Educational Platform ATLAS open data

Sue Cheatham
on behalf the
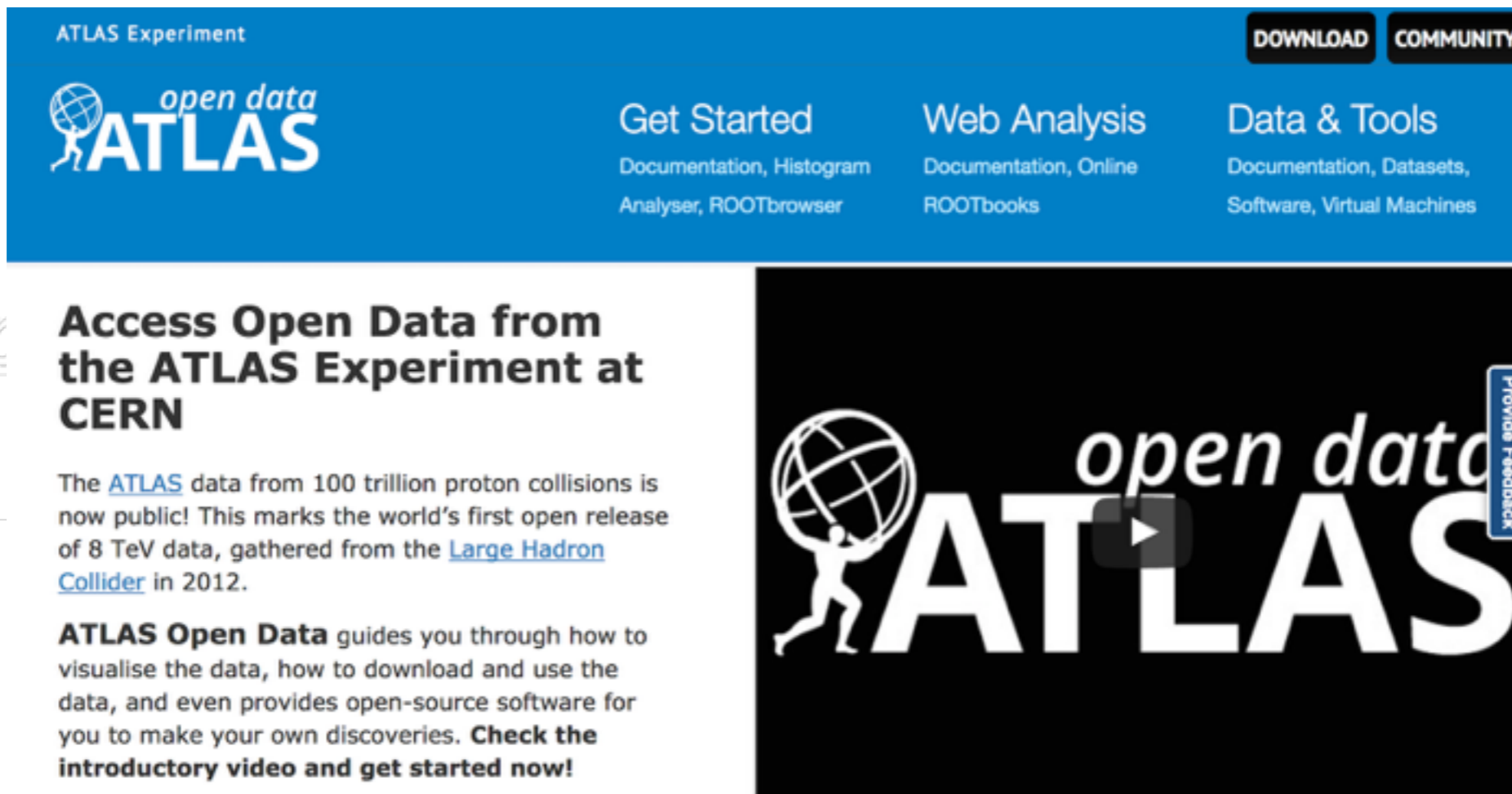**ATLAS Data&Tools Outreach Group**

UK Teacher Programme
13-17 December 2016

ATLAS
OPEN
DATA

EXPLORE 8TEV DATA ON
ATLAS' COMPREHENSIVE
EDUCATIONAL PLATFORM

HTTP://OPENDATA.ATLAS.CERN

ATLAS open data

- **Data**: 1 fb^{-1} of 8 TeV ATLAS data released (about 100 trillion proton-proton collisions)

- **Tools**: Analysis tools for educational use

- **Documentation**: To explain the different concepts and guide the user to look at the data

- **Target Audience**: Currently aimed at University students

- **Forum**: Questions and feedback welcome

Get Started

- Data visualisation 4 processes: H→WW, WW, top pair, Z

- Make cuts with cursor on one variable and immediately see the effects on the other variables

- Expected number of events for 1 fb^{-1} shown, along with significance of Higgs signal

