This project has received funding from the European Union’s Horizon 2020 Research and Innovation programme under GA No 101004730.

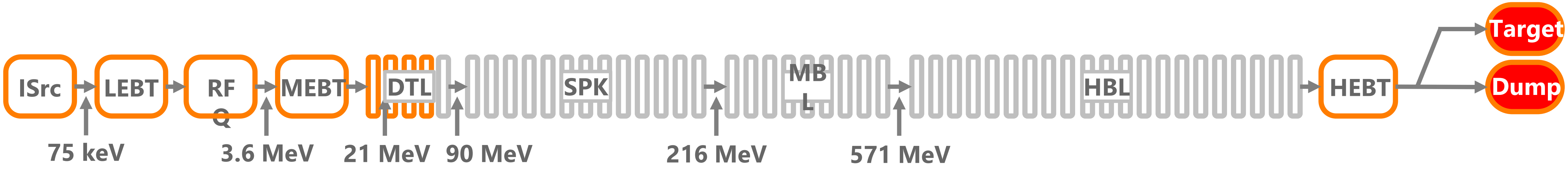
Overview of AI Activities for the ESS Accelerator

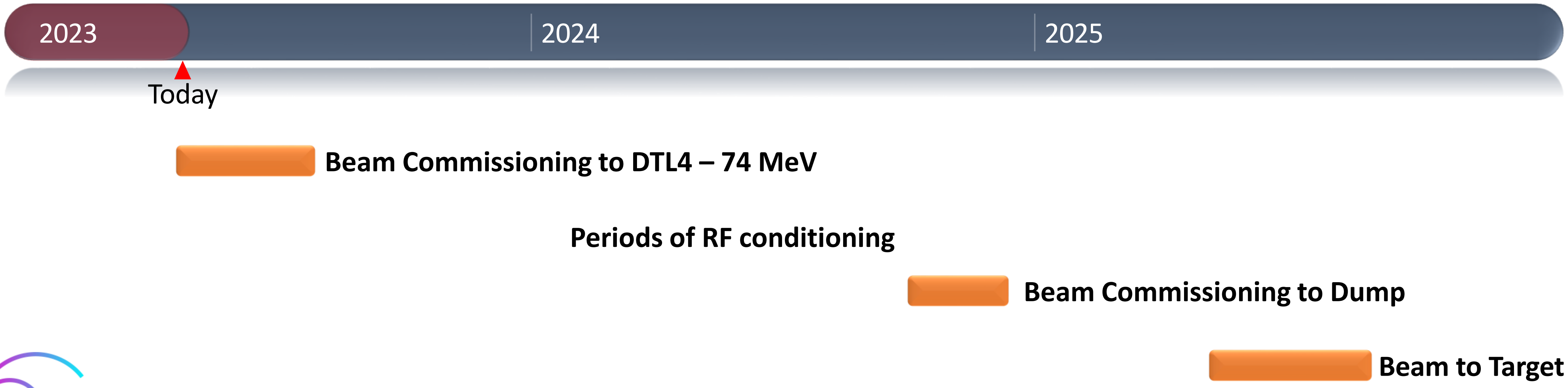
2nd I.FAST Annual meeting, 2023-04-20

Thomas Shea, Irena Dolenc Kittelmann, Karin Rathsman, Natalia Milas
and the ESS AI collaborators

iFAST

European Spallation Source Accelerator





>95% Availability

Examples of ML Studies at ESS

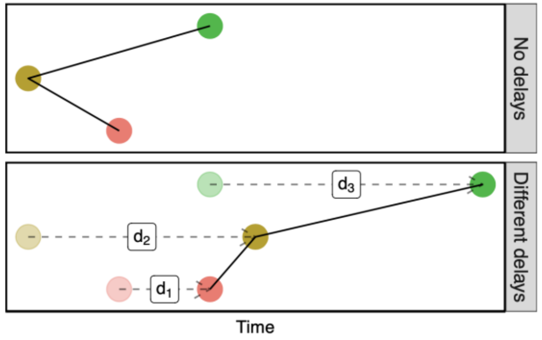
AI/ML activities for accelerator are ongoing in 3 parts of the organization: Controls, Beam Physics, Diagnostics

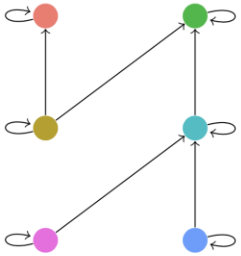
- Root cause analysis of alarm cascades – Controls
- Forecast temperature (water, oil…) and its impact on the machine – Controls
- Anomaly detection for target wheel – Controls
- Tuning the DTL in 1-3 shots – Beam Physics
- Data sharing for ML – Controls, Beam Physics, Diagnostics
- Low latency platform – Diagnostics
- Errant Beam identification/prediction – Diagnostics

