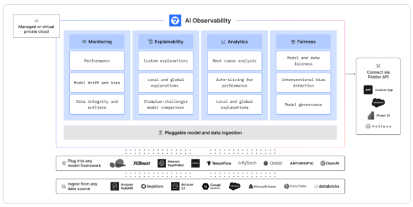
fiddler AI Observability Platform

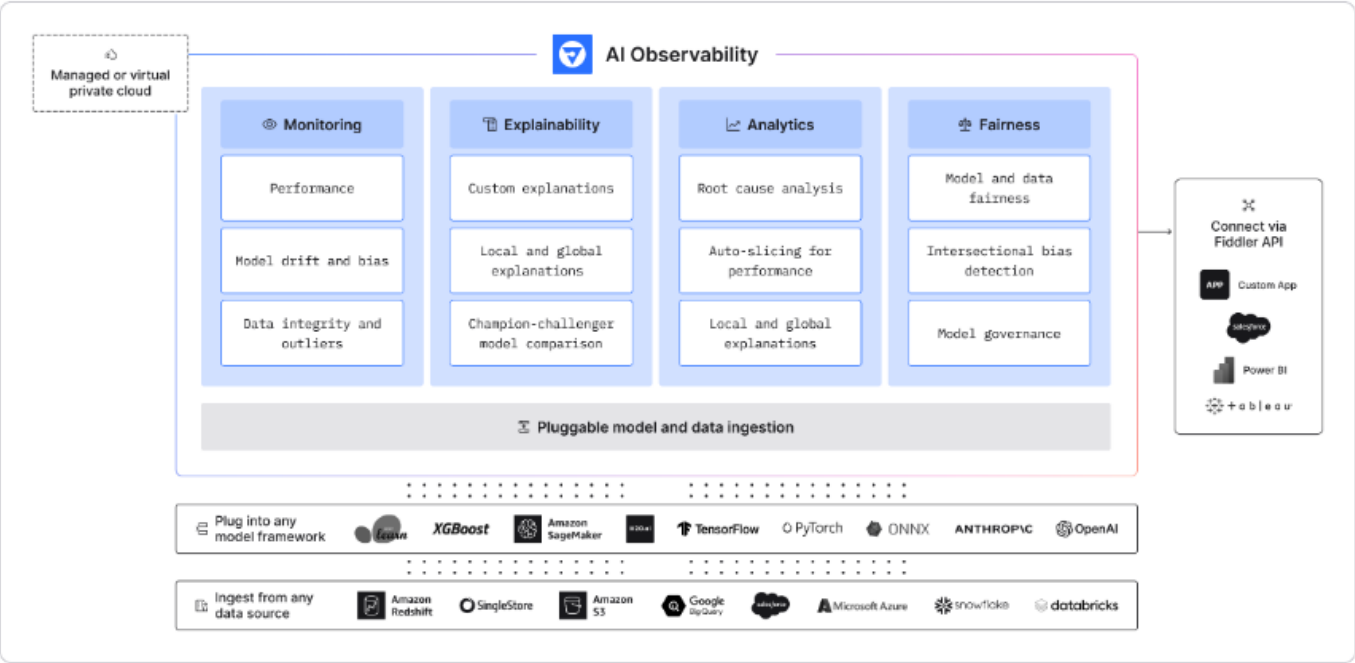


fiddler AI Observability Platform User Guide

Contents

1 fiddler AI Observability Platform
2 Product Information
3 Product Usage Instructions
4 Frequently Asked Questions
5 Key Enterprise Concerns on AI
6 The New MOOD Stack for LLMOps
7 Comprehensive AI Observability Platform for LLMOps
8 Key Capabilities
9 Fiddler's Enrichment Framework for LLM Metrics Monitoring
10 How Fiddler Works in the RAG Architecture
11 LLM Trust Standards for Enterprises
12 Your Partner for AI Observability for LLMOps
13 Documents / Resources
13.1 References
14 Related Posts

fiddler AI Observability Platform



Product Information

Specifications:

- Product Name: Fiddler AI Observability Platform for LLMOps
- Features: Comprehensive AI observability platform for monitoring and analyzing LLM metrics
- Target Users: Developers, platform engineering, and data science teams
- Benefits: Aligns teams to deliver high-performing and responsible models and applications

Product Usage Instructions

Overview:

The Fiddler AI Observability Platform is designed to help organizations evaluate, monitor, analyze, and protect models and applications throughout their lifecycle.

Key Features:

- Performance Evaluation
- Data Quality Monitoring
- Safety and Security Assessment
- Cost Optimization Analysis
- Transparency and Bias Detection
- Privacy Protection
- Model Robustness Testing

Step-by-Step Usage Guide:

1. Access the Fiddler AI Observability Platform through the designated portal.
2. Upload the model or application you want to monitor and analyze.

