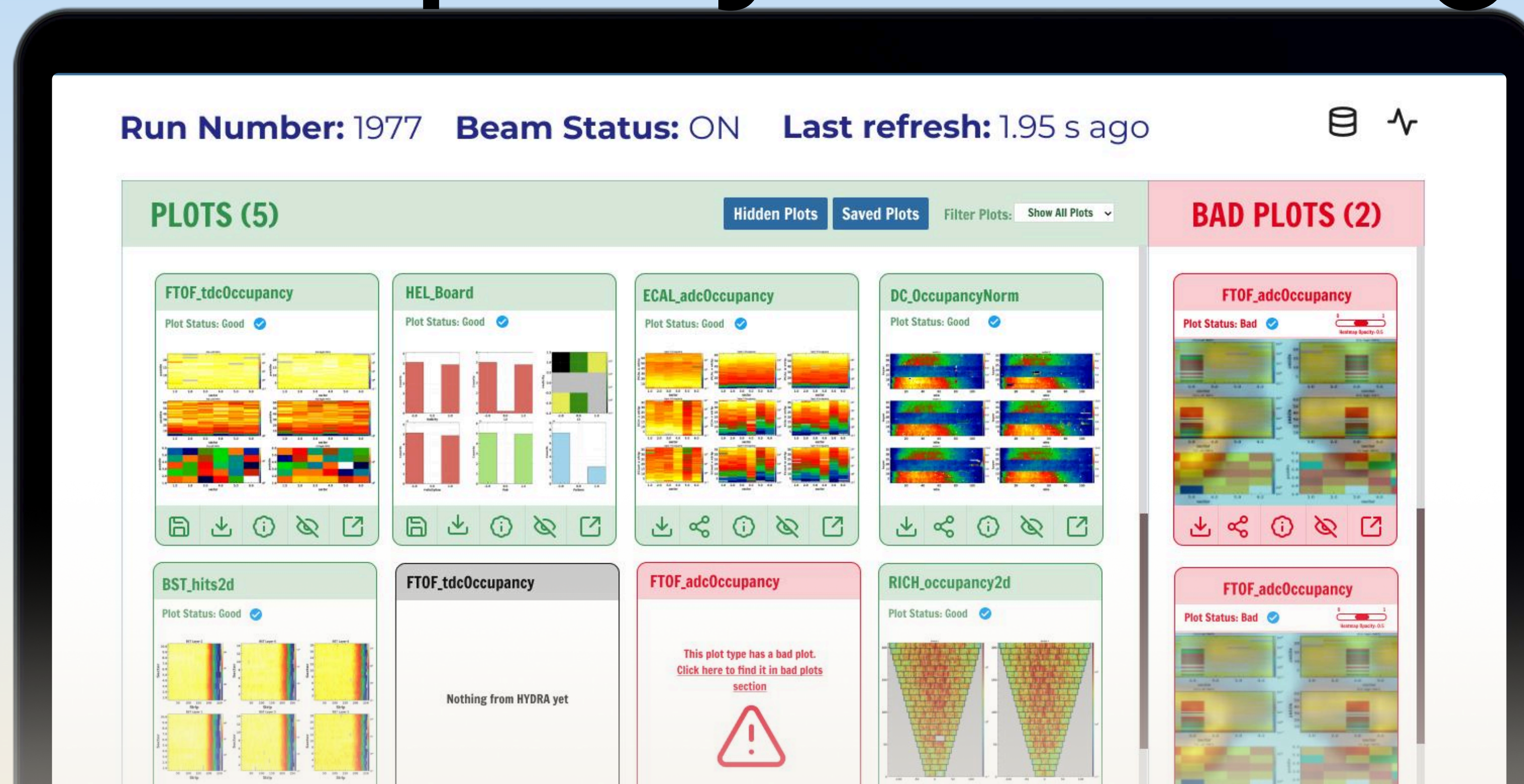
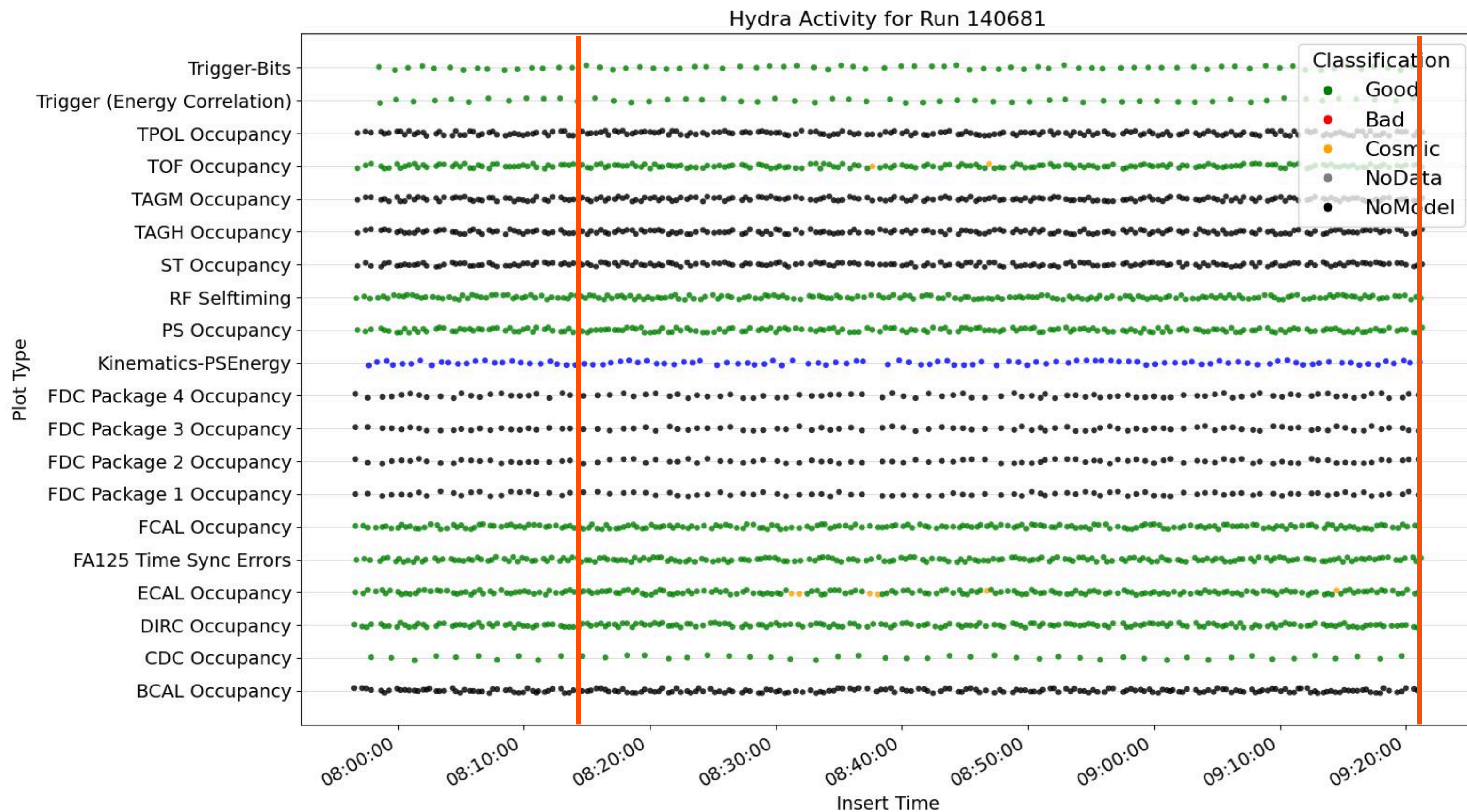


AI for data quality monitoring



The challenge

Human-based data quality monitoring does not scale, yet we do it anyways.



