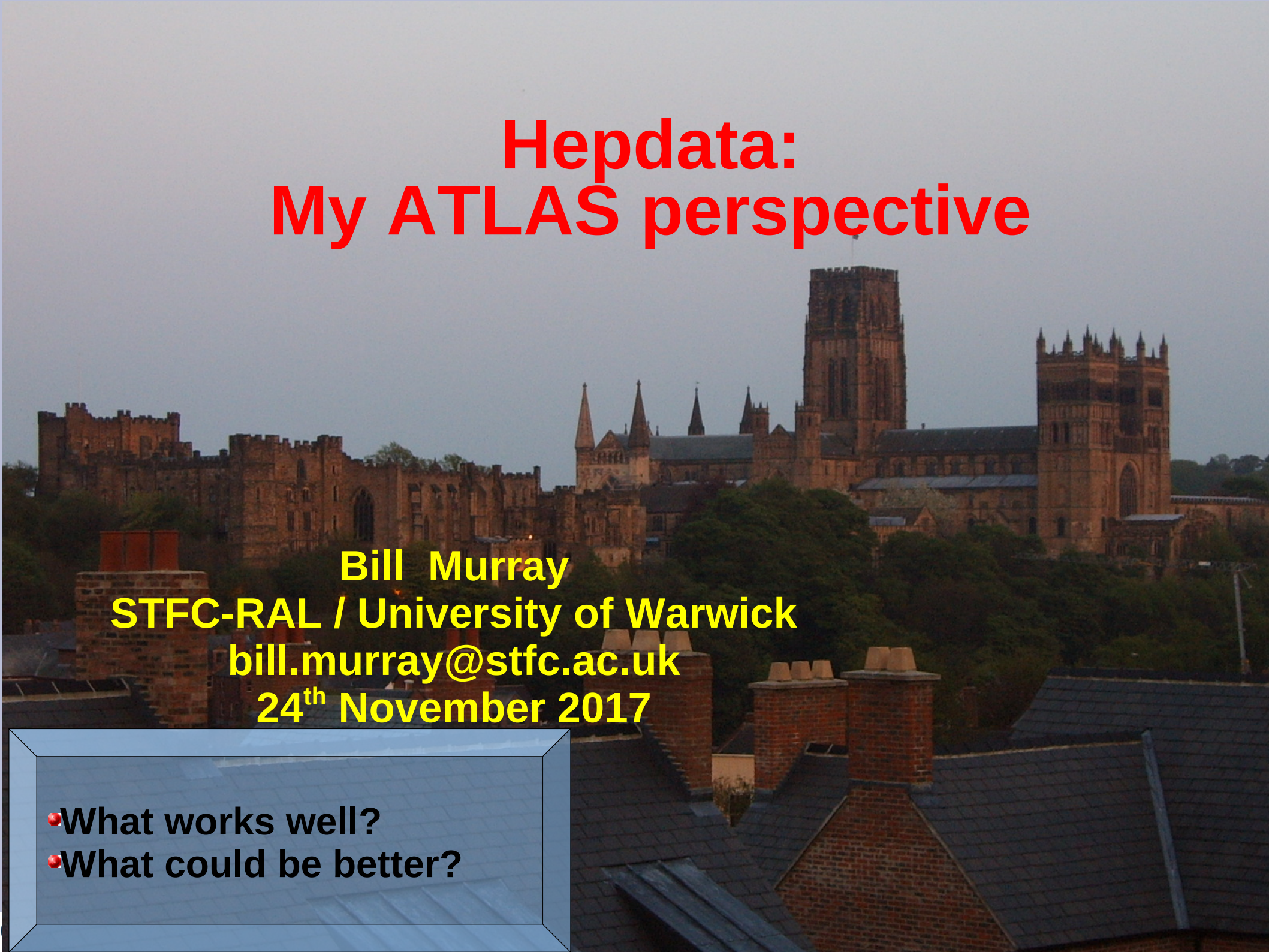
Hepdata:
My ATLAS perspective



Bill Murray
STFC-RAL / University of Warwick
bill.murray@stfc.ac.uk
24th November 2017

- What works well?
- What could be better?

HepData uage

- ATLAS is a big user
  - See table right
- And a satisfied customer
  - At the `ATLAS new physics search reinterpretation workshop' in July the talk 'What to we already do for SUSY' has the word 'hepdata' in large font in top left on 13/19 slides.
    - And argues the approach works
    - "Everyone agrees hepdata is fantastic and we should use it for everything under the sun"
  - Exotics and Higgs less developed, but same idea
  - Desire to make publishing analysis easier
    - But SUSY seem happy with code use
  - Debate on full/simplified likelihoods

ATLAS	346
CMS	220
CDF	209
ZEUS	169
ALICE	164
H1	146
OPAL	145
D0	127
L3	112
DELPHI	101
ALEPH	83
CLEO	67
TASSO	65
LHCB	59
THE ATLAS	1

Data entry: much improved

- Personal comment: some of the support tables uploaded to hepdata are not so well scrutinised
  - They are not well scrutinised in experiment review
  - Mistakes do happen
  - The visualisation tools are very useful
- But this is much improved.

