opendata CERN

Explore more than **two petabytes** of open data from particle physics!

Portal | Content | Website

The CERN Open Data portal is the access point to a growing range of open data and other content produced through the research performed at CERN.

The data is made available under the CC0 waiver.

Experiment data policies
All four LHC experiments have approved data preservation and access policies which state that they will make their data available.

Content on the portal
• Datasets of collision data, simulated events and simplified derived datasets
• Documentation
• Analysis examples, software frameworks and tools
• Supplemental material

Currently available
• More than 2 petabyte of datasets from experiments
• 7000 records
• 800000 files

Matter Antimatter Differences (B meson decays to three hadrons) - Data Files
This record contains the data set that is made available for the analysis described in the Project Notebook. The data was collected using the LHCb experiment in 2011 at a centre-of-mass energy of ...

Dataset | Collision | LHCb

Getting Started with ALICE Open Data
Learn how to use the ALICE virtual machine to have a first look at ALICE events and use analysis tools: How do I start the ALICE software?...

Documentation | Guide | ALICE | Getting Started

Filter by type
Dataset	2066
Collision	131
Derived	1010
Simulated	925
Documentation	64
About	9
Activities	19
Authors	5
Guide	22
Help	2
Policy	6
Report	1
Environment	26
Condition	9
VM	12
Validation	5
Glossary	33
News	11
Software	42
Analysis	17
Framework	4
Tool	16
Validation	5
Supplementaries	2701
Configuration	58
Configuration HLT	213
Configuration LHE	242
Configuration RECO	149
Configuration SIM	313
Luminosity	3
Trigger	1723

Filter by experiment
|	
ALICE	26
ATLAS	115
CMS	3918
LHCb	11
OPERA	835

Research | Example

Research level analysis example:
Higgs boson decay to four leptons
The open data analysis example recreates the official CMS analysis that was used to discover the Higgs boson in 2012.