- Documentation describes separate samples and suggests cuts inspired by Higgs Analysis
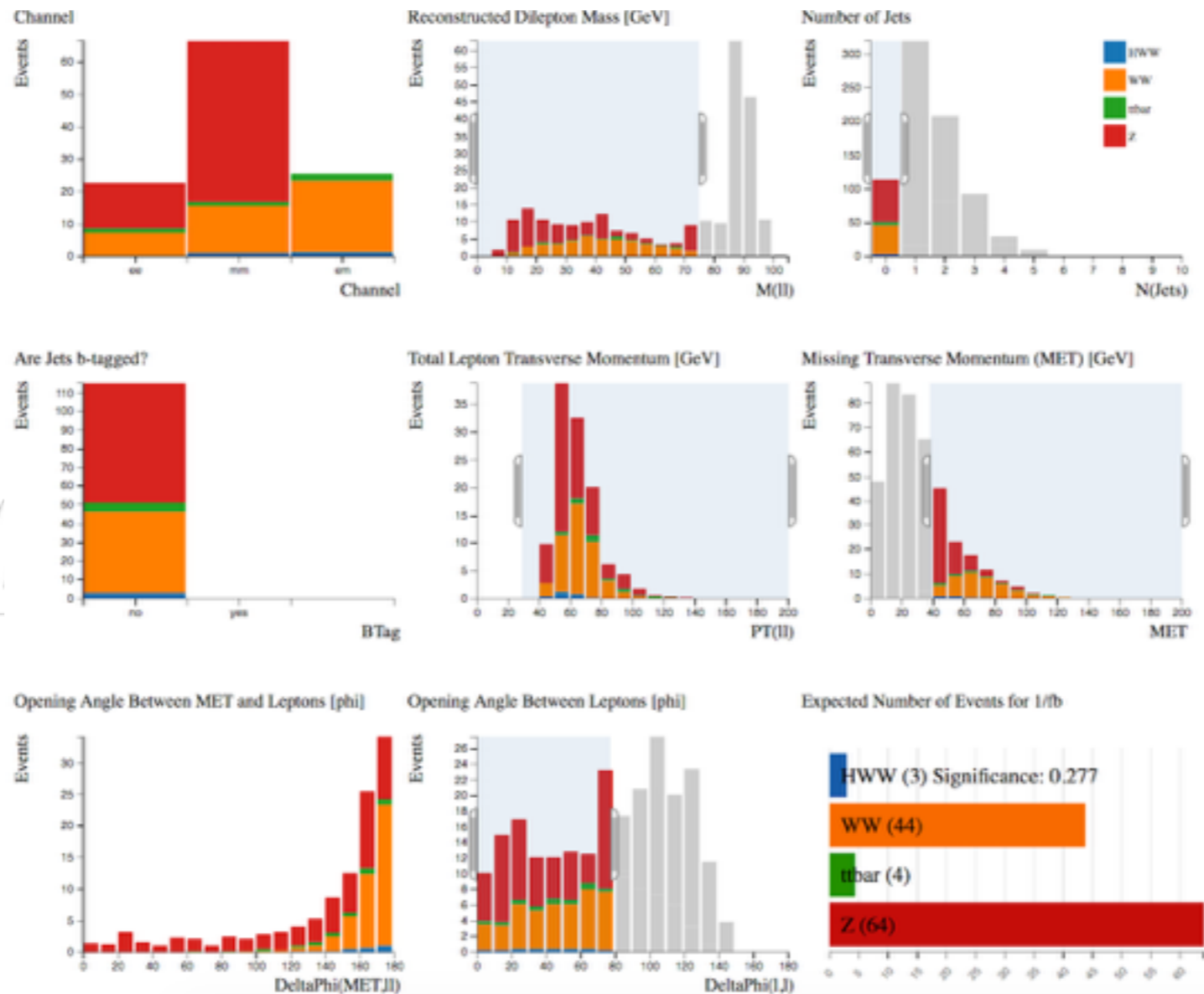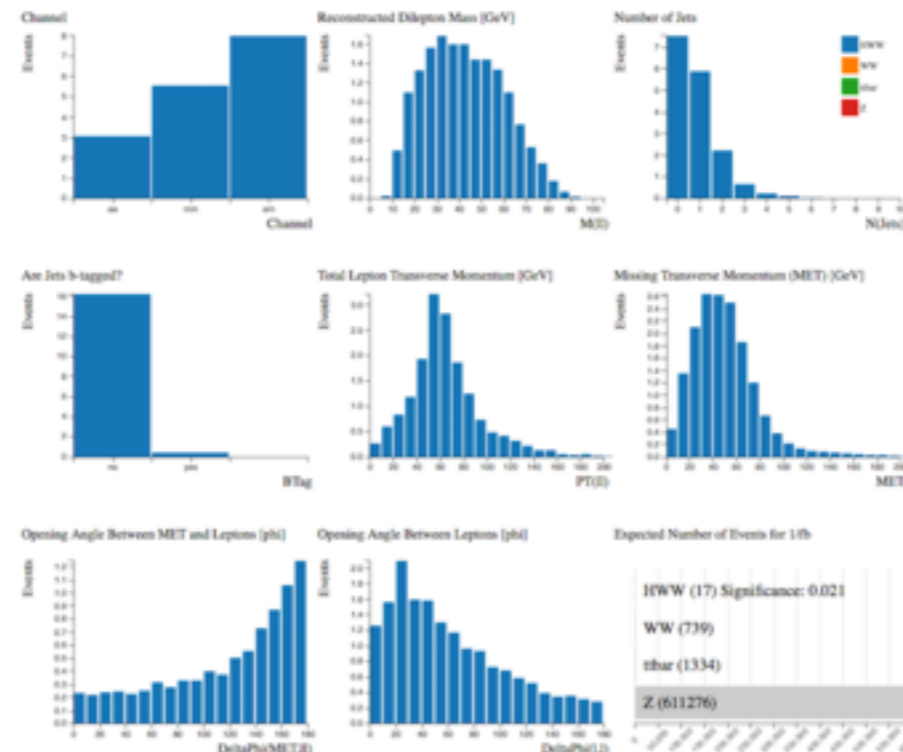
Get Started

Analyses:
• Explains the signal and three background processes

Histogram Analyser:
• Making cuts with your cursor
• The histograms explained one by one

Separate Signals
• Simulated data allows us to look at the distributions and understand the differences between the signal and background
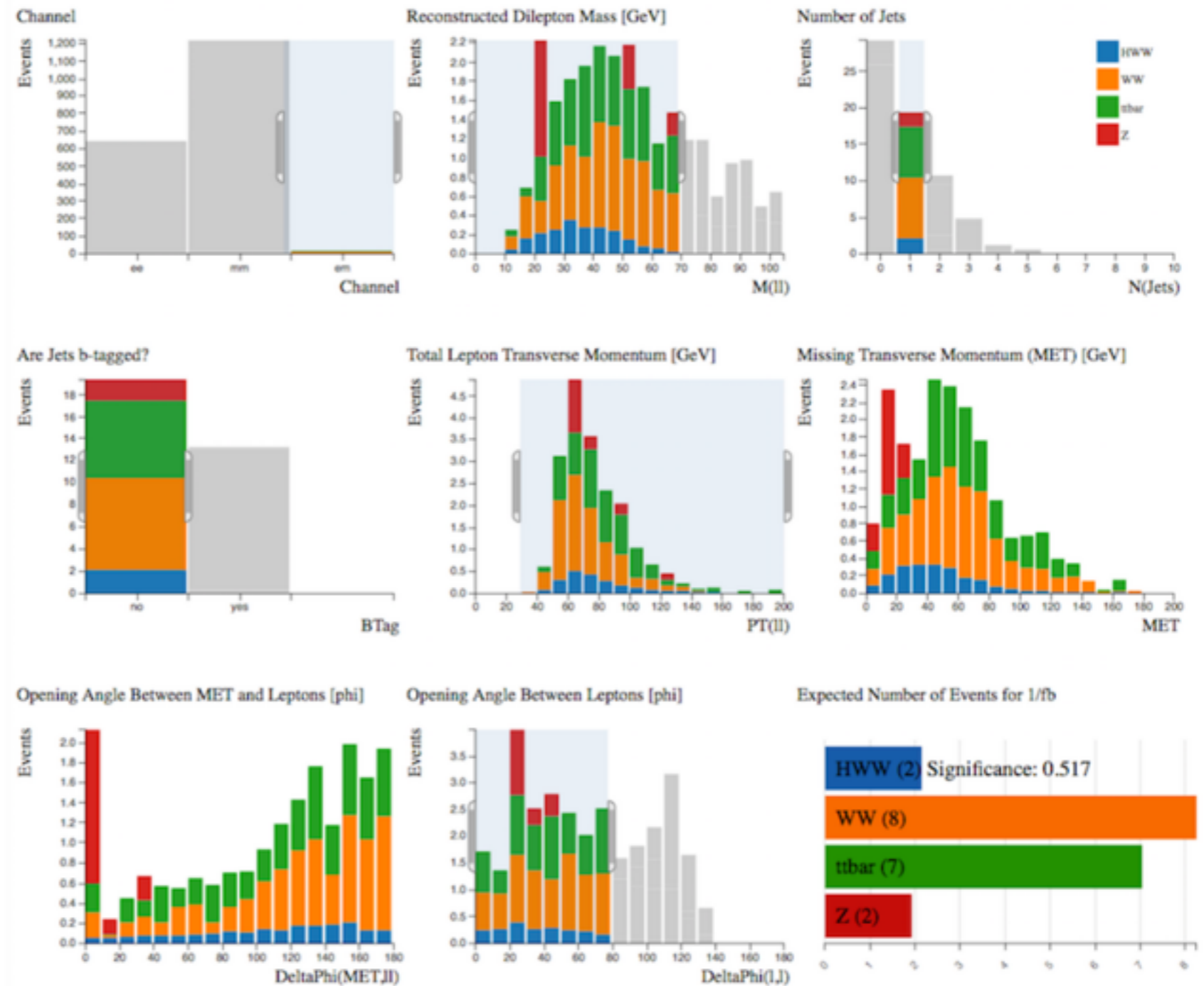
Find the Higgs

