

Amenazas Emergentes

Construyendo confianza en un panorama digital en evolución.

Mauricio Romero
eddyesco@amazon.com

LATAM – NOLA – Sector Público



La seguridad es ahora un imperativo empresarial

91% Organizaciones que informan al menos un incidente o violación cibernética

56% Organizaciones que informan que sufrieron consecuencias relacionadas en una medida moderada o grande.

70% Organizaciones que informan que la seguridad cibernética estaba en la agenda de su junta directiva de forma regular, ya sea mensual o trimestralmente

Deloitte 2023 Global Future of Cyber Survey

"Hoy en día, estamos viendo el surgimiento de nuevas actitudes poderosas en lo que respecta a la cibernética. Los líderes están mirando la cibernética a través de una lente nueva y nítida, una que revela el valor comercial inherente que puede surgir al incorporar la cibernética. No solo en toda la empresa, sino como parte crucial de una poderosa estrategia de crecimiento".

Emily Mossburg,
Líder Cibernética Global de Deloitte

Panorama de seguridad de las telecomunicaciones

- Existe una expansión de las redes de telecomunicaciones impulsadas por tendencias como BYOD, nuevos tipos de dispositivos conectados y redes de comunicación más amplias.
- Las empresas de telecomunicaciones son atacadas directamente o como un medio para afectar industrias como finanzas, servicios públicos, atención médica y gobierno.
- Cada vez cobra mas relevancia a estas empresas considerar estándares de seguridad de industria (PCI, HIPPA) aplicar controles de seguridad (SOC, ISO) y estándares nacionales de gobierno e industria.

Algunos datos relacionados:



Un tercio de todos los ataques de denegación de servicio distribuido (DDOS) se dirigieron a las empresas de telecomunicaciones *



Se espera que el tamaño del mercado de ciberseguridad de telecomunicaciones alcance los USD 82,64 mil millones para 2030.



En 2021, el 76% de los 500 principales ataques de ciberseguridad se lanzaron contra la infraestructura de telecomunicaciones *.

Voz del cliente...

- Se enfrentan a barreras para mantener su postura de seguridad, ya que **carecen de visibilidad de las amenazas cibernéticas** en su patrimonio digital.
- Los actores de **amenazas** están utilizando **IA, aprendizaje automático y otras tecnologías** para lanzar ataques cada vez más sofisticados.
- Necesitan **consolidar herramientas y plataformas de seguridad** para obtener información procesable de los datos de seguridad.
- La escasez global de **talento cibernético** agrava los desafíos para detectar y responder rápidamente a las amenazas cibernéticas.

Las amenazas cibernéticas están evolucionando rápidamente. El robo de identidad, el fraude, el ransomware, el espionaje y el lavado de dinero están en constante crecimiento.



\$4.45M Costo promedio de una violación de datos en 2023.



51% de las organizaciones que planean aumentar las inversiones en seguridad como resultado de la infracción.

Sources: Gartner, Global Enterprise Cyber Security Industry Research Report 2023.
IBM Cost of Data Breach Report 2023

