



Graeme Watt (IPPP Durham)
PrecisionSM meeting, 3rd June 2020

<https://hepdata.net>

Email: info@hepdata.net

 Follow @HEPData

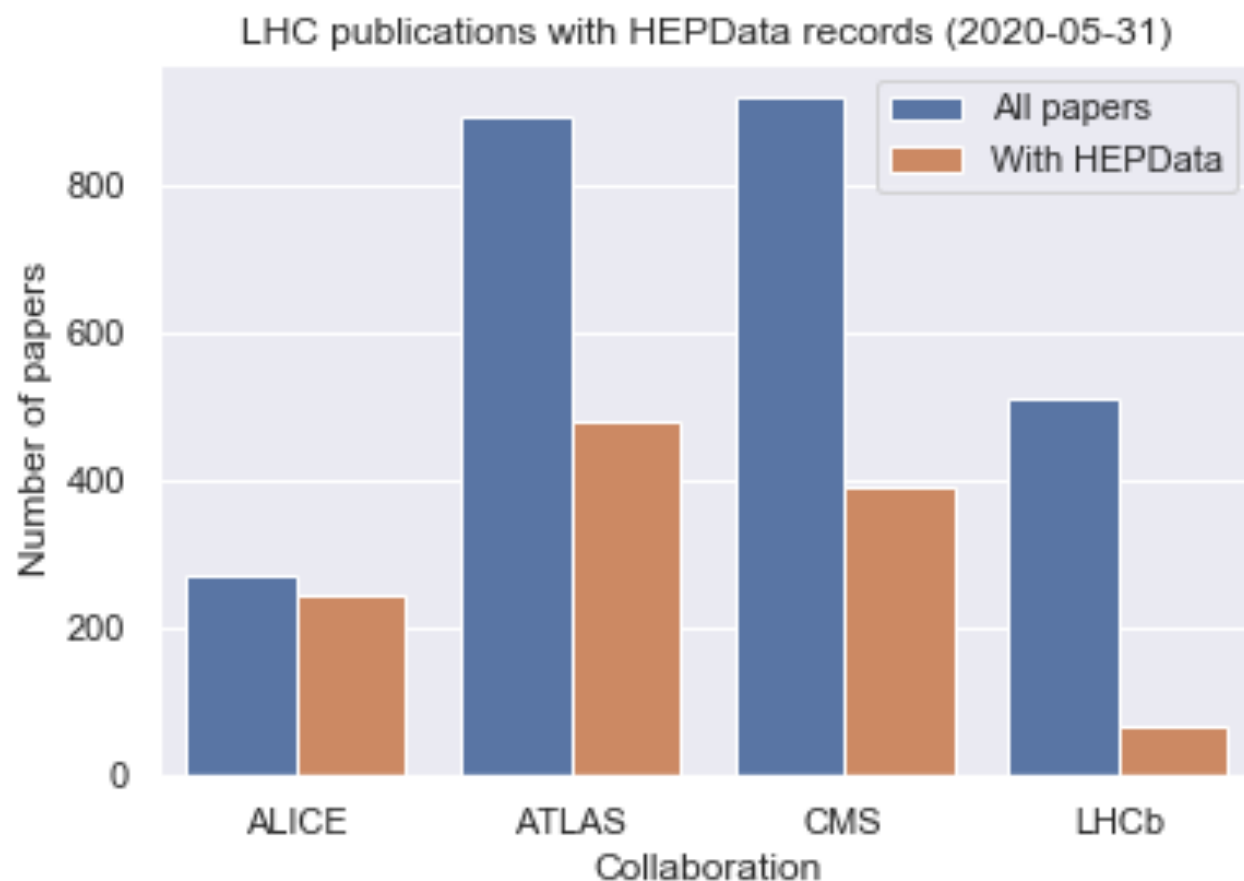
Code: <https://github.com/HEPData>

What is HEPData?