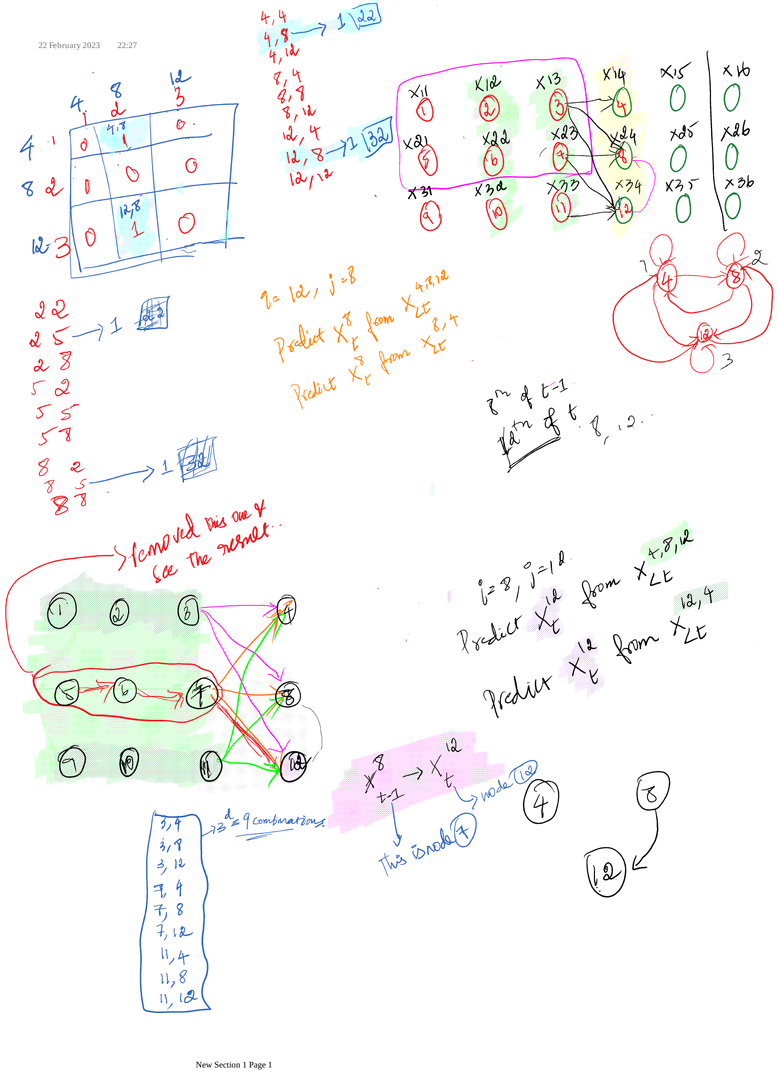
Alarms: Student project

Root cause analysis of alarm cascades.

- Post doc student, a phd student and a master student at automatic controls at Lund University.
- Data from the cryogenics system.



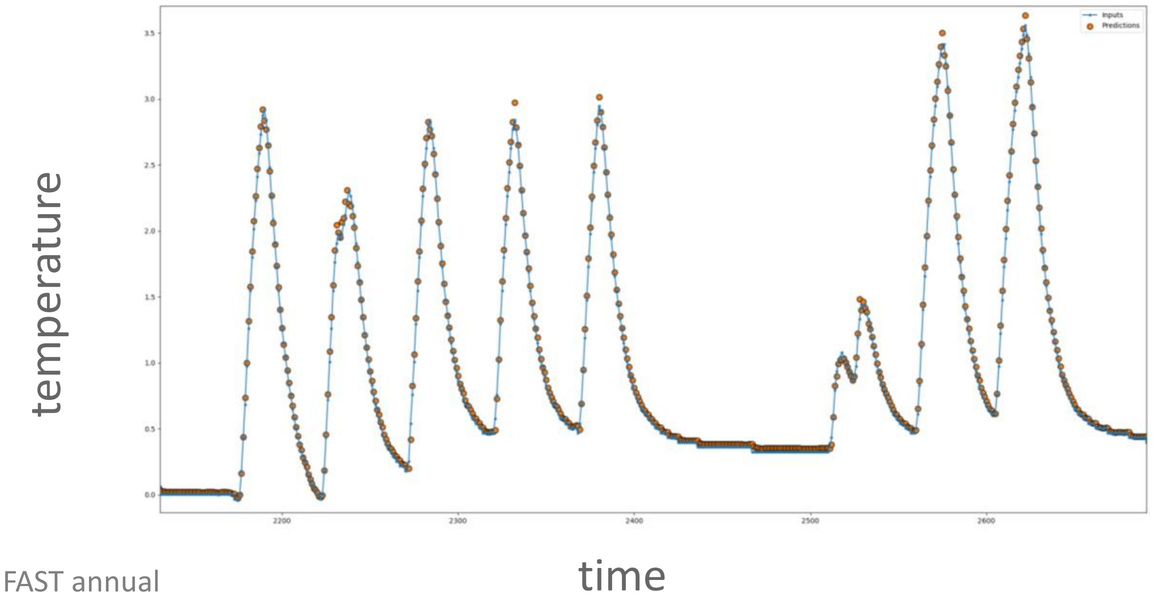




Temperature Prediction: Algorithms

Machine learning algorithms developed in other fields (natural language processing, robotics, image analysis and trading...) are applicable to process controls and does not necessary need to be developed.

- Build up the competence incrementally and make it work. (You don't start playing Bach double concerto as a newbie on violin and you don't need to build a violin to play.)
- Use the algorithms and realize that the bottleneck is not the algorithms.

