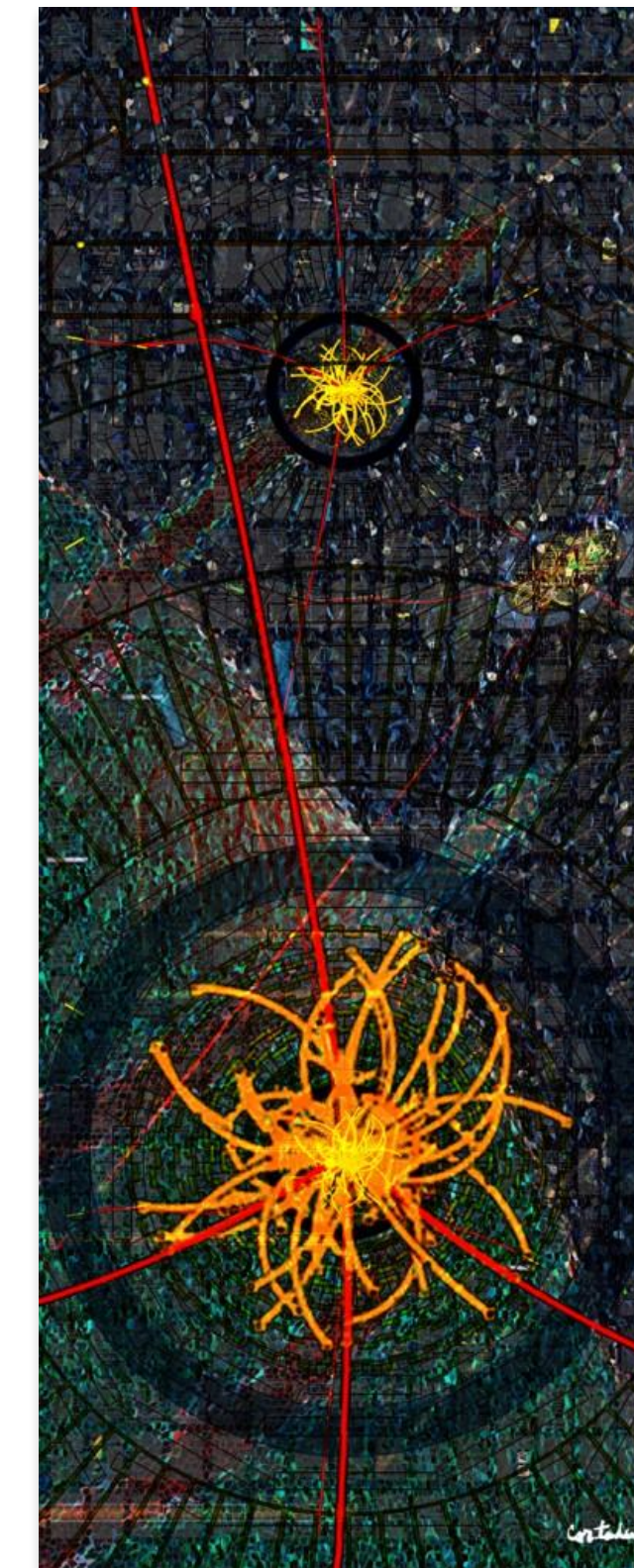
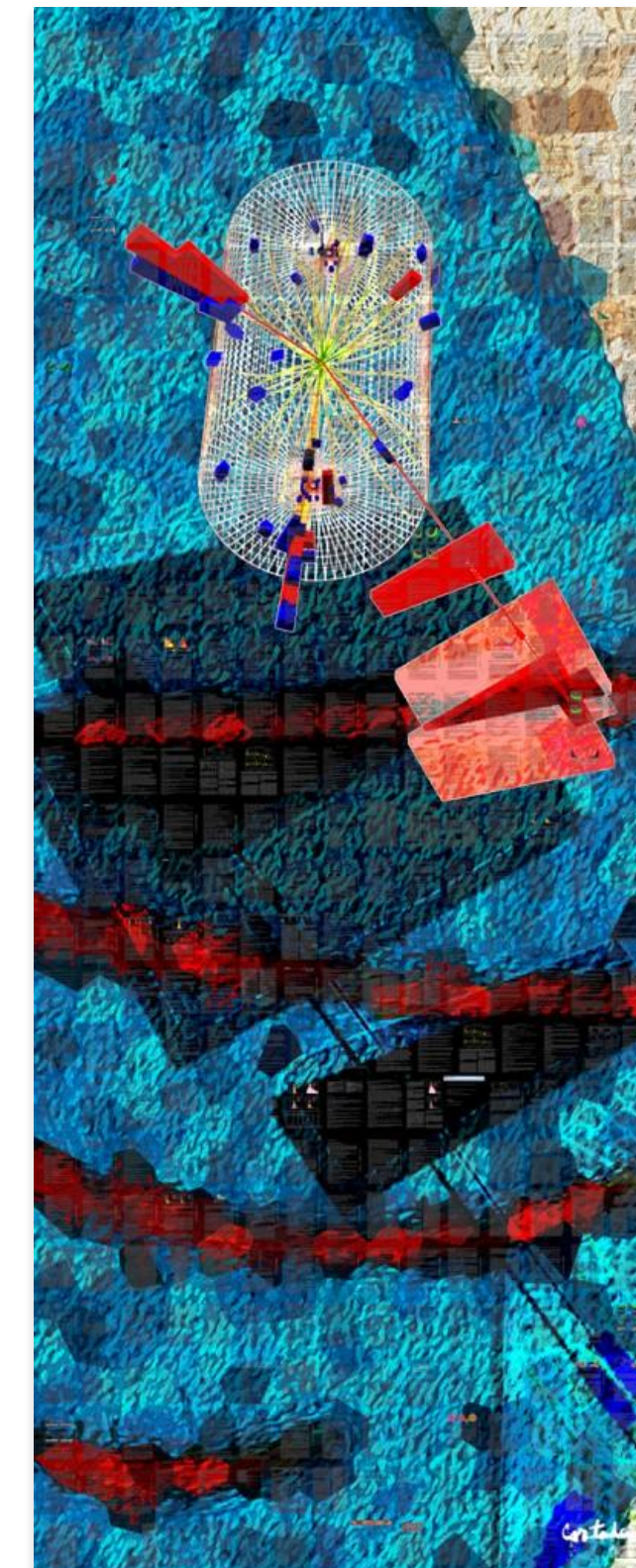