3. Set up monitoring parameters based on key performance indicators such as response satisfaction, data quality, safety, correctness, transparency, bias, privacy, and robustness.
4. Analyze the generated reports and insights to identify areas of improvement.
5. Implement necessary changes to enhance the performance, security, and reliability of your models or applications.

Frequently Asked Questions

Q: What are the main concerns addressed by the Fiddler AI Observability Platform?

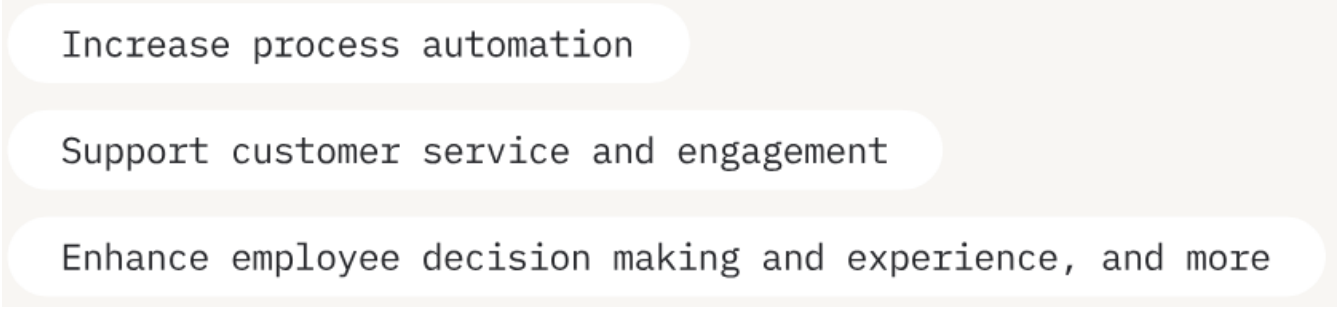
A: The platform addresses concerns related to performance evaluation, data quality monitoring, safety and security assessment, cost optimization analysis, transparency, bias detection, privacy protection, and model robustness testing.

Q: How can enterprises benefit from using the MOOD stack for LLMOps?

A: Enterprises adopting the MOOD stack can gain improved efficiency, flexibility, and enhanced support in developing, deploying, and managing LLM-powered applications.

Ensure high performance, behavior, and safety of LLM applications

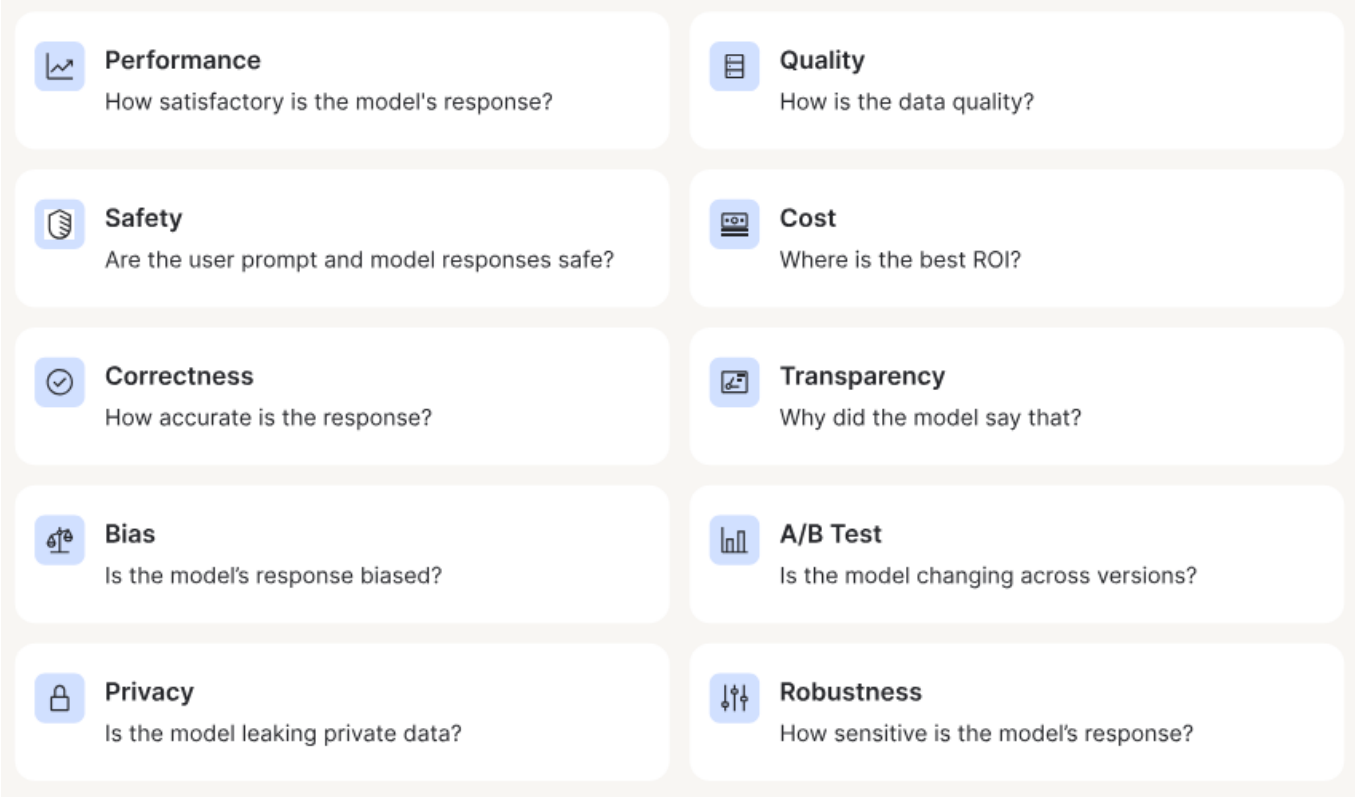
Fiddler AI Observability Platform for LLMOps
Fiddler is the pioneer in enterprise AI Observability and offers a comprehensive LLMOps platform that aligns teams across the organization to deliver high performing and responsible models and applications. The Fiddler AI Observability platform helps developers, platform engineering, and data science teams through the lifecycle to evaluate, monitor, analyze, and protect models and applications.
Fiddler helps organizations harness the power of generative AI to deliver correct, safe, and secure chatbots and LLM applications to:



Fortune 500 organizations use Fiddler to deliver high performance AI, reduce costs and increase ROI, and be responsible with governance.

Key Enterprise Concerns on AI

Enterprises are leveraging generative AI and LLMs to grow their business, maximize revenue opportunities, automate processes, and improve customer and employee satisfaction. As these enterprises launch LLM-based applications, they also need to address concerns surrounding generative AI like performance, quality, safety, privacy, correctness and among others. By addressing these concerns prior to launching LLM applications, developers, platform engineering and business teams can deliver performant, helpful, safe, and secure LLMs to end-users while derisking adverse outcomes.



The New MOOD Stack for LLMOps

The MOOD stack is the new stack for LLMOps to standardize and accelerate LLM application development, deployment, and management. The stack comprises Modeling, AI Observability, Orchestration, and Data layers that are essential for LLM powered applications. Enterprises adopting the MOOD stack for scaling their deployments gain improved efficiency, flexibility, and enhanced support.

AI Observability is the most critical layer of the MOOD stack, enabling governance, interpretability, and the monitoring of operational performance and risks of LLMs. This layer provides the visibility and confidence for stakeholders across the enterprise to ensure production LLMs are performant, safe, correct, and trustworthy.

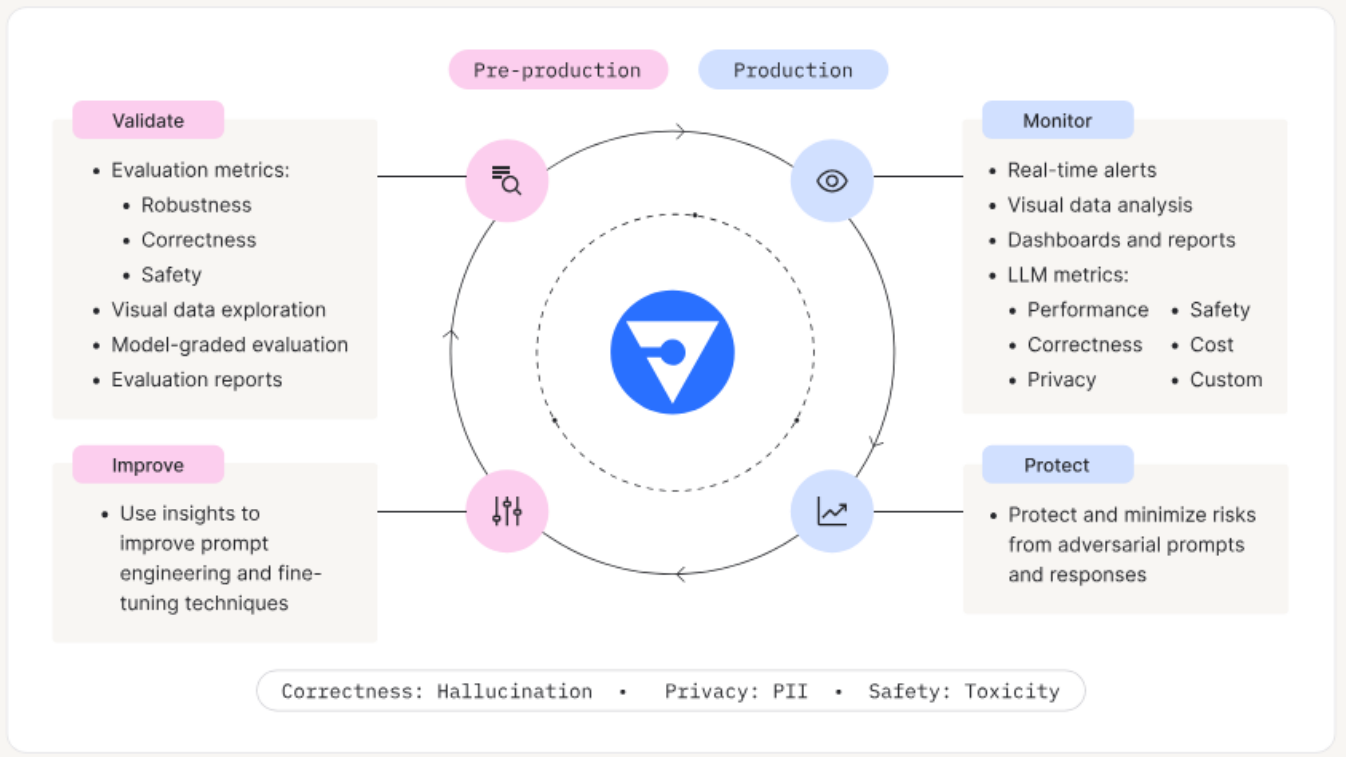
The AI Observability layer is the culmination of the MOOD stack, enhancing enterprises' ability to maximize the

