Safe AI & Data Protection for Credit Unions & Community Banks

Deploying AI is a shared responsibility: Eltropy provides the platform, safeguards, and transparency, while the financial institution (FI) sets the policies, controls what data the AI can access, and oversees outcomes. This guide describes how the platform protects FI data and the controls available to the institution.

Quick Reference

What does the AI do, and what data does it process?	Section 1: AI Functionality & Scope
Is our data used to train AI models? Are prompts retained?	Section 2: Customer Data & Model Training
What safeguards and guardrails constrain the AI?	Section 3: AI Safety & Guardrails
How can we review AI decisions and trace what the AI did?	Section 4: Transparency & Explainability
Where does our data live? Is it protected and kept separate from other clients'?	Section 5: Data Protection by Design
What controls do we have? Can AI features be disabled?	Section 6: Shared Responsibility & Controls
How is the AI tested, governed, and updated over time?	Section 7: AI Governance & Compliance

1. AI FUNCTIONALITY & SCOPE

This section describes what Eltropy's AI does: the types of AI the platform uses, the content it generates, and the categories of data it processes.

Types of AI Used

Eltropy is a digital conversation platform for credit unions and community banks with AI-enabled member communications across text, chat, voice, video and email. The platform combines several types of AI, each suited to different tasks: generative AI and large language models (LLMs), natural language processing (NLP), machine learning, agentic AI, and deterministic, rule-based workflows. These are applied selectively by risk. Lower-risk, informational tasks may rely on generative AI and NLP, while higher-risk tasks-such as those affecting accounts, balances, or money movement-use more deterministic, rule-based logic with human oversight.

What Content the AI Generates

The AI generates conversational text and synthesized voice (text-to-speech) responses-not images, video, or software code- and it is limited to the workflows your institution approves. These outputs run through the safeguards in Section 3: AI Safety & Guardrails, with human review and escalation set by your configuration (see Section 6: Shared Responsibility & Controls).

What Data the AI Processes

The AI processes the following categories of data to respond to inquiries and support staff:

- Natural-language prompts and queries submitted by members, customers, or staff.
- Authenticated member and customer personal information, such as names, mobile numbers, and email addresses.
- Financial reference numbers, such as account or loan numbers and balances.
- Text extracted from FI-provided materials, including website sitemaps (.XML) and policy documents (.PDF).
- ***Excluded data***: Raw payment card industry (PCI) cardholder data—such as full card numbers and CVV codes—is tokenized by Eltropy's third-party payment service providers and is never transmitted to or processed by the AI models.

2. CUSTOMER DATA & MODEL TRAINING

Eltropy does not use customer data to train, fine-tune, or improve third-party AI models (e.g., OpenAI, Google, Microsoft models). This is a contractual commitment enforced at the model provider level.

How It Works

Eltropy uses AI models from Microsoft Azure, OpenAI, and Google Gemini. When members, customers, or staff interact with Eltropy's AI, their queries pass through the system to the model to generate a response, and that data is processed for your institution's use only.

How Your Data Is Handled by Model Providers

Your data is not used to train or improve any foundation model or third-party product, and is not available to other customers or to OpenAI, Google, or Microsoft for their own product development. Customer prompts are processed in memory on a per-request basis and discarded immediately after execution.

What About Fine-Tuning?

Fine-tuning is a common practice where vendors adjust AI behavior based on customer interactions. Where fine-tuning occurs, it is performed solely for your institution's benefit: it is specific to your deployment and used only to improve AI's performance for you. It is not used to train or improve third party foundation models, and it is not shared with or used to benefit other customers. Your data stays yours.

3. AI SAFETY & GUARDRAILS

Eltropy's AI doesn't rely on a single safeguard. Instead, we layer protections at four levels—from the foundation of the model itself, through filters and rules, to application design and user education.

© 2026 Eltropy, Inc. All Rights Reserved.

Layer 1: Model Foundation
Before any data reaches Eltropy's application, we vet the underlying AI model (Azure OpenAI, Google Gemini)-assessing its training methods, data sources, and built-in safety measures. The Azure OpenAI and Gemini models include their own safeguards for hallucination prevention and prompt-injection defense, which Eltropy builds on with the guardrails described below. This is our first line of defense against bias baked into the model itself.

Layer 2: Programmable Guardrails (Input & Output Filters)
Everything a member or customer types passes through Eltropy's input pipeline before reaching the AI model. Automated filters screen for profanity and harmful content, while PII and sensitive data are extracted through a separate pipeline so they are not stored in conversation history. Prompt-injection attacks are defended through application-level input sanitization, controlled prompt templates, and filtering of untrusted user or external inputs. After the model responds, its output passes through similar filters to catch any sensitive data that may have been inadvertently included.

Layer 3: Application Design
Eltropy's application architecture includes built-in safeguards: retrieval-augmented generation (RAG) grounds AI answers in approved knowledge rather than allowing free reasoning, which guards against hallucination; confidence thresholding ensures low-confidence answers escalate to a human rather than misleading the member or customer; output validation checks that responses make sense.

Layer 4: Product Positioning & User Education
You stay in control of how the AI shows up for your members and customers. The platform can disclose when someone is interacting with AI and offers escalation paths to your staff-you decide how disclosure appears and define those paths, so no one is confined to an AI interaction. Staff training is a shared effort: Eltropy is transparent about how the platform works, and you bring the deep knowledge of your members and your team.

4. TRANSPARENCY & EXPLAINABILITY

AI decisions can be reviewed and traced. Every interaction is logged, AI answers include citations and sources, and decision trails are retained for review and compliance.

