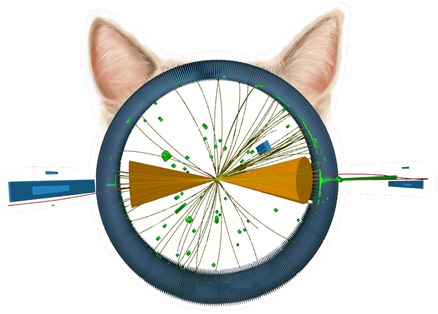
Common Analysis Tools in CMS

Clemens Lange (Paul Scherrer Institute) on behalf of the CMS Collaboration

Common Analysis Tools (CAT) group mandate

Take ownership of the **development, maintenance and documentation of analysis tools** that are considered of common interest to the collaboration.
Provide a **forum to discuss developments** of new analysis tools, and offer guidance such that some of these **tools can become centrally maintained** and handled as in point 1.

Three subgroups:
- Data processing tools
- Workflow orchestration and analysis preservation
- Statistical inference tools



Promoting data processing tools

Several data processing frameworks are supported through the CAT group with different approaches based on:
- CMSSW (Mini- and NanoAOD)
- ROOT's RDataFrame (NanoAOD)
- awkward arrays (NanoAOD)

Frameworks are required to provide **documentation, user support, container images,** and **analysis examples**.

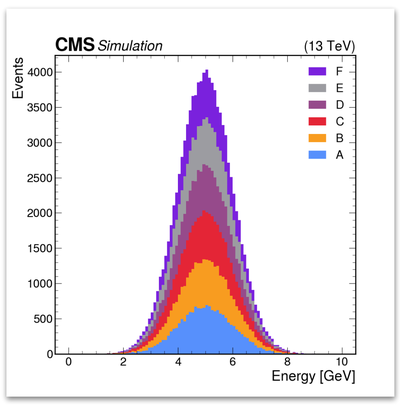
Training the collaboration

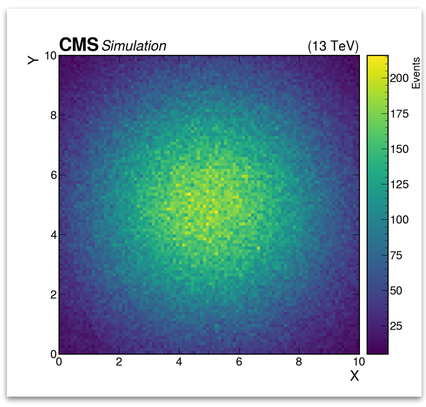
To provide newcomers with guidance, an **"Analysis 101"** is provided to explain how physics analyses are typically performed and brought through the internal review process to publication.

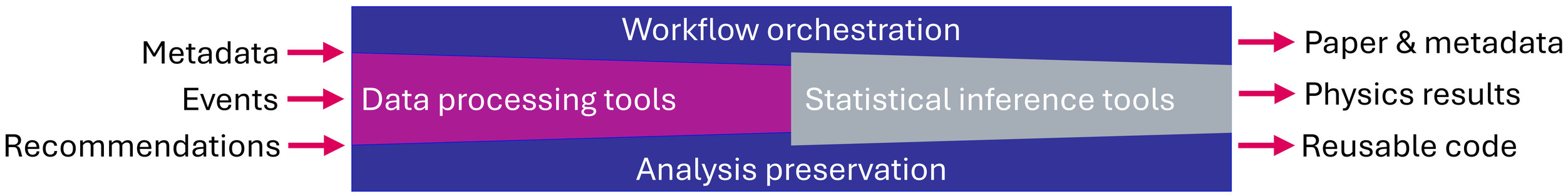
Regular **training events** on analysis tools are organised, partly based on material by the HEP Software Foundation Training Working Group.

Common plot styles

CMS plots should be **easily recognisable** and also be **readable for colour vision impaired people**.
A new common plot style with various colour scheme options has been developed (for use with both **PyROOT** and **matplotlib**).

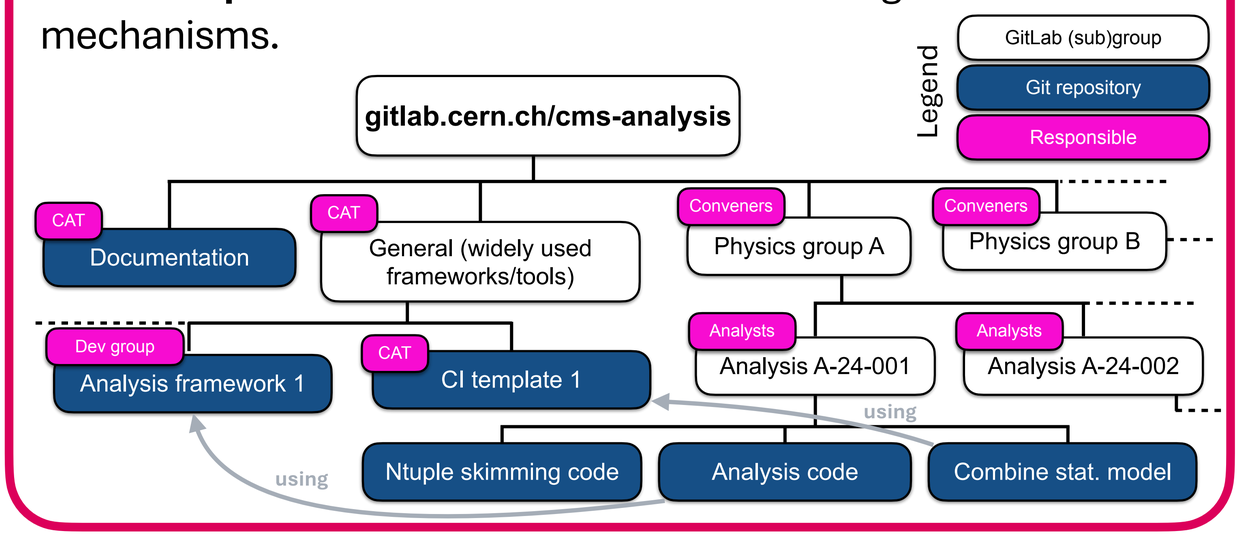






Making analysis preservation easy

Using the GitLab installation at CERN, a **central analysis code area** is provided for anyone in CMS, following the structure of the physics groups.
This area is enhanced with templates and services for **automatic code quality and functionality checks** as well as **seamless data access capabilities** for use with continuous integration mechanisms.



Rediscovering the Higgs boson with public statistical models

The **CMS Combine tool** is used for statistical inference by the overwhelming majority of CMS physics analyses.
This tool has recently been released to the public under a **free and open source license** and is already used outside the collaboration.
Along with the software, the **full statistical model** used for the **Higgs boson discovery** is now publicly available.
Furthermore, a mechanism to release statistical models for all CMS analyses on the CERN Document Server has been established.

Scan to view available models