



PyHEP 2018
Sofia, Bulgaria

PyHEP 2019
Abingdon, U.K.



PyHEP 2020 workshop experience and 2021 plans

Eduardo Rodrigues, for the PyHEP organising committee
University of Liverpool

PyHEP 2020
Virtual, online



PyHEP = "Python in HEP" series of workshops

- ❑ **Started in 2018, recognising the increasing importance of Python in Particle Physics**
 - There are several conferences & workshops on C&SW but nothing existed with Python as first-class

The **PyHEP workshops** are a series of workshops initiated and supported by the **HEP Software Foundation** (HSF) with the aim to provide an environment to discuss and promote the usage of Python in the HEP community at large. Further information is given on the **PyHEP WG website**.

- ❑ **Workshop format – unchanged for in-person and virtual events:**

- Only plenary sessions
- Very informal, lots of time for (lively) discussions
- Bring together users and developers
- Educative, not just informative

[PyHEP - Python in HEP \(hepsoftwarefoundation.org\)](https://hepsoftwarefoundation.org)

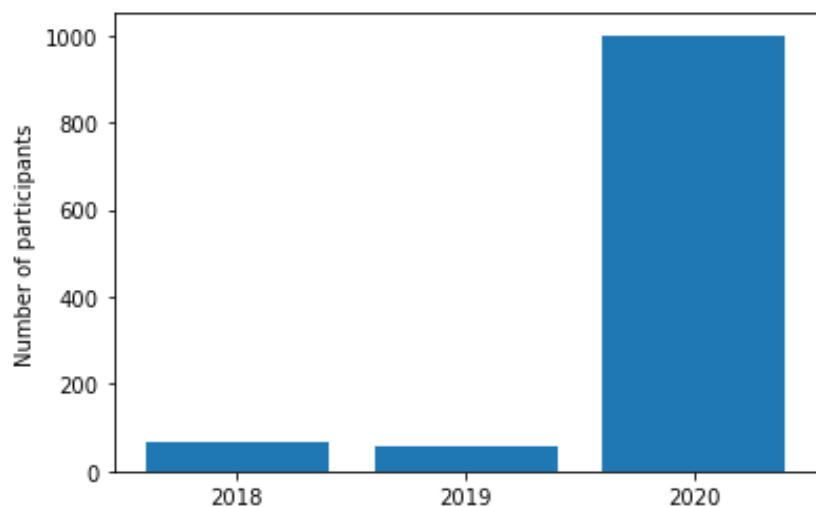
- ❑ **In-person seemed the adequate format but we had to run online in 2020 ...**
... and we learned from this experience ...

<i>Workshop</i>	<i>Location</i>	<i>Date</i>
PyHEP 2021	Virtual workshop	July 5-9, 2021
PyHEP 2020	Virtual workshop	July 13-17, 2020
PyHEP 2019	Abingdon, U.K.	October 16-18, 2019
PyHEP 2018	Sofia, Bulgaria	July 7-8, 2018

[Figures have active links.]

What changed when moving to a virtual event in 2020

- ❑ Event spread over a week (5 days) rather than 3 (1.5) days in 2019 (2018)
- ❑ Shorter sessions per day, 3-hour long at most
- ❑ Sessions organised in 2 time zones!
- “Pacific-friendly” and “Atlantic-friendly”
- ❑ Workshop became a truly global event with participants from all over the world
- ❑ No registration fees
- ❑ Impressive level of interest with 1000 registrations (limited to)
(72, 55 in previous years)



[Information taken from the 408/1000 responses received from the pre-workshop survey.]

[Plot from Jim Pivarski]

PyHEP 2020 – Indico page, organising team, sponsors

PyHEP 2020 (virtual) Workshop

13-17 July 2020

US/Central timezone

Overview

Call for Abstracts

Timetable

Registration

Participant List

Poster

Surveys

Code of conduct

EDI statement

Workshop photos

Contact us

✉ pyhep2020-organisation...

<https://indico.cern.ch/e/PyHEP2020>

Organising Committee

Eduardo Rodrigues - University of Liverpool (Chair)

Ben Krikler - University of Bristol (Co-chair)

Jim Pivarski - Princeton University (Co-chair)

Matthew Feickert - University of Illinois at Urbana-Champaign

Local organisation

Chris Tunnell - Rice University

Peter Onyisi - The University of Texas at Austin

Sponsors

The event is kindly sponsored by



PyHEP 2020 – sessions & presentations

- ❑ **Sessions spread in “Atlantic”- and “Pacific”-friendly time zones to accommodate Asia, Americas and Europe**

- **Atlantic:** ~3h with 30-min breaks; **Pacific:** ~1h, no breaks. E.g.,

Atlantic: 15h00 - 18h00 CET, 06h00 - 09h00 PDT, 18h30 - 21h30 IST, 21h00 - 24h00 CST, 22h00 - 01h00+1 JST

Pacific: 15h00 - 16h00 PDT, 00h00 - 01h00+1 CET, 03h30+1 - 04h30+1 IST, 06h00+1 - 07h00+1 CST, 07h00+1 - 08h00+1 JST

- **Quite a bit more work and more demanding for the session chairs**

- **The Pacific sessions ended up far less popular than the Atlantic sessions. Decision not to replicate this year**

- ❑ **We strongly encouraged “notebook presentations” + Binder for an interactive experience**

- ❑ **Notebooks and related material made available in public GitHub repositories with a [Binder](#) launch button (all presentational material posted on workshop agenda, including repo. links)**



- ❑ **We used both the [Binder Federation](#) and the [CERN Binder Hub](#) resources (for those with CERN accounts)**
- ❑ **Got in touch with the Binder team to have [resources allocated to talk repositories at the relevant time](#) !**
 - **It worked very well – thank you MyBinderTeam**
 - **Binder was a leitmotif during the workshop:**



