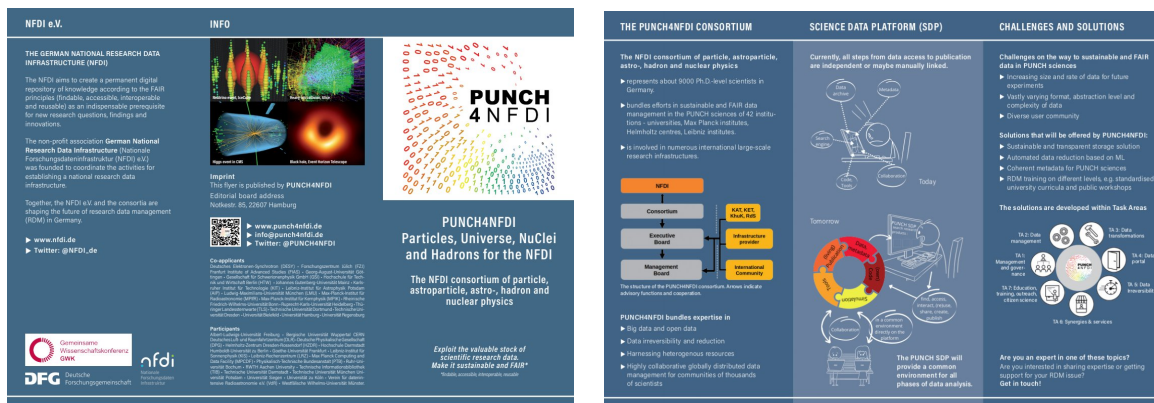


Figure 3: The new PUNCH4NFDI flyer.

## 2. NFDI and related topics

**Base4NFDI:** The proposal for Base4NFDI received a funding recommendation from the DFG. A Base4NFDI office has been set up at Goettingen, and the Base4NFDI team (consisting of designated task area leaders and members of the editorial team for the Base4NFDI proposal) is very active and meeting regularly in order to be prepared for a positive funding recommendation, expected for 4 November.

Two aspects are particularly relevant for PUNCH4NFDI:

- Base4NFDI – if funded – will open calls for basic service (initialisation, design, implementation, integration). The first of these calls is supposed to open already in February 2023. PUNCH4NFDI will discuss potential contributions to such basic services, and ideas are welcome!
- In order to evaluate these basic services proposals, a Base4NFDI Technical Expert Committee (TEC) will be installed, consisting of politically well connected technical experts. The selection of these experts is done by the NFDI Consortia Assembly, and care has to be taken that the interests of PUNCH4NFDI and the physics-related consortia are regarded in the composition of the committee.

**NFDI sections:** The work in all sections and task forces is progressing.

- The section “Training & Education” was especially engaged in the support of proposals for the BMBF call “Datenkompetenzzentrum in der Wissenschaft”.
- The task fore “Monitoring” will soon publish a white paper to provide a common basis for reporting and reference regarding some questions of cross-consortia relevance in the template for the interim reports of the DFG.
- A new NFDI section – Industry engagement – is about to be established.

In general, many meetings and contributions to conferences and other events are organised by NFDI sections or NFDI colleagues who ask for contributions by all consortia. In order to keep track of all events and to spread the news in PUNCH4NFDI internally, we started recently to collect these events in the PUNCH intranet<sup>8</sup>.

