

# ROOT 2021

---

Axel Naumann, CERN  
2021-01-06, ROOT Godparents



# Outline

---

- 2020 Achievements
- Team Situation
- 2021's Theme
- 2021 Main Work Areas



# Main 2020 Achievements

---

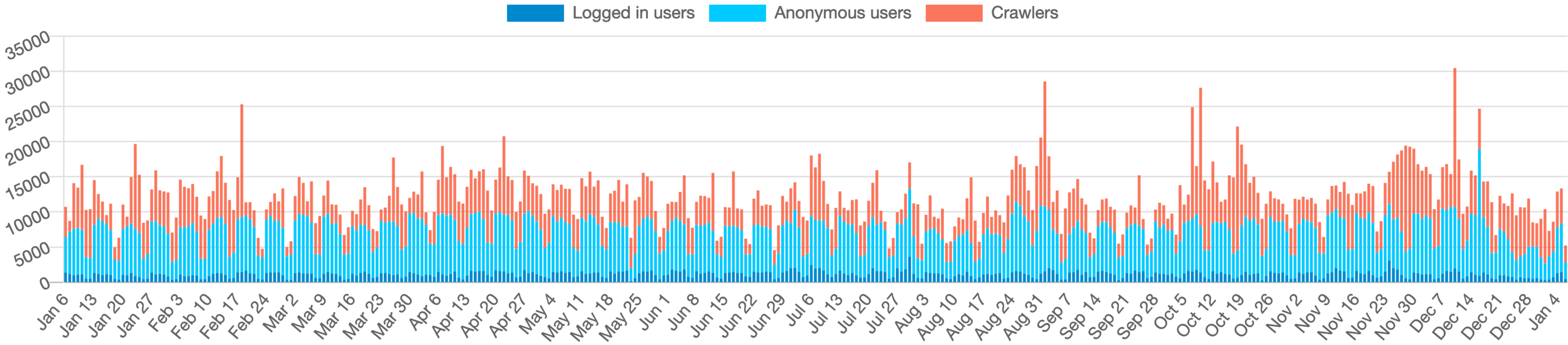
- Presentations at ICHEP, ATLAS, CMS, LHCb, ALICE, EIC, CEPC,....: ROOT and our progress remains relevant
- Major progress with RNTuple, proving relevance and validity of the approach
- WebGUI now "minimum viable product"
- Analysis interface RDataFrame now battle tested: number of forum posts (tags) as high as for tree, hist
- Upgraded PyROOT, significantly improved RooFit, RNNs, REve, plus much more!



