

Masterclass Upgrades and Open Data

Yiota Foka

GSI

IPPOG Meeting, Pisa, 20 April 2018

Aims and Disclaimer(s)

Masterclasses

We have our Masterclasses that we run the last 8 or so years and we are happy with.

There are some cases we wish to develop new Masterclasses.

It is natural to review what we have and base new developments on our experience and try to improve optimising effort.

CERN Open Data

LHC data restricted for 5 years, then open.

Desirable to identify synergies and commonalities.

Presentation should had been given by Tibor Simko who could not make it (possibility to invite him in next IPPOG meeting)

Here discuss ideas, not technical solutions



Can CERN open source software handle... 1 PB of LHC data?

Tibor Šimko

@tiborsimko

ITTF · 16 March 2018

Some Examples of possible Masterclass Developments

Development of Masterclass for CBM (future experiment at GSI/FAIR)

NOTE: miniCBM setup summer 2018

Development of Masterclass for HADES (running experiment at GSI, sister of CBM)

Development of further Masterclasses for ALICE

EXAMPLE: LHCb Do re-implement for ALICE (started from ALICE Vo)

Development of CMS Do Masterclasses

Development of ATLAS RAA Masterclass

possibility to combine it with ALICE RAA (and cover low and high pt range)

Development of J/Psi ALICE, CMS....

Development of similar observables for STAR at RHIC etc

Based on tracks, decays and invariant mass analysis

Aim

Explore possibility to implement Masterclass observables for different experiments in flexible and economic way

EXAMPLE

RAA Masterclass, based on tracks

Read in ALICE tracks, plot RAA (cover low pt)

Read ATLAS tracks, plot RAA (cover high pt) at the same histogram

EXAMPLE

Vo or Do Masterclass, based on decays and invariant mass

Read in decay products, calculate invariant mass of pairs

LHCb Do Masterclass could be implemented for ALICE and CMS

**Need of well structured modular framework
And common definition of data format**

ALICE Masterclasses Evolution

Strangeness (by Pawel Debski at CERN)

RAA (by Frederike Bock at GSI)

J/Psi (by Steffen Weber at GSI)

Unified Package (by Christian Holm Christensen Niels Bohr Institute)

To facilitate users and avoid ROOT installation
use of Virtual Machine (or Docker Containers)

Future Plans

CERN Summer Student Proposal from ALICE

Supervisors: Redmer Alexander Bertens, Friederike Bock

This summer student project is aimed at improving and expanding the current ALICE masterclasses

and at developing a **general, experiment independent framework** for displaying detector geometry and reading in and manipulating open data.

**Feedback, coordination with CMS, LHCb, ATLAS
ROOT and Open Data groups**

CERN Open Data portal

- disseminating public particle physics data
 - datasets, software, VMs, configuration, documentation, and more
- LHC collaboration data policies
 - restricted → embargo period (~5 years) → open
- users
 - education: general public, high-school students, masterclasses
 - research: data scientists, physicists
- timeline
 - launched in November 2014 (Invenio v2)
 - major upgrade in December 2017 (Invenio v3)

*Developed by CERN-IT and CERN-SIS
in close collaboration with LHC experiments*



CERN Open Data portal

opendata
CERN

About

Explore more than **1 petabyte**
of open data from particle physics!

Start typing...

Search

search examples: [collision datasets](#), [keywords: education, energy, ITeV](#)

Explore

- [datasets](#)
- [software](#)
- [environments](#)
- [documentation](#)

Focus on

- [ATLAS](#)
- [ALICE](#)
- [CMS](#)
- [LHCb](#)

Get started

<http://opendata.cern.ch/>

Information discovery

The screenshot shows the OpenData CERN interface. At the top left is the 'opendata CERN' logo. A search bar is located at the top center. On the top right, there is an 'About' link. Below the search bar, there are filters for 'Sort by' (Most recent, asc) and 'Display' (detailed, 20 results). A pagination bar shows 'Found 3778 results.' with page numbers 1 through 9. On the left side, there is a 'Filter by type' section with a tree view of categories and their counts. The main content area displays two search results, each with a title, a description, and a 'See the description of the sim...' link. Below each result are buttons for 'Dataset', 'Simulated', and 'CMS'.

opendata CERN

Search

About

Filter by type

- Dataset 997
 - Collision 100
 - Derived 173
 - Simulated 723
- Documentation 56
 - About 8
 - Activities 19
 - Authors 3
 - Guide 16
 - Help 2
 - Policy 4
 - Report 1
- Environment 19
 - Condition 5
 - VM 11
 - Validation 3
 - Glossary 22
 - News 9
- Software 33
 - Analysis 16
 - Framework 4
 - Tool 8
 - Validation 5
- Supplementaries 2642
 - Configuration 917
 - Luminosity 3
 - Trigger 1722

Sort by: Most recent asc

Display: detailed 20 results

Found 3778 results.

