Learning the Learner,
Using Machine Learning to Track the Performance of Machine Learning Algorithms

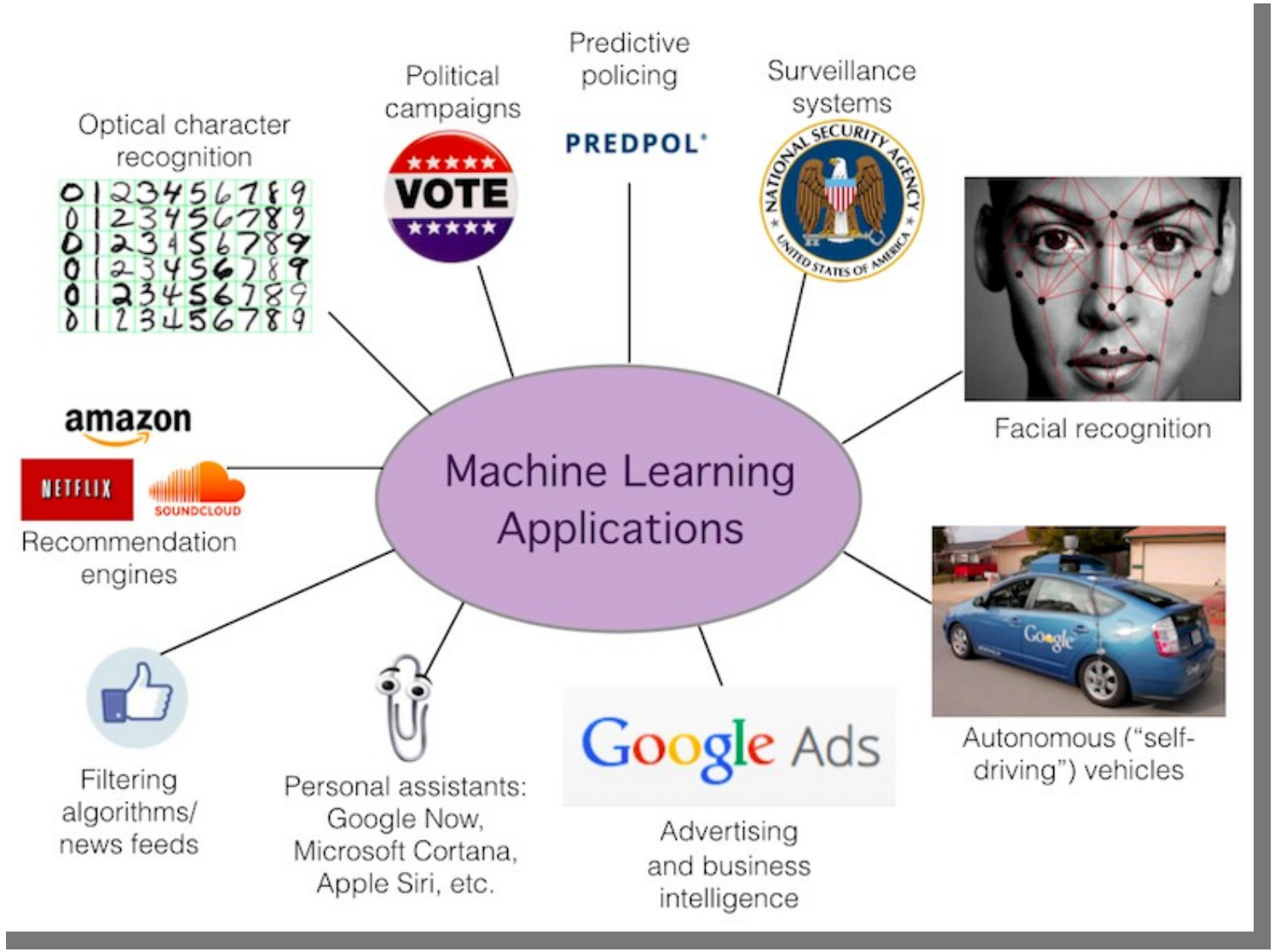
Ira Cohen, Chief Data Scientist

6th June, 2016

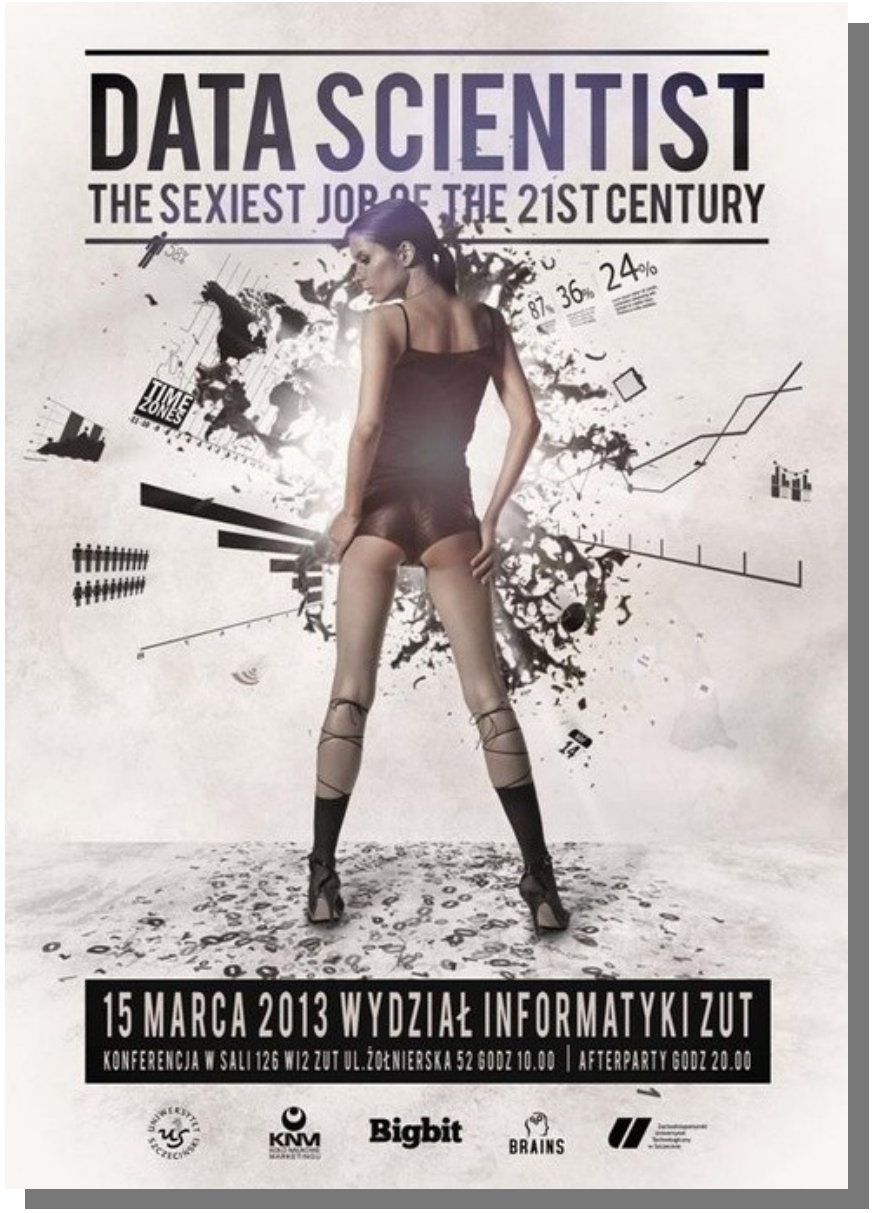
Gartner Cool Vendor 2016

Machine Learning used to be about publishing papers at NIPS and ICML…

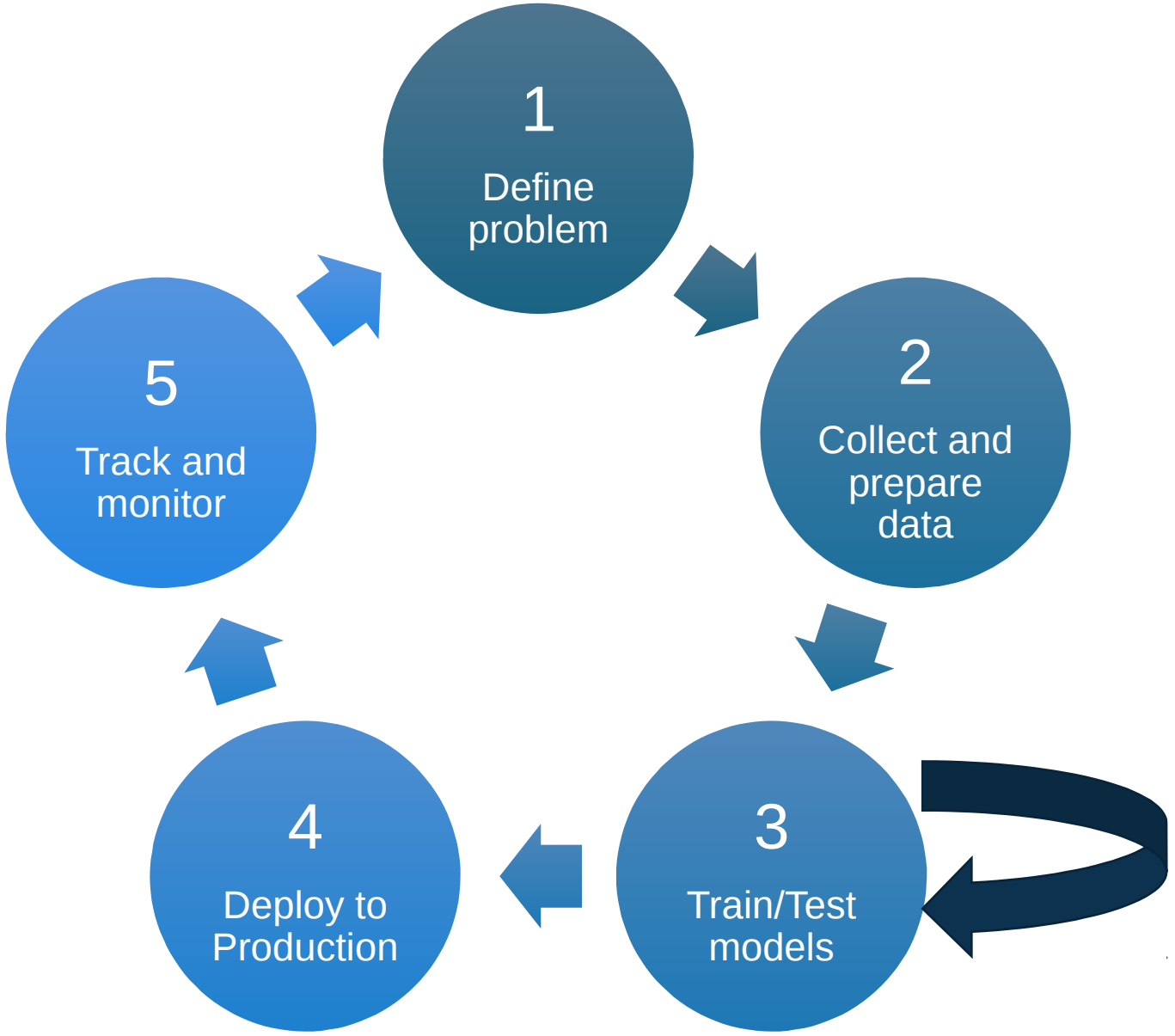
The coming of age of Machine Learning…



and now it is sexy...

