GenerativeAgent

ASAPP

Security & Safety FAQ

At ASAPP, we prioritize safety, security, and privacy with our GenerativeAgent product. This FAQ covers four essential sections:

01

GenerativeAgent Product Security

Customer data remains within its original geographic region adhering to data sovereignty, protected by robust encryption and stringent access controls. Real-time oversight is maintained through Amazon CloudWatch and AWS CloudTrail.

02

GenerativeAgent Product Safety

ASAPP's orchestration enhances safety and implements guardrails to ensure that outputs generated by GenerativeAI are accurate, secure, and traceable back to their source of truth. Automated quality checks and human oversight enhance reliability.

03

Data Handling

Customer data is handled by comprehensive data controls and sensitive data is appropriately classified and managed with robust data redaction techniques. These measures ensure that customer data is managed with the highest standards of security and privacy.

04

Foundational Security Controls

We utilize encryption, role-based access control, multi-factor authentication, and continuous monitoring with advanced intrusion detection systems and Security Information and Event Management (SIEM). To promptly address security gaps, we also conduct regular vulnerability management and third-party penetration tests by reputable organizations.

GenerativeAgent Product Security

01. What platform do you use and what are its key characteristics?

The ASAPP AI platform and LLMs are self-hosted on AWS. We ensure that no input or output text from inference requests is used to train the LLMs for other customers. AI model deployments occur within an AWS account and are securely accessed by GenerativeAgent either through an Internet Gateway or directly via a Private Network VPC link. Additionally, foundational models are either hosted locally on AWS, or accessed through the cloud environments of partner LLM providers where we guarantee a zero data retention policy.

02. How does the product handle customer data?

For certain use cases, the foundational models are hosted on a partner LLM provider's cloud, but we guarantee zero data retention outside of our AWS environment. Customer data, including prompts, information used for prompt refinement, AI model responses, and customized models, is stored within the specified AWS region where the API request is initiated by the application. The LLMs operate either locally in a stateless manner within the ASAPP AI Platform, hosted on AWS, or on the partner LLM provider's cloud, with strict data protection measures in place. Consequently, in production, customer data is never retained outside the AWS environment.

03. How is customer data handled in terms of AI model service enhancement?

We never use data from one customer to train models for other customers. However, we do have the capability to improve customer-specific models using their own data, allowing these models to improve themselves based on the customer's unique feedback and data.

04. What type of encryption do you use for data protection?

We protect customer data by encrypting it during transit using at least TLS 1.2. For encryption at rest, to secure our models and customer data hosted on ASAPP's AWS account, we use AES 256 with AWS Key Management Service (KMS) keys.

05. How do you ensure the security of the AI models you use?

The AI models' security is maintained through a combination of encryption, private connectivity options such as AWS PrivateLink, and strong identity and access management protocols.

06. How do you monitor and audit AI model usage?

We use Amazon CloudWatch and AWS CloudTrail for comprehensive monitoring and logging. CloudWatch helps in tracking usage metrics and creating custom dashboards, while CloudTrail logs API activity, aiding in troubleshooting and meeting audit and compliance requirements. Additionally, we provide audit logs for any changes made by a customer's users to the configuration of their implementation of GenerativeAgent, allowing customers to track all modifications and ensure transparency.

Moreover, our model management and deployment processes are designed to be robust and reliable. All model deployments are versioned, auditable, and support rapid rollbacks if needed. This ensures that we maintain the highest standards of accountability and can quickly address any issues that arise.

07. What measures are implemented to prevent unauthorized access to customer data?

Multiple layers of security controls are in place, including encryption for data both in transit and at rest (for when data needs to persist), identity and access management via IAM, multi-factor authentication (MFA), AWS native Security tools such AWS GuardDuty, Cloud Security Posture Management with CIS benchmarking, Attack Surface Management, and detailed API activity logging through AWS CloudTrail, all designed to safeguard against unauthorized access and data breaches.

08. Which LLM providers do you use?

ASAPP's GenerativeAgent leverages a combination of self-hosted, 3rd party proprietary, and open-source language models to ensure high quality and compliance. This allows the system to dynamically incorporate and use the best available models while maintaining robust security and operational standards.

09. How do you ensure secure API connections and user authentication?

ASAPP employs an API proxy layer that handles authentication. Authentication is performed at the user level, meaning the authentication token used by GenerativeAgent only permits the retrieval of data for the authenticated user. This design ensures that the system cannot make API calls for other users or systems, thereby enhancing security and data privacy.