< 1 2 3 4 5 6 7 8 9 >

/TtJets_MSDecays_scaleup_mt172_5_7TeV-madgraph-tauola/Summer11LegDR-PU_S13_START53_LV6-v1/AODSIM

Simulated dataset TtJets_MSDecays_scaleup_mt172_5_7TeV-madgraph-tauola in AODSIM format for 2011 collision data (SM Systematic Variations)

See the description of the sim...

Dataset Simulated CMS

/Vector1MtoZZto4L_M-125p6_7TeV-JHUGenV3-pythia6/Summer11LegDR-PU_S13_START53_LV6-v1/AODSIM

Simulated dataset Vector1MtoZZto4L_M-125p6_7TeV-JHUGenV3-pythia6 in AODSIM format for 2011 collision data (SM Inclusive)

See the description of the simulated dataset nam...

Dataset Simulated CMS

/VBFHiggs0PToGG_M-125p6_7TeV-JHUGenV4-pythia6-tauola/Summer11LegDR-PU_S13_START53_LV6-v1/AODSIM

Explore a variety of data types

Visualise detector events

open**data**
CERN

Search

About

Need HELP?

iSpy WebGL DoubleMu.jg:Events/Run_167674/Event_255544818 [3 of 25]

Detector

- Pixel Barrel
- Pixel Endcap (+)
- Pixel Endcap (-)
- Tracker Inner Barrel
- Tracker Outer Barrel
- Tracker Inner Detector (+)
- Tracker Inner Detector (-)
- Tracker Endcap (+)

CMS Experiment at the LHC, CERN
Data recorded: 2011-Jun-25 00:15:00.683123 GMT
Run / Event / LS: 167674 / 255544818 / 209

Click on a name under "Provenance", "Tracking", "ECAL", "HCAL", "Muon", and "Physics" to view contents in table

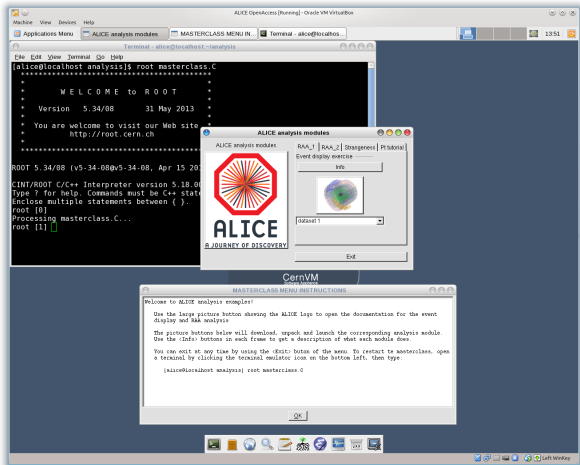
Interactive event display for high-level derived datasets

Visualise histograms



Interactive histogram plotting for high-level derived datasets

Virtual machines



Install CernVM virtual machines to explore primary datasets

Analysis examples

opendata
CERN

Search

About

Higgs-to-four-lepton analysis example using 2011-2012 data

Jomhari, Nur Zulaiha; Geiser, Achim; Bin Anuar, Afiq Aizuddin;

Cite as: Jomhari, Nur Zulaiha; Geiser, Achim; Bin Anuar, Afiq Aizuddin; (2017). Higgs-to-four-lepton analysis example using 2011-2012 data. CERN Open Data Portal. DOI:10.7483/OPENDATA.CMS.JKB8.RR42

Software Analysis CMS Accelerator CERN/LHC

Description

This research final example is a strongly simplified reimplement of parts of the original CMS Higgs to four lepton analysis published in [Phys.Lett. B716 \(2012\) 30-61, arXiv:1207.7235](#).

The published reference plot which is being approximated in this example is https://inspirehep.net/record/1124338/files/H4l_mass_3.png. Other Higgs final states (e.g. Higgs to two photons), which were also part of the same CMS paper and strongly contributed to the Higgs boson discovery, are not covered by this example.

The example consists of different levels of complexity. The highest level minimal understanding of the content of this paper and of the meaning, educational exercises. The lower levels might also be interesting for ed with the linux operating system and the ROOT analysis tool.

Use with

The example uses legacy versions of the original CMS datasets in the A publication due to improved calibrations. It also uses legacy versions o but not identical to, the ones in the original publication. These legacy d in many later CMS publications.

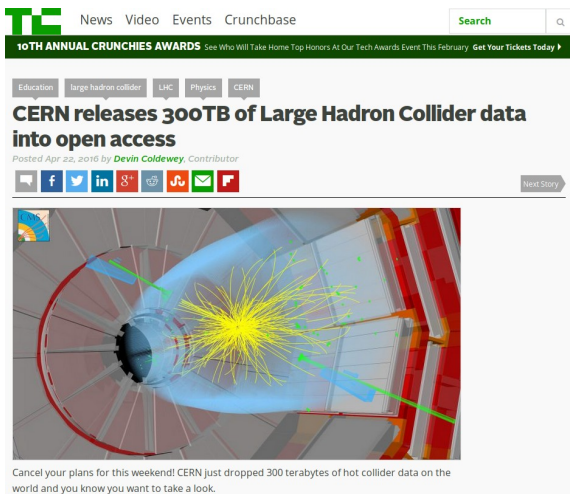
[/DoubleElectron/Run2011A-12Oct2013-v1/AOD](#)
[/DoubleMu/Run2011A-12Oct2013-v1/AOD](#)

Left plot: CMS Preliminary, $\sqrt{s} = 7 \text{ TeV}, L = 5.05 \text{ fb}^{-1}$, $\sqrt{s} = 8 \text{ TeV}, L = 5.28 \text{ fb}^{-1}$. Legend: Data (black dots), Z+X (green), Z γ , ZZ (blue), m_H=126 GeV (red). Y-axis: Events / 3 GeV. X-axis: m_H [GeV].

