CSURFACE

WHITEPAPER · CREDENTIAL LEAKAGE

Credenciales filtradas · el intervalo entre la exposición y el descubrimiento

Credential Leakage Monitor: por qué la monitorización continua de bases de datos de brechas, la dark web y repositorios públicos supera sistemáticamente cualquier auditoría puntual — con validación activa para eliminar falsos positivos.

CAPACIDAD	PÚBLICO	EDICIÓN
Credential Leakage Monitor	CISO · IAM Lead · SOC	2026

LECTURA	CLASIFICACIÓN	FUENTE
7 páginas	Público	Plataforma CSURFACE

RESUMEN EJECUTIVO

La ventana de reacción se mide en horas

La tesis de este whitepaper, en una página — lo que los ejecutivos necesitan saber antes del detalle técnico.

BRECHAS POR CREDENCIAL

30%

de los incidentes comienzan con una credencial comprometida

TIEMPO HASTA LA DETECCIÓN

94 días

es el promedio entre la filtración y el descubrimiento

DETECCIÓN CON MONITORIZACIÓN CONTINUA

< 24h

de la filtración a la alerta validada

Alrededor del **30% de las violaciones de datos comienzan con una credencial comprometida** (Verizon DBIR). El problema central es que las credenciales se filtran sin ser percibidas. El intervalo promedio entre la filtración de una credencial y su descubrimiento por parte de la organización afectada es de **94 días**.

Noventa y cuatro días representan una amplia ventana operativa para el atacante: tiempo suficiente para concluir el reconocimiento, escalar privilegios, exfiltrar datos y, en muchos casos, cerrar la operación sin evidencias inmediatas. La credencial filtrada constituye el vector de acceso que precede al incidente, disponible meses antes de que cualquier mecanismo de detección interno sea activado.

Este whitepaper presenta el **Credential Leakage Monitor** de CSURFACE: monitorización continua de bases de datos de brechas públicas, la dark web y repositorios públicos de código para detectar credenciales corporativas filtradas en horas. La operación es enteramente externa, sin agente ni instalación, e incluye validación activa que confirma la vigencia de cada credencial antes de la emisión de cualquier alerta.

La tesis. La defensa de credenciales requiere monitorización continua. Las credenciales se filtran de forma continua y, por eso, necesitan ser monitorizadas de forma continua. Reducir la ventana de descubrimiento de trimestres a horas es lo que convierte un incidente de meses en una respuesta contenida en el plazo de un día hábil.

Lo que cubre este whitepaper

- Por qué se filtran las credenciales — y por qué la detección tradicional es lenta.
- Las tres fuentes de filtración que realmente importan para una organización.
- Por qué la validación activa es lo que separa una alerta útil del ruido.
- El workflow de respuesta recomendado y un caso de uso representativo.

EL ESCENARIO · PARTE 01

Las credenciales se filtran continuamente — la detección no acompaña

Cómo las credenciales corporativas llegan a las manos equivocadas y por qué los controles tradicionales tardan en percibirlo.

Whitepaper 2026
CSURFACE

LOS CINCO CAMINOS DE LA FILTRACIÓN DE CREDENCIALES

CRÍTICO Brecha de SaaS de terceros
El colaborador reutilizó la contraseña corporativa; el servicio sufre una brecha y la credencial va al mercado.

CRÍTICO Phishing exitoso
La credencial se escribe en una página falsa y se envía directamente al atacante.

ATENCIÓN Repositorios públicos
Commit accidental de un archivo .env, una clave de nube o una contraseña en un archivo de configuración.

CRÍTICO Infostealers
Malware en el portátil personal exfiltra todas las contraseñas guardadas en el navegador.

ATENCIÓN Acción de insider
Un colaborador descontento comparte o vende accesos a terceros.

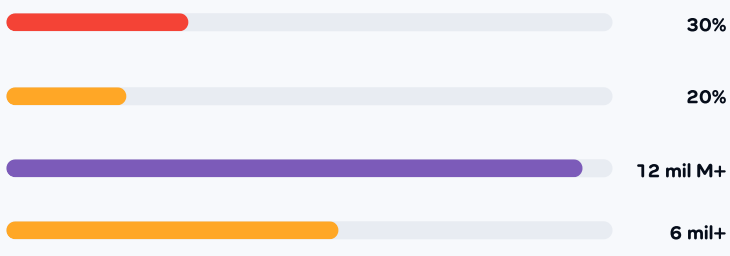
LO QUE CSURFACE OBSERVA · DATOS PROPIOS DE LA PLATAFORMA

170
credenciales corporativas filtradas identificadas en 16 empresas analizadas por la plataforma

37%
de las empresas analizadas tenían al menos una credencial corporativa expuesta

La escala del problema

Ninguno de estos cinco caminos atraviesa el perímetro que la organización controla — y ninguno activa una alerta interna. La credencial se vuelve disponible externamente, en fuentes accesibles a agentes adversarios, sin que se genere señal interna alguna.