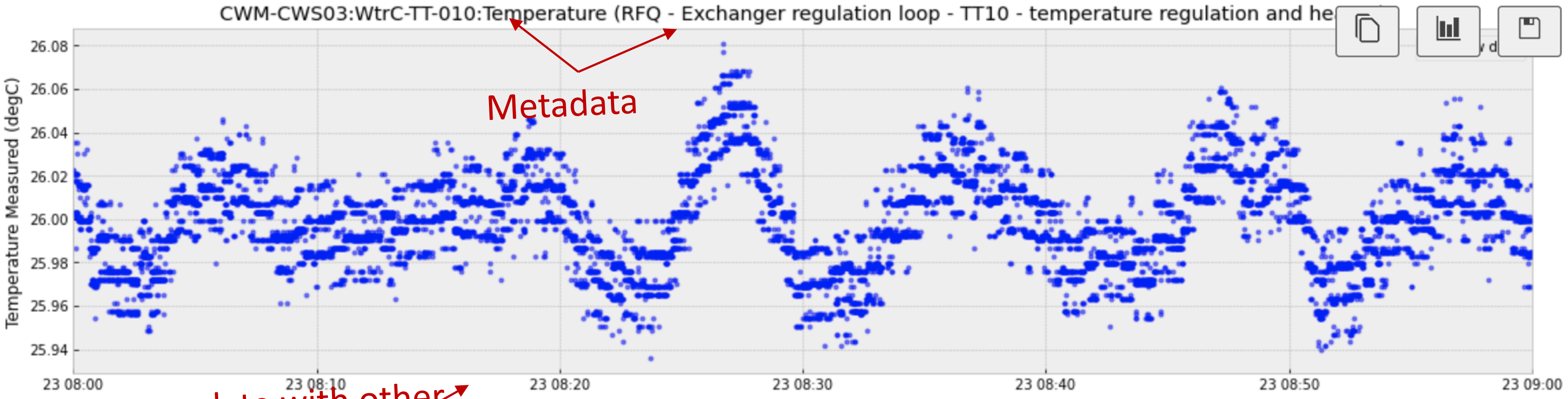


...LSTM (algorithm developed two decades ago and in your smart phone) solved the heat equation and could predict oil temperature ahead of time

Temperature Impact –
Detuning RF Structure: Data Preparation

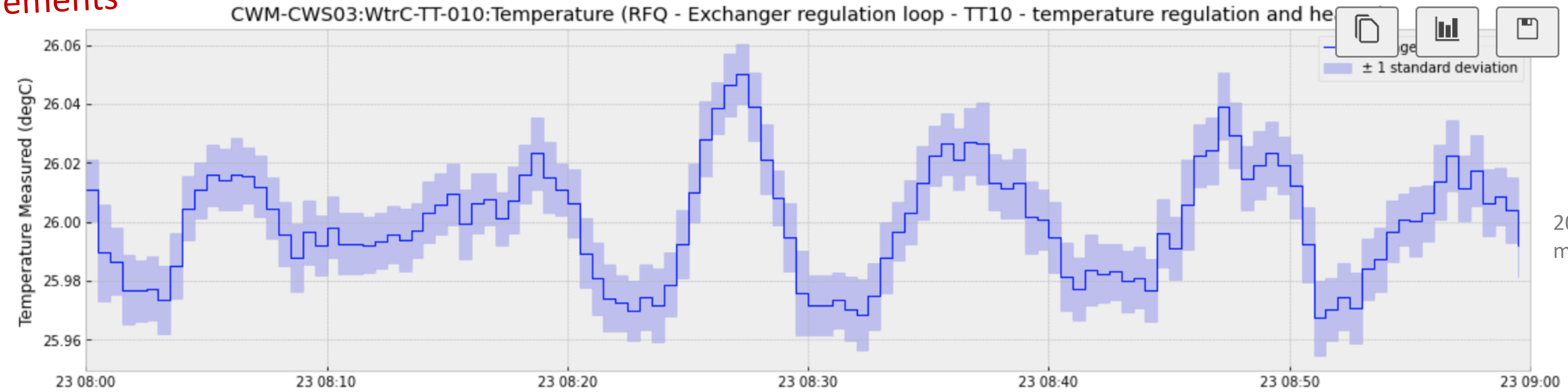
Time averaging and other aggregation functions are sometimes useful in the data preparation/reduction process

