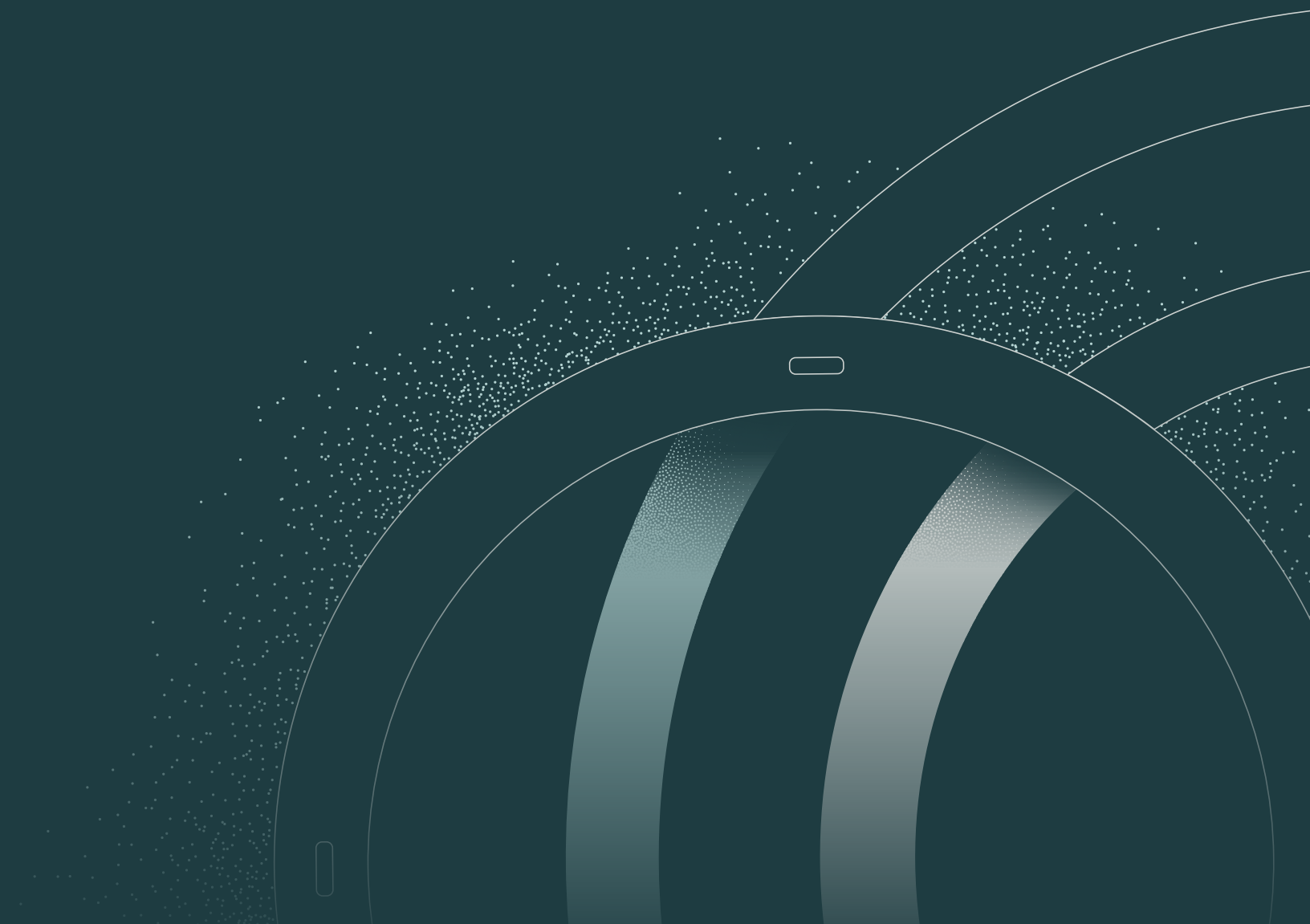
censys

CUSTOMER WORKBOOK

How To Keep Up With Evolving AI Practices



INTRODUCTION

How To Keep Up With Evolving AI Practices

This workbook is designed to help security teams use **Censys Attack Surface Management (ASM)** to proactively manage the risks introduced by the next generation of AI. Leveraging the principles from the Ultimate Guide to Detection Engineering, this guide focuses on finding what others miss across all 65,535 ports.