Fuentes: Verizon DBIR, SpyCloud, HaveIBeenPwned, GitHub Secret Scanning Trends.

El costo de no detectar. El costo promedio de una violación de datos en Brasil alcanzó los **€3,90M** en 2025 (IBM Cost of a Data Breach 2025), y la sanción administrativa de la RGPD puede alcanzar el 2% de la facturación, limitada a €20M. Cada día eliminado de la ventana de descubrimiento reduce la probabilidad de que la filtración evolucione hacia un incidente — impacto que la calculadora de riesgo de CSURFACE traduce en pérdida anual esperada (ALE).

Por qué la detección es lenta. Las auditorías de contraseñas son puntuales (anuales o semestrales); el MFA reduce el riesgo pero no cubre las cuentas de servicio ni *legacy*; el DLP observa la salida de datos, no el regreso de una credencial a la venta en un foro; y el SIEM solo detecta el uso anómalo después de que el atacante ya está dentro. Ninguno de estos controles observa el lado externo — exactamente donde la credencial filtrada queda expuesta.

LAS TRES FUENTES · PARTE 02

Dónde monitorizar: tres fuentes que importan

Bases de datos de brechas, la dark web y repositorios públicos — qué revela cada fuente y qué hace CSURFACE con ella.

LAS TRES FUENTES MONITORIZADAS EN PARALELO

1. **Bases de datos de brechas**
   Cruce del dominio contra bases agregadas de filtraciones públicas.

2. **Dark web**
   Monitorización de mercados y foros que comercializan credenciales recientemente robadas.

3. **Repositorios públicos**
   Rastreo de nuevos commits en busca de secrets del entorno del cliente.

01. Bases de datos de brechas

Más de **12 mil millones de credenciales** ya han sido expuestas en incidentes públicos — desde filtraciones históricas hasta colecciones consolidadas. Fuentes públicas como HaveIBeenPwned mantienen bases agregadas de esos eventos.

Qué hace CSURFACE. Realiza el cruce continuo del dominio corporativo del cliente contra esas bases. Cuando una credencial vinculada al dominio aparece en cualquier brecha — incluso de un SaaS de terceros — la organización es notificada.

02. Dark web

Mercados, foros y canales de mensajería comercializan credenciales recientemente extraídas por infostealers, con actualización diaria de los lotes disponibles. Organizaciones de perfil relevante llegan a tener dosieres de acceso disponibles para adquisición inmediata.

Qué hace CSURFACE. Monitoriza comunidades conocidas a través de socios de inteligencia de amenazas e identifica la venta activa de credenciales asociadas al dominio del cliente.

03. Repositorios públicos de código

En plataformas como GitHub, GitLab y Bitbucket, los desarrolladores publican inadvertidamente archivos **.env**, claves de acceso de nube y contraseñas en archivos de configuración — del orden de miles de secrets expuestos por día.

Qué hace CSURFACE. Realiza el rastreo continuo de nuevos commits públicos en busca de patrones asociados al entorno del cliente: claves de nube, tokens, hostnames internos y otros secrets.

Por qué tres fuentes, y no una. Cada fuente cubre un camino distinto de filtración: la base de datos de brechas revela la reutilización de contraseñas; la dark web revela la credencial recién robada a la venta; el repositorio público revela el secret olvidado en el código. Monitorizar solo una de ellas deja dos tercios del problema invisibles.

VALIDACIÓN Y WORKFLOW · PARTE 03

Validación activa: alertar solo sobre lo que aún funciona

Detectar una credencial filtrada es la mitad del trabajo. La otra mitad es confirmar si sigue siendo una amenaza.

Por qué importa la validación activa

Una parte significativa de las filtraciones registradas es históricamente antigua: la contraseña ya fue sustituida, la clave ya fue revocada, la cuenta ya fue desactivada. Emitir alertas sobre esas credenciales produce ruido operativo — y la acumulación de ruido compromete la capacidad del equipo de responder a las alertas realmente críticas.

CSURFACE ejecuta una validación automática, segura y no destructiva para confirmar la vigencia de cada credencial detectada. El resultado: la organización recibe notificaciones **exclusivamente** sobre credenciales que permanecen activas, con una reducción acentuada de falsos positivos.

Cómo se valida la credencial, por tipo

SAML / OAuth
Una solicitud de refresh token; si es aceptada, la cuenta está activa.

SMTP / correo
Un intento de autenticación sin envío de mensaje.

Claves de nube
Una llamada de solo lectura de identidad, que falla si la clave fue revocada.