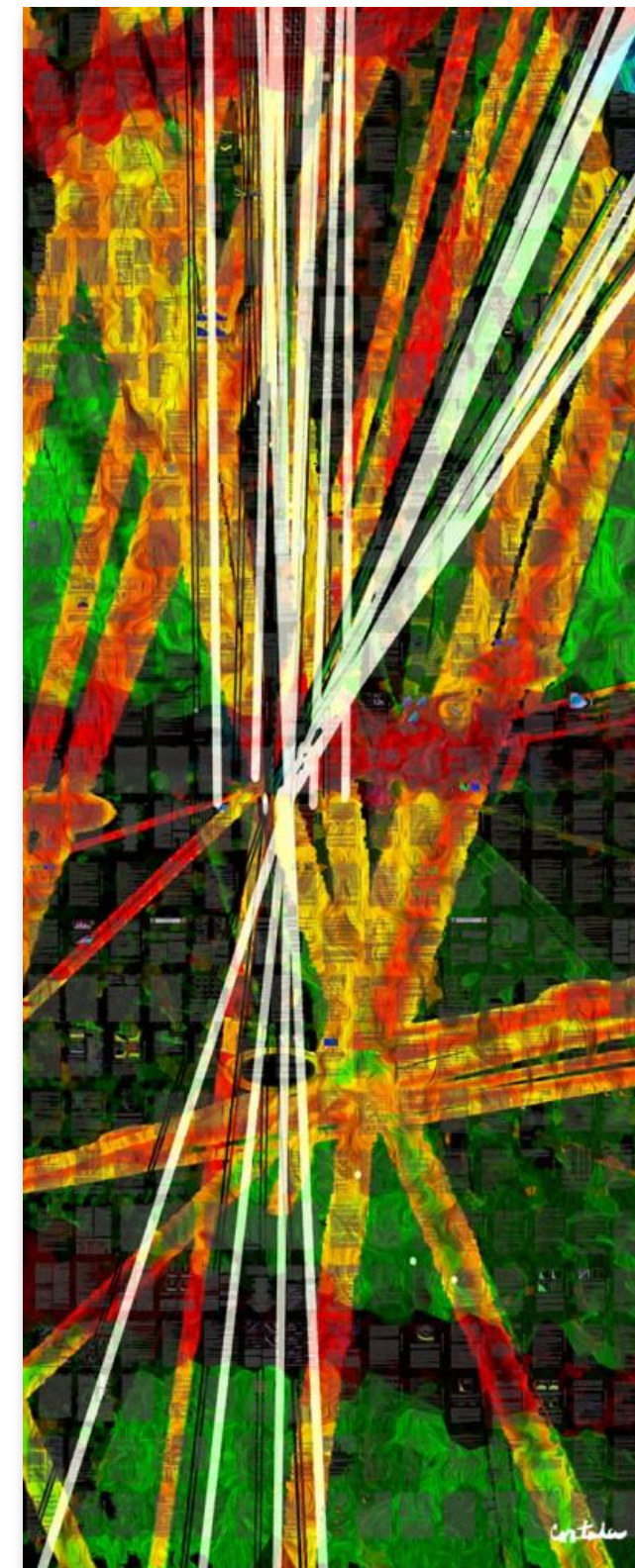
