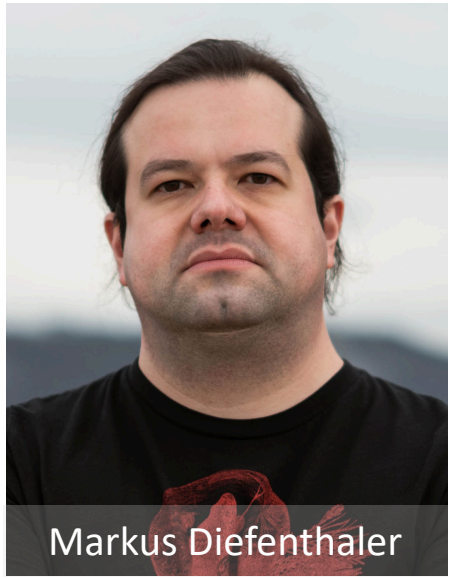
Update from EICUG Software Working Group

EICUG Software Working Group: Conveners


Andrea Bressan


Cristiano Fanelli


Markus Diefenthaler


Tanja Horn


Torre Wenaus

Liaison to EICUG SC

Many thanks to EICUG SC for support.

Mailing list (135 members): eicug-software@eicug.org

Meetings (16 meetings in 2022): https://indico.bnl.gov/category/301/

EIC Software Efforts

2016 – 2020 EIC Software Consortium (ESC)

2018 – now EICUG Software Working Group (SWG)

2019 – 2021 Yellow Report Initiative

2021 – 2022 Detector Collaboration Proposals

2022 – now EPIC Computing and Software (CompSW) Working Group
EPIC Simulation, Production, and QA (SimQA) Working Group

EPIC CompSW and EICUG SWG

EPIC CompSW

Sylvester's Picture of Salt & Pepper

EICUG SWG

- EPIC Collaboration and EIC project have strict timeline with well-defined deliverables from short to long term.
- This defines the scope and deliverables of EPIC CompSW.

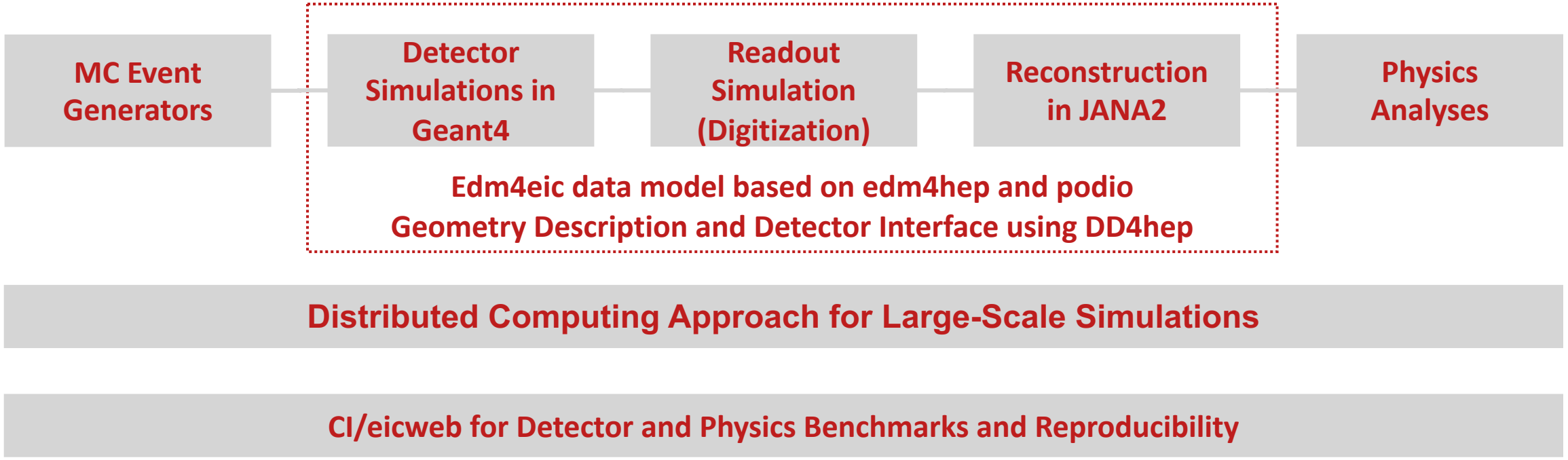
- EICUG fosters the community.
- EICUG SWG is a forum for discussion and forward-looking projects.
- Ideal platform for cross-collaboration:
  - Experient and theory
  - EPIC and 2nd Detector
  - Interdisciplinary collaborations (data scientists, HEP, …)

