

Fiddler AI Observability and Security Platform

Increase performance, safety, and trust in LLM and ML applications

Fiddler is the pioneer in AI Observability and Security, providing an all-in-one platform that enables Data Science, Engineering, Trust & Safety, and Security teams to move beyond experimentation and confidently deploy LLM applications and ML models in production. With enterprise-grade safeguards, real-time monitoring, and actionable insights, Fiddler helps teams improve LLM and ML deployments while protecting AI applications from safety and security risks.

Fiddler aligns teams with guardrails, monitoring, and rich diagnostics to protect LLM and ML applications, uncover issues affecting model outcomes, and establish a framework for responsible AI. AI Observability is reliant not only on metrics but also on how actionable model insights are when something eventually goes wrong.




Fortune 500 organizations use Fiddler to establish a foundation of trust, safety, and security for their LLM and ML deployments, delivering high performance AI to reduce costs and increase ROI, and be responsible with governance.






How Fiddler AI Observability and Security Works

Getting started with Fiddler is easy. Fiddler is model, framework, and data-agnostic.

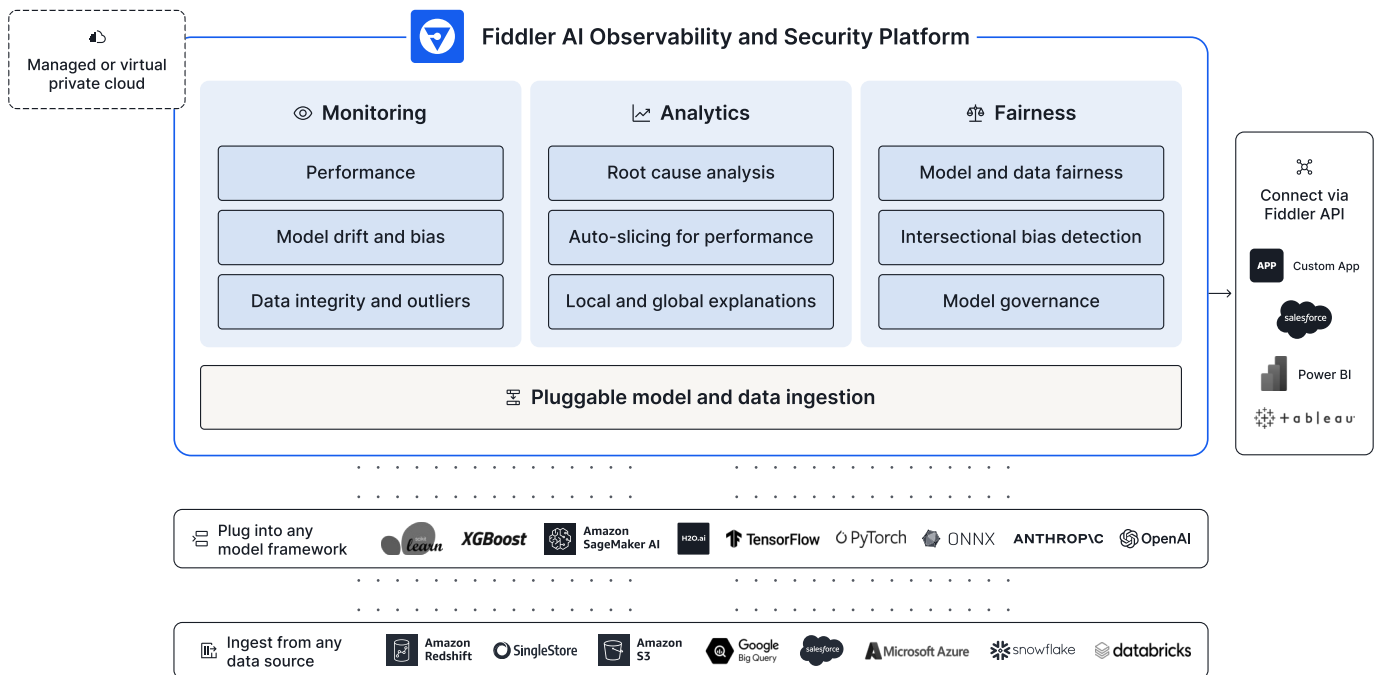
Scale ML Model Deployments

-  Upload training and production data from any local or cloud data source of your choice, including Snowflake, AWS, Google, PostgreSQL, or Databricks.
-  Publish data with real-time event streaming or schedule batches of events that fit your business needs.
-  Monitor performance, detect bias, and analyze models trained using Amazon SageMaker AI, TensorFlow, PyTorch, SAS, Python, H2O.ai, Scikit Learn, and more.

Launch LLM Applications into Production

-  Upload inference data from any LLM application of your choice — from any LLM model such as OpenAI or Llama or a tool call/vector database retrieval.
-  Publish data with real-time event streaming or schedule batches of events that fit your business needs.
-  Moderate prompts and responses in real time, monitor for hallucination, toxicity, PII leakage, prompt injection attacks and other LLM metrics. Analyze models deployed in Amazon Bedrock, GCP Vertex AI, Microsoft Azure, and more.

The All-In-One AI Observability and Security Platform



Predictive Models

Monitoring

Fiddler helps you reduce costs and increase efficiencies by tracking how all your ML deployments are performing in pre-production and production. Mitigate model performance issues before they impact your business.