What Gets Logged

- Interaction Records: Member and customer inquiries, AI responses, outcomes, and timestamps are retained so you can review what the AI did and why.
- Decision Trails: When the AI takes an action (e.g., reverses a fee), that decision is logged with verification of authority and the reasoning behind it.
- Escalations: Any time the AI escalates a question or decision to staff, that escalation is documented.

Access for Your Teams
Your compliance and legal teams can access activity logs and decision trails to review AI behavior and outcomes. You can examine individual interactions, monitor performance metrics, and confirm that the AI is operating within the boundaries you set. Both Eltropy and your institution review these logs and monitor for bias or unintended outcomes; Eltropy's bias review process and mitigations are described in Section 7: AI Governance & Compliance.

5. DATA PROTECTION BY DESIGN

Customer data—and the personal information your members and customers share through interactions—is protected across the data lifecycle: upload, storage, and processing. Eltropy applies a "privacy by design" approach, meaning data protection is incorporated into the system's design rather than added afterward.

Storage & Encryption
Customer data and institutional documents are stored in U.S.-based Amazon Web Services (AWS) data centers.

- Data at rest: Encrypted with AES-256
- Data in transit: Encrypted with TLS 1.2
- Access controls: Limited to authorized Eltropy personnel only
- Activity logs: Changes made to the AI agents configuration by Users of the platform are logged and stored for up to 90 days
- Tenant Isolation: Customer data is not commingled with other clients' data.

Each customer instance is logically isolated across Eltropy's storage and application layers, with a unique branded URL and dedicated document stores that restrict the AI to that tenant's authorized files.

PII & Sensitive Data Handling
Eltropy uses AI to extract the data needed to answer inquiries and support staff. Because this extraction runs through a separate pipeline, the extracted PII and sensitive data-names, addresses, phone numbers, government IDs, SSNs, and payment card data-are not stored in conversation history.

© 2026 Eltropy, Inc. All Rights Reserved.

What Eltropy Employees Can See

Eltropy teams have access only to what's necessary to operate and support the platform, governed by role-based access controls (RBAC).

6. SHARED RESPONSIBILITY & CONTROLS

The AI does not operate independently. Responsibilities are divided between your institution and Eltropy. Eltropy provides the platform, guardrails, and oversight. Your institution sets the policies and approved knowledge and decides what the AI can access, what it can recommend, and what requires human approval and validation.

What Your Institution Controls

- Knowledge Base: You decide which documents, policies, and information the AI has access to
- Approved Actions: You define what the AI can do autonomously vs. what requires approval or escalation.
- Configuration: You configure system prompts, tone, escalation rules, and business logic*
- Testing & Validation: You test scenarios before launch and monitor outcomes continuously
- Review & Monitoring: You have full access to AI decision logs and activity records to review and monitor outcomes

**AI capabilities are supplemental to the platform and are configurable by use case. In general, your institution can enable or disable specific AI features through system administration, and alternative configurations—such as integrating a Bring Your Own Bot (BYOB) engine—may be available depending on your deployment. The specific features that can be toggled, and the level at which they can be controlled, depend on the use case and deployment; Eltropy can confirm the available options with you during onboarding.*

What Eltropy Controls

- Platform Security: Infrastructure, data encryption, access controls, disaster recovery
- Guardrails: Input/output filters, bias detection, PII handling, prompt injection defense
- Oversight: Continuous monitoring, incident response, bias reviews
- Compliance: Regulatory adherence, SOC 2 audits, third-party assessments

7. AI GOVERNANCE & COMPLIANCE

Eltropy's controls are independently verified. Eltropy undergoes independent audits and maintains documented governance demonstrating that the controls outlined above are implemented and effective.

Audit Report

Eltropy maintains SOC 2 Type II, with an annual audit by an independent firm verifying controls over security, confidentiality, and availability.

Testing & Monitoring

Eltropy tests, validates, and verifies AI before changes reach customers: when an upstream vendor issues a model upgrade, Eltropy runs functionality tests against an internal, customer-specific "golden dataset" to confirm the AI is performing as expected before release. This validation runs with every upstream release. Disparate impact and fairness reviews are conducted by the AI model providers on each model they release. Eltropy commissions third-party penetration tests of AI endpoints annually, alongside quarterly internal tests; the due diligence package includes the latest results.

Monitoring in Production

In production, AI behavior is monitored through application error tracking, confidence thresholding that automatically escalates low-certainty queries to live agents, and output similarity comparison.

Addressing Bias, Unfair Outcomes, and Inappropriate Content

Addressing bias, unfair outcomes, and inappropriate content is a shared responsibility between your institution and Eltropy. Eltropy relies on third-party foundational Large Language Models (LLMs) provided by Microsoft Azure OpenAI and Google Cloud (Gemini). These foundational models are pre-trained by OpenAI and Google on public datasets and aligned for neutrality using Reinforcement Learning from Human Feedback (RLHF). On Eltropy's side, these risks are further mitigated by grounding responses through retrieval-augmented generation (RAG) in your approved sitemaps and documents, enforcing automated profanity and harmful-content filters, and maintaining mandatory human-in-the-loop review for consequential actions. By default, Eltropy restricts AI to assistive workflows, and because the platform is configurable, your institution can tailor these settings to its own policies and risk tolerance.

Change Management and Notifications

AI model and feature changes follow Eltropy's formal change management procedure and are delivered on a monthly release cycle; your institution can view upcoming changes through the Eltropy application. Before deployment, you receive a dedicated user acceptance testing (UAT) staging environment to review and test updates, and can opt into new model versions and upgrades. When a specific model is scheduled to be deprecated, Eltropy notifies your institution ahead of time so you can transition to a newer supported model before the older version is retired.

© 2026 Eltropy, Inc. All Rights Reserved.