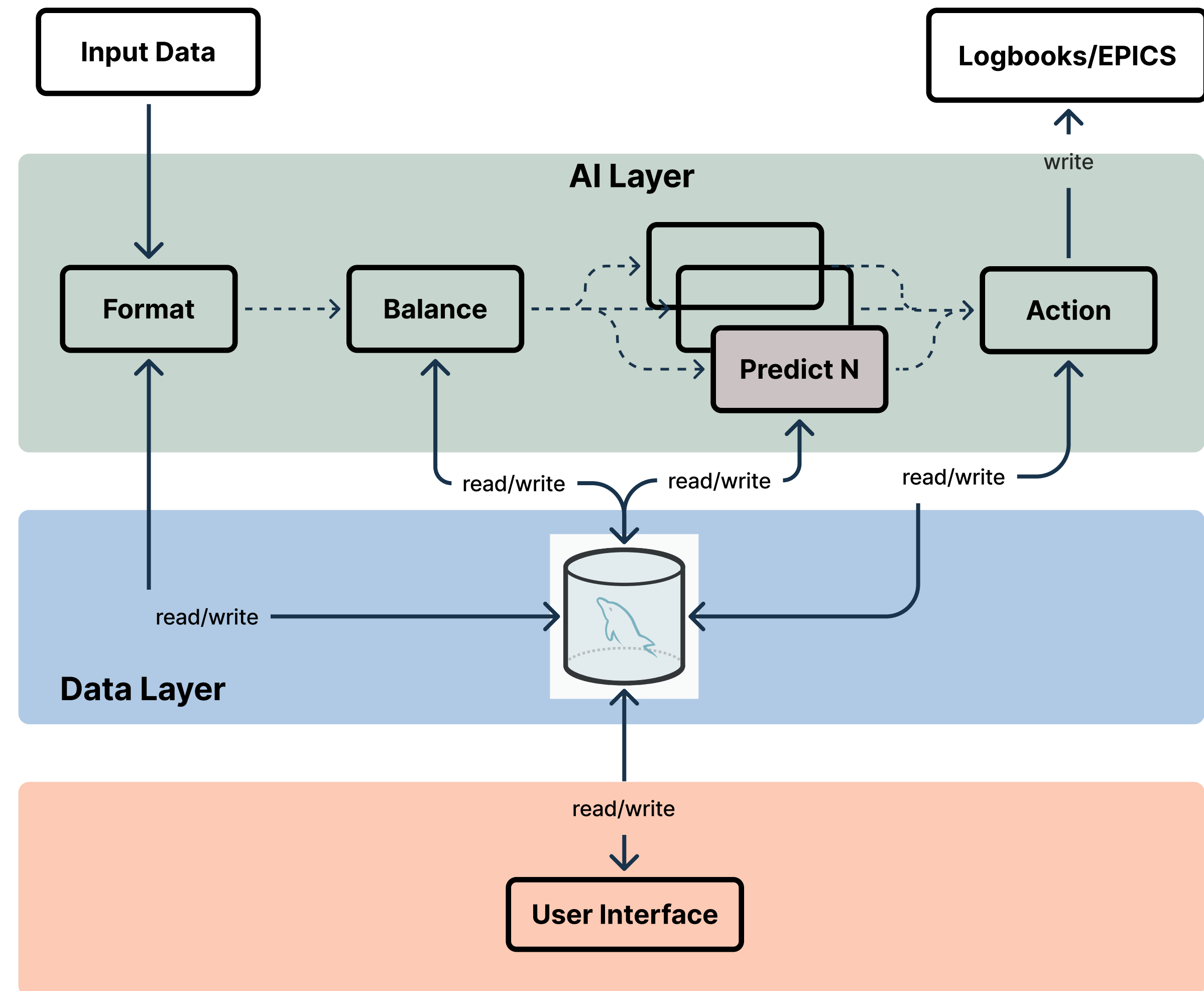
when shift crew posted plots to the logbook



Hydra provides a framework to organize, train models, and act on image-based* data. From anywhere.

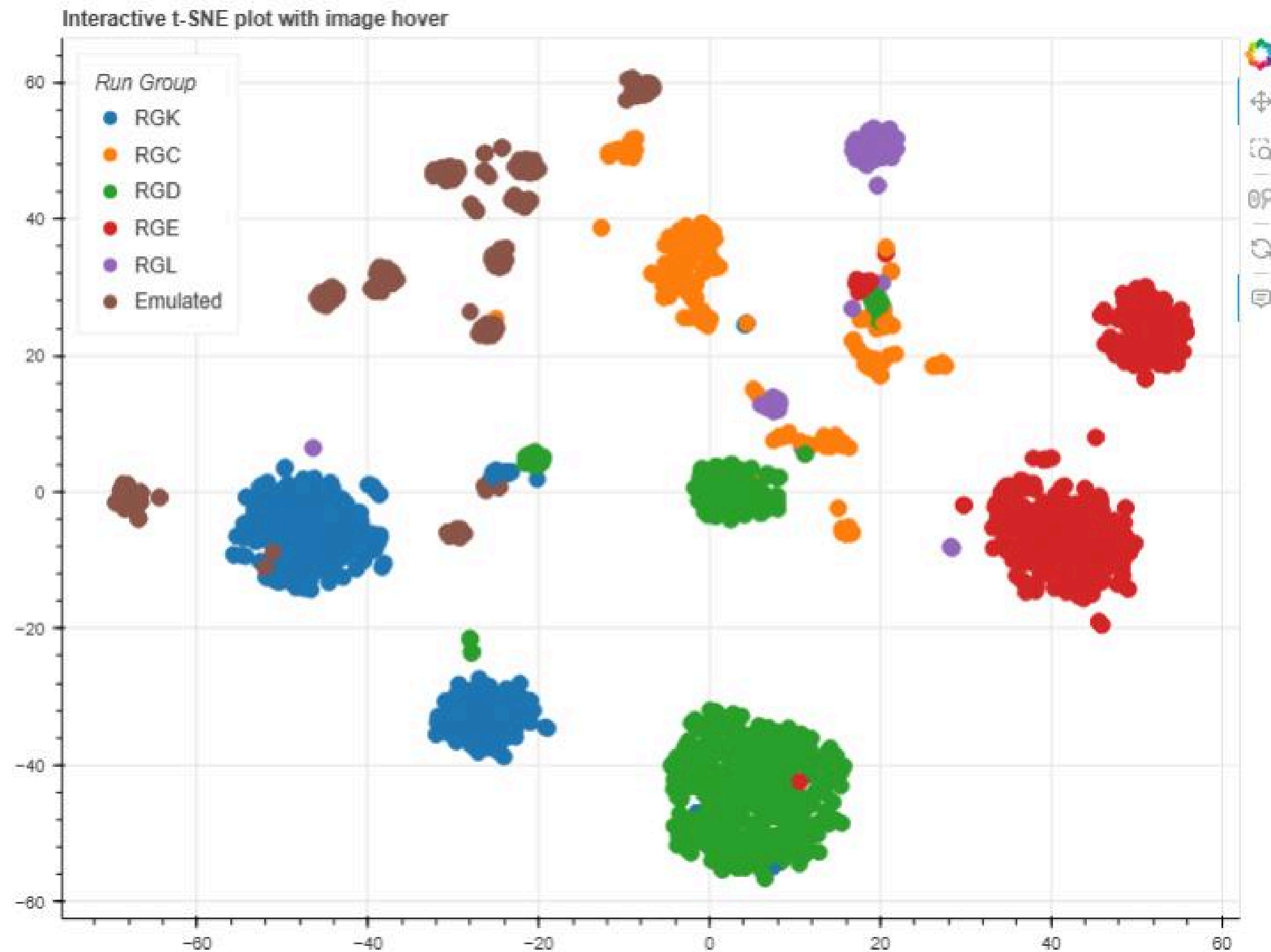
Python, Flask, and React
You can have any model you want, we default to InceptionV3 and VGG16