# 2020 in Numbers: Forum Visitors

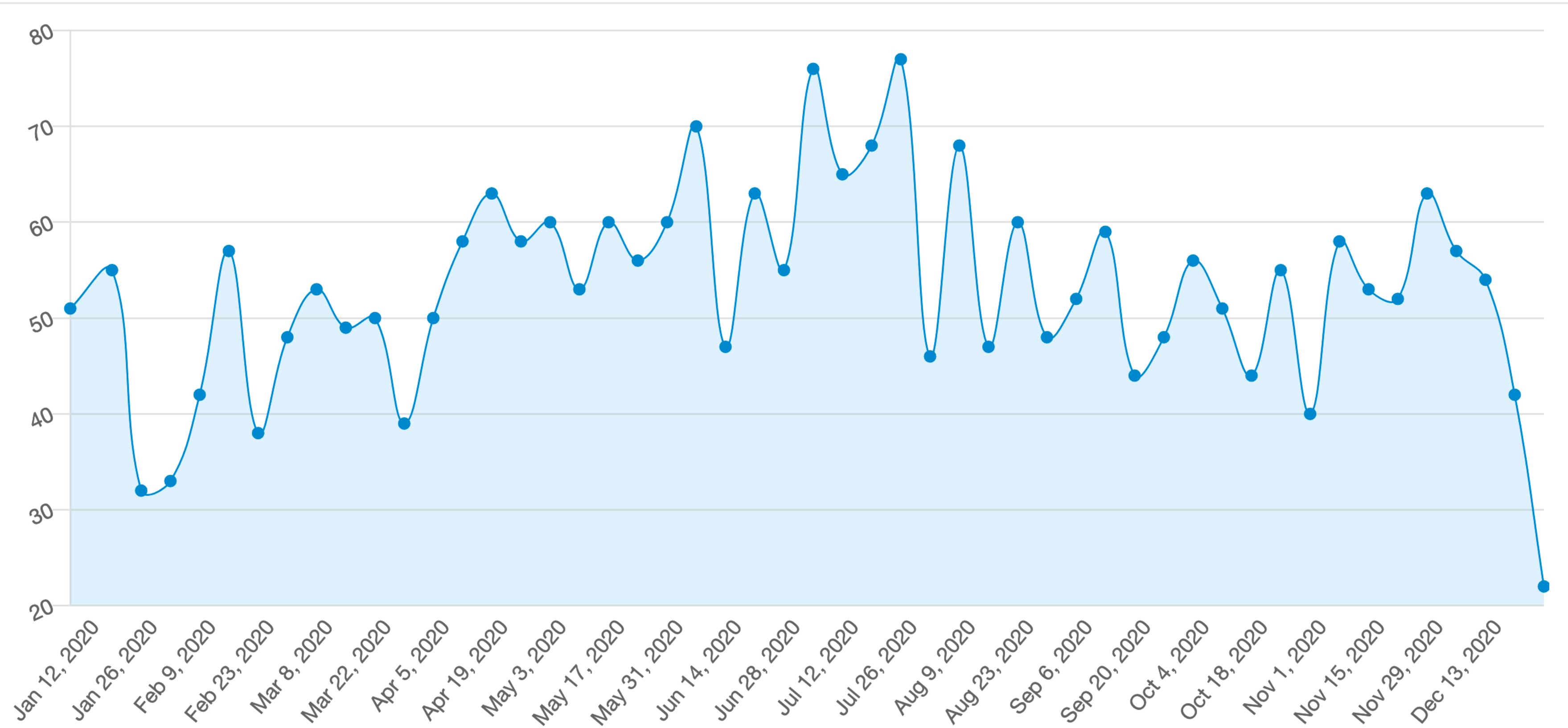
- Number of human visitors: about 8000/day!

