PAUL SCHERRER INSTITUT

PSI

Reinterpretation Tools

An overview, challenges, and how to get involved

Clemens Lange (Paul Scherrer Institute PSI)

LHCP 2023, Belgrade, Serbia

25th May 2023

Introduction

- The (annual) meeting for LHC analysis reinterpretation is the “forum on the interpretation of the LHC results for BSM studies”
- Last forum took place in December 2022: Indico agenda
- This talk is largely based on updates presented at this workshop

- Reports originating from this workshop series:
  - Reinterpretation of LHC Results for New Physics: Status and Recommendations after Run 2 (arXiv:2003.07868)
  - Publishing statistical models: Getting the most out of particle physics experiments (arXiv:2109.04981)

PAUL SCHERRER INSTITUT

Recasting tools

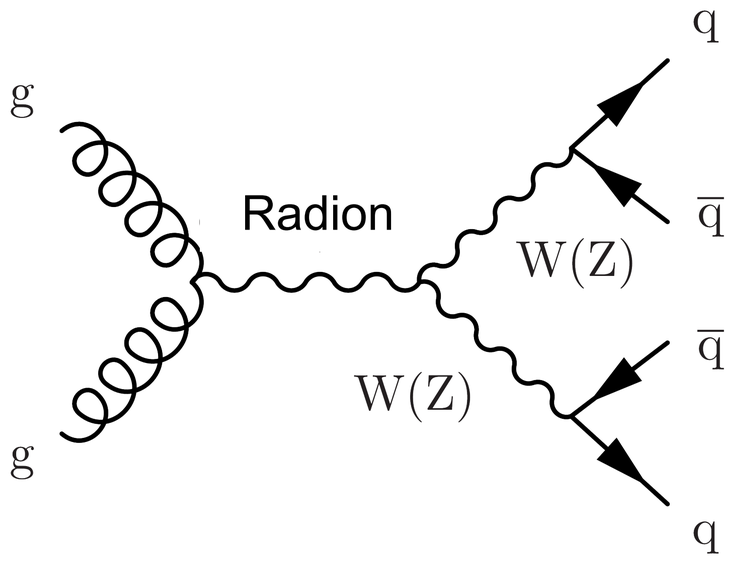
A largely complete list of recasting tools can be found at https://twiki.cern.ch/twiki/bin/view/LHCPhysics/RecastingTools