Evolución de las amenazas

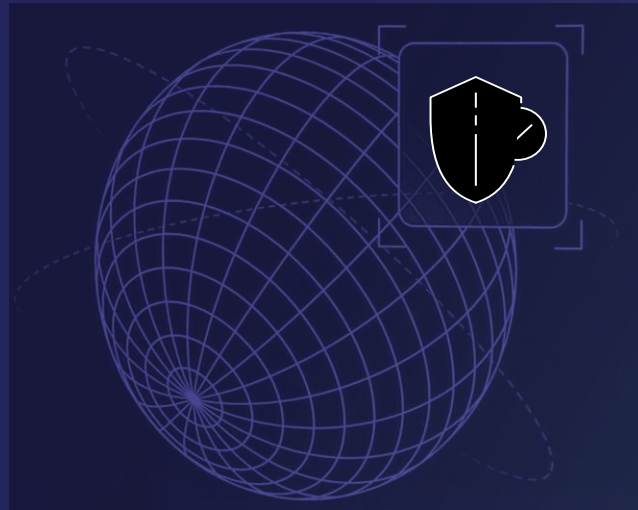
Ataques puntuales

Actores individuales

Desde un garaje

Vectores simples

Habilidades raras



Ataques sostenidos

Grupos organizados

Infraestructura global

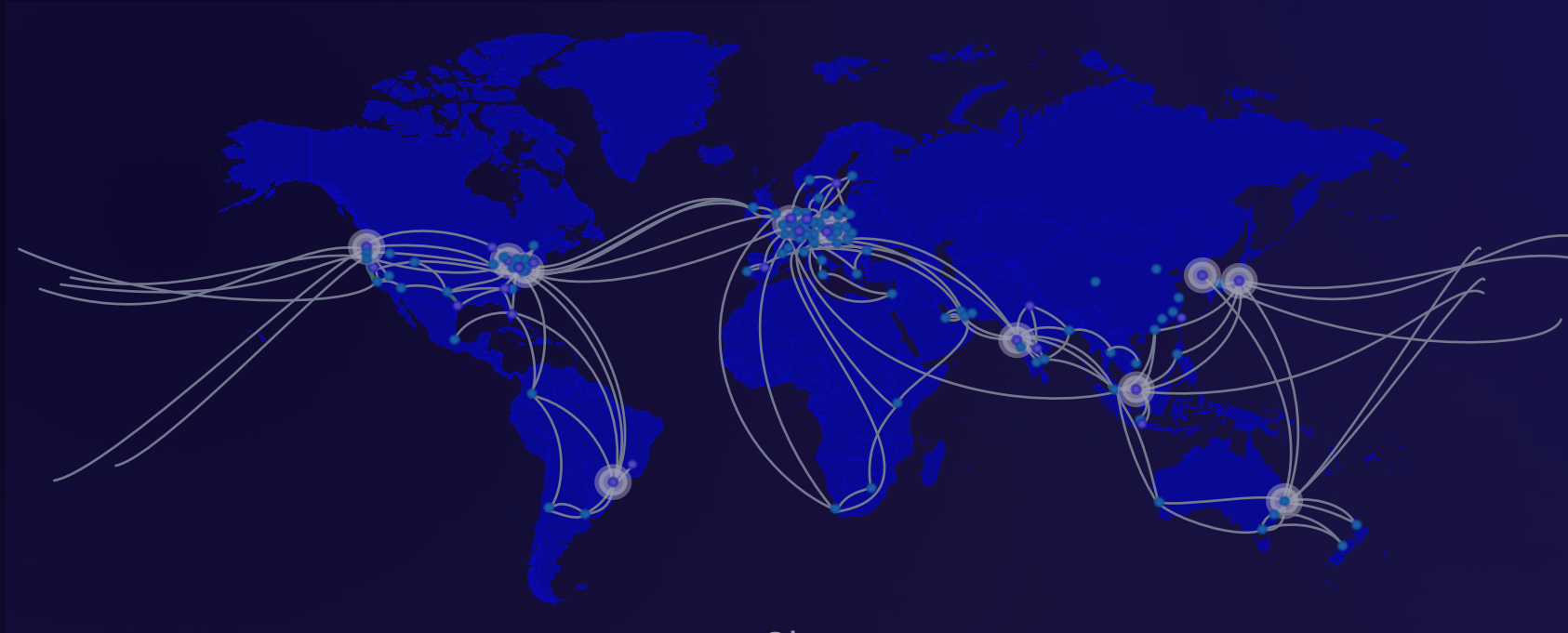
Múltiples vectores

Acelerados con IA

Escala de peticiones y amenazas globales

Más de **1 billón de solicitudes de DNS** en todas las regiones de AWS todos los días

124.000 nuevos dominios maliciosos detectados de media diaria



Inspeccionado más de
360T
eventos en promedio por día

Observa
750M+
Interacciones de amenazas
potenciales diarias

Cada día
400M
Actividades clasificadas
como maliciosas diarias

Tendencias de la industria que observamos para DDoS

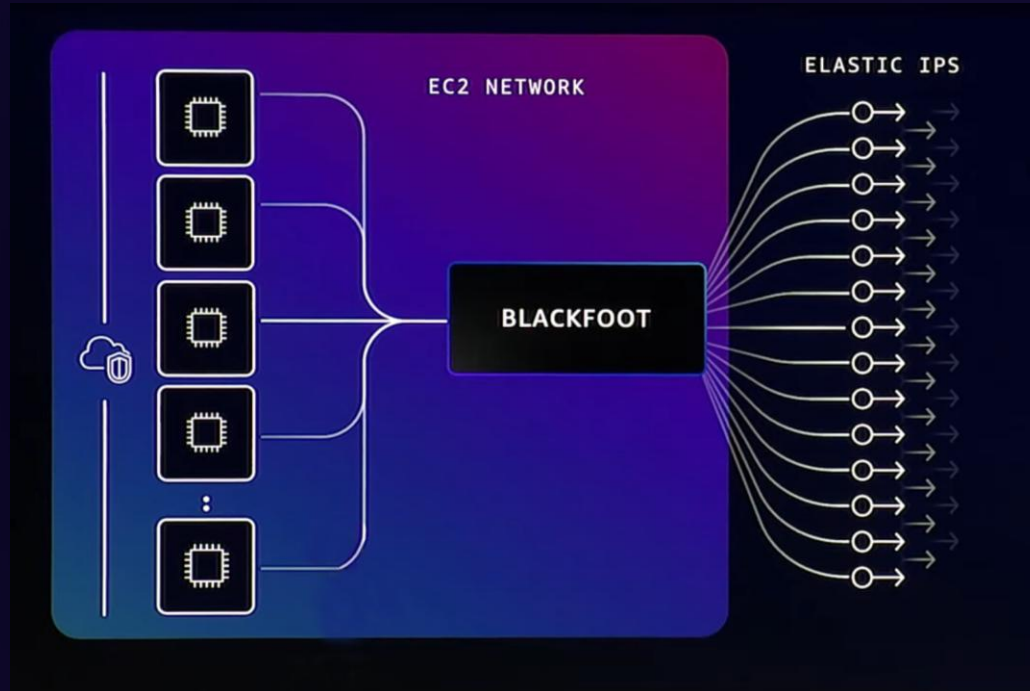
- **Extorsión:** Negocios nativos digitales como el comercio electrónico o los juegos en línea
- **Competencia:** individuo/empresa utiliza el evento DDoS contra el sitio web de la competencia para apoderarse de sus clientes. Esto es común en la industria del juego.
- **Hacktivismo:** objetivos DDoS con fines políticos o sociales, por ejemplo: Operación Megaupload, Universal Music Group, Departamento de Justicia de EE. UU., RIAA
- **Guerra cibernética:** actor estatal lanza evento DDoS contra oponente

