Artificial Intelligence and Machine Learning Data Sheet

Skyhigh Security is a pioneering leader in cybersecurity powered by artificial intelligence (AI) and machine learning (ML), and is committed to deploying AI/ML in a safe, transparent and responsible way - consistent with our customers' high expectations and in compliance with all regulatory requirements, including the European Union's Artificial Intelligence Act (EU AI Act) and other evolving laws and regulations worldwide.

The purpose of this document is to provide our customers with information on how Skyhigh Security uses AI/ML in our products and services, and to help our customers understand and assess the impact of Skyhigh's data processing on their overall AI/ML compliance posture.

For more information about Skyhigh AI vision and guiding principles, see the AI section and FAQs on our Skyhigh Trust Center.

Overview of AI/ML in Skyhigh Solutions

Skyhigh Security uses Artificial Intelligence (AI) and Machine Learning (ML) in its Security Service Edge (SSE) platform to enhance overall security operations, improve detection accuracy, and streamline administrative workflows across various areas. Our solutions are focused exclusively on Secure Access, Threat and Data Protection, ensuring they operate without human and social biases.

These capabilities enable us to:

- **Increase Efficiency.** AI streamlines Security Operations Center (SOC) workflows by automating alert analysis, prioritizing incidents, and reducing false positives, resulting in faster incident response and reduced analyst fatigue.
- **Reduce Complexity**. AI minimizes human errors, creates patterns for identifying sensitive data, improves classification accuracy, develops context-aware expressions, and instantly explains new and changing sensitive data types.
- **Lower Risk**. Elevates protection against insider threats by detecting and responding to abnormal user behavior, proactively identifying risks before they escalate. With intelligent anomaly detection it enhances threat visibility and reduces false positives, enabling faster and more accurate responses.

- **Automation:** Automates threat intelligence by synthesizing signals from sensors across the globe. This allows the system to rapidly update its threat knowledge, ensuring that new threats are dynamically identified and blocked instantly.

The following key areas demonstrate how Skyhigh Security has utilized AI and ML to develop impactful solutions:

Automate Content Classification with ML-driven Auto Classifiers

Skyhigh uses ML models to process all DLP-scanned content and automatically determine its classification.

Here's a breakdown of how this process works:

- **Classification Process**. The ML models analyze both textual and image-based content, determining its classification based on predefined categories. Different models are specifically designed for various types of content, utilizing a combination of statistical methods and neural network techniques.
- **Data Privacy**. User data is not used in the training of these models, nor retained for that purpose. The models are trained independently using a separate Skyhigh corpus, ensuring data privacy and compliance.
- **Operational Boundaries**. These models operate exclusively within the confined boundary of the Skyhigh solution, meaning that no content is transmitted outside of these boundaries.

This capability is integrated with DLP, allowing you to automatically detect and classify various types of sensitive files based on Skyhigh pre-trained AI and ML models. For details, see ML Auto Classifiers.

For common queries on ML Auto Classifiers, see FAQs on ML Auto Classifiers.

Simplify DLP Classification with AI Regex Builder/Generator

Skyhigh leverages an AI model to create complex regular expressions for DLP classification, making it easier for administrators who lack expertise in regex syntax.

Here's a breakdown of how this process works:

- **Classification Process**. Use the DLP Classification Regex Building Assistant or AI RegEx Generator that seamlessly constructs and comprehends complex Google RE2-compliant regular expressions through a conversation-based interface.
- **Data Privacy**. All your queries remain confidential and are not used for training. However, responses may rely on standard external large language models that are independently trained.

- **Privacy Notice for Data Input**. Refrain from entering confidential, personal, or sensitive data into the query field. Information entered is transmitted to an external AI service to generate the required answers. The AI-generated answers are only to aid the expression building.

This capability is integrated with DLP, allowing you to simplify the task of building complex expressions and minimize the chances of inaccuracies. For details, see AI RegEx Generator for Custom Advanced Patterns.

Minimize Potential False Positive Incidents with ML

Skyhigh uses ML models to identify and predict incidents that are most likely to be false positives and ensures data integrity and privacy in user incident prediction models.