*for right now



Hydra

Almost* everything you need to organize, train, and act on data



Organize and Label

001

Classify

Identify anomalies with supervised and unsupervised techniques

002

Report

Evaluate and track detector *and* model performance

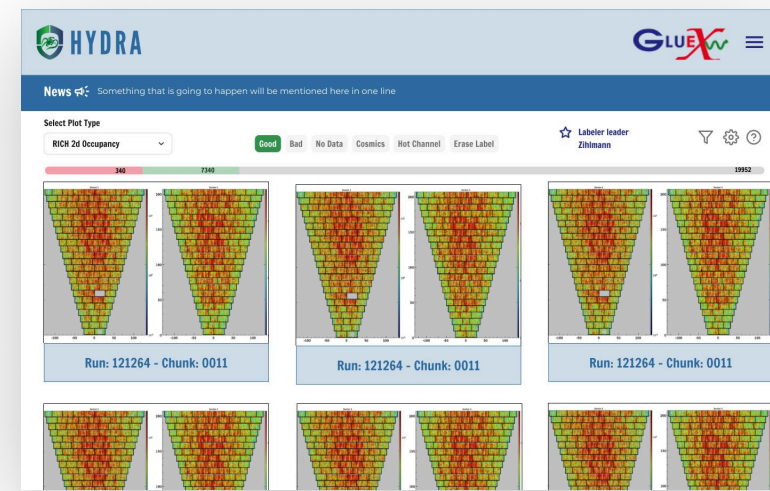
003

Act

Surface insights and operational next steps

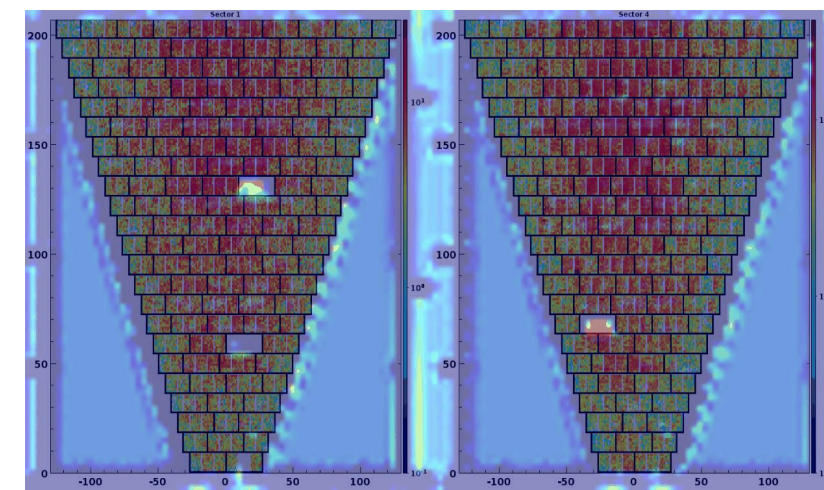
004

Key features



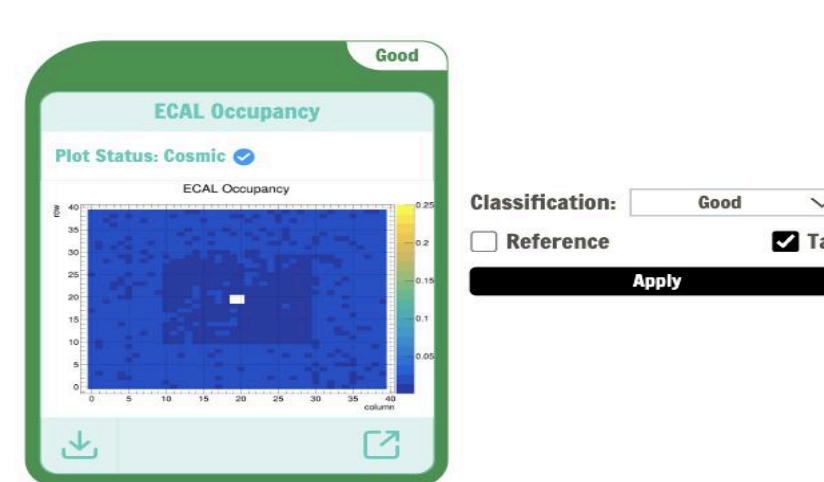
User Interface

Abstract detailed knowledge of AI/ML and SQL away from general end users



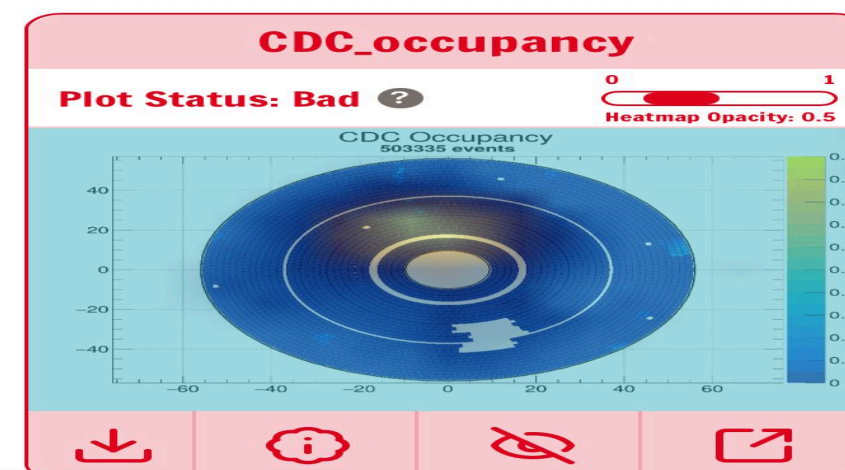
Explainability Techniques

Gradient-weighted Class Activation Maps to indicate important regions of images for classification



Live Labeling

Shift crews can now interact with Hydra to live label, update reference plots, or override classifications



Visual + Audible Alarms

Front end indicators to indicate problematic plots, and EPICS integration for audible alarms to the shift crew

Hydra in GlueX



News

Kinematics-PSEnergy and Trigger-Bits have newly trained models. Their performance is being monitored