EPIC Software

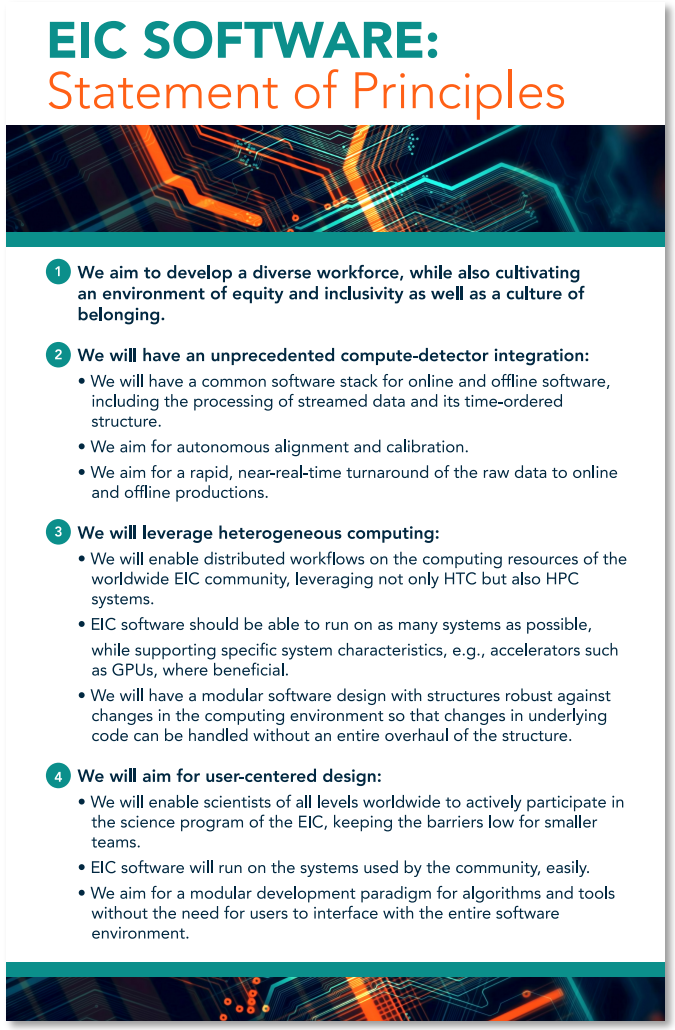
Results from the “State of Software” surveys: Commonality; one software stack.

The design of the **modular simulation and reconstruction toolkit** for the development of the EPIC detector and the EPIC science program is based on the **EIC Software: Statement of Principles** and a **decision-making process involving the wider EIC community**.

Simulation



EIC Software Vision and Guiding Principles for EIC Software R&D



EIC SOFTWARE:
Statement of Principles

1 We aim to develop a diverse workforce, while also cultivating an environment of equity and inclusivity as well as a culture of belonging.

