Punch 2.0 use case: LHC open data

Arnulf Quadt

LHC open data

FSP ATLAS
Erforschung von
Universum und Materie

The FAIR principles imply that interested users outside the PUNCH or science community , e.g. high school students or teachers, can actually analyse (LHC) open data. This requires substantial compute resources, which PUNCH1.0 now provides (via the "EXPLORE" service). Further work is needed to continue providing such compute resources and services beyond 2030.

PUNCH2.0 could transfer its present "EXPLORE" service, which is based on university based WLCG resources, to NHR and continue the offer, hence ensuring the realisation of FAIR principles including a complete example analysis infrastructure.

PUNCH uses external (LHC open) data from large data lakes/mass storage facilities. Such a compute model can be realised for future services on NHR as a contribution to federated infrastructures, and requires some further development work along with user support during a subsequent service stage.

LHC open data

FSP ATLAS
Erforschung von
Universum und Materie

Institute: U Goettingen

Topics for PUNCH2.0 SDP+DRP:

Integration of Compute Resources for (LHC) Open Data Processing
This proposal aims to integrate computing resources for processing/analyzing LHC Open Data. This includes:
- Analysis frameworks and extensive documentation.
- Implementation of FAIR (Findable, Accessible, Interoperable, Reusable) principles to ensure sustainable and reproducible workflows.
- Providing a helpdesk for user support.

What type of user is served?

Mostly beginner and interested users outside the PUNCH/science community profile.
The project benefits primarily:
- Beginners and early-career researchers interested in LHC Open Data analysis.
- Users outside the traditional PUNCH and physics communities who wish to engage in data-driven research.
- Educators and students require accessible computational resources for training and research.

How does the suggestion contribute to the PUNCH infrastructure / services portfolio?

- Strengthens the service aspect
- Can utilize the existing contacts between PUNCH and NHR
- Strengthen the implementation of FAIR principles
- Experience in user support and running help desks
- Enhances accessibility to PUNCH infrastructure's compute resources for LHC Open Data analysis.

Does it build on/makes use of existing PUNCH infrastructure or not
- It does as a compute resources service "EXPLORE" has just been finished for (LHC) open data analysis using university based WLCG resources. However, they will fade out in 2030.
- This project ensures continuity and future scalability by transitioning to sustainable infrastructure solutions, i.e. NHR.

LHC open data

FSP ATLAS
Erforschung von
Universum und Materie

How does it integrate into SDP or DRP?

- It would extend the already existing individual parts to a full guided tour of the user
- It supports structured access to compute resources, frameworks, and documentation, forming a coherent and user-friendly workflow.

Does it build/extend community based infrastructure?

Yes.

How is the suggested service sustainable (extensible beyond PII?)

- As long as federal funding agencies continue funding NHR resources, the PUNCH consortium and here the users interested in (LHC) open data analysis will be able to use such services and users support.
- The PUNCH consortium will maintain its commitment to providing computational services for Open Data analysis.
- The service model can be adapted to future computational infrastructures and emerging research needs.

Ten questions

FSP ATLAS
Erforschung von
Universum und Materie

Which elements do they involve? Connection to TA / pillars and requirements on them

→ requires Punch to connect to NHR ressources (technical and administrative questions to be addressed/settleed/solved)
→ connects to "federated instrastructure", can connect to "SDP+DRP"

Which problems do they solve?

→ integration to NHR ressources is very attractive for "federated infrastructure"
→ strengthens Punch2.0 application
→ establishes FAIR principles in reality (CERN does not do this!)

Which gaps do they fill?
→ "federated infrastructure" based on many local clusters, does not scale
access to NHR resources could scale also to short-term large needs
In case of needs
→ FAIR principles in particle physics only theoretically offered.

How could they be generalized / what is generic?

Ten questions

FSP ATLAS
Erforschung von
Universum und Materie

How could they be generalized / what is generic?
→ access to NHR resources for Punch2.0 project is generic,
application to LHC open data is only a small demonstration use case

End-to-end FAIR use case?
→ yes with this project, no without

Connection to DRP? – a digital output of a research process
(which category: publication, data set, software, …)
→ "EXPLORE" tool could be eintegrated in DRP,
Need to work on some technical aspects (AAI) and some formal questions
steuerliche Fragen, rechtliche Fragen, Absprache mit NHR

How is the use case viewed from outside of PUNCH?
→ access to NHR would strengthen Punch2.0 application to referees & funding agency
→ service is highly welcome by LHC/particle physics community

Ten questions

FSP ATLAS
Erforschung von
Universum und Materie

Define input – procedure – output? Specify workflow.
→ access to NHR requires contact/negotiations for Punch2.0 projeect with NHR management
→ LHC open data/EXPLORE as use case requires technical work
Punch2,0 access to NHR, user (outside PUNCH, general public) idntification and authorisation, job submission to NHR, existing frameworks from LHC open data, user accounting/monitoring, access to LHC open data from NHR clusters, output provision to users (outside PUNCH, general public)

Operation model? - what is required for the use case, and how to organize that?
→ provided by NHR application for entire PUNCH2.0 project
→ Punch2.0 sub-provides resources (and user support) for general public

Sustainability?
→ automatically given by NHR as long as federal government provides funding for NHR
→ mor sustainability than any resources model used at present