Beam Status: OFF | Run Number: 131458 | Last Update: 5/8/2025 2:27:06 PM

1.1 s ago

Hidden Plots

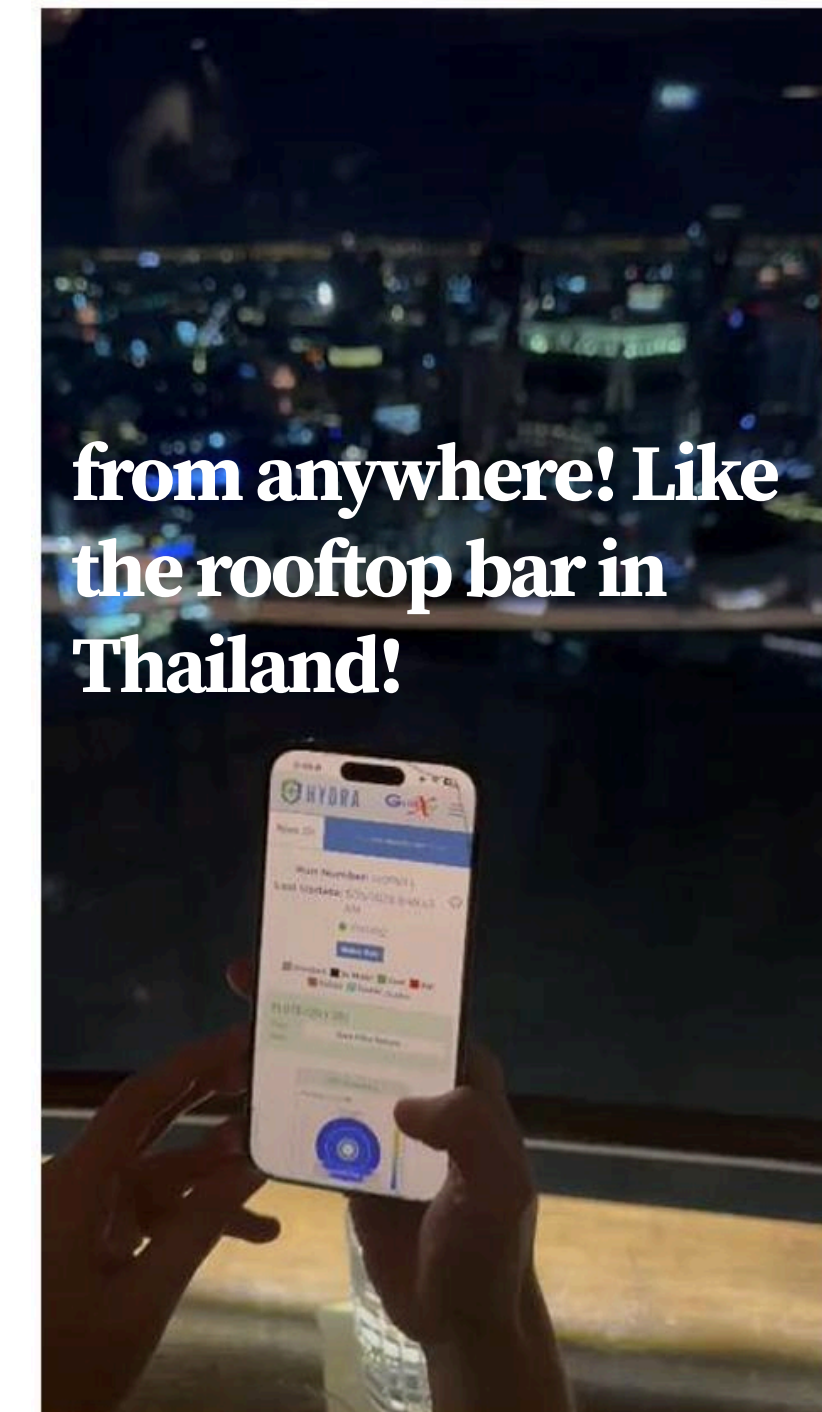
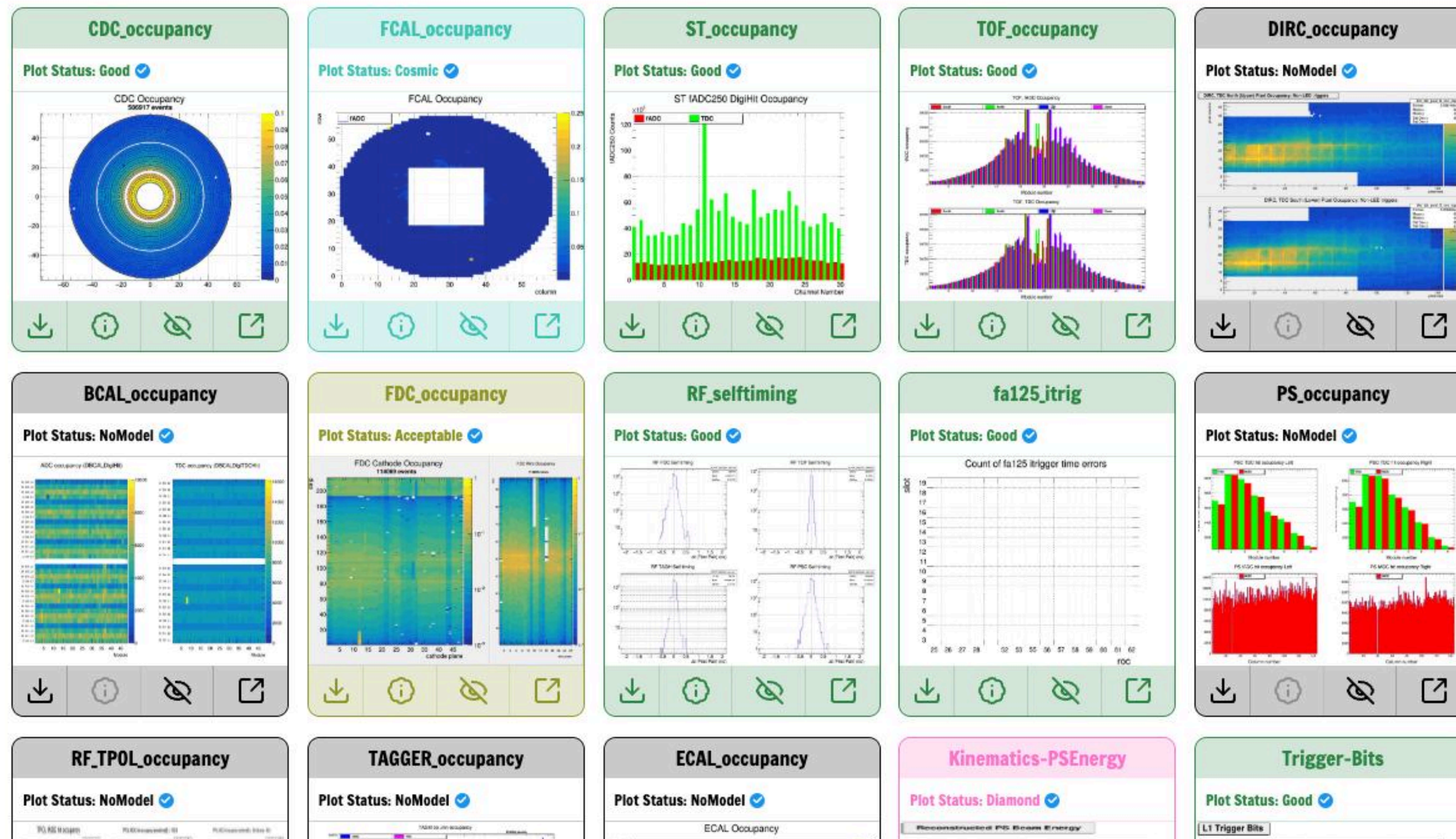
No Model Good Bad NoData Cosmic LED [See More](#)

[See Grafana run history](#)

LOTS (15 / 15)

Filter Plots: [Open Filter Options](#)

BAD PLOTS (0 / 0)



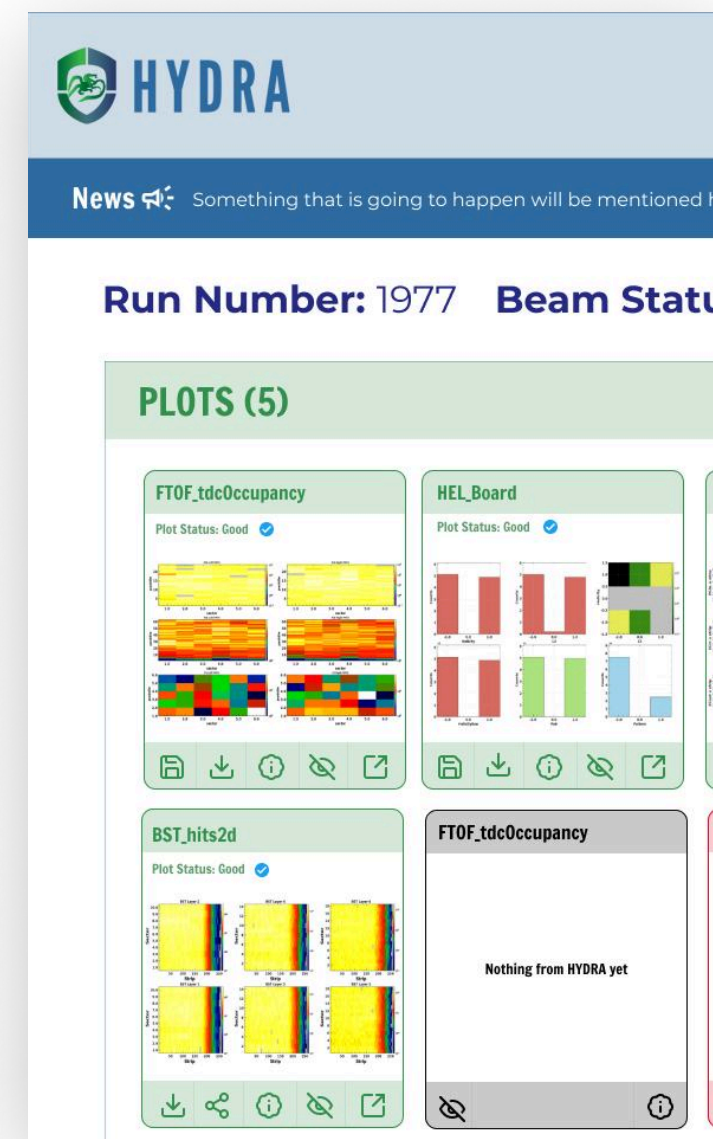
from anywhere! Like the rooftop bar in Thailand!

User Interface



Image Labeler

Efficiently label thousands of images used for training a model.



