



# U Michigan: Research Semester Abroad



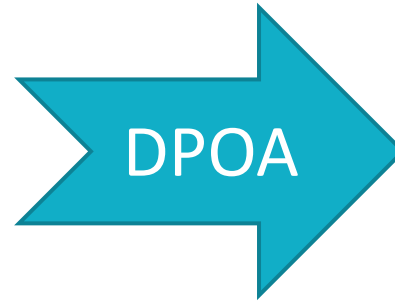
~Final Talks~

Diyaselis M. Delgado López





# On-Demand Distributed Workflows for Physics Analysis at the CMS Experiment



**reana**

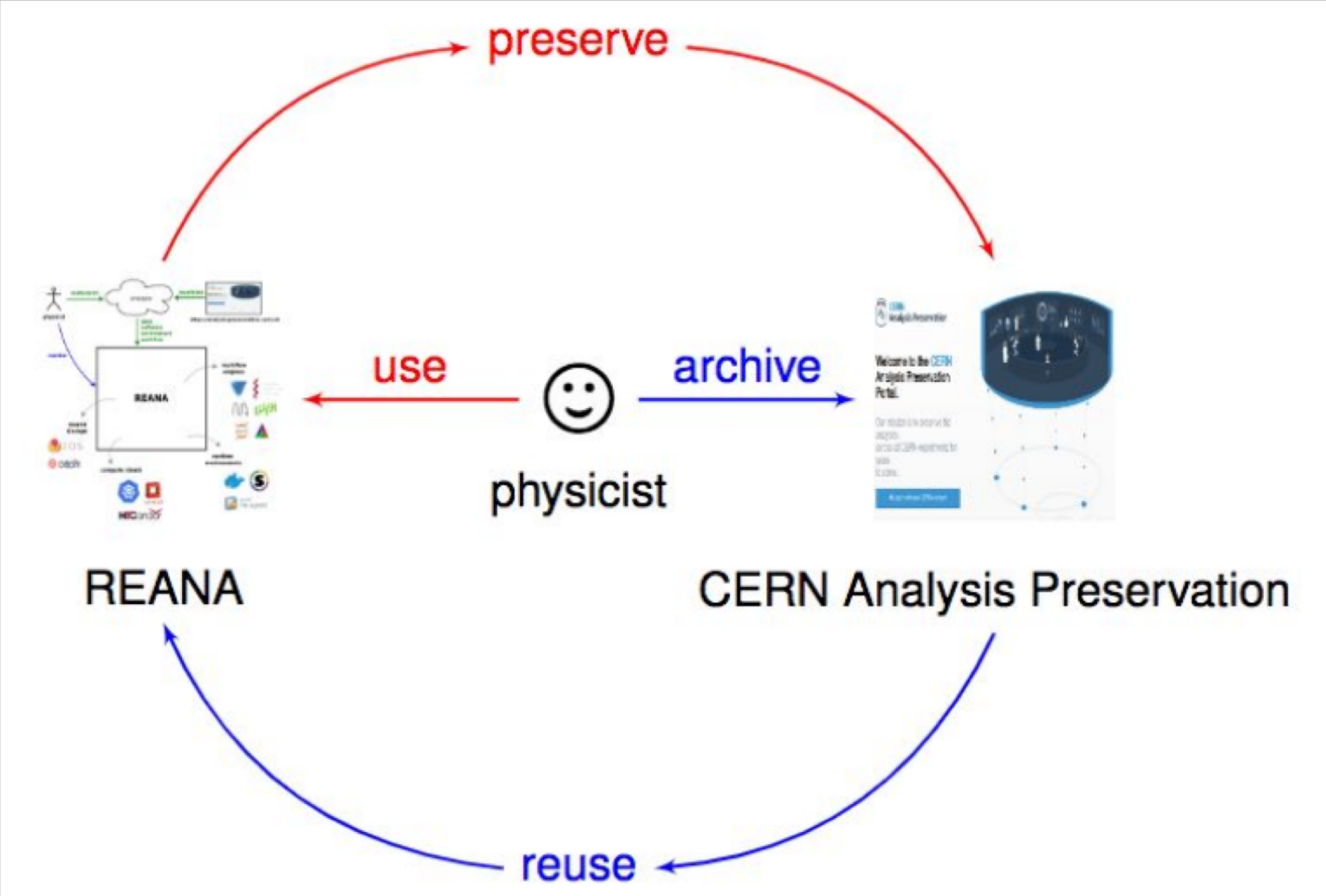
- Reusable and reproducible research data analysis platform



**CERN**  
ANALYSIS PRESERVATION

- Addresses the need for the long-term preservation of the data analysis process

CASE STUDIES IN REPRESENTATION OF CMS





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

# Project Analysis



Uses CMS Open  
Data inputs and  
available CMS  
Software  
environment

# Project Analysis



CAP preserves  
the analysis  
structure and  
metadata

# Project Analysis



ReANA, captures  
the analysis  
workflow and runs  
the commands to  
obtain the output



## Structure your analysis

- Input data
- Analysis code
- Containerized environments
- Computational workflows

## Structure your analysis

Input data

2011 – 2012 RAW data and MC

Analysis code

Containerized environments

Computational workflows

## Structure your analysis

Input data

2011 – 2012 RAW data and MC

Analysis code

Process data and plot

Containerized environments

Computational workflows

## Structure your analysis

Input data

2011 – 2012 RAW data and MC

Analysis code

Process data and plot

Containerized environments

CMSSW image from Docker

Computational workflows

## Structure your analysis

❑ Input data

2011 – 2012 RAW data and MC

❑ Analysis code

Process data and plot

❑ Containerized environments

CMSSW image from Docker

❑ Computational workflows

CWL (Common workflow)

# Structure your analysis

❑ Input data

2011 – 2012 RAW data and MC

❑ Analysis code

Process data and plot

❑ Containerized environments

CMSSW image from Docker

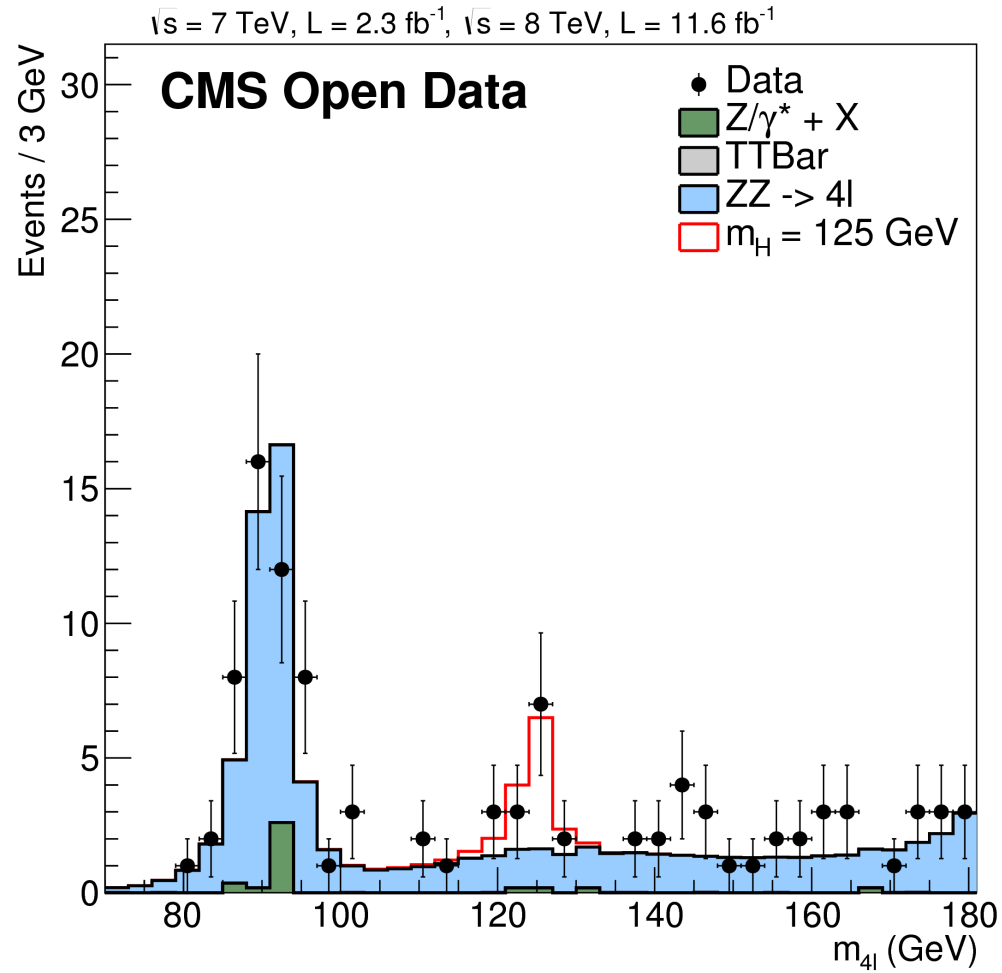
❑ Computational workflows

CWL (Common workflow)