- Documentation explains idea of making cuts
- Steps through cuts to find Higgs

Higgs boson + 1 jet
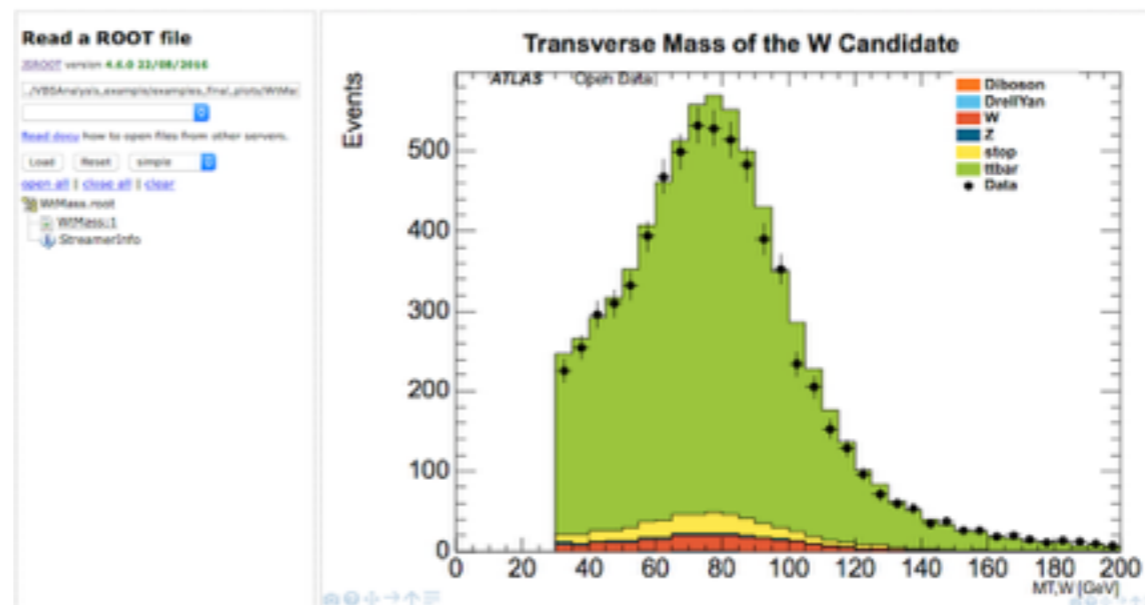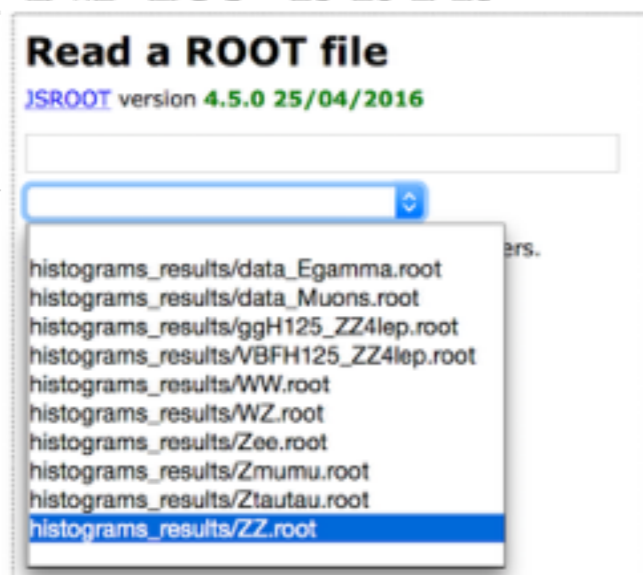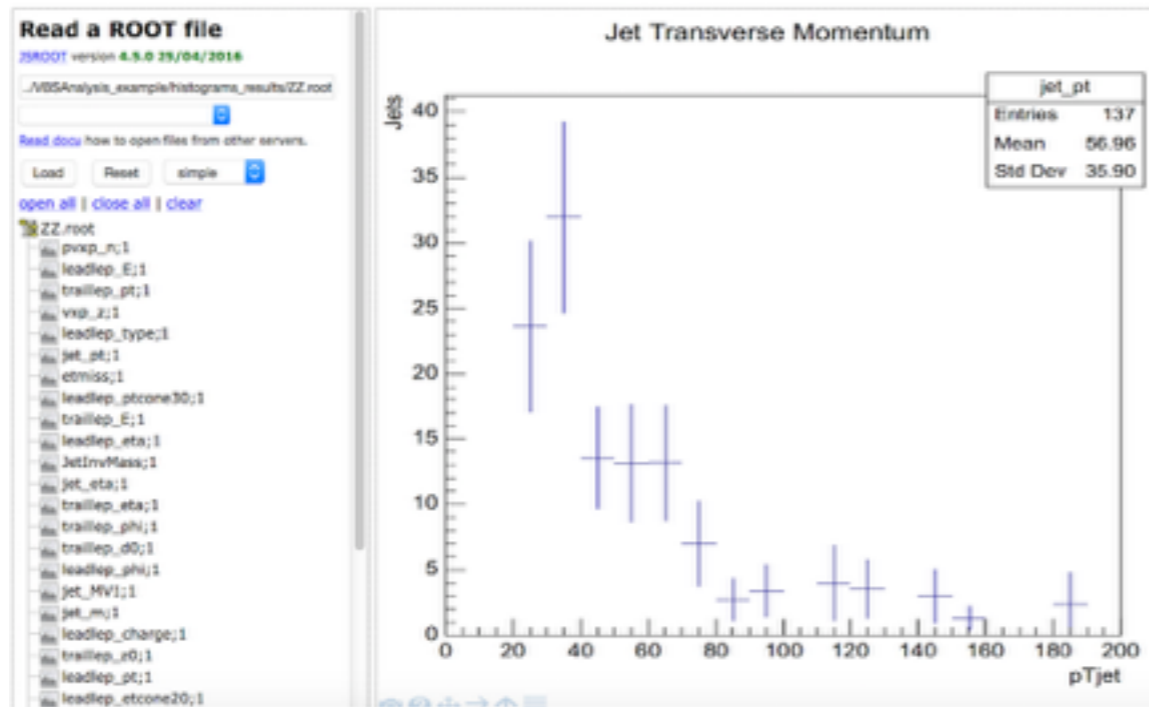
Select:

- Number of Jets = 1
- no b-tagged jets
- electron-muon channel only
- Reconstructed Dilepton Mass < 70 GeV
- Total Lepton Transverse Momentum > 30 GeV
- Opening angle between leptons < 80

ROOTbrowser

- A web based tool for displaying and analysing data
- Documentation defines ntuple variables
- Histograms explained and data-simulated data agreement discussed

Data

- Data and simulated data: ROOT tree format
- 1 fb^{-1} of data
- egamma ~ 33.6 M events + muons ~ 33.8 M events
- Datasets available to be downloaded individually or bulk download
- Also available on the CERN open data portal

Set of Data samples

File type	Name	Description	Last modified	Size	# Events
√	DataEgamma.root	ATLAS 2012 data Egamma-string sample for 2016 open data release	21-Jul-2016 16:00	746,3Mb	7917590
√	DataMuons.root	ATLAS 2012 data Muons-string sample for 2016 open data release	21-Jul-2016 16:00	619,8Mb	7028084
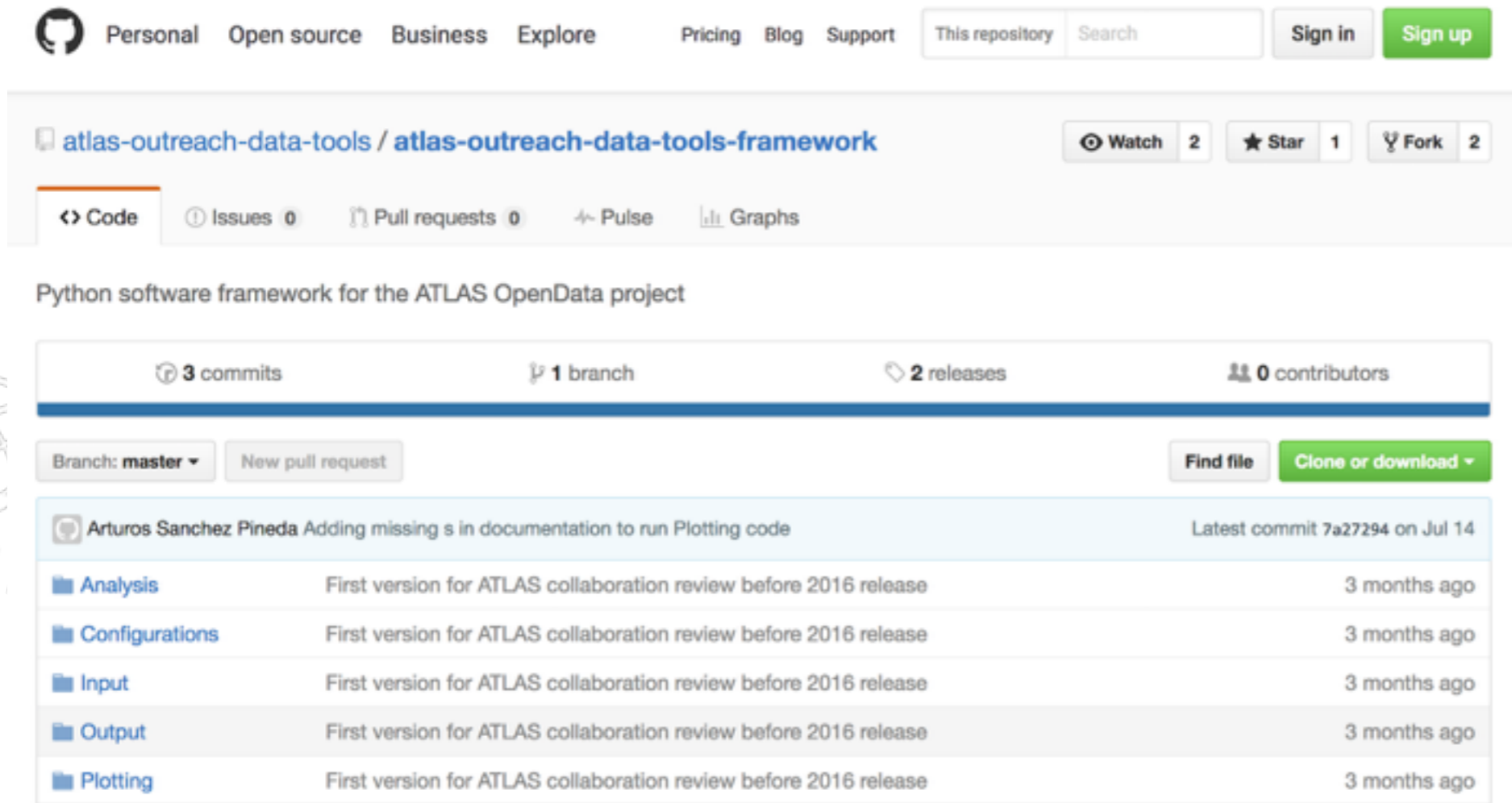
Set of MonteCarlo (MC) samples

File type	Name	Description	Last modified	Size	# Events
√	mc_105985.WW.root	Diboson process *WW*	21-Jul-2016 16:00	64,7Mb	500000
√	mc_105986.ZZ.root	Diboson process *ZZ*	21-Jul-2016 16:00	19,8Mb	125000
√	mc_105987.WZ.root	Diboson process *WZ*	21-Jul-2016 16:00	69,5Mb	500000
√	mc_110090.stop_tchan_top.root	Single top *t*-channel top	21-Jul-2016 16:00	21,6Mb	150000
√	mc_110091.stop_tchan_antitop.root	single top *t*-channel antitop	21-Jul-2016 16:00	14,5Mb	150000

Tools

- Analysis software to run and modify
- Seven analyses: H→WW, WW, ZZ, top pair, Z, W, Z'
- Documentation describes analysis steps to follow and produce histograms
- Available in GitHub repository, as zip file or on the CERN open data portal

Data & Tools
Virtual Machines

Virtual Machines to run on any operating system (5 available)

VM Small Version: 1.7 GB

Lubuntu in conjunction with ROOT and 10% of the data.

