Capacitación y concienciación anti-phishing

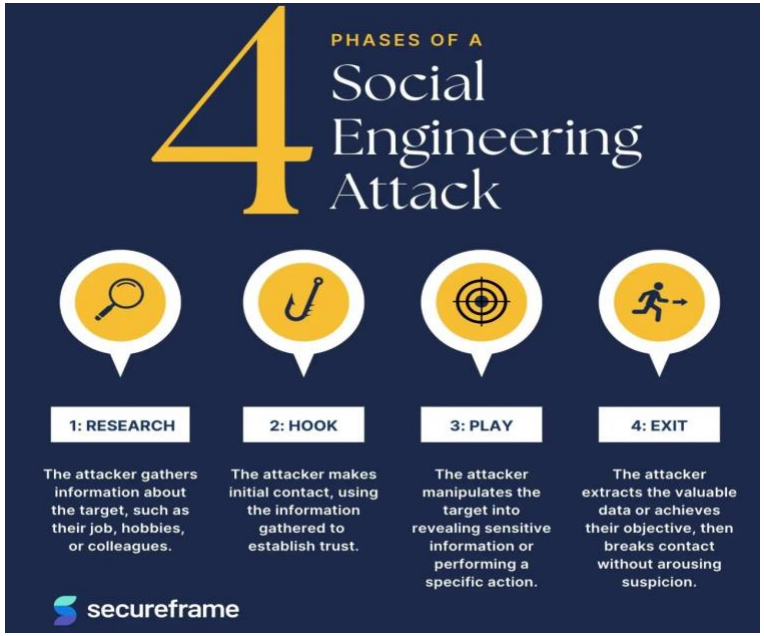
Alejandro Corletti Estrada
acorletti@darfe.es

Ingeniería social

La ingeniería social, en el contexto de ciberseguridad, son técnicas de manipulación utilizadas por ciberdelincuentes para obtener información confidencial, acceso a sistemas, o inducir a una persona a realizar acciones que pongan en riesgo su seguridad. Se basa en la psicología y el comportamiento humano, más que en conocimientos técnicos.

Según **Secureframe**
(https://secureframe.com/es-es/blog/social-engineering-statistics)

Los ingenieros sociales están utilizando tácticas cada vez más personalizadas para ganar confianza y evitar sospechas. La clonación de voz y la tecnología deepfake hacen posible que los actores de amenazas se hagan pasar por sus objetivos de maneras aún más convincentes. En un caso de alto perfil, la voz creada por IA de un director de banco se utilizó para engañar a un gerente de banco y transferir 35 millones de dólares a los actores de amenazas.
Los ataques de ingeniería social son una amenaza especialmente peligrosa para las organizaciones precisamente por el elemento humano. Los errores cometidos por usuarios legítimos son más difíciles de detectar, predecir y remediar. En muchos casos, las víctimas ni siquiera se dan cuenta de que han sido engañadas.



