ADC-B3

How CattleEye leverage AI/ML on AWS to make farming more sustainable

Jonah Craig

Startups Solutions Architect
Amazon Web Services (AWS)

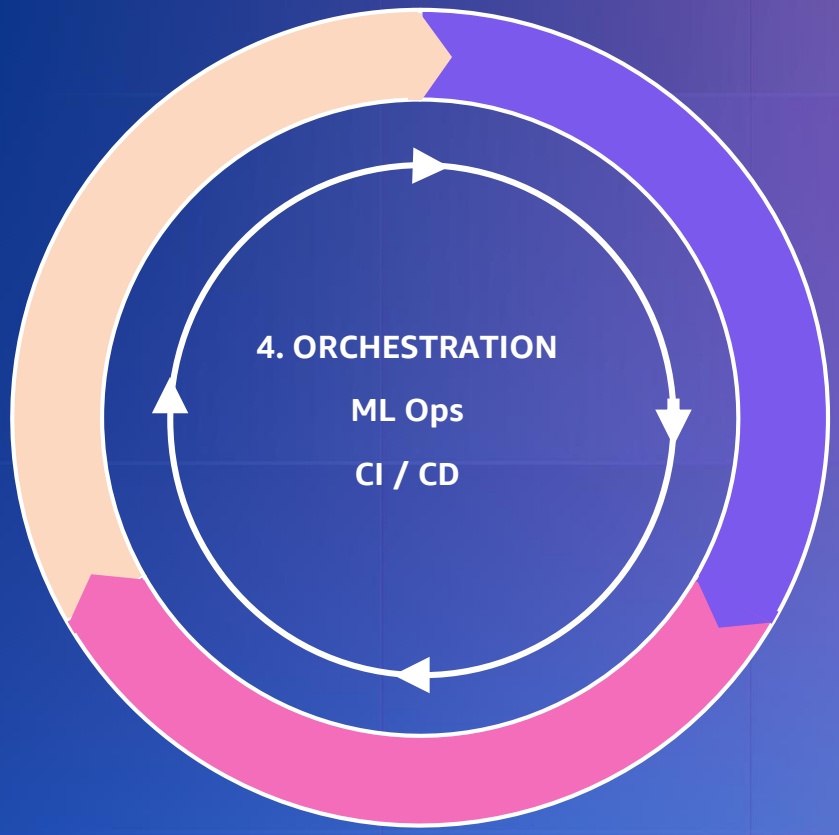
Adam Askew

Chief Technology Officer
CattleEye

The end-to-end ML Process

The Machine Learning loop

1: Data Preparation



2: Model Training

3: Model Deployment & Monitoring

CattleEye high-level ML challenges

Quick Iteration

- By leveraging capabilities of SageMaker, CattleEye can easily create and adopt new approaches to their ML workloads staying on the cutting edge