• Data taken with the CMS detector in 2011-2012
• Analysis based on CMS software
• Full analysis processes about 70 TB of data
• Full analysis takes about a month on single machine, simplified version about 10 minutes
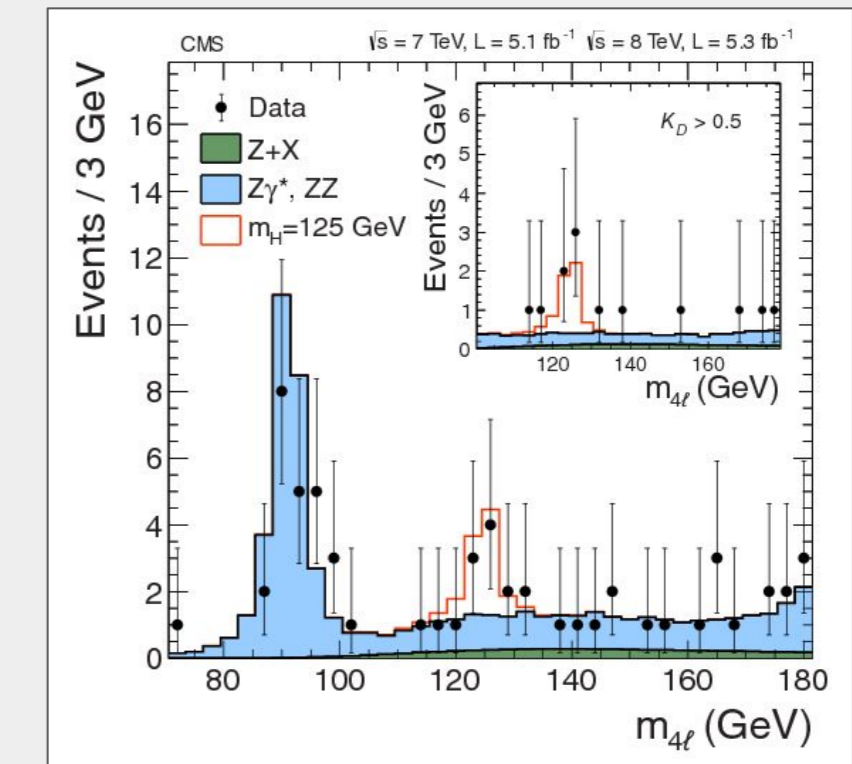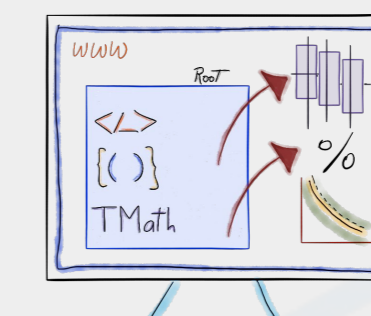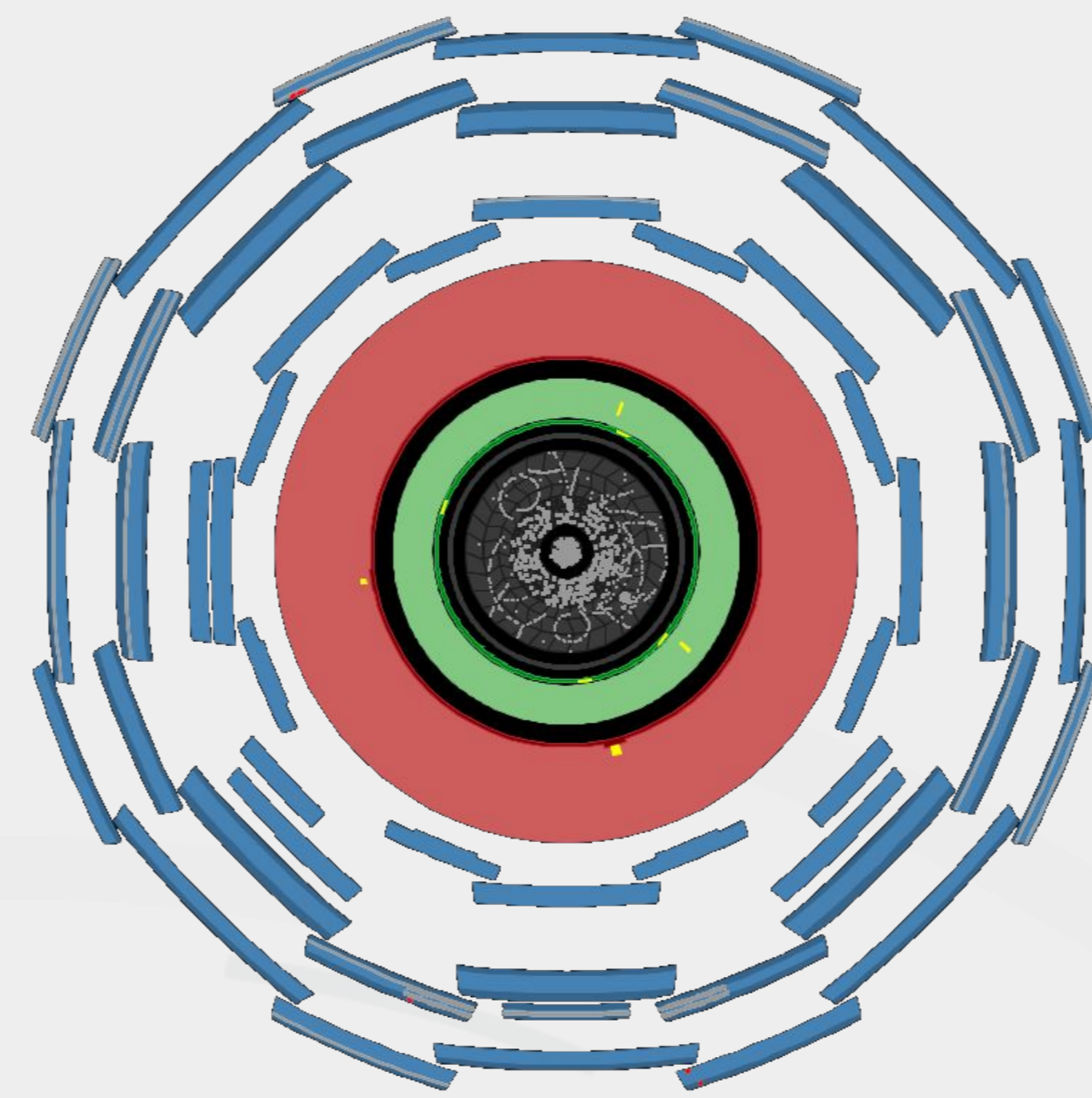
Published CMS result
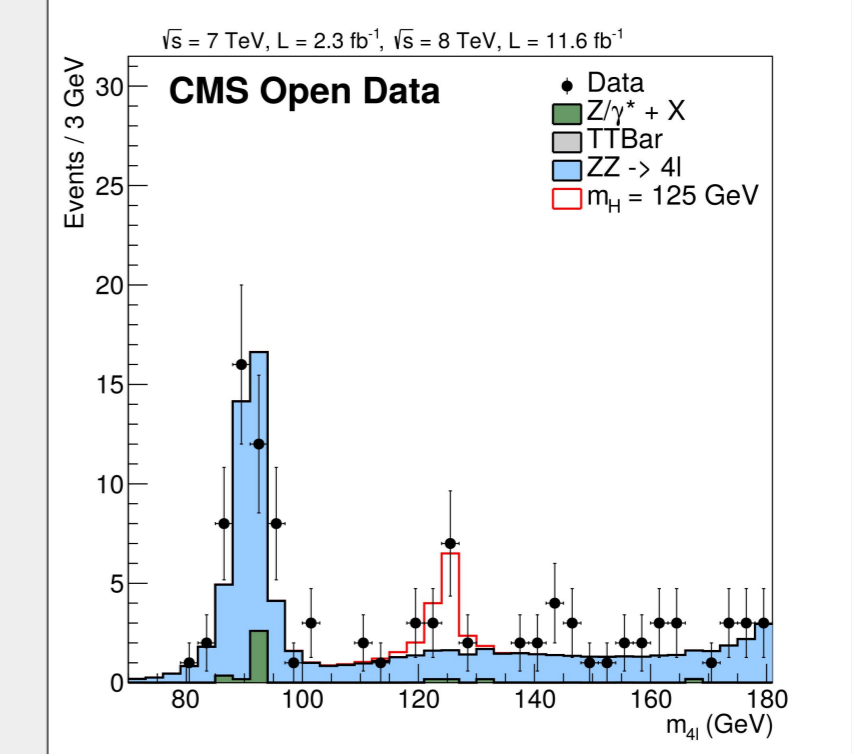
Open data analysis