New ATLAS results page: HepData

ExoticsPublicResults < AtlasPublic < TWiki - Mozilla Firefox

File Edit View History Bookmarks Tools Help

ExoticsPublicResults < At | Search for dark matter produ

https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ExoticsPublicResults

Search

Safe gMail Listen Live Telegraph science Gnome Extensions BBC Science

Summary Plots [Link to full list of EXOT summary plots]

Summary Plots

Filter Documents

Select the desired keywords to filter the results.

Selections within a section row are combined with a logical OR, while selections among different section rows are combined with a logical AND.

Global Selections	Show All · Deselect All · Show Latest 20
CM Energy	7 TeV · 8 TeV · **13 TeV**
Exotic physics	BSM resonance · Contact interactions · **Dark matter** · Dark sector · Vector-like quarks · Strong gravity · Leptoquarks · Heavy vector triplets · Compositness · Other new particle searches · Heavy neutrinos
Signature *'OR' logic within row* *'AND' between rows*	W · Z · Photon · WW · WZ · ZZ · Di-photon · Vphoton · VH · VVV · Single top · Top pair · >=3 tops · Higgs · Di-Higgs 0 lepton · 1 lepton · 2 leptons · 2 leptons (same charge) · >=3 leptons MET · Taus · Photons · Charged tracks 0 jets · 1 jet · 2 jets · >=3 jets · All hadronic · c-jets · b-jets
Analysis characteristics	ISR · MVA / machine learning · EFT · High luminosity upgrade studies · Statistical combination · VBF · BSM reinterpretation · Long-lived massive particles
Min luminosity :	0 fb^{-1} · Filter by minimum integrated luminosity
Date :	YYYY-MM-DD · Filter by date: ≤ · ≥

Filtered results: [Papers Confnotes Pubnotes]

Papers: (11)

Short Title	Journal reference	Date	√s (TeV)	L	Links
MET + jet search 13 TeV 2016	Submitted to JHEP	09-NOV-17	13	36 fb^{-1}	Documents | 1711.03301 | Inspire Internal
WIMP DM pair + HF quarks; 0, 2 leptons	Submitted to EPJC	31-OCT-17	13	36 fb^{-1}	Documents | 1710.11412 | Inspire HepData | Internal
Search Zh(125), h(125)->invisible and Z+MET	Submitted to PLB	31-AUG-17	13	36 fb^{-1}	Documents | 1708.09624 | Inspire Internal
jet+MET cross section 13 TeV 2015	Eur. Phys. J. C 77 (2017) 765	11-JUL-17	13	3.2 fb^{-1}	Documents | 1707.03263 | Inspire Rivet | HepData | Internal
MET + H->bb search 13 TeV 2016	Phys. Rev. Lett. 119 (2017) 181804	05-JUL-17	13	36.1 fb^{-1}	Documents | 1707.01302 | Inspire HepData | Internal
Search h(125)->gamgam + Missing-ET	Submitted to PRD	13-JUN-17	13	36.1 fb^{-1}	Documents | 1706.03948 | Inspire Internal
MET + photon search 13 TeV 2016	Eur. Phys. J. C 77 (2017) 393	12-APR-17	13	36.1 fb^{-1}	Documents | 1704.03848 | Inspire HepData | Internal
MET+Higgs search 13 TeV 2015	Phys. Lett. B 765 (2016) 11	15-SEP-16	13	3.3 fb^{-1}	Documents | 1609.04572 | Inspire Internal
	Phys. Lett. B 763				Documents | 1608.02372 | Inspire

Lets explore this one

HepData Link clear (usually)

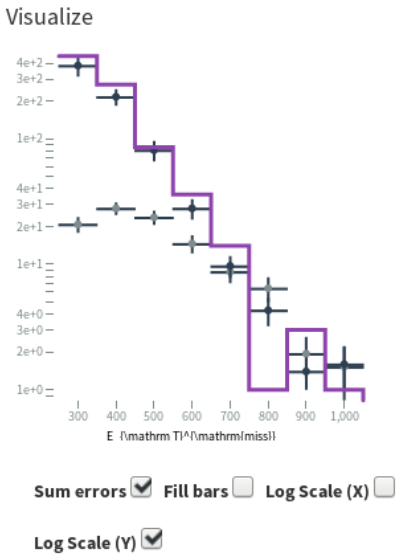
Science & Technology Facilities Council
Rutherford Appleton Laboratory