Scaling Globally

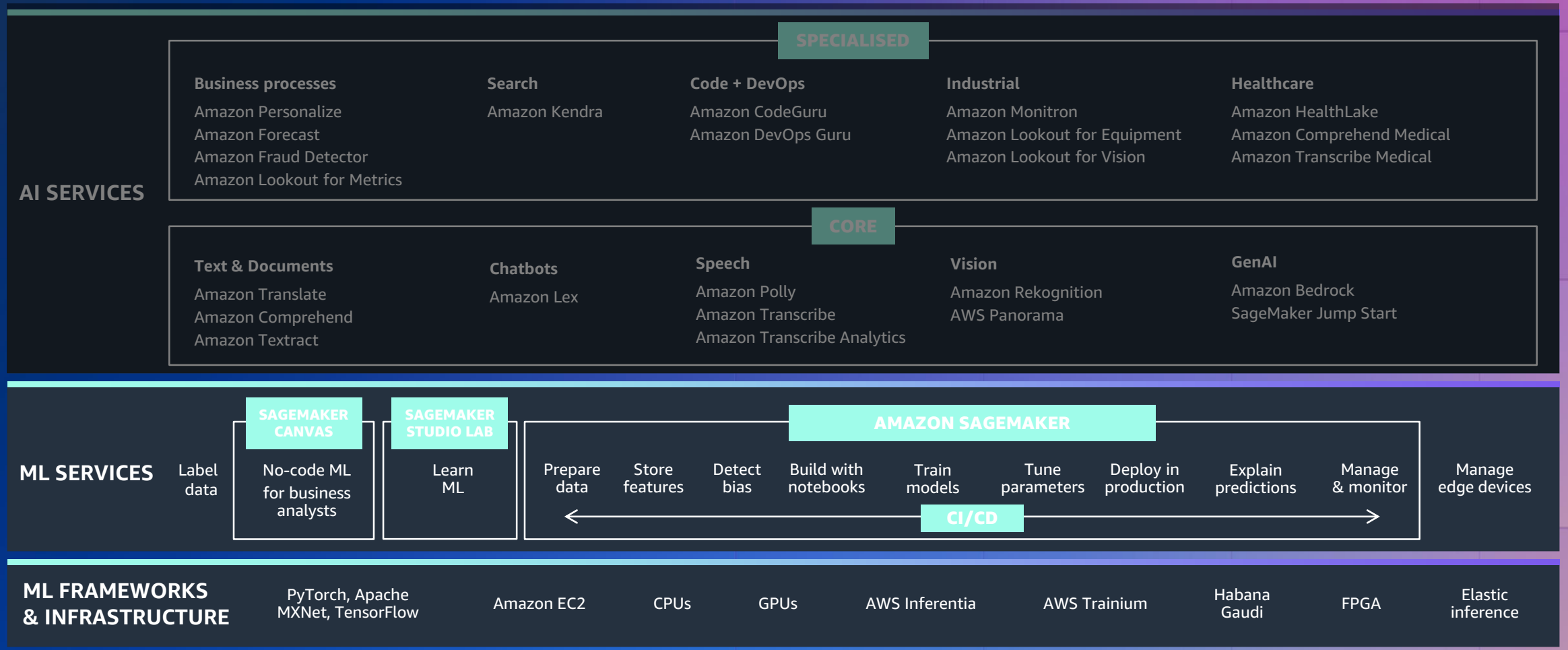
- CattleEye work with farms across the globe
- AWS provides global infrastructure for ML workloads allowing Startups to scale into new regions

Cost of Delivery

- Leveraging Batch Inference on SageMaker
- CPU vs GPU optimisation

The AWS AI/ML stack

Broadest and most complete set of machine learning capabilities



What is Amazon SageMaker?

Amazon SageMaker

Purpose-built tools so you can be 10x more productive

Amazon SageMaker Studio notebooks

Access ML data
Connect to many data sources such as Amazon S3, Apache Spark, Amazon Redshift, CSV files, and more

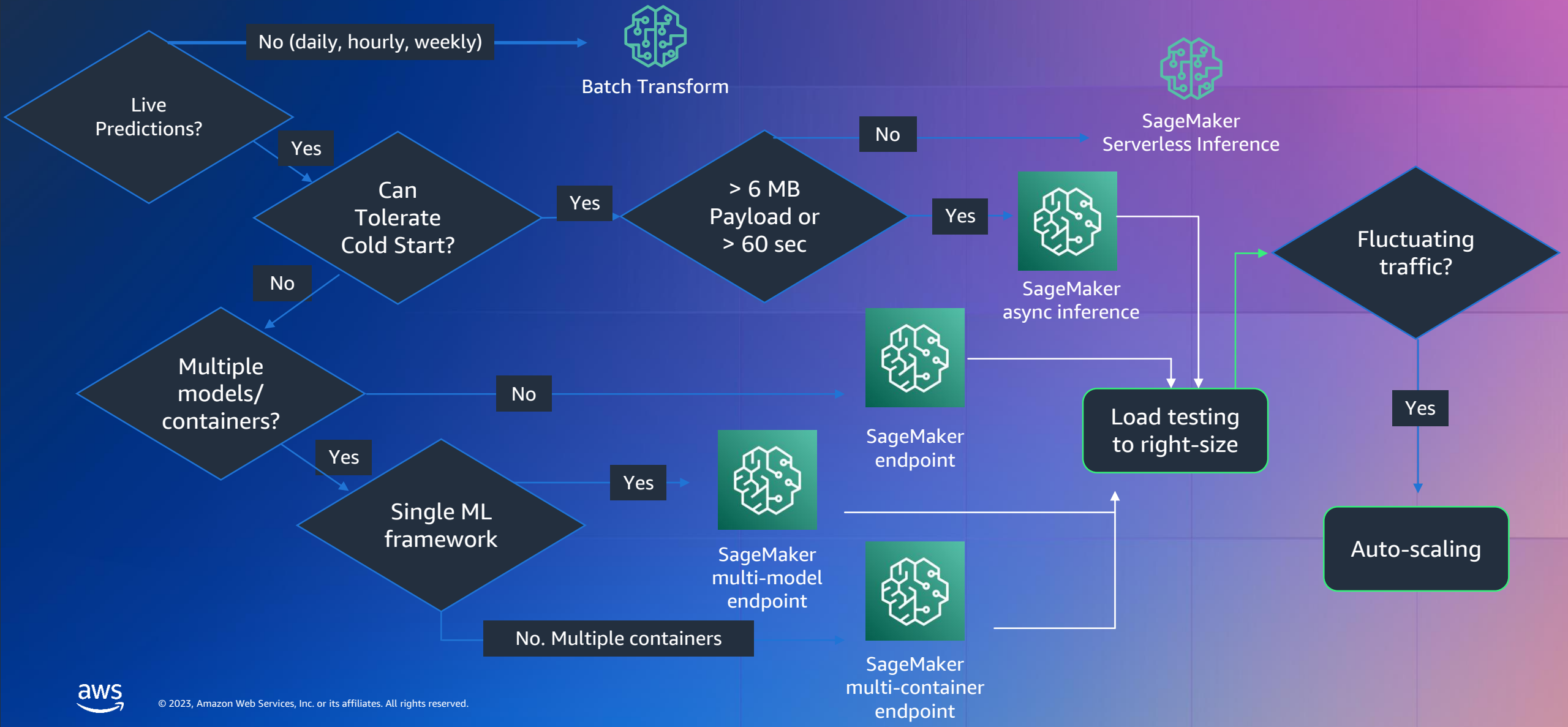
Prepare data
Transform data to browse data sources, explore metadata, schemas, and write queries in popular languages