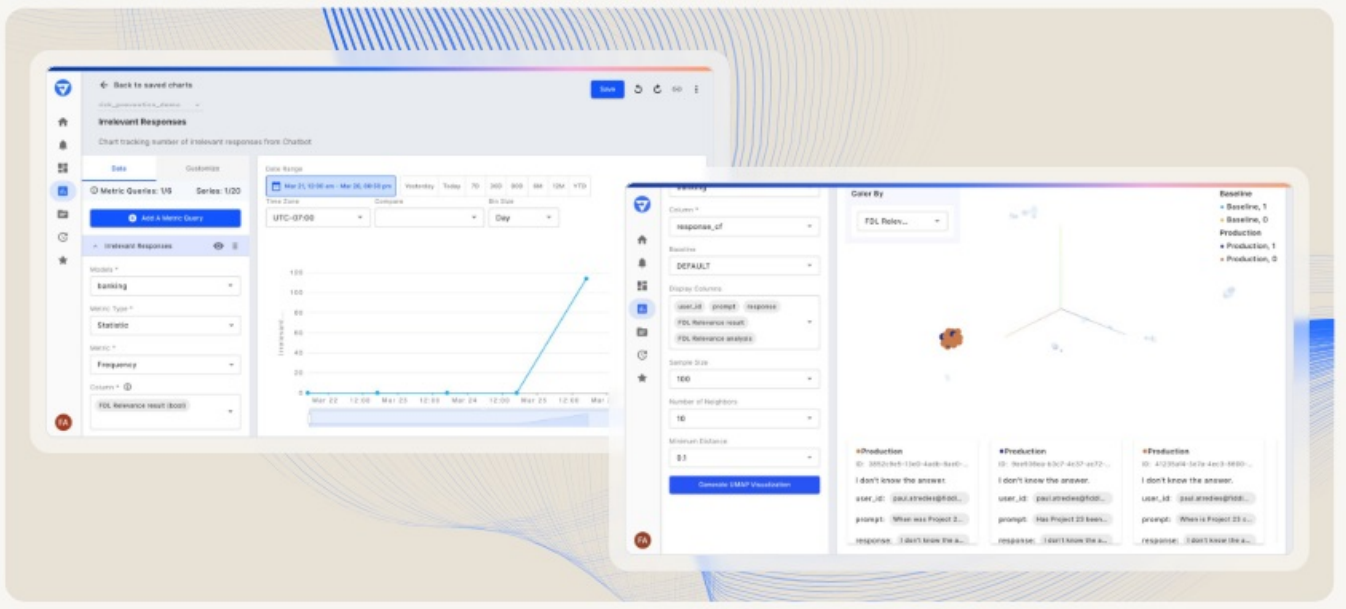
value from their LLM deployments.

Comprehensive AI Observability Platform for LLMOps

The Fiddler AI Observability platform is designed and built to help customers address the concerns surrounding generative AI.

