# AI in ePIC and the Mystery of the AI WG

- There is a common misconception that an **AI WG** exists within ePIC. **It does not**.
- AI is a **cross-cutting theme** that touches **all areas of activity** and is embedded across multiple WGs.
- Our metric for success is the **integration of AI into simulation, reconstruction, and analysis workflows**, a high bar that requires moving from prototypes to solutions that are general, maintainable, and scalable. Our role is to **coordinate and support AI developments** within ePIC.
- At this stage, it is especially important to **provide the AI and data infrastructure** needed to enable and scale these developments effectively.

Software & Computing Organization
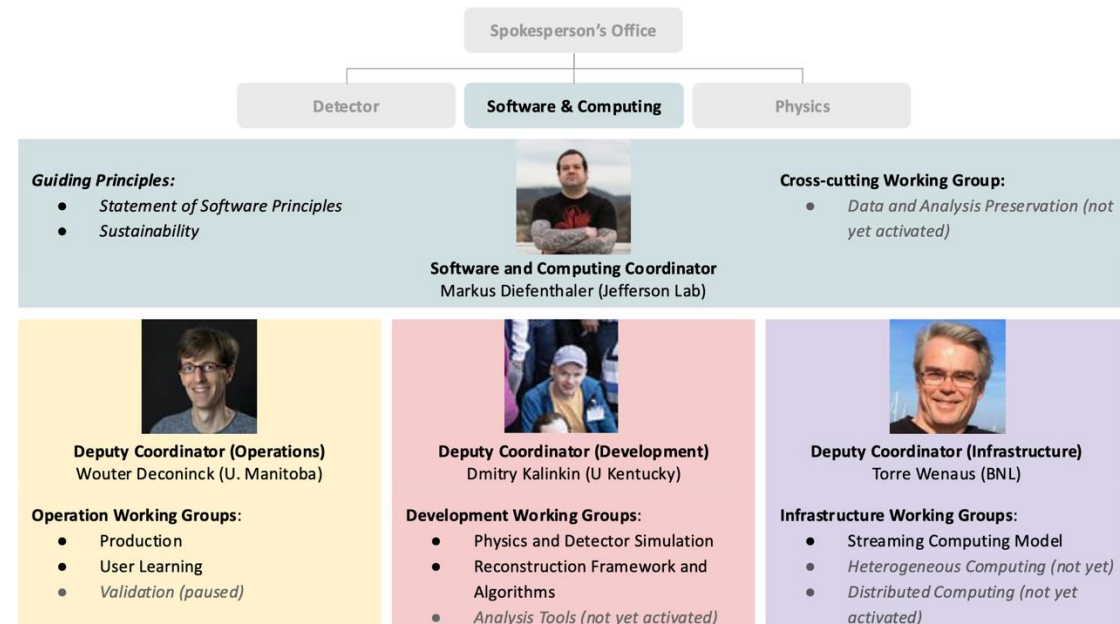
Our Three Priorities and the Role of AI

Software and Simulations for the (pre)TDR

Onboarding and Community Building

Computing Model and Testbeds

- **Simulations:** Fast and full simulations in Geant4, as well as ML models for specific components (e.g., optical photon simulation). These efforts will inform our strategy to accelerate simulations while substantially reducing the compute resources required.
- **Reconstruction and Analysis:** It will be interesting to explore where AI algorithms can outperform traditional methods. However, this is not our current priority. We welcome all types of reconstruction approaches, with a focus on addressing existing gaps.
- **Production:** Autonomous monitoring and validation of simulation production will improve turnaround time, increase engagement in simulation workflows, and serve as a first step toward autonomous experimentation and control (see Computing Model and Testbeds).

- **AI-Assisted Software Development:** Wouter will guide a discussion.

- **See next slide**.

Streaming DAQ and Computing Milestones

FY25	FY26	FY27	FY28	FY29	FY30	FY31	
PicoDAQ	MicroDAQ	MiniDAQ	Full DAQ-v-1	Production DAQ			DAQ
Streaming Orchestration			Streaming Challenges				
AI-Empowered Streaming Data Processing			Analysis Challenges				Computing
				Distributed Data Challenges			
AI-Driven Autonomous Alignment and Calibration			Self-Driven ePIC Experiment				AI

- **Compute**-**Detector** **Integration**:

  - Joint deliverables between **DAQ** and **computing** to develop integrated systems for detector readout, data processing, and ultimately physics analysis.

  - **Key role of AI**: Empowering data processing and enabling autonomous experimentation and control.
    - A priority are **AI agents for calibration workflows**.

A Word on Data

- **Data Model**:
  - AI workflows are **supported** through containerized environments including **JAX**, **PyTorch**, and **TensorFlow**, **ONNX** for inference, and the **ePIC data model.**
  - The data model enables data-driven API design, promoting modularity through standardized interfaces across simulation, reconstruction, and analysis tasks.

- **Open Data**: Simulation data is publicly accessible from both Python and ROOT, with Python serving as a bridge to the broader data science community and tools.

- We must discuss priorities around:

  - **Adopt Rucio as the Default Data Management**: Fully transition to Rucio for managing simulation productions and MC and background input. This includes establishing Rucio as the default interface through which the collaboration accesses and interacts with simulated data.
  - **Conclude Metadata Discussions**: Finalize the ongoing discussions on metadata, with an immediate focus on incorporating MC and background configuration settings as a component of the data (model).
  - **Evaluate and Advance FAIR Data Principles at ePIC**: Assessment of the implementation of FAIR (Findable, Accessible, Interoperable, Reusable) data principles at ePIC. Based on the findings, identify and prioritize the next steps to enhance compliance and utility.