## Consolidated Pageviews ?



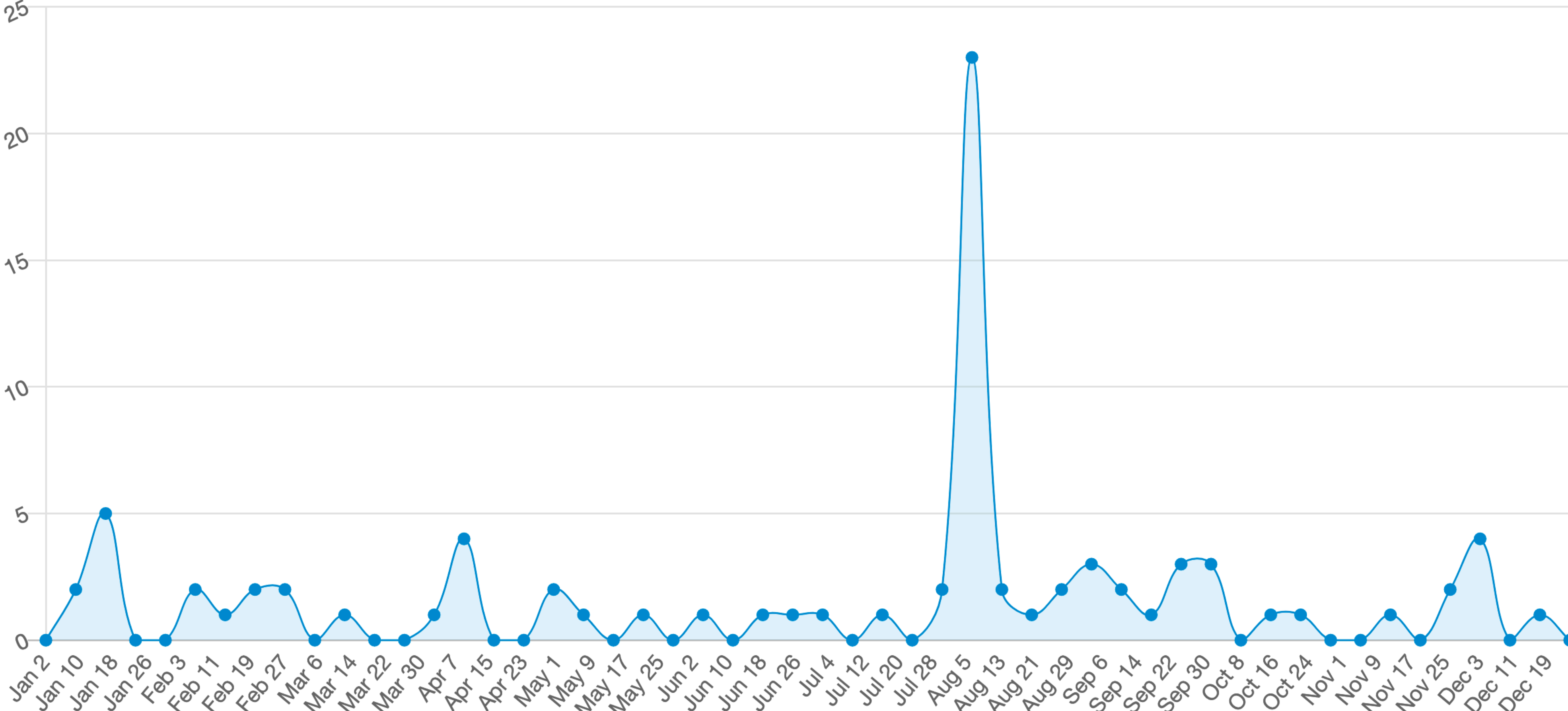
# 2020 in Numbers: Forum Topics

- About 50 topics (questions) / week!