Build ML models
Optimised with 150+ popular open-source models and frameworks such as TensorFlow and PyTorch

Train and tune ML model
Correct performance problems in real time

Deploy and monitor results
Create, automate, and manage end-to-end ML workflows to improve model quality

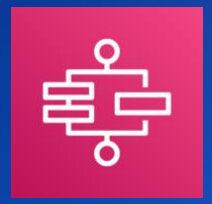
How to choose your deployment strategy



Machine Learning orchestration on AWS

Step Functions

- Integration with multiple AWS Services
- Integration with training, tuning, processing, and endpoints



SageMaker Pipelines

- Natively integrated with SageMaker Studio
- Define, visualise and manage pipelines in UI or Python
- Supports specific SageMaker features such as parallelism and custom dependencies

Adam Askew

Chief Technology Officer
CattleEye

CattleEye - Autonomous Livestock Monitoring

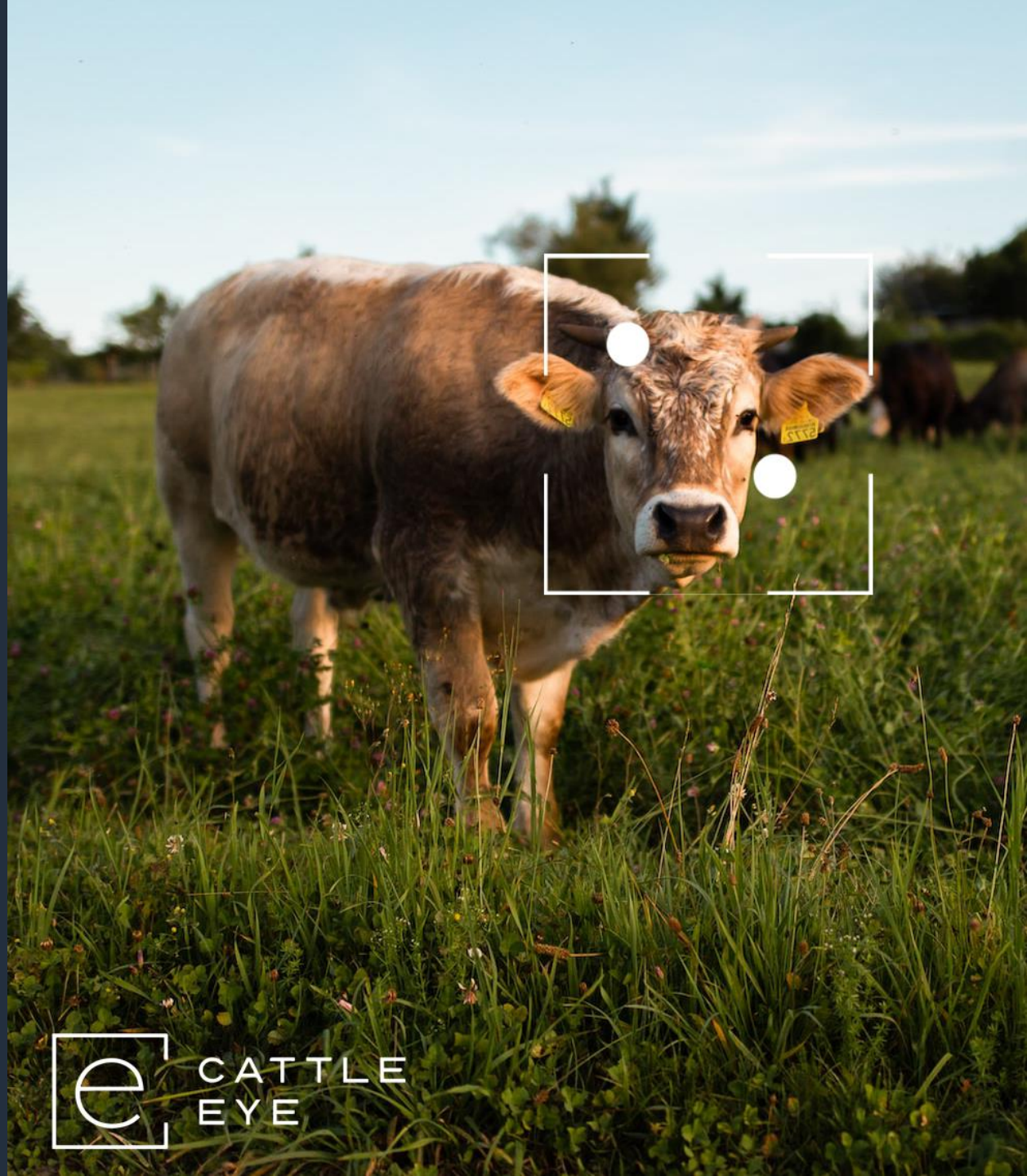
Founded 2019

Graduates of AWS Activate & Nvidia Inception

Generating 3+ Million insights per month

AWS Services

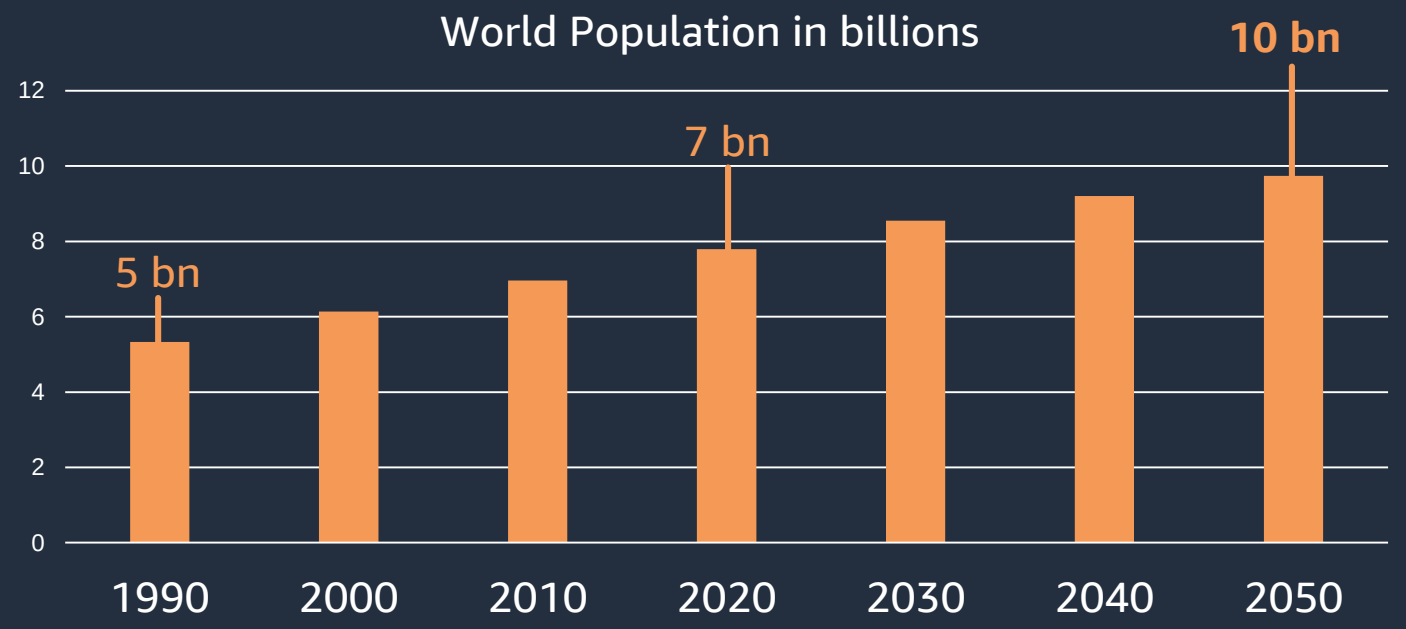
Sagemaker, AWS Batch, Step Functions, Event Bridge, Aurora, API Gateway, Fargate, Kinesis, S3, AWS Panorama