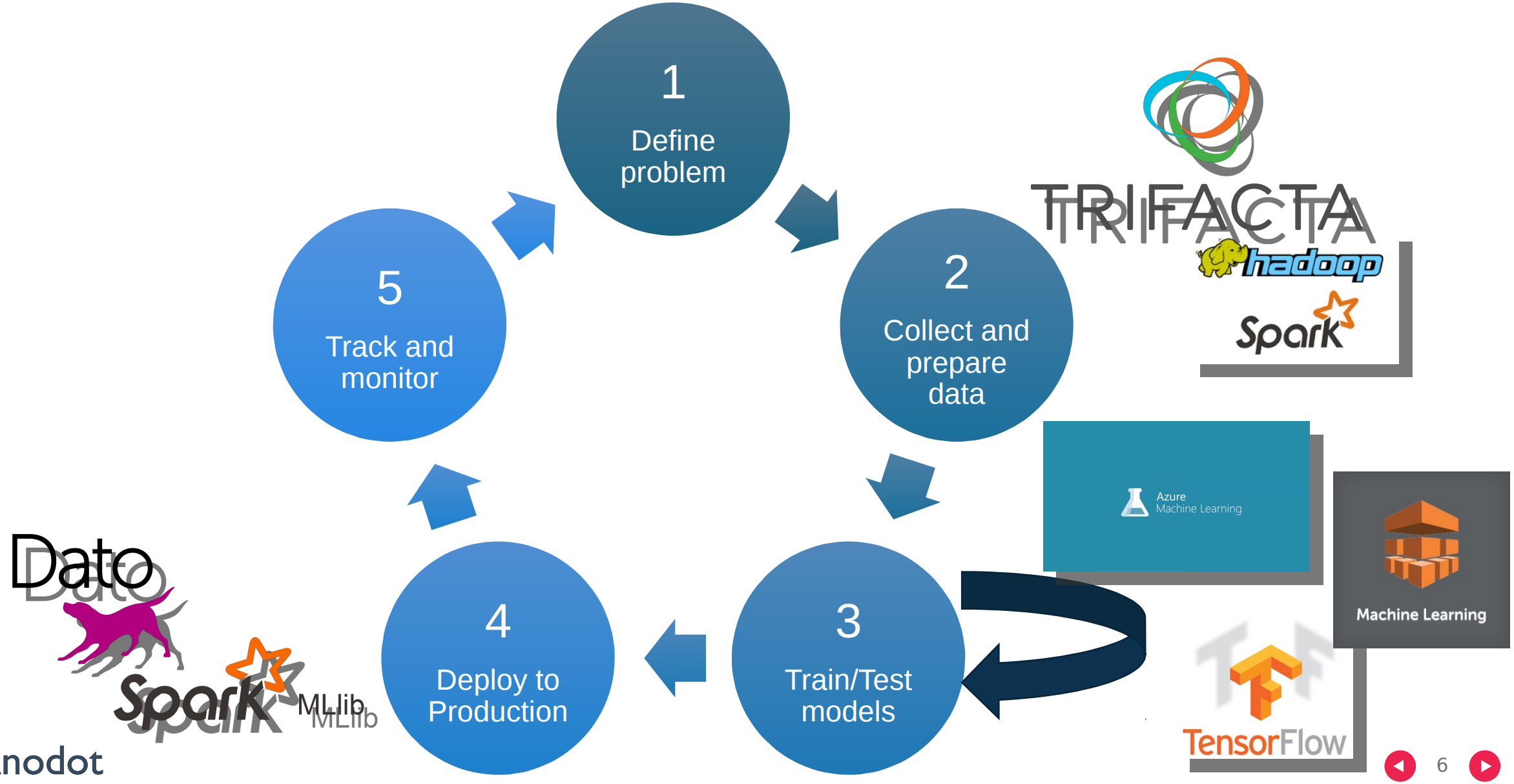


The Practical Machine Learning Process

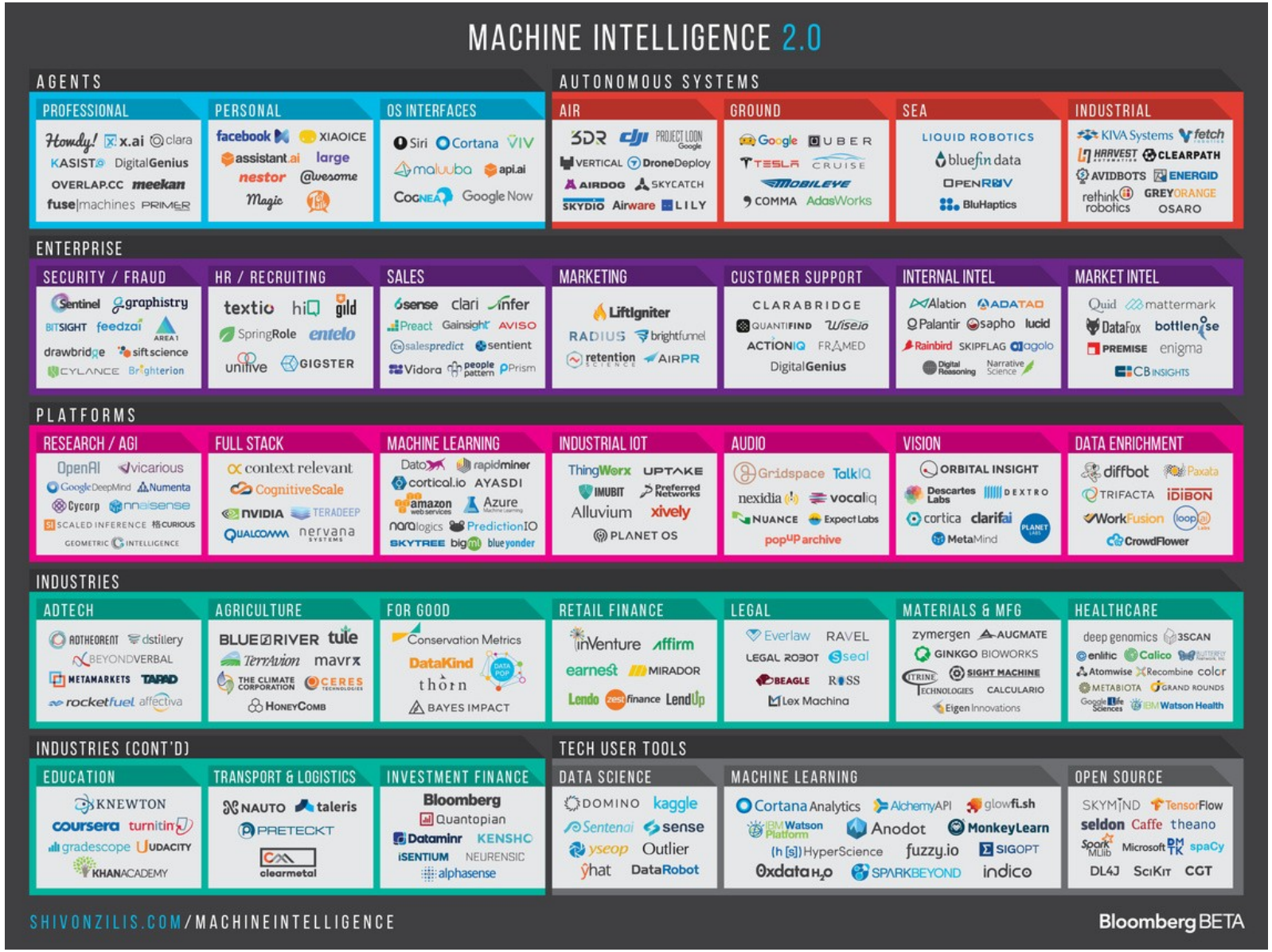


Anodot

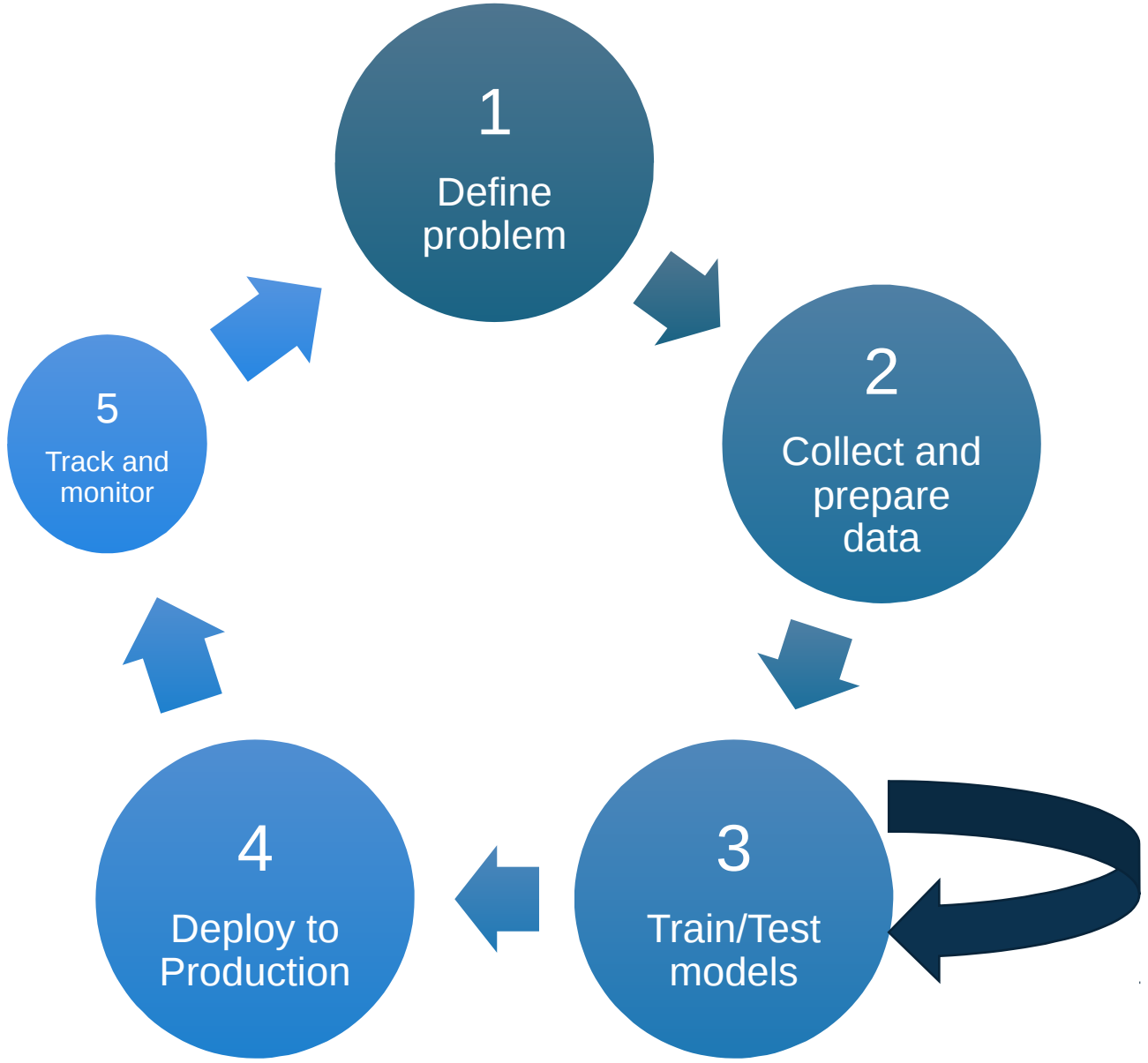
The Practical Machine Learning Process: The race for automation…



Lots of tools/platforms/solutions !