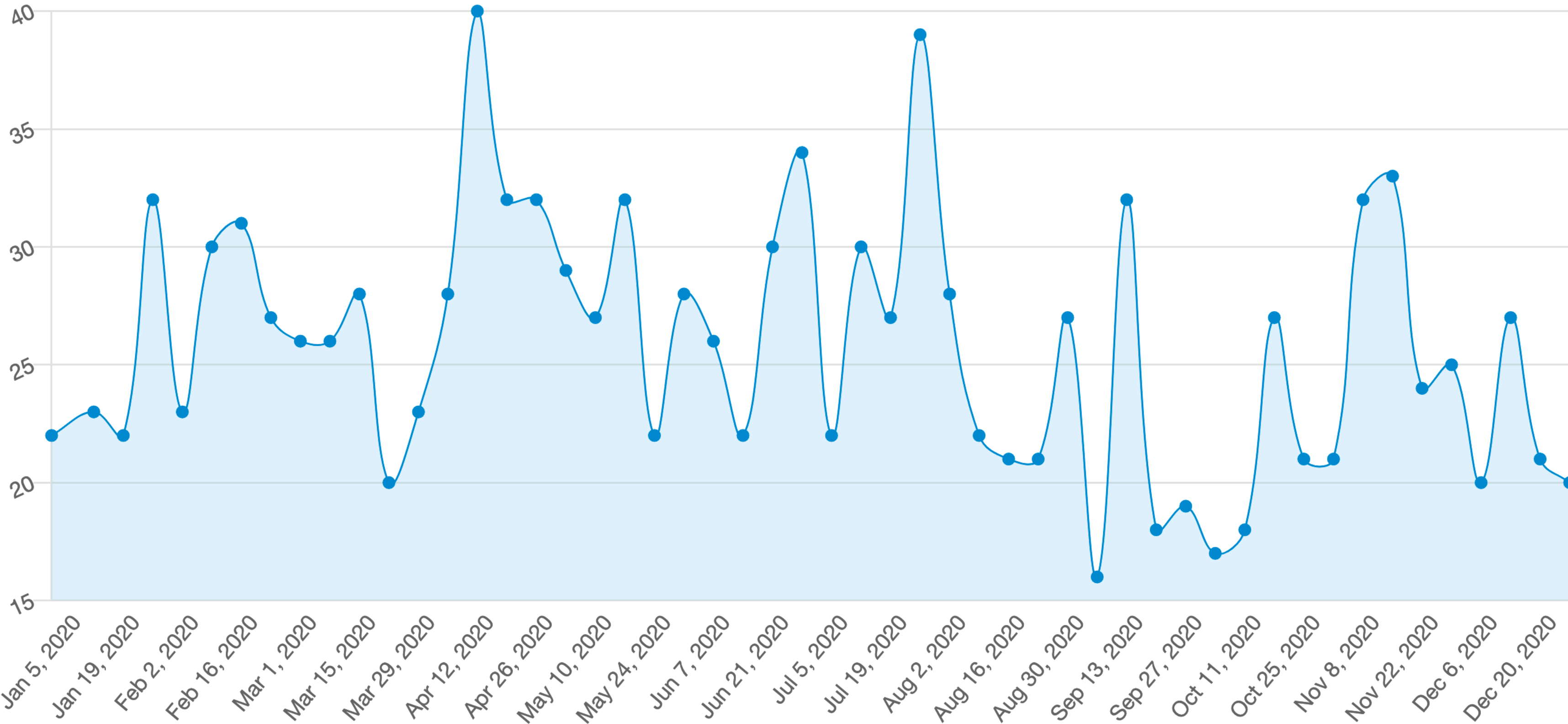
# 2020 in Numbers: Reply-Rate

- $<5$  / week. Includes announcements etc.