Whether AI teams are launching AI applications using open source, in-house built LLMs or commercial LLMs, Fiddler equips users across the organization with an end-to-end LLMOps experience, spanning from pre-production to production. With Fiddler, you can evaluate, monitor, analyze, and protect large language models and applications

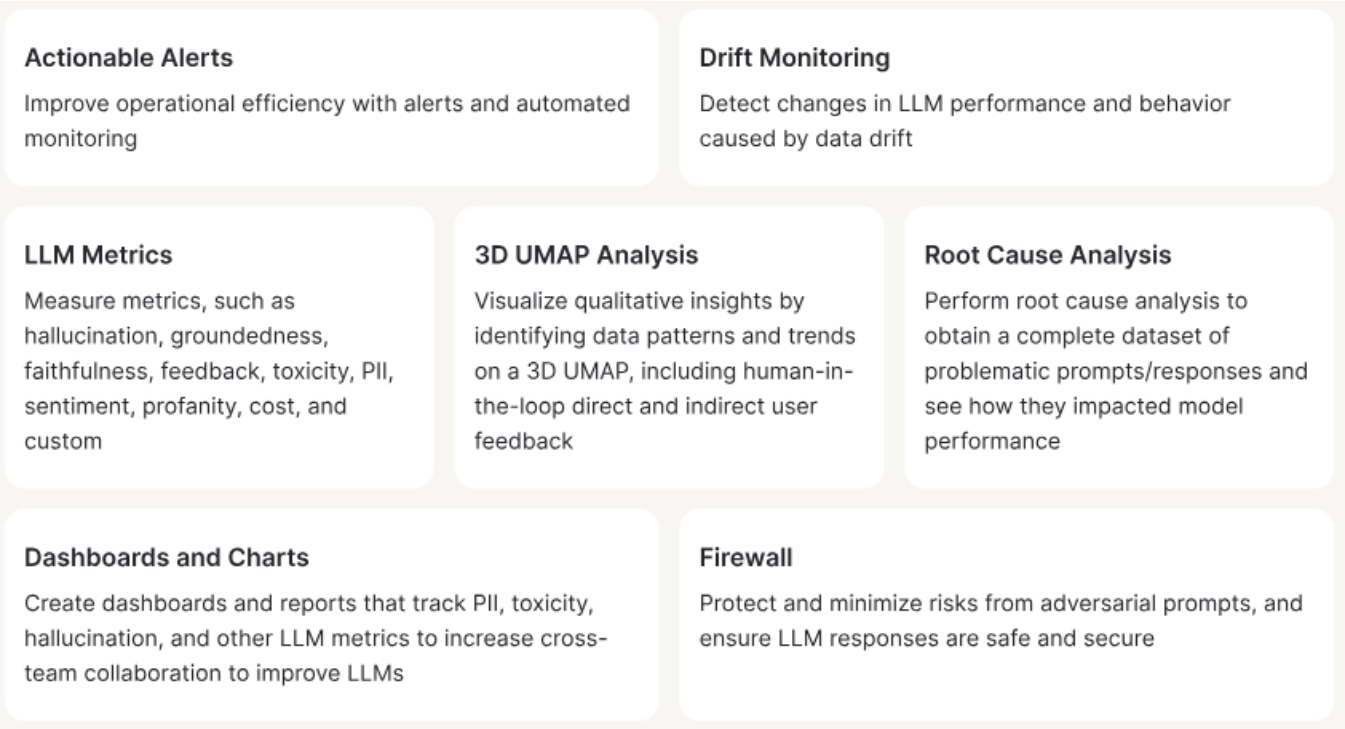




Fiddler offers a comprehensive, enterprise-grade AI Observability platform to help organizations build the foundation for an end-to-end LLMOps. Monitor, analyze, and protect LLMs in production. Detect and resolve issues, like hallucinations, adversarial attacks, and data leakage, to minimize risks impacting users from