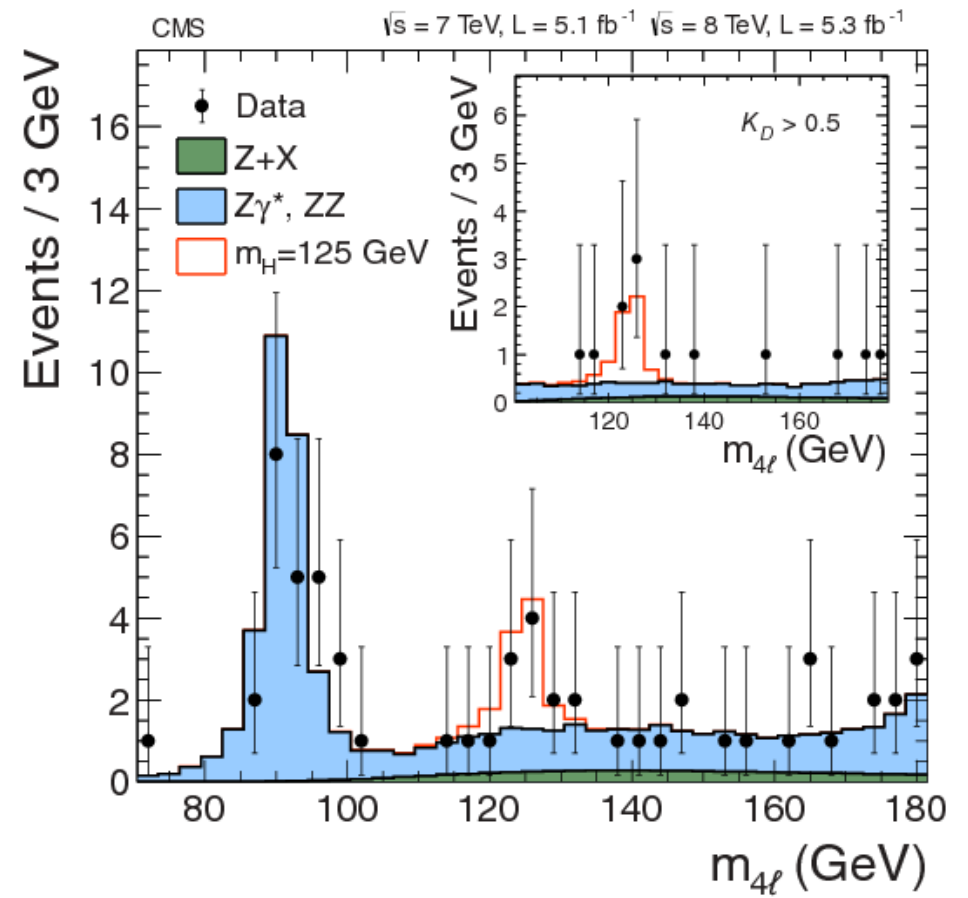
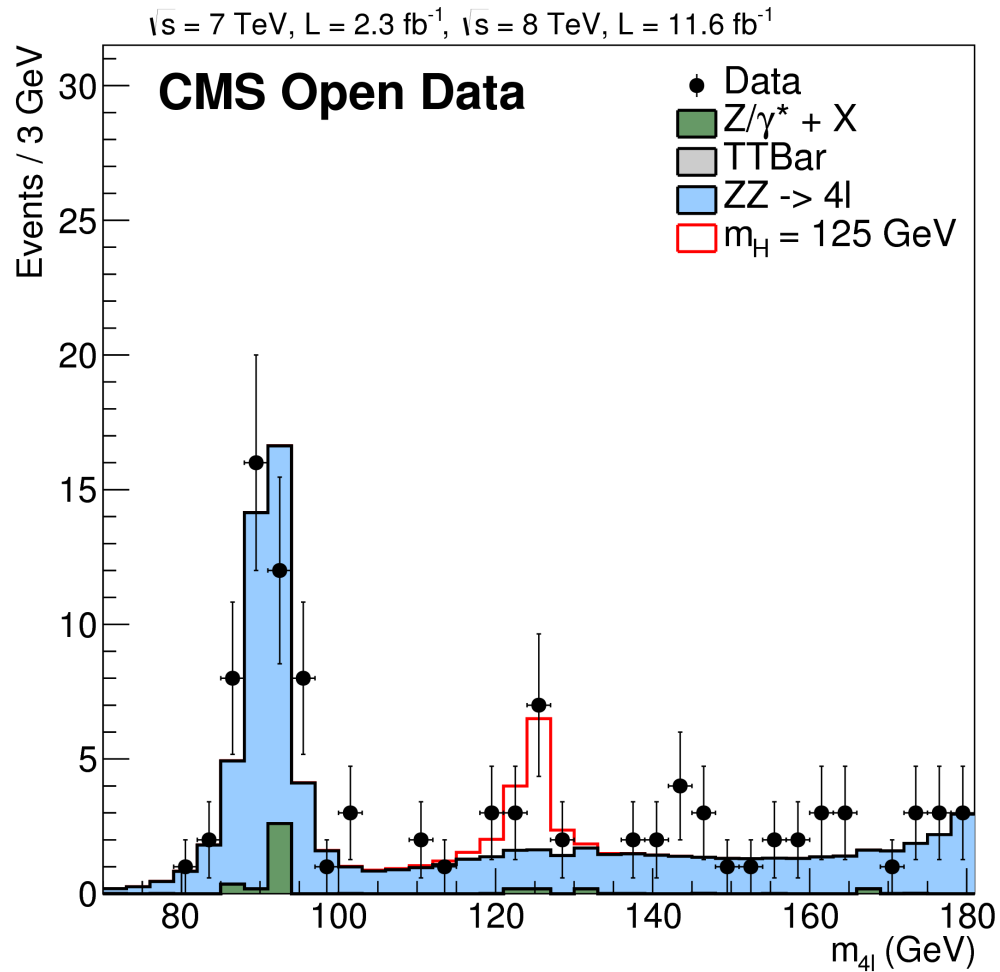
OUTPUT



# Higgs simplified reimplementation



# Higgs simplified reimplementation







# Reprocessing AOD from 2010 - 2012 RAW samples

The objective is to compare the outputs of reprocessed AOD files for 2010-2012 RAW samples with CMS Open Data VM results.

# Selected RAW Samples\*



## 2010 DATASETS

/MinimumBias/Run2010B-v1/RAW  
/Electron /Run2010B-v1/RAW  
/Mu /Run2010B-v1/RAW  
/Jet /Run2010B-v1/RAW



## 2011 DATASETS

/DoubleElectron/Run2011A-v1/RAW  
/SingleElectron /Run2011A-v1/RAW  
/DoubleMu /Run2011A-v1/RAW  
/SingleMu /Run2011A-v1/RAW  
/Jet /Run2011A-v1/RAW