<sup>8</sup><https://intra.punch4nfdi.de/?md=/docs/NFDI/schedule.md>

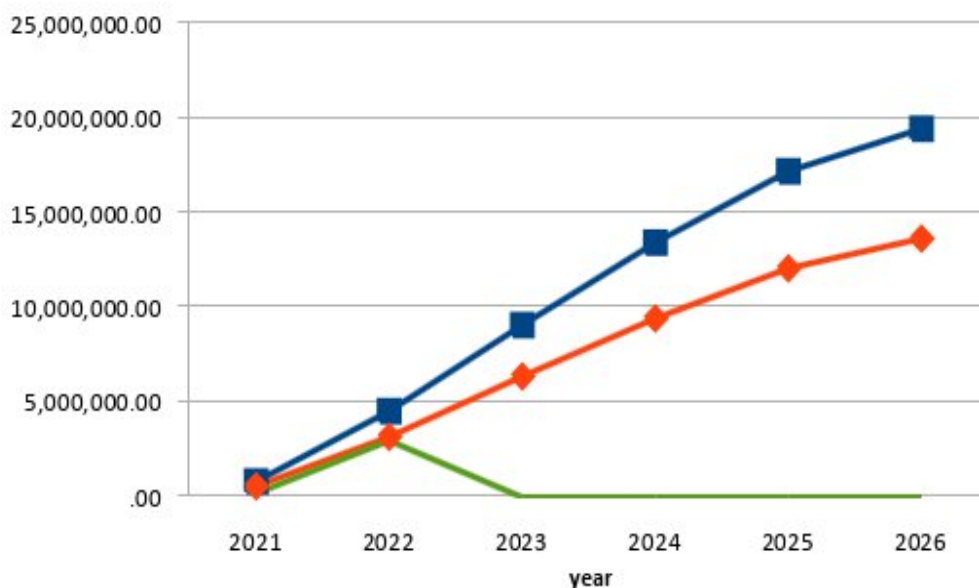


Figure 4: Spending profile (accumulated) of the PUNCH4NFDI funds. Blue - proposal; orange - DFG grant; green - actual spending.

### 3. Formal, legal and financial topics

Figure 4 above shows the financial situation of the PUNCH4NFDI Consortium as a function of time, completely comprising the years 2021 and 2022. By filling open positions, and through the redistribution of funds between partners, we were able to compensate the underspending still observed in the first half of 2022.

The PUNCH4NFDI **Executive Board** welcomes Olaf Kaczmarek as new representative for the Committee for Hadron and Nuclear Physics (KHuK). Olaf works in the field of strongly interacting matter under extreme conditions with lattice methods at the University Bielefeld. He succeeds Kilian Schwarz, whom the Executive Board warmly thank for his tireless work and valuable contributions in the last years.

Changes in lead positions on the working level comprise the handover from Marcus Brüggem (Hamburg University) to Joseph Mohr (LMU) in task area 3 and from Jim Hinton to Makarim Bouyahiaoui (both MPIfK) in work package 4 of task area 6.

### 4. Status of the consortium

#### Reporting in PUNCH4NFDI

The quarterly reports are a permanent setup now. By end of October, the third quarterly report will be published. A special significance have the statements from our boards – SAB, IRB and UC – as kind of external feedback from experts.

## 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!



**Johanna Rätz (University of Bonn, TA7):** I studied math and physics to become a teacher at the University of Bonn and recently submitted my master thesis about a new teaching concept on modeling and simulation as scientific ways of gaining knowledge. Since April I am working at PUNCH on WP7.3 "Public Outreach" and will focus on school-related public outreach.

**Tim Oelkers (HTW, started mid-May, TA5):** I am a master student at HTW Berlin and study Applied Computer Science. I participate as a student assistant in the research project „Memory-centric Architectures". The project has access to a corresponding prototype at HP Labs. Within TA 5 (WP 4), I will analyze the scaling and performance behavior of Big Data frameworks on this new computing infrastructure.



**Carsten Burgard (TU Dortmund, TA7):** I have a PhD in Physics, received from the University of Freiburg. I'm with the TU Dortmund right now, working with Prof. Kevin Kröninger on Data analysis for ATLAS, focusing on topics of combinations of differential cross-section measurements of Higgs, Standard Model and Top Physics in the context of interpretations in terms of Effective Field Theories. In the context of PUNCH4NFDI, I am part of WP 7.2, focusing on ways of improving the teaching of data science as a part of university physics curricula, giving particular weight to the FAIR principles.

Contact: [carsten.burgard@uni-dortmund.de](mailto:carsten.burgard@uni-dortmund.de)



**Christiane Schneide (DESY, TA1):** received her Ph.D. in astrophysics in 2017 from Hamburg Observatory. After a postdoc position at Leuphana University Lüneburg, she started as project manager for PUNCH4NFDI in June. In this position, she is responsible for the internal and external communication of the consortium, the organisation of meetings, the reporting and cooperation with other consortia and NFDI units.