# 2020 in Numbers: New Forum Users

- Consistently > 20 new users / week



# 2020 in Numbers: Code and releases

---

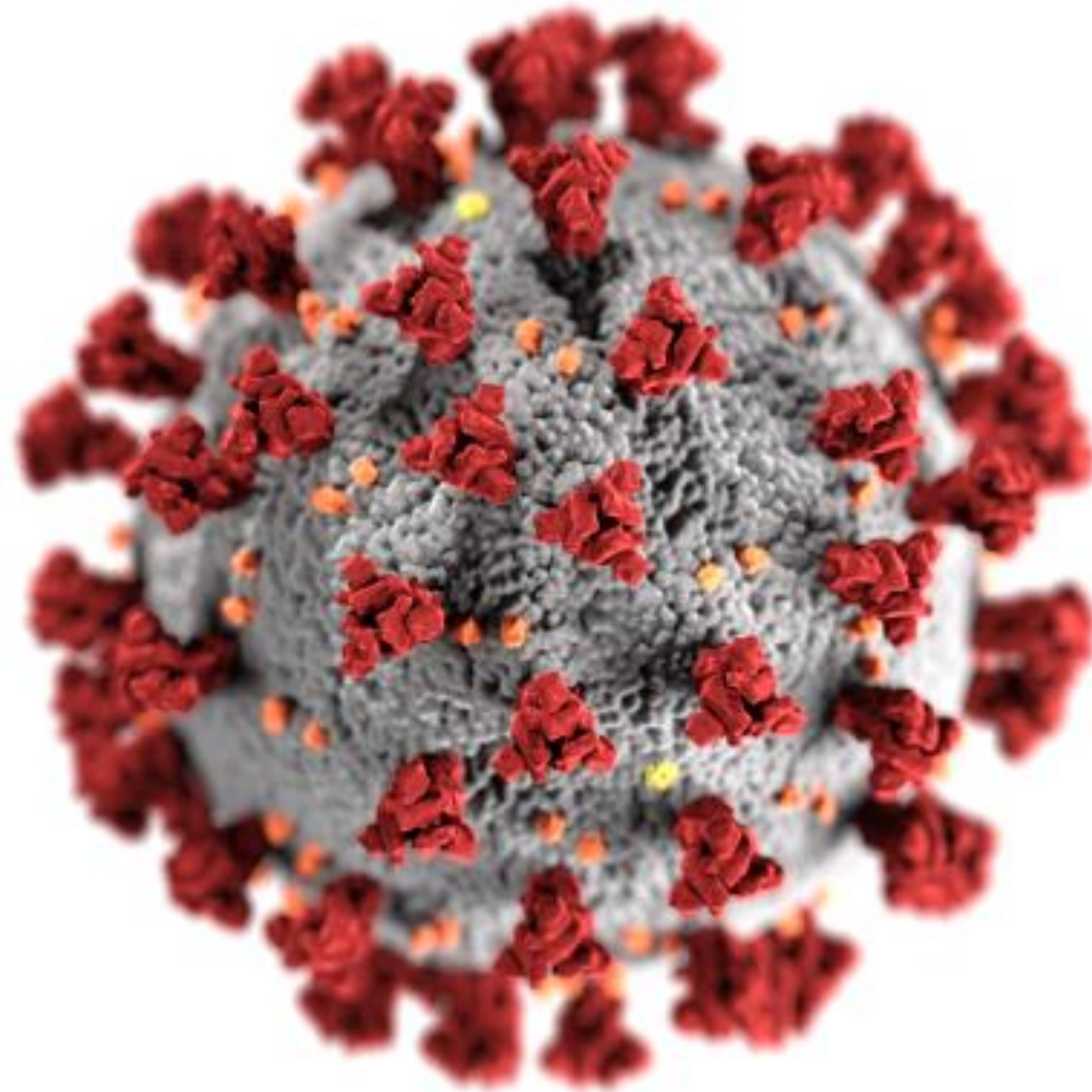
- Released v6.20 in February, v6.22 in June
  - Always supporting newest compilers and distro versions
- 2020's diff is 32MB, with about 750'000 lines
- From about 100 contributors





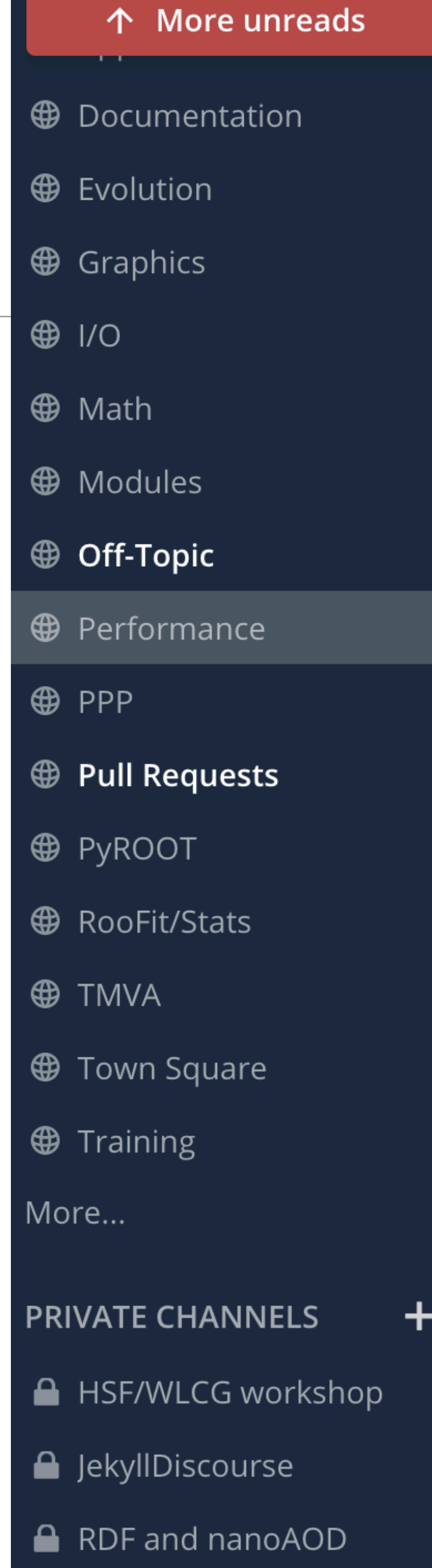
All of that despite of...

