

# INSPIRE

## A service for HEP

Benefits and Costs

Kirsten Sachs

Bibliothekskommission, Hamburg, 28.11.2022

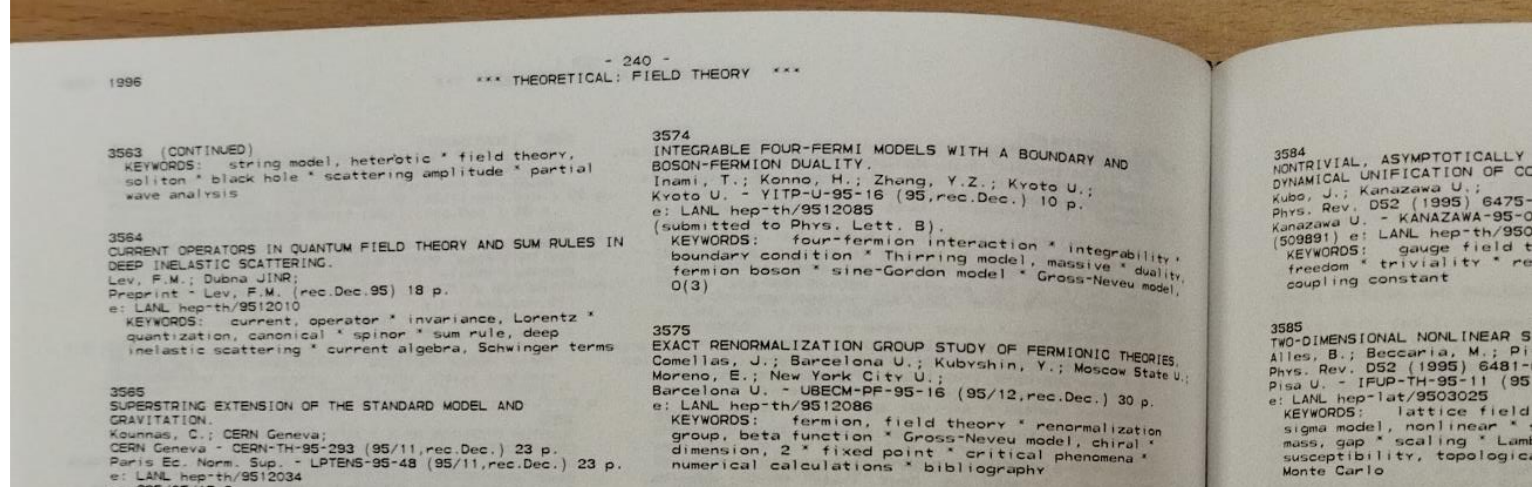
HELMHOLTZ



# History

## Some background

- 1963 HEPIndex
- 1974 **SPIRES-HEP** by **SLAC** & **DESY**  
access via SLAC account
- Mid Remote access via email
- 1980s **Fermilab** joined
- Dec. First web-server outside Europe
- 1991 Close collaboration with arXiv
- 2011 - Transition SPIRES → INSPIRE
- 2013 **CERN** joined
- 2014 - **IHEP**, **IN2P3** joined
- 2019 No curation at SLAC anymore
- 2020 Transition to Invenio3
- 2022 Expansion to QIS



- Literature database with focus on High Energy Physics
  - incl. gray literature: proceedings, books, thesis
  - used by ~100% of HEP physicists daily
- Close HEP community allows dedicated service
  - ~ 10000 active users : authors == readers
- 1.5M records in literature
- Jobs, Conferences, Authors, Experiments, Institutions
- Close collaboration with:
  - arXiv (preprint server)
  - NASA-ADS (literature for astrophysics)  
everything NASA wants, will expand to e.g. earth science

# What are we doing - Highlights

## What is the benefit

Receive metadata feeds for most content

Value added by identification & enrichment

### Preprint → Journal Publication

Very Hairy Inflation [*arXiv:2112.13861*] ==

General double monodromy inflation

[*10.1103/PhysRevD.105.103527*]

match & merge is time-consuming when  
metadata are not identical (title & authors)

(Mis-)Matches are communicated  
from one database to the other

### Name → Author

M.Meyer, U. Hamburg →

Mareike Meyer (0000-0003-2436-8195)

Manuel Meyer (0000-0002-0738-7581)

### Reference → Citation

Most of the time this is relatively easy. However,

*J. High Energ. Phys.* 2022, 62 (2022).