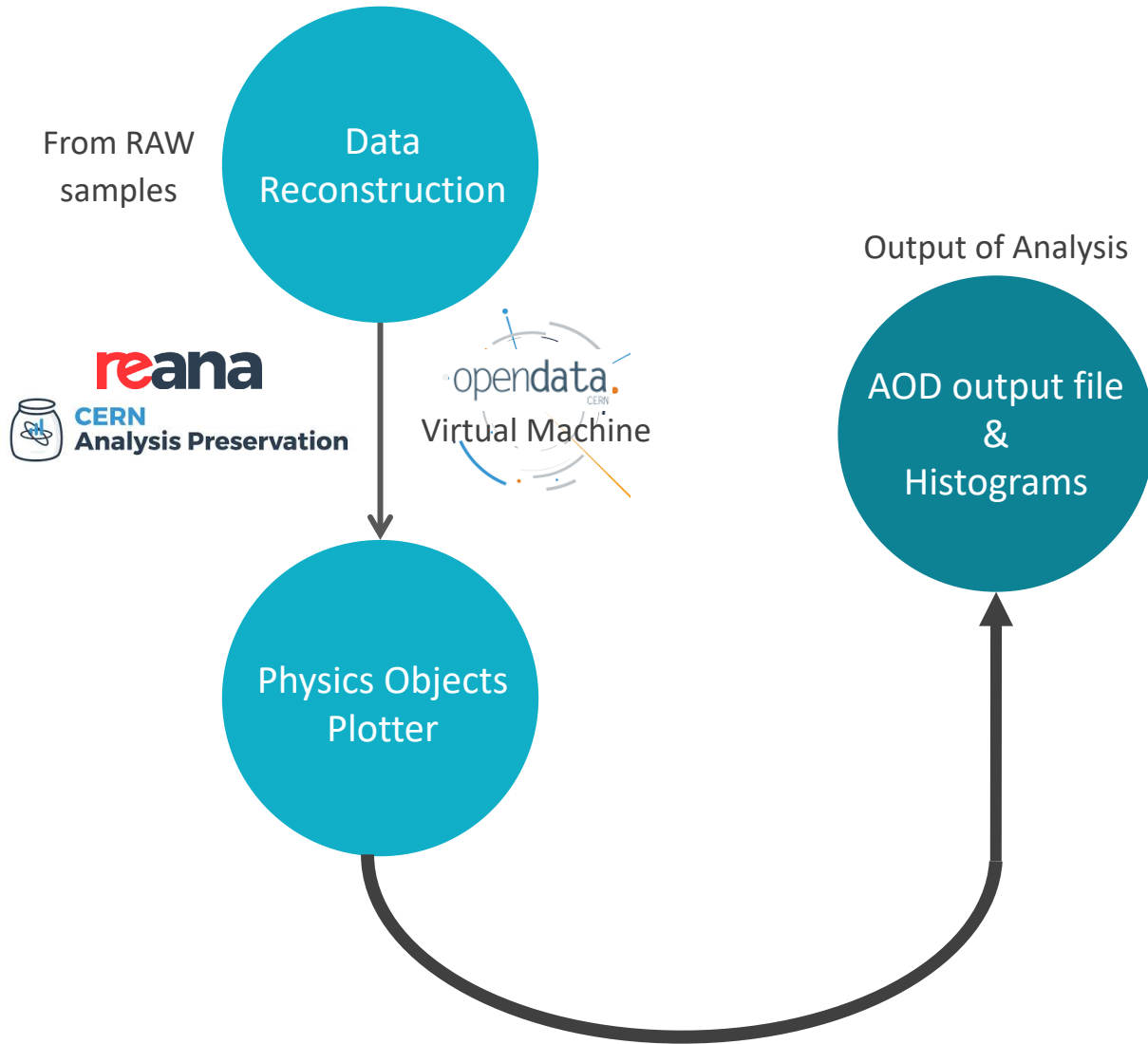


## 2012 DATASETS

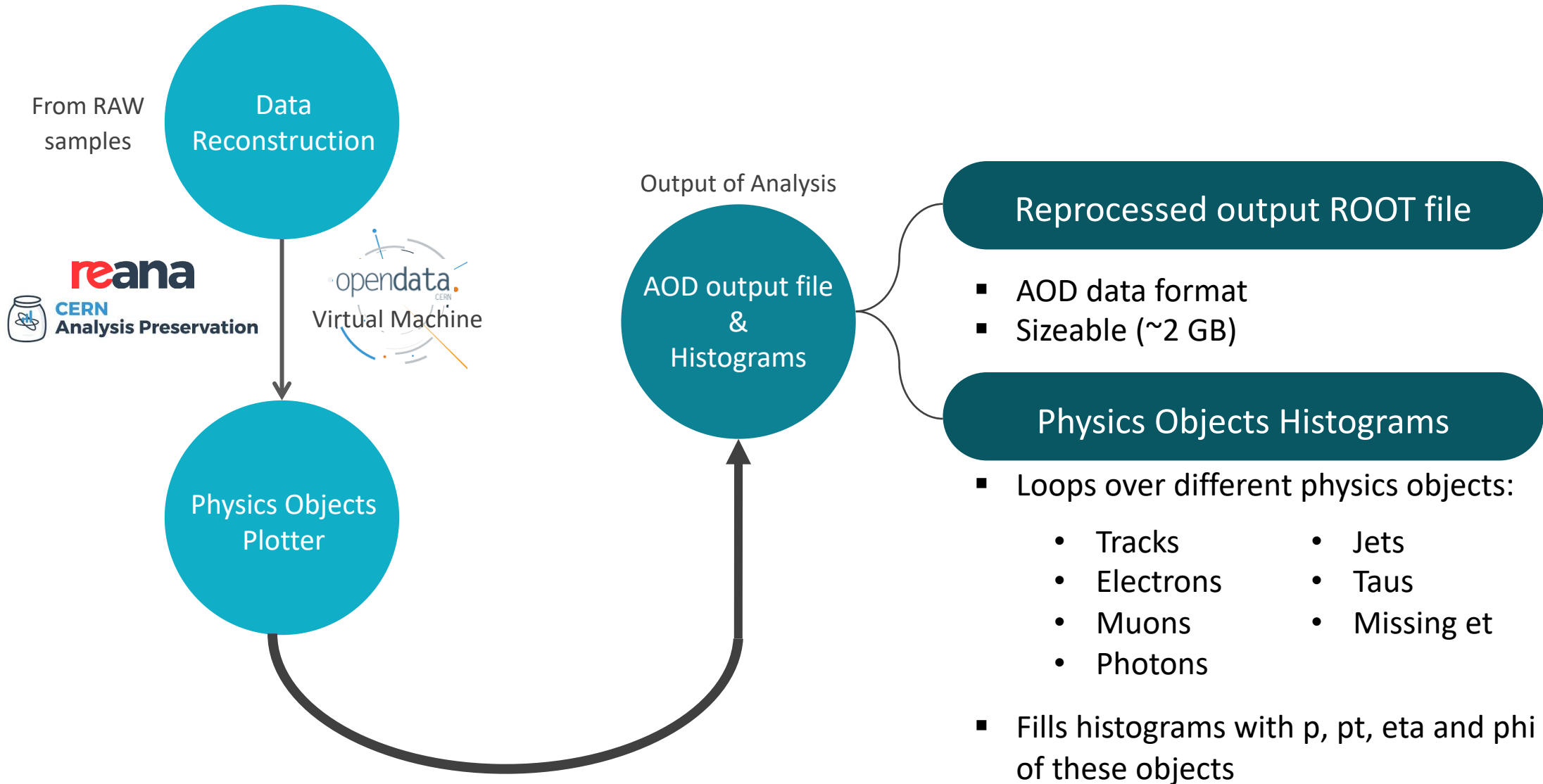
/DoubleElectron/Run2012B-v1/RAW  
/SingleElectron /Run2012B-v1/RAW  
/DoubleMuParked/Run2012B-v1/RAW  
/SingleMu /Run2012B-v1/RAW  
/JetHT/Run2012B-v1/RAW