Can be downloaded fairly quickly. Remaining 90% data downloaded in background if required.

Documentation: Software book

Explains the different virtual machines & how to choose which one suits your needs

How to setup environment for small VM

How to run and look at data

Explains plots

Details event selection

Defines variable names

relevant to all VMs

Virtual Machines

- Step by step instructions on how to set up a VM
- Documentation describes steps to follow, to run the software and produce histograms
- Histogram variables explained and differences between data and simulated data discussed

Data & Tools

Chapters added to explain what happens in the code:
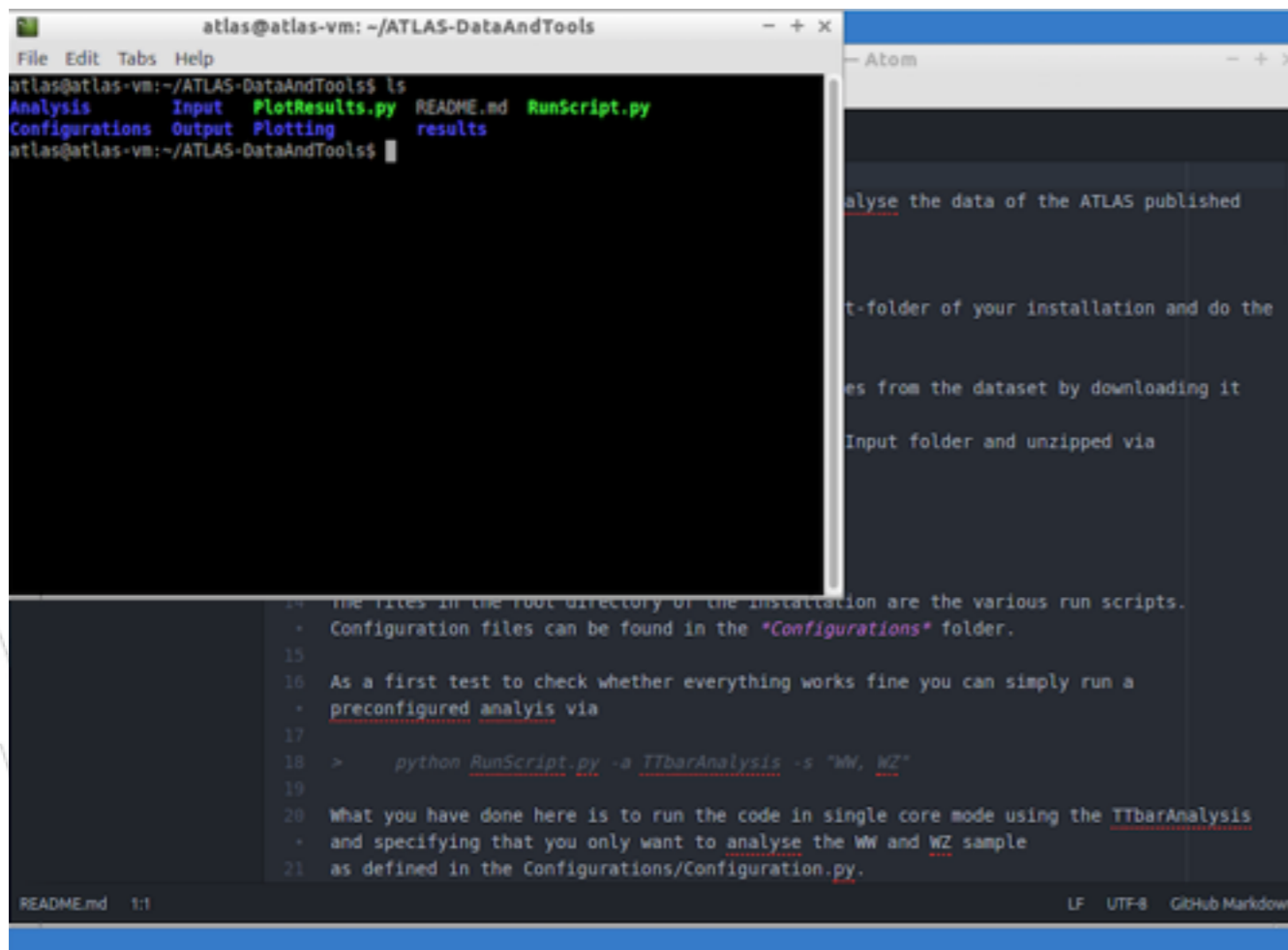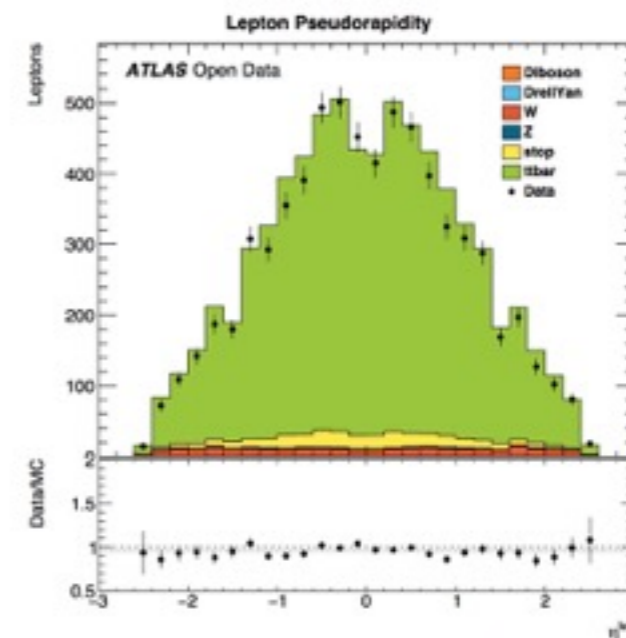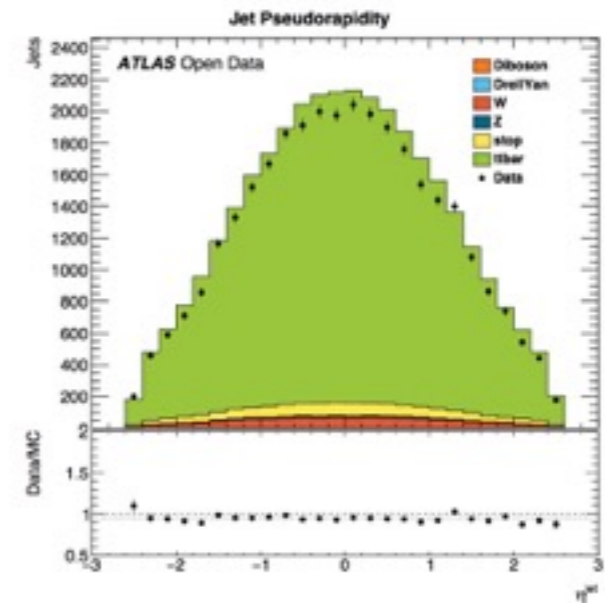- **production of histograms and plots**
- **code snippets to explain event selection**

Event selection

The events in the dataset ntuples have been selected according to a selection criteria. The variables are defined in variable names.