[https://doi.org/10.1007/JHEP10\(2022\)062](https://doi.org/10.1007/JHEP10(2022)062)

is the official citation, but the issue (month) is missing!  
We have to search for author & title (or use DOI)

### Two **Physics** Journals

Physics 4 (2022) 12 → 10.3390/physics4010002 – MDPI

Physics 4 (2011) 15 → 10.1103/Physics.4.15 – APS

Btw: Citation numbers have an error!

And we have no Monte Carlo to estimate how large it is.



M.Meyer

Bob or Robert Smith, Cheng Li

INSPIRE HEP

authors

LiteratureAuthorsJobsSeminars

Mareike Meyer (Hamburg U.)

Experiments: CERN-LHC-CMS  
Author Identifier: Mareike.Meyer.1

edit

Research works (684)

Cited By

Date of paper

684 results | cite all | claim

20132022

Number of authors

Single author4

10 authors or less4

Exclude RPP

Exclude Review of Particle Physics684

Geoneutrino Detection and Other Non-Standard Neutrino Phenomena of Borexino

BOREXINO Collaboration • Sandra Zavatarelli (Osaka U.)  
Published in: Moscow Univ.Phys.Bull. 77 (2022) 2, 4  
Conference on Elementary Particle Physics

DOI

cite

edit

claim

First Detection of Solar Neutrinos from the Borexino Experiment

BOREXINO Collaboration • N. Rossi (Gran Sasso) et al.  
Published in: Moscow Univ.Phys.Bull. 77 (2022) 2, 4  
Conference on Elementary Particle Physics

DOI

cite

edit

claim

Search for new particles in an extended kinematic range at  $\sqrt{s} = 13$  TeV

CMS Collaboration • Armen Tumasyan (Yerevan Ph

INSPIRE HEP

authors

LiteratureAuthorsJobsSeminarsConferencesMore...

Manuel Meyer (Hamburg U., Inst. Exp. Phys. II)

Experiments: ALPS, FERMI-LAT, HESS, CTA  
Author Identifier: Manuel.Meyer.1  
PhD Advisor: Dieter Horns

2021-present  
JUNIOR, Hamburg U., Inst. Exp. Phys. II

2019-2021  
POSTDOC, Erlangen - Nuremberg U., ECAP

2019-2019  
POSTDOC, DESY

Show all positions (6)

Updated on Aug 31

Research works (121)

Cited By

Date of paper

121 results | cite all | claim

20092022

Number of authors

Single author3

10 authors or less46

Constraints on axionlike particles from a combined analysis of three flaring *Fermi* flat-spectrum radio quasars

James Davies (Oxford U.), Manuel Meyer (Hamburg U.), Garret Cotter (Oxford U.) (Nov 7, 2022)  
e-Print: 2211.03414 [astro-ph.HE]  
Note: 18 pages, 14 figures; submitted to Physical Review D

pdf

cite

edit

claim

reference search

0 citations

Interpretation of multi-TeV photons from GRB221009A

Ali Baktash (Hamburg U.), Dieter Horns (Hamburg U.), Manuel Meyer (Hamburg U.) (Oct 13, 2022)  
e-Print: 2210.07172 [astro-ph.HE]

AUTHORS	
Meyer, M	4k
Meyer, Michael	281
Meyer, M R	190
Meyer, Michael R	184
Meyer, M A	105
Meyer, Martin	103
Meyer, Matthias	94
Meyer, M S	72
Meyer, Manuel	58
Meyer, M J	50
Meyer, Manfred	45
Meyer, Michael T	44
Meyer, M Wilhelm	31
Meyer, Max F	30
Meyer, Markus	22
Meyer, Michael A	21
Meyer, M K	19
Meyer, Markus R	19
Meyer, Mathias	18
Meyer, Miriah	17
Meyer, Mike	16
Meyer, M L	15
Meyer, Madeleine	14
Meyer, Martin J	14
Meyer, Mathieu	14
Meyer, Mervin	14
Meyer, Morten	14
Meyer, Michael B	12
Meyer, Michael G	12
Meyer, M C	11