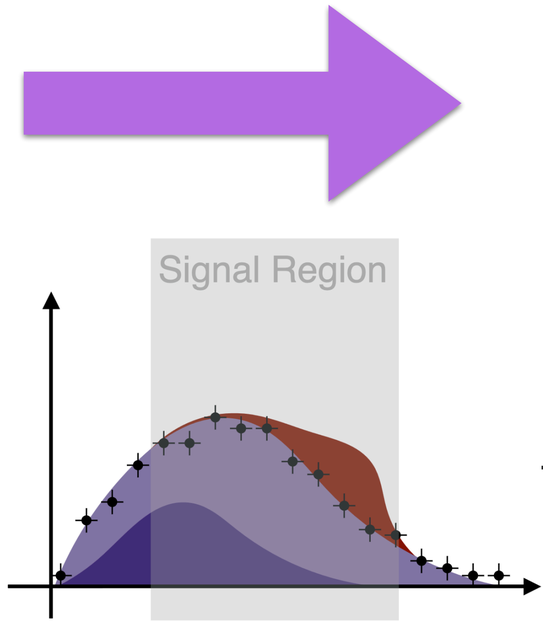
Recasting: the idea

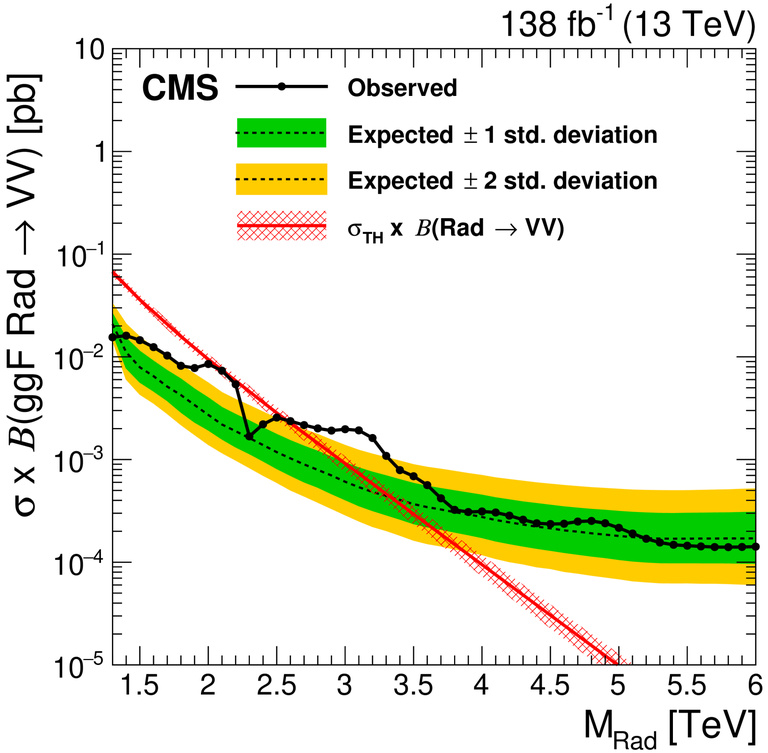
arXiv:2210.00043

> Experiments typically choose benchmark models when performing their measurements/searches

Model | **Analysis** | **Limit/Measurement**



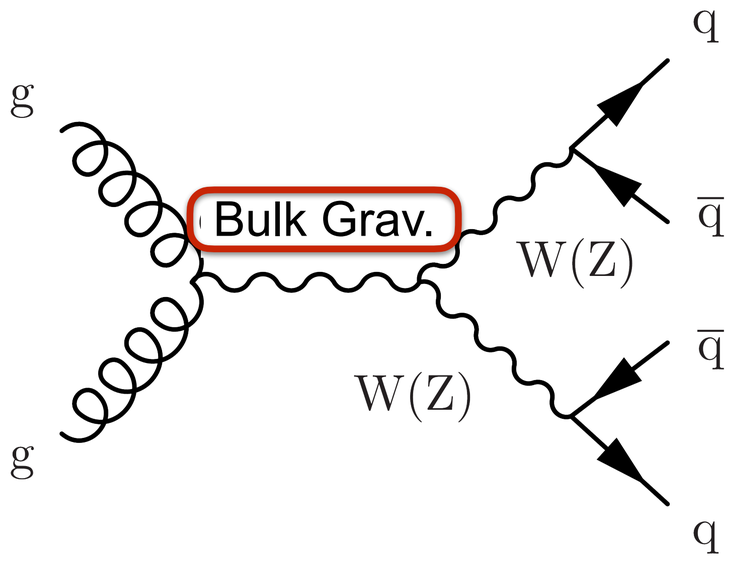


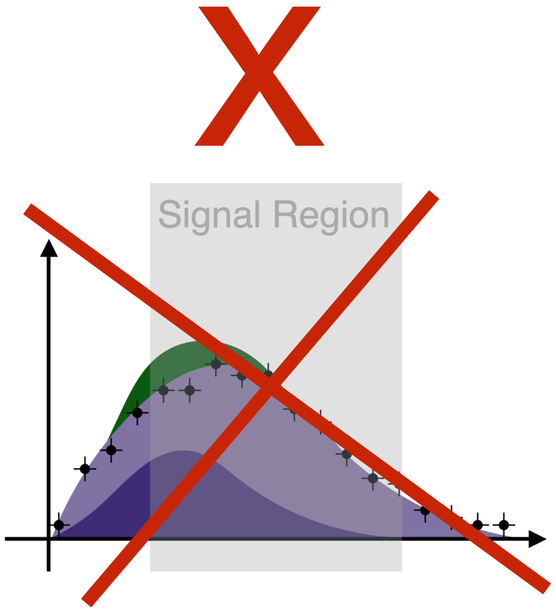


Recasting: the idea

- Experiments typically choose benchmark models when performing their measurements/searches
- What if your favourite model has not been used in the analysis?

