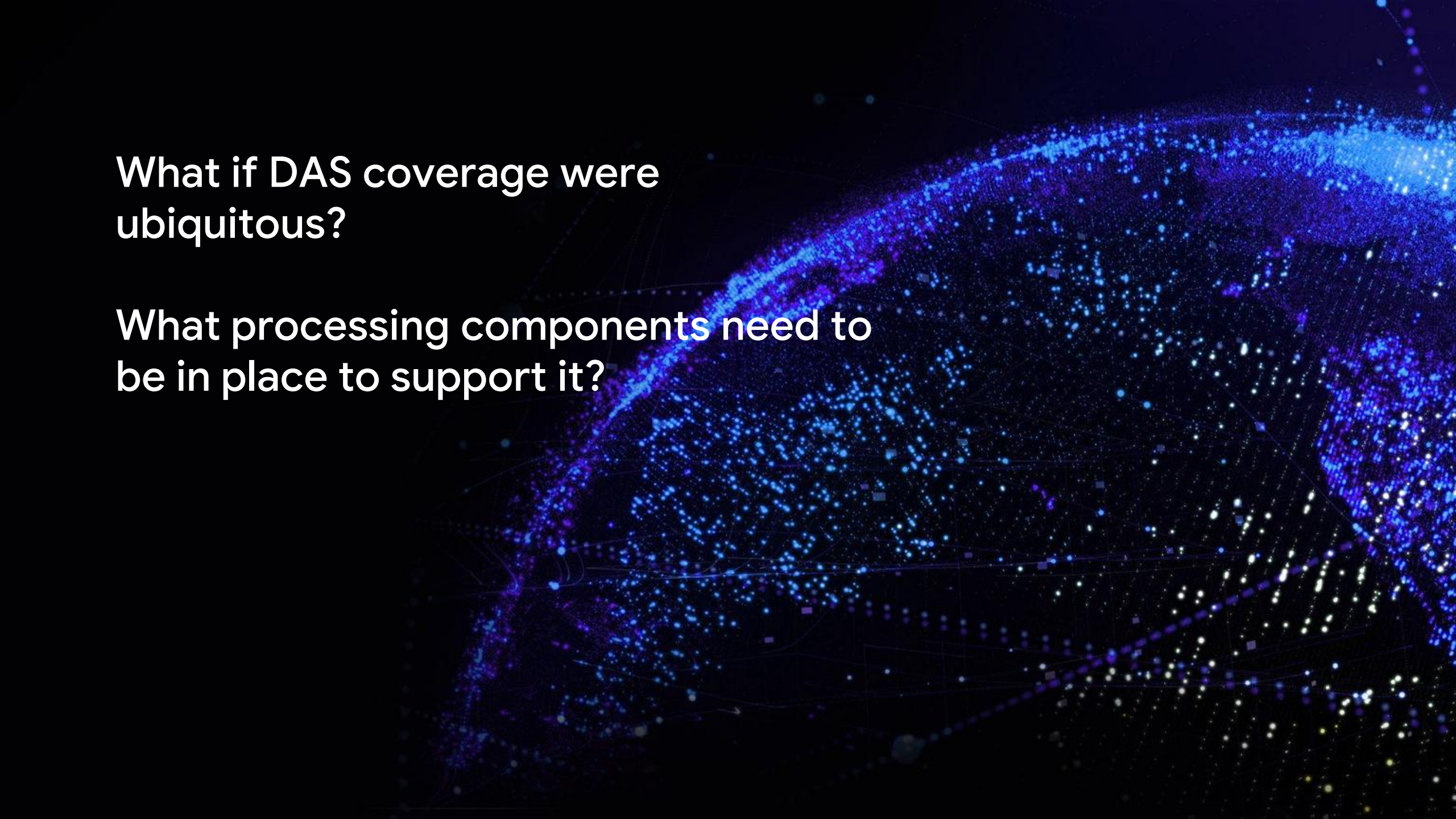

Ubiquitous DAS
Bob Clapp


What if DAS coverage were ubiquitous?
What processing components need to be in place to support it?