---



# Communication, communication, communication

- Mattermost for low-latency communication
- Hundreds of messages per day, across the whole team
- Video conference meetings
  - Daily stand-up meetings (20 mins)
  - Once per week topical meetings: I/O, PPP, RooFit, TMVA,...
  - Weekly team meetings



# Communication, communication, communication

- VC on a special day with a special guest



# Main 2020 Achievements, behind the scenes

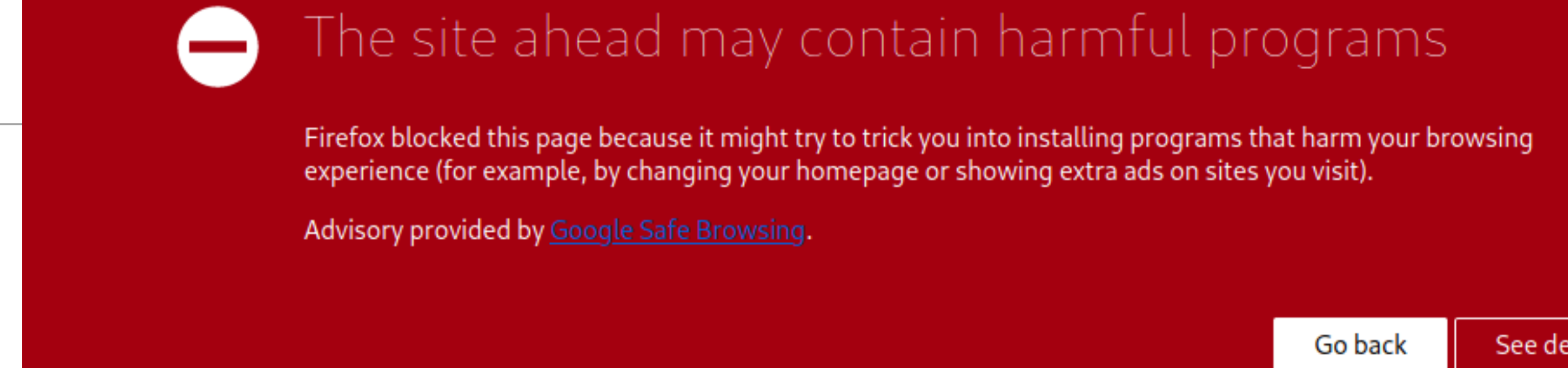
---

- New website
  - Got rid of complex and fragile Drupal system
- GitHub Issues
  - Moved away from slow, uncooperative Jira
- Moved to new root.cern server
- Moved to new Jenkins (continuous integration) server



# The Google "Attack"

- End October, suffered from de-facto DOS attack by Google
- Claimed to anyone visiting <https://root.cern> that it was distributing malicious software
- Caused by virus scanner false positives
- Immediate work-around, but work still ongoing for more sustainable solution
- See <https://root.cern/blog/false-positives/> for some background



