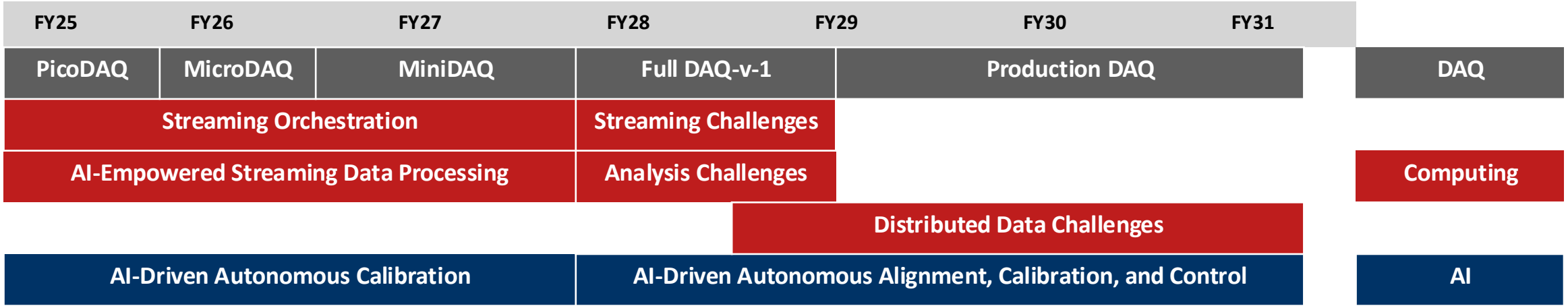
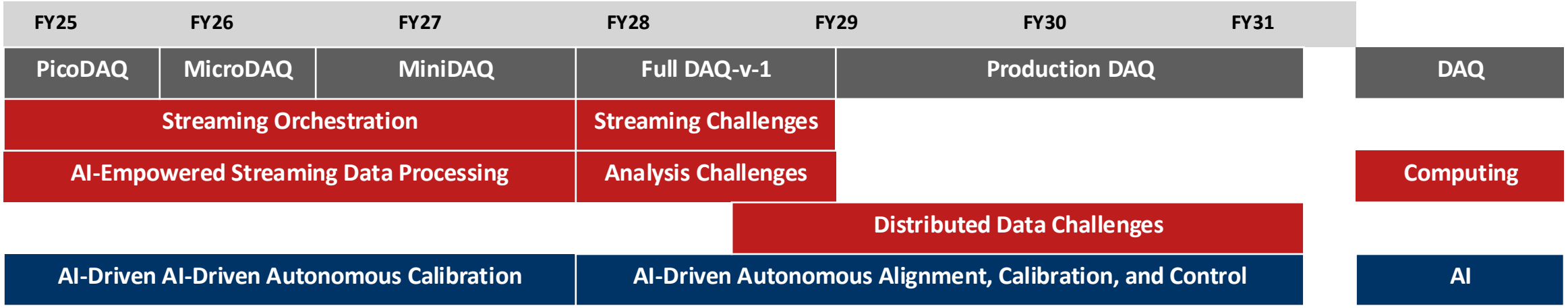
Streaming DAQ and Computing Milestones



- **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(/ML)**: Empowering data processing and enabling autonomous experimentation and control.

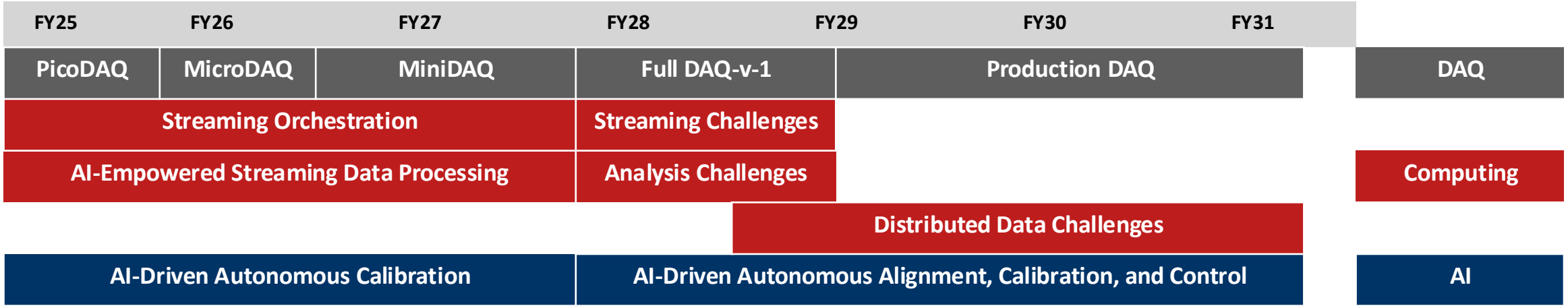
Streaming DAQ and Computing Milestones



AI-Driven Autonomous Calibration

- Progress continues on understanding calibration workflows in collaboration with subsystem experts, with a focus on identifying timelines and interdependencies.
- The strategy for autonomy involves algorithms for change detection and agentic workflows.
- **FY28 Q1 Goal**: Autonomous calibration of one detector system using simulated streaming data.
- Milestones and deliverables linked to:
  - Streaming Orchestration Milestones and Deliverables (slide 3)
  - Streaming Data Processing Milestones and Deliverables (slide 4)