From the PyHEPConf Twitter account

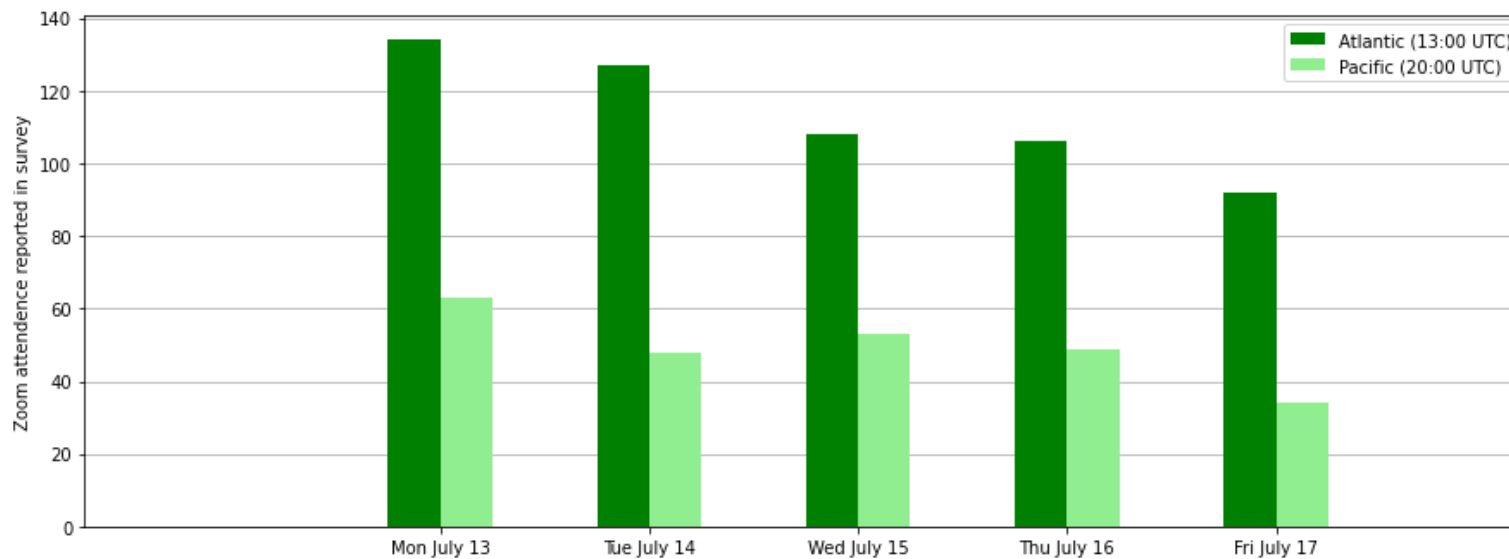
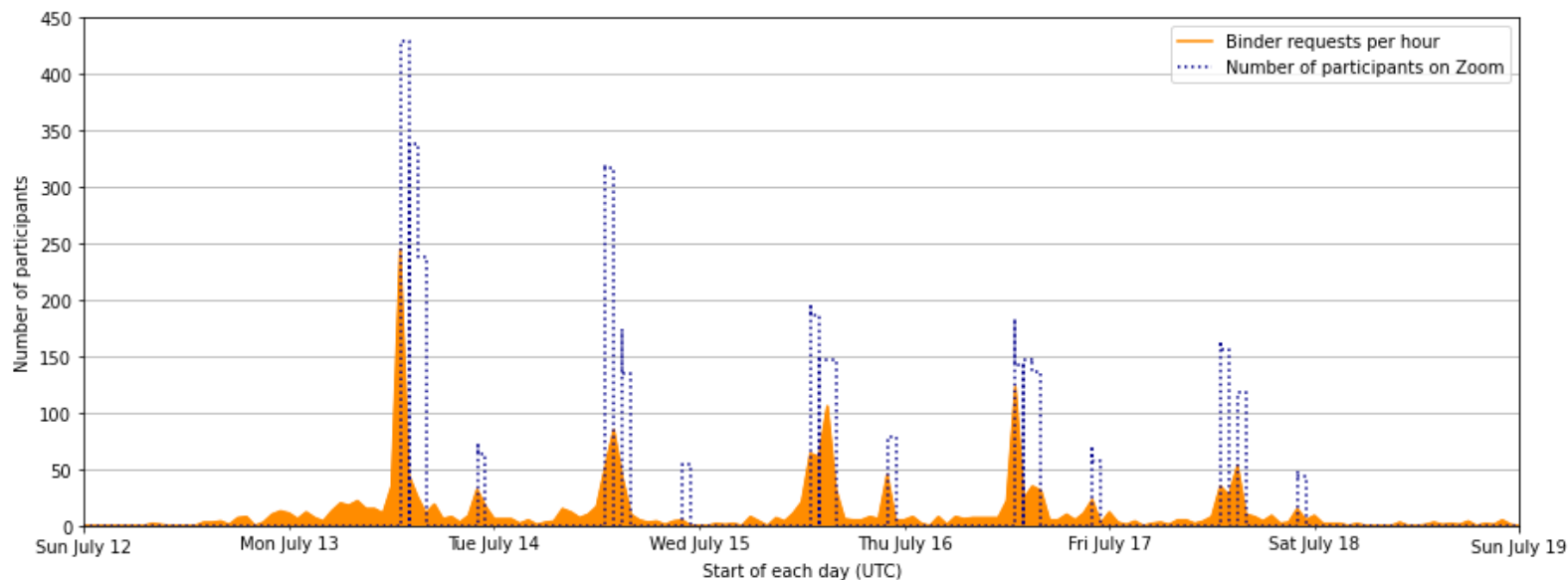
PyHEP 2020 – session attendance & Binder usage

- ❑ Session participants

- ❑ Binder requests during sessions

⇒ Clear correlation !