Metadata



Metadata

Time as index to correlate with other measurements



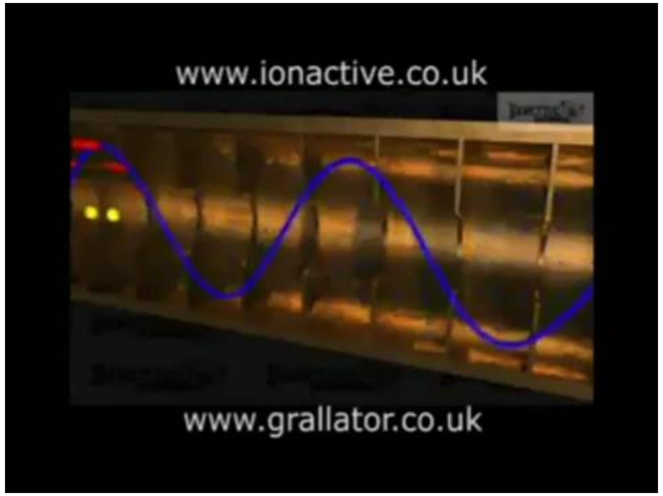
Anomaly detection: Student project

Anomaly detection target wheel

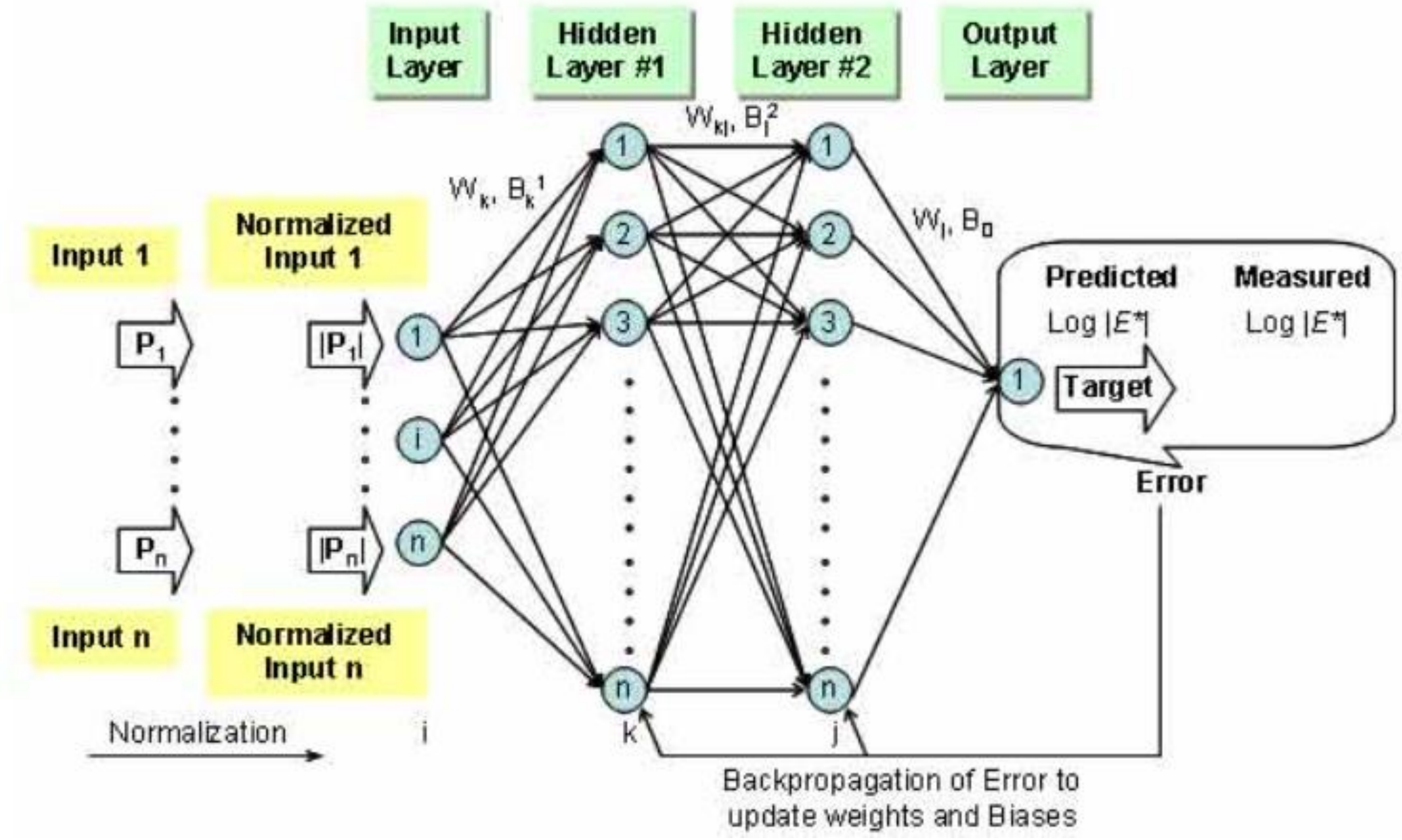
- The neutron source: sectored, rotating tungsten target
- Bachelor in software engineering from Chalmers in Gothenburg, supervised in the target division and ICS
- Focus on Software workflows for machine learning.

Developing a Machine Learning (ML)-based model for RF tuning of the DTL

- **RF tuning:**
  - Having the bunches of the proton beam arrived in a good timing with the RF field
  - Keeping the bunches synched with the field while they are travelling within the cavity
- **Challenges:**
  - The RF amplitude and phase experienced by the passing bunch do not correspond exactly to the ones set by the RF generator and controller (the RF modulator and the klystron) due to losses in the wave guides and some other construction-related limitations
- **Solution:**
  - Using some diagnostics being sensitive to time of flight of the bunches, like the phase of signals in the Beam Position Monitors (BPMs). There 6 of them in DTL1.



Artificial Neural Networks (ANN) models and the design parameters for the model for single-shot