The standard event based selection criteria are:

- A single electron or muon trigger has fired;
- The primary vertex has at least 5 tracks;
- There is at least one good lepton with $p_T > 25$ GeV;
- Leptons are required to be isolated (Both ptcone30 and etcone20 < 0.15);
- The event passes the Good Run List (GRL);
- A veto exists on events containing bad jets.

The standard event based selection code is

```
12   def isGoodLepton(Lepton):
13       if (abs(Lepton.pdgId()) == 11 and isGoodElectron(Lepton)): return True;
14       if (abs(Lepton.pdgId()) == 13 and isGoodMuon(Lepton)): return True;
15       return False;
16
17   def isGoodElectron(Lepton):
18       if not Lepton.isTight(): return False
19       if not Lepton.pt() > 25: return False
20       if not Lepton.isoetconerel20() < 0.15: return False
```

Take a look at the data

1. **Run an analysis** using **RunScript.py** to produce histograms of individual variables eg jet_pt. The histograms are written to your **results** folder.
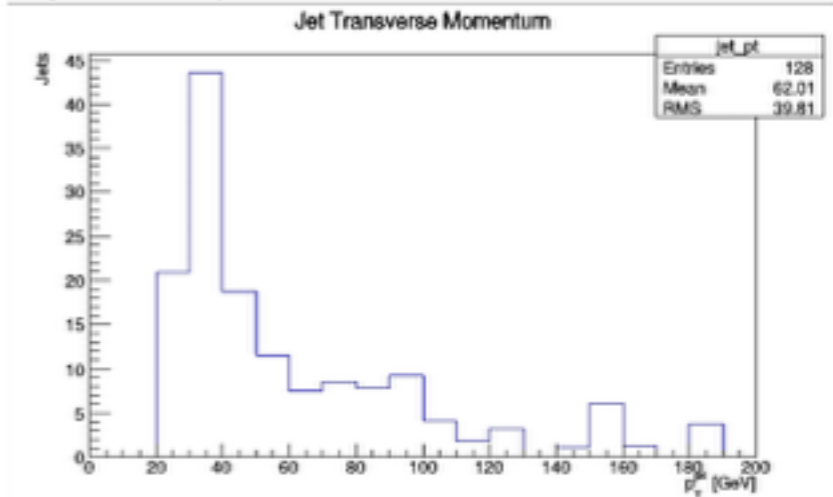
2. **Plot the results** using **PlotResults.py** to scale the histograms, colour them in and stack them. Combined plots are produced, for all the simulated and real data. These plots are written to your **Output** folder.

ATLAS open data

4.4k visits

Countries | World-Wide | Visits
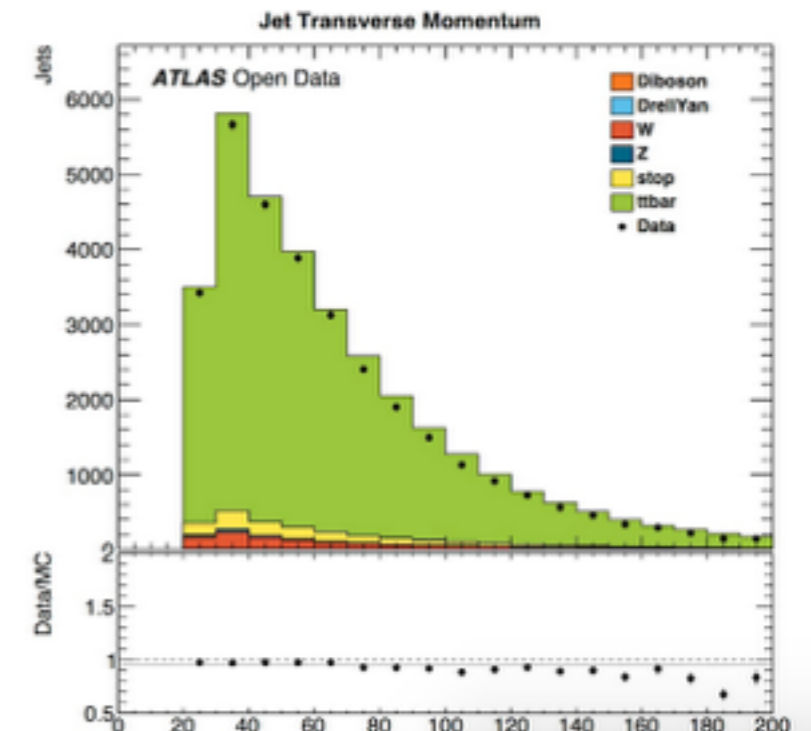
- Achieving international impact
- Release peak: will continue to re-advertise when new data or tools available
- Bounce rate and page visit length average for a website

Visits Activity

Top Social Sources 480 Total Visits

f Facebook 330 Visits

2,855 DIRECT | 1,824 REFERRING DOMAINS | 480 SOCIAL SOURCES | 1 CAMPAIGN

● Sessions ● % New Sessions

Oct 9, 2016 - Oct 15, 2016
● Sessions: 198
● % New Sessions: 48.48%

Sessions	Users	Pageviews
5,700	3,345	19,192

Pages / Session	Avg. Session Duration	Bounce Rate	% New Sessions
3.37	00:04:10	51.58%	57.74%

■ New Visitor ■ Returning Visitor

42.1% 57.9%

Bounce Rate is the percentage of visitors that abandon the site after the first interaction with the site 40 to 55 % is average

The average page visit ~1 minute

Summary

- Outreach is both central and essential to the scientific process

- ATLAS open data aims to make ATLAS data and tools accessible to all

- Broad target audience:

  Currently aimed at University students

  Interesting to high school teachers and students

- Development of tools ongoing

- Testing of website and tools ongoing with small groups volunteers

- Feedback always useful and very welcome

ATLAS Outreach

ATLAS public website

Engage public: news stories, evergreen content and opportunity to learn more about ATLAS
Core message: ATLAS, Physics, Collaboration, Detector/Technology

Social Media

Engage new audiences
Connects people with ATLAS
Brings traffic to the website

Facebook likes: 20 k
Twitter followers: 39 k
Google+ followers: 123 k

ATLAS LIVE AT ICHEP
Join our Facebook Live event
Thursday, 4 Aug 16:45 CST (23:45 CEST)
www.facebook.com/ATLASexperiment
#ICHEP2016

ATLAS Experiment at CERN
August 30 at 3:02pm · 🌐
👍 Like Page

An exceptional summer: 12 stories about ATLAS summer students and their supervisors

https://atlas-public.web.cern.ch/.../atlas.../exceptional-summer

For many studen...