- Unique *open-access* repository for high-level **data** from about 9000 **HEP** (“*hep-ex*” or “*nuc1-ex*”) papers.
- *Complementary* to other HEP information providers, e.g. INSPIRE-HEP (literature), PDG (particle properties), CERN Open Data (event-level data), Zenodo (files).
- Based in **I**nstitute for **P**article **P**hysics **P**henomenology (**IPPP**) at Durham University (UK), going back to 1970s.
- Transition in 2017 to hepdata.net site, hosted at CERN. Partnership with CERN Scientific Information Service. J. Phys.: Conf. Ser. 898 102006 [arXiv:1704.05473]
- 11th May 2020 → new cluster (same as inspirehep.net).

Search from inspirehep.net

- New INSPIRE (blog) shares many features with HEPData.
- INSPIRE publication records link to hepdata.net records.
- Search INSPIRE for publication records with HEPData.
- LHC publications with HEPData (Jupyter Notebook):



(Old) INSPIRE search query:

- `hep-ex` or `nucl-ex`
- Published in a journal
- Not conference paper

ALICE: **90%**
ATLAS: **54%**
CMS: **42%**
LHCb: **13%**

Search from hepdata.net

HEPData

hepdata.net/search/?q=reactions%3A"E%2B+E-+--+>+PI%2B+PI-+GAMMA"

HEPData

About Submission Help Sign in

reactions:"E+ E- --> PI+ PI- GAMMA" Search Reset search Advanced

Max results Sort by Reverse order Showing 3 of 3 results

Date

1997 2009

Collaboration

KLOE 2

CMD-2 1

Subject_areas

HEP Experiment 3

Phrases

Cross Section 3

E+ E- Scattering 3

Exclusive 3

« 1 »

Rivet Analysis **Measurement of $\sigma(e^+e^- \rightarrow \pi^+\pi^-\gamma)$ and extraction of $\sigma(e^+e^- \rightarrow \pi^+\pi^-)$ below 1-GeV with the KLOE detector**

The KLOE collaboration Aloisio, A. ; Ambrosino, F. ; Antonelli, A. ; *et al.*

Phys.Lett. B606 (2005) 12-24, 2005.

Inspire Record 655225 DOI 10.17182/hepdata.41901

We have measured the cross section $\sigma(e^+e^- \rightarrow \pi^+\pi^-\gamma)$ at an energy

0 data tables match query

Rivet Analysis **Measurement of $\sigma(e^+e^- \rightarrow \pi^+\pi^-\gamma(\gamma))$ and the dipion contribution to the muon anomaly with the KLOE detector**

The KLOE collaboration Ambrosino, F. ; Antonelli, A. ; Antonelli, M. ; *et al.*

Phys.Lett. B670 (2009) 285-291, 2009.

Inspire Record 797438 DOI 10.17182/hepdata.57088

Record from hepdata.net

HEPData

hepdata.net/record/ins1203852

HEPData

Search HEPData

Search

About

Submission Help

Sign in

Browse all

Aad, Georges et al.

Last updated on 2014-04-30 20:20

Accessed 1055 times

Cite

JSON

Hide Publication Information

Measurement of ZZ production in pp collisions at $\sqrt{s} = 7$ TeV and limits on anomalous ZZZ and $ZZ\gamma$ couplings with the ATLAS detector

The ATLAS collaboration

Aad, Georges , Abajyan, Tatevik , Abbott, Brad , Abdallah, Jalal , Abdel Khalek, Samah , Abdelalim, Ahmed Ali , Abidinov, Ovsat , Aben, Rosemarie , Abi, Babak , Abolins, Maris

JHEP 1303 (2013) 128, 2013.

https://doi.org/10.17182/hepdata.62535

Journal

INSPIRE

HepData

Resources

Rivet Analysis

Abstract (data abstract)

CERN-LHC. Measurements of the cross section for ZZ production using the $4l$ and $2l2\nu$ decay channels in proton-proton collisions at a centre-of-mass energy of 7 TeV with 4.6 fb^{-1} of data collected in 2011. The final states used are 4 electrons, 4 muons, 2 electrons and 2 muons, 2 electrons and missing transverse momentum, and 2 muons and missing transverse momentum (MET).

