HSF Highlights 2022 and Plans 2023

Graeme Stewart and **Benedikt Hegner**, for the HSF

SFT Group Meeting, 2022-02-13

Background and Usual Disclaimer

The HSF was formed in 2015, with the role to catalyse, enable and guide common efforts in the community. To first order, the HSF itself does not own or distribute resources.

So every successful project or effort is owned by the people doing the work!

So, when talking about HSF highlights, we are talking about great work done as well by others.

HSF

- Coordination Team for oversight and driving overall engagement, organising workshops
  - Modest sized group of motivated individuals who contribute to general running of HSF
  - Ex-officio members from experiments and WLCG as stakeholders
  - **SFT Members: Benedikt Hegner, Pere Mato and Graeme Stewart**
- Working Groups for key areas of HEP activity
  - Working Group Conveners from SFT:
    - **Frameworks - Benedikt Hegner**
    - **Software Tools and Packaging - Valentin Volkl**
- The HSF's main role here is one of an information conduit and meeting point
  - Forum for technical comments and discussion
  - Inform about and encourage cooperation across experiments and regions
  - Motivate the publication of summary documents or papers for future use

A More Normal Year...

- In 2022 we had a gradual easing of restrictions related to COVID-19
  - Finally again sizeable in-person conferences: ICHEP in July, ACAT in November
  - Experiments resumed many in-person events
- All of this helped re-introduce a much needed face-to-face / coffee time / beer dimension to activities
  - This is essential to the long term health of our community
  - Student activities are particularly valuable, e.g., the CERN summer student programme
- However, the world is not the same as it was before
  - Virtual participation is now accepted for almost all events
    - Balance costs vs. quality of interaction
  - Increasing awareness of environmental costs of travel

Having established communication channels, the HSF was a core ingredient to stay connected through the peak times of Covid-19

HSF - a place for International Efforts and Cooperation

- The HSF continues to support and foster common projects
  - This encourages diverse R&D!
  - There is now real success in attracting funding to this area
  - Recognised links to other main players (WLCG, LHCC, IRIS-HEP, EPPSU, SWIFT-HEP, Experiments)
- HSF offers an excellent place to present work, discuss successes (and disappointments!) and to help lead community activities

Ensuring consistency and coherency stays a continuous task and challenge for 2023

Joining forces with the NP community

- Communication between HEP and nuclear physics community sometimes very limited
  - Despite tackling increasingly similar problems
- In 2021 we joined the organization of the Software and Computing Roundtable of BNL and JLab
- For 2022 we pushed for “recruitment” of convenors from the NP community
  - Got nicely reflected in talks and work in the WGs
- Similar engagement we have seen in e.g. discussions between Key4HEP and EIC.
  - Very open mindset on both sides!

We will continue to work in this direction of removing historical boundaries

Community Advocacy

- We continued to advocate for software in the community, with several talks at conferences and events
  - The HEP Software Foundation, SMARTHEP kick-off workshop, 24 November 2022, Benedikt Hegner
  - Sustainability and future of software frameworks, JENA Symposium, 5 May 2022, Graeme A Stewart
  - HEP Software Foundation and Software Project R&D, SWIFT-HEP Meeting, 24 March 2022, Graeme A Stewart
  - Software and Computing R&D, 30th International Symposium on Lepton Photon Interactions at High Energies, 14 January 2022, Graeme A Stewart
- In addition, the HSF submitted several papers and LOIs to the US Snowmass process, particularly in the Computing Frontier
  - Many US HSF colleagues involved
- We regularly give input to the LHCC together with WLCG

HSF Workshops and Events

We had a rich programme of workshops in 2022, organised, in many cases, with other partners:

- HSF Detector Simulation on GPU Community Meeting
  - https://indico.cern.ch/event/1123314/ (summary talk by Witek Pokorski)
- Analysis Ecosystem II Workshop
  - https://indico.cern.ch/event/1123314/ (summary talk by Enrico Guiraud)
- PyHEP 2022 Workshop
  - https://indico.cern.ch/event/1150631/ (summary talk by Graeme Stewart amd Vincenzo Padulano)
- HSF - IRIS-HEP Workshop on Software Citations
  - https://indico.cern.ch/event/1212344/
- Future Trends in Nuclear Physics Computing
  - https://indico.bnl.gov/event/15089/
- MC4EIC
  - https://indico.bnl.gov/event/17608/

HSF/IRIS-HEP Workshop: Software Citation and Recognition

- Workshop organised on Software Citation and Recognition
  - Review status of citation in HEP
  - Give credit to software developers and maintainers
  - Provide better and more sustainable software
  - Support for reproducibility
- Key principles developed by Force11 group
  - Importance, Credit and attribution, Unique identification, Persistence, Accessibility, Specicifity
  - Group then had task forces which helped to develop
    - Citation Format File standard (CITATION.cff)
    - CodeMeta
      - Metadata standard for software, a richer description of software

What to cite?