See More

Thanks to CERN's Summer Student Programme, every year dozens of university students spend their holidays working at the ATLAS experiment.

12 STORIES ABOUT ATLAS SUMMER STUDENTS AND THEIR SUPERVISORS

An exceptional summer
For many students, summer means sun and beach volleyball. For some, though, it is an opportunity to learn at ATLAS! Thanks to CERN's Summer Student Programme, every year dozens of university students come to ATLAS to spend their holidays in...
ATLAS.CERN

Picturing Particles 🕐 ↗
PICTURING PARTICLES
The evolution of event displays
▶ 🔊 0:05 / 2:09

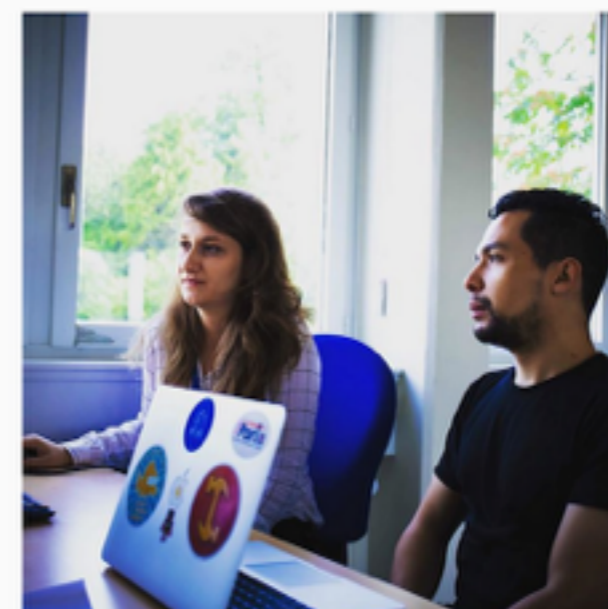**Kate Shaw** shared **ATLAS Experiment at CERN**'s live video.
August 4 ·

Interview with Marumi Kado, ATLAS Physics Coordinator Live from ICHEP 2016!

1,310 Views
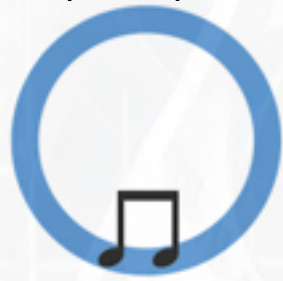**ATLAS Experiment at CERN** was live.
August 4 · 🌐
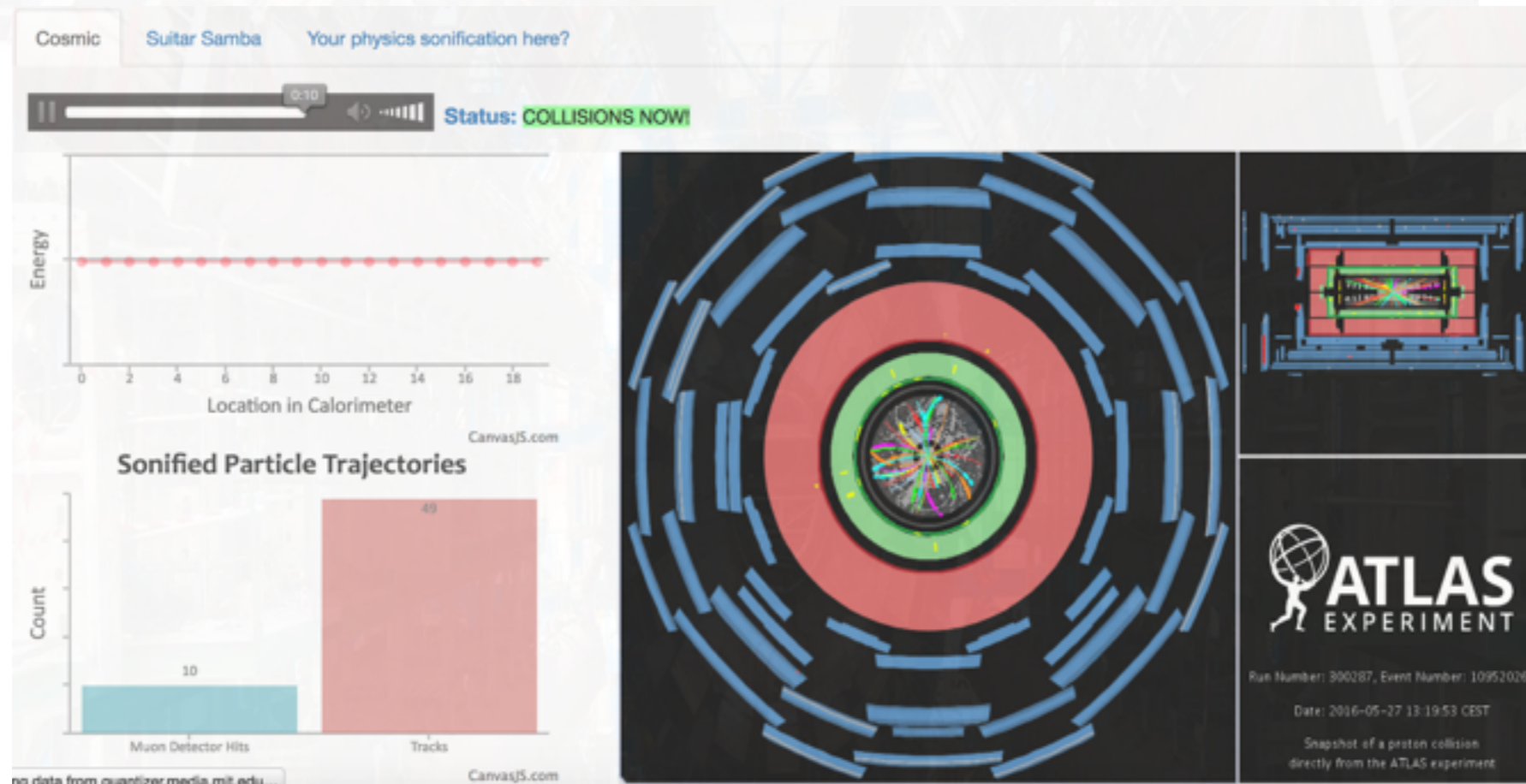ICHEP Interview with Marumi Kado #facebooklive #ichep2016
👍 Like Page

ATLAS
EXPERIMENT

ATLAS Experiment at CERN
@ATLASexperiment

Instagram

ATLAS
EXPERIMENT

http://quantizer.media.mit.edu/

Website:
~ 6300 visits
~ 4500 users

Soundcloud:
pre-recorded music
tracks
~3200 plays

Data sonification:
- Property of sound, such as the pitch, mapped to a physical property, such as speed.
- Various different software packages available for turning numbers into sounds.