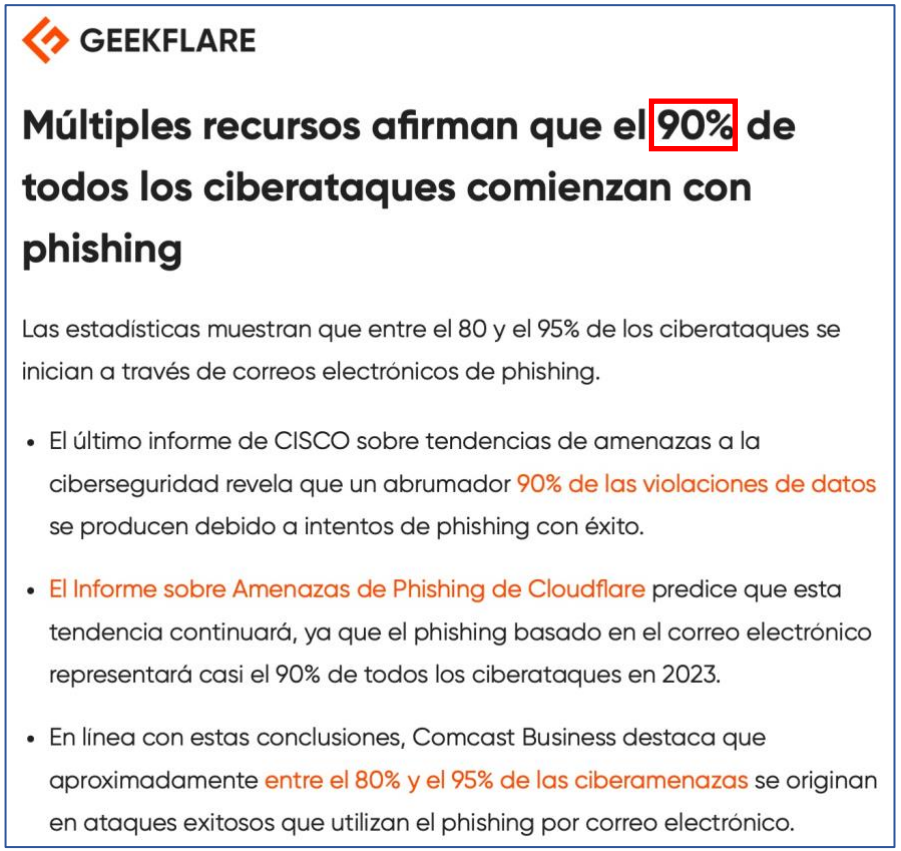
Phishing

El **phishing** es uno de los ataques de ingeniería social más comunes.

El atacante usa correo electrónico, mensajes en redes sociales, mensajería instantánea o mensajes de texto para engañar a la víctima de forma que rebele información confidencial, descargue y abra un archivo o visite una web fraudulenta.

Normalmente tienen las siguientes características:

- Tratan de atraer la atención o generar curiosidad. Por ejemplo: "*Acceso indebido identificado en su cuenta bancaría, si no ha sido usted, utilice el link de abajo para asegurarla*".
- Sensación de urgencia. Esto trata de reducir la posibilidad de que el atacado se dé cuenta de que "algo raro ocurre".
- Uso de **URLs** (Unifrom Resource Locator) recortadas (ocultan La URL real) o enlaces incrustados (redirecciones a páginas webs ocultas en parte del texto o imágenes). Esto trata de evitar que los usuarios puedan ver que la dirección a la que van a acceder no es la verdadera o que tiene un nombre sospechoso.
- Suelen tener un asunto engañoso para tratar de emular que el mensaje viene de un origen fiable.
- Copia de logos, textos o imágenes corporativas de la fuente.



GEEKFLARE

Múltiples recursos afirman que el 90% de todos los ciberataques comienzan con phishing

Las estadísticas muestran que entre el 80 y el 95% de los ciberataques se inician a través de correos electrónicos de phishing.

- El último informe de CISCO sobre tendencias de amenazas a la ciberseguridad revela que un abrumador 90% de las violaciones de datos se producen debido a intentos de phishing con éxito.
- El Informe sobre Amenazas de Phishing de Cloudflare predice que esta tendencia continuará, ya que el phishing basado en el correo electrónico representará casi el 90% de todos los ciberataques en 2023.
- En línea con estas conclusiones, Comcast Business destaca que aproximadamente entre el 80% y el 95% de las ciberamenazas se originan en ataques exitosos que utilizan el phishing por correo electrónico.

Dentro de phishing, podemos diferenciar dos tipos que son más peligrosos aún:

Spear Phishing o **Phishing dirigido**: mucho más peligroso que el phishing normal. Este tipo de phishing requiere más esfuerzo de los atacantes. Su principal diferencia con el phishing normal es que los atacantes estudian antes contra que victimas van a realizar el ataque, de forma que el e-mail o mensaje de texto que mandan puede estar dirigido usando datos personales o reales de las víctimas. Esto hace que las victima lo consideren con una probabilidad mucho más alta mensajes legítimos y caigan ante ellos, también haciendo que sea más difícil detectarlo como spam. Hoy en día con la IA (como veremos más abajo), este tema es aún mucho más peligroso.

Whaling: son un Spear Phishing dirigido a directivos o individuos con información crítica (de ahí el nombre, ataque de ballenero). Utilizan las técnicas de Spear Phishing, pero normalmente se camuflan como correos de negocio crítico y urgente.

