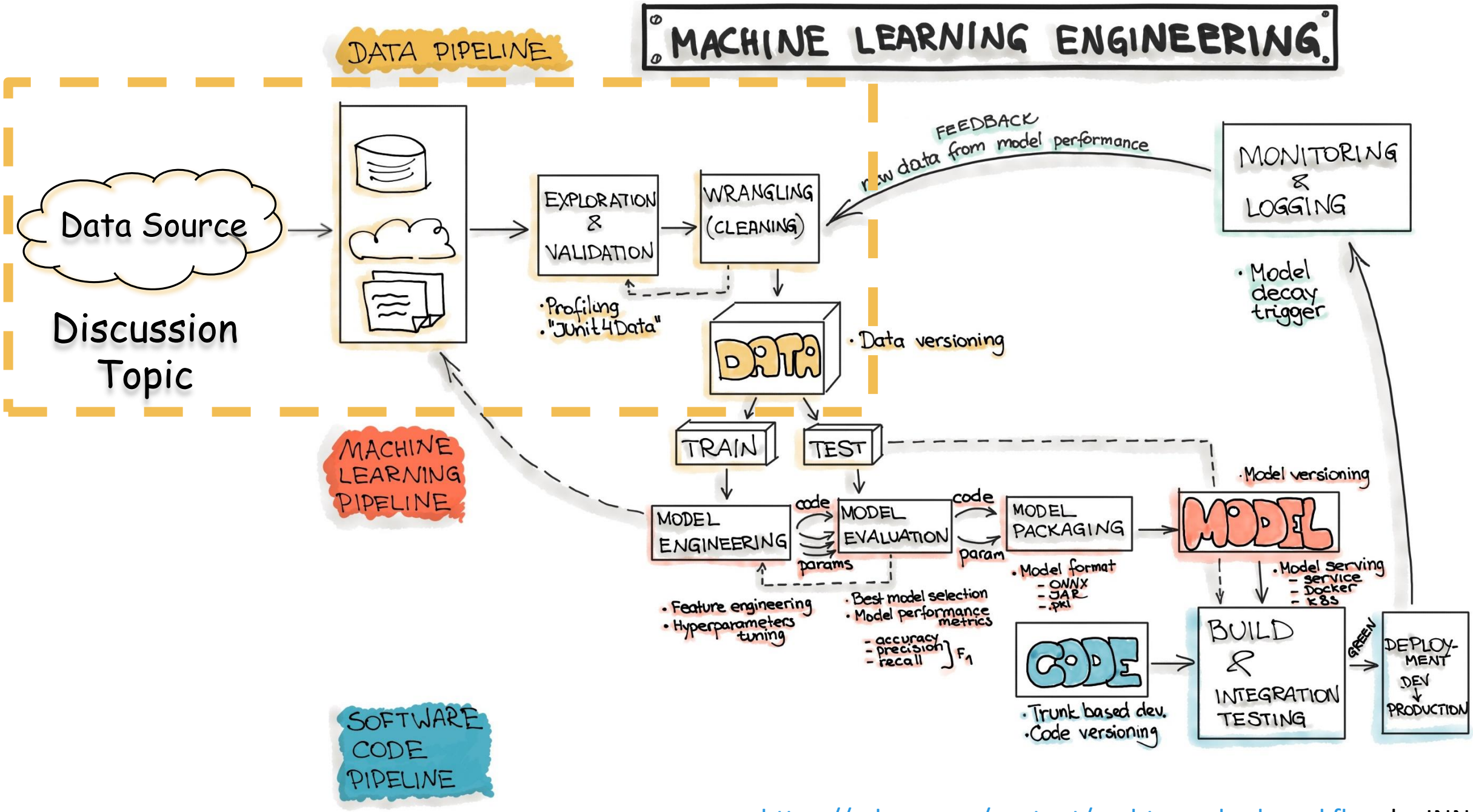
Discussion Section: Accelerator ML Challenges

ADAM CARPENTER AND MATT POELKER

AI IN THE ACCELERATOR

NOV 12, 2021



https://ml-ops.org/content/end-to-end-ml-workflow by INNOQ

Story of the C100 RF waveforms and Fault Classifier

Initial the waveform capture was needed to analyze fault modes in the C100 cryomodules.

Hardware and software work to capture data from faults

} Data Source Creation

Fault happens, waveforms saved to disk

SME investigated and labeled faults from operational running

} Data Set Creation

Data Sources and Sets

Data Source: System capable of generating raw data for data sets

- Can be specialized or general, hardware or software
  - RF waveform sampling system
  - NDX radiation detectors
  - MYA Archiver Database*
  - CEBAF Element Database*

Data Set: Collections of processed data

- Gathered from data sources and prepared for a particular purpose
  - Labeled harvester C100 RF fault waveforms
  - Field Emission Gradient Scans

Bottlenecks and Limitations

What hurdles prevent progress on developing data sources or data sets?

- Example: Streaming C100 waveform data concerns
  - Data rate / throughput could overwhelm IOCs NICs/CPUs or network infrastructure
  - ~300 Mbps across all C100 FCC IOCs at most
  - Will play some caching and reduction games to avoid any potential problems
- Need new/improved DAQ hardware systems?
- EPICS 3 vs EPICS 7?
- IT Infrastructure?
- Workforce/organizational limitations?
- Test stand availability?

[Bonus Question] How do we organize and maintain data sets?

Discussion Data Goals

Data sources

- Existing or not
- Point of contact
- Roadblocks

Data sets

- Existing or not
- Point of contact
- Roadblocks

Needed Improvements to support dataset and ML

We have a Google Live Notes up (link on Indico Overview). Please add your thoughts as we go.

- Increasing accelerator reliability
  - The downtime manager is a good source for prioritization information
  - Major themes are
    - Identifying incipient failures
    - Reducing time lost to component failures
    - Speeding trip recovery or reducing trip occurrence
- Increasing accelerator availability
  - Major themes are
    - Reducing time needed for beam studies/machine setup
    - Identifying and responding to off-normal accelerator configurations
- Improving energy reach
  - Major themes are
    - Distributing the gradient we have
    - Restoring the gradient we've "lost"
- Reducing workload for routine tasks
  - AI/ML is the next step in the general automation journey. Are there smaller tasks, appropriate for AI/ML, that happen enough to matter in aggregate person hours spent?