Education | Outreach | Example

How can we overcome geographical distances to allow anyone interested in particle physics to learn?

• ATLAS collaboration launched a comprehensive educational platform to guide university-level students and teachers on how to use the data and analysis tools

• Provides simplified ATLAS Open Data which corresponds to 100 trillion proton-proton collisions

• ATLAS datasets and analysis tools have been optimised to fit on a USB memory stick

open data ATLAS

TMVA

INTERNATIONAL MASTERCLASSES
hands on particle physics

Education | Outreach | Example

Educational physics analysis examples
• Examples on high-school and university level
• Working with real physics data
• Supported by analysis code running out-of-the-box

Growing collection of analyses
• Rediscover particle resonances in a wide energy range up to the Z mass
• Study decays of a Higgs boson into two tau leptons
• And many more!
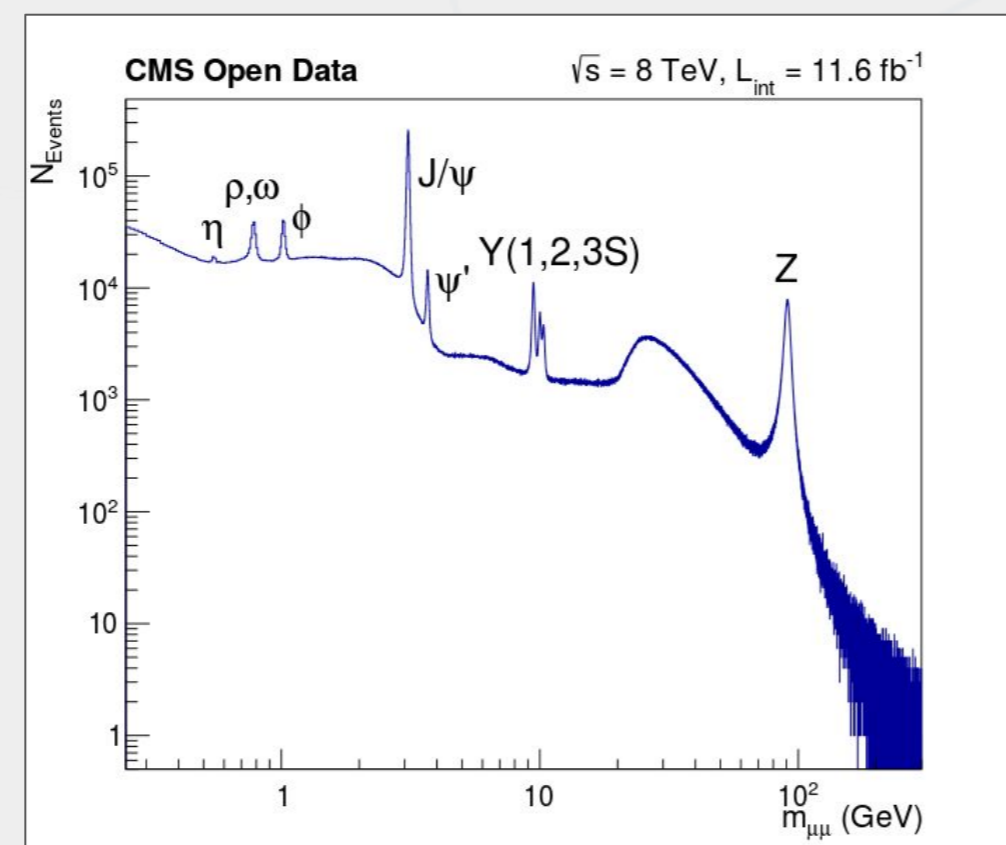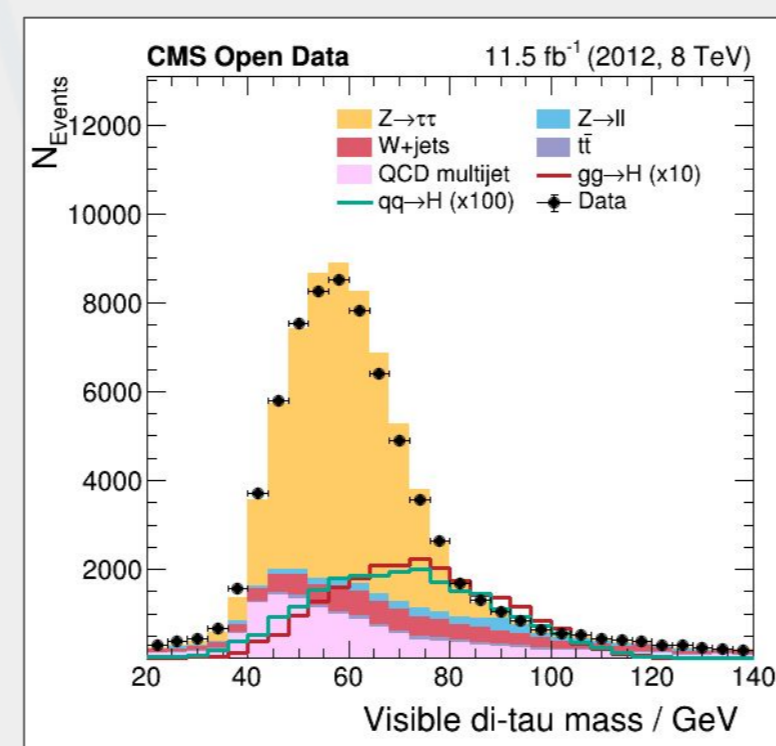
Portal | Technology