The Practical Machine Learning Process – The overlooked 5th step…



Anodot

Monitoring and Tracking

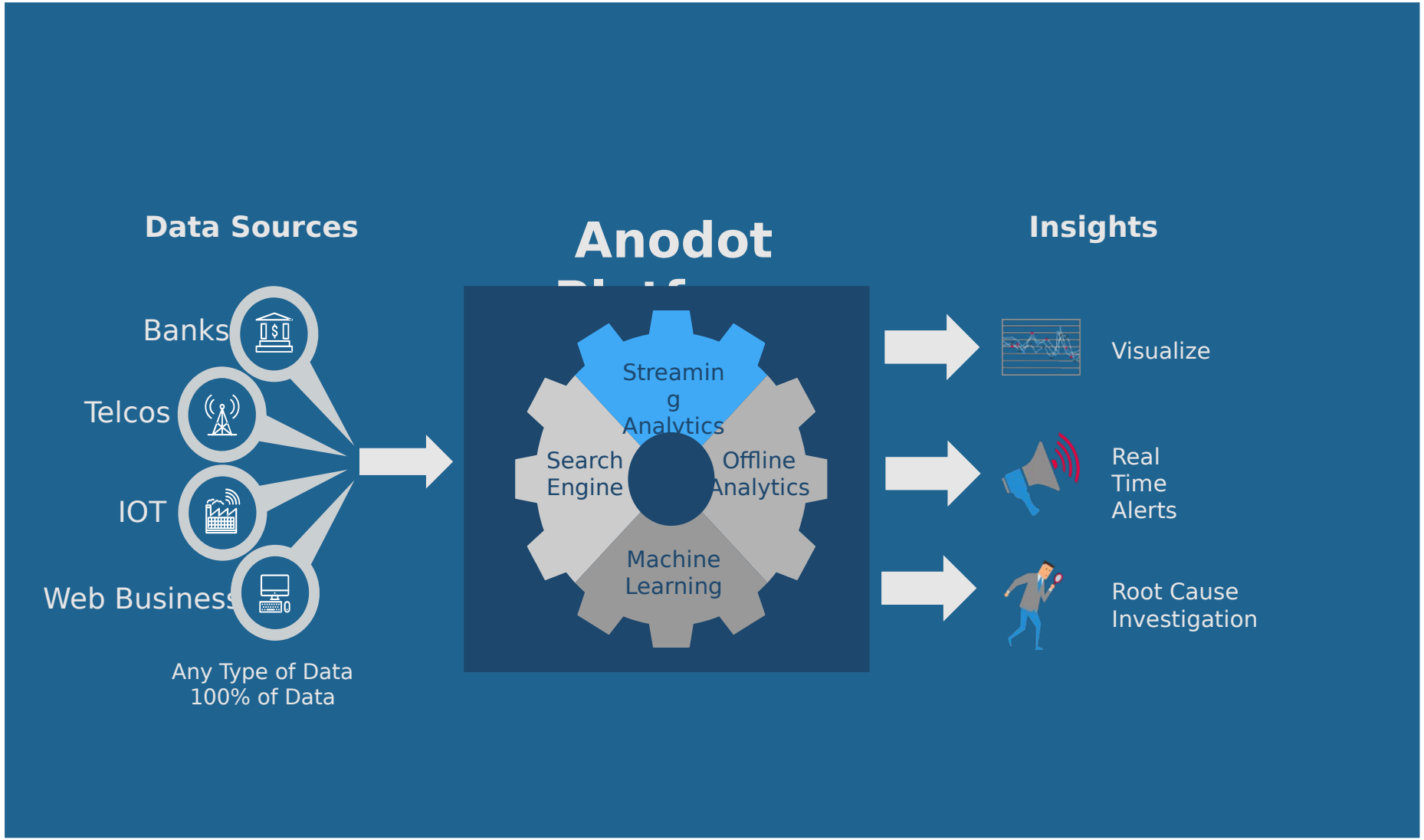
But How to Track and Monitor?
Manually is **NOT** an option…

Step 1	Step 2	Step 3
Define performance metrics and expected behavior	Collect metrics continuously in production	Track for unexpected changes in metrics

Anomaly Detection

Anodot

Anodot's Business Incident Detection Platform



Detecting Business Incidents: Anomaly Detection

Anomaly detection: Detecting the Unknowns Saves Time + Money

Web Services

Detecting business incidents + unknown business opportunities

Industrial IoT

Proactive Maintenance
Detecting issues in factories/machines

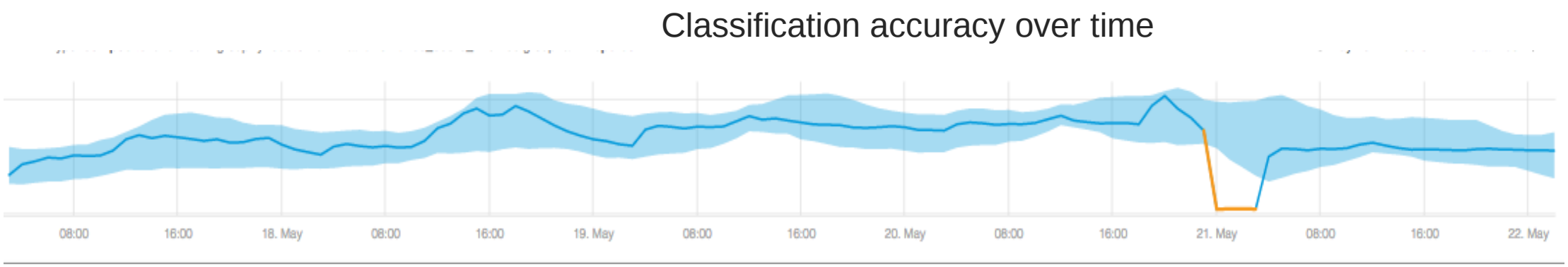
Security

Detection of unknown breach/attack patterns

Machine Learning

Closing the “Machine Learning” loop
Tracking and detecting ”unknowns” not modeled during training

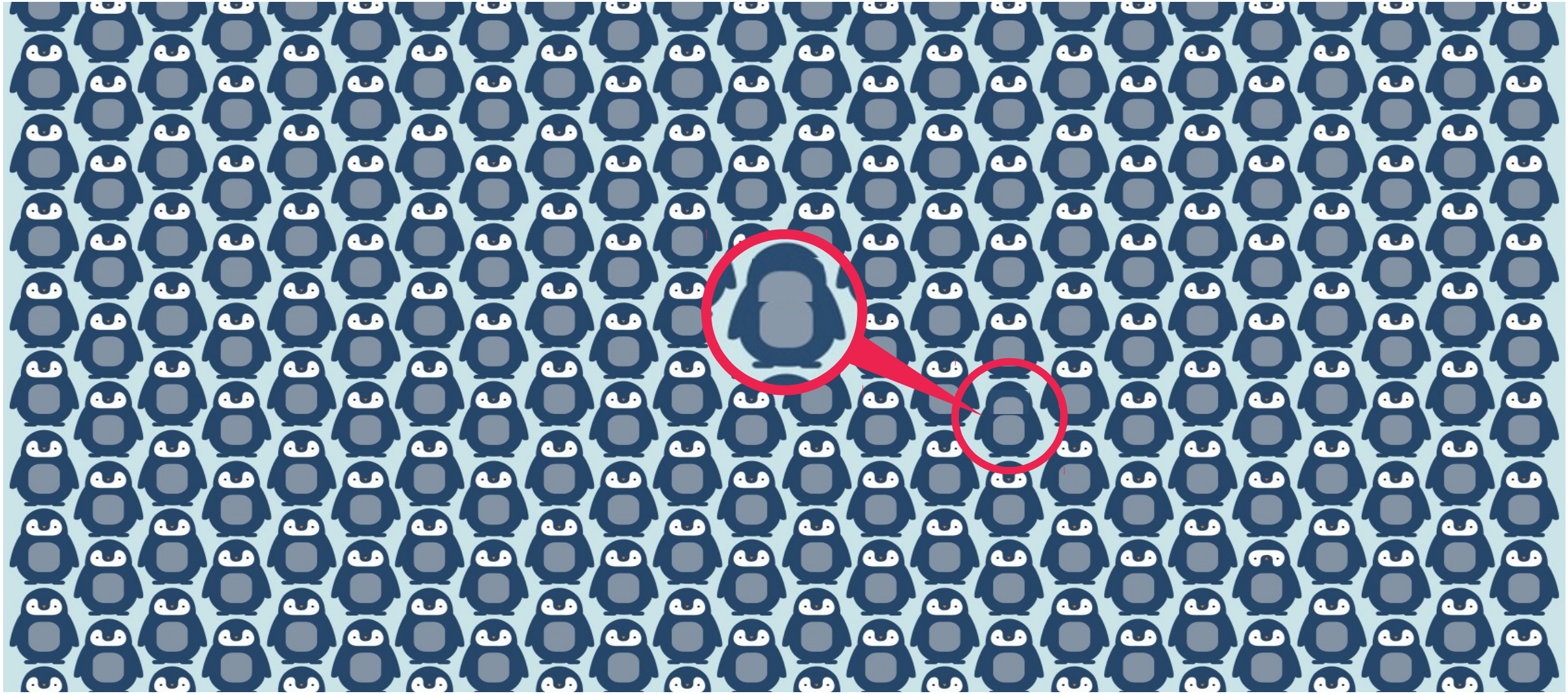
Detecting Unknowns of ML in Production: Anomaly Detection



What is Anomaly Detection?