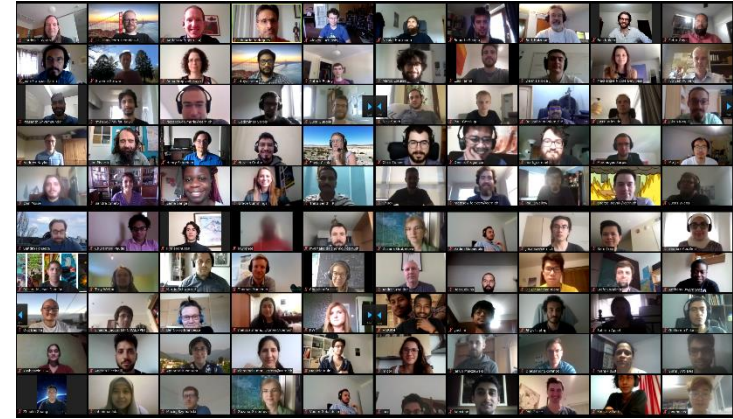
- ❑ Number of participants per day & time zone, as reported by those who filled in the post-workshop survey - “Atlantic” time zone suited most



Study by Jim Pivarski

PyHEP 2020 – videoconferencing

- ❑ **Zoom video conferencing system**
 - With capacity for 1000 participants
 - Public room but PIN provided via email
- ❑ All recorded. No pre-recordings



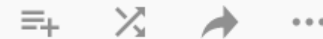
“Workshop photo”
from close-up Atlantic session

- ❑ (HSF has its own YouTube channel, with several playlists)
- ❑ **All presentations got recorded and captioned**
(captioning thanks to sponsors, see later)
- ❑ Later uploaded to the [HSF YouTube channel](#)
to a dedicated playlist [“PyHEP 2020 Workshop”](#)



PyHEP 2020 Workshop

32 videos • 945 views • Last updated on 19 Jul 2020



Talks, tutorials and keynotes from the PyHEP 2020 Workshop,
<https://indico.cern.ch/e/pyhep2020>

PyHEP 2020 – communication & interactions

❑ Slack channels



- Various channels:
- By topic, mapping to sessions, discussions encouraged here
- Announcements, for actual announcements
- Random, used to encourage community spirit and add social context

❑ Communication also on



[@PyHEPConf](https://twitter.com/PyHEPConf)

[#PyHEP2020](https://twitter.com/PyHEP2020)



❑ Questions & answers with slido



- (AFAIK we were among the very first to try Slido in HEP)
- Used *slido* to crowd-source questions, to prioritise the most popular ones upvoted by participants
- Session chair shares link to questions at end of presentation
- Most popular ones get answered/discussed
- At end of Q&A all questions are copied to Slack in the appropriate topical channel
⇒ participants can continue to discuss and exchange
- A few polls also run via slido

- ❑ **Post-conference:** participants encouraged to join the PyHEP Gitter channel(s)

PyHEP 2020 – Slack for discussion during *and* after sessions

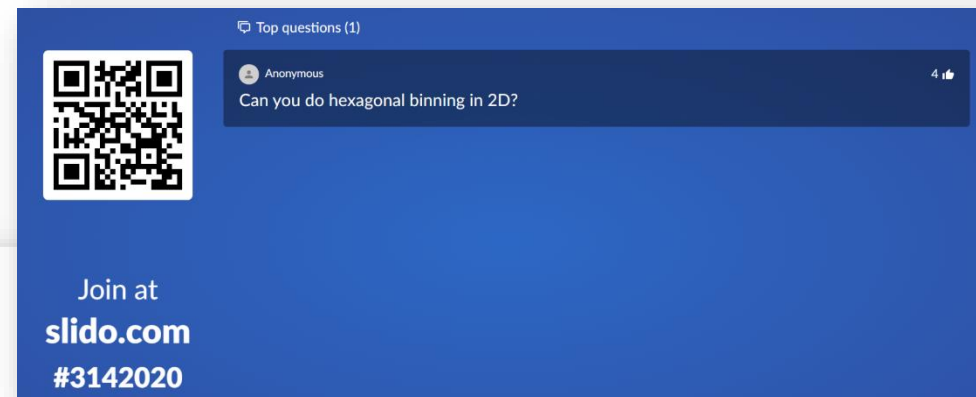
The screenshot displays a Slack workspace for PyHEP2020. On the left sidebar, the 'Channels' section is highlighted with a green box and contains the following channels: # announcements, # favorite-tools, # organising-team, # random, # session-chairs, # talk-question-and-answer, # topic-analysis-fundamentals, # topic-analysis-platforms-systems, # topic-automatic-differentiation, # topic-fitting-statistics, # topic-hep-ecosystem, # topic-performance, # tutorial-high-performance-python, # twitter, and # zenodo-organisers. The main content area shows the #announcements channel with a search bar at the top. A message from [Org] Jim Pivarski at 4:55 PM is visible, stating 'Here is mine (Atlantic session)'. Below the message are two large group photos of conference attendees. A green box on the right side of the image contains the following text:

- Several general and topical channels
- A few channels for organisers and session chairs

PyHEP 2020 logistics – slido for Q&A post-talk sessions




As actually seen by participants:

Was slido a success? Yes !



slido

Event summary report PyHEP2020

 Active users 181	 Questions 182	 Poll votes 195
Engagement score 978	Likes / dislikes 483 / -54	Polls created 5
Engagement per user 5.4	Anonymous rate 34%	Votes per poll 39

With 413 joined participants in total



- ❑ **No (standard) proceedings per se**
- ❑ **All presentational material posted on workshop agenda**
and later given a DOI with Zenodo, in a dedicated “pyhep2020 community” – formal citation, replaces proceedings
- Indico contains links to slides, notebook repositories, Binder launch buttons, YouTube recordings

- ❑ **With Zenodo + Binder, all code from the workshop should be reproducible into the future**
⇒ **“living workshop proceedings”!**