Takeaway:

El volumen de amenazas está aumentando y la naturaleza de las amenazas está evolucionando para centrarse más en las aplicaciones

Escala de las amenazas en la red de AWS

PROTECCION E INTELIGENCIA DE RED DE HONEYPOT DINÁMICO



13T

Flujos por hora

10K+

Madpots concurrentes
en cualquier momento

90s/3min

Inicia escaneo / completa exploit

12.5%

IP bloqueadas cambian
cada 3 minutos

2.4T

Solicitudes de análisis
maliciosas evitadas en 6 meses

Hallazgos de secuencias de ataque en los principales escenarios de amenazas

Identidad comprometida

Credenciales de IAM utilizadas para realizar múltiples solicitudes sospechosas que se alinean con comportamientos de ataque conocidos, como persistencia, evasión y creación de cargas de trabajo

Datos comprometidos

Secuencia de acciones que sugieren el compromiso o la exfiltración de datos en uno o más buckets de S3 mediante credenciales potencialmente utilizadas indebidamente

Host comprometido

Secuencia detectada en una instancia EC2, como ejecuciones de comandos sospechosas seguidas de la ejecución de archivos maliciosos y conexión con un grupo de criptominería conocido

Tres enfoques de seguridad de IA generativa - 1

- Tres formas de pensar en la IA generativa + seguridad



Seguridad de la IA generativa

¿Cómo protejo mis aplicaciones empresariales que aprovechan la IA generativa?

MITRE ATLAS™

MITRE ATLAS™ (ADVERSARIAL THREAT LANDSCAPE FOR ARTIFICIAL-INTELLIGENCE SYSTEMS)

MITRE ATLAS™ es una base de conocimientos viva y accesible a nivel mundial de las tácticas y técnicas de los adversarios contra los sistemas habilitados para IA basada en observaciones de ataques del mundo real y demostraciones realistas de equipos y grupos de seguridad de IA.

Reconnaissance&	Resource Development&	Initial Access&	ML Model Access	Execution&	Persistence&	Privilege Escalation&	Defense Evasion&	Credential Access&	Discovery&	Collection&	ML Attack Staging	Exfiltration&	Impact&
5 techniques	7 techniques	6 techniques	4 techniques	3 techniques	3 techniques	3 techniques	3 techniques	1 technique	4 techniques	3 techniques	4 techniques	4 techniques	6 techniques
Search for Victim's Publicly Available Research Materials	Acquire Public ML Artifacts	ML Supply Chain Compromise	ML Model Inference API Access	User Execution&	Poison Training Data	LLM Prompt Injection	Evade ML Model	Unsecured Credentials&	Discover ML Model Ontology	ML Artifact Collection	Create Proxy ML Model	Exfiltration via ML Inference API	Evade ML Model
Search for Publicly Available Adversarial Vulnerability Analysis	Obtain Capabilities&	Valid Accounts&	ML-Enabled Product or Service	Command and Scripting Interpreter&	Backdoor ML Model	LLM Plugin Compromise	LLM Prompt Injection		Discover ML Model Family	Data from Information Repositories&	Backdoor ML Model	Exfiltration via Cyber Means	Denial of ML Service
Search Victim-Owned Websites	Develop Capabilities&	Evade ML Model	Physical Environment Access	LLM Plugin Compromise	LLM Prompt Injection	LLM Jailbreak	LLM Jailbreak		Discover ML Artifacts	Data from Local System&	Verify Attack	LLM Meta Prompt Extraction	Spamming ML System with Chaff Data
Search Application Repositories	Acquire Infrastructure	Exploit Public-Facing Application&	Full ML Model Access						LLM Meta Prompt Extraction		Craft Adversarial Data	LLM Data Leakage	Erode ML Model Integrity
Active Scanning&	Publish Poisoned Datasets	LLM Prompt Injection											Cost Harvesting
	Poison Training Data	Phishing&											External Harms
	Establish Accounts&												

Source: <https://atlas.mitre.org/>



OWASP® Top for 10 Large Language Models

- Algunos artículos requieren algo más que contramedidas técnicas

LLM01

Prompt Injection

Esto manipula un lenguaje grande (LLM) a través de entradas astutas, causando acciones no deseadas por parte del LLM. Las inyecciones directas sobrescriben los avisos del sistema, mientras que las indirectas manipulan las entradas de fuentes externas.