IMPORTANTE: Antes de hacer click en cualquier enlace, comprueba su legitimidad examinando detenidamente la URL. Pasa el ratón por la parte superior de la URL. Si esta dirección difiere de la que se muestra, **no hagas click en ella.**

Entrando en más categorías de phishing, podemos mencionar también:

Vishing

El vishing es una combinación de las palabras voz y phishing. Se refiere a estafas de phishing que se hacen por teléfono. Se intenta engañar para que se revele información crucial de carácter financiero o personal. El vishing funciona igual que el phishing, pero se lleva a cabo mediante tecnología de voz. Muchas los atacantes utilizan sistemas de IVR (Respuesta de Voz Interactiva) para simular que son centros de atención al usuario o ayuda al usuario y piden a la víctima que llame a un número gratuito, otras veces directamente llaman ellos simular un agente.

Normalmente tienen las siguientes características:

- Uso de información correcta. Usan métodos informados para obtener información del atacado para dar legitimidad a su engaño.
- Urgencia.
- Spoofing telefónico
- Simulación de Service Desk: la llamada trata de simular muchas veces un centro de servicio al usuario legítimo y podemos escuchar al fondo y ruido de call center, obviamente no es real sino un efecto de sonido incorporado a la llamada.

Smishing

El Smishing es un ataque similar al Phishing o Vishing salvo que utiliza mensajes **SMS** para tratar de engañar a la víctima. Comparte la mayoría de características del phishing.

Scams de Voz (Scam: estafa, timo)

Llamadas telefónicas con objeto de estafar o sonsacar datos a la víctima haciéndose pasar por representantes de organizaciones legítimas. Se podría considerar un tipo de phishing, pero la verdad es que los voice scams existen desde antes de la creación del concepto de pshishing moderno.

Características típicas:

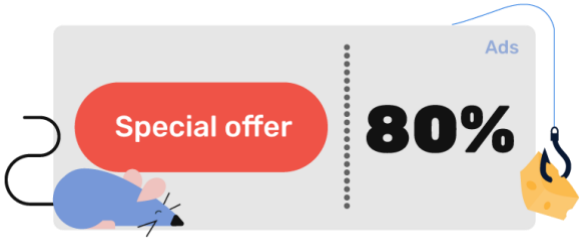
- Suplantación de organizaciones bancarias, de seguros, salud o estatales.
- Uso de guiones profesionales.
- Urgencia.


GEEKFLARE

el 40% cae en la falsa urgencia de los correos electrónicos de RRHH, saltándose una verificación crucial

La seguridad del correo electrónico es una prioridad absoluta para todo tipo de empresas, ya que los piratas informáticos se aprovechan de nuestro comportamiento rutinario y de nuestra urgencia. El último análisis de KnowBe4 muestra que el 40% de estos mensajes engañosos dirigidos a la manipulación imitan a mensajeros de RR.HH. con el objetivo de tratar a los empleados en términos de estabilidad laboral o falta de tiempo.

Baiting o Uso de Cebo (Bait: cebo)



Las técnicas de Baiting se basan en ofrecer un premio a la víctima, o que consiga algo. Esto puede estar representado como colgar en la web una falsa aplicación para descarga con software malicioso o que el atacante deje tiradas unidades de USB con malware, el objetivo es que algún incauto los recoja y lo conecte a su computador, infectándolo en el proceso. Durante la pandemia, se puso de moda también los códigos GR falsos, con enlaces a páginas fraudulentas.

Desgraciadamente los ataques de Baiting son bastante efectivos, dado que los seres humanos suelen bajar la guardia cuando están alegres al tener la sensación a haber logrado algo o conseguido un activo gratuito.

Scareware (Scare: asustar)

El Scareware es un tipo de ataque y un tipo de software. El Scareware se característica por intentar asustar a la víctima con falsos mensajes de amenaza, urgencia o virus para que llame a un número, mande un correo o realice una acción. La forma de que conseguir que la víctima reciba estos mensajes es normalmente comprar direcciones de web similares a sitios legítimos, de forma que la gente que se equivoca al escribir la dirección legítima acaba en ellos e inmediatamente son bombardeados por pop-ups o banners que tratan de simular mensajes de sistema.