# Team Situation: Long-Term Contributors

---

- Continuity provided by team members with indefinite contracts:
  - Bertrand, Jakob (50%, IC since 2020!), Lorenzo, Olivier, Philippe (50%), AN
  - Hope to get Danilo back in 2021
- And long-term key contributors:
  - Enric, Oksana (50%), share of Vassil



# Team Situation

---

- Fellow Stephan Hageböck (RooFit) moved to IT; successor Jonas Rembser to arrive in February
- Massimiliano Galli (student) left in May; DOCT contract for Stefan Wunsch (RDF, TMVA, Python) will expire soon; IRIS-HEP fellow Max Orok
- Ineffective summer student program, mostly effective GSoC contributions
- Enrico Guiraud now fellow; new CERN EP R&D fellow Javier Lopez (a clang expert!), working on I/O; 50% of fellow Jonas Hahnfeld, shared with simulation



2021 Themes



# Not This

---

Math

I/O

Graphics

Analysis

RooFit

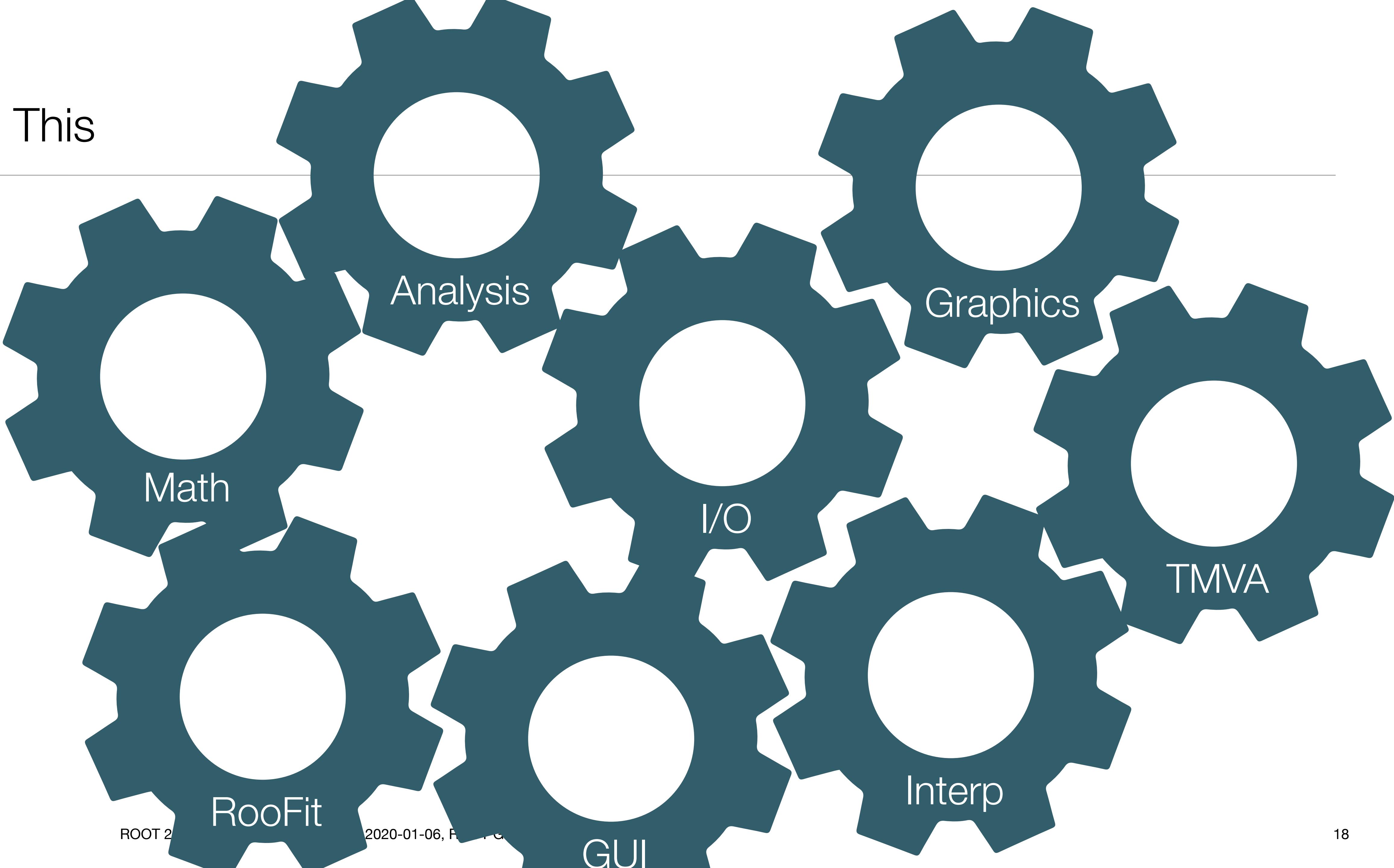
TMVA

Interpreters

GUI



But This



# Meta-Goal: Symbiotic Evolution. Current Examples

---

- JSROOT + WebGUI and REve
- RNTuple + RDataFrame
- PyROOT + TTree = dataframe.AsNumpy()
- Cling + TMVA for fast BDT inference



# Meta-Goal: Symbiotic Evolution. Future Examples

---

- Mechanism used for RooFit AVX2 etc compute libs to be re-used for compression?
- TMVA-GUI based on WebGUI? RNTuple in WebGUI's RBrowser?
- PyRDF combining RDataFrame, new PyROOT, RNTuple caching?
- Clad (cling-based automatic differentiation) employed by RooFit?
- Cling + TMVA for fast ONNX / DNN inference



# 2021 Main Focus Points

# Analysis

---

- Fast-path from RNTuple to analysis results:  
million events / second into histograms, for actual analyses
- With systematics variation
- With convenient, multi-threaded, distributed compute model:  
RDataFrame + PyRDF



# RNTuple

---

- Exposure of RNTuple: integration of RNTuple into CMSSW to test-drive nanoAODs, PODIO