LLM02

Insecure Output Handling

Esta vulnerabilidad ocurre cuando se acepta una salida de LLM sin escrutinio, exponiendo los sistemas backend. El uso indebido puede tener consecuencias graves como XSS, CSRF, SSRF, escalada de privilegios o ejecución remota de código.

LLM03

Training Data Poisoning

Esto ocurre cuando se manipulan los datos de entrenamiento de LLM, introduciendo vulnerabilidades o sesgos que comprometen la seguridad, la eficacia o el comportamiento ético.

LLM04

Model Denial of Service

Los atacantes provocan operaciones que consumen muchos recursos en los LLM, lo que lleva a la degradación del servicio o a altos costos. La vulnerabilidad se magnifica debido a la naturaleza intensiva en recursos de los LLM y la imprevisibilidad de las entradas de los usuarios.

LLM05

Supply Chain Vulnerabilities

El ciclo de vida de las aplicaciones de LLM puede verse comprometido por componentes o servicios vulnerables, lo que lleva a ataques de seguridad. El uso de conjuntos de datos de terceros, modelos preentrenados y complementos puede agregar vulnerabilidades.

LLM06

Sensitive Information Disclosure

Los LLM pueden revelar inadvertidamente datos confidenciales en sus respuestas, lo que lleva a un acceso no autorizado a los datos, violaciones de la privacidad y violaciones de seguridad. Es crucial implementar la desinfección de datos y políticas estrictas de usuario para mitigar esto.

LLM07

Insecure Plugin Design

Los complementos LLM pueden tener entradas inseguras y un control de acceso insuficiente. Esta falta de control de aplicaciones hace que sean más fáciles de explotar y puede tener consecuencias como la ejecución remota de código.

LLM08

Excessive Agency

Los sistemas basados en LLM pueden emprender acciones que conducen a consecuencias no deseadas. El problema surge de la funcionalidad, los permisos o la autonomía excesivos otorgados a los sistemas basados en LLM.

LLM09

Overreliance

Los sistemas o las personas que dependen demasiado de los LLM sin supervisión pueden enfrentar información errónea, falta de comunicación, problemas legales y vulnerabilidades de seguridad debido a contenido incorrecto o inapropiado generado por los LLM.

LLM10

Model Theft

Esto implica el acceso no autorizado, la copia o la exfiltración de modelos LLM patentados. El impacto incluye pérdidas económicas, ventaja competitiva comprometida y acceso potencial a información confidencial.

Tres enfoques de seguridad de IA generativa - 2

- Tres formas de pensar en la IA generativa + seguridad



**Seguridad de la IA
generativa**



**IA generativa
en la seguridad**

¿Cómo utilizar la IA generativa?
para minimizar las vulnerabilidades,
las amenazas y los riesgos?