2 We will have an unprecedented compute-detector integration:

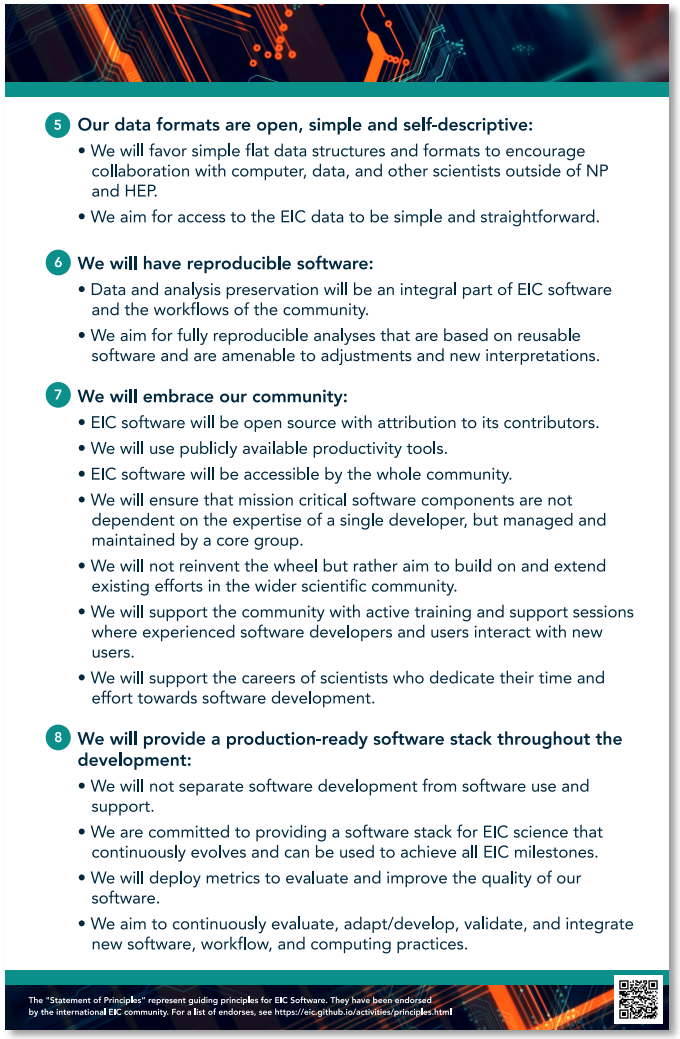
- We will have a common software stack for online and offline software, including the processing of streamed data and its time-ordered structure.
- We aim for autonomous alignment and calibration.
- We aim for a rapid, near-real-time turnaround of the raw data to online and offline productions.

3 We will leverage heterogeneous computing:

- We will enable distributed workflows on the computing resources of the worldwide EIC community, leveraging not only HTC but also HPC systems.
- EIC software should be able to run on as many systems as possible, while supporting specific system characteristics, e.g., accelerators such as GPUs, where beneficial.
- We will have a modular software design with structures robust against changes in the computing environment so that changes in underlying code can be handled without an entire overhaul of the structure.

4 We will aim for user-centered design:

- We will enable scientists of all levels worldwide to actively participate in the science program of the EIC, keeping the barriers low for smaller teams.
- EIC software will run on the systems used by the community, easily.
- We aim for a modular development paradigm for algorithms and tools without the need for users to interface with the entire software environment.



5 Our data formats are open, simple and self-descriptive:

- We will favor simple flat data structures and formats to encourage collaboration with computer, data, and other scientists outside of NP and HEP.
- We aim for access to the EIC data to be simple and straightforward.

6 We will have reproducible software:

- Data and analysis preservation will be an integral part of EIC software and the workflows of the community.
- We aim for fully reproducible analyses that are based on reusable software and are amenable to adjustments and new interpretations.

7 We will embrace our community:

- EIC software will be open source with attribution to its contributors.
- We will use publicly available productivity tools.
- EIC software will be accessible by the whole community.
- We will ensure that mission critical software components are not dependent on the expertise of a single developer, but managed and maintained by a core group.
- We will not reinvent the wheel but rather aim to build on and extend existing efforts in the wider scientific community.
- We will support the community with active training and support sessions where experienced software developers and users interact with new users.
- We will support the careers of scientists who dedicate their time and effort towards software development.

8 We will provide a production-ready software stack throughout the development:

- We will not separate software development from software use and support.
- We are committed to providing a software stack for EIC science that continuously evolves and can be used to achieve all EIC milestones.
- We will deploy metrics to evaluate and improve the quality of our software.
- We aim to continuously evaluate, adapt/develop, validate, and integrate new software, workflow, and computing practices.

The "Statement of Principles" represent guiding principles for EIC Software. They have been endorsed by the international EIC community. For a list of endorses, see https://eic.github.io/activities/principles.html

Webpage: https://eic.github.io/activities/principles.html

Focus on Common Tools

- Software in very early life stage:

> 8 **We will provide a production-ready software stack throughout the development:**
> - We will not separate software development from software use and support.
> - We are committed to providing a software stack for EIC science that continuously evolves and can be used to achieve all EIC milestones.
> - We will deploy metrics to evaluate and improve the quality of our software.
> - We aim to continuously evaluate, adapt/develop, validate, and integrate new software, workflow, and computing practices.