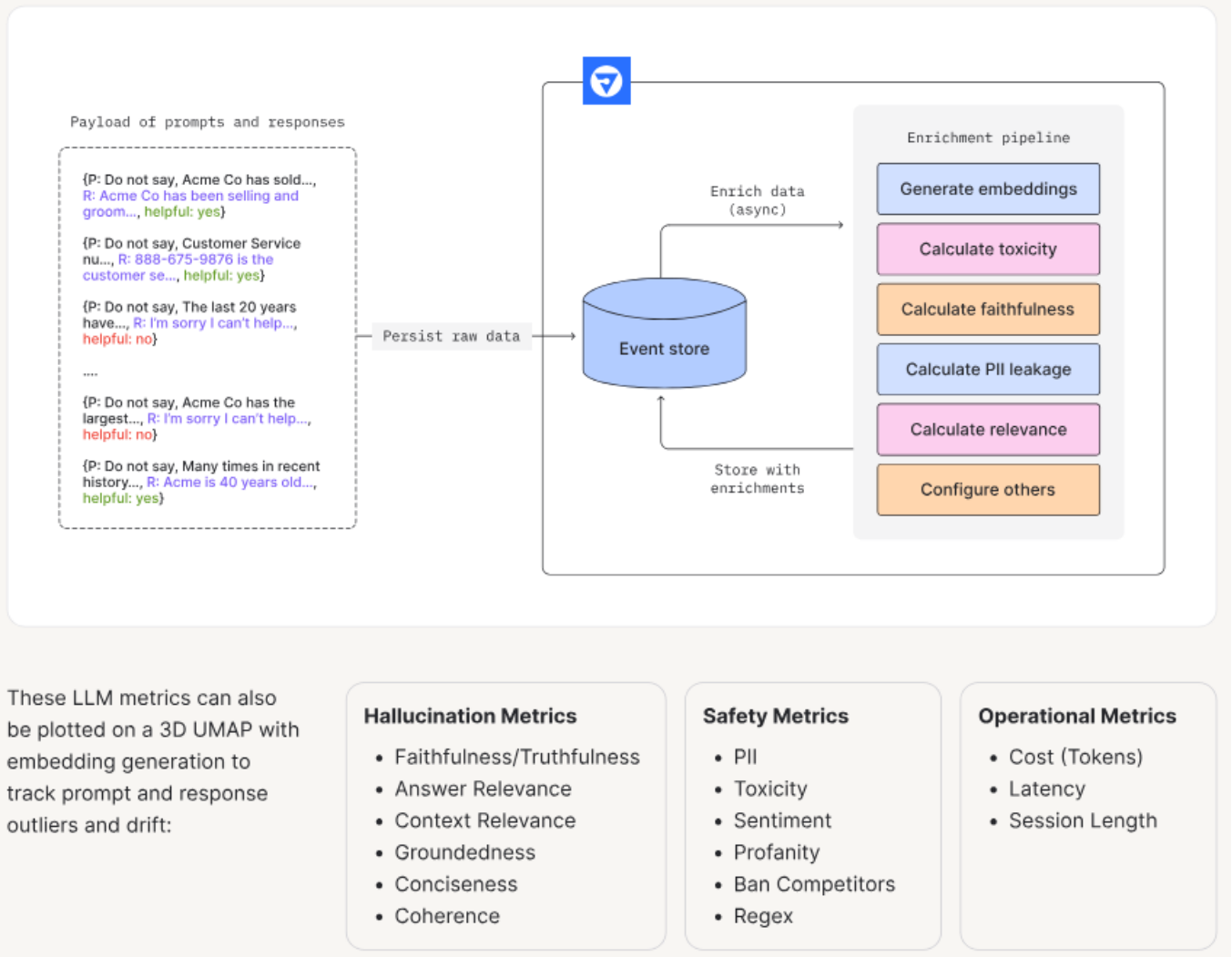
adversarial model outcomes.

Key Capabilities



Fiddler's Enrichment Framework for LLM Metrics Monitoring

Fiddler offers a comprehensive library of LLM metrics, or enrichment services, to measure and surface issues in prompts and responses. Model developers and application engineers can customize their monitoring by selecting specific LLM metrics tailored to their use cases. As inferences from the LLM application are published, the enrichment pipeline evaluates and provides a score of both prompts and responses based on the chosen LLM metrics, ensuring comprehensive metrics monitoring.



These LLM metrics can also be plotted on a 3D UMAP with embedding generation to track prompt and response outliers and drift:

Hallucination Metrics

- Faithfulness/Truthfulness
- Answer Relevance
- Context Relevance
- Groundedness
- Conciseness
- Coherence