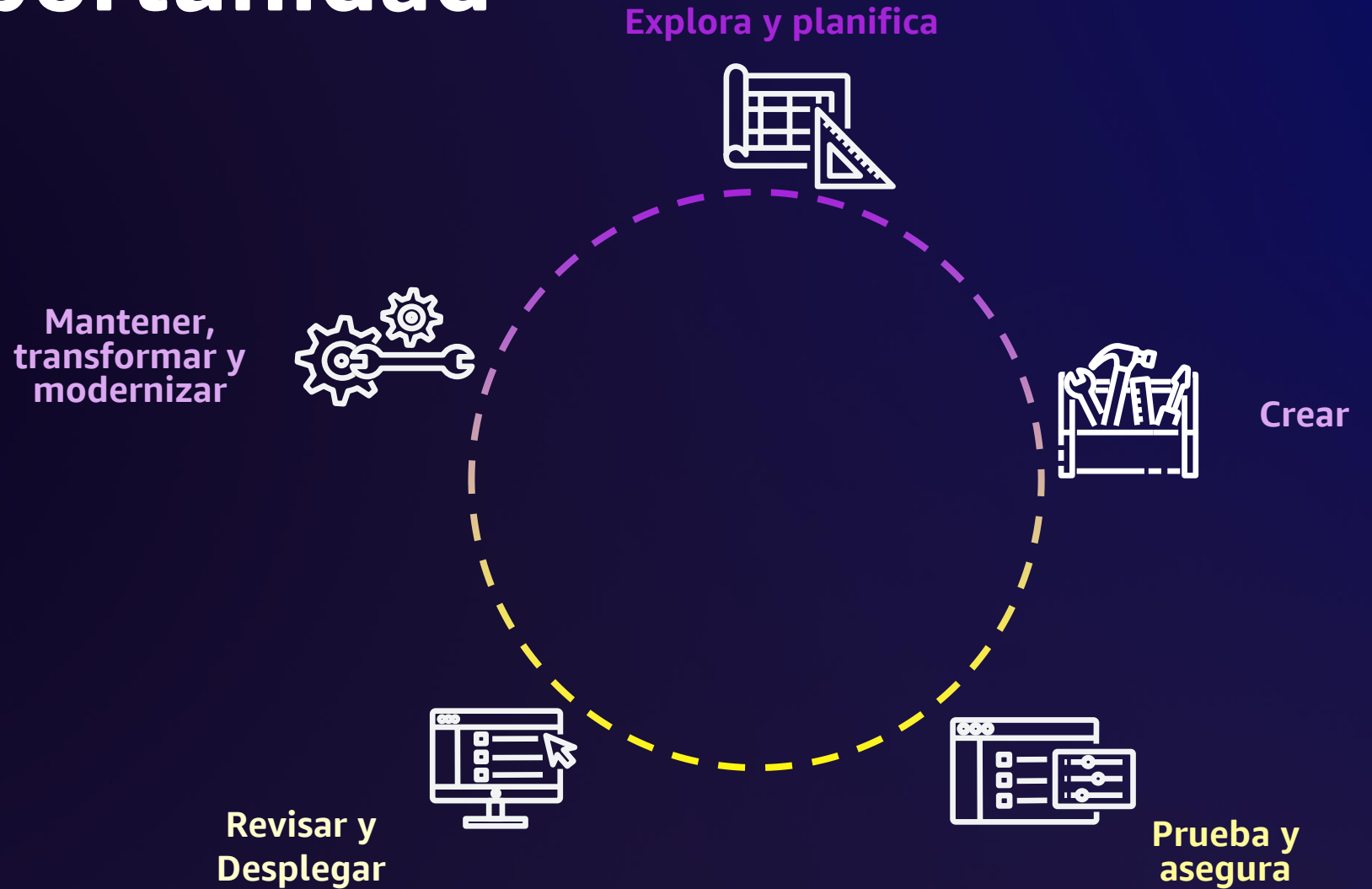
La IA generativa puede empoderar a los equipos



Source: [IDC - New Use Cases for Generative AI in Security Analytics](#)

Amenaza y oportunidad

IA Generativa en el ciclo de vida del software



Tres enfoques de seguridad de IA generativa - 3

- Tres formas de pensar en la IA generativa + seguridad



Seguridad de la IA generativa



IA generativa en la seguridad



Seguridad frente a amenazas generativas impulsadas por IA

¿Cómo protegerme contra los actores de amenazas que emplean la IA generativa?

Potenciales amenazas impulsadas por IA

UNA LISTA NO EXHAUSTIVA



Denegación de servicio

- SYN Floods
- Reflection Attacks
- Web Request Floods



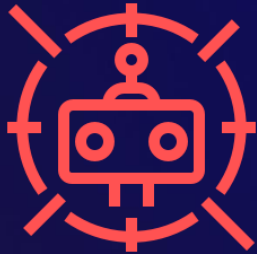
Vulnerabilidades de aplicaciones

- SQL Injection
- Cross-site Scripting (XSS)
- OWASP Top 10
- Common Vulnerabilities and Exposures (CVE)
- Privilege Escalation
- Data Exfiltration



Buenos bots

- Search Engine Crawling
- Website Health Monitoring
- Vulnerability Scanning



Bad Bots

- Comment Spam
- SEO Spam
- Fraud
- Account Takeover

Threat Intelligence como estrategia

COMO GUARDDUTY APROVECHA IA PARA PREVENIR LOS ATAQUES IMPULSADOS POR AI Y ESCALA GLOBAL



Alerta temprana

Una red neuronal proporciona inteligencia de amenazas de 182k nuevos dominios maliciosos por día a Amazon GuardDuty



Tácticas emergentes

Observamos más de 100 millones de amenazas potenciales todos los días en todo el mundo, con aproximadamente 500,000 de esas actividades observadas clasificadas como maliciosas.



Prevención automática

Activamos protecciones automatizadas con una precisión del 99,998 % en AWS Shield, Amazon Virtual Private Cloud (2,6T), Amazon S3 (24B) y AWS WAF

// Todo falla, todo el tiempo"

Puntos únicos
de falla

Carga excesiva

Configuración
incorrecta y errores

Amenazas

Destino
compartido



Werner Vogels CTO, Amazon.com

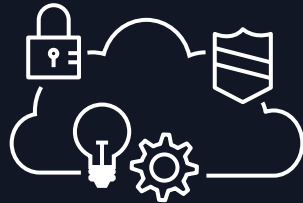
Resiliencia

CAPACIDAD DE UNA CARGA DE TRABAJO PARA
RECUPERARSE DE INTERRUPCIONES DE INFRAESTRUCTURA O SERVICIO

El modelo mental

Alta disponibilidad

Resistencia a fallas comunes a través de mecanismos de diseño y operación en un sitio primario