Tokens de API
Una consulta a un endpoint de identidad del tipo /me o /whoami.

Validación no destructiva. Toda verificación confirma la vigencia de la credencial sin alterar datos, enviar mensajes ni disparar bloqueos en la cuenta objetivo. El objetivo es responder a una única pregunta — ¿esta credencial permanece activa? — antes de movilizar al equipo de seguridad para la investigación.

WORKFLOW DE RESPUESTA RECOMENDADO · DE LA ALERTA AL RIESGO RECALCULADO

1. **Notificación inmediata**
   Alerta vía Slack, correo o SMS con contexto: fuente, fecha de la filtración, credencial parcial y resultado de la validación.
2. **Activación del IdP**
   Integración opcional con Okta, Microsoft Entra ID o JumpCloud fuerza el restablecimiento de contraseña y revoca sesiones activas.
3. **Ticket para el SOC**
   Ticket automático con la evidencia reunida, incluyendo el historial reciente de acceso de la cuenta, para investigación.
4. **Panel actualizado**
   La exposición de la organización se recalcula y se refleja en el panel, manteniendo la priorización alineada con el riesgo vigente.

Las integraciones con IdP son opcionales — la detección y la validación operan de forma 100% externa, sin agente.

APLICACIÓN Y RESULTADOS · PARTE 04

De la alerta a la respuesta en horas

Un caso de uso representativo, los indicadores de resultado y el encuadre de cumplimiento.

DETECCIÓN TRADICIONAL

94 días

promedio entre la filtración y el descubrimiento — ventana suficiente para que el atacante complete el ciclo de compromiso.

→

CON EL MONITOR DE CSURFACE

< 24h

de la filtración a la alerta validada — SLA típico de notificación en menos de 4 horas.

Caso de uso · medios, 2.500 colaboradores

Una empresa de medios puso en monitorización el dominio raíz y cuatro plataformas SaaS conocidas. En la **primera semana**, el cruce con las bases de datos de brechas reveló el siguiente embudo:

ATENCIÓN **247 credenciales expuestas**
del dominio, encontradas en brechas públicas.

CRÍTICO **38 aún válidas**
confirmadas por la validación activa de la plataforma.

CRÍTICO **12 de colaboradores actuales**
forzadas a restablecimiento inmediato de contraseña.

ATENCIÓN **26 de excolaboradores**
cuentas huérfanas en SaaS cerradas; offboarding revisado.

En el **mes 2**, una nueva alerta identificó una clave de nube activa publicada en un repositorio público. La validación confirmó la vigencia de la clave, que tenía permiso de lectura sobre un bucket que contenía datos de clientes. La revocación se completó en 4 horas a partir del disclosure — antes de cualquier explotación registrada.

Indicadores de resultado

INDICADOR	BASELINE	CSURFACE
Tiempo hasta la detección	94 días	**< 24 horas**
Falsos positivos	Altos	**Bajo**
Notificación tras la exposición	Puntual	**SLA < 4 horas**
Cobertura de fuentes	Una, si la hay	**Tres fuentes**

Rangos observados en monitorizaciones conducidas por la plataforma CSURFACE. El resultado varía según la dispersión de la superficie de cada organización.

Encuadre de cumplimiento. La detección rápida sostiene la comunicación oportuna a la ANPD prevista en la RGPD (Art. 46), la monitorización de credenciales exigida por la Resolución BACEN nº 4.893, la gestión de acceso de la ISO 27001 (A.9) y la identificación y autenticación del PCI DSS (Requisito 8).

PRÓXIMOS PASOS

La defensa de credenciales comienza con un único dominio

El Credential Leakage Monitor es una de las capacidades de la plataforma CSURFACE — una plataforma de Continuous Exposure Management. En una arquitectura integrada, reúne el descubrimiento continuo de la superficie externa con Machine Learning, el análisis de la cadena digital de proveedores, la validación de explotabilidad, la inteligencia de amenazas con priorización dinámica y la monitorización de credenciales filtradas. Opera enteramente de forma externa — sin agente y sin instalación —, con integraciones opcionales de nube, WAF y CIEM disponibles para organizaciones que requieran una cobertura en profundidad.

Reciba el análisis preliminar

Indique el dominio corporativo y reciba, de forma gratuita, una evaluación de las credenciales ya expuestas en fuentes públicas asociadas a su organización.

Use la calculadora de riesgo

Estime la pérdida anual esperada (ALE) y el VaR de su sector, con metodología FAIR calibrada por benchmark de mercado.

Lea el próximo whitepaper

Inteligencia de Amenazas y Priorización Dinámica — cómo las credenciales comprometidas alimentan la priorización del riesgo.

Vea qué credenciales ya se han filtrado — monitorización iniciada con el dominio raíz

Recibir análisis preliminar gratuito