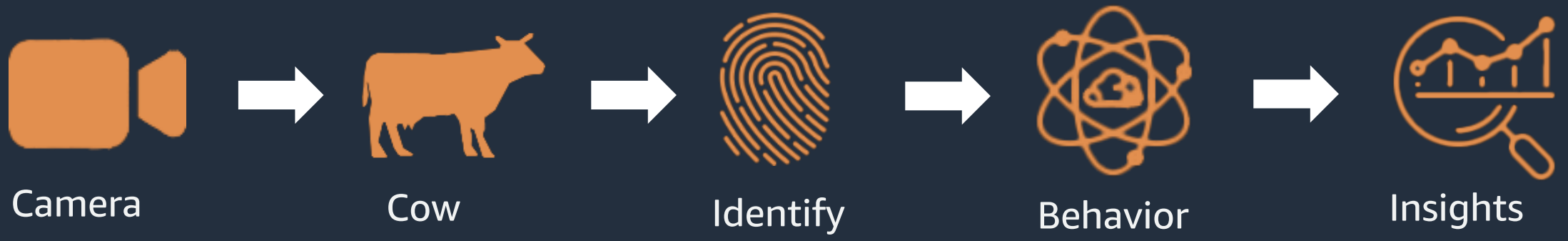
Feeding 10Bn people sustainably will be a challenge

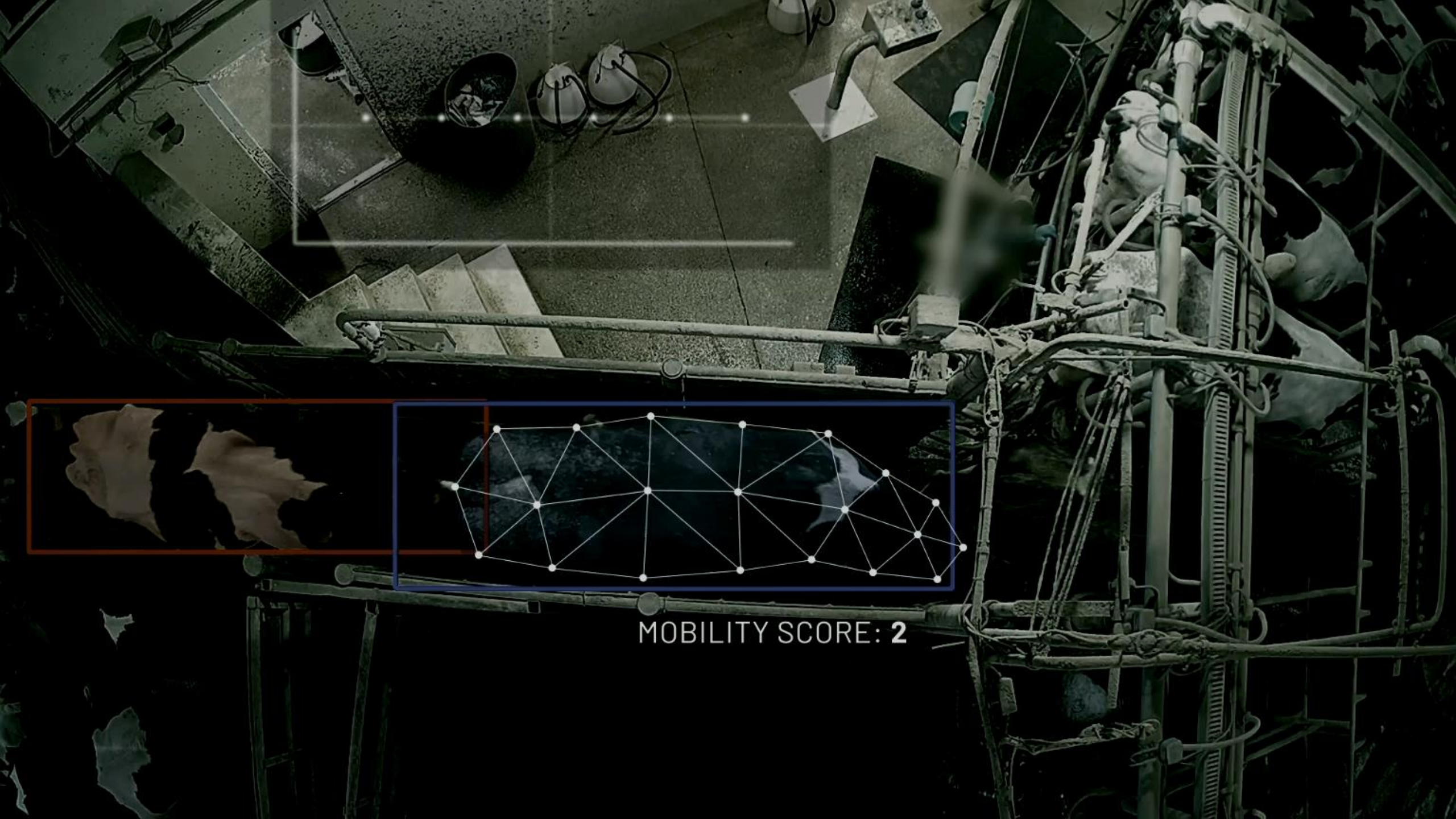
By 2050:

- World population of 10 Billion
- 70% more food needed
- Finite land availability
- GHG emissions must be reduced by 90%



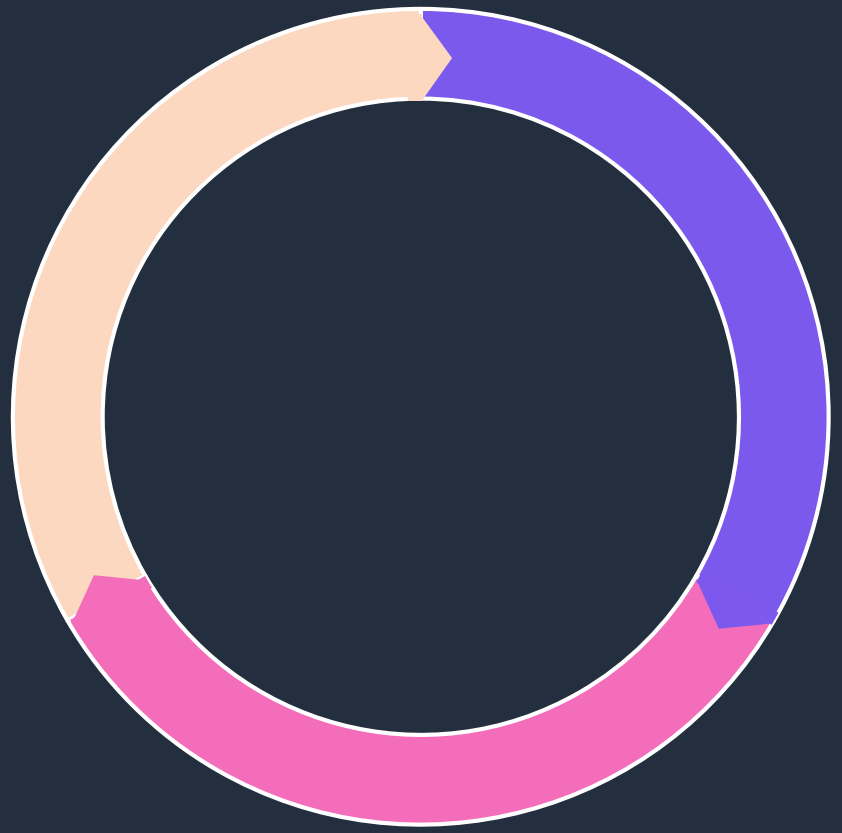
To reach these targets farms need better data and that requires accurate monitoring




MOBILITY SCORE: 2

Model training loop

1: Data Preparation



2: Model Training

3: Model Deployment & Monitoring

Data preparation

Storage – Data Lake

Multi-modal data storage:

- Videos
- System generated results
- Expert annotations
- Animal health data

Amazon S3, AWS Glue, Amazon Athena

Preparation

Collect data from the data lake.

Pre-process and prepare the data.

Version the dataset.

