ICFA Data Lifecycle Panel: Goals for the upcoming workshops

ICFA Data Lifecycle panel meeting - August 13, 2024

Kati Lassila-Perini
Helsinki Institute of Physics - Finland

From the last meeting

- Collected input for from CERN Open data contributors from different experiments for FAIR / Open data specific sessions in major HEP conferences
    - See <u>replies</u> attached to the July meeting
    - Views vary but in general in favour.

- Very good response
    - Understanding it as Open data contributors valuing the chance to express their voices.

Upcoming workshops

4th DPHEP Collaboration Workshop

DPHEP

2–3 Oct 2024
ZOOM
Europe/Zurich timezone

Overview	**Data Preservation in High Energy Physics DPHEP**
Call for Abstracts	This is the fourth DPHEP Collaboration Workshop.
Timetable	It will also host a Collaboration Board meeting.
Registration	*The goals of this workshop are:*
Participant List	

1. *Perform a site-experiment round-table to capture the current situation including common problems & solutions HEP-wide*
2. *Provide an update on the changing (or changed) landscape, e.g. FAIR data management (plans), reproducibility, sustainability of data repositories, value of data lifecycle*
3. *Status reports of the transverse services and ongoing common developments and their outlook*
4. *Issue a status report document to e submitted to the ICFA Data Lifecycle panel.*

Please note that the agenda has to be adjusted to include new proposals/speakers. Proposals welcome.

HSF Training Pre-CHEP Workshop

19–20 Oct 2024
Europe/Warsaw timezone

Overview

Workshop Organisers
- ✉ valeriia.lukashenko@cer...
- ✉ pivarski@princeton.edu
- ✉ alexander.moreno@uan...

The HSF Training Working Group are delighted to organise a workshop on Training in HEP in advance of the main CHEP2024 conference in Krakow.

Training for software and computing is critical in high-energy physics and this workshop is open to everyone.

The workshop will take place over two half days, Saturday PM and Sunday AM.

Registration and payments will be handled via the main CHEP conference website - be sure to click the pre-CHEP workshop option and organise your travel accordingly!

More details about the event will follow shortly.

DPHEP:https://indico.cern.ch/event/1432766/
HSF: https://indico.cern.ch/event/1410343/

3

Our goals

- Collect input from the enablers of the field

  - Many committee-type initiatives lack this aspect, we should make sure that we reach out to those involved

  - This input can be used when formulating recommendations and guidelines

  - DPHEP to reach Open Data contributors (also beyond CERN)

  - HSF training WS to reach SW and training contributors.

- Draft for the DPHEP document:

  - https://docs.google.com/document/d/1Y7O2zARI0BXGBBIlOrem4JzEjTOX3FyVjKJbrPy-9tc/edit?usp=sharing

Discussion

Next meetings on October 8th (skipping September unless urgent issues)

Review the input from DPHEP
Plan the HSF training WS actions

Mission

Mission

The mission of the panel is to enhance global coordination on all aspects of the data lifecycle including acquisition, processing, distribution, storage, access, analysis, simulation, preservation, management, software, workflows, computing and networking in particle physics, with a focus on open science and FAIR practices.

In order to achieve this, the panel will

A. address all aspects of the data lifecycle, encompassing the efforts and expertise from previous panels, and relating to and building on activities of other relevant bodies and committees;
B. encourage global cooperation on the above topics in particle physics and with neighbouring fields;
C. discuss strategic questions and recommend to the community future directions;
D. encourage engagement with and profit from industry expertise in data management solutions, in artificial intelligence, and in systems competence;
E. develop ideas and strategies for the workforce development and for professional recognition mechanisms within the topical areas of the panel.

Mandate 1

Mandate

1. Address the data lifecycle within a structured and integrated systems approach in HEP
    1.1. Formulate recommendations on organisation, technology, standards, outreach, education for past/current/future experiments.
    1.2. Connect regional and local activities in the field and encourage international cooperation, aiming at stimulating active participation from the global HEP community.
    1.3. Raise awareness of open science and the FAIR principles applied to data, software and workflows, and stimulate relevant developments.
    1.4. Assess the openness and FAIRness of the field.
    1.5. Encourage transfer of knowledge
    1.6. Support the ongoing projects and collaborations started within the "Data Preservation in High Energy Physics" collaboration (DPHEP) and the "Standing Committee on Interregional Connectivity" (SCIC).

Mandate 2

Mandate (cont)

2. Improve the awareness for the importance of the data lifecycle in HEP
   - 2.1. Work out and communicate the motivation of FAIR (findability, accessibility, interoperability, and reusability) principles and open science and encourage its dissemination.
   - 2.2. Organise workshops, formulate recommendations and cookbooks, issue global reports
   - 2.3. Contribute to the training and education on open science issues in all world regions, employing in particular the facilities of the large laboratories in the field.
   - 2.4. Help in sharing expertise and existing solutions; catalyse new common projects; promote collaboration.

Mandate 3

Mandate (cont)

3. Encourage and foster connections to other fields of science, to industry and to open science initiatives in order to profit from their expertise and competence in the following fields:
   3.1. Big and distributed data management.
   3.2. Data management systems.
   3.3. Artificial intelligence.
   3.4. Open science processes.
   3.5. Data preservation systems.
   3.6. Reach out to neighbouring fields such as astro(particle) physics, hadron physics, and accelerator science, but also to the communities of photon and neutron science and others with large data volumes and related data challenges (genomic, public health, smart city, …)

Mandate 4 & 5

Mandate (cont)

4. Help in organising practical support and act as point of contact for practical issues in the field of data, software, workflows and computing
    4.1. Support the ongoing projects and co-operations started within DPHEP in order to maintain data sets that (can) still produce science, keep track on parked data sets
5. Improve recognition of the nature and value of work on the data lifecycle in researchers' CVs and support their career development.