Status PUNCH4NFDI TA6

Kilian Schwarz

TA6 WP1 - Deliverables

- D-TA6-WP1-1 (01 Jan 2022):
  - Marketplace noticeboard: being set up at DESY
- D-TA6-WP1-2 (01 Jan 2023): physics.tools
  - Exchange platform for archives, software, services: done and published
- D-TA6-WP1-3 (01 Jan 2023):
  - Synergies on cross cutting topics with NFDI: done and published
- D-TA6-WP1-4 (continuous):
  - Services and tools to be made available
- D-TA6-WP1-5 (continuous):
  - Collaboration with NFDI

TA6 WP1 – updates and highlights

Physics.tools, a search engine for software referenced in publications

Now also part of the **Base4NFDI** proposal, in the **nfdi.software** group

➔**Working to host the service on a machine in Heidelberg University**

- The software database:
  - physics.tools feeds on a database containing software extracted from papers on arxiv
  - ➔ **Extended the database from 2020-2022 to 2010-2022**
  - ➔ **Working to extend with the missing years and software from private repositories**
  - ➔ **Working to extend the database with referenced data, datasets, catalogues, etc.**
- The search engine:
  - Works if you know the name of the software you are looking for
  - We want it to work with keywords, too
  - ➔ **Implementing a Retrieval Augmented Generation-like pipeline, interfacing with an LLM from Mistral AI**
  - ➔ **Working prototype to be deployed and discussed with the nfdi.software group**

TA6 WP2 - Deliverables

- D-TA6-WP2-1
  - Prototype PUNCH AAI (31 Mar 2022): done
  - Basic PUNCH AAI (31 Dec 2023): Unity dev requests #1, #2, #3
  - Extended PUNCH AAI (30 Sep 2026)
- D-TA6-WP2-2
  - Coordination with NFDI, national, international stake holders
    - Draft design (31 Dec 2022): done, published
    - Complete design (31 Dec 2024): done, published
- D-TA6-WP2-3
  - PUNCH AAI group management
    - Prototype (31 Dec 2022): done
    - Revised version (31 Dec 2024): in progress
    - Full group management (30 Sep 2026)

TA6 WP2 – updates and highlights

- Documentation see PUNCH AAI requirements document
- Established direct contact to Unity in close collaboration with IAM4NFDI and FZJ
  - Group information has been added to Tokens
  - Filter groups embedded in tokens in progress
  - Offer for granular authorisations still pending
- Successful development work: Indico access rights can now be controlled based on groups in PUNCH AAI

TA6 WP3 - Deliverables

- D-TA6-WP3-1 (31 Jul 2022)
  - Reference guide on publishing data: living document
- D-TA6-WP3-2 (31 Dec 2022)
  - Reference guide for publishing software: living document
- D-TA6-WP3-3 (31 Dec 2023)
  - (Dynamic) Metadata frameworks for PUNCH-SDP: declared finished
  - Document (10 pages) by Tim Oelkers
- D-TA6-WP3-4 (31 Mar 2024)
  - Effelsberg data: data need to be converted, ongoing
  - Updated completion date: 31 Dec 2025
- D-TA6-WP3-5 (31 Dec 2024)
  - Converter for FITS/ROOT formats: done, being published
- D-TA6-WP3-6 (31 Dec 2025)
  - Metadata extensions: on track

TA6 WP3 – updates and highlights

TA6 WP4 - Deliverables

- D-TA6-WP4-1: survey of PUNCH tools
  - Initial overview (31 Jul 2022): done, being published
  - Final list (30 Jun 2026)
- D-TA6-WP4-2 (30 Jan 2023)
  - Reference repository with CI: done, being published
- D-TA6-WP4-4 (30 Jun 2026)
  - Data analysis examples
- D-TA6-WP4-5 (30 Jul 2023)
  - Software platform: done, merged with Physics tool

TA6 WP4 – updates and highlights

See highlights WP1

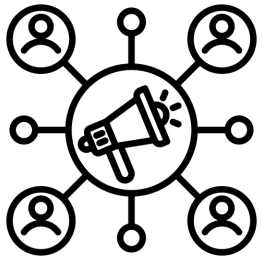
TA6 WP5 - Deliverables

- D-TA6-WP5-2 (30 Sep 2022)
  - Dynamic disk cache: done and published
- D-TA6-WP5-3 (30 Dec 2025)
  - Memory based computing
- D-TA6-WP5-4 (31 Dec 2025)
  - Interfaces to supercomputer, GPU, GoeGrid
- D-TA6-WP5-5 (31 Dec 2024)
  - COBalD/TARDIS: done, being published
- D-TA6-WP5-6 (30 Sep 2022)
  - MultiCloud resources: done, being published
- D-TA6-WP5-7 (30 Sep 2026)
  - Standard analysis software in JupyterHub
- D-TA6-WP5-8 (31 Dec 2023)
  - Services via API: ➔ changed to continuous
- D-TA6-WP5-10 (30 Jun 2024)
  - FTS and Rucio: in progress

TA6 WP5 – updates and highlights

See following slides

From Development to Production: **EXPLORE** @ GAU Göttingen for CERN Open Data



Aim:

Provide access to **GoeGrid resources for CERN Open Data analysis** to users **without CERN/university affiliation** (TA6-WP5-4).

Technical Overview: will be shared in the **TA2 status report** later today.

Key Operational Points:

- **Service Testing & Optimization:**
  After **alpha/beta testing (12 testers)**, including with HEP Masterclass students, the service is **now fully optimized** for scalable, reproducible CERN Open Data analysis.
- **User Authentication & Registration:**
  Custom registration system for independent authentication. Users register with a valid email and SSH key pair—**no third-party identity provider needed.**
- **Dedicated Resources & Access:**
  CERN Open Data analysis resources provided to **Public** via the **University of Göttingen**
- **Physics Analysis & Tutorials:**
  Ready-to-use tutorials for HZZ, TTbar, and Hyy analyses. User's guide through job execution, result generation, and data visualization. Hosted in a public PUNCH GitLab repository for easy access
- **Register at**: Register to EXPLORE

EXPLORE *supports researchers, teachers, students, and HEP enthusiasts with resources for scalable, reproducible analysis of CERN Open Data.*

EXPLORE Service @ GAU Göttingen
Promoting!

ATLAS Week & Open Data Weekly Meetings

- **ATLAS Open Data Weekly Meetings:** Regular Updates
  - Contribution on 18.07.2024
  - Contribution on 12.12.2024
- **ATLAS Week Outreach:** The service was presented during the **ATLAS Week Outreach parallel session** *(Speaker: Miguel Ángel García Ruíz on behalf of the ATLAS Open Data team)*.
- ATLAS WEEK Outreach Parallel Session February 19, 2025

Advertising & Outreach

- **GAU Newsletters:** Promoting the service within academic networks (**Approximately 2,000 recipients**.).
- **Email/Letter Campaign:** Targeting **Lower Saxony High Schools (# 59 Gymnasien**) to expand accessibility.