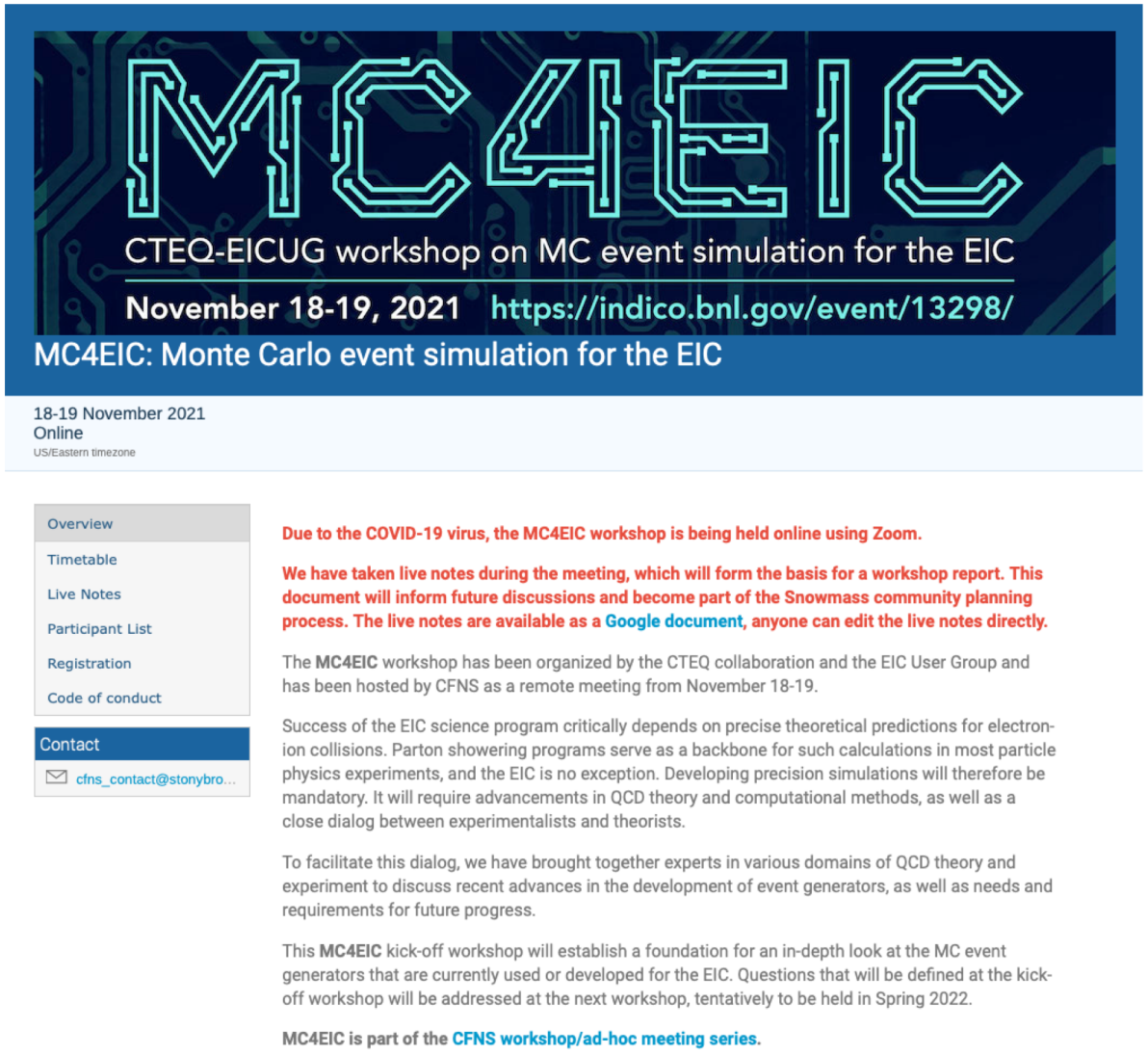
- **Focus on common software tools**
  - Avoid duplication of efforts.
  - Team up on challenges, e.g., heterogeneous computing.