- ❑ **Recordings on YouTube are in some way an alternative to proceedings**

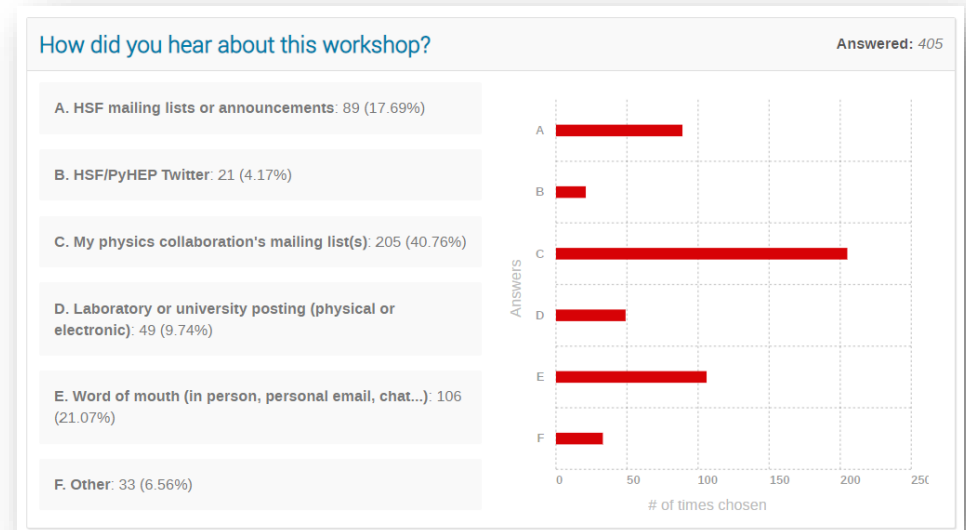
PyHEP 2020 – sponsors

- ❑ The Python Software Foundation is a sponsor from the onset, which we greatly appreciate !
- ❑ Sponsors for PyHEP 2020 covered the only expenses we had, i.e. the captioning of the recordings – thank you!
 - Covered in kind by the PSF and IRIS-HEP
 - Fermilab provided the Zoom system for 1000 participants – not the standard configuration in July 2020
 - Other sponsorships “kept warm” for the next workshop iteration ...



PyHEP 2020 – Stepping back

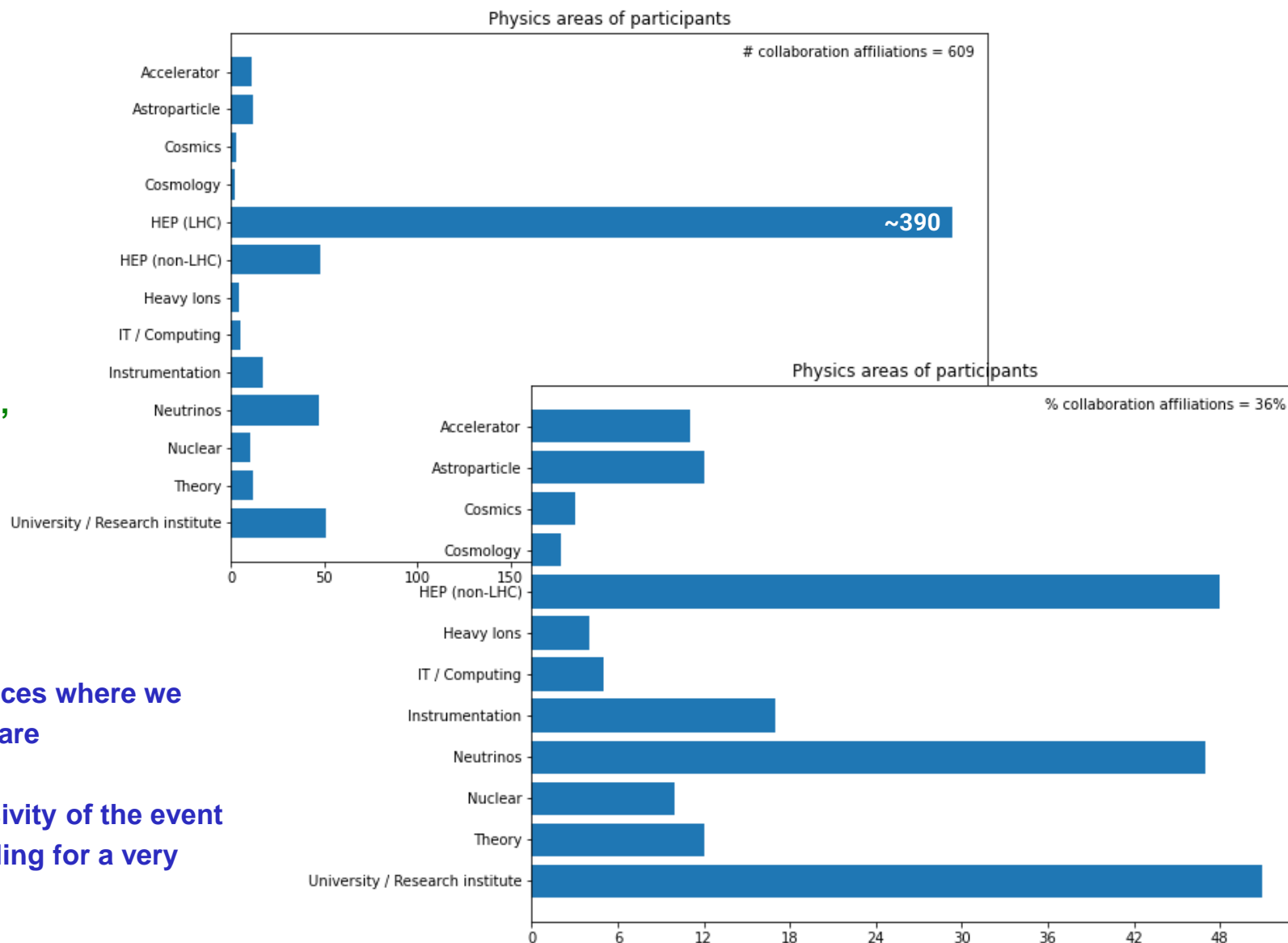
- ❑ The broadcasting of the event to as many communities as possible is a game-changer
- ❑ Things worked really well overall !
(Trend from post-workshop survey feedback)
- ❑ But engagement from several *organisers* during the sessions is crucial
 - Chair the session, take care of activity in Slack/slido, ...
- ❑ But some usual suspects say they miss the in-person events
- ❑ When going virtual:
 - Organisation can start much later and is less cumbersome
 - Is easier - no venue / catering to sort out
 - More time to focus on actual event content
 - No registration fees makes it more inclusive