- **Performance Monitoring:** Track your predictive model's performance with out-of-the-box metrics, including accuracy, recall, precision, F1-score, regression, mean absolute error (MSE), and mean average precision (MAP) for binary classification, multi-class classification, regression, and ranking models.
- **Unstructured Monitoring:** Detect drift in natural language processing (NLP), computer vision, and deep learning models to improve model performance.
- **3D UMAP Visualizer:** Gain contextual insights into complex data drift by locating and identifying drift in high dimensional spaces.
- **Data Drift:** Easily monitor data drift, and compare data distributions between baseline and production datasets to assess how shifts in data impact model outcomes.
- **Prediction Drift and Impact:** Use popular drift metrics like Jensen-Shannon Divergence (JSD) and Population Stability Index (PSI) to uncover any data drift and help calculate which drifting features are impacting your model's predictions.
- **Data Integrity:** Uncover data integrity issues in your data pipeline causing models to underperform, such as missing data, range violations, and data type mismatches.
- **Feature Quality:** Pinpoint key features attributing to model outcomes for further deep dive into the model and inputs/outputs.
- **Class Imbalance:** Detect changes in low-frequency predictions due to class imbalance in each stage of your ML workflow.
- **Ground Truth Updates:** Update ground truth labels in a delayed, asynchronous fashion.
- **Alerts:** Configure and receive real-time alerts to identify and troubleshoot high-priority issues caused by performance, data drift, data integrity, and traffic metrics.

Analytics

Fiddler's model analytics tool provides you with rich model diagnostics to perform root cause analysis and gain actionable insights to create a continuous feedback loop for MLOps.

- **Dashboards:** Increase business alignment and confidence in decision-making by enabling teams across the organization to glean insights and connect ML metrics to business KPIs in a unified view.
- **Insights:** Build custom reports with the insights you need to gain deep understanding of your models and their impact on business outcomes, from monitoring metrics, feature impact, correlation, and distribution to partial dependence plot (PDP) charts.
- **Root Cause Analysis:** Drill down on problem areas to uncover the root cause of underperforming segments.
- **Segment Analysis:** Drill down into specific segments for targeted analysis, and find underperforming cohorts.
- **Model Validation:** Evaluate your model's performance and validate it before deploying it into production.
- **Fiddler Report Generator:** Create and share custom reports for periodic risk and compliance reviews.

