

TA4: The PUNCH Portal and the (Dynamic) Research Product

PUNCH Kickoff Workshop, 15 October 2021

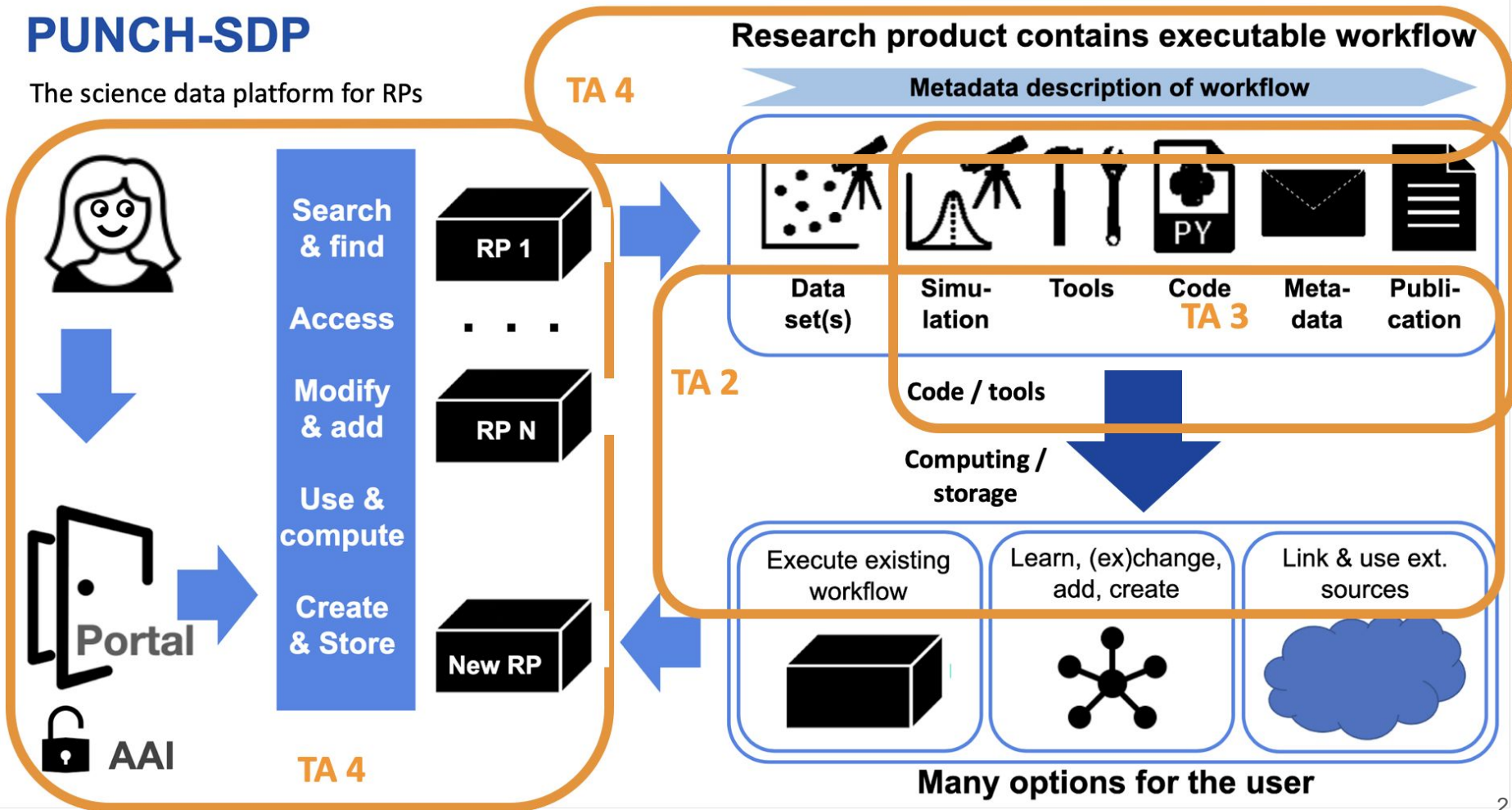
Particles, Universe, NuClei and Hadrons for the NFDI

P. Bechtle (U Bonn), Harry Enke (AIP)
for the whole of Task Area 4



PUNCH-SDP

The science data platform for RPs



TA4 : Data portal and Research Product

The main goals of TA4 are

- the PUNCH4NFDI data portal as central element (hub) and main external access point for the PUNCH and broader communities to the PUNCH4NFDI science data platform services
- the creation of the Digital / Dynamic Research Product (DRP), combining the available advanced developments in our communities for data, software, workflows, pipelining