Here's a breakdown of how this process works:

- **Data Training and Isolation**. The ML models are trained based on users' incident metadata. User data isolation and security are maintained through strict tenant separation. Each user's data remains completely segregated, ensuring that analytics and predictions for one tenant are exclusively based on their data, without any cross-tenant data sharing or influence on outcomes.
- **Operational Boundaries**. These models operate exclusively within the confined boundary of the Skyhigh solution, meaning that no content is transmitted outside of these boundaries.

This capability is integrated with DLP, allowing you to automatically detect and classify sanctioned DLP incidents that may be false positives. It also provides actionable insights, including the statistics for ML-driven Potential False Positives and a comprehensive list of potential false positive incidents. For details, see About ML-Driven Potential False Positives.

Enhanced Threat Detection and Remediation with UEBA

The Skyhigh CASB Threat Protection offers real-time detection and remediation of threats and anomalies, ensuring enhanced security, compliance, and data governance across sanctioned cloud services. As AI becomes integral to applications, sensitive data faces growing vulnerabilities, demanding robust security at every level. To address these challenges, Skyhigh Security delivers purpose-built solutions that empower organizations to maintain control, prevent data leaks, and safeguard users in the evolving AI era.

Here's a breakdown of how this process works:

- **Data Training and Isolation**. Skyhigh models train on tenant-specific data to derive threat vectors and pinpoint anomalies. User data isolation and security are maintained through strict tenant separation. Each user's data remains completely segregated,

ensuring that analytics and predictions for one tenant are exclusively based on their data, without any cross-tenant data sharing or influence on outcomes.
- **Operational Boundaries.** These models operate exclusively within the confined boundary of the Skyhigh solution, meaning that no content is transmitted outside of these boundaries.

Enhanced Content Classification with GAM

Skyhigh processes all Generalized Additive Models (GAM) scanned content through its ML models to automatically determine the classification of the content.

Here's a breakdown of how this process works:

- **GAM Model Techniques**. Different models are specifically designed for various attack vectors, utilizing a combination of statistical methods and neural network techniques.
- **Data Privacy**. User data is not used in the training of these models, nor retained for that purpose. The models are trained independently using a separate Skyhigh corpus, ensuring data privacy and compliance.
- **Operational Boundaries**. These models operate exclusively within the confined boundary of the Skyhigh solution, meaning that no content is transmitted outside of these boundaries.

Skyhigh Support Portal: Virtual Agent

The Skyhigh Support Portal utilizes a sophisticated Virtual Agent to streamline customer assistance and improve resolution times through advanced AI/ML capabilities. This system is designed with a focus on accuracy, contextual relevance, and strict security standards.

Core AI capabilities and Governance

Core AI Capabilities. The Virtual Agent leverages Federated RAG (FRAG™) to pull localized context from internal knowledge bases, forums, and CRMs, significantly reducing LLM hallucinations. It utilizes LLM Intelligence for natural language understanding and contextual memory, providing case summarization to condense customer interaction histories and autonomous support agents for first-line resolution, ticket deflection, and smart human handoffs.

Governance & Security. Data processing operates within single-tenant virtual private clouds to ensure isolation. Access controls strictly respect user permission hierarchies, and Model Context Protocols (MCPs) are employed to maintain rigorous technical and contextual boundaries.

Regulatory Compliance

Regulatory requirements related to AI/ML continue to evolve worldwide. Skyhigh Security

complies with applicable laws in every jurisdiction where we do business, and is committed to supporting our customers' compliance journeys as well.

In general, Skyhigh Security does not engage in activities that are deemed "high risk" from an AI/ML perspective. For example, Skyhigh Security AI/ML engines are not used in any way that could fundamentally violate human rights or manipulate human behavior, allow social scoring, or make automated decisions that could result in discrimination. Rather, our Skyhigh Security AI/ML is used solely by our customers' security analysts to detect, manage and respond to potential cybersecurity threats, and for no other purposes. Automated decision rules made by our AI/ML solutions are configured by our customer security analysts who understand and fully expect that such decisions are implemented via AI/ML. Further, the results of automated decisions typically result only in the quarantine of suspicious data and/or turning off access or administrative privileges to prevent proliferation of potentially malicious activity – which in no way affect fundamental human rights. In all events, our customers' human security analysts can transparently review and understand such automated decisions, and make corrections where appropriate. (For more information about cybersecurity in the context of high risk systems, see Recital 15 of the EU AI Act.)