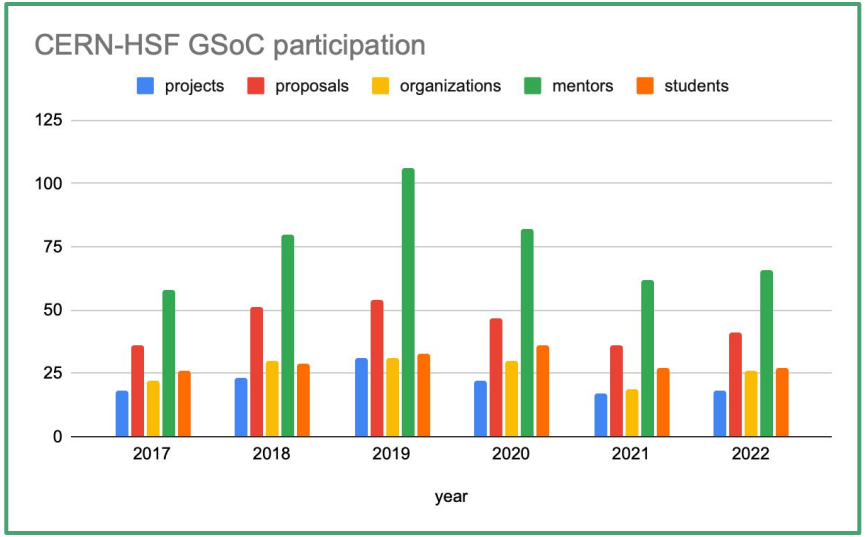
- An academic paper written about the software ?
  - This is the traditional approach, currently giving the most academic credit
  - There is a serious issue with ancestor papers picking up all citations
    - E.g., the 2003 Geant4 paper gets most citations - even though the code today is almost completely different and all the recent authors and contributors are missing
- The software itself ?
  - E.g., the Zenodo DOI
  - Not well rewarded academically
  - Does it tell anything about the software itself?
- A combination of the two ?
  - E.g. the Journal of Open Source Software (JOSS)
  - Combining code with a short paper
  - Code and repository is reviewed as well for best-practice standards

Citations and Recognition Outcomes

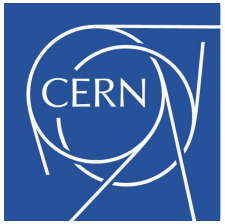
- Developers
  - If you want your software properly cited, put the citation everywhere...
    - In the README, in the documentation, on the distribution page (PyPI)
    - And make this a single source of truth!
  - Adopt a citation format file
    - CITATION.cff - first version can be easily generated via a webpage
  - Make sure you keep things up to date
- Experiments
  - Desire for consistency in citation recommendations - possibly curated by the HSF?
- Zenodo / Inspire
  - Better support for software citation coming this year, when automatically harvestable (e.g., from Zenodo INSPIRE HEP, CERN open data)
  - Will track citations ➜ credit for authors
- **Workshop conclusions in preparation**

GSoC 2022

- We participated as an umbrella organisation again in GSoC
  - Programme improved last year: allowed to have short and long projects: 175 or 350 total coding hours
- Key numbers
  - 26 Organizations
  - 18 HSF projects
  - 27/41 proposals got a student (⅔)
  - 21/27 successful student projects
    - Record failure rate (22%)!
    - 2 withdrawn, 4 bad performance
    - Relaxed participation requirements (?)
- Student blogs available