Model **Analysis** **Limit/Measurement**





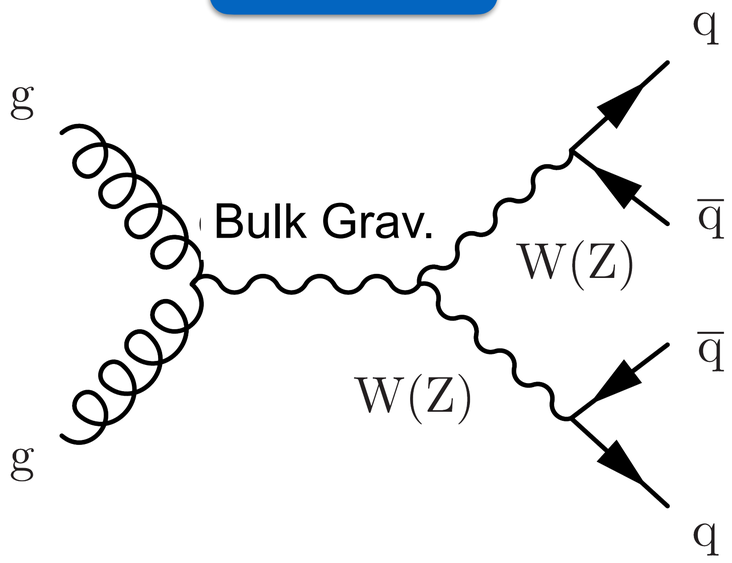
?

Recasting: the idea

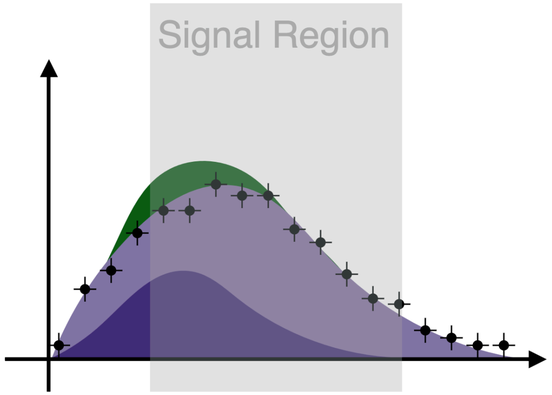
arXiv:2210.00043
*this is actually not a recast

- Experiments typically choose benchmark models when performing their measurements/searches
- What if your favourite model has not been used in the analysis?
- Recasting is (one of) the solution(s)