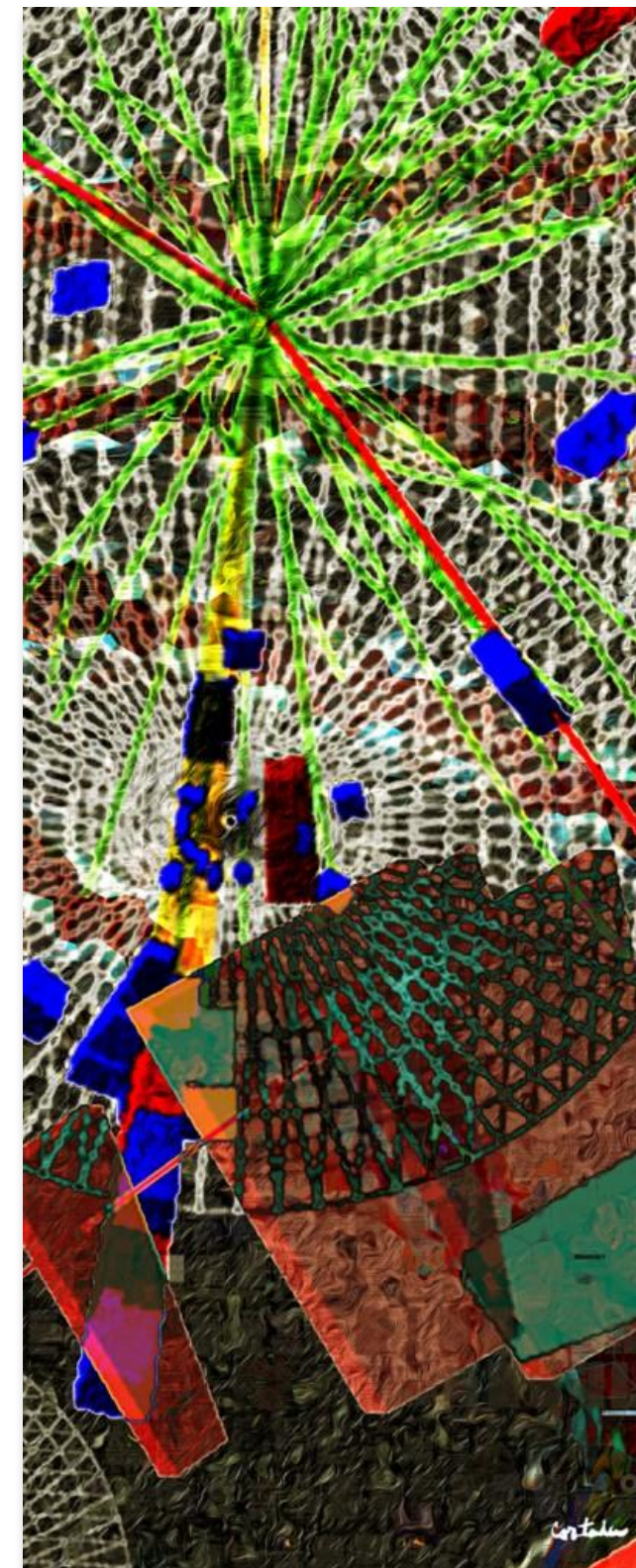
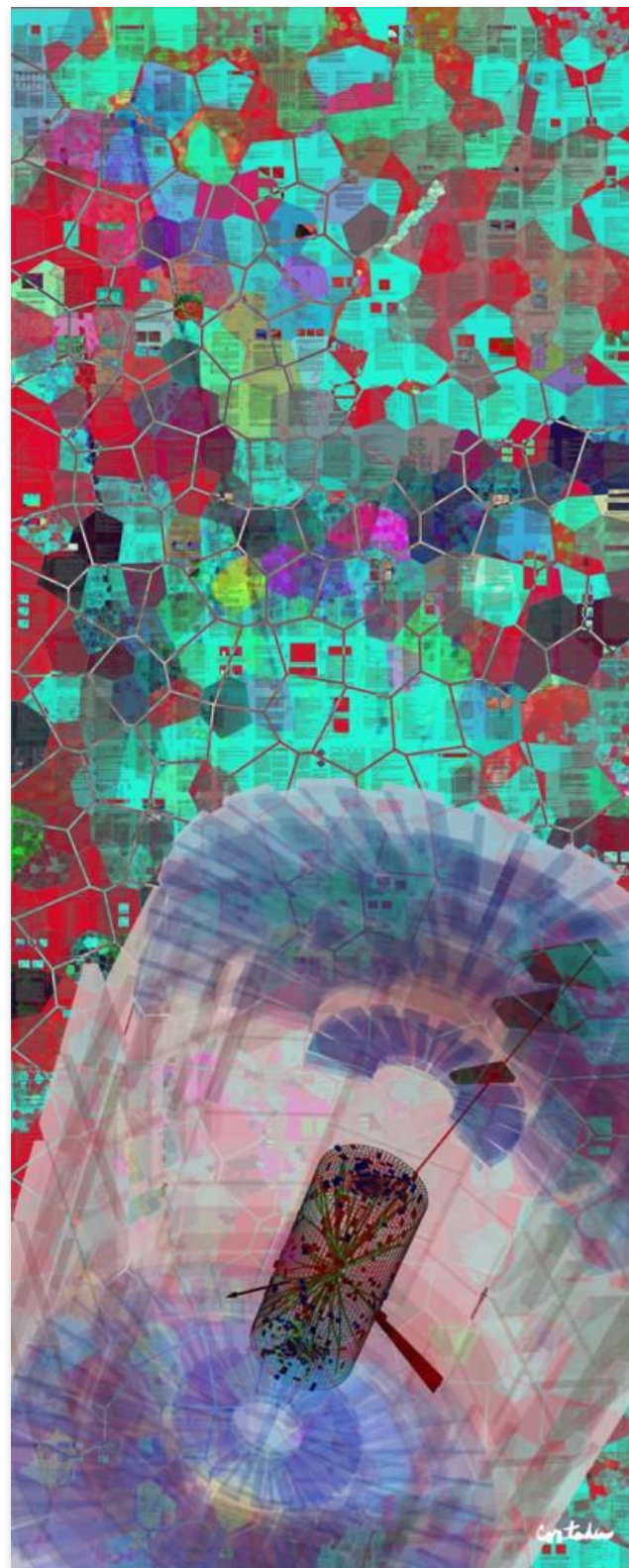