Core Topics

TOPIC 1

Addressing "Mythos" Risks (Frontier AI Capabilities)

THE RISK

The window of exposure has shrunk to near zero

With the emergence of frontier models like Claude Mythos, AI can now autonomously identify zero-day vulnerabilities and generate working exploits. For defenders, this means the "window of exposure" has shrunk to near zero.

In this new era, pre-emptive attack surface reduction has never been more important.

Organizations must prioritize gaining a granular understanding of everything that is exposed, recognizing that in this new era, a CVE may never actually be published for a vulnerability. By focusing on footprint reduction and maintaining a clear, continuous understanding of what is exposed—and evaluating potential blast radius—security teams can stay ahead.

In this model, comprehensive inventory management often becomes more critical than traditional risk assessment.

Step-by-Step Defense

1. **Enable the Censys AI Assistant**
   Use Censys Assistant to match the speed of AI-driven attacks.
2. **Conduct "Resolution Hour" Triage**
   When a new critical CVE is announced, ask the Assistant: *"Which of my assets are vulnerable to [CVE-ID]?"*
3. **Automate Exploit Context**
   Navigate to the **Risks** page in **ASM**. Prioritize assets where Censys indicates **Active Exploitation** is occurring in the wild, as these are the primary targets for autonomous AI agents.
4. **Configure Saved Query Alerts**
   Set up Slack, Teams, or notification automation to enable within-the-hour alerts to your most-used messaging system.

TOPIC 2

Governing Ephemeral Cloud-Based AI Resources

THE RISK

Shadow AI instances are invisible to your security controls

Data scientists often spin up "Shadow AI" instances in AWS, GCP, or Azure to test models. These ephemeral assets often bypass corporate security controls and may contain sensitive training data.

Step-by-Step Defense

1 **Configure Cloud Connectors**

Ensure Cloud Connectors for all providers (AWS, Azure, GCP) are active.

2 **Audit via Cloud Tags**

Use the ASM Inventory to filter by cloud metadata fields. Search for keywords like `ai-test`, `gpu-instance`, or `jupyter` in the asset tags to identify unmanaged AI experiments. Use this ASM QL query to help identify potential Shadow AI instances across cloud providers:

```
(host.cloud: {aws, google, azure, "Amazon Aws", "Google Cloud", "Microsoft Azure", "Microsoft Corporation"} or cloud.aws.tags.value: * or cloud.gcp.tags: * or cloud.azure.tags: *) and (cloud.aws.tags.value: {"ai-test", "gpu-instance", "jupyter"} or cloud.gcp.tags.value: {"ai-test", "gpu-instance", "jupyter"} or cloud.azure.tags.value: {"ai-test", "gpu-instance", "jupyter"} or tags: {"ai-test", "gpu-instance", "jupyter"} or host.tags: {"ai-test", "gpu-instance", "jupyter"})
```

3 **Configure Saved Query**

Configure saved queries with the cloud-based keywords necessary for review or use the example provided above.

4 **Configure Alerts**

Configure automated alerting of the saved query with Slack, Teams, or notification automation to enable within the hour alerts to your most used messaging systems for faster triage and review. above.

TOPIC 3

Uncovering Hidden AI Tools on Non-Standard Ports

THE RISK

Most scanners are blind to AI tools hiding on non-standard ports

Many AI orchestration tools (like Ollama, Xinference, or local LLM UIs) default to non-standard ports (e.g., 11434). Most scanners only check the "Top 1000" ports, leaving these tools invisible to security.

Step-by-Step Defense

1 **Leverage Full-Port Visibility**

Unlike traditional tools, Censys ASM scans all **65,535 ports**. Full Port coverage insights specific to your attack surface can be viewed in the Insights and Trends dashboard in ASM. Filter on non-standard ports to review.

2 **Hunt for AI Signatures**

Use the ASM Query Builder to search for common AI service fingerprints:

- `host.services.software.vendor: "Ollama"`
- `host.services.http.response.body: "Xinference"`
- `host.services.port: {11434, 8080, 7860}` – Common ports for AI web UIs

(Continued on next page)

3 Example Queries

Use these ASM QL queries to hunt for AI services across your attack surface.