Servicios principales, objetivos de diseño para cumplir con los objetivos de disponibilidad

Recuperación ante desastres

Volver al funcionamiento dentro de destinos específicos en un sitio de recuperación para fallas que HA no puede manejar



Objetivos de copia de seguridad y recuperación, abastecimiento de datos, recuperación gestionada

Mejora continua



CI/CD, observabilidad, más allá de las pruebas hacia patrones de ingeniería del caos



Estrategia de defensa

DEFENSA EN PROFUNDIDAD + ZERO-TRUST

Policies, Procedures & Awareness

Network & Edge Protection

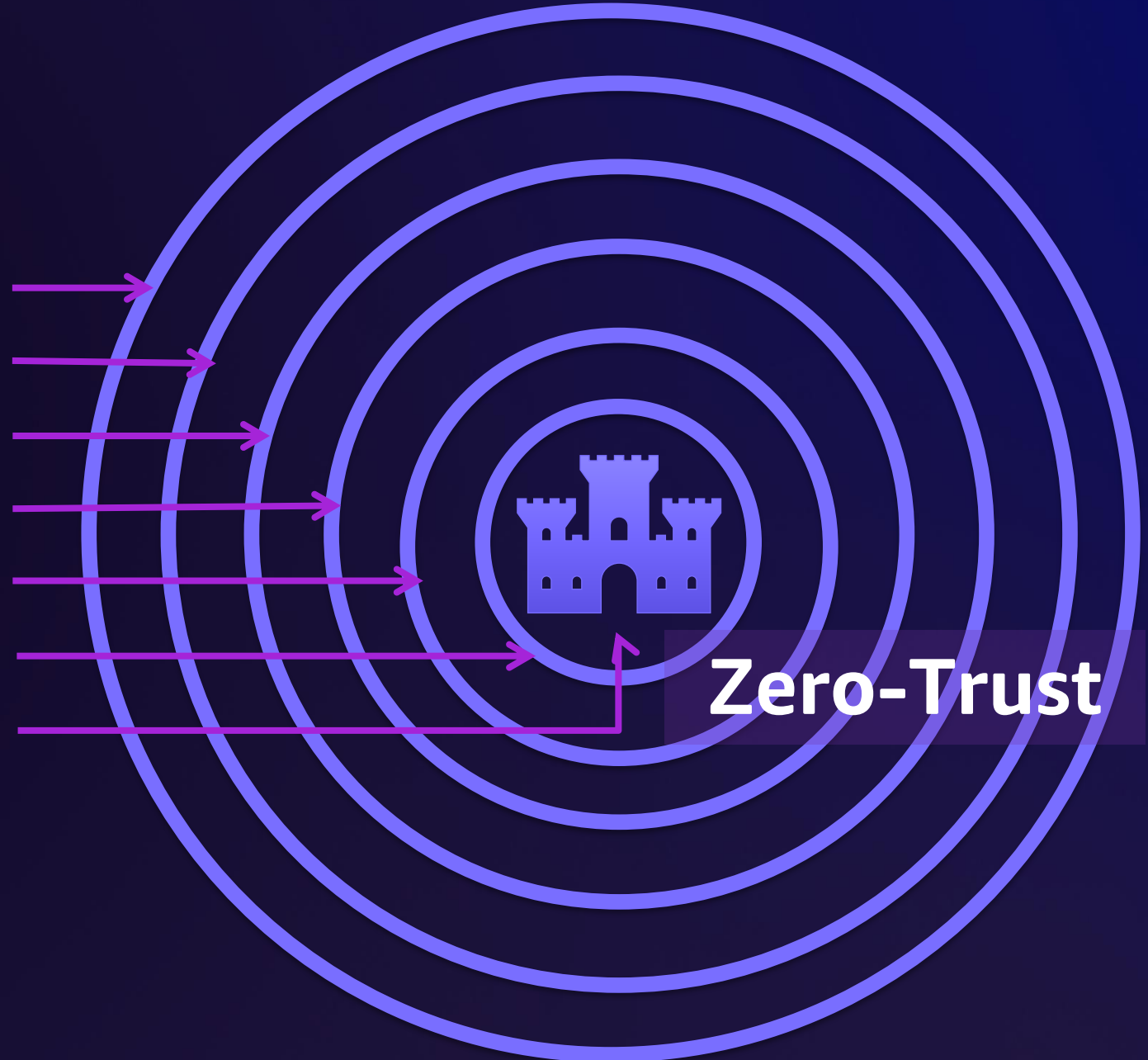
Identity & Access Management

Threat Detection & Incident Response

Infrastructure Protection

Application Protection

Data Protection



Matriz de alcance de seguridad de IA generativa



“

Brindar servicios seguros y operativamente excelentes para nuestros clientes es, y siempre será, nuestra principal prioridad.”