\*TO BE RELEASED TO OPEN DATA PORTAL

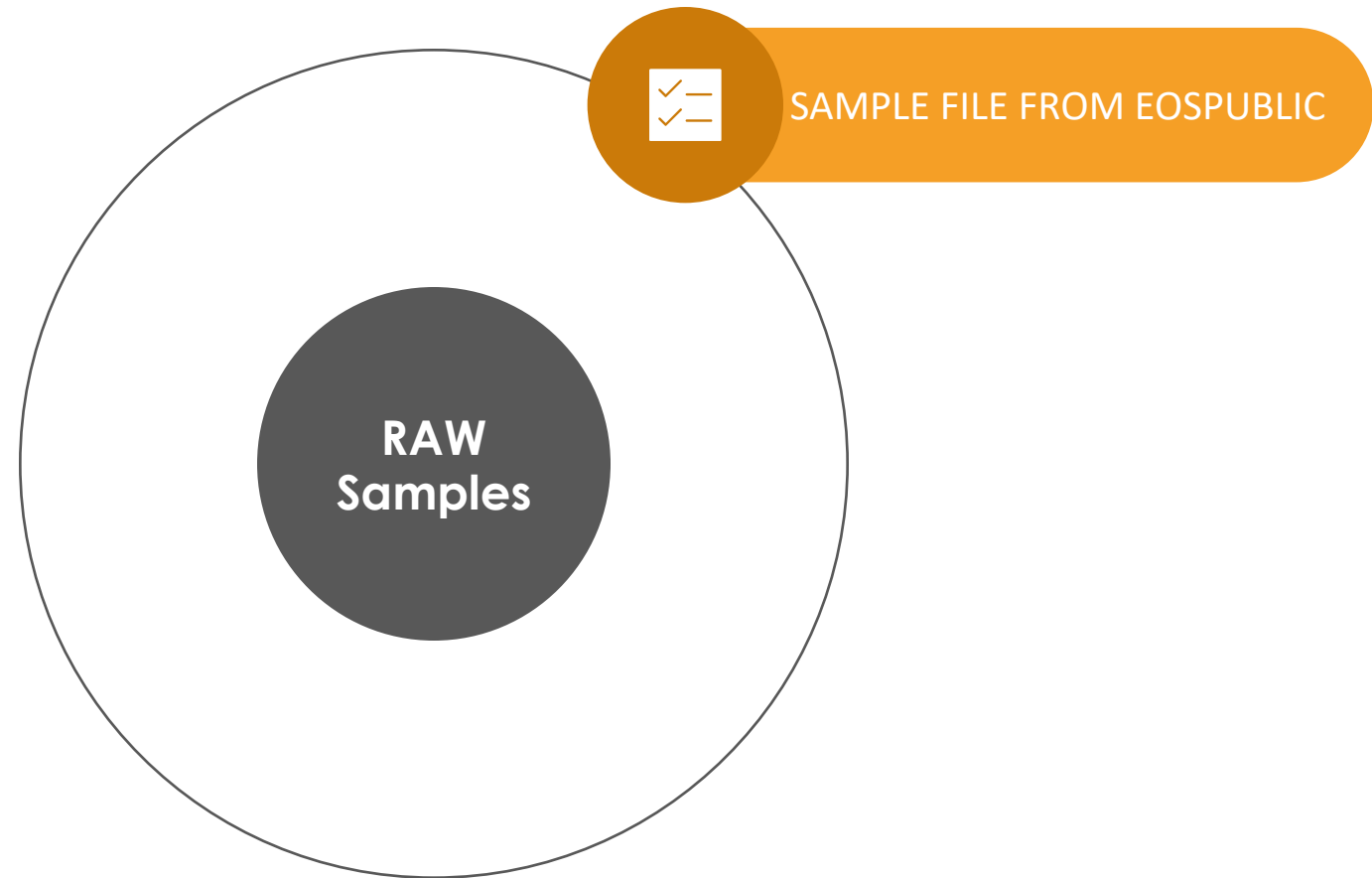
# Project Analysis



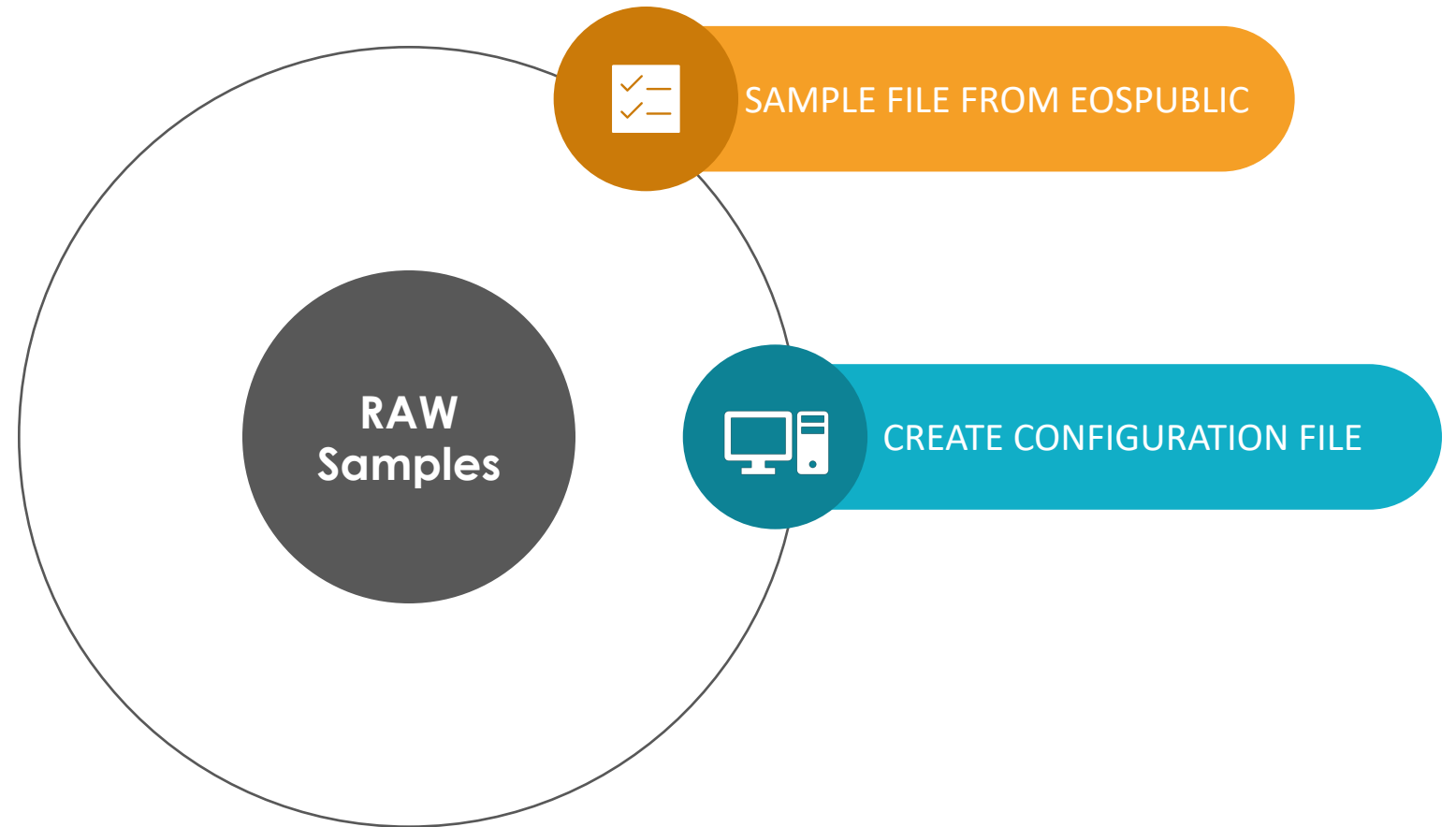
# Project Analysis



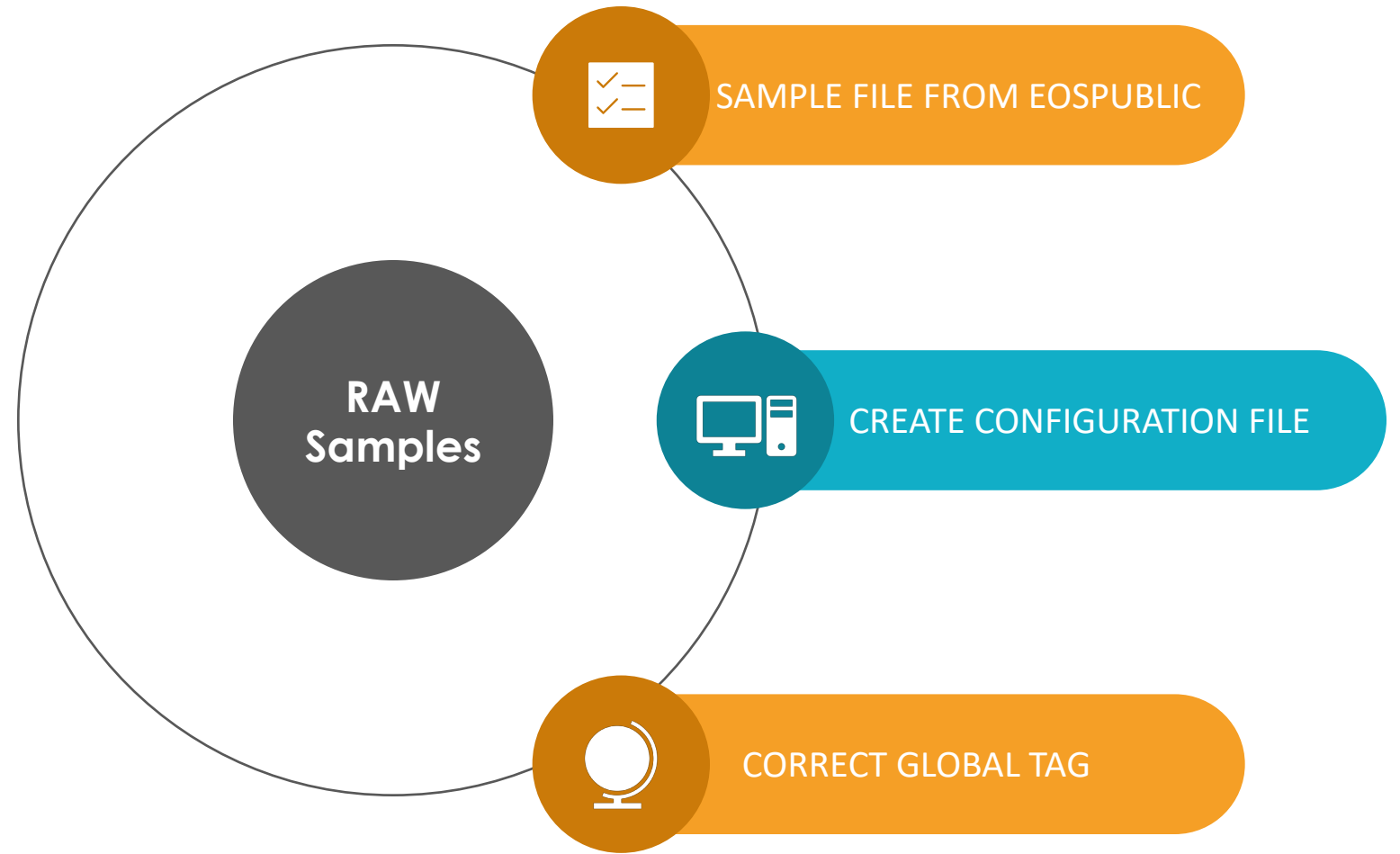
# Data Reconstruction



# Data Reconstruction

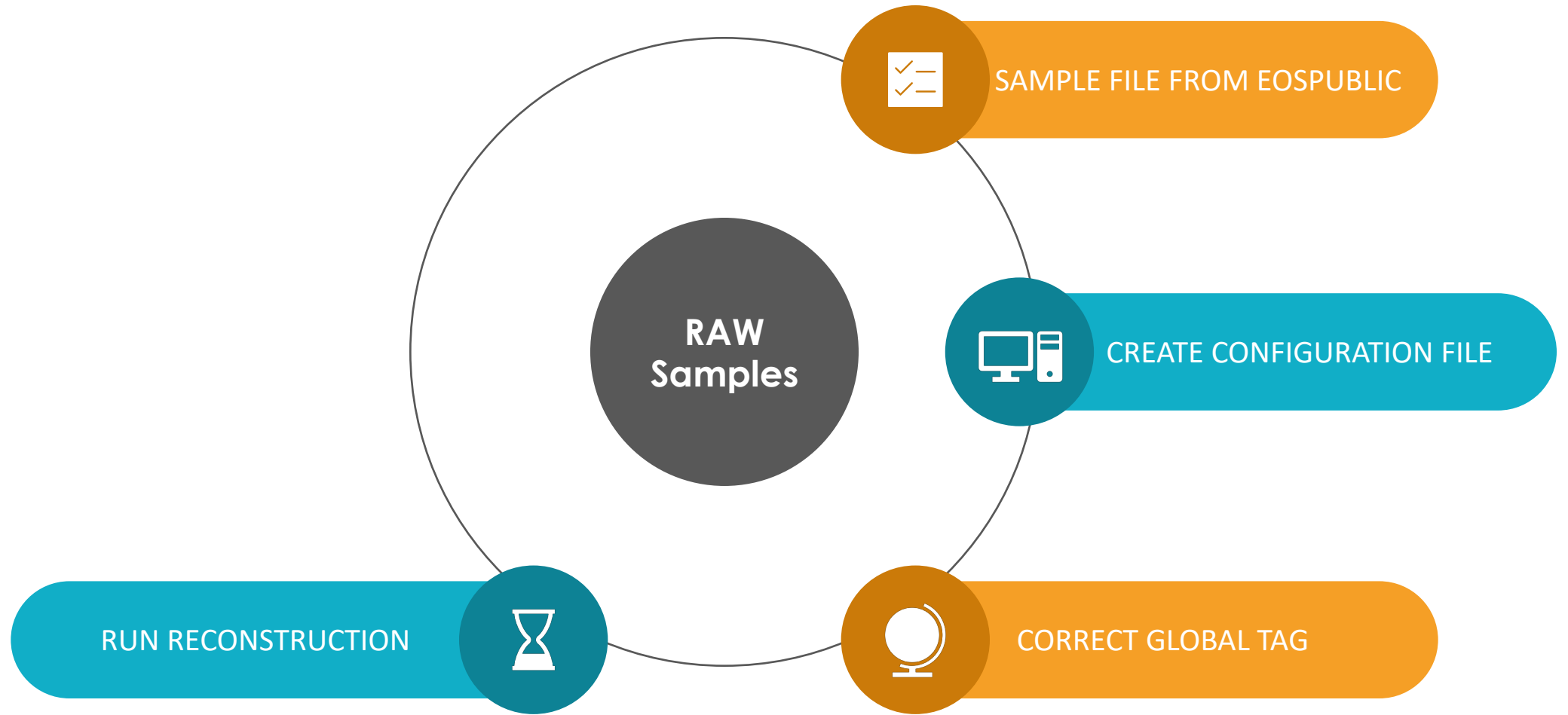