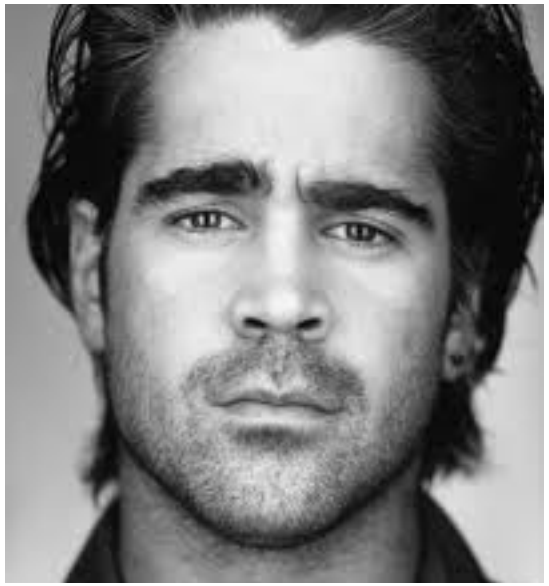
Anodot

Find the Anomaly



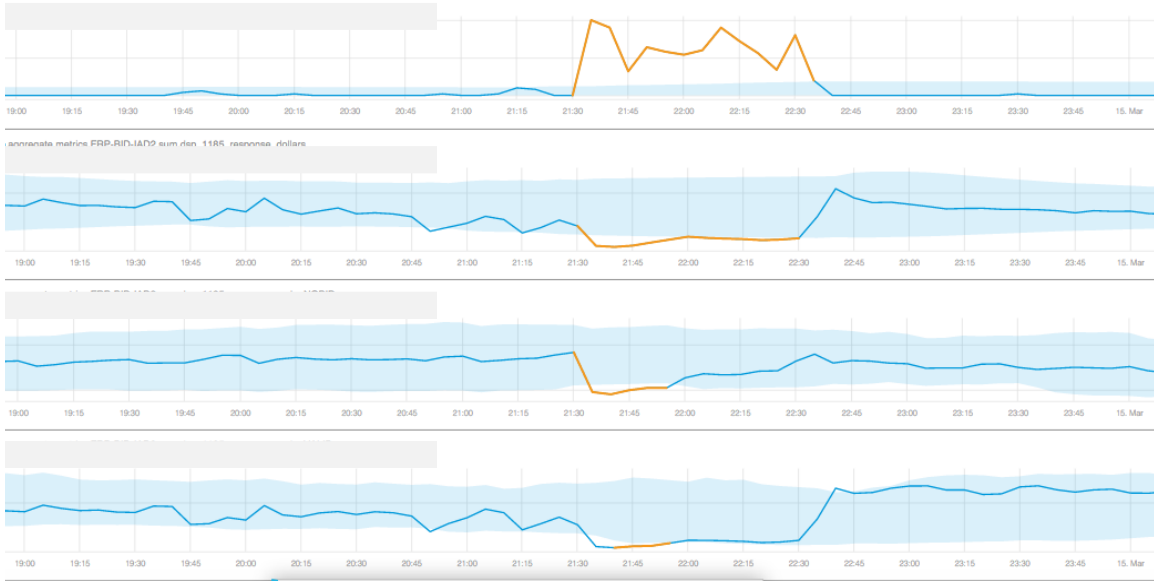
Anomaly Detection

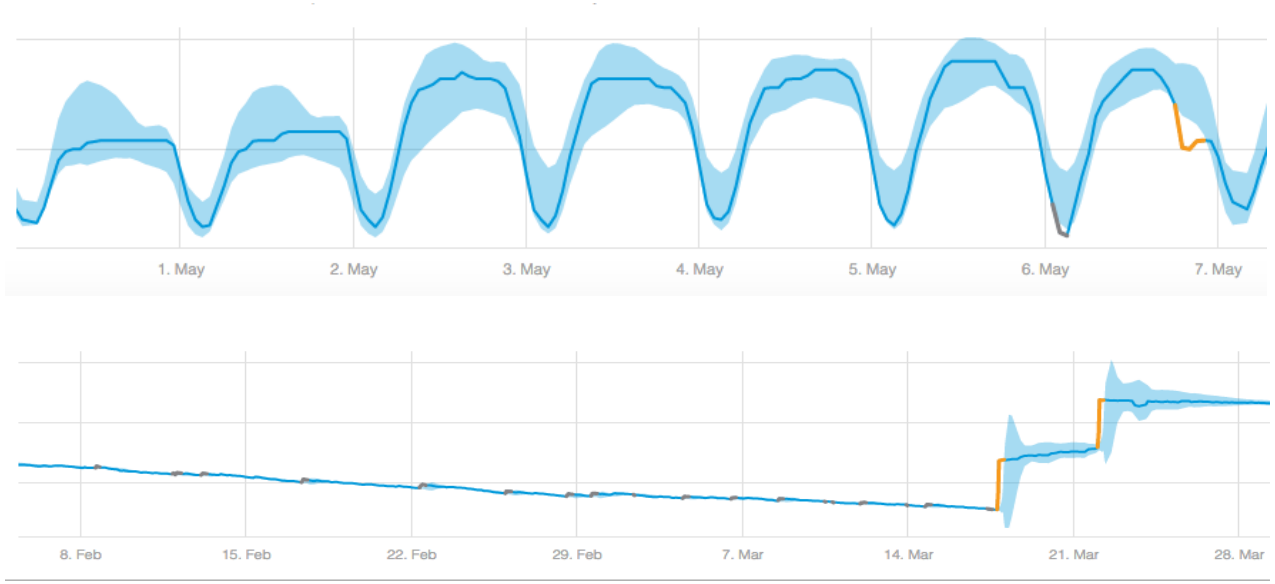
- Ill posed problem
- What is an anomaly?



Anomaly Detection in Time Series Signals



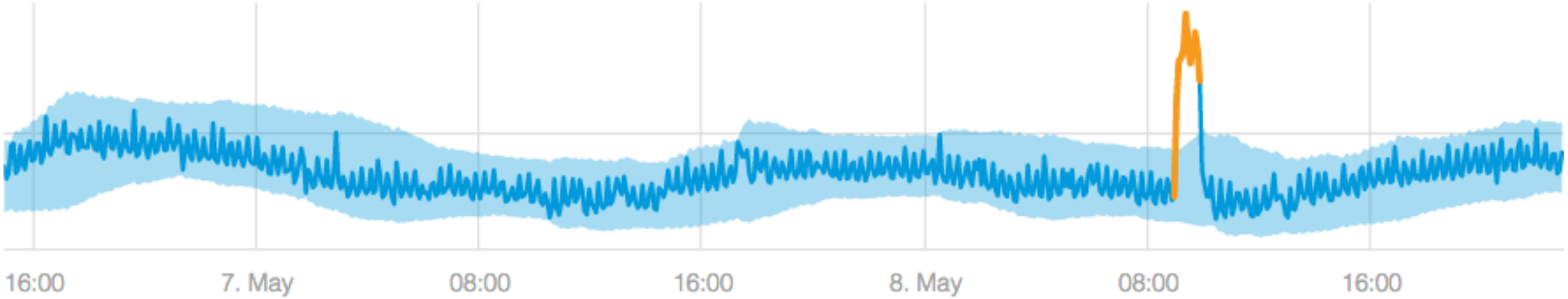


Unexpected change of temporal pattern of one or more time series signals.