# U.S. CMS Operations Program Feedback

Oliver Gutsche  
IRIS-HEP Steering Board Meeting  
26. November 2019





## Disclaimer

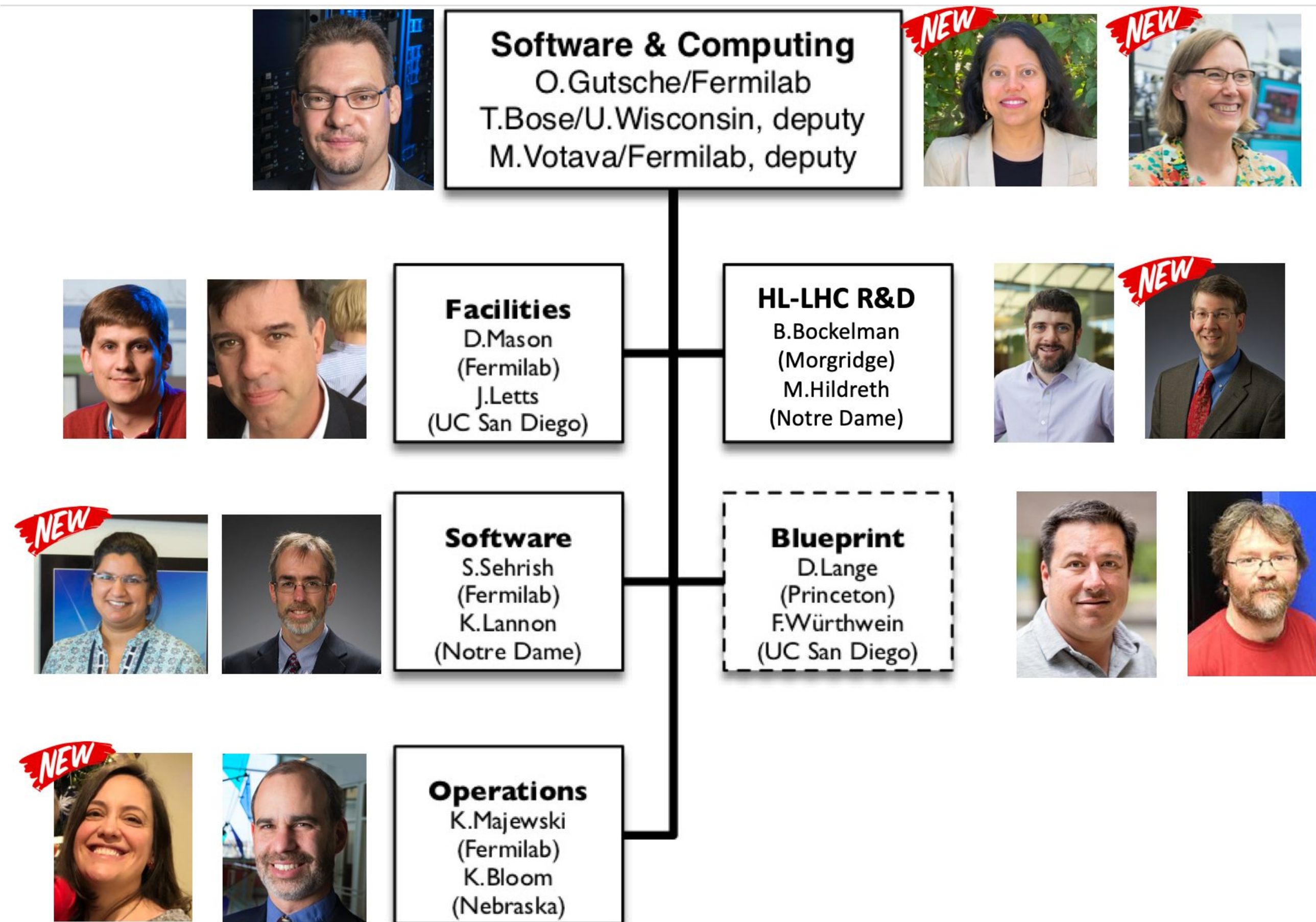
- I am speaking for the U.S. CMS S&C Operations Program
- I cannot speak for the whole U.S. CMS Collaboration
  - The S&C Operations Program has in the past enabled and driven new capabilities and guided the Collaboration
  - The S&C Operations Program will continue enabling the U.S. CMS Collaboration with the help of IRIS-HEP