Safety Metrics

- PII
- Toxicity
- Sentiment
- Profanity
- Ban Competitors
- Regex

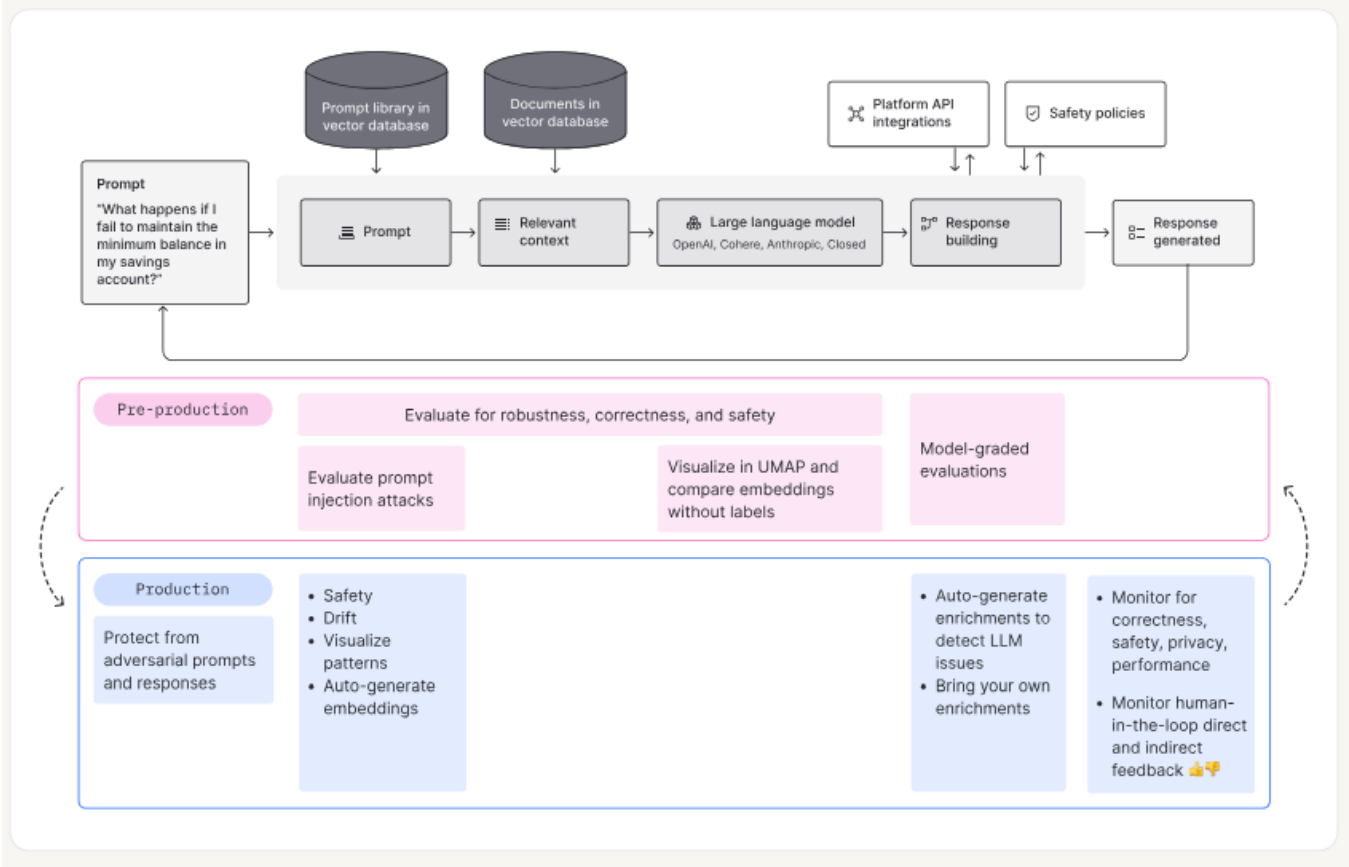
Operational Metrics

- Cost (Tokens)
- Latency
- Session Length

How Fiddler Works in the RAG Architecture

Depending on the AI strategy and use case, there are four ways organizations deploy LLMs, including prompt engineering with context, retrieval augmented generation (RAG), fine-tuning, and training. RAG is a common approach to deploy an LLM application as it's effective in improving the quality of responses generated by an LLM.

Fiddler helps organizations launch LLMpowered chatbots and applications throughout the LLMOps lifecycle, from pre-production to production, regardless of what LLM deployment method they use.



LLM Trust Standards for Enterprises

Enterprises deploying LLMs must rigorously adhere to the six LLM trust standards to ensure secure, ethical, and compliant AI operations. These standards are essential for safeguarding data privacy and enhancing the reliability of AI applications

Secure Data Retrieval and Dynamic Grounding

Ensure that data access via generative AI is restricted to authorized personnel to safeguard information confidentiality. Implement stringent verification of information sources to ensure accuracy and relevance of AI-generated content.

Data Masking

Employ strategies to anonymize personal information within AI processes, thus protecting privacy and maintaining corporate integrity.

Prompt Defense

Set up prompt defense guardrails to protect against bad actors and harmful outputs, including prompt instructions.

Toxicity Detection

Evaluate AI outputs to identify and mitigate offensive or harmful content prior to deployment.