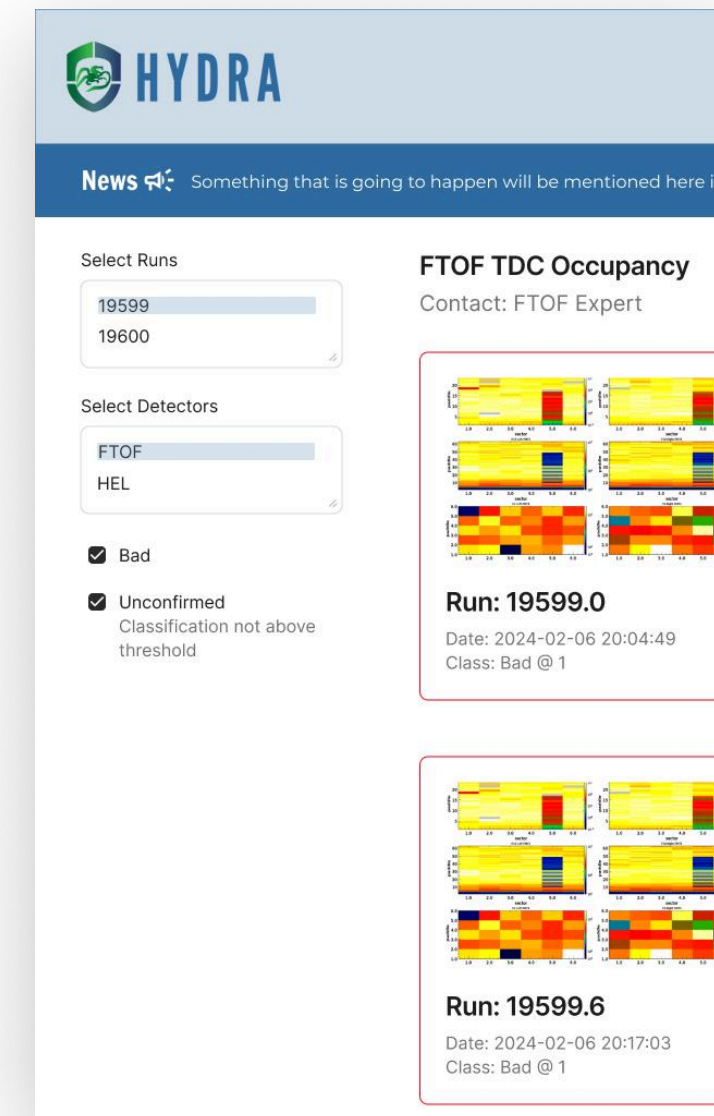
Run

See predictions in near real-time. This page continuously updates with new images during an experiment.



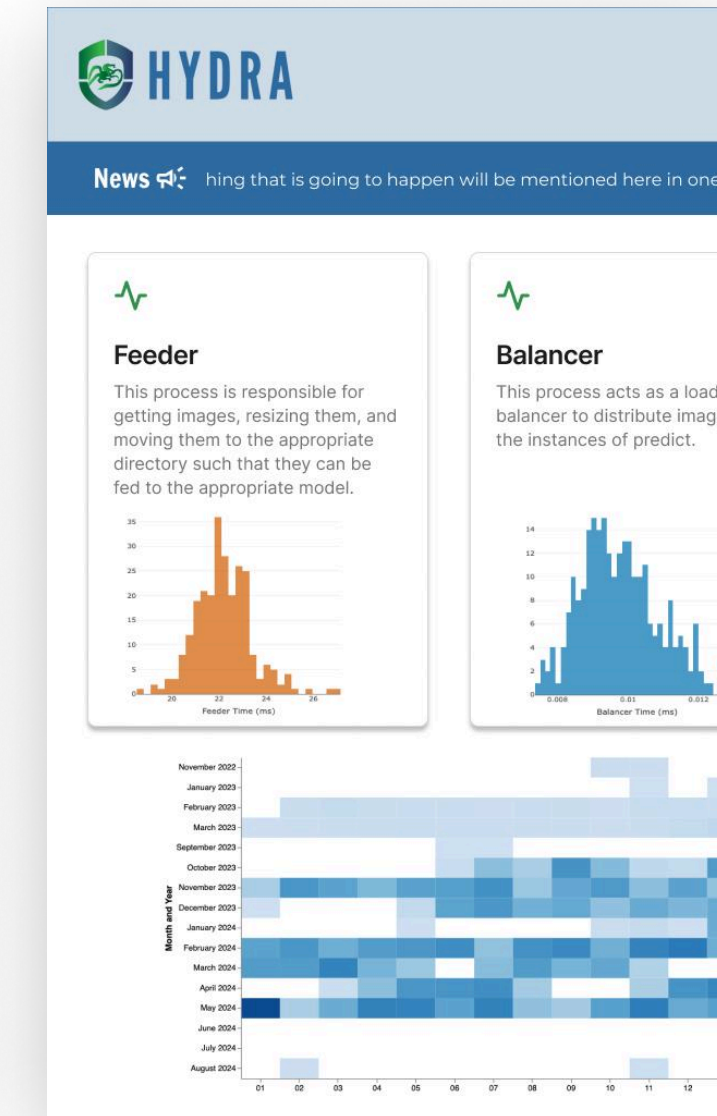
Library

Contains information useful to evaluate a given model's training and performance.



Log

Displays problematic and potentially problematic images from a trailing 24 hour window.



Status

Primarily used by administrators to monitor system performance.



Grafana

Displays all predictions over time. Trend analysis on predictions can indicate when it is time to retrain a model.

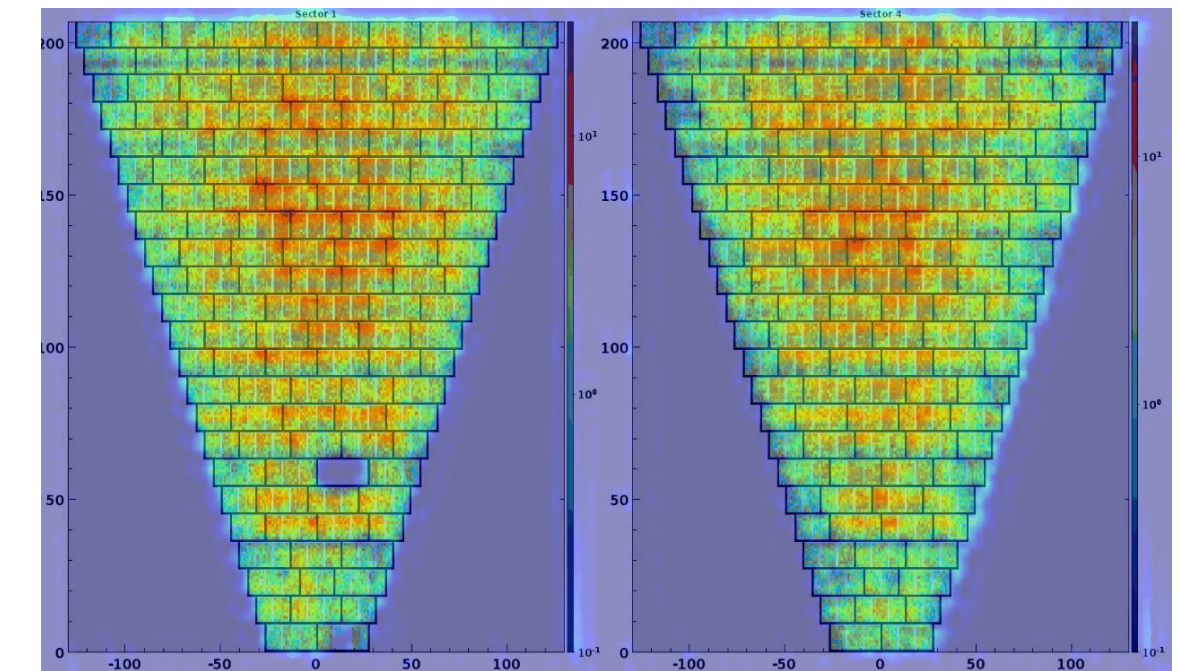
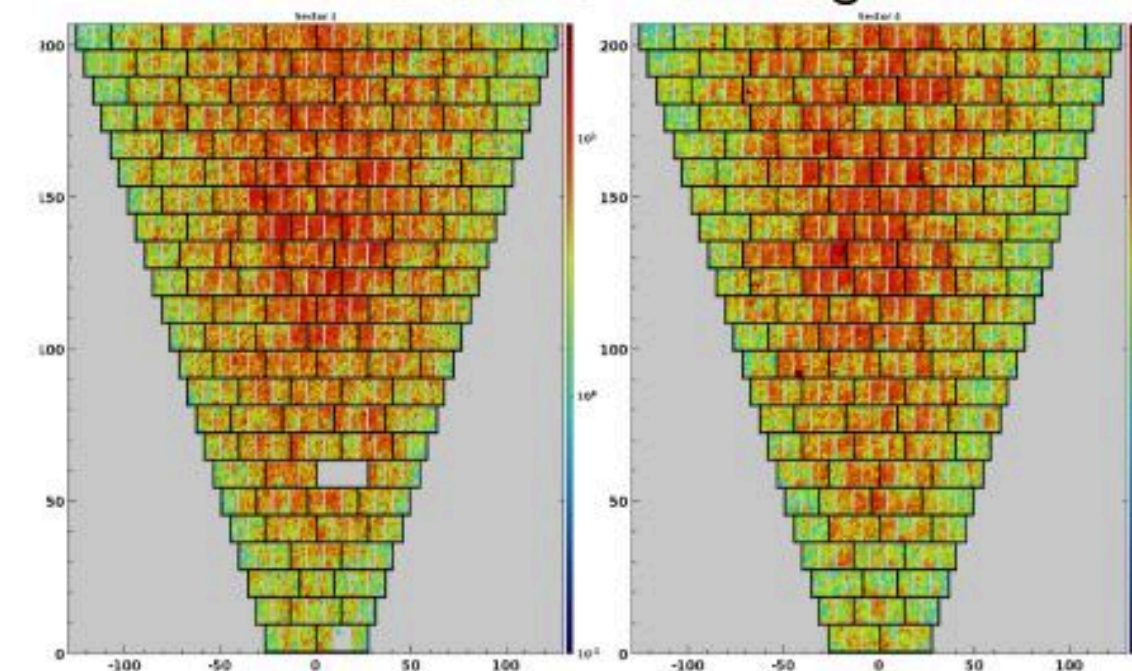
Why is the image bad?