# Data Reconstruction

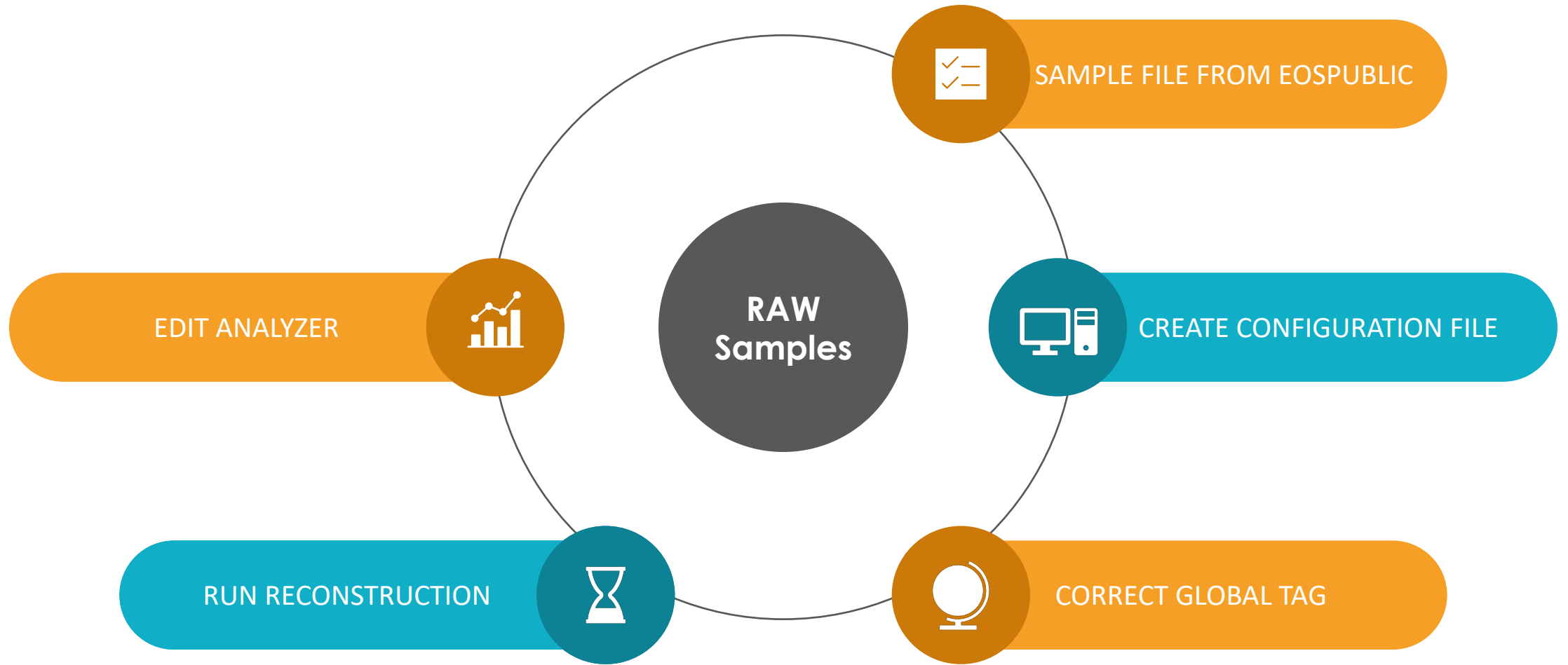




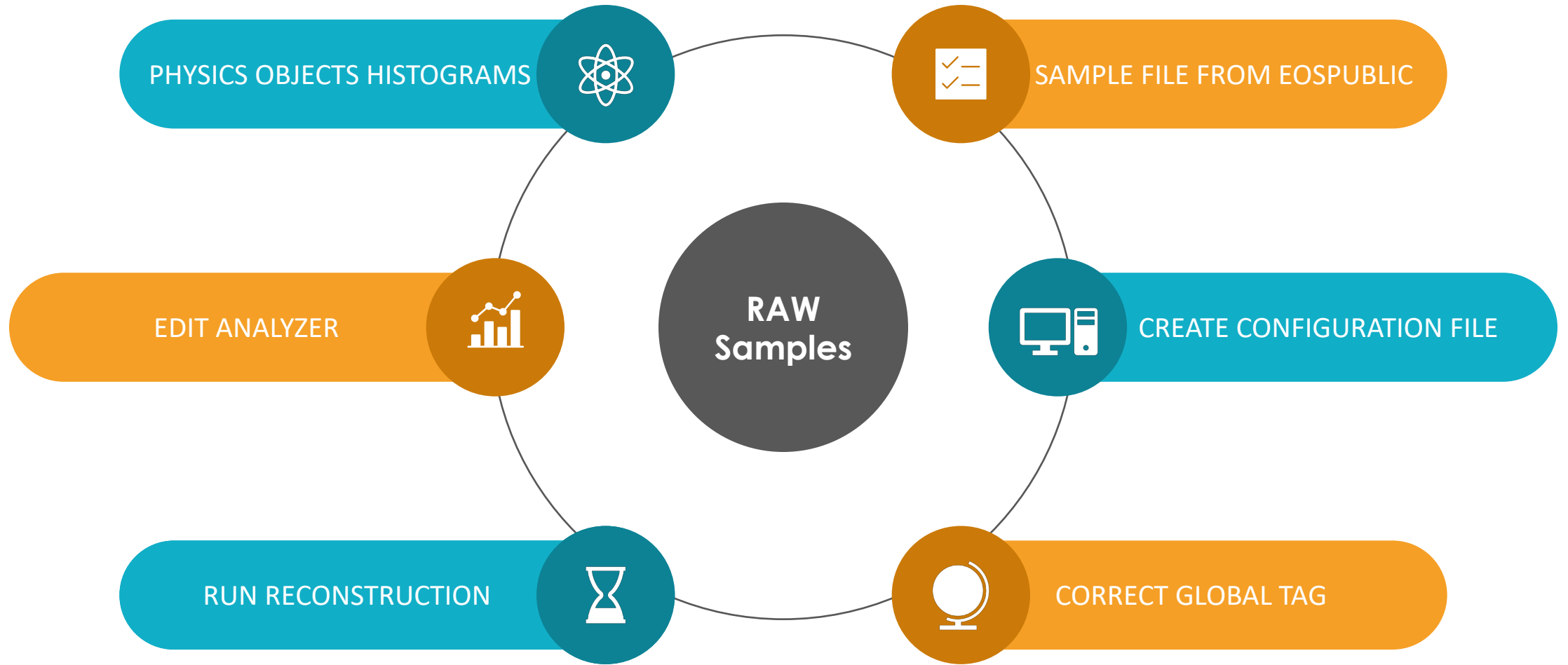
# Data Reconstruction



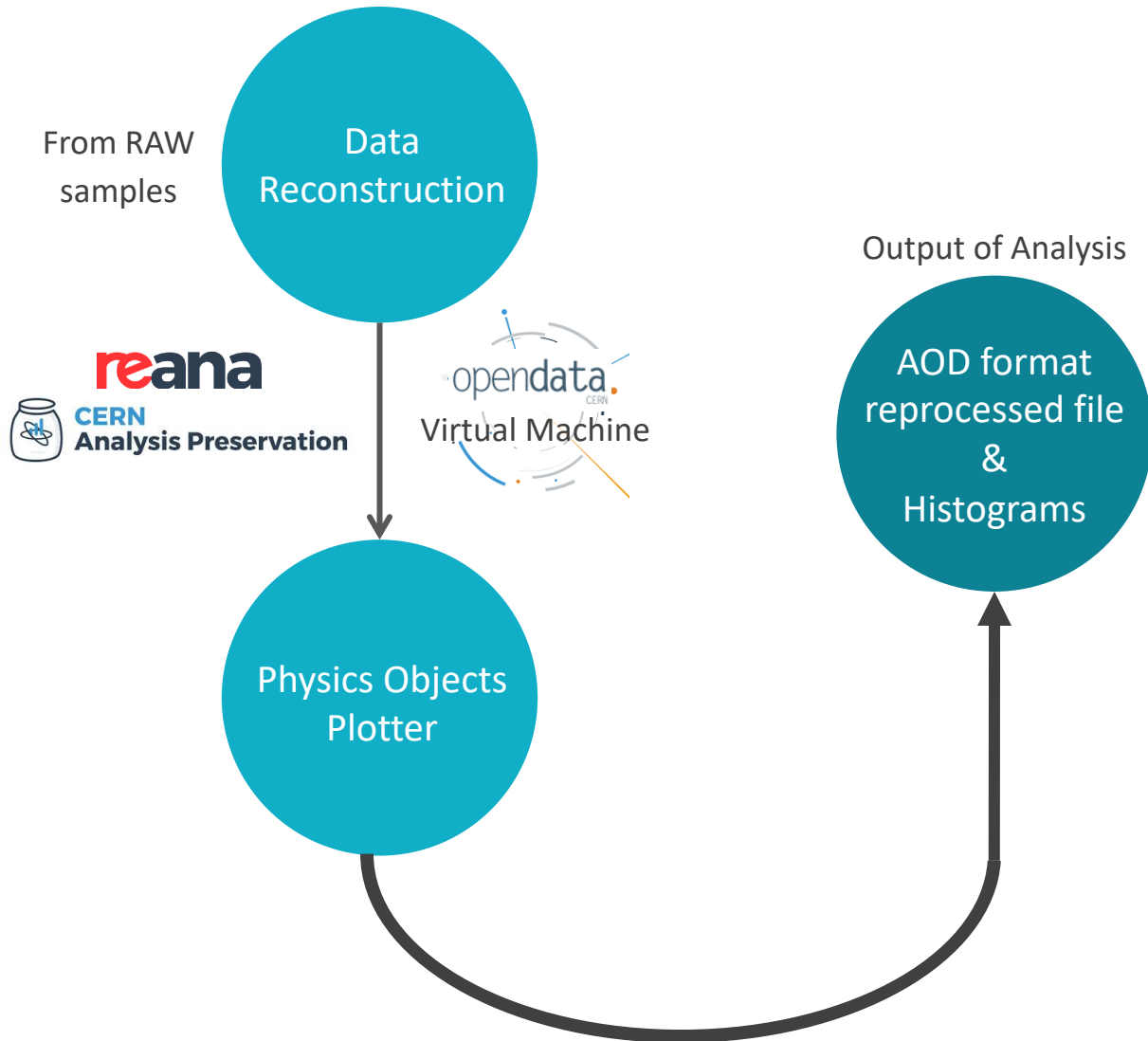
# Data Reconstruction



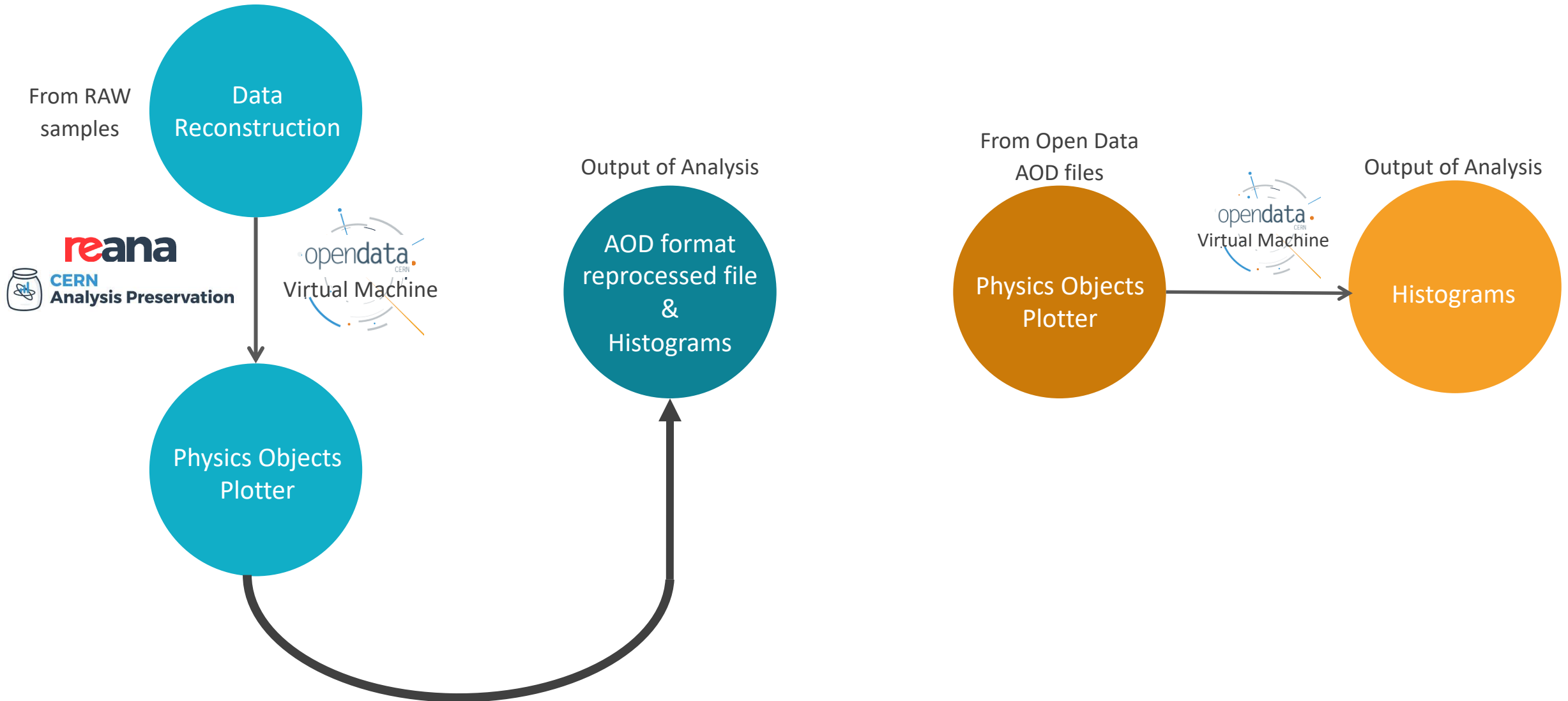
