ThreatCanvas
Security Whitepaper

ThreatCanvas by SecureFlag is an advanced, Threat Modeling automation solution designed to significantly enhance the security posture of organizations.

This whitepaper provides an in-depth overview of the stringent security measures and protocols integral to ThreatCanvas, ensuring the highest standards of data confidentiality, integrity, and availability for customers.

Introduction to ThreatCanvas

ThreatCanvas is an automated, AI-powered platform developed by SecureFlag for identifying, assessing, and mitigating potential security threats in digital infrastructures. ThreatCanvas offers comprehensive and customized threat modeling to meet the unique security requirements of each organization.

Data Security and Privacy

Encryption at Rest
All customer data stored within the ThreatCanvas system is encrypted at rest using industry-standard encryption protocols, safeguarding data from unauthorized access.

Encryption in Transit
Data transmitted to and from ThreatCanvas is secured using robust encryption protocols, such as TLS 1.2 and 1.3, ensuring secure and private data transfer.

Customer Data Segregation
ThreatCanvas employs a logical data separation architecture to ensure that data is segregated between different customers, guaranteeing independent storage and management of each customer's data.

SecureFlag implements rigorous access protocols to ensure that customer data, specifically Threat Model data, is accessible on a need to know basis. These protocols are designed to enforce the principle of least privilege, ensuring that staff access is limited to what is strictly necessary for their job functions.

Data Storage
ThreatCanvas does not retain user prompts or attachments; data is processed temporarily in memory and discarded immediately afterward. Users have the option to save their Threat Model diagrams to their library on SecureFlag. Customers have the ability to disable this saving feature if preferred.

All data is securely stored within the European Union, ensuring compliance with the EU's stringent data protection laws and regulations for enhanced security and privacy.

Data Accessibility and Visibility Settings

- Private Visibility: By default, threat model data created by a user is set to *private*, meaning it is only accessible to the user who created it.
- Team Visibility: Users have the option to set their threat model data to *team* visibility, allowing access to users within the same team on SecureFlag.
- Organization Visibility: For broader access, users can set their threat model data to *organization* visibility, making it accessible to all users within the same organization on SecureFlag.

ThreatCanvas supports designated collaborators who are granted read and write access to shared models.
Organization Administrators have read and write access to all Threat Models created by users within the organization, while Team Managers have similar access to Threat Models created by users in their respective teams.

Data Usage and Accessibility

No Customer Data Used for Model Training
The AI model powering ThreatCanvas is provided by Anthropic and hosted through Amazon Web Services Bedrock. The model is trained using diverse data sources, excluding customer data.
Content processed through Amazon Bedrock is not used to improve the base models and is not shared with any model providers.

Robust Model Design
The AI model used by ThreatCanvas is provided by Anthropic and it is designed with security as a paramount concern. It incorporates features to prevent biases, ensure fairness, and avoid any form of manipulation.

Continuous Monitoring and Updating
The AI model used by ThreatCanvas is continuously monitored for anomalies and updated to adapt to new threats and vulnerabilities, ensuring it remains effective against evolving security challenges.

Data Integrity and Model Validation
Regular validation checks are conducted to ensure the integrity of the model's outputs. This includes rigorous testing against known threat scenarios to verify accuracy and reliability.

AI Ethics and Governance
Anthropic and Amazon Web Services adhere to strict ethical guidelines in the development and deployment of its AI model, ensuring responsible and transparent AI practices.

Compliance, Standards, and Information Security Management

ISO 27001 Certification
SecureFlag's commitment to information security is demonstrated through its ISO 27001 certification.

Information Security Management System (ISMS)
SecureFlag's robust ISMS underscores its systematic approach to managing and protecting information security risks.

Amazon Web Services (AWS) Bedrock
Amazon Bedrock is in scope for common compliance standards such as Fedramp Moderate, Service and Organization Control (SOC), International Organization for Standardization (ISO), Health Insurance Portability and Accountability Act (HIPAA) eligibility, and the General Data Protection Regulation (GDPR). Amazon Bedrock is included in the scope of the SOC 1, 2, 3 reports; ISO Compliance for the ISO 9001, ISO 27001, ISO 27017, ISO 27018, ISO 27701, ISO 22301, and ISO 20000 standards. Amazon Bedrock is CSA Security Trust Assurance and Risk (STAR) Level 2 certified.

Risk Management and Mitigation
ThreatCanvas proactively manages risks, continuously updating its security measures to address emerging threats and adapting swiftly to new vulnerabilities through its AI capabilities.

Regular Security Assessments by Third Parties
To further bolster its security posture, SecureFlag performs regular security assessments delivered by independent third parties. These assessments are critical in identifying and addressing potential vulnerabilities, ensuring that ThreatCanvas remains resilient against evolving security threats. This external validation adds an additional layer of assurance for customers, reinforcing the robustness of SecureFlag's security measures.

Incident Response and Recovery
SecureFlag has a comprehensive incident response plan for immediate containment, investigation, and recovery, ensuring minimal impact and rapid restoration of normal operations in the event of a security incident.