Matt Garman
Chief Executive Officer
Amazon Web Services



Gracias



Feedback

<https://pulse.aws/survey/WQURD8TF>

Mauricio Romero

eddyesco@amazon.com

LinkedIn: RomeroGT

Apéndice

Recursos adicionales

Securing generative AI

Generative AI applications are increasingly becoming essential to how we work and live. They are used in a wide range of applications, from content creation to customer support. As these applications become more widespread, it's crucial to ensure they are secure and reliable. This means protecting the data they use, the models they run on, and the outputs they generate. In this presentation, we'll explore the challenges of securing generative AI and share some best practices for doing so.

One of the biggest challenges is protecting the data used to train the models. This data is often sensitive and can be a valuable asset for the organization. It's important to ensure that this data is stored securely and that access is restricted to only those who need it. Additionally, it's important to ensure that the data is not used for anything other than training the models.

Another challenge is protecting the models themselves. These models are often complex and can be difficult to understand. This makes it difficult to identify and fix vulnerabilities. It's important to ensure that the models are tested thoroughly and that any vulnerabilities are addressed as quickly as possible.

Finally, it's important to protect the outputs of the models. These outputs can be sensitive and can be used in a variety of ways. It's important to ensure that the outputs are protected and that they are not used in a way that could harm the organization or its customers.

By following these best practices, organizations can ensure that their generative AI applications are secure and reliable. This will help them to get the most out of these powerful tools and to protect their data and their customers.

Keep your data private

Generative AI applications often use large amounts of data to train their models. This data is often sensitive and can be a valuable asset for the organization. It's important to ensure that this data is stored securely and that access is restricted to only those who need it. Additionally, it's important to ensure that the data is not used for anything other than training the models.

One way to ensure data privacy is to use a secure storage solution. This could be a cloud storage service or a local storage solution. The important thing is to ensure that the data is encrypted and that access is restricted. Additionally, it's important to ensure that the data is not shared with anyone who does not need it.

Another way to ensure data privacy is to use a secure transmission method. This could be a secure web connection or a secure network connection. The important thing is to ensure that the data is encrypted and that access is restricted. Additionally, it's important to ensure that the data is not shared with anyone who does not need it.

By following these best practices, organizations can ensure that their generative AI applications are secure and reliable. This will help them to get the most out of these powerful tools and to protect their data and their customers.

Secure the model and its output

Generative AI models are often complex and can be difficult to understand. This makes it difficult to identify and fix vulnerabilities. It's important to ensure that the models are tested thoroughly and that any vulnerabilities are addressed as quickly as possible.

One way to secure the model is to use a secure environment. This could be a cloud environment or a local environment. The important thing is to ensure that the model is protected from unauthorized access and that it is not used for anything other than its intended purpose.

Another way to secure the model is to use a secure output method. This could be a secure web connection or a secure network connection. The important thing is to ensure that the output is encrypted and that access is restricted. Additionally, it's important to ensure that the output is not shared with anyone who does not need it.

By following these best practices, organizations can ensure that their generative AI applications are secure and reliable. This will help them to get the most out of these powerful tools and to protect their data and their customers.

Customer privacy for your business with a full AI/ML ecosystem

When all pieces are in place to build a secure AI/ML ecosystem, you can ensure that your data is protected and that your models are secure. This will help you to get the most out of your generative AI applications and to protect your data and your customers.

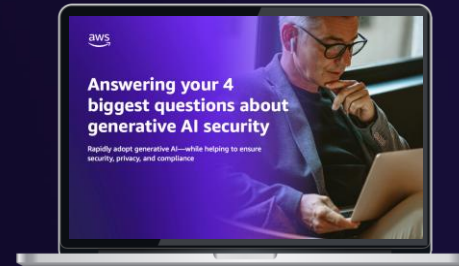
By following these best practices, organizations can ensure that their generative AI applications are secure and reliable. This will help them to get the most out of these powerful tools and to protect their data and their customers.

```

graph LR
    A[AI/ML model] --> B[AI/ML output]
    B --> C[AI/ML application]
    C --> D[AI/ML service]
    D --> E[AI/ML provider]
    C --> F[AI/ML provider]
  
```

The diagram illustrates the components and data flow of an AI/ML ecosystem. It starts with an 'AI/ML model' which produces an 'AI/ML output'. This output is then used by an 'AI/ML application'. The application interacts with an 'AI/ML service' and an 'AI/ML provider'. The 'AI/ML service' also interacts with the 'AI/ML provider'. The 'AI/ML application' is shown as a central component that receives input from the 'AI/ML output' and sends output to the 'AI/ML service' and 'AI/ML provider'. The 'AI/ML service' is shown as a component that receives input from the 'AI/ML application' and sends output to the 'AI/ML provider'. The 'AI/ML provider' is shown as the final component in the ecosystem, receiving input from both the 'AI/ML application' and the 'AI/ML service'.