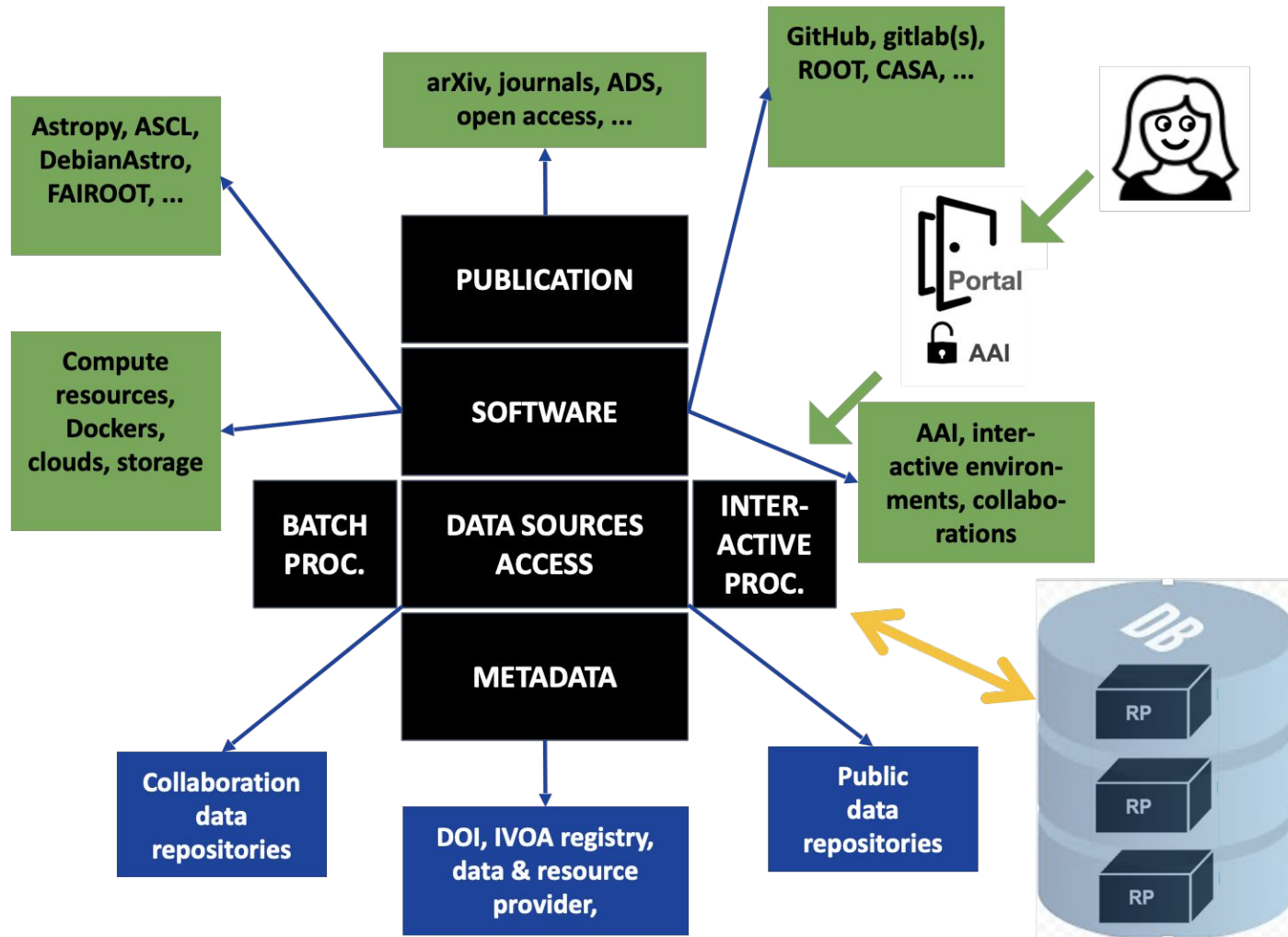
TA4-WP1 Digital/Dynamic research products and their catalogue(s)

TA4-WP2 Mapping and collating metadata

TA4-WP3 Implementation of interfaces (API) to data services and infrastructures

TA4-WP4 Build and operate the data portal with appropriate microservices

Digital Research Product



- Improves reproducibility and re-usability
- guided by FAIR
- Access via portal
- Interaction with other RPs
- Interfaces to tools and infrastructures
- Built on available developments of the communities our