## Outline

- Discuss the IRIS-HEP activities
  - Discuss how the U.S. CMS Operations program interacts with the IRIS-HEP activities
  - Give indications of input from the U.S. CMS Collaboration perspective
- Summarize with general remarks



# U.S. CMS S&C Operations Program







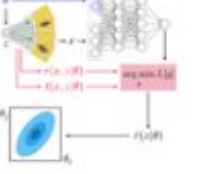











- Refreshed team

- Most relevant for IRIS-HEP: new HL-LHC R&D area, lead by Brian and Mike
- Asking Mike to be included in executive board meetings instead of Oli



## AS Projects

 <p><b>ADL Benchmarks</b></p> <p>Functionality benchmarks for analysis description languages <a href="#">More information</a></p>	 <p><b>AmpGen</b></p> <p>Generation and fitting for multibody hadron decays <a href="#">More information</a></p>	 <p><b>Awkward Array</b></p> <p>Manipulate arrays of complex data structures <a href="#">More information</a></p>	 <p><b>DecayLanguage</b></p> <p>Describe and convert particle decays <a href="#">More information</a></p>
 <p><b>Functional ADL</b></p> <p>Functional Analysis Description Language <a href="#">More information</a></p>	 <p><b>Histogram projects</b></p> <p>Histogramming efforts <a href="#">More information</a></p>	 <p><b>MadMiner</b></p> <p>Likelihood-free Inference <a href="#">More information</a></p>	 <p><b>Particle</b></p> <p>Pythonic particle information <a href="#">More information</a></p>
 <p><b>ROOT on Conda Forge</b></p> <p>Use ROOT in Conda through Conda-Forge <a href="#">More information</a></p>	 <p><b>Scikit-HEP</b></p> <p>pythonic analysis tools <a href="#">More information</a></p>	 <p><b>awesome-hep</b></p> <p>A curated list of awesome high energy and particle physics software <a href="#">More information</a></p>	 <p><b>exploratory-ml</b></p> <p>Analysis Reinterpretation <a href="#">More information</a></p>
 <p><b>PPX</b></p> <p><b>ppx</b></p> <p>cross-platform Probabilistic Programming eXecution protocol <a href="#">More information</a></p>	 <p><b>pyhf</b></p> <p>Differentiable Likelihoods <a href="#">More information</a></p>	 <p><b>recast</b></p> <p>Analysis Reinterpretation <a href="#">More information</a></p>	 <p><b>uproot</b></p> <p>Read and write ROOT files in Python <a href="#">More information</a></p>

## U.S. CMS Collaboration

- Analysis hot topic: Collaboration very interested in exploring alternatives to ROOT that are based on industry solutions
- Prominent example: Coffea
  - Using awkward array, uproot etc. → IRIS-HEP products

## U.S. CMS S&C Operations Program

- Facilitate development of analysis facilities
  - Using Coffea as first user-facing frontend
  - Backend: exploring different possibilities, interested in data delivery, parallelization (Spark, Dask, ... ) and modern deployment capabilities

## General observation

- Knowledge of IRIS-HEP solutions (other than what is used for Coffea prototypes) limited in U.S. CMS Collaboration
  - How can we better understand how IRIS-HEP solutions are being used and are useful for the U.S. CMS Collaboration?



# Blueprint Activity

- Concentrating on multi-day workshops
  - Generating a lot of momentum, very useful for the community
  - Disadvantage: need to travel, takes a lot of time

## Blueprint Dashboard

Topic / Title	Areas	Dates	Location	Host / Lead(s)	Summary Report
<b>Completed</b> (also see the top-level <a href="#">Indico page</a> )					
<a href="#">Analysis Systems R&amp;D on Scalable Platforms</a>	AS, SSL	June 21-22	NYU	K. Cranmer, R. Gardner	<a href="#">report.pdf</a>
<a href="#">Fast Machine Learning &amp; Inference</a>	IA,, SSL	Sept 10-13	FNAL	N. Tran, P. Harris	
<a href="#">A Coordinated Ecosystem for HL-LHC Computing R&amp;D</a>	All	Oct 23-25	CUA	A. Dominguez	
<b>Scheduled</b> (also see the top-level <a href="#">Indico page</a> )					
<b>2020</b>					
<b>2019</b>					
<b>Other Proposed Workshops</b>					
<b>2020</b>					
Software Training, Education & Workforce Development	SSC	Fall 2019?	TBD	P. Elmer, S. Malik	
Strengthening Connections Between Theory & Experiment	AS	?	NYU	K. Cranmer, M. Neubauer	
Analysis Systems & Software Ecosystem	AS	?	NCSA?	Neubauer?	
Intelligent & Accelerated Big Data Delivery	DOMA			C. Maltzahn, B. Bockelman	
Analysis Preservation & Open Access Data	SSC, AS			M. Hildreth	