Contact: [christiane.schneide@desy.de](mailto:christiane.schneide@desy.de)

**Benoit Roland (KIT, TA2):** I started to work at KIT in May 2022, at the Steinbuch Computing Center (SCC), in the Matterminers group of Manuel Giffels. I am involved in the Compute4Punch project, working on the integration of compute resources into a common infrastructure, and providing suitable access points to the community. My activities have been focused until now on the development and deployment of Compute4Punch login nodes whose access is granted to via OIDC token. To illustrate the use of the Compute4Punch storage and compute resources, I have implemented in our framework the CERN Open Data Higgs to 4 leptons CMS analysis<sup>9</sup> developed by Achim Geiser et al. at DESY. My contributions have been realized in close collaboration with Manuel Giffels and with the continuous support of various SCC members.

## 5. Communication and collaborative tools

The management and task area 2 are currently testing OpenProject as project management tool for PUNCH4NFDI. Possible implementations are an individual instance for our consortium or an attachment to the NFDI instance that is used by various other consortia.

In collaboration with all task areas, the consortium is currently setting up a page for the publication of first (preliminary) results and tools available for the public. The release will be announced on our website<sup>10</sup>.

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

## 6. Upcoming events and excitements

- In collaboration with the WissKon-Netzwerk, task area 7 organises a monthly Journal Club on science communication (in german). Next date is on **26 October** at 4 p.m.<sup>11</sup>
- The 5<sup>th</sup> general PUNCH4NFDI meeting will take place on **17 January 2023** as a purely virtual meeting – save the date!
- 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](mailto:info@punch4nfdi.de).
- For a complete list of TA and other working meetings, see the Indico category <https://indico.desy.de/category/741/>
- NFDI InfraTalk<sup>12</sup> and ToolTalks<sup>13</sup> are now both on Mondays, 4-5 p.m.<sup>14</sup>.

<sup>9</sup><https://opend0ata.cern.ch/record/550>

<sup>10</sup><http://www.punch4nfdi.de/>

<sup>11</sup><https://www.nawik.de/wisskon/veranstaltungen/journal-club-okt22/>

<sup>12</sup><https://www.youtube.com/playlist?list=PL08nwOdK76QlnmEB659qokiWN3AC-kqFS>

<sup>13</sup><https://www.youtube.com/playlist?list=PL08nwOdK76Qm5QqvehnECo477Llfe6Pk>

<sup>14</sup><https://www.nfdi.de/termine/>

## 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”).

- [Annual Meeting of the Astronomische Gesellschaft – Bremen, 12-16 September 2022](#): Presentations on PUNCH matters were given by Thomas Schörner, Harry Enke, and Kilian Schwarz. In addition, a poster booth was staffed mainly by TA6<sup>15</sup>.



Figure 5: Pictures of the splinter meeting and posters at the annual meeting of the Astronomische Gesellschaft.

- [The 39<sup>th</sup> International Symposium on Lattice Field Theory \(Lattice 2022\) – Bonn, 8-13 August 2022](#): A presentation was given by Frithjof Karsch<sup>16</sup> and two posters were presented by Olaf Kaczmarek<sup>17</sup> and Nils Meyer, Daniel Richtmann and Tilo Wettig<sup>18</sup>.
- Several presentations and posters were presented at our General Meeting. Please refer to the Indico page<sup>19</sup> of the event to access the files.

<sup>15</sup>[https://www.punch4nfdi.de/sites/sites\\_custom/site\\_punch4nfdi/content/e144069/e144104/e173878/2022\\_0915.PUNCH4NFDI.AG2022.TS.pdf?preview=preview](https://www.punch4nfdi.de/sites/sites_custom/site_punch4nfdi/content/e144069/e144104/e173878/2022_0915.PUNCH4NFDI.AG2022.TS.pdf?preview=preview)

<sup>16</sup><https://indico.hiskp.uni-bonn.de/event/40/contributions/852/>

<sup>17</sup>[https://indico.hiskp.uni-bonn.de/event/40/contributions/701/attachments/432/712/SIMULATeQCD\\_Lattice\\_2022.pdf](https://indico.hiskp.uni-bonn.de/event/40/contributions/701/attachments/432/712/SIMULATeQCD_Lattice_2022.pdf)

<sup>18</sup>[https://indico.hiskp.uni-bonn.de/event/40/contributions/698/attachments/438/723/poster\\_pdflatex.pdf](https://indico.hiskp.uni-bonn.de/event/40/contributions/698/attachments/438/723/poster_pdflatex.pdf)

<sup>19</sup><https://indico.desy.de/event/34710/>