DESY. | INSPIRE | Kirsten Sachs, 28.11.2022

Page 4

# NASA-ADS

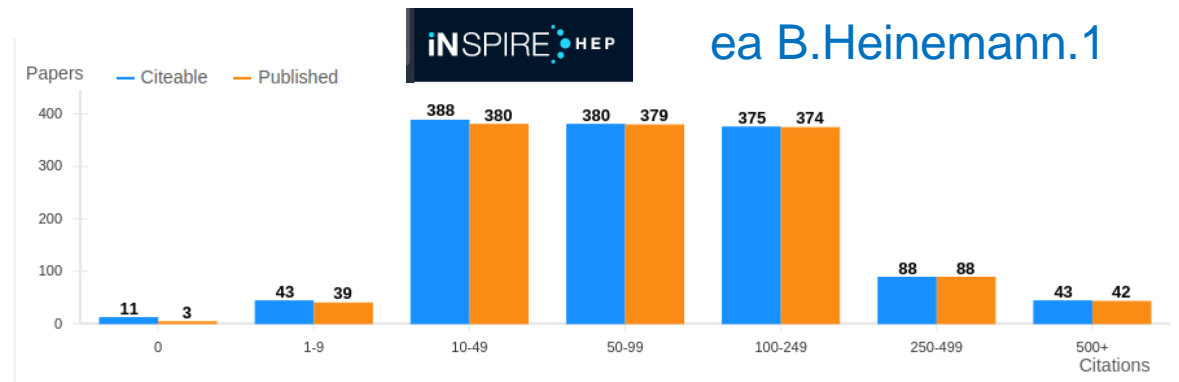
## What is different

- From NASA for NASA
- More content: 17M records (1.5M in INSPIRE)
- Less (gray) HEP content compared to INSPIRE
- Affiliations from feeds, not for preprints
- Started to add standardized affiliations recently
- Search for stars & galaxies
- No search for experiments, conferences, reactions
- No author profiles; only facet for name variants

Due to the larger corpus author disambiguation is more difficult, almost impossible

## Heinemann, Beate:

INSPIRE: 1339 / ADS: 1434



Track reconstruction at the LUXE experiment using quantum algorithms #1

Arianna Crippa (DESY, Zeuthen and Humboldt U., Berlin), Lena Funcke (MIT, Cambridge, CTP and IAIFI, Cambridge), Tobias Hartung (Northeastern U.), Beate Heinemann (DESY and Freiburg U.), Karl Jansen (DESY, Zeuthen) et al. (Oct 24, 2022)

e-Print: [2210.13021](#) [hep-ex]

Note: 7 pages, 6 figures, Proceedings of the Connecting The Dots workshop 2022 (CTD2022)

pdf cite edit claim reference search 0 citations



Author: "Heinemann, Beate" not aff: "DESY"  
344 records

- 1 2022arXiv221013021C 2022/10   
[Track reconstruction at the LUXE experiment using quantum algorithms](#)  
Crippa, Arianna; Funcke, Lena; Hartung, Tobias *and 8 more*
- 2 2022arXiv220912908B 2022/09   
[Using nonlinear Breit-Wheeler to test nonlinear vacuum birefringence](#)  
Borysov, O.; Heinemann, B.; Ilderton, A. *and 2 more*
- 3 2022NucFu..62d2022S 2022/04 cited: 3   
[Experimental confirmation of efficient island divertor operation and successful neoclassical transport optimization in Wendelstein 7-X](#)  
Sunn Pedersen, Thomas; Abramovic, I.; Agostinetti, P. *and 494 more*
- 4 2022NucFu..62d2006S 2022/04 cited: 2   
[Progress from ASDEX Upgrade experiments in preparing the physics basis of ITER operation and DEMO scenario development](#)  
Stroth, U.; Aguiam, D.; Alessi, E. *and 408 more*
- 5 2022JPhCS2244a2051W 2022/04