“What about the image does *the model* think is bad?”

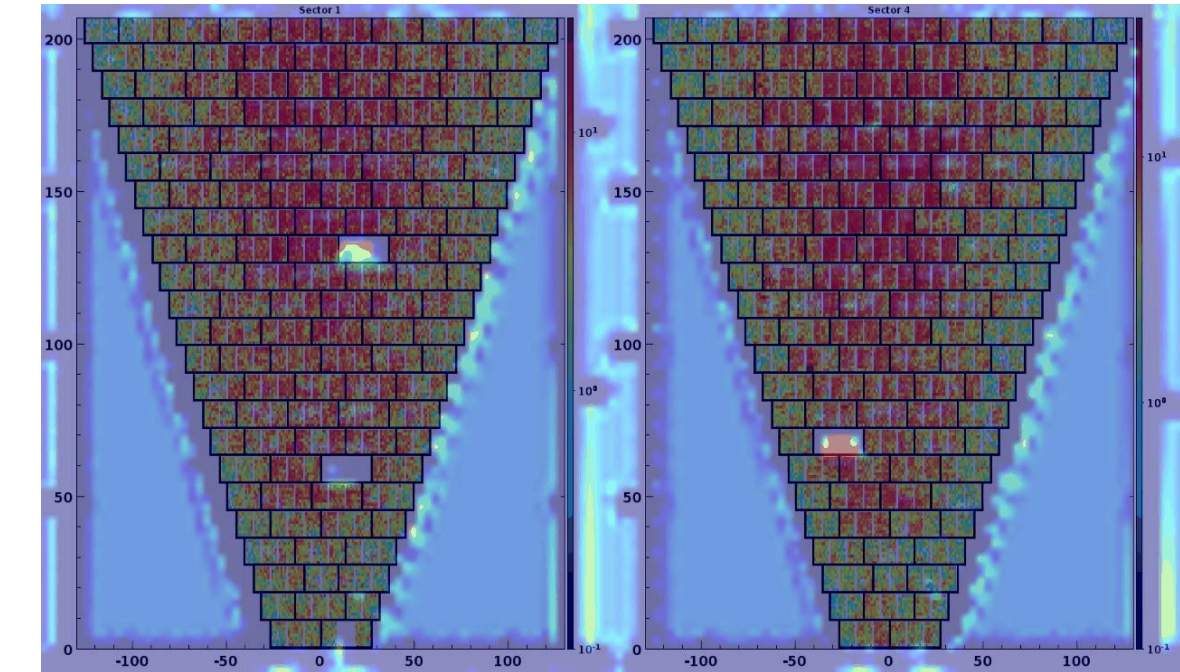
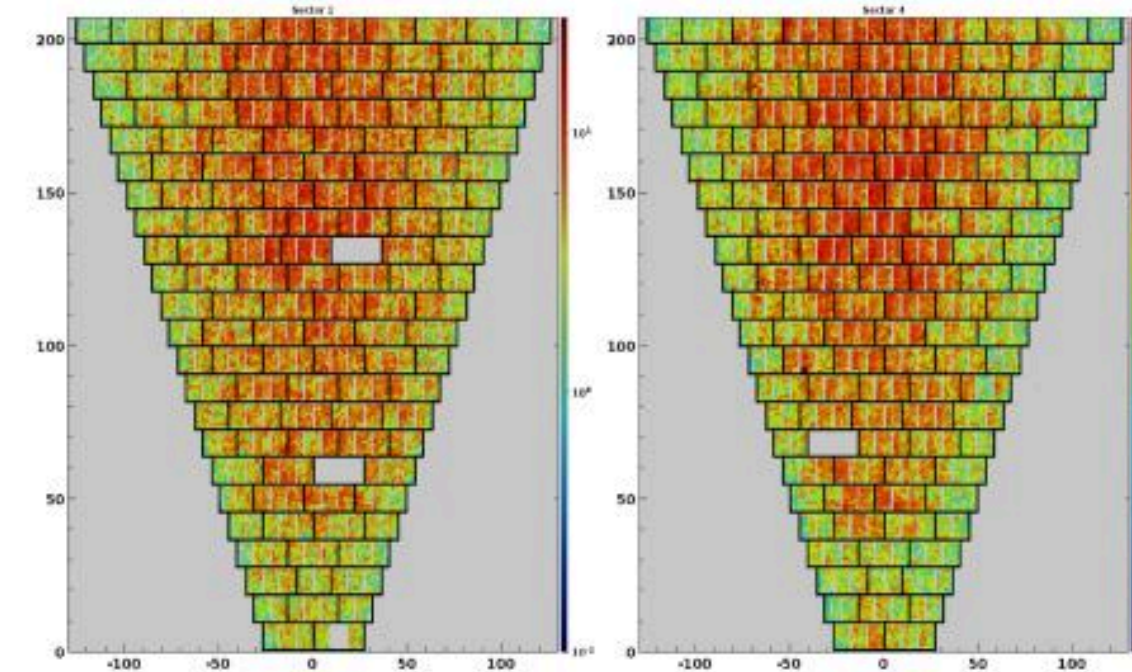
Bright spots indicate important regions of the image for the given classification.

These are very sensitive to how well the model is trained.

this is a normal image

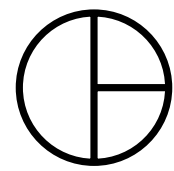


this is a bad image



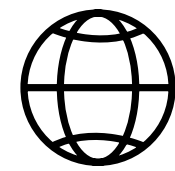
Hydra

Use cases



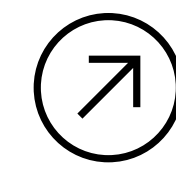
Browse

We help experiments organize and access image-based data from anywhere in the world.