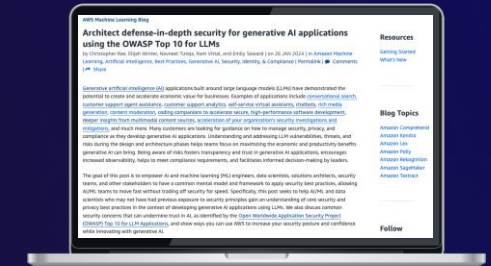
How AWS Securing the Future of Generative AI



Answering your 4 biggest questions about generative AI Security

The future of generative AI is secure - With AWS

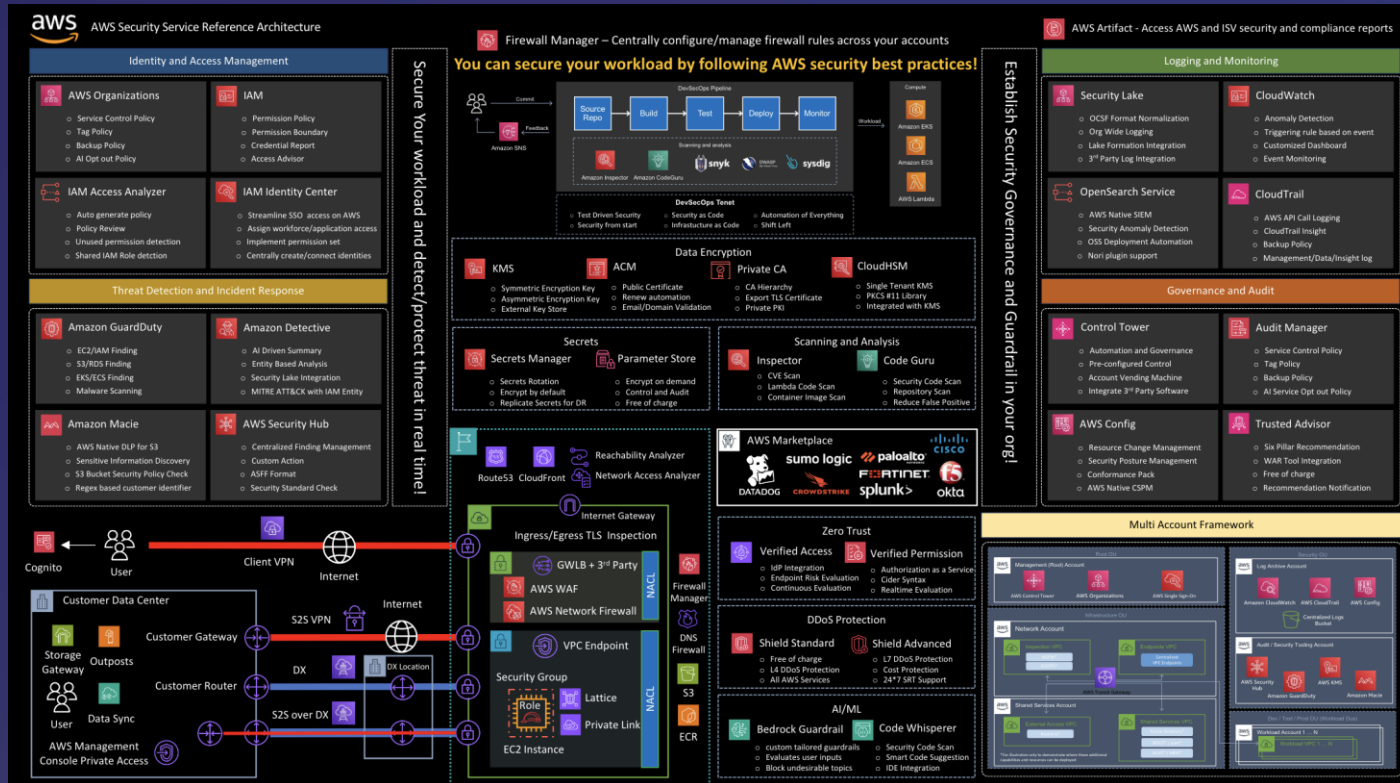
Accelerate machine learning innovation through security



- Securing generative AI: An introduction to the Generative AI Security Scoping Matrix
- Architect defense-in-depth security for generative AI applications using the OWASP Top 10 for LLMs
- Securing generative AI: Applying relevant security controls
- Automate and enhance your code security with AI powered services

<https://aws.amazon.com/ai/generative-ai/security/>

Recurso: Arquitectura de referencia de seguridad



- Definir arquitectura de seguridad estado objetivo
- Revisar e implementar seguridad ya existente
- Implementar infraestructura como código de inicio
- Impulsar discusiones sobre gobernanza de seguridad





bitly

<https://bit.ly/46kaDGX>

Otros recursos disponibles



Top re:Inforce 2025
Announcements Blog



AWS Security
Activation Days



Session Replays on
AWS Events YouTube

Become an AWS Security Champion



AWS LEARNING PLAN

Earn the AWS Digital Knowledge Security Champion Badge

Deepen your security skills with digital learning on demand and hands-on interactive content

Join the **AWS learners community** and get exclusive benefits



TRAIN NOW