GSoC 2023

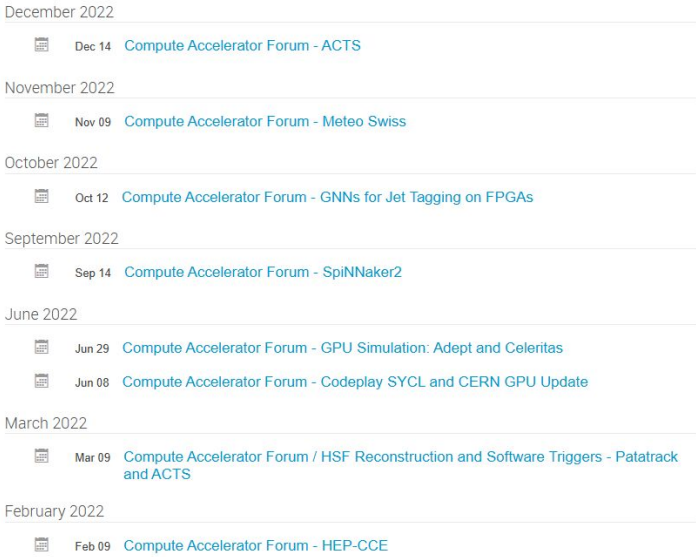


- Same rules as for 2022, but even more open
  - Mix of medium (175 hours) and large (350 hours) projects
  - Flexible project duration
  - “Program open to students **and** to beginners in open source software development”
- Deadline for project submission passed
  - Waiting for acceptance/go-ahead by Google now
- Student selection procedure starts Feb 22
  - April 4 deadline for students to submit their proposals
  - May 4 accepted projects announced

- More details available here

Compute Accelerator Forum

- Compute Accelerator Forum is a series of meetings touching on many diverse topics in the shift towards extreme parallelisation and non-CPU architectures
- In 2022 the CAF was an established part of the Software landscape here at CERN and in NHEP



Will continue the series with the same rhythm

Working Groups

- In addition to dedicated workshops and events we have many working groups and activity areas
  - Led by enthusiasts and advocates for common work and solutions
- These were active in 2022 and held many useful discussions in their field of expertise
  - See their Indico categories for more details on the meetings and topics covered
- Had two (Jan 19, Feb 2) planning meetings in January

Working Groups

- Data Analysis
- Detector Simulation
- Frameworks
- Physics Generators
- PyHEP - Python in HEP
- Reconstruction and Software Triggers
- Software Developer Tools and Packaging
- HSF Training

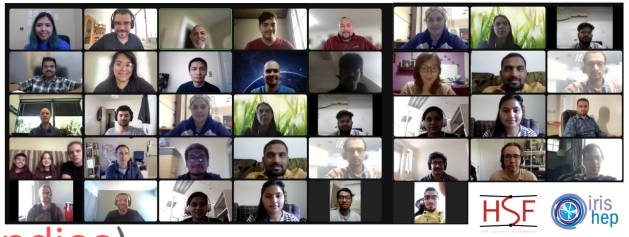
Activity Areas

- Analysis Facilities Forum
- Conditions Databases
- Differentiable Computing
- Season of Docs
- Google Summer of Code
- intelligent Data Delivery Service
- Licensing
- Reviews
- Visualisation

Training WG

Webpage, Indico: Internal, Events, C++

Sudhir Malik, Michel Hernandez Villanueva, Wouter Deconinck, Kilian Lieret



Achievements unlocked:

- April 21-22: Matplotlib Training (indico)
- July 13-15: Software Carpentry Training (indico)
- July 25: Matplotlib Training Hackathon (indico)
- September 6: Containerization Training Hackathon (indico)
- Participation in ICHEP, PyHEP, Sustainable HEP conferences
- September 28-30: Software Carpentry Training (indico)
- October 11-13: Advanced C++ Training (indico)
- January 16-21: Analysis Preservation Training (indico)

Upcoming Quests:

- February 8-10: Software Carpentry training (indico)
- May 15-19: C++ Training - The American Edition @JLab (TBC)

Frameworks

Webpage, Indico

Nathan Brei, **Benedikt Hegner**, Kyle Knoepfel, Vakho Tsulaia

- 2022
  - New features in C++20
  - “Small” experiment frameworks
    - Mu2e
    - NOvA
    - JANA2
- 2023
  - Integration of accelerators into experiment frameworks
  - New language features
  - New parallelization libraries

Software Tools and Packaging

Webpage, Indico

Patrick Gartung, Henry Schrei, **Valentin Volkl**

- 2022
  - A very diverse set of events
    - Build-time-optimizations, HPC I/O profiling, conda
    - Invitations to talk about Spack
- 2023
  - Continue Spack Discussions with contributions from Spack Developers and Experiments
  - Python packaging
  - Cross-WG activities
    - C++ evolution
    - Training material

Summary

- In 2022 HSF colleagues organised many events and discussions on important topics for High Energy and Nuclear Physics
  - Reinforcing the role of the HSF as a place where the community can gather for discussions and exchange of ideas
  - This feeds back into the work tackled by the software projects
  - We try to encourage diverse R&D, but also very practical solutions that deliver for the experiments
- In 2023 already look forward to another active year
  - A highlight will be the first in-person WLCG-HSF workshop since 2019, co-located with CHEP
  - Topics will be Analysis Facilities and Heterogeneous Computing
  - PyHEP and PyHEP Dev already in preparation
  - The working groups discussed and presented their plans in January

Backup