THE UNIVERSITY OF WARWICK

www.hepdata.net/record/80080

- "Search for dark matter produced in association with bottom or top quarks in √s = 13 TeV pp collisions with the ATLAS detector"
- Record clear, well laid out
- 63 data tables in this example
  - Cut flows
  - Efficiencies
  - Results plots
    - (I am used to errors on data but that's philosophy)
- Very nice
- Someties WWW hangs when a table is requested
  - OK, ask again

$\sqrt{s}$	13000.0 GeV		
	Data	Total Bkg	(1000,35)
$E_{\mathrm{T}}^{\mathrm{miss}}$	**Events / 100 GeV**		
250 - 350	456	383.6 ±70.8 stat+syst	20.805 ±3.00816
350 - 450	271	216.2 ±32.8 stat+syst	27.8835 ±3.1619
450 - 550	86	81.2 ±15 stat+syst	23.6675 ±3.10245
550 - 650	36	27.9 ±5.2 stat+syst	14.6047 ±2.41529
650 - 750	14	9.7 ±1.9 stat+syst	8.7116 ±1.62698
750 - 850	1	4.3 ±1.1 stat+syst	6.43781 ±1.46634
850 - 950	3	1.4 ±0.4 stat+syst	1.9378 ±0.703358
950 - 1050	1	1.6 ±0.6 stat+syst	1.52518 ±0.703185



Example: poor table

- Sometimes automatic binning of tabular data makes unhelpful choice
  - Here it would be nice to be able to separate the columns
  - This is nit-picking

Cmenergies

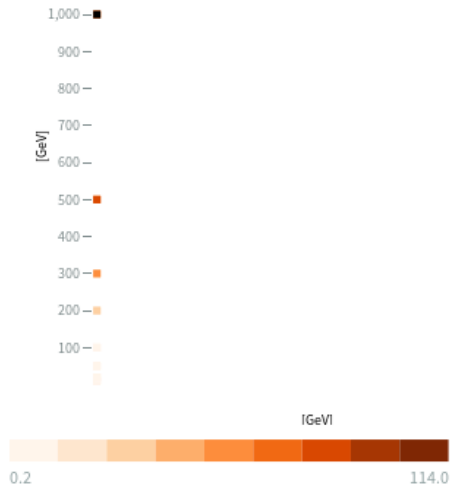
13000.0

Reactions

P P --> DM tt | P P --> DM bb

$\sqrt{s}$		13000 GeV			
$m(a)$ [GeV]	$m(\chi)$ [GeV]	Acceptance SRb2-bin1 [x10^{-5}]	Acceptance SRb2-bin2 [x10^{-5}]	Acceptance SRb2-bin3 [x10^{-5}]	Acceptance SRb2-bin4 [x10^{-5}]
10	1	0.24	0.246	0.345	0.391
20	1	0.38	0.36	0.559	0.534
50	1	0.906	0.831	1.6	1.95
100	1	1.71	2.56	3.98	4.96
200	1	5.1	7.8	11.8	16.1
300	1	7.55	9.86	20.3	38.4
500	1	13.2	18.3	45.3	71.7
1000	1	40	30.3	63.6	114

Visualize



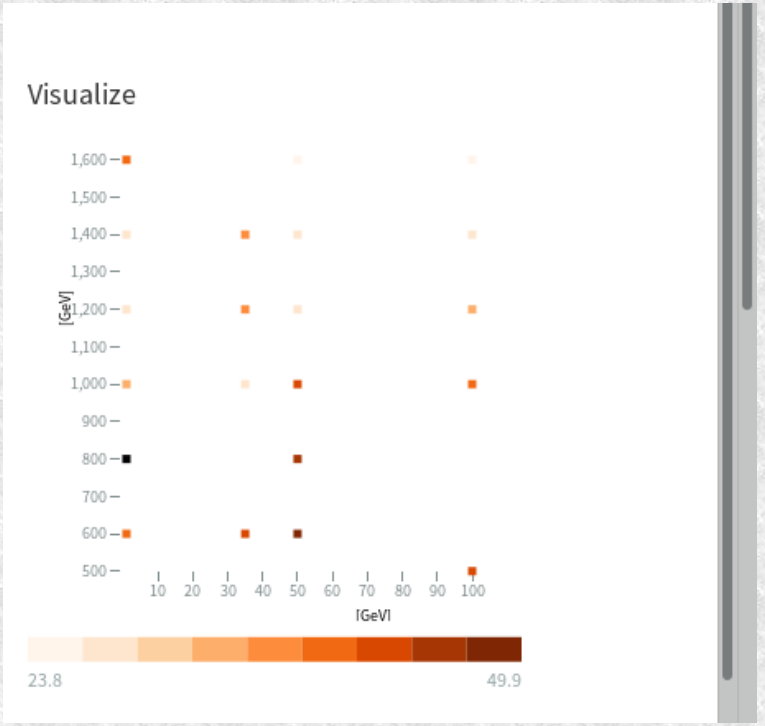
Brushing Enabled?