Right plot: CMS Open Data, $\sqrt{s} = 7 \text{ TeV}, L = 2.3 \text{ fb}^{-1}$, $\sqrt{s} = 8 \text{ TeV}, L = 11.6 \text{ fb}^{-1}$. Legend: Data (black dots), Z γ + X (green), TTBar (grey), ZZ \rightarrow 4l (blue), m_H = 126 GeV (red). Y-axis: Events / 3 GeV. X-axis: m_H [GeV].

Run realistic physics analysis examples

Press coverage



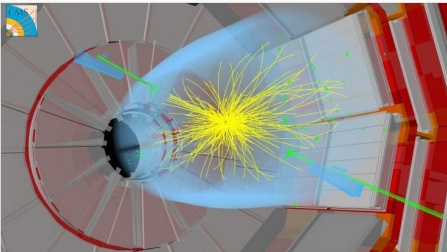
The image shows a screenshot of a TechCrunch news article. At the top left is the TechCrunch logo (TC) and navigation links for News, Video, Events, and Crunchbase. A search bar is on the top right. Below the navigation is a green banner for the '10TH ANNUAL CRUNCHIES AWARDS'. The article title is 'CERN releases 300TB of Large Hadron Collider data into open access'. It is categorized under Education, large hadron collider, LHC, Physics, and CERN. The author is Devin Coldewey, and it was posted on April 22, 2016. Social media sharing icons for Facebook, Twitter, LinkedIn, Google+, YouTube, Email, and Print are visible. A 'Next Story' button is on the right. The main image is a 3D visualization of a particle collision at the LHC, showing a central yellow starburst of particles and green tracks extending outwards. Below the image is a short paragraph of text.

Education large hadron collider LHC Physics CERN

CERN releases 300TB of Large Hadron Collider data into open access

Posted Apr 22, 2016 by [Devin Coldewey](#), Contributor

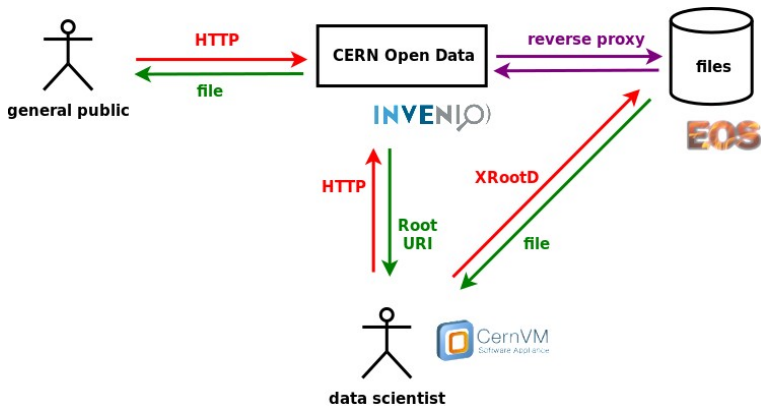
Next Story



Cancel your plans for this weekend! CERN just dropped 300 terabytes of hot collider data on the world and you know you want to take a look.

Open data releases are widely covered by world-wide media

Education vs research use patterns




Typical HTTP and XRootD access scenarios


Conclusions



Invenio


 <http://inveniosoftware.org>

 <http://github.com/inveniosoftware>

 @inveniosoftware



CERN Open Data

 <http://opendata.cern.ch>

 <http://github.com/cernopendata>

CERN IT H. Hirvonsalo, D. Rodriguez, T. Šimko · **CERN SIS** S. Dallmeier-Tiessen, S. Feger, P. Fokianos, A. Lavasa, I. Tsanaktsidis, A. Trisovic, A. Trzcinska · **ALICE** M. Gheata, C. Grigoras, M. Zimmermann · **ATLAS** K. Cranmer, L. Heinrich, A. Sanchez Pineda, D. Rousseau, F. Socher · **CMS** A. Calderon, E. Carrera, A. Geiser, A. Huffman, K. Lassila-Perini, T. McCauley, A. Rao, A. Rodriguez Marrero · **LHCb** S. Amerio, C. Burr, B. Couturier, S. Neubert, C. Parkes, S. Roiser · **CERN CernVM** J. Blomer · **CERN EOS** L. Mascetti, H. Rousseau · **CERN OpenShift** A. Lossent, A. Peon · **DASPOS** M. Hildreth · **DPHEP** J. Shiers