- RNTuple features: conversion from TTree, friends + chains, fast merging, MT-writing, finalize user data-type changes ("schema evolution")
- Benchmarking: into ML tools, vs HDF5, caching, object store



# WebGUI + Graphics

---

- `root --web -e 'TH1F h; h.Fill(0.3); h.Draw()'`
- TBrower to become RBrowser until the release, for wider exposure
- Review of "web-graphics programming model", C++ and Python





# Interpreters / JIT

---

- Cling
  - Finish ongoing llvm 9 upgrade
  - Followed by llvm 11 upgrade: requires C++14!
  - Debug symbols for fitted code, performance review
- PyROOT
  - API for user pythonization of their C++ classes



# Analysis

---

- Generator interface for ML training out of RDataFrame
- Support for parameter variation (systematics) within single event loop
- Support for categories, such as multiple data vs MC samples and their weights
- Bulk-processing multiple events at once in RDF
- PyRDF inclusion in ROOT
  - Workshop to review and define feature set



# Framework-Oriented

---

- Speedup of compression algorithms
- Lossy compression, such as Accelogic, ZFP
- TBufferFile > 1GB
- Speed-up of concurrent writing into TBufferFile



# ROOT Workshop Alternatives

---

- Decided against virtual ROOT workshop
  - Low-barrier (social) contact between users and devs is a prerequisite
- Instead: topical events
- Ongoing discussion on how to improve user feedback during COVID



“The LHCC ... encourages CERN to ensure continuing core support for ROOT ...”

– CERN/LHCC-2020-003, LHCC-141, February 2020

# Thank You!

---

- Only possible with **your** commitment to ROOT!
  - Our bottleneck remains the lack of hands.
- Aim: prove that ROOT deserves your investment
  - Tweaking state-of-the-art technology to our very own HEP challenges
  - Making it production-grade and sustainable