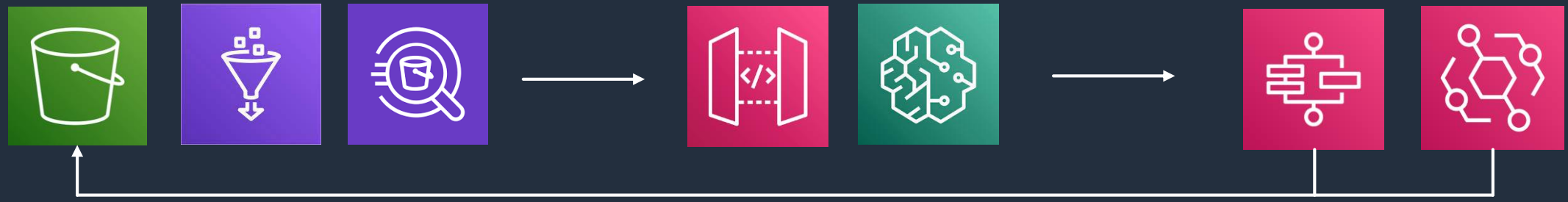
Amazon S3 versioning, Amazon SageMaker Processing Job, Amazon API Gateway

Automation

Scheduled creation of new dataset versions.

Orchestrate the job running and version checking.

Amazon EventBridge Scheduler , AWS Step Functions



Model training

Custom model

CattleEye custom models for :

Cow detection, Identification, Mobility, Lameness, Body Condition

Training process

Models are pretrained with weights from previous version

Training performed on GPU, although production inference on CPU

SageMaker Training Job

Validation / readiness

Base line metrics and new metrics from the training job are compared.

Amazon SageMaker Model Registry

Model deployment and monitoring

Model deployment

Once the model is approved for release, a production artifact is built and stored.

SageMaker model registry

Model Execution

Videos are processed in batch processing for low cost high throughput. This comes at the cost of high latency.

Models are all processed on EC2 CPUs while using Spot to keep cost low.

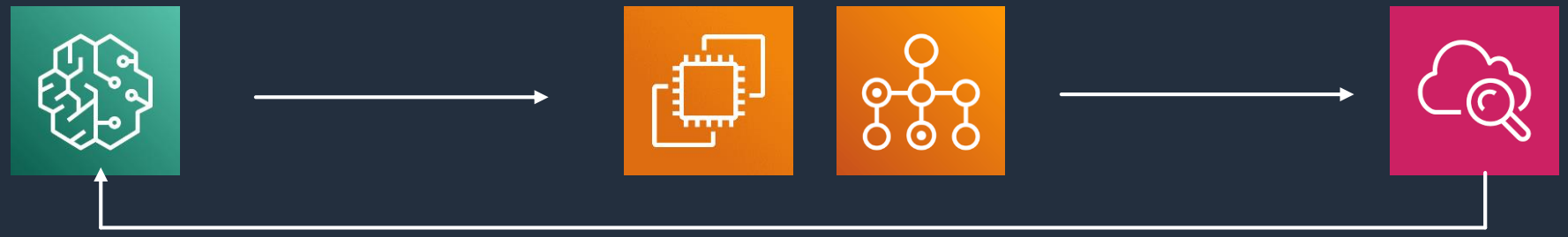
AWS Batch, Amazon EC2 Spot

Model Evaluation

New models are executed across the previous weeks data to ensure consistency of results.

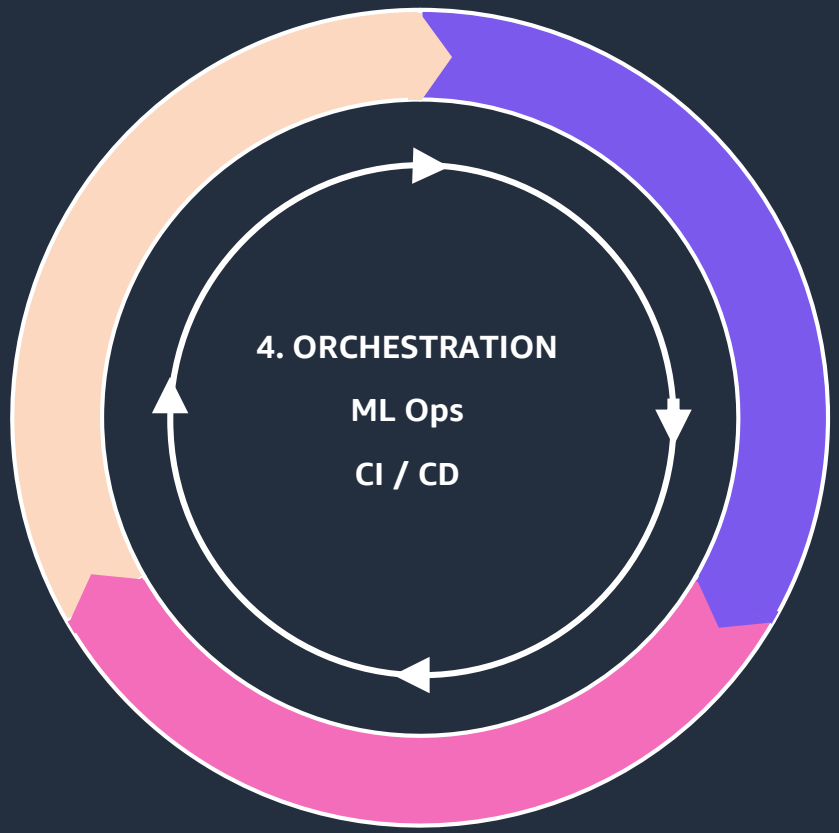
Confidence metrics and others are monitored to continuously evaluate the current models