# What does it cost

## Resources

CERN	DESY	Fermilab	IHEP	IN2P3	SLAC
5 FTE	3.1* FTE * incl. 1 temporary	3.2 FTE	2.5 FTE	2 FTE	0 FTE
•development •tech.operation •curation(CERN) •user support	•harvesting •selection •matching •merging •curation •user support •conferences	•author profiles •jobs •curation (QIS) •user support	•curation •chinese authors and affiliations •jobs	•curation (FR)	no curation since 10/2018  tech. support for invenio1 until 10/2021

### Only @ DESY:

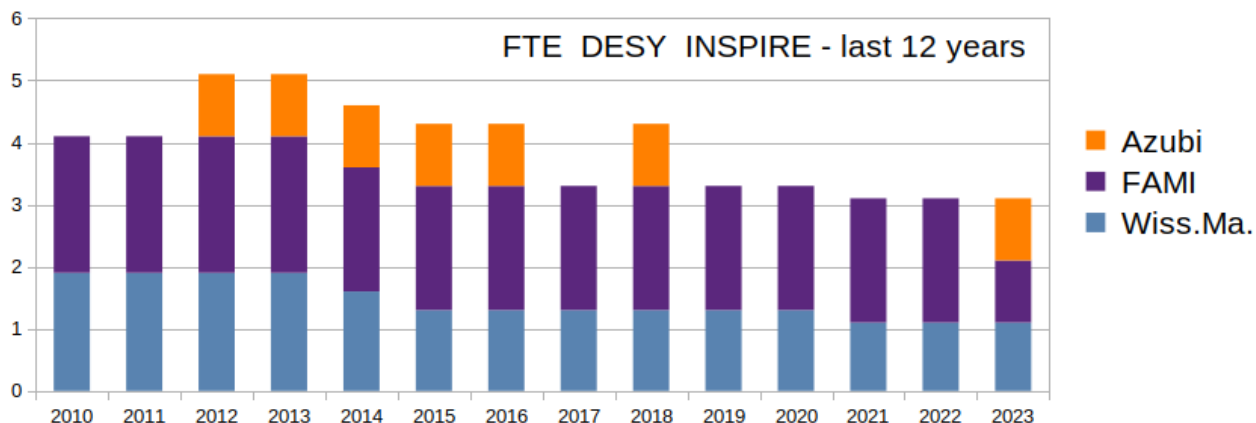
- General curation, others for dedicated (their) literature
- Physics know-how, needed for selection by content
- Harvesting of many small publishers, thesis server, etc. (APS, Elsevier harvested by CERN)

Different competence, limited resources ⇒ it would be very difficult if one of CERN, DESY, Fermilab drops out

# End of INSPIRE@SLAC – decreasing resources

## What did we do?

- DESY took over, we had 3 curators at that time
- No curation of references for arXiv preprints (back now)
- Less gray literature  
e.g. we used to search for proceedings
- More automatization
- author.xml  
author-IDs and affiliations provided by authors



## Consequences:

- Data quality went down slightly
- Less new content
- The easier problems are automatized.  
What is left over is more difficult,  
needs more time for development  
or needs human curation.  
Or we are willing to accept more mistakes,  
reducing the data quality further.

If this goes on,  
a critical drop of quality might be reached.

Without high quality data,  
INSPIRE becomes pointless

# Is it worth it?

To be answered by users and funders

## The obvious:

- Search for literature by
  - Publication note, identifier
  - Author, Title
  - Content (keywords, subject, fulltext)
- Get link to fulltext
  - Preprint server (Open Access)
  - Publisher
- Citations
  - Articles
  - Numbers (with error)

## The special:

- Author profiles
  - A persons literature list, out of the box
- List of publications incl. preprints
  - from Institute of the last year / 2 weeks
- Experiments
- Conferences
- ...

**INSPIRE is a service  
for the HEP community.  
by CERN, DESY, Fermilab, IHEP, IN2P3**



# Thank you

# Other Collections – INSPIRE only

## Jobs, Experiments, Conferences

The image displays four overlapping screenshots of the INSPIRE HEP website, illustrating various collections and search results.