Otras veces el scareware es software que normalmente simula ser un antivirus o antimalware, cuya única característica es mostrar siempre que nuestro sistema tiene muchos errores o riesgo, pidiéndonos que llamemos a un número para asistencia o compremos una licencia. Otras veces el scareware es directamente en sí mismo un virus informático.

Según el **Informe de Cloudflare sobre las amenazas de phishing en 2023**

A nivel global, se estima que el **90%** de los ciberataques exitosos comienzan con el **phishing** por correo electrónico. El impacto financiero medio de una violación de datos causada por phishing se disparó hasta los 4,76 millones de dólares en 2024. Además, el **83%** de las organizaciones fueron víctimas de ataques de phishing exitosos, y el **54%** de estos ataques resultaron en la violación de datos de los clientes.
The Cloudflare BlogMetaComplianceGeekfla (https://blog.cloudflare.com/es-es/2023-phishing-report/?utm_source=chatgpt.com/)

Factor humano: El DBIR 2022 de **Verizon** revela una verdad aleccionadora: el **82%** de las violaciones de datos implican la manipulación humana a través del phishing o el robo de credenciales.

La creciente sofisticación de los ataques de phishing, impulsada por el uso de inteligencia artificial (**IA**) y técnicas de ingeniería social, ha hecho que las campañas sean más personalizadas y difíciles de detectar. Los ciberdelincuentes utilizan herramientas de IA para crear correos electrónicos fraudulentos altamente personalizados, lo que aumenta la probabilidad de éxito de sus ataques. **Financial Times**

Ante este panorama, es fundamental que las organizaciones implementen programas de concienciación y capacitación en anti-**phishing**.

Herramientas de código abierto como **GoPhish** permiten simular ataques de phishing para educar a los empleados y fortalecer la postura de seguridad de la empresa. Estas simulaciones ayudan a identificar vulnerabilidades humanas y a desarrollar estrategias efectivas para mitigar los riesgos asociados al phishing.

Importancia de la IA en los ataques de phishing

📊 Estadísticas y tendencias clave en 2025

- **Aumento de quejas por ciberataques impulsados por IA**: Se estima que en 2025 se registrarán aproximadamente 1,31 millones de quejas relacionadas con ciberataques potenciados por IA, con pérdidas potenciales que podrían alcanzar los 18.600 millones de dólares. ESED
- **Incremento global de intentos de phishing**: Según un informe de Kaspersky, los intentos de phishing han aumentado un 617% a nivel global desde 2022, impulsados por el uso de IA para generar mensajes más creíbles y dirigidos a usuarios específicos. ITware Latam
- **Eficacia de correos electrónicos generados por IA**: Un estudio académico reveló que los correos electrónicos de phishing generados completamente por IA alcanzaron una tasa de clics del **54%**, comparable a los creados por expertos humanos y significativamente superior al 12% de correos genéricos. arXiv
- **Fraude financiero y deepfakes**: Más del 50% del fraude financiero en 2025 involucra el uso de IA y deepfakes, lo que ha llevado a que el 90% de las instituciones financieras adopten soluciones basadas en IA para detectar y prevenir estos delitos. Feedzai

🧠 Técnicas emergentes potenciadas por IA

- **Spear phishing automatizado**: La IA permite a los atacantes recopilar información específica sobre sus víctimas y generar mensajes personalizados que imitan el estilo de comunicación de personas de confianza, aumentando la probabilidad de éxito de los ataques. Whalemate
- **Suplantación de identidad con deepfakes**: Los ciberdelincuentes utilizan IA para crear videos y audios falsos que imitan a ejecutivos o empleados, facilitando fraudes como el compromiso del correo electrónico empresarial (BEC).
- **Evasión de filtros de seguridad**: Los modelos de lenguaje avanzados permiten generar correos de phishing que evaden los filtros tradicionales de spam y detección de amenazas, aumentando la efectividad de los ataques. Hopla! Tech

Importancia de la Capacitación y concienciación anti-phishing

Como hemos visto, este creciente e imparable volumen de ataques de phishing, potenciando su credibilidad por medio de IA, está generando campañas más personalizadas y difíciles de detectar.

Es importante tener en cuenta que este tipo de ataques, contrariamente a lo que se piensa, tiene mayor impacto en gente joven que en adultos mayores, tal cual lo demuestra otro de los análisis realizado por Geekflare, que presentamos a la derecha.