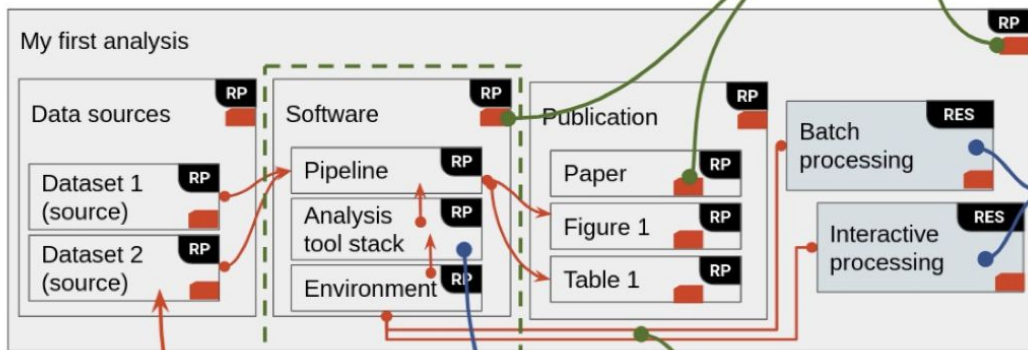
Handshakes

WP1: Domain independent structure

- basic properties and typology of research products
- composition and unpacking

WP2: Metadata

- bridges to connect to domain specifics



WP4: Portal

- Operate (distributed) DB servers
- Web-Server, DB-UI

WP3: Interfaces

TA3: Analysis tools

- Provide tools
- descriptive metadata

TA2: Infrastructure

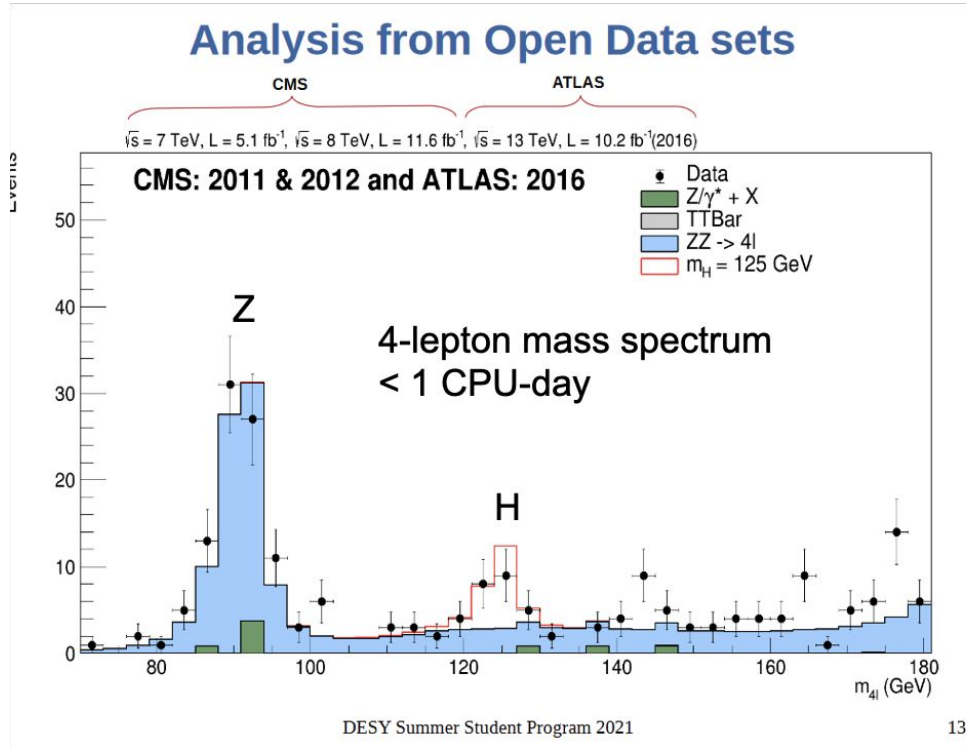
- compute elements
- Virtual machines for DB servers?

TA2

TA4 WP3

Teaser:

- Convert available CMS, ATLAS and ZEUS data (others to follow) to common analysis format (WP3-3) (many CPU-months, make output available to users)
 - -> analyze them with the **same** Root or Python script on your laptop
- (remote data access; Jupyter notebook possible, (not yet))