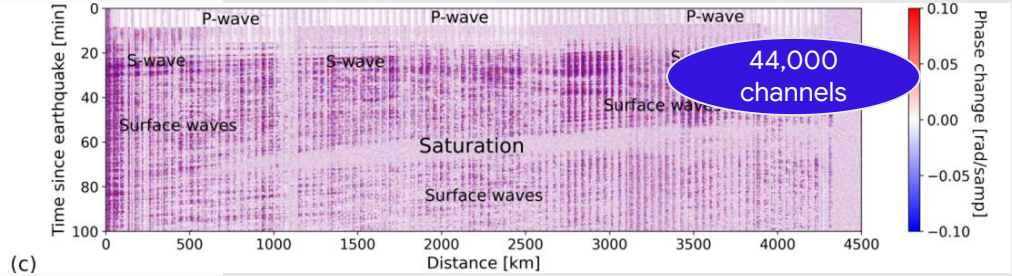
Ingredients for ubiquity are being worked on

- Price of interrogators is (will be) more affordable with use of less expensive components and volumes. Especially if fit for purpose trade offs are applied.

- Compute price is dropping and hardware footprint is reducing.

- Interrogator specs, e.g. range, noise floor, .. are improving.

terra15.com.au

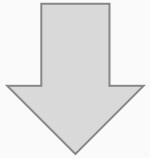


M. Mazur et al., “Submarine Cable Deep-Ocean Observation of Mega-Thrust Earthquake and Tsunami with 44,000 100-m Spaced Sensors,” ArXiv e-prints, Art. No. 2509.24813, Sep. 2025.

Ubiquity leads to massive volumes of data generated

In the US, 9.3M miles of fiber to are expected to be available by 2028*

Assume: 1000 Hz acquisition, 5m spacing, 2 byte precision



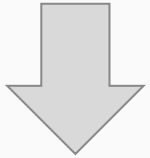
517,000 TB / day

* Benton Institute for Broadband and Society

Ubiquity leads to massive volumes of data generated

In the US, 9.3M miles of fiber to are expected to be available by 2028*

Assume: 1000 Hz acquisition, 5m spacing, 2 byte precision



517,000 TB / day

Solutions that need to be built

1. Machine learning
1. Edge processing
1. Compression

* Benton Institute for Broadband and Society

Drowning in data. Thirsty for insights. Solution: ML

Noisy data problem: **Terabytes / day** from a single cable, but mostly “noise”.

Requirements:

- Balance sensitivity and specificity - lots of false triggers make manual reviews infeasible
- Quickly adaptable to new sites and new noise regimes
- Real-time or near real-time detection

Operational constraints:

- Equipment cost
- Data storage cost
- Bandwidth limitations, especially for remote sites
- Space and power budget limitations on a platform

3%

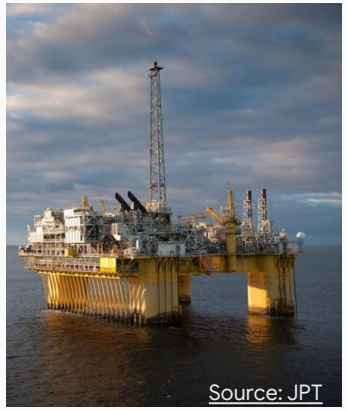
In a typical site, less than 3% of data contains a useful signal.

Real-time edge microseismic monitoring system

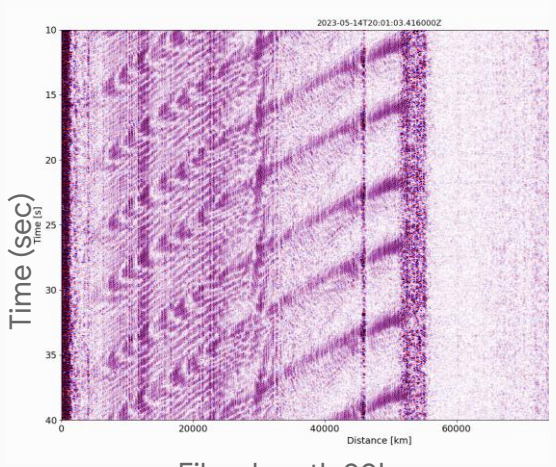
TotalEnergies

An operational solution has been developed to manage the huge DAS data flow

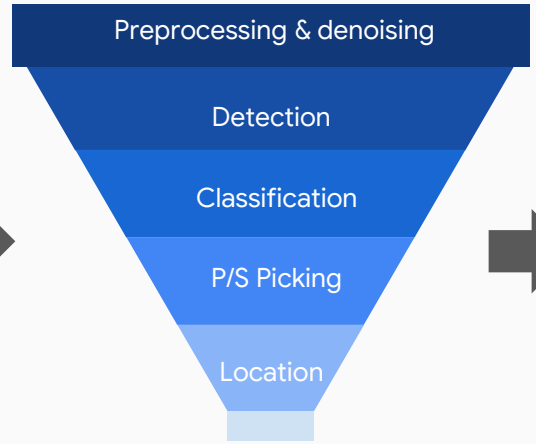
On the platform (edge processing)