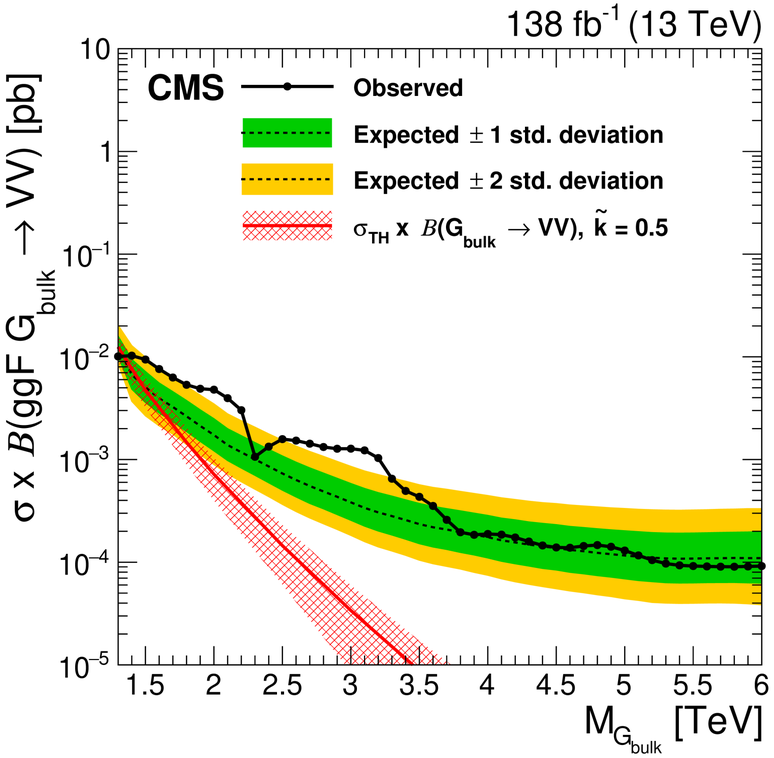
Model



Re-run actual analysis,
Run simplified (particle-level) analysis,
…



Limit/Measurement



Rivet

- Rivet: “Robust Independent Validation of Experiment and Theory”
- Originally developed as a **toolkit for the validation of Monte Carlo event generators**…
  - MC tuning → Professor, New Physics limits → Contur
- … now tool with **most widespread use by LHC experiments**
  - >1000 analyses from LEP, LHC and other experiments integrated
- **Preserves complete particle-level analyses in form of code**
  - Often directly integrated into experiment software, see e.g. CMSSW interface
- Initially only suitable for **unfolded measurements**
- Features such as detector smearing and jet algorithms (via FastJet) allow adding of **BSM analyses**

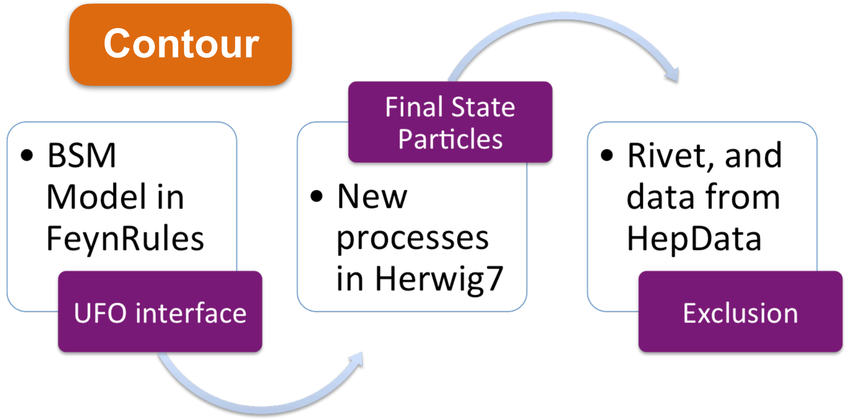
Rivet: adoption

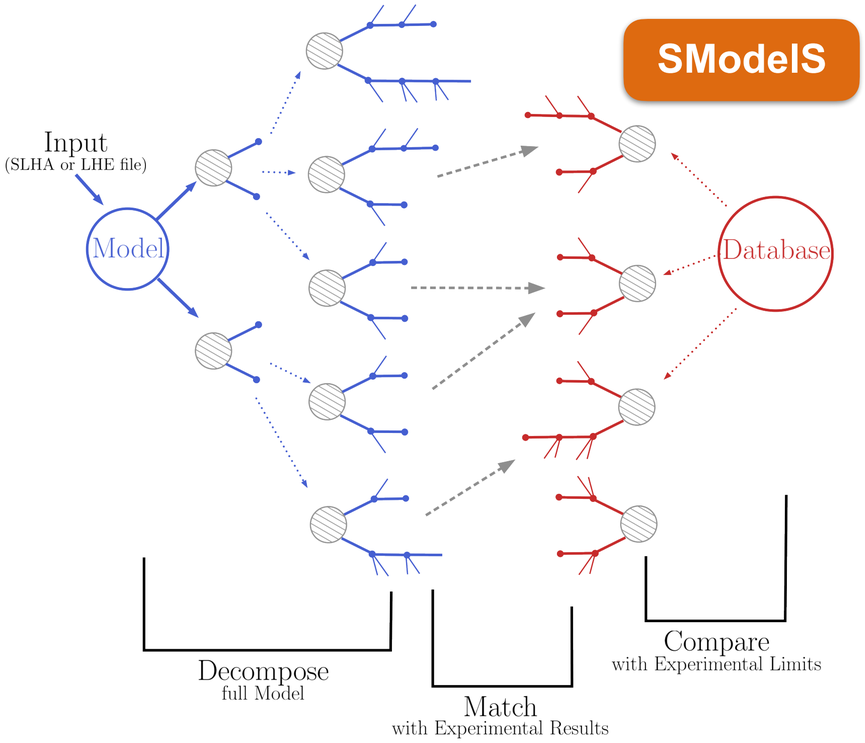
Key	ALICE	ATLAS	CMS	LHCb	Forward	HERA	e^+e^- (≥ 12 GeV)	e^+e^- (≤ 12 GeV)	Tevatron	RHIC	SPS	Other
Rivet wanted (total):	290	337	465	180	17	476	703	513	1118	477	56	3
Rivet REALLY wanted:	36	39	92	15	0	15	1	1	9	2	5	1
Rivet provided:	**29/319 = 9%**	**190/527 = 36%**	**104/569 = 18%**	**17/197 = 9%**	**8/25 = 32%**	**34/510 = 7%**	**192/895 = 21%**	**347/860 = 40%**	**58/1176 = 5%**	**8/485 = 2%**	**4/60 = 7%**	**131/134 = 98%**