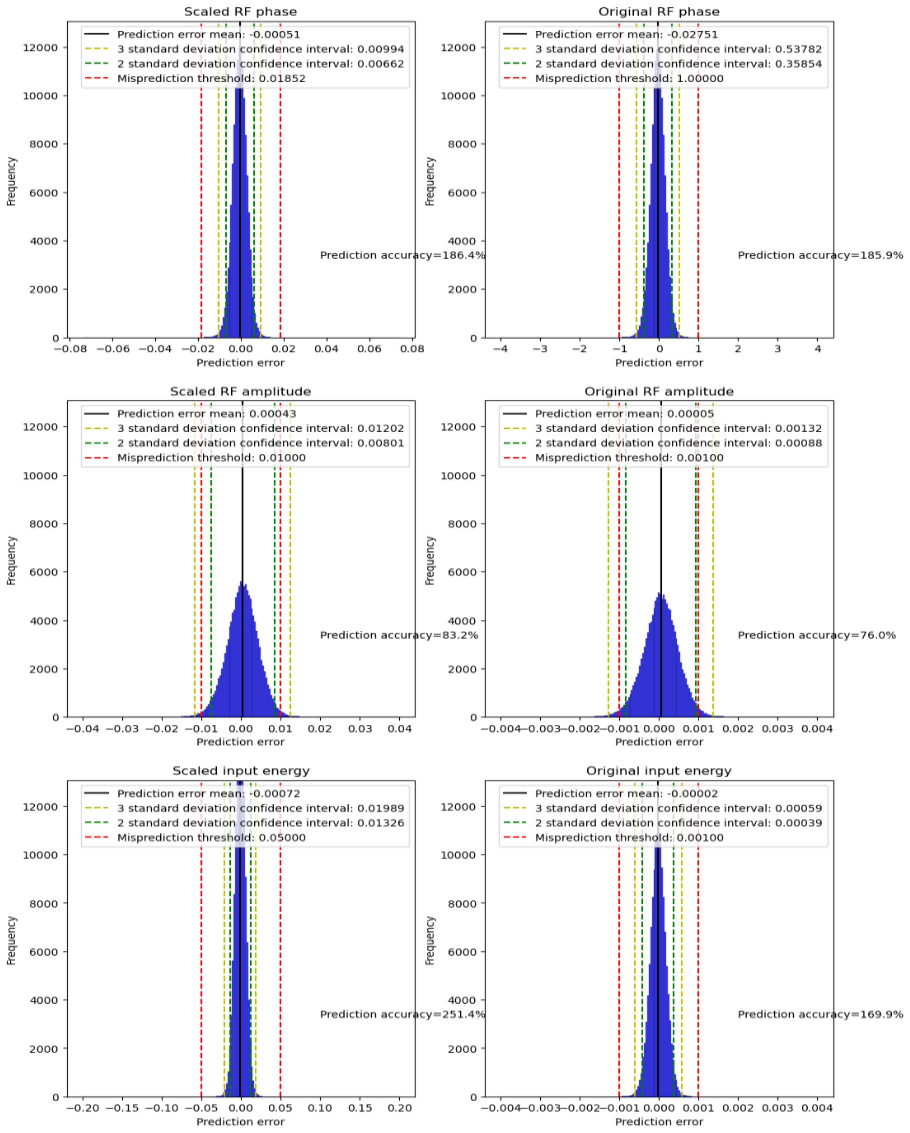
Schematic of how ANN models work, source: https://www.fhwa.dot.gov/publications/research/infrastructure/pavements/ltpp/10035/005.cfm

Layer widths	16-32-16-16-16-8
Initial LR	2e-3
Batch size	256
Optimization algorithm	ADAM
Number of epochs	300
Loss function	MSE loss
Activation function	ReLU

- Model structure and hyperparameter values for the best model in the single-shot measurement scenario
- For other measurement scenarios, they are the same, but the layer widths

Triple-shot measurement

- Histogram of the prediction error



RF phase
Prediction error: 0.54º
Requirement: 1º

RF amplitude
Prediction error: 1.32%
Requirement: 1%

Input energy
Prediction error: 0.06%
Requirement: 0.1%

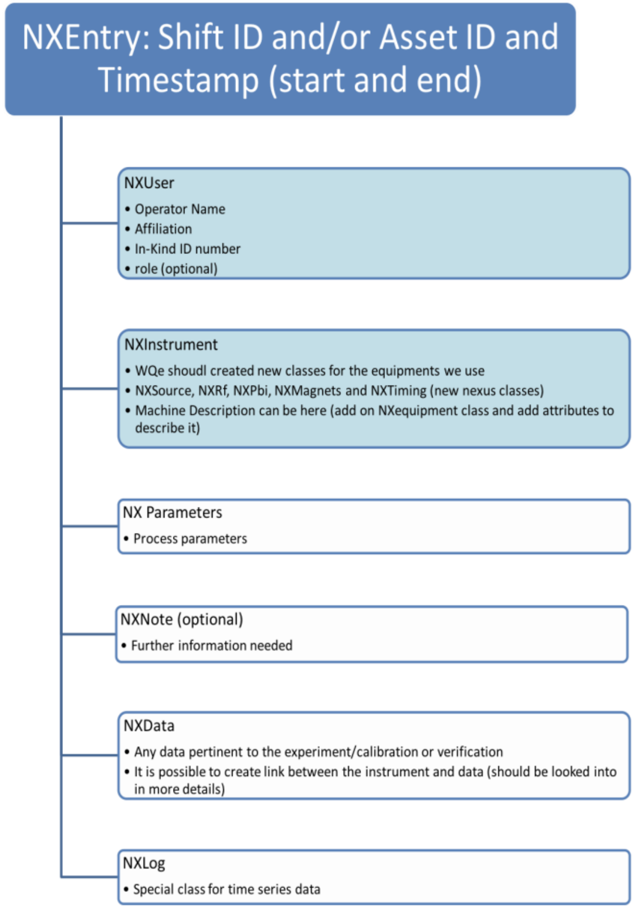
- Meets requirements for Phase and input energy; nearly for amplitude
- It is the best measurement scenario compared to the single-shot and double-shot ones
- The best model meets the requirements for the RF phase and input energy, but not for the RF amplitude by 99.7% confidence
- This is the model that is going to be deployed in ESS control system to be used while RF phasing or tuning of the machine

Data Sharing

- Example exchange format
  - NeXus (https://www.nexusformat.org) based on HDF5
  - Standard in the neutron and X-ray communities
  - First example: agreement with ORNL on emittance data structure.
  - For IFAST Task 10.6, we build on this to cover ML-related data