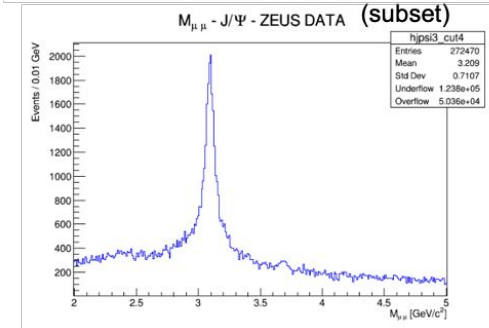
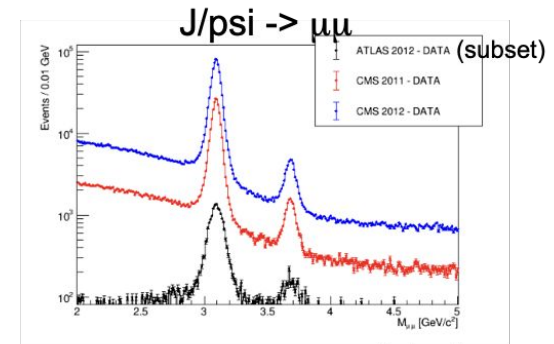


21.9.21

A. Geiser, DPOA meeting

DESY summer student projects 2021:

A. Bal, A. Geiser, L. Olivi,
R. Schwenzer, M. Velasco, Y. Yang



TA4 Information and conclusion

- Information about organization, ongoing work, meetings etc:

[TA4 Intranet](#)

<https://intra.punch4nfdi.de/?md=/docs/TA4/overview.md>

- Regular Meetings:

Every Monday 15:00 <https://indico.desy.de/category/741/>

- We are looking forward to enhancing interaction and cooperation,
- Everybody is welcome!

TA4 WP1: The Dynamic Research Product

– **D-TA4-WP1-1 (31 Dec 2022): Prototype** Design and implementation of the PUNCH DRP database on the infrastructure provided by TA 2 (*we might need stage the roll out: start with simple DB server first*).

Trans-domain description of RPs

– **D-TA4-WP1-2 (31 Dec 2023)** : Dynamic ingestion and curation processes of **selected** existing RPs from different PUNCH subcommunities; demonstration and routine processing.

– **D-TA4-WP1-3 (31 Dec 2024)**: Enrichment of DRPs with descriptive metadata; demonstrator project to allow user access to SFB 1245 and astroparticle legacy data.

– **D-TA4-WP1-4 (30 Sep 2026)**: Definition of the functionality of DRPs and their interfaces; prototype demonstrators and consistent integration for running them on the PUNCH-SDP; continuously integrate needs of PUNCH community.

Prototyping of new catalogue+DRP approach vs routine adaption of science data portal

TA4 WP2: Metadata

Next steps:

- Compile list and definitions of available metadata
- Define prototype metadata schema for parts of the communities not yet using large-scale established metadata schemes
- Organize different layers of metadata – from the data types upward up to the interface to the DRP workflow
- track data provenance
- Harvesting of Metadata: synergies with INSPIRE/HEPDATA activities, esp. via DESY

Deliverables:

- D-TA4-WP2-1 – year 2 Provide prototype metadata schema and prototype data format.
- D-TA4-WP2-2 – year 4 Define metadata layers and minimal extensions for existing metadata standards.
- D-TA4-WP2-3 – year 5 Define cross-layer connections for transformations of digital objects.

TA4 WP3: Interfaces

Goals:

- Implementation of interfaces for the functionality of the PUNCH science data portal respecting the FAIR principles.
- Interfaces for access management to the central platform, its decentralized branches and external resources linked into the platform (WP3-1/2);
- interfaces to the content and format of initially heterogeneous data both inside and outside the platform, enabling their homogeneous treatment with common tools provided by other task areas and work packages (e.g. TA3, other TA4 WPs).

Formal Deliverables

- D-TA4-WP3-1 — 221009 - 260930: Technical interfaces to external resources
- D-TA4-WP3-2 — 221009 - 260930: Integration of platform services and interfaces
- D-TA4-WP3-3 — 221009 - 260930: Interfaces allowing combined analysis of data sets from different sources and experiments

Next Steps:

- start working on integration with WP4 (portal), WP2 (related metadata), extend to other subcommunities
- develop towards demonstrator project(s) for PUNCH4NFDI platform within first year, ideally with publishable physics result; extend to other subcommunities.

TA4-WP4: Build and operate the data portal with appropriate microservices

Goals:

- integration into PUNCH AAI
- instantiation of the Science Data Portal as access point for users of the PUNCH4NFDI services and facilities
- working environment for DRPs
- uploading own data for combining with other data collections
- searching for and selecting from data collections with certain physical properties

Deliverables:

- D-TA4-WP4-1 —by 30.06.2023: Working prototype portal (web interface) to access and use the platform
- D-TA4-WP4-2 —by 30.06.2024: Published Research Product examples with stored data and interoperable analysis workflows
- D-TA4-WP4-3 — by 30.06.2026: Feature complete portal service
- D-TA4-WP4-4 —by 30.06.2026: Published and interoperable Digital Research Products using the full range of services, including combined analysis of data sets from different sources and experiments