- While tool with largest adoption, large number of analyses missing (as of September 2022)
  - e.g. 36 and 18% of “wanted” ATLAS and CMS analyses, respectively
- Experimentalists: writing a “plugin” still means additional work as part of publication process
- Theorists: ideally use **both** ATLAS and CMS plugins when available

Other tools

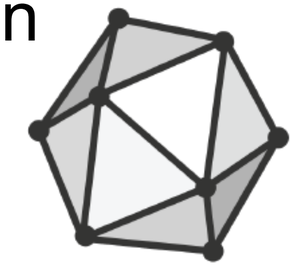
- Contour (Constraints On New Theories Using Rivet)
  - Can replace FeynRules/Herwig7 by other tools
- MadAnalysis 5: originally for phenomenological studies, recast analyses using Delphes using Python/ C++
  - Also features an analysis database with tens of analyses
- SModelS: “decompose BSM collider signatures presenting a Z_2-like symmetry into Simplified Model Spectrum”
  - Simplified results database based on ~160 analyses
- GAMBIT: global fits with support for parameter scans
- SimpleAnalysis: Simplified ATLAS SUSY analysis framework
- CheckMATE, HiggsBounds, …





Rivet: challenges

- Physics analyses become more and more complicated
- In particular, challenging to implement analyses at particle level when machine-learning (ML) techniques are used
  - So far mostly solved using efficiency maps



- ONNX Runtime (with Open Neural Network Exchange format) seems to be common denominator for exchange and use of ML models
  - See e.g. report on ATLAS experience and dedicated session at Reinterpretation Forum 2022

Old analyses are unlikely to be added since analysis code not preserved

- Experiments need to make sure to add plugins at publication time

Analysis description languages

- Idea: agree on common "**domain specific language**" to specify collider physics analysis, i.e. describe analysis objects and selection cuts
  - Avoid ambiguity of descriptions in paper publications
  - Declarative approach: focus on physics content

Available tools:

- ADL/CutLang V2: text-based language with its own runtime interpreter (CutLang) and transpiler (adl2tnm)
- FuncADL: Python-based, inspired by functional programming and query languages

See e.g. ACAT 2021 presentation for overview/comparison

Simplified to full likelihoods

- CMS has released some simplified likelihoods since 2017, ATLAS 2019:

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \mathcal{P}(\text{data}|\mu \cdot s(\boldsymbol{\theta}) + b(\boldsymbol{\theta})) \cdot p(\tilde{\boldsymbol{\theta}}|\boldsymbol{\theta}),$$

Full likelihood

$$\mathcal{L}_S(\mu, \boldsymbol{\theta}) = \prod_{i=1}^{N} \frac{(\mu \cdot s_i + b_i + \theta_i)^{n_i} e^{-(\mu \cdot s_i + b_i + \theta_i)}}{n_i!} \cdot \exp\left(-\frac{1}{2}\boldsymbol{\theta}^T \mathbf{V}^{-1} \boldsymbol{\theta}\right),$$