Anomaly Detection Methods

Anodot

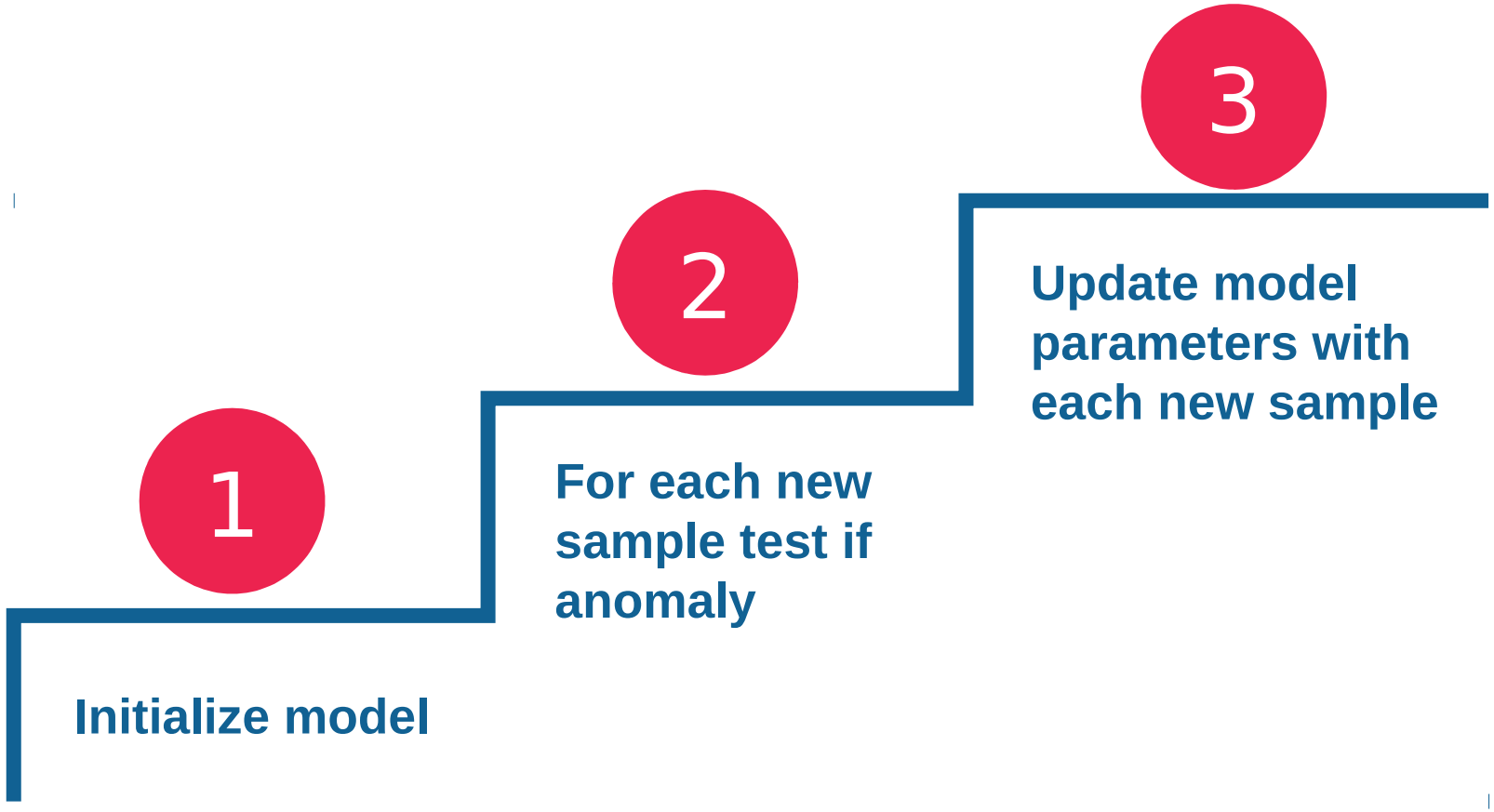
Anomaly Detection: General Scheme



General scheme

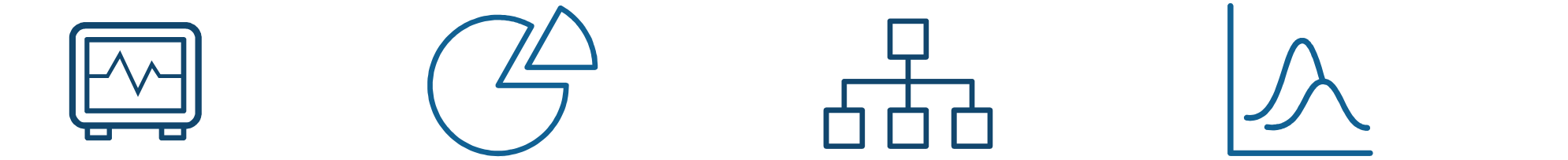
Step 1	Step 2	Step 3
Model the normal behavior of the metric(s) using a statistical model	Devise a statistical test to determine if samples are explained by the model.	Apply the test for each sample. Flag as anomaly if it does not pass the test

Online Anomaly Detection Algorithms



Example Online Models/Algorithms

Batch Anomaly Detection



1. Collect historical samples
2. Segment samples to similarly behaving segments
3. Cluster segments according to some similarity measure
4. Mark as anomalies segments that are in small or no clusters

Example Batch Anomaly Detection Methods

Multi-model distributions:

- Gaussian models
- Generalized mixture models

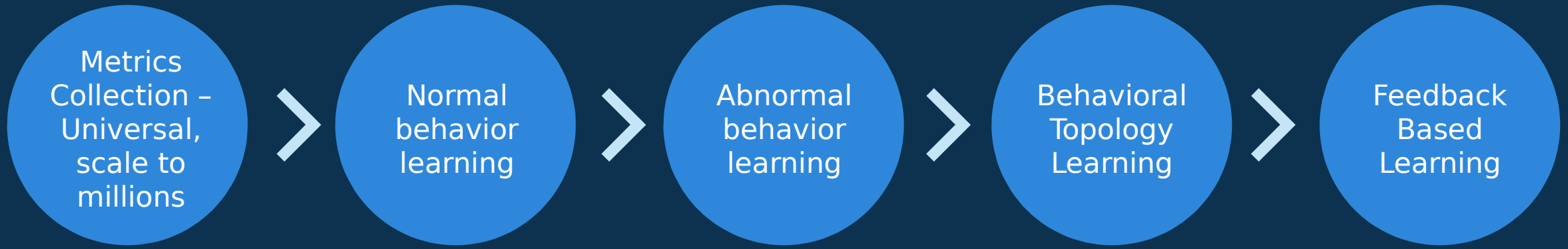
One sided SVM

PCA

Clustering methods (K-Means, DBScan, Mean-Shift)
MOST COMMON IN USE

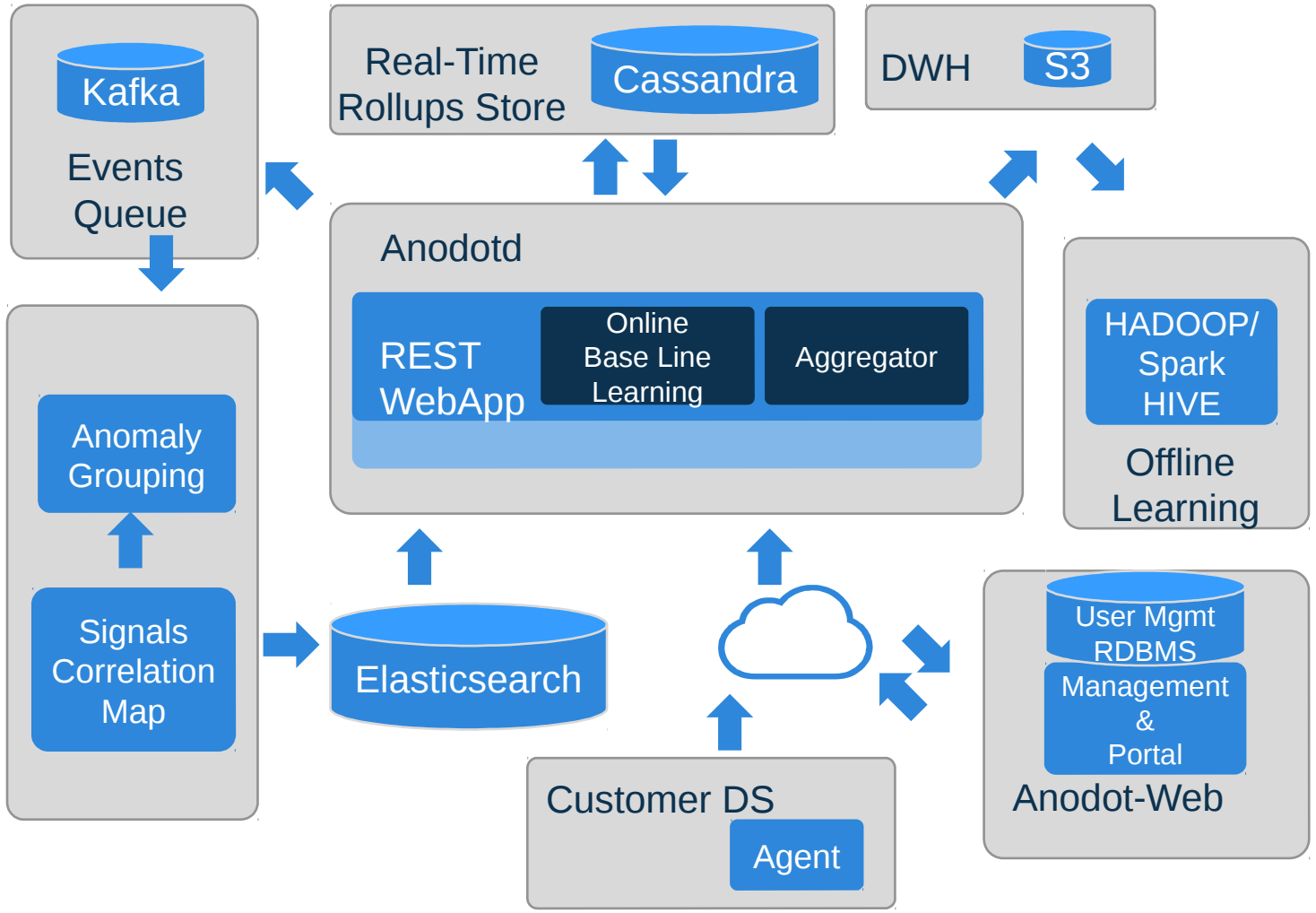
Hidden Markov Models

Anodot's Automatic Anomaly Detection



1 2 3 4 5

Large Scale Anomaly Detection System Architecture



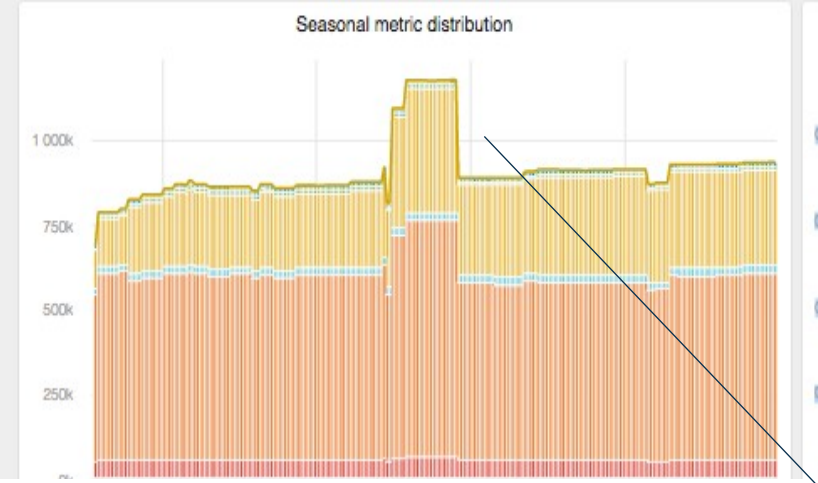
- 3.3 billion daily samples
- 75,000,000 metrics
- 150,000,000 models
  - updated with each sample
- 300,000,000 correlation links
  - Updated daily
- 7,000,000 seasonal models
  - Updated daily
- 30 types of learning algorithms
  - Metric classification, seasonality detection, trend, baseline models, clustering algos, LSH, …
- And counting…

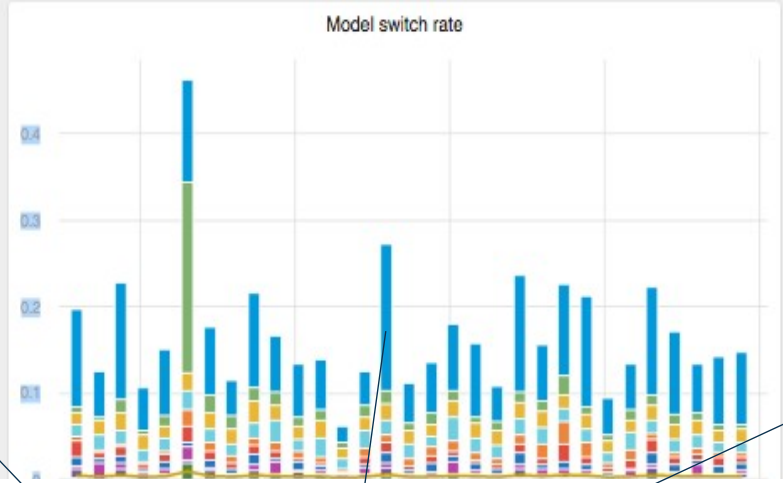
Tracking the performance of the algorithms: With Metrics

