

# Achieving Trusted AI in Government

Ethical, Transparent, and Explainable AI

## The prototype to production chasm

Government agencies face an avalanche of algorithms from a myriad of sources, both internal and external, as development teams act on their enthusiasm for the potential of AI to address critical challenges. However, AI is often developed without governance, controls, or clear target outcomes. What cannot be trusted or explained cannot be used, and unused innovation fails to achieve results.

Millions have been spent by government research agencies and others to build prototypes that never advance into production due to the lack of model trustworthiness.

## Cross the chasm with TruEra

TruEra's AI Quality solutions help models traverse the chasm, by introducing necessary transparency and explainability while providing independent verification and validation. TruEra works across all model types, no matter how models were originally created or where they reside.

TruEra's AI Quality solutions are both model and development platform agnostic. TruEra's model quality metrics include accuracy, explainability, bias, stability, reliability. TruEra provides model comparison benchmarking, collaboration, and reporting. TruEra can be used for development, evaluation and post-deployment Monitoring. Containerized and highly portable, TruEra solutions can be used across a myriad of government use cases and in any environment where real-time, comprehensive model intelligence is required.

## Key benefits



### Experience fast, complete visibility

Proven to be faster and more accurate than current solutions



### Deploy in a variety of environments

On-premises, private cloud, hybrid, or air-gapped networks



### Enable fast action - know where to look and what to fix

With more reliable analyses of root causes, your data science and MLOps teams stay focused.



### Demonstrate transparency that leads to greater trust

Reporting and analytical capabilities demonstrate the fairness and effectiveness of AI initiatives.

# Key capabilities

## Performance visibility

### Model quality dashboards

For each machine learning model, TruEra provides a comprehensive dashboard of key metrics. Top level diagnostics include:

- Model performance (actual or estimated accuracy, decision trends)
- Model score drift (distribution and stability score)
- Feature drift (model inputs)
- Bias: disparate impact
- Data quality trends (rule violations)

### Model overwatch dashboard

A quick summary of 100s of models, highlighting which need quick attention and which are performing at expectation.

## Featuring analyses unique to TruEra

- Drivers of model instability
- Estimated AUC, with and without data labels
- Drivers of model bias

## Bias Tracking

### Model score disparity analysis

Quickly understand the impact of protected features on model outputs.

### Feature contributions of bias impact analysis

Get visibility into the feature contributions to overall model bias for protected classes, to analyze whether variable treatment is justified or unjustified. This is unique to TruEra.

## Diagnosis

### Root cause analysis

TruEra helps you answer questions like:

- Which features are contributing to model stability issues?
- Which features are contributing to bias?
- Which features have egregious data quality problems?
- Which segments are most impacted?
- Are there data quality changes from baseline behavior?

# Key capabilities

## Alerts and notifications

Alerts on any monitoring diagnostic metric in the dashboard, such as:

- Model stability
- Feature influence/stability
- Feature drift
- Bias
- Segment comparisons for all of the above

### Alert configuration

Set triggers, evaluation frequency, messaging, tags, user notifications, channels, and error handling, among others.

### Notifications across 19 communication channels.

Use the most popular communications tools, such as:

- Email
- Webhooks
- Pager Duty
- Slack

Tracking of issue drivers and common resolutions with annotations for alerts, anomalies, outcomes, and resolutions.



