Security and Data Protection White Paper

Version: 1.0
Effective Date: 31/11/2024
Last Updated: 05/10/2025
Contact: admin@reapapp.io

Introduction

At REAP ("Company", "We", "Us", "Our"), we are committed to delivering an AI-powered widget that enhances website user experiences through intelligent, real-time responses while prioritizing security and privacy. Our widget is trained on publicly available data, hosted in Amazon Web Services (AWS), a SOC 2-compliant cloud platform, and collects only the user's email address to prevent spamming robots. No other personal data, such as names, is collected. This white paper outlines our security and data protection practices, demonstrating how we safeguard the limited data we handle, ensure availability, maintain compliance with regulations like GDPR and CCPA, and build trust without requiring extensive certifications like SOC 2.

Scope

This white paper applies to the AI widget, which:

- Uses only public, non-sensitive data for training.
- Operates within AWS infrastructure.
- Collects only the user's email address to prevent spam and robot abuse, with no other personal data required.
- Is embedded on third-party websites to provide general knowledge responses.
- Extends to related services, including Android and iOS mobile apps, employees, contractors, and third-party providers handling data.

Security and Data Protection Principles

1. Data Handling and Training

- **Public Data Only**: The AI model is trained exclusively on publicly available datasets, such as open-domain web content, official company websites, and publicly licensed corpora. No personally identifiable information (PII), confidential data, or regulated data is used in training.
- **Data Integrity**: Training data is sourced from reputable, verified repositories (e.g., official exchange APIs, public records) and undergoes automated validation to ensure accuracy and non-sensitivity.
- **AI Knowledge Base**: The widget retrieves data from a vector-based knowledge store containing structured public company information, sourced exclusively from official public records. No private or user-submitted data is stored in the knowledge base.

- **Minimal User Data Collection**: The widget collects only the user's email address to verify legitimate usage and prevent spamming robots. No names or other personal data are collected or stored, minimizing privacy risks.

2. Email Data Protection

- **Purpose-Limited Collection**: Email addresses are collected solely to authenticate users and prevent automated spam or abuse via one-time password (OTP) sign-in. They are not used for marketing, profiling, or other purposes.
- **Secure Storage**: Email addresses are stored in an encrypted AWS Relational Database Service (RDS) instance (PostgreSQL/MySQL) using AES-256 encryption, accessible only to authorized processes.
- **Data Minimization**: Email addresses are retained only for session-based validation (typically 30 days) and are automatically deleted unless required for ongoing spam prevention, in accordance with AWS data lifecycle policies.
- **Anonymization**: Email addresses are hashed using SHA-256 for processing, reducing the risk of exposure in case of a breach.
- **Authentication**: The widget uses OTP sign-in, eliminating password storage. Authentication tokens are stored in memory (not persisted), and refresh tokens are stored in HttpOnly Secure Cookies for web users, protected from JavaScript access.

3. Infrastructure Security

- **AWS Hosting**: All AI model data, widget operations, and email storage are hosted in AWS, compliant with SOC 2, ISO 27001, and FedRAMP. AWS's Shared Responsibility Model ensures robust physical and infrastructure security.
- **Encryption**: Data at rest (including email addresses) is encrypted using AWS Key Management Service (KMS) with AES-256. Data in transit is secured via HTTPS/TLS 1.3.
- **Access Controls**: AWS Identity and Access Management (IAM) enforces least-privilege access. Only authorized personnel with multi-factor authentication (MFA) can access the AWS environment, with access logged and audited via AWS CloudTrail.
- **Network Security**: The widget operates within a Virtual Private Cloud (VPC) with strict firewall rules. AWS Web Application Firewall (WAF) and Cloudflare Security (Attack & Defense options) protect against SQL injection, cross-site scripting (XSS), DDoS attacks, and bot traffic.
- **Monitoring**: AWS CloudWatch, GuardDuty, and Security Hub monitor security events, providing real-time alerts for suspicious activity, including brute force attempts and unusual API activity.

4. Widget Runtime Security

- **Input Validation**: User inputs, including email addresses, are sanitized to prevent prompt injection, SQL injection, or malicious payloads.
- **Static Responses**: The widget generates responses based solely on its pre-trained model, with email collection limited to spam prevention, minimizing runtime data risks.

- **Content Filtering**: Responses are filtered to exclude inappropriate or harmful content, ensuring a safe user experience.
- **User Interaction Logging**: User questions and AI-generated responses are logged for quality review and to generate reports for companies whose data is referenced, stored securely in AWS with access restricted via IAM roles.

5. Availability and Reliability

- **High Availability**: The widget is deployed across multiple AWS Availability Zones, ensuring uptime and resilience against regional outages.
- **Monitoring**: AWS CloudWatch monitors widget performance, email validation, and logs issues in real time, with automated alerts for rapid response.
- **Disaster Recovery**: Encrypted backups of model data and email storage are maintained in AWS S3, with a recovery time objective (RTO) of under 4 hours.

6. Secure Development Practices

- **Code Security**: Widget code, including email validation and OTP logic, undergoes static analysis, peer review, and mandatory code approval before deployment.
- **Dependency Management**: Third-party libraries are updated regularly and scanned for vulnerabilities using AWS Inspector.
- **CI/CD Pipeline**: A secure continuous integration and deployment pipeline, hosted in AWS CodePipeline, ensures updates are tested and deployed safely, with mandatory 2FA for all accounts.

7. Cookie and Tracking

- **Minimal Tracking**: The widget uses no cookies for session storage; user sessions are stored in memory. PostHog analytics uses cookies to track user interactions for performance improvements, with user opt-out available via the cookie settings panel.
- **Security Tracking**: Cloudflare's Attack & Defense options use tracking techniques (e.g., bot detection, behavioral analysis) to prevent malicious activities, essential for platform protection and non-disableable.
- **User Control**: Users can opt out of non-essential PostHog tracking cookies via the widget's cookie settings page or browser settings.

8. Employee and Contractor Access

- **Employee Restrictions**: Only authorized personnel can access user data, with role-based access control (RBAC) enforced. Employees complete bi-annual security training and sign non-disclosure agreements (NDAs).
- **Contractor and Third-Party Providers**: Contractors and third-party providers sign NDAs and comply with our Data Protection Agreement (DPA). Access is restricted via strict IAM permissions and monitored via AWS CloudTrail.

Compliance and Assurance

Our setup aligns with industry best practices and data protection regulations (e.g., GDPR, CCPA, Colombian GDPR equivalent) without requiring formal certifications like SOC 2, given the minimal data collection and robust AWS infrastructure:

- **AWS Compliance**: We leverage AWS's SOC 2, ISO 27001, and FedRAMP certifications for infrastructure security.
- **Privacy Compliance**: Email collection adheres to data minimization and consent requirements. Personal data is retained for 10 years after the last interaction or as required by law, with secure deletion via AWS lifecycle policies.
- **Transparency**: Clients can request documentation of security controls, AWS compliance reports, and data handling policies under an NDA.
- **Audits**: Internal security audits occur quarterly, with annual third-party penetration testing to validate controls, including email data protection and AI model accuracy.

User Rights and Request Handling

- **User Rights**: Users have the right to:
    - **Access**: Request a copy of their stored email address.
    - **Correction**: Update their email address.
    - **Deletion**: Request deletion of their email address (except where legally required).
    - **Objection**: Withdraw consent for non-essential processing (e.g., PostHog analytics).
- **Request Process**: Users can submit requests via admin@reapapp.io. Queries are processed within 10 business days, and deletion requests within 15 business days.
- **Consent Mechanism**: Users provide explicit consent for email collection via the widget interface, with options to review, modify, or delete data via the settings page.

Incident and Breach Response

- **Monitoring and Detection**: AWS CloudTrail, Security Hub, CloudWatch, and GuardDuty, combined with Cloudflare's Attack & Defense options, continuously monitor for unauthorized access, DDoS attacks, and anomalies. Automated logging captures security events across all environments.
- **Response Plan**:
    1. **Containment**: Isolate affected systems, revoke compromised credentials, and block suspicious access via AWS IAM and Cloudflare.
    2. **Assessment**: Conduct forensic analysis using AWS and PostHog logs to determine scope and entry points.
    3. **Notification**: Inform affected users within 72 hours, as legally required, and coordinate with internal stakeholders.
    4. **Remediation and Recovery**: Patch vulnerabilities, rotate credentials, enhance monitoring, and conduct a security audit to validate fixes.

- **MFA Enforcement**: All access attempts require MFA/2FA, ensuring strong identity verification.

AI and Knowledge-Based Processing

- **Transparency**: The widget uses AI to structure, summarize, and categorize public content (e.g., newsfeed items) sourced from official exchange APIs and company websites. Users' questions are answered using a vector-based knowledge store of public data, with no independent decision-making.
- **Data Handling**: No private or user-submitted data is stored in the AI knowledge base. Questions and responses are logged for quality review and reporting, stored securely in AWS.
- **Audits**: AI models are regularly audited to ensure factual accuracy, prevent bias, and align with official sources, with manual reviews of logged interactions.

Risk Management

- **Low-Risk Profile**: Collecting only email addresses for spam prevention, using public training data, and hosting in AWS significantly reduces privacy and confidentiality risks.
- **AI-Specific Risks**: Model biases and unintended outputs are mitigated through training data validation, response filtering, and regular audits. Email validation is isolated from AI response generation.
- **Third-Party Risk**: Exclusive use of AWS minimizes third-party dependencies, reducing supply chain vulnerabilities.
- **Spam Prevention Risks**: Rate limiting, bot detection, and OTP authentication prevent abuse while protecting legitimate users' email data.

Client Benefits

- **Trust and Privacy**: Minimal email collection, secure AWS hosting, and transparent practices ensure user trust with low privacy risks.
- **Cost Efficiency**: Clients benefit from a secure, reliable widget without the overhead of vendor compliance costs, ideal for small to medium-sized websites.
- **Scalability**: AWS's global infrastructure ensures seamless scaling to meet demand.
- **Ease of Integration**: The widget's lightweight data footprint and robust security simplify embedding without complex vendor assessments.

Conclusion

REAP's AI widget delivers value with minimal security and privacy risks. By using public data for training, collecting only email addresses for spam prevention via OTP authentication, and hosting in AWS, we provide a secure, reliable, and compliant solution. Our practices align with GDPR, CCPA, and industry standards, leveraging AWS's compliance and our rigorous controls

to ensure trust without formal certifications like SOC 2. We offer transparent documentation and ongoing support for clients requiring additional assurance.

Contact

For questions, data-related requests, or detailed security documentation, contact:
📧 **Email**: admin@reapapp.io

Policy Review and Updates

This white paper is reviewed annually to ensure compliance with evolving regulations. Significant changes will be communicated via email and updated on our website.

REAP - Protecting Your Data, Always.