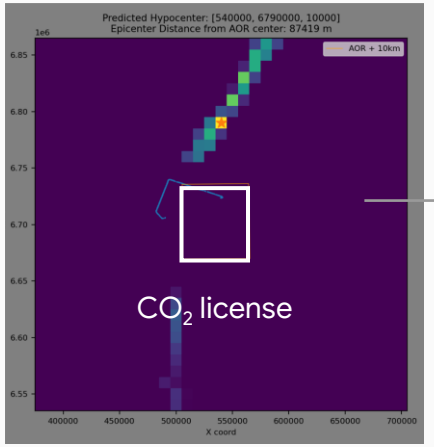


Seabed fiber optic cable



Fiber length 90km





Outside License

Inside License

Offshore DAS Acquisition

Continuous Datastream

~1.7 TB/day

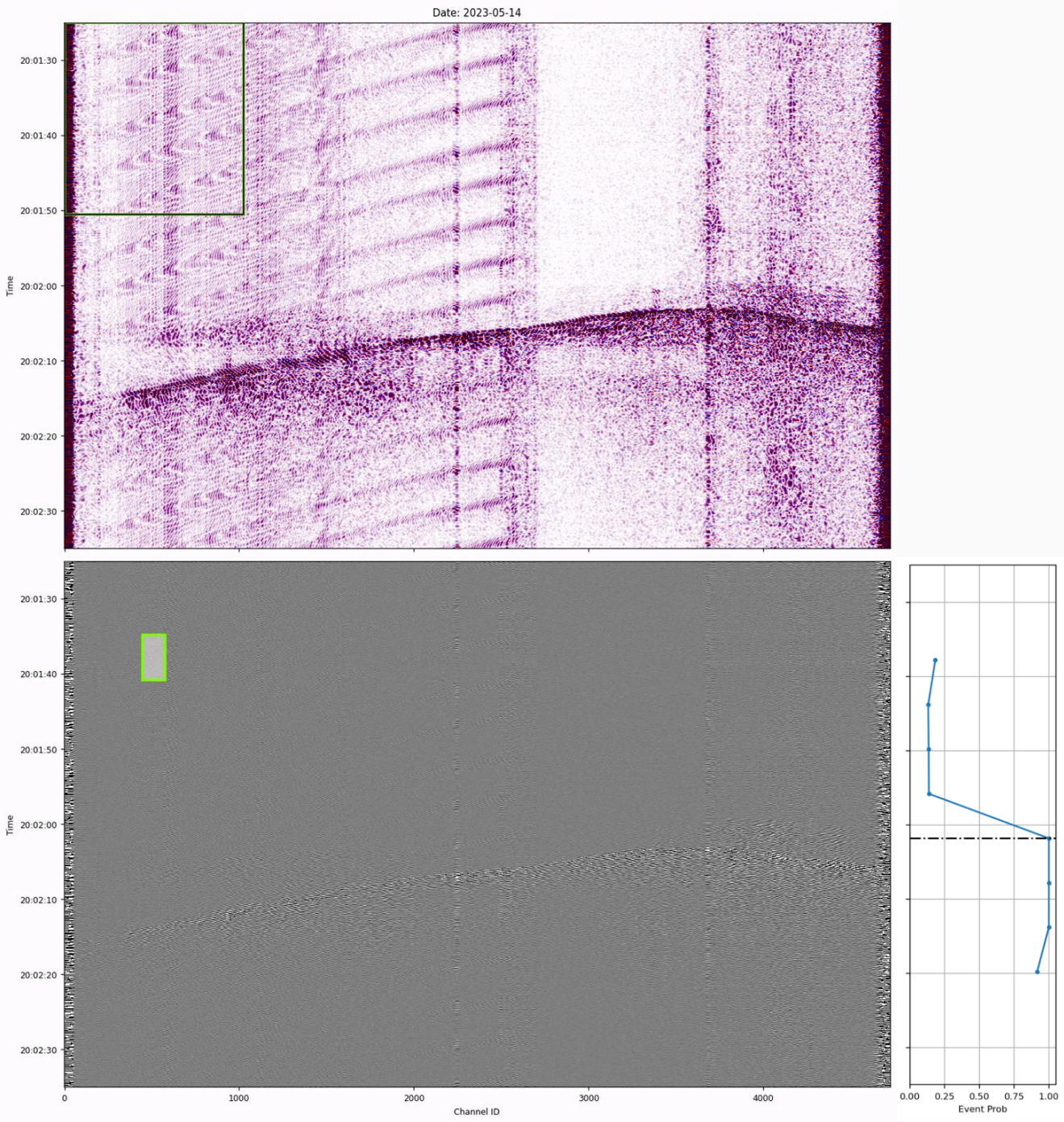
Automatic AI Detection & Location of Earthquakes