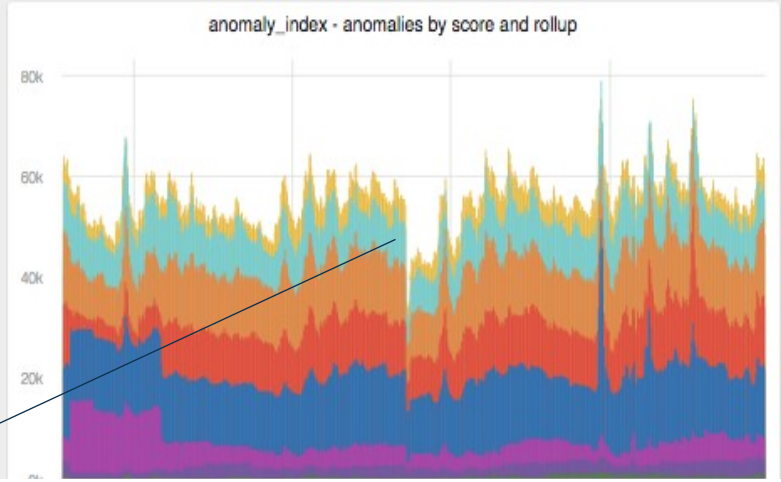
Seasonality detection distribution

Model switching rate and distribution

Anomaly quality scores

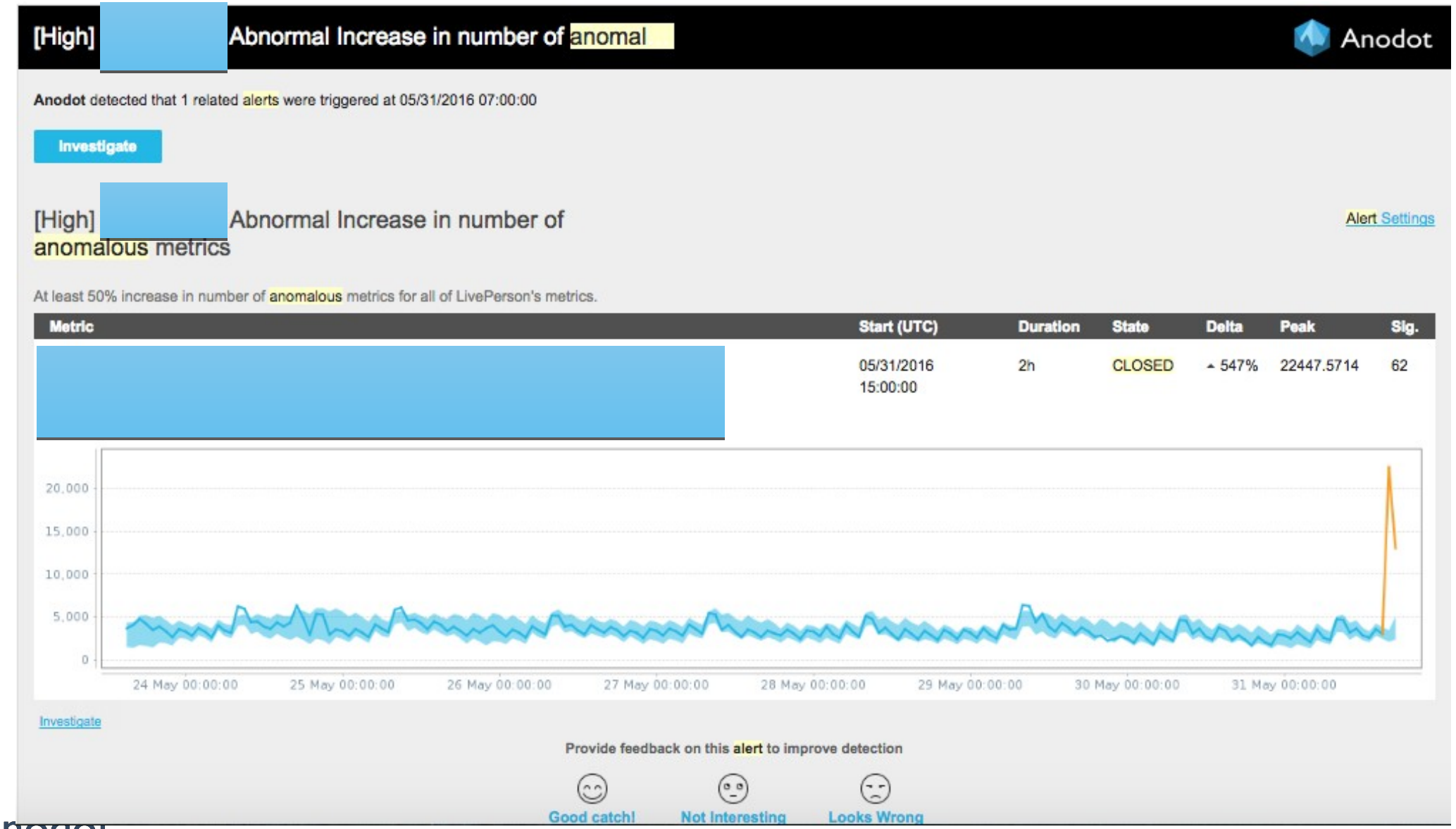




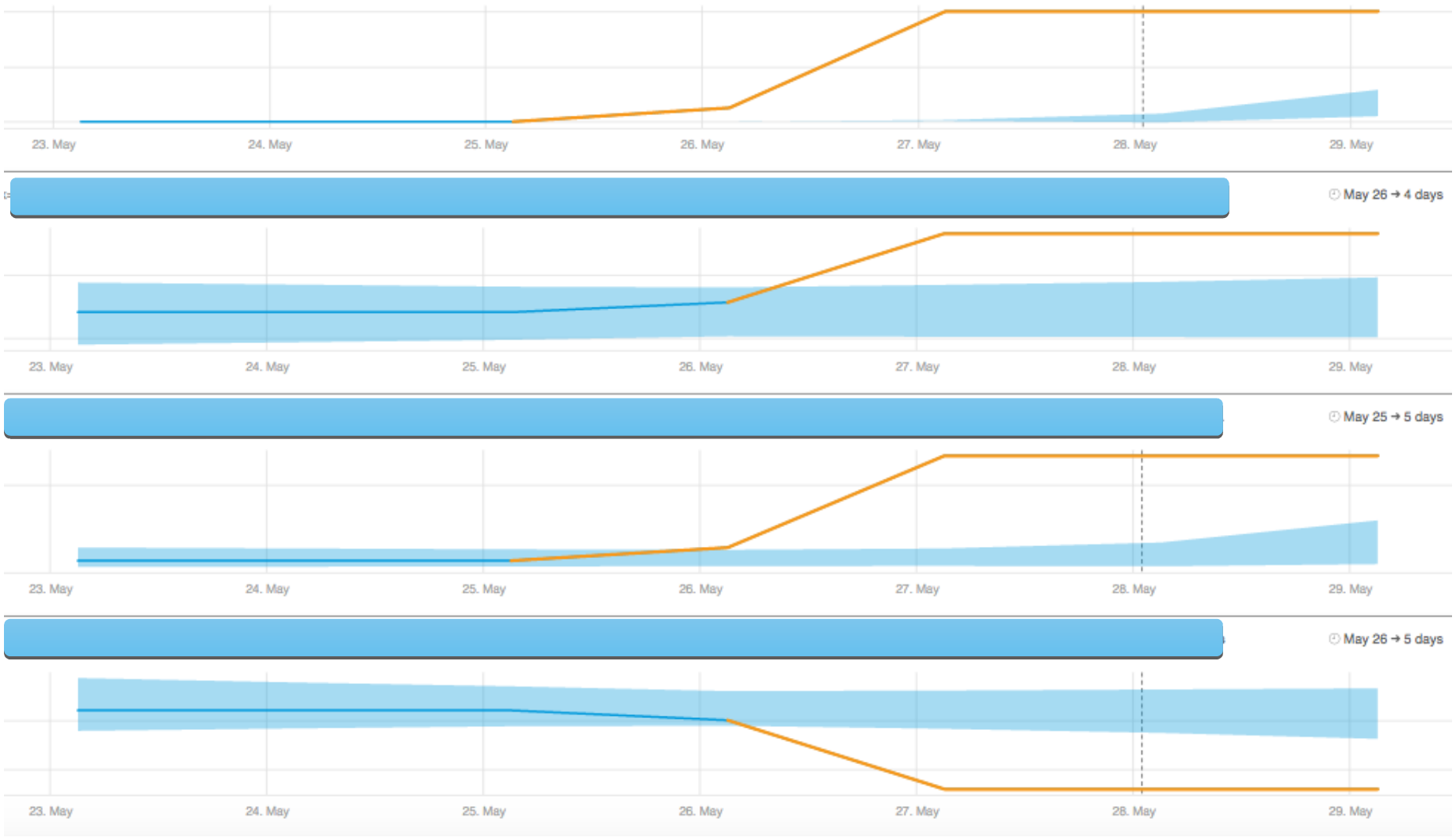


Anomaly as a result of deployment

Alert when results are Abnormal: New type of model required