Offline DQM

Identify changes due to calibrations, detailed data quality and run period assessments.

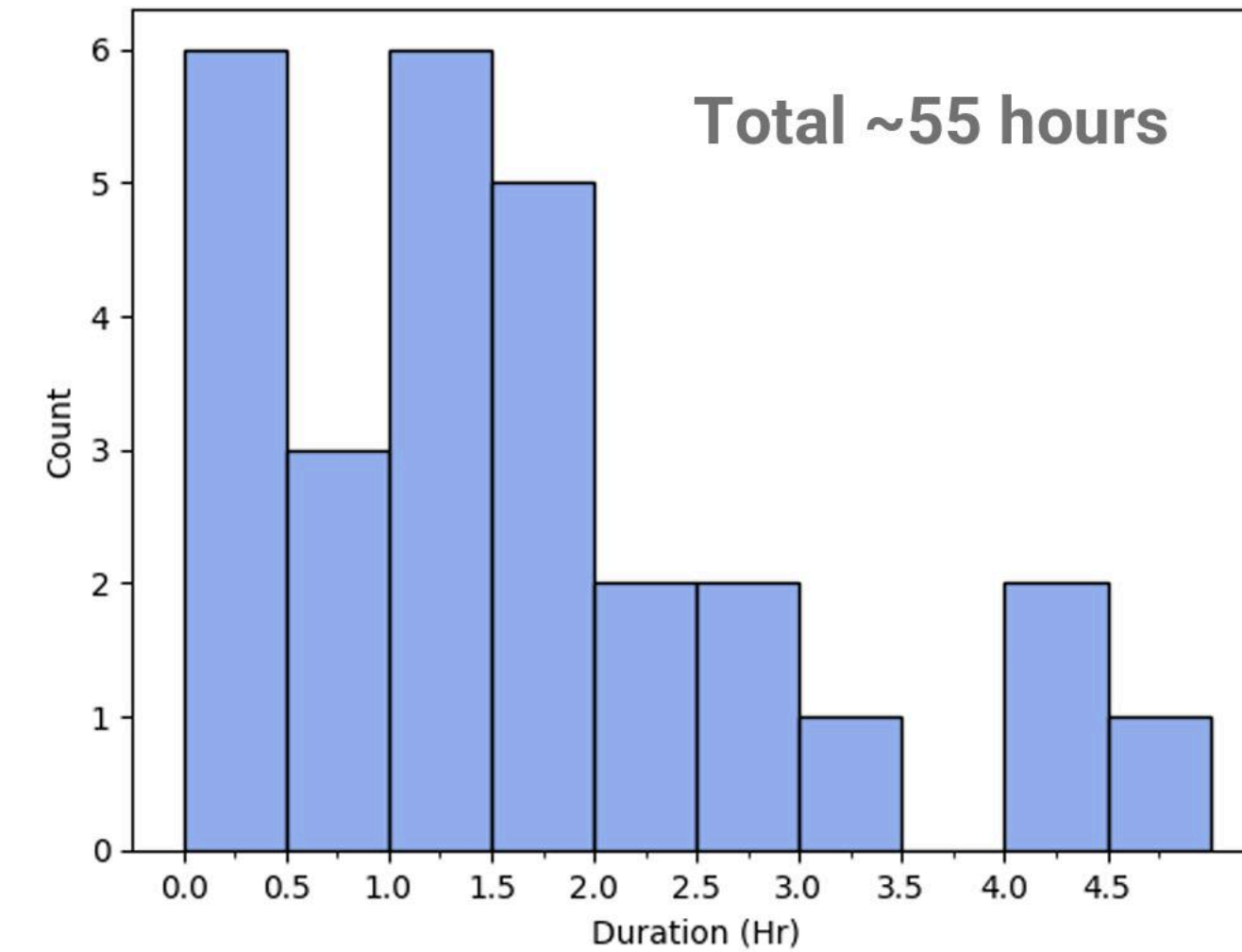


Online DQM

Consistently, frequently, and more accurately monitor incoming data from an experiment

Indirect human performance measurements

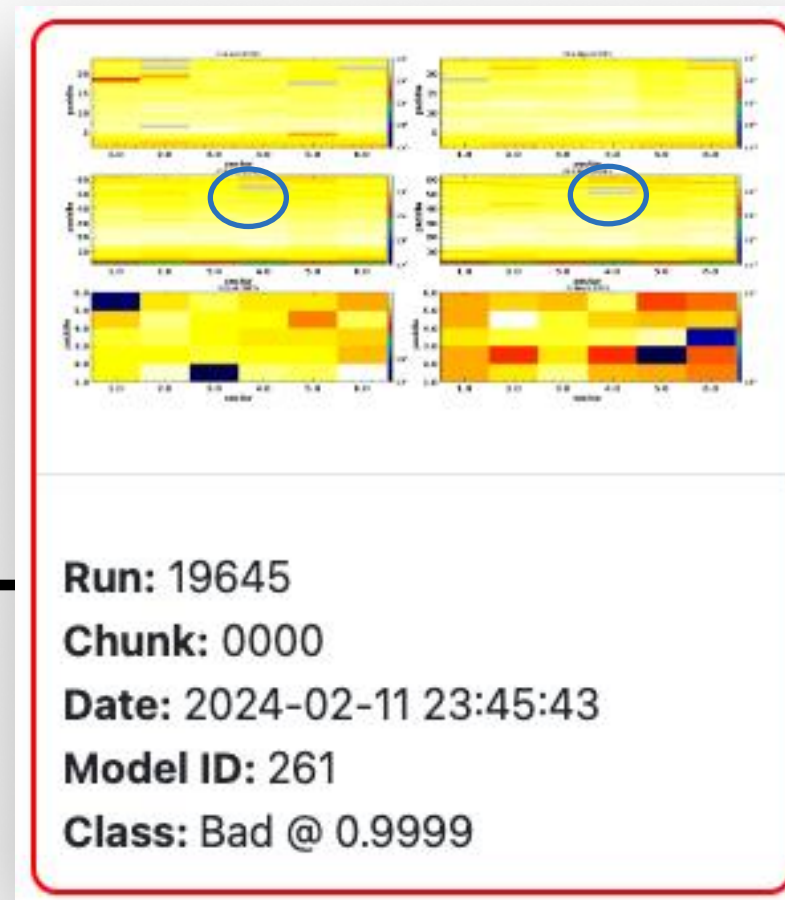
Comparing Hydra's detection time to posts in the logbook, we can indirectly measure how much time is lost due to waiting on humans to discover issues. For one particular run period in Hall-B, this was approximately 55 PAC hours.



Total response time: 7 hours 44 minutes

2/12/2024 00:02:00

Shift crew submits monitoring images to logbook. Problems with detector indicated.



2/11/2024 23:45:43

Hydra alarms for problem with Forward Time of Flight detector

2/12/2024 07:30

Shift crew first becomes aware of problem with detector.

Stats and momentum

Hydra is deployed at Jefferson Lab for the experimental halls, the ePIC collaboration, and the CBM and HADES experiments at GSI.

7+

Deployments across all use cases

>14M

Images in Hydra

>1M

Labeled images across all deployments

>95%

Accuracy on all images for deployed models

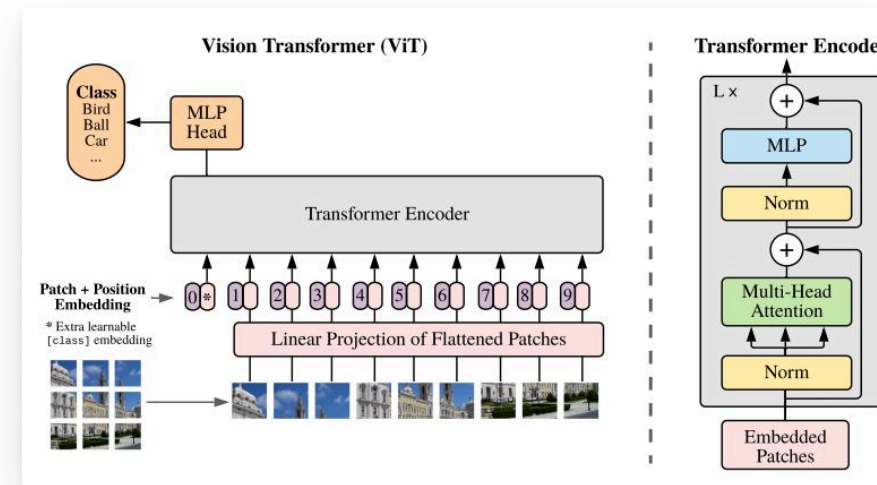
What's next

Development projects



Updated user interface and experience

New front end designs, chat with Hydra



Expanded model capabilities

Autoencoders, Siamese models, vision transformers, and continual learning

Community Engagement

AI-based monitoring is applicable across the DOE complex and industry



Chimera

Knowing what we've learned from 2019 on, how would we rebuild Hydra for the future?