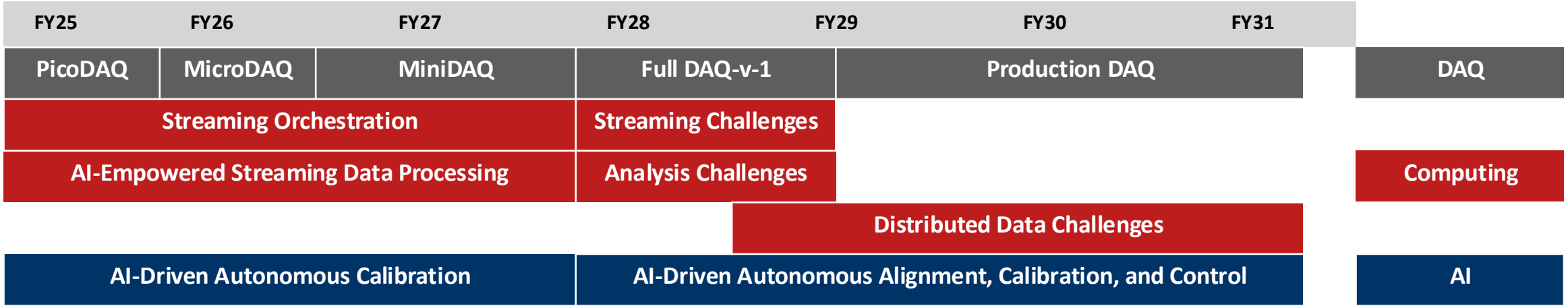
Streaming DAQ and Computing Milestones



Streaming Orchestration Milestones and Deliverables

- ✓ **Requirement documents** for streaming orchestration developed.
- **FY28 Q1 Goal**: Deliver a functional testbed for calibrating one detector system using simulated streaming data.
- Progress is ongoing in testbed development:
  - We are evaluating streaming orchestration using **PanDA + Rucio** (slides 15–20).
  - We have demonstrated streaming data processing using **EJFAT** (slides 21–22).
  - Additional prototypes under consideration: LHCb Allen, SPADI Alliance.

Streaming DAQ and Computing Milestones



Streaming Data Processing Milestones and Deliverables

- ✓ **JANA2 enables data processing at the timeframe, event, and sub-event levels.**
- **FY28 Q1 Goal**: Achieve streaming data reconstruction with high efficiency in identifying physics collision events in simulations, including varying levels of background. This includes an AI/ML challenge focused on developing algorithms for distinguishing physics events from background.
- Progress is ongoing in streaming data reconstruction (slides 13–14).

AI-Driven Autonomous Calibration

Carlos (Muñoz Camacho) presented on Oct. 28 on the **Backward ECal (EEEMCal) Calibration**. The EEEMCal is a good example. Carlos will provide the necessary calibration scripts and integrate them into JANA2/EICrecon. As part of this integration, the data flow and required input data will be defined.

To achieve autonomy in this calibration case, we have three distinct tasks (**AI aspects** are highlighted in **dark blue**):

1. **Calibration Logic**: We need **software that detects when a new calibration is required** for the EEEMCal and updates the state machine accordingly.
2. **Calibration Integration Into State Machine**: The state machine must be connected with a calibrations/conditions database system to track calibration status and link to calibration data. Clarify who is reading and writing.
3. **Calibration Execution and Validation**: The **calibration script must recognize** when it has derived new constants, **validate these constants**, and then register them in the calibrations/conditions database.

There are also three boundary conditions:

- **The Online Condition**: All of this must function in an online context, meaning while data is still being accumulated and not yet available offline. We will start with file-based workflows to make immediate progress, while developing support for streaming.
- **The Human Condition**: We aim to automate these processes as much as possible. However, it is essential to clearly define where human oversight is required. This includes questions such as: When should a human intervene in the loop? Which decisions must be reviewed or approved manually? While this is a crucial discussion, this is a separate conversation. Our priority is to design and implement the autonomous system and then determine how and where to introduce the human-in-the-loop component.
- **The Cybersecurity Condition:** Cybersecurity requirements will be addressed in a later implementation phase.

Questions to the DSCs

- **From slide 2**: Progress continues on understanding calibration workflows in collaboration with subsystem experts, with a focus on identifying timelines and interdependencies. In simple terms: Are there any showstoppers that could prevent calibrating the ePIC data within two weeks?

Information to Provide to the DSCs:

- Markus will prepare a slide explaining what is meant by AI-Driven Autonomous Calibration.
- Jeff will prepare a slide outlining the different levels of calibration.
- Marco and Taku will draft a set of questions for the DSCs:
  - What is the minimum amount of statistics required to calibrate your system?
  - Can the necessary calibration data be collected during regular data taking, or is a dedicated calibration run required?
  - On which other calibrations does your system's calibration depend?
- We propose using the spreadsheet compiled by Jin to update and organize this information:
  - Marco and Taku will revise and complete the entries as needed.
  - They will also prepare simple instructions for the DSCs on how to provide the revised information for the spreadsheet.