# Data Reconstruction



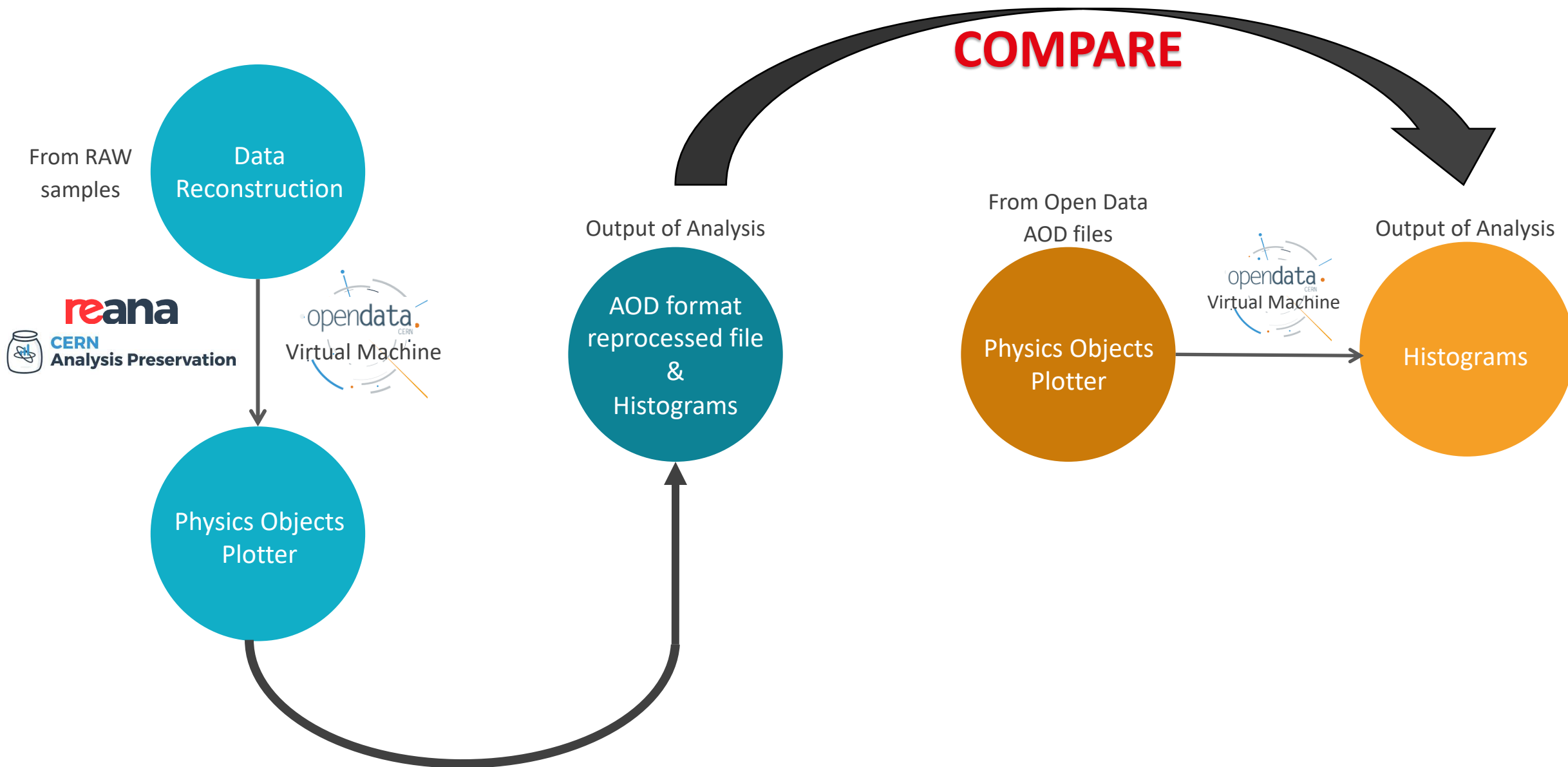
# Project Analysis



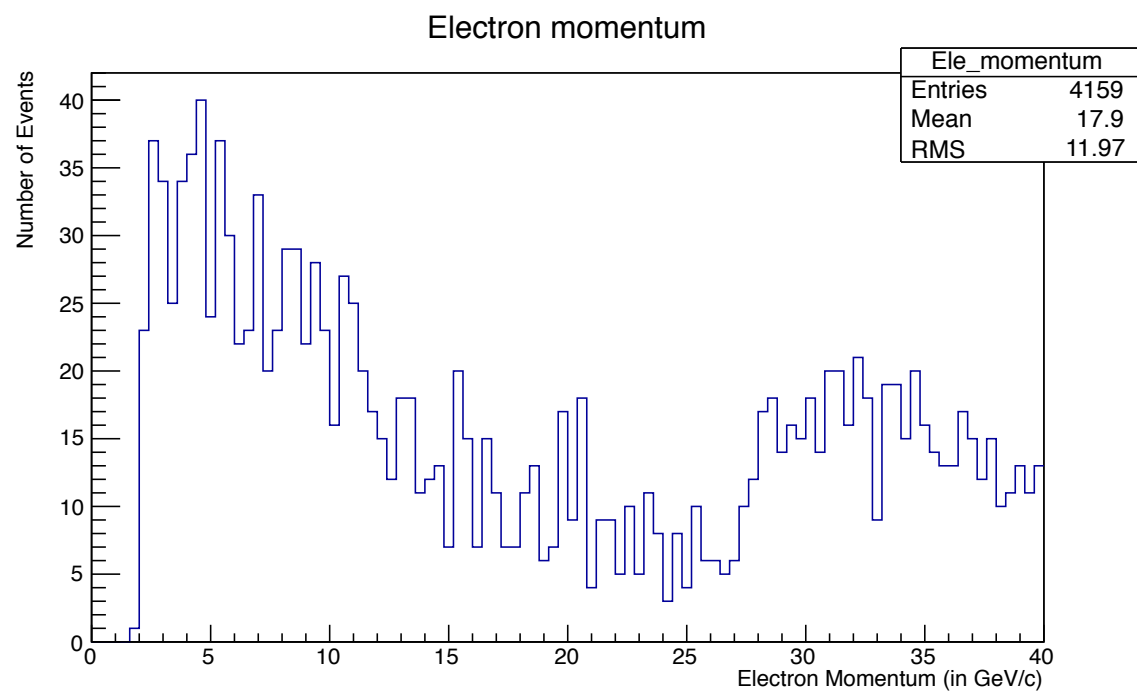
# Project Analysis



# Project Analysis

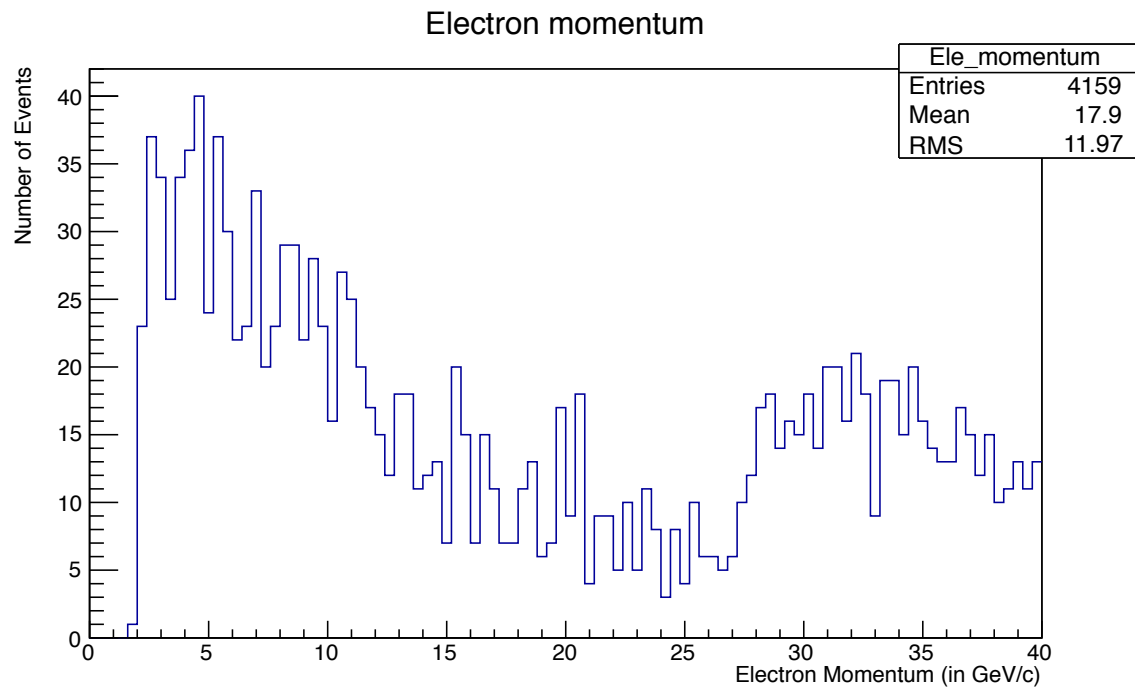