Citizen science project

Higgs Hunters

- A collaboration between Oxford, Birmingham and New York Universities
- The first particle physics venture on Zooniverse, a collection of web-based citizen science projects
- Invites online volunteers to participate in classifying off-centre vertices
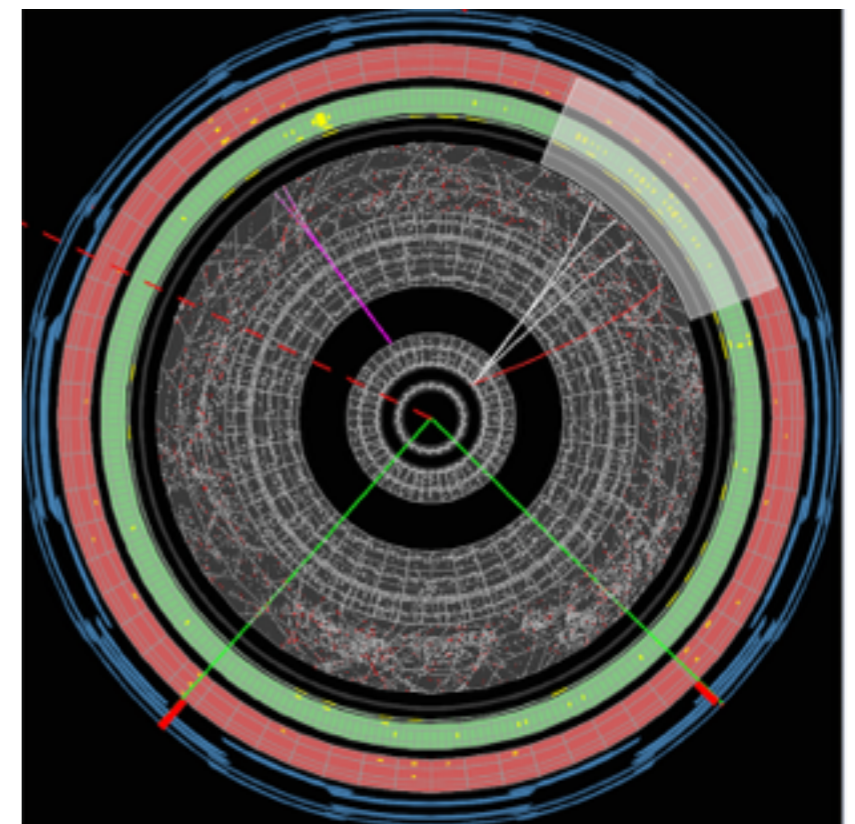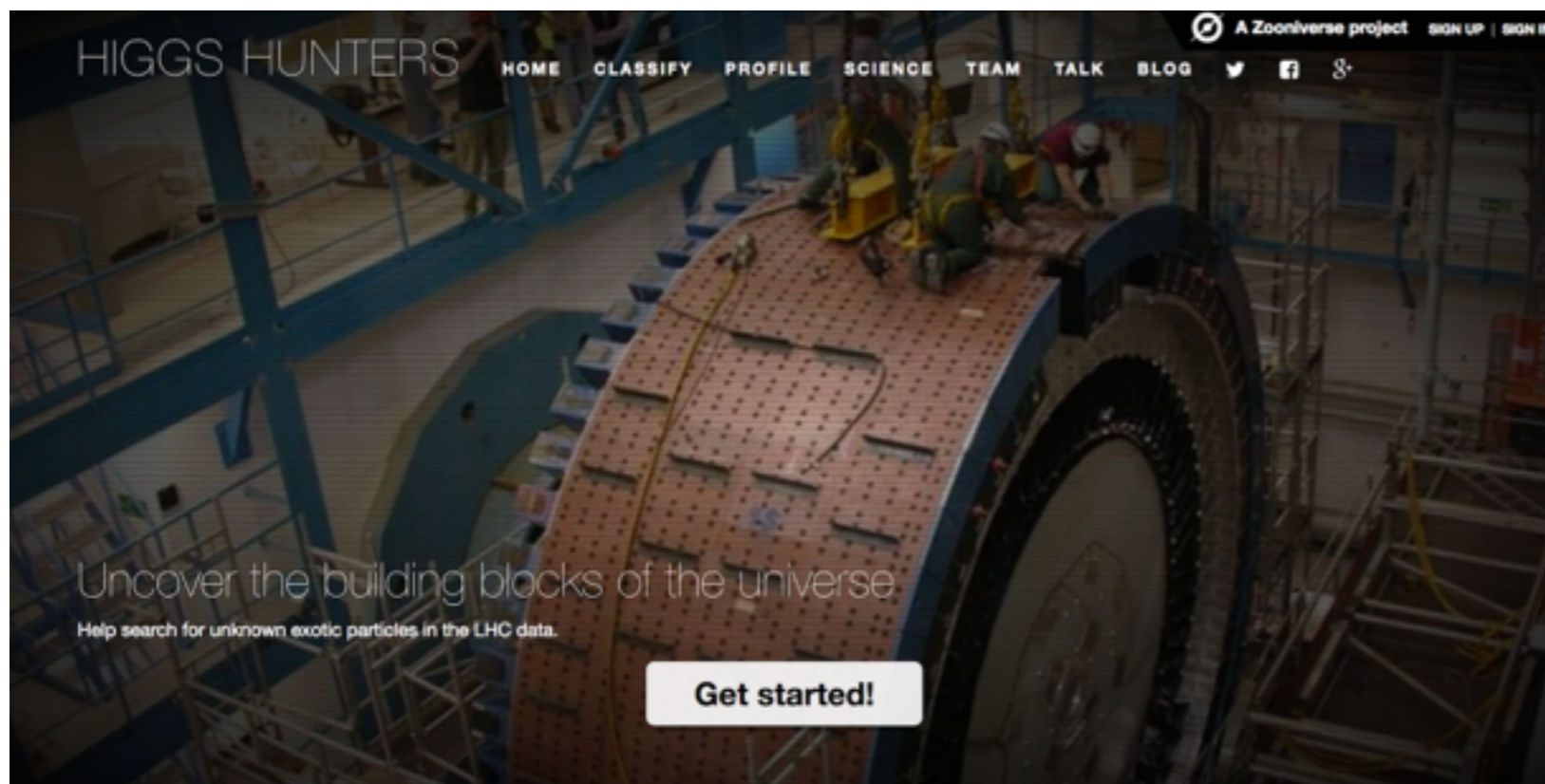
- More than 30,000 volunteers from 179 countries participated, classifying 980,000 features of interest on about 39,000 distinct images.
- Non-expert volunteers are capable of identifying the decays of long- lived particles with an efficiency and fake-rate comparable to that of the ATLAS algorithms.

Virtual Visits

Digital Communications Award 2012

Virtual visits continue to be popular.
Excellent resource to connect with school and university students.

In 2016, around 70 visits from more than 14 countries

Higgs Branching Ratios