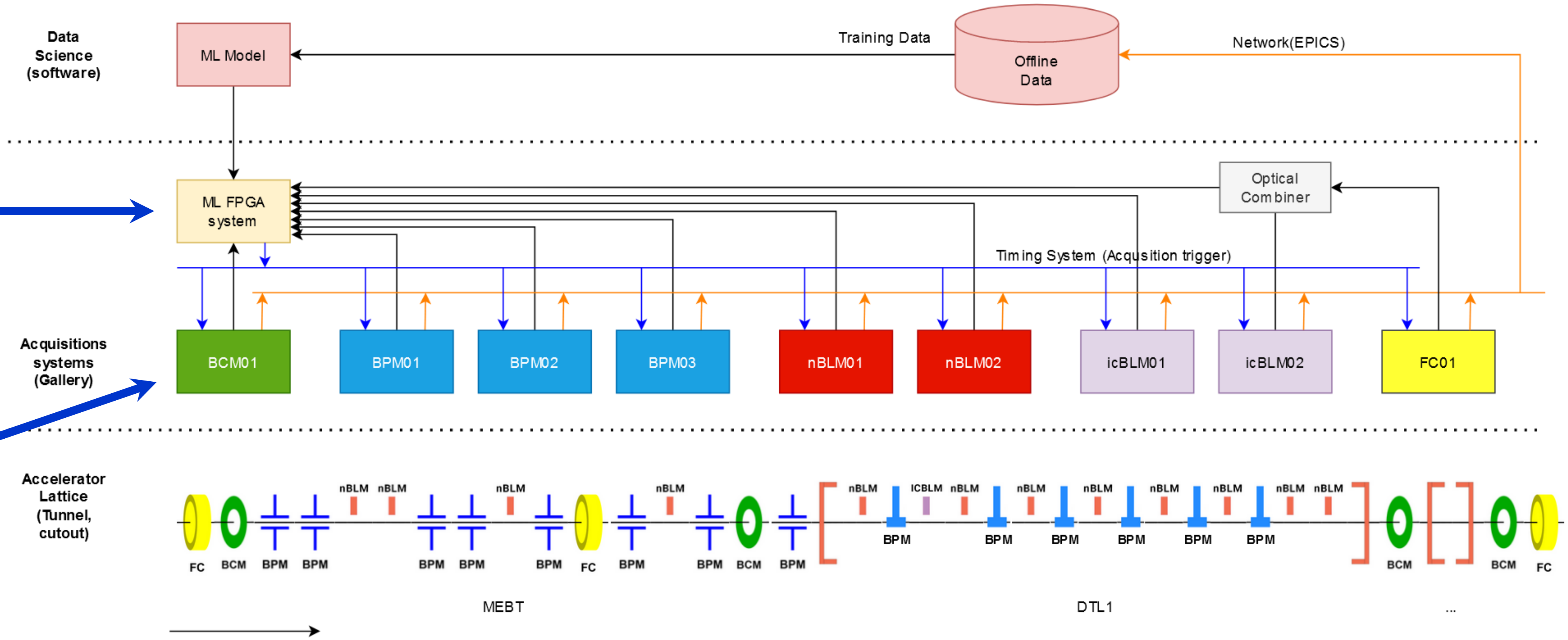
- ESS Controls, Lund University initiatives on data sharing, data quality, and openness:
  - https://portal.research.lu.se/en/publications/ess-control-system-data-lab-executive-summary
  - https://portal.research.lu.se/en/publications/tools-and-ecosystems-for-open-control-systems-data-at-ess

NeXus data format
(example ESS diagnostics data)



Low Latency ML Prototype – ESS Normal Conducting Linac, Low Energy Section

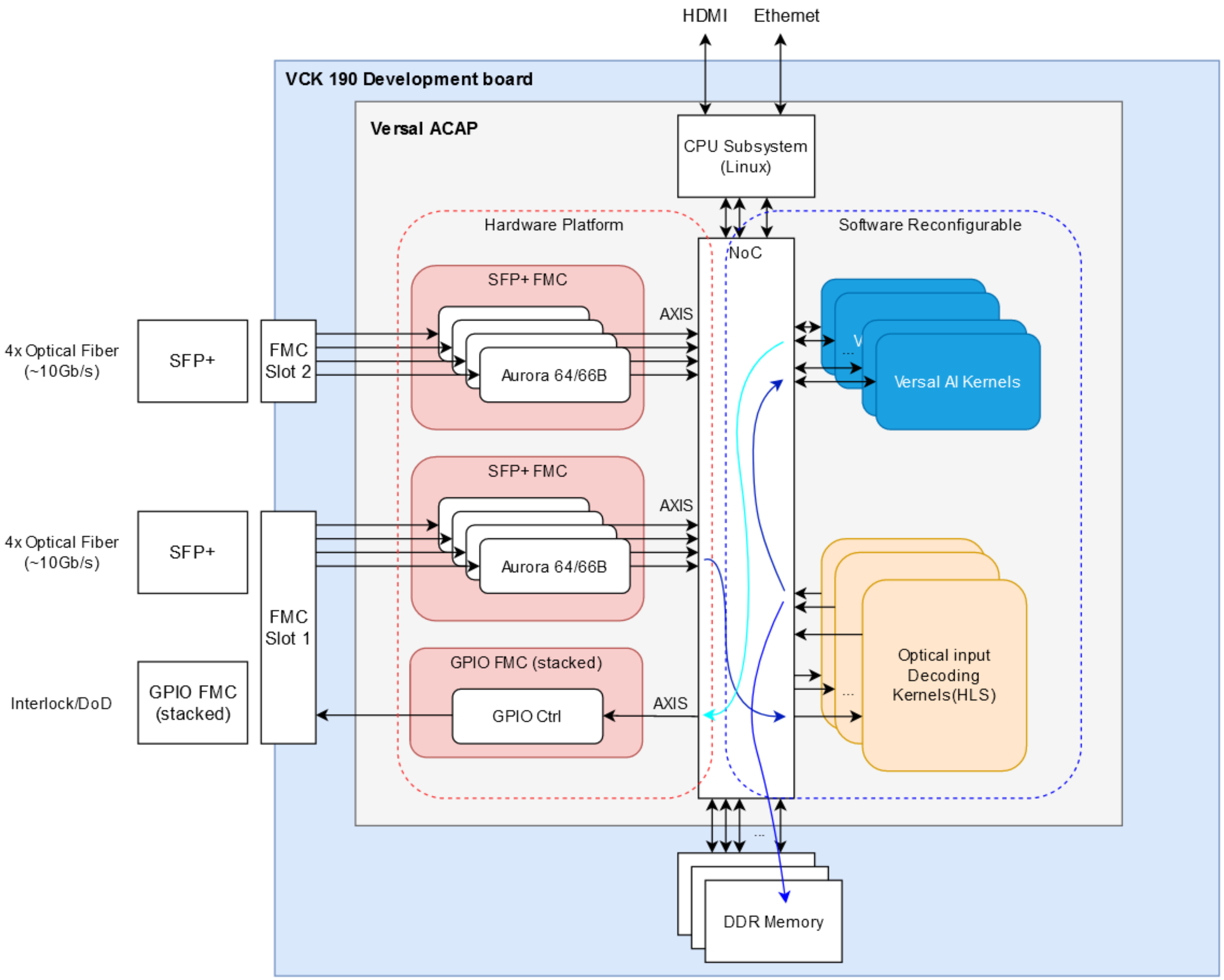
- System covering Normal Conducting Linac (NCL)
- 2nd layer: ML FPGA System to detect “off normal” events and request readout via timing system (DoD extraction).
- 1st layer: network of FPGAs to acquire and process signal from detector channels.
- System could connect to beam interlock system (not shown).