## Reconstruction from RAW samples

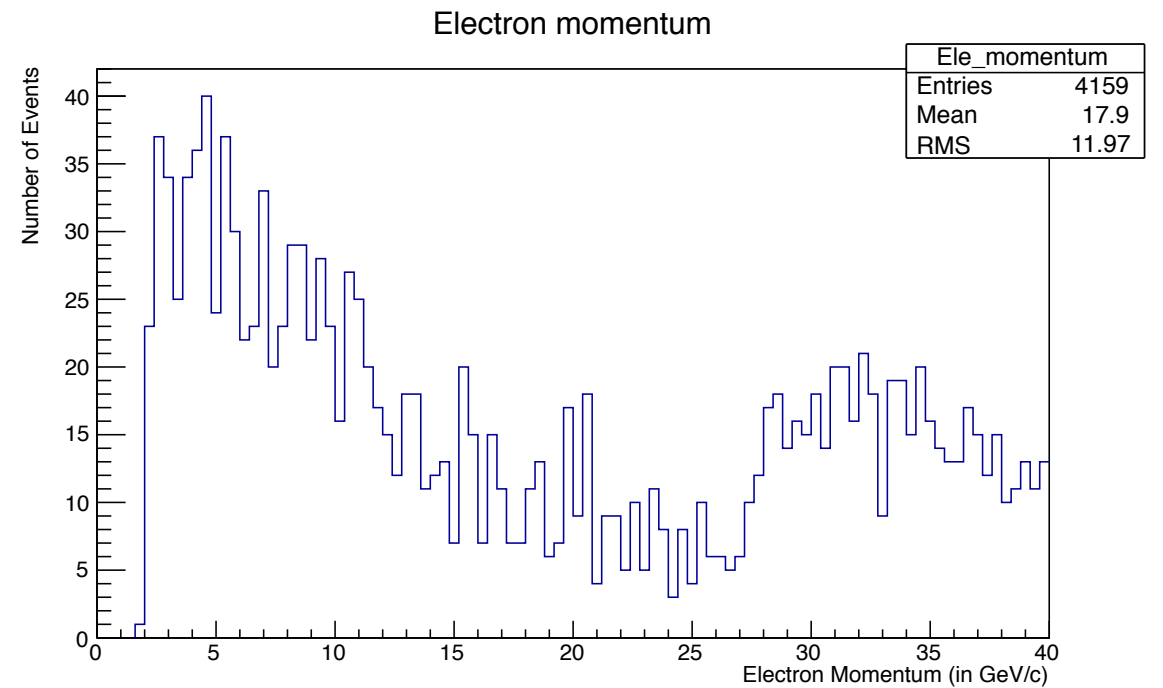


# Run2011A: SingleElectron

Reconstruction from  
RAW samples



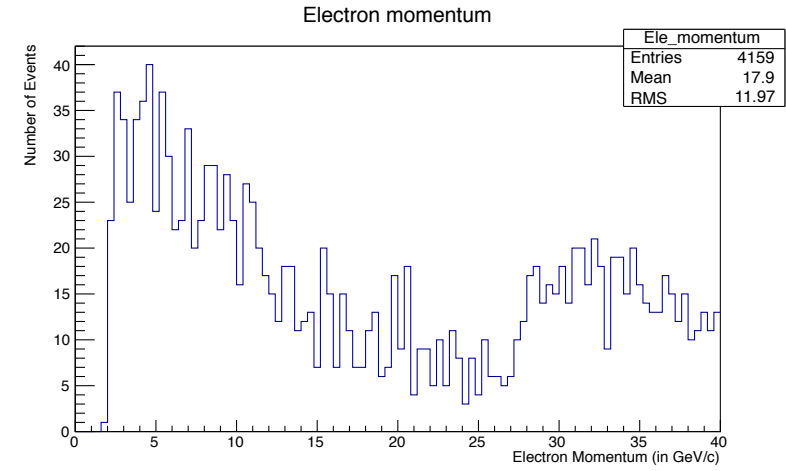
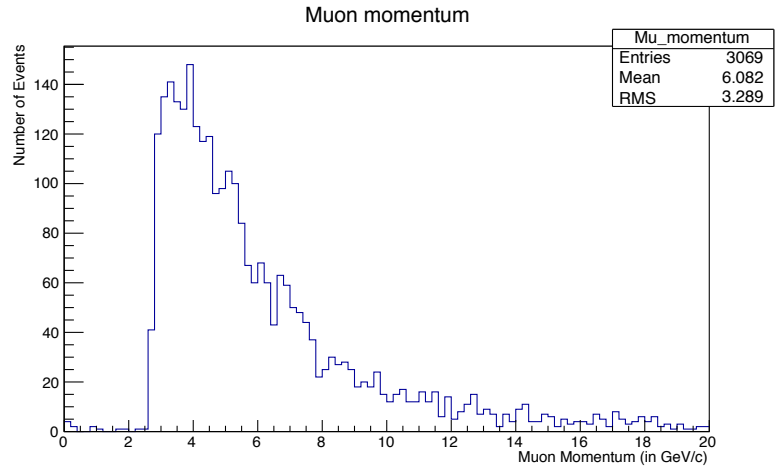
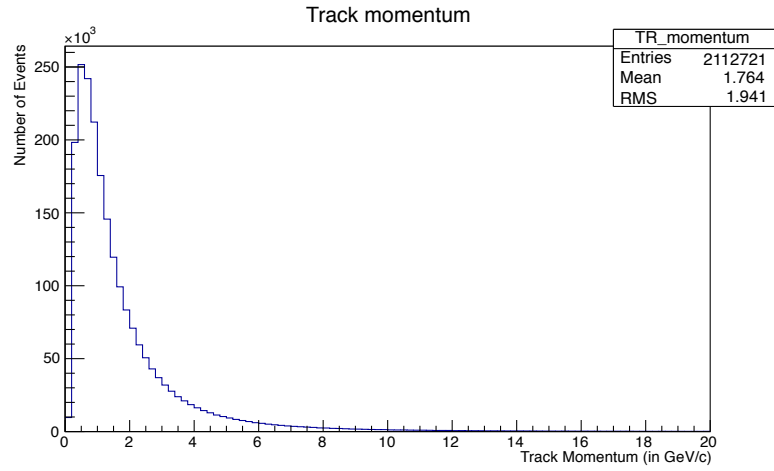
CMS Open Data  
AOD files