~52 GB/day

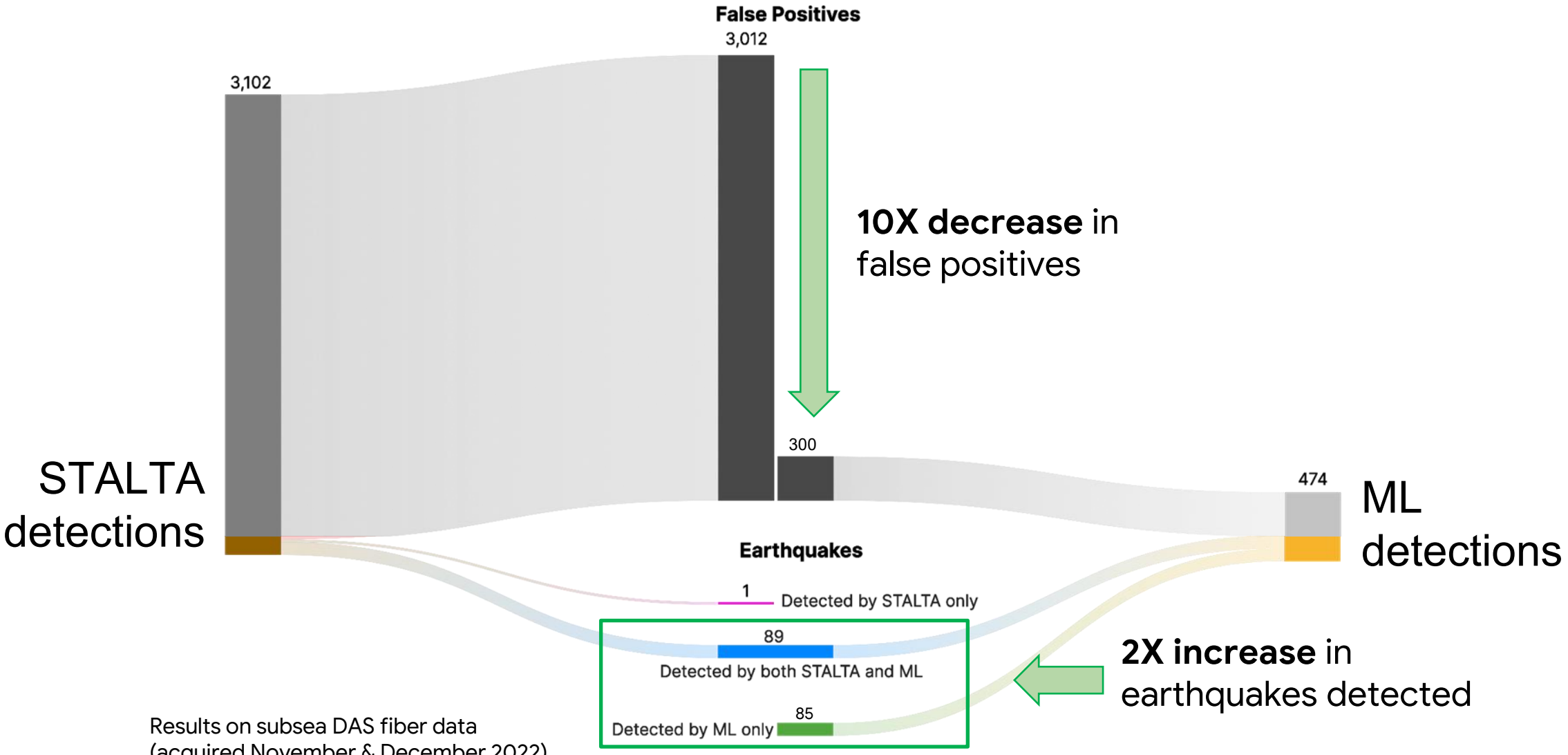
Validation of Localized Events

ML MODEL INFERENCE IN ACTION

- Two-minute time windows are divided into patches
- Each patch is evaluated by the ML model (earthquake, anomaly, background)
- Post-process the detection patches
- Derive detection curve from the patches



