# Geant4 Input to Community Input on the EPPSU

**Ben Morgan** (The University of Warwick)
*On behalf of the Geant4 Collaboration*

WARWICK
THE UNIVERSITY OF WARWICK

GEANT4
A SIMULATION TOOLKIT

# Geant4: Prime Objectives

1. Provide **long-term support** (LTS) to experiments using the toolkit
2. Provide **long-term maintenance and sustainability** (M&S) of the toolkit for experiments
3. **Improve** the **physics models** for precision and energy range coverage to meet experiment requirements
4. **Improve** the overall **technical performance** of the toolkit

***Implicit in what follows: Geant4's challenges are ultimately defined by the needs of the experiments using it and how to deliver these***

# Overall Development Challenges

- The HSF submission to the 2020 LHCC HL-LHC Computing Review noted three topics for Geant4 development:

  *"Firstly, it is necessary to continue to modernise and refactor the simulation code by using more optimal, modern programming techniques.*

  *Secondly, different fast simulation techniques, where tracking of individual particles in given detectors is replaced by some sort of parameterisation and precision is traded to achieve higher event throughput.*

  *Finally, the use of compute accelerators (like GPUs or FPGAs) is the third emerging path that needs to be undertaken and that requires intense R&D effort."* HSF-DOC-2020-01

- **First item ongoing and will remain a core topic**, but challenging as:
  - *Experiments want **stability of API/physics**, but modernization eventually leads to **breaking API** backward compatibility (so **co-working on cost/benefit vital**).*
  - *Availability of personnel at sufficient FTE at current funding level (**more later**)*

# Physics Modelling Challenges: Electromagnetic

- **Increased need for precision** => include higher-order corrections to all EM physics processes (*simple sentence hides complexity of task*)
  - *Some progress already with implementation of* $e^+e^-\rightarrow\gamma\gamma\gamma$ *in Geant4 11.3*
  - *Implementations for rare EM processes will also be needed.*
- High-granularity calorimeters => more precise tracking, may require new algorithmic methods in physics.
- Potential use of MC Event Generators to calculate cross-sections or final states **during transport loop** is one specific area to explore
  - *Similarly, use of Event Generators for decay processes/final states during transport*

- **Overarching all of these**: though EM models have solid theoretical base, **validation against experimental data essential** as need for precision increases.

# Physics Modelling Challenges: Hadronic

- Development of **nuclear de-excitation, neutronics at low energies** expected to continue, important collaboration and common topic with **Nuclear Physics** experts.

- **Most important development for HEP** will be possibility to use FLUKA hadronic physics as part of an overall Geant4 Physics List
  - *Will require FLUKA-CERN v5 in C++, which is expected by Run 4.*

- Extension of FTF/QGS string models above 5-10TeV, or development of models like EPOS/DPMJET valuable for **cosmic-ray experiments or FCC-hh**, for example.

- **Key overall challenge**: phenomenological nature of models means **ongoing validation** against thick/thin target **experiment data is essential** to the improvement and development of models such as
  - *Multi-parameter and energy transition range tuning for FTF/QGS/BERT models*
  - *More detailed use of models according to particle types, energy, target material*
  - *More specific tunings by experiments against, e.g. test beam data, important as well, with co-working essential to assist in understanding validations here*

# Geometry Description/Navigation Challenges

- Geant4's Geometry Modeller is used reliably for O($10^{25}$) geometrical steps per year by LHC experiments.
- Key requirements:
  - *robustness for complex detector geometries*
  - *optimal performance on current and future CPU and GPUs*
  - *a platform to investigate other hardware architectures or approaches*
- Today's approaches:
  - *Geant4 navigation engine use the **solid modelling** approach*
  - *VecGeom R&D on **surface modelling** on GPU for AdePT (and ORANGE within Celeritas) is a promising base for GPU geometry navigation.*
- We foresee the need to retain, refine and further develop **both solid-based and surface-based modelling** in mature and well-tested implementations, to ensure robust and performant use on CPUs and GPUs.

# Fast/Biasing Simulation Techniques

*"...Secondly, different fast simulation techniques, where tracking of individual particles in given detectors is replaced by some sort of parameterisation and precision is traded to achieve higher event throughput."* HSF-DOC-2020-01

- Challenges for application of any ML/Biasing model (not only Geant4-provided):
  - *Efficient event processing that allows ML GPU inference with large batches of particles*
  - *Models that balance speed/accuracy for granular EM **and** Hadronic calorimeters (High-Lumi and Future colliders)*
  - *How best to handle geometrically/kinematically distinct regions of a detector.*
  - *Lack of personpower/attractiveness of work: not only model design, but also integration, validation within experimental frameworks*