## General remarks

- Would like to talk about overall R&D strategy for HL-LHC → synchronization between stakeholders
  - Workshops like the last CUA meeting are valuable
  - Regular coordination discussions over video could be added
- Would like to talk about evolution of IRIS-HEP
  - Blueprint could be the time/place where we come together and evolve the program of IRIS-HEP






## Request

- Adding video blueprint meetings



# Data Organization, Management and Access (DOMA)

## DOMA Projects

 <p><b>Caching Analysis Data</b> Cached-based placement of analysis datasets. <a href="#">More information</a></p>	 <p><b>Intelligent Data Delivery Service</b> Delivering Data. Better. <a href="#">More information</a></p>	 <p><b>Modeling Data Workflows</b> Modeling HL-LHC Data flows <a href="#">More information</a></p>	 <p><b>ServiceX</b> Delivering columnar data on demand <a href="#">More information</a></p>
 <p><b>Third Party Copy</b> Envisioning a new way to move LHC data <a href="#">More information</a></p>			

## U.S. CMS S&C Operations Program

- Facilitate data lake concepts: bulk movement (Rucio), caching (XCache), close coordination with workflow management and analysis data delivery
- A lot of activity in IRIS-HEP, OSG and WLCG context, operations program closely following and participating

## U.S. CMS Collaboration

- DOMA efforts not sufficiently visible in U.S. CMS Collaboration (the same for international CMS Collaboration)

## Request



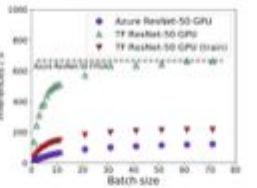
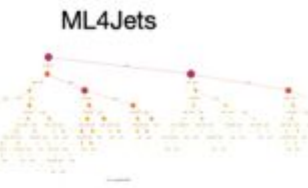
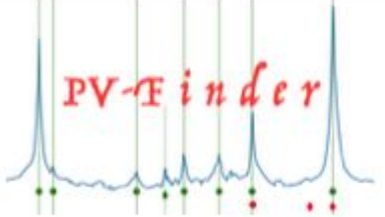

- Promote IRIS-HEP DOMA activities in U.S. CMS and CMS Collaborations and increase awareness of developments going on



## U.S. CMS S&C Operations Program

- Facilitate reconstruction and ML exploiting hardware capabilities (CPU, GPU, FPGA (vectorization, data structures, etc.) to reduce the needs of reconstruction in HL-LHC
  - Framework is enabling offloading, looking into portability solutions and I/O and EDM with CCE

### IA Projects

 <p><b>ACTS</b></p> <p>Development of experiment-independent, inherently parallel track reconstruction.</p> <p><a href="#">More information</a></p>	 <p><b>FastPID</b></p> <p>Fast PID simulation for LHCb</p> <p><a href="#">More information</a></p>	 <p><b>ML on FPGAs</b></p> <p>Fast inference of deep neural networks on FPGAs</p> <p><a href="#">More information</a></p>	 <p><b>ML4Jets</b></p> <p>Machine learning for jets</p> <p><a href="#">More information</a></p>
 <p><b>PV-Finder</b></p> <p>CNNs to find primary vertices</p> <p><a href="#">More information</a></p>	 <p><b>exploratory-ml</b></p> <p>Analysis Reinterpretation</p> <p><a href="#">More information</a></p>		

## U.S. CMS Collaboration

- mkFit efforts well known to Collaboration (same for international CMS)
- ML on FPGAs also discussed and known to Collaboration
- Other IA projects less known: their capabilities are not explored/used a lot in the U.S. CMS Collaboration




## Request

- Promote IA activities other than mkFit and ML on FPGAs in U.S. CMS Collaboration and increase awareness of new capabilities



# Open Science Grid (OSG-LHC)

## OSGLHC Projects

 <p><b>Caching Analysis Data</b> Cached-based placement of analysis datasets. <a href="#">More information</a></p>	 <p><b>Modeling Data Workflows</b> Modeling HL-LHC Data flows <a href="#">More information</a></p>	 <p><b>Third Party Copy</b> Envisioning a new way to move LHC data <a href="#">More information</a></p>
--	---	---

## U.S. CMS S&C Operations Program

- Facilitates the computing infrastructure for the U.S. CMS Collaboration: OSG-LHC is integral part of the operation of the infrastructure
- Rely on OSG-LHC to evolve the computing infrastructure together with the community at large, and integrate and deploy new developments from other IRIS-HEP areas
  - Success story: OSG-LHC provides XCache installation software

## U.S. CMS Collaboration

Many successes of OSG-LHC are not known and are not credited to OSG-LHC

- Maybe because they are considered “normal” operation and therefore taken as granted?
- Example: many networking activities listed on the OSG-LHC page might not be known well in the Collaboration

