

Open data in CMS

Paavo Rikkilä

November 25, 2016

CMS Open Data

- CMS has published and will publish openly available data it collects from particle collisions
- suitable embargo period
- data are published on portal <http://opendata.cern.ch/>



Figure: CMS logo, © CMS

Four different preservation levels

- Level 1: Provide additional documentation
- Level 2: Preserve the data in a simplified format
- Level 3: Preserve the analysis level software and data format
- Level 4: Preserve the reconstruction and simulation software and basic level data

⇒ CMS open data on the portal are level 2 or level 3 data

Where the data can be used?

- Research

- ▶ data are in the same format as used in CMS (AOD, Analysis Object Data)
- ▶ CMS-specific analysis framework is provided as a Virtual Machine image

- Education

- ▶ data in derived and simplified formats
- ▶ access to data with VM or e.g. with Jupyter Notebook

Open Data for high school

- simplified format (CSV files)
- no ROOT – analysis with e.g. Python in Jupyter Notebook
- course for high school teachers in March at Viikin Normaalikoulu, Helsinki

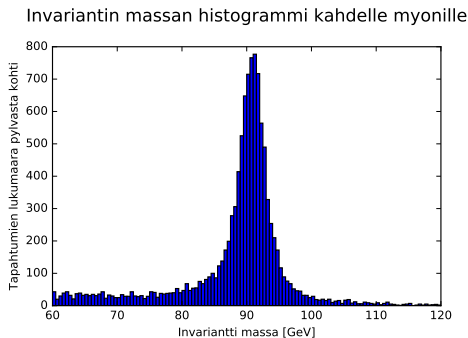


Open Data for university physics

- for particle physics courses, laboratory works, etc.
- for Bachelor's theses
 - ▶ students can easily explore real data and research tools
 - ▶ it takes only 10 minutes to start the VM and the first easy example

Conclusion

- lots of different possibilities for using the data
- course for high school teachers in Helsinki, also under planning in Jyväskylä
- for university physics I would like to hear your opinions and ideas



Thank you for your attention!