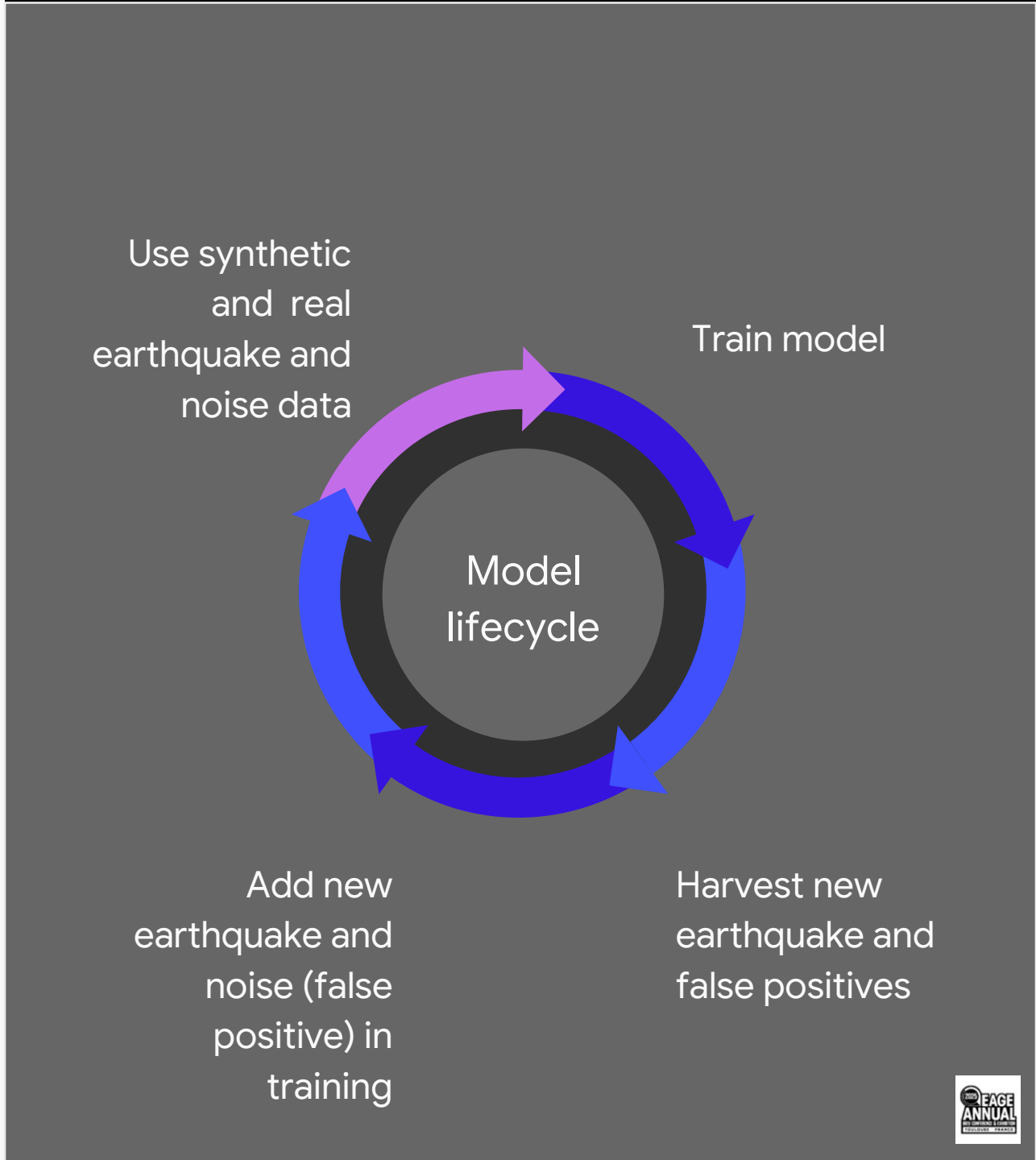
Compared to STALTA, the ML system has 10X fewer false detections while identifying 2X more earthquakes



Results on subsea DAS fiber data
(acquired November & December 2022)

