



AI Governance Statement

Artificial Intelligence Usage, Data Handling, and Oversight

Version: 1.1 Date: August 2026 Classification: Public

Executive Summary

ScamZero uses AI to analyze potentially fraudulent content submitted by residents. This document explains what AI we use, how data is handled, and our governance practices.

- **AI Provider:** Cloudflare Workers AI (model: Gemma 4, open weights, run on Cloudflare’s network)
- **Data Usage:** Ephemeral processing only - no data retained by AI provider
- **Training:** Submitted content is NOT used to train AI models
- **Oversight:** Human review process for edge cases and quality assurance

AI Technology Used

Cloudflare Workers AI

ScamZero runs its analysis on Cloudflare Workers AI, using the open-weights Gemma 4 model. Inference happens on Cloudflare’s network. Google publishes the model weights; Google does not receive or process submissions.

Attribute	Details
Provider	Cloudflare, Inc.

Attribute	Details
Model	Gemma 4 (open weights, run on Cloudflare Workers AI)
Access Tier	Enterprise / Paid API
Certifications	SOC 2 Type II, ISO 27001, FedRAMP

Why Workers AI?

- **Multimodal:** Can analyze text, images, and screenshots
- **Published data policy:** Cloudflare states that it does not retain Workers AI inputs or outputs and does not use them to train any model
- **Compliance:** Cloudflare maintains SOC 2 Type II and ISO 27001 covering the infrastructure the model runs on, and its published Workers AI data policy states that inputs are not retained and are not used to train any model
- **Performance:** High accuracy for scam detection patterns

How AI Is Used

Scam Analysis Workflow

When a resident submits content for analysis:

1. A resident submits text, email, or screenshot via the portal
2. Content is sent to Cloudflare Workers AI for analysis
3. AI evaluates content against known scam patterns
4. Risk assessment and educational guidance are generated
5. Results are displayed to the resident immediately

AI Role: Advisory Only

The AI provides **educational guidance** to help residents recognize potential scams. It does not make financial decisions, block transactions, or take automated actions on anyone’s accounts.

What AI Analyzes

- Suspicious text messages or emails
- Screenshots of potentially fraudulent communications
- Website URLs for phishing indicators
- Caller ID information and phone scam patterns

Data Handling & Privacy

Data Flow

Stage	Data Handling	Retention
Resident submission	Received by ScamZero API	Deleted on first view; 1 hour max; 15 min if no link requested
AI Processing	Sent to Cloudflare Workers AI	Not retained by the model service
Analysis Results	Stored in ScamZero database	Deleted on first view; 1 hour max; 15 min if no link requested

Cloudflare's Workers AI Data Policy

Under Cloudflare's Workers AI data policy:

- ✓ Submitted content is **not used to train any model**
- ✓ Data is processed only to provide the requested service
- ✓ Data is not retained after the API request completes
- ✓ No human review of submitted content by the model service — and none is possible at ScamZero either, because nothing is retained to review

Key Privacy Commitment

Resident-submitted content is never used to train AI models. Neither ScamZero nor Cloudflare uses this data for model improvement or training purposes.

AI Governance & Oversight

Human Oversight

ScamZero maintains human oversight of AI operations:

- **Prompt Engineering:** AI instructions are carefully crafted and reviewed
- **Output Review:** Regular sampling of AI responses for quality assurance
- **Edge Cases:** Human review process for ambiguous or unusual content
- **Feedback Loop:** User feedback incorporated to improve prompts

Bias & Fairness Considerations

- AI is used for scam pattern recognition, not decisions about individuals
- No automated decisions affecting anyone's accounts or access
- Regular review of AI outputs for consistency and accuracy
- Clear disclosure to residents that AI is being used

Limitations Acknowledged

We acknowledge that AI systems have limitations:

- May not detect all scams, especially novel patterns
- Results are advisory and should not replace the person’s own judgment
- Members are encouraged to contact their bank or their local police department

Regulatory Compliance

Relevant Frameworks

Framework	Relevance to ScamZero AI
FERPA	No student education record is ever received; resident-submitted content only
CCPA/State Privacy Laws	Residents control what they submit; clear disclosure provided
EU AI Act (if applicable)	Advisory tool, not high-risk automated decision-making
NIST AI Risk Management Framework	Cloudflare maintains SOC 2 Type II and ISO 27001 covering the infrastructure the model runs on

Transparency

ScamZero is committed to AI transparency:

- ✓ Residents are informed that AI is used for analysis
- ✓ Clear explanation of what AI does and doesn't do

- ✓ This governance document publicly available
- ✓ Privacy policy describes AI data handling

Ongoing Commitments

- **Monitoring:** Continuous monitoring of AI performance and accuracy
- **Updates:** Regular updates to AI prompts based on emerging scam patterns
- **Compliance:** Ongoing monitoring of AI regulatory developments
- **Documentation:** This statement reviewed and updated at least annually

Questions?

For AI governance inquiries, contact security@scamzero.com

ScamZero

AI Governance Statement

Version 1.1 | August 2026