- **Engage with the wider NHEP community**
  - **AIWG**: 5 meetings and AI4EIC workshop
  - **Detector collaborations**
    - EPIC and 2nd detector efforts
  - Geant4 collaboration
  - HEP Software Foundation
  - **Key4Hep**: Meeting with CERN and DESY in Sept.
  - **Theory community**
    - MC event generators (MCEGs), next slides

Start building a MCEG Community for the EIC

Satellite workshop during POETIC 8

MC4EIC in 2021



Goals of MC4EIC

- Establish a foundation for in-depth look at event generators currently used or developed for the EIC.
- Understand precision level to be satisfied by event generators in order to match experimental analysis requirements.
- Highlight areas in need of cross-talk between theory and experiment.
- Establish benchmarks for MCEG development.

210 participants.

Workshop report has been input to Snowmass white paper on "Event Generators for High-Energy Physics Experiments"

Event Generators for the EIC

arXiv:2203.11110

Submitted to the US Community Study
on the Future of Particle Physics (Snowmass 2021)

Event Generators for High-Energy Physics Experiments

We provide an overview of the status of Monte-Carlo event generators for high-energy particle physics. Guided by the experimental needs and requirements, we highlight areas of active development, and opportunities for future improvements. Particular emphasis is given to physics models and algorithms that are employed across a variety of experiments. These common themes in event generator development lead to a more comprehensive understanding of physics at the highest energies and intensities, and allow models to be tested against a wealth of data that have been accumulated over the past decades. A cohesive approach to event generator development will allow these models to be further improved and systematic uncertainties to be reduced, directly contributing to future experimental success. Event generators are part of a much larger ecosystem of computational tools. They typically involve a number of unknown model parameters that must be tuned to experimental data, while maintaining the integrity of the underlying physics models. Making both these data, and the analyses with which they have been obtained accessible to future users is an essential aspect of open science and data preservation. It ensures the consistency of physics models across a variety of experiments.

CP3-22-12 DESY-22-042 FERMILAB-PUB-22-116-SCD-T IPPP/21/51
JLAB-PHY-22-3576 KA-TP-04-2022 LA-UR-22-22126 LU-TP-22-12
MCNET-22-04 OUTP-22-03P P3H-22-024 PITT-PACC 2207 UCI-TR-2022-02

Monte Carlo Simulation of

- electron-proton (ep) collisions,
- electron-ion (eA) collisions, both light and heavy ions,
- including higher order QED and QCD effects,
- including a plethora of spin-dependent effects.

Common challenges, e.g. with HL-LHC: **High-precision QCD measurements require high-precision simulations**.

Unique challenges MCEGs for electron-**ion** collisions and **spin-dependent** measurements, including novel QCD phenomena (e.g., GPDs or TMDs).
Will result in deeper understanding of QCD factorization and evolution, QED radiative corrections, hadronization models etc.

MC4EIC in 2022

ongoing

1) Following up directly on previous workshops, we have had **in-depth reports on the precision of foreseen measurements and the related MCEG** needs.
2) This will frame a discussion of **MCEGs that are currently being developed for the EIC**. Developers are presenting today the status of their MCEG project and the thrust of future R&D. We will discuss where we can work together.
3) Developing precision simulation will require advancements in QCD theory and computational methods, as well as a close dialogue between experimentalists and theorists. We would like to conclude the workshop with a discussion on how to facilitate this dialogue and **work together on cross-cutting aspects between theory and experiment in NHEP**.

After This Talk: Detailed Update from AIWG

AI/ML already has an **important presence in EIC** with many prototypes, e.g., for detector optimization or reconstruction methods using ML.

To explore and develop the full potential of **AI/ML for the EIC**, we as a community need to **move from prototyping to production** and add promising AI/ML solutions into our workflows.

Promising candidate: Detector optimization using ML.

- EICUG fosters the community.
- EICUG SWG is a forum for discussion and forward-looking projects.
- Ideal platform for cross-collaboration:
  - Experient and theory
  - EPIC and 2nd Detector
  - Interdisciplinary collaborations (data scientists, HEP, …)