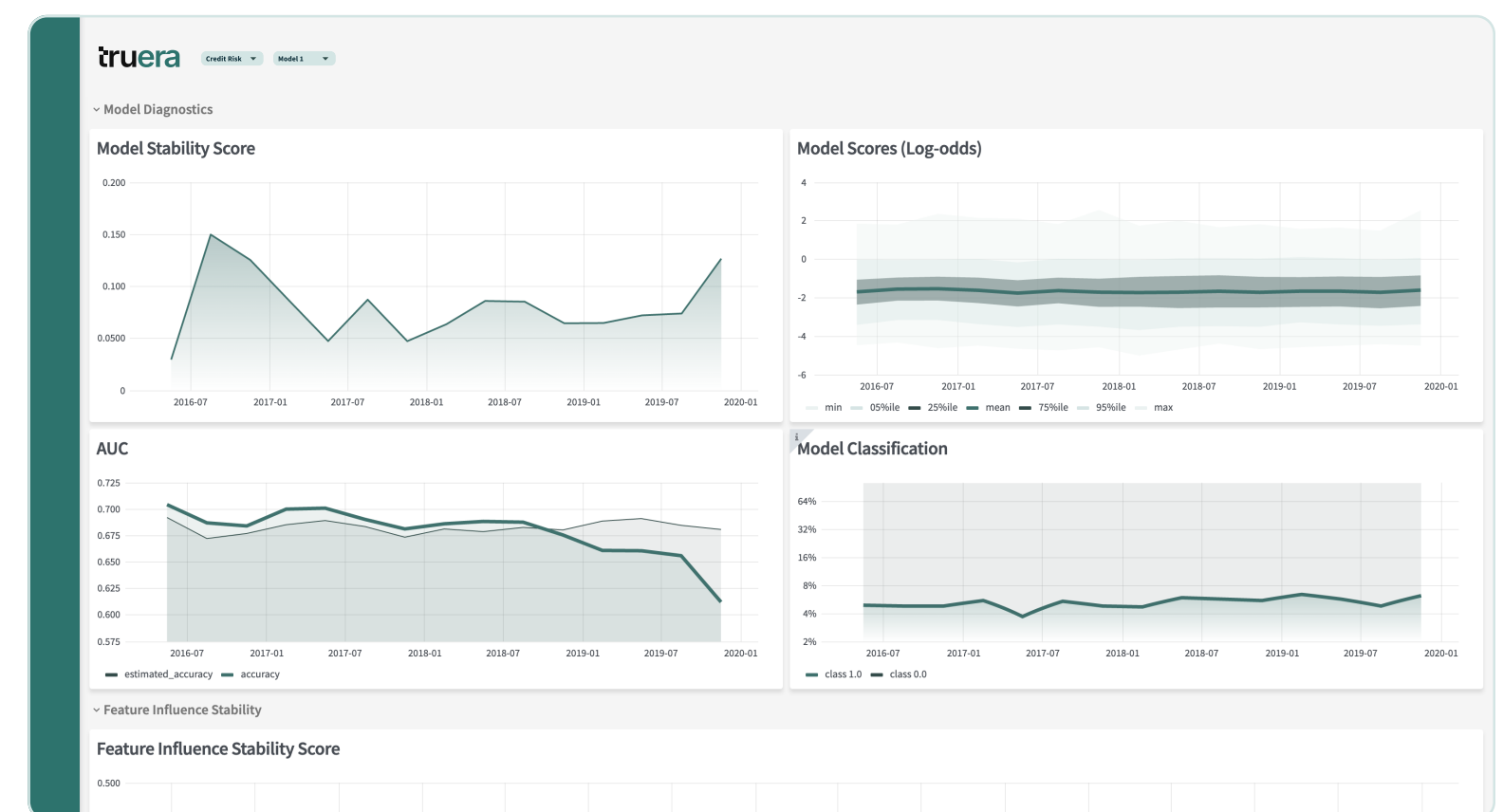
We see TruEra as an essential partner... in how we build and operationalize higher-quality, trusted AI models faster and more efficiently.

- Vishu Ramachandran, Global Head, Retail Banking, Standard Chartered

## Integration and workflow

TruEra is quick and easy to deploy, even in complex enterprise environments.

- **Deploy on-premises or on your cloud, including public cloud** providers like AWS, Google, or Azure.
- **Fast setup and easy user administration** with SSO integration and local TruEra authentication.
- **Easy integration** to common model serving infrastructure such as Sagemaker and Algorithmia.
- **Scale quickly** for high volumes of ML models
- **Easily export data via APIs** for integrations to applications or business intelligence solutions, such as Tableau and Looker.



TruEra Monitoring tracking a live production model for performance against key metrics, such as stability and accuracy (AUC).