Simplified likelihood

$\mathbf{V}$: covariance matrix

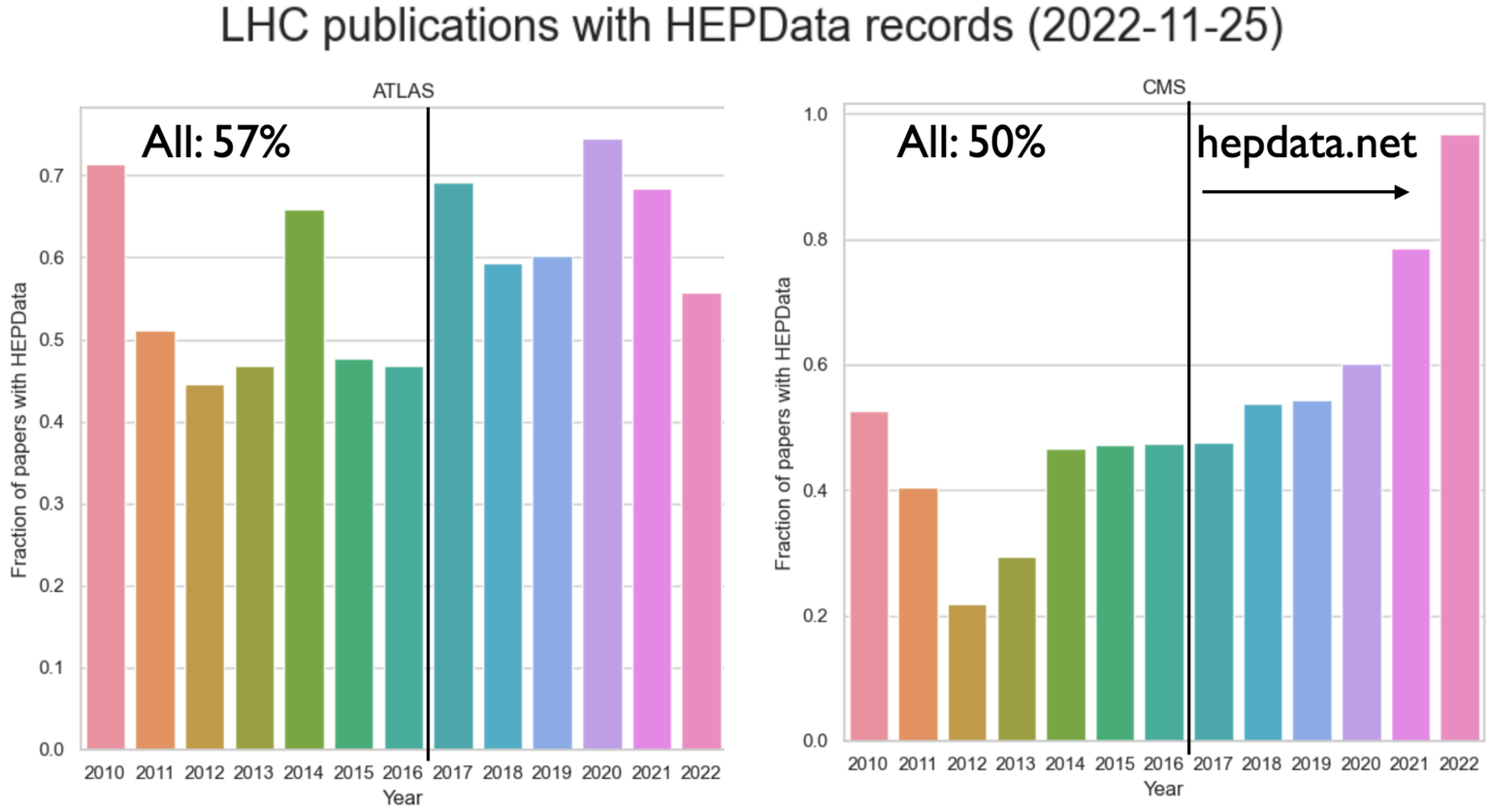
  - See also “Simplified likelihood for the re-interpretation of public CMS results” (CMS-NOTE-2017-001)
  - Earlier (since 2013) also ATLAS and CMS likelihood scan results

- ATLAS has published some **full likelihoods** in JSON format for use with pyHF since 2019
  - See “Reproducing searches for new physics with the ATLAS experiment through publication of full statistical likelihoods” (ATL-PHYS-PUB-2019-029)