Download All

YAML

YODA

ROOT

CSV

Table 2

Data from Page 20 of preprint 10.17182/hepdata.62535.v1/t2

The measured total cross sections. The first systematic uncertainty is the combined systematic uncertainty excluding luminosity, the second is the...

Table 3

Data from Figure 8A

10.17182/hepdata.62535.v1/t3

Normalized ZZ fiducial cross section (multiplied by 10^6 for readability) in bins of the leading reconstructed dilepton p_T for the...

Table 4

Data from Figure 8B

10.17182/hepdata.62535.v1/t4

Normalized ZZ fiducial cross section (multiplied by 10^6 for

Table 3

10.17182/hepdata.62535.v1/t3

Data from Figure 8A

Normalized ZZ fiducial cross section (multiplied by 10^6 for readability) in bins of the leading reconstructed dilepton p_T for the 4 lepton channel. The first systematic uncertainty is detector systematics, the second is background systematic uncertainties.

Resources

https://www.hepdata.i

JSON

cmenergies

7000.0

observables

DSIG/DPT

phrases

Inclusive

Single Differential Cross

Transverse Momentum

reactions

P P --> Z0 Z0 X

Visualize

11,000

10,000

9,000

8,000

7,000

6,000

5,000

RE

P P --> Z0 < LEPTON+ LEPTON- -> Z0 < LEPTON+ LEPTON- -> X

SQRT(S)

7000.0 GeV

Leading dilepton PT [GEV]

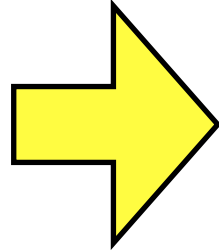
10**6 * 1/SIG(fiducial) * D(SIG(fiducial))/DPT [GEV**-1]

0.0 - 60.0

7000.0 ±1100.0 stat ±79.0 sys,detector ±15.0 sys,background

Data output formats

YAML: native
HEPData format.



submission.yaml
+ YAML data files for each table

- JSON: JavaScript Object Notation.
- CSV: comma-separated values.
- YODA: for inclusion in a Rivet analysis.
- ROOT: binary .root file.

NEW

- New *command-line interface* (CLI) to search and download.
- Work by Giuseppe De Laurentis just in last two weeks.

```
hepdata-cli find 'reactions:"E+ E- --> PI+ PI- GAMMA"' -i inspire
```

```
hepdata-cli download 655225 797438 451091 -f csv -i inspire
```

Code: <https://github.com/HEPData/hepdata-cli>

Modes of data entry

1. Manually harvested from “hep-ex” / “nucl-ex”. HEPData staff extracted tables from `.tex` source.
2. Data points directly submitted by experiments.
 - Pre-2014: no guidelines on preferred format.
 - Early 2014: encourage standard “input” format.
 - Late 2014: introduce online submission system.
 - Early 2017: allow submissions from hepdata.net.