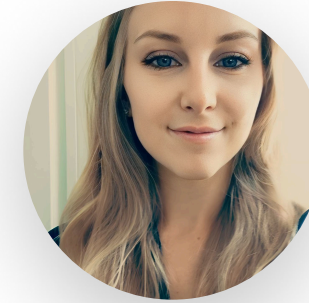
Meet the team



David Lawrence
EPSCI Lead



Thomas Britton
EPSCI



Torri Jeske
EPSCI



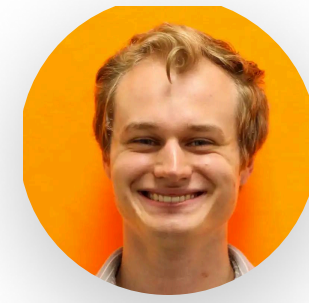
Nataliia Matsiuk
Information + Records



Raiqa Rasool
EPSCI



Darren Upton
ODU Physics Grad Student
Hall B Service Work



Jordan O’Kronley
UT Physics Graduate Student
Hall C Service Work

Join Hydra

We really want to collaborate and not reinvent wheels.

▪ code.jlab.org/epsci/hydra

▪ hydrateam@jlab.org

The screenshot displays the HYDRA monitoring interface. At the top, the HYDRA logo is on the left, and a 'News' button is in the center. On the right, a status message reads: 'Kinematics-PSEnergy and Trigger-Bits have newly trained models. Their performance...'. Below this, the current system status is shown: 'Beam Status: OFF | Run Number: 131458 | Last Update: 5/8/2025 2:27:06 PM'. A legend identifies plot statuses: No Model (black), Good (green), Bad (red), NoData (brown), Cosmic (cyan), and LED (purple). The main area is titled 'PLOTS (15 / 15)' and contains a grid of 15 individual plot cards. Each card shows a specific metric with its plot status and a set of control icons (download, info, refresh, zoom, share). The plots include: CDC_occupancy (Good, circular heatmap), FCAL_occupancy (Cosmic, circular heatmap), ST_occupancy (Good, bar chart), TOF_occupancy (Good, histogram), BCAL_occupancy (NoModel, two heatmaps), FDC_occupancy (Acceptable, two heatmaps), RF_selftiming (Good, four histograms), fa125_itrigr (Good, line plot), RF_TPOL_occupancy (NoModel), TAGGER_occupancy (NoModel), ECAL_occupancy (NoModel), and Kinematics-PSEnergy (Diamond, histogram).

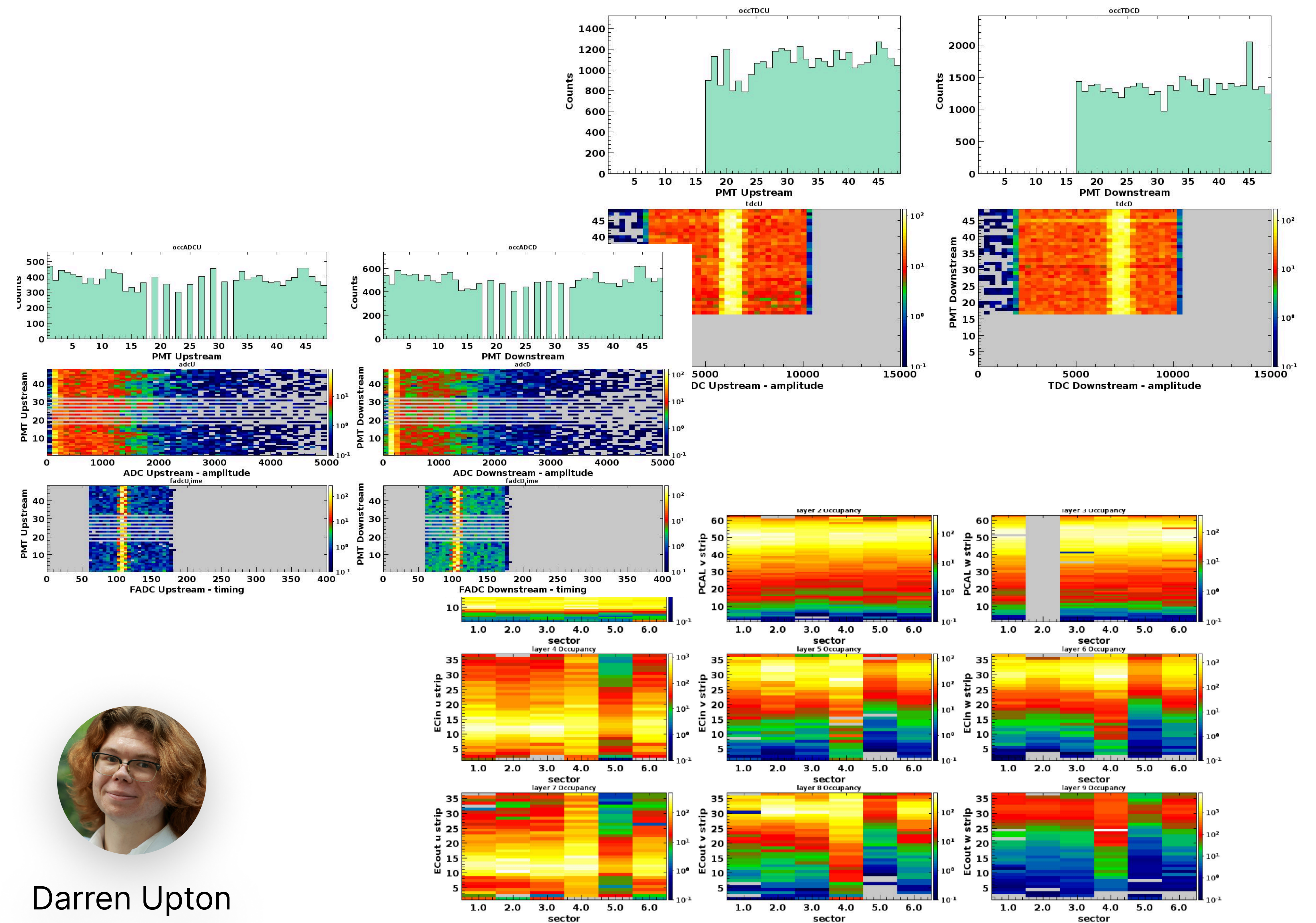
Summary Statistics

	Hall D	Hall B	Hall A	Hall C	FAIR/GSI	Total
Year deployed	2019	2022	2024	2024	2026	
Labeled Images / Total Images	637,885 / 12,243,320	352,804 / 747,554	5440 / 83,376	7800 / 465,853	2002 / 2002	1,005,949 / 13,542,105
Active Models	19	61	-	6	-	86
Batch / frequency	16 / min	81 / 3-5min	42 / run	92 / run	-	-
Runs with 1 bad image	1,822	1,081	-	33	-	-
Total classified	2,834,043*	2,324,870	-	2475	-	5,161,388

*since record keeping began

What if we don't have any "Bad" images?

Our graduate students have developed software to emulate various failure modes in real data by modifying databases and regenerating the images.



Darren Upton

We can simulate bad plots to use in training.

Siamese Models

Siamese models learn to compare two inputs (in our case images) and determine how *similar* they are.

Think of this as Hydra's equivalent to the shift crew comparing images to their reference.



Jordan O'Kronley

UT Physics Graduate Student
Image clustering and siamese models