Rosengren, Dolenc Kittelmann, Shea

FPGA Platform

- Versal “Adaptable Compute Acceleration Platform, ACAP”
  - The NoC and Versal AI kernels in the diagram is made of dedicated hardware (not FPGA logic)
- The design can be divided into two parts (compare with tool diagram)
  - The “Hardware Platform” part is made by FPGA designer
  - The “Software Reconfigurable” part can be reconfigured in the SW compile flow with quick turnaround.
- Data can be streamed directly from input to the AI kernels for processing.



Low-latency applications of interest

- ML based Intelligent Trigger for Data-On-Demand (DoD)
  - Goal:
    - Identify "off-normal/interesting" and "normal" events
    - Trigger DoD* acquisition for each "off normal" event

    Data-On-Demand (DoD) definition: partially or fully processed highly detailed data, buffered and then extracted from the FPGA level on event occurrence (on demand)
  - Relevant systems: BLM, BCM, BPM

- ML based machine protection
  - Goal: inhibit beam production when dangerous conditions are recognised/predicted
  - Relevant systems: BLM and BCM; also BPM
  - Can be considered as specific version of "intelligent trigger" application
    - "Intelligent trigger" that is trained to recognise dangerous conditions, triggers DoD extraction (post mortem) and inhibits beam production.

Dolenc Kittelmann, Shea, 2nd I.FAST WP10 meeting, 2023-01-17

ML model selection: ESS experimental data

- ESS beam commissioning run in July 2022 (2 weeks):
  - Beam to DTL1 (21MeV): covering only part of the linac relevant for the ML prototype.
  - BPM, BCM, partly nBLM data acquired.
- Current focus: data curation
  - Goal: extract stable run data.
  - Idea:
    - Check beam stability prior to each beam trip.
    - Extract relevant data sets (BCM, BPM) to desired HDF5 file if beam stability condition is satisfied for certain time period before the trip – time period and stability conditions to be tuned.
- Two types of offline analysis studies planned with cleaned data:
  - Clustering.
  - Binary classification.

Model selection: ESS experimental data

- Clustering study:
  - Goal: find types of beam trips.
  - ML based analysis to explore whether the data can be grouped into several clusters of different beam trip types and normal operation.
- Binary classification:
  - Goal: prepare analysis tools, explore selected ML models for Intelligent trigger application.
  - Asses predictive power and limitation of a set of ML models.
    - Model dependence on coverage along the linac; waveform sampling rate, size and position with respect to the beam pulse window etc.
- Next ESS commissioning run in mid-April to mid-July 2023
  - Beam to DTL4 (74 MeV): covering full ML prototype.
  - Planned dedicated stability runs to collect relevant data and exercise selected model(s) offline (perform the same studies as on data from previous beam commissioning run).
  - Prototype tests – in lab and real conditions towards end of commissioning run.