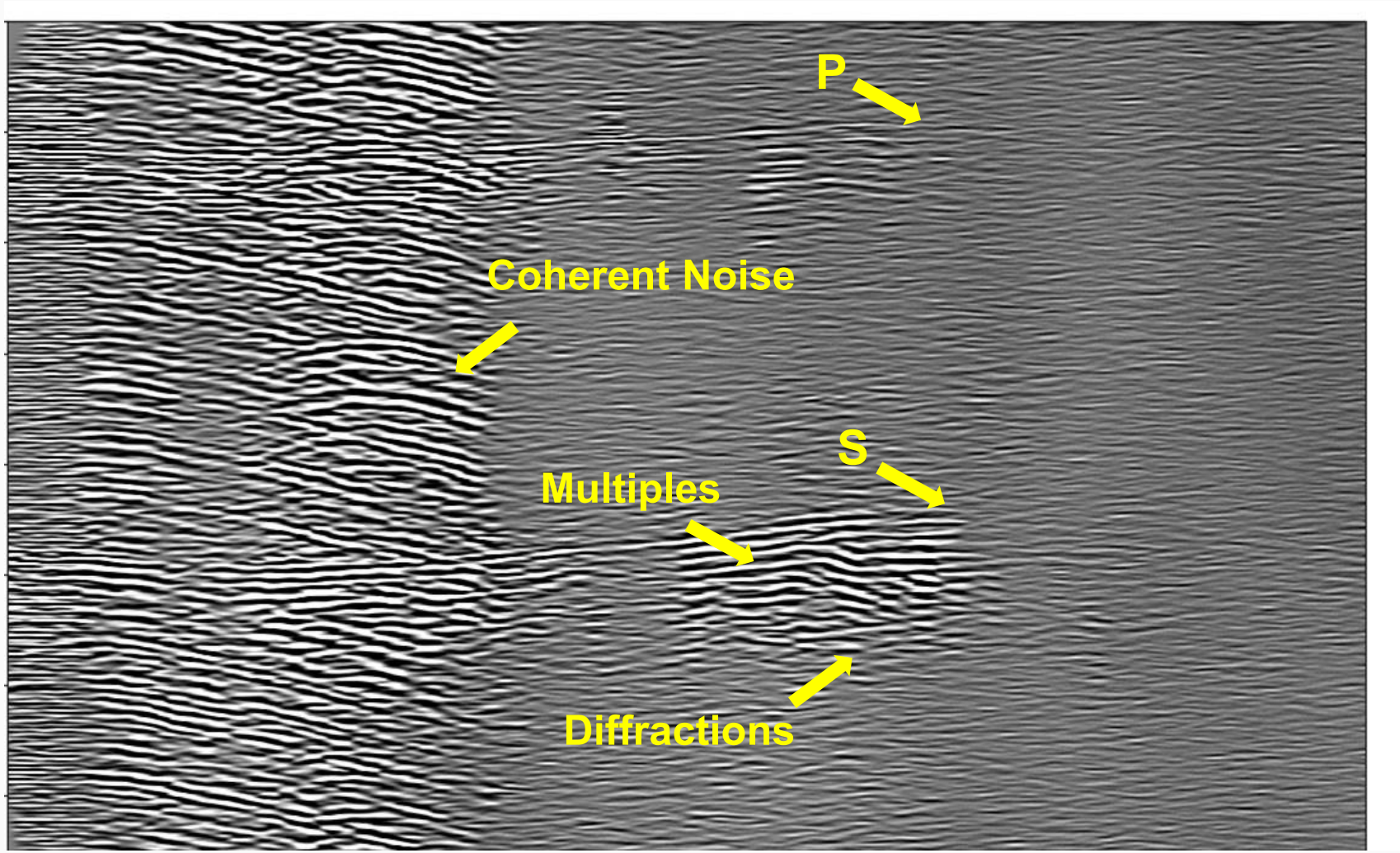
ML Flywheel: Infrastructure to wrangle terabytes of data is imperative

- Manual and ML-assisted labeling tools to facilitate initial training / active learning
- Synthetic data generation to bootstrap detector / PS picker training and augmentation
- Ability to run models on large amounts of continuous data, not just the existing event catalog (if it exists) to assess performance and harvest fainter events
- ML experiment tracking
- Hyperparameter tuning
- Parallelization and autoscaling



SYNTHETICS AS A DATA BOOTSTRAP AND MODEL PERFORMANCE BOOSTER

- Cheap to generate thousands of examples.
- Bootstrapping ML models when real events are scarce
- Augment real events to cover broader magnitude and hypocenter ranges.
- Synthetics need to look realistic. Otherwise ML models don’t work.




Thank you
X, The
Moonshot
Factory