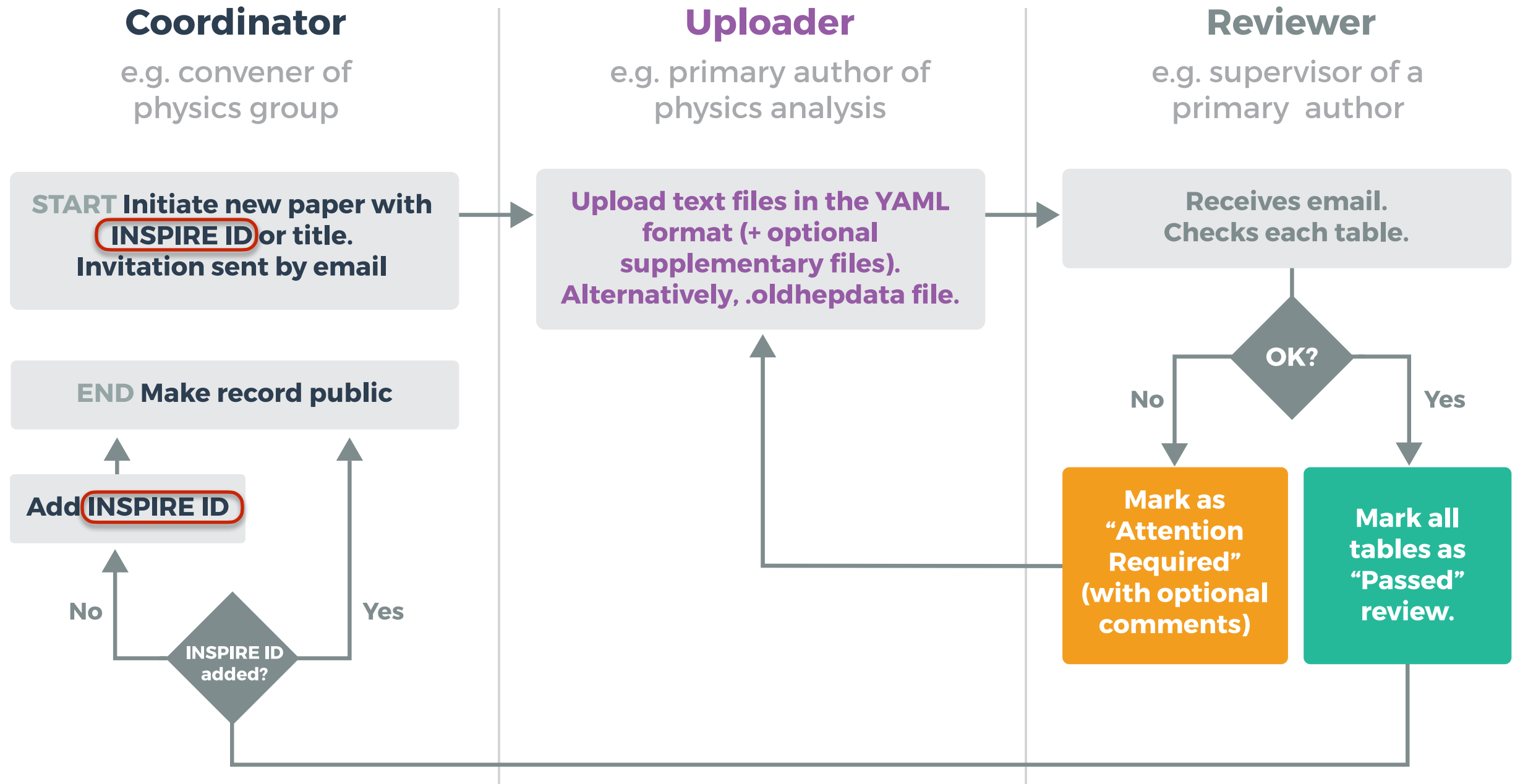
Mode 1. now phased out in favour of mode 2.