- **Specific for Geant4-integrated, general approach to fast/biased sim**:
  - *Lack of interest from major LHC experiments (who can afford dedicated fast simulation teams unlike **smaller experiments**) leading to decreased funding and **potential lack of future support***
  - *Co-working with experiments to identify/implement where biasing or general ML can give benefit (e.g. Russian Roulette in CMS/ATLAS, Woodcock Tracking in ATLAS)*

# Value of General Fast/Biasing Simulation

*"...Secondly, different fast simulation techniques, where tracking of individual particles in given detectors is replaced by some sort of parameterisation and precision is traded to achieve higher event throughput."* HSF-DOC-2020-01

- Geant4-provided detector-independent workflow for fast calo simulation released
  - *Workflow used for data production for training and for integration of models back within Geant4*
  - *Wrappers of DD4hep: ddfastsim for voxelisation, DDFastShowerML for integration within simulation*
  - *This workflow and VAE model is implemented and tested by LHCb (CHEP'24) showing satisfactory accuracy*
  - *Involvement in organising CaloChallenge 2022, showing how many models may be used in this workflow.*
- Demonstration of **generalisation capabilities of ML models** trained within this workflow that **drastically reduces computing needs** (need of adaptation on top of pre-trained models):
  - *MetaHEP approach based on the VAE model*
  - *Generalisation of the CaloDiT model (ML4Jets'24)*
- **Challenge for Geant4 with ML is prescience: what balance between experiment-specific and general techniques will arise in response to overall ML landscape, and how should Geant4 adapt to this whilst retaining expertise to support experiments large and small?**
- **Equally, how best to support Biasing/Optimization methods in cooperation with experiments? e.g. ATLAS achieved 1.8x throughput from Run2-3 through combination of these**

# R&D Topics and Challenges

*"... Finally, the use of compute accelerators (like GPUs or FPGAs) is the third emerging path that needs to be undertaken and that requires intense R&D effort."* HSF-DOC-2020-01

- Major progress on using GPUs for e/γ transport by the AdePT and Celeritas projects - **see their talks later for their challenges and perspectives**
  - ***Optical photon transport on GPU*** *another key area where R&D is ongoing, with interest from LHCb and the Intensity Frontier/Low Background domain*

- Important that R&D is **not just for GPUs** - core algorithms/methods also vital
  - *Reducing expensive calls (math, MFP/MSC), variance reduction (e.g. Woodcock Tracking), optimized data/algorithms (G4HepEm)*
  - *We note the synergy/feedback loop with GPU work, especially in HEP/HPC knowledge sharing!*

- **From Geant4 perspective, challenge over next decade will be twofold**
  - ***If/how to integrate successful outcomes****, even partial, into production for experiments to deliver performance, validation, ease of use,* ***portability****.*
  - *If integrated into Geant4,* ***how to ensure LTS and M&S*** *given technology evolution and funding timescales*

# Sustainability of Personnel: M&S vs R&D

- **Key challenge: Today's R&D may be tomorrow's LTS/M&S**
  - *Sustaining successful R&D outcomes into Geant4 production releases requires **sufficient personnel with expertise** be available for M&S tasks over 5-10 years*
  - ***Continuity/overlap** of personnel critical here, alongside **training opportunities** to keep pace with rapid technological changes that may impact, e.g. GPU code.*
    - *R&D funding can sometimes have a cliff-edge…*
  - *Also an **infrastructure** problem to ensure availability of testing, performance, and physics validation on quickly changing heterogeneous CPU/GPU systems.*

- **Balance in personnel/funding needed** to ensure **M&S** for experiments is not compromised, alongside **R&D** to meet their new/future challenges
  - *Ideally **not zero-sum** for funding or overall FTE when need arises.*
  - *Again, co-working/consultation with experiments vital*

# Sustainability of Personnel in General

- *"There is a strong consensus throughout the field that the decline in detector simulation investment, including the Geant4 toolkit, must be reversed, with funding not just restored, but increased above historical levels"*: Snowmass 2021 report

- **This is perhaps Geant4's (or any common scientific software) greatest challenge over the EPPSU timescale**
  - *Generational change underway, but **limited funding opportunities or stable career paths** for the new generation, making it less attractive despite the **interesting work** and challenges.*
  - ***Short-term/piecemeal funding** means small FTE fractions spread over many people and short time periods.*
  - *Continuity/overlap of personnel over longer periods vital for training and **knowledge transfer**.*
  - *Technology evolution requires a **balance of expertise across HEP/HPC**, which may have different funding bodies and programmes which may not be connected.*
- Contributions by user community a valuable resource and **always welcome**, but cannot be the **sole mechanism** of Geant4's LTS and M&S for experiments
  - ***Not just coding** - **significant** M&S effort required for **performance measurement** and most critically, **physics validation**, to ensure changes are fit for use in production*

# Comments, Questions?

WARWICK
THE UNIVERSITY OF WARWICK

GEANT4
A SIMULATION TOOLKIT