AI SERVICE FINGERPRINTS

```
host.services.http.response.body: {Ollama, Xinference} or web_entity.instances.http.response.body: {Ollama, Xinference} or
host.services.port: {11434, 8080, 7860} or web_entity.instances.port: {11434, 8080, 7860}
```

SERVICE DETECTION FOR AI/ML WORKFLOWS

```
(host.services.port: {8888, 6006, 5000, 7077, 8080, 8265} or host.services.software.product: {Jupyter, TensorBoard, MLflow,
Spark, Ray} or host.services.service_name: {Jupyter, TensorBoard, MLflow, Spark, Ray} or web_entity.port: {8888, 6006, 5000,
7077, 8080, 8265} or web_entity.instances.port: {8888, 6006, 5000, 7077, 8080, 8265} or web_entity.instances.software.
product: {Jupyter, TensorBoard, MLflow, Spark, Ray} or web_entity.instances.service_name: {Jupyter, TensorBoard, MLflow,
Spark, Ray})
```

4 Configure Saved Query

Configure saved queries with the cloud-based keywords necessary for review or use the example provided above.

5 Configure Alerts

Configure automated alerting of the saved query with Slack, Teams, or notification automation to enable within the hour alerts to your most used messaging systems for faster triage and review.

PLATFORM BONUS

Leverage the Censys Platform

1. Run exploratory queries looking for AI/ML tooling on nonstandard ports

```
(host.services: (labels.value: "AI" and not port: 80 and not port: 443))
or (web.labels.value: "AI" and not web.port: 80 and not web.port: 443)
```

2. Tune the query to your attack surface
3. Click **Save as Collection**
4. Enable Alerting to get notified the moment a new AI tool is exposed on your network

TOPIC 4

The Internet-First Defense: Pre-empting AI Exploitation

THE STRATEGY

Detect AI exposures at the point of attacker contact

Waiting for endpoint execution is often too late. Use Censys to detect these AI exposures at the point of attacker contact—the public internet—before they can be exploited by frontier models.

Step-by-Step Pre-emption

1. **Eliminate the Blind Spot**

   Identify every AI-related service exposed to the public internet using the `services.software.vendor: "AI/ML"` query.

2. **Verify the Attack Path**

   If an AI service is visible to Censys, it is visible to an autonomous attacker. Prioritize shutting down these public gateways immediately.

3. **Continuous Internet Discovery**

   Set up automated alerts via saved query alerting for any new "Internet-Facing AI" assets. This ensures that a developer opening an AI port is notified before an external AI agent can scan and exploit it.

   EXAMPLE QUERY

   Here's an ASM QL query to identify every AI-related service exposed to the public internet using the AI/ML vendor classification:

```
host.services.software.vendor="AI/ML" or web_entity.instances.software.vendor="AI/ML"
```

TOPIC 5

Secure AI-to-ASM Integration and True Automation

THE RISK

New automation, new attack surface

As you build your own AI-driven security workflows, "hardcoding" API keys or giving AI agents broad access can create new vulnerabilities.

Step-by-Step Secure Access

1

Deploy the ASM MCP Server

Use the Censys ASM Model Context Protocol (MCP) Server to give your AI agents secure, structured access to your attack surface data.

2

Define Agent Boundaries

By using the MCP server, you ensure the AI only interacts with verified asset data, preventing it from hallucinating infrastructure that doesn't exist.

THE STRATEGY

Move beyond manual triage with SOC-integrated automation

Integrate the Censys ASM API into your Security Operations Center (SOC) automated workflows to eliminate the lag between discovery and response.