Submission system on hepdata.net

STRONG-2020

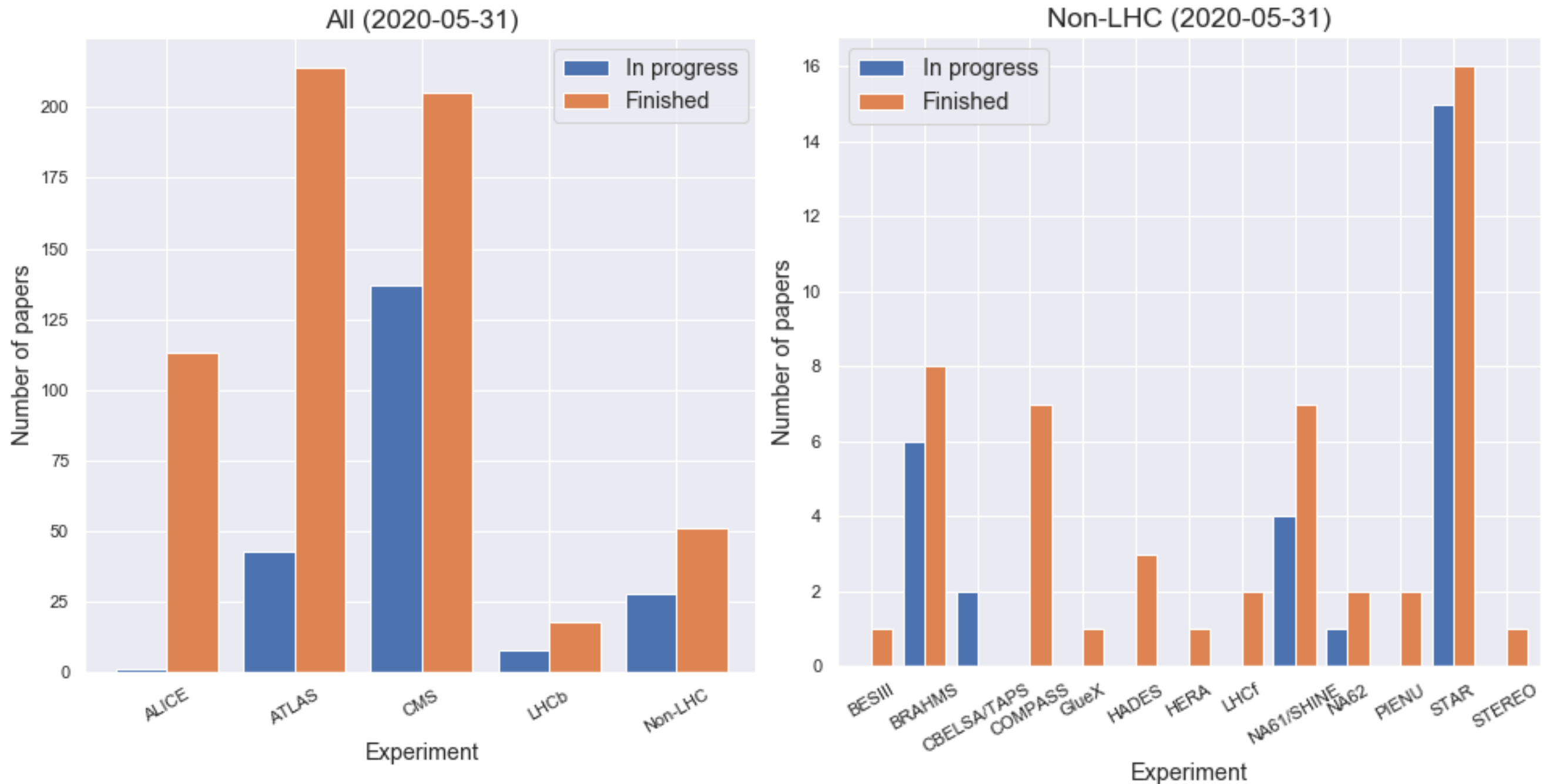
alberto.lusiani@cern.ch

<https://hepdata.net/submission>



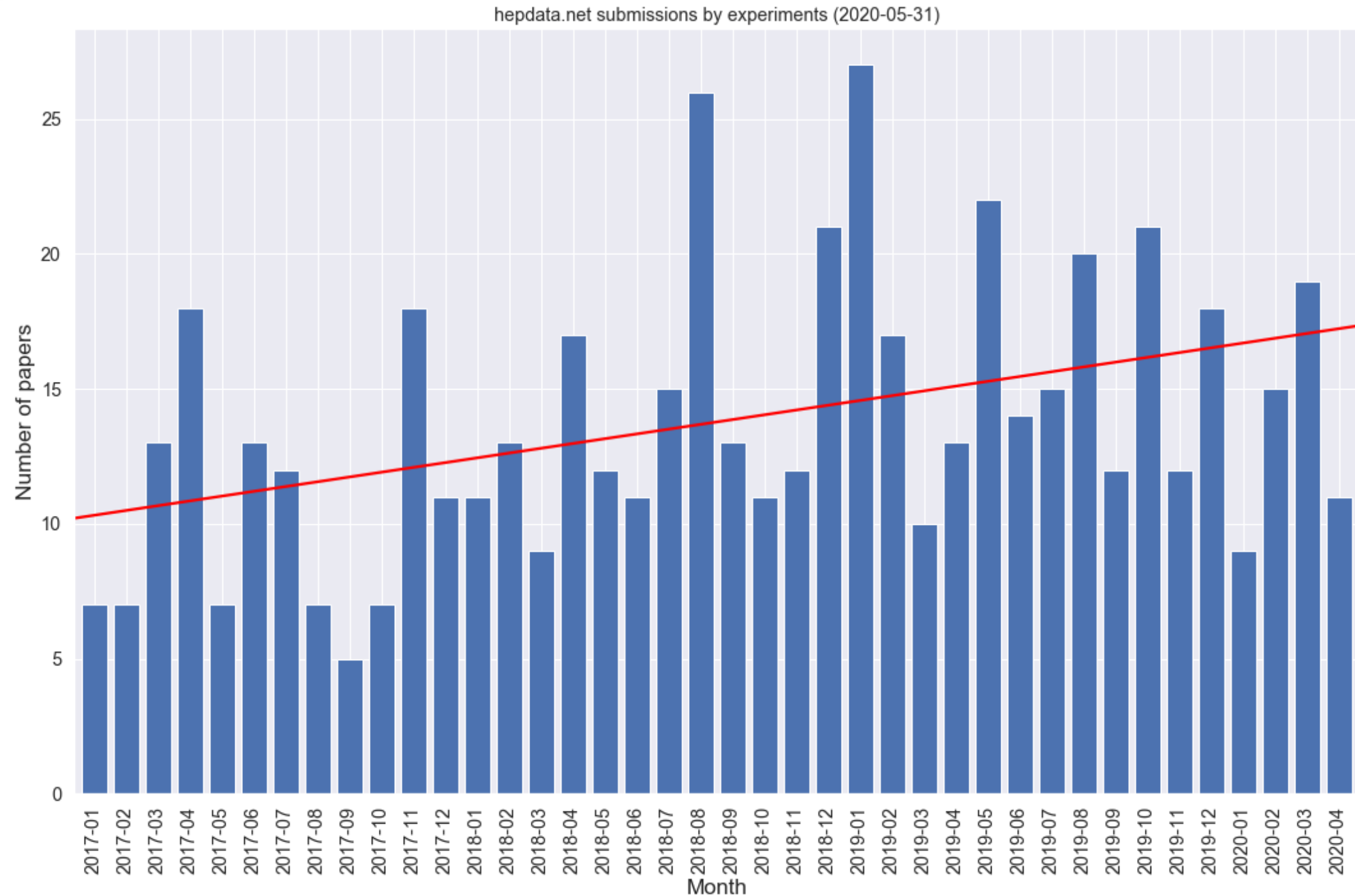
- Also a Sandbox for any user to test uploads without special privileges.

Submission system usage (01/2017-)



- Total of **601** finished submissions, comprising ALICE (113), ATLAS (214), CMS (205), LHCb (18), Non-LHC (51).

Submissions per month from 2017 to 2020



- Year-on-year growth: **125** in **2017**, **171** in **2018**, **201** in **2019**.

submission.yaml file

- Separate YAML “document” for each data table, e.g.

