luma

AI Governance & Oversight

How we govern the AI systems in our platform, from change control through production monitoring

Change Management

Every change to prompts, models, or agent behavior follows a controlled software engineering process. No ad-hoc edits are made in production.

- All prompts and model selections are version-controlled and changes are restricted to engineering staff
- Changes move through a staged pipeline:

development → staging → production

(peer-reviewed, minimum two approvers)

- Production releases occur only during published deployment windows
- Larger changes are validated in an isolated pre-production environment and rolled out gradually via canary release (by customer or traffic percentage) before full deployment

Evaluation & Quality Assurance

Every release is tested against defined conversational scenarios before it reaches customers.

- Automated test suites run against every staging deployment, covering accuracy, expected behavior, and edge cases
- Independent security-authored guardrail tests run in the same pipeline to validate safety boundaries
- A dedicated evaluation team reviews test and conversation performance results and must sign off before any release proceeds

Production Monitoring

AI performance is monitored continuously in real time, not just at release.

- Live dashboards track call volume, latency, session health, error and timeout rates, intent handling, hand-off rates, and escalation rates
- Automated alerts notify our engineering and on-call teams the moment anomalous behavior is detected

Auditability & Traceability

Every interaction can be fully reconstructed, including what an AI agent saw, what it did, and why.

- Complete message history and every tool/action call for each engagement is retained and available to customers
- A Post-Call Analysis is generated for every engagement, summarizing what happened
- All behavioral changes are traceable through version control; end-user customizations are captured through the same audit trail

Human Oversight & Clinical Guardrails

AI agents operate within clearly defined boundaries and are built to hand off to a human when needed.

- Every agent includes escalation logic at both the platform and agent level, triggered by defined behaviors or conditions
- On escalation, the agent stops and the outcome is routed per policy: notify staff, trigger an EHR signal, or end the engagement with a defined message
- All agents honor Do-Not-Contact (DNC) flags without exception

Security and trust

Luma follows internationally-recognized security and privacy standards.

This overview reflects our current governance controls and is updated as our processes evolve.

Updated 09.21.2026

Learn more about how Luma can help your organization at info@lumahealth.io

luma™