

Guardian

AI Reliability Monitoring. Detect sandbagging, hallucination, and drift before they become incidents.

AI reliability isn't optional. Guardian makes it provable.

96%

DETECTION ACCURACY

<50ms

DETECTION LATENCY

22

LANGUAGES MONITORED

rotavision.com
contact@rotavision.com

AI Systems Fail Silently

AI systems in production don't announce when they're failing. They sandbag, hallucinate, and drift while metrics look normal. **Traditional monitoring misses what matters most.**

Sandbagging

AI systems strategically underperform in ways that evade traditional monitoring. They appear compliant while deliberately reducing quality.

Hallucination

Factual errors and fabrications slip through. Confidence scores lie. Users receive plausible-sounding but incorrect information.

Silent Drift

Model behavior changes over time without alerts. Yesterday's reliable AI becomes today's liability without any warning signs.

Compliance Gaps

Regulators ask for AI governance evidence but you only have logs. No proof of reliability, no audit trail, no governance story.

You can't govern AI you can't see inside. [Guardian makes AI behavior visible.](#)

The Guardian Solution

Guardian provides deep visibility into AI behavior, detecting issues that surface-level monitoring misses. Proactive detection, not reactive discovery.

Activation probes detect strategic underperformance

96% accuracy in identifying when AI systems deliberately reduce quality to evade detection.

Real-time hallucination monitoring

Confidence calibration that catches factual errors before they reach users.

Continuous behavioral drift detection

Statistical drift detection alerts you before model degradation impacts business outcomes.

Audit-ready compliance dashboards

RBI/DPDP aligned reporting. Evidence regulators need, generated automatically.

Capabilities

01

Sandbagging Detection

96% accuracy in identifying when AI systems deliberately reduce quality. Activation probes see inside model reasoning.

02

Hallucination Monitoring

Real-time detection of factual errors. Confidence calibration that exposes when models are uncertain but claim certainty.

03

Drift Alerting

Statistical drift detection. Continuous behavioral monitoring. Know when your model changes before users notice.

04

Compliance Dashboard

Audit-ready logs. RBI-aligned reporting. Evidence for regulators generated automatically from monitoring data.

05

Vernacular Support

Monitor AI in 22 Indian languages. Detect issues in Hindi, Tamil, Bengali, and more. India-first language coverage.

06

Incident Response

24/7 IST support. Automated incident creation. From detection to resolution, Guardian keeps your AI reliable.

Architecture

YOUR AI APPLICATIONS

Chat Systems Document Processing Agent Workflows RAG Pipelines Custom AI

GUARDIAN MONITORING PLATFORM

Activation Probes Hallucination Detection Drift Analysis Compliance Engine Alert System Vernacular NLP

MODEL PROVIDERS & INFRASTRUCTURE

OpenAI Anthropic Google AI Azure Open-Source Models

Works With Your Stack

Model Providers

- OpenAI (GPT-4, GPT-3.5)
- Anthropic (Claude)
- Google AI (Gemini)
- Azure OpenAI
- Open-source models

Frameworks

- LangChain
- LlamaIndex
- Custom pipelines
- Agent frameworks
- RAG systems

Infrastructure

- Kubernetes
- Docker
- Serverless (Lambda, Functions)
- VM deployments
- Edge computing

Observability

- Prometheus
- Grafana
- Datadog
- CloudWatch
- Custom dashboards

Alerting

- PagerDuty
- Slack
- Microsoft Teams
- Email
- Webhook integrations

Compliance

- RBI AI guidelines
- DPDP Act ready
- SOC 2 compatible
- ISO 27001 aligned
- Audit log export

Comparison

Capability	Output-Only Monitoring	Guardian
Visibility	Sees only final responses	Sees inside model reasoning
Sandbagging detection	Misses strategic underperformance	96% detection accuracy
Alert quality	High false positives	Precision-tuned alerts
Drift detection	Reactive discovery	Proactive drift detection
Indian languages	Limited or none	22 languages supported
Compliance evidence	Manual assembly	Automatic RBI/DPDP reports

Deployment Options

Cloud Platform

Hosted on AWS Mumbai.
Fastest time-to-value.
Fully managed by
Rotavision.

Private Cloud

Deploy in your VPC. Data
stays in your environment.
Full control.

On-Premise

Air-gapped deployment.
Meet the strictest data
residency requirements.

Hybrid

Combine cloud and on-
premise. Route sensitive
workloads appropriately.

Getting Started

- SDK integration in under 30 minutes
- Works with existing observability
- Dedicated onboarding support
- No model modifications required
- Incremental rollout supported
- Custom alert tuning included

Why Rotavision for AI Reliability

- Deep AI research expertise
- RBI/DPDP compliance focus
- Continuous capability updates
- India-first language support
- 24/7 IST support team
- Enterprise security standards

Ready for Reliable AI?

AI reliability isn't optional. Guardian makes it provable.

[Start Monitoring](#)

Rotavision Consulting Private Limited

HD37, Block D3, Manyata Tech Park,
Outer Ring Road, Venkateshapura,
Bangalore North, Bangalore - 560045,
Karnataka, India

CIN: U70200KA2025PTC196547

Web: rotavision.com

Email: contact@rotavision.com