Dolenc Kittelmann, Shea, 2nd I.FAST WP10 meeting, 2023-01-17

Retrospective and Outlook

- Focus now on long commissioning run. Specific beam studies are being planned including stability runs; always archiving detailed pulse waveforms as well as metadata.
- Continue simulations, analysis of existing datasets, and student projects
- Our data acquisition, analysis, archiving tools are not optimal for this work – data management effort dominates AI/ML effort. Cross-division initiative launched last year.
- Will continue to strengthen collaborations – within ESS and beyond

- Consider participation in a proposal development this year*
  - "HORIZON24-INFRATEC-01-01 WP3 - "Call - Developing, consolidating and optimising the European research infrastructures landscape, maintaining global leadership (2023)"
  - 62 million Euros for 6 or 7 projects, call closing March 2024
  - Opportunity for several million Euro, data-focussed collaboration across several RIs (initiative coordinated by Adnan Ghribi, GANIL/CEA/CNRS; 12 May 2023 meeting after IPAC)

*suggestion from C. Darve, ESS

ESS AI Collaborators

Tomaz Bryz (ESS, Controls)

Per Runesson (Lund University, Computer Science)

Emma Soderberg (Lund University, Computer Science)

Per Andersson (GoalArt AB and Lund University)

Jan Eric Larsson (GoalArt AB and Lund University)

Anna Hall (Big Science Sweden)

Paul Dadzie, (DVEL AB)

Jim Anderson (DVEL AB)

Karlis Berkolds (RTU)

Kaj Rosengren (ESS Diagnostics)

Roxana Tarkeshian (ESS Diagnostics)

Rafael Baron (ESS Diagnostics)

Clement Derrez (ESS Diagnostics; currently with Hilti)

Elena Donegani (ESS Diagnostics)

Yngve Levensen (ESS Beam Physics)

Cosylab

Søren Wegen Mogensen (Lund University, Automatic Controls)

Vishnu Pradheep Raveendran (Master Student, Lund)

Alessandro Bregoli (PhD Student, Milano)

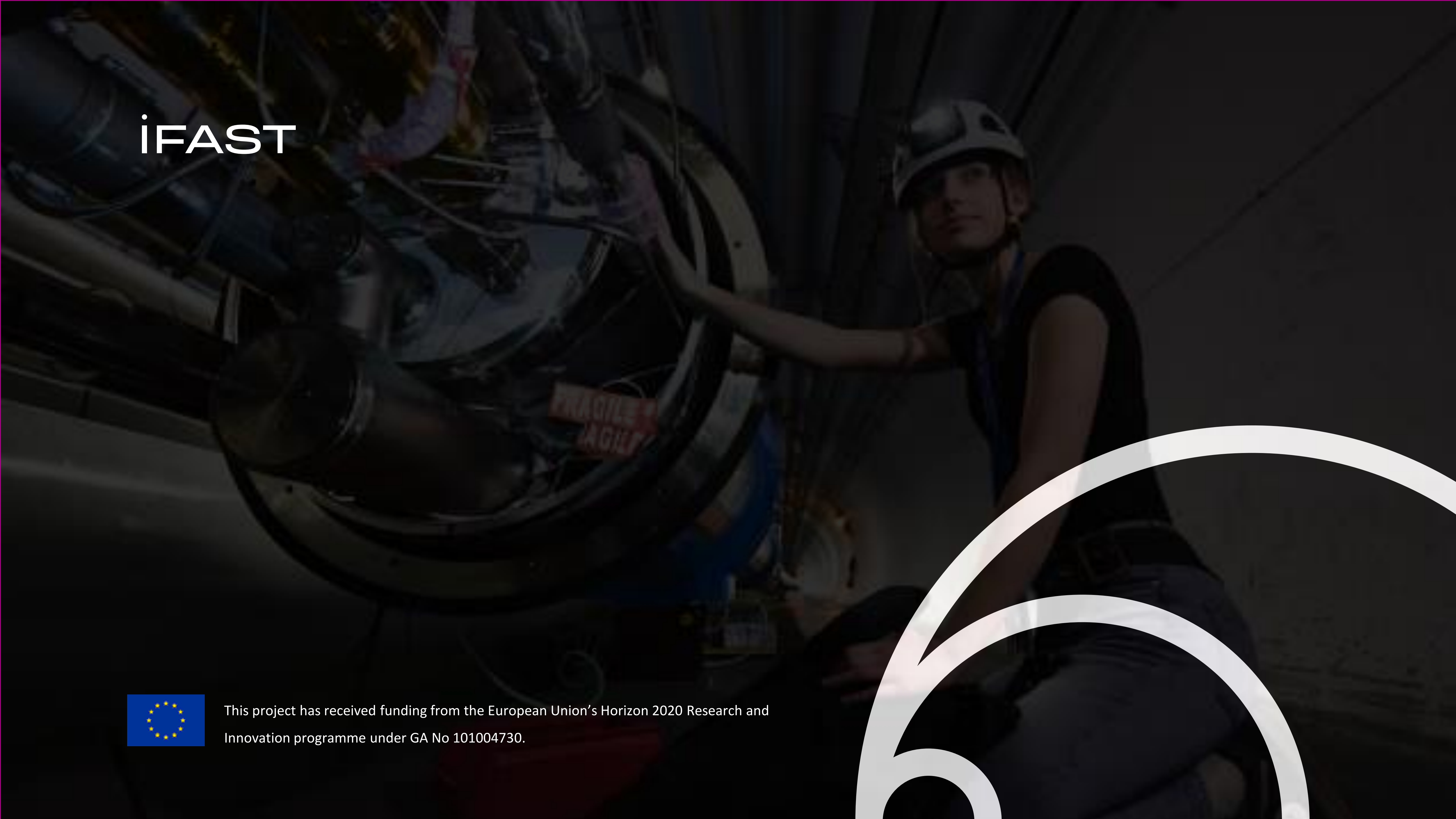
Per Nilsson (ESS, Cryogenics)

Amin Hosseini (Master Student, Lund University)

Bo Bernhardsson (Lund University, Automatic Controls)

Vernita Gouws (Student, Gothenburg)

Jamie Arriagada (ESS, Target Division)



iFAST

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under GA No 101004730.