- Jobs Collection:** The top-left screenshot shows the 'Jobs' collection. It features a search bar, navigation tabs (Literature, Authors, Jobs, Seminars, Conferences, More...), and a list of job opportunities. Filters include 'Select Job Filters' (hep-ph, PostDoc, Europe, open) and 'Subscribe to mailing list'. Results include 'Postdoctoral Research Associate in Theoretical Particle Physics - Collider and Low Energy Phenomenology (10358)' and 'Postdoctoral Position in Cosmology (Madrid, IFT - E)'. A 'Papers per year' bar chart is visible at the bottom.
- Conferences Collection:** The top-right and middle-right screenshots show the 'Conferences' collection. The top-right screenshot displays a list of upcoming conferences, including '12th edition of the international QCD@LHC workshop (QCD@LHC2022)' and 'Supernovae in the Gravitational Wave Detection Era'. The middle-right screenshot shows a search result for '19th Workshop of the LHC Higgs Working Group'.
- Experiments Collection:** The bottom-left screenshot shows the 'Experiments' collection. It features a search bar, navigation tabs, and a list of experiments. The 'ICECUBE' experiment is highlighted, with details about its proposal, successor experiments (ICECUBE-PINGU, IceCube-Gen2), and associated articles.
- Search Results:** The bottom-right screenshot shows a search result for '31st Linear Accelerator Conference (LINAC22)'. It includes a list of results (231 results), a citation summary, and a list of related articles, such as 'Prototype HB650 Transportation Validation for the PIP-II Project' and 'Run 2 of the Advanced Plasma Wakefield Experiment (AWAKE) at CERN'.

# Tickets

## Feedback

I really appreciate your hard work on maintaining a very nice and useful site. I use it on a daily basis for years now and I have never found any flaws or errors.

It would be really nice if you put my affiliation in that publication. I don't understand why you didn't do that in the first place. There is also no tool to do it myself either. Do it.

I am trying to import my publications from INSPIRE to ORCID but I faced many problems, unfortunately I need to fix this issue in a couple of days maximum since it is needed for an important application...

You show some bibliometric data in the listings of the inspirehep.net paper database, such as the number of papers, number of citations, h-index, etc

This is very helpful :)

I would like to suggest you to also show another important number: the average number of authors in the papers, which is very important for several indicators, such as the size of the collaborations, the hm\_index etc. Are you planning to show any of these numbers in the future?

Dear Madame/Sir:

on 29 Sep 2022 I noticed a #citation drop of 76 in general and for Journals by 65 using: find a blumlein or bluemlein.

What has been the reason for this?

Best Johannes Bluemlein

NB: total #citations 20000

# Collaboration agreement

CERN, DESY, Fermilab, IHEP, IN2P3 and SLAC

Excerpts:

This Agreement establishes a common understanding among the Parties of the collaborative effort required for the further development, maintenance and operation of INSPIRE.

It is expressly acknowledged that this Agreement is not legally binding and each Party's involvement in the Project is on a "best-effort" basis and might depend on such factors as resources made available by funding agencies that are not signatories of this Agreement.

The Project is managed by a "Board of Directors".

The Project is advised by an independent "Advisory Board".

List of tasks in Appendix

## Board of Directors

One person per institute

Meets once per month (VC)

## Advisory Board

Leading physicists, member of ADS

Pre Corona: one in-person meeting per year

Now: ~2h VC on specific topic

## AAHEP

Meeting of arXiv, ADS & INSPIRE

Every ~2 years meeting of information providers, incl. publisher

# What does it cost

@ DESY

## Resources

- 2 Scientists (1.1 FTE)
- 2 FAMI (2 FTE):
  - 1 permanent
  - 1 finished apprentice - max. 2 years

## Curation needs resources

CERN and IN2P3 do their own publications  
incl. 3000+ author papers from ATLAS & CMS

Fermilab has 1 FTE for QIS

The bulk of general work is done at DESY

## Knowledge

- harvesting
  - many small publisher
  - large publishers @CERN (APS, Elsevier)
- selection (needs scientists)
- keywords (needs scientists)

## Re-distribution of work

When DESY can't fulfill its commitment  
it will be difficult for other labs to compensate.

Missing know-how and resources