10. What safety measures does ASAPP's GenerativeAgent product have in place to prevent prompt injection attacks?

ASAPP's GenerativeAgent product employs multiple layers of defense to prevent prompt injection attacks. Firstly, the system includes input safety mechanisms to detect and block attempts by users to exploit the system, thus preventing most forms of abuse such as prompt leaks, tampering code injection, and "jailbreak" scenarios.

Furthermore, all user-generated traffic is routed through an advanced Web Application Firewall (WAF) that detects and prevents injection attacks, including SQL injection. Collectively, these measures ensure robust protection against prompt injection attacks, maintaining the security and integrity of the system.

GenerativeAgent Product Safety

11. How does GenerativeAgent ensure the safety and integrity of inputs and outputs?

GenerativeAgent employs Input/Output Safety evaluators to ensure the safety and integrity of data processed by the AI systems. These evaluators are responsible for filtering malicious inputs and ensuring that the AI's outputs are safe. Specifically, the evaluators prevent sensitive data leakage and profanity by analyzing inputs and outputs in real-time.

12. What measures does GenerativeAgent have to address and mitigate AI hallucinations?

GenerativeAgent addresses and significantly reduces the risk of AI hallucinations through several measures, including the use of guardrails and safety agents that analyze the traffic and outputs to identify and prevent the generation of inaccurate or misleading information.

Additionally, both automated quality assurance testing and manual red teaming focus on identifying hallucinations. A human-in-the-loop escalation process enables agents to quickly intervene and rectify any AI-generated inaccuracies. We also ensure that responses are firmly grounded in customer data, including content, policies, APIs, and other pertinent information.

13. Does GenerativeAgent have controls in place to ensure the AI's actions remain within defined parameters?

GenerativeAgent includes two mechanisms to ensure the conversation and actions remain within defined parameters: (1) ASAPP-managed scope safety layer ensures that GenerativeAgent responds appropriately to customer input that is out-of-scope of your business, constraining the conversation itself; (2) Customer administrators define which actions GenerativeAgent can take in backend systems, such as looking up information or taking actions on behalf of customers, and GenerativeAgent is restricted to said actions for any given use case.

14. Does GenerativeAgent use mechanisms to verify critical actions before they are executed?

GenerativeAgent utilizes a Confirmation Workflow mechanism that requires explicit user confirmation prior to executing any actions that alter customer data. This protocol ensures significant operations are verified, enhancing the reliability of the AI's actions. Additionally, a human-in-the-loop framework is employed for scenarios requiring oversight to guide, correct, and validate system outputs, improving accuracy, reliability, and safety.
This approach is especially critical when the system encounters highly nuanced or sensitive decisions that automated processes alone may not adequately manage.

Data Handling

15. In which AWS regions is your solution hosted, and where is customer data physically stored?

The ASAPP solution is hosted in AWS US regions. All customer data is stored within an AWS US region, and our primary data centers and backup locations are based in the United States East zones.

16. Can customer data be accessed by authorized individuals from outside the US?

Customer data may be accessed outside the US based on the principle of Least Privilege. However, ASAPP employs a comprehensive set of data protection measures to safeguard customer data. These measures include, but are not limited to, secure data categorization, encryption both in transit and at rest, access controls, monitoring, and adherence to regulatory requirements.

17. What steps are taken to securely manage internal employee user access?

We configure user accounts using AWS Identity and Access Management (IAM), assigning permissions based on specific roles to minimize access. Enhanced security is also achieved through multi-factor authentication (MFA) and using SSL/TLS for secure communication with AWS services. Furthermore, we have the option to enforce Virtual Private Network (VPN) for user access.

18. Can I choose to have my data locked down only to US employees?

Yes, having customer data locked down only to US resources is feasible but subject to additional fees.

19. Aside from AWS, who are the sub processors that ASAPP needs to access my data?

ASAPP uses sub-processors to leverage specialized services and infrastructure that enhance the performance and security of its platform - See the sub processors list here.

Customer data that goes to ASAPP's sub processors is processed in the United States.

---

20. Do you redact customer PII data prior to storing it? If you don't, why not? And can you add custom redaction rules?

ASAPP offers a sophisticated redaction engine capable of redacting any sensitive information, including PII and PCI, in real-time prior to any persistence. Customers are in full control of redaction rules, with configuration for different entities. This mechanism is part of ASAPP's comprehensive security measures to protect data privacy and comply with industry standards.

ASAPP also offers the flexibility to implement custom redaction rules specified by the customer. This redaction process can be tailored to capture and redact specific data elements such as names and addresses, ensuring that all sensitive information is properly managed and handled in compliance with our customers' requirements.