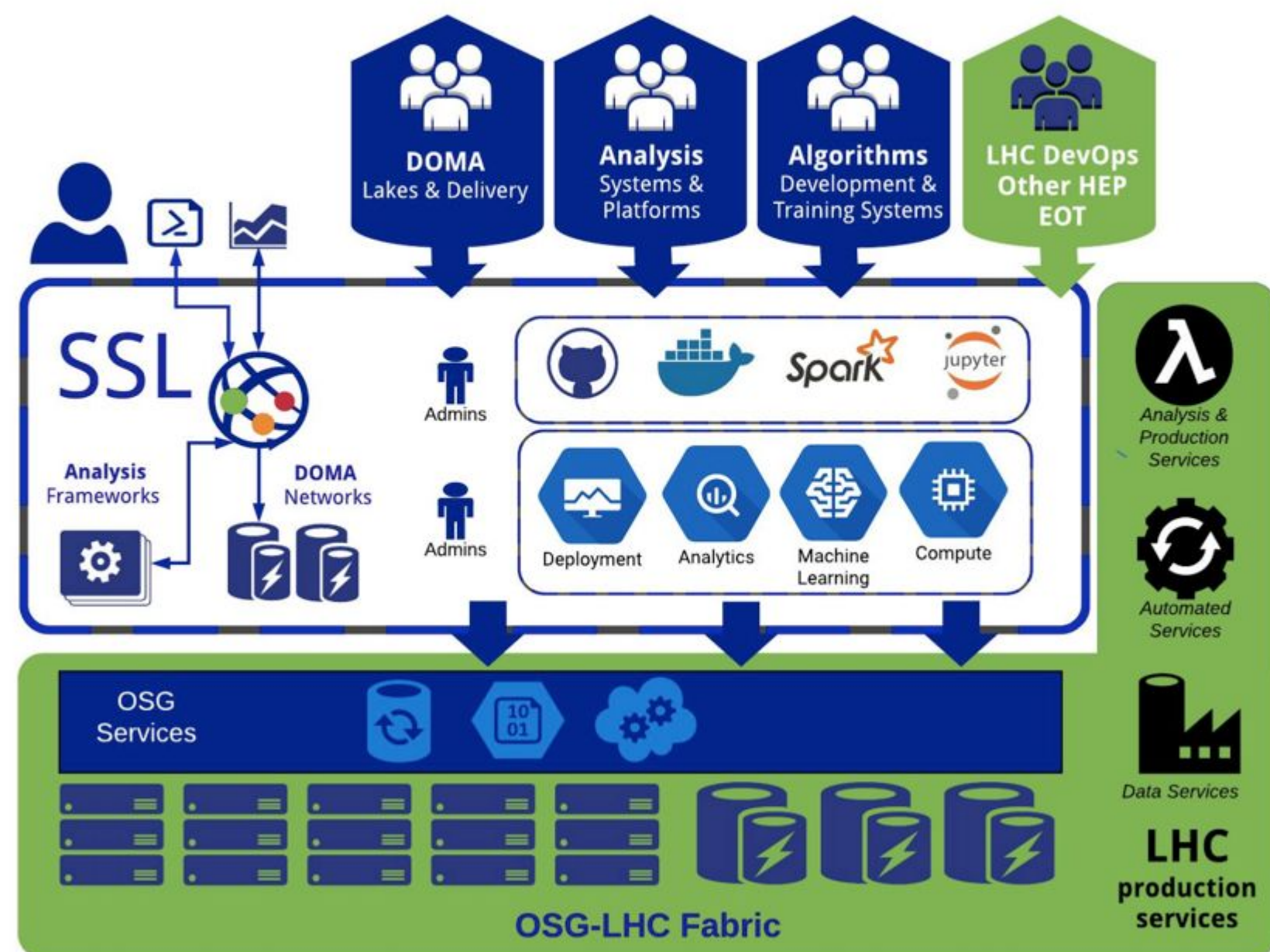
## Request

- Promote all activities of OSG-LHC and raise awareness of importance and successes of OSG-LHC in the U.S. CMS Collaboration



## Scalable Systems Laboratory

Together with the OSG-LHC, the Scalable Systems Laboratory (SSL) is designed to be the primary integration path to deliver the output of IRIS-HEP R&D activities into the distributed and scientific production infrastructure of the experiments.



## U.S. CMS S&C Operations Program

- Facilitate deployment of new infrastructure and interest in flexible deployment capabilities
  - Starting to work with U.S. CMS T2 sites on Xcache deployments and analysis facility deployments

## U.S. CMS Collaboration

- Concepts and Capabilities of SSL not well known in U.S. CMS Collaboration (same for international CMS Collaboration) beyond what is discussed in WLCG/OSG context
  - Well publicized in ATLAS (see listed talks)