Taken from the pre-workshop survey (408 respondents)

On PyHEP 2021 plans

- ❑ Largely unchanged since 2020
- ❑ But no sessions in 2 time zones
- ❑ Plan to live stream to YouTube (atop Zoom) to avoid any limit
- ❑ Again seeing much interest
- ❑ Seeing new communities joining, in particular the neutrino one!
- ❑ Maybe for next year
 - Go for a hybrid event
 - Arrange large meeting rooms for places where we know large fractions of participants are (CERN, Fermilab?, ...)
 - This keeps “global-ness” and inclusivity of the event while providing some in-person feeling for a very large fraction of the participants
 - To be discussed ...



Thank you for listening

And see you at PyHEP 2021 ;-) !

<https://indico.cern.ch/e/PyHEP2021>

Workshop agenda (1/2)

Keynotes

- ❑ Rubin Observatory: the software behind the science (Nate Lust)
- ❑ Python & HEP: a perfect match, in theory (David Straub)

Tutorials

- ❑ Uproot & Awkward Arrays (Jim Pivarski)
- ❑ Jagged physics analysis with Numba, Awkward, and Uproot on a GPU (Joosep Pata)
- ❑ Ganga: flexible virtualization for user-based large computations (Ulrik Egede)
- ❑ A prototype U.S. CMS analysis facility (Oksana Shadura)
- ❑ Columnar analysis at scale with Coffea (Mat Adamec)
- ❑ Introduction to automatic differentiation (Lukas Heinrich)
- ❑ High-performance Python (Henry Schreiner)
- ❑ Model-building & statistical inference with zfit and hepstats (Jonas Eschle)
- ❑ pyhf: accelerating analyses and preserving likelihoods (Matt Feickert)
- ❑ ThickBrick: optimal event selection and categorization in HEP (Prasanth Shyamsundar)

*Typically
45 minutes*

Workshop agenda (2/2)

Talks

- NanoEvents object (Nick Smith)
- TITANIA: how to structure detector monitoring (Jakub Kowalski, Maciej Witold Majewski)
- A new PyROOT for ROOT 6.22 (Enric Tejedor Saavedra)
- Resample: bootstrap and jackknife from Python (Hans Dembinski)
- Design pattern for analysis automation using Luigi (Marcel Rieger)
- ServiceX: on-demand data transformation & delivery (Kyungeon Choi)
- Integrating Coffea and WorkQueue (Cami Carballo)
- High granularity calorimeter (HGICAL) test beam analysis using Jupyter (Matteo Bonanomi)
- neos: physics analysis as a differentiable program (Nate Simpson)
- SModelS: a tool for interpreting simplified-model results (Wolfgang Waltenberger)
- TensorFlow-based maximum likelihood fits for high-precision Standard Model measurements at CMS (Josh Bendavid)
- Error computation in iminuit and MINUIT: how HESSE and MINOS work (Hans Dembinski)
- zfit with TensorFlow 2.0: dynamic and compiled HPC (Jonas Eschle)
- Machine learning for signal-background separation of nuclear interaction vertices in CMS (Anna Kropivnitskaya)
- The boost-histogram package (Henry Schreiner)
- Providing Python bindings for complex and feature-rich C and C++ libraries (Martin Schwinzerl)
- Integrating GPU libraries for fun and profit (Adrian Oeftiger)
- mplhep: bridging Matplotlib and HEP (Andrzej Novak)
- ROOT preprocessing pipeline for machine learning with TensorFlow (Matthias Komm)
- Integrated data acquisition in Python (Charles Burton)

**Typically
20+10 minutes**

PyHEP 2020 – overview of organisational aspects

□ Sessions & presentations

- Spread in sessions for “Atlantic”- and “Pacific”-friendly time zones
- We strongly encouraged notebook presentations, available in public Github repositories with a [Binder](#) launch button
- All presentational material posted on workshop agenda and later given a DOI with [Zenodo](#), in a dedicated “[pyhep2020 community](#)” – formal citation, replaces proceedings
- All talks got recorded, captioned and later uploaded to the [HSF YouTube channel](#) – dedicated playlist “[PyHEP 2020 Workshop](#)”



□ Zoom video conferencing system



- With capacity for 1000 participants
- Public room but PIN provided via email

□ Slack channels



- Various channels:
 - By topic, mapping to sessions, discussions encouraged here
 - Announcements, for actual announcements
 - Random, used to encourage community spirit and add social context

□ Questions & answers with slido



- Used *slido* to crowd-source questions, to prioritise the most popular ones upvoted by participants
- Session chair shares link to questions at end of presentation
- Most popular ones get answered/discussed
- At end of Q&A all questions are copied to Slack in the appropriate topical channel
 - ⇒ participants can continue to discuss and exchange
- A few polls also run via slido

