

# Bringing the Research Product to Life: The PUNCH4NFDI Science Data Platform ...

... the central deliverable of the PUNCH4NFDI Consortium – an ecosystem for the entire lifecycle of digital research products, a FAIR and federated infrastructure for access to data and to the resources for analysing, transforming and exploiting them.

A scientist pursuing an exciting research question or wanting to repeat or improve an existing workflow



The authentication and authorisation infrastructure PUNCH-AAI is built on the Helmholtz AAI. It ensures authentication of users (usually via email) and grants authorisation based on suitably designed groups and roles. Services that are so far connected to the AAI are the PUNCH intranet, the DESY Sync&share cloud environment, the Gitlab infrastructure, and the access to the Compute4Punch and Storage4PUNCH facilities.



Containers are our way of organising the job traffic, i.e. the workflows in which research questions are tackled and data analyses are performed. They offer a means of virtualisation and abstraction from the concrete computing environment and are the tool of choice for federated heterogeneous infrastructures.

With e.g. SciTrace, first attempts are successfully made to capture workflows on "research products" using containers.



Digital research products (DRPs) contain the components the scientist wants to work:

- pointers to the data
- Workflow descriptions used to obtain the data and for the desired transformations
- the necessary transformation code
- Links to publications
- Description of the interfaces to required infrastructures
- ...

DRPs are living objects that contain their own history.



The stuff you need in order to do your job as a scientist:

- Data & code registry
- Transformations and algorithms
- Container registry

Organised via **Gitlab**



Above sea level: User control and user choices



Below sea level: technology implementations and choices transparent to the user



Data – old and new – are treasures. They are in constant danger of being lost or becoming useless. Making data FAIR – findable, accessible, interoperable and reusable is the central objective of the NFDI and of PUNCH4NFDI.


In the process of their work, scientists derive new from old data and enrich them – or the digital research products the core of which they form.



We are your useful minions, toiling away invisibly for you: as workflow managers like REANA or others, as data providers, as technology choosers, or in the Compute4PUNCH and Storage4PUNCH facilities.



The computing facility Compute4PUNCH is a federated infrastructure that transparently integrates heterogeneous resources like grid centres, HPC and HTC centres, etc.



The Storage4PUNCH facilities provide caching and persistent storage services for the workflows carried out on the platform.



You could think of several different container technologies and workflow management systems – the platform will deal with them all transparently!



They call me data loss and negligence. "Ich bin der Geist, der stets verneint."



Making the results of our work persistent is an important step ... to quote a slightly modified yet still important contribution to art history: "Every bit is sacred ...". In the end, new data and new research products need to be made known to our **Gitlab** registry in an automatised way.



The PUNCH4NFDI Consortium

Mail: [info@punch4nfdi.de](mailto:info@punch4nfdi.de)

Twitter: [@PUNCH4NFDI](https://twitter.com/PUNCH4NFDI)

Web: [www.punch4nfdi.de](http://www.punch4nfdi.de)