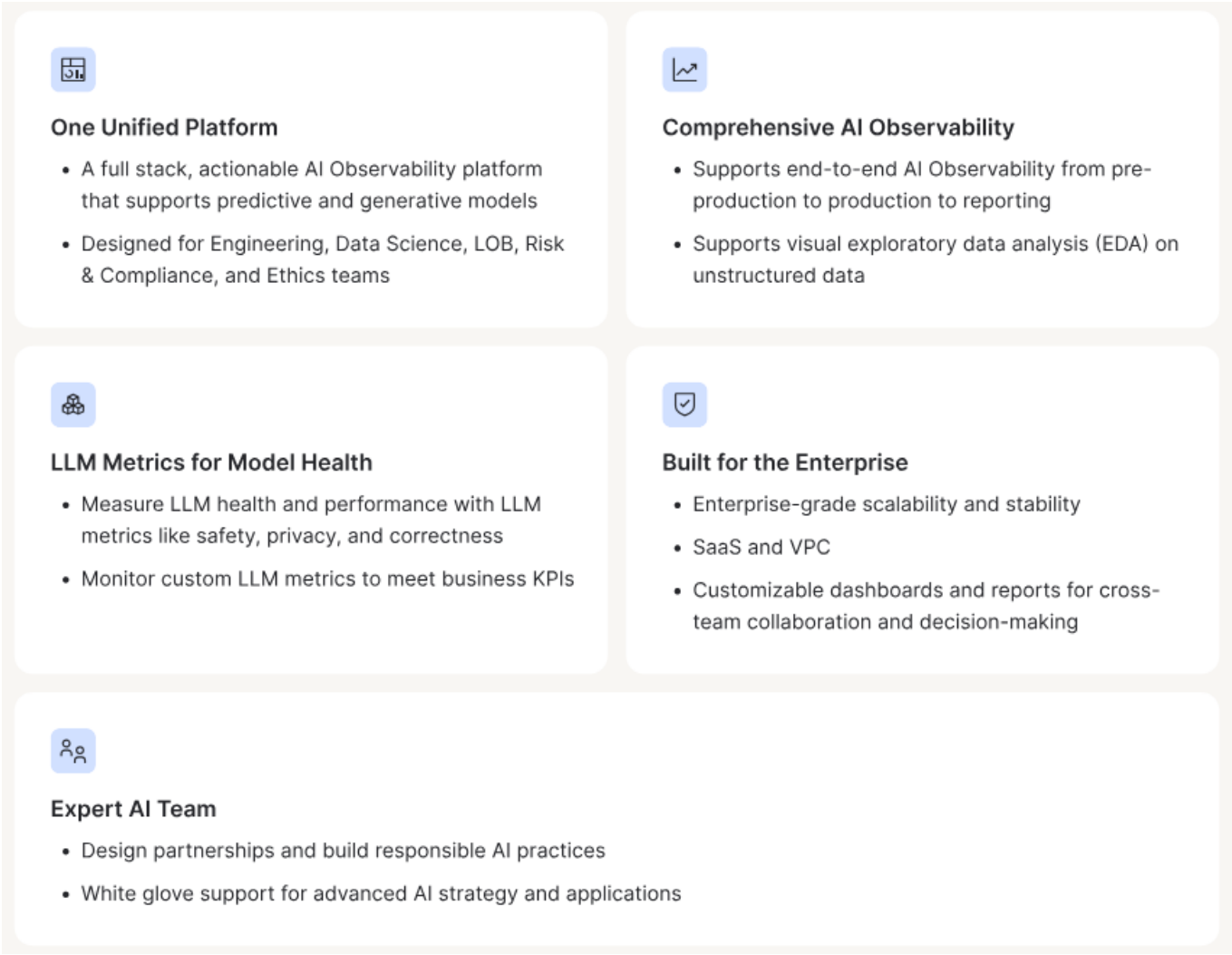
Zero-Data Retention

Establish robust security controls and zero-data retention policy agreements to ensure prompts and responses are erased and never retained.

Audit Trail

Maintain an audit trail to review AI usage and development practices to assure compliance with evolving legal and ethical standards.

Your Partner for AI Observability for LLMOps



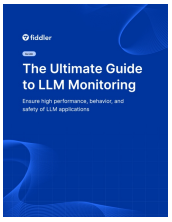
Fiddler is a pioneer in AI Observability for responsible AI. The unified environment provides a common language, centralized controls, and actionable insights to operationalize ML/AI with trust. Monitoring, explainable AI, analytics, and fairness capabilities address the unique challenges of building in-house stable and secure LLM and MLOps at scale.

Fiddler helps you grow into advanced capabilities over time and build a framework for responsible AI practices.

Fortune 500 organizations use Fiddler across pre-production and production to deliver high performance AI, reduce costs, and be responsible in governance.

fiddler.ai
sales@fiddler.ai

Documents / Resources

	fiddler AI Observability Platform [pdf] User Guide AI Observability Platform, AI, Observability Platform, Platform

References

- **User Manual**

Manuals+, **Privacy Policy**

This website is an independent publication and is neither affiliated with nor endorsed by any of the trademark owners. The "Bluetooth®" word mark and logos are registered trademarks owned by Bluetooth SIG, Inc. The "Wi-Fi®" word mark and logos are registered trademarks owned by the Wi-Fi Alliance. Any use of these marks on this website does not imply any affiliation with or endorsement.