Amazon CloudWatch Metrics



Model training loop

1: Data Preparation



2: Model Training

3: Model Deployment & Monitoring

ML Ops

CI / CD

All Data Science code is stored in GitHub.

GitHub actions are used to run some code testing and to kick off AWS Step Functions for extra processes.

GitHub, GitHub Actions, AWS Step Functions

Continuous data collection

Low confidence scores and fresh annotations are continuously added to our data lake.

Actions required human in the loop are assigned to work teams for annotation.

Amazon SageMaker Ground Truth, V7 Labs

Orchestration

AWS Step Functions are used through the training, deployment and production to orchestrate workloads and handle failures and retries.

All infrastructure is created and maned by AWS CDK

AWS Step Functions,

How do we do this sustainably?

$2.70

Cost per lame cow per day **

10% = **$120**

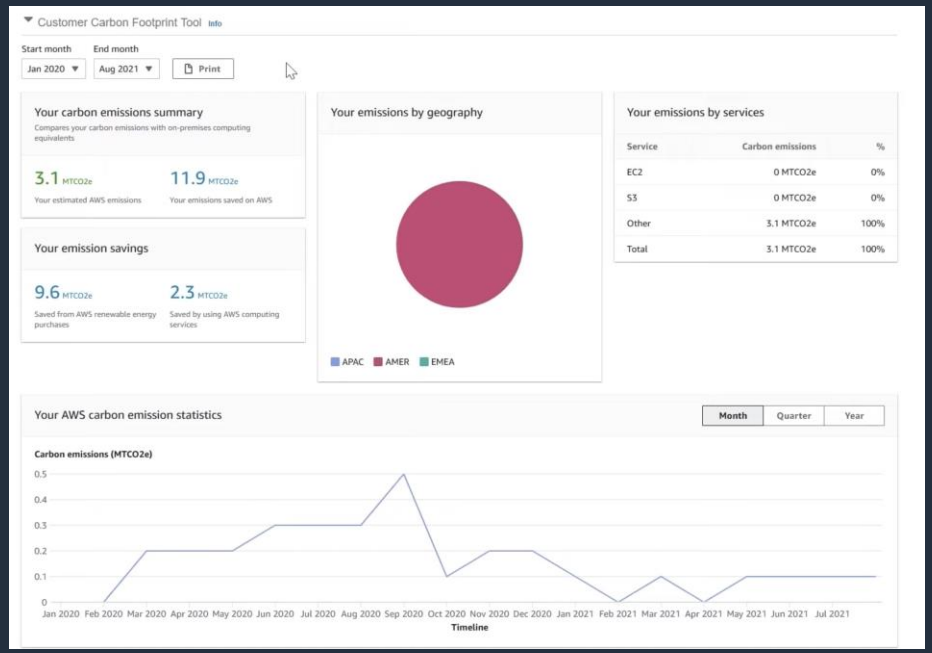
Decrease in lameness

Savings per cow per year

0.57 tons

Reduction in carbon footprint per cow per year

Carbon Footprint Tool



How do we do this sustainably?

OUR PROGRESS

By the Numbers

5x

AWS infrastructure is up to 5 times more energy efficient than typical European data centers

2.4 Billion

Liters of water are returned to communities each year from replenishment projects completed or underway

90%

In 2022, 90% of the electricity consumed by Amazon was attributable to renewable energy sources

CattleEye's upcoming projects

Edge

Support low bandwidth farms.

Offer low latency products.

Enable 24/7 monitoring.

AWS Panorama, AWS IoT Greengrass

Gen AI

Foundation models, chatbots, text generation, synthetic training data

Amazon Bedrock, Amazon SageMaker JumpStart,

Amazon Bedrock

Monitoring

Bias detection, explainability reports

SageMaker Clarify,

SageMaker Clarify

Thank you!

Jonah Craig

Startups Solutions Architect
Amazon Web Services (AWS)

Adam Askew

Chief Technology Officer
CattleEye

Please complete the session survey