The CERN Open Data portal is built around following technologies (and many more).

INVENIO — **Invenio** Digital repository software framework that allows to build and run your own digital repositories

EOS Elastic and scalable disk-base storage system providing a low-latency storage infrastructure for open data

CernVM Complete, portable and easy to configure user environment for developing and running LHC data analysis

doi — **DOI** Infrastructure for the registration and use of persistent identifiers for use on digital networks

Preservation | Reproducibility

Reproducible research data analysis platform
Platform to ensure preservation and reproducibility of analysis results

• **Flexible** Run many computational workflow engines
• **Scalable** Support for remote compute clouds
• **Reusable** Containerize once, reuse elsewhere, natively in the cloud
• **Free** Free software using MIT licence

COMMON WORKFLOW LANGUAGE

reana
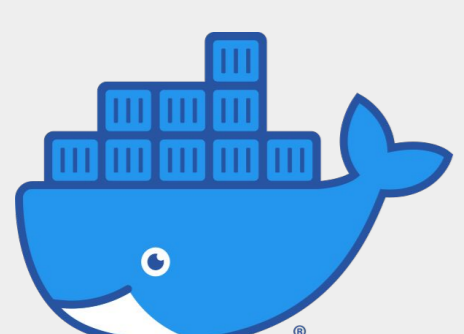
Research | Publications | Open science

The CERN Open Data effort strengthens open science resulting for example in publications in peer-reviewed journals.

Fast and Accurate Simulation of Particle Detectors Using Generative Adversarial Networks

Authors — Authors and a...
Pasquale Musella, Francesco Pandolfi

Jet substructure studies with CMS open data
Aashish Tripathee, Wei Xue, Andrew Larkoski, Simone Marzani, and Jesse Thaler
Phys. Rev. D **96**, 074003 – Published 3 October 2017

ALICE | ATLAS EXPERIMENT | CMS | LHCb
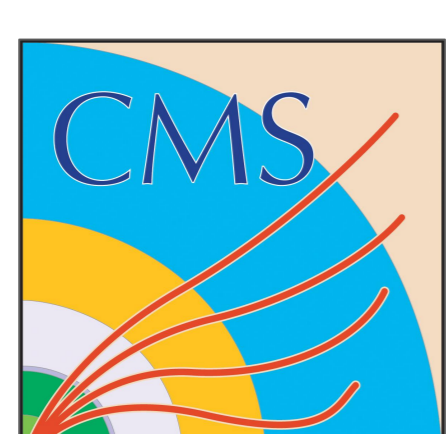
Get in touch
opendata.cern.ch

@cernopendata
github.com/cernopendata
opendata-support@cern.ch

open data | CERN