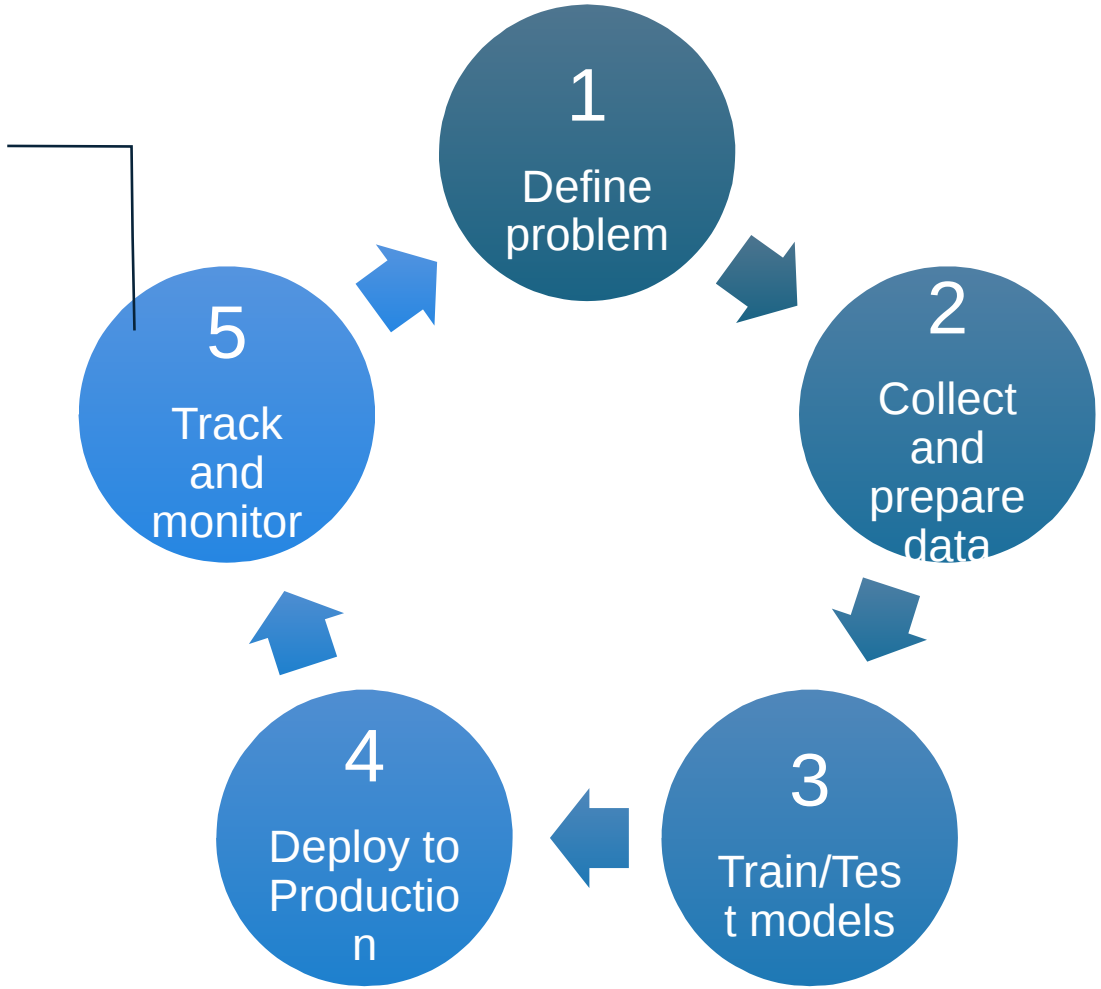


Alert when results are Abnormal:
Anomaly in Seasonality Detection Distribution



Summary

- Automated anomaly detection
  - Closing the loop of the machine learning process
- The first step in full automation of the learning process




Anodot
Thank you
Gartner
Cool Vendor 2016
ira@anodot.com