□ Communication also on



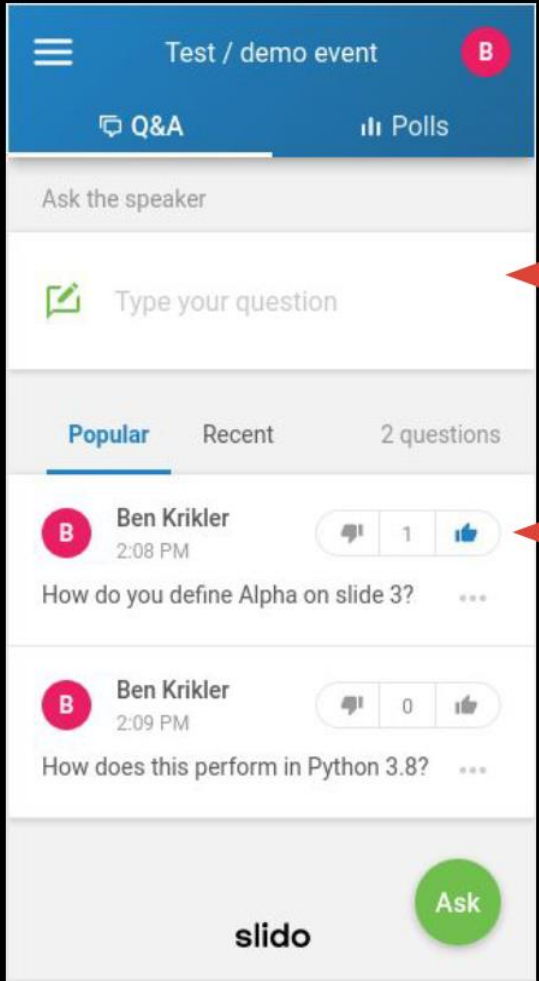
PyHEP 2020 – how does slido work for Q&As

slido

✓ Easy to use

✓ Works with your live video

✓ No app downloads



PyHEP2020: Asking questions

Click here to enter a new question

Up and downvote existing questions

When asking a question set your name
*it helps us find you on slack
no account needed*

Ask the speaker

Type your question

Popular Recent 2 questions

Ben Krikler 2:08 PM 1

How do you define Alpha on slide 3?

Ben Krikler 2:09 PM 0

How does this perform in Python 3.8?

216

Ben Krikler SEND

slido Ask

PyHEP 2020 – slido at work for Q&As and polls

As actually seen by participants



Top questions (1)

Anonymous

Can you do hexagonal binning in 2D?

4

Join at
slido.com
#3142020



Join at
slido.com
#3142020

slido

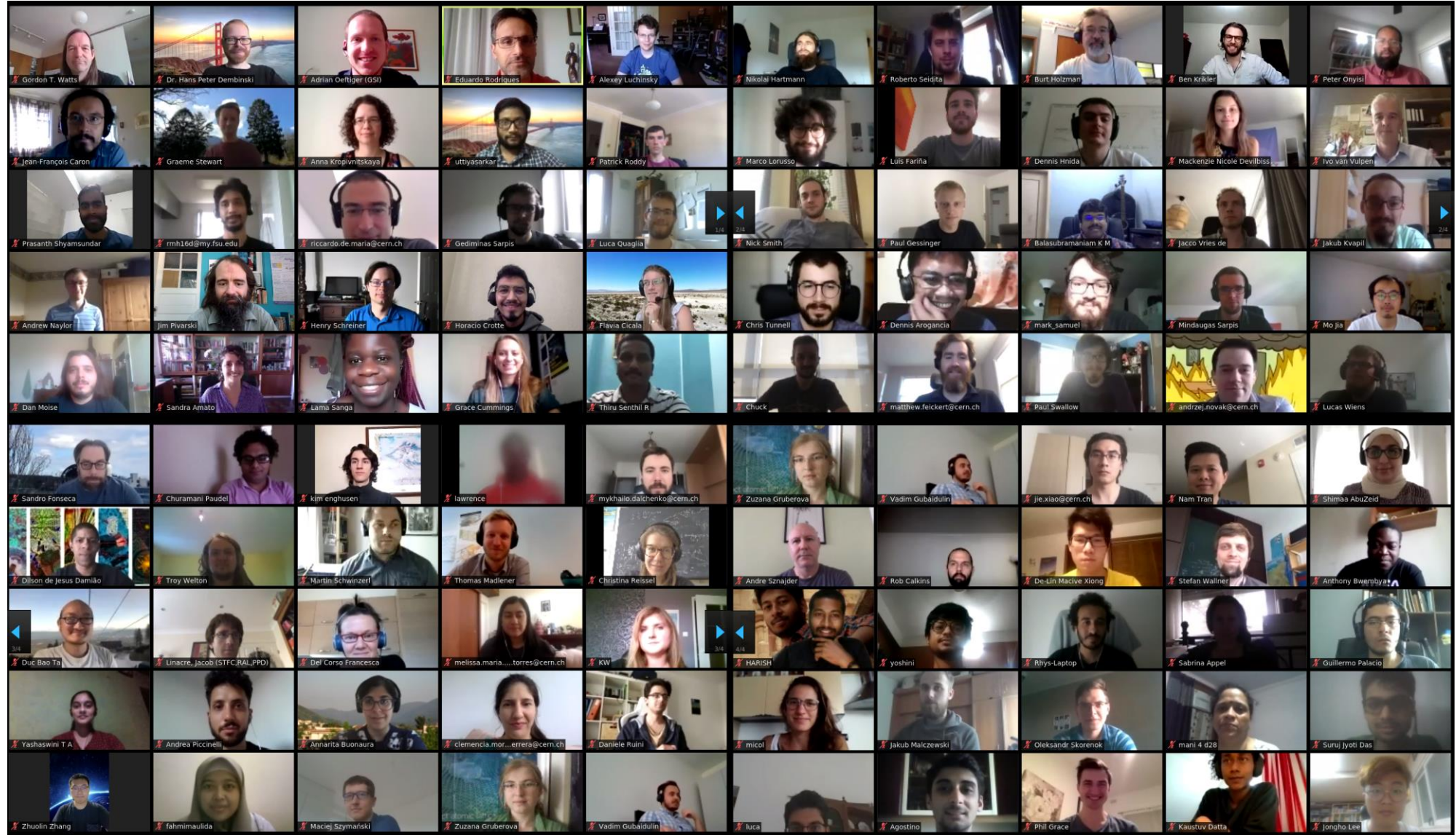
Active poll

What one or two words would be the "buzz word" for this workshop?