However even though the use of AI/ML in Skyhigh Security solutions is itself not high risk, we recognize that many Skyhigh Security customers do directly engage in high risk activities, including the management or support of critical infrastructure, and the processing of highly sensitive data. We also recognize that Skyhigh Security solutions are integral to our customers' overall security posture. As a leader in responsible AI, Skyhigh Security has implemented various controls and processes that align with the requirements for high-risk AI systems. This commitment is designed to support our customers in meeting their security and compliance requirements, as further detailed below.

Risk Management System. Skyhigh Security has established a risk management system that spans the lifecycle of our AI/ML systems, from initial design through deployment and post-market monitoring. This system includes (1) proactive identification of risks to safety and fundamental rights, (2) assessment of the severity and probability of identified risks, and (3) technical and organizational measures to reduce risks to an acceptable level.

Data Governance and Management. Skyhigh Security has appropriate data governance and management practices that cover how we train, validate and test our AI/ML models.

- **Data Quality** We implement comprehensive data quality checks to help ensure the accuracy, completeness, and consistency of datasets.
- **Data Relevance and Representativeness:** We meticulously curate datasets to be relevant to the intended purpose of the AI system and sufficiently representative to mitigate bias.
- **Data Collection Practices:** We collect data lawfully and ethically, adhering to privacy and data protection regulations worldwide. We use

anonymization and pseudonymization techniques where appropriate.

- **Use Limitation:** We do NOT use customer data (e.g., systems information, IP addresses, email addresses, host names, file paths, log information, etc.) to train our AI models. Our AI models are computation models that do NOT self-mutate or train based on customer data. Though we use customer detections and events for inference to provide security-related insights, such data does not update our AI models. All such insights are based on information about attacks – NOT on information about our customers.
- **Data Labeling:** We establish clear protocols for data-labeling to help ensure consistency and accuracy.

Privacy and Data Protection

- **Compliance**: Our AI/ML data handling practices comply with all applicable privacy and data protection requirements worldwide, including the EU General Data Protection Regulation (GDPR) and other EU and Member State legislation. To the extent our products and services process personal data (e.g., usernames and file paths) in the course of detecting potentially malicious events, we adhere to strict privacy and security policies and procedures to protect all such information. We collect and process only the data necessary for the intended purpose of our AI systems, and have implemented robust technical and organizational security measures to protect all data used in our AI/ML processes from unauthorized access and manipulation. For data residency compliance, all AI/ML inference requests originating in the EU are processed using regional cloud services, ensuring the data remains within the European Union.
- Customer Control: Customers using Skyhigh Security solutions maintain significant control over their data usage. This includes, where appropriate, options to customize data flows for specific environments and the ability to choose when to enable or disable AI/ML functionality. For detailed information about customization and opt-in/out choice for AI inference and data usage, see our relevant product documentation, as features vary by product. Or, contact your Skyhigh Security support team directly with any questions.

  For more information on how our various products and services collect, store, use, secure and delete customer data, see our Privacy Data Sheet on the Skyhigh Security Trust Center.

Technical Documentation

We maintain comprehensive technical documentation for all Skyhigh Security AI/ML systems, including:

- A description of the AI system's general characteristics, capabilities, and limitations
- Information pertaining to the data utilized for training, validation, and testing;
- A description of the design specifications, development process, and algorithms employed

- Information regarding the risk management system and its implementation;
- Results of conformity assessments and testing; and
- Instructions for use, including information on human oversight.

This confidential and proprietary documentation is regularly updated, and may be made available to relevant regulatory authorities upon request.

Record-Keeping

Skyhigh Security AI/ML systems are engineered to facilitate our customers' automatic logging of events to help ensure the traceability of operations. Such logs may enable:

- Recording of decisions rendered by the AI system.
- Monitoring of key performance indicators.
- Tracking of human interventions and overrides.

Customers may maintain logs, with Skyhigh Security support as needed, for durations meeting regulatory and audit requirements.

Transparency and Provision of Information to Customers

Skyhigh Security is committed to providing customers with clear and comprehensible information regarding our AI/ML systems. In general, the intended purpose of our AI/ML systems is to help our customers detect, manage and respond to potential cybersecurity threats. As with any cybersecurity solution, there are inherent limitations to the ability of our AI/ML systems to detect and respond to all possible threats, as malicious actors continuously evolve their tactics and attack techniques.