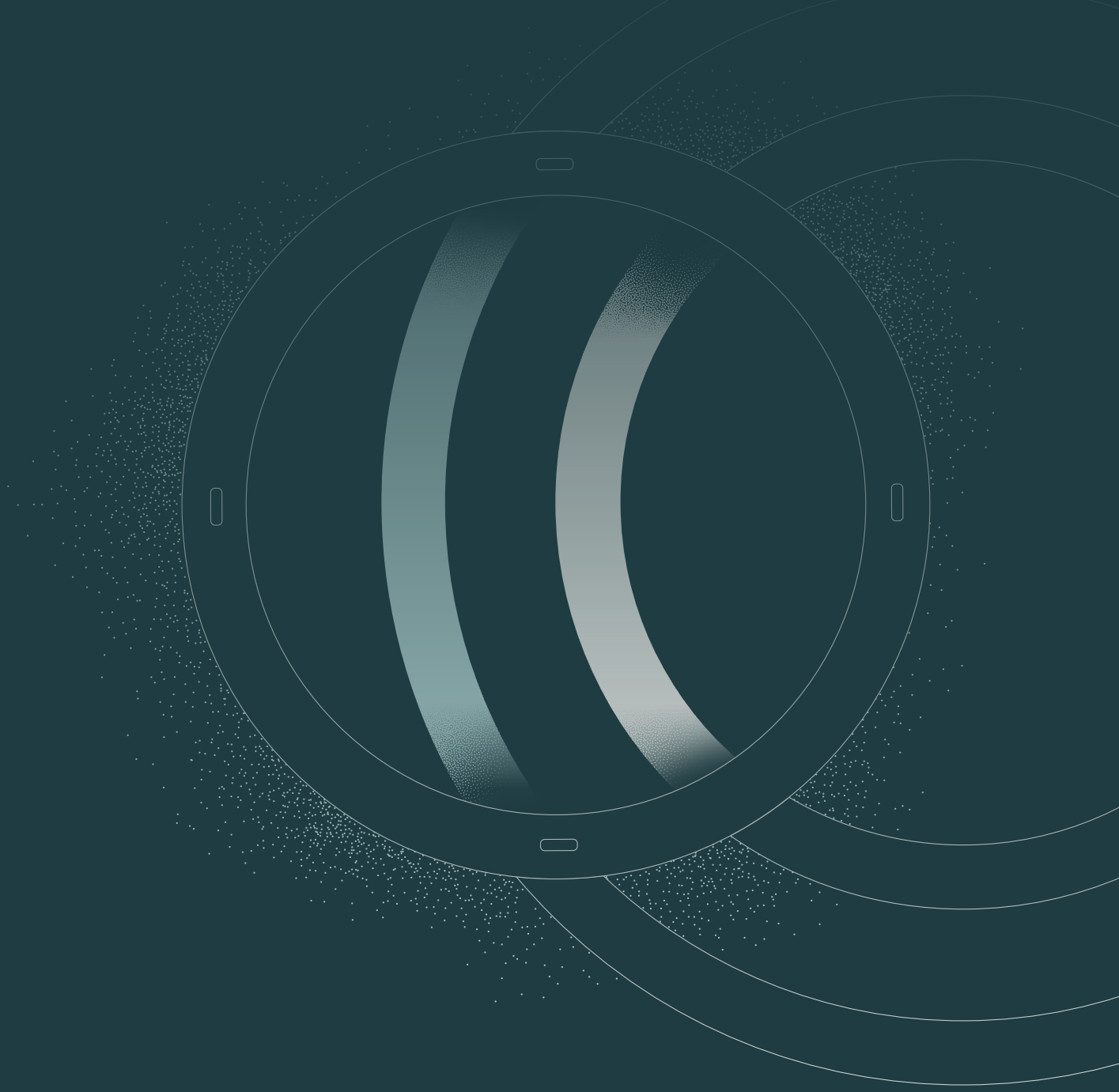
Step-by-Step True Automation

1. **Establish API Connectivity**
   Generate a Censys API key and integrate it with your SOAR platform (e.g., Splunk, Google SecOps or SentinelOne)

2. **Automate "New Asset" Ticket Creation**
   Set up a webhook or saved query with alerting in Censys ASM to trigger every time a new AI-related asset is discovered. Leverage existing integrations or create a Jira or ServiceNow ticket with the asset details provided with the saved query. *Please note: Webhooks could potentially lead to a high volume of tickets created.*

3. **Programmatic Risk Remediation**
   Use the `/risks` endpoint to pull high-priority AI vulnerabilities. Configure your automation to update firewall rules or trigger a vulnerability scan (like Tenable or Qualys) the moment a new risk is flagged.

4. **Continuous Validation**
   Schedule an automated weekly report via the API that compares "Discovered AI Assets" vs. "Approved AI Assets" to maintain a Zero Trust posture.



censys

Censys is the authority for Internet intelligence and insights. Delivering the most complete, accurate, and up-to-date global map of Internet infrastructure, Censys provides industry leading solutions for attack surface management, threat hunting, and proactive incident response. Global governments, Fortune 500 companies, and security providers around the world trust Censys to uncover risks faster, respond more effectively, and prevent breaches before they happen.

VISIT
censys.com

CONTACT
hello@censys.com