X Axis m(χ)m(χ) [GeV]

Y Axis m(a)m(a) [GeV]

Example dubious data

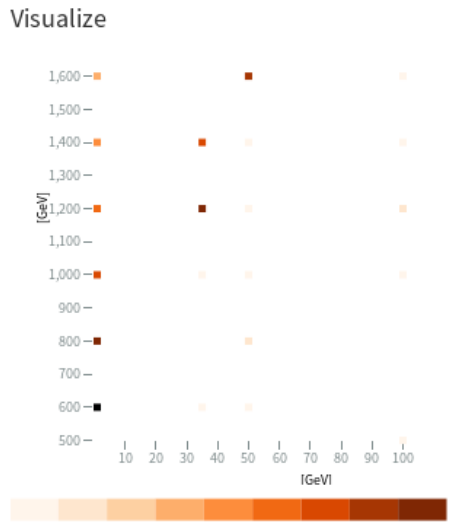
- A table from this paper
- Table 40: “Data from table 19, Efficiency of the SRb1 selection of the b-FDM model signal samples
- This particular table seems to me either
  - Structure meaningful?
    - Insufficient granularity
  - Structure reflects MC stats?
    - Not given - unhelpful
  - Wrong
- But it reflects perfectly what ATLAS supplied



Example wrong data

- A table from this paper
- Table 39: “Data from table 19, Efficiency of the SRb1 selection of the b-FDM model signal samples
- This table is simply wrong
- (600,50) is a factor 100 different between ATLAS and HepData table
  - And some others too

$\sqrt{s}$		13000 GeV
$m(\phi_b$ [GeV]	$m(\chi)$ [GeV]	Acceptance. SRb1 [$\times 10^{-4}$]
600	1	9.19
800	1	9.17
1000	1	5.32
1200	1	3.7
1400	1	2.71
1600	1	2.29
600	50	0.265
800	50	0.527



$m(\phi_b)$ [GeV]	$m(\chi)$ [GeV]	acc. SRb1	eff. SRb1
600	1	$9.19 \cdot 10^{-4}$	$4.01 \cdot 10^{-1}$
800	1	$9.17 \cdot 10^{-4}$	$4.99 \cdot 10^{-1}$
1000	1	$5.32 \cdot 10^{-4}$	$3.53 \cdot 10^{-1}$
1200	1	$3.70 \cdot 10^{-4}$	$3.06 \cdot 10^{-1}$
1400	1	$2.71 \cdot 10^{-4}$	$2.83 \cdot 10^{-1}$
1600	1	$2.29 \cdot 10^{-4}$	$3.91 \cdot 10^{-1}$
600	50	$2.65 \cdot 10^{-3}$	$4.70 \cdot 10^{-1}$
800	50	$5.27 \cdot 10^{-3}$	$4.55 \cdot 10^{-1}$
1000	50	$3.41 \cdot 10^{-3}$	$4.22 \cdot 10^{-1}$
1200	50	$1.83 \cdot 10^{-3}$	$3.00 \cdot 10^{-1}$

ATLAS ongoing

- ATLAS task force
  - Reps from all physics groups
- New ATLAS guidlines/tips twiki
  - How to use
  - What to store
  - Pitfalls
- Currently Hepdata upload timing varies with physics group
  - Recommendation to to in parallel with submission to Archive
  - Change after journal review if necessary
- HepData should go hand in hand with Rivet
  - Still in discussion

Preliminary Recommendations

- General: Everything that is useful for a theoretician, i.e. all final results (also tables which are already in the paper)

- Measurements
  - Fiducial Definition
  - Measured values / cross-sections and uncertainties
  - Correlations
    - Correlation matrices
    - Effect of each systematic separately
    - Measurement variations
  - Correction factors between particle-level definitions

- Searches
  - model dependent limits (table of expected +/- 1, 2 sigma, observed limits vs mass)
  - model independent limits (table of observed limits vs mass)
  - acceptance / efficiency vs mass (or whatever other variable is relevant)
  - final distribution if it is useful: data counts, total background yield, statistical error on the background, systematic error on the background

ATLAS Request:

- HEPData is in pretty good shape.
- One thing bothers us:
  - Implementing a review process for the HEPData entries before the publication:
    - i.e. each ATLAS author can check what we put into HEPData before we publish the entry.
  - Could use CDS for this, where the HEPData tar-file could be stored. This requires that an ATLAS member downloads this file, and uploads it to his private HEPData Sandbox. There are two issues:
    - The Sandbox link can be seen by everybody, when known - it would be better to have a “real” private version
    - The procedure is quite complicated, since it involves one download and one upload. It would be somehow good, if the HEPData access could be linked somehow to the CDS access rights…
- Can HEPData access use CERN CDS rights?