Fairness

Build transparent, accountable, and ethical practices for your business with responsible AI. Fiddler increases visibility in AI governance with continuous monitoring, while detecting and mitigating bias in datasets and predictive models.

- **Algorithmic Bias Detection:** Detect algorithmic bias using powerful visualizations and metrics
- **Intersectional Bias Detection:** Discover potential bias by examining multiple dimensions simultaneously (e.g. gender, race, etc.).
- **Model Fairness:** Obtain fairness information by comparing model outcomes and model performance for each subgroup of interest.
- **Dataset Fairness:** Check for fairness in your dataset before training your model by catching feature dependencies and ensuring your labels are balanced across subgroups.
- **Fairness Metrics:** Use out-of-the-box fairness metrics, such as disparate impact, demographic parity, equal opportunity, and group benefit, to help you increase transparency in your models.

Generative AI Models and Applications

Fiddler Trust Service for LLM Monitoring and Guardrails

The Fiddler Trust Service enables enterprises to efficiently and securely score and monitor LLM applications at scale. It is powered by proprietary, fine-tuned Fiddler Trust Models, designed for task-specific, high-accuracy scoring of LLM prompts and responses with low latency. These models provide guardrails to detect and moderate risky prompts and responses while enabling comprehensive LLM monitoring and rich diagnostics for GenAI use cases. By scoring critical trust-related dimensions such as hallucinations, toxicity, PII leakage, and prompt injection attacks, Fiddler safeguards LLM applications and end users.

- **Fiddler Trust Models:** Quickly calculate Trust Scores on user prompts and LLM responses for LLM metrics monitoring.
- **Fiddler Trust Scores:** Evaluate multiple trust-related dimensions of prompts and responses, including: faithfulness, legality, hateful, racist, sexist, violent, harassing, sexual, harmful, unethical, jailbreaking content.
- **Fiddler Guardrails:** Moderates prompts and responses based on predefined thresholds using Fiddler Trust Models' scoring to protect LLM applications from security threats like prompt injection and adversarial attacks.

Monitoring and Analytics

Monitor drift of LLM-based embeddings in production and detect issues as soon as drift happens to minimize risks impacting users from adversarial model outcomes.

- **Actionable Alerts:** Get early warnings on performance of embeddings
- **Dashboards and Charts:** Pinpoint performance issues for deeper analysis on LLM outputs. Measure metrics, such as toxicity, costs, and safety.
- **Drift Monitoring:** Continuously detect dips in performance caused by data drift.
- **3D UMAP Visualizer:** Gain contextual insights into complex data drift by locating and identifying drift in high dimensional spaces.

Enterprise-level Scale and Security

Fiddler helps companies securely operationalize and scale model deployment.



Scalability

Accelerate the deployment of LLMs and ML models that require large volumes of data ingestion. Gain deeper insights with scalable baseline dataset ingestion.



Single Pane of Glass

Safeguard, monitor, and analyze your models within a unified platform with a single pane of glass. Easily view all your models in production and track what's most important to your business.



Choose Your Deployment

Choose where you deploy your models to meet your company's needs — it can be in the Fiddler cloud, your own cloud, within a virtual private cloud (VPC), or even in airgapped environments.



Security

Fiddler provides SOC 2 Type 2 security and HIPAA compliance. Users across the business will receive level-specific permissions to access protected environments through role-based access control (RBAC), and SSO.

Fiddler is the all-in-one AI Observability and Security platform for responsible AI. Monitoring and analytics capabilities provide a common language, centralized controls, and actionable insights to operationalize production ML models, GenAI, and LLM applications with trust. An integral part of the platform, the Fiddler Trust Service provides quality and moderation controls for LLM applications. Powered by cost-effective, task-specific, and scalable Fiddler-developed trust models — including air-gapped deployment for secure environments — it delivers the fastest guardrails in the industry.

Fortune 500 organizations use Fiddler to scale LLM and ML deployments to deliver high performance AI, reduce costs, and be responsible in governance.