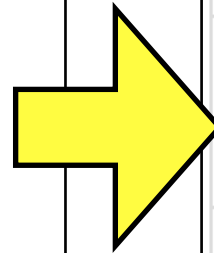
```
---
# This is Table 3.
name: "Table 3"
location: Data from Figure 8A
description: Normalized ZZ fiducial cross section (multiplied by 10^6 for
keywords: # used for searching, possibly multiple values for each keyword
- {name: reactions, values: [P P --> Z0 Z0 X]}
- {name: observables, values: [DSIG/DPT]}
- {name: cmenergies, values: [7000.0]}
- {name: phrases, values: [Inclusive, Single Differential Cross Section,
data_file: data3.yaml
additional_resources:
- {description: Image file, location: figFigure8A.png}
- {description: Thumbnail image file, location: thumb_figFigure8A.png}
```

- Validated against submission_schema.json

YAML data file

- “Independent” and “dependent” variables:

```
independent_variables:
- header: {name: Leading dilepton PT, units: GEV}
  values:
  - {low: 0, high: 60}
  - {low: 60, high: 100}
  - {low: 100, high: 200}
  - {low: 200, high: 600}
dependent_variables:
- header: {name: 10**6 * 1/SIG(fiducial) * D(SIG(fiducial))/DPT, units: GEV**-1}
  qualifiers:
  - {name: RE, value: P P --> Z0 < LEPTON+ LEPTON- > Z0 < LEPTON+ LEPTON- > X}
  - {name: SQRT(S), units: GEV, value: 7000}
  values:
  - value: 7000
    errors:
    - {symerror: 1100, label: stat}
    - {symerror: 79, label: 'sys,detector'}
    - {symerror: 15, label: 'sys,background'}
  - value: 9800
    errors:
    - {symerror: 1600, label: stat}
    - {symerror: 75, label: 'sys,detector'}
    - {symerror: 15, label: 'sys,background'}
  - value: 1600
    errors:
    - {symerror: 490, label: stat}
    - {symerror: 41, label: 'sys,detector'}
    - {symerror: 2, label: 'sys,background'}
  - value: 80
    errors:
    - {symerror: 60, label: stat}
    - {symerror: 2, label: 'sys,detector'}
    - {symerror: 0, label: 'sys,background'}
```



RE	P P --> Z0 < LEPTON+ LEPTON- > Z0 < LEPTON+ LEPTON- > X
SQRT(S)	7000.0 GeV
Leading dilepton PT [GEV]	$10^{**6} * 1/SIG(fiducial) * D(SIG(fiducial))/DPT [GEV^{**-1}]$
0.0 - 60.0	7000.0 ± 1100.0 stat ± 79.0 sys,detector ± 15.0 sys,background
60.0 - 100.0	9800.0 ± 1600.0 stat ± 75.0 sys,detector ± 15.0 sys,background
100.0 - 200.0	1600.0 ± 490.0 stat ± 41.0 sys,detector ± 2.0 sys,background
200.0 - 600.0	80.0 ± 60.0 stat ± 2.0 sys,detector

- Validated against data_schema.json

Submission documentation

- Documentation at hepdata-submission.readthedocs.io. Includes example Python scripts ([simple](#), [complicated](#)).
- HEPData YAML files checked against [JSON schema](#) by [validation code](#) during submission. [Offline validation script](#).
- [hepdata_lib](#) package by *Clemens Lange* and *Andreas Albert*. Library to read in text/ROOT and write HEPData YAML. https://github.com/HEPData/hepdata_lib
- Similar Python package by *Christian Holm Christensen*. <https://gitlab.com/cholmcc/hepdata>

Summary

Email: info@hepdata.net
Forum: hepdata-forum.cern.ch

 Follow @HEPData

- hepdata.net site replaced previous hepdata.cedar.ac.uk in 2017.
- Responsibility for data submission now held by experiments.
- Submission system widely used by LHC collaborations.
Non-LHC experiments welcome to upload their data.
- [Alberto Lusiani](#) is HEPData Coordinator for [STRONG-2020](#).
- See hepdata.net/submission for overview of procedure.
- Documentation and software tools available to assist upload.
- [hepdata-cli](#) to search/download from the command line.