More details: “publishing statistical models” session at RiF 2022

HEPData

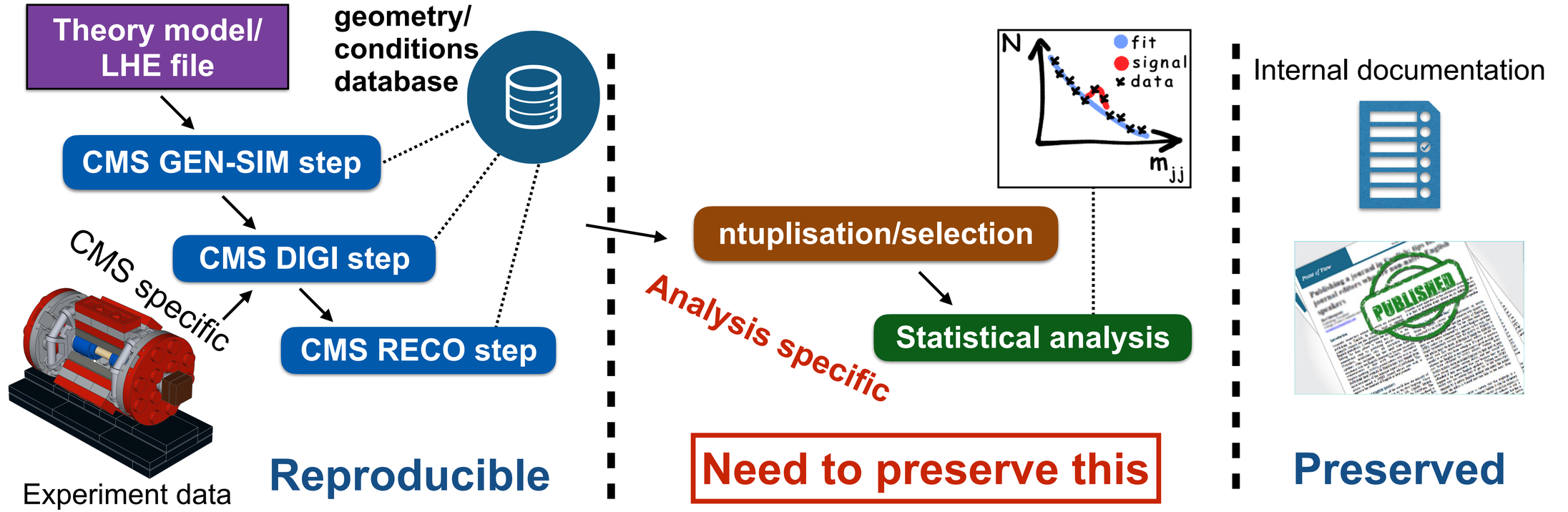
- Open-access repository for **tabular high-level data** from almost 10,000 HEP publications: https://www.hepdata.net/
- Figures and tables from papers in YAML format
  - HEPData entry creation now part of approval process in CMS



See "News from HEPData"

Open Data + REANA

- LHC experiments agreed on common Open Data and CERN-wide Open Science Policy
  - CMS already provides all of Run-1 plus fraction of 2015 data and accompanying simulation
  - Expect open data for scientific purposes from others soon
- With (research-level) Open Data, could rerun full analyses
  - Otherwise, recasting only possible within experiment
- Challenge: preserve analysis code (required in some ATLAS groups) as container images
- Then could rerun, e.g. using REANA (connect analysis steps using workflow languages)



Next LHC Reinterpretation Forum (RiF)

- 8th Forum on the interpretation of the LHC results for new physics: 29th August - 1st September; IPPP Durham, UK: Indico agenda
- Focus of this workshop will lie on recent advances in:
  - Storage and (re)use of theoretical predictions, including event samples
  - Communication and reuse of statistical and machine-learned models
  - Combined/global interpretation of searches and measurements
- Abstract deadline: 16th July 2023


PAUL SCHERRER INSTITUT
PSI

Further reading

- Reinterpretation of LHC Results for New Physics: Status and Recommendations after Run 2 (arXiv:2003.07868)
- Publishing statistical models: Getting the most out of particle physics experiments (arXiv:2109.04981)
- Data and Analysis Preservation, Recasting, and Reinterpretation (Snowmass 2021 contribution, arXiv:2203.10057)
- Reinterpretation and Long-Term Preservation of Data and Code (Snowmass 2021 contribution, arXiv:2209.08054)
- RAMP seminars: https://indico.cern.ch/category/14155/