0 2 5



PyHEP 2020 – "workshop photo" @ end of last Atlantic session



PyHEP 2020 – "workshop photo" @ end of last Pacific session



PyHEP 2020 stats – background of participants ...

If you're involved in physics, what area(s) do you study?

Answered : 405 You can answer this AND the area of computing (below) or only one, depending on what you do.

A. General physics (student): 53 (8.48%)

B. High-energy collider physics: 295 (47.20%)

C. Neutrino physics: 52 (8.32%)

D. Physics of nuclei or exotic atoms: 14 (2.24%)

E. Precision frontier: 28 (4.48%)

F. Direct dark matter searches: 32 (5.12%)

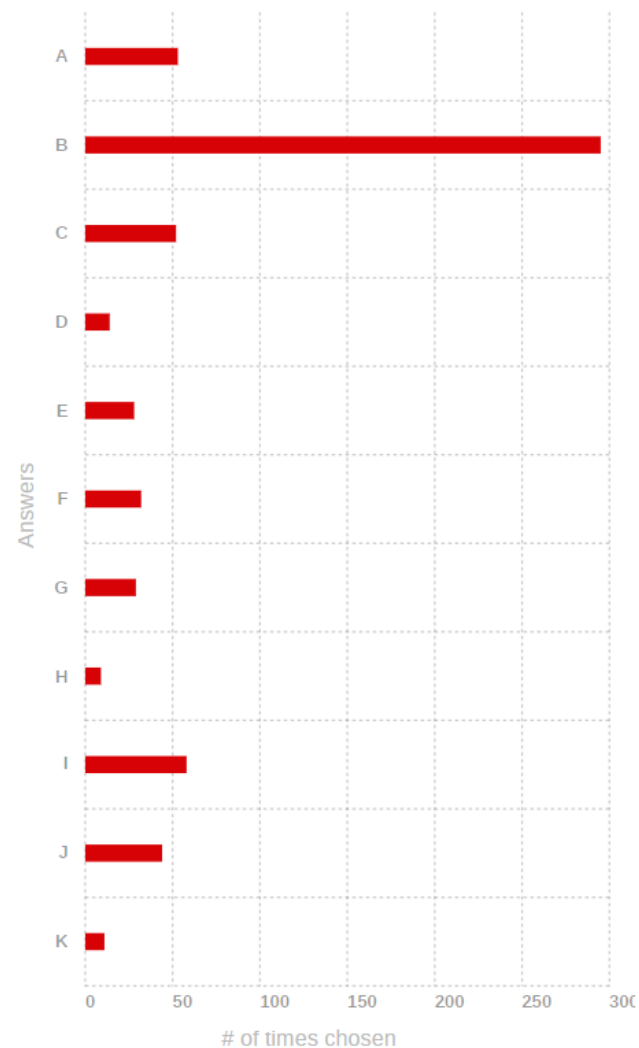
G. Astroparticle physics: 29 (4.64%)

H. Astronomy: 9 (1.44%)

I. Theory/simulations: 58 (9.28%)

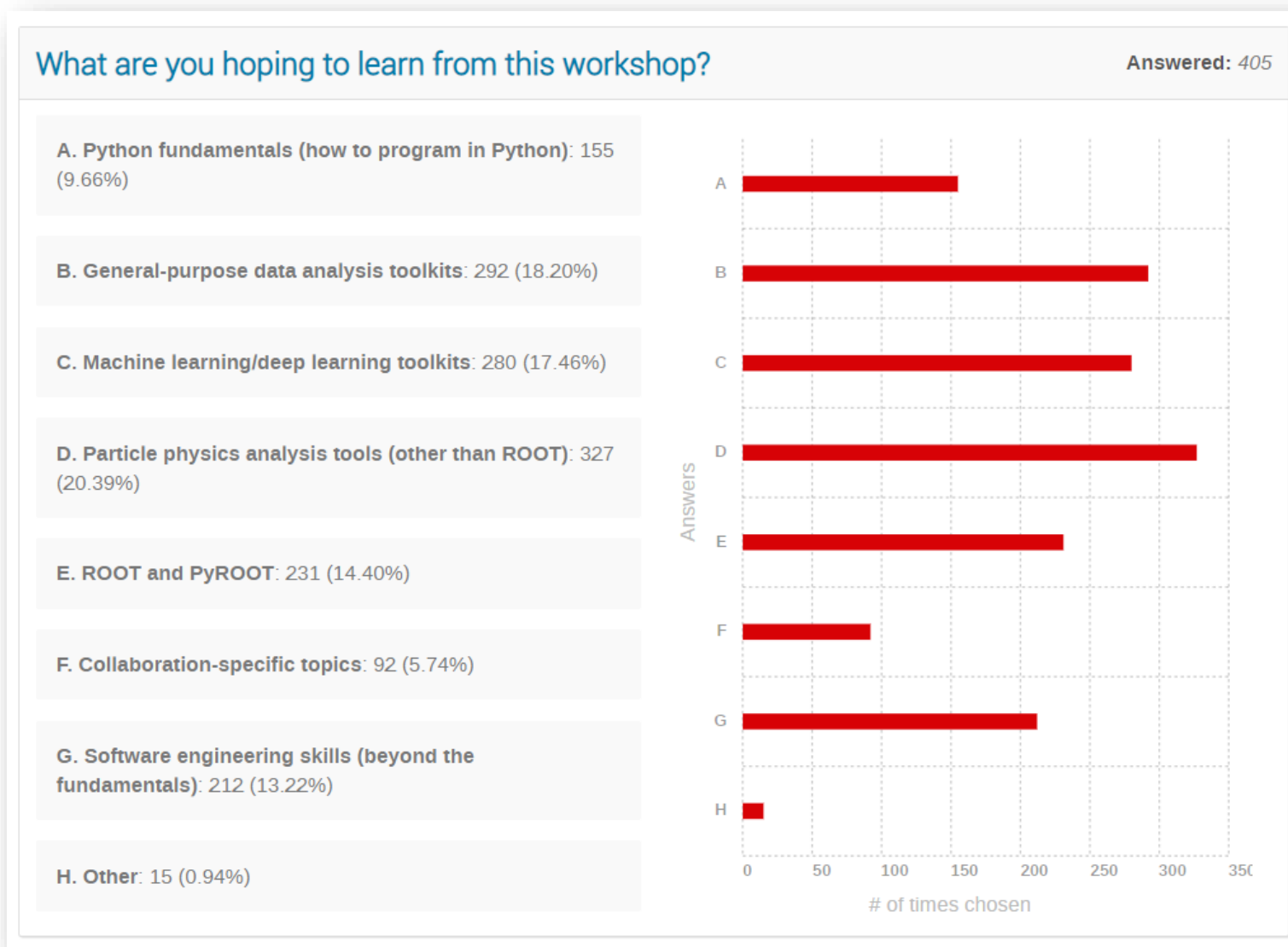
J. Instrumentation: 44 (7.04%)