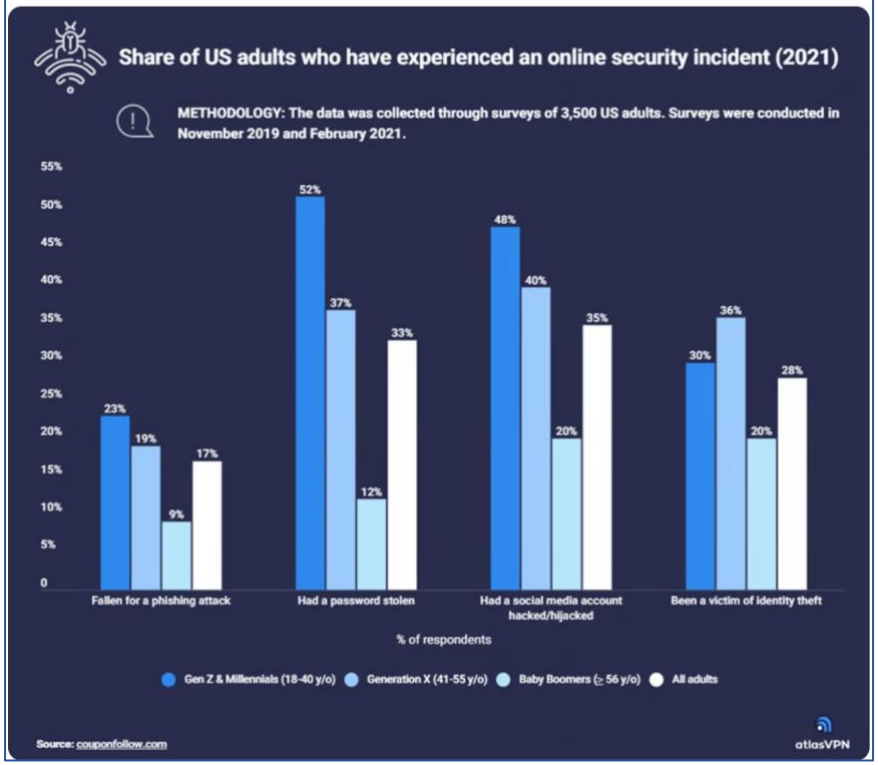
Esta imagen, nos presenta una nueva visión de esta realidad, en la cual si prestamos atención, podemos ver que en cada una de las gráficas, los mayores porcentajes lo tienen las dos columnas de la izquierda que se corresponden a las edades de: 18-40 años y 41-55 respectivamente, es decir la masa de la base activa empresarial.

Todos los estudios demuestran que la mejor forma de combatir este tipos de ataques, que reiteramos va en crecimiento exponencial y son el **90%** de los ciberataques exitosos comienzan con el **phishing** por correo electrónico, es por medio de la:

“Concienciación y capacitación”

GEEKFLARE

Sorprendentemente, los Millennials y los Gen Zers de EE.UU. son víctimas de correos electrónicos de phishing con más frecuencia que sus homólogos de más edad. De hecho, el 23% de los de la Generación Z y Millennial han sido engañados por estos correos fraudulentos en el pasado.



Volviendo a los análisis de Geekflare, otra de sus estadísticas, nos pone de manifiesto que los resultados de formación pueden llegar a reducir los efectos de este tipo de ataque a un 5,4% luego de 12 meses de un plan de formación adecuado. Lo podemos ver en detalle en la imagen de la derecha.

Por todas las razones expuestas, es que este artículo, lo hemos pensado como la base teórica de una serie de videos sobre "**Capacitación y concienciación**", empleando la herramienta Open-Source: **Gophish.**

En nuestro [Canal Youtube], bajo el ciclo "[Aprendiendo Ciberseguridad paso a paso]".

GEEKFLARE

Las cifras del informe KnowBe4 2023 Phishing by Industry Benchmarking Report ponen de relieve los resultados de la formación en materia de concienciación sobre seguridad para los empleados. Según el informe, antes de la formación, cerca del 33,2% de los empleados no eran capaces de superar una prueba de phishing. Esto se ha reducido al 18,5% tras un periodo de 90 días de formación y de nuevo al 5,4% tras un año de formación.