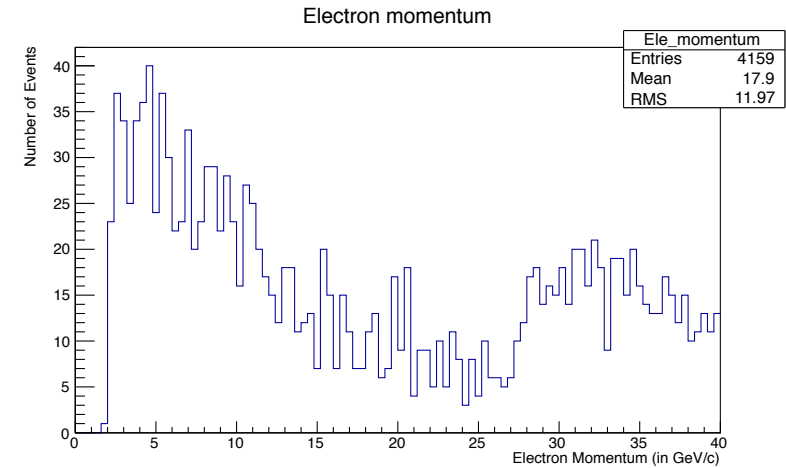
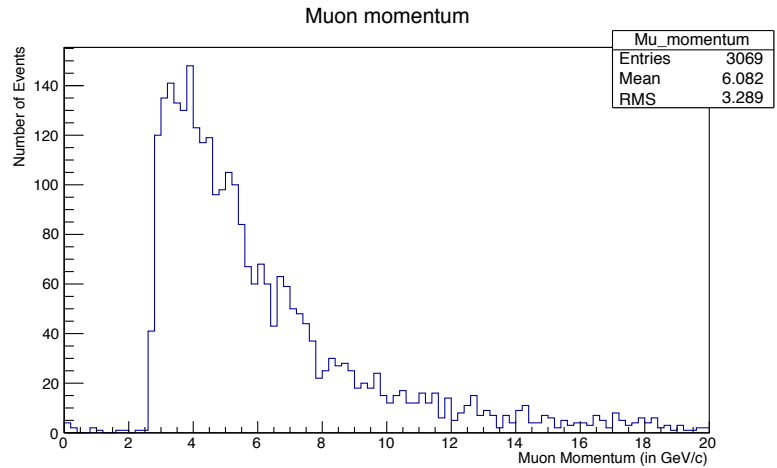
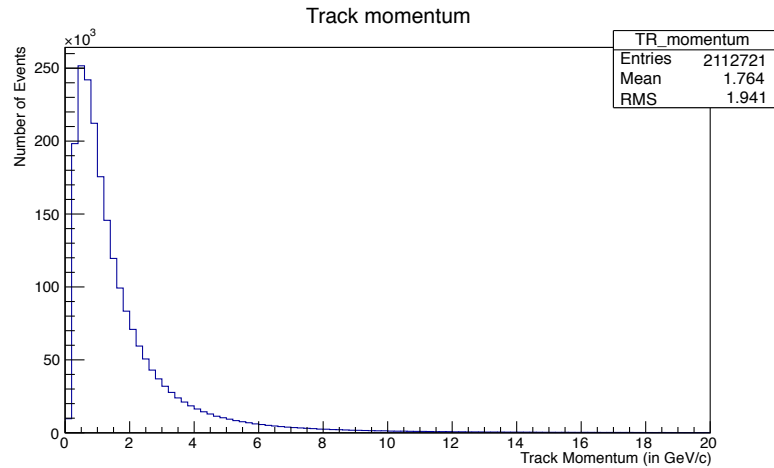


# Run2011A: SingleElectron

Reconstruction from  
RAW samples

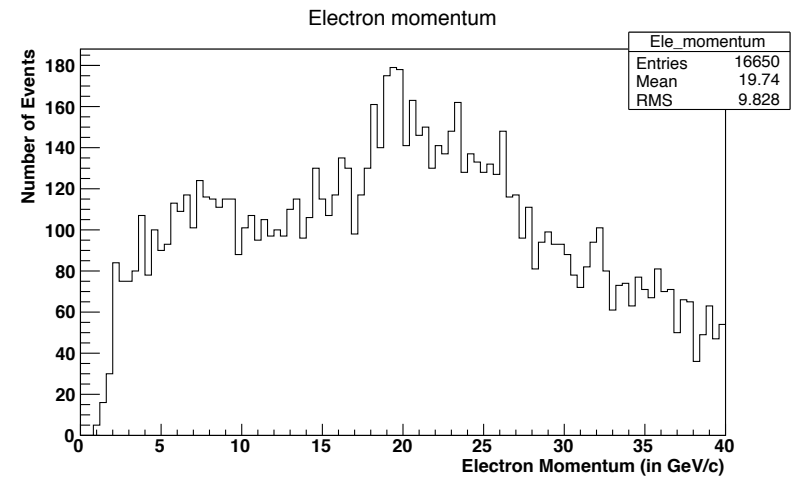
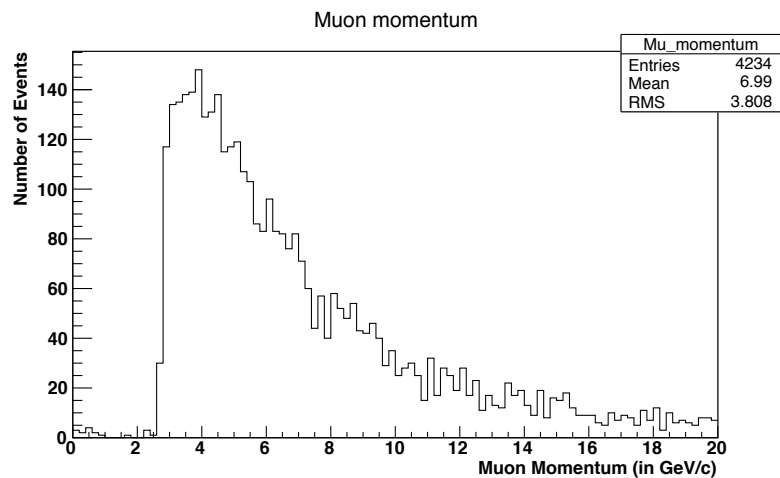
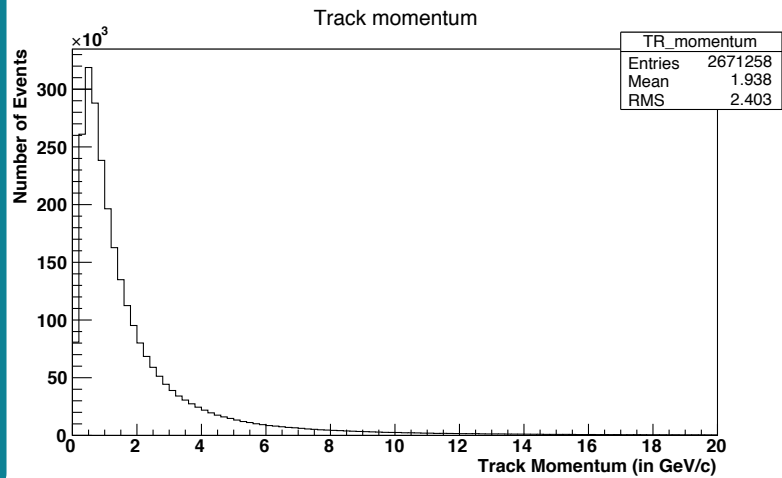


CMS Open Data  
AOD files

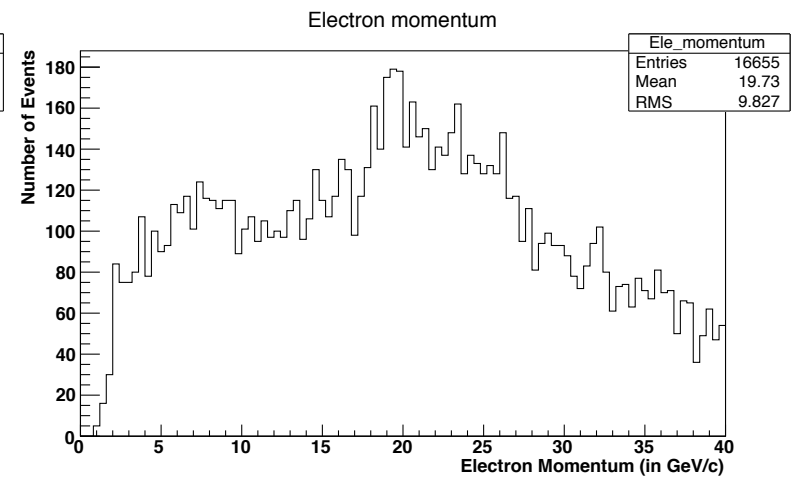
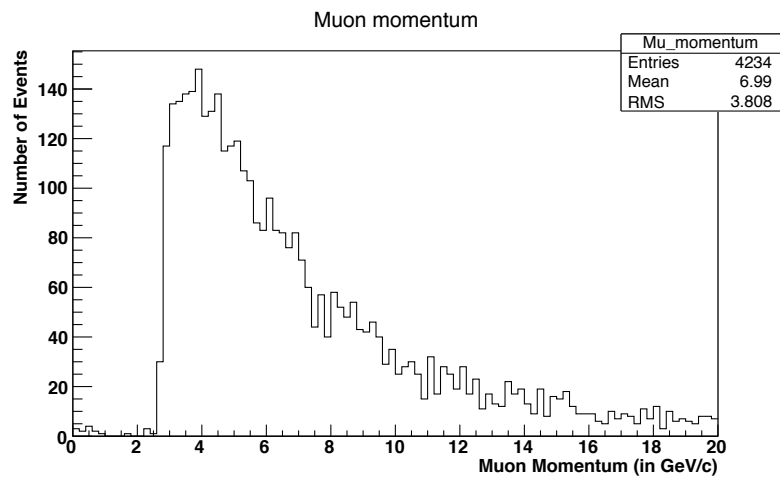
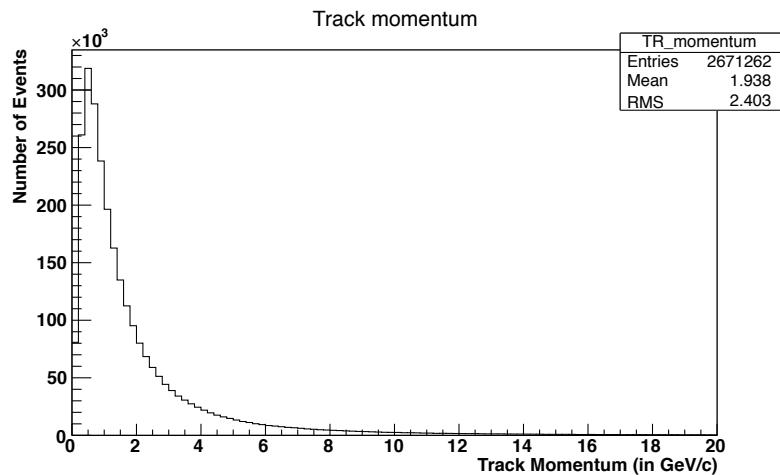


# Run2010B: Electron

Reconstruction from  
RAW samples



CMS Open Data  
AOD files



## Plots Summary

All RAW samples were reconstructed successfully, and have one-to-one match with the original AOD.

## Plots Summary

All RAW samples were reconstructed successfully, and have one-to-one match with the original AOD.

The logo for ReANA, featuring the word "reana" in a lowercase, sans-serif font. The "re" is in red and the "ana" is in dark blue.

- Workflows are ready  
Problems with reading condition data from /cvmfs are under investigation by the ReANA team (symbolic links)
- Docker container must be able to read from cvmfs also



Thank You