K. Other, not listed above: 11 (1.76%)



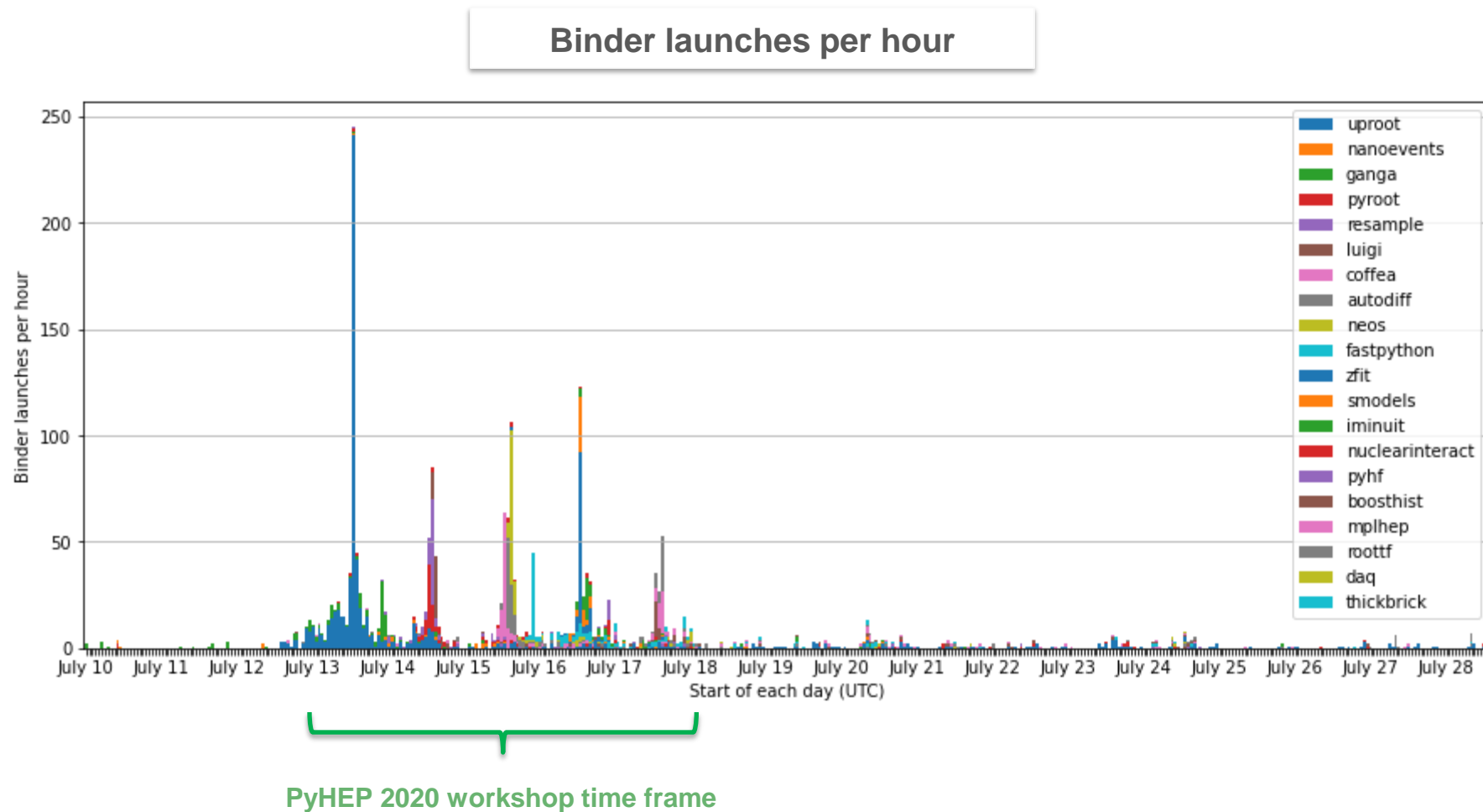
Taken from the pre-workshop survey (408 respondents)

PyHEP 2020 stats – ... and their hopes



Taken from the pre-workshop survey (408 respondents)

PyHEP 2020 stats – Jupyter notebook presentations & Binder usage



Study by Jim Pivarski