---

21. Will you use customer data to train your models for other clients?

No. Any data that ASAPP collects may only be used to train models that support your engagement. This data is not shared with anyone outside of ASAPP and its approved subprocessors, and it is used solely to enhance specifically your usage of our services.

Foundational Security Controls

22. How does ASAPP ensure data encryption?

ASAPP implements industry standard data encryption both at rest and in transit over public networks to safeguard data integrity and security. ASAPP ensures data encryption through multiple robust and industry-standard mechanisms. Data in transit is encrypted using TLS 1.2 or higher, providing secure communication over public networks. For data at rest, ASAPP relies on Amazon's Key Management Service (KMS) using AES 256 encryption.

23. How does ASAPP restrict data access?

ASAPP restricts data access through a combination of policies, processes, and technological controls. Central to this approach is implementing the principle of least privilege and a need-to-know basis for granting access. Identity and access management is centrally managed via an Identity Provider (IDP), where access requests undergo a formal submission, review, and approval process ensuring that permissions align with job responsibilities.

Enhanced security is achieved through Multi-Factor Authentication (MFA) mechanisms, requiring users to authenticate using a combination of credentials and secondary factors such as software or hardware tokens, with mandatory registration of devices within ASAPP's central asset management system. Role-Based Access Control (RBAC) is employed to assign specific permissions based on defined roles and responsibilities, thus ensuring users have the minimum level of access necessary.

24. What measures does ASAPP take for intrusion detection and prevention?

ASAPP has an intrusion detection system along with Denial of Service (DoS) and Web Application Firewall (WAF) protections. We also use next-generation antivirus and anomaly detection tools.

ASAPP employs a multi-faceted approach for intrusion detection and prevention to ensure the security of its environment. The company uses a combination of host and network intrusion detection systems (IDS) across its environment. Intrusion detection tools are deployed with comprehensive coverage, alerting the Security Engineering team to any detected threats.

In addition to these IDS measures, ASAPP employs endpoint detection and response (EDR) solutions on all end-user devices. These EDR tools monitor both traditional signature-based antivirus protections as well as behavioral, machine learning-based detection methodologies to prevent malware and other suspicious activities.

ASAPP's Security Information and Event Management (SIEM) system further enhances monitoring by collecting and analyzing logs from various sources. This system allows for automated alerting and real-time analysis of security events, ensuring timely detection and response to any anomalies.

25. How does ASAPP handle vulnerability management?

ASAPP handles vulnerability management through a robust program enforced by its Security Engineering team, guided by a comprehensive Vulnerability Management Policy. This policy establishes criteria for categorizing vulnerabilities as critical, high, medium, or low, each with specific resolution timelines. The vulnerability management includes both internal and external assessments: internal scans encompass the whole internal and external environment, while annual penetration testing is conducted by reputable third-party entities.

Continuous monitoring of new vulnerabilities is a key aspect, facilitated by a dedicated security engineering team leveraging various scanning tools. The program is reviewed and updated annually to align with industry standards. Detailed records of identified and resolved vulnerabilities are maintained to enhance future strategies and incident responses. This process ensures that vulnerabilities are tracked from discovery to remediation, with categorized tickets raised for their resolution, followed by re-scanning to confirm effective resolution.

26. What are ASAPP's protocols for incident detection and response?

ASAPP's protocols for incident detection and response are governed by its Incident Management Response Policy. The policy outlines a comprehensive and methodical approach to handling security incidents, which includes 24/7 monitoring. ASAPP employs a Security Information and Event Management (SIEM) system for continuous analysis, correlation, and triage of security events. The policy specifies the procedures for incident detection, categorization, and resolution phases, including registration of the incident for tracking purposes.

27. How does ASAPP ensure business continuity?

ASAPP conducts business continuity through a robust Business Continuity and Disaster Recovery (BCDR) policy. This policy includes defined recovery objectives and redundancy strategies to maintain high availability and rapid recovery of services. ASAPP leverages cloud service providers like Amazon Web Services (AWS) for its infrastructure, ensuring deployment across multiple Availability Zones to handle failovers and disruptions effectively. Additionally, the company conducts annual tests of its BCP-DR program to validate and update plans based on test results.

28. Does ASAPP conduct independent security assessments?

Yes, ASAPP conducts independent security assessments through periodical third-party penetration tests. These tests include external and internal network penetration testing, as well as application security tests. Identified vulnerabilities are categorized by severity, reported to stakeholders, and prioritized for remediation.