## Request

- Find new ways to promote SSL capabilities in U.S. CMS Collaboration



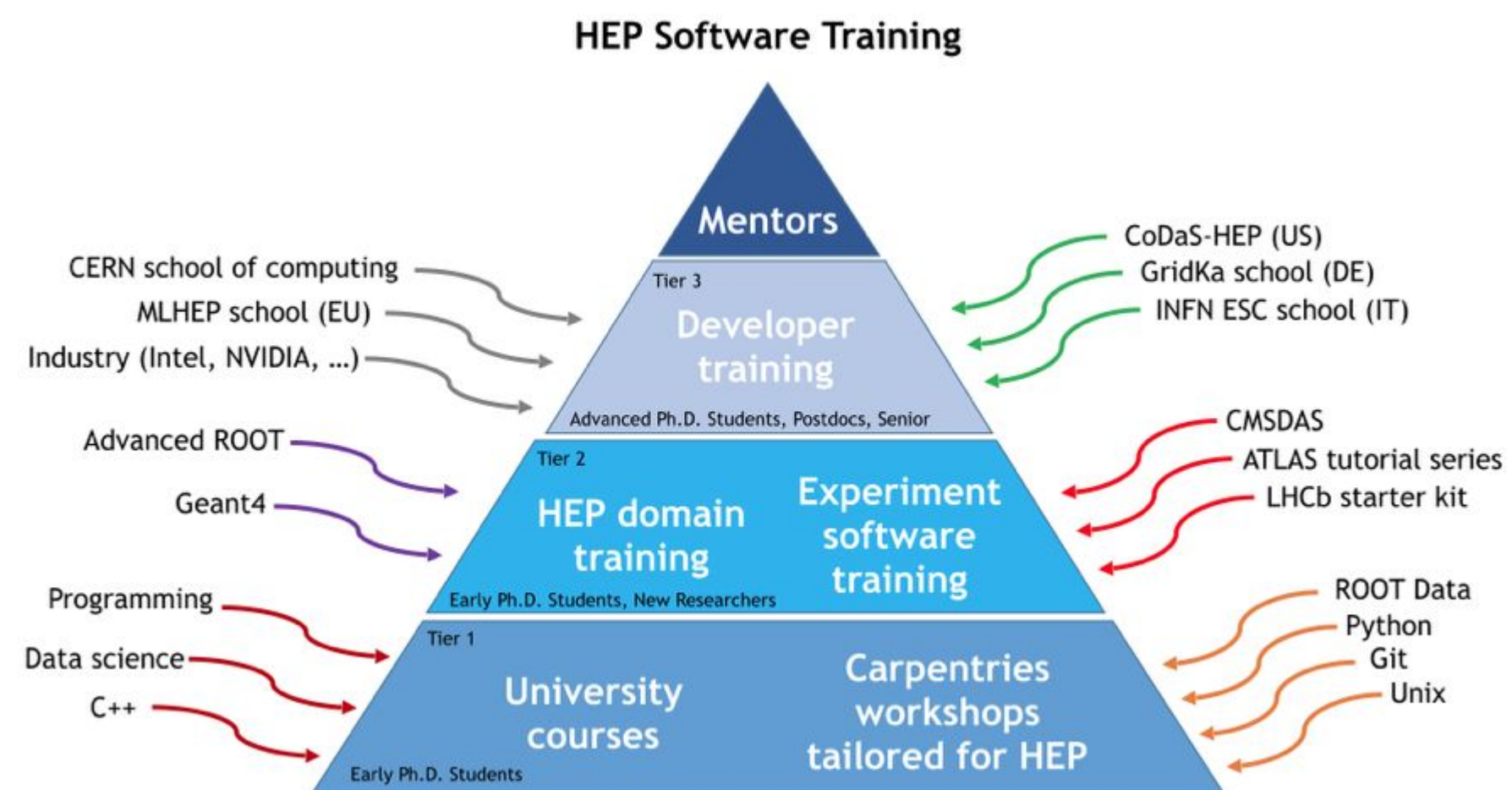
# Training, Education and Outreach

## U.S. CMS S&C Operations Program

- Does not have a major mandate to facilitate Training and Education
- Outreach was not a high priority for the S&C Operations Program so far

## U.S. CMS Collaboration

- Benefits a lot from Education and Training activities, from CMSDAS to CoDaS-HEP schools



## Suggestions

- Programming training aimed at advanced/heterogeneous architectures are in high demand, increase in possibilities would be well received by Collaboration
- Co-location of more training programs with U.S. CMS Collaboration events (like Collaboration meeting, etc.)
- Integration of more advanced training content from IRIS-HEP into CMSDAS and similar activities would make it more impactful for the U.S. CMS Collaboration



- U.S. CMS S&C Operations Program facilitates many new capabilities for the U.S. CMS Collaboration with the help of IRIS-HEP
  - Integration and communication is working
  - Would like to have more video blueprint meetings, especially talking about HL-LHC R&D strategy and evolution of IRIS-HEP
    - Need to slowly start talking about sustainability of IRIS-HEP products/solution and how we can maintain them in the future  
→ longterm topic
- U.S. CMS Collaboration benefits from many IRIS-HEP products/solutions
  - There are though cases where the U.S. CMS Collaboration (same for international CMS) is not aware of IRIS-HEP activities
  - Increasing promotion and information of the U.S. CMS Collaboration would help making the Collaboration aware of opportunities (maybe also to get involved and test/use new products/solutions)
    - IRIS-HEP could let the U.S. CMS Collaboration know where opportunities and needs are to get involved (or give examples)