For more information about our AI system characteristics and limitations, computational and hardware resources needed for our AI systems, and mechanisms that allow customers to properly collect, store and interpret logs, see our Skyhigh Security Trust Center.

Human Oversight

It is critical that our customers' human security analysts can understand, interpret, and intervene (when necessary) in the automated decisions and actions of our Skyhigh Security products and services.

As discussed above, automated decision rules made by our AI/ML solutions are configured by our customer administrators and security analysts who understand and expect that such decisions are implemented via AI/ML. In addition, human security analysts (Skyhigh Security or customer) can transparently review, understand, and correct these automated decisions and actions.

For instance, Skyhigh Security solutions enable human oversight through Human-in-the-Loop (HITL) mechanisms. We provide customer security analysts with intuitive control panels,

dashboards and alerts that enable them to review, validate and override AI-generated recommendations or automated actions.

In addition, Skyhigh Security AI/ML solutions use Retrieval Augmented Generation (RAG) for step-by-step processing, citing sources for human security analysts to ensure accuracy, reliability, and to prevent AI "hallucinations."

Customers always have access to standard documentation and training focused on effectively interacting with and overseeing AI/ML features across all Skyhigh Security offerings.

Accuracy, Robustness, and Cybersecurity

Skyhigh Security has implemented appropriate measures to ensure the high performance and security of our AI/ML systems.

- **Accuracy:** We continuously maintain and enhance accuracy in threat detection and response through rigorous testing and validation using diverse datasets.
- **Robustness:** We have designed our AI models to exhibit resilience to errors, inconsistencies, and adversarial attacks, including comprehensive testing for data poisoning, model evasion, and other adversarial techniques.
- **Cybersecurity:** We have implemented robust cybersecurity measures to safeguard our AI systems (as well as related training data, models, and underlying infrastructure) from unauthorized access, manipulation, or attacks.

Importantly, our AI/ML products have customer-visible telemetry that provide details related to the operations of the models themselves. For example, our ML-based anomaly detection provides a variety of metrics around the scoring and confidence, which help users understand how anomalous a given security event is. Similarly, our LLM-based detection products provide a confidence score; and the LLM provides detailed reasoning on exactly why it is sure or unsure of its decision.

Quality Management System

Skyhigh Security integrates the development and deployment of its AI/ML systems within its existing comprehensive Quality Management System (QMS), which includes:

- Organizational structure and defined responsibilities.
- Established processes for design, development, testing, and deployment.
- Rigorous risk management procedures.
- Comprehensive documentation and record-keeping protocols.
- Processes for continuous improvement.

Ongoing Monitoring

To help identify any emerging risks and to ensure continued compliance, Skyhigh Security engages in ongoing monitoring of the performance and behavior of our AI/ML systems. This

includes:

- Performance monitoring (e.g., analysis of false positive/negative rates)
- Incident reporting and analysis
- Collection of real-world performance data
- Regular review and updates of AI models based on new data and evolving threat landscapes.

Supporting the AI Journey of Our Customers

Among the most concerning developments for security teams is the rise of agentic cyberattacks that use AI to operate independently, make decisions, and adapt in real time with unprecedented speed and scale. Customers need next-generation, AI-driven cybersecurity solutions to combat these threats within a complex operational and regulatory environment. At Skyhigh, we are committed to supporting our customers as they navigate the complex requirements of an AI landscape that is evolving faster than ever. Guided by our foundational Skyhigh SecurityAI principles, we are dedicated to protecting fundamental human rights – and building a safer, more secure, and more trustworthy digital future powered by responsible AI.

Contact Information

For more information about Skyhigh Security's AI/ML practices or our controls and measures to comply with applicable laws, please contact your designated Skyhigh Security representative. Or reach out to our privacy and AI compliance team at privacy@trellix.com

About This Data Sheet

Please note that the information provided in this document concerning technical or professional subject matter is for general awareness only, may be subject to change, and does not constitute legal or professional advice, a warranty of fitness for a particular purpose, or a guarantee compliance with applicable laws. This Data Sheet is reviewed and updated on an annual, or as needed, basis.