29. Is ASAPP SOC 2 Type II and PCI DSS compliant?

Yes. ASAPP undergoes periodic external audits for certifications such as SOC 2 Type II and PCI DSS, ensuring compliance with industry standards and best practices.

30. How does ASAPP prioritize data privacy?

ASAPP prioritizes privacy incorporating “privacy by design” as a main pillar for its Privacy and Security posture. This approach is embedded across company operations; this is conducted through a proactive approach, including the redaction of sensitive data and conducting thorough privacy impact assessments.

31. What data types does ASAPP redact?

ASAPP performs automated redaction of Personally Identifiable Information (PII), Payment Card Information (PCI), and various other types of customer data. Customers have control over the redaction process, allowing for customization to meet specific requirements for masking and obfuscation. Redaction occurs before the data is stored at rest.

32. How does ASAPP ensure secure software development?

ASAPP ensures secure software development through a comprehensive Secure Software Development Lifecycle (SDLC) framework, integrated within its continuous integration/continuous deployment (CICD) process. This framework includes repository vulnerability scanning, mandatory secure code training based on OWASP Top 10 vulnerabilities, security design reviews, threat modeling, and both automated and manual code vulnerability assessments.

Scans are conducted automatically and periodically reviewed. There’s a structured peer review process ensuring code cannot be merged without approval. High-risk applications or those dealing with sensitive data undergo formal threat modeling. Moreover, an annual application penetration test is performed to identify, grade, and remediate vulnerabilities based on defined severity levels, ensuring robust and secure software development practices.

33. How does ASAPP ensure compliance with data protection regulation?

ASAPP ensures compliance with data protection regulations, through a comprehensive privacy program overseen by our Chief Legal Officer who also serves as our Data Protection Officer (DPO). The DPO, in conjunction with a dedicated Security and Privacy Assurance team, manages ASAPP's privacy posture, defines data protection efforts, and ensures compliance with both internal and external requirements. This includes conducting privacy impact assessments, facilitating data subject access requests, and aligning data handling practices with regulatory obligations, ensuring robust data protection and privacy standards. Additionally, our infrastructure is GDPR-ready, ensuring compliance with stringent data protection regulations.

34. How does ASAPP handle third-party vendor compliance?

ASAPP handles third-party vendor compliance through a comprehensive Third-Party Risk Management Program. This program involves performing detailed security and privacy due diligence on vendors prior to engagement and ongoing monitoring of vendor performance against established security and contractual requirements. Our assessments include considerations of usage of LLMs by third-parties and respective risks of exposing any of ASAPP data (for generic vendors) or our customer data (for subprocessors) as part of our overall due diligence. This involves a coordinated effort between ITS, Security & Privacy Assurance, Finance, and Legal teams.

Security risk assessments are conducted as part of the vendor evaluation process. These assessments review key artifacts such as due diligence questionnaires covering data privacy, certifications, identity and access management, encryption standards, and more. Third-party compliance reports and vendor policies are also scrutinized.

35. Does ASAPP provide employee training on data security?

Yes, ASAPP provides regular training to employees on data handling, privacy regulations, and security awareness programs.

36. How does ASAPP handle data retention and deletion?

ASAPP ensures secure data storage in accordance with customer's regulatory and business requirements and offers flexible retention policies. Data is securely and irreversibly deleted when no longer needed.

37. What data obfuscation techniques does ASAPP use?

ASAPP uses a proprietary redaction service that combines techniques such as P-Filtering, Named Entity Recognition (NER), and regex-based redaction to achieve high levels of data obfuscation, including obfuscation of different categories of sensitive data (i.e. PII, NPI, PCI, etc.). Customizations to different data elements are available to ensure adequate obfuscation based on customer's needs.

38. Is ASAPP compliant with AI risk management frameworks?

ASAPP is adopting the NIST AI Risk Management Framework (AI-RMF) to align with best practices and ensure the responsible and transparent use of AI.

ASAPP

ASAPP's Security and Privacy Assurance team has the mission of enabling customer trust. Please feel free to reach us in case you require our support:

security@asapp.com | privacy@asapp.com

About ASAPP

ASAPP is an artificial intelligence cloud provider committed to solving how enterprises and their customers engage. Inspired by large, complex, and data-rich problems, ASAPP creates state-of-the-art AI technology that covers all facets of the contact center. Leading businesses rely on ASAPP's AI Cloud applications and services to multiply agent productivity, operationalize real-time intelligence, and